CALCUL

DES

PROBABILITÉS

PAR

J. BERTRAND,

DE L'ACADÉMIE FRANÇAISE,

SECRETAIRE PERPETUEL DE L'ACADEMIE DES SCIENCES.

> Facile videbis hunc calculum esse sæpe non minus nodosum quam jucundum
>
> DANIEL BERNOULLI.

DEUXIÈME ÉDITION

CONFORME A LA PREMIÈRE.

PARIS,

GAUTHIER-VILLARS, IMPRIMEUR-LIBRAIRE

DU BUREAU DES LONGITUDES, DE L'ÉCOLE POLYTECHNIQUE,

Quai des Grands-Augustins, 55.

1907

PRÉFACE.

Le Calcul des probabilités est une des branches les plus attrayantes des Sciences mathématiques et cependant l'une des plus négligées. Le beau Livre de Laplace en est peut-être une des causes. Deux opinions, en effet, se sont formées, sans rencontrer presque de contradicteurs : on ne peut bien connaître le Calcul des probabilités sans avoir lu le livre de Laplace ; on ne peut lire le livre de Laplace sans s'y préparer par les études mathématiques les plus profondes.

La seconde de ces propositions est incontestable, et le *Traité analytique du Calcul des probabilités* commence par deux cents pages, au moins, dans lesquelles l'exposition des théories mathématiques qui doivent servir au calcul des chances est complètement indépendante de toute application ultérieure. Laplace, après avoir trouvé des méthodes nouvelles, devait leur donner la préférence : les problèmes sont choisis et les solutions proposées de manière à mettre en évidence l'utilité des fonctions génératrices.

J'ai cherché dans ce Livre, résumé de Leçons faites au Collège de France, à faire reposer les résultats les plus utiles et les plus célèbres du Calcul des probabilités sur les démonstrations les plus simples. Bien peu de pages, je crois, pourront embarrasser un lecteur familier avec les éléments de la Science mathématique. Si le signe $\int$ s'introduit quelquefois; il suffit presque toujours d'en connaître la définition.

Je me suis efforcé, à l'occasion de chaque question, de marquer avec précision le degré de certitude des résultats et les limites nécessaires de la Science. La plupart des réflexions suggérées par l'étude approfondie des questions souvent controversées ont été proposées dans un Travail dégagé de toute intervention des signes algébriques, imprimé déjà depuis plusieurs années. Il servira d'Introduction à l'exposé complet des théories.

LES LOIS DU HASARD.

Comment oser parler des lois du hasard ? Le hasard n'est-il pas l'antithèse de toute loi? En repoussant cette définition, je n'en proposerai aucune autre. Sur un sujet vaguement défini on peut raisonner sans équivoque. Faut-il distraire le chimiste de ses fourneaux pour le presser sur l'essence de la matière ? Commence-t-on l'étude du transport de la force par définir l'électricité ?

I.

Le mot *hasard,* intelligible de soi, éveille dans l'esprit une idée parfaitement claire. Quand un joueur de tric-trac jette les dés, s'ils ne sont pas pipés, s'il ne sait ni ne veut amener aucun point plutôt qu'aucun autre, le coup est l'œuvre du hasard. Les grands noms de Pascal, de Fermat, de Huygens décorent le berceau du Calcul des hasards. On est injuste en oubliant Galilée. Un amateur du jeu, qui observait les coups et discutait les chances, lui proposa, comme cinquante ans plus tard le chevalier de Méré à Pascal, une contradiction et un doute. Au jeu de *passe-dix*, on jette trois dés et l'on gagne si la somme des points surpasse 10. Les chances sont égales ; les combinaisons qui passent 10 forment la moitié du nombre total. L'ami de Galilée, très familier avec les dés, s'étonnait de gagner par le point 11 plus souvent que par le point 12 et de voir sortir 10 plus souvent que 9. Ces quatre points arrivent cependant chacun de six manières, et pas davantage. Pourquoi 12 est-il plus rare que 11 ? Faut-il nier l'expérience ou douter du calcul ? Il faut les accorder en faisant mieux le compte. Les cas que l'on dénombre ne sont pas pareils ; 4, 4, 4, par exemple, qui donne 12, n'est pas comparable à 4, 5, 2, qui donne 11 ; la première de ces combinaisons est unique, chacun des trois dés doit amener 4 ; 4, 5, 2, au contraire, représentent six combinaisons, par la même raison qu'avec trois lettres distinctes on peut écrire six mots différents. Attentif à tout circonstancier, Galilée, au lieu de six chances, en montre distinctement vingt-sept pour le point 11, vingt-cinq seulement pour le point 12. Le calcul, le compte, pour parler mieux, s'accorde, comme toujours, avec l'expérience des joueurs. Galilée n'en faisait aucun doute. Quoique ce grand géomètre Jacques Bernoulli, pour avoir établi la loi sur des preuves, ait pris un rang élevé entre les plus illustres, la conviction universelle des joueurs a précédé ses profonds travaux. Quand un dé lui montrait trop souvent la même face, Panurge, qui s'y connaissait, pour y voir biffe et piperie, n'invoquait rien que l'évidence. Ainsi faisait l'ami de Galilée : en comptant 1080 fois le point 11 contre 1000 fois le point 12, il devinait une cause et voulait la connaître.

Un jour, à Naples, un homme de la Basilicate, en présence de l'abbé Galiani, agita trois dés dans un cornet et paria d'amener rafle de 6 ; il l'amena sur-le-champ. Cette chance est possible, dit-on ; l'homme réussit une seconde fois, et l'on répéta la même chose ; il remit les dés dans le cornet, trois, quatre, cinq fois, et toujours rafle de 6. « *Sangue di Bacco!* s'écria l'abbé, les dés sont pipés! » et ils l'étaient. Pourquoi l'abbé jurait-il? Toute combinaison n'est-elle pas possible? Elles le sont toutes, mais inégalement. Galilée nous en avertit. Commençons, pour aller pas à pas, par jeter deux dés ensemble ou deux fois un seul dé, — les deux cas n'en font qu'un. Si deux joueurs parient, l'un pour deux 6, l'autre 6 et 5, les chances, pour eux, sont inégales. *Sonnez* représente l'une des trente-six combinaisons possibles; le 6 et 5 en réunit deux. Si l'un arrive deux fois plus que l'autre, faudra-t-il accuser le hasard de partialité? attribuer au point 6 une antipathie occulte pour son semblable? Cette interprétation n'est pas à craindre.

Si, prenant soixante dés, on compare la réunion des soixante 6, équivalente à trente *sonnez* de suite, avec la combinaison qui contient chacun des six points précisément dix fois, les nombres par leur immensité se dérobent à l'imagination, et l'esprit troublé par une telle abondance cherche les causes d'un mystère qui n'existe pas.

Avec soixante dés, pour amener soixante fois 6, une seule combinaison est possible : chaque dé doit montrer le point 6. Dix 6, au contraire, et dix fois chacun des autres points, peuvent se distribuer et s'arranger avec tant de variété que, si chacun des arrangements possibles était préparé dans une boîte d'un décimètre carré sans que, dans aucune boîte, les mêmes dés présentassent les mêmes faces, la cent-millionième partie de celles que la combinaison désignée enveloppe sous un même nom pourrait couvrir un million de fois la surface de la terre sans y laisser aucun vide. Jeter les soixante dés à la fois, c'est charger le hasard de désigner une des boîtes, et si, dans cette abondance, les combinaisons peu nombreuses ne se montrent jamais, est-ce lui qui les exclut? La boîte qui contient les soixante 6, toutes celles même qui en contiendraient plus de cinquante, sont introuvables dans la masse comme des gouttes d'eau désignées dans l'Océan.

Sur le Pont-Neuf, pendant une journée ou pendant une heure, on peut prédire résolument que les passants de taille inférieure à deux mètres l'emporteront par le nombre. Le pont écarte-t-il les géants? Quand, au jeu de dés, on annonce quelles combinaisons prévaudront, c'est, comme pour les passants du Pont-Neuf, une question d'arithmétique; les combinaisons qu'on ose exclure forment, dans le nombre total, si les épreuves sont nombreuses, une proportion beaucoup moindre que, parmi les Parisiens, les hommes de six pieds de haut.

Buffon, qui, ce jour-là, manqua de patience, fit jeter une pièce de monnaie en l'air 4040 fois; il obtint 2048 fois face au lieu de 2020. Un tel écart n'a rien d'inattendu. Le jeu étudié par Buffon était moins simple que pile ou face. Quelques millions d'épreuves ne pourraient ni en révéler ni en infirmer la loi. La pièce jetée en l'air est jetée de nouveau et de nouveau encore, s'il le faut, jusqu'à l'arrivée de face. Buffon, ayant amené face 2048 fois, a joué 2048 parties.

Un paradoxe singulier rend ce jeu, — ce *problème de Saint-Pétersbourg*, c'est le nom qu'on lui donne, — mémorable et célèbre. Pierre joue avec Paul; voici les conditions : Pierre jettera une pièce de monnaie autant de fois qu'il sera nécessaire pour qu'elle montre le côté face. Si cela arrive au premier coup, Paul lui donnera un écu; si ce n'est qu'au second, deux écus; s'il faut attendre un troisième coup, il en donnera quatre, huit au quatrième, toujours en doublant. Tels sont les engagements de Paul. Quels doivent être ceux de Pierre? La Science, consultée par Daniel Bernoulli, donne pour réponse : Une somme infinie. Le *parti* de Pierre, c'est le mot consacré, est au-dessus de toute mesure.

Les géomètres ont interprété de plusieurs façons et désavoué, comme excessive, la réponse irréprochable de la théorie du jeu. D'Alembert écrivait en 1768 : « Je connais jusqu'à présent cinq ou six solutions au moins de ce problème dont aucune ne s'accorde avec les autres et dont aucune ne me paraît satisfaisante. » Il en ajoute une sixième ou septième, la moins acceptable de toutes. L'esprit de d'Alembert, habituellement juste et fin, déraisonnait complètement sur le Calcul des probabilités.

Buffon, pour expliquer le paradoxe de Saint-Pétersbourg, allègue que posséder ne sert de rien si l'on ne peut jouir. « Un mathématicien, dans ses calculs, — ce sont les propres paroles de Buffon, — n'estime l'argent que par sa quantité, c'est-à-dire par la valeur numérique ; mais l'homme moral doit l'estimer par les avantages et les plaisirs qu'il peut procurer. » On promet à Pierre de doubler son gain à chaque coup qui retarde l'arrivée de face, on ne peut doubler que ses écus. Pierre ne demande rien de plus, Buffon peut en être certain. « L'accroist de chevance, avait dit avant lui Montaigne, n'est pas l'accroist d'appétit au boire, manger et dormir... » ; chacun peut allonger la liste. Daniel Bernoulli, réduisant cette distinction en formule, oppose à la richesse mathématique une richesse morale que l'or accroît, mais si lentement, que toutes les unités, jusqu'à la dernière, procurent un égal contentement.

Cette théorie condamne tous les jeux de hasard. Le conseil de ne jouer jamais, si excellent qu'il soit, ne peut être proposé pour une théorie du jeu. Supposons en présence deux disciples de Bernoulli. « Si je gagne, dirait Pierre, qui est pauvre, en proposant à Paul une partie d'écarté, votre enjeu de 3^{fr} payera mon dîner. — Repas pour repas, répondrait Paul, vous me devrez 20^{fr} en cas de perte, car tel sera le prix de mon souper. — Si je perdais 20^{fr}, s'écrierait Pierre effrayé, je ne dînerais pas demain; vous pouvez, sans en venir là, perdre $10\,000^{fr}$, déposez-les contre mes 20^{fr}; l'avantage, Daniel Bernoulli l'affirme, restera de votre côté. » — Ils ne s'entendront pas.

Ceux qui suivent Condorcet et Poisson, sans contester la bonne foi de Paul, tiennent ses engagements pour nuls. Si le hasard amenait pile soixante quatre fois, Paul devrait payer autant d'écus que le sultan des Indes ne put donner de grains de blés à l'inventeur du jeu d'échecs Une telle promesse est téméraire; si riche qu'on le suppose, Paul, ruiné dès le trentième coup, ne pourra plus payer double. Ne comptant plus sur ses promesses, Pierre ne doit pas les payer, et le calcul règle le droit de Paul à 15 écus.

On propose à 50 personnes possédant chacune 20 millions et pas davantage d'organiser une loterie à 20 millions le billet. Le gagnant deviendra l'homme le plus riche du monde, les 49 autres seront ruinés. Les 50 vigésimillionnaires acceptent. Ils sont peu sensés, mais équitables. La justice et la raison sont choses distinctes. Au jeu de Saint-Pétersbourg, tout aussi bien qu'à cette loterie, les espérances doivent être payées; il ne s'agit plus d'un seul, mais d'un nombre illimité de milliards. Le problème imaginé par Daniel Bernoulli dissimule ingénieusement cette énorme mise. L'Algèbre, en la dégageant, met la chance à son juste prix.

Les conditions d'un jeu peuvent être équitables et dangereuses; iniques dans d'autres cas, mais acceptables. Est-il déraisonnable, malgré le o, le double o et le refait, de risquer 5^{fr} à la roulette ou au trente-et-quarante ?

Quant au problème de Saint-Pétersbourg, il faut approuver absolument et simplement la réponse réputée absurde. Pierre possède, je suppose, un million d'écus et les donne à Paul en échange des promesses convenues. Il est fou ! dira-t-on. Le placement est aventureux, mais excellent ; l'avantage infini est réalisable. Qu'il joue obstinément, il perdra 1 partie, 1000, 1000 millions, 1 million de milliards peut-être ; qu'il ne se rebute pas, qu'il recommence un nombre de fois que la plume s'userait à écrire, qu'il diffère surtout le règlement des comptes, la victoire, pour lui, est certaine, la ruine de Paul inévitable. Quel jour ? quel siècle ? On l'ignore ; avant la fin des temps certainement, le gain de Pierre sera colossal.

Une fourmi transporte un grain de poussière de la cime du mont Blanc dans la plaine, retourne sur la hauteur, descend une nouvelle charge et recommence toujours. Après combien de voyages aura-t-elle comblé les vallées et nivelé la chaîne des Alpes ? Le premier écolier, sans consulter l'arénaire d'Archimède, fera le calcul sans erreur. Le dessein de la fourmi dépasse ses forces, s'écrieront des gens sages ; elle mourra à la peine. Condorcet et Poisson ne sont pas moins sages. Pierre est un imprudent ; il entreprend, au delà de son crédit, une opération beaucoup trop longue ; il est aussi certain pourtant de ruiner Paul que la fourmi de niveler la Suisse.

Dans un problème plus célèbre et plus grave, la vie humaine servait d'enjeu. L'inoculation, avant la vaccine, était, contre la variole, le meilleur parti qu'on pût prendre ; mais 1 inoculé sur 200 mourait des suites de l'opération. Quelques-uns hésitaient ; Daniel Bernoulli, géomètre impassible, calculait doctement la vie moyenne, la trouvait accrue de trois ans et déclarait par syllogisme l'inoculation bienfaisante. D'Alembert, toujours hostile à la théorie du jeu, qu'il n'a jamais comprise, repoussait, avec grande raison cette fois, l'application qu'on en voulait faire : « Je suppose, dit-il, que la vie moyenne d'un homme de trente ans soit trente autres années et qu'il puisse raisonnablement espérer de vivre encore trente ans en s'abandonnant à la nature et en ne se faisant pas inoculer. Je suppose ensuite qu'en se soumettant à cette opération la vie moyenne soit de trente-quatre ans. Ne semble-t-il pas que, pour apprécier l'avantage de l'inoculation, il ne suffit pas de comparer la vie moyenne de trente-quatre ans à la vie moyenne de trente, mais le risque de 1 sur 200, auquel on s'expose, de mourir dans un mois, par l'inoculation, à l'avantage éloigné de vivre quatre ans de plus au bout de soixante ans? »

On argumente mal pour vider de telles questions : supposons que l'on puisse, par une opération, accroître la vie moyenne, non plus de quatre, mais de quarante ans, à la condition qu'une mort immédiate menacera le quart des opérés : un quart des vies sacrifié pour doubler les trois autres, le bénéfice est grand. Qui voudra le recueillir? Quel médecin fera l'opération? Qui se chargera, en y invitant 4000 habitants robustes et bien portants d'une même commune, de commander pour le lendemain 1000 cercueils? Quel directeur de collège oserait annoncer à 50 mères, qu'empressé à accroître la vie moyenne de ses 200 élèves, il a joué pour eux ce jeu avantageux et que leurs fils sont les perdants? Les parents les plus sages acceptaient 1 chance sur 200; aucun, sur la foi d'aucun calcul, ne s'exposerait à 1 chance sur 4.

Un jeu, sans blesser la justice, peut causer de grands dommages; il peut être périlleux d'y échanger les chances de perte et de gain, les règles que doivent suivre ceux qui veulent commettre cette imprudence n'en reçoivent aucun changement.

Un ingénieur calcule la charge capable d'abaisser de $0^{m},50$

le tablier d'un pont. L'épreuve est inutile, imprudente, dangereuse : le poids calculé est-il moins juste ? Il est mauvais de trop charger un pont, mauvais aussi de jouer trop gros jeu. Cela ne change ni la théorie du jeu ni celle de l'élasticité.

Revenons au théorème de Bernoulli.

S'il pleut un jour entier sur la place du Carrousel, tous les pavés seront également mouillés. Sous une forme simplifiée, mais sans rien retrancher, c'est là le théorème de Bernoulli. Il pourrait se faire assurément, lorsque tout alentour la pluie tombe à torrents, qu'un certain pavé restât sec. Aucune goutte n'a pour lui de destination précise, le hasard les disperse, il peut les porter toutes sur les pavés voisins ; personne ne le supposera sérieusement.

Telle est la puissance des grands nombres. Le hasard a des caprices, jamais on ne lui vit d'habitudes. Si 1000 gouttes tombent sur 1000 pavés, chaque pavé n'aura pas la sienne ; s'il en tombe 1000 millions, chaque pavé recevra son million ou bien peu s'en faudra. Si l'on jette deux dés 36 millions de fois, le double-six, au lieu de 1 million de fois, pourrait ne se présenter que 100000, et peut-être n'arriver jamais. Une telle exclusion soumise au calcul, d'après notre façon de parler, est déclarée impossible.

L'analogie va à l'identité. Considérons en effet, sur la place, pendant la pluie, un carré de $0^{m},6$ de côté. Partageons la base, aussi bien que la hauteur, en 6 parties, portant chacune un numéro d'ordre ; découpons le carré, par des parallèles aux côtés, en 36 cases égales désignées chacune par les deux numéros placés en tête des bandes auxquelles elle appartient ; une case répondra à 6,6 ; une autre à 5,6 ; une troisième à 6,5 ; elles auront mêmes noms que les coups possibles avec deux dés. Chaque goutte de pluie tombant sur le carré représente un coup de dés. Le hasard, dans les deux épreuves, décide entre les mêmes points. A la fin de la journée, la pluie a également mouillé les 36 cases, les dés ont amené les 36 points également : où est la différence ?

Rien ne manque au rapprochement et le même tempérament est nécessaire aux deux assertions trop précises. Il serait fort étrange que les pavés, quoique mouillés également, n'eussent pas reçu dans le cours d'une journée quelques centaines de gouttes en plus ou en moins ; de même, sur quelques millions de coups de dés, quelques points se montrent sans doute un peu plus, d'autres un peu moins souvent.

Les rapports sont certains, non les différences, et c'est malheureusement la différence qui ruine. On joue 100 parties à un jeu de hasard, l'enjeu est 20^{fr}; il est peu probable, mais possible, que l'on perde 65 parties. La perte de 30 louis représente 30 pour 100 du nombre des parties jouées.

Au lieu de 100 parties, on en joue 10000; une perte de 30 pour 100, c'est-à-dire de 6500 parties, doit être tenue pour impossible. 5150 parties perdues supposeront, d'après le calcul, une fortune aussi adverse que 65 sur une série unique de 100 parties; la perte correspondante, 300 louis, représente 3 pour 100 du nombre des parties jouées.

Sur 1 million de parties, une perte de 3 pour 100 supposerait, contre les lois du hasard, un dérèglement qui jamais ne s'est vu ; 3 pour 1000 représente une chance défavorable équivalente à celle des deux hypothèses précédentes. Trois parties sur 1000, pour 1 million de parties, feraient une perte de 3000 louis; un jeu égal devient à la longue dangereux. Non seulement les lois du hasard permettent la ruine du joueur, elles la prédisent. Tout joueur se ruinera si le temps ne lui manque pas. Lagrange, Laplace et Ampère l'ont démontré; leurs raisonnements n'ont corrigé personne, ils intéressent tout le monde.

Si deux joueurs jouent sans cesse jusqu'à la ruine de l'un d'eux, le moins riche sera probablement vaincu. Le rapport du nombre des parties gagnées ou perdues différera de moins en moins de l'unité, mais la différence augmentera, comme nous l'avons dit ; tantôt l'un sera en perte, tantôt l'autre. La différence, petite d'abord, deviendra grande. La perte, dans ses oscillations, frappera chacun des deux joueurs alternativement ; quand elle dépassera la fortune du perdant, la ruine pour lui sera consommée. Le danger menace surtout, on le comprend, le moins riche des deux joueurs. L'homme qui joue sans limite et sans cesse accepte tous les adversaires, dont l'ensemble, sans changer son sort, peut recevoir un nom collectif : *le public*, qui n'est jamais ruiné, ruine les imprudents qui l'attaquent.

Tout change quand les conditions du jeu sont inégales. Le moindre avantage fait pencher la balance. Pour le joueur que les conditions favorisent, le gain augmente sans limite. Au trente-et-quarante, par exemple, l'avantage du banquier est un peu plus de 0,6 pour 100. Si l'on joue 100 parties, en évaluant à 1000^{fr} la somme des enjeux pour chacune d'elles, l'avantage réservé au banquier par les règles du jeu représente 600^{fr}. Les accidents du hasard produiront un écart dont la valeur moyenne, indiquée par le calcul, est 8000^{fr}. Le banquier, sur une série de 100 parties, a donc chances égales, à très peu près, de perdre ou de gagner. La perte moyenne, c'est tout son avantage, est un peu moindre que le gain moyen.

Sur 10000 parties, en supposant toujours l'enjeu de 1000^{fr}, l'avantage ménagé au banquier par les règles du jeu représente 60000^{fr}. L'écart moyen, dix fois plus grand seulement pour un nombre centuple de parties, est 80000^{fr}. La perte du banquier sur 10000 parties sera donc un événement très ordinaire; mais, en ce cas, la valeur moyenne de la somme perdue sera 20 000$^{\text{fr}}$, tandis que, dans l'hypothèse plus vraisemblable du gain, la valeur moyenne est 140 000$^{\text{fr}}$.

Sur 1 million de parties, le bénéfice régulier, équivalent à l'avantage réservé au banquier, serait 6 millions; l'écart moyen en plus ou en moins, 800 000$^{\text{fr}}$ seulement; s'il gagne moins de 5 millions, le banquier a eu du malheur; un gain inférieur à 4 millions serait invraisemblable, et il y a plus de 10000 à parier contre 1 que son gain ne s'abaissera pas au-dessous de 2 millions.

La loi de Bernoulli, quand elle est mise en défaut, révèle une cause perturbatrice du hasard.

Tels se montrent souvent les résultats du suffrage universel. Supposons 10 millions d'électeurs. Attribuons 6 millions de votes à un parti, celui de la majorité. 4 millions seulement à la minorité. On forme 1000 collèges de 10000 électeurs chacun : tout candidat qui réunira plus de 5000 suffrages sera élu. L'opinion approuvée par les quatre dixièmes des votants serait représentée proportionnellement par 400 députés sur 1000. Les lois du hasard ne lui accordent rien. Sur 1000 représentants, pas un seul pour elle. Le calcul réduit à zéro, pour ainsi dire, la vraisemblance de toute autre hypothèse. Supposons, pour donner une idée des chiffres, que, saisissant l'occasion pour tenter la chance, un joueur s'engage, dans les conditions électorales supposées, à payer autant de millions qu'il se trouvera de députés de la minorité vainqueurs dans la lutte. On ne pourrait pas, en échange de ses promesses, c'est la réponse rigoureuse, sinon acceptable, du calcul, lui offrir équitablement plus d'un centime.

Ce centime pourrait lui coûter cher. Les minorités, même beaucoup moindres, obtiennent quelques représentants. Les électeurs n'étant pas associés par le sort, les influences locales triomphent des lois du hasard. C'est avec grande défiance qu'il faut, sur les traces de Condorcet, *éclairer les Sciences morales et politiques par le flambeau de l'Algèbre*.

Les étoiles, sur la voûte céleste, semblent semées sans ordre et sans loi; 3000 environ, pour qui a la vue bonne, brillent au-dessus du notre horizon. Ptolémée, dans son Catalogue, n'en inscrivait que 1020. Un astronome dont le nom est resté obscur sans injustice, l'archevêque Mitchell, a fait d'une idée ingénieuse et juste une application trop hardie. Si le hasard distribuait sur la voûte du ciel 3000 points brillants, quelle serait la distance moyenne de chacun d'eux à son voisin le plus proche? Le problème est intéressant; Mitchell ne le résout pas; mais, remarquant dans la constellation du Dragon deux étoiles situées à 3′ l'une de l'autre, il trouve que contre un tel rapprochement on pourrait, *a priori*, parier 80 contre 1; dirigeant ensuite ses calculs sur le groupe des Pléiades, Mitchell conclut à 500000 chances contre 1 pour qu'une cause, en dehors du hasard, ait rapproché les six étoiles.

En proposant la mesure précise d'assertions aussi vagues, on peut compromettre la Science. Si Mitchell, soupçonnant entre les étoiles un lien mécanique, avait tiré avantage de leur rapprochement singulier, s'il avait déclaré vraisemblable, très vraisemblable, presque certain, qu'une cause particulière a troublé pour elles les lois générales, il serait sans reproche, mais la précision du chiffre $\frac{1}{500000}$ ne peut trouver d'approbateurs. Les appréciations sans chiffres n'engagent à rien, un chiffre engage la Science, et c'est sans aucun droit.

L'application du calcul aux questions de ce genre est une illusion et un abus.

« Les motifs de croire que, sur 10 millions de boules blanches mêlées à 1 noire, ce ne sera pas la noire que je tirerai du premier coup est de même nature, a écrit Condorcet, que le motif de croire que le Soleil ne manquera pas de se lever demain. » L'assimilation n'est pas permise : l'une des probabilités est objective, l'autre subjective. La probabilité de tirer la boule noire du premier coup est $\frac{1}{10000000}$, ni plus ni moins. Quiconque l'évalue autrement se trompe. La probabilité pour que le Soleil se lève varie d'un esprit à l'autre. Un philosophe peut, sans être fou, annoncer sur la foi d'une science que le Soleil va bientôt s'éteindre ; il est dans son droit comme Condorcet dans le sien ; tous deux l'excéderaient en accusant d'erreur ceux qui pensent autrement. L'assimilation à une urne est le procédé de démonstration. Une urne contient des boules blanches, peut-être aussi des noires ; on y fait 1 million de tirages, tous donnent des boules blanches ; quelle est la probabilité pour qu'un nouveau tirage amène une noire ? Le calcul répond : $\frac{1}{1000000}$. « On a vu, conclut Condorcet, 1 million de fois le Soleil se lever du côté de l'orient, quelle est la probabilité pour qu'il manque demain ? La question n'est-elle pas la même ? » Elle est différente. L'urne, dans le premier cas, est invariable ; qui peut, dans le second, savoir le train des choses ?

Paul, sur la foi de Condorcet, veut parier que le Soleil se lèvera demain. La théorie fixera les enjeux. Paul recevra 1^{fr} si le Soleil se lève et donnera 1 million s'il fait défaut. Pierre accepte le pari. Au lever de chaque aurore, il perd 1^{fr} et le paye. La chance pour lui diminue chaque jour, puisque le Soleil compte un lever de plus. Paul consciencieusement augmente son enjeu ; consciencieusement aussi, Pierre continue à lui payer 1^{fr}. Les conventions demeurent équitables. Les parieurs voyagent, on parcourt vingt contrées, de l'occident à l'orient, Pierre perd toujours ; il poursuit sa chance cependant, conduit Paul vers le nord ; on franchit le cercle polaire ; le Soleil reste un mois au-dessous de l'horizon ; Paul perd 30 millions, croit l'ordre de nature perverti et soupçonne que l'urne est changée.

Tarquin l'ancien, rebelle aux prétentions de l'augure Accius Nævius, osa, dit-on, le mettre au défi. Ce que je pense est-il possible ? demanda le roi. L'augure accepta l'épreuve. « Tu peux donc couper cette pierre ? » Nævius prit un rasoir et coupa le caillou. Avec une très louable impartialité, Condorcet a cherché la chance de vérité. Le point de départ de son calcul est le nombre des cailloux que, depuis l'invention des rasoirs, on n'a pas réussi à couper, et, sans répondre du détail des chiffres, il évalue à $\frac{1}{1000000}$ la probabilité de l'anecdote. Il est un peu naïf. Un caillou que l'on coupe comme un radis est un caillou miraculeux ou un faux caillou. La saine philosophie dont il se vante repousse tout miracle ; l'accord fait sous main entre Nævius et le roi sauverait la vraisemblance. Pour résoudre le problème, au lieu de compter des cailloux, il faut comparer, si on le connaît, le nombre des princes capables d'imposture à celui des augures complaisants et des historiens sans critique.

Le hasard, à tout jeu, corrige ses caprices. Les irrégularités mêmes ont leur loi.

Supposons qu'à un jeu de pur hasard une série de parties ait été jouée. Précisons, pour plus de clarté : le jeu est pile ou face ; la série de 100 parties. Pour chacune, on marque la différence entre le nombre des gains et le nombre normal 50. Si l'on a gagné 44 ou 56 fois, on marque 6 dans les deux cas. Chaque série, de cette manière, se trouve caractérisée par un nombre que nous appellerons l'*écart* ; supposons obtenus 1 million d'écarts. Le hasard décide leur grandeur, comme si l'on puisait 1 million de fois dans un sac contenant des boules de loto. La différence est grande cependant : tandis que toutes les boules sortiront également ou peu s'en faut, les petits écarts seront les plus nombreux. Chacun se présentera, à la longue, un nombre de fois proportionnel à la probabilité que l'on peut calculer ; la régularité des résultats peut recevoir une forme apparente et visible. Marquez sur une ligne droite, à distances égales et petites, les chiffres 0, 1, 2, 3, ... représentant les écarts possibles. Par chacun de ces points élevons une hauteur égale au nombre de fois que l'écart s'est produit ; les extrémités de ces lignes feront paraître une courbe, toujours de même forme ; le sommet correspond au point 0 ; l'abaissement, à partir de ce point, très lent d'abord, s'accroît suivant une loi prévue par le calcul. Si quelques irrégularités déparent le dessin, doublez, décuplez le nombre des épreuves, l'exactitude des prédictions est à peine croyable.

Les grands nombres régularisent tout. La moyenne de tous les écarts peut être prédite avec confiance : elle sera 4 si la série est de 100 épreuves, 40 si elle est de 10 000. La même certitude s'attache à la moyenne des carrés des écarts, à celle de leurs cubes, de leur quatrième puissance. Pour des séries de 100, par exemple, la moyenne des carrés est 25. Ces prédictions sont sûres. N'est-ce pas, pour ainsi parler, miracle de voir un hasard aveugle dicter des résultats exactement prévus ?

Aidée de ces théorèmes singuliers, la dextérité des géomètres a su, chose merveilleuse, rencontrer sur ces voies détournées une solution de la quadrature du cercle. Si, dans une série d'épreuves suffisamment nombreuses, on divise la moyenne des carrés des écarts par le carré de la moyenne, le quotient est, à très peu près, la moitié de la surface du cercle de rayon unité. Avec de la patience, le succès est certain.

Beaucoup de joueurs, préoccupés de cette régularité nécessaire dans les moyennes, cherchent, dans les coups qui précèdent celui qu'ils vont jouer, une indication et un conseil. Ce n'est pas bien entendre les principes. La Science, à ces chimères, ne reste pas sans réponse. La décision du bon sens suffit, elle est nette et claire : à quoi bon la traduire en Algèbre ? Le préjugé est opiniâtre. Les géomètres perdraient à le combattre leur temps et leurs formules.

L'illusion repose sur un sophisme : on allègue la loi de Bernoulli comme certaine ; elle n'est que probable. Sur 20000 épreuves, dit-on, à la roulette, la noire ne peut pas sortir plus de 10500 fois, l'assertion de la Science est formelle. Si les 10000 parties ont donné 6000 noires, les 10000 suivantes ont donc contracté une dette envers la rouge. On fait trop d'honneur à la roulette : elle n'a ni conscience ni mémoire. En supposant qu'à une rencontre inouie succédera, pour la réparer, un nouvel écart de la règle, on n'efface pas l'invraisemblance, on la redouble.

La certitude des lois de Bernoulli est celle d'un chasseur très adroit, qui, connaissant son arme, est certain d'abattre une bête féroce à dix pas. La bête se présente, il la manque ; en la voyant, furieuse, se ruer et l'assaillir, doit-il rester impassible, confiant dans la certitude de l'avoir tuée ?

II.

Le hasard, sans choisir, régularise tout ; la raison en est que, si toutes les combinaisons, dont le nombre est immense, étaient présentes matériellement, les moins nombreuses deviendraient introuvables. Le hasard reste libre, mais la carte est forcée.

Appliquée aux dés, aux cartes, au jeu de rouge et de noire, aux numéros pairs ou impairs, à pile ou face, la théorie des chances est indiscutable : rien n'y altère la rigueur des preuves ; l'Algèbre exécute plus rapidement les dénombrements qu'avec de la patience et du temps on pourrait faire sur ses doigts. Tous les arrangements sont également possibles ; que les plus nombreux se présentent, il n'y a pas de sujet d'étonnement.

La Physique, l'Astronomie, les phénomènes sociaux, semblent, dans plus d'un cas, régis par le hasard. Peut-on comparer la pluie ou le beau temps, l'apparition ou l'absence des étoiles filantes, la santé ou la maladie, la vie ou la mort, le crime ou l'innocence à des boules blanches ou noires tirées d'une même urne ? Le même désordre apparaît dans les détails, cache-t-il la même uniformité dans les moyennes ? Retrouvera-t-on dans les écarts les traits connus et la physionomie des effets du hasard ?

Tout événement qui alterne avec son contraire est comparable aux boules blanches ou noires puisées dans un sac ; le sac est-il toujours le même ? est-il ouvert ? Une force intelligente, se proposant une fin, intervient-elle dans une mesure petite ou grande pour corriger les caprices du sort ? Le raisonnement ne peut devancer l'expérience ; les observations, soigneusement discutées, condamnent, en même temps que les sceptiques rebelles à tout rapprochement, les esprits absolus qui prétendent tout soumettre au calcul.

L'empreinte du hasard est marquée, très curieusement quelquefois, dans les nombres déduits des lois les plus précises. Une Table de logarithmes en témoigne. Pour 10000 nombres successifs, dans les Tables à 10 décimales de Véga, je prends la septième figure du logarithme : rien dans ce choix n'est laissé au hasard. L'Algèbre gouverne tout, une loi inflexible enchaîne tous les chiffres. Si l'on compte cependant les résultats, on aura, à très peu près, sur 10000, 1000 fois le chiffre 0, 1000 fois le chiffre 1 et ainsi des autres ; la formule se conforme aux lois du hasard. Vérification faite, sur 10000 logarithmes, le septième chiffre s'est trouvé 990 fois égal à 0 ; 997 fois à 1 ; 993 fois à 2 ; 1012 fois à 4. En partageant les 10000 nombres en dix séries et prenant pour chacune les moyennes des écarts, j'entends la différence entre le nombre des apparitions de l'un des chiffres et le nombre normal 100, et les comparant à la moyenne du carré des écarts, le rapport des nombres, qui, d'après les lois du hasard, devrait être 1,570796, moitié du nombre que les géomètres désignent habituellement par la lettre π, se trouve égal à 1,561 ; le même calcul fait à l'aide du chiffre 1 donne 1,598, et la moyenne de ces deux résultats est 1,579. Les trois premiers chiffres sont exacts.

La marque du hasard semble visible. Pouvait-on cependant le mieux tenir à l'écart? Nos lois expriment une propriété commune aux combinaisons les plus nombreuses ; elles se vérifient quand on ne choisit pas, il ne suffit pas de choisir pour s'y soustraire.

Le partage des naissances entre les deux sexes a été étudié sur plus de 200 millions d'enfants. Depuis près de deux siècles le nombre des garçons a dépassé celui des filles ; aucun pays ne fait exception ni aucune époque. Le rapport varie peu : le nombre des garçons, pour 100 filles, est compris, pour un grand nombre de naissances, entre 104 et 108. On s'est demandé si cette supériorité observée chez toutes les races, dans les villes comme à la campagne, au midi comme au nord, chez les plus pauvres comme chez les plus riches, est une loi de l'humanité ou un accident fortuit.

A notre époque et pour notre état social, l'évidence est complète ; ni les calculs ne sont nécessaires ni les raisonnements. Ils le sont pour un second problème. Les variations observées d'une année à l'autre pour un même pays, d'une province à l'autre pour une même année, sont-elles assimilables aux résultats capricieux du hasard? Peut-on voir dans la constance approchée du rapport un témoignage suffisant de la loyauté du jeu? Je précise la question : une urne, toujours la même, contient des boules noires et blanches; on y puise une boule au moment de chaque naissance. Pourrait-on sans invraisemblance représenter par le nombre de boules de chaque couleur la proportion variable des naissances? Le nombre des noires, bien entendu, l'emporte sur celui des blanches dans la proportion qui convient au succès.

Les écarts de la moyenne produits par le hasard sur 1 million d'épreuves, pour un événement dont la probabilité diffère peu de $\frac{1}{2}$, ont pour valeur moyenne 400. De plus grands écarts sont possibles assurément, mais leur probabilité diminue rapidement. On peut parier 1000 contre 1 pour un écart moindre que 1600. La probabilité d'un écart supérieur à 2000 est $\frac{1}{100000000}$. Telles sont les indications du calcul.

2000 naissances masculines en plus sur 1 million accroîtraient de moins d'un centième le rapport du nombre de garçons à celui des filles. Les rapports extrêmes fournis par la Statistique, 1,04 et 1,08, diffèrent trop l'un de l'autre pour permettre l'assimilation pure et simple aux effets du hasard. Les conditions ne peuvent donc être, en tout temps et en tout pays, identiquement les mêmes, mais la variation est petite. Pendant l'année 1837, le nombre des garçons nés à Paris est descendu à 10074 pour 10000 filles. Dans les hasards d'un tirage au sort dont les conditions seraient invariables, sur un nombre d'épreuves égal à celui des naissances annuelles à Paris, on pourrait parier plus de 1 million contre 1 dant les résultats, on aura, à très peu près, sur 10000, 1000 fois le chiffre 0, 1000 fois le chiffre 1 et ainsi des autres ; la formule se conforme aux lois du hasard. Vérification faite, sur 10000 logarithmes, le septième chiffre s'est trouvé 990 fois égal à 0 ; 997 fois à 1 ; 993 fois à 2 ; 1012 fois à 4. En partageant les 10000 nombres en dix séries et prenant pour chacune les moyennes des écarts, j'entends la différence entre le nombre des apparitions de l'un des chiffres et le nombre normal 100, et les comparant à la moyenne du carré des écarts, le rapport des nombres, qui, d'après les lois du hasard, devrait être 1,570796, moitié du nombre que les géomètres désignent habituellement par la lettre π, se trouve égal à 1,561 ; le même calcul fait à l'aide du chiffre 1 donne 1,598, et la moyenne de ces deux résultats est 1,579. Les trois premiers chiffres sont exacts.

La marque du hasard semble visible. Pouvait-on cependant le mieux tenir à l'écart? Nos lois expriment une propriété commune aux combinaisons les plus nombreuses ; elles se vérifient quand on ne choisit pas, il ne suffit pas de choisir pour s'y soustraire.

Le partage des naissances entre les deux sexes a été étudié sur plus de 200 millions d'enfants. Depuis près de deux siècles le nombre des garçons a dépassé celui des filles ; aucun pays ne fait exception ni aucune époque. Le rapport varie peu : le nombre des garçons, pour 100 filles, est compris, pour un grand nombre de naissances, entre 104 et 108. On s'est demandé si cette supériorité observée chez toutes les races, dans les villes comme à la campagne, au midi comme au nord, chez les plus pauvres comme chez les plus riches, est une loi de l'humanité ou un accident fortuit.

A notre époque et pour notre état social, l'évidence est complète ; ni les calculs ne sont nécessaires ni les raisonnements. Ils le sont pour un second problème. Les variations observées d'une année à l'autre pour un même pays, d'une province à l'autre pour une même année, sont-elles assimilables aux résultats capricieux du hasard? Peut-on voir dans la constance approchée du rapport un témoignage suffisant de la loyauté du jeu? Je précise la question : une urne, toujours la même, contient des boules noires et blanches ; on y puise une boule au moment de chaque naissance. Pourrait-on sans invraisemblance représenter par le nombre de boules de chaque couleur la proportion variable des naissances? Le nombre des noires, bien entendu, l'emporte sur celui des blanches dans la proportion qui convient au succès.

Les écarts de la moyenne produits par le hasard sur 1 million d'épreuves, pour un événement dont la probabilité diffère peu de $\frac{1}{2}$, ont pour valeur moyenne 400. De plus grands écarts sont possibles assurément, mais leur probabilité diminue rapidement. On peut parier 1000 contre 1 pour un écart moindre que 1600. La probabilité d'un écart supérieur à 2000 est $\frac{1}{100000000}$. Telles sont les indications du calcul.

2000 naissances masculines en plus sur 1 million accroîtraient de moins d'un centième le rapport du nombre de garçons à celui des filles. Les rapports extrêmes fournis par la Statistique, 1,04 et 1,08, diffèrent trop l'un de l'autre pour permettre l'assimilation pure et simple aux effets du hasard. Les conditions ne peuvent donc être, en tout temps et en tout pays, identiquement les mêmes, mais la variation est petite. Pendant l'année 1837, le nombre des garçons nés à Paris est descendu à 10074 pour 10000 filles. Dans les hasards d'un tirage au sort dont les conditions seraient invariables, sur un nombre d'épreuves égal à celui des naissances annuelles à Paris, on pourrait parier plus de 1 million contre 1 qu'une telle anomalie ne se produira pas. Que s'est-il passé en 1837 ? On doit s'attendre à l'ignorer toujours. Dans plusieurs départements, depuis le commencement du siècle, le nombre des naissances annuelles des filles a surpassé exceptionnellement celui des garcons. L'anomalie a moins d'importance que l'écart observé à Paris, elle se rapporte à des nombres cinq fois moindres.

La recherche des causes est délicate et obscure. Il est à regretter, dit M. Quetelet après de longues et patientes recherches, qu'on ait si peu de documents pour s'éclairer.

L'âge des parents joue sans doute un grand rôle. Cette explication semble la meilleure. Si on ne l'accepte qu'avec doute, c'est que, masquée par le hasard, l'influence reste mal connue ; l'âge moyen du père et celui de la mère varient peu dans un même pays. La variation des âges peut cependant expliquer, en partie au moins, les anomalies observées.

Allons plus avant et cherchons dans les effets troublés les traits généraux du hasard.

La quadrature du cercle déduite approximativement du nombre des naissances ne laisse guère subsister de doutes. En appliquant la formule des écarts aux 86 départements pendant l'année 1878 et prenant dans l'*Annuaire du Bureau des Longitudes* les écarts entre le nombre des naissances de garçons correspondant à 10000 filles pour chacun des départements et la moyenne pour la France entière, et la comparant à la moyenne de leurs carrés, au lieu du quotient 1,57 prévu par la théorie, on obtient 1,75. La petitesse de l'erreur paraît digne d'attention.

La recherche des causes est le grand problème : on le transforme sans le résoudre. En enchaînant les inconnues aux inconnues, la Science s'agrandit et s'élève. Si chaque effet n'avait qu'une seule cause, les énoncés au moins seraient faciles. La complication est plus grande. Dans le monde immense des faits, les parentés existent à tous les degrés. L'énumération des observations révèle les liens quand les nombres sont grands. La discussion est délicate, le bon sens la dirige, le calcul prononce.

L'inventeur d'un système associe, je suppose, la chute de la pluie à un phénomène astronomique ; il a observé vingt fois, sans une seule exception, qu'une pluie plus ou moins forte suivait le phénomène indiqué ; ce rapprochement est digne d'attention. Mais c'est à Brest qu'on a observé ; les jours sans pluie, à Brest, sont une rare exception. Que vaut alors la démonstration? Au Caire, elle serait décisive.

Il faut rapprocher, dans les cas semblables, le nombre des coïncidences observées de celui qui le remplacerait probablement, si tout était réglé par le hasard. Si deux phénomènes se présentent chacun 9 jours sur 10, les coïncidences, même très fréquentes, ne prouvent rien. Si chacun d'eux revient deux fois par an seulement, la coïncidence, plusieurs fois observée, sera difficilement attribuée au hasard. Difficilement : l'indication est vague ! elle doit l'être. Quand les géomètres, dans les cas semblables, ont donné un chiffre précis, ils ne tenaient aucun compte de la probabilité *a priori* du rapprochement qu'on a voulu faire, ou ils l'évaluaient, ce qui revient au même, tout à fait au hasard. Une comète a précédé la mort de César. Quelque nombreux et bien constatés que fussent les événements de ce genre, oserait-on croire, sur la foi du calcul, que *telles âmes sont tant nobles et héroïques que de leur délogement et trépas nous est certains jours devant donné signification des cieux?*

Un géomètre a trouvé une démonstration nouvelle du théorème de Bernoulli. J'en examine le principe, j'en parcours les calculs, j'en vérifie quelques-uns, et, n'apercevant aucune objection et aucune méprise, je déclare avec confiance l'exactitude de la méthode.

Le même auteur propose une démonstration du célèbre théorème énoncé par Fermat. J'examine le principe, je parcours les calculs, j'en vérifie quelques-uns, et, n'apercevant aucune objection et aucune méprise, je continue à chercher la faute. Pourquoi cette différence? Si les cas sont identiques, l'inégalité est-elle juste? Les cas sont différents. L'auteur qui démontre le théorème de Bernoulli enfonce une porte ouverte, il ne peut guère trébucher au passage. Celui qui démontre le théorème de Fermat suit un sentier sans issue connue; les chances d'une chute, d'après l'expérience du passé, y surpassent 100 contre 1 pour les plus habiles.

Toujours exact et précis dans l'énoncé des règles, Laplace n'a pas manqué d'introduire cette probabilité *a priori* comme point de départ et base nécessaire du calcul. Quelles que soient les conditions du problème, elle entre comme facteur, presque toujours inconnu, dans la formule qui la résout. L'illustre auteur de la *Théorie analytique des probabilités* a plus d'une fois cependant donné des chiffres précis qu'il faudrait changer avec l'hypothèse arbitrairement adoptée sur la probabilité *a priori*. Quand il assigne 1826214 à parier contre 1 comme mesure de la probabilité pour que le Soleil se lève demain, l'affirmation, quelles que soient les atténuations qui la suivent, repose sur une pure illusion.

Le rapport du nombre des décès à la population n'a pas été moins soigneusement étudié que celui des naissances. Les Compagnies d'assurances ont intérêt à le connaître et à en grossir l'évaluation. La Statistique le montre à peu près constant. Les variations, quoique petites, sont supérieures à celles du rapport des naissances des deux sexes. L'assimilation à des boules tirées d'une urne de composition invariable n'est donc pas acceptable. La vicissitude des événements règle sans cesse la composition de l'urne. Tantôt c'est le choléra qui passe et y verse des boules noires. Ce sont les eaux plus pures et plus fraîches qui apportent des boules blanches. C'est la disette qui rend les maladies plus abondantes et plus graves, la guerre qui accroît les mauvaises chances dans l'urne sans cesse renouvelée.

M. Dormoy, dans un livre savant et bien composé sur la théorie des assurances, a cherché curieusement dans les documents de la Statistique la confirmation de la loi des écarts. Il introduit, sous le nom de *coefficient de divergence*, le rapport de l'écart observé à l'écart moyen prévu par le calcul.

Un phénomène semble régulier, les chiffres qui le résument, sans être constants, varient peu d'une année à l'autre. On peut composer une urne qui, sous l'influence du hasard, représentera en moyenne, dans un grand nombre de tirages, par les boules noires amenées, la loi de l'arrivée de l'événement. On nomme *écart*, pour l'urne, la différence avec la moyenne annoncée par le calcul. L'écart, pour l'événement, est la différence entre le chiffre relatif à une année et la moyenne générale. Si le hasard règle le phénomène, le coefficient de divergence différera peu de l'unité. Un rapport plus grand révèle, s'il se maintient, l'influence d'une cause perturbatrice. Un coefficient de divergence plus petit que l'unité ferait deviner, au contraire, une action régulatrice qui, surveillant pour ainsi dire le hasard, amoindrit les inégalités et en efface le caractère. Tel est le cas d'un observateur trop avisé qui, dans les cas douteux, altère et corrige les observations pour en accroître la vraisemblance.

Pour les naissances des filles et des garçons, le coefficient de divergence à été 1,17 pour la France entière, de 1832 à

1841, et 1,38 de 1851 à 1864. Il confirme pour ces périodes la supposition d'une probabilité constante.

Le rapport du nombre des naissance naturelles au nombre total des naissances est moins régulier. Le coefficient de divergence, de 1817 à 1826, est égal à 15; pour le rapport du nombre des mariages à la population, le coefficient de divergence, de 1829 à 1848, s'est élevé à 25.

Le rapport du nombre des décès à la population a pour coefficient de divergence 86 ! Les anomalies sont continuelles. Le coefficient ne porte que sur des écarts, il faut le remarquer. Le nombre des décès pendant une année étant supposé pour la France entière égal à 1 million et au trente-sixième de la population, l'assimilation des Tables mortuaires annuelles aux tirages faits 36 millions de fois dans une urne contenant 1 boule noire et 35 boules blanches peut être tentée. Le nombre des boules noires, comme celui des décès, différera peu de 1 million ; mais, tandis que l'écart moyen, pour le nombre des boules noires, sera égal à 800, celui des décès sera 86 fois plus grand ; 86 fois 800 font 68 800, c'est moins de 2 pour 1000 de la population. Une épidémie produisant à Paris 4000 décès pour une année pourrait, pour le département de la Seine, expliquer le coefficient 86. Le choléra de 1849 a fait périr 20000 Parisiens.

Les lois du hasard sont invariables, ce sont les conditions du jeu qui changent. Poisson, pour les plier à tous les accidents, a cru compléter l'œuvre de Bernoulli en énonçant sa *loi des grands nombres*.

Pour que le hasard régularise l'arrivée d'un événement et que, sur un grand nombre d'épreuves, les rapports soient certains, aussi bien que la loi des écarts, il faut que la probabilité soit constante. Poisson supprime cette condition.

Un cas fictif très simple montrera la portée du nouveau principe. Une urne contient des boules numérotées, on y fait une série de tirages ; mais, en remettant chaque fois la boule qu'on a tirée, on néglige d'agiter et de faire le mélange : les chances, peu à peu, deviennent inégales ; certaines boules sortent plus souvent que les autres, la théorie semble mise en défaut. Continuez, dit Poisson ; pour prolongé que soit le désordre, il est embrassé lui-même dans la loi des grands nombres ; certaines boules sont dessus, vous les verrez dessous un autre jour ; l'homme peu soigneux à faire le mélange aura un successeur plus consciencieux ou dont la négligence, qu'il faut prévoir, profitera à des combinaisons nouvelles ; tout à la longue se compensera. Citons ses propres paroles : « Les choses de toute nature sont soumises à une loi universelle qu'on peut appeler la *loi des grands nombres*. Elle consiste en ce que, si l'on observe des nombres très considérables d'événements de même nature, dépendant de causes qui varient irrégulièrement, tantôt dans un sens, tantôt dans un autre, c'est-à-dire sans que leur variation soit progressive dans aucun sens déterminé, on trouvera entre ces nombres des rapports à très peu près constants ; pour chaque nature de choses, les rapports auront une valeur spéciale dont ils s'écarteront de moins en moins à mesure que la série des événements observés augmentera davantage et qu'ils atteindraient, s'il était possible de prolonger cette série à l'infini. »

Tel est le résumé fait par Poisson lui-même d'une découverte qui se distingue bien peu des lois du hasard, et à laquelle il a, à peu près seul je crois, attaché une grande importance.

III.

Aucune mesure n'est certaine, mille opérations successives donnent mille résultats différents. Non que l'observateur, de mieux en mieux instruit, corrige ses défauts et s'avance vers la perfection. Il n'en va pas ainsi. Les derniers résultats ne ressemblent en rien à une limite dont on s'approcherait par continuel progrès ; les évaluations, tantôt trop petites, tantôt trop grandes, se succèdent en confusion et sans ordre comme des boules blanches ou noires puisées dans une urne.

Bessel, après un siècle écoulé, comparait les observations de Bradley aux résultats connus d'une théorie devenue certaine. En classant les différences, dont le désordre est complet, il trouva, sur 470 observations, 94 erreurs inférieures à $\frac{1}{10}$ de seconde, 88 comprises entre $\frac{1}{10}$ et $\frac{2}{10}$, puis, successivement, entre $\frac{2}{10}$ et $\frac{3}{10}$, entre $\frac{3}{10}$ et $\frac{4}{10}$, ..., jusqu'à 1″, la plus grande des erreurs commises par Bradley, les nombres décroissants 78, 58, 51, 36, 26, 14, 10, 7 et 8. Si les plus petits sont les plus nombreux, l'honneur n'en revient ni à ce grand observateur Bradley, ni aux constructeurs des instruments de Greenwich ; leur excellence fait la petitesse, non la loi des erreurs ; un instrument médiocre, un observateur moins soigneux, remplaceraient les dixièmes de seconde par des secondes, les secondes peut-être par des minutes ; à cela près tout resterait pareil. La courbe des erreurs en s'élevant conserverait la même forme.

L'origine des erreurs est très diverse. Les unes sont fortuites, l'enchaînement en est infini ; c'est tantôt l'air agité par le vent, tantôt un ébranlement du sol, un nuage qui passe, un rayon de soleil qui trouble l'observateur, tantôt une attention précipitée ou distraite ; le hasard décide, mille causes imprévues se réunissent, ajoutent quelquefois leurs effets, quelquefois les retranchent, suspendent ou reprennent leur action : tout est incertain, tout change, sans inclination dans aucun sens.

Il n'en va pas ainsi des causes permanentes ; c'est une balance mal construite, les fils d'une lunette mal placés, un mètre trop court, un chronomètre trop rapide. Les mesures prises sous de telles influences n'entourent plus la valeur exacte, mais une autre, souvent fort différente ; une nouvelle série de mesures, sous l'influence permanente des mêmes causes, se groupera autour de la même moyenne.

Tout observateur soigneux étudie les erreurs constantes et les corrige sans retrancher la cause ; rien ne trompe moins qu'une balance trompeuse. Qu'importe que les bras soient inégaux, pourvu qu'on le sache ? Qu'un gramme ait 999 milligrammes, un décimètre 99 millimètres, l'observation réduite conserve toute sa valeur. Toute mesure est comparable à un jeu ; les erreurs possibles en plus ou en moins sont les chances de gain ou de perte ; les erreurs constantes changent les règles du jeu, les erreurs fortuites laissent le jeu équitable.

La loi que doivent suivre, d'après une ingénieuse théorie, et que suivent à très peu près, quand elles sont nombreuses, les erreurs corrigées de toute inclination fixe, a été proposée par Gauss. L'histoire en est singulière. En proposant en 1809 une hypothèse sur la théorie des erreurs, l'illustre auteur ne prétendait nullement établir la vérité, mais la chercher. Laplace, par une voie différente, sans beaucoup de rigueur à son tour, avait obtenu la même formule qui, très voisine souvent de la vérité, pourrait s'en éloigner sans démentir la Science.

Le principe de Gauss est fort simple : quand une grandeur a été mesurée plusieurs fois, les erreurs constantes étant écartées (la précaution est nécessaire), entre plusieurs résultats également dignes de confiance, la moyenne est, en l'absence de tout autre renseignement, la valeur la plus probable. Les conséquences de cet axiome sont belles et imprévues, mais incertaines; Gauss en convient volontiers. Le rapprochement des observations peut affaiblir la confiance en quelques-unes d'elles. Si quatre pesées successives ont donné 20^{mg}, puis 27^{mg}, 26^{mg} et 28^{mg}, on se décidera sans doute, quelles que soient les circonstances, à écarter la première mesure pour adopter la moyenne des suivantes. Quoi qu'il en soit, Gauss, sur ce fondement, établit ingénieusement une formule que l'expérience confirme. Le hasard, quand les épreuves sont nombreuses, amenant chaque événement en raison de sa probabilité, il suffit, pour juger la formule, de faire mesurer un grand nombre de fois une grandeur que l'on connaît très exactement à l'avance.

La probabilité des erreurs suit, d'après la formule, précisément la loi des écarts dans les épreuves répétées. La rencontre n'est pas fortuite, Laplace l'a expliquée. Les erreurs constantes étant écartées, les accidents fortuits troublent seuls chaque épreuve, ils sont analogues aux tirages faits dans une urne. Laplace développe ce rapprochement, le rend précis, transforme le problème, et retrouve la formule de Gauss.

Cette admirable et très simple formule s'étend à toutes les grandeurs, s'applique à tous les instruments, régit toutes les observations et embrasse tous les procédés de mesure; les différences, d'un cas à l'autre, si grandes qu'elles puissent être, se résument dans un nombre caractéristique représentant la *précision*, l'*erreur probable*, le *poids de l'observa-*

tion; peu importe le nom, un seul nombre connu permet de calculer toutes les chances et de prédire, sur un grand nombre d'épreuves, la distribution certaine des écarts.

Si l'on caractérise une série de mesures par l'*erreur probable* qu'il y a chance d'atteindre ou de ne pas atteindre, en prenant cette erreur pour unité, la probabilité d'une erreur double diffère peu de $\frac{1}{10}$, celle d'une erreur quintuple s'abaisse à $\frac{1}{1000}$; pour une erreur dix fois plus grande que l'erreur probable, le nombre donné par la formule vaut une déclaration d'impossibilité.

L'instrument, il ne faut pas l'oublier, est, aussi bien que l'observateur, supposé sans défaut; on n'accepte en lui que des défaillances, des accidents fortuits qu'aucune cause constante n'incline dans aucun sens.

Les épreuves du tir, soit au canon, soit à la carabine, mettent en évidence les effets du hasard; les erreurs fortuites ont pour origine, outre le coup d'œil plus ou moins juste et les distractions du pointeur, le poids variable du projectile, les inégalités de sa structure, le tassement irrégulier de la poudre, les courants, les vibrations, l'humidité des couches d'air traversées; c'est pour cela que, sans changer en rien les conditions du tir, on voit les coups s'écarter les uns des autres, en se groupant autour d'un point central, autour du but lui-même, si les erreurs constantes sont écartées.

Un savant professeur, M. Jauffret, a défini, par une image fort nette, les lois de distribution des coups, identiques, d'après le théorème de Bernoulli, à celles des probabilités. Si, visant pendant un long temps un même but placé sur le sol, on arrête chaque boulet au point même de sa chute, l'amas des projectiles présentera l'aspect d'une cloche dont la base circulaire aurait le but pour centre; un tireur plus adroit rétrécirait la cloche et la rendrait plus haute ; une moindre précision donnerait naissance à un solide moins élevé, s'abaissant plus lentement vers le sol.

N'est-il pas merveilleux et presque incroyable qu'on puisse, par le raisonnement seul, prédire ainsi la disposition des boulets sans connaître l'adresse du pointeur ni demander la précision de l'arme ?

Les formules, a dit Poinsot, ne donnent que ce qu'on y a mis. Aucun raisonnement ne fait davantage ; le dernier anneau d'une chaîne de déductions est, pour qui sait l'y voir, tout entier dans les hypothèses. Nous avons expressément supposé, il ne faut pas l'oublier, qu'il n'existe dans l'arme ni dans la maladresse du pointeur aucune cause d'erreur constante ; il n'y a donc pas plus de chance, c'est l'hypothèse même, de tirer à droite plutôt qu'à gauche, trop près plutôt que trop loin. Faut-il s'étonner que le but se trouve au centre des divers points atteints dans une longue série d'épreuves ? Si plus de la moitié se trouvait à droite, on en conclurait qu'une cause les y porte, et ce serait une erreur constante.

Un doute peut s'élever encore. Les erreurs constantes sont celles que l'on peut corriger, la maladresse est une cause fortuite, un tireur maladroit atteint bien rarement le but ; au lieu de le cacher sous le sommet d'un dôme de projectiles, ne le laisserait-il pas au centre d'un grand vide ? Diogène pensait ainsi : « Un jour, voulant s'esbattre, il visita les archers qui tiroient à la butte ; entre iceux, un étoit tant fautier, impérit et maladroit, que lorsqu'il estoit en ranc de tirer, tout le peuple spectateur s'escartoit de peur d'être par lui féru. Diogènes l'avoit un coup vu si perversement tirer, que la flesche tomba plus d'un trabut loin de la butte ; au second coup, le peuple, loin de côté et d'autre, s'escartant, il accourut et se tint en pied, jouxte le blanc, affirmant cetuy lieu être le plus sûr et que l'archer fériroit tout autre lieu, le blanc seul être en seureté de traict. » La plaisanterie fit rire. Il n'aurait pas fallu recommencer souvent; les gouttes d'eau, guidées par le hasard, n'épargnent à la longue aucun pavé. Pourquoi les projectiles, non moins nombreux, c'est l'hypothèse, éviteraient-ils le point vers lequel, adroitement ou non, on s'étudie à les diriger tous?

Dans la formule de probabilité des erreurs, la rigueur, nous l'avons avoué, n'a pas été mise; l'axiome supposé est loin d'être évident; les conséquences sont comme lui discutables. Dans les concours de tir à la carabine, chaque tireur ayant droit à un certain nombre de balles, on décide du mérite de chacun par la distance moyenne de ses balles au but. La théorie consultée prescrirait une autre règle : la plus petite moyenne du carré des écarts caractérise mieux le plus adroit. La décision, je crois, a été prise pour l'armée belge. Le changement est de petite conséquence, et sur un grand nombre d'épreuves toutes les méthodes s'accorderaient; toutes deux, la seconde surtout, traitent trop sévèrement le tireur, si adroit qu'il se soit montré, dont un coup s'est égaré des autres. Supposons, pour donner des chiffres simples, qu'un tireur ayant placé neuf balles à la distance moyenne 1 du but, la dixième s'en écarte à la distance 10. D'après la première règle, la moyenne générale étant 1,9, il sera préféré à celui dont toutes les balles seraient à la distance 2; cela paraît juste. La seconde règle, celle qui s'appuie sur la loi de probabilité des écarts, placerait avant lui le tireur dont toutes les balles seraient à la distance 3. Peut-être vaudrait-il mieux, sans tant raffiner, s'en tenir à la vieille méthode, qui réserve le prix à qui le plus souvent touche la mouche, sans rechercher l'écart des balles moins heureuses.

La formule de Gauss déclare, pour ainsi parler, certains cas impossibles. N'invite-t-elle pas par là, quand ils se présentent, à se défier un peu d'elle? Les cas exceptionnels échappent à toute règle. Le bon sens ne perd jamais ses droits : opposer à l'évidence une formule *démontrée*, c'est à peu près comme si, pour refuser à un homme le droit de vivre, on alléguait devant lui un acte de décès authentique.

La moyenne d'un grand nombre de mesures, quand on écarte les erreurs constantes, est une mesure plus précise que celles qui l'ont fournie ; l'erreur probable est diminuée, et la précision augmente comme la racine carrée du nombre des épreuves.

Fourier connaissait ou soupçonnait cette règle : pour prendre la hauteur de la pyramide de Chéops, il fit simplement mesurer par des soldats les 203 marches de ce gigantesque escalier. « Vos hommes manquent d'habitude, disait-on : les surfaces sont irrégulières, les arêtes sont inclinées ; aucune précision n'est possible, et l'erreur commise sur chaque marche sera multipliée par 203. — Elle le sera par 14 seulement, répondit-il résolument, car 14 est la racine carrée de 203. » La comparaison avec une mesure plus exacte aurait pu le contredire ; on ne la fit pas.

Entre les grandeurs inconnues enchaînées par les formules, la Science, dans chaque problème, choisit pour la déterminer directement la plus accessible aux mesures. Pour peser l'obélisque il n'existe pas de balance ; une chaîne d'arpenteur donnerait très lentement et très mal la distance de Paris à Rome. La théorie fournit des équations, on les accepte toutes, chacune est irréprochable ; l'Algèbre dégage les inconnues ; les chiffres malheureusement se contredisent toujours. Que doit-on faire? Entre des mesures discordantes on prend la moyenne : pour des équations, ce mot n'a pas de sens; à chacune, cependant, il faut un rôle; la méthode des moindres carrés enseigne et prescrit la meilleure combinaison.

Cette méthode, inventée par Gauss, proposée pour la première fois par Legendre, a procuré plus d'une déception.

La masse de Jupiter, déduite par Newton de l'étude des satellites, corrigée peu à peu par les progrès des observateurs, calculée de nouveau par Bouvard à l'aide des perturbations de Saturne, semblait fixée à $\frac{1}{1070}$ de celle du Soleil. Les principes du calcul des chances permettaient de parier, suivant Laplace, 999 308 contre 1 que l'erreur n'est pas la centième partie de la valeur trouvée. Quelle ostentation de consciencieux savoir! C'est 999 308$^{\text{fr}}$, ni plus ni moins, que l'on peut risquer contre 1$^{\text{fr}}$. On aurait eu tort de risquer dix sous; on les aurait perdus; les perturbations de Junon l'ont prouvé. Sans contester ce témoignage irréprochable de la petite planète, Poisson maintenait les principes. « Les calculs de Laplace, dit-il, ont donné, avec une précision voisine de la certitude, une masse plus petite qu'elle n'est réellement. Cela ne provient d'aucune inexactitude dans les formules dont il a fait usage; il y a lieu de croire que la masse de Jupiter, un peu trop petite, résulte de quelques termes fautifs dans l'expression des perturbations. » Poinsot, son spirituel adversaire, pour transformer l'apologie en épigramme, ne change rien au trait que l'accent : « Après avoir calculé la probabilité d'une erreur, il faudrait calculer la probabilité d'une erreur dans le calcul. »

Peut-on, par des combinaisons habiles, s'assurer sur les résultats d'observations imparfaites, puisées à des sources douteuses? On le peut, répond la théorie, *pourvu qu'on n'ait pas à craindre d'erreurs constantes*. Le calcul échouera, répond le bon sens. Les deux réponses sont d'accord.

Lorsqu'en 1761, après soixante-dix années d'attente, les astronomes de tous les pays distribuèrent sur la portion du globe désignée par Halley plus de cent observateurs du passage de Vénus, — la crainte du mauvais temps et l'émulation du zèle pour la Science en accrurent ainsi le nombre, — on croyait la méthode infaillible, et deux observateurs soigneux, Halley l'avait prouvé, pouvaient sans aucun associé donner la parallaxe exacte au centième de seconde. 60 observations, au lieu de 2, faisaient espérer par leurs combinaisons 1770 déterminations identiques. La déception fut grande; les résultats variaient entre 7″ et 11″. En combinant 15 observations européennes avec celle du Cap de Bonne-Espérance, Short trouva une moyenne de 8″,47. L'observation de Tobolsk, combinée avec 15 autres, donnait 9″,56; en en supprimant 4, il restait 8″,69. Ces 4 observations, 2 de Stockholm et 2 de Tornéa, comparées à celle de Tobolsk, auraient donné plus de 11″. L'opération était à refaire. Rien ne fut épargné en 1769, le succès fut pareil. En combinant les observations sans règle et sans méthode, les calculateurs du XVIIIe siècle n'en purent montrer que l'incertitude. Encke, en 1822, voulut reprendre dans leur ensemble les résultats des deux expéditions, et, par un prodigieux travail, appliquant dans toutes ses prescriptions la méthode des moindres carrés, il obtint 8″,5776. L'erreur probable était 0″,0370.

Cette expression d'*erreur probable* exige une explication : l'erreur probable est celle qu'il y a chance égale d'atteindre ou de ne pas atteindre; de celle-là, nous l'avons dit, on déduit toutes les autres. Contre une erreur huit fois plus grande il n'y a pas, dit la théorie, 1 chance sur 1 million. C'est justement celle-là qui s'est produite. La parallaxe, aujourd'hui bien connue, surpasse le résultat d'Encke de huit fois son erreur probable. Tous ces calculs devaient être stériles, rien ne garantissait contre les causes constantes, et le nombre des observations douteuses n'était pas assez grand pour assurer une compensation.

IV.

Tout semblait débattu sur les universaux, et tout oublié. M. Quetelet, sans réveiller ce vieux problème, a cru sérieusement le résoudre, et, dans un livre riche de faits judicieusement recueillis, a voulu définir et préciser le mot *homme* indépendamment des hommes particuliers considérés comme accidents. Sans discussion ni subtilités, le patient auteur attribue à son type, par définition, la moyenne de chaque élément variable d'un homme à l'autre. En relevant, par exemple, les tailles de 20 000 soldats, on a trouvé pour moyenne $1^m,75$; telle est la taille de l'homme moyen; autour d'elle, dans la série des mesures, se groupent les tailles plus grandes ou plus petites, exactement graduées suivant la loi des écarts. Rien ne distingue les tailles des conscrits des mesures qu'un observateur très maladroit aurait prises 20 000 fois de suite sur un même homme de $1^m,75$, avec des instruments bien grossiers, il faut le supposer, mais corrigés de toute erreur constante.

M. Quetelet, dans ce rapprochement, voit une identité; nos tailles inégales sont pour lui le résultat des mesures très mal prises par la nature sur un modèle immuable, qui, seul, révèle tout son savoir. $1^m,75$ est la taille normale; pour avoir un peu plus, on n'en est pas moins homme, mais ce qui manque ou dépasse pour chacun est erreur de nature et monstruosité.

Abailard, si habile à raisonner des choses, aurait réduit l'argument en forme, mais on ne remue plus de telles subtilités. M. Quetelet, sur ce vieux champ de bataille des écoles, n'a rencontré ni défenseurs ni adversaires.

La thèse a cependant plus d'un inconvénient. L'homme idéal, dit-on, représente en toute chose la moyenne de l'humanité. Cela paraît très simple et très clair, mais ces détails, définis par règle et par compas, comment s'ajustent-ils? La hauteur de la tête, par exemple, pourra, pour l'homme moyen, se calculer par deux méthodes : on peut prendre la moyenne des longueurs, ou, pour chaque individu, le rapport de la tête à la hauteur du corps, puis la moyenne de ces rapports. Les résultats sont différents : comment les accorder?

Grave difficulté et inévitable écueil! Pour le montrer avec évidence, cherchons entre deux sphères la sphère moyenne; l'une a pour rayon 1; nous choisirons les unités de manière à représenter également la surface et le volume par 1. La seconde sphère a, je suppose, pour rayon 3, pour surface 9 et pour volume 27; ces chiffres sont forcés. Les moyennes 2, 5 et 14 sont incompatibles; une sphère de rayon 2 aurait pour surface 4 et pour volume 8 très exactement; aucune concession n'est possible, nulle sphère n'est difforme. Un homme malheureusement peut l'être, et M. Quetelet en profite; en associant le poids moyen de 20 000 conscrits à leur hauteur moyenne, on fera l'homme type ridiculement gros et, quoi qu'en ait pensé Reynolds, un mauvais modèle pour un peintre. Cet artiste éminent, dans ses leçons sur les Beaux-Arts, avait, avant Quetelet, signalé dans l'homme moyen le type de la beauté parfaite. Si tel était le cas, a dit Sir John Herschel, la laideur serait l'exception. Je n'en aperçois pas la raison. Aucun trait de la beauté parfaite ne serait rare; distribués sans convenance, ils seraient sans mérite. L'harmonie fait la grâce. Le hasard appellerait sans doute peu d'élus, et, n'en déplaise à Sir John Herschel, dans les assemblages incohérents, si la laideur formait l'exception, le grotesque deviendrait la règle.

Dans le corps de l'homme moyen, l'auteur belge place une âme moyenne. Il faut, pour résumer les qualités morales, fondre vingt mille caractères en un seul. L'homme type sera donc sans passions et sans vices, ni fou ni sage, ni ignorant ni savant, souvent assoupi : c'est la moyenne entre la veille et le sommeil; ne répondant ni oui ni non; médiocre en tout. Après avoir mangé pendant trente-huit ans la ration moyenne d'un soldat bien portant, il mourrait, non de vieillesse, mais d'une maladie moyenne que la Statistique révélerait pour lui.

V.

L'application du calcul aux décisions judiciaires est, dit Stuart Mill, le scandale des Mathématiques. L'accusation est injuste. On peut peser du cuivre et le donner pour or, la balance reste sans reproche. Dans leurs travaux sur la théorie des jugements, Condorcet, Laplace et Poisson n'ont pesé que du cuivre.

La réunion, quelle qu'elle soit, qui peut juger bien ou mal, est remplacée dans leurs études par des urnes où l'on puise des boules blanches ou noires. « On peut, dans plusieurs cas, — a dit Laplace, le plus grand des trois, le moins imprudent, et incomparable aux deux autres, — résoudre des questions qui ont beaucoup d'analogie avec les questions qu'on se propose, et dont les solutions peuvent être regardées comme des approximations propres à nous guider et à nous garantir des erreurs et des dangers auxquels les mauvais raisonnements nous exposent. Une approximation bien conduite est toujours préférable aux raisonnements les plus spécieux. »

Rien n'est plus sage : les bonnes approximations valent mieux que les mauvais raisonnements; mais il n'y a, malgré cela, moyen ni apparence de les réduire en acte pour rendre la justice meilleure que les juges. On peut assurément supposer le nombre des boules noires égal à celui des jugements mal rendus, les deux problèmes n'en restent pas moins fort différents et, pour tout dire, sans analogie.

Un juge, supposons-le, se trompe une fois sur dix. Condorcet et Poisson l'assimilent à une urne contenant 9 boules blanches et 1 noire. Le sort des accusés resterait-il le même?

Sur 1000 épreuves, la boule noire sortira 100 fois, tout comme, sur 1000 jugements, 100 seront mal rendus. Les nombres se ressemblent, tout le reste diffère. Quand un juge se trompe, c'est que le cas sans doute est complexe et ardu. On condamne à coup sûr le coupable qui avoue, on acquitte en hésitant celui que l'on n'a pu convaincre; les 100 boules noires de l'urne se montreront le même nombre de fois, mais tout autrement. Condorcet répondrait peut-être que pour la société, qui seule l'intéresse, le dommage et l'alarme resteraient les mêmes et qu'ils dépendent du nombre des crimes impunis et des innocents déclarés coupables. Mais une autre objection est sans réplique : l'indépendance des tirages est supposée; les urnes, dans les calculs, échappent à toute influence commune. Les juges, au contraire, s'éclairent les uns les autres, les mêmes faits les instruisent, les mêmes témoignages les troublent, les mêmes sollicitations les tourmentent, la même éloquence les égare, c'est sur les mêmes considérants qu'ils font reposer la vérité ou l'erreur. L'assimilation est impossible.

« Condorcet a pris possession de l'univers moral pour le soumettre au calcul. » C'est la louange qu'on lui a donnée; on s'est demandé si c'est après l'avoir lu. Dans son Livre sur *La probabilité des jugements*, il se propose d'abord deux problèmes. Premièrement : Quel est, pour chaque jugement et pour chaque juge, la probabilité de rencontrer juste? En second lieu : Quelle est la probabilité d'erreur à laquelle la société peut se résigner sans alarmes?

La première question lui semble facile.

« Je suppose, dit Condorcet, que l'on ait choisi un nombre d'hommes véritablement éclairés et qu'ils prononcent sur la vérité ou sur la fausseté de la décision. Si, parmi les décisions de ce tribunal d'examen, on n'a égard qu'à celles qui ont obtenu une certaine pluralité, il est aisé de voir qu'on peut, sans erreur sensible, les regarder comme certaines. »

C'est un concile infaillible, tout simplement, qu'il définit et prétend convoquer. Sans douter, il hésite; non que les hommes véritablement éclairés soient rares, gardons-nous de le croire, mais leur temps est précieux; pour l'épargner, Condorcet propose une seconde méthode dont Poisson, plus tard, n'a pas aperçu l'illusion. La probabilité d'erreur étant supposée pour un juré, on peut, en augmentant leur nombre, la diminuer sans limite pour l'ensemble. L'instrument est trouvé, on n'a plus qu'à choisir. « Que l'on compte, dit Condorcet, combien il périt de paquebots sur le nombre de ceux qui vont de Calais à Douvres, et qu'on n'ait égard qu'à ceux qui sont partis par un temps regardé comme bon par les hommes instruits dans la navigation. Il est clair qu'on aura, par ce moyen, la valeur d'un risque que, pour les autres comme pour soi, on peut négliger sans imprudence. »

Préfère-t-on le danger de périr au Pont-Saint-Esprit, quand on descend le Rhône de Lyon à Avignon? Les honnêtes gens s'y exposent sans frayeur. Veut-on, pour le faire court, la probabilité $\frac{1}{144768}$? Il ne faut que dire oui. Je n'invente ni n'exagère. Dans une assemblée de 65 votants, on exigera la majorité de 9 voix. Deux conditions seulement sont supposées : chaque juge, isolément, ne doit se tromper qu'une fois sur cinq. En jugeant la même cause, le raisonnement proposé le suppose, ils ne doivent pas non plus être exposés aux mêmes chances d'erreurs.

Lorsque, 8 ans plus tard, Condorcet préférait le poison à une justice suspecte, s'il eût pu s'assurer en des juges courageux et honnêtes, il n'en aurait pas exigé 65.

Laplace aborde très modestement le problème des jugements : « La probabilité des décisions d'une assemblée dépend, dit-il, de la pluralité des voix, des lumières et de l'impartialité des juges. Tant de passions et d'intérêts particuliers mêlent si souvent leur influence qu'il est impossible de soumettre le résultat au Calcul des probabilités. » Il l'y soumet pourtant, et Poisson, en fondant, dans son Livre, sur des principes certains, des applications à peine douteuses, a cru suivre son illustre exemple. Laplace cherche d'abord, pour les assemblées, le meilleur système de vote. Il est rare que l'on puisse, en répondant oui ou non, exprimer toute son opinion. Plusieurs propositions, presque toujours, sont relatives aux mêmes objets. Le calcul, suivant Laplace, ne conseille pas de les mettre aux voix successivement. Voici ce qu'il faut faire : chaque votant recevra un nombre illimité de boules, et l'on passera, pour recueillir les votes, autant d'urnes qu'il y a d'opinions en présence, en invitant chaque votant à verser dans chaque urne un nombre de boules proportionnel à la probabilité qu'il attribue à la proposition correspondante. Docile à la théorie du probabilisme, chacun résistera à la tentation de verser sa provision tout entière dans l'urne favorable à l'opinion qui lui agrée le plus.

Les assemblées n'ont pas tenté l'épreuve; elles cherchent le sûr, comme Pascal, le probable ne leur suffit pas.

Laplace, reprenant une idée de Condorcet, cherche dans le compte des votes concordants ou discordants des divers juges la chance qu'ils ont de prononcer juste. Se séparant pourtant de Condorcet sur un point de grande importance, il fait varier cette probabilité d'une cause à l'autre, mais la fait, dans chaque cause, égale pour tous les juges; la seule donnée introduite est le nombre des juges favorables à chaque opinion. Si un jury de douze nègres prononce sur le vol d'une banane, la probabilité de bien juger sera, d'après la formule, précisément la même, à majorité égale, que pour douze conseillers à la Cour de cassation décidant une question de droit.

La probabilité, dans les calculs de Poisson, reste la même pour toutes les causes; il n'ignore pas qu'elle peut varier, mais il croit obtenir, sans doute, une de ces approximations bien conduites dont parle Laplace.

Une urne contient des boules noires ou blanches; la proportion est inconnue; il suffira, pour la découvrir, de faire un grand nombre de tirages. Le rapport du nombre des boules blanches sorties au nombre total des tirages fera connaître leur proportion dans l'urne. La vérité, malheureusement, aussi différente de l'erreur que la couleur blanche l'est de la noire, ne s'en distingue pas si facilement.

Supposons, en second lieu, deux urnes en présence. On ignore la proportion des boules noires ou blanches, et, à chaque tirage, on fait connaître, non la couleur des boules, mais leur accord seulement ou leur désaccord. On ne pourra par de telles épreuves, si souvent qu'elles soient répétées, déterminer la composition des urnes, mais seulement renfermer le doute dans des limites plus ou moins étroites.

En consultant trois urnes au lieu de deux, le problème se résout exactement. Si, tirant une boule de chacune, on sait quelles urnes s'accordent à donner même couleur, l'épreuve, suffisamment répétée, fera connaître, avec telle probabilité qu'on voudra, la composition des trois urnes, sans distinguer toutefois les cas où les noires seraient changées en blanches, et réciproquement.

Poisson et Cournot substituent aux trois urnes les trois juges d'un même tribunal. Si Pierre, Paul et Jacques prononcent sur un grand nombre d'affaires, on pourra, sans savoir si leurs décisions sont justes ou injustes, connaître leurs différences d'opinion. La formule qui révèle les boules blanches des urnes s'appliquera aux chances de bien juger, en repoussant toutefois, pour chaque magistrat, la probabilité de se tromper plus d'une fois sur deux. Mieux vaudrait sans cela, après avoir vu, lu, relu, paperassé et feuilleté les pièces du procès, jouer, comme faisait Bridoye, la sentence à trois dés.

Les deux problèmes assimilés par Poisson sont, en réalité, très différents. Si Pierre et Paul s'accordent souvent contre Jacques, il peut se faire qu'ils aient, sur certains cas douteux, une opinion pareille et que, en la repoussant, Jacques comprenne mieux la loi. Peut-être Pierre et Paul montrent-ils pour certains plaideurs une même indulgence, pour d'autres une égale rigueur. Pour être plus éclairé, plus droit, plus impartial, Jacques alors serait diffamé par la formule. Si Paul, quand un de ses collègues a opiné le premier, n'a pas la hardiesse de le contredire, la formule y verra une preuve de son mérite. Est-elle digne de confiance? Sans s'arrêter à des difficultés aussi visibles, Poisson n'a pas craint d'assigner, pour un juré pris au hasard, la probabilité de décider juste. D'après l'ensemble des documents interprétés par ses calculs, chaque juré, en France, se trompe une fois sur trois. C'est beaucoup : Condorcet n'en demanderait pas davantage. Quelques centaines de ces jurés sans lumières lui suffiraient pour promettre, au nom de la Science, aux accusés innocents toute la sécurité d'un joyeux touriste qui, par un temps serein, s'embarque sur une mer sans écueils.

VI.

L'action libre des êtres humains, celle aussi des animaux, quoi qu'en ait dit Descartes, mêle à l'enchaînement des effets et des causes un élément inaccessible au calcul. La liberté du choix produit, à parler rigoureusement, les seuls cas fortuits.

Les lois du hasard étendent plus loin leur domaine. Un homme agite un cornet, lance les dés, doucement ou avec force, à droite ou à gauche, use sans contrainte de son libre arbitre; il amène *sonnez* 1 fois sur 36.

On substitue au bras de chair des organes de cuivre et d'acier. Une machine jette les dés, les ramasse, les lance encore, mue par la force aveugle d'un ressort entretenue par d'autres ressorts. Tout est déterminé; un géomètre calcule à l'avance la succession des points. La formule donne *sonnez* 1 fois sur 36.

Tous les soldats d'une nombreuse armée sont appelés tour à tour à dire un nombre moindre que 7, le premier venu. Dans leurs réponses, inscrites deux par deux, on rencontre deux 6 1 fois sur 36.

D'où vient cela? Les lois du hasard gênent-elles la liberté des efforts musculaires? règlent-elles l'ordonnance d'un mécanisme aveugle? troublent-elles le caprice de 100000 imaginations qui les ignorent? Il n'en est pas ainsi. Si l'on influence la volonté de ces hommes; si le mécanicien, rebelle à la loi de Bernoulli, prend plaisir à la mettre en défaut; si le joueur de dés s'y applique avec ou sans adresse, toutes nos assertions seront fausses. A tout effort le hasard est docile; sans souci de la règle, il suit les gros bataillons.

Le hasard est sans vertu : impuissant dans les grandes affaires, il ne trouble que les petites. Mais, pour conduire les faits de nature à une fin assurée et précise, il est, au milieu des agitations et des variétés infinies, le meilleur et le plus simple des mécanismes. Les vapeurs s'élèvent, les vésicules se forment, les nuées s'épaississent, les vents les dispersent, les mêlent, les entre-choquent, engendrent la tempête et la pluie; le hasard conduit tout sans surveillance ni délibération aucune, et, précisément parce qu'il est aveugle, il remplit le lit de tous les fleuves, arrose toutes les campagnes et donne à chaque brin d'herbe sa ration nécessaire de gouttes d'eau.

TABLE DES MATIÈRES.

CHAPITRE III.

ESPÉRANCE MATHÉMATIQUE.

CHAPITRE IV.

THÉORÈME DE JACQUES BERNOULLI.

CHAPITRE V.

DÉMONSTRATIONS ÉLÉMENTAIRES DU THÉORÈME DE BERNOULLI.

CHAPITRE VI.

LA RUINE DES JOUEURS.

CHAPITRE VII.

PROBABILITÉ DES CAUSES.

CHAPITRE VIII.

LOI DES ERREURS D'OBSERVATION.

CHAPITRE IX.

ERREURS DE SITUATION D'UN POINT.

CHAPITRE X.

LA THÉORIE DES MOYENNES.

CHAPITRE XI.

COMBINAISON DES OBSERVATIONS.

CHAPITRE XII.

LES LOIS DE LA STATISTIQUE.

CHAPITRE XIII.

PROBABILITÉ DES DÉCISIONS.

FIN DE LA TABLE DES MATIÈRES.

CALCUL

DES

PROBABILITÉS.

CHAPITRE I.

ÉNUMÉRATION DES CHANCES.

> On estime la probabilité d'un événement par le nombre des cas favorables divisé par le nombre des cas possibles. La difficulté ne consiste que dans l'énumération des cas.
>
> LAGRANGE.

1. Définition de la probabilité. L'égalité des chances est supposée dans la définition. — 2. Exemple d'une énumération incorrecte. — 3. Autre exemple. — 4. Le nombre des cas ne doit pas être infini. Contradiction résultant de l'oubli de cette condition. — 5. Second exemple. — 6 Troisième exemple. — 7. Quatrième exemple. — 8, 9, 10, 11, 12, 13. Solution de quelques problèmes par l'énumération des chances. — 14. Prétendu paradoxe du chevalier de Méré. — 15. Combien faut-il tenter de coups pour obtenir une probabilité donnée de produire au moins une fois un événement dont la probabilité est connue? — 16. Problème du jeu de rencontre. — 17. Problemes relatifs aux tirages de boules numérotées sans les remettre après chaque tirage. — 18. Problème relatif au dépouillement d'un scrutin de ballottage. — 19. Une urne contient des boules numérotées, quelle est la probabilité pour que sur n tirages la somme des points tirés ait une valeur donnée. — 20. Application au cas de trois dés.

1. La probabilité d'un événement est estimée par l'énumération des cas favorables, rapprochée de celle des cas possibles.

On parie, en jetant un dé, qu'il montrera le point 4. Le dé a six faces : six cas sont possibles, un seul est favorable. La probabilité est $\frac{1}{6}$. C'est une définition.

On jette deux dés; les six points du premier, en s'associant aux six du second, peuvent former trente-six combinaisons : la probabilité d'amener une d'entre elles, double deux par exemple, est $\frac{1}{36}$.

La probabilité d'amener 3 et 4 est $\frac{2}{36}$: chacun des dés pouvant donner 3 lorsque l'autre donne 4, il y a deux combinaisons favorables 3 et 4, 4 et 3; on leur donne le même nom, 3 et 4, mais elles sont réellement distinctes.

La probabilité d'un événement est le rapport du nombre des cas favorables au nombre total des cas possibles. Une condition est sous-entendue : tous les cas doivent être également possibles. La définition, sans cette restriction, n'aurait aucun sens. Il peut se faire que l'événement arrive, il se peut aussi qu'il n'arrive pas; ce sont deux cas possibles, un seul est favorable. Toute probabilité serait donc $\frac{1}{2}$. L'erreur est grossière. D'Alembert a élevé l'objection et refusé de passer outre.

Avant de compter les chances, il faut constater qu'elles ont même vraisemblance.

2. Trois coffrets sont d'apparence identique. Chacun a deux tiroirs, chaque tiroir renferme une médaille. Les médailles du premier coffret sont en or; celles du deuxième coffret en argent; le troisième coffret contient une médaille d'or et une médaille d'argent.

On choisit un coffret; quelle est la probabilité pour trouver, dans ses tiroirs, une pièce d'or et une pièce d'argent?

Trois cas sont possibles et le sont également puisque les trois coffrets sont d'apparence identique.

Un cas seulement est favorable. La probabilité est $\frac{1}{3}$.

Le coffret est choisi. On ouvre un tiroir. Quelle que soit la médaille qu'on y trouve, deux cas seulement restent possibles. Le tiroir qui reste fermé pourra contenir une médaille dont le métal diffère ou non de celui de la première. Sur ces deux cas, un seul est favorable au coffret dont les pièces sont différentes. La probabilité d'avoir mis la main sur ce coffret est donc $\frac{1}{2}$.

Comment croire, cependant, qu'il suffira d'ouvrir un tiroir pour changer la probabilité et de $\frac{1}{3}$ l'élever à $\frac{1}{2}$?

Le raisonnement ne peut être juste. Il ne l'est pas en effet.

Après l'ouverture du premier tiroir deux cas restent possibles. Sur ces deux cas, un seul est favorable, cela est vrai, mais les deux cas n'ont pas même vraisemblance.

Si la pièce qu'on a vue est en or, l'autre peut être en argent, mais on aurait avantage à parier pour qu'elle soit en or.

Supposons, pour en faire paraître l'évidence, qu'au lieu de trois coffrets on en ait trois cents. Cent contiennent deux médailles d'or, cent deux médailles d'argent et cent une médaille d'or et une médaille d'argent. Dans chaque coffret on ouvre un tiroir, on voit par conséquent trois cents médailles. Cent d'entre elles sont en or et cent en argent, cela est certain; les cent autres sont douteuses, elles appartiennent aux coffrets dont les pièces ne sont pas pareilles : le hasard en règlera le nombre.

On doit s'attendre, en ouvrant les trois cents tiroirs, à y voir moins de deux cents pièces d'or : la probabilité pour que la première qui se présente appartienne à l'un des cent coffrets dont l'autre pièce est en or est donc plus grande que $\frac{1}{2}$.

3. Supposons, pour second exemple, que Pierre et Paul jouent aux boules; celui qui placera la boule la plus rapprochée du but gagnera. Ils sont également habiles; mais Pierre a deux boules à jeter, Paul n'en a qu'une. Quelle est la probabilité pour que Pierre gagne?

Sur trois boules jetées par des joueurs également habiles, Pierre en a deux. La probabilité de gagner est pour lui $\frac{2}{3}$.

Ne pourrait-on pas dire cependant :

Chacune des boules de Pierre peut être meilleure ou moins bonne que la boule de Paul; quatre cas sont donc possibles. Sur les quatre, un seul fait perdre Pierre, celui où ses deux boules sont l'une et l'autre moins bonnes que celle de Paul, les trois autres cas lui sont favorables. La probabilité de gagner, pour Pierre, est $\frac{3}{4}$.

L'énumération est exacte, mais les cas n'ont pas même vraisemblance.

Paul a de bons et mauvais coups. Si la boule qu'il a lancée l'emporte sur la première boule de Pierre, il est à croire, sans rien savoir de plus, qu'elle n'est pas parmi les mauvaises. La chance pour qu'elle soit moins bonne que la seconde boule de

Pierre est diminuée. Parmi les quatre cas possibles, ceux dans lesquels Pierre vaincu dans un coup est vainqueur dans l'autre sont moins vraisemblables que ceux dans lesquels ses deux boules ont le même sort.

4. Une remarque encore est nécessaire : l'infini n'est pas un nombre; on ne doit pas, sans explication, l'introduire dans les raisonnements. La précision illusoire des mots pourrait faire naître des contradictions. Choisir *au hasard,* entre un nombre infini de cas possibles, n'est pas une indication suffisante.

On demande, par exemple, la probabilité pour qu'un nombre, entier ou fractionnaire, commensurable ou incommensurable, choisi *au hasard* entre 0 et 100, soit plus grand que 50. La réponse semble évidente : le nombre des cas favorables est la moitié de celui des cas possibles. La probabilité est $\frac{1}{2}$.

Au lieu du nombre, cependant, on peut choisir son carré. Si le nombre est compris entre 50 et 100, le carré le sera entre 2500 et 10000.

La probabilité pour qu'un nombre choisi *au hasard* entre 0 et 10000 surpasse 2500 semble évidente : le nombre des cas favorables est les trois quarts du nombre des cas possibles. La probabilité est $\frac{3}{4}$.

Les deux problèmes sont identiques. D'où vient la différence des réponses? Les énoncés manquent de précision.

Les contradictions de ce genre peuvent être multipliées à l'infini.

5. On trace *au hasard* une corde dans un cercle. Quelle est la probabilité pour qu'elle soit plus petite que le côté du triangle équilatéral inscrit?

On peut dire : si l'une des extrémités de la corde est connue, ce renseignement ne change pas la probabilité; la symétrie du cercle ne permet d'y attacher aucune influence, favorable ou défavorable à l'arrivée de l'événement demandé.

L'une des extrémités de la corde étant connue, la direction doit être réglée par le hasard. Si l'on trace les deux côtés du triangle équilatéral ayant pour sommet le point donné, ils forment entre eux et avec la tangente trois angles de 60°. La corde, pour être plus grande que le côté du triangle équilatéral, doit se trouver dans celui des trois angles qui est compris entre les deux autres. La probabilité pour que le hasard entre trois angles égaux qui peuvent le recevoir le dirige dans celui-là semble, par définition, égale à $\frac{1}{3}$.

On peut dire aussi : si l'on connaît la direction de la corde, ce renseignement ne change pas la probabilité. La symétrie du cercle ne permet d'y attacher aucune influence, favorable ou défavorable à l'arrivée de l'événement demandé.

La direction de la corde étant donnée, elle doit, pour être plus grande que le côté du triangle équilatéral, couper l'un ou l'autre des rayons qui composent le diamètre perpendiculaire, dans la moitié la plus voisine du centre. La probabilité pour qu'il en soit ainsi semble, par définition, égale à $\frac{1}{2}$.

On peut dire encore : choisir une corde au hasard, c'est en choisir au hasard le point milieu. Pour que la corde soit plus grande que le côté du triangle équilatéral, il faut et il suffit que le point milieu soit à une distance du centre plus petite que la moitié du rayon, c'est-à-dire à l'intérieur d'un cercle quatre fois plus petit en surface. Le nombre des points situés dans l'intérieur d'une surface quatre fois moindre est quatre fois moindre. La probabilité pour que la corde dont le milieu est choisi au hasard soit plus grande que le côté du triangle équilatéral semble, par définition, égale à $\frac{1}{4}$.

Entre ces trois réponses, quelle est la véritable? Aucune des trois n'est fausse, aucune n'est exacte, la question est mal posée.

6. On choisit au hasard un plan dans l'espace; quelle est la probabilité pour qu'il fasse avec l'horizon un angle plus petit que $\frac{\pi}{4}$?

On peut dire : tous les angles sont possibles entre o et $\frac{\pi}{2}$, la probabilité pour que le choix tombe sur un angle inférieur à $\frac{\pi}{4}$ est $\frac{1}{2}$.

On peut dire aussi : par le centre d'une sphère, menons un rayon perpendiculaire au plan en question. Choisir le plan au hasard, c'est choisir au hasard le point où cette perpendiculaire perce la sphère.

Pour que l'angle du plan avec l'horizon soit plus petit que $\frac{\pi}{4}$, il faut que la perpendiculaire coupe la sphère dans l'intérieur d'une zone dont la surface est

$$2\pi R^2\left(1-\cos\frac{\pi}{4}\right)=4\pi R^2\sin^2\frac{\pi}{8}.$$

Le rapport de la surface de cette zone à celle de la demi-sphère est

$$2\sin^2\frac{\pi}{8}=0,29.$$

La probabilité est donc 0,29.

Cette question, comme la précédente, est mal posée et les deux réponses contradictoires en sont la preuve.

7. On fixe au hasard deux points sur la surface d'une sphère; quelle est la probabilité pour que leur distance soit inférieure à $10'$?

Le premier point peut être supposé connu, la position qu'il occupe, quelle qu'elle soit, ne change rien à la probabilité cherchée.

Le grand cercle qui réunit les deux points peut être également supposé connu, les chances possibles sont les mêmes dans toutes les directions. Si l'on partage ce cercle en 2160 arcs de $10'$, de manière que le point commun soit un point de division, les points situés dans les deux arcs séparés par le point donné remplissent seuls la condition demandée : la probabilité est donc $\frac{2}{2160}=\frac{1}{1080}$.

On peut dire aussi : le premier point étant connu, pour que le second soit à une distance moindre que $10'$, il faut qu'il soit situé dans une zone dont la surface est

$$4\pi R^2\sin^2 5'=4\pi R^2\sin^2\frac{\pi}{2160}.$$

Le rapport de la surface de cette zone à celle de la sphère est

$$0,0000042308=\frac{1}{236362},$$

plus de deux cents fois plus petite que $\frac{1}{1080}$.

Les probabilités relatives à la distribution des étoiles, en les supposant semées *au hasard* sur la sphère céleste, sont impossibles à assigner si la question n'est pas précisée davantage.

8. L'application directe de la définition, lorsque les cas possibles sont en nombre déterminé et d'égale vraisemblance, est souvent, au contraire, un problème facile appartenant à la théorie des combinaisons.

Problème I. — *On jette une pièce de monnaie n fois de suite. Quelle est la probabilité pour que pile et face se succèdent dans un ordre assigné?*

Pile et face, à chaque épreuve, sont également possibles. Toutes les successions présentent des chances égales. Leur nombre, pour n épreuves, est 2^n, car chaque jet nouveau présente deux cas possibles qui, l'un et l'autre, peuvent s'associer à chacune des combinaisons précédentes, dont le nombre est par conséquent doublé. Parmi ces 2^n combinaisons, une seule est demandée. La probabilité est $\frac{1}{2^n}$. L'arrivée de face ou de pile est prise ici comme le type d'un événement dont la probabilité est $\frac{1}{2}$.

La probabilité de voir, à la roulette, la rouge et la noire se succéder dans un ordre assigné serait exactement la même.

Quelle est, par exemple, la probabilité pour que, la rouge se montrant au premier coup, la couleur change sans interruption à chacun des 29 coups suivants?

La probabilité est

$$\frac{1}{2^{30}} = 0{,}0000000009313\,3.$$

La probabilité pour que, pendant les trente premiers coups, la rouge et la noire se succèdent sans que la même couleur sorte deux fois de suite est double de la précédente. La suite alternée peut commencer par rouge ou par noire. Cela fait deux cas possibles.

9. Problème II. — *Quelle est la probabilité pour obtenir avec un dé à six faces n fois de suite le point* 3?

A chaque jet, six cas sont possibles et peuvent s'adjoindre à chacune des combinaisons précédentes, dont le nombre est ainsi sextuplé. Le nombre des combinaisons possibles sur n épreuves est donc 6^n; une seule est demandée, elle a pour probabilité $\frac{1}{6^n}$.

10. Problème III. — *Quelle est la probabilité pour amener avec deux dés une somme de points égale à 7?*

Le nombre des cas possibles quand on jette deux dés est 36. La somme 7 peut être obtenue par 6 et 1, 5 et 2, 4 et 3. Chacune de ces manières représente deux cas; 6 et 1, par exemple, peut résulter de 6 donné par le premier dé avec 1 donné par le second, ou de 1 donné par le premier avec 6 donné par le second.

Sur 36 cas possibles, 6 sont favorables; la probabilité est $\frac{6}{36}$ ou $\frac{1}{6}$.

11. Problème IV. — *Quelle est la probabilité pour amener avec deux dés n fois de suite le point 7?*

Le nombre des combinaisons possibles sur n épreuves faites avec deux dés est 36^n.

Chacune des six combinaisons qui, à chaque épreuve, peuvent amener la somme 7, en s'associant à toutes les précédentes, multiplie par 6 le nombre des cas favorables. Le nombre total de ces cas pour les n épreuves est donc 6^n. La probabilité, rapport du nombre des cas favorables à celui des cas possibles, est

$$\frac{6^n}{36^n} = \frac{1}{6^n}.$$

12. Problème V. — *Quelle est la probabilité, en jetant deux dés, pour avoir une somme de points égale à 8?*

La somme 8 peut être comme la somme 7, obtenue de trois manières : 6 et 2, 5 et 3, 4 et 4; mais ces trois manières représentent cinq cas et non six. Le doublet 4 et 4 ne peut se produire, en effet, que si les deux dés donnent 4. Le nombre des cas possibles est 5 et la probabilité est $\frac{5}{36}$.

13. Problème VI. — *Quelle est la probabilité, en jetant deux dés trois fois de suite, pour amener une fois au moins un doublet.*

Le nombre des combinaisons possibles sur trois épreuves est

$$36^3 = 46656.$$

A chaque coup, six combinaisons sont des doublets, trente n'en sont pas. Le nombre total des combinaisons qui ne contiennent aucun doublet est donc, sur trois épreuves, 30^3 ou 27000. Le nombre des combinaisons qui contiennent un doublet au moins est

$$36^3 - 30^3 = 19656.$$

La probabilité demandée est

$$\frac{19656}{46656} = 0{,}421296.$$

14. Problème VII. — *Quelle est la probabilité pour qu'en jetant deux dés n fois de suite on amène* sonnez *au moins une fois?*

A chaque fois que l'on jette deux dés, le nombre des associations de points possibles est 36; 35, sur ces 36, sont autres que *sonnez*.

Le nombre des combinaisons possibles en n coups est 36^n; le nombre de celles dont *sonnez* est exclu est 35^n; le nombre des combinaisons qui contiennent une ou plusieurs fois *sonnez* est

$$36^n - 35^n.$$

La probabilité demandée est

$$\frac{36^n - 35^n}{36^n} = 1 - \left(\frac{35}{36}\right)^n.$$

Si l'on veut que cette probabilité soit $\frac{1}{2}$, il faut déterminer n par l'équation

$$1 - \left(\frac{35}{36}\right)^n = \frac{1}{2}.$$

On en déduit

$$n = \frac{l2}{l36 - l35} = 24,605.$$

Si l'on parie d'amener *sonnez* en 24 coups, les chances de perte l'emportent sur celles de gain. C'est le contraire si l'on accorde 25 coups.

Le problème précédent, proposé à Pascal par le chevalier de Méré, a été l'occasion des premières recherches sur le Calcul des probabilités.

« Je n'ai pas le temps, écrivait Pascal à Fermat, de vous envoyer la démonstration d'une difficulté qui étonnait fort M. de Méré; car il a un très bon esprit, mais il n'est pas géomètre. C'est, comme vous savez, un grand défaut. Il me disait donc qu'il avait trouvé difficulté sur les nombres pour cette raison : Si l'on entreprend de faire 6 avec un dé, il y a avantage de l'entreprendre par 4 coups. Si l'on entreprend de faire *sonnez* avec deux dés, il y a désavantage de l'entreprendre en 24 coups, et néanmoins 24 est à 36, qui est le nombre des faces des deux dés comme 4 est à 6, qui est le nombre des faces d'un dé.

» Voilà quel était son grand scandale et qui lui faisait dire hautement que les propositions n'étaient pas constantes et que l'Arithmétique se dément. »

La réponse ne pouvait embarrasser ni Pascal ni Fermat.

Il semble que l'exemple d'un dé à deux faces, c'est-à-dire d'une pièce de monnaie, aurait suffi pour ouvrir les yeux de Méré. Sur 1 coup, avec une pièce de monnaie, on peut parier sans avantage ni désavantage, d'amener une fois pile. Avec un dé à six faces, Méré le savait, il y a avantage à entreprendre d'amener un point donné en 4 coups. La proportion est donc en défaut, et la petitesse des nombres permet d'en apercevoir aisément la raison.

Lorsque le nombre des faces d'un dé s'accroît et lorsque l'on augmente le nombre des tirages, le nombre total des cas possibles et celui des cas qui contiennent un point donné ne croissent proportionnellement ni au nombre des faces ni à celui des coups.

On peut s'en assurer soit par l'examen des cas les plus simples, soit par la recherche d'une formule. La première méthode était seule accessible au chevalier de Méré. Pascal, en voulant lui expliquer l'autre, a aperçu son grand défaut : il n'était pas géomètre.

15. Problème VIII. — *La probabilité d'un événement est p, combien faut-il tenter d'épreuves pour que la probabilité de voir l'événement se produire au moins une fois dépasse une fraction donnée r?*

Supposons que, dans une urne, soient contenues m boules blanches et n boules noires, m et n étant telles que

$$\frac{m}{m+n}=p.$$

Il faut chercher combien de tirages doivent être tentés pour que la probabilité d'amener une boule blanche soit plus grande que r.

La boule sortie est remise dans l'urne après chaque épreuve, de telle sorte que la probabilité reste, à chaque tirage, égale à p.

Le nombre des combinaisons possibles, sur k épreuves, est

$$(m+n)^k;$$

le nombre de celles qui ne contiennent pas de boules blanches est n^k.

Le nombre des combinaisons contenant une boule blanche au moins est

$$(m+n)^k-n^k.$$

La probabilité demandée est donc

$$\frac{(m+n)^k-n^k}{(m+n)^k}=1-\left(\frac{n}{m+n}\right)^k=1-(1-p)^k.$$

Le nombre k des épreuves à tenter pour que cette probabilité soit égale à r est donné par l'équation

$$1-(1-p)^k=r,$$

d'où

$$k=\frac{l(1-r)}{l(1-p)}. \tag{1}$$

Cette valeur de k n'est pas, en général, un nombre entier. Pour un nombre d'épreuves plus petit que k, la probabilité de voir une boule blanche sortir sera moindre que r; elle surpassera r si le nombre des épreuves est plus grand que k.

Le chevalier de Méré se scandalisait de ne pas voir k inversement proportionnel à p.

« Pourquoi, disait-il, pour $p = \frac{1}{36}$, k n'est-il pas six fois plus grand que pour $p = \frac{1}{6}$? »

La formule (1) sert de réponse pour les géomètres; elle montre que, pour de petites valeurs de p, la proportionnalité supposée par de Méré s'éloigne peu de la vérité.

On a, en effet,

$$l(1-p) = -p - \frac{p^2}{2} - \frac{p^3}{3} - \dots.$$

Si p est très petit, on peut supposer

$$l(1-p) = -p.$$

Le nombre k des épreuves est donc, à très peu près, pour une valeur donnée de r, proportionnel à $\frac{1}{p}$, pourvu que p soit très petit.

On a, en supposant les logarithmes népériens, ce qui est permis dans la formule (1), dont cela ne change pas les rapports,

$$\begin{aligned}
l\tfrac{1}{2} &= -\, l2 = -\,0,6931.\\
l\tfrac{1}{3} &= -\, l3 = -\,1,0986,\\
l\tfrac{1}{10} &= -\, l10 = -\,2,3025,\\
l\tfrac{1}{100} &= -\, l100 = -\,4,6051,\\
l\tfrac{1}{1000} &= -\, l1000 = -\,6,9077,\\
l\tfrac{1}{10000} &= -\, l10000 = -\,9,210.
\end{aligned}$$

La formule (1) donne donc les théorèmes suivants :

Si un événement a pour probabilité $\frac{1}{N}$, N étant assez grand pour qu'on puisse négliger $\frac{1}{N^2}$ et prendre $-\frac{1}{N}$ pour le logarithme népérien de $1-\frac{1}{N}$, le nombre des épreuves qu'il faut tenter pour acquérir les probabilités $\frac{1}{2}$, $\frac{2}{3}$, $\frac{9}{10}$, $\frac{99}{100}$, $\frac{999}{1000}$, $\frac{9999}{10000}$ de voir l'événement se produire au moins une fois est :

Pour la probabilité	$\frac{1}{2}$..........	0,69 N
»	$\frac{2}{3}$.........	1,098 N
»	$\frac{9}{10}$......	2,3025 N
»	$\frac{99}{100}$......	4,605 N
»	$\frac{999}{1000}$.....	6,907 N
»	$\frac{9999}{10000}$.....	9,210 N

On peut s'expliquer sans calcul comment, avec le nombre des épreuves, la probabilité s'accroît pour devenir rapidement une quasi-certitude.

Le chevalier de Méré aurait accepté sans étonnement le premier chiffre du Tableau. Il y a chance égale, il l'avait reconnu, pour que *sonnez*, dont la probabilité est $\frac{1}{36}$, arrive ou n'arrive pas en 24 coups, environ. 24 est les $\frac{2}{3}$ de 36, comme 0,69N est, à peu près, les $\frac{2}{3}$ de N.

La proportion n'est pas exacte, mais approchée.

Quand le nombre des épreuves est 9,2N, on peut le partager en 14 groupes, environ, de 0,69N chacun. Celui qui a parié pour l'arrivée de l'événement une fois au moins en 9,2N coups est dans le même cas que s'il était admis à renouveler quatorze fois une épreuve qui donne probabilité égale pour gagner ou pour perdre.

Ne pas voir une seule fois l'arrivée de l'événement est, pour lui, aussi peu vraisemblable que de perdre quatorze fois de suite à pile ou face.

16. Problème IX. — *Quelle est la probabilité pour qu'en tirant n boules de suite dans une urne qui en contient* μ, *marquées* 1, 2, 3, ..., μ, *et remettant la boule sortie après chaque tirage, k numéros désignés soient sortis au moins une fois?*

Lorsque k est égal à 1, un seul numéro étant désigné, le nombre des combinaisons qui le contiennent est

$$\mu^n - (\mu - 1)^n. \tag{2}$$

Il faut, en effet, retrancher du nombre total des combinaisons, μ^n, le nombre des combinaisons qui restent possibles quand le numéro désigné est laissé en dehors.

Pour avoir le nombre des combinaisons qui contiennent deux numéros désignés, il faut retrancher de (2) le nombre des combinaisons qui, ne contenant pas le second, contiennent le premier ; ce nombre est donné par la même formule, dans laquelle μ est remplacé par $\mu - 1$,

$$(\mu - 1)^n - (\mu - 2)^n.$$

Le nombre des combinaisons qui contiennent une ou plusieurs fois deux numéros désignés est donc

$$\mu^n - 2(\mu - 1)^n + (\mu - 2)^n. \tag{3}$$

Pour avoir le nombre des combinaisons qui contiennent trois numéros désignés, il faut, de (3), retrancher le nombre des combinaisons qui, ne contenant pas le troisième, contiennent les deux premiers. Ce nombre s'obtient en remplaçant dans (3) μ par $\mu - 1$; il est

$$(\mu - 1)^n - 2(\mu - 2)^n + (\mu - 3)^n.$$

Le nombre des combinaisons qui contiennent une ou plusieurs fois trois numéros désignés est

$$\mu^n - 3(\mu - 1)^n + 3(\mu - 2)^n - (\mu - 3)^n. \tag{4}$$

La méthode est générale. Le nombre des combinaisons qui contiennent k numéros désignés est

$$\mu^n - k(\mu - 1)^n + \frac{k(k-1)}{1.2}(\mu - 2)^n \ldots \pm (\mu - k)^n.$$

En divisant ce nombre par μ^n, nombre des combinaisons possibles, on aura la probabilité demandée.

Si l'on suppose $k = \mu$, on aura la probabilité pour que n tirages fassent sortir tous les numéros

$$1 - \mu\left(1 - \frac{1}{\mu}\right)^n + \frac{\mu(\mu - 1)}{1.2}\left(1 - \frac{2}{\mu}\right)^n - \ldots. \tag{5}$$

La suite contient $\mu + 1$ termes.

Si n et μ sont grands, on a approximativement

$$\left(1-\frac{1}{\mu}\right)^n = e^{-\frac{n}{\mu}},$$
$$\left(1-\frac{2}{\mu}\right)^n = e^{-\frac{2n}{\mu}},$$
$$\dots\dots\dots\dots\dots,$$
$$\left(1-\frac{h}{\mu}\right)^n = e^{-\frac{hn}{\mu}}.$$

Ces formules d'approximation sont applicables aux premiers termes de (5); les autres sont négligeables, et l'on peut représenter approximativement la probabilité pour que tous les numéros soient sortis en n tirages par

$$\left(1-e^{-\frac{n}{\mu}}\right)^{\mu}.$$

17. Problème X. — *Une urne contient μ boules marquées* $1, 2, 3, \ldots, \mu$. *On les tire successivement, sans les remettre dans l'urne quand elles sont sorties. Quelle est la probabilité pour que, parmi k boules désignées, aucune ne sorte à un rang égal au numéro dont elle est marquée?*

Les combinaisons dans lesquelles une boule désignée n'est pas à son rang sont au nombre

$$1.2.3\ldots\mu - 1.2.3\ldots(\mu-1). \tag{6}$$

Le nombre de celles dans lesquelles deux boules désignées ne sont ni l'une ni l'autre à leur rang s'obtiendra en retranchant de (6) le nombre de celles dans lesquelles, la seconde boule étant à son rang, la première, dans l'arrangement des $\mu-1$ autres, ne serait pas au sien. Ce nombre s'obtient en remplaçant dans (6) μ par $\mu-1$; il est

$$1.2.3\ldots(\mu-1) - 1.2.3\ldots(\mu-2).$$

Le nombre des combinaisons dans lesquelles deux boules désignées ne sont ni l'une ni l'autre à leur rang est, par conséquent,

$$1.2.3\ldots\mu - 2.1.2.3\ldots(\mu-1) + 1.2.3\ldots(\mu-2). \tag{7}$$

Pour avoir le nombre des combinaisons dans lesquelles trois boules désignées ne sont pas à leur rang, il faut retrancher du nombre (7) le nombre des combinaisons dans lesquelles, la troisième étant à son rang, les deux autres ne sont pas aux leurs dans les $\mu - 1$ boules restantes. Ce nombre s'obtient en remplaçant dans (7) μ par $\mu - 1$; il est

$$1.2.3\ldots(\mu-1) - 2.1.2.3\ldots(\mu-2) + 1.2.3\ldots(\mu-3). \tag{8}$$

La différence avec (7) est

$$1.2.3\ldots\mu - 3.1.2.3\ldots(\mu-1) + 3.1.2.3\ldots(\mu-2) - 1.2.3\ldots(\mu-3).$$

La méthode est générale. Le nombre des combinaisons dans lesquelles k boules ne sont pas à leur place est

$$\left\{\begin{array}{l} 1.2.3\ldots\mu - k.1.2.3\ldots(\mu-1) \\ \quad + \dfrac{k.k-1}{1.2}\,1.2.3\ldots(\mu-2) - \ldots \pm 1.2.3\ldots(\mu-k). \end{array}\right. \tag{9}$$

La probabilité pour qu'aucune des boules désignées ne sorte à son rang est le quotient de (9) par le nombre $1.2.3\ldots\mu$ des combinaisons possibles.

En supposant $k = \mu$, on obtient la probabilité pour que, dans le tirage de toutes les boules, aucune ne sorte à son rang; elle est

$$1 - 1 + \frac{1}{1.2} - \frac{1}{1.2.3} + \frac{1}{1.2.3.4} - \ldots. \tag{10}$$

Le nombre des termes dont la loi est évidente est celui des numéros de l'urne augmenté de un.

La série (10), si on la prolongeait indéfiniment, représenterait $\frac{1}{e}$. Si donc μ n'est pas un très petit nombre entier, la probabilité devient, quel que soit le nombre des boules, très voisine de $\frac{1}{e}$,

$$\frac{1}{e} = 0,36787944.$$

La probabilité pour qu'un numéro au moins sorte à son rang est

$$1 - \frac{1}{e} = 0,63212055$$

Voici, du reste, les valeurs, pour

$$\mu = 1, 2, \ldots, 10, 11,$$

de la probabilité p_μ, à moins d'une demi-unité du septième ordre décimal :

$$\begin{aligned}
p_1 &= 0.0000000,\\
p_2 &= 0,5000000,\\
p_3 &= 0,3333333,\\
p_4 &= 0.3750000,\\
p_5 &= 0,3666667,\\
p_6 &= 0.3680556,\\
p_7 &= 0.3678571,\\
p_8 &= 0,3678819,\\
p_9 &= 0,3678792,\\
p_{10} &= 0,3678795,\\
p_{11} &= 0,3678794.
\end{aligned}$$

On voit donc que, *dès que* μ *est égal ou supérieur à* 9, *la différence entre* p *et* $\frac{1}{e}$ *n'atteint pas un millionième.*

18. Problème XI. — *Pierre et Paul sont soumis à un scrutin de ballottage; l'urne contient* m *bulletins favorables à Pierre,* n *favorables à Paul;* m *est plus grand que* n, *Pierre sera élu. Quelle est la probabilité pour que, pendant le dépouillement du scrutin, les bulletins sortent dans un ordre tel que Pierre ne cesse pas un seul instant d'avoir l'avantage.*

Le nombre des combinaisons possibles est celui des arrangements de $m+n$ lettres, parmi lesquelles m semblables entre elles représentent le nom de Pierre et les n autres, semblables aussi, représentent le nom de Paul.

Le nombre de ces arrangements est

$$\frac{1.2.3\ldots(m+n)}{1.2.3\ldots m.1.2.3\ldots n}.$$

Cherchons le nombre des arrangements dans lesquels Paul, à un certain moment, aura le même nombre de voix que Pierre.

Parmi ces arrangements, il faut compter tous ceux qui commencent par le nom de Paul. Leur nombre est celui des arrangements de $m+n-1$ lettres, parmi lesquelles m semblables entre elles représentent le nom de Pierre, et les $n-1$ autres celui de Paul. Ce nombre est

$$\frac{1.2.3\ldots(m+n-1)}{1.2.3\ldots m.1.2.3\ldots(n-1)}; \tag{11}$$

ils forment la moitié, précisément, du nombre que nous cherchons.

Nous allons démontrer, en effet, que tous les dépouillements commençant par le nom de Pierre, et dans lesquels l'égalité des suffrages se produit à un certain moment, correspondent un à un, d'une manière parfaitement déterminée, aux arrangements de $m+n-1$ lettres dont le nombre est représenté par la formule (11).

Représentons par la lettre A les bulletins qui portent le nom de Pierre et par B ceux qui portent le nom de Paul.

Considérons une combinaison commençant par A et dans laquelle, à un certain moment, quand on appelle tous les termes à partir du premier, le nombre des B est égal à celui des A. Arrêtons-nous dans cette énumération faite de gauche à droite la première fois que cette égalité se produit. Enlevons le groupe déjà appelé, transportons-le à la fin de la combinaison, après avoir enlevé le B qui le termine nécessairement. Nous formerons la combinaison de $m+n-1$ lettres, conjuguée de celle que nous avons choisie.

Par exemple, la combinaison

AABBABAB,

qui contient quatre A et quatre B, donnera

ABABAAB,

qui contient quatre A et trois B.

La combinaison de $m+n-1$ lettres, déduite, comme nous l'avons dit, de l'une des combinaisons commençant par A dont nous voulons faire le compte, permet, quand elle-même est donnée, de retrouver celle dont on l'a déduite. Il suffit d'y appeler les lettres en commençant par la droite jusqu'au moment, qui ne peut manquer de se produire, puisque les A sont en majorité, où le nombre des A surpassera d'une unité celui des B. On enlèvera alors le groupe ainsi défini pour en faire le commencement de la combinaison, en le transportant à la gauche après l'avoir séparé des termes non transportés, qui gardent leur ordre, par la lettre B ajoutée à la fin.

Le nombre total des dépouillements dans lesquels le nombre des lettres B, à un certain moment, atteint l'égalité est donc

$$2\frac{1.2.3\ldots(m+n-1)}{1.2.3\ldots m.1.2.3\ldots n-1}.$$

En le divisant par le nombre total des arrangements distincts

$$\frac{1.2.3\ldots(m+n)}{1.2.3\ldots m.1.2.3\ldots n},$$

le quotient

$$(12)\qquad \frac{2n}{m+n}$$

est la probabilité pour que Pierre, pendant la durée du scrutin, perde, à un certain moment, l'avantage.

La probabilité pour que Pierre garde toujours l'avantage est

$$(13)\qquad 1-\frac{2n}{m+n}=\frac{m-n}{m+n}.$$

Cette ingénieuse démonstration est due à M. André.

19. Problème XII. — *Une urne contient μ boules numérotées* 1, 2, 3, ..., μ. *Quelle est la probabilité pour que, dans n tirages, la somme des numéros sortis soit égale à k?*

On suppose qu'après chaque tirage la boule sortie soit remise dans l'urne.

Le nombre des combinaisons possibles est μ^n, puisque chacun des μ numéros peut sortir à chacun des n tirages.

Le nombre des cas favorables à l'événement demandé est le nombre des manières de former une somme égale à k avec n nombres inférieurs à $\mu+1$. Cette somme est, évidemment, le coefficient de t^μ dans le développement de

$$(1)\qquad (t+t^2+\ldots+t^\mu)^n$$

ordonné suivant les puissances de t. Par quelque voie que l'on effectue ce développement, les coefficients des puissances de t feront connaître les numérateurs des diverses probabilités dont le dénominateur commun est μ^n.

On a

$$t + t^2 + \ldots + t^\mu = \frac{t^{\mu+1} - t}{t - 1}.$$

L'expression (1) peut donc s'écrire

$$t^n (t^\mu - 1)^n (t - 1)^{-n}.$$

Les développements du deuxième et du troisième facteur se feront par la formule du binome.

Le troisième donnera un développement illimité; mais la multiplication fera disparaître tous les termes dans lesquels l'exposant est négatif.

Le Tableau suivant a été formé par cette méthode. On y a inscrit en regard de chaque somme demandée, et pour un nombre de tirages inférieur à huit, le nombre des combinaisons qui la produisent en supposant $\mu = 6$.

Les tirages peuvent être remplacés par des dés, en nombre égal à n, que l'on jette à la fois sur le tapis. La somme des points est inscrite dans la première colonne verticale, le nombre des dés en tête de chaque autre colonne, et le nombre des combinaisons sur la colonne horizontale commençant par la somme demandée.

Nombre demandé	1 dé.	2 dés.	3 dés.	4 dés.	5 dés.	6 dés.	7 dés.	8 dés.
1......	1	0	0	0	0	0	0	0
2......	1	1	0	0	0	0	0	0
3......	1	2	1	0	0	0	0	0
4......	1	3	3	1	0	0	0	0
5......	1	4	6	4	1	0	0	0
6......	1	5	10	10	5	1	0	0
7......	0	6	15	20	15	6	1	0
8......	0	5	21	35	35	21	7	1
9......	0	4	25	56	70	56	28	8
10......	0	3	27	80	126	126	84	36
11......	0	2	27	104	205	252	210	120
12......	0	1	25	125	305	456	462	330
13......	0	0	21	140	420	756	917	792

Nombre demandé.	1 dé.	2 dés.	3 dés.	4 dés.	5 dés.	6 dés.	7 dés.	8 dés
14......	0	0	15	146	540	1161	1667	1708
15......	0	0	10	140	651	1666	2807	3368
16......	0	0	6	125	735	2247	4417	6147
17......	0	0	3	104	780	2856	6538	10480
18......	0	0	1	80	780	3431	9142	16808
19......	0	0	0	56	735	3906	12117	25488
20.....	0	0	0	35	651	4221	15267	36688
21......	0	0	0	20	540	4332	18327	50288
22......	0	0	0	10	420	4221	20993	65808
23......	0	0	0	4	305	3906	22967	82384
24......	0	0	0	1	205	3431	24017	98813
25......	0	0	0	0	126	2856	24017	113688
26......	0	0	0	0	70	2247	22967	125588
27.....	0	0	0	0	35	1666	20993	133288
28......	0	0	0	0	15	1161	18327	135954
29......	0	0	0	0	5	756	15267	133288
30......	0	0	0	0	1	456	12117	125588

20. Problème XIII. — *Combien faut-il jeter de fois trois dés pour avoir la probabilité p d'amener une fois au moins le point* 15?

La probabilité z d'amener 15 en jetant trois dés est, d'après le Tableau précédent,

$$\frac{10}{216} = 0,0462963.$$

La probabilité d'amener 15 une fois au moins en n coups est

$$1-(1-z)^n.$$

Le nombre n est donc donné par l'équation

$$(1-z)^n = 1-p;$$

on trouve, en faisant

$$p = \tfrac{1}{2} \ldots\ldots\ldots\ldots \quad n = 14,62$$
$$p = \tfrac{3}{4} \ldots\ldots\ldots\ldots \quad n = 29,24$$
$$p = \tfrac{9}{10} \ldots\ldots\ldots\ldots \quad n = 48,57$$

En jetant quinze fois les trois dés, la probabilité d'amener 15 est plus grande que $\frac{1}{2}$; elle est plus grande que $\frac{3}{4}$ en les jetant trente fois et que $\frac{9}{10}$ en les jetant quarante-neuf fois.

21. Le jeu du passe-dix, autrefois fort en vogue, donnait aux deux adversaires chance égale de gagner.

Pierre jetait trois dés et gagnait si la somme des points surpassait 10; il perdait dans le cas contraire.

Le Tableau montre que, dans le cas de trois dés, les points supérieurs sont précisément en même nombre que les points non supérieurs à 10; 11 a autant de chances que 10, 12 que 9, 13 que 8, etc.

La même symétrie se rencontre pour les points obtenus avec cinq ou sept dés. Avec quatre, six ou huit, il existe un maximum pour lequel il n'y a pas de point conjugué.

La raison en est simple. Sur un dé, le point 6 et le point 1 sont sur des faces opposées, le point 5 est opposé à 2 et 4 à 3. Si donc on jette sur une table $2n+1$ dés, la somme des points marqués par les faces posées sur la Table, ajoutée à celle des points marqués par les dés, sera $7(2n+1)$. Si l'une des deux sommes est plus grande que $7n+3$, l'autre sera au plus égale à $7n+3$. Celui qui parierait d'amener un point supérieur à $7n+3$ a donc pour lui la moitié des chances; lorsque $2n+1=3$, n est égal à 1 et $7n+3$ est égal à 10. Si l'on jouait avec cinq dés, il faudrait au passe-dix, pour laisser les chances égales, substituer le passe-dix-sept.

C'est pour cette raison que les derniers chiffres du Tableau ont été supprimés. Veut-on savoir, par exemple, le nombre des manières d'amener 44 avec 8 dés. Ce nombre est le même pour $28+16$ et pour $28-16$, ou 12; il est donc 330.

CHAPITRE II.

PROBABILITÉS TOTALES ET PROBABILITÉS COMPOSÉES.

> Un des points les plus importants de la théorie des probabilités et celui qui prête le plus aux illusions est la manière dont les probabilités augmentent ou diminuent par leurs combinaisons mutuelles
>
> LAPLACE

22. Les solutions de problèmes isolés ne font pas une théorie. — 23. Théorème des probabilités totales. — 24 Probabilités composées. — 25 Cas où le premier événement influe sur la probabilité du second. — 26. Les théorèmes ne sont démontrés que dans les cas pour lesquels la probabilité est définie; on complète les définitions. Les théorèmes deviennent généraux. — 27. Erreur commise dans la théorie du tir à la cible. — 28. Erreur commise dans la théorie des gaz. — 29 Erreur commise dans l'appréciation des pronostics sur le temps. — 30. Problème relatif aux tirages faits dans une urne Fausseté d'un raisonnement en apparence très plausible. — 31. Probabilité du brelan au jeu de la bouillotte. — 32. Avantage du banquier au jeu de trente-et-quarante. — 33. Études sur le jeu du baccarat — 34. Problème de la poule.

22. La solution des problèmes précédents n'a pas exigé de principes nouveaux. Le compte des cas favorables rapproché de celui des cas possibles appartient à la théorie des combinaisons. L'ingénieux et bizarre Cardan, bon géomètre et grand ami des dés, connaissait pour chaque coup le nombre exact des chances; il l'a publié sans devenir, pour cela, l'inventeur de la Science du hasard.

Une Science doit enchaîner les cas simples aux cas composés et reposer sur des principes. Ceux du Calcul des probabilités sont aussi simples que féconds.

Deux théorèmes fondamentaux rencontrent des applications à chaque page pour ainsi dire de la théorie des chances. Nous les démontrerons d'abord dans les cas où la probabilité a été définie. L'étude des autres cas ne peut précéder, évidemment, la définition du sens exact attaché aux mots qui y figurent.

23. *Probabilités totales.* — La probabilité d'un événement étant, par définition, le rapport du nombre des cas favorables à celui des cas possibles, *si l'on partage les cas favorables en plusieurs groupes, la probabilité de l'événement sera la somme des probabilités pour qu'il appartienne à chacun des groupes.* On ajoute en effet les fractions de même dénominateur en ajoutant les numérateurs.

La probabilité d'amener avec trois dés une somme de points supérieure à 14 est la somme des probabilités pour obtenir 15, 16, 17 ou 18.

Le choix des groupes est arbitraire, sous la seule condition, bien entendu, d'y enfermer tous les cas possibles sans qu'aucun s'y rencontre deux fois.

Si, par exemple, pour calculer la probabilité d'amener *sonnez* une fois *au moins* en 20 coups, on faisait la somme des probabilités calculées pour chaque coup, on appliquerait mal le principe. Le *sonnez* obtenu à un coup n'empêche pas celui du coup suivant; on compterait vingt fois, de cette manière, la série de coups dans laquelle *sonnez* arrive vingt fois.

La probabilité d'amener avec deux dés le point 3 ou le point 4 n'est pas la somme de la probabilité pour amener 3 et de celle pour amener 4. On peut, en effet, les amener tous deux; et le point 3 et 4, compté une première fois comme contenant 3, ne doit pas l'être une seconde comme contenant 4.

24. *Probabilités composées.* — Un événement composé est défini par le concours de plusieurs événements simples que le hasard doit amener successivement ou simultanément. Le nombre N des cas possibles, lorsque les événements simples sont indépendants, est le produit $\mu_1 \mu_2 \ldots \mu_k$ des nombres de cas possibles dans chacun d'eux, chaque cas de l'un des groupes pouvant s'associer avec tous les cas des autres groupes. Le nombre des cas favorables au concours des événements simples dont la réunion forme l'événement composé est, par une raison semblable, le produit des nombres de cas m_1, m_2, ..., m_k favorables à l'arrivée de chacun d'eux.

La probabilité de l'événement composé est donc

$$\frac{m_1 m_2 m_3 \ldots m_k}{\mu_1 \mu_2 \mu_3 \ldots \mu_k} = \left(\frac{m_1}{\mu_1}\right)\left(\frac{m_2}{\mu_2}\right)\cdots\left(\frac{m_k}{\mu_k}\right);$$

c'est le produit des probabilités des événements simples.

25. *La probabilité d'un événement composé est le produit des probabilités des événements simples dont il exige la réunion.* Un cas doit être prévu, c'est celui où la probabilité du second événement est influencée par la manière dont se produit le premier. Si, par exemple, une urne contient deux boules blanches et deux boules noires, quelle est la probabilité, en tirant deux boules de suite sans les remettre dans l'urne, pour obtenir les deux boules blanches? La probabilité d'en obtenir une au premier tirage est $\frac{1}{2}$; mais, au second, elle est douteuse : égale à $\frac{2}{3}$ si le premier tirage a enlevé une boule noire, elle sera $\frac{1}{3}$ seulement si la boule disparue est blanche. Quelle est celle de ces deux fractions qui doit servir de multiplicateur? C'est la seconde, évidemment. Dans le premier des deux cas supposés, l'événement demandé est devenu impossible. Peu importe la probabilité des épreuves qui viennent ensuite.

La probabilité d'un événement composé est le produit de la probabilité du premier événement par la probabilité qu'acquiert le second *quand on sait que le premier est arrivé.*

On a souvent commis des erreurs graves en oubliant les derniers mots de cet énoncé.

26. Les deux théorèmes précédents sont et doivent être incomplètement démontrés. La probabilité, en effet, n'a été définie que pour une classe très restreinte d'événements. Il en existe d'autres, incertains comme eux, dans lesquels l'énumération des cas ne peut rien apprendre. Les principes leur sont-ils applicables? Sont-ils dès à présent démontrés pour eux ?

Les principes sont applicables.

Ils ne sont pas encore démontrés. Comment le seraient-ils? Les probabilités dont ils donnent la mesure n'ont pas même été définies.

Quelle est la probabilité pour que la Seine soit gelée à Paris dans le courant de l'année 1995?

Pour qu'un médecin appelé près d'un malade sache découvrir la nature, la cause et le remède du mal?

Pour qu'un homme âgé de quarante ans, aujourd'hui bien portant, atteigne l'âge de soixante ans?

Il faut compléter la définition; tous ces cas lui échappent.

La probabilité d'un événement, quelle qu'en soit la nature, est dite égale à une fraction donnée p, lorsque celui qui attend l'événement pourrait échanger indifféremment les craintes ou les espérances, les avantages ou les inconvénients attachés à l'arrivée de cet événement contre les conséquences supposées identiques de la sortie d'une boule puisée dans une urne dont la composition fait naître une probabilité égale à p.

Comment, dans des cas tels que ceux qu'on a cités, justifier une telle assimilation? Ce n'est pas en ce moment la question. Si l'assimilation est impossible, on ne le fera pas. Si on la fait, il faudra la justifier.

La définition, dans les deux cas, reste irréprochable.

Lorsque l'assimilation sera faite et justifiée, on en acceptera les conséquences.

On aura le droit de dire, par exemple :

Si, après avoir appelé un médecin, on évalue à $\frac{9}{10}$ la probabilité pour qu'il vienne et à $\frac{1}{3}$ la probabilité pour qu'il procure, s'il vient, la guérison du malade; sans discuter ces chiffres, celui qui les admet peut ajouter : La probabilité pour que le malade soit visité et guéri par le médecin est, *pour moi*,

$$\frac{9}{10} \times \frac{1}{3} = \frac{3}{10}.$$

Celui qui accepte, en effet, les probabilités $\frac{9}{10}$ et $\frac{1}{3}$ est, par définition, dans le même doute que si, en présence de dix urnes indiscernables dont neuf renferment une boule blanche et deux boules noires, il cherchait la probabilité pour mettre la main sur l'une des neuf urnes dont la composition est connue et en tirer une boule blanche. L'identité des deux problèmes est admise; elle fait partie de l'énoncé. Si on allègue l'impossibilité de mesurer en chiffres les probabilités dont nous parlons, l'objection serait aussi peu fondée que si, évaluant la longueur d'un champ d'apparence rectangulaire à 300^m et la largeur à 100^m, on contestait le droit d'ajouter, indépendamment de toute vérification, ces mesures, si douteuses qu'elles soient, et ces appréciations assignent au champ une surface de trois hectares.

La règle des probabilités totales et celle des probabilités composées s'appliqueront à toutes les combinaisons de probabilités simples supposées évaluées en nombres. L'évaluation sera plus ou moins judicieuse, plus ou moins justifiée, les conséquences vaudront ce qu'elle vaut elle-même, mais sans introduire aucun doute nouveau.

Les précautions prescrites dans l'application des deux principes sont indispensables, bien entendu, quand on les transporte aux cas nouveaux.

Il ne faudrait pas dire, par exemple :

La probabilité pour que la Seine soit gelée à Paris en 1995 est la somme des probabilités pour qu'elle soit gelée pendant chacun des mois qui composent l'année.

Elle peut geler, en effet, en plusieurs mois différents, et (23) l'application du principe des probabilités totales n'est pas permise.

Si l'on a évalué la probabilité pour qu'il gèle demain à $\frac{1}{3}$, celle pour qu'il tombe de la neige a $\frac{1}{4}$, la probabilité pour voir à la fois de la glace et de la neige n'est pas $\frac{1}{12}$. La gelée, en effet, si elle se présente, change la probabilité de la neige.

Nous citerons trois exemples intéressants d'erreurs commises par l'oubli de ces conditions nécessaires dans l'énoncé des principes.

27. Problème XV. — *On tire à la cible. L'arme, sans être parfaite, ne présente aucun défaut systématique; les déviations ont en tous sens même probabilité. L'hypothèse est-elle réalisable? On le suppose.*

Quelle est la probabilité pour que le point frappé soit à une distance du but comprise entre r et $+ dr$?

Les données sont insuffisantes, cela semble évident. On a résolu cependant le problème par une fausse application des principes.

Rapportons le point où frappe la balle à deux axes de coordonnées ayant pour origine le centre de la cible, c'est-à-dire le point que l'on veut atteindre. Soient $\varphi(x)\,dx$ la probabilité inconnue pour que l'abscisse du point frappé tombe entre x et $x+dx$, $\varphi(y)\,dy$ la probabilité pour que l'ordonnée tombe entre y et $y+dy$. La probabilité pour que la balle frappe le rectangle $dx\,dy$, dont les coordonnées sont x et y, est, d'après le principe des probabilités composées,

$$\varphi(x)\,\varphi(y)\,dx\,dy.$$

Cette probabilité, d'après notre hypothèse, ne doit dépendre que de la distance du point frappé à l'origine, et l'on doit avoir

$$\varphi(x)\,\varphi(y) = \mathrm{F}(x^2+y^2).$$

Cette équation suffit pour déterminer la fonction φ. On en déduit, en prenant les dérivées successivement par rapport à x et à y,

$$\frac{\varphi'(x)}{x\,\varphi(x)} = \frac{\varphi'(y)}{y\,\varphi(y)}.$$

La fraction $\dfrac{\varphi'(x)}{x\,\varphi(x)}$ est par conséquent constante, et l'on en conclut que la fonction $\varphi(x)$, qui doit s'annuler quand x est infini, est de la forme

$$\varphi(x) = \mathrm{G}e^{-k^2x^2}.$$

Ce résultat, fort remarquable, n'est pas, malheureusement, acceptable.

La connaissance de la valeur de x changerait, en effet, la probabilité de celle de y et le facteur par lequel il faudrait multiplier $\varphi(x)\,dx$, pour obtenir la probabilité d'un écart y dans un sens, associé à un écart x dans l'autre, serait une fonction inconnue de x et de y, très différente peut-être de $\varphi(y)$.

La déviation de la balle dépend, en effet, du soin plus ou moins grand et plus ou moins habile avec lequel le coup a été préparé. Si l'on a réussi sous un certain point de vue, il y a plus de chances pour que le coup soit bon et que tous les écarts soient petits en même temps. La démonstration précédente ne tient aucun compte de cette remarque; les probabilités y sont traitées comme indépendantes.

28. Un physicien justement célèbre, Maxwell, a proposé dans l'étude des gaz un raisonnement dont l'illusion est semblable.

Les molécules d'un gaz, suivant une théorie que nous n'avons pas à discuter, se meuvent en tous sens avec de grandes vitesses. Les directions sont réglées par le hasard aussi bien que les vitesses, mais toutes les vitesses ne sont pas également probables; la vitesse maxima et la vitesse moyenne varient avec la température. C'est la probabilité pour qu'une molécule ait une vitesse donnée qu'on espère découvrir, sans introduire d'autres conditions.

Soit $\varphi(x)$ la probabilité pour que la composante parallèle à l'axe des X de la vitesse d'une molécule prise au hasard soit x. La probabilité pour que les trois composantes soient x, y, z parallèlement aux trois axes sera, d'après le théorème des probabilités composées, proportionnelle à

$$\varphi(x)\,\varphi(y)\,\varphi(z).$$

Mais la probabilité pour qu'une molécule ait une vitesse donnée, sans que l'on indique dans quel sens, doit être une fonction de cette vitesse, puisque toutes les directions sont supposées également possibles. On doit donc avoir

$$\varphi(x)\,\varphi(y)\,\varphi(z) = \mathrm{F}(x^2+y^2+z^2),$$

et cette condition suffit pour déterminer la forme de la fonction φ. On en déduit, comme (27),

$$\varphi(x) = \mathrm{G}e^{-k^2x^2}.$$

La démonstration n'est pas acceptable. Le principe des probabilités composées n'a pas été correctement appliqué. Si la composante de la vitesse d'une molécule parallèlement à l'axe des X est x, la valeur de x supposée connue influe sur la probabilité pour que la vitesse composée parallèlement à l'axe des Y soit y. Si, par exemple, x est égal à la vitesse maxima, le mouvement est certainement dirigé parallèlement à l'axe des X, et la probabilité de y est nulle.

La conclusion obtenue, indépendamment des objections que la théorie peut faire naître, ne mérite donc aucune confiance.

29. On pourrait multiplier les exemples; nous en citerons un troisième :

Un météorologiste annonce chaque soir le temps qu'il fera le lendemain. La probabilité pour qu'il pronostique juste est supposée égale à p.

Un second observateur fait, de son côté, des prédictions dont l'exactitude a pour probabilité p'.

Les deux observateurs s'accordent pour prédire qu'il pleuvra demain. Quelle est la probabilité pour qu'ils se trompent tous les deux?

La probabilité pour que le premier se trompe est $1 - p$.

La probabilité pour que le second se trompe est $1 - p'$.

La probabilité pour qu'ils se trompent tous les deux est une probabilité composée, mais elle n'est pas mesurée par le produit $(1 - p)(1 - p')$.

La probabilité composée est le produit de $(1 - p)$ par la probabilité pour que le second observateur se trompe, quand on sait que le premier a fait un faux pronostic.

Les données du problème laissent cette probabilité complètement inconnue.

Si les deux observateurs ont reçu les mêmes leçons, s'ils ont adopté les mêmes principes, en présence des mêmes faits ils porteront le même jugement. Si l'un se trompe, l'autre se trompera aussi; le second facteur du produit sera l'unité. L'accord certain des deux prédictions ne diminue pas la chance d'erreur.

Si les deux méthodes sont différentes, les conclusions pourront l'être aussi, sans pour cela devenir indépendantes. Certains symptômes sont nécessairement appréciés de la même manière, et leur nombre inconnu laisse le problème insoluble.

Si l'un des météorologistes annonce qu'il pleuvra, l'autre qu'il ne pleuvra pas, la probabilité pour qu'ils disent juste tous deux n'est pas pp' : elle est nulle.

30. Problème XVI. — *Une urne contient trois boules marquées* 1, 2 *et* 3. *On en tire deux successivement, en remettant*

dans l'urne, après le premier tirage, la boule qui en est sortie.

Quelle est la probabilité pour que le plus grand numéro sorti soit 2?

Pour que 2 soit le plus grand des numérss sortis, il faut que 3 ne se soit pas montré et que l'on n'ait pas tiré deux fois le n° 1.

Pour qu'à l'une des épreuves 3 ne sorte pas, la probabilité est $\frac{2}{3}$. Pour qu'il ne sorte ni au premier ni au second tirage, elle est $\frac{4}{9}$. Pour que 1 sorte deux fois, la probabilité est $\frac{1}{9}$; $\frac{8}{9}$ est par conséquent la probabilité pour qu'il ne sorte pas à l'un et l'autre tirage.

L'événement demandé semble donc composé de deux autres dont les probabilités sont $\frac{4}{9}$ et $\frac{8}{9}$; il ne faut pas cependant faire le produit de ces deux fractions. Il faut (25) multiplier $\frac{4}{9}$ par la probabilité pour que 1 ne sorte pas 2 fois, *lorsque l'on sait que 3 n'est pas sorti*. Cette probabilité est $\frac{3}{4}$; le numéro 3 étant écarté, il n'en reste en effet que deux. La probabilité pour que 1 sorte est $\frac{1}{2}$, pour qu'il sorte deux fois $\frac{1}{4}$, et pour qu'il ne sorte pas deux fois, par conséquent, elle est $\frac{3}{4}$.

La probabilité demandée est

$$\frac{9}{4} \times \frac{3}{4} = \frac{1}{3}.$$

On aurait pu dire : Pour que le plus grand des numéros sortis soit 2, il faut que 2 soit sorti et que 3 ne le soit pas.

La probabilité de n'amener 2 à aucune des deux épreuves est $\frac{4}{9}$, celle de l'amener une fois au moins est, par conséquent, $\frac{5}{9}$.

La probabilité $\frac{5}{9}$ doit être multipliée par la probabilité de ne pas amener 4, *sachant que 2 est sorti*.

Si 2 est sorti au premier tirage, la probabilité de ne pas amener 3 au second est $\frac{2}{3}$.

Si 2 est sorti au second tirage, la probabilité de ne pas avoir amené 3 au premier est $\frac{2}{3}$. Acceptons donc cettte probabilité $\frac{2}{3}$ qui convient aux deux cas, nous trouverons pour la probabilité demandée

$$\frac{5}{9} \times \frac{2}{3} = \frac{10}{27}.$$

Le désaccord avec le résultat précédent est un avertissement. Il n'était pas permis, comme nous l'avons fait, de partager la sortie supposée de 2 en deux cas distincts et de supposer successivement 2 sorti à la première ou à la seconde épreuve. Il a pu sortir à toutes deux ; notre calcul fait entrer deux fois en compte le cas où les deux tirages auraient, l'un et l'autre, amené le point 2.

31. Problème XVII. — *Probabilité des brelans au jeu de la bouillotte.*

La bouillotte se joue avec vingt cartes. On enlève d'un jeu de trente-deux cartes les sept, les valets et les dix.

Chacun des quatre joueurs reçoit trois cartes. On retourne la treizième.

Un joueur a brelan lorsque ses trois cartes sont de même espèce, trois rois, trois as, etc. Il a brelan carré lorsque toutes trois sont de même espèce que la retourne.

Soit p_i la probabilité pour que i joueurs désignés aient des brelans. On déduit du théorème des probabilités composées

$$p_1 = \frac{20.3.2}{20.19.18} = 0,017544,$$
$$p_2 = p_1 \frac{16.3.2}{17.16.15} = 0,0004128,$$
$$p_3 = p_2 \frac{12.3.2}{14.13.12} = 0,000013608,$$
$$p_4 = p_3 \frac{8.3.2}{11.10.9} = 0,00000065981.$$

On peut avoir brelan en effet quelle que soit la première carte; la probabilité pour avoir une première carte est $\frac{20}{20}$, celle pour que la deuxième soit de même espèce que la première est $\frac{3}{19}$, et pour que la troisième soit de même espèce que les deux autres la probabilité est $\frac{2}{18}$.

Cette probabilité p_1 calculée pour un premier joueur à qui l'on donnerait trois cartes, sans s'occuper des autres, convient à l'un quelconque des trois, car la manière de distribuer les cartes est indifférente.

Lorsqu'un premier joueur a brelan, la probabilité pour qu'un second l'ait aussi est changée. Il faut d'abord que la première carte rende le brelan possible. La probabilité est $\frac{16}{17}$, car elle ne doit pas être celle qui complète le brelan déjà fait; pour que la deuxième carte soit pareille à la première, la probabilité est $\frac{3}{16}$, et pour que la troisième soit pareille aux deux autres elle est $\frac{2}{15}$.

On en conclut

$$p_2 = p_1 \cdot \frac{16}{17}\frac{3}{16}\frac{2}{15};$$

les autres formules se démontrent de la même manière.

p_1, p_2, p_3, p_4 sont des probabilités totales. Si l'on représente par ϖ_i la probabilité pour que i joueurs désignés aient des brelans et qu'ils soient seuls à en avoir, on aura les relations

$$(1)\qquad \left\{\begin{aligned} &p_4 = \varpi_4, \qquad p_3 = \varpi_3 + \varpi_4, \qquad p_2 = \varpi_2 + 2\varpi_3 + \varpi_4,\\ &p_1 = \varpi_1 + 3\varpi_2 + 3\varpi_3 + \varpi_4.\end{aligned}\right.$$

Il est clair, en effet, que la probabilité pour qu'un joueur ait brelan se compose de la probabilité ϖ_1 pour qu'il l'ait seul, de la probabilité $3\varpi_2$ pour qu'il l'ait en même temps que l'un ou l'autre de ses adversaires, de la probabilité $3\varpi_2$ pour qu'il l'ait avec deux d'entre eux, ce qui fait trois cas, et de la probabilité ϖ_4 pour que les quatre joueurs aient brelan.

On déduit des équations (1)

$$\varpi_4 = p_4, \qquad \varpi_3 = p_3 - p_4, \qquad \varpi_2 = p_2 - 2p_3 + p_4,$$
$$\varpi_1 = p_1 - 3p_2 + 3p_3 - p_4$$

et, par conséquent,

$$\begin{aligned} \varpi_4 &= 0{,}000000659,\\ \varpi_3 &= 0{,}000012948,\\ \varpi_2 &= 0{,}000386244,\\ \varpi_1 &= 0{,}016345764.\end{aligned}$$

La probabilité pour qu'il n'y ait aucun brelan est

$$Q = (1 - R_1)(1 - R_2)(1 - R_3)(1 - R_4),$$

R_i étant la probabilité d'avoir brelan pour le joueur de rang i quand les précédents n'en ont pas.

On a évidemment

$$R_4 = \varpi_1, \qquad R_3 = \varpi_1 + \varpi_2, \qquad R_2 = \varpi_1 + 2\varpi_2 + \varpi_3,$$
$$R_1 = \varpi_1 + 3\varpi_2 + 3\varpi_3 + \varpi_4,$$

La probabilité R_4 pour que le quatrième joueur ait brelan, les autres ne l'ayant pas, ne diffère pas en effet de ϖ_1.

R_3, probabilité pour que le troisième joueur ait brelan, les deux premiers ne l'ayant pas, est une probabilité totale. Il peut avoir brelan tout seul, la probabilité est ϖ_1 ; ou l'avoir en même temps que le quatrième joueur, la probabilité est ϖ_2. Les autres formules se justifient par des raisons semblables.

On en conclut

$$\begin{aligned}
R_1 &= p_1 = 0{,}017554,\\
R_2 &= p_1 - p_2 = 0{,}0171312,\\
R_3 &= p_1 - 2p_2 + p_3 = 0{,}016782,\\
R_4 &= p_1 - 3p_2 + 3p_3 - p_4 = 0{,}1634576,\\
Q &= 0{,}93395.
\end{aligned}$$

32. Problème XVIII. — *Quelle est la probabilité de la chance favorable réservée au banquier dans le jeu de trente et quarante?*

Le jeu de trente et quarante se joue avec un grand nombre de cartes, réunion de divers jeux complets mêlés en un seul, bien entendu.

On abat une à une assez de cartes pour obtenir une somme de points supérieure à 30, les figures étant comptées pour 10 et les autres cartes pour le nombre de points qui s'y trouvent marqués.

Une seconde épreuve suit la première.

Le hasard donne ainsi deux sommes toutes deux plus grandes que 30 et egales au plus à 40.

Le joueur parie pour celle des deux sommes qu'il choisit et gagne si elle est plus grande que l'autre.

Si les sommes sont égales, le coup est nul. Un seul cas est excepté, celui où l'on a deux fois 31.

Ce refait de 31 est le seul avantage réservé au banquier; il a droit, dans ce cas, à la moitié des mises.

Nous supposerons le nombre des cartes assez grand pour que, pendant les deux premiers tirages, on puisse négliger l'influence des cartes sorties sur les probabilités des divers points. S'il y a huit jeux, par exemple, et par conséquent trente-deux sept, la sortie d'un sept diminue la probabilité d'en voir sortir un second. Il serait difficile de dire quelles sont les cartes dont la sortie accroît ou diminue la probabilité du refait de 31. Nous négligerons cette très petite influence.

Soit P_n la probabilité pour que, en abattant successivement les cartes, la somme prenne à un certain moment la valeur n.

La somme 1 ne peut se produire qu'au premier coup, si l'on abat un as. La probabilité pour cela est $\frac{1}{13}$; on a

$$P_1 = \frac{1}{13}.$$

La somme 2 peut se produire de deux manières : 2 au premier coup ; as au premier, as au second. On a

$$P_2 = \frac{1}{13} = \frac{1}{13^2} = \frac{1}{13}\left(1 + \frac{1}{13}\right).$$

La somme 3 peut se produire de trois manières : 3 au premier coup ; passer par la somme 2 obtenue en une ou en deux fois et la faire suivre d'un as ; commencer par un as, puis amener un 2.

On en déduit

$$P_3 = \frac{1}{13} + \frac{1}{13} P_2 + \frac{1}{13^2} = \frac{1}{13}\left(1 + \frac{1}{13}\right)^2.$$

On a, en général, tant que n est inférieur à 10,

$$P_n = \frac{1}{13} + \frac{1}{13} P_1 + \frac{1}{13} P_2 + \ldots + \frac{1}{13} P_{n-1}.$$

On trouve ainsi

$$\begin{aligned}
P_1 &= 0{,}076927,\\
P_2 &= 0{,}082846,\\
P_3 &= 0{,}089212,\\
P_4 &= 0{,}096075,\\
P_5 &= 0{,}108465,\\
P_6 &= 0{,}111424,\\
P_7 &= 0{,}117995,\\
P_8 &= 0{,}129226,\\
P_9 &= 0{,}139166.
\end{aligned}$$

A partir du point 10, les conditions changent; quatre cartes différentes, en effet, les dix et les trois figures, peuvent amener dix points. On a

$$P_{10} = \frac{4}{13} + \frac{1}{13}(P_1 + P_2 + P_3 + \ldots + P_9) = 0,380641.$$

Au-dessus de $n = 10$, la formule évidente

$$P_n = \frac{1}{13}(P_{n-1} + P_{n-2} + \ldots + P_{n-9}) + \frac{4}{13} P_{n-10}$$

donne

$$\begin{array}{ll}
P_{11} = 0,120218, & P_{22} = 0,142640, \\
P_{12} = 0,124908, & P_{23} = 0,145089, \\
P_{13} = 0,129613, & P_{24} = 0,147362, \\
P_{14} = 0,134304, & P_{25} = 0,149543, \\
P_{15} = 0,139398, & P_{26} = 0,151192, \\
P_{16} = 0,143167, & P_{27} = 0,152728, \\
P_{17} = 0,147143, & P_{28} = 0,154272, \\
P_{18} = 0,151978, & P_{29} = 0,155382, \\
P_{19} = 0,156021, & P_{30} = 0,168488, \\
P_{20} = 0.213024, & P_{31} = 0,148218. \\
P_{21} = 0,140033, &
\end{array}$$

La probabilité d'obtenir 31 étant P_{31}, celle de l'obtenir deux fois est

$$(P_{31})^2 = 0,0219686.$$

Poisson, dans un Mémoire de lecture difficile, a trouvé, pour valeur approchée de la probabilité du refait de 31,

$$0,021967.$$

Il tient compte de l'influence des cartes déjà passées sur la probabilité de celles qui suivent et suppose huit jeux réunis. L'influence est petite, on le voit par la concordance du résultat exact du calcul fait en supposant le nombre des jeux infini avec le résultat approché de l'autre.

Il ne faut pas croire que l'influence augmente lorsque, le jeu continuant, les cartes deviennent moins nombreuses. Il s'agit, en effet, de calculer l'*avantage* du banquier et non son bénéfice, variable avec le hasard des coups. Le calcul doit être fait avant que la première carte soit abattue ; si l'on doit, avant d'épuiser les huit jeux, faire cinquante *tailles*, les probabilités de la cinquantième sont, à ce moment, identiquement les mêmes que celles de la première.

Les cartes sorties n'auraient d'influence que si un joueur avait assez de mémoire pour se les rappeler toutes, et assez d'habileté pour calculer en quelques secondes leur influence sur la probabilité d'un refait. Le mérite serait grand, l'avantage bien petit.

33. Problème XIX. — *Est-il avantageux pour le ponte ou pour le banquier de demander une carte au jeu du baccarat quand il a le point* 5?

L'étude mathématique du jeu de baccarat est compliquée par la nécessité d'énumérer les cas, toujours nombreux, en faisant pour chacun un petit calcul.

Le ponte reçoit deux cartes, le banquier en prend deux. Le ponte a droit de demander une carte qui s'adjoint aux deux premières, ou de s'y tenir en gardant son jeu.

Le banquier a les mêmes droits, mais il a l'avantage, avant de prendre sa décision, de savoir si l'adversaire a demandé une carte et de connaître celle qu'il a reçue.

Chaque carte vaut, comme au trente et quarante, le nombre des points marqués sur elle. Les figures valent 10. Le gagnant est celui qui a le point le plus fort, les chiffres des dizaines ne comptant dans aucun cas : 11 vaut 1 ; 12 vaut 2 ; 23, si l'on a trois cartes, vaut 3.

Le jeu se termine immédiatement si l'un des joueurs reçoit, quand on donne les cartes, l'un des points 8 ou 9. Il *abat* et gagne si l'adversaire n'a pas un point meilleur. Dans ce cas, il n'est pas donné de cartes nouvelles.

Tous les points, à l'exception de 10 (ou zéro) ont même probabilité quand on donne les cartes : la probabilité de 10 est $\frac{25}{169} = 0,1479$, celle de chacun des autres $\frac{16}{169} = 0,09467$. La démonstration ne présente aucune difficulté.

Quand un joueur demande une carte, quel que soit le point qu'il ait, il a probabilité $\frac{4}{13}$ de le conserver en recevant un dix et probabilité $\frac{1}{13}$ de le changer pour un quelconque des autres points, qui deviennent tous également probables.

Nous traiterons une seule question.

Lorsque le ponte a reçu le point 5, est-il avantageux pour lui de demander une carte?

Il faut résoudre quatre problèmes :

Le ponte ayant 5 et ne demandant pas de carte, quelle est pour lui la probabilité de gagner et quelle est celle de faire coup nul, lorsque le banquier, ignorant qu'il a le point 5, sait qu'il a l'habitude, quand il a ce point, de ne pas demander de carte?

Le ponte ayant 5 et ne demandant pas de carte, quelles sont pour lui les probabilités de gagner ou de faire coup nul, lorsque le banquier, ignorant qu'il a le point 5, croit qu'il a l'habitude, quand il a ce point, de demander une carte?

Le ponte ayant 5 et demandant une carte, quelles sont pour lui les probabilités de gagner ou de faire coup nul, lorsque le banquier connaît son habitude de demander une carte dans cette circonstance ou lorsqu'il croit savoir qu'il n'en demande pas?

Résolvons le premier problème.

Le ponte a 5. Le banquier l'ignore. En le voyant s'y tenir, il apprend qu'il a 5, 6 ou 7. Ces trois cas ont chacun pour probabilité $\frac{1}{3}$. Huit suppositions peuvent être faites : le banquier peut avoir 0, 1, 2, 3, 4, 5, 6 ou 7. Les probabilités des huit hypothèses étaient $\frac{25}{169}$ pour le point 0 et $\frac{16}{169}$ pour les autres, au moment où les cartes ont été données. Mais on n'a pas abattu ; les points 8 et 9 ne sont plus possibles. Les probabilités deviennent

$$\frac{25}{137} = 0,18248$$

et

$$\frac{16}{137} = 0,01679.$$

Que fera le banquier? S'il a 0, 1, 2, 3, 4, il prendra une carte.

Il pourra hésiter s'il a 5 ou 6, et il faut résoudre cette question incidente : Dans les circonstances supposées, c'est-à-dire le ponte s'y étant tenu et ayant l'habitude de ne pas tirer à 5, le banquier doit-il tirer à 6? doit-il tirer à 5?

Nous ferons le calcul pour le cas où le banquier a le point 6 ; il se compose de deux autres :

Quelle est la probabilité de gagner pour le banquier qui, ayant 6, demande une carte dans les conditions supposées?

Quelle est la probabilité s'il ne demande pas de carte?

Si le banquier ne tire pas ayant 6, il a probabilité $\frac{1}{3}$ de perdre le ponte ayant 7, $\frac{1}{3}$ de gagner le ponte ayant 5, $\frac{1}{3}$ de faire coup nul.

Si le banquier, ayant 6, demande une carte, il acquiert la probabilité $\frac{1}{13}$ d'avoir chacun des neuf points autres que 6 et $\frac{4}{13}$ d'avoir celui-là.

La probabilité de perdre contre l'adversaire dont le point est 5, 6 ou 7 doit s'évaluer par l'énumération des cas :

Il y a la probabilité $\frac{5}{13}$ de perdre avec les points 0, 1, 2, 3, 4 ; la probabilité $\frac{1}{13}$ pour avoir, avec le point 5, probabilité $\frac{1}{13}$ de faire coup nul et $\frac{2}{3}$ de perdre ; la probabilité $\frac{4}{13}$ pour avoir, avec le point 6, probabilité $\frac{1}{3}$ de gagner, $\frac{1}{3}$ de faire coup nul et $\frac{1}{3}$ de perdre; probabilité $\frac{1}{13}$ pour acquérir, avec le point 7, probabilité $\frac{2}{3}$ de gagner, $\frac{1}{3}$ de faire coup nul ; probabilité $\frac{2}{13}$, enfin de gagner avec le point 8 ou avec le point 9.

La probabilité de gagner est, d'après cette énumération,

$$\frac{4}{13}\,\frac{1}{3}+\frac{1}{13}\,\frac{2}{3}+\frac{2}{13}=\frac{4}{13};$$

celle de faire coup nul

$$\frac{1}{13}\,\frac{1}{3}+\frac{4}{13}\,\frac{1}{3}+\frac{1}{13}\,\frac{1}{3}=\frac{2}{13}.$$

La probabilité de gagner était $\frac{1}{3}$; elle devient $\frac{4}{13}$: elle a diminué.

Celle de faire coup nul a également diminué ; le banquier, dans les conditions supposées, ne doit pas tirer à 6.

Un calcul fondé sur les mêmes principes et tout aussi facile montre que, dans les conditions supposées, le banquier doit tirer à 5.

Sans énumérer les cas possibles, qui deviennent plus nombreux quand le ponte a reçu une carte supposée connue, nous nous bornerons à donner les résultats.

Le ponte se tient à 5, le banquier sachant qu'il a cette habitude.

Le sort du ponte qui a 5 est le suivant :

Probabilité	de gagner...........	0,444694
»	de faire coup nul....	0,085907
»	de perdre....	0,469400

Le ponte qui a 5 demande une carte, le banquier sait que telle est son habitude.

Le sort du ponte qui a 5 est le suivant :

Probabilité	de gagner...........	0,444348
»	de faire coup nul....	0,120935
»	de perdre...........	0,434717

Tels sont les résultats du calcul lorsque le ponte ne cherche pas à tromper le banquier sur ses habitudes.

Si le ponte qui a 5 s'y tient, faisant croire au banquier qu'il a l'habitude de demander des cartes, le sort est le suivant :

Probabilité	de gagner...........	0,489612
»	de faire coup nul....	0,094890
»	de perdre...........	0,415497

Si le ponte, ayant 5, demande une carte en faisant croire au banquier qu'il a l'habitude de s'y tenir, son sort est le suivant :

Probabilité	de gagner...........	0,447113
»	de faire coup nul....	0,126463
»	de perdre...........	0,426424

L'étude n'est pas complète. Si le banquier ignore ce que fait le ponte quand il a 5, doit-il tirer à 6? Quelle est, dans cette indécision, la chance du ponte qui a 5?

Il est impossible de la calculer : elle dépend de la chance qu'il y a, quand il aura pris son parti, pour que le banquier se trompe ou devine juste en se demandant ce qu'il fait quand il a 5.

En résumé, le ponte doit-il se tenir à 5, doit-il tirer?

Si, sans jouer au plus fin, dès le début de la partie, il déclare franchement ses habitudes, il doit tirer à 5.

Si les conventions du jeu permettent la ruse, il doit se tenir à 5, en faisant croire au banquier, s'il le peut, qu'il a l'habitude de tirer.

34. Problème XX. — *$n+1$ joueurs A_1, A_2, ..., A_{n+1} jouent aux conditions suivantes : A_1 et A_2 jouent une première partie ; A_3 remplace le perdant; A_4, A_5, ... luttent successivement contre le perdant de la partie précédente. La* poule *continue ainsi jusqu'à ce qu'un joueur ait gagné successivement tous les autres, et par conséquent n parties de suite. Quelle est la probabilité pour que la partie se termine au coup de rang x?*

Le joueur qui gagne la poule au coup de rang x est entré au jeu au coup de rang $x-n+1$; il a gagné ce coup et les $n-1$ suivants. On peut partager cette hypothèse en $n-1$ autres, en distinguant les cas d'après le nombre des parties gagnées par l'adversaire du joueur dont nous parlons au moment où celui-ci est entré au jeu.

Soit p_1 la probabilité pour que ce premier adversaire ait gagné une seule partie, la probabilité pour que cette hypothèse se présente et fasse gagner la poule au coup de rang x est

$$p_1 \frac{1}{2^n};$$

mais on a évidemment

$$y_{x-1} = p_1 \frac{1}{2^{n-1}}.$$

Car, si la partie se termine au coup de rang $x-1$, le joueur qui la gagne avait une partie gagnée déjà au coup de rang $x-n+1$, et il en a ensuite gagné $n-1$ sans interruption.

La probabilité pour que la poule se termine au coup de rang x, le joueur qui la gagne ayant rencontré d'abord un adversaire qui n'avait gagné qu'une seule partie, est, d'après cela,

$$\frac{1}{2} y_{x-1}.$$

Soit p_2 la probabilité pour que le gagnant de la poule au coup de rang x ait eu pour premier adversaire un joueur ayant gagné deux parties déjà, le terme correspondant de la valeur de y_x sera

$$p_2 \frac{1}{2^n}.$$

On a, évidemment,

$$y_{x-2} = p_2 \frac{1}{2^{n-2}};$$

par conséquent,

$$y_{x-2} \frac{1}{2^2}$$

est le second terme de la valeur de y_x.

Le raisonnement successivement appliqué à chacun des cas donnera

$$(1) \qquad y_x = \frac{1}{2} y_{x-1} + \frac{1}{4} y_{x-2} + \frac{1}{8} y_{x-3} + \ldots + \frac{1}{2^{n-1}} y_{x-n+1}.$$

On a, évidemment,

$$y_1 = 0, \qquad y_2 = 0, \qquad y_3 = 0, \qquad \ldots, \qquad y_{n-1} = 0, \qquad y_n = \frac{1}{2^{n-1}}.$$

L'équation (1), à l'aide de ces données initiales, fera connaître y_x pour $x = n + 1$ et pour les valeurs suivantes.

35. Nous donnerons, pour terminer ce Chapitre, l'énoncé de quelques problèmes empruntés pour la plupart à Moivre et dont la solution est facile.

Problème XXI. — *Probabilité pour obtenir une fois le point 1, et une fois seulement, en jetant quatre dés.*

$$\frac{500}{1296} = 0,3858.$$

Problème XXII. — *Pierre parie que, en jetant cinq dés, il obtiendra une fois le point 1 et pas davantage. Quelle est la probabilité de gagner?*

$$\frac{4651}{7776} = 0,59813.$$

Problème XXIII. — *Quelle est la probabilité pour que, sur cent essais successifs avec un seul dé, on obtienne une fois au moins une succession de cinq as sans interruption?*

$$0,01026.$$

PROBLÈME XXIV. — *Pierre parie d'amener en trois coups, avec deux dés, le point 5 et le point 7, chacun une fois. Quelle est sa probabilité pour gagner?*

$$\frac{31}{324} = 0,0956.$$

PROBLÈME XXV. — *Pierre entreprend d'obtenir le point 7 avec deux dés avant qu'aucun autre ne se soit produit deux fois. Quelle est la probabilité de gagner?*

$$\frac{7303}{13860} = 0,5269.$$

PROBLÈME XXVI. — *Pierre jette trois pièces de monnaie, Paul en jette deux; celui des deux qui amènera le plus de* faces *gagnera. Si les nombres sont égaux, on recommencera l'épreuve jusqu'à ce qu'elle donne un résultat. Quelle est la probabilité pour que Pierre gagne?*

$$\frac{8}{11}.$$

PROBLÈME XXVII. — *Pierre lance un dé et recommence autant de fois qu'il faudra pour amener, soit deux fois le point* 1, *soit l'un des points* 2 *ou* 3. *Quelle est la probabilité d'amener deux fois le point* 1?

$$\frac{1}{9}.$$

PROBLÈME XXVIII. — *Pierre et Paul jouent avec deux dés, Pierre gagnera par le point* 7, *Paul par le point* 6. *Paul joue le premier et ils jettent les dés alternativement jusqu'à ce que l'un d'eux ait amené le point qui le fait gagner. Quelle est la probabilité de Pierre?*

$$\frac{31}{61}.$$

PROBLÈME XXIX. — *Pierre et Paul jouent avec deux dés; Pierre jette les dés deux fois. S'il obtient le même point, il* *gagne. Si les points sont différents, il continue à jeter les dés jusqu'à ce qu'il ait obtenu l'un des deux premiers points. Il gagne si c'est le premier. Quelle est la probabilité de gagner?*

$$\frac{721}{1296} = 0,5563.$$

PROBLÈME XXX. — *On jette en l'air cinq pièces de monnaie. Quelle est la probabilité pour qu'elles montrent trois piles et deux faces?*

$$\frac{5}{16}.$$

PROBLÈME XXXI. — *Pierre et Paul jouent à un jeu d'adresse. Le gagnant, après chaque partie, marque un point. Le jeu devient égal lorsque Pierre, plus habile que Paul, lui rend deux points sur trois, qui font gagner la partie. Quelle est, à chaque partie, la probabilité du gain pour Pierre?*

$$\frac{1}{\sqrt[3]{2}} = 0,7937.$$

PROBLÈME XXXII. — *Pierre et Paul jouent aux conditions énoncées dans le problème précédent. Pierre peut, sans désavantage ni avantage, rendre un point sur trois. Quelle est pour lui la probabilité p de gagner une partie?*

Le rapport $\frac{1-p}{p} = z$ est donné par l'équation

$$1 + 4z = z^4 + 4z^3 + 6z^2.$$

On en déduit approximativement

$$z = 0,627, \qquad p = 0,614.$$

PROBLÈME XXXIII. — *Une loterie contient un très grand nombre de billets; le nombre des lots est le quarantième du nombre des billets. Combien faut-il prendre de billets pour avoir probabilité $\frac{1}{2}$ de gagner ou de ne pas gagner un lot?*

27 ne suffisent pas, 28 donnent au gain une chance plus grande que $\frac{1}{2}$.

Problème XXXIV. — *Dans une loterie de* 40 000 *billets, il y a trois lots. Pierre prend* 8 000 *billets. Quelle est la probabilité d'avoir un lot?*

$$\frac{61}{125} = 0,488, \quad \text{à peu près.}$$

Quelle est celle de les avoir tous les trois?

$$\frac{1}{125} = 0,008, \quad \text{à peu près.}$$

Problème XXXV. — *Combien Pierre devrait-il prendre de billets, à la loterie indiquée au problème précédent, pour avoir chance* $\frac{1}{2}$ *de gagner un des lots?*

8252.

Problème XXXVI. — *Une loterie, sur* 100 000 *billets, doit donner* 10 000 *lots. Combien faut-il prendre de billets pour avoir chance* $\frac{1}{2}$ *de gagner un lot?*

Le calcul donne 6,6. Avec 7 billets, la probabilité de gagner un lot, au moins, est 0,52172; avec 6 billets, elle est 0,46857.

Problème XXXVII. — *Une loterie, sur* 1 000 *billets, promet trois lots. Quelle est la probabilité, avec trente billets, de gagner un lot?*

$$\frac{8714976}{99700200} = 0,0874.$$

Problème XXXVIII. — *On a dans une urne trente boules, dix blanches, dix noires, dix rouges; on en prend trois au hasard. Quelle est la probabilité pour avoir une boule de chaque couleur?*

$$\frac{100}{406} = 0,2463.$$

Problème XXXIX. — *En combien de coups, avec trois dés, peut-on parier, avec chance égale, d'amener trois as deux fois?*

361 coups.

PROBLÈME XL. — *En combien de coups, avec trois dés, peut-on parier, avec chance égale, d'amener deux fois le point* 15?

45 coups.

PROBLÈME XLI. — *En combien de coups, avec un seul dé, peut-on parier, avec chance égale, de voir les six faces?*

13 coups.

PROBLÈME XLII. — *En combien de coups, avec deux dés, peut-on parier, avec chance égale, d'amener tous les doublets?*

79 coups.

CHAPITRE III.

ESPÉRANCE MATHÉMATIQUE.

Hoc utrobique utar fundamento : nimirum, in aleæ ludo tanti æstimandum est cujusque sortem, seu expectationem ad aliquid obtinendum, quantum si habeat, possit denuo ad similem sortem sive expectationem pervenire, æqua conditione certans.

HUYGENS.

36. Définition de l'espérance mathématique. — 37. Assertion exagérée de Poisson. — 38. La recherche de l'espérance mathématique et celle de la probabilité sont deux problèmes distincts. — 39. Exemple de la simplification d'un problème par la recherche directe de l'espérance mathématique. — 40. Second exemple. — 41. Troisième exemple. — 42. Quatrieme exemple dans lequel la recherche de l'espérance mathématique fait connaître la probabilité. — 43. Calcul d'une espérance mathématique déduite des probabilités des divers cas possibles. — 44. Problème sur le jeu de dés. — 45. Discussion de la formule obtenue. — 46. La valeur probable d'une fonction n'est pas déterminée par celle des grandeurs dont elle dépend. — 47. Exception relative aux sommes et aux produits quand les facteurs sont indépendants. — 48. Paradoxe de Saint-Pétersbourg. — 49. Insuffisance des explications proposées par Condorcet et par Poisson. — 50. La réponse du calcul est parfaitement raisonnable et n'a besoin d'aucune justification. — 51. Insignifiance de l'explication proposée par Daniel Bernoulli et devenue célèbre sous le nom de *théorie de l'espérance morale*.

36. Le *sort* des joueurs a été la première préoccupation des créateurs de la théorie du hasard. On le nomme aujourd'hui *espérance mathématique,* il est la valeur équitable des avantages encore incertains que font espérer les conditions du jeu.

L'espérance mathématique, dans un jeu équitable, est, pour chaque joueur, égale à sa mise qui, livrée au jeu, ne lui appartient plus. Cette égalité traduit la définition : le joueur échange sa mise contre une espérance mathématique. Si l'équivalence n'existe pas, le jeu n'est pas équitable.

L'espérance mathématique de celui qui a la probabilité p de recevoir la somme S est mesurée par le produit pS.

Si n personnes, en effet, ont sur la somme S des droits égaux, elles peuvent équitablement la partager et prendre chacune $\frac{S}{n}$ ou la tirer au sort et accepter pour part la probabilité $\frac{1}{n}$ d'obtenir S. Si quelques-uns des ayants-droit, au nombre de m, s'associent, on pourra leur offrir équitablement, soit le partage donnant à leur association $\frac{mS}{n}$, soit la probabilité $\frac{m}{n}$ de recevoir S. Les deux offres sont donc équivalentes.

37. Poisson a écrit : « Si le gain espéré par une spéculation est 60000^{fr} et que $\frac{1}{3}$ soit la probabilité de l'obtenir, la personne qui devra recevoir cette somme éventuelle pourra considérer le tiers de 60000^{fr} comme un bien qu'elle possède et que l'on devrait comprendre dans l'inventaire de sa fortune. »

Poisson va trop loin. Le plaideur engagé dans un procès qu'il a neuf chances sur dix de perdre et qui, en cas de gain, doit lui rapporter 1 million, mentirait en disant qu'il possède 100000^{fr}. Un homme prudent, sur une telle garantie, refuserait de lui prêter 500^{fr}. Son espérance mathématique *vaut* 100000^{fr}, mais vraisemblablement il ne trouvera pas d'acheteur.

Cette confusion entre une espérance mathématique et la certitude d'une somme équivalente a fait naître de grandes difficultés. Nous aurons à y revenir.

38. Si la somme espérée est connue, la recherche de la probabilité et celle de l'espérance mathématique forment un même problème. Il en est autrement lorsque les conditions du jeu impliquent la possibilité de gagner ou de perdre, suivant les cas, plusieurs sommes différentes.

Si des événements ayant pour probabilités $p_1, p_2, \ldots, p_n$ donnent droit aux sommes $S_1, S_2, \ldots, S_n$, l'espérance mathématique est

$$p_1 S_1 + p_2 S_2 + \ldots + p_n S_n.$$

Celui qui doit, suivant les cas encore incertains, recevoir l'une ou l'autre des sommes $S_1, S_2, \ldots, S_n$, peut vendre ses droits et, par des marchés équitables, échanger ses chances contre les sommes $p_1 S_1, p_2 S_2, \ldots, p_n S_n$ qui donneront droit à n acheteurs différents de toucher en son lieu et place, le premier la somme S_1, si c'est celle qui lui échoit, le second la somme $S_2, \ldots$, et enfin le dernier la somme S_n.

Le règlement des comptes ne fera naître aucune difficulté puisque deux acheteurs, d'après les conventions, n'auront jamais à réclamer la même somme.

L'espérance mathématique est connue quand les probabilités des divers cas possibles ont été calculées. Mais il est plus aisé, quelquefois, de la chercher directement sans s'occuper des termes qui la composent.

39. Problème XLIII. — *Pierre et Paul jouent au jeu de rencontre. Une urne contient μ numéros marqués $1, 2, 3, \ldots, \mu$. Paul tire successivement les μ boules et s'engage à donner à Pierre 1fr chaque fois qu'un numéro sortira à son rang. Quelle est l'espérance mathématique de Pierre?*

S'il fallait évaluer la probabilité des divers cas possibles, le problème, sans être difficile, exigerait de longs calculs.

En nommant p_i la probabilité pour qu'il y ait i rencontres, l'espérance demandée est

$$p_1 + 2p_2 + 3p_3 + \ldots + \mu p_\mu.$$

La somme est plus aisée à calculer que chacun des termes qui la composent.

L'espérance mathématique de Pierre est, à chaque tirage, $\frac{1}{\mu}$. La somme espérée est en effet 1fr, et la probabilité de l'obtenir $\frac{1}{\mu}$.

La probabilité au moment où le tirage va se faire dépend des numéros sortis antérieurement. Si le numéro i est sorti déjà, la probabilité de le voir au $i^{\text{ième}}$ tirage est nulle. S'il n'est pas sorti, elle est $\frac{1}{\mu - i + 1}$. Mais c'est au début du jeu qu'il faut la calculer, et la probabilité est alors $\frac{1}{\mu}$.

L'espérance mathématique de Pierre, étant pour chaque tirage $\frac{1}{\mu}$, est, pour les μ tirages, $\mu\frac{1}{\mu}$, c'est-à-dire 1^{fr}.

40. Problème XLIV. — *Une urne contient des boules noires et des boules blanches; la probabilité de la sortie d'une boule blanche est p, celle de la sortie d'une boule noire est q. On fait μ tirages en remettant chaque fois dans l'urne la boule qui en est sortie. Pierre recevra 1^{fr} chaque fois qu'une boule blanche sortira, précédée et suivie par une boule noire. Quelle est, pour l'ensemble des μ tirages, l'espérance mathématique de Pierre?*

La probabilité pour que le tirage de rang i donne à Pierre le droit de recevoir 1^{fr} est pq^2, car il faut le concours de trois événements dont les probabilités sont q, p et q. L'espérance mathématique étant, pour chaque tirage, égale à pq^2, elle est, pour l'ensemble des μ tirages, égale à μpq^2.

L'espérance mathématique est la même que si les μ tirages se faisaient dans une urne donnant à la sortie d'une boule blanche la probabilité pq^2, et que Pierre dût recevoir 1^{fr} pour chaque boule blanche sortie, sans qu'aucune condition fût imposée à celle qui la précède ni à celle qui la suit. Les deux jeux donnent à Pierre des avantages équivalents, mais non identiques. Dans le premier, sur μ tirages, la plus grande somme qu'il puisse gagner est $\frac{\mu}{2}$, dans le second elle est μ. Si l'on voulait, pour obtenir l'espérance mathématique, calculer les termes qui la composent et, pour cela, chercher la probabilité pour que sur μ tirages Pierre eût un nombre donné n de francs à recevoir, la difficulté des deux problèmes serait fort inégale. La considération, parfaitement rigoureuse, de l'espérance mathématique dispense de les résoudre. Pierre, en effet, peut, pour chacun des tirages, vendre la chance qu'il a de gagner 1^{fr} sans avoir à tenir compte de l'influence exercée par les chances de l'un des coups sur celles du suivant. Il est certain que, si l'acheteur du coup de rang i est favorisé par le sort, celui du coup suivant ne le sera pas; la boule qui lui appartient sera noire. Peu leur importe : la chance qu'ils achètent, au moment où ils la payent, n'est ni augmentée ni diminuée.

41. Problème XLV. — *Une urne contient un grand nombre n de boules marquées des numéros* 1, 2, 3, ..., n. *On fait successivement μ tirages, en remettant chaque fois dans l'urne la boule qui en est sortie. Pierre recevra* 1^{fr} *chaque fois que la série des numéros écrits dans leur ordre de sortie présentera un maximum ou un minimum. Quelle est l'espérance mathématique de Pierre?*

Si l'on considère un tirage désigné, la probabilité pour que le numéro correspondant soit dans la suite un maximum ou un minimum est $\frac{2}{3}$. Si l'on compare en effet le numéro sorti à celui qui le précède et à celui qui le suit, il sera maximum s'il est le plus grand des trois : la probabilité pour cela est $\frac{1}{3}$, et minimum s'il est le plus petit : la probabilité est également $\frac{1}{3}$; en écartant comme ayant une probabilité négligeable, si n est très grand, le cas où deux numéros consécutifs seraient égaux, la probabilité pour que le numéro dont le rang est donné soit maximum ou minimum est $\frac{2}{3}$. L'espérance mathématique de Pierre pour chacun des termes de la série est donc $\frac{2}{3}$; elle est pour les μ termes de la série $\frac{2}{3}\mu$.

Le raisonnement peut, comme le précédent, laisser un doute qu'il faut discuter.

Si, dans la suite, un numéro est maximum ou minimum, cela change la probabilité pour que le suivant le soit aussi.

La probabilité pour que le vingtième terme de la suite soit maximum est $\frac{1}{3}$.

La probabilité pour que le vingt et unième soit maximum est aussi $\frac{1}{3}$.

La probabilité pour qu'ils le soient tous deux n'est pas $\frac{1}{9}$, elle est nulle.

L'objection serait fondée si nous calculions les probabilités. Quand il s'agit des espérances mathématiques, elle perd toute sa force.

Pierre, en effet, peut vendre à des acheteurs différents chacun des francs qu'il peut espérer; tous les marchés sont équitables. Chaque acheteur a bien réellement probabilité $\frac{2}{3}$ de recevoir 1^{fr}; peu lui importe que son gain, s'il est obtenu, diminue les chances d'un autre : les conventions partielles n'en sont pas moins équitables. Pierre a pu vendre équitablement ses droits $\frac{2}{3}\mu^{\text{fr}}$. Telle est donc son espérance mathématique.

42. Si, une urne contenant deux boules blanches et une boule noire, on fait successivement μ tirages, l'espérance mathématique de celui qui doit recevoir 1^{fr} par boule blanche sortie sera $\frac{2}{3}\mu$ comme dans le cas précédent; là s'arrête la ressemblance des deux problèmes. La probabilité d'extraire de l'urne un certain nombre de boules blanches n'est pas égale à celle d'avoir un même nombre de maxima ou minima.

La probabilité pour que toutes les boules soient noires est $(\frac{1}{3})^\mu$, celle pour qu'elles soient toutes blanches $(\frac{2}{3})^\mu$. Les probabilités de n'avoir ni maximum ni minimum et celle de n'avoir que des maxima et des minima sont très différentes.

L'égalité entre les espérances mathématiques n'implique nullement celle des diverses probabilités qui figurent dans leur expression.

43. Problème XLVI. — *On trace sur un plan indéfini des lignes parallèles équidistantes. Une aiguille est lancée au hasard sur le plan. Pierre recevra* 1^{fr} *par rencontre de l'aiguille avec une des parallèles. Quelle est l'espérance mathématique de Pierre?*

L'espérance mathématique de Pierre est proportionnelle à la longueur de l'aiguille et indépendante de sa forme, droite ou courbe.

Il n'en est pas de même de la probabilité de la rencontre, et ce problème, comme le précédent, montre la différence entre le calcul de la probabilité et celui de l'espérance mathématique.

Une aiguille est droite; on la courbe; la probabilité de la rencontre est changée: l'aiguille ne pouvait, si elle est plus petite que la distance des deux parallèles, procurer deux rencontres à la fois; elle le peut quand, sans changer sa longueur, on l'a pliée en courbe. On peut même être certain, si la courbe est fermée, qu'une première rencontre en rend une seconde nécessaire.

Les conditions du problème sont donc changées. L'espérance mathématique de celui qui doit recevoir 1^{fr} par rencontre reste la même. Chaque élément de l'aiguille donne, en effet, une espérance indépendante des autres éléments qui lui sont attachés. Celui qui doit recevoir 1^{fr} par rencontre (c'est la même chose que 1^{fr} s'il y a rencontre, quand une seule rencontre est possible) peut vendre à des acheteurs différents les droits résultant pour lui de chaque élément de l'aiguille, et la valeur de ces droits reste la même, soit que l'aiguille soit droite ou courbe.

Soient a la distance de deux parallèles, l la longueur de l'aiguille, plus petite que a. Si la probabilité d'une rencontre est p, p est aussi l'espérance mathématique de Pierre. Remplaçons l'aiguille par un cercle de même longueur. Le rayon R de ce cercle sera $\frac{l}{2\pi}\cdot$ La probabilité pour que le cercle rencontre une des parallèles est $\frac{2R}{a}\cdot$ Si l'on partage, en effet, la distance a des deux parallèles entre lesquelles tombe le centre du cercle en n parties, il y a chance égale pour qu'il tombe sur chacune d'elles, et il faut, pour qu'il y ait rencontre, qu'il soit à une distance moindre que R de l'une ou l'autre des deux lignes. Le nombre des divisions favorables est $\frac{2Rn}{a}$; leur nombre total est n, et la probabilité de la rencontre est, par conséquent,

$$\frac{2R}{a}.$$

Si le cercle rencontre une des parallèles, il la rencontre deux fois; l'espérance mathématique est donc

$$\frac{4R}{a} = \frac{2l}{a\pi}.$$

Telle est la probabilité pour la rencontre de l'aiguille.

L'ingénieuse substitution d'un cercle à une aiguille rectiligne est due à M. Émile Barbier.

44. Problème XLVII. — *Pierre a trois pièces de 5^{fr}, Paul en a deux; ils conviennent que chacun jettera ses pièces. Celui qui obtiendra le plus grand nombre de faces prendra les cinq pièces. Le jeu est-il équitable?*

Pierre expose 15^{fr} et Paul 10^{fr} seulement; mais Pierre a plus grande chance de gagner. La compensation est-elle exacte?

Dans ce cas, comme dans un grand nombre d'autres, la méthode la plus simple pour calculer l'espérance mathématique est de chercher la probabilité de chacun des cas possibles.

Pierre peut, avec ses trois pièces, amener 0, 1, 2 ou 3 fois face. Les probabilités sont $\frac{1}{8}$, $\frac{3}{8}$, $\frac{3}{8}$ et $\frac{1}{8}$.

Paul peut avoir 0, 1 ou 2 fois face; les probabilités sont $\frac{1}{4}$, $\frac{1}{2}$ et $\frac{1}{4}$.

Si Pierre n'a pas face une seule fois, sa probabilité de gagner est 0, celle de perdre est $\frac{3}{4}$.

S'il a face une fois, sa probabilité de gagner est $\frac{1}{4}$, celle de perdre $\frac{1}{4}$.

S'il a face deux fois, sa probabilité de gagner est $\frac{3}{4}$, celle de perdre 0.

S'il a face trois fois, sa probabilité de gagner est 1.

La probabilité de Pierre est donc

$$\left(\frac{1}{8}\times 0+\frac{3}{8}\times\frac{1}{4}+\frac{3}{8}\times\frac{3}{4}+\frac{1}{8}\times 1\right)=\frac{16}{32}=\frac{1}{2},$$

celle de Paul est

$$\left(\frac{1}{8}\times\frac{3}{4}+\frac{3}{8}\times\frac{1}{4}+\frac{3}{8}\times 0+\frac{1}{8}\times 0\right)=\frac{6}{32}=\frac{3}{16}.$$

La somme des deux probabilités n'est pas égale à l'unité, parce que la partie peut être nulle; mais, s'il est convenu qu'on recommencera jusqu'à ce qu'un des joueurs ait gagné, les probabilités $\frac{1}{2}$ et $\frac{3}{16}$ représentent deux fractions de même dénominateur ayant pour numérateur les nombres de cas favorables, dont le rapport reste constant quand le nombre des parties augmente. Il faut donc, pour avoir les deux probabilités, partager l'unité en deux parties qui soient dans le rapport de $\frac{1}{2}$ à $\frac{3}{16}$; elles sont

$$\frac{8}{11} \quad \text{et} \quad \frac{3}{11}.$$

L'espérance mathématique de Pierre est

$$25\times\frac{8}{11}=\frac{200}{11};$$

elle est plus grande que sa mise.

L'espérance mathématique de Paul est

$$25 \times \frac{3}{11} = \frac{75}{11};$$

elle est plus petite que sa mise.

Le jeu est avantageux à Pierre.

45. Problème XLVIII. — *Un nombre n de joueurs ayant déposé chacun 1^{fr} jettent un nombre μ de dés. L'enjeu total appartiendra à celui qui amènera la plus grande somme de points, ou sera partagé entre ceux qui auront amené le même nombre de points, plus grand que celui des autres joueurs.*

Pierre joue le premier, il amène k points. Quelle est, avant que les adversaires aient joué, son espérance mathématique?

Les n^{fr} qui forment l'enjeu total appartiendront à Pierre si aucun des adversaires n'amène un point égal ou supérieur à k; ils seront partagés entre $m+1$ joueurs si, aucun n'amenant un point supérieur à k, m des adversaires amènent k.

La probabilité d'amener avec μ dés un point donné est connue par le Tableau donné (19), qu'il serait aisé d'étendre au cas d'un plus grand nombre de dés. Nommons q_k la probabilité d'un point supérieur à k et p_k celle d'un point précisément égal à k; la probabilité d'un point inférieur à k sera

$$r_k = 1 - p_k - q_k.$$

Pour que Pierre ait part à l'enjeu, il faut qu'aucun des points ne soit supérieur à k; la probabilité pour qu'il en soit ainsi est $(1-q_k)^{n-1}$. Soient p'_k la probabilité, dans cette hypothèse, d'amener le point k et r'_k celle d'amener un point inférieur.

La probabilité pour que Pierre partage avec i de ses adversaires est

$$\frac{1.2.3.\ldots(n-1)}{1.2.3\ldots i.1.2.3\ldots(n-1-i)} p'^{i}_k r'^{n-1-i}_k.$$

L'espérance mathématique de Pierre est

$$(1-q_k)^{n-1} n \left[r'^{n-1}_k + \frac{1}{2}(n-1) r'^{n-1}_k p'_k \right.$$
$$\left. + \frac{1}{3}\frac{(n-1)(n-2)}{1.2} r'^{n-2}_k p'^{2}_k + \ldots + \frac{1}{n} p'^{n-1}_k \right],$$

égale évidemment à

$$\frac{(1-q_k)^{n-1}}{p'_k}[(r'_k+p'_k)^n - r'^n_k],$$

et, à cause de $r'_k + p'_k = 1$, puisque r'_k et p'_k sont les probabilités de deux événements contraires,

$$(1) \qquad \frac{(1-q_k)^{n-1}}{p'_k}(1 - r'^n_k);$$

il reste à calculer r'_k et p'_k.

Le rapport de r'_k à p'_k est le même que celui de r_k à p_k. L'hypothèse faite qu'aucun numéro supérieur à k n'est sorti diminue en effet le nombre des cas possibles, sans changer celui des cas favorables à l'arrivée d'un nombre de points égal ou inférieur à k. Les numérateurs de r'_k et de p'_k sont, avec un même dénominateur devenu plus petit, ceux de r_k et de p_k. On a donc

$$\frac{r'_k}{p'_k} = \frac{r_k}{p_k}$$

et, puisque $r'_k + p'_k = 1$,

$$r'_k = \frac{r_k}{r_k + p_k},$$

$$p'_k = \frac{p_k}{r_k + p_k}.$$

46. Un problème intéressant peut être résolu :

Pierre ayant obtenu le point k, quel est le nombre d'adversaires le plus favorable à ses intérêts, c'est-à-dire quel est celui qui lui laisse la plus grande espérance mathématique?

En accroissant le nombre des joueurs, on accroît la somme à partager, mais on diminue la chance d'y avoir part et celle de l'avoir tout entière. Il faut déterminer n de manière à rendre (1) maximum.

En égalant à zéro la dérivée par rapport à n, on obtient

$$(2) \qquad r'^n_k = \frac{l(1-q_k)}{l r'_k + l(1-q_k)}.$$

Cette équation fera connaître la valeur de n.

Si Pierre a amené le point le plus élevé possible, on a $q_k = 0$;

r'^{n}_{k} est nul et n infini. Pierre, dans ce cas, certain d'avoir une part, a avantage à accroître sans limite le nombre de ses adversaires.

L'espérance mathématique de Pierre ne grandit pas indéfiniment. La formule (1), quand on y suppose $q_k = 0$ et n infini, se réduit à

$$\frac{1}{p'_k}.$$

L'enjeu total n devant être partagé entre tous ceux qui auront le point maximum, pour que chaque part soit $\frac{1}{p'_k}$, il faut que le nombre des partageants soit np'_k. p'_k, quand k est le plus grand des points possibles, ne diffère pas de p_k. np_k est le nombre probable de ceux d'entre les n adversaires qui amèneront le point maximum dont la probabilité est p_k.

Le Tableau suivant donne les résultats relatifs au cas où l'on joue avec trois dés. Pierre ayant amené au premier coup l'un des points inférieurs au maximum 18, la première colonne donne le point obtenu, les deux suivantes donnent les probabilités désignées dans la formule par p_k et q_k; le nombre d'adversaires le plus avantageux pour chaque valeur de k est inscrit dans la quatrième colonne; la cinquième donne la valeur de $(1 - q_k)^n$, probabilité pour que Pierre ait une part.

k.	q_k.	p_k.	n.	$(1 - q_k)^n$.
17	$\frac{1}{216}$	$\frac{3}{216}$	98,15600	0,634140
16	$\frac{4}{216}$	$\frac{6}{216}$	31,4127	0,555902
15	$\frac{10}{216}$	$\frac{10}{216}$	13,4234	0,529248
14	$\frac{20}{216}$	$\frac{15}{216}$	6,5174	0,530862
13	$\frac{35}{216}$	$\frac{21}{216}$	3,2913	0,558870
12	$\frac{56}{216}$	$\frac{25}{216}$	1,64045	0,611215
10	$\frac{81}{216}$	$\frac{27}{216}$	»	»
9	$\frac{108}{216}$	$\frac{27}{216}$	»	»

Si Pierre ayant lancé six dés a obtenu cinq 6 et un 5, on aura

$$q_k = \frac{1}{6^6},$$

$$p_k = \frac{6}{6^6}.$$

Le nombre d'adversaires le plus avantageux est 15 144 et la probabilité de recevoir une part, dans ce cas,

$$0,723035.$$

47. Lorsque les valeurs probables de plusieurs grandeurs sont connues, on ne connaît pas pour cela la valeur probable d'une fonction donnée de ces grandeurs.

La valeur probable d'une grandeur inconnue a est, par définition, l'espérance mathématique de celui qui devrait recevoir une somme égale à a.

Si a peut prendre, selon les décisions du hasard, les valeurs α_1, α_2, ..., α_n, et que, pour chacune d'elles, les probabilités soient p_1, p_2, ..., p_n, la valeur probable de a est

$$p_1\alpha_1 + p_2\alpha_2 + \ldots + p_n\alpha_n.$$

La valeur probable de a^2 est

$$p_1\alpha_1^2 + p_2\alpha_2^2 + \ldots + p_n\alpha_n^2;$$

elle est très différente du carré de la valeur probable de a.

Si deux grandeurs a et b ont pour valeurs probables A et B, la valeur probable de $\frac{a}{b}$ n'est pas $\frac{A}{B}$. On peut même remarquer que, si o est une des valeurs *possibles* de b, si peu probable qu'elle soit, la valeur probable de $\frac{a}{b}$ est infinie.

48. Deux cas importants sont à noter.

La valeur probable d'une somme est la somme des valeurs probables des parties de la somme.

C'est la conséquence immédiate de la définition.

La valeur probable d'un produit, *quand les facteurs sont indépendants*, est le produit des valeurs probables des facteurs.

Soient a et b deux grandeurs. La première peut recevoir les valeurs $\alpha_1, \alpha_2, \ldots, \alpha_n$, dont les probabilités sont $p_1, p_2, \ldots, p_n$. La seconde peut avoir les valeurs $\beta_1, \beta_2, \ldots, \beta_m$, dont les probabilités sont $q_1, q_2, \ldots, q_m$. Les valeurs probables sont

$$p_1\alpha_1 + p_2\alpha_2 + \ldots + p_n\alpha_n,$$
$$q_1\beta_1 + q_2\beta_2 + \ldots + q_m\beta_m;$$

leur produit est la somme des termes tels que

$$p_i q_{i'} \alpha_i \beta_{i'},$$

qui représentent les produits de toutes les valeurs possibles $\alpha_i\beta_{i'}$ du produit ab par leurs probabilités $p_i q_{i'}$.

Si l'arrivée de l'événement a exerçait une influence sur celle de b, la probabilité de l'association de α_i avec $\beta_{i'}$ ne serait pas (23) le produit $p_i q_{i'}$.

Si, par exemple, b est égal à a^2, les valeurs possibles de a étant $\alpha_1, \alpha_2, \ldots, \alpha_n$, celles de b sont $\alpha_1^2, \alpha_2^2, \ldots, \alpha_n^2$. Si a est égal à α_i, il est certain que b sera α_i^2.

49. *Paradoxe de Saint-Pétersbourg.* — On a opposé les indications du bon sens aux décisions de la théorie et pour condamner les principes allégué leurs conséquences.

La discussion, jusqu'ici, n'a pas dissipé la méprise. La théorie, pourtant, est irréprochable; il n'est pas juste de lui opposer l'absurdité de ses conseils : elle n'en donne pas. Il est déraisonnable d'exposer au jeu une forte somme, indélicat d'accepter un risque qui peut rendre insolvable. Le Calcul des probabilités doit ignorer ces sages appréciations.

On joue, c'est l'hypothèse. A-t-on tort ou raison? La question n'est pas posée.

On cherche les conditions équitables du jeu sans se demander si elles sont raisonnables, ni établir aucune relation entre cette question, que l'on ne veut pas aborder, et le problème à résoudre.

Le parti raisonnable, si les risques sont grands, est de ne pas jouer.

On peut, en acceptant des conventions strictement équitables, faire un acte de folie ou commettre une escroquerie. La remarque, pour être incontestée, ne change pas la théorie du jeu équitable. La confusion est analogue à celle que faisait Poisson (37), en portant une espérance mathématique dans l'inventaire d'une fortune.

Pierre, pour toute fortune, possède 100000^{fr}, il veut avoir une chance de gagner 100 millions.

Rien n'est plus facile, répond sans s'émouvoir le géomètre qu'il consulte. Si le jeu est équitable, vous aurez 999 chances sur 1000 de perdre vos 100000^{fr}.

La réponse doit s'arrêter là. Si Pierre trouve 999 personnes désireuses comme lui de gagner 100 millions et résignées comme lui à la presque certitude de perdre 100000^{fr}, ils organiseront une loterie de 1000 billets à 100000^{fr} le billet.

Le jeu sera équitable, la théorie le déclare, le bon sens le démontre, et Pierre, cependant, sera très justement interdit comme insensé.

Le problème de Saint-Pétersbourg est devenu célèbre par l'ingénieuse combinaison des conventions faites pour dissimuler l'énormité des mises, ou, pour être plus exact, des engagements.

Pierre et Paul jouent aux conditions suivantes :

Pierre jette une pièce de monnaie. Si elle montre face, il donnera 1^{fr} à Paul. Si elle montre pile, il jettera la pièce de nouveau. Si face arrive au second coup il donnera 2^{fr} à Paul. La pièce sera jetée jusqu'à la première arrivée de face qui termine la partie. S'il a fallu la jeter n fois, Paul recevra 2^{n-1}^{fr}.

Quelle est l'espérance mathématique de Paul?

Elle est infinie.

Paul, en effet, recevra, selon le nombre des coups qui seront joués, une somme égale à l'un des termes de la suite illimitée

$$1,\quad 2,\quad 4,\ 8,\quad \ldots,\quad 2^n,\quad \ldots.$$

Les probabilités pour lui de recevoir ces différentes sommes sont

$$\frac{1}{2},\quad \frac{1}{4},\quad \frac{1}{8},\quad \ldots,\quad \frac{1}{2^{n+1}},\quad \ldots.$$

L'espérance mathématique de Paul est, par conséquent,

$$1\times\frac{1}{2}+2\times\frac{1}{4}+4\times\frac{1}{8}+\ldots+2^n\times\frac{1}{2^{n+1}}+\ldots;$$

elle se compose d'un nombre infini de termes qui sont tous égaux à $\frac{1}{2}$.

Quel que soit le prix dont il paye la promesse qui lui est faite, le marché sera avantageux à Paul.

Qui voudrait cependant risquer 100^{fr} à un tel jeu?

50. Les réponses proposees au prétendu paradoxe sont nombreuses. D'Alembert écrivait à Lagrange : « Votre Mémoire sur les jeux me fait désirer beaucoup que vous nous donniez une solution du problème de Pétersbourg, qui me paraît impossible en admettant les principes connus. »

Les conditions du jeu impliquent contradiction, ont dit Condorcet et Poisson. Pierre prend des engagements qu'il ne peut tenir. Si face ne se présente qu'au centième coup, le gain de Paul représentera une masse d'or plus grosse que le soleil. Pierre le trompe en la lui promettant.

L'observation est juste, mais n'éclaircit rien. Si l'on joue des centimes au lieu de francs, des grains de sable au lieu de centimes, des molécules d'hydrogène au lieu de grains de sable, la crainte d'être insolvable peut diminuer sans limite. La théorie ne doit pas faire la différence. Elle ne suppose pas non plus qu'avant chaque jet de la pièce on dépose la mise. Quelle que soit la dette de Pierre, la plume peut l'écrire, on réglera les comptes sur le papier; la théorie triomphera s'ils confirment ses prescriptions. Le hasard très probablement, on peut dire très certainement, finira par favoriser Paul. Quel que soit le prix dont il paye, à chaque partie, la promesse de Pierre, le jeu, s'il est tenace, l'enrichira sans limite. Pierre, insolvable ou non, lui devra une somme immense.

51. Si une machine pouvait jeter 100000 pièces par seconde et enregistrer les résultats, Paul, en payant 1000^{fr} par partie, s'endetterait sans doute d'une centaine de millions par seconde; il n'en ferait pas moins, après quelques millions de milliards de siècles, un bénéfice colossal. Les conditions du jeu le favorisent, et la théorie a raison.

On peut, sans calculs, rendre l'assertion évidente :

Supposons, pour ne pas aborder de trop grands nombres, que Pierre s'engage à faire 1 milliard de parties.

Cherchons quelle somme Paul peut espérer recevoir, sans présumer pour lui le hasard bénévole.

Sur 1 milliard de parties, il faut s'attendre à en voir 500 millions pour lesquelles Paul ne recevra que 1^{fr}. On ne peut s'étonner d'amener face une fois sur deux. Si le nombre est moindre, Paul aura du bonheur; nous ne voulons pas lui en supposer.

La seconde moitié des parties jouées commencent toutes par l'arrivée de pile. Le second coup amènera face *vraisemblablement* 250 millions de fois : une fois sur deux. Paul, pour chacune de ces parties recevra 2^{fr}; il doit, de ce chef, espérer 500 millions.

Les 250 millions de parties restantes commencent par deux fois pile; pour la moitié d'entre elles, on doit s'y attendre, face arrivera au troisième coup, Paul recevra 4^{fr}; c'est encore 500 millions qu'il peut raisonnablement espérer. On peut renouveler trente fois le même raisonnement et voir clairement que l'ensemble des parties doit *vraisemblablement* rapporter à Paul, d'après les conditions convenues, une quinzaine de milliards. Le hasard, évidemment, peut le favoriser et accroître la somme, ou le maltraiter et la diminuer; mais, en donnant 15^{fr} par partie, il a des chances sérieuses de ne rien perdre.

Si Pierre doit faire 100 parties seulement, les chances sont différentes; Paul, en donnant 15^{fr} par partie, aurait grande chance de se trouver en perte. Les conditions du jeu lui seraient toujours avantageuses; mais son avantage résulterait des gros bénéfices possibles, dont la chance est petite.

Si, au contraire, au lieu de 1 milliard de parties, on en devait faire 1000 milliards, Paul, au lieu de 15^{fr}, pourrait donner 20^{fr} par partie, avec chance sérieuse de recouvrer 20000 milliards, sans compter, bien entendu, pour rien dans cette appréciation sommaire la possibilité de gagner des sommes immenses, dont le calcul exact a dû tenir compte.

52. La réponse la plus singulière faite au prétendu paradoxe est celle de Daniel Bernoulli qui, le premier, a tiré de l'oubli cette question autrefois proposée par son cousin Nicolas.

100 millions, suivant Daniel Bernoulli, ajoutés à une fortune déjà acquise de 100 millions, ne suffisent pas pour la doubler. Quels avantages nouveaux peuvent-ils procurer.

Il substitue, en conséquence, à l'espérance mathématique l'espérance morale, dans le calcul de laquelle une fortune dépend non du nombre d'écus dont elle se compose mais des satisfactions qu'elle procure

Le problème étant ainsi posé, Bernoulli a l'audace de le résoudre. La solution est simple. Un accroissement dx ajouté à une fortune x vaut $\frac{dx}{x}$. Celui dont la fortune était a et devient b gagne un avantage mesuré par

$$\int_a^b \frac{dx}{x} = lb - la.$$

Jamais compte n'a été ni ne sera réglé de la sorte; mais, grâce à d'illustres approbations, la théorie de l'espérance morale n'a pas moins contribué à la célébrité de Daniel Bernoulli que ses admirables travaux de Physique.

53. Buffon a accru par son éloquence l'importance, je veux dire le retentissement de l'idée de Bernoulli.

La théorie de l'espérance morale est devenue classique, jamais le mot ne put être plus exactement employé : on l'étudie, on l'enseigne, on la développe dans des livres justement célèbres. Le succès s'arrête là, on n'en a jamais fait et n'en pourra faire aucun usage.

L'importance d'une somme d'argent diminue avec la fortune de celui qui la reçoit.

« L'avare, dit Buffon, est comme le mathématicien : tous deux estiment l'argent par sa quantité numérique. L'homme sensé n'en considère ni la masse ni le nombre. Il n'y voit que les avantages qu'il peut en tirer. Il raisonne mieux que le mathématicien. L'écu que le pauvre a mis à part pour payer un impôt de nécessité et l'écu qui complète les sacs d'un financier n'ont pour l'avare et le mathématicien que la même valeur. Celui-ci les comptera par unités égales, l'autre se les appropriera avec un plaisir égal, au lieu que l'homme sensé comptera l'écu du pauvre pour un louis et l'écu du financier pour un liard. »

Un commentateur, qui n'est pas sans mérite, Quételet, a ajouté, pour faire mieux comprendre une théorie qu'il expose et approuve :

« Ainsi 1000^{fr} pour celui qui ne possède que 2000^{fr} ont la même importance que 500000^{fr} pour celui qui possède 1 million. »

Si, par ce fier dédain de la fortune lorsque le nécessaire est assuré, on échappe au reproche d'avarice, c'est pour en mériter un plus grave. Celui qui possédant un million en acquiert un second changera fort peu, pas du tout peut-être, les habitudes de sa vie.

Est-ce là, pour qui n'est pas avare, le seul fruit de la richesse?

Si l'homme sensé dont parle Buffon n'est pas un cynique égoïste, il pourra, sans thésauriser, faire bon usage des millions qu'on lui suppose. On pourra les doubler, les décupler et les doubler encore, sans ralentir la progression constante du bien qu'il peut faire. N'a-t-il pas une famille à enrichir, des misères à soulager, de grandes œuvres à créer ou à faire naître? Il évitera, s'il est sage, de jouer gros jeu, même à des conditions équitables; mais, s'il ne porte pas, si riche qu'il soit, 100000^{fr} par jour à la roulette, la crainte de la perte l'arrêtera beaucoup plus que le mépris du gain.

CHAPITRE IV.

THEORÈME DE JACQUES BERNOULLI

Hoc igitur est illud problema, quod evulgandum hoc loco proposui, postquam jam per vicennium pressi, et cujus tum novitas, tum summa utilitas cum pari conjuncta difficultate omnibus reliquis hujus doctrinæ capitibus pondus et pretium superaddere potest

JACOBUS BERNOULLI

53. Régularité observée des résultats du hasard — 54. Probabilité des épreuves répétées — 55. Evénements dont la probabilité est maximum. — 56. Valeur approchée du produit 1 2.3. .n. — 57. Probabilité maxima dans une série d'épreuves. — 58. Probabilité d'un événement peu différent du plus probable. — 59. Fiction d'un écart représenté par une variable continue — 60. Première vérification. — 61 Seconde vérification. — 62. Calcul exact de la valeur probable du carre de l'écart; elle ne diffère pas de la valeur approchée. — 63. Troisième vérification. — 64. Calcul exact de la valeur probable de l'écart, elle n'est pas égale à la valeur approchée. — 65. La probabilite pour que l'écart soit inférieur à une limite donnée est donnée par une intégrale que l'on a réduite en Table. — 66 La probabilité d'un écart absolu inférieur à une limite fixe tend vers zéro; quand le nombre des épreuves augmente, c'est l'écart relatif qui tend vers zéro. — 67. La probabilité d'un écart α, sur μ épreuves, dépend de $\frac{\alpha}{\sqrt{\mu}}$. exemples numériques. — 68 Écart probable et écart moyen; leur rapport. — 69 Représentation du nombre probable d'arrivées en introduisant un facteur dont la valeur détermine la confiance méritée par la formule. — 70. Ce qu'on doit entendre par jouer plus ou moins gros jeu, expression de laquelle dépendent les chances de perte sur un grand nombre de parties — 71 Application du théorème de Bernoulli aux chances électorales. — 72. Différences entre les conditions réelles et les données du problème précédent. — 73. Le théorème de Bernoulli suppose la probabilité d'un événement invariable, il suppose aussi que cette probabilité ait une valeur objective; remarques sur cette question. — 74. Exemple d'une serie d'épreuves faites avec probabilité variable.

53. Le hasard corrige le hasard. Une vague expérience révèle la justesse de cette maxime à ceux mêmes qui en ignorent la rigueur. Le mot *rigueur* n'est pas exagéré. Les résultats de l'action libre du hasard sont prédits avec certitude, sans gêner en rien ses caprices.

La certitude n'est pas celle d'un théorème de Géométrie. Il faut l'accepter dans le sens qu'indique Jacques Bernoulli, véritable inventeur de l'un des plus admirables résultats de la Science mathématique : *In usu vitæ civili ubi moraliter certum pro absolute certo habetur.*

Vous sortez sans abri par un violent orage, vous serez mouillé, cela est *certain*.

La certitude du théorème de Bernoulli est de même sorte. La ressemblance va à l'identité. La même objection peut s'appliquer avec autant, osons dire avec aussi peu de raison dans les deux cas.

La pluie, dites-vous, me mouillera, qu'en savez-vous? Chaque goutte est dirigée par le hasard, aucune n'a de destination; rien ne prouvant pour aucune d'entre elles qu'elle ira s'abattre sur le promeneur, de quel droit affirmer de l'ensemble ce qui est incertain pour chacune?

Sans que l'assertion soit certaine à la manière du carré de l'hypoténuse, on peut la produire avec entière confiance. Il y a plus : si deux promeneurs sortent ensemble et marchent sans se quitter sous la même pluie, non seulement ils seront mouillés tous deux, mais ils le seront également. Si l'un d'eux accusait le hasard de lui avoir donné la plus grosse part, il ne rencontrerait pas plus de créance que s'il affirmait n'avoir rien reçu.

Les événements fortuits ressemblent aux gouttes de pluie. Pourvu qu'ils soient assez nombreux, le hasard les distribue équitablement entre tous les cas possibles, sans en favoriser aucun. Tel est le théorème que nous allons démontrer avec précision et dont l'importance même justifiera plusieurs démonstrations.

La première repose sur plusieurs propositions, dont chacune par elle-même est de grande importance.

54. Problème XLVII. — *La probabilité d'un événement est p, celle de l'événement contraire est q; on fait μ épreuves dans les mêmes conditions. Quelle est la probabilité pour que le premier événement se présente n fois, le second $\mu - n$ fois?*

Si l'ordre dans lequel les événements doivent se succéder était assigné, la probabilité demandée serait (24) le produit de n facteurs égaux à p et de $\mu - n$ égaux à q

$$p^n q^{\mu-n}. \tag{1}$$

L'ordre restant indéterminé, l'événement peut être décomposé en autant d'autres, de probabilités égales, qu'il y a de combinaisons possibles de μ objets dont n sont égaux à A et $\mu - n$ à B. Ce nombre est

$$\frac{1.2.3\ldots\mu}{1.2.3\ldots n.1.2.3\ldots\mu - n}.$$

La probabilité demandée est donc

$$\frac{1.2.3\ldots\mu}{1.2.3\ldots n.1.2.3\ldots\mu - n} p^n q^{\mu-n}; \tag{2}$$

c'est le terme général du développement de $(p+q)^\mu$.

Si l'on écrit

$$(p+q)^\mu = p^\mu + \mu p^{\mu-1} q + \frac{\mu.\mu-1}{2} p^{\mu-2} q^2 + \ldots + q^\mu,$$

le premier terme représente la probabilité pour que l'événement dont la probabilité est q ne se présente pas une seule fois; le deuxième, la probabilité pour que cet événement arrive une fois; le troisième, pour qu'il arrive deux fois, etc.

La somme des k premiers termes est la probabilité pour que cet événement, dont la probabilité est q, arrive au plus $k - 1$ fois sur μ épreuves.

La somme de tous les termes est la probabilité pour que l'événement arrive au plus μ fois, c'est la certitude, et la somme des termes est égale en effet à l'unité, puisque $p + q = 1$.

55. Problème XLVIII. — *La probabilité d'un événement est p, celle de l'événement contraire est q; on fait μ épreuves dans les mêmes conditions. Quel est, pour chacun des deux événements, le nombre d'arrivées le plus probable?*

Il faut trouver pour quelle valeur de n, μ étant donné, l'expression (2)

$$\frac{1.2.3\ldots\mu}{1.2.3\ldots n\; 1.2.3\ldots(\mu-n)} p^n q^{\mu-n}$$

acquiert la plus grande valeur.

Si l'on range les expressions (2), c'est-à-dire les termes du développement de $(p+q)^\mu$, suivant l'ordre des valeurs décroissantes de n, le rapport d'un terme au précédent est

$$\frac{n+1}{\mu-n}\,\frac{q}{p}.$$

La valeur cherchée de n doit rendre ce rapport plus grand que l'unité, et le rapport suivant, obtenu en changeant n en $n-1$, doit être plus petit que l'unité.

Nous avons donc, pour déterminer n,

$$\frac{n+1}{\mu-n}\,\frac{q}{p} > 1,$$

$$\frac{n}{\mu-n+1}\,\frac{q}{p} < 1,$$

$$(n+1)q > \mu p - np,$$

$$nq < \mu p - np + p$$

et, à cause de $p+q=1$,

$$n > \mu p - q.$$

$$n < \mu p + p.$$

Le nombre entier n est compris, d'après ces formules, entre deux limites dont la différence $p+q$ est égale à l'unité.

Il est donc déterminé.

On peut dire, en négligeant la fraction, que μp est le nombre le plus probable d'arrivées pour l'événement dont la probabilité est p; le nombre d'arrivées le plus probable pour l'événement dont la probabilité est q est $\mu - \mu p = \mu q$.

La combinaison dont la probabilité est la plus grande est donc celle dans laquelle les événements se produisent en nombre proportionnel à leurs probabilités.

56. Pour rendre possible l'application des expressions trouvées pour la probabilité, il faut transformer les produits qui y figurent. Le nombre des facteurs, quand les épreuves sont nombreuses, rendrait les calculs interminables.

La transformation repose sur la célèbre formule de Stirling

$$1.2.3\ldots n = e^{-n} n^n \sqrt{2\pi n}.$$

Les deux membres ne sont pas égaux, leur différence grandit sans limite quand n augmente; mais leur rapport tend vers l'unité, et cela suffit. Nous devons d'abord démontrer cette formule.

Posons

$$(1) \qquad 1.2.3\ldots n = n^n \varphi(n),$$

et cherchons, lorsque n est grand, une valeur approchée de $\varphi(n)$.

On a rigoureusement, en changeant dans (1) n en $n+1$ et divisant la seconde équation par la première,

$$\frac{\varphi(n+1)}{\varphi(n)} \frac{(n+1)^{n+1}}{n^n} = n+1;$$

par conséquent,

$$\varphi(n+1) = \varphi(n)\left(1+\frac{1}{n}\right)^{-n}.$$

$\left(1+\frac{1}{n}\right)^{-n}$, lorsque n est grand, diffère peu de $\frac{1}{e}$.

Si donc on pose

$$\varphi(n) = e^{-n}\psi(n),$$

la fonction $\psi(n)$, par le changement de n en $n+1$, se multipliera par un facteur peu différent de l'unité. Nous pourrons écrire

$$1.2.3\ldots n = n^n e^{-n}\psi(n),$$

avec la certitude que $\psi(n)$ varie lentement quand n augmente.

$\psi(n)$ varie lentement, mais il n'est pas constant et ne tend pas à le devenir. On a, en effet,

$$\frac{\psi(n+1)}{\psi(n)} = e\left(1+\frac{1}{n}\right)^{-n}.$$

Posons, pour évaluer le second membre,

$$e\left(1+\frac{1}{n}\right)^{-n} = z;$$

on en déduit

$$lz = 1 - nl\left(1 + \frac{1}{n}\right)$$

et, en remplaçant $l\left(1 + \frac{1}{n}\right)$ par les deux premiers termes de son développement en série, c'est-à-dire en négligeant $\frac{1}{n^3}$,

$$lz = 1 - n\left(\frac{1}{n} - \frac{1}{2n^2}\right) = \frac{1}{2n}.$$

Par conséquent

$$z = e^{\frac{1}{2n}}$$

et, en remplaçant $e^{\frac{1}{2n}}$ par les deux premiers termes de son développement,

$$z = 1 + \frac{1}{2n}.$$

Nous pouvons écrire, en négligeant $\frac{1}{n^2}$,

$$\frac{\psi(n+1)}{\psi(n)} = 1 + \frac{1}{2n}.$$

On a au même degré d'approximation

$$\sqrt{\frac{n+1}{n}} = \sqrt{1 + \frac{1}{n}} = 1 + \frac{1}{2n};$$

la fonction $\psi(n)$ varie donc, quand on néglige $\frac{1}{n^2}$, suivant la même loi que $\sqrt{n}$.

Si l'on pose

$$\psi(n) = \sqrt{n}\,F(n),$$

par conséquent

$$1.2.3\ldots n = n^n e^{-n}\sqrt{n}\,F(n),$$

la fonction $F(n)$ variera très lentement, le rapport $\frac{F(n+1)}{F(n)}$ se réduisant à l'unité quand on néglige $\frac{1}{n^2}$.

On pourrait continuer le même mode d'approximation pour chercher la valeur approchée de $F(n)$, comme nous avons cherché celles de $\varphi(n)$ et de $\psi(n)$; mais il est permis de s'arrêter, parce que, d'après l'évaluation obtenue pour le rapport $\frac{F(n+1)}{F(n)}$, on peut démontrer que $F(n)$ tend vers une valeur constante par laquelle on peut dès lors le remplacer si n est très grand.

$\frac{F(n+1)}{F(n)}$ se réduisant à l'unité quand on néglige $\frac{1}{n^2}$, on peut poser

$$\frac{F(n+1)}{F(n)} = 1 + \frac{\lambda_1}{n^2}.$$

λ_1 ne grandissant pas indéfiniment avec n, nous en déduirons, en changeant la valeur de n,

$$\frac{F(n+2)}{F(n+1)} = 1 + \frac{\lambda_2}{(n+1)^2},$$
$$\frac{F(n+3)}{F(n+2)} = 1 + \frac{\lambda_3}{(n+2)^2},$$
$$\dots\dots\dots\dots\dots\dots,$$
$$\frac{F(n+p)}{F(n+p-1)} = 1 + \frac{\lambda p}{(n+p-1)^2},$$

$\lambda_1, \lambda_2, \dots, \lambda_p$ n'augmentant pas sans limite avec p. En multipliant les équations précédentes, on a

$$(3) \qquad \frac{F(n+p)}{F(n)} = \left(1 + \frac{\lambda_1}{n^2}\right)\left[1 + \frac{\lambda_2}{(n+1)^2}\right]\dots\left[1 + \frac{\lambda p}{(n+p-1)^2}\right];$$

le second membre, et par conséquent le premier qui lui est égal, tend vers l'unité lorsque, n étant de plus en plus grand, on donne à p une valeur quelconque, grande ou petite. On a, en effet,

$$(4) \qquad l\left(1 + \frac{\lambda}{n^2}\right) = \frac{\lambda}{n^2}(+\theta),$$

θ tendant vers zéro lorsque n augmente. Le logarithme du second membre de (3) est

$$l\left(1 + \frac{\lambda_1}{n^2}\right) + l\left[1 + \frac{\lambda_2}{(n+1)^2}\right] + \dots + l\left[1 + \frac{\lambda p}{(n+p-1)^2}\right]$$

et, à cause de (4),

$$\frac{\lambda_1}{n^2}(1+\theta_2)+\frac{\lambda^2}{(n+1)^2}(1+\theta_3)+\ldots+\frac{\lambda p}{(n+p-1)^2}(1+\theta_p).$$

Cette expression tend vers zéro quand n augmente, parce que la série dont le terme général est $\frac{1}{n^2}$ est convergente: la somme des termes qui, dans cette série, suivent $\frac{1}{n^2}$ est donc très petite, quelque loin qu'on la prolonge, et elle reste très petite quand les termes sont multipliés par des facteurs qui ne grandissent pas sans limite.

Nous pouvons donc poser, en nommant G la limite constante vers laquelle tend $F(n)$,

$$1.2.3\ldots n = \mathrm{G}n^n e^{-n}\sqrt{n}.$$

Le théorème, sous cette forme, était connu de Moivre; Stirling a trouvé la valeur de la constante.

On a, d'après un théorème célèbre de Wallis,

$$\frac{\pi}{2}=\frac{2}{1}\,\frac{2}{3}\,\frac{4}{3}\,\frac{4}{5}\cdots\frac{2n}{2n-1}\,\frac{2n}{2n+1},$$

que l'on peut écrire

$$\frac{\pi}{2}=\frac{2^{4n}(1.2.3\ldots n)^4}{(1.2.3\ldots 2n)^2(2n+1)}.$$

Si, dans cette formule, nous remplaçons $1.2.3\ldots n$ et $1.2.3\ldots 2n$ par leurs valeurs déduites du théorème de Moivre,

$$1.2.3\ldots n = \mathrm{G}e^{-n}n^n\sqrt{n},$$
$$1.2.3\ldots 2n = \mathrm{G}e^{-2n}(2n)^{2n}\sqrt{2n},$$

elle devient

$$\frac{\pi}{2}=\mathrm{G}^2\frac{n}{2(2n+1)}$$

et, lorsque n devient infini,

$$\mathrm{G}=\sqrt{2\pi}.$$

Nous pouvons écrire enfin

$$(5)\qquad 1.2.3\ldots n = e^{-n} n^n \sqrt{2\pi n}.$$

Pour éprouver l'exactitude de cette formule, véritablement indispensable dans les calculs de probabilités, faisons $n = 20$; nous trouverons

$$1.2.3.4.5\ldots 20 = 2432902008176640000,$$
$$e^{-20} 20^{20} \sqrt{40\pi} = 2422786385510400000.$$

Le rapport de ces deux nombres est 1,00417.

57. Les probabilités de deux événements contraires étant p et q, la combinaison la plus probable, sur μ épreuves, est celle dans laquelle le nombre d'arrivées du premier événement est μp et celle du second μq.

Cette probabilité maxima est

$$(6)\qquad \frac{1.2.3\ldots\mu}{1.2.3\ldots\mu p . 1.2.3\ldots\mu q} p^{\mu p} q^{\mu q}.$$

L'application de la formule de Stirling donne pour valeur approchée

$$\frac{e^{-\mu}\mu^{\mu}\sqrt{2\pi\mu}\, p^{\mu p} q^{\mu q}}{e^{-\mu p}(\mu p)^{\mu p}\sqrt{2\pi\mu p}\, e^{-\mu q}(\mu q)^{\mu q}\sqrt{2\pi\mu q}},$$

ce qui se réduit, en ayant égard à la condition $p + q = 1$, à

$$(7)\qquad \frac{1}{\sqrt{2\pi\mu pq}}.$$

Cette probabilité, la plus grande de toutes, contient $\sqrt{\mu}$ en diviseur. Elle tend vers zéro, quand le nombre des épreuves augmente.

58. Cherchons la valeur approchée du terme dans lequel l'exposant de p est $\mu p - h$, en supposant ce terme assez voisin du terme maximum pour que $\frac{h}{\mu}$ et même $\frac{h}{\sqrt{\mu}}$ soient petits.

L'expression exacte est

$$\frac{1.2.3\ldots\mu.p^{\mu p-h}q^{\mu q+h}}{1.2.3\ldots(\mu p-h).1.2.3\ldots(\mu q+h)}. \tag{8}$$

Le théorème de Stirling la réduit à

$$\frac{1}{\sqrt{2\pi\mu pq}}\,\frac{1}{\left(1-\dfrac{h}{\mu p}\right)^{\mu p-h+\frac{1}{2}}\left(1-\dfrac{h}{\mu q}\right)^{\mu q+h+\frac{1}{2}}}.$$

On a, en négligeant $\frac{1}{\mu^3}$,

$$l\left(1-\frac{h}{\mu p}\right)^{\mu p-h+\frac{1}{2}}=\left(\mu p-h+\frac{1}{2}\right)\left(-\frac{h}{\mu p}-\frac{h^2}{2\mu^2p^2}\right),$$

$$l\left(1+\frac{h}{\mu q}\right)^{\mu q+h+\frac{1}{2}}=\left(\mu q+h+\frac{1}{2}\right)\left(\quad\frac{h}{\mu q}-\frac{h^2}{2\mu^2q^2}\right)$$

et, par conséquent, en négligeant $\frac{1}{\mu^2}$,

$$l\left(1-\frac{h}{\mu p}\right)^{\mu p-h+\frac{1}{2}}\left(1+\frac{h}{\mu q}\right)^{\mu q+h+\frac{1}{2}}=\frac{h^2}{2\mu}\left(\frac{1}{p}+\frac{1}{q}\right)+\frac{1}{2}\frac{h}{\mu}\left(\frac{1}{q}-\frac{1}{p}\right).$$

On a

$$\frac{1}{p}+\frac{1}{q}=\frac{1}{pq},$$

et l'on conclut

$$\left(1-\frac{h}{\mu p}\right)^{\mu p-h+\frac{1}{2}}\left(1+\frac{h}{\mu q}\right)^{\mu q+h+\frac{1}{2}}=e^{\frac{h^2}{2\mu pq}}e^{\frac{h}{2\mu}\left(\frac{1}{q}-\frac{1}{p}\right)};$$

le second facteur diffère fort peu de l'unité; on peut le supprimer et remplacer le terme (8) par

$$\frac{1}{\sqrt{2\pi\mu pq}}e^{-\frac{h^2}{2\mu pq}}. \tag{9}$$

La démonstration de cette formule suppose h petit par rapport à μ; on l'étend cependant à toutes les valeurs de h, sans qu'il en résulte, lorsque μ est grand, aucune erreur sur les chiffres.

La fonction

$$e^{-\frac{h^2}{2\mu pq}}$$

décroît, en effet, lorsque h augmente, avec une telle rapidité que sa valeur numérique n'a plus d'influence sur les chiffres conservés. La probabilité dans l'expression de laquelle on l'introduit est elle-même d'une extrême petitesse et, quoique la formule acceptée n'en soit plus la valeur approchée, la substitution de deux quantités négligeables l'une et l'autre est sans inconvénient.

On étend même l'expression approchée de la probabilité à des valeurs de h pour lesquelles cette probabilité n'existe pas; h, en effet, ne peut être plus grand que μp, et on le fait varier sans que cela exerce d'influence sur les chiffres, entre $-\infty$ et ∞.

59. Par une fiction qui rendra les calculs plus faciles, nous remplacerons le nombre entier h par une variable continue, en assignant à un *écart* compris entre z et $z + dz$ la probabilité

$$\frac{1}{\sqrt{2\pi\mu pq}} e^{-\frac{z^2}{2\mu pq}} dz. \tag{10}$$

Ce mot *écart* doit être défini. On considère la valeur μp du nombre d'arrivées de l'événement dont la probabilité est p comme une valeur normale, la plus probable de toutes, et les autres sont définies par leur différence avec celle-là. Cette différence prend le nom d'*écart*.

Si, par exemple, on fait 10000 épreuves à pile ou face et que pile se présente 5021 fois, l'*écart* sera 21. Si l'on jette deux dés 36000 fois et que *sonnez* se présente 995 fois, l'*écart* sera -5.

La substitution de la fonction continue (10) à l'expression qui convient aux écarts entiers, les seuls possibles, est comparable à celle d'une courbe à un polygone.

Supposons, par exemple, $\mu = 1000$, $p = q = \frac{1}{2}$.

La formule (9) devient, pour $h = 40$,

$$\frac{\sqrt{2}}{\sqrt{1000\pi}} e^{-3,2} = 0,0010285;$$

pour $h = 60$,

$$\frac{\sqrt{2}}{\sqrt{1000\pi}} e^{-7,2} = 0,00001883;$$

pour $h = 100$,

$$\frac{\sqrt{2}}{\sqrt{1000\pi}} e^{-20} = 0,0000000052006.$$

L'influence d'un nombre aussi petit dans les formules est insignifiante; on peut le supprimer s'il existe, l'introduire s'il n'existe pas, le remplacer par un autre de même ordre, les chiffres utiles du résultat n'en seront pas modifiés.

60. Quelques épreuves justifieront la confiance accordée à la formule (10).

La probabilité d'une erreur comprise entre z et $z + dz$ étant

$$\frac{1}{\sqrt{2\pi\mu pq}} e^{-\frac{z^2}{2\mu pq}} dz,$$

il est nécessaire que la somme des probabilités de toutes les erreurs possibles représente la certitude; on doit donc avoir

$$\frac{1}{\sqrt{2\pi\mu pq}} \int_{-\infty}^{\infty} e^{-\frac{z^2}{2\mu pq}} dz = 1.$$

Si l'on pose, en effet,

$$\frac{z}{\sqrt{2\mu pq}} = t,$$

cette intégrale devient

$$\frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-t^2} dt,$$

égale à l'unité en vertu d'un théorème très connu.

L'épreuve réussit mieux qu'on n'était en droit de l'espérer; toute formule approchée doit en effet laisser craindre une erreur. Le résultat, ici, est rigoureusement exact.

61. Cherchons, d'après la formule (10), la valeur probable du carré de l'écart, c'est-à-dire l'espérance mathématique de celui qui doit recevoir une somme égale à z^2. Cette espérance est la somme des produits de chaque valeur de z^2 par sa probabilité : elle a pour expression, d'après notre formule,

$$\frac{1}{\sqrt{2\pi\mu pq}}\int_{-\infty}^{\infty} e^{-\frac{z^2}{2\mu pq}} z^2\,dz.$$

Posons

$$\frac{z}{\sqrt{2\mu pq}} = t,$$

la formule devient

$$(11)\qquad \frac{2\mu pq}{\sqrt{\pi}}\int_{-\infty}^{\infty} e^{-t^3} t^2\,dt = \mu pq.$$

62. Nous allons chercher directement la valeur exacte de l'espérance mathématique dont la formule (11) donne la valeur approchée. Posons

$$(p+q)^\mu = p^\mu + A_1 p^{\mu-1} q + A_2 p^{\mu-2} q^2 + \ldots + A_{q\mu} p^{p\mu} q^{q\mu} + \ldots.$$

Les termes de cette formule représentent rigoureusement les probabilités dont (9) donne la valeur approchée.

La probabilité qui correspond à un écart h est

$$A_n p^{\mu-n} q^n,$$

où l'on a

$$n = \mu q - h,$$
$$h = \mu q - n.$$

L'espérance mathématique de celui qui attend une somme égale à h^2 est donc

$$(11)\qquad \Sigma(\mu q - n)^2 A_n p^{\mu-n} q^n,$$

la somme Σ s'étendant à toutes les valeurs de n.

L'expression (11) peut s'écrire

$$(12)\qquad \Sigma\,\mu^2 q^2 A_n p^{\mu-n} q^n - 2\,\Sigma\,\mu q n A_n p^{\mu-n} q^n + \Sigma\, n^2 A_n p^{\mu-n} q^n.$$

On a

$$\Sigma\,\mu^2 q^2 A_n p^{\mu-n} q^n = \mu^2 q^2,$$

puisque l'on multiplie par $\mu^2 q^2$ tous les termes d'une somme égale à l'unité.

L'identité

$$(13)\qquad (p+q)^\mu = \Sigma A_n p^{\mu-n} q^n$$

donne, en égalant les dérivées par rapport à q, préalablement multipliées par q,

$$\mu(p+q)^{\mu-1} q = \Sigma n A_n p^{\mu-n} q^n,$$

et, en remplaçant $(p+q)$ par l'unité dans l'équation précédente, qui est une identité, on en déduit

$$\Sigma n A_n p^{\mu-n} q^n = \mu q$$

et, par conséquent,

$$\Sigma \mu q n A_n p^{\mu-n} q^n = \mu^2 q^2.$$

L'identité (13), différentiée par rapport à q, donne, en multipliant ensuite les deux membres par q,

$$\mu(\mu-1)(p+q)^{\mu-2} q^2 + \mu(p+q)^{\mu-1} q = \Sigma n^2 A_n p^{\mu-n} q^n$$

et, à cause de $p+q=1$,

$$\Sigma n^2 A_n p^{\mu-n} q^n = \mu(\mu-1) q^2 + \mu q.$$

Ces formules réduisent la somme (12) à

$$\mu^2 q^2 - 2\mu^2 q^2 + \mu(\mu-1) q^2 + \mu q = \mu q(1-q) = \mu p q.$$

C'est le résultat obtenu (61). La substitution des valeurs approchées aux valeurs exactes n'a donné aucune erreur.

Il ne peut pas toujours en être ainsi. Faisons une troisième épreuve.

63. Cherchons la valeur probable de l'écart z considéré en grandeur absolue; cela est nécessaire. car sans cela les termes négatifs détruiraient les termes positifs et le résultat serait nul.

La valeur probable de z, d'après la formule adoptée pour représenter la probabilité d'une erreur z, est

$$\frac{2}{\sqrt{2\pi\mu pq}}\int_0^\infty z e^{-\frac{z^2}{2\mu pq}}\,dz.$$

L'intégration s'étend de o à l'infini seulement, parce que nous ne considérons pas les valeurs négatives, dont on tient compte cependant en introduisant le multiplicateur 2.

Posons

$$\frac{z}{\sqrt{2\mu pq}} = t.$$

L'intégrale devient

$$\frac{2\sqrt{2\mu pq}}{\sqrt{\pi}}\int_0^\infty e^{-t^2}t\,dt = \frac{\sqrt{2\mu pq}}{\sqrt{\pi}};$$

telle est l'expression approchée de la valeur probable de z.

64. Calculons exactement la valeur probable de l'écart.

La probabilité d'un écart égal à z est le terme du développement de $(p+q)^\mu$, dans lequel l'exposant de p est $\mu p + z$ et celui de q, $\mu q - z$.

Il faut multiplier ce terme

$$A_z p^{\mu p+z} q^{\mu q-z}$$

par z, afin de former l'espérance mathématique de celui qui attend une somme égale à z. On a, à cause de $p+q=1$,

$$z = q(\mu p + z) - p(\mu q - z);$$

$\mu p + z$ et $\mu q - z$ sont les exposants de p et de q dans le terme de $(p+q)^\mu$. Or, multiplier un terme

$$q^\alpha p^\beta$$

par

$$\beta q - \alpha p = pq\left(\frac{\beta}{p} - \frac{\alpha}{q}\right),$$

c'est prendre la dérivée par rapport à p, en retrancher la dérivée par rapport à q et multiplier la différence par pq.

Nous devons donc prendre les termes de développement de

$(p+q)^\mu$ pour lesquels z est positif, c'est-à-dire

$$p^\mu + \mu p^{\mu-1} q + \frac{\mu(\mu-1)}{1.2} p^{\mu-2} q^2 + \ldots + \frac{1.2.3\ldots\mu}{1.2.3\ldots\mu p . 1\ 2.3\ldots\mu q} p^{\mu p} q^{\mu q},$$

et retrancher la dérivée par rapport à q de la dérivée par rapport à p, pour multiplier ensuite le résultat par pq.

Les termes, on le voit aisément, se détruisent deux à deux : la dérivée par rapport à q du second est égale à la dérivée du premier par rapport à p ; la dérivée du troisième par rapport à q est la dérivée du second par rapport à p, et ainsi de suite ; il ne reste que la dérivée du dernier par rapport à p.

$$\frac{1.2.3\ldots\mu}{1.2.3\ldots\mu p\ 1.2.3\ldots\mu q} p^{\mu p} q^{\mu q}\, \mu q p.$$

C'est le produit par μpq du terme maximum, c'est-à-dire, d'après la valeur approchée (57) de ce terme,

$$\frac{\mu pq}{\sqrt{2\mu\pi pq}} = \frac{\sqrt{\mu pq}}{\sqrt{2\pi}}.$$

Il faut doubler ce résultat, puisque nous n'avons pris que les valeurs positives de z : il s'accorde avec l'expression obtenue (63).

65. Nous pouvons donc accepter avec confiance l'expression (9) pour la probabilité d'un écart égal à z.

La probabilité pour que z soit inférieur à une limite donnée α, c'est-à-dire pour qu'il soit compris entre $-\alpha$ et $+\alpha$, sera

$$\frac{1}{\sqrt{2\mu\pi pq}} \int_{-\alpha}^{\alpha} e^{-\frac{z^2}{2\mu pq}}\, dz$$

ou, en posant $\dfrac{z}{\sqrt{2\mu pq}} = t$,

$$\varpi_\alpha = \frac{2}{\sqrt{\pi}} \int_0^{\frac{\alpha}{\sqrt{2\mu pq}}} e^{-t^2}\, dt. \tag{14}$$

Cette probabilité s'obtiendra dans chaque cas particulier à l'aide de la Table des valeurs de la fonction

$$\frac{2}{\sqrt{\pi}}\int_0^t e^{-t^2}\,dt = \Theta(t),$$

que l'on trouvera à la fin du Volume.

$\Theta(t)$, on peut le voir à l'inspection de la Table, tend rapidement vers l'unité; on a

$$\Theta(3) = 0,9999779,$$
$$\Theta(3,70) = 0,9999998.$$

66. On peut, sans consulter la Table des valeurs de la fonction Θ, déduire de la formule (9) d'importantes conséquences.

L'écart zéro est le plus probable. Il a cependant une probabilité infiniment petite lorsque le nombre des épreuves devient infini.

La formule qui représente la probabilité d'un écart inférieur à α s'annule pour $\alpha = 0$. Il faut remarquer qu'en substituant une variable continue à l'écart, qui nécessairement est entier, nous devons, pour représenter l'écart nul, prendre tout l'intervalle entre $\alpha = 0$ et $\alpha = 1$. Nous n'avons plus alors une probabilité rigoureusement nulle, qui devrait diminuer la confiance inspirée par la formule.

Si l'on assigne à α une valeur déterminée, quelque grande qu'elle soit, la probabilité pour qu'il ne la dépasse pas tend vers zéro lorsque μ augmente sans limite.

On peut regarder comme certain, par conséquent, que, le nombre des épreuves croissant indéfiniment, l'*écart* croîtra sans limite. La probabilité pour qu'il reste au-dessous d'une limite donnée est infiniment petite.

Ce résultat fixe le sens du beau théorème de Bernoulli; l'écart *relatif* est de plus en plus petit, l'écart *absolu* de plus en plus grand.

67. La certitude, quand les épreuves se multiplient, de voir l'écart absolu grandir sans limite peut se démontrer *a priori*, indépendamment de toute formule.

Le résultat le plus probable de μ épreuves successives est l'arrivée μp fois de l'événement dont la probabilité est p. Cette probabilité maximum, quoique plus grande que toutes les autres, tend vers zéro; elle est (57) proportionnelle à $\frac{1}{\sqrt{\mu}}$.

Les autres, qui sont plus petites, tendent *a fortiori* vers zéro; et, si l'on en prend un nombre désigné, quel qu'il soit, leur somme ne peut manquer de tendre vers zéro lorsque, ce nombre restant fixe, μ augmente indéfiniment.

68. Il faut insister, en donnant un exemple, sur cette distinction entre l'écart absolu, qui doit grandir, et l'écart relatif, qui tend vers zéro.

On joue à pile ou face. Combien doit-on faire d'épreuves pour que la probabilité d'obtenir pile ou face, sans préciser lequel, un million de fois au moins de plus que l'autre surpasse 0,01?

Il faut déterminer μ par l'équation

$$\frac{2}{\sqrt{\pi}}\int_0^{\frac{1\,000\,000}{\sqrt{\frac{\mu}{2}}}} e^{-t^2}\,dt = 0,99.$$

On trouve dans la Table de la fonction $\Theta(t)$

$$\Theta(1,83) = 0,99.$$

Nous poserons donc

$$\frac{1\,000\,000}{\sqrt{\frac{\mu}{2}}} = 1,83;$$

on en déduit

$$\mu = 597\,211 \text{ millions.}$$

69. La probabilité ϖ_α d'un écart plus petit que α est, en général,

$$\Theta\left(\frac{\alpha}{\sqrt{2\mu pq}}\right);$$

il dépend du rapport $\frac{\alpha}{\sqrt{\mu}}$ et tend rapidement vers l'unité quand ce rapport augmente. Si l'écart relatif $\frac{\alpha}{\mu}$ reste constant, $\frac{\alpha}{\sqrt{\mu}} = \sqrt{\mu}\,\frac{\alpha}{\mu}$ augmentera sans limite, et la probabilité d'un écart moindre que α se rapprochera de la certitude.

Cherchons, par exemple, le nombre d'épreuves qu'il faut tenter à la roulette pour avoir 99 à parier contre 1 de voir le zéro sortir plus d'une fois sur quarante en moyenne. Le nombre de sorties le plus probable est une sur trente-sept; si le zéro sort moins d'une fois sur quarante, l'écart est négatif et plus grand que

$$\frac{\mu}{37} - \frac{\mu}{40} = \mu(0,002027).$$

Nous avons

$$\Theta(1,65) = 0,98.$$

La probabilité d'un écart plus grand que α, si l'on pose

$$\frac{\alpha}{\sqrt{2\mu pq}} = 1,65,$$

est donc 0,02; le signe étant donné, elle est 0,01.

Nous poserons donc

$$\frac{\alpha}{\mu} = 0,002027,$$

$$\frac{\alpha}{\sqrt{2\mu pq}} = 1,65.$$

En prenant

$$p = \frac{1}{37}, \qquad q = \frac{36}{37},$$

on trouvera

$$\mu = 34\,848.$$

Quelque petit que soit l'écart relatif assigné, on pourra faire un nombre d'épreuves assez grand pour rendre la probabilité de ne pas l'atteindre aussi grande que l'on voudra.

70. Deux écarts remarquables ont reçu des noms qu'il faut connaître.

On a

$$\Theta(0,4769363) = \frac{1}{2}.$$

L'écart α défini par l'équation

$$\frac{\alpha}{\sqrt{2\mu pq}} = 0,476936$$

a donc probabilité égale d'être ou de ne pas être surpassé.

Cet écart

$$0,47693\sqrt{2\mu pq}$$

se nomme l'*écart probable* sur μ épreuves.

L'écart moyen est la valeur probable de l'écart; il a pour valeur (64)

$$\frac{\sqrt{2}}{\sqrt{\pi}}\sqrt{\mu pq} = 0,79789\sqrt{\mu pq}.$$

Le rapport de l'écart probable à l'écart moyen est 0,8463. Ils sont l'un et l'autre proportionnels à la racine carrée du nombre des épreuves.

71. On représente souvent le nombre d'arrivées de l'événement dont la probabilité est p, sur μ épreuves successives, par

$$\mu p \pm \rho\sqrt{2\mu pq}.$$

En assignant à ρ une valeur numérique donnée, la probabilité pour que le nombre d'arrivées soit compris entre les limites indiquées par cette formule est un nombre fixe, indépendant de μ, de p et de q.

Dans une série quelconque d'épreuves, il y a un à parier contre un de voir le nombre d'arrivées de l'événement dont la probabilité est p compris dans les limites

$$\mu p \pm 0,4769\sqrt{2\mu pq};$$

neuf à parier contre un qu'il sera dans les limites

$$\mu p \pm 1,163\sqrt{2\mu pq},$$

et mille à parier contre un qu'il ne sortira pas des limites

$$\mu p \pm 2,327\sqrt{2\mu pq}.$$

72. La formule (14) permet de définir avec précision ce qu'on doit entendre par *jouer plus ou moins gros jeu*.

Pierre joue 1 000 parties, l'enjeu est 1^{fr}; Paul en joue 100, mais l'enjeu est 10^{fr}.

Tous deux peuvent perdre 1000^{fr}. Ils courent les mêmes risques, mais n'y sont pas également exposés. Supposons un jeu équitable; la probabilité de gagner est p, celle de perdre est q. La mise de Pierre est a, celle de son adversaire est b, et l'on a

$$pb = qa.$$

Si Pierre gagne $\mu p + h$ parties, il en perdra $\mu q - h$; son bénéfice sera

$$(\mu p + h)b - (\mu q - h)a = h(a+b).$$

Il est proportionnel à h.

La probabilité pour que la somme perdue ou gagnée soit inférieure à $h(a+b) = \mathrm{S}$ est donc

$$\frac{2}{\sqrt{\pi}} \int_0^{\frac{h}{\sqrt{2\mu pq}}} e^{-t^2}\, dt = \Theta\left[\frac{\mathrm{S}}{(a+b)\sqrt{2\mu pq}}\right].$$

La situation faite aux joueurs par les conditions du jeu et par le nombre des parties est déterminée par le produit $(a+b)\sqrt{\mu pq}$.

Pierre expose 10^{fr} par partie et joue 20000 parties, avec chance égale de gagner ou de perdre.

Paul expose également 10^{fr}, mais sa chance de gain est $\frac{1}{36}$ seulement; il doit, en cas de gain, recevoir 350^{fr}. Combien Paul doit-il faire de parties, à ce jeu équitable, pour courir les mêmes risques que Pierre?

Le nombre μ des parties est donné par l'équation

$$20\sqrt{5000} = 360\sqrt{\frac{\mu.35}{36^2}},$$

$$\mu = \frac{20\,000}{35} = 571.$$

La situation de Paul, qui joue 571 parties, serait exactement la même que celle de Pierre qui en joue 20000, si les formules étaient rigoureuses; elles ne sont qu'approchées. Il ne faudrait pas les appliquer à de petits nombres.

73. Supposons que, dans un pays qui compte 10 millions d'électeurs, on en désigne 20000 par un tirage au sort, pour leur faire élire un représentant. En supposant que le pays soit partagé entre deux opinions, 4500000 d'un côté et 5500000 de l'autre, quelle est la probabilité pour que le candidat élu appartienne à la minorité.

Le problème est identique à celui-ci :

La probabilité d'un événement est 0,45, celle de l'événement contraire 0,55; on fait entre eux 20000 épreuves, quelle est la probabilité pour que l'écart dépasse 1000 *en faveur de l'événement le moins probable*.

La combinaison la plus probable donnerait à la minorité 9000 électeurs sur 20000. Pour que le hasard la transforme en majorité, il faut un écart supérieur à 1000, dont la probabilité s'évaluera en divisant par 2 l'indication fournie par la formule (14), qui se rapporte à un écart inférieur à α dans un sens ou dans l'autre.

La probabilité pour avoir un écart supérieur à 1000 est

$$1 - \Theta\left(\frac{1000}{\sqrt{10000}}\right) = 1 - \Theta(10).$$

$\Theta(10)$ diffère tellement peu de l'unité que l'événement peut être regardé comme absolument impossible. La probabilité de gagner deux quines de suite à la loterie serait beaucoup plus grande.

L'expérience semble démentir la théorie. Les minorités sont représentées, et, sans qu'il y ait aucune relation entre leur nombre dans le pays et celui de leurs élus, celui-là est loin d'être nul.

Le calcul précédent suppose, en effet, que, tous les électeurs étant inscrits sur une seule liste, on y prenne 20000 noms au hasard pour composer le collège électoral. Il n'en est pas ainsi; les électeurs d'un même département, d'un même arrondissement quelquefois ou d'une même ville ayant des intérêts communs, ne pouvant manquer de subir les mêmes influences, ne sont nullement assimilables à un groupe de votants désignés indépendamment les uns des autres sur le pays tout entier.

74. Lorsque la probabilité d'un événement est connue, on peut prédire avec presque certitude la valeur approchée du nombre d'arrivées de cet événement sur un nombre connu d'épreuves.

Il faut cependant faire d'importantes réserves.

Le théorème reste vrai quand la probabilité de l'événement est inconnue. Le rapport du nombre des arrivées de l'événement au nombre total des épreuves s'approchera certainement de cette probabilité inconnue, si l'on augmente sans cesse le nombre des épreuves. Si, par exemple, on puise dans une urne contenant des boules blanches et des boules noires, en remettant après chaque tirage la boule qui est sortie, le rapport du nombre de boules blanches sorties au nombre des tirages convergera vers la probabilité assignée par la composition de l'urne.

Mais deux conditions sont nécessaires : la probabilité ne doit pas changer pendant les épreuves, et elle doit avoir une valeur déterminée.

Un événement incertain n'a-t-il pas toujours une probabilité déterminée connue ou inconnue?

Il faut se garder de le croire.

Quelle est la probabilité pour qu'il pleuve demain?

Elle n'existe pas.

Non pas parce qu'elle varie d'un jour à l'autre avec l'état du ciel et la direction des vents; mais parce que dans aucune circonstance elle n'a de valeur *objective*, la même pour tous ceux qui l'évaluent sans se tromper.

Il pleuvra ou il ne pleuvra pas, l'un des deux événements est certain, *dès à présent,* et l'autre impossible. Les forces physiques dont dépend la pluie sont aussi bien déterminées, soumises à des lois aussi précises que celles qui dirigent les planètes.

Oserait-on demander la probabilité pour qu'il y ait éclipse de lune le mois prochain?

Un homme est âgé de quarante ans, quelle est la probabilité pour qu'il vive dans dix ans?

Ni l'événement cette fois, ni l'événement contraire n'ont, dès à présent, certitude égale à celle d'une éclipse, mais la probabilité n'a pas davantage, pour cela, de valeur *objective*, indépendante des renseignements connus et du bon jugement de celui qui en fait usage.

Le roi de Siam a quarante ans, quelle est la probabilité pour qu'il vive dans dix ans? Elle est autre pour nous que pour ceux qui ont interrogé son médecin, autre pour le médecin que pour ceux qui ont reçu ses confidences; très différente enfin pour des conjurés qui prendraient leurs mesures pour l'étrangler le lendemain.

Toutes ces probabilités sont subjectives, il n'en existe pas d'objective. L'esprit se refuse à concevoir une urne contenant des boules blanches et des boules noires que l'on puisse composer chaque matin de telle sorte que la chance de vie pendant la journée puisse être remplacée, pour un individu désigné, par le tirage d'une boule dans cette urne.

Les cas précédents ne peuvent être pris pour exemple de l'application du théorème de Bernoulli. Que le roi de Siam meure ou vive, il est impossible de renouveler l'épreuve.

Il faut considérer, au lieu d'un événement isolé, une classe d'événements définis tous ensemble sans distinction. La réponse alors est moins évidente.

Quelle est la probabilité pour qu'il pleuve dix fois à Paris pendant le mois de janvier? On ne dit pas de quelle année.

Pour qu'un habitant de Paris âgé de quarante ans vive encore dans dix ans? On ne désigne pas l'habitant.

Ces deux questions, semblables en apparence à celles qui ont été posées d'abord, en sont réellement très différentes. Nous réunissons en effet, sans distinguer entre eux, tous les jours du mois de janvier de toutes les années et tous les hommes âgés de quarante ans, au lieu de désigner un seul jour et de s'occuper d'un seul homme.

Supposons que les progrès de la Science permettent la prédiction des jours de pluie avec une certitude égale à celle des éclipses. Le Tableau dressé à l'avance soustrairait le problème à la théorie du hasard, mais pour l'y faire rentrer s'il s'agit d'un jour indéterminé du mois de janvier choisi dans une année indéterminée.

L'assimilation à une urne devient possible alors. Le nombre des boules est le nombre des journées de janvier pendant un grand nombre de siècles ; celui des boules blanches, le nombre des jours de pluie.

Une différence subsiste. Dans les épreuves relatives aux jours de pluie, on ne peut ni agiter les boules ni les remettre dans l'urne. En termes plus clairs, l'événement observé un jour n'est pas sans influence sur celui du jour suivant.

S'il s'agit des chances de mort pour un Français âgé de quarante ans, la question est moins évidente encore. On ne peut plus supposer le tableau des événements dressé à l'avance avec certitude ni accepter, *a priori*, l'assimilation avec les tirages dans une urne.

Il n'est pas permis d'affirmer la régularité de la proportion des décès dans une grande ville ou même dans un grand pays.

Cette régularité est révélée par la statistique; elle est très remarquable, mais nullement nécessaire.

Le rapport du nombre des décès à celui des survivants est, pour un âge donné, à peu près invariable.

Le rapport du nombre des naissances masculines à celui des naissances féminines est constant à très peu près, à toutes les époques et dans tous les pays.

Il serait impossible de le démontrer *a priori*.

Le nombre d'hectolitres de blé récoltés dans un département varie d'une année à l'autre.

Pourquoi celui des naissances est-il invariable?

On ne saurait en dire la raison.

La constance des rapports, quand elle est constatée, donne-t-elle le droit d'assimiler les chances des décès et celles des naissances pour l'un et l'autre sexe à des tirages faits dans une urne de composition invariable?

La conséquence n'est nullement permise.

Les tirages au sort, si l'urne est composée pour cela, seront d'accord, en moyenne, avec la statistique. C'est une vérité identique.

Mais de l'égalité des moyennes peut-on conclure celle des chances d'écart? Ce serait admettre ce qui est en question.

Substituons aux hommmes d'une même ville les moutons d'un même troupeau et faisons la statistique de l'abattoir; un marché fait pour un grand nombre d'années règle le nombre des victimes, la régularité est parfaite. On peut composer l'urne dans laquelle on tirera avec certitude, pour un nombre suffisant d'épreuves, un nombre moyen de boules noires égal à celui des moutons sacrifiés chaque jour.

L'assimilation n'ira pas plus loin. Le nombre des boules noires variera d'un jour à l'autre; le nombre des moutons sera invariable.

On fait une expérience sur les dés. On les introduit dans un cornet, on agite, on les lance et on note le coup. L'épreuve répétée 20000 fois confirme les prévisions théoriques. Tous les cas se produisent proportionnellement à leurs probabilités. L'expérimentateur se fatigue; une seule chose importe après tout, c'est de mettre le hasard à l'épreuve. Il dicte les points sans jeter les dés, il s'interdit le choix, le hasard parle par sa bouche.

Obtiendra-t-il les mêmes résultats? Aura-t-il 1000 doublets environ sur 6000 épreuves? Le nombre des points impairs sera-t-il peu différent de celui des points pairs? Il serait téméraire de l'affirmer. L'improvisateur des coups de dés peut se dire : il y a longtemps que je n'ai appelé le double six, l'impartialité veut qu'il ait son tour. Le sort n'a pas de tels scrupules, et il pourra arriver qu'une trop grande conformité aux proportions prévues trahisse l'intervention d'une cause perturbatrice.

Si l'on jette deux dés 6000 fois de suite, le nombre le plus probable des doublets est 1000, l'écart probable est $0,47693\sqrt{\frac{5000}{3}} = 19,4706$ et l'écart moyen $\frac{\sqrt{2}}{\sqrt{\pi}}\sqrt{5000} = 23,0329$.

Si, en faisant 1000 séries de 6000 coups, tous les écarts sont inférieurs à 20; si la valeur moyenne des écarts, au lieu d'être 23, est inférieure à 10, on pourrait affirmer avec une certitude presque infaillible l'existence d'une cause régulatrice corrigeant les écarts du hasard.

75. Lorsque la probabilité d'un événement est variable d'une épreuve à l'autre, le théorème de Bernoulli n'est plus applicable.

La généralisation proposée par Poisson sous le nom de *loi des grands nombres* manque non seulement de rigueur, mais de précision. Les conditions supposées dans l'énoncé échappent par le vague à toute appréciation mathématique. On peut, dans un cas simple et digne d'intérêt, appliquer le théorème de Bernoulli, malgré la variation des chances pendant les épreuves.

Supposons une urne contenant un nombre λ de boules blanches ou noires, la probabilité d'en extraire une boule blanche est p, celle d'extraire une boule noire est q.

On fait μ tirages sans remettre dans l'urne les boules qui en sortent. Si λ et μ sont de grands nombres, il est très probable que le rapport du nombre des boules blanches sorties à celui des boules noires différera peu de $\frac{p}{q}$. En ne remettant pas les boules, on change à chaque épreuve la probabilité de choisir une boule blanche ; mais ce changement est, en quelque sorte, un régulateur de la proportion prévue par le théorème de Bernoulli. Quand la proportion d'une des couleurs est inférieure à ce rapport normal, la probabilité pour elle augmente et les épreuves suivantes ont plus de chance pour corriger l'irrégularité.

On peut préciser cet indication.

Soient λ le nombre total des boules, λp celui des boules blanches, λq celui des boules noires. La probabilité de tirer sur μ épreuves $\mu p - k$ boules blanches et $\mu q + k$ boules noires est, comme on le voit aisément,

$$\frac{1.2.3\ldots\mu}{1.2.3\ldots\mu p - k.1.2.3\ldots\mu q + k}$$
$$\times \frac{1.2.3\ldots\lambda p.1.2.3\ldots\lambda q.1.2.3\ldots(\lambda-\mu)}{1.2.3\ldots[(\lambda-\mu)q+k]1\;2.3\ldots\lambda\;1.2.3\ldots[(\lambda-\mu)p-k]}.$$

On a trouvé (58) approximativement

$$\frac{1.2.3\ldots n}{1.2.3\ldots(np-k)(1.2\;3\ldots nq+k)} p^{np-k} q^{nq+k} = \frac{1}{\sqrt{2\pi\mu pq}} e^{-\frac{k^2}{2npq}}$$

et, par conséquent,

$$\frac{1.2\;3\ldots n}{1.2.3\ldots(np-k)1.2.3\ldots(nq+k)} = \frac{1}{\sqrt{2\mu\pi pq}} e^{-\frac{k^2}{2npq}} p^{k-np} q^{-k-nq}.$$

En remplaçant les fractions

$$\frac{1.2.3\ldots\mu}{1.2.3\ldots(\mu p-k)1.2.3\ldots(\mu q+k)},$$
$$\frac{1.2.3\ldots(\lambda-\mu)}{[1.2.3\ldots(\lambda-\mu)p-k][1.2.3\ldots(\lambda-\mu)q+k]},$$
$$\frac{1.2.3\ldots\lambda p.1.2.3\ldots\lambda q}{1.2.3\ldots\lambda}$$

par leurs valeurs déduites de cette formule, on trouve

$$\frac{1}{\sqrt{2\pi\mu pq}}\sqrt{\frac{\lambda}{\lambda-\mu}}\, e^{-\frac{k^2}{2pq}\frac{\lambda}{\mu(\lambda-\mu)}}.$$

C'est précisément la formule qui convient au cas où l'on ne remet pas les boules, dans laquelle le nombre μ des tirages est multiplié par $\frac{\lambda-\mu}{\lambda}$.

Si l'on tire la moitié des boules sans les remettre, la probabilité d'un même écart absolu entre le nombre des boules blanches et le nombre le plus probable sera la même que si l'on tirait un quart seulement en remettant les boules.

CHAPITRE V.

DÉMONSTRATIONS ÉLÉMENTAIRES DU THÉORÈME DE BERNOULLI.

L'emploi du calcul est comparable à celui d'un instrument dont on connaît exactement la précision

FOURIER.

76. Le calcul dont parle Fourier est celui des moyennes.

Lorsqu'un événement douteux peut se présenter de plusieurs manières entre lesquelles prononce le hasard et qu'une certaine grandeur résulte de chaque épreuve, la valeur probable définie (47) différera de moins en moins de la moyenne des valeurs obtenues.

La probabilité d'une différence supérieure à un nombre donné, si petit qu'il soit, tend vers zéro quand le nombre des épreuves augmente.

Soient $\lambda_1, \lambda_2, \ldots, \lambda_n$ les valeurs possibles d'une grandeur, p_1, $p_2, \ldots, p_n$ leurs probabilités. La valeur probable de la grandeur est par définition

$$p_1\lambda_1 + p_2\lambda_2 + \ldots + p_n\lambda_n = G.$$

Si l'épreuve renouvelée un grand nombre de fois donne successivement les valeurs x_1, x_2, ..., x_μ, la moyenne arithmétique

$$\frac{x_1+x_2+\ldots+x_\mu}{\mu}$$

différera peu, si μ est grand, de la valeur probable G.

Rien n'est certain, bien entendu, mais la valeur probable du carré de la différence

$$\frac{x_1+x_2+\ldots+x_\mu}{\mu}-G$$

tend vers zéro lorsque μ augmente. Lorsque la valeur probable d'une grandeur tend vers zéro, on n'en peut rien conclure. De grandes valeurs positives peuvent avoir grande probabilité, aussi bien que les grandes valeurs négatives qui les compensent; il en est autrement quand la valeur probable d'un carré est très petite. Toutes les valeurs possibles étant de même signe, aucune ne peut à la fois n'être pas très petite et ne pas avoir une probabilité très petite. La probabilité pour que la valeur surpasse un nombre donné tend nécessairement vers zéro.

Nous considérerons, dans le cas actuel, l'expression essentiellement positive

$$\left(\frac{x_1+x_2+\ldots+x_\mu}{\mu}-G\right)^2, \tag{1}$$

pour démontrer que sa valeur probable tend vers zéro.

L'expression (1) peut être remplacée, en adoptant une notation bien connue, par

$$\frac{\Sigma x_i^2+2\Sigma x_i x_{i'}}{\mu^2}-2\frac{G\Sigma x_i}{\mu}+G^2. \tag{2}$$

λ_1, λ_2, ..., λ_n désignant, comme on l'a supposé, les valeurs possibles de x_1, x_2, ..., x_n et p_1, p_2, ..., p_n leurs probabilités, la valeur probable de x_i^2 est

$$p_1\lambda_1^2+p_2\lambda_2^2+\ldots+p_n\lambda_n^2.$$

Représentons-la par H; x_1^2, x_2^2, ..., x_μ^2 ont même valeur probable, puisque rien ne les distingue. La valeur probable de Σx_ι^2 est donc μH.

La valeur probable de $x_\iota x_{\iota'}$ est (48) le produit de la valeur probable G de x_ι par la valeur égale G de $x_{\iota'}$.

La valeur probable de x_ι est G.

La valeur probable de la somme (2) est donc

$$\frac{H}{\mu} + \frac{\mu(\mu-1)}{\mu^2}G^2 - \frac{2\mu G^2}{\mu} + G^2,$$

qui se réduit à

$$\frac{H - G^2}{\mu}. \tag{3}$$

Quels que soient H et G, elle tend vers zéro lorsque μ augmente.

77. Le théorème précédent a de nombreuses applications.

Pierre joue à un jeu de hasard. S'il gagne, il recevra une somme b, enjeu de son adversaire ; s'il perd, il donnera une somme a. La probabilité de gagner est p, celle de perdre est q. Le gain probable à chaque partie est

$$pb - qa = G.$$

La valeur probable du carré du gain est

$$pb^2 + q(-a)^2 = pb^2 + qa^2 = H.$$

Si donc, après μ parties, on nomme X le gain, positif ou négatif, résultant de ces μ parties, la valeur probable de

$$\left(\frac{X}{\mu} - G\right)^2 \tag{4}$$

est

$$\frac{H - G^2}{\mu} = \frac{pb^2 + qa^2 - (pb - qa)^2}{\mu} = pq\,\frac{(a+b)^2}{\mu}.$$

En multipliant l'expression (4) par μ^2, la valeur probable sera multipliée par μ^2, et

$$\mu pq(a+b)^2$$

est, par conséquent, la valeur probable de

$$[X - \mu(pb - qa)]^2. \tag{5}$$

Si l'on représente ce carré (5) par ε^2, on aura

$$X = \mu(pb - qa) \pm \varepsilon. \tag{6}$$

Le gain X fait sur μ parties à un jeu quelconque est donc représenté par la somme de deux termes : l'un, complètement connu, proportionnel au nombre des parties jouées; l'autre ε, qui dépend du hasard et dont le signe même est inconnu.

Un tel énoncé, si l'on n'ajoutait rien, serait insignifiant. Ignorer dans une somme la valeur et le signe de l'un des termes, c'est ne rien savoir sur la somme. Mais, sans connaître ε, nous avons sur lui un renseignement. La valeur probable de ε^2 est $\mu(a+b)^2pq$, elle est proportionnelle à μ; celle de ε est donc de l'ordre de $\sqrt{\mu}$ et le second terme de (6) deviendra de plus en plus petit par rapport au premier lorsque μ croîtra sans limite.

Nous comparons, il est vrai, la valeur certaine du premier terme à la valeur probable du second. Nos conclusions peuvent être en défaut; mais, lorsque μ augmente, la probabilité pour qu'il en soit ainsi tend vers zéro. Il faudrait que le hasard rendît ε^2 infiniment grand par rapport à sa valeur probable.

Si ϖ désigne la probabilité d'une valeur supérieure à α, il en résulte dans la valeur probable de cette grandeur essentiellement positive un terme $\varpi\alpha^2$, qui deviendrait à lui seul plus grand que la somme dont il fait partie, si, lorsque α augmente, ϖ ne devenait pas très petit.

En réduisant l'évaluation du gain fait en μ parties au premier terme

$$X = \mu(pb - qa), \tag{7}$$

on pourra affirmer avec une certitude croissante que l'erreur *relative* tend vers zéro lorsque μ augmente sans limite.

78. Le cas où le jeu est équitable doit être traité à part.

On a alors

$$pb - qa = 0,$$

et le théorème est en défaut.

Le premier terme de la valeur (6) de X étant nul, le gain fait en μ parties se réduit à

$$X = \pm \varepsilon.$$

On ne peut rien affirmer sur son signe. Quelque grand que soit μ, Pierre, à un jeu équitable, peut perdre ou gagner. La valeur probable du *carré* du gain positif ou négatif est proportionnelle au nombre des parties. On peut donc tenir pour certain qu'après un grand nombre de parties la valeur moyenne $\frac{X}{\mu}$ du gain fait à chaque partie par celui des joueurs que le hasard a favorisé sera très petite. La valeur probable de $\frac{X^2}{\mu^2}$, carré du gain moyen par partie, est en effet

$$\frac{\mu pq(a+b)^2}{\mu^2} = \frac{pq(a+b)^2}{\mu};$$

elle tend vers zéro quand μ augmente.

79. Le théorème démontré (77) peut s'appliquer aux épreuves successives entre deux événements contraires.

Soient p et q les deux probabilités. Leur somme est égale à l'unité; une série de μ épreuves amenant, dans un ordre réglé par le hasard, l'un ou l'autre de ces deux événements peut donner lieu à un jeu équitable. Pierre gagnera la partie si l'événement dont la probabilité est p se présente; sa mise sera égale à p, celle de l'adversaire égale à q.

On aura, en adoptant les notations du théorème (77),

$$G = pq + q(-p) = 0,$$
$$H = pq^2 + qp^2 = pq.$$

Si, sur μ parties, Pierre en gagne m et en perd $\mu - m$, son gain sera

$$mq - (\mu - m)p = m - \mu p.$$

La valeur probable de

$$\left(\frac{m - \mu p}{\mu}\right)^2 = \left(\frac{m}{\mu} - p\right)^2 \tag{8}$$

est donc (76)

$$\frac{\mathrm{H}}{\mu} = \frac{pq}{\mu};$$

elle tend vers zéro lorsque μ augmente. On peut donc affirmer avec une certitude croissante que la différence entre le rapport $\frac{m}{\mu}$ et la probabilité p tend vers zéro lorsque le nombre des épreuves augmente. C'est le théorème de Bernoulli.

La valeur probable de l'expression (8) a été obtenue déjà sous une autre forme.

La valeur probable (62) du carré de l'écart sur μ épreuves est μpq.

Cet écart est la difference $m - \mu p$. La valeur probable de $(m - \mu p)^2$ étant μpq, celle de $\left(\frac{m}{\mu} - p\right)^2$ est $\frac{pq}{\mu}$.

C'est le même théorème.

80. Nous donnerons du théorème de Bernoulli une démonstration plus simple encore que la précédente.

Supposons deux événements contraires ayant pour probabilités p et q. Sur μ épreuves faites entre eux, les nombres d'arrivées les plus probables sont μp et μq; nommons *écart* la différence entre le nombre d'arrivées de l'événement dont la probabilité est p et le nombre le plus probable μp.

On promet à Pierre une somme égale à la valeur absolue de l'écart, et non plus à son carré comme on l'a supposé dans la démonstration précédente. Sans calculer l'espérance mathématique de Pierre, nous allons démontrer que, le nombre μ des épreuves grandissant indéfiniment, cette espérance est très petite par rapport à μ.

Désignons par $\varphi(\mu)$ l'espérance mathématique de Pierre.

Si l'on double le nombre des épreuves, $\varphi(\mu)$ deviendra $\varphi(2\mu)$, très différent de $2\varphi(\mu)$.

Partageons, en effet, les 2μ épreuves en deux séries de μ. L'écart peut, dans les deux séries, avoir le même signe ou des signes différents. Dans le premier cas, l'écart total est la somme des deux écarts partiels; dans le second, il est leur différence. L'espérance mathématique, si l'on sait que le premier cas se produit, est $2\varphi(\mu)$;

dans le second, elle est évidemment plus petite. Désignons-la par $2\alpha\varphi(\mu)$, α étant plus petit que l'unité.

Soient p_1 la probabilité pour que, sur μ épreuves, l'écart soit positif, q_1 pour qu'il soit négatif; la probabilité pour que, dans les deux séries de μ épreuves, les deux signes de l'écart soient semblables est $p_1^2+q_1^2$; pour qu'ils soient différents, elle est $2p_1q_1$. On en conclut

$$\varphi(2\mu) = 2(p_1^2+q_1^2)\varphi(\mu)+4p_1q_1\alpha\varphi(\mu).$$

On a

$$p_1^2+q_1^2+2p_1q_1 = 1;$$

par conséquent

$$p_1^2+q_1^2+2\alpha q_1p_1 < 1.$$

$\varphi(2\mu)$ est donc plus petit que $2\varphi(\mu)$.

La somme

$$p_1^2+q_1^2+2\alpha p_1q_1$$

ne pourrait tendre vers l'unité lorsque μ augmente que dans deux cas : ou bien l'une des probabilités p_1 ou q_1 tend vers zéro, ou bien α tend lui-même vers l'unité.

La première hypothèse est impossible. Si, en effet, p_1 tend vers zéro, il en résulterait que, sur un nombre d'épreuves indéfiniment croissant, on pourrait regarder comme certain que l'écart sera d'un certain signe. En jouant alors au jeu équitable, défini (79), l'un des joueurs, après un grand nombre de parties, serait certain de gagner.

La seconde hypothèse est également inadmissible.

L'espérance mathématique de celui qui attend la différence de deux écarts ne peut être la même que s'il attend leur somme.

Nous pouvons donc écrire

$$\varphi(2\mu) = 2G_1\varphi(\mu),$$

G_1 étant plus petit que l'unité et ne s'en approchant pas indéfiniment.

On en conclut

$$\begin{aligned}
\varphi(4\mu) &= 2G_1\varphi(2\mu),\\
\varphi(8\mu) &= 2G_3\varphi(4\mu),\\
&\ldots\ldots\ldots\ldots\ldots\ldots,\\
\varphi(2^n\mu) &= 2G_n\varphi(2^{n-1}\mu)
\end{aligned}$$

et, en multipliant ces équations,

$$\varphi(2^n \mu) = 2^n G_1 G_2 \ldots G_n \varphi(\mu).$$

On en conclut

$$\frac{\varphi(2^n \mu)}{2^n \mu} = G_1 G_2 \ldots G_n \frac{\varphi(\mu)}{\mu}.$$

Les facteurs G_1, G_2, ..., G_n étant plus petits que l'unité et ne tendant pas vers l'unité, leur produit tend vers zéro. En posant $2^n \mu = z$, on a, lorsque z grandit sans limite,

$$\lim \frac{\varphi(z)}{z} = 0.$$

L'espérance mathématique de celui qui attend une somme égale à l'écart absolu sur z épreuves est donc très petite par rapport au nombre des épreuves, et l'on peut regarder, par conséquent, comme certain qu'après un nombre indéfiniment croissant d'épreuves l'écart est infiniment petit par rapport au nombre des épreuves.

CHAPITRE VI.

LA RUINE DES JOUEURS.

Si le nombre des parties est indéfini, l'avantage du joueur le plus riche serait infini si sa fortune pouvait l'être c'est pour cela que c'est courir a une ruine certaine que de jouer indifféremment avec tous ceux qui se rencontrent.

AMPERE

81. Lorsqu'un joueur joue indéfiniment a un jeu équitable, sa ruine tôt ou tard est certaine. La proposition semble contradictoire, elle ne l'est pas. — 82. Lorsque deux joueurs luttent indéfiniment, quelles que soient les conditions du jeu, l'un des deux doit finir par se ruiner. — 83. Calcul numérique. — 84. La perte peut entraîner la ruine avant la fin du nombre convenu de parties. Cela accroît les chances de ruine. — 85. Deux manières d'énoncer le problème de la ruine des joueurs. — 86. Cas où Pierre possédant m francs joue indéfiniment à un jeu équitable. — 87. La valeur probable du nombre des parties est infinie. Il n'y a pas contradiction. — 88. Démonstration de l'énoncé précédent. — 89. Calcul des chances de ruine dans un nombre donné de parties. — 90. Exemples numériques. — 91. Cas où deux joueurs ont des fortunes données. Chance de ruine de chacun. — 92. Cas où le jeu n'est pas équitable. — 93. Autre manière d'obtenir le même résultat. — 94. Cas où les deux joueurs ont même fortune et exposent la même mise. — 95. Probabilité pour qu'un joueur qui joue indéfiniment finisse par se ruiner. Trois cas peuvent se présenter. — 96. Le cas où la ruine n'est pas certaine est celui où le joueur a un avantage. — 97. Exemple numérique. — 98. Probabilité d'être ruiné précisément après un nombre donné de coups. — 99. Valeur approchée de la probabilité. — 100. Probabilité pour que la ruine soit postérieure au $\mu^{\text{ième}}$ coup. — 101. Valeur maxima de la probabilité. — 102. Valeur maxima de la valeur approchée. — 103. Valeur probable du nombre des parties jouées avant la ruine de l'un des joueurs. — 104. Cas où le jeu n'est pas équitable. Théorème de M. Rouché. — 105. Évidence apparente du théorème. — 106. Insuffisance de la démonstration. — 107. Réduction du cas où le jeu est équitable au cas général. — 108. Cas où les fortunes sont égales en commençant le jeu. — 109. Cas où deux joueurs luttant l'un contre l'autre peuvent l'un et l'autre être ruinés. Insuffisance d'un raison-

nement qui semble fort simple. — 110. Cas où Pierre et Paul possèdent chacun 2^{fr}. — 111 Cas où ils possèdent 3^{fr}. — 112. Cas général. — 113. Examen d'une combinaison proposée pour accroître les chances de gain.

81. Le jeu ruine ceux qui s'y livrent. Il n'y a exception que pour les joueurs auxquels les conditions acceptées accordent un avantage.

Le fermier des jeux à Monte-Carlo peut accroître sans crainte le nombre des coups. La menace ne s'adresse qu'aux pontes.

Lorsque le jeu est équitable, la ruine tôt ou tard est certaine.

La proposition semble contradictoire. En ruinant l'un des joueurs, le jeu enrichit l'autre; en s'exposant à perdre une fortune, on a l'espoir de la doubler.

Cela n'est pas douteux ; mais, quand la fortune est doublée, le théorème s'y applique avec la même certitude : elle peut doubler encore, centupler peut-être, tout sera emporté à la fois par un caprice du hasard. En combien de temps ? Nul ne le sait ; la probabilité augmente avec le nombre des parties et converge vers la certitude. C'est cette progression extrêmement lente, il faut le déclarer tout d'abord, à l'étude de laquelle est consacré ce Chapitre.

82. Lorsque deux joueurs luttent constamment l'un contre l'autre, quelles que soient leurs fortunes et les conditions, équitables ou non, du jeu qu'ils renouvellent sans cesse, l'un des deux finira par ruiner l'autre ; la probabilité pour qu'ils puissent faire un nombre donné de parties tend vers zéro quand ce nombre augmente.

Lorsque deux joueurs, en effet, font un grand nombre de parties, la probabilité de la répartition la plus probable entre les parties gagnées et perdues par l'un d'eux tend vers zéro quand le nombre des parties augmente. Elle est (57) inversement proportionnelle à la racine carrée du nombre des parties.

Si la probabilité de la combinaison la plus probable tend vers zéro, il en est de même, à plus forte raison, de toute autre combinaison désignée et, par conséquent, aussi d'un ensemble de combinaisons quel qu'il soit, dont le nombre ne croîtrait pas indéfiniment avec le nombre des parties.

Quelles que soient les mises des deux joueurs, on peut assigner une répartition des pertes et des gains, telle que la compensation soit parfaite et que chaque joueur, à la fin, se retrouve avec sa fortune primitive. Si, prenant pour point de départ ce mode de répartition, on accroît le nombre des parties gagnées par l'un des joueurs, le gain sera pour lui, pour chaque partie gagnée de plus, et par conséquent pour chaque partie perdue de moins, égal à la somme des deux mises, et tout écart de cette répartition qui équilibre les pertes et les gains ruinera l'un ou l'autre joueur, si, multiplié par la somme des mises, il donne un produit plus grand que la fortune du plus riche.

En caractérisant les combinaisons par le nombre total des parties gagnées par l'un des joueurs, le nombre de celles qui, ne ruinant aucun des deux joueurs, permettent la continuation du jeu est donc indépendant du nombre des parties jouées, et la probabilité pour que l'une ou l'autre de ces combinaisons se produise tend vers zéro quand le nombre des parties augmente.

83. Supposons, par exemple, que l'on joue à pile ou face, l'enjeu étant 1^{fr} la partie. Si le nombre des arrivées de pile est égal à celui des arrivées de face, il n'y aura ni perte ni gain.

Cherchons combien il faut faire de parties pour que la probabilité d'une perte de 100000^{fr} pour l'un des joueurs soit égale à 0,999.

Pour que, en jouant 1^{fr} la partie, l'un des joueurs perde 100000^{fr}, il faut que l'écart, dans un sens ou dans l'autre, soit plus grand que 50000.

On entend par *écart*, il n'est pas inutile peut-être de le rappeler, la différence entre le nombre des parties gagnées et le nombre relatif à la combinaison qui rend les pertes nulles. Cette combinaison correspond, dans le cas actuel, à l'égalité des pertes et des gains.

Si μ est le nombre des parties jouées, la probabilité de la combinaison la plus probable est (57)

$$\frac{1}{\sqrt{2\pi\mu pq}},$$

et, dans le cas actuel, puisque p et q sont égaux à $\frac{1}{2}$,

$$\frac{1}{\sqrt{\frac{\pi\mu}{2}}}.$$

La probabilité pour que le hasard amène une des 100 000 combinaisons pour lesquelles la perte est inférieure à 100 000$^{\text{fr}}$ est plus petite que

$$\frac{100\,000}{\sqrt{\frac{\pi\mu}{2}}}.$$

Si donc on a

$$\frac{100\,000}{\sqrt{\frac{\pi\mu}{2}}} < \frac{1}{1000}, \tag{1}$$

la probabilité d'une perte moindre que 100 000$^{\text{fr}}$ sera plus petite que $\frac{1}{1000}$ et, par conséquent, celle d'une perte plus grande surpassera 0,999.

L'inégalité (1) donne

$$\mu > \frac{10^{16}.2}{\pi}.$$

7 millions de milliards de parties suffiraient pour donner la probabilité demandée.

Si deux joueurs pouvaient faire ce nombre immense de parties, il y aurait pour chacun d'eux probabilité presque égale à $\frac{1}{2}$ de perdre plus de 100 000$^{\text{fr}}$ et probabilité égale de les gagner.

Cette évaluation numérique pourra rassurer ceux que la certitude de ruine effrayait plus qu'il ne faut.

La probabilité d'un écart reste la même (58), lorsque l'écart diminue ou augmente proportionnellement à la racine carrée du nombre des parties. Si l'on divise le nombre des parties par 1 million, en les réduisant à 7 milliards, la somme dont la perte par l'un des joueurs aura 999 chances sur 1000 sera réduite à 100$^{\text{fr}}$.

La méthode précédente ne donne qu'une limite. Le calcul exact est facile.

La probabilité pour qu'en jouant à un jeu équitable, les mises étant a et b et les probabilités de gagner à chaque partie p et q, la somme perdue ou gagnée soit inférieure à une limite S a été donnée (72); elle est

$$\frac{2}{\sqrt{\pi}}\int_0^{\frac{S}{(a+b)\sqrt{2\mu pq}}} e^{-t^2}\,dt = \Theta\left[\frac{S}{(a+b)\sqrt{2\mu pq}}\right];$$

en la retranchant de l'unité, on aura la probabilité pour que la perte de l'un des joueurs soit plus grande que S.

On trouve dans la Table

$$\Theta(0{,}09) = 0{,}10.$$

Si l'on pose

$$\frac{S}{(a+b)\sqrt{2\mu pq}} = 0{,}09,$$

on en déduira le nombre μ des parties nécessaires pour que la probabilité d'une perte supérieure à S soit égale à $\frac{9}{10}$.

Soient

$$p = \frac{1}{2}, \qquad q = \frac{1}{2}, \qquad a = 1, \qquad b = 1;$$

on trouvera

$$\mu = \frac{10000\,S^2}{162} = 62{,}4\,S^2.$$

Si l'on suppose $S = 100000$, on aura

$$\mu = 624000 \text{ millions}.$$

En prenant $S = 100$, on a

$$\mu = 624000.$$

Il faut faire 624000 parties à 1fr pour que la probabilité de perdre ou de gagner 100fr soit égale à $\frac{9}{10}$.

84. Les chiffres précédents, donnés sans commentaire, feraient naître une idée très fausse sur les chances de ruine.

Nous supposons le nombre des parties convenu à l'avance. On réglera à la fin : telle a été notre hypothèse. La chance de perte n'a rien alors de bien effrayant, il y a une chance sur dix pour qu'après 624000 parties jouées tout se réduise à une perte inférieure à 100^{fr}.

La chance de ruine que nous voulons étudier est bien différente. On doit à chaque partie déposer la mise ; si donc, à un certain moment, la perte du joueur dépasse sa fortune, il ne sera pas admis à continuer et perdra toute chance de se relever. La probabilité qu'il faut connaître n'est pas celle de la perte finale, mais celle de la perte maximum à l'un des instants de la série de parties.

85. Le problème peut être posé de deux manières. On peut considérer deux joueurs jouant l'un contre l'autre à des conditions données, et chercher les chances de ruine pour chacun d'eux et les probabilités relatives à la durée du jeu, c'est-à-dire au nombre des parties jouées avant la ruine de l'un d'eux. On peut, et c'est une étude très différente, étudier le sort du premier joueur sans se préoccuper du second, supposer qu'il change d'adversaire, qu'il en trouve toujours un disposé à jouer aux mêmes conditions, ou, ce qui revient au même, que celui contre lequel il entreprend la lutte ait une fortune infinie.

Dans le premier cas, nous l'avons vu (82), le jeu doit finir tôt ou tard par la ruine de l'un des joueurs. La probabilité pour que le nombre des parties nécessaires atteigne une limite donnée tend vers zéro quand cette limite augmente.

Dans le second cas, quand un seul des deux joueurs peut être ruiné, la ruine est également certaine si le jeu est équitable ; la probabilité pour que le nombre des parties dépasse toute limite est infiniment petite.

On peut le démontrer sans calcul. Supposons que Pierre possède m francs et qu'il ait résolu de risquer une même somme à un jeu équitable, tant qu'il pourra déposer sa mise.

Nous pouvons supposer qu'une première lutte s'établisse entre Pierre et un adversaire de fortune égale. Le jeu est équitable. Pierre a chance $\frac{1}{2}$ de sortir vainqueur de cette lutte : il possédera alors $2m$ francs. Supposons-lui alors un second adversaire possédant comme lui $2m$ francs, Pierre a chance $\frac{1}{2}$ d'être ruiné par lui, mais chance $\frac{1}{2}$ aussi de le ruiner et de posséder $4m$ francs.

Pierre luttera alors contre un adversaire possédant $4m$ francs, et il aura chance $\frac{1}{2}$, en le ruinant, d'en posséder $8m$. La série est indéfinie. On voit que, pour échapper à toutes les chances de ruine, Pierre devrait avoir autant de bonheur que si, jouant sans cesse à pile ou face, il ne perdait jamais une seule partie. Une telle persistance doit être évidemment considérée comme impossible et Pierre, tôt ou tard, se ruinera.

L'assimilation avec le jeu de pile ou face est évidente. Pour en effacer toute différence, il faut supposer aux parties de pile ou face des durées croissantes. Si l'une d'elles ne finissait pas, la démonstration perdrait toute sa force. Un tel hasard (82) doit être tenu pour impossible.

86. La démonstration précédente, en montrant la ruine du joueur inévitable, n'apprend rien sur les probabilités relatives au nombre de parties qui doivent l'amener. Nous pouvons, dès à présent, démontrer que la valeur probable de ce nombre de parties est infinie. La contradiction semble choquante.

La ruine est certaine, dit-on, et la valeur probable du nombre des parties qui la procurent est infinie. Si le nombre des parties est infini on ne pourra pas les jouer, la ruine ne s'accomplira pas : elle n'est donc pas certaine.

La ruine est certaine. Il ne faut pas confondre le nombre des parties qui, vraisemblablement, seront jouées avec le nombre probable des parties; il faut surtout ne pas oublier ce que nous entendons par *certitude*. Quand on dit la ruine est *certaine*, tôt ou tard, on n'entend pas affirmer qu'après un nombre de parties, si grand qu'il soit, la ruine est assurée comme un théorème de Géométrie. S'il en était ainsi, le nombre probable des parties ne pourrait pas, évidemment, surpasser et serait certainement très loin d'égaler cette limite infranchissable.

La ruine est certaine, cela veut dire : la probabilité pour que le nombre des parties qui s'accompliront surpasse une limite donnée tend vers zéro quand cette limite augmente.

La valeur probable du nombre de parties est infinie, cela veut dire : l'espérance mathématique de celui qui doit recevoir autant de francs qu'on jouera de parties est infinie.

Les propositions ainsi comprises ne sont nullement contradictoires.

Supposons que Pierre, se mettant au jeu avec 1^{fr} seulement, c'est le cas le plus défavorable, soit décidé à jouer, sans interruption, jusqu'à ce que ce franc soit perdu. La probabilité pour qu'il le perde, *tôt ou tard,* est une certitude; mais, quelle que soit la limite qu'on voudra assigner, il y a possibilité pour que le nombre des parties jouées la surpasse. La probabilité pour qu'il en soit ainsi tend vers zéro quand la limite augmente. En disant que le nombre des parties ne *peut* pas être infini, on ne veut pas dire autre chose.

Si l'on promet à Paul 1^{fr} par partie que jouera Pierre avant d'avoir perdu le franc qu'il possède en entrant au jeu, l'espérance mathématique de Paul est, par définition, le nombre probable des parties. Il faudra, pour la calculer, multiplier chaque nombre possible par la probabilité correspondante. Or les nombres *possibles* vont à l'infini : il n'y a rien de contradictoire à annoncer qu'en les multipliant par des probabilités de plus en plus petites qui, suivant notre façon de parler, expriment des impossibilités, la somme des produits augmente sans limite.

Le résultat est analogue au paradoxe de Saint-Pétersbourg, dans lequel nous avons rencontré déjà une espérance mathématique rendue infinie par l'énormité des sommes dont la probabilité paraissait assez petite pour qu'on n'y attachât aucun prix.

87. Démontrons que la durée probable du jeu est infinie pour un joueur qui change d'adversaire et veut risquer la même mise, à un jeu équitable, tant qu'il aura possibilité de la mettre en jeu.

Considérons d'abord deux joueurs, Pierre et Paul, possédant chacun $2m$ francs. Ils luttent à un jeu équitable jusqu'à la ruine de l'un deux. Soit $\varphi(2m)$ le nombre probable des parties qu'ils joueront, $\varphi(m)$ désignant, par la même notation, le nombre probable des parties quand les deux joueurs possèdent chacun m francs. La lutte entre Pierre et Paul pourra, sans que rien soit changé aux chances de chacun, commencer par deux épreuves distinctes. Paul, dans une première série de parties, risquera m francs contre m francs de Pierre, et dans une seconde série il exposera la seconde moitié de son avoir contre la seconde moitié de celui de Pierre.

Deux cas pourront se présenter : ou les deux luttes, quand elles seront terminées, auront le même vainqueur, qui aura ruiné son adversaire ; ou chacun gagnera l'une des deux séries, et les joueurs se retrouveront dans la situation primitive, possédant chacun $2m$ francs.

On en conclut

$$\varphi(2m) = 2\varphi(m) + \frac{1}{2}\varphi(2m). \tag{2}$$

Le nombre des parties se compose, en effet, des nombres de parties faites dans les deux séries, et dont chacune a pour valeur probable $\varphi(m)$, et de plus, éventuellement, du nombre de celles qu'il faudra faire encore si, à la suite des deux séries, on se retrouve dans la situation primitive, ce dont la probabilité est $\frac{1}{2}$.

L'équation (2) donne

$$\varphi(2m) = 4\varphi(m). \tag{3}$$

Supposons maintenant que Pierre, possédant m francs, ait résolu de jouer sans limite contre tout adversaire qui se présentera. On peut régler son jeu de la manière suivante : il luttera d'abord contre un premier adversaire possédant comme lui m francs. S'il le ruine, il possédera $2m$ francs; on lui opposera un second adversaire de fortune égale. S'il ruine le second, il possédera $4m$ francs et pourra lutter, à chances égales, contre un adversaire ayant comme lui $4m$ francs. Le jeu continuera jusqu'à la rencontre d'un adversaire qui, à ce jeu toujours égal, réussira à le ruiner.

On voit tout d'abord, comme on l'a remarqué déjà (85), que la ruine de Pierre est certaine; il ne pourrait l'éviter qu'en étant toujours favorisé par le hasard dans une série indéfinie d'épreuves pour chacune desquelles la probabilité est $\frac{1}{2}$.

Le nombre probable des parties est, d'après l'analyse précédente,

$$\varphi(m) + \frac{1}{2}\varphi(2m) + \frac{1}{4}\varphi(4m) + \frac{1}{8}\varphi(8m) + \ldots \tag{4}$$

Pierre, en effet, est certain de jouer la première série, pour laquelle le nombre probable des parties est $\varphi(m)$: il a probabilité $\frac{1}{2}$ de jouer la seconde : il suffit pour cela qu'il sorte vainqueur de la première ; il a probabilité $\frac{1}{4}$ de jouer la troisième, car il suffit qu'il gagne les deux premières, etc. En ayant égard à la relation (3), la somme (4) devient

$$\varphi(m) + 2\varphi(m) + 4\varphi(m) + 8\varphi(m) + \ldots;$$

elle est infinie quel que soit $\varphi(m)$.

89. Les chances de ruine d'un joueur, à un jeu équitable, se calculent aisément lorsque, le nombre des parties étant fixé à l'avance, on doit les faire, quoi qu'il arrive, et régler les comptes à la fin.

p et q désignant les probabilités de perte et de gain à chacune des μ parties, la probabilité pour que le nombre des parties perdues surpasse $\mu q + h$, c'est-à-dire pour que la perte surpasse le produit de h par la somme des mises, est

$$\frac{1}{\sqrt{2\pi\mu pq}} \int_h^\infty e^{-\frac{z^2}{2\mu pq}}\, dz. \tag{5}$$

Si l'on pose

$$\frac{z^2}{2\mu pq} = t^2,$$

elle devient

$$\frac{1}{\sqrt{\pi}} \int_0^\infty c^{-t^2}\, dt - \frac{1}{\sqrt{\pi}} \int_0^{\frac{h}{\sqrt{2\mu pq}}} e^{-t^2}\, dt$$

Le premier terme est égal à $\frac{1}{2}$, le second tend vers zéro quand μ augmente.

La probabilité pour qu'un joueur, après un nombre croissant de parties, fasse un gain supérieur à une somme donnée, tend vers $\frac{1}{2}$, quelle que soit cette somme, quand le nombre des parties augmente sans limite.

La probabilité pour qu'il soit en perte d'une somme égale tend aussi vers $\frac{1}{2}$, et la probabilité pour que la perte ou le gain restent, pour un nombre croissant de parties, inférieurs à un nombre donné tend vers zéro, quel que soit le nombre. Le jeu, on ne doit pas l'oublier, est supposé équitable.

90. La probabilité d'une perte donnée pour un nombre donné

de parties diminue rapidement quand la perte assignée est grande.

Supposons que l'on joue 1000 parties, p et q étant l'un et l'autre égaux à $\frac{1}{2}$. La probabilité pour l'un des joueurs de perdre un nombre de parties supérieur à $500 + h$, par conséquent d'être en perte de $2mh$, m étant la mise, est

$$\frac{1}{2} - \frac{1}{\sqrt{\pi}} \int_0^{\frac{h\sqrt{2}}{\sqrt{1000}}} e^{-t^2}\, dt.$$

Le Tableau suivant donne la probabilité pour que le joueur qui possède $2h$ francs soit ruiné en jouant mille parties.

$$A = \frac{1}{2} - \frac{1}{\sqrt{\pi}} \int_0^{\frac{h\sqrt{2}}{\sqrt{1000}}} e^{-t^2}\, dt.$$

h.	A.	h	A.	h.	A.
1......	0,475	18......	0,129	35......	0,014
2......	0,450	19......	0,115	36. ...	0,012
3......	0,425	20......	0,103	37......	0,010
4......	0,400	21... ..	0,093	38.. . .	0,009
5......	0,377	22......	0,083	39	0,007
6......	0,353	23.... .	0,073	40......	0,006
7....	0,330	24......	0,065	41	0,005
8... ..	0,307	25......	0,058	42......	0,004
9......	0,285	26......	0,051	43	0,004
10......	0,264	27......	0,045	44......	0,003
11... ..	0,244	28	0,039	45......	0,003
12......	0,225	29......	0,034	46......	0,002
13... ..	0,206	30......	0,029	47... ..	0,002
14......	0,188	31......	0,026	48......	0,002
15.. ...	0,172	32......	0,022	49... ..	0,001
16......	0,156	33......	0,019	50.	0,001
17......	0,142	34......	0,016	82... ..	0,000004

Le Tableau suivant donne le nombre des parties qu'il faut jouer pour que, les deux adversaires ayant à chaque partie probabilité $\frac{1}{2}$ de gagner, et l'enjeu étant 1^{fr}, celui des deux qui sera favorisé par le hasard ait une probabilité $\frac{1}{2}$ de gagner une somme supérieure à un nombre donné :

Bénéfice assigné dont la probabilité est $\frac{1}{2}$.	Nombre des parties.
50	5493
100	21975
200	87900
400	251603
600	791108
1000	2197522

Les parties, comme dans le cas précédent, seront jouées, quoi qu'il arrive. Si le jeu devait se terminer dès qu'un joueur a obtenu le bénéfice demandé ou dès qu'il ne peut plus déposer sa mise, les résultats seraient très différents.

91. Problème XLIX. — *Pierre et Paul font un nombre illimité de parties à un jeu dont les conditions sont équitables; leurs fortunes sont m et n. Quelle est, pour chacun d'eux, la probabilité de ruiner l'autre?*

Le jeu devant se prolonger jusqu'à la ruine de l'un des joueurs peut être assimilé à une seule partie dans laquelle celui qui risque m francs devrait, s'il est vainqueur, en obtenir $m+n$. L'espérance mathématique doit être égale à la mise, et, si l'on nomme p la probabilité pour que Pierre ruine son adversaire, l'équation

$$p(m+n) = m$$

donne

$$p = \frac{m}{m+n}.$$

On peut calculer directement cette probabilité. Soit y_x la probabilité pour que Pierre ruine Paul au moment où il possède x francs et où Paul, par conséquent, en possède $m+n-x$. On pourra écrire

$$y_x = p y_{x+b} + q y_{x-a}. \tag{5}$$

Lorsque Pierre, en effet, possède x francs, il peut, à la fin de la partie suivante, selon qu'il la gagne ou qu'il la perd, posséder

$x+b$ ou $x-a$ francs; il y a donc probabilité p pour que y_x se change en y_{x+b}, et probabilité q pour qu'il devienne y_{x-a}. L'équation (5) en est la conséquence.

pb étant égal à qa, puisque le jeu est équitable, la solution générale de l'équation (5) est

$$y_x = \alpha x + \beta,$$

α et β étant arbitraires.

Cette valeur de y_x satisfait, on le vérifie immédiatement, et, renfermant deux constantes arbitraires, elle est la solution la plus générale.

On a, pour déterminer les constantes,

$$y_0 = 0, \qquad y_{m+n} = 1,$$

on en déduit

$$\beta = 0, \qquad \alpha = \frac{1}{m+n}$$

et, par conséquent,

$$y_x = \frac{x}{m+n},$$

qui s'accorde avec la solution précédente.

Cette solution peut donner lieu à une difficulté. Si les enjeux a et b ne sont pas égaux à l'unité, l'un des joueurs pourra être forcé de cesser le jeu avant d'avoir tout perdu, possédant encore une somme inférieure à la mise exigée. Nous négligeons cette petite somme, qui, cependant, mettrait la formule en défaut dans le cas où elle formerait une partie notable de la fortune du joueur.

92. Problème L. — *Les conditions restant celles du problème précédent, on ne suppose plus les conditions du jeu équitables. Quelle est, pour chacun des joueurs, la probabilité de ruiner l'autre?*

Un ingénieux artifice de Moivre permet de déduire la solution de la théorie de l'espérance mathématique.

Donnons à Pierre, au lieu des m francs qu'il possède, m jetons de valeur $\alpha, \alpha^2, \ldots, \alpha^m$; α sera choisi ultérieurement. Remplaçons également les n francs de Paul par n jetons de valeur α^{m+1}, α^{m+2}, ..., α^{m+n} qui continuent la progression.

Le jeu se faisant dans les conditions supposées, la mise de Pierre sera a jetons, au lieu de a francs, et celle de Paul b jetons. On conviendra qu'à chaque partie Pierre exposera toujours les jetons dont la valeur est exprimée par les plus hautes puissances de α, et Paul, au contraire, ceux dont la valeur est représentée par les plus petites.

La série des jetons restant toujours la même, la séparation après chaque partie se fera en un point différent, mais Pierre aura toujours les premiers termes de la série, et Paul tous les suivants.

La chance, pour chaque joueur, de perdre tous ses jetons est indépendante de la valeur qu'on leur attribue et, par conséquent, du choix de α. Le jeu sera équitable si l'on pose

$$\frac{\alpha^{x+1}+\alpha^{x+2}+\ldots+\alpha^{x+a}}{\alpha^{x+a+1}+\alpha^{x+a+2}+\ldots+\alpha^{x+a+b}}=\frac{p}{q}.$$

Cette équation se réduit, quel que soit x, à

$$\frac{\alpha^a-1}{\alpha^{a+b}-\alpha^a}=\frac{p}{q},$$

c'est-à-dire, $p+q$ étant égal à 1,

$$(6)\qquad p\alpha^{a+b}-\alpha^a+q=0.$$

Le jeu, grâce à cet artifice, étant devenu équitable, l'espérance mathématique de chaque joueur doit être égale à sa mise ; et, si l'on nomme P_m la probabilité pour que Pierre, qui a m jetons, ruine Paul, qui en a n, on aura

$$P_m(\alpha+\alpha^2+\ldots+\alpha^{m+n})=\alpha+\alpha^2+\ldots+\alpha^m,$$

$$P_m=\frac{\alpha^m-1}{\alpha^{m+n}-1}.$$

93. Le calcul de P_m peut se faire directement.

Soit y_x la probabilité, au moment où Pierre possède x francs, pour qu'il finisse par ruiner Paul, on aura, comme (91),

$$y_x=py_{x+b}+qy_{x-a};$$

pb n'étant plus égal à qa, l'intégrale de cette équation est, comme on le vérifie aisément,

$$y_x = C_1 + C_2 \alpha^x, \tag{7}$$

C_1 et C_2 étant des constantes arbitraires et α satisfaisant à la condition

$$1 = p\alpha^b + q\frac{1}{\alpha^a},$$

c'est-à-dire à l'équation (6) déjà obtenue

$$p\alpha^{a+b} - \alpha^a + q = 0.$$

Les constantes C_1 et C_2 se détermineront par les conditions évidentes

$$y_0 = 0, \qquad y_{m+n} = 1,$$

et l'on trouve

$$y_m = \frac{\alpha^m - 1}{\alpha^{m+n} - 1}:$$

c'est le résultat déjà obtenu (92).

L'équation (6) a pour racine $\alpha = 1$ qui, évidemment, ne convient pas ou qui, plutôt, sert à former le premier terme de la formule (7), $C_1 1^x$.

94. Si l'on suppose $a = b = 1$ et $m = n$, la formule donne un résultat signalé par Huygens dans un cas particulier.

On trouve

$$y_m = \frac{p^m}{p^m + q^m}$$

et

$$1 - y_m = \frac{q^m}{p^m + q^m}.$$

Les chances de ruine pour les deux joueurs qui possèdent chacun m francs, et exposent 1^{fr} par partie à un jeu dans lequel les probabilités de gagner chaque partie sont p pour l'un et q pour l'autre, sont dans le rapport de p^m à q^m. On peut le démontrer directement.

Les deux joueurs ayant même fortune et les enjeux étant égaux, les successions de perte et de gain qui peuvent ruiner Pierre correspondent une à une aux successions de gain et de perte qui peuvent ruiner Paul; il suffit de changer les pertes en gains, et réciproquement, pour passer d'une série à l'autre. Dans l'une des séries, le nombre des parties surpassera de m celui des gains; dans l'autre, ce sera le contraire. Les probabilités des deux combinaisons sont donc entre elles dans le rapport de

$$\frac{p^{m+h}q^h}{q^{m+h}p^h} = \frac{p^m}{q^m}.$$

Ce rapport est celui des probabilités totales dont les termes sont en même nombre.

95. Problème LI. — *Pierre joue à un jeu équitable ou non, mais dans des conditions invariables d'une partie à l'autre, contre tout adversaire qui se présente. Quelle est la probabilité pour qu'il finisse par se ruiner?*

La solution de ce problème, déjà résolu en partie (85), peut se déduire des résultats précédents.

La chance de ruine est la même, évidemment, pour Pierre que s'il luttait contre un adversaire de fortune infinie.

Lorsque, Pierre possédant m francs, son adversaire en possède n, la probabilité pour que Pierre soit ruiné est (92)

$$\frac{\alpha^n - 1}{\alpha^{m+n} - 1}, \tag{8}$$

α étant la racine de l'équation

$$q\alpha^{a+b} - \alpha^b + p = 0;$$

p et q sont les probabilités de gain pour chaque joueur à chaque partie, a la mise de Pierre, b celle de son adversaire. Il faut, dans cette formule, supposer n infini. Trois cas peuvent se présenter : α est plus petit que l'unité, égal à l'unité ou plus grand que l'unité.

Si α est plus petit que l'unité, l'expression (8), en y supposant n infini, se réduit à l'unité. Il est certain que Pierre sera ruiné.

Si α est égal à l'unité, la formule prend la forme $\frac{0}{0}$, elle a pour valeur le rapport des dérivées de ses termes par rapport à α

$$\frac{n}{m+n},$$

égal à l'unité quand n est infini.

Dans ce cas, comme dans le précédent, la ruine de Pierre est certaine.

Lorsque α est plus grand que l'unité, la formule (8) a pour limite $\frac{1}{\alpha^m}$. La probabilité de la ruine de Pierre dans sa lutte contre un adversaire de fortune infinie n'est égale, dans ce cas, ni à zéro, ni à l'unité.

96. Il importe de chercher à quelles hypothèses correspondent les trois valeurs de α.

L'équation

$$q\,\alpha^{a+b} - \alpha^b + p = 0$$

a, dans tous les cas, la racine $\alpha = 1$. Pour que l'autre racine réelle et positive soit égale à l'unité, il faut que l'équation ait deux racines égales et que, par conséquent, α soit racine de l'équation dérivée

$$q(a+b)\alpha^{a+b-1} - b\alpha^{b-1} = 0.$$

Cette racine double étant égale à l'unité, on doit avoir

$$q(a+b) = b.$$

Le jeu, par conséquent, est équitable.

Pour celui qui joue indéfiniment à un jeu équitable, α est égal à l'unité et la ruine est certaine. Ce résultat a été obtenu (85).

Le seul cas où la ruine ne soit pas certaine est celui de α plus grand que l'unité : le jeu alors n'est pas équitable et les conditions favorisent celui des joueurs dont la fortune est limitée.

L'avantage, quelque petit qu'il soit, fait disparaître la certitude de ruine.

C'est le cas du banquier dans les jeux publics. Dans tout autre, sa ruine serait certaine.

Un avantage est pour lui juste et nécessaire; il importe seulement de ne pas l'exagérer. La chance de ruine

$$\frac{1}{\alpha^m},$$

lorsque α est plus grand que 1, est petite. On peut, dans les conditions où se placent habituellement les maisons de jeu, la considérer comme nulle.

97. Supposons, comme à la roulette ordinaire,

$$p = \frac{19}{37}, \qquad q = \frac{18}{37}.$$

Les mises étant supposées égales à 1^{fr}, α est donné par l'équation

$$q\alpha^2 - \alpha + p = 0:$$

une des racines, comme toujours, est égale à l'unité. C'est l'autre qu'il faut prendre; on a

$$\alpha = \frac{p}{q} = \frac{19}{18}.$$

La chance de ruine du banquier est donc

$$\left(\frac{18}{19}\right)^n,$$

n étant le rapport de l'avoir du banquier à la mise totale de l'un des coups.

Si l'on suppose $n = 1000$, on a

$$\left(\frac{18}{19}\right)^{1000} < \frac{1}{10^{24}}.$$

98. Problème LII. — *Pierre joue à un jeu dans lequel il a à chaque partie la probabilité p pour gagner et pour perdre la probabilité q. L'enjeu est 1^{fr} pour chacun des deux adversaires. Quelle est la probabilité pour que Pierre, qui possède m francs, soit ruiné précisément après avoir fait μ parties, de telle sorte que la $\mu^{\text{ième}}$ partie lui enlève son dernier franc?*

Pour que Pierre ait perdu m francs en μ parties, il faut qu'il ait perdu $\frac{\mu+m}{2}$ parties et gagné $\frac{\mu-m}{2}$.

Le nombre μ étant tel que ces deux fractions soient des nombres entiers, c'est-à-dire de même parité que m, la probabilité pour que, sur μ parties, Pierre en gagne $\frac{\mu-m}{2}$, est (54)

$$(9) \qquad \frac{1.2.3\ldots\mu}{1.2.3\ldots\frac{\mu-m}{2}.1.2.3\ldots\frac{\mu+m}{2}} p^{\frac{\mu-m}{2}} q^{\frac{\mu+m}{2}}.$$

Cette probabilité est plus grande que celle que nous cherchons. On compte, en effet, comme séries de parties faisant perdre Pierre en μ coups, toutes celles dans lesquelles, à la fin de la $\mu^{\text{ième}}$ partie, il est en perte de m francs. On doit exclure celles qui, avant de procurer la ruine de Pierre au $\mu^{\text{ième}}$ coup, l'ont procurée déjà à un coup antérieur.

Pierre une fois ruiné, en effet, le jeu doit cesser; il n'est pas admis à exposer l'argent qu'il n'a pas.

Le problème résolu (18) nous fait connaître le rapport du nombre des combinaisons qui ruineront Pierre en μ coups au nombre total de celles qui assurent en μ coups, joués quoi qu'il arrive, la perte de ses m francs.

Si l'on considère, en effet, l'une de ces combinaisons et que l'on range les pertes et les gains dans l'ordre où ils se sont produits, en les appelant en commençant par la dernière partie, l'excès du nombre des pertes sur celui des gains étant m, le rapport du nombre des combinaisons dans lesquelles, à aucun moment, les pertes et les gains ne seront en même nombre, au nombre total des combinaisons est (18) $\frac{m}{\mu}$, m étant la différence et μ la somme des nombres de parties gagnées et perdues. Le nombre des cas qui procurent la ruine de Pierre doit donc être multiplié par la fraction $\frac{m}{\mu}$. La probabilité pour que Pierre soit ruiné pour la première fois au $\mu^{\text{ième}}$ coup est

$$(10) \qquad \frac{m}{\mu} p^{\frac{\mu-m}{2}} q^{\frac{\mu+m}{2}} \frac{1.2.3\ldots\mu}{1.2.3\ldots\frac{\mu-m}{2}.1.2.3\ldots\frac{\mu+m}{2}}.$$

99. L'expression précédente peut être remplacée, lorsque p et q sont égaux à $\frac{1}{2}$, par une valeur approchée très simple. Le facteur qui multiplie $\frac{m}{\mu}$ est alors le terme dont le rang s'écarte de $\frac{m}{2}$ du plus grand, dans le développement de $(p+q)^\mu$; il a été calculé (58).

La probabilité de ruiner en μ coups précisément celui qui joue à un jeu équitable à 1fr la partie, et qui possède m francs, est

$$(11) \qquad \frac{m\sqrt{2}}{\mu\sqrt{\mu}\sqrt{\pi}} e^{-\frac{m^2}{2\mu}}.$$

100. La probabilité pour que la ruine s'accomplisse en μ coups précisément permet de calculer celle pour qu'elle ait lieu après le μ^{ieme} coup.

Cette probabilité est la somme des valeurs que prend l'expression (10) quand on y remplace successivement μ par les valeurs $\mu+2$, $\mu+4$, ... de même parité, seuls nombres possibles de parties qui puissent procurer la ruine.

Si $\varphi(\mu)$ désigne l'expression (10), la somme

$$\varphi(\mu)+\varphi(\mu+2)+\varphi(\mu+4)+\ldots$$

prolongée indéfiniment est la probabilité pour que la ruine de Pierre soit postérieure au $\mu^{\text{ième}}$ coup. Cette somme, pour de grandes valeurs de μ, peut être remplacée par

$$\frac{1}{2}\int_\mu^\infty \varphi(z)\,dz;$$

c'est-à-dire, d'après la valeur approchée de $\varphi(z)$,

$$\frac{m}{\sqrt{2\pi}}\int_\mu^\infty e^{-\frac{m^2}{2z}}\frac{dz}{z\sqrt{z}}.$$

La probabilité pour que Pierre soit ruiné avant le $\mu^{\text{ième}}$ coup est, par conséquent,

$$1-\frac{m}{\sqrt{2\pi}}\int_\mu^\infty e^{-\frac{m^2}{2z}}\frac{dz}{z\sqrt{z}}.$$

Si l'on pose $\frac{m^2}{2z} = t^2$, cette expression devient

$$1 - \frac{2}{\sqrt{\pi}} \int_0^{\frac{m}{\sqrt{2\mu}}} e^{-t^2}\, dt.$$

La Table des valeurs de la fonction

$$\frac{2}{\sqrt{\pi}} \int_0^t e^{-t^2}\, dt = \Theta(t)$$

se trouve à la fin du Volume.

On en déduit, en appliquant la formule au cas d'un joueur qui possède 100fr et dont la mise est de 1fr par partie, avec probabilité $\frac{1}{2}$ de gagner ou de perdre :

Probabilité d'être ruiné

Avant d'avoir fait	2000	parties............	0,026
»	4000	»	0,114
»	10000	»	0,3154

Le Tableau suivant, calculé à l'aide de la formule (10), que pour des petits nombres la formule approchée remplace mal, donne la probabilité de perdre 10fr en un nombre précis de parties inférieur à 100.

$$z_{2x} = \frac{1.2.3.4\ldots(2x-1).10}{1.2.3\ldots(x-5)\,1.2.3\ldots(x+5)}\left(\frac{1}{2}\right)^{2x}.$$

z_{10}...	0,00097656	z_{42}...	0,00901427	z_{72}...	0,00654715
z_{12}...	0,00244141	z_{44}...	0,00886698	z_{74}..	0,00640106
z_{14}...	0,00396730	z_{46}...	0,00870864	z_{76}...	0,00625906
z_{16}...	0,00534057	z_{48}...	0,00854268	z_{78}...	0,00612100
z_{18}...	0,00648498	z_{50}...	0,00837182	z_{80}...	0,00598692
z_{20}...	0,00739290	z_{52}...	0,00819821	z_{82}...	0,00585677
z_{22}...	0,00808596	z_{54}...	0,00802353	z_{84}...	0,00573047
z_{24}...	0,00859558	z_{56}...	0,00784911	z_{86}...	0,00560795
z_{26}...	0,00895373	z_{58}...	0,00767597	z_{88}...	0,00548910
z_{28}...	0,00918936	z_{60}...	0,00750490	z_{90}...	0,00537382
z_{30}...	0,00932720	z_{62}...	0,00733653	z_{92}...	0,00526203
z_{32}...	0,00938780	z_{64}...	0,00717130	z_{94}...	0,00515361
z_{34}...	0,00938780	z_{66}...	0,00700953	z_{96}...	0,00504846
z_{36}...	0,00934067	z_{68}...	0,00685150	z_{98}...	0,00494647
z_{38}...	0,00925727	z_{70}...	0,00669733	z_{100}..	0,00484754
z_{40}...	0,00914618				

101. Il est intéressant de chercher pour quel nombre μ de parties la probabilité de voir la ruine du joueur s'accomplir au μ^{ieme} coup précisément a la valeur maxima. L'expression de cette probabilité, en supposant les enjeux égaux à 1^{fr} et la probabilité de gagner chaque partie égale à $\frac{1}{2}$, est

$$\frac{1.2.3\ldots\mu}{1.2.3\ldots\frac{\mu-m}{2}\,1.2.3\ldots\frac{\mu+m}{2}}\frac{m}{\mu}\left(\frac{1}{2}\right)^{\mu};$$

si l'on change μ en $\mu+2$, elle se multiplie par

$$\frac{(\mu+1)(\mu+2)}{\left(\frac{\mu-m}{2}+1\right)\left(\frac{\mu+m}{2}+1\right)}\left(\frac{1}{2}\right)^2\frac{\mu}{(\mu+2)},$$

c'est-à-dire

$$\frac{\mu(\mu+1)}{(\mu+2)^2-m^2}.$$

L'expression augmente avec μ, tant que l'on a

$$\mu(\mu+1) > (\mu+2)^2 - m^2,$$

c'est-à-dire

$$3\mu < m^2 - 4,$$

et la probabilité maxima correspond à

$$\mu = \frac{m^2-4}{3},$$

si $m = 100$.

On a

$$\frac{m^2-4}{3} = \frac{9996}{3} = 3332.$$

La probabilité maxima est égale à $0,0000925$.

102. Si l'on remplaçait l'expression (10) par la valeur approchée

$$\frac{m\sqrt{2}}{\mu\sqrt{\mu\pi}}e^{-\frac{m^2}{2\mu}},$$

la valeur maxima s'obtiendrait en égalant à zéro la dérivée par rapport à μ; on trouve

$$\mu = \frac{m^2}{3}.$$

La différence des résultats est, par sa petitesse, une vérification de la formule approchée.

103. Problème LIII. — *Pierre et Paul jouent à un jeu de hasard. La probabilité de gagner chaque partie est p pour Pierre et q pour Paul. L'enjeu de Pierre est a francs, celui de Paul b francs. Pierre possède m francs, Paul n francs; le jeu est équitable. Quelle est la valeur probable du nombre des parties qui seront jouées avant la ruine de l'un des joueurs?*

En nommant y_x cette valeur probable lorsque Pierre possède x francs, c'est-à-dire l'espérance mathématique de celui qui aurait promesse de recevoir 1fr par partie jouée, on aura

$$y_x = 1 + p y_{x+b} + q y_{x-a}. \tag{12}$$

Il est clair, en effet, que le nombre des parties jouées comprend d'abord la partie par laquelle on commence, qui certainement aura lieu, et qu'après cette partie l'espérance mathématique cherchée est devenue y_{x+b} ou y_{x-a}, selon celui des deux joueurs qui a gagné.

La probabilité pour que la valeur de l'espérance mathématique devienne y_{x+b} étant p, et pour qu'elle devienne y_{x-a} étant q, l'équation (12) est la conséquence immédiate des principes.

La solution générale de l'équation (12) doit contenir deux constantes arbitraires, et ne peut évidemment en contenir davantage.

Posons

$$y = \alpha x^2 + \beta x + \gamma;$$

en supposant $pb = qa$, l'équation sera satisfaite si l'on pose

$$\alpha = -\frac{1}{ab},$$

β et γ restant arbitraires.

La solution générale est donc

$$y_x = -\frac{x^2}{ab} + \beta x + \gamma.$$

Les conditions évidentes

$$y_0 = 0,$$
$$y_{m+n} = 0$$

donnent

$$\gamma = 0,$$
$$\beta = \frac{m+n}{ab}$$

et

$$y_m = \frac{mn}{ab}.$$

Le nombre probable des parties est donc proportionnel au produit des fortunes des deux joueurs; il devient infini (88) lorsque l'une des fortunes est infinie.

104. M. Rouché a étendu la solution précédente au cas où le jeu n'est pas équitable. Il a résolu le problème suivant :

Problème LIV. — *Pierre et Paul jouent aux conditions énoncées dans le problème précédent; mais le jeu n'est pas équitable, la différence $pb - qa$ n'est pas nulle. Trouver la valeur probable du nombre des parties qui précéderont la ruine de l'un des joueurs.*

L'équation

$$y_x = 1 + p y_{x+b} + q y_{x-a}$$

définit, comme dans le cas précédent, le nombre probable des parties qui restent à jouer lorsque Pierre possède x francs et Paul $m+n-x$ francs.

On satisfait à cette équation, quelles que soient les constantes qui y figurent, excepté dans le cas traité précédemment, en posant

$$y_x = C\alpha^x + C'x + C'', \tag{13}$$

α étant la racine de l'équation

$$p\alpha^{a+b} - \alpha^a - q = 0.$$

et en prenant

$$C' = \frac{-1}{(a+b)p - a}.$$

Ces valeurs sont données par la substitution de (13) dans l'équation, en écrivant qu'elle devient identique. C et C″ restent arbitraires. Quelles que soient leurs valeurs, l'équation est satisfaite. On les déterminera par les conditions

$$y_0 = 0, \qquad y_{m+n} = 0.$$

En nommant P la probabilité calculée (92) pour que Pierre finisse par ruiner Paul, on trouve, en effectuant les calculs,

$$y_m = \frac{(m+n)P - m}{pb - qa}; \tag{14}$$

résultat élégant qui peut s'énoncer ainsi :

Le nombre probable des parties est égal au rapport de l'avantage total de l'un des joueurs à l'avantage du même joueur dans chaque partie.

$(m+n)P$ est en effet l'espérance mathématique du joueur qui a probabilite P de posséder l'enjeu total $(m+n)$, m est la fortune de ce joueur, et le numérateur est l'avantage qui résulte pour lui de la décision prise de continuer le jeu indéfiniment. Le dénominateur est, pour chaque partie jouée, l'excès de son espérance mathématique sur sa mise.

105. Il semble facile de démontrer ce théorème directement. Supposons $pb - aq$ positif, le jeu est avantageux au premier joueur. Soit n le nombre de parties qui seront jouées. Si p_1, p_2, ..., p_μ sont les probabilités pour que le jeu finisse en x_1, x_2, ..., x_μ parties, le nombre probable des parties est

$$p_1 x_1 + p_2 x_2 + p_3 x_3 + \ldots + p_\mu x_\mu. \tag{15}$$

Celui qui aura droit à une somme égale à (15) pourrait conclure avec des acheteurs différents des marchés équitables pour leur vendre en détail les avantages qui, suivant les cas, pourront pour lui résulter du jeu. Si le nombre des parties est x_1, un acheteur recevrait la promesse du bénéfice correspondant; il devra pour cela payer

$$p_1 x_1 + (pb - aq),$$

puisque chaque partie jouée équivaut pour Pierre à un avantage $pb - aq$.

Les ventes simultanées faites à des acheteurs différents ne peuvent faire naître aucune difficulté pour le règlement des comptes. Quel que soit, en effet, le nombre des coups joués, l'un des acheteurs se substituera au vendeur, et les autres n'auront rien à réclamer. La somme payée en échange de la totalité du gain espéré sera

$$(p_1 x_1 + p_2 x_2 + \ldots + p_\mu x_\mu)(pb - qa).$$

Cette somme est l'excès, sur la fortune de Pierre, de l'espérance mathématique résultant pour lui de la détermination de continuer le jeu jusqu'à la ruine de l'un des joueurs, à un jeu inégal dont les conditions lui sont avantageuses. Cette espérance mathématique est le produit de l'enjeu $(m + n)$ par la probabilité P de le gagner, et l'avantage du joueur est l'excès de cette espérance mathématique sur la somme qu'il possédait avant d'entrer au jeu, mais qui, une fois le jeu commencé, ne lui appartient plus.

On peut donc écrire

$$(p_1 x_1 + p_2 x_2 + \ldots + p_\mu x_\mu)(pb - qa) = \mathrm{P}(m + n) - m$$

et, par conséquent,

$$p_1 x_1 + p_2 x_2 + \ldots + p_\mu x_\mu = \frac{\mathrm{P}(m + n) - m}{pb - qa}.$$

C'est le théorème de M. Rouché.

106. Nous avons plusieurs fois signalé des raisonnements plausibles, qui, lorsqu'on y regarde de près, manquent de rigueur et conduisent à des conclusions fausses. Celui qui précède conduit à une formule exacte. On peut cependant élever contre lui une objection fondée.

Lorsqu'un joueur est admis à jouer une partie inégale, dont les conditions lui sont favorables, la probabilité de gagner la mise b de l'adversaire étant p, et celle de perdre une mise égale à a étant q, l'avantage de jouer une partie dans ces conditions a pour valeur équitable $pb - qa$; le droit de jouer un nombre x de parties doit être payé $x(pb - qa)$; c'est ce qu'il vaut.

Mais quand ce nombre x, au lieu d'être donné, est désigné comme le nombre de parties jouées jusqu'à la ruine de l'un des joueurs, ces parties, quoique le détail des pertes et des gains soit inconnu, ne présentent pas les mêmes chances que si l'on connaissait seulement leur nombre fixé à l'avance et les conditions du jeu. Si, par exemple, le nombre de ces parties est assez petit pour que la fin du jeu, que par hypothèse elles procurent, ait exigé la perte continuelle du second joueur, le droit, pour le premier, de jouer chaque partie ne vaut plus $pb - qa$, il vaut b.

Le calcul de M. Rouché était donc nécessaire, et le raisonnement, quoique très spécieux, qui conduit au résultat exact, n'est cependant pas rigoureux.

107. La démonstration (105) ne s'applique pas au cas où le jeu est équitable. L'expression (14) prend la forme $\frac{0}{0}$; on peut en trouver la vraie valeur.

La valeur probable du nombre des parties est

$$\frac{P(m+n)-m}{pb-qa};$$

on a (92)

$$P = \frac{1-\alpha^m}{1-\alpha^{m+n}},$$

α étant racine de l'équation

$$p\alpha^{a+b} - \alpha^a + q = 0.$$

Si l'on suppose $pb - qa = \varepsilon$, ε étant infiniment petit, α diffère infiniment peu de l'unité. Posons

$$\alpha = e^h.$$

La valeur de P devient

$$\frac{1-e^{mh}}{1-e^{(m+n)h}} = \frac{mh + \frac{1}{2}m^2h^2}{h(m+n) + \frac{h^2}{2}(m+n)^2 + \dots}$$

et, en négligeant le carré de h,

$$P = \frac{m}{m+n}\left(1 - \frac{hn}{2}\right);$$

l'expression du nombre probable de parties devient, par la substitution de cette valeur de P,

$$-\frac{hmn}{\varepsilon};$$

le rapport $\frac{h}{\varepsilon}$ est indépendant de m et de n : le nombre probable des parties est donc

$$Cmn,$$

C étant une constante; et, comme le nombre des parties est égal à l'unité quand on a $m = a$, $n = b$, il faut supposer $C = \frac{1}{ab}$. On retrouve le résultat déjà obtenu.

108. Dans le cas où les deux joueurs possèdent au début la même somme, on peut trouver la valeur probable du nombre des parties par une méthode très différente des précédentes.

Soit $\varphi(m)$ le nombre probable des parties lorsque chaque joueur possède m francs. Supposons que l'avoir de chacun soit doublé, le nombre probable des parties deviendra $\varphi(2m)$. Si l'on fait dans l'avoir de chaque joueur deux parts égales à m, on peut supposer que chacun expose d'abord, dans deux luttes séparées, la moitié m de ses francs contre la moitié de ceux de son adversaire. Après cette première série de parties, de deux choses l'une, l'un des joueurs a gagné deux fois, et l'autre est ruiné, ou bien chacun a gagné une série et ils se retrouvent tous deux avec $2m$ francs. La valeur probable du nombre des parties est alors, comme au début, $\varphi(2m)$. La probabilité pour que le premier joueur ruine son adversaire, lorsque tous deux possèdent m francs, est (92)

$$\frac{1}{1+\alpha^m}$$

et, pour qu'il soit ruiné,

$$\frac{\alpha^m}{1+\alpha^m}.$$

La probabilité pour que l'un des joueurs gagne une série et perde l'autre est le produit de ces deux probabilités, qu'il faut doubler puisqu'on ne dit pas dans quel ordre les événements doivent se succéder; on doit donc avoir

$$(16) \qquad \varphi(2m) = 2\varphi(m) + \frac{2\alpha^m}{(1+\alpha^m)^2}\varphi(2m),$$

α étant (92) la racine de l'équation

$$p\alpha^{a+b} - \alpha^a + q = 0.$$

L'équation (16) donne

$$(17) \qquad \varphi(2m)\frac{1+\alpha^{2m}}{(1+\alpha^m)^2} = 2\varphi(m).$$

Cette équation, si on la suppose vraie pour toute valeur de m, permet de déterminer la fonction de φ. Nous nous bornons à mentionner ce problème, qui n'intéresse pas le Calcul des probabilités.

109. La probabilité dans un nombre donné de parties de la ruine d'un joueur, dont l'adversaire est infiniment riche, a été donnée par Lagrange, puis par Ampère dans le Mémoire sur la théorie du jeu, lequel a été son début dans la Science.

Lorsque deux joueurs luttent l'un contre l'autre et que chacun peut ruiner son adversaire, le problème est très différent. La solution suivante, qui se présente d'abord, n'est pas exacte.

Soient m et n les fortunes des deux joueurs. Supposons le jeu équitable : la probabilité pour chaque joueur de gagner une partie est $\frac{1}{2}$, l'enjeu est 1^{fr}.

La probabilité pour que le premier joueur ruine le second est $\frac{m}{m+n}$, pour qu'il soit ruiné lui-même elle est $\frac{n}{m+n}$.

Si l'on sait que l'un des joueurs doit être ruiné, on peut, sans changer les chances, supposer à l'adversaire une fortune infinie. Si donc on nomme $\varphi(m, \mu)$ la probabilité pour qu'un joueur qui possède m francs soit ruiné, en μ coups précisément, par un adversaire dont la fortune est infinie, la probabilité pour que la partie engagée entre deux joueurs, dont l'un possède m francs et l'autre n francs, se termine en μ coups précisément sera

$$\frac{m}{m+n}\varphi(n, \mu)+\frac{n}{m+n}\varphi(m, \mu).$$

Le raisonnement n'est pas exact.

Si l'on sait que Pierre a été ruiné par un adversaire dont la fortune est finie, on en peut conclure que le hasard ne l'a pas favorisé : les combinaisons qui, débutant par un grand nombre de parties gagnées, auraient ruiné son adversaire doivent être exclues, celles dans lesquelles il gagne au début plus souvent qu'il ne perd sont rendues moins probables. La probabilité pour que la ruine se soit produite en μ coups n'est plus égale à $\varphi(m, \mu)$.

110. Problème LV. — *Pierre et Paul possèdent chacun* 2^{fr}, *ils jouent jusqu'à la ruine de l'un d'eux. La probabilité de gagner chaque partie étant* $\frac{1}{2}$ *et l'enjeu égal à* 1^{fr}, *quelle est la probabilité pour que le jeu se termine précisément en* 2μ *parties?*

Le nombre des parties doit être évidemment pair. Si le jeu n'est pas terminé, chaque joueur possédera 2^{fr}; car, en un nombre pair de parties, la perte de chacun est un nombre pair; si donc elle n'était pas nulle, le perdant serait ruiné.

Soit y_μ la probabilité pour que le jeu ne soit pas terminé en 2μ parties, on aura

$$y_{\mu+1}=\frac{1}{2}\,y_\mu. \tag{18}$$

Il est clair en effet que, si le jeu n'est pas terminé en 2μ parties, ce dont la probabilité est y_μ, pour qu'il ne le soit pas par les deux parties qui suivent, il faut que chacun des joueurs gagne une partie et perde l'autre. La probabilité pour qu'il en soit ainsi est $\frac{1}{2}$.

De l'équation (18), on conclut

$$y_\mu=\left(\frac{1}{2}\right)^\mu C, \tag{19}$$

et comme, pour $\mu=1$, on a $y_\mu=\frac{1}{2}$, C est égal à l'unité. La probabilité pour que le jeu ne soit pas terminé après 2μ parties est donc $(\frac{1}{2})^\mu$. Pour que le jeu se termine précisément en 2μ parties, il faut qu'il ne soit pas terminé en $2\mu-2$, ce dont la probabilité est $(\frac{1}{2})^{\mu-1}$, et que le même joueur perde les deux parties suivantes, ce dont la probabilité est $\frac{1}{2}$. Le produit $(\frac{1}{2})^{\mu}$ est la probabilité pour que le jeu se termine à la $2\mu^{\text{ième}}$ partie.

111. Problème LVI. — *Pierre et Paul possèdent chacun* 3^{fr}, *ils jouent dans les conditions définies dans l'énoncé précédent. Quelle est la probabilité pour que l'un des deux soit ruiné précisément au coup de rang* μ?

Le nombre de parties doit être impair. Après $2\mu+1$ parties, si le jeu n'est pas terminé, la perte et le gain seront un nombre impair, il doit être moindre que 3, il est donc l'unité : l'un des joueurs possédera 2^{fr} et l'autre 4^{fr}. Soit y_μ la probabilité pour que le jeu ne soit pas terminé au $(2\mu+1)^{\text{ième}}$ coup, on aura

$$y_{\mu+1} = \frac{3}{4} y_\mu. \tag{20}$$

Si, en effet, le jeu n'est pas terminé au $(2\mu+1)^{\text{ième}}$ coup, ce dont la probabilité est y_μ, il faut, pour qu'il ne le soit pas au $(2\mu+3)^{\text{ième}}$, ou que chacun des joueurs perde une partie et gagne l'autre, ou que celui qui a conservé 4^{fr} perde les deux parties; la probabilité pour que l'un ou l'autre de ces événements se produise est $\frac{3}{4}$.

De l'équation (20), on conclut

$$y_\mu = C\left(\frac{3}{4}\right)^{\mu}.$$

On a d'ailleurs

$$y_1 = \frac{3}{4};$$

car, pour que la partie soit terminée en trois coups, il faut que le même joueur gagne les trois premières parties, ce dont la probabilité est $\frac{1}{4}$. Il faut donc prendre $C=1$, et l'on a

$$y_\mu = \left(\frac{3}{4}\right)^{\mu}.$$

Pour que la partie se termine précisément au coup $2\mu+1$, il faut qu'elle ne le soit pas au coup $2\mu-1$ et que le joueur qui possède 2^{fr} seulement perde les deux parties suivantes, ce dont la probabilité est $\frac{1}{4}$.

La probabilité pour que le jeu se termine au coup de rang $2\mu + 1$ précisément est donc

$$\frac{1}{4}\left(\frac{3}{4}\right)^{\mu-1}.$$

112. M. Rouché a traité avec beaucoup d'élégance et de bonheur le cas général.

PROBLÈME LVII. — *Pierre et Paul jouent l'un contre l'autre avec des probabilités égales. Ils possèdent chacun* n *francs avant d'entrer au jeu; à chaque partie le perdant donne* 1fr *au gagnant, et le jeu ne cesse que lorsque l'un quelconque des deux joueurs est ruiné. Quelle est la probabilité* P *pour que le jeu se termine précisément à la fin d'une partie de rang assigné?*

Si, après μ parties, le jeu doit encore se continuer, c'est que l'état des fortunes est alors l'un des suivants :

$$(1, 2n-1), \quad (2, 2n-2), \quad \ldots, \quad (n, n);$$

la notation $(i, 2n-i)$ indiquant que l'un quelconque des deux joueurs possède i francs et, par suite, l'autre $2n-i$ francs.

Désignons par $\varphi_i(\mu)$ la probabilité pour que, après μ parties, le jeu ne soit pas encore terminé et que l'état des fortunes soit $(i, 2n-i)$. On aura, par le principe de la probabilité totale,

$$\varphi_i(\mu+1) = p_1^i\varphi_1(\mu) + p_2^i\varphi_2(\mu) + \ldots + p_n^i\varphi_n(\mu), \tag{21}$$

p_k^i désignant la probabilité de passer en une partie de l'état $(k, 2n-k)$ à l'état $(i, 2n-i)$.

Or on voit avec un peu d'attention que p_n^{n-1} est égal à 1, et que pour les autres valeurs des indices p_k^i est égal à $\frac{1}{2}$ ou à zéro, suivant que la valeur absolue de $i-k$ est égale à 1 ou diffère de 1. D'après cela, la formule (21) donne les n relations

$$\left\{\begin{array}{l} 2\varphi_n\;(\mu+1) = \varphi_{n-1}(\mu), \\ 2\varphi_{n-1}(\mu+1) = \varphi_{n-2}(\mu) + 2\varphi_n(\mu), \\ 2\varphi_{n-2}(\mu+1) = \varphi_{n-3}(\mu) + \varphi_{n-1}(\mu), \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots, \\ \quad 2\varphi_2(\mu+1) = \varphi_1(\mu) + \varphi_3(\mu), \\ \quad 2\varphi_1(\mu+1) = \varphi_2(\mu); \end{array}\right. \tag{22}$$

d'où il faut tirer $\varphi_1(\mu)$; car, cette quantité étant connue, on aura, pour la probabilité cherchée,

$$P = \frac{1}{2}\varphi_1(\mu - 1).$$

Or, si l'on adjoignait aux équations (22) celles qu'on obtient en changeant μ en $\mu + 1$ dans la seconde, μ en $\mu + 1$, $\mu + 2$ dans la troisième, ..., μ en $\mu + 1$, $\mu + 2$, ..., $\mu + n - 1$ dans la dernière, on aurait $\frac{1}{2}n(n+1)$ relations qui, par l'élimination des $\frac{1}{2}n(n+1) - 1$ quantités φ dont l'indice diffère de 1, conduiraient à une équation de la forme

$$\varphi_1(\mu + n) + A_1\varphi_1(\mu + n - 1) + \ldots + A_n\varphi_1(\mu) = 0. \tag{23}$$

D'ailleurs, si $a_1, a_2, \ldots, a_n$ désignent les racines de l'équation caractéristique que l'on obtient en faisant dans (23) $\varphi_1(\mu) = a^\mu$, l'expression générale de $\varphi_1(\mu)$ sera $\Sigma c_k a_k^\mu$, et, par suite, la valeur de P_μ sera $\frac{1}{2}\Sigma c_k a_k^{\mu-1}$, les constantes $c_1, c_2, \ldots, c_n$ étant déterminées par les conditions initiales du problème.

Cette détermination se fait élégamment de la façon suivante : d'abord, ni Pierre ni Paul ne peuvent être ruinés avant la fin de la $n^{\text{ième}}$ partie; d'autre part, la probabilité pour que le jeu cesse juste après la $n^{\text{ième}}$ partie, c'est-à-dire la probabilité pour que Pierre ou Paul perde n fois de suite, est

$$\frac{1}{2^n} + \frac{1}{2^n} = \frac{1}{2^{n-1}}.$$

De là résultent les relations linéaires

$$\Sigma c_k = 0, \quad \Sigma c_k a_k = 0, \quad \ldots, \quad \Sigma c_k a_k^{n-2} = 0, \quad \Sigma c_k a_k^{n-1} = \frac{1}{2^{n-2}}, \tag{24}$$

qui permettraient d'obtenir $c_1, c_2, \ldots, c_n$ et, par suite, P_μ en fonction des racines $a_1, a_2, \ldots, a_n$. Mais ce que nous voulons, c'est l'expression de P_μ en fonction des coefficients de l'équation caractéristique.

Or, si l'on pose

$$(z - a_1)(z - a_2)\ldots(z - a_n) = \theta(z)$$

et si l'on ajoute les équations (24) après les avoir multipliées respectivement par les coefficients de t_0, t_1, t^2, ..., t^{n-1} dans le quotient

$$\frac{\theta(z)-\theta(t)}{z-t},$$

on obtient

$$\Sigma\frac{c_k\theta(z)}{z-a_k}=\frac{1}{2^{n-2}},$$

ou, en multipliant par

$$\frac{z^{\mu-1}}{\theta(z)}$$

et désignant par $E(z)$ la partie entière du premier membre,

$$E(z)+\Sigma\frac{c_k a_k^{\mu-1}}{z-a_k}=\frac{1}{2^{n-2}}\frac{z^{\mu-1}}{\theta(z)}.$$

On voit par là que le coefficient de $\frac{1}{z}$ dans le développement du premier membre est $2P_\mu$; la probabilité cherchée P_μ est donc égale au coefficient de $\frac{1}{z}$ dans le développement de

$$\frac{z^{\mu-1}}{2^{n-1}\theta(z)}.$$

Il reste à trouver l'équation caractéristique. Au lieu de la chercher par le procédé laborieux ci-dessus indiqué, nous l'obtiendrons rapidement comme il suit : les relations (22) montrent que, si l'on pose

$$\varphi_1(\mu)=a^\mu,$$

on a

$$\varphi_k(\mu)=a^\mu\psi_k(a),\qquad \varphi_k(\mu+1)=a^{\mu+1}\psi_k(a);$$

de là résultent les équations

$$\begin{aligned}
x\psi_n+\psi_{n-1}&=0,\\
2\psi_n+x\psi_{n-1}+\psi_{n-2}&=0,\\
\psi_{n-1}+x\psi_{n-2}+\psi_{n-3}&=0,\\
\cdots\cdots\cdots\cdots&\cdots,\\
\psi_3+x\psi_2+\psi_1&=0,\\
\psi_2+x\psi_1&=0,
\end{aligned}$$

où x désigne la quantité $-2a$; et l'élimination des ψ donne

$$V_n \equiv \begin{vmatrix} x & 1 & 0 & & & & 0 \\ 2 & x & 1 & & & & 0 \\ 0 & 1 & x & 1 & & & 0 \\ \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot & & & \cdot\cdot \\ 0 & & & & 1 & x & 1 \\ 0 & & & & 0 & 1 & x \end{vmatrix} = 0.$$

Les valeurs évidentes

$$V_1 = x, \qquad V_2 = x^2 - 2,$$

et l'échelle de relation

$$V_n = x V_{n-1} - V_{n-2},$$

que l'on obtient immédiatement en développant le déterminant par rapport aux deux dernières lignes, prouvent que la fonction V_n est celle que l'on rencontre dans la théorie de la division du cercle en parties égales et qui a pour expression

$$V_n(x) = x^n - n x^{n-2} + \frac{n(n-3)}{1.2} x^{n-4} + \ldots$$
$$+ (-1)^r \frac{n(n-r-1)\ldots(n-2r+1)}{1.2\ldots r} x^{n-2r},$$

r désignant le plus grand nombre entier contenu dans $\frac{1}{2}n$.

Donc P_μ est le coefficient de $\frac{1}{z}$ dans le développement de

$$\frac{(-1)^n . 2 z^{\mu\ 1}}{V_n(-2z)};$$

ou, ce qui revient au même, en posant

$$-2z = \frac{1}{y},$$

P_μ est le coefficient de $y^{\mu-n}$ dans le développement de

$$\frac{(-1)^{\mu-n}}{2^{\mu-1}} \cdot \frac{1}{1 + A_1 y^2 + A_2 y^4 + \ldots},$$

A_i désignant le coefficient de x^{n-2i} dans $V_n(x)$ et, par suite, étant égal à zéro quand l'indice i surpasse $\frac{1}{2}n$.

Ce quotient a pour expression

$$\frac{(-1)^{\mu-n}}{2^{\mu-1}}(1+B_1 y^2+B_2 y^4+\ldots+B_k y^{2k}+\ldots),$$

B_k désignant le déterminant d'ordre k

$$B_k=(-1)^k\begin{vmatrix} A_1 & 1 & 0 & \ldots & 0 \\ A_2 & A_1 & 1 & \ldots & 0 \\ A_3 & A_2 & A_1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ A_k & A_{k-1} & A_{k-2} & \ldots & A_1 \end{vmatrix}.$$

On voit par là que, si $\mu-n$ est impair, P_μ est nul, et que, si $\mu-n$ est pair et égal à $2k$, on a

$$P_\mu=\frac{B_k}{2^{\mu-1}}.$$

Une vérification s'offre ici d'elle-même. Pour $n=2$ ou $n=3$, le déterminant B_k se réduit à son terme principal, et l'on retrouve immédiatement les résultats obtenus directement aux nos 110 et 111.

113. On peut rattacher au problème de la ruine des joueurs l'espérance, acceptée souvent, d'accroître les chances de gain par une ingénieuse disposition des mises et du nombre des parties jouées. Toutes ces combinaisons sont illusoires. Si les conditions du jeu rendent pour chaque partie l'espérance mathématique égale à la mise, l'égalité existera, quoi qu'on fasse, quel que soit le nombre des parties.

Nous nous bornerons à examiner un procédé plausible pour accroître les chances de gain en laissant celles de perte petites.

Un joueur entre au jeu avec la résolution de continuer tant que le sort lui sera favorable et de se retirer après sa première perte. Le nombre des parties qu'il jouera peut être illimité et le bénéfice immense; la perte, au contraire, sera nécessairement petite, égale tout au plus à la mise pour une seule partie.

Lorsque l'on compare, cependant, les chances de perte à celles de gain, on les trouve, comme cela doit être, équivalentes, lorsque les conditions du jeu, à chaque partie, sont équitables.

Soient

p la probabilité de gagner une partie;
a la somme à recevoir;
q la probabilité de perdre;
b la somme à payer dans ce cas;

on a

$$p + q = 1.$$

Le joueur a une probabilité q de perdre la somme b; une probabilité pq de gagner $a - b$; une probabilité p^2q de gagner $2a - b$, etc.; une probabilité p^nq de gagner $na - b$. Son espérance mathématique, au moment où il se met au jeu, est donc

$$pq(a - b) + p^2q(2a - b) + \ldots + p^nq(na - b) + \ldots - qb,$$

c'est-à-dire

$$pqa(1 + 2p + 3p^2 + \ldots) - pqb(1 + p + p^2 + \ldots) - qb$$

et, en remplaçant les deux séries par leurs valeurs $\frac{1}{(1-p)^2}$ et $\frac{1}{(1-p)}$,

$$\frac{pqa}{(1-p)^2} - \frac{pqb}{1-p} - qb = \frac{pa}{q} - b,$$

c'est-à-dire zéro si le jeu est équitable.

L'avantage prétendu de la combinaison se réduit à accroître la valeur possible du gain, en diminuant proportionnellement la probabilité de l'obtenir.

CHAPITRE VII.

PROBABILITÉ DES CAUSES.

Un jour, à Naples, un homme de la Basilicate, en présence de l'abbé Galiani, agita trois dés dans un cornet et paria d'amener rafle de 6, il l'amena sur-le-champ. Cette chance est possible, dit-on, l'homme réussit une seconde fois, et l'on répéta la même chose, il remit les dés dans le cornet trois, quatre, cinq fois, et toujours rafle de 6 « *Sangue di Barco!* s'écria l'abbé, les dés sont pipés! » et ils l'étaient

DIDEROT.

114. Étudier les faits pour remonter aux causes est le but le plus élevé de la Science. Notre curiosité est ici moins ambitieuse. Nous n'aurons dans ce Chapitre aucune loi de la nature à discuter, aucune énigme à résoudre. Les causes sont pour nous des accidents qui ont accompagné ou précédé un événement observé. Le mot n'implique pas qu'au sens philosophique l'événement soit un effet produit par la cause.

Pierre a parié d'amener avec trois dés un point supérieur à 16 ; il a gagné : tel est l'événement. Le point amené peut être 17 ou 18 : telles sont les causes possibles du succès.

On a, sous un même nom, réuni plusieurs cas distincts : ce sera, par exemple, un point plus grand que 16 qui peut être 17 ou 18 ; la sortie, au loto, d'un numéro divisible par 5 qui peut être 5, 10, 15, ... ; le tirage, dans une urne, d'une boule de couleur qui peut être rouge, verte ou jaune ; la retourne, dans une partie de cartes, d'une figure qui peut être roi, dame ou valet. L'événement s'est produit, on le sait ; les causes qui, dans ces différents cas, restent possibles sont les manières diverses dont il a pu se présenter. On sait qu'une rivière a débordé en Espagne : c'est à la Géographie, non à la Météorologie, qu'il appartient, par l'énumération des cours d'eau, de faire connaître la diversité possible des causes.

De tels cas sont les plus simples. Les causes dont on cherche la probabilité ne peuvent, si elles agissent, produire qu'un seul événement dont l'arrivée, supposée connue, diminue le dénominateur de la probabilité sans en changer le numérateur. On sait, par exemple, qu'en donnant les cartes on a retourné une figure. Quelle est la probabilité pour qu'elle soit un roi ? Le numérateur de la probabilité est le même qu'avant le renseignement donné : c'est le nombre des rois ; mais le dénominateur, nombre des cartes possible, a diminué : les figures seules doivent y être comptées.

On sait qu'au loto le numéro sorti est divisible par 5, on demande la probabilité pour qu'il soit 25. Le numérateur de cette probabilité est 1, comme avant le renseignement donné ; mais le dénominateur, qui était 90, nombre total des numéros, est devenu 18, nombre des multiples de 5.

Le numérateur de la probabilité, dans d'autres cas, est changé, en même temps que le dénominateur, par la connaissance de l'événement observé. Quelquefois aussi, les cas restés possibles ne sont pas également vraisemblables.

Deux urnes, par exemple, sont d'aspect identique : l'une contient une boule blanche et une boule noire ; l'autre, dix boules noires et une blanche. On choisit une des urnes, on en fait sortir une boule, elle est blanche ; quelle est la probabilité d'avoir choisi la première urne ?

Deux causes sont possibles : la première urne et la seconde urne ; mais, contrairement à ce qui avait lieu dans les cas précédents, aucune de ces deux causes, supposée véritable, ne rend certain l'événement observé. On pourrait considérer comme causes possibles les deux boules blanches qui ont pu sortir, l'une de la première urne, l'autre de la seconde ; mais ces boules n'ont pas même vraisemblance. La seconde, associée à dix autres boules, a moins de chances de sortir que la première et sortira certainement moins souvent si l'épreuve est renouvelée un grand nombre de fois.

115. Le problème général peut s'énoncer comme il suit :

Diverses causes $E_1, E_2, \ldots, E_n$ *ont pu produire un événement observé. Les probabilités de ces causes, lorsque le résultat n'était pas encore connu, étaient* $\varpi_1, \varpi_2, \ldots, \varpi_n$. *L'événement se produit; la cause* E_i, *lorsqu'on est certain que c'est elle qui* agit, *donne à l'événement la probabilité* p_i. *Quelle est la probabilité de chacune des causes qui sont, on l'admet, les seules possibles*?

Le type des problèmes dont nous parlons peut être représenté par une urne contenant des boules blanches et des boules noires. L'événement est la sortie d'une boule ; elle est blanche, on le sait. Mais chaque boule est marquée par un des numéros 1, 2, 3, ..., n. Quelle est la probabilité pour que la boule blanche sortie soit marquée d'un numéro donné ? Ces numéros représentent ici ce que l'on nomme les *causes* possibles de l'événement, sans avoir rien de commun, bien entendu, avec l'idée de *causalité*.

Si μ désigne le nombre total des boules, μ_i le nombre de celles qui sont marquées i et, dans ces μ_i, m_i le nombre des boules blanches, le nombre total des boules blanches est

$$m_1 + m_2 + m_3 + \ldots + m_\mu.$$

Elles sont toutes également possibles, puisque, placées dans la même urne, chacune, considérée individuellement, a chance égale de sortie. La probabilité pour que la boule blanche que l'on a tirée et dont on n'a pas vu le numéro soit marquée d'un i est

$$\frac{m_i}{m_1 + m_2 + \ldots + m_\mu}.$$

Telle est la solution du problème. Il reste à l'exprimer en fonction des données. On a

$$\frac{m_i}{\mu_i} = p_i, \qquad \frac{\mu_i}{\mu} = \varpi_i;$$

par conséquent,

$$m_i = p_i \mu_i = p_i \varpi_i \mu.$$

La probabilité est donc, en supprimant le facteur μ,

$$\frac{p_i \varpi_i}{p_1 \varpi_1 + p_2 \varpi_2 + \ldots + p_n \varpi_n}.$$

Le dénominateur est le même pour toutes les valeurs de i, et les probabilités des diverses causes sont proportionnelles, par conséquent, aux produits de la probabilité de chacune, avant l'événement (ϖ_i), par la probabilité qu'elle donne à l'événement (p_i) quand on la suppose certaine.

116. La démonstration peut se faire autrement.

La probabilité cherchée est celle pour que l'événement qui est arrivé, on le sait, soit dû à la cause représentée par l'indice i.

La probabilité pour que, avant l'épreuve, l'événement en question se produisît et fût dû à la cause désignée est un événement composé, et cela de deux manières :

1° Il faut que la cause soit mise en jeu ;
2° Il faut qu'elle produise l'événement.

Ou bien :

1° Il faut que l'événement se produise;
2° Il faut que, étant produit, il soit dû à la cause désignée.

On en déduit deux expressions de la même probabilité

$$\varpi_i p_i = (p_1\varpi_1 + p_2\varpi_2 + \ldots + p_n\varpi_n)x,$$

et, par conséquent, la probabilité x pour que l'événement, étant produit, soit dû à la cause désignée par l'indice i est celle qui a été obtenue (115)

$$x = \frac{p_i\varpi_i}{p_1\varpi_1 + p_2\varpi_2 + \ldots + p_n\varpi_n}. \tag{1}$$

117. Problème LVIII. — *Une urne contient* μ *boules : les unes sont blanches, les autres noires, on ignore en quelle proportion. On tire* k *boules, en remettant à chaque fois la boule sortie. Il ne sort que des boules blanches. Quelle est la probabilité pour que l'urne ne contienne que des boules blanches?*

La question est mal posée.

On ignore, dit l'énoncé, la proportion dans l'urne des boules blanches et des boules noires. Toutes les hypothèses sont possibles. Il faudrait dire, en outre, quelle est, *a priori,* la probabilité de chacune. Si, toutes les combinaisons possibles ayant été préparées dans des urnes d'apparence identique, le hasard a décidé entre elles, les conditions sont autres que si l'on a puisé au hasard dans une urne de composition convenue, pour composer avec les boules ainsi tirées l'urne nouvelle dont nous parlons.

Nous admettrons d'abord, pour préciser la question, que toutes les compositions de l'urne soient, *a priori,* également possibles. Toutes restent possibles après l'épreuve, à l'exception d'une réunion de boules noires, mais les probabilités ne sont plus égales.

La combinaison dans laquelle, sur μ boules, le nombre des blanches est n donne à l'événement observé la probabilité

$$\left(\frac{n}{\mu}\right)^k.$$

Les probabilités désignées par ϖ_1, ϖ_2, ..., ϖ_n dans l'énoncé général sont supposées égales entre elles; en les supprimant comme facteur commun dans la formule (1), on trouve la probabilité pour que le nombre des boules blanches soit n

$$\frac{n^k}{1+2^k+3^k+\ldots+\mu^k}.$$

La probabilité pour que toutes les boules de l'urne soient blanches est donc

$$\frac{\mu^k}{1+2^k+3^k+\ldots+\mu^k}.$$

Si l'on suppose, par exemple, $\mu=5$, $k=6$, après avoir tiré six fois de suite une boule blanche d'une urne qui contient cinq boules, la probabilité pour que les cinq boules soient blanches est

$$\frac{5^6}{1+2^6+3^6+4^6+5^6}=0,76163.$$

118. Si, au lieu de supposer toutes les combinaisons également possibles *a priori*, on avait composé l'urne en tirant au sort, à pile ou face par exemple, la couleur de chaque boule, le problème serait très différent.

Les hypothèses possibles sur la composition de l'urne, au lieu d'être également vraisemblables *a priori*, ont les probabilités suivantes :

5 blanches ou 5 noires,

$$1\left(\frac{1}{2}\right)^5=0,03125;$$

4 blanches et 1 noire ou 4 noires et 1 blanche,

$$5\left(\frac{1}{2}\right)^5=0,15625;$$

3 blanches et 2 noires ou 3 noires et 2 blanches,

$$10\left(\frac{1}{2}\right)^5=0,3125.$$

Les probabilités désignées par $\varpi_1, \varpi_2, \ldots, \varpi_n$ dans la formule (1) sont proportionnelles aux nombres 1, 5, 10, 10, 5, 1, et la probabilité, quand six fois de suite on a extrait une boule blanche, pour que les cinq boules de l'urne soient blanches est

$$\frac{5^6}{5+10.2^6+10.3^6+5.4^6+5^6} = 0{,}35479.$$

119. Si, après avoir extrait les boules de l'urne, on ne les remettait pas, les résultats seraient différents.

Il est clair, d'abord, qu'on ne peut, dans cette hypothèse, extraire plus de cinq boules, et que, si on les extrait toutes les cinq, il n'y a plus de problème.

Supposons donc que l'événement observé soit la sortie de quatre boules blanches; les probabilités, *a priori*, des diverses compositions de l'urne étant proportionnelles à 1, 5, 10, 10, 5, 1, cherchons la probabilité pour que la cinquième boule qui reste dans l'urne, la seule que l'on n'ait pas vue, soit blanche.

L'événement observé est la sortie de quatre blanches. Les hypothèses possibles lui donnent pour probabilités : 1, $\frac{1}{5}$, 0, 0, 0, 0.

La formule (1) devient, en y substituant les valeurs de ϖ_i et de p_i,

$$\frac{1}{1+5\times\frac{1}{5}+0} = \frac{1}{2}.$$

Ce résultat est évident *a priori*.

Après avoir vu quatre des boules, sachant que pour les cinq la couleur a été tirée au sort, on n'a acquis sur la dernière aucun renseignement. Les circonstances pour elle sont les mêmes que si, lorsque l'on procédait à la formation de l urne, le hasard avait désigné la couleur blanche aux quatre premières épreuves. On ne devrait y voir aucune raison pour qu'il la désignât une cinquième fois.

120. Problème LVIII *bis*. — *Une urne contient des boules noires ou blanches en proportion inconnue. On y fait μ tirages, en remettant dans l'urne, après chaque tirage, la boule qui en est sortie. On a obtenu m boules blanches et n boules noires. Quelle est la composition la plus probable de l'urne?*

L'énoncé, comme celui du problème précédent, n'est pas suffisamment précis.

Toutes les hypothèses sur la composition de l'urne étaient possibles avant l'épreuve. L'étaient-elles également? Il est nécessaire de le dire. Nous le supposerons d'abord.

Soit x la probabilité assignée à la sortie d'une boule blanche par la composition de l'urne.

La probabilité de l'événement observé, la sortie de m boules blanches et de n noires, est

$$x^m(1-x)^n.$$

Le nombre des combinaisons qui peuvent se présenter est indépendant de x, et la probabilité de l'événement observé, que l'on connaisse ou non l'ordre de sortie des boules, est proportionnelle à

$$x^m(1-x)^n.$$

Ce produit doit remplacer la probabilité désignée par p_i dans la formule (1); les probabilités ϖ_i sont supposées égales entre elles, et la probabilité de chaque valeur supposée pour x, proportionnelle au produit $p_i\varpi_i$, est, dans le cas actuel, proportionnelle à

$$x^m(1-x)^n.$$

La valeur de x la plus probable rendra ce produit maximum.

En égalant la dérivée à zéro, on trouve

$$\frac{x}{m} = \frac{1-x}{n}.$$

La composition la plus probable est celle qui rend les probabilités de sortie des boules blanches ou noires proportionnelles aux nombres de fois qu'elles se sont montrées.

121. Chaque hypothèse sur la valeur de x a une probabilité. Nous devons en chercher la loi. Il ne peut être question d'assigner la valeur de l'une de ces probabilités. Toutes les hypothèses ayant été supposées possibles et leur nombre étant infini, la probabilité de l'une d'elles, rigoureusement désignée, est 0; mais la probabilité pour que x soit compris entre z et $z+dz$ est proportionnelle à dz.

La loi des probabilites est celle des valeurs de la fonction

$$x^m(1-x)^n,$$

à laquelle elles sont proportionnelles.

Pour étudier cette fonction dans le voisinage du maximum, posons

$$x = \frac{m}{m+n} + \varepsilon,$$

$$1-x = \frac{n}{m+n} - \varepsilon.$$

Le maximum étant

$$\left(\frac{m}{m+n}\right)^m \left(\frac{n}{m+n}\right)^n = p^m q^n,$$

la valeur voisine sera, en posant $x = p - \varepsilon$, $1 - x = q + \varepsilon$,

$$(p-\varepsilon)^m (q+\varepsilon)^n = p^m q^n \left(1 - \frac{\varepsilon}{p}\right)^m \left(1 + \frac{\varepsilon}{q}\right)^n;$$

$p^m q^n$ étant indépendant de ε, la probabilité est proportionnelle au produit

$$(2) \qquad \left(1 - \frac{\varepsilon}{p}\right)^m \left(1 + \frac{\varepsilon}{q}\right)^n$$

ou à

$$1 - \left(\frac{m}{p} - \frac{n}{q}\right)\varepsilon + \left[\frac{m(m-1)}{p^2} - \frac{2mn}{pq} + \frac{n(n-1)}{q^2}\right]\frac{\varepsilon^2}{2},$$

en négligeant les puissances de ε supérieures à la seconde. Or, à cause de

$$\frac{m}{p} = \frac{n}{q} = \frac{m+n}{p+q} = m+n,$$

l'expression précédente peut être remplacée par

$$1 - \frac{(m+n)}{2pq}\varepsilon^2$$

ou, au même degré d'approximation, par

$$e^{-\frac{\varepsilon^2(m+n)}{2pq}},$$

et, par conséquent, la probabilité pour que la composition de l'urne donne à la sortie d'une boule blanche la probabilité

$$p = \frac{m}{m+n} + \varepsilon$$

et à celle d'une boule noire

$$q = \frac{n}{m+n} - \varepsilon$$

est proportionnelle à

$$e^{-\frac{\varepsilon^2(m+n)}{2pq}},$$

et peut être représentée par

$$\mathrm{G}e^{-\frac{\varepsilon^2(m+n)}{2pq}},$$

G étant indépendant de ε.

Cette formule équivaut à celle qui a été trouvée (58). On a, dans les deux cas, obtenu n fois sur μ épreuves un événement dont la probabilité est p. La différence ε, égale à $\frac{n}{\mu} - p$, est remplacée par h égale à $n - p\mu$. La probabilité d'une valeur désignée de h est proportionnelle à

$$e^{-\frac{h^2}{2\mu pq}}.$$

La seule différence des deux théorèmes consiste en ce que, dans un cas (58), p est donné exactement, le doute porte sur la valeur de n; dans la formule actuelle (121), n est donné exactement, le doute porte sur la valeur de p.

122. La formule précédente est déduite d'une hypothèse qui se réalisera rarement. Toutes les probabilités désignées par x ont, en général, *a priori*, des valeurs inégales.

Problème LIX. — *Une urne contient* N *boules. On a tiré au sort la couleur noire ou blanche, avec probabilité* $\frac{1}{2}$ *pour chacune des boules. Sur* μ *tirages, faits dans l'urne ainsi com-*

posée, on obtient m boules blanches et n noires. Quelle est la composition la plus probable de l'urne?

La probabilité pour que, dans une urne ainsi composée, le nombre des boules blanches soit $\frac{N}{2} - z$ est approximativement (58), si N est grand et z petit,

$$\frac{\sqrt{2}}{\sqrt{\pi N}} e^{-\frac{2z^2}{N}}.$$

La probabilité de la sortie d'une boule blanche est, dans cette hypothèse,

$$\frac{1}{2} - \frac{z}{N}.$$

Posons $\frac{z}{N} = y$; la probabilité d'une valeur désignée de y est proportionnelle à

$$e^{-\frac{2z^2}{N}} = e^{-2Ny^2}. \tag{4}$$

La probabilité de l'événement observé est proportionnelle à

$$\left(\frac{1}{2} - y\right)^m \left(\frac{1}{2} + y\right)^n. \tag{5}$$

Les probabilités désignées par ϖ_i et p_i dans la formule générale (116) doivent être remplacées par (4) et (5).

La probabilité de la cause, c'est-à-dire de la valeur y, est proportionnelle au produit

$$e^{-2Ny^2} \left(\frac{1}{2} - y\right)^m \left(\frac{1}{2} + y\right)^n. \tag{6}$$

En égalant à zéro la dérivée du logarithme, on obtient, pour déterminer la valeur de y qui rend (6) maximum, l'équation

$$-2Ny - \frac{m}{1-2y} + \frac{n}{1+2y} = 0$$

ou, comme y est petit,

$$-2Ny - m(1+2y) + n(1-2y) = 0,$$

d'où

$$y = \frac{n-m}{2(N+m+n)}.$$

La plus grande probabilité correspond à

$$y = \frac{n - m}{2(\mathrm{N} + m + n)}.$$

La composition la plus probable de l'urne donne à la sortie d'une blanche la probabilité

$$\frac{1}{2} - \frac{n - m}{2(\mathrm{N} + m + n)} = \frac{\mathrm{N} + 2m}{2(\mathrm{N} + m + n)}. \tag{7}$$

Cette fraction est comprise entre $\frac{1}{2}$ et $\frac{m}{m+n}$. On aurait pu le prévoir. Avant le tirage d'aucune boule, les chances pour les deux couleurs étaient égales, le rapport le plus vraisemblable était celui qui donne la probabilité $\frac{1}{2}$. Si le tirage indique pour l'une des couleurs la proportion de m à $m + n$, c'est une raison pour croire au même rapport dans l'ensemble des boules. Si ces deux indications ne s'accordent pas, la probabilité la plus plausible est entre les deux.

Si N est très grand, la formule (7) est très voisine de $\frac{1}{2}$, quels que soient les nombres m et n; si, au contraire, m et $m + n$ sont très grands, elle est voisine de $\frac{m}{m+n}$, quel que soit N.

Ces conclusions du calcul pouvaient également se prévoir.

Si le nombre N est très grand, on a fait, pour choisir les couleurs des boules de l'urne, un très grand nombre d'épreuves, donnant chacune à la couleur blanche une probabilité $\frac{1}{2}$. Il est certain, d'après le théorème de Bernoulli, que le rapport du nombre des boules blanches à celui des boules de l'urne diffère peu de $\frac{1}{2}$. Cette certitude est assez grande pour ne pas être notablement amoindrie par les couleurs, quelles qu'elles soient, de quelques boules tirées de l'urne. Si cependant, après un nombre immense d'essais, on trouve entre le nombre des boules blanches et le nombre des boules sorties un rapport très différent de $\frac{1}{2}$, on se trouvera en présence de deux certitudes inconciliables.

Nous adoptons, on le voit, le sens vulgaire du mot *certitude*. On tire au sort mille fois entre la couleur blanche et la couleur noire, en leur donnant des probabilités égales. Dans l'urne contenant les boules dont les couleurs sont ainsi désignées se trouveront, à très peu près, autant de boules blanches que de boules noires :

on peut le tenir pour certain. Sur $m+n$ tirages, on obtient m boules blanches; le rapport du nombre des boules blanches au nombre total diffère peu de $\frac{m}{m+n}$, *on peut aussi le tenir pour certain.* Les deux rapports cependant sont très inégaux. On est évidemment dans un cas exceptionnel, possible assurément, mais fort rare.

Supposons, par exemple, $N = 1000$. Dans l'urne composée de 1000 boules, contenant vraisemblablement, d'après la manière dont elles ont été choisies, 500 blanches environ, on fait 4 tirages. On tire 4 boules blanches. La composition la plus probable de l'urne, d'après la formule (7), est telle que le rapport du nombre des boules blanches au nombre total soit

$$\frac{1008}{2008} = \frac{126}{251}.$$

La démonstration supposant un grand nombre d'épreuves faites dans l'urne n'est plus applicable, il est vrai, au cas où le nombre $m+n$ se réduit à 4. Le résultat est cependant peu différent du véritable. On a vu 4 boules blanches; amenées par le hasard, elles sont presque certainement différentes; on ne sait rien sur les 996 autres. Le nombre des boules blanches le plus vraisemblable est, pour cette portion de l'urne, 498; cela fait, en tout, 502 pour le nombre le plus probable des boules blanches et pour probabilité la plus vraisemblable

$$\frac{502}{1000} = \frac{251}{500}.$$

Si, dans la même urne composée de 1000 boules, on a fait 40000 tirages et obtenu 22000 boules blanches, le rapport le plus probable donné par la formule (7) est

$$\frac{45000}{81000} = \frac{5}{9},$$

peu différent du rapport $\frac{22}{40}$ indiqué par le résultat du tirage; les motifs qu'on avait d'abord de croire à un rapport voisin de $\frac{1}{2}$ se trouvent en quelque sorte annulés par les 40000 épreuves qui les contredisent. Tous ces chiffres, nous devons le répéter, sont possibles, mais absolument invraisemblables.

La probabilité pour que, sur 1000 boules dont la couleur a été désignée par le sort, avec chance égale pour blanc et pour noir, le nombre des blanches soit inférieur à 550 est (65)

$$\frac{1}{2}+\frac{1}{\sqrt{\pi}}\int_0^{\frac{50}{\sqrt{500}}} e^{-t^2}\,dt = \frac{1}{2}+\frac{1}{2}\Theta(\sqrt{5}).$$

Deux cas, en effet, sont possibles : ou les noires sont en majorité, ce dont la probabilité est $\frac{1}{2}$; ou l'écart est positif et compris entre 0 et 50.

On a

$$\frac{50}{\sqrt{500}} = \sqrt{5} = 2{,}2361,$$

$$\Theta(2{,}23) = 0{,}99838,$$

$$\frac{1}{2}+\frac{1}{2}\Theta(2{,}23) = 0{,}9991.$$

Il y a plus de mille à parier contre un, *a priori*, qu'un tel écart ne se produira pas. Nos hypothèses cependant le rendent probable. La sortie de 22000 blanches sur 40000 tirages dans une urne contenant nombre égal de blanches et de noires présenterait une anomalie plus singulière encore. Si la probabilité de tirer 1 boule blanche est $\frac{1}{2}$, la probabilité d'en obtenir moins de 22000 sur 40000 tirages est

$$\frac{1}{2}+\frac{1}{\sqrt{\pi}}\int_0^{\frac{2000}{\sqrt{20000}}} e^{-t^2}\,dt = \frac{1}{2}+\frac{1}{2}\Theta(14).$$

$\Theta(14)$ est tellement voisin de l'unité, que l'événement doit être considéré comme certain, l'événement contraire comme impossible. On pourrait renouveler l'essai des milliards de milliards de fois sans avoir chance d'obtenir, sur 40000 épreuves, 22000 fois un événement dont la probabilité est $\frac{1}{2}$.

123. On a assimilé sans raison au problème précédent des questions en réalité fort différentes.

Lorsque les observations démentent des prévisions dont la probabilité semblait grande, on présume, naturellement, l'influence d'une cause perturbatrice et l'on est conduit à chercher la probabilité de son existence.

La question est insoluble. On n'a pas, d'une part, les données nécessaires. Le dilemme, d'autre part : ou il existe une cause, ou il n'en existe pas, n'a pas la netteté promise par la forme de l'énoncé.

Qu'entend-on, si l'on dit : il y a une cause?

On a jeté une pièce de monnaie 1000 fois, elle a montré face 510 fois. Quelle est la probabilité pour que cet écart soit dû au hasard ou pour qu'il résulte de l'imperfection de la pièce?

Les hypothèses possibles sont en nombre infini.

La pièce peut être parfaite.

Elle peut donner à l'arrivée de face une probabilité quelconque plus grande ou plus petite que $\frac{1}{2}$.

L'événement observé, l'arrivée de 510 fois face sur 1000 coups, est compatible avec toutes les hypothèses : il se peut que, la pièce étant parfaite, le hasard ait amené ce petit écart; que, la pièce favorisant l'arrivée de face de manière à rendre l'écart le plus probable plus petit que 10, le hasard ait complété la différence; que la pièce rende probable un écart plus grand, beaucoup plus grand même que 10, ou que, inégale en sens opposé, elle donne probabilité à l'arrivée de pile plus fréquente que celle de face, et que le hasard cependant ait amené l'excès 10.

C'est précisément, dira-t-on peut-être, parce que tant d'hypothèses sont possibles qu'il y a lieu de chercher la probabilité de chacune.

La recherche ne peut aboutir : les données sont insuffisantes. La solution varie, en effet, avec la probabilité *a priori* de telle ou telle imperfection de la pièce, et cette probabilité n'est pas connue.

Si l'expérience est faite dans un pays où la fabrication des monnaies a une grande perfection, les grands écarts, *a priori,* sont presque impossibles, et, parmi les petits, ceux qui favorisent face ont même probabilité que ceux qui favorisent pile.

Si les pièces, par leur relief exagéré, favorisent toutes le même résultat, le problème est autre que si, par un autre accident de la fabrication, elles favorisaient le résultat contraire. Il faut remplacer par une hypothèse les renseignements qui font défaut.

L'hypothèse adoptée est inouïe.

Toutes les probabilités, depuis 0 jusqu'à 1, données par la pièce à l'arrivée de face sont supposées, *a priori*, également vraisemblables.

On a cherché quelquefois, non la probabilité de chaque hypothèse, mais la probabilité pour que la chance donnée à l'arrivée de face soit plus grande que $\frac{1}{2}$. Surpasse-t-elle $\frac{1}{2}$ de un cent-millionième seulement, il faudra donner à cette différence imperceptible le nom de *cause* et laisser croire, d'après la dénomination adoptée, que l'écart observé est dû à cette imperfection de la pièce.

Il n'est pas inutile de traiter, pour ne laisser aucun doute, un cas célèbre pris pour exemple par Poisson.

124. Buffon a jeté une pièce de monnaie 4040 fois et obtenu 2048 fois face.

Poisson a cherché la probabilité pour que la pièce de Buffon donnât à l'arrivée de face une probabilité plus grande que celle de pile.

Avant de résoudre la question, il semble naturel de chercher si cet écart de 28, qui substitue 2048 fois face au chiffre probable 2020, est assez invraisemblable par lui-même pour rendre suspecte la pièce qui l'a donné.

La probabilité d'un écart h pour un événement dont la probabilité est p est

$$\frac{1}{\sqrt{2\mu\pi pq}}\, e^{-\frac{h^2}{2\mu pq}},$$

μ désignant le nombre des épreuves et q la probabilité $1-p$ de l'événement contraire.

La probabilité d'un écart moindre que h, en valeur absolue, est

$$\frac{2}{\sqrt{2\mu\pi pq}}\int_0^h e^{-\frac{z^2}{2\mu pq}}\,dz = \frac{2}{\sqrt{\pi}}\int_0^{\frac{h}{\sqrt{2\mu pq}}} e^{-t^2}\,dt = \Theta\left(\frac{h}{\sqrt{2\mu pq}}\right).$$

Il faut, dans cette formule, faire $h = 28$, $\mu = 4040$. On peut, sans erreur sensible, remplacer pq par $\frac{1}{4}$; on aura

$$\frac{h}{\sqrt{2\mu pq}} = \frac{28}{\sqrt{2020}} = 0,6236.$$

La Table des valeurs de la fonction Θ donne

$$\Theta(0,62) = 0,619.$$

La probabilité de l'événement contraire, c'est-à-dire la probabilité pour que, la pièce étant parfaite, l'écart soit égal ou supérieur à 28, est donc 0,38.

Si l'on recommençait 1000 fois l'expérience de Buffon, *avec des pièces parfaites,* on obtiendrait 380 fois environ un écart supérieur à 28. Si donc le hasard est la cause du résultat obtenu, il n'y a pas sujet d'étonnement.

Résolvons cependant le problème.

Soient $\frac{1}{2}+z$ la probabilité donnée par la pièce de Buffon à l'arrivée de face; $\frac{1}{2}-z$ celle qu'elle donnait, par conséquent, à l'arrivée de pile. *Toutes les valeurs de* z, *entre* $-\frac{1}{2}$ *et* $\frac{1}{2}$, *sont supposées également probables* a priori.

La probabilité de l'événement observé était, avant l'épreuve, proportionnelle au produit

$$\left(\frac{1}{2}+z\right)^{2048}\left(\frac{1}{2}-z\right)^{1992}.$$

En prenant le logarithme de ce produit, remplaçant

$$l\left(\frac{1}{2}+z\right) \quad \text{par} \quad l\frac{1}{2}+l(1+2z) = l\frac{1}{2}+2z-2z^2,$$

$$l\left(\frac{1}{2}-z\right) \quad \text{par} \quad l\frac{1}{2}+l(1-2z) = l\frac{1}{2}-2z-2z^2,$$

on voit qu'en supprimant un facteur constant dont la présence ne change rien, la probabilité de l'événement était, pour une petite valeur de z, proportionnelle à

$$e^{-8080z^2+112z}.$$

La probabilité pour que z soit positif est donc proportionnelle à

$$\int_0^\infty e^{-8080z^2+112z}\,dz$$

et, pour qu'il soit négatif, à

$$\int_0^\infty e^{-8080z^2-112z}\,dz;$$

on a

$$\int_0^\infty e^{-\alpha^2z^2+2\beta z}\,dz = e^{\frac{\beta^2}{\alpha^2}}\int_0^\infty e^{-\left(\alpha z-\frac{\beta}{\alpha}\right)^2}\,dz$$

et, en posant $\alpha z - \frac{\beta}{\alpha} = t$,

$$\frac{1}{\alpha}e^{\frac{\beta^2}{\alpha^2}}\int_{-\frac{\beta}{\alpha}}^\infty e^{-t^2}\,dt = \frac{\sqrt{\pi}}{2\alpha}e^{\frac{\beta^2}{\alpha^2}} + e^{\frac{\beta^2}{\alpha^2}}\frac{\sqrt{\pi}}{2\alpha}\Theta\left(\frac{\beta}{\alpha}\right);$$

on a également

$$\int_0^\infty e^{-\alpha^2z^2-2\beta z}\,dz = \frac{\sqrt{\pi}}{2\alpha}e^{\frac{\beta^2}{\alpha^2}} - \frac{\sqrt{\pi}}{2\alpha}e^{\frac{\beta^2}{\alpha^2}}\Theta\left(\frac{\beta}{\alpha}\right).$$

Le rapport de la probabilité pour que z soit positif à celle pour qu'il soit négatif est donc

$$\frac{1+\Theta\left(\frac{56}{\sqrt{8080}}\right)}{1-\Theta\left(\frac{56}{\sqrt{8080}}\right)} = 4,263.$$

Si l'on désigne par p' et q' ces deux probabilités, on déduit de

$$\frac{p'}{q'} = 4,263,$$

$$p' = \frac{4,263}{5,263} = 0,81;$$

Poisson a trouvé 0,81043.

125. Une difficulté pourrait s'élever, le principe du calcul étant admis, sur la formule d'approximation employée. Après avoir désigné la probabilité cherchée par $\frac{1}{2}+z$, et annoncé que toutes les valeurs de z seraient traitées comme également vraisemblables *a priori*, nous avons négligé les puissances de z supérieures à la seconde.

Cela est permis. Lorsque z en effet n'est pas petit, la probabilité de l'événement observé peut être considérée comme nulle, aussi bien que l'exponentielle qui la remplace. C'est pour la même raison que, z étant compris entre 0 et $\frac{1}{2}$, nous pouvons étendre les intégrations de 0 à ∞.

126. Si Buffon, au lieu de jeter la pièce 4040 fois, l'avait jetée 1 fois seulement et qu'il eût obtenu face, l'événement, on en conviendra sans peine, n'apprendrait rien sur la qualité de la pièce.

Cherchons cependant, en appliquant les mêmes principes, la probabilité pour que la pièce ait une tendance à favoriser face.

Soit x la probabilité que la pièce donne à l'arrivée de face. La probabilité de l'événement observé étant x et les probabilités de toutes les hypothèses étant supposées égales *a priori*, la probabilité de chaque valeur de x est proportionnelle à x; la probabilité pour que x soit compris entre $\frac{1}{2}$ et 1 est proportionnelle à

$$\int_{\frac{1}{2}}^{1} x\,dx = \frac{1}{2} - \frac{1}{8} = \frac{3}{8}$$

et, pour qu'il soit compris entre 0 et $\frac{1}{2}$, à

$$\int_{0}^{\frac{1}{2}} x\,dx = \frac{1}{8};$$

le rapport est 3 et les probabilités dont la somme est l'unité sont $\frac{3}{4}$ et $\frac{1}{4}$.

Une telle conséquence suffirait pour condamner le principe.

127. La régularité du rapport des naissances masculines et féminines a beaucoup occupé les géomètres. On a commis, en étudiant des anomalies toujours petites, des erreurs semblables à celles que je viens de signaler. On a assimilé les naissances à des tirages au sort faits dans une urne de composition constante, dans laquelle le rapport du nombre des boules blanches à celui des boules noires différerait peu de celui des nombres de naissances indiqué par la Statistique.

Une telle substitution n'est légitime que si les écarts observés sont compris dans les limites et suivent les lois que la théorie montre certaines dans une suite d'épreuves réglées par le hasard.

S'il arrivait que, dans la France entière, le rapport, d'une année à l'autre, présentât de trop grandes variations ; ou si, au contraire, il se maintenait dans de trop étroites limites, il faudrait conclure, avec grande probabilité, qu'une cause intervient pour régler le hasard ou pour le troubler.

La régularité de la proportion à Londres entre les années 1629 et 1710 a été admirée comme un miracle par un savant, novice encore à la théorie du hasard. Nicolas Bernoulli, digne héritier de son oncle Jacques et éditeur de son beau Livre, montra au contraire dans les chiffres signalés la confirmation des principes. Admettant l'assimilation des naissances à un tirage au sort, sur 14000 naissances annuelles, tel était le chiffre moyen pour la ville de Londres, il est très vraisemblable que l'excès du chiffre des garçons sur la valeur moyenne ne surpassera pas une fois en cent ans 163 ; c'est l'écart le plus grand que l'on ait observé à Londres.

La probabilité d'un écart inférieur à une limite λ entre le nombre des naissances masculines, sur 14000 enfants nés annuellement, et le nombre supposé le plus probable, 7200, a pour expression très approchée

$$\frac{2}{\sqrt{\pi}}\int_0^{\frac{\lambda}{\sqrt{2\mu pq}}} e^{-t^2}\,dt = \Theta\left(\frac{\lambda}{\sqrt{2\mu pq}}\right).$$

On suppose, bien entendu, que, par une règle de trois, on ramène toujours le nombre des naissances à 14000.

Si l'on fait $\mu = 14000$, $\lambda = 163$, le produit pq pouvant être remplacé par $\frac{1}{4}$, on trouve pour probabilité

$$\Theta(1,949) = 0,994\,;$$

il y a donc plus de cent à parier contre un, chaque année, pour que le hasard n'amène pas l'anomalie dont Arbuthnot admirait la petitesse et qui s'est produite une fois seulement en cent ans.

128. On peut se demander jusqu'où devrait aller la régularité pour qu'il y eût lieu de s'en étonner. Cherchons, pour préciser la question, quel est l'écart qu'il y a dix mille à parier contre un de franchir une fois au moins en cent ans.

Nous avons vu (15) que, si la probabilité d'un événement est $\frac{1}{n}$, il y a dix mille à parier contre un que, sur $9,2\,n$ épreuves, l'événement se produira au moins une fois. Si l'on a

$$9,2\,n = 100, \qquad \text{c'est-à-dire} \qquad \frac{1}{n} = 0,092,$$

il y aura dix mille à parier contre un pour que l'événement dont la probabilité est $\frac{1}{n}$ se produise une fois au moins sur cent épreuves. La probabilité pour que l'écart, pendant une année, sur 14000 naissances soit plus grand que λ est

$$1 - \frac{2}{\sqrt{\pi}} \int_0^{\frac{\lambda}{\sqrt{2\mu pq}}} e^{-t^2} dt = 1 - \Theta\left(\frac{\lambda}{\sqrt{2\mu pq}}\right).$$

Déterminons λ de telle sorte que cette probabilité soit 0,092 et, par conséquent,

$$\Theta\left(\frac{\lambda}{\sqrt{2\mu pq}}\right) = 0,908.$$

La Table donne

$$\frac{\lambda}{\sqrt{2\mu pq}} = 1,19.$$

On peut remplacer $2\mu pq$ par 7000; on en déduit

$$\lambda = 99.$$

Si, dans un siècle, l'écart n'avait pas une seule fois dépassé 99; si, sur 14000 naissances annuelles, le nombre des garçons s'était maintenu entre 7100 et 7100, une cause régulatrice serait presque certaine; il y a dix mille à parier contre un, *a priori*, pour que le hasard, sur cent épreuves tentées dans la même urne, ne maintienne pas une telle régularité.

129. Buffon a signalé une commune de Bourgogne dans laquelle, sur 2000 baptêmes enregistrés pendant cinq ans, le nombre des filles a surpassé de 20 celui des garçons. Il est né dans cette commune, sur 2000 enfants, 39 garçons de moins que le chiffre normal. La formule donne 0,92 pour la probabilité d'un écart, en plus ou en moins, inférieur à 39; 0,08 est, par conséquent, celle d'une anomalie au moins égale à celle qu'a signalée Buffon; les statisticiens la rencontreraient souvent, s'ils la cherchaient.

130. Laplace a trouvé, pendant le XVIII^e^ siècle, la proportion des garçons aux filles plus petite à Paris que dans l'ensemble du pays, $\frac{25}{24}$ au lieu de $\frac{22}{21}$, chiffre normal adopté alors, auquel les documents nouveaux et plus nombreux ont substitué $\frac{18}{17}$.

Quelle est, se demande Laplace, la probabilité pour que cette différence *soit due à une cause?*

« A Paris, dit-il, les baptêmes des enfants des deux sexes s'écartent un peu du rapport de 22 à 21. Depuis 1745, époque à laquelle on a commencé à distinguer les sexes sur les registres des naissances, jusqu'à la fin de 1784, on a baptisé dans cette capitale 393386 garçons et 377555 filles. Le rapport de ces deux nombres est à peu près celui de 25 à 24; il paraît donc qu'à Paris une cause particulière rapproche de l'égalité les baptêmes des deux sexes.

» Si l'on applique à cet objet le Calcul des probabilités, on trouve qu'il y a 238 à parier contre 1 en faveur de l'existence de cette cause. »

Laplace supprime les détails. Ni ses calculs ne sont rapportés, ni les principes sur lesquels ils reposent.

131. Faut-il croire à un écart fortuit ou affirmer l'existence d'une cause? Les données ne sont pas suffisantes. Comment se prononcer de la même manière si les études antérieures ont appris que le rapport varie très rarement, ou si l'on constate, partout où les documents sont nombreux, des écarts comparables à ceux dont s'étonne Laplace? La solidité plus ou moins grande de la règle qui se trouve en défaut doit faire apprécier différemment les conséquences.

Il n'est pas inutile d'insister.

Si l'on assimile la distribution des naissances entre les deux sexes à des tirages faits dans une urne, la probabilité pour que le hasard produise, sans l'intervention d'aucune cause perturbatrice, sur 770941 naissances, un écart relatif égal ou supérieur à $\frac{22}{43} - \frac{25}{49} = 0,00142$, est

$$1 - \frac{2}{\sqrt{\pi}} \int_0^{0,00142\sqrt{770\,941}\sqrt{2}} e^{-t^2} dt = 1 - \Theta(1,762) = 0,013.$$

Si donc on dresse des listes pendant un temps suffisant, le hasard seul, on peut l'affirmer, produira 13 fois sur 1000 environ, dans un sens ou dans l'autre, un écart égal ou supérieur à celui dont s'est préoccupé Laplace.

Si des causes autres que le hasard amènent aussi des anomalies, le statisticien, en classant par groupes de 770000 naissances les registres de tous les temps et de tous les pays, trouvera un certain nombre de rapports anomaux égaux ou supérieurs à celui de Paris. Parmi ceux-là, quelques-uns seront dus au hasard, 13 sur 1000 environ, cela peut être tenu pour certain, si les chiffres sont suffisamment grands. D'autres écarts seront dus à des causes; nous en saurions à peu près le nombre, si la statistique était faite en retranchant du nombre total le nombre probable de ceux que le hasard a produits. Nous n'avons qu'un seul fait : est-il dû au hasard? Il serait téméraire, impossible même, d'en rien dire sans accepter sur les probabilités *a priori* quelque convention arbitraire.

132. Le possesseur d'un chronomètre a remarqué un retard de 1^s, quand la température de la chambre dont le chronomètre ne sort pas s'élève de 10°.

L'observation a été renouvelée vingt fois.

Quelles sont les probabilités pour que la chaleur soit la cause du ralentissement et pour que le concours des deux faits soit fortuit?

Le possesseur d'un chronomètre a remarqué une avance de 1^s le lendemain de chaque jour où les artilleurs se sont exercés au champ de tir voisin.

L'observation a été renouvelée vingt fois.

Quelles sont les probabilités pour que l'ébranlement causé par le tir ait changé la marche du chronomètre et pour que le concours soit fortuit?

Le possesseur d'un chronomètre a remarqué un ralentissement de 1^s, chaque fois que la planète Mars passe au méridien entre minuit et 1^h du matin.

L'observation a été renouvelée vingt fois.

Quelles sont les probabilités pour que la planete influe sur le chronomètre et pour que le concours soit fortuit?

Les problèmes sont identiques. Les réponses ne peuvent cependant être les mêmes : n'est-ce pas une raison pour reconnaître les données insuffisantes?

133. Les habitants de Saint-Malo s'étaient persuadé, il y a un siècle, que, dans leur ville, le nombre des décès à l'heure de la marée haute était plus grand qu'à marée basse.

Admettons le fait.

Supposons que, sur les côtes de la Manche, on ait remarqué une plus grande proportion de naufrages par le vent du nord-ouest que par aucun autre.

Les chiffres recueillis à l'appui des deux remarques étant supposés en même nombre et inspirant même confiance, on sera loin d'en déduire les mêmes conséquences.

Lorsqu'on sera conduit à accepter comme une certitude l'influence du vent de nord-ouest sur les naufrages, les gens prudents exigeront des preuves nouvelles pour reconnaître seulement vraisemblable l'influence de la marée sur la dernière heure des Malouins.

Les problèmes, cette fois encore, sont identiques; l'impossibilité d'accepter une même réponse montre la nécessité de faire intervenir la probabilité *a priori* de la cause qu'on veut apprécier.

134. Le fermier d'une maison de jeu a installé une roulette nouvelle. L'instrument, sur 10000 coups, a amené la rouge 5300 fois et 4700 fois la noire. L'acheteur refuse le payement et demande une indemnité : les joueurs ont remarqué les sorties plus fréquentes de la rouge et en ont profité. Un procès s'engage. On allègue le Calcul des probabilités. Jamais, dit le fermier, machine bien construite n'a donné un tel écart. 300 coups sur 10000 ne peuvent être l'effet du hasard. La probabilité de la rouge n'est pas $\frac{1}{2}$ comme elle devrait. Peu importe, répond le mécanicien, la statistique des parties jouées : on ne peut pas garantir les caprices du hasard; la machine, construite par d'excellents ouvriers, a été vérifiée avec soin. Aucune pièce n'est imparfaite; on ne montre ni roue mal centrée, ni cases inégales, ni nivellement défectueux. Le tribunal nomme un expert. Quelle décision doit-il conseiller?

L'écart observé est un indice. Quelle en est l'importance? L'application du principe (**115**) suppose des données qu'on n'a pas.

Les probabilités des diverses hypothèses, que nous nommons les *causes,* sont proportionnelles au produit de leur probabilité *a priori* par la probabilité qu'elles donnent à l'événement.

L'événement est l'arrivée de 5300 fois rouge sur 10000 épreuves. La cause inconnue, c'est la valeur de la probabilité, donnée par la machine, à l'arrivée de la couleur rouge.

La probabilité *a priori* désignée par ϖ_i dans la formule (**115**), est complètement inconnue; en supposant, comme on l'a fait dans des cas analogues, toutes les valeurs également probables, on proposerait une hypothèse inacceptable. Si, réellement, la roulette favorise la rouge, un très petit écart est plus probable qu'un grand, un très grand est impossible. Les grands écarts sont, en outre, d'autant moins probables que le mécanicien a meilleure réputation et qu'il a employé de meilleurs ouvriers.

L'expert doit répondre :

Les faits connus de la cause ne permettent pas l'évaluation des probabilités : il manque l'appréciation *a priori* de la probabilité qu'on veut connaître *a posteriori.*

Le calcul serait sans issue.

Il faut simplifier la question.

La machine, suivant l'une des parties, donne à la sortie de la rouge une probabilité peu différente de 0,53.

Le résultat de 10000 épreuves ne permet pas, suivant lui, d'en douter.

La machine, suivant l'adversaire, donne, comme elle doit, à la sortie de la rouge une probabilité voisine de 0,500. Le soin apporté à la construction ne permet de croire à aucun défaut grave.

Précisons les deux dires :

La probabilité de la rouge, dans l'un des systèmes. serait comprise entre 0,499 et 0,501 ; suivant l'autre, entre 0.529 et 0,531.

Laissons de côté le cas très possible où les plaideurs se tromperaient tous deux, et cherchons le rapport des probabilités de leurs assertions.

Soient x la probabilité pour que la roulette donne à la couleur rouge une chance comprise entre 0,529 et 0,531 ; y la probabilité pour que la chance soit comprise entre 0,499 et 0,501, $x+y$ n'est pas égal à l'unité. Le rapport $\frac{x}{y}$, sans faire connaître les deux probabilités, apportera un renseignement utile.

Soient ϖ_1 la probabilité *a priori* pour que la machine, d'après ce qu'on savait avant sa mise en œuvre, en tenant compte de sa bonne apparence, de la bonne renommée du constructeur et de l'apparente sincérité de ses déclarations, donne à la sortie de la couleur rouge une valeur comprise entre 0,499 et 0,501 ; ϖ_2 la probabilité, évaluée toujours avant l'épreuve, pour que, malgré les garanties énumérées, elle donne une probabilité comprise entre 0,529 et 0,531.

ϖ_1 et ϖ_2 sont inconnus.

Soient p_1 et p_2 les probabilités que les deux hypothèses sur le mérite de la roulette donneraient à l'événement observé, on aura (115)

$$\frac{x}{y}=\frac{\varpi_1 p_1}{\varpi_2 p_2}.$$

En nommant X la probabilité de la sortie de la rouge, la probabilité qu'elle donne à l'événement observé est proportionnelle à

$$X^{5300}(1-X)^{4700}.$$

Si l'on suppose

$$X_1=0,500 \qquad \text{et} \qquad X_2=0,530,$$

on trouve

$$\frac{X_1{}^{5300}(1-X_1)^{4700}}{X_2{}^{5300}(1-X_2)^{4700}}=0,0000000 1508.$$

Le rapport varie peu quand X_1 et X_2 restent dans les limites assignées. On peut donc supposer

$$\frac{p_1}{p_2}=0,000000015;$$

par conséquent,

$$\frac{x}{y} = \frac{15}{1000000000}\frac{\varpi_1}{\varpi_2} = \frac{1}{66666666}\frac{\varpi_1}{\varpi_2}.$$

$\frac{\varpi_1}{\varpi_2}$ est inconnu. Si le constructeur est très habile, ce rapport est très grand; il est invraisemblable qu'une pièce détestable sorte d'ateliers dignes de confiance; mais, quelle que soit la confiance, en divisant le rapport qui la mesure par 66666666, il est à craindre que le quotient soit petit.

Si vous croyez qu'un défaut tel que celui qu'on soupçonne ne peut se produire qu'une fois sur un million, il restera 66 à parier contre 1 qu'il s'est produit cette fois-là.

135. Les restrictions proposées peuvent s'appliquer aux conséquences déduites par Mitchell du rapprochement fréquent de deux étoiles dans le ciel. Deux hypothèses sont possibles :

Pour être aperçues dans la même direction, deux étoiles n'ont rien de commun, leur vraie distance est immense;

Les deux étoiles, au contraire, sont voisines dans l'espace; c'est pour cela qu'elles sont rapprochées dans le ciel.

En comptant les étoiles de 1^re^, de 2^e^ et de 3^e^ grandeur et les supposant indépendantes les unes des autres, Mitchell a calculé le nombre probable des groupes dont la distance est inférieure à une limite donnée.

Il ne faudrait pas croire que toutes les distances angulaires soient, *a priori*, également probables : les petites le sont moins que les grandes. Considérons une première étoile, peu importe la position qu'elle occupe. Si une seconde étoile est placée *au hasard*, pour qu'elle se trouve à une distance angulaire de la première comprise entre θ et $\theta + d\theta$, il faut que le hasard la place dans une zone comprise entre deux cercles ayant pour pôles la première étoile et pour rayons sphériques θ et $\theta + d\theta$. La surface de cette zone est d'autant plus grande que θ s'approche davantage de $\frac{\pi}{2}$; elle est proportionnelle au sinus de l'angle θ dont on cherche la probabilité. La probabilité pour qu'une étoile se trouve à une distance inférieure à γ d'une étoile donnée est le rapport de la zone à une base terminée par le petit cercle dont γ est le rayon sphérique à la surface de l'hémisphère sur laquelle on suppose l'étoile placée au hasard. Ce rapport, égal à $1 - \cos \gamma$, peut être représenté, si γ est petit, par $\frac{\gamma^2}{2}$.

Si, par exemple, on prend

$$\gamma = 10' = \frac{\pi}{1080} = 0{,}0029089,$$

on aura

$$\frac{\gamma^2}{2} = 0{,}0000042307 = \frac{1}{236362}.$$

Le nombre des étoiles des trois premières grandeurs étant égal à 230, on peut former $\frac{230 \cdot 229}{2} = 26\,335$ combinaisons deux à deux. La probabilité pour que le hasard procure deux étoiles à distance inférieure à $10'$ est celle d'obtenir 1 boule blanche en 26335 tirages dans une urne contenant une seule blanche et 236362 noires. Cette probabilité est

$$1 - \left(1 - \frac{1}{236\,362}\right)^{26335} = 0{,}103403.$$

L'ingénieux argument de Mitchell ne peut pas cependant fournir d'évaluation numérique. La recherche plus ou moins rigoureusement faite de la probabilité pour que le hasard ait produit les groupes observés n'est pas le seul élément de la question. C'est le seul, cependant, que l'on fasse intervenir. Si l'on trouvait, en étudiant le ciel, trois étoiles de 1^{re} grandeur ayant, à 1^s près, la même ascension droite, la probabilité pour que le hasard produise un tel rapprochement est plus petite assurément que la formation fortuite du groupe des Pléiades.

En conclura-t-on, avec la même vraisemblance, que ce rapprochement doit avoir une cause et que ces étoiles, dont la déclinaison diffère de $40°$ ou $50°$ peut-être, sont solidaires et forment un système?

Personne, assurément, n'y songera, et la raison en est que la probabilité *a priori*, désignée par ϖ_i dans nos formules, est trop petite.

Comment définir, d'ailleurs, la singularité dont on juge le hasard incapable?

Les Pléiades semblent plus rapprochées les unes des autres qu'il n'est *naturel*. L'assertion est digne d'intérêt ; mais, si l'on veut traduire la conséquence en chiffres, les éléments font défaut. Faut-il, pour préciser cette idée vague de rapprochement, chercher le plus petit cercle qui contienne le groupe? la plus grande des distances angulaires? la somme des carrés de toutes les distances? l'aire du polygone sphérique dont quelques-unes des étoiles sont les sommets et qui contient les autres dans son intérieur? Toutes ces grandeurs, dans le groupe des Pléiades, sont plus petites qu'il n'est vraisemblable. Laquelle d'entre elles donnera la mesure de l'invraisemblance? Si trois étoiles forment un triangle équilatéral, faut-il faire entrer cette circonstance, assurément peu probable *a priori*, au nombre de celles qui révèlent une cause?

L'intervention du hasard dans la formation de l'univers est inacceptable. Une loi règle tout, cela n'est pas douteux. Quelle est cette loi? voilà la question. Il n'est pas admissible qu'on demande s'il y en a une et que l'on évalue la probabilité de la réponse.

Si l'étude du ciel suggérait plusieurs lois, on pourrait les mettre en balance et non choisir l'une d'elles, celle du groupement par attraction mutuelle, pour l'opposer à l'ensemble des autres, réunies sous le nom vague de *hasard*.

Lorsque plusieurs étoiles semblent voisines sur la voûte céleste, leur proximité dans l'espace est la première explication qui se présente. Ne peut-on en imaginer d'autres?

Si l'ensemble des étoiles forme un réseau régulier, si le Soleil en est un sommet, les lois géométriques de cette immense cristallisation peuvent exiger des alignements. La Terre, voisine du Soleil, aperçoit, pour ce motif, beaucoup d'étoiles dans une même direction. L'hypothèse, dira-t-on, ne mérite pas examen. Raison de plus, si on la condamne par l'appréciation des probabilités *a priori*, pour ne négliger dans aucun cas le rôle indispensable qu'elles doivent jouer.

136. On a attaché beaucoup d'importance à la recherche de la probabilité des événements futurs, déduite des événements observés comme corollaire de la probabilité des causes.

Le problème peut s'énoncer ainsi :

Plusieurs causes peuvent produire un même événement dont la probabilité dépend à la fois de la probabilité pour que chaque cause agisse et de la probabilité que, dans ce cas, elle donne à l'événement.

Si les probabilités des causes sont ϖ_1, ϖ_2, ..., ϖ_n, et celles qu'elles donnent à l'événement $p_1, p_2, \ldots, p_2$, la probabilité, *a priori*, pour que l'événement se produise, est

$$p_1\varpi_1 + p_2\varpi_2 + \ldots + p_n\varpi_n.$$

On fait un certain nombre d'épreuves, dans des conditions telles que la même cause, on ignore laquelle, a agi dans toutes. Si les causes sont des urnes de compositions différentes, que le hasard peut désigner pour qu'on y fasse le tirage, une même urne a servi à tous les tirages. La connaissance des résultats obtenus change les probabilités, qui ne sont plus, pour les différentes causes, ϖ_1, ϖ_2, ..., ϖ_n. On demande la probabilité pour qu'une nouvelle épreuve, faite dans les mêmes conditions que les précédentes, procure l'arrivée d'un événement désigné.

Une urne, par exemple, contient des boules noires et des boules blanches en proportion inconnue. *Toutes les suppositions sont également possibles.*

On a fait μ tirages; il est sorti m boules blanches et n noires. Quelle est la probabilité pour que le $(\mu + 1)^{\text{ième}}$ tirage amène une boule blanche?

Soit x la probabilité donnée par la composition de l'urne à la sortie d'une boule blanche. *Toutes les valeurs de x sont,* a priori, *également probables.*

La probabilité qu'une valeur x donne à l'événement observé est

$$x^m(1-x)^n.$$

La probabilité pour que x soit compris entre x et $x + dx$ peut donc être représentée par

$$Gx^m(1-x)^n\,dx,$$

G étant indépendant de x. La probabilité pour que x soit compris entre o et 1 est la certitude. On doit donc avoir

$$1 = G\int_0^1 x^m(1-x)^n\,dx = G\frac{\Gamma(m+1)\Gamma(n+1)}{\Gamma(m+n+2)};$$

donc

$$G = \frac{\Gamma(m+n+2)}{\Gamma(m+1)\Gamma(n+1)},$$

et la probabilité pour que le rapport du nombre des boules blanches au nombre total des boules soit compris entre x et $x+dx$ est

$$\frac{\Gamma(m+n+2)}{\Gamma(m+1)\Gamma(n+1)}x^m(1-x)^n\,dx,$$

c'est-à-dire

$$\frac{1.2.3\ldots m+n+1}{1.2.3\ldots m.1.2.3\ldots n}x^m(1-x)^n\,dx.$$

La probabilité pour qu'une nouvelle épreuve amène une boule blanche est la somme des produits des diverses valeurs de la probabilité x par la probabilité pour que chacune soit la véritable, c'est-à-dire

$$\frac{\Gamma(m+n+2)}{\Gamma(m+1)\Gamma(n+1)}\int_0^1 x^{m+1}(1-x)^n\,dx = \frac{\Gamma(m+n+2)\Gamma(m+2)\Gamma(n+1)}{\Gamma(m+1)\Gamma(n+1)\Gamma(m+n+3)};$$

on a

$$\frac{\Gamma(m+n+2)}{\Gamma(m+n+3)} = \frac{1}{m+n+2},$$

$$\frac{\Gamma(m+2)}{\Gamma(m+1)} = +1,$$

et la probabilité demandée est

$$\frac{m+1}{m+n+2}.$$

137. Les applications faites de cette formule ont été presque toutes sans fondement.

On a osé chercher la probabilité pour que le Soleil se lève demain.

Pour chercher la probabilité d'un événement, il faut accepter son contraire comme possible. Une urne est donc supposée qui contient des boules blanches et des boules noires. La probabilité d'en tirer 1 boule blanche représente celle de voir le Soleil se lever. Jamais il n'a manqué : cela dure depuis six mille ans. L'urne, consultée 2 191 500 fois, a donné 2 191 500 boules blanches.

La formule donne pour la probabilité d'un nouveau tirage semblable aux précédents

$$\frac{2\,191\,501}{2\,191\,502} = 0{,}999999543.$$

Est-il besoin d'insister sur l'insignifiance d'un tel calcul?

Représentons-nous le premier homme au premier coucher du Soleil. Il devrait, pendant sa première nuit, s'il raisonne comme Condorcet, assigner la valeur $\frac{2}{3}$ à la probabilité de le revoir. S'il comprend la question et s'il se la pose, les chances pour lui seront beaucoup moindres. Le Soleil a disparu ; s'il est éteint, qui le rallumera? S'il est tombé dans la mer, comment en sortirait-il?

Sans avoir cependant la science parfaite que lui supposent les théologiens, le premier homme se persuadera sans doute, après cent apparitions, que le Soleil tourne autour de la Terre; rien alors ne doit lui faire craindre un arrêt brusque, l'absence du Soleil au cent et unième jour ne l'étonnerait pas moins que celle des objets éclairés chaque matin par ses premiers rayons ; faudra-t-il, en invoquant la formule, supposer qu'après cent jours la probabilité de les revoir est $\frac{101}{102}$? L'absurdité serait précisément la même.

Le lever du Soleil, après une année, sera pour le premier homme une certitude.

Si le temps doit la confirmer et l'accroître, c'est par la découverte des lois astronomiques et non par le succès renouvelé d'un même jeu de hasard.

CHAPITRE VIII.

LOI DES ERREURS D'OBSERVATION.

In der Astronomie ist die Praxis eine Aufgabe der Wahrscheinlichkeits-Rechnung, die Theorie eine Aufgabe der hoheren Mechanik

BESSEL.

138. Postulatum de Gauss. — 139. Autre hypothèse faite implicitement. — 140. Cas dans lequel le postulatum est rigoureusement exact. — 141. Conséquence des suppositions acceptées. — 142. Comparaison du résultat avec une formule connue; accord apparent, mais non réel. — 143. Autres contradictions résultant de la loi admise. — 144. Conséquence d'un autre mode de combinaison des mesures. — 145. L'observation confirme la règle de Gauss, les erreurs constantes étant écartées. — 146. Méthode de vérification. — 147. Résultats de Bradley discutés par Bessel. — 148. Détermination du paramètre k; diverses formules. — 149. Vérifications possibles. — 150. Valeur probable du carré de l'erreur commise en adoptant la première formule. — 151. Valeur probable pour la seconde. — 152. Comparaison des deux résultats. — 153. Les formules peuvent être remplacées par d'autres qui sont préférables. — 154. Autre modification accroissant la confiance méritée par la formule. — 155. Autres méthodes pour calculer k. — 156. Groupement des observations deux par deux; valeur probable de la plus grande erreur. — 157. Autre démonstration du résultat. — 158. Valeur probable du carré de la plus grande des erreurs considérées deux à deux. — 159. Groupement des erreurs trois par trois. — Distinction nécessaire entre l'erreur véritable et l'erreur présumée. — 160. Cas plus général. — 161. Discussion de la démonstration précédente. — 162. Expression qui caractérise la précision d'un système de mesures — 163. Ce qu'on entend par *poids* et par *précision*. — 164 Si la loi de probabilité était autre, ces deux mots n'auraient plus de sens précis. — 165. Est-il permis d'écarter les mesures rendues suspectes par leur différence avec la moyenne? — 166. Valeur probable de la plus petite des erreurs commises — 167. Combinaison d'observations qui n'inspirent pas égale confiance. — 168. Partage d'une grandeur en plusieurs parties mesurées séparément. — 169. Évaluation donnée par Fourier en Égypte. — 170. Discussion relative à la détermination de la constante k déduite d'un système d'observations.

138. La loi rigoureuse de probabilité des erreurs d'observation varie sans doute avec la grandeur mesurée, comme avec le choix de l'instrument et l'habileté de l'observateur; elle est inaccessible aux géomètres.

Euler, Bernoulli, Lagrange et Laplace ont fait des hypothèses démenties par les faits et mal justifiées par des preuves sans vraisemblance.

Gauss, plus heureux, a déduit d'un raisonnement fort simple une loi que la démonstration laisserait douteuse, mais que les conséquences justifient.

Lorsque plusieurs mesures d'une grandeur inspirent une confiance égale, la valeur la plus probable est la moyenne de celles qu'on a obtenues.

Tel est le postulatum de Gauss. Tous les observateurs, avant qu'il fût pris pour base de la théorie, y avaient adhéré en en faisant usage.

S'il suffisait d'admettre une règle aussi plausible, la théorie serait parfaite.

A la condition énoncée, il faut, malheureusement, en ajouter plusieurs autres qu'on ne dit pas. Nous devons signaler d'abord une différence essentielle entre la valeur la plus probable d'une grandeur et la meilleure valeur à adopter.

La valeur la plus probable est celle dont la probabilité est la plus grande. Peu importent les autres. Elles doivent toutes, cependant, diriger le choix à faire. S'il est utile d'accroître la probabilité des petites erreurs, il est désirable aussi de diminuer celle des grandes. S'attacher seulement à choisir la valeur la plus probable, c'est imiter le joueur qui, pouvant espérer un grand nombre de gains différents et craindre un grand nombre de pertes, prendrait ses décisions de manière à accroître la chance de gagner le plus gros lot, sans aucunement se soucier des autres.

En disant : « En présence de plusieurs mesures d'une même grandeur, le parti le meilleur est d'adopter la moyenne », et : « La moyenne entre plusieurs mesures est la valeur la plus probable », on énonce deux propositions différentes. On a eu tort de les confondre.

Supposons, pour donner un exemple, que l'on cherche l'origine probable d'une plante d'espèce connue cueillie en France.

Dans la liste des 100 localités où l'espèce se rencontre, ou en trouve trois dans la même commune du Cantal et les autres dans 97 communes différentes du Finistère.

La commune la plus probable est en Auvergne ; l'origine probable de la plante, la Bretagne.

139. Après avoir proposé le postulatum qui a été admis sans difficulté, Gauss représente par $\varphi(\Delta)d\Delta$ la probabilité pour que l'erreur d'une mesure soit comprise entre Δ et $\Delta + d\Delta$: la fonction $\varphi(\Delta)$ est l'inconnue qu'il veut déterminer.

Ici encore s'élève une objection. La probabilité d'une erreur Δ est-elle une fonction de Δ?

Ne dépend-elle pas de la grandeur mesurée? Si l'on fait une pesée, si l'on mesure un angle, lorsque le poids est un nombre exact de milligrammes, lorsque l'angle contient un nombre exact de secondes, n'a-t-on pas plus de chances d'évaluer juste que s'il faut ajouter une fraction? Si cette fraction, que l'instrument ne donne pas, est exactement égale à $\frac{1}{2}$, n'a-t-on pas, en l'évaluant, moins de chances d'erreur que si elle est 0,27 ?

En disant, sans explication : Soit $\varphi(\Delta)d\Delta$ la probabilité d'une erreur Δ, on s'écarte, dès les premiers mots, de la rigueur promise en quelque sorte par la forme géométrique de la démonstration.

140. Il est un cas où, le postulatum étant rigoureusement démontrable, la conclusion cependant n'est qu'approchée. C'est une preuve décisive contre l'exactitude de la théorie.

Supposons que la grandeur à mesurer soit le rapport du nombre des boules blanches au nombre total des boules, dans une urne de composition inconnue.

On tire un certain nombre de boules; nous le désignerons par μ.

Sur ces μ boules, m sont blanches.

La fraction $\frac{m}{\mu}$ peut être considérée comme une mesure du rapport cherché. L'instrument qui l'a donnée est d'autant plus précis que le nombre des tirages est plus grand. L'opération recommencée n fois donne les mesures successives

$$\frac{m_1}{\mu}, \quad \frac{m_2}{\mu}, \quad \ldots, \quad \frac{m_n}{\mu}.$$

La valeur la plus probable du rapport cherché déduite des $n\mu$ tirages est

$$\frac{m_1 + m_2 + \ldots + m_n}{n\mu};$$

elle est, on le voit, la moyenne entre les valeurs successivement obtenues par n mesures qui méritent même confiance. Si la démonstration que nous allons donner était irréprochable, la conclusion devrait, dans ce cas, s'appliquer en toute rigueur.

141. Soient $x_1, x_2, \ldots, x_n$ les valeurs d'une même grandeur données par n mesures successives, faites dans les mêmes conditions et dignes de la même confiance.

La valeur la plus probable de la grandeur inconnue est, on le suppose, la moyenne

$$X = \frac{x_1 + x_2 + \ldots + x_n}{n}.$$

Si z est la valeur véritable, les erreurs successivement commises ont été $z - x_1$, $z - x_2$, ..., $z - x_n$. La probabilité, avant l'épreuve, de ce concours de mesures, c'est-à-dire celle de l'événement observé, est proportionnelle au produit

$$\varphi(z - x_1)\varphi(z - x_2)\ldots\varphi(z - x_n. \tag{1}$$

La probabilité pour que la valeur exacte soit z est proportionnelle à (1); la valeur la plus probable est celle qui rend ce produit maximum : elle doit donc rendre nulle sa dérivée et, par conséquent, aussi celle de son logarithme, et l'on doit avoir

$$\frac{\varphi'(z - x_1)}{\varphi(z - x_1)} + \frac{\varphi'(z - x_2)}{\varphi(z - x_2)} + \ldots + \frac{\varphi'(z - x_n)}{\varphi(z - x_n)} = 0. \tag{2}$$

Cette équation, d'après le postulatum, doit être satisfaite par la valeur

$$z = \frac{x_1 + x_2 + \ldots + x_n}{n}.$$

Si l'on pose

$$\frac{\varphi'(u)}{\varphi(u)} = F(u), \tag{3}$$

l'équation (2) peut s'énoncer de la manière suivante :

La somme des valeurs de la fonction $F(u)$ doit être nulle lorsque celle des valeurs de la variable est elle-même égale à zéro.

Si l'on considère deux valeurs u_1 et u_2 de la variable, on doit avoir

$$F(u_1) + F(-u_1) = 0;$$

la fonction est impaire.

Si l'on en considère trois, u_1, u_2 et $-u_1 - u_2$, ont doit avoir

$$F(u_1) + F(u_2) = F(u_1 + u_2).$$

On déduit de cette équation bien connue la forme de la fonction F; elle est proportionnelle à la variable.

L'équation (3) donne

$$\frac{\varphi'(u)}{\varphi(u)} = Cu.$$

On en déduit, en intégrant,

$$l\varphi(u) = \frac{C u^2}{2} + c'$$

et, en remarquant que $\varphi(u)$ doit s'annuler quand u est infini,

$$\varphi(u) = G e^{-k^2 n^2},$$

G et k étant des constantes.

On peut déterminer le facteur G.

La probabilité d'une erreur comprise entre z et $z + dz$ étant

$$G e^{-k^2 z^2} dz,$$

celle d'une erreur comprise entre $-\infty$ et $+\infty$ est

$$\int_{-\infty}^{+\infty} e^{-k^2 z^2} dz;$$

elle représente la certitude. Il faut l'égaler à l'unité. On en déduit, en se rappelant l'équation bien connue

$$\int_{-\infty}^{+\infty} e^{-h^2 z^2} dz = \frac{1}{k}\sqrt{\pi},$$

$$G = \frac{k}{\sqrt{\pi}},$$

et la loi de probabilité des erreurs est représentée par

$$\frac{k}{\sqrt{\pi}} e^{-k^2 z^2} dz, \tag{4}$$

une seule constante servant à distinguer et à caractériser les cas.

142. La vérification dont nous avons parlé (**140**) semble confirmer le résultat.

Si, sur μ tirages faits dans une urne, on a rencontré m boules blanches, la probabilité pour que la composition de l'urne donne

$$\frac{m}{\mu} + z$$

pour rapport du nombre des boules blanches au nombre total des boules est (**121**)

$$\frac{\mu}{\sqrt{2\pi m(\mu - m)}} e^{-\frac{z^2 \mu^2}{2m(\mu-m)}}. \tag{5}$$

Le théorème semble confirmé; il est mis en défaut : la formule (5) est seulement approchée, elle devrait être rigoureusement exacte.

143. La règle des moyennes, il importe d'insister sur ce point, n'est ni démontrée ni exacte. S'il était admis que la moyenne entre plusieurs mesures fût toujours la valeur la plus probable, il en résulterait des contradictions. Quand on mesure une grandeur, on mesure, par cela même, toutes les fonctions de cette grandeur, son carré par exemple, ou le logarithme du nombre qui la représente. Pourquoi la valeur la plus probable du carré ne serait-elle pas la moyenne des valeurs obtenues pour le carré, et la valeur probable du logarithme, la moyenne des logarithmes?

La valeur la plus probable d'une grandeur dont les mesures successives sont $x_1, x_2, \ldots, x_n$ serait alors représentée par l'une ou l'autre des formules

$$\left\{ \begin{array}{l} \sqrt{\dfrac{x_1^2 + x_2^2 + \ldots + x_n^2}{n}}, \\ \sqrt[n]{x_1, x_2, \ldots, x_n}; \end{array} \right. \tag{6}$$

le carré de la première est, en effet, la moyenne des carrés, et le logarithme de la seconde, la moyenne des logarithmes.

Il ne faut pas, pour écarter l'objection, faire une distinction entre les grandeurs directement mesurées et celles qui résultent d'un calcul. Un mécanicien pourrait, bien aisément, annexer à une balance une aiguille marquant le carré ou le logarithme du poids. Ce carré ou ce logarithme deviendrait alors la grandeur mesurée. Le postulatum admis dans un cas devient donc impossible dans les autres.

La seconde supposition indispensable à la démonstration, la possibilité d'exprimer la probabilité d'une erreur en fonction de cette erreur seulement, implique également contradiction. Une grandeur étant égale à X, on lui trouve, en la mesurant, la valeur x_1; l'erreur est $X - x_1$. L'erreur sur le carré est $X^2 - x_1^2$. Si l'on pose

$$X - x_1 = z,$$

on aura

$$X^2 - x_1^2 = 2x_1 z + z^2;$$

la probabilité d'une erreur z étant $\varphi(z)$, celle de l'erreur $2x_1 z + z^2$ sur le carré sera aussi $\varphi(z)$. Elle n'est pas une fonction de l'erreur commise.

144. L'importance du résultat justifie l'étude et la discussion minutieuses des détails. Nous résoudrons encore le problème suivant :

Si Gauss avait adopté, au lieu de la moyenne, un autre mode de combinaison des mesures, quelle loi des erreurs en aurait-il déduite?

En supposant que

$$\varphi(x_1, x_2, \ldots, x_n) \tag{7}$$

représente la valeur la plus probable d'une grandeur dont les valeurs successivement obtenues sont $x_1, x_2, \ldots, x_n$, quelle est la loi correspondante pour la probabilité des erreurs?

En général, nous allons le démontrer, il n'existe pas de loi, exprimant la probabilité d'une erreur Δ en fonction de Δ, qui puisse justifier l'adoption de la fonction φ. Le problème ne peut être résolu que pour certaines formes très particulières.

En désignant par z la valeur de la grandeur mesurée et par $\psi(\Delta)$ la fonction de Δ à laquelle est proportionnelle la probabilité d'une erreur Δ, la valeur la plus probable z doit rendre maximum le produit

$$\psi(z-x_1)\,\psi(z-x_2)\ldots\psi(z-x_n).$$

La dérivée par rapport à z de ce produit doit être nulle et, si l'on pose

$$\frac{\psi'(u)}{\psi(u)} = \mathrm{F}(u),$$

on doit avoir

$$(8) \qquad \mathrm{F}(z-x_1)+\mathrm{F}(z-x_2)+\ldots+\mathrm{F}(z-x_n)=0,$$

z étant lié à x_1, x_2, ..., x_n par la relation

$$z=\varphi(x_1, x_2, \ldots, x_n).$$

Pour que l'équation (8) puisse être satisfaite par une fonction F, il faut que les valeurs $z-x_1$, $z-x_2$, ..., $z-x_n$ de la variable, auxquelles correspondent les valeurs de la fonction dont la somme doit être nulle, ne soient pas indépendantes les unes des autres.

Les n fonctions de n variables

$$\varphi-x_1, \quad \varphi-x_2, \quad \ldots, \quad \varphi-x_n$$

doivent être liées par une relation. Il faut et il suffit, pour cela, que leur déterminant fonctionnel soit nul.

On doit avoir

$$\begin{vmatrix} \dfrac{d\varphi}{dx_1}-1 & \dfrac{d\varphi}{dx_1} & \dfrac{d\varphi}{dx_1} & \ldots & \dfrac{d\varphi}{dx_1} \\ \dfrac{d\varphi}{dx_2} & \dfrac{d\varphi}{dx_2}-1 & \ldots & \ldots & \dfrac{d\varphi}{dx_2} \\ \ldots & \ldots\ldots & \ldots & \ldots & \ldots \\ \dfrac{d\varphi}{dx_n} & \dfrac{d\varphi}{dx_n} & \ldots & \ldots & \dfrac{d\varphi}{dx_n}-1 \end{vmatrix} = 0,$$

ce qui donne l'équation linéaire

$$\frac{d\varphi}{dx_1}+\frac{d\varphi}{dx_2}+\ldots+\frac{d\varphi}{dx_n}=1,$$

dont l'intégrale est

$$(9)\qquad \varphi = \frac{x_1 + x_2 + \ldots + x_n}{n} + \mathrm{F}(x_2 - x_1,\ x_3 - x_1,\ \ldots,\ x_n - x_1).$$

Telle est la seule forme possible de la fonction φ, pour laquelle la recherche, telle qu'on l'a faite, puisse donner une loi de probabilité.

La forme obtenue pour φ peut se définir simplement. La fonction φ est telle qu'en donnant à toutes les mesures prises un même accroissement α, elle augmente elle-même de α. Cette condition est nécessaire et suffisante pour que la fonction ait la forme (9).

Si l'on prenait

$$(10)\qquad \left\{\begin{array}{l} \varphi(x_1, x_2, \ldots, x_n) = \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2}, \\ \varphi(x_1, x_2, \ldots, x_n) = \sqrt[n]{x_1, x_2, \ldots, x_n}, \end{array}\right.$$

la condition ne serait pas remplie.

Aucune loi de probabilité des erreurs, *en fonction de l'erreur commise,* ne peut donc justifier l'adoption des formules (10).

145. Malgré les objections précédentes, la formule de Gauss doit être adoptée. L'observation la confirme : cela doit suffire dans les applications. Les conséquences minutieusement étudiées se sont toujours trouvées d'accord avec les faits. Il est bien entendu que les erreurs constantes sont en dehors des formules; chaque observateur doit étudier son instrument et sa méthode pour les écarter. La moyenne, évidemment, ne peut donner que la mesure entachée de l'erreur systématique inhérente à l'instrument. Si l'on pèse avec de faux poids, si les fils de la lunette sont mal placés, les moyennes ne pourront pas faire les corrections; c'est à la grandeur mesurée accrue de l'erreur constante que se rapportent alors les formules.

146. Disons d'abord comment la vérification de la formule a été faite.

Bessel, le premier, a fait la comparaison des erreurs commises dans de nombreuses observations avec les conséquences de la formule de Gauss.

Le théorème de Bernoulli permet d'affirmer que, dans une série suffisamment nombreuse, si l'on partage les erreurs en groupes bien définis, chaque groupe se présentera un nombre de fois à peu près proportionnel à sa probabilité.

Bessel a étudié quatre cents observations de Bradley portant sur les coordonnées, parfaitement connues aujourd'hui, d'une même étoile. Les erreurs ont été classées par ordre de grandeur; on a compté le nombre de celles qui sont inférieures à $0'',4$ en valeur absolue, le nombre de celles qui sont comprises entre $0'',4$ et $0'',8$, etc.

On a comparé ces différents nombres avec le Tableau des nombres probables de chaque groupe d'erreurs, d'après la loi de probabilité admise

$$\frac{k}{\sqrt{\pi}} e^{-k^2 z^2}.$$

Le calcul exige la connaissance de la constante k. Les méthodes pour la calculer sont nombreuses; leur accord numérique est une des meilleures vérifications de la théorie.

Nous reviendrons sur cette importante question. Bornons-nous à faire remarquer en ce moment que chacun des calculs qui vont donner, en fonction de k supposé connu, le nombre probable des erreurs d'un certain groupe, suffit, par la condition d'égaler le résultat du calcul à celui de l'expérience, pour fournir une valeur de k. Si l'on n'avait qu'un groupe, la détermination de la constante par la condition de satisfaire à l'égalité qu'on veut vérifier formerait un cercle vicieux. Mais les égalités sont nombreuses. Chacune d'elles déterminera une valeur de k; on adoptera la moyenne et les vérifications ultérieures auront toute leur valeur.

Nous aurons d'ailleurs à revenir sur ce problème très important de la détermination de k.

La probabilité d'une erreur comprise entre z et $z+dz$ étant

$$\frac{k}{\sqrt{\pi}} e^{-k^2 z^2}\, dz,$$

celle d'une erreur plus petite que α, en valeur absolue, sera

$$\frac{2k}{\sqrt{\pi}} \int_0^{\alpha} e^{-k^2 z^2}\, dz = \frac{2}{\sqrt{\pi}} \int_0^{k\alpha} e^{-t^2}\, dt = \Theta(k\alpha).$$

Cette formule donnera la probabilité pour qu'une erreur soit comprise entre o et α, entre o et 2α, o et 3α, etc. La différence de deux termes consécutifs sera la probabilité pour qu'elle soit comprise entre deux multiples.

147. Le Tableau suivant a été formé par M. Nikolaus Wuich, professeur à l'École d'Artillerie de Vienne.

Les erreurs sont évaluées dans la première colonne par leur rapport à l'*erreur probable*, c'est-à-dire à celle qu'il y a probabilité $\frac{1}{2}$ de surpasser ou de ne pas surpasser. La moitié des erreurs sur un grand nombre de mesures doivent être par définition, pour ainsi dire, plus grandes, l'autre moitié plus petites que l'erreur probable. On voit en effet dans le Tableau, en regard de 1, le nombre 5000 indiquant que la moitié des erreurs sont inférieures à 1. On lit en regard de 2, 8227; sur 10000 erreurs, par conséquent, il y en a 8227 plus petites que le double de l'erreur probable; 9570, d'après le Tableau, sont plus petites que le triple, 9930 que le quadruple et 9993 que le quintuple.

Sur 10000 observations, nombre probable des erreurs plus petites que mh, h étant l'erreur probable :

m.	Nombre sur 10000.	m.	Nombre sur 10000.	m.	Nombre sur 10000.	m.	Nombre sur 10000.
0,00...	0000	0,15...	806	0,30...	1604	0,45...	2385
0,01...	54	0,16...	859	0,31...	1656	0,46...	2436
0,02...	108	0,17...	913	0,32...	1709	0,47...	2488
0,03...	161	0,18...	966	0,33...	1761	0,48...	2539
0,04...	215	0,19...	1020	0,34...	1814	0,49...	2590
0,05...	269	0,20...	1073	0,35...	1866	0,50...	2641
0,06...	323	0,21...	1126	0,36...	1919	0,51...	2691
0,07...	377	0,22...	1180	0,37...	1971	0,52...	2742
0,08...	430	0,23...	1233	0,38...	2023	0,53...	2793
0,09...	484	0,24...	1286	0,39...	2075	0,54...	2843
0,10...	538	0,25...	1339	0,40...	2127	0,55...	2893
0,11...	591	0,26...	1392	0,41...	2179	0,56...	2944
0,12...	645	0,27...	1445	0,42...	2230	0,57...	2994
0,13...	699	0,28...	1498	0,43...	2282	0,58...	3044
0,14...	752	0,29...	1551	0,44...	2334	0,59...	3093

$m.$	Nombre sur 10000.	$m.$	Nombre sur 10000.	$m.$	Nombre sur 10000.	$m.$	Nombre sur 10000.
0,60...	3143	0,96...	4827	1,62...	7255	2,80...	9411
0,61...	3192	0,97...	4871	1,64...	7313	2,85...	9454
0,62...	3242	0,98...	4914	1,66...	7371	2,90...	9495
0,63...	3291	0,99...	4957	1,68...	7428	2,95...	9534
0,64...	3340	1,00...	5000	1,70...	7485	3,00...	9570
0,65...	3389	1,02...	5085	1,72...	7540	3,05...	9603
0,66...	3438	1,04...	5170	1,74...	7594	3,10...	9635
0,67...	3487	1,06...	5254	1,76...	7648	3,15...	9664
0,68...	3535	1,08...	5337	1,78...	7701	3,20...	9691
0,69...	3584	1,10...	5419	1,80...	7753	3,25...	9716
0,70...	3632	1,12...	5500	1,82...	7804	3,30...	9740
0,71...	3680	1,14...	5581	1,84...	7854	3,35...	9762
0,72...	3728	1,16...	5660	1,86...	7904	3,40...	9782
0,73...	3775	1,18...	5739	1,88...	7952	3,45...	9800
0,74...	3823	1,20...	5817	1,90...	8000	3,50...	9818
0,75...	3870	1,22...	5894	1,92...	8047	3,55...	9834
0,76...	3918	1,24...	5971	1,94...	8093	3,60...	9848
0,77...	3965	1,26...	6046	1,96...	8138	3,65...	9862
0,78...	4012	1,28...	6121	1,98...	8183	3,70...	9870
0,79...	4059	1,30...	6194	2,00..	8227	3,75...	9886
0,80...	4105	1,32...	6267	2,05...	8332	3,80...	9896
0,81...	4152	1,34...	6339	2,10...	8433	3,85...	9906
0,82...	4198	1,36...	6410	2,15...	8530	3,90...	9915
0,83...	4244	1,38...	6480	2,20...	8622	3,95...	9923
0,84...	4290	1,40...	6550	2,25...	8709	4,00...	9930
0,85...	4336	1,42...	6618	2,30...	8792	4,10...	9943
0,86...	4381	1,44...	6686	2,35...	8870	4,20...	9954
0,87...	4427	1,46...	6753	2,40...	8945	4,30...	9963
0,88...	4472	1,48...	6818	2,45...	9016	4,40...	9970
0,89...	4517	1,50...	6883	2,50...	9082	4,50...	9976
0,90...	4562	1,52...	6947	2,55...	9146	4,60...	9981
0,91...	4606	1,54...	7011	2,60...	9205	4,70...	9985
0,92...	4651	1,56...	7073	2,65...	9261	4,80...	9988
0,93...	4695	1,58...	7134	2,70...	9314	4,90...	9991
0,94...	4739	1,60...	7195	2,75...	9364	5,00...	9993
0,95...	4783						

La valeur de k est inversement proportionnelle à l'erreur probable. Si l'on a posé, en effet,

$$\frac{1k}{\sqrt{\pi}}\int_0^{\lambda} e^{-k^2z^2}\,dz = \frac{2}{\sqrt{\pi}}\int_0^{k\lambda} e^{-t^2}\,dt = \frac{1}{2},$$

on trouvera

$$k\lambda = 0,47693$$

et, par conséquent, λ désignant l'erreur probable, la valeur correspondante de k est

$$k = \frac{0,47693}{\lambda}.$$

148. Bessel a trouvé par la discussion des 470 observations de Bradley, en réduisant les nombres par une proportion à l'hypothèse de 100 observations seulement, les nombres suivants pour les erreurs variant de $0'',4$:

Déclinaison d'une étoile.

$$k = 0'',43539.$$

Erreurs.		Nombre des erreurs.	Nombre calculé.
$0'',0$ à	$0'',4$	22,0	19,5
0,4	0,8	19,3	18,3
0,8	1,2	18,3	16,2
1,2	1,6	9,3	13,6
1,6	2,0	9,0	10,6
2,0	2,4	7,7	7,9
2,4	2,8	3,3	5,5
2,8	3,2	5,0	3,6
3,2	3,6	2,7	2,2
3,6	4,0	1,3	1,3
4,0	...	2,0	1,4

L'étude des observations d'ascension droite de la même étoile a donné :

$$k = 0'',2283.$$

Erreurs.		Nombre des erreurs.	Nombre calculé.
$0'',0$	à $0'',1$	38,0	33,5
0,1	0,2	28,0	28,0
0,2	0,3	17,7	19,2
0,3	0,4	8,0	10,9
0,4	0,5	4,7	5,1
0,5	0,6	2,0	2,0
0,6	0,7	1,0	0,7
0,7	0,8	0,3	0,0

Cent déterminations de l'ascension droite de l'étoile polaire faites à l'observatoire de Königsberg, de 1813 à 1815, ont donné :

$$k = 1'',3093.$$

Erreurs.		Nombre des erreurs.	Nombre calculé.
$0'',0$	à $0'',4$	25	24,9
0,4	0,8	22	21,9
0,8	1,2	19	18,2
1,2	1,6	11	13,7
1,6	2,0	9	9,5
2,0	2,4	8	6,0
2,4	2,8	2	3,4
2,8	3,2	3	1,8
3,2	3,6	1	0,9
3.6	...	0	0,6

149. Nous devons dire maintenant comment on détermine le paramètre k.

Deux cas peuvent se présenter : si les mesures ont porté sur une grandeur exactement mesurée avant ou après les observations qu'on discute, toutes les erreurs sont connues avec certitude. Si, au contraire, les mesures ont été prises sur une grandeur qui n'est connue que par elles, les erreurs sont incertaines comme la grandeur elle-même; et, en donnant ce nom d'*erreur* à la différence entre chaque valeur observée et la moyenne, on ne peut avoir qu'une évaluation vraisemblable et approchée.

Nous traiterons d'abord le premier cas, auquel se ramène le second.

La détermination approchée de k peut se faire d'un grand nombre de manières. Il suffit, en effet, de calculer la valeur probable d'une fonction quelconque de l'erreur commise dans une observation. La moyenne des valeurs de cette fonction, dans une série suffisamment nombreuse, en vertu du théorème de Bernoulli, différera peu de la valeur probable. En les égalant, on obtiendra la valeur de k.

On peut ainsi déduire la valeur approchée de k de la valeur probable de l'erreur prise en valeur absolue, du carré de l'erreur, du cube et, en général, d'une puissance quelconque de l'erreur.

La probabilité d'une erreur z étant

$$\frac{k}{\sqrt{\pi}} e^{-k^2z^2},$$

la valeur probable de z, considéré comme essentiellement positif, est

$$\frac{2k}{\sqrt{\pi}} \int_0^{\infty} z e^{-k^2z^2} dz = \frac{1}{k\sqrt{\pi}}.$$

La valeur probable de z^2 est

$$\frac{2k}{\sqrt{\pi}} \int_0 z^2 e^{-k^2z^2} dz = \frac{1}{2k^2};$$

celles de z^3 et de z^4,

$$\frac{2k}{\sqrt{\pi}} \int_0^{\infty} e^{-k^2z^2} z^3 dz = \frac{1}{k^3\sqrt{\pi}},$$

$$\frac{2k}{\sqrt{\pi}} \int_0^{\infty} e^{-k^2z^2} z^4 dz = \frac{3}{4k^4}.$$

En admettant que, sur un grand nombre d'épreuves, la moyenne des valeurs d'une grandeur diffère peu de la valeur probable, en nommant $e_1, e_2, \ldots, e_n$ les erreurs successivement commises dans n observations, on pourra écrire

$$\frac{e_1 + e_2 + \ldots + e_n}{n} = \frac{S_1}{n} = \frac{1}{k\sqrt{\pi}},$$

$$\frac{e_1^2 + e_2^2 + \ldots + e_n^2}{n} = \frac{S_2}{n} = \frac{1}{2k^2},$$

$$\frac{e_1^3 + e_2^3 + \ldots + e_n^3}{n} = \frac{S_3}{n} = \frac{1}{k^3\sqrt{\pi}},$$

$$\frac{e_1^4 + e_2^4 + \ldots + e_n^4}{n} = \frac{S_4}{n} = \frac{3}{4k^4}.$$

150. Chacune de ces équations donne une valeur de k; elles doivent s'accorder et s'accordent, en effet, dans les applications, d'une manière très satisfaisante. On en déduit ces relations remarquables. S_1, S_2, S_3, S_4 désignant la somme des erreurs, prises en valeur absolue, la somme des carrés, la somme des cubes, la somme des quatrièmes puissances, on doit avoir approximativement

$$\frac{\frac{S_2}{n}}{\left(\frac{S_1}{n}\right)^2} = \frac{\pi}{2},$$

$$\frac{\frac{S_3}{n}}{\left(\frac{S_1}{n}\right)^3} = \pi,$$

$$\frac{\frac{S_4}{n}}{\left(\frac{S_1}{n}\right)^4} = \frac{3\pi^2}{4}.$$

Ces formules singulières, dont le premier membre est fourni par le hasard, méritent tant de confiance qu'un calculateur à qui des observations sont remises et qui trouve ces égalités en défaut peut tenir pour certain qu'on a retouché et altéré les résultats immédiats de l'expérience.

151. Il importe d'évaluer la confiance méritée par les valeurs de la constante k que nous venons d'obtenir. Cette appréciation

consistera dans le calcul de la valeur probable de l'erreur commise en adoptant une des équations précédentes.

Lorsque nous écrivons, par exemple,

$$\frac{e_1+e_2+\ldots+e_n}{n}-\frac{1}{k\sqrt{\pi}}=0, \tag{11}$$

l'équation n'est pas rigoureuse. Nous savons que, pour de grandes valeurs de n, la moyenne des erreurs *diffère peu* de $\frac{1}{k\sqrt{\pi}}$, qui est la valeur probable. Quel que soit n, une grande différence est *possible*, cela est évident; mais elle est peu probable.

Nous aurons une appréciation de la confiance méritée par l'équation (11) en cherchant la valeur probable de

$$\left(\frac{e_1+e_2+\ldots+e_n}{n}-\frac{1}{k\sqrt{\pi}}\right)^2. \tag{12}$$

On comprend la nécessité d'élever l'expression au carré. Elle peut être positive ou négative : sa valeur probable est nulle, cela ne prouve nullement la précision des mesures; les grandes valeurs peuvent être détruites dans la somme par des termes de signes contraires.

On a

$$\left(\frac{e_1+e_2+\ldots+e_n}{n}-\frac{1}{k\sqrt{\pi}}\right)^2=\frac{1}{n^2}\Sigma e_i^2+\frac{2}{n^2}\Sigma e_i e_i'-\frac{2}{nk\sqrt{\pi}}\Sigma e_i+\frac{1}{k^2\pi}.$$

La valeur probable de chacun des termes du second membre est connue.

Le carré d'une erreur e_i^2 a pour valeur probable $\frac{1}{2k^2}$; l'erreur e_i, qui est prise ici en valeur absolue, a pour valeur probable $\frac{1}{k\sqrt{\pi}}$, et le produit $e_i e_{i'}$ de deux erreurs a, par conséquent (48), pour valeur probable $\frac{1}{\pi k^2}$. La valeur probable de l'expression (12) se réduit, en ayant égard au nombre des termes de chaque somme, à

$$\frac{1}{2nk^2}+\frac{n.n-1}{n^2}\left(\frac{1}{k^2\pi}\right)-\frac{2}{k^2\pi}+\frac{1}{k^2\pi}=\frac{1}{2nk^2}\left(1-\frac{2}{\pi}\right);$$

elle tend vers zéro lorsque n augmente.

152. Cherchons la valeur probable de

$$(13)\qquad \left(\frac{e_1^2+e_2^2+\ldots+e_n^2}{n}-\frac{1}{2k^2}\right)^2,$$

cette valeur étant calculée, bien entendu, de même que la précédente, avant les opérations faites et sans qu'on puisse prévoir ni présumer quelles sont les plus grandes erreurs.

L'expression (13) peut s'écrire

$$\frac{1}{n^2}\Sigma e_i^4+\frac{2}{n^2}\Sigma e_i^2 e_i'^2-\frac{1}{k^2 n}\Sigma e_i^2+\frac{1}{4k^4}.$$

La valeur probable s'obtiendra en remplaçant e_i^4, quel que soit i, par sa valeur probable $\frac{3}{4k^4}$, $e_i^2 e_i'^2$ par $\frac{1}{4k^4}$ et e_i^2 par $\frac{1}{2k^2}$. On a ainsi pour valeur probable de (12)

$$\frac{3}{4nk^4}+\frac{n.n-1}{n^2}\frac{1}{4k^4}-\frac{1}{2k^4}+\frac{1}{4k^4}=\frac{1}{2nk^4}.$$

152. On peut se demander quelle est, entre ces deux évaluations d'erreur à craindre, celle qui justifie le mieux la formule correspondante.

Il ne faut pas se borner à comparer les résultats $\frac{1}{2nk^2}\left(1-\frac{2}{\pi}\right)$ et $\frac{1}{2nk^4}$ pour donner l'avantage au plus petit. En admettant que ces expressions donnent une *appréciation* de l'erreur à craindre, ces erreurs ne portent pas sur la même inconnue. La première est commise dans l'évaluation de $\frac{1}{k\sqrt{\pi}}$; l'autre, dans celle de $\frac{1}{2k^2}$. Ce qu'il faut assurer évidemment, c'est la petitesse de l'erreur commise sur k. Soit y cette erreur, l'erreur sur $\frac{1}{k}$ sera $-\frac{1}{k^2}y$, et sur $\frac{1}{k^2}$ elle sera, en négligeant bien entendu y^2, $-\frac{2}{k^3}y$; le carré de l'erreur commise sur $\frac{1}{k\sqrt{\pi}}$ est donc $\frac{1}{\pi k^4}y^2$, et le carré de l'erreur commise sur $\frac{1}{2k^2}$ est $\frac{1}{k^6}y^2$.

Les résultats obtenus doivent donc s'énoncer de la manière suivante :

La valeur probable de $\frac{1}{\pi k^4}y^2$ est $\frac{1}{2nk^2}\left(1-\frac{2}{\pi}\right)$;

La valeur probable de $\frac{1}{k^6}y^2$ est $\frac{1}{2nk^4}$.

Par conséquent, en adoptant la première formule, celle qui déduit k de la moyenne des erreurs, la valeur probable de y^2 est

$$\frac{\pi k^2}{2n}\left(1-\frac{2}{\pi}\right)=\frac{k^2}{2n}\cdot 1{,}3141,$$

et, en adoptant la seconde, c'est-à-dire en déduisant k de la moyenne des carrés des erreurs, la valeur probable de y^2 est

$$\frac{k^2}{2n}.$$

La seconde formule doit être préférée.

153. Les formules précédentes peuvent être remplacées par d'autres qui donnent une plus grande chance d'exactitude.

Nous résoudrons d'abord le problème suivant :

Déterminer la valeur de λ qui rend minima la valeur probable de

$$\left(\frac{1}{k}-\lambda\frac{e_1+e_2+\ldots+e_n}{n}\right)^2. \tag{14}$$

Cette expression peut s'écrire

$$\frac{1}{k^2}-\frac{2\lambda}{k}\frac{\Sigma e_i}{n}+\lambda^2\frac{\Sigma e_i^2}{n^2}+\frac{2\lambda^2}{n^2}\Sigma e_i e_i'. \tag{15}$$

En remplaçant les erreurs e_i et leurs carrés par les valeurs probables, on a, pour valeur probable de (15),

$$\frac{1}{k^2}-\frac{2\lambda}{k^2\sqrt{\pi}}+\frac{\lambda^2}{2k^2n}+\frac{\lambda^2(n-1)}{n}\frac{1}{k^2\pi}$$

ou

$$\frac{\lambda^2}{2k^2}\left(\frac{1}{n}+\frac{1}{\pi}-\frac{2}{n\pi}\right)-\frac{2\lambda}{k^2\sqrt{\pi}}+\frac{1}{k^2}.$$

Le minimum correspond à

$$\lambda=\frac{2}{\sqrt{\pi}\left(\frac{1}{n}+\frac{2}{\pi}-\frac{2}{n\pi}\right)}=\frac{2\sqrt{\pi}}{2-\left(\frac{\pi-2}{n}\right)}.$$

Il convient donc, pour diminuer les chances d'erreur, de remplacer la formule

$$\frac{1}{k} = \sqrt{\pi}\left(\frac{e_1 + e_2 + \ldots + e_n}{n}\right)$$

par

(16) $$\frac{1}{k} = \frac{\sqrt{\pi}}{1 + \frac{\pi - 2}{2n}}\left(\frac{e_1 + e_2 + \ldots + e_n}{n}\right).$$

154. Nous déterminerons le facteur λ par la condition que la valeur probable de l'expression

(17) $$\left(\frac{1}{k^2} - \lambda\frac{e_1^2 + e_2^2 + \ldots + e_n^2}{n}\right)^2$$

soit la plus petite possible.

Cette expression peut s'écrire

$$\frac{1}{k^4} - \frac{2\lambda}{nk^2}\Sigma e_i^2 + \frac{\lambda^2}{n^2}\Sigma e_i^4 + \frac{2\lambda^2}{n^2}\Sigma e_i^2 e_{i'}^2.$$

En faisant le calcul comme au n° 153, on trouve pour valeur probable de cette expression

$$\frac{1}{k^4} - \frac{\lambda}{k^4} + \frac{\lambda^2}{nk^4}\left(\frac{1}{2} + \frac{n}{4}\right).$$

Le minimum correspond à

$$\lambda = \frac{n}{2\left(\frac{1}{2} + \frac{n}{4}\right)} = \frac{2n}{n+2}.$$

On diminuera donc la valeur probable du carré de l'erreur à craindre, en substituant à la formule

$$\frac{1}{k^2} = \frac{2(e_1^2 + e_2^2 + \ldots + e_n^2)}{n}$$

la formule, probablement plus exacte,

(18) $$\frac{1}{k^2} = \frac{2(e_1^2 + e_2^2 + \ldots + e_n^2)}{n+2}.$$

155. Les déterminations de la valeur approchée de k sont en nombre infini.

On peut poser le problème d'une manière différente.

Les observations sont faites, on connaît les erreurs, quelle est la valeur la plus probable de k, la valeur probable de k et celle de k^2?

On sait que la valeur la plus probable n'est pas la valeur probable et que la valeur probable du carré n'est nullement égale au carré de la valeur probable.

Soient $e_1, e_2, \ldots, e_n$ les erreurs successivement commises, k la valeur inconnue du paramètre caractéristique de la série d'observations.

L'événement observé étant l'arrivée successive de n erreurs e_1, $e_2, \ldots, e_n$, la probabilité de cet événement était, avant l'épreuve, proportionnelle au produit

$$\frac{k^n}{\pi^{\frac{n}{2}}} e^{-k^2(e_1^2+e_2^2+\ldots+e_n^2)}$$

des probabilités simples.

La valeur de k la plus probable est celle qui rend ce produit maximum.

Le maximum correspond à celui du logarithme

$$n l k - k^2(e_1^2+e_2^2+\ldots+e_n^2),$$

c'est-à-dire à la valeur de k qui annule la dérivée

$$\frac{n}{k} - 2k(e_1^2+e_2^2+\ldots+e_n^2),$$

$$\frac{1}{2k^2} = \frac{e_1^2+e_2^2+\ldots+e_n^2}{n}.$$

C'est une des formules déjà obtenues.

Au lieu de la valeur la plus probable de k, on peut chercher la valeur probable, qui doit, en général, être préférée; la valeur probable est l'espérance mathématique de celui qui, après les mesures prises, attendrait une somme égale à k.

La probabilité d'une valeur de k est proportionnelle à

$$k^n e^{-k^2(e_1^2+e_2^2+\ldots+e_n^2)}.$$

En désignant la somme des carrés des erreurs par S_2, on peut la représenter par

$$G k^n e^{-k^2 S_2}.$$

On déterminera G par la condition

$$G \int_0^\infty k^n e^{-k^2 S_2}\, dk = 1,$$

qui exprime que la probabilité pour que k soit compris entre o et ∞ est égale à l'unité. On a

$$\int_0^\infty k^n e^{-k^2 S_2}\, dk = \frac{\Gamma\left(\frac{n+1}{2}\right)}{2 S_2^{\frac{n+1}{2}}};$$

par conséquent,

$$G = \frac{2 S_2^{\frac{n+1}{2}}}{\Gamma\left(\frac{n+1}{2}\right)}.$$

La probabilité d'une valeur k du paramètre est

$$\frac{2 S_2^{\frac{n+1}{2}}}{\Gamma\left(\frac{n+1}{2}\right)} k^n e^{-S_2 k^2}.$$

La valeur probable de k est la somme des produits obtenus en multipliant chaque valeur de k par la probabilité qui lui correspond

$$\frac{2 S_2^{\frac{n+1}{2}}}{\Gamma\left(\frac{n+1}{2}\right)} \int_0^\infty k^{n+1} e^{-S_2 k^2}\, dk = \frac{2 S_2^{\frac{n+1}{2}}}{\Gamma\left(\frac{n+1}{2}\right)} \frac{\Gamma\left(\frac{n}{2}+1\right)}{2 S_2^{\frac{n+1}{2}+\frac{1}{2}}},$$

qui se réduit à

$$k = \frac{1}{\sqrt{S_2}} \frac{\Gamma\left(\frac{n}{2}+1\right)}{\Gamma\left(\frac{n}{2}+\frac{1}{2}\right)}.$$

Si l'on suppose n très grand, en remplaçant la fonction $\Gamma(x)$ par sa valeur approchée

$$\Gamma(x) = e^{-x} x^{x-1} \sqrt{2\pi x},$$

on trouvera

$$k = \frac{1}{\sqrt{S_2}} \frac{e^{-\frac{n}{2}-1}\left(\frac{n}{2}+1\right)^{\frac{n}{2}}\sqrt{2\pi\left(\frac{n}{2}+1\right)}}{e^{-\frac{n}{2}-\frac{1}{2}}\left(\frac{n}{2}+\frac{1}{2}\right)^{\frac{n}{2}-\frac{1}{2}}\sqrt{2\pi\left(\frac{n}{2}+\frac{1}{2}\right)}}.$$

$\sqrt{\dfrac{\frac{n}{2}+1}{\frac{n}{2}+\frac{1}{2}}}$ peut, n étant très grand, être remplacé par l'unité et, en supprimant le facteur commun $e^{-\frac{n}{2}-\frac{1}{2}}$, on peut écrire

$$k = \frac{1}{\sqrt{S_2}} e^{-\frac{1}{2}} \sqrt{\frac{n}{2}+1} \left(1 + \frac{1}{n+1}\right)^{\frac{n}{2}}.$$

Le produit $e^{-\frac{1}{2}}\left(1+\frac{1}{n+1}\right)^{\frac{n}{2}}$ a pour limite l'unité, et l'on voit aisément que, pour de grandes valeurs de n, le rapport

$$\frac{S_2}{n}$$

a pour limite $\frac{1}{2k^2}$.

La formule nouvelle est donc d'accord avec celles qui donnent

$$\frac{1}{2k^2} = \frac{S_2}{n},$$
$$\frac{1}{2k^2} = \frac{S_2}{n+2},$$

lorsque n devient très grand.

156. La loi de probabilité étant admise, on peut chercher à l'avance la probabilité d'une combinaison quelconque des résultats des mesures. La valeur moyenne de la quantité calculée, si le nombre des mesures devient grand, différera peu de sa valeur probable.

Nous résoudrons quelques problèmes dont les solutions, remarquables par leur simplicité, permettent d'élégantes vérifications.

Si, après avoir fait un grand nombre d'observations, on les groupe deux par deux en chargeant le hasard de les associer, et que, dans chaque groupe, on choisisse la plus grande des deux erreurs, la valeur probable de cette erreur peut se calculer : elle est le produit par $\sqrt{2}$ de la valeur probable d'une erreur prise au hasard. La probabilité pour que, sur deux erreurs prises au hasard, l'une soit comprise entre z et $z + dz$ et l'autre plus petite que z est

$$2\,\frac{2k}{\sqrt{\pi}}e^{-k^2z^2}\,dz\int_0^z \frac{2k}{\sqrt{\pi}}e^{-k^2z^2}\,dz.$$

Le premier facteur est, en effet, la probabilité pour que l'erreur, positive ou négative, soit comprise entre z et $z + dz$; le second, la probabilité pour que la seconde erreur soit plus petite que z. On multiplie par 2, parce que l'une ou l'autre des erreurs cherchées peut être la plus grande et égale à z.

La valeur probable de z s'obtiendra en multipliant la probabilité par z et faisant la somme pour toutes les valeurs de z

$$2\int_0^\infty \frac{2k}{\sqrt{\pi}}e^{-k^2z^2}\,dz.z\int_0^z \frac{2k}{\sqrt{\pi}}e^{-k^2z^2}\,dz.$$

En intégrant par parties et remarquant que le terme intégré disparaît aux deux limites, cette expression devient

$$\frac{4}{\sqrt{\pi}}\int_0^\infty e^{-2k^2z^2}\,dz = \frac{\sqrt{2}}{k\sqrt{\pi}};$$

c'est le produit par $\sqrt{2}$ de la valeur probable d'une erreur.

La valeur probable d'une grandeur quelconque différant peu, sur un grand nombre d'épreuves, de la moyenne arithmétique des valeurs fournies par le hasard, l'application de ce théorème à une série de mesures d'une grandeur connue donnera et a donné à plusieurs reprises une vérification des formules.

157. Le théorème précédent, remarqué expérimentalement par M. le commandant Delauney, peut s'obtenir plus directement.

Soient x_1 et x_2 deux mesures prises successivement, $\frac{x_1 + x_2}{2}$ peut être considéré comme une mesure.

En nommant e_1 et e_2 les erreurs commises sur e_1 et e_2, le carré de l'erreur sur cette mesure nouvelle sera

$$\frac{e_1^2 + e_2^2 + 2e_1e_2}{4}. \tag{19}$$

La valeur probable de e_1e_2 est nulle, puisque les valeurs positives de e_1 et de e_2 ont même probabilité que les valeurs négatives. La valeur probable de (19) est donc

$$2\frac{\frac{1}{2k^2}}{4} = \frac{1}{4k^2},$$

c'est précisément la valeur probable du carré de l'erreur pour un système de mesures dans lequel le paramètre caractéristique k serait remplacé par $k\sqrt{2}$.

La valeur probable de l'erreur commise sur $\frac{x_1 + x_2}{2}$ est donc $\frac{1}{k\sqrt{2}\sqrt{\pi}}$. Mais cette erreur peut être la somme ou la différence des deux erreurs commises sur x_1 et x_2 : la somme, si ces erreurs sont de même signe; la différence, si elles sont de signes contraires. Les deux suppositions ont chacune pour probabilité $\frac{1}{2}$; on a donc, en nommant G et P la valeur probable de la plus grande et celle de la plus petite des deux,

$$\frac{1}{2}(G + P) + \frac{1}{2}(G - P) = \frac{1}{\sqrt{2}k\sqrt{\pi}},$$

c'est-à-dire

$$G = \frac{\sqrt{2}}{k\sqrt{\pi}}.$$

On a aussi

$$\frac{1}{2}(G + P) = \frac{1}{k\sqrt{\pi}};$$

par conséquent,

$$P = \frac{1}{k\sqrt{\pi}}(2 - \sqrt{2}).$$

158. On peut, en chargeant le hasard de grouper les erreurs deux à deux comme dans le calcul précédent, calculer la valeur probable du carré de la plus grande.

L'expression de cette valeur, obtenue par un raisonnement tout semblable, est

$$2\frac{2k}{\sqrt{\pi}}\int_0^\infty z^2 e^{-k^2z^2}\,dz\int_0^z \frac{2k}{\sqrt{\pi}} e^{-k^2z^2}\,dz.$$

En intégrant par parties, faisant porter l'intégration sur $ze^{-k^2z^2}\,dz$ et remarquant que le terme intégré disparaît aux limites, cette expression devient

$$\frac{4}{\pi}\int_0^\infty e^{-k^2z^2}\,dz\int_0^z e^{-k^2z^2}\,dz + \frac{4}{\pi}\int_0^\infty z e^{-2k^2z^2}\,dz,$$

$e^{-k^2z^2}\,dz$ étant la différentielle de l'intégrale par laquelle il est multiplié dans le premier terme. On a

$$\frac{4}{\pi}\int_0^\infty e^{-k^2z^2}\,dz\int_0^z e^{-k^2z^2}\,dz = \frac{1}{2k^2};$$

on a aussi

$$\int_0^\infty z e^{-2k^2z^2}\,dz = \frac{1}{4k^2},$$

et la valeur probable du carré de la plus grande erreur est

$$\frac{1}{2k^2}\left(1+\frac{2}{\pi}\right).$$

159. On peut grouper les erreurs trois par trois, en chargeant le hasard de les associer, et chercher la valeur probable du carré de la plus grande des trois.

Un raisonnement identique aux précédents donne pour cette valeur probable

$$\frac{24k^3}{\pi\sqrt{\pi}}\int_0^\infty z^2 e^{-k^2z^2}\,dz\left(\int_0^z e^{-k^2z^2}\,dz\right)^2.$$

Intégrons par parties, en faisant porter l'intégration sur $ze^{-k^2z^2}\,dz$; en posant, pour simplifier l'écriture,

$$\int_0^z e^{-k^2z^2}\,dz\,\varpi(z),$$

et n'écrivant pas le terme intégré, qui s'annule aux deux limites, nous aurons

$$(20)\qquad \frac{12k}{\pi\sqrt{\pi}}\int_0^\infty e^{-k^2z^2}\varpi(z)^2\,dz+\frac{24k}{\pi\sqrt{\pi}}\int_0^\infty e^{-k^2z^2}z\,\varpi(z)\,\varpi'(z)\,dz;$$

mais on a

$$\varpi'(z)=e^{-k^2z^2},$$

$$\int e^{-k^2z^2}\varpi(z)^2\,dz=\frac{\varpi(z)^3}{3},$$

$$\int_0^\infty z\,\varpi(z)\,\varpi'(z)\,e^{-k^2z^2}\,dz=\int_0^\infty e^{-2k^2z^2}z\,\varpi(z)=\int_0^\infty\frac{e^{-2k^2z^2}}{2k^2}\varpi'(z)\,dz$$

$$=\int_0^\infty\frac{e^{-3k^2z^2}}{2k^2}\,dz=\frac{\sqrt{\pi}}{2k^3\sqrt{3}}.$$

La substitution de ces résultats dans l'expression (20) donne, pour la valeur probable cherchée,

$$\frac{1}{2k^2}\left(1+\frac{2\sqrt{3}}{\pi}\right).$$

160. Dans toutes les formules précédentes, que les observations soient faites ou projetées seulement, ce sont les erreurs véritables qui sont introduites.

Il arrivera très souvent (dans l'immense majorité des cas, on peut le dire) que, la grandeur étant déterminée par les observations seulement, les erreurs, évaluées par la différence de chaque mesure avec la moyenne, différeront des erreurs véritables qui restent inconnues.

En remplaçant l'erreur véritable par l'erreur présumée, on obtiendra une valeur approchée de la constante k. On a proposé, dans ce cas, des règles de correction fort simples, dont la démonstration, malheureusement, peut laisser des doutes.

Le problème est celui-ci :

$x_1, x_2, \ldots, x_n$ étant les mesures successivement prises d'une grandeur X, si l'on adopte la valeur

$$X=\frac{x_1+x_2+\ldots+x_n}{n},$$

l'erreur probable est d'autant moindre que n est plus grand, pourvu que l'on ait pu, comme nous le supposons toujours dans ce Chapitre, écarter les erreurs constantes.

La formule

$$\frac{e_1^2+e_2^2+\ldots+e_n^2}{n}=\frac{1}{2k^2} \tag{21}$$

sera remplacée par

$$\frac{(X-x_1)^2+(X-x_2)^2+\ldots+(X-x_n)^2}{n}=\frac{1}{2k_1^2}. \tag{22}$$

Cette valeur de $\frac{1}{2k_1^2}$ n'est pas, je n'ose pas dire exacte (il ne faut espérer dans aucun cas qu'elle le soit), mais conforme à la formule qu'on voudrait appliquer.

La formule (22) donnera $\frac{1}{2k_1^2}$ toujours plus petit que $\frac{1}{2k^2}$ déduit de (21). La correction la plus plausible consiste, nous allons le montrer, à multiplier la formule (22) par $\frac{n}{n-1}$; ce qui, lorsque n est grand, en changera très peu la valeur.

L'expression (22) proposée pour représenter $\frac{1}{2k^2}$ peut être transformée par l'introduction de l'erreur E commise en adoptant la moyenne X pour valeur de la grandeur mesurée.

Les différences $X-x_1$, $X-x_2$, ..., $X-x_n$ sont rigoureusement égales aux différences $E-e_1$, $E-e_2$, ..., $E-e_n$, et la formule (22) peut s'écrire

$$\frac{(E-e_1)^2+(E-e_2)^2+\ldots+(E-e_n)^2}{n},$$

identique à

$$\frac{e_1^2+e_2^2+\ldots+e_n^2}{n}+\frac{nE^2-2E(e_1+e_2+\ldots+e_n)}{n}; \tag{23}$$

mais E, étant l'erreur commise sur la moyenne, est égale à la moyenne des erreurs, et la somme $e_1+e_2+\ldots+e_n$ est égale à nE; l'expression (22) se réduit donc à

$$\frac{1}{2k_1^2}=\frac{e_1^2+e_2^2+\ldots+e_n^2}{n}-E^2. \tag{24}$$

Le premier terme représente la valeur de $\frac{1}{2k^2}$ telle que la formule (21) doit la donner. Cette valeur est diminuée de E^2; par conséquent, k_1^2 est plus grand que k^2, et la substitution des erreurs présumées aux erreurs véritables, en accroissant la valeur de k^2, conduit à attribuer aux observations une précision plus grande qu'il ne faut.

On a exactement

$$E = \frac{e_1 + e_2 + \ldots + e_n}{n};$$

par conséquent,

$$(25) \qquad E^2 = \frac{e_1^2 + e_2^2 + \ldots + e_n^2}{n^2} + \frac{2}{n^2}\Sigma e_i e_{i'}.$$

Les erreurs e_i, $e_{i'}$ sont les unes positives, les autres négatives, et, les mesures n'exposant, c'est l'hypothèse, à aucune erreur constante, la valeur probable du produit $e_i e_{i'}$ est nulle. On se permet pour cette raison de supprimer le second terme de la formule (25). Il peut en résulter une erreur notable; une grandeur dont la valeur probable est petite est *certainement* petite elle-même, quand elle est essentiellement positive; mais, quand elle peut, comme ici, changer de signe, la valeur probable étant nulle, cela prouve seulement que les valeurs positives ont même probabilité que les négatives. Elles peuvent être fort grandes. Quoi qu'il en soit, les observateurs, sur le conseil et à l'exemple de Gauss, suppriment le second terme de (25) et $\frac{1}{2k_1^2}$ devient

$$\frac{1}{2k_1^2} = \frac{1}{2k^2} - \frac{1}{2nk^2} = \frac{1}{2k^2}\left(\frac{n-1}{n}\right);$$

on conseille, en conséquence, pour calculer $\frac{1}{2k^2}$, de diviser la somme des carrés des erreurs présumées par $n-1$ et non par n, dénominateur prescrit pour la somme des carrés des erreurs réelles.

161. La formule précédente prend, dans un cas plus général, une forme différente, très élégante et très souvent appliquée.

Soient $X_1, X_2, \ldots, X_p$, p grandeurs de même espèce mesurées par les mêmes procédés, avec le même instrument et les mêmes chances d'erreur. Chaque mesure a été reproduite plusieurs fois :

la première, n_1 fois; la deuxième, n_2 fois, ...; la dernière, n_p fois, les erreurs successivement commises étant $e'_1, e'_2, \ldots, e'_n$ pour la première mesure; $e''_1, e''_2, \ldots, e''_{n_2}$ pour la deuxième; $e_1^{(p)}$, $e_2^{(p)}, \ldots, e_{n_p}^{(p)}$ pour la dernière. La formule (21), évidemment applicable, donnera

$$(26)\qquad \frac{1}{2k^2} = \frac{e_1'^2 + e_2'^2 + \ldots + e_{n_1}'^2 + e_1''^2 + e_2''^2 + \ldots + e_{n_2}''^2 + \ldots + e_1^{(p)2} + \ldots + e_p^{(p)2}}{n_1 + n_2 + \ldots + n_p}.$$

Supposons que, les valeurs véritables des grandeurs mesurées restant inconnues, on remplace chacune d'elles par la moyenne des résultats obtenus en la mesurant. Si $E^{(i)}$ est l'erreur ainsi commise sur l'une d'elles, les erreurs apparentes seront

$$(E^i - e_1^i), \quad (E^i - e_2^i), \quad \ldots, \quad (E_i^i - e_{n_i}^{(i)}),$$

et la substitution de ces erreurs apparentes aux erreurs véritables remplacera, comme on l'a vu (160), la formule (26) par

$$(27)\qquad \frac{1}{2k^2} - \frac{\dfrac{e_1'^2 + e_2'^2 + \ldots + e_{n_i}'^2}{n_1} + \dfrac{e_1''^2 + e_2''^2 + \ldots + e_{n_i}''^2}{n_2} + \ldots}{n_1 + n_2 + \ldots + n_p}.$$

Mais $e'_1, e'_2, \ldots, e'_p$ étant les erreurs commises dans la mesure d'une même grandeur et le paramètre de précision étant k, on a approximativement

$$\frac{e_1'^2 + e_2'^2 + \ldots + e_n'^2}{n_1} = \frac{1}{2k^2}.$$

La même réduction peut se faire pour les fractions relatives aux autres grandeurs, et la formule (27) devient

$$\frac{1}{2k^2} - \frac{p\,\dfrac{1}{2k^2}}{N} = \frac{1}{2k^2}\left(\frac{N-p}{N}\right),$$

N désignant le nombre total $n_1 + n_2 + \ldots + n_p$ des mesures prises.

La formule conseillée dans ce cas est, d'après cela,

$$\frac{1}{2k^2} = \frac{\Sigma(X_i - x_i)^2}{N - p},$$

$\Sigma(X_i - x_i)^2$ étant la somme des carrés de toutes les erreurs *présumées* et $N - p$ l'excès du nombre total des mesures sur le nombre des grandeurs mesurées.

Si les erreurs réellement commises étaient connues, il faudrait diviser la somme de leurs carrés par N et non par $N - p$.

La démonstration suppose, malheureusement, que l'on puisse regarder comme nulle la somme des produits des erreurs considérées deux à deux.

162. Les erreurs $e_1, e_2, \ldots, e_n$ étant approximativement connues dans chaque application de la méthode, on pourra, dans chaque cas, vérifier approximativement aussi l'exactitude de l'hypothèse faite en supprimant le terme $\Sigma e_i e_{i'}$.

Cette vérification, si n est grand, entraînerait, il est vrai, des calculs hors de proportion avec l'importance du résultat.

Il est digne de remarque que, dans l'expression

$$\Sigma e_i e_{i'},$$

dont les termes, les uns positifs, les autres négatifs, ont tous, *a priori*, si on les considère isolément, même valeur probable, le nombre des termes négatifs doit, le plus souvent, l'emporter sur celui des termes positifs.

Soit 2μ le nombre des épreuves. Le cas le plus probable est celui où, parmi les erreurs, il s'en trouverait μ positives et μ négatives; le nombre des produits $e_i e_{i'}$ dont le signe est négatif se trouve alors égal à μ^2 et le nombre des produits positifs est $\mu(\mu - 1)$. La différence est μ. C'est une raison, dans ce cas, si l'on ne sait rien de plus, pour présumer que cette somme arbitrairement négligée sera négative.

Si l'on nomme $\mu - \alpha$ le nombre des erreurs positives et $\mu + \alpha$ celui des erreurs négatives, le nombre des produits $e_i e_{i'}$ obtenus en multipliant une erreur positive par une erreur négative sera

$$\mu^2 - \alpha^2;$$

celui des produits positifs est

$$\frac{1}{2}(\mu + \alpha)(\mu + \alpha - 1) + \frac{1}{2}(\mu - \alpha)(\mu - \alpha - 1) = \mu^2 + \alpha^2 - \mu.$$

Le nombre des produits négatifs sera le plus grand, si l'on a

$$\alpha^2 < \mu - \alpha^2,$$
$$\alpha < \sqrt{\frac{\mu}{2}}.$$

La probabilité pour qu'une erreur soit positive est $\frac{1}{2}$; par conséquent, la probabilité pour que, sur 2μ erreurs prises au hasard, l'*écart*, différence entre le nombre des erreurs positives et le nombre le plus probable μ, soit plus petit que $\sqrt{\frac{\mu}{2}}$ est (65)

$$\frac{2}{\sqrt{\pi}}\int_0^{\frac{1}{\sqrt{2}}} e^{-t^2}\,dt = 0,68.$$

La probabilité pour que le nombre des produits négatifs l'emporte sur celui des produits positifs est donc plus grande que $\frac{2}{3}$.

163. La formule

$$\frac{1}{n}\Sigma\left[x_i - \frac{x_1+x_2+\ldots+x_n}{n}\right]^2 = \frac{1}{2k^2} \tag{28}$$

caractérise la précision d'un système de mesures dans lesquelles une grandeur a été successivement trouvée égale à $x_1, x_2, \ldots, x_n$. On peut lui donner une forme plus élégante et plus commode. x_i et $x_{i'}$ étant deux quelconques de n mesures, l'expression est identiquement égale à

$$\frac{1}{n^2}\Sigma(x_i - x_{i'})^2. \tag{29}$$

Les deux formules (28) et (29) sont composées des mêmes termes.

On a, identiquement aussi,

$$\frac{1}{n^2}\Sigma(x_i - x_{i'})^2 = \frac{x_1^2+x_2^2+\ldots+x_n^2}{n} - \left(\frac{x_1+x_2+\ldots+x_n}{n}\right)^2.$$

La valeur de $\frac{1}{2k^2}$ est donc donnée par la différence entre la moyenne des carrés des mesures prises et le carré de leur moyenne. Si toutes les mesures sont égales, l'expression se réduit à zéro. La précision doit être supposée infinie.

164. Lorsqu'on a pris plusieurs mesures d'une même grandeur, calculé la moyenne, déterminé la valeur de la constante spécifique k, cette constante est souvent nommée la *précision* et k^2 le *poids* de chaque observation.

Il importe d'expliquer ces locutions.

La précision d'une mesure est dite α fois plus grande que celle d'une autre mesure, lorsque la probabilité d'une erreur comprise entre z et $z+dz$ pour une mesure du premier système est la même que celle d'une erreur comprise entre αz et $\alpha(z+dz)$ pour le second. Si, par exemple, les erreurs commises sur les minutes, en mesurant un angle, ont mêmes probabilités que les erreurs sur les secondes commises en se servant d'un autre instrument, la précision du second système de mesures est dite 60 fois plus grande que celle du premier.

Le poids d'une observation est dit β fois plus grand que celui d'une autre observation, lorsque les conséquences que l'on peut déduire sur la valeur de la grandeur mesurée par une observation du premier système équivalent à celles que l'on peut déduire de β observations du second système donnant toutes le même résultat.

Si β n'est pas entier et qu'il soit égal à une fraction $\frac{m}{n}$, il faudra que m observations concordantes du premier système puissent être remplacées par n observations identiques du second.

Le système d'observations qui donne à l'erreur z une probabilité proportionnelle à

$$\frac{k}{\sqrt{\pi}} e^{-k^2 z^2},$$

aura k pour précision et k^2 pour poids, si l'on prend pour unités la précision et le poids d'une observation d'un système dans lequel la probabilité d'une erreur z serait proportionnelle à e^{-z^2}.

La probabilité d'une erreur comprise entre z et $z+dz$ étant, en effet, supposée égale à

$$\frac{1}{\sqrt{\pi}} e^{-z^2}\,dz,$$

en remplaçant z par kz, la probabilité d'une erreur comprise entre

kz et $kz + k\,dz$ sera

$$\frac{k}{\sqrt{\pi}} e^{-k^2z^2}\,dz;$$

c'est précisément celle d'une erreur comprise entre z et $z + dz$ dans le premier système.

Si l'on suppose $k^2 = \frac{m}{n}$, la probabilité d'une erreur z étant proportionnelle à

$$e^{-z^2},$$

celle de m erreurs égales à z, commises dans m observations successives, sera proportionnelle à

$$e^{-mz^2}.$$

La probabilité de n erreurs égales à z, commises dans n observations successives faites dans le second système de mesures, sera proportionnelle à

$$(e^{-k^2z^2})^n = e^{-mz^2}.$$

Les probabilités des deux systèmes d'erreurs sont donc proportionnelles et par conséquent égales pour toutes valeurs de z, si l'on a

$$nk^2 = m;$$

par conséquent,

$$k^2 = \frac{m}{n}.$$

Le poids du premier système d'observations, celui du second étant pris pour unité, est donc égal à $\frac{m}{n}$, c'est-à-dire à k^2.

165. Il est intéressant de remarquer que, si la loi de probabilité n'avait pas une forme toute spéciale, les mots *poids* et *précision*, dont les physiciens font souvent usage, ne pourraient pas avoir de sens exact et précis.

Supposons deux systèmes d'observations. Soient, dans le premier, $\varphi(z)\,dz$ la probabilité pour que l'erreur commise soit comprise entre z et $z + dz$ et $\psi(z)\,dz$ la probabilité d'une même erreur dans le second système. Pour que le rapport des précisions soit h, il faut et il suffit que le rapport $\frac{\varphi(hz)}{\psi(z)}$ soit constant, et, pour que le rapport des poids d'une observation dans les deux systèmes soit égal à k, il faut que le rapport

$$\frac{\varphi(z)^k}{\psi(z)}$$

soit constant.

Pour que le système caractérisé par la fonction $\varphi(z)$ donne, par rapport à un autre système quel qu'il soit, caractérisé par $\psi(z)$, un poids déterminé et une précision définie, il faut que le rapport

$$\frac{\varphi(hz)}{\varphi(z)^k}$$

soit indépendant de z.

Posons donc

$$\varphi(hz) = G\,\varphi(z)^k, \tag{30}$$

G étant indépendant de z, et déterminons les formes possibles de la fonction $\varphi(z)$.

On peut remplacer l'équation (30) par

$$l\,\varphi(hz) = l\,G + kl\,\varphi(z)$$

et aussi, en prenant les dérivées, par

$$\frac{h\,\varphi'(hz)}{\varphi(hz)} = \frac{k\,\varphi'(z)}{\varphi(z)}.$$

En posant

$$\frac{\varphi'(u)}{\varphi(u)} = F(u),$$

la fonction F est définie par la condition

$$F(hz) = \lambda\,F(z),$$

λ représentant le rapport $\frac{k}{h}$.

Soit $\lambda = h^\mu$.

On aura

$$F(hz) = h^\mu\,F(z). \tag{31}$$

La fonction F est donc multipliée par h^μ quand la variable est multipliée par h. On aperçoit la solution

$$F(z) = Hz^{\mu},$$

H étant une constante; on en déduit

$$\frac{\varphi'(z)}{\varphi(z)} = F(z) = Hz^{\mu},$$

$$l\,\varphi(z) = H\frac{z^{\mu+1}}{\mu+1} + C,$$

$$\varphi(z) = C'e^{-az^{\mu+1}};$$

$\varphi(z)$ devant rester invariable quand on change z en $-z$, l'exposant $\mu+1$ doit être pair.

Le cas de $\mu = 1$ correspond à la loi de Gauss.

On a dans ce cas

$$\frac{k}{h} = h.$$

Le poids k est le carré de la précision h. L'équation (31) n'admet pas d'autres solutions; mais la démonstration, d'ailleurs facile, est sans intérêt pour le Calcul des Probabilités.

166. Entre plusieurs mesures d'une même grandeur, celles qui s'écartent le plus de la moyenne sont présumées avec raison, presque avec certitude, être moins bonnes que les autres : il semble naturel de les écarter.

La question est délicate. Les observateurs ne se permettent une telle suppression qu'après avoir reconnu une cause vraisemblable et intrinsèque à l'infériorité de la mesure suspecte. On croirait manquer à la sincérité due en omettant une mesure, d'ailleurs irréprochable, par cela seul qu'elle s'écarte du résultat présumé.

Il est incontestable que, si l'on est certain d'avoir opéré de la même manière et avec le même soin, toutes les observations ont le même droit à exercer leur influence sur la moyenne adoptée. Mais, si l'on opérait toujours de la *même manière*, on obtiendrait toujours le même résultat.

L'assimilation des erreurs fortuites à des tirages au sort dans une urne composée de manière à donner à chaque erreur la probabilité qui lui convient est une fiction, non une réalité. Si l'urne existait et qu'elle fût composée avec une perfection infinie, s'il arrivait qu'une série de tirages faits dans les conditions normales démentissent l'ensemble des autres, il faudrait les conserver assurément. Mais, si celui qui tirait les boules vient déclarer qu'il a négligé de les agiter, peut-être aussi de plonger la main dans les parties inférieures; s'il a pris et remis sa boule toujours à la surface, les tirages ainsi faits seront à rejeter. Ils le seront également si, le tireur n'avouant pas sa négligence, on le soupçonne d'être peu soigneux; peut-être aussi, la question est délicate je le répète, si la seule raison de le soupçonner est l'écart observé quand il a tiré les boules.

Cherchons les conséquences de l'abandon des mesures présumées les moins bonnes.

La question se pose de la manière suivante :

On a pris n mesures d'une même grandeur; tout semble régulier, la moyenne est obtenue; la constante caractéristique k, déduite de toutes les observations, a une valeur très grande; aucun indice n'éveille la défiance. On est en droit de regarder comme très probable la valeur de la moyenne; comme très probable aussi la valeur obtenue pour k. La probabilité d'une erreur z est regardée avec confiance comme égale à

$$\frac{k}{\sqrt{\pi}} e^{-k^2 z^2}\, dz.$$

Dans ces conditions, je calcule l'erreur λ, telle que la probabilité pour qu'une erreur soit inférieure à λ ait une valeur p arbitrairement choisie. Il suffit de résoudre l'équation

$$(32) \qquad p = \frac{2}{\sqrt{\pi}} \int_0^{k\lambda} e^{-t^2}\, dt = \Theta(k\lambda).$$

λ étant ainsi déterminé, je supprime parmi les n observations celles dont la différence avec la moyenne est plus grande que λ; il en restera np à peu près, dont une moitié environ sera plus grande, l'autre moitié plus petite que la moyenne. On recommencera les calculs en considérant ces observations comme les seules.

Supposons, pour justifier la hardiesse de ce parti, que, dans le laboratoire où les mesures sont prises, se trouve un surveillant d'une rare habileté qui, sans observer lui-même, suive des yeux les observateurs, et qui, très au courant des méthodes, très instruit des détails relatifs à chaque instrument, des erreurs de division, des inégalités des pas de vis, des petites imperfections des poids, de l'influence de la température, etc., déclare, après chaque mesure, son opinion en un seul mot : bien ou mal.

Personne ne contestera que, si ce surveillant est, comme nous l'avons supposé, très habile, la suppression des mesures qu'il a condamnées accroîtrait d'autant plus la confiance qu'il en aurait écarté davantage.

Quoi qu'il en soit, acceptant le principe, nous remplacerons les observations $x_1, x_2, \ldots, x_n$ par $y_1, y_2, \ldots, y_m$, m différant peu de np. Si ces m observations sont données à un calculateur soigneux, il n'aura pas besoin d'en connaître l'origine pour les déclarer suspectes. Le rapport de la moyenne du carré des erreurs au carré de la moyenne pourra différer beaucoup de $\frac{\pi}{2}$ (149).

Pour aucune valeur de k, les erreurs des observations maintenues ne s'accorderont avec la formule

$$\frac{k}{\sqrt{\pi}} e^{-k^2z^2}.$$

Le calculateur prévenu renoncera à la formule et prendra la moyenne

$$\frac{y_1 + y_2 + \ldots + y_m}{m}$$

pour représenter la grandeur inconnue.

Soient $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ les erreurs successives. La valeur du carré de l'erreur commise sur la moyenne sera

$$\left(\frac{\varepsilon_1 + \varepsilon_2 + \ldots + \varepsilon_m}{m}\right)^2 = \frac{\varepsilon_1^2 + \varepsilon_2^2 + \ldots + \varepsilon_m^2}{m^2} + \frac{2\Sigma\varepsilon_i\varepsilon_{i'}}{m^2}.$$

Négligeons $\varepsilon_i\varepsilon_{i'}$, dont la valeur probable est nulle. La valeur probable du carré de l'erreur moyenne s'obtiendra en divisant par m la moyenne

$$\frac{\varepsilon_1^2 + \varepsilon_2^2 + \ldots + \varepsilon_m^2}{m},$$

peu différente, très probablement, de la valeur probable du carré des erreurs maintenues.

Cette moyenne sera plus petite évidemment qu'avant la suppression des observations rejetées; mais il faudra, pour obtenir le carré de l'erreur probable, la diviser par m, et la fraction qu'elle remplace devait être divisée par n.

La valeur probable du carré z^2 d'une erreur plus petite que λ est, lorsque les autres ont été écartées,

$$\frac{2}{p\sqrt{\pi}}\int_0^\lambda z^2 e^{-k^2z^2}\,dz. \tag{33}$$

Le nombre des erreurs ayant été multiplié par la fraction p, la probabilité de chacune d'elles, dans l'intervalle où elles sont commises, est divisée par p. Le numérateur de la fraction qui représente la probabilité reste le même, en effet, et le dénominateur, nombre des erreurs commises, est multiplié par p. En intégrant par parties, l'expression (33) devient

$$-\frac{2k}{p\sqrt{\pi}}\left(z\frac{e^{-k^2z^2}}{2k^2}\right)_0^\lambda+\int_0^\lambda\frac{2k}{p\sqrt{\pi}}\,dz\,\frac{e^{-k^2z^2}}{2k^2};$$

c'est-à-dire, d'après l'équation (32) qui définit λ,

$$-\frac{k\lambda e^{-k^2\lambda^2}}{k^2p\sqrt{\pi}}+\frac{1}{2k^2}=\frac{1}{2k^2}\left(1-\frac{2k\lambda e^{-k^2\lambda^2}}{p\sqrt{\pi}}\right).$$

En divisant par m, égal approximativement à np, on a, pour représenter le carré de l'erreur à craindre,

$$\frac{1}{2nk^2p}\left(1-\frac{2k\lambda e^{-k^2\lambda^2}}{p\sqrt{\pi}}\right),$$

en posant, selon la notation habituelle,

$$\frac{2}{\sqrt{\pi}}\int_0^t e^{-t^2}\,dt=\Theta(t).$$

$\Theta(k\lambda)$ est, par définition, égal à p, et la valeur probable du carré de l'erreur devient

$$\frac{1}{2nk^2}\left[\frac{\Theta(k\lambda)-\dfrac{2k\lambda}{\sqrt{\pi}}e^{-k^2\lambda^2}}{\Theta(k\lambda)^2}\right].$$

La fonction

$$\frac{\Theta(t) - \frac{2t}{\sqrt{\pi}} e^{-t^2}}{[\Theta(t)]^2}$$

est nulle pour $t = 0$; elle augmente avec t et devient égale à l'unité pour une valeur infinie de t.

Voici quelques-unes de ses valeurs :

t.	
0,10	0,091
0,20	0,133
0,30	0,198
0,40	0,246
0,50	0,293
0,60	0,339
0,70	0,379
0,80	0,416

Le multiplicateur de $\frac{1}{2nk^2}$, évalué approximativement, peut diminuer sans limite. On ne doit pas oublier que, pour appliquer la formule, il faut que les observations conservées, de même que les observations faites, soient en grand nombre.

167. La loi de probabilité des erreurs dont l'exactitude a été vérifiée expérimentalement en quelque sorte (**145**) permet de prévoir, lorsque les épreuves sont nombreuses, tous les détails relatifs à la série des erreurs commises.

Nous donnerons un dernier exemple en cherchant la valeur probable de la plus petite des erreurs commises dans un grand nombre de mesures successives.

Soit z la plus petite des erreurs commises sur n mesures. La probabilité pour qu'une erreur désignée soit égale à z et soit la plus petite de toutes est

$$\frac{2k}{\sqrt{\pi}} e^{-k^2z^2}\, dz \left(1 - \frac{2k}{\sqrt{\pi}} \int_0^z e^{-k^2z^2}\, dz\right)^{n-1}.$$

Le premier facteur exprime la probabilité pour que l'erreur soit z ; le second, pour que toutes les autres soient plus grandes que z. Chacune des n mesures pouvant produire cette erreur minima, il faut multiplier l'expression par n; puis, pour avoir la valeur probable de l'erreur z dont elle représente la probabilité, il faut multiplier par z et intégrer entre o et ∞.

La valeur probable de la plus petite des n erreurs est donc

$$\frac{2nk}{\sqrt{\pi}}\int_0^\infty z\,e^{-k^2z^2}\,dz\left(1-\frac{2k}{\sqrt{\pi}}\int_0^z e^{-k^2z^2}\,dz\right)^{n-1}. \tag{34}$$

Intégrons par parties, en faisant porter l'intégration sur le facteur qui multiplie z, dont l'intégrale est

$$-\left(1-\frac{2k}{\sqrt{\pi}}\int_0^z e^{-k^2z^2}\,dz\right)^n,$$

et, en remarquant que le terme intégré est nul aux deux limites, l'expression (34) devient

$$\int_0^\infty\left(1-\frac{2k}{\sqrt{\pi}}\int_0^z e^{-k^2z^2}\,dz\right)^n dz. \tag{34'}$$

Lorsque z est petit, $\frac{2k}{\sqrt{\pi}}\int_0^z e^{-k^2z^2}\,dz$ diffère peu de kz, et, lorsque z cesse d'être petit, la puissance n placée sous le signe d'intégration est négligeable. On peut donc, en désignant par λ une valeur arbitraire de z, remplacer (34') par

$$\int_0^\lambda (1-kz)^n\,dz = \frac{1}{(n+1)k}\left[1-(1-k\lambda)^{n+1}\right].$$

n étant grand et λ n'étant pas très petit, le second terme de la parenthèse est négligeable, et la valeur probable approchée de la plus petite erreur est

$$\frac{1}{(n+1)k}.$$

168. Si l'on calcule par une méthode semblable la valeur probable de la seconde des erreurs classées par ordre de grandeur, on la trouve double de la plus petite; la troisième est triple, et ainsi de suite, tant que les formules d'approximation sont applicables. Il peut sembler étrange que, en classant les erreurs par ordre de grandeur absolue, la valeur probable de la seconde ne soit pas égale à celle de la première. Tout est symétrique, en effet, entre les erreurs positives et les erreurs négatives. Pourquoi la valeur probable de la plus petite erreur négative diffère-t-elle de celle de la plus petite erreur positive?

Elle n'en diffère pas.

La singularité apparente provient d'une confusion qui s'établit entre la valeur probable d'une grandeur et la valeur qu'il est probable de lui voir prendre.

Pierre et Paul doivent se partager un héritage. On sait que l'une des parts, on ignore laquelle, sera double de l'autre.

Les deux parts ont même valeur probable, égale pour chacune à la moitié de la somme totale.

La valeur probable de la plus grande part est double de la valeur probable de la plus petite.

169. La valeur probable $\frac{1}{(1+n)k}$ de la plus petite erreur est plus petite que la valeur probable de l'erreur commise sur la moyenne.

La valeur probable du carré de l'erreur commise sur la moyenne

$$\frac{e_1 + e_2 + \ldots + e_n}{n},$$

égale à la valeur probable de

$$\frac{e_1^2 + e_2^2 + \ldots + e_n^2}{n^2},$$

est, en effet,

$$\frac{1}{2k^2 n}.$$

Le carré de l'erreur commise sur la moyenne étant de même ordre que $\frac{1}{n}$, celle de l'erreur est de l'ordre $\frac{1}{\sqrt{n}}$, beaucoup plus grande par conséquent que la valeur probable des plus petites erreurs.

On accroîtrait beaucoup la précision d'une mesure si l'on pouvait découvrir les évaluations les meilleures et les adopter de préférence à la moyenne générale; il en serait de même si l'on pouvait prendre la moyenne des observations les meilleures. Elles sont malheureusement confondues avec les autres sans que rien les décèle avec certitude.

170. La règle acceptée comme postulatum prescrit, entre plusieurs mesures d'une même grandeur, d'adopter la moyenne. Cette règle suppose deux hypothèses : la première, que nous avons faite constamment dans l'étude de la théorie des erreurs, est l'absence de toute erreur constante ; la seconde, que toutes les observations méritent la même confiance et qu'elles aient été faites, par exemple, par le même observateur, suivant la même méthode, avec le même instrument.

Lorsque cette seconde condition n'est pas remplie, il faut apprécier ces valeurs relatives.

Supposons que les constantes caractéristiques de chaque système de mesures soient connues : soient $k_1, k_2, \ldots, k_n$ ces constantes, les poids des observations sont $k_1^2, k_2^2, \ldots, k_n^2$.

Supposons que l'on ait α_1 mesures prises dans le premier système, α_2 dans le deuxième, ..., α_n dans le $n^{\text{ième}}$.

Ramenons toutes ces mesures à d'autres dont le poids soit l'unité et supposons cette unité choisie de telle sorte que $k_1^2, k_2^2, \ldots, k_n^2$ soient des nombres entiers.

Chaque mesure de poids k_i^2 sera remplacée par le système équivalent de k_i^2 mesures ayant chacune un poids égal à l'unité. Toutes les valeurs mises en présence ayant ainsi même poids, il suffira de prendre la moyenne. α_n désignant le nombre des mesures de poids k_n^2 et x_1 la moyenne de ces mesures, on obtiendra ainsi la valeur

$$\frac{\alpha_1 k_1^2 x_1 + \alpha_2 k_2^2 x_2 + \ldots + \alpha_n k_n^2 x_n}{\alpha_1 k_1^2 + \alpha_2 k_2^2 + \ldots + \alpha_n k_n^2}. \tag{35}$$

On peut énoncer ce résultat en disant que chaque mesure est multipliée par le poids de l'observation correspondante et que la somme des produits est divisée par la somme des multiplicateurs.

Le poids de la valeur (35) est

$$\alpha_1 k_1^2 + \alpha_2 k_2^2 + \ldots + \alpha_n k_n^2.$$

La moyenne entre plusieurs observations équivaut à une observation unique ayant pour poids la somme des poids des observations partielles.

171. Il arrive quelquefois que, pour mesurer une grandeur, on la partage en plusieurs parties $a_1, a_2, \ldots, a_n$; on mesure chacune d'elles et l'on fait la somme.

Supposons que toutes ces mesures aient le même poids k^2.

Quel est le poids de la somme $a_1 + a_2 + \ldots + a_n$?

Soient $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ les erreurs, inconnues bien entendu, commises sur chacune des mesures, l'erreur sur la somme est

$$\varepsilon_1 + \varepsilon_2 + \ldots + \varepsilon_n;$$

on a

$$(\varepsilon_1 + \varepsilon_2 + \ldots + \varepsilon_n)^2 = \Sigma \varepsilon_1^2 + \Sigma \varepsilon_i \varepsilon_{i'}.$$

Cherchons la valeur probable des deux membres. Celle des produits tels que $\varepsilon_i \varepsilon_{i'}$ est nulle, puisque les erreurs positives ont, par hypothèse, même probabilité que les erreurs négatives.

Les valeurs probables des termes tels que ε_i^2 sont égales (148) à $\frac{1}{2k^2}$ et, par conséquent, puisqu'il y a n termes, la valeur probable du carré de l'erreur commise est

$$\frac{n}{2k^2}.$$

Si la probabilité d'une erreur z commise sur la somme était représentée par

$$\frac{k'}{\sqrt{\pi}} e^{-k'^2 z^2},$$

la valeur probable du carré d'une erreur serait

$$\frac{1}{2k'^2}.$$

Si donc nous posons

$$\frac{1}{2k'^2} = \frac{n}{2k^2},$$

$$k' = \frac{k}{\sqrt{n}},$$

k' est la précision d'une mesure qui donnerait au carré de l'erreur commise même valeur probable que la somme étudiée.

La confiance méritée par la moyenne de n mesures est inversement proportionnelle à $\sqrt{n}$.

172. L'armée française, pendant l'expédition d'Égypte, voulut déterminer la hauteur de la pyramide de Chéops. Les assises successives, formant pour ainsi dire les 203 marches d'un gigantesque escalier par lequel on s'élève au sommet, furent mesurées successivement par les soldats du Génie. L'erreur commise 203 fois, disait-on, sera multipliée par 203. L'erreur à craindre, répondit Fourier, est multipliée par 14, racine carrée de 203. On ignore sur quels principes l'illustre savant faisait reposer cette règle, que personne avant lui n'avait proposée.

173. Je terminerai ce Chapitre par l'examen d'une question très délicate. On a pris plusieurs mesures d'une même grandeur $x_1, x_2, \ldots, x_n$; elles ne s'accordent pas. Une formule démontrée (163) fait connaître l'évaluation de la précision des observations, déduite de la somme des carrés des différences entre les valeurs obtenues ou, ce qui revient au même, de l'excès de la moyenne de la somme des carrés sur le carré de la moyenne. Il faut appliquer cette règle avec beaucoup de prudence; elle suppose, en effet, que le mérite des observations soit, *a priori*, inconnu et que la discordance qu'elles présentent soit le seul indice dont on dispose pour leur appréciation.

Il en sera très rarement ainsi. Si les observations sont faites par un observateur habile avec un instrument excellent, elles méritent plus de confiance, lors même qu'elles s'accorderaient moins, que des observations plus concordantes d'un observateur mediocre. Il n'est pas admissible, par exemple, qu'en détachant sur les registres de Bessel quelques observations d'une même étoile on prétende en déduire une appréciation de son habileté et du mérite de son instrument.

La constante désignée par k, prise pour inconnue dans les formules trouvées (163), est souvent assez bien connue à l'avance pour que des renseignements nouveaux n'autorisent pas à en changer la valeur.

Le problème alors devient très différent.

La constante qui caractérise la précision du système d'observations étant connue, quelle doit être l'influence de l'accord plus ou moins grand des mesures sur la confiance méritée par la moyenne?

Si, par exemple, on a mesuré les trois angles d'un triangle indépendamment les uns des autres, leur somme se trouvant exactement égale à deux droits, faut-il, pour cette raison, accorder plus de confiance aux mesures que si la somme avait surpassé de 0,25 la valeur qu'elle doit avoir?

La réponse doit varier suivant les cas. Si l'observateur est inconnu aussi bien que l'instrument, on devra appliquer la formule trouvée (163), et la concordance des mesures déterminera la précision présumée des observations.

Si la constante appelée k^2 est assez bien connue, *a priori*, pour qu'on écarte l'idée d'en changer la valeur, la concordance des observations ne doit pas accroître la confiance qu'elles inspirent.

La conclusion semble inacceptable.

Si les observations sont discordantes, si leurs différences dépassent l'erreur probable de l'instrument, peut-on leur accorder autant de confiance qu'avant cet indice défavorable?

Si un observateur, quelle que soit son habileté reconnue, donne du même angle trois mesures très différentes, on n'échappera pas à l'idée que, ce jour-là, soit fatigue, soit oubli d'une précaution nécessaire, il a moins bien observé que de coutume. C'est sur cette conviction que repose la défiance éveillée par le résultat, et c'est elle, précisément, que nous écartons dans l'énoncé du problème.

L'application du Calcul des probabilités à l'étude des erreurs d'observation repose sur une fiction dont il ne faudrait pas faire une réalité. Les erreurs sont supposées tirées au sort dans une urne dont la composition est définie par la loi de probabilité acceptée.

Si l'observateur est malade, si l'on a dérangé les fils de sa lunette, exposé à l'humidité les poids de sa balance, changé son thermomètre habituel, il fait, ce jour-là, le tirage dans une autre urne, les résultats échappent à toute théorie.

Nous écartons, dans le calcul suivant, toutes les suppositions de ce genre. L'observateur s'est appliqué comme de coutume, son instrument est en bon état, les erreurs sont soumises aux chances habituelles; la constante k est connue et ne peut être changée par le succès de quelques observations nouvelles.

Il faut admettre, en outre, pour que les hypothèses n'impliquent pas contradiction, que les écarts n'aient rien d'exceptionnel.

174. Le calcul, lorsque ces réserves sont légitimes, justifie les conclusions qui précèdent.

Supposons que n mesures d'une même grandeur aient été prises. Soient $y_1, y_2, \ldots, y_n$ les erreurs commises successivement. Chacune d'elles est inconnue, bien entendu.

La probabilité d'une erreur comprise entre z et $z+dz$ étant, pour l'observation de rang i,

$$\frac{k_i}{\sqrt{\pi}} e^{-k_i^2 z^2}\, dz,$$

et les valeurs de k_i étant connues, la probabilité du concours des erreurs supposées sera

$$(36) \qquad \frac{k_1 k_2 \ldots k_n}{\pi^{\frac{n}{2}}} e^{-k_1^2 y_1^2 - k_2^2 y_2^2 - \ldots - k_n^2 y_n^2}\, dy_1\, dy_2 \ldots dy_n.$$

Soit z la moyenne des diverses mesures, prise en ayant égard à l'erreur probable de chacune d'elles, c'est-à-dire (171) en attribuant à chaque détermination de la même grandeur un poids inversement proportionnel à la valeur probable du carré de l'erreur.

On aura

$$(37) \qquad z = \frac{k_1^2 y_1 + k_2^2 y_2 + \ldots + k_n^2 y_n}{k_1^2 + k_2^2 + \ldots + k_n^2}.$$

Posons

$$\begin{aligned} y_1 &= z + \alpha_1, \\ y_2 &= z + \alpha_2, \\ &\ldots\ldots\ldots\ldots, \\ y_n &= z + \alpha_n \end{aligned}$$

et substituons dans l'expression de la probabilité (36), considérée comme un élément d'intégrale multiple, les variables z, α_1, $\alpha_2, \ldots, \alpha_{n-1}$ à $y_1, y_2, \ldots, y_n$; α_n n'est pas une variable indépendante, car on a identiquement

$$k_1^2\alpha_1 + k_2^2\alpha_2 + \ldots + k_n^2\alpha_n = 0.$$

L'élément

$$dy_1\, dy_2 \ldots dy_n$$

doit être remplacé, comme on sait, par le produit

$$dz\, d\alpha_1\, d\alpha_2 \ldots d\alpha_{n-1} \left[\mathrm{D}\, \frac{y_1\, y_2 \ldots y_n}{z . \alpha_1\, \alpha_2 \ldots \alpha_{n-1}} \right],$$

$$\mathrm{D}\, \frac{y_1\, y_2 \ldots y_n}{z . \alpha_1\, \alpha_2 \ldots \alpha_{n-1}}$$

étant le déterminant fonctionnel des variables $y_1, y_2, \ldots, y_n$ par rapport à celles qu'on leur substitue.

Ce déterminant est

$$\begin{vmatrix}
1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
\cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot \\
1 & 0 & 0 & \cdot\cdot & \cdot\cdot & \cdot\cdot & 1 & 0 \\
1 & \dfrac{-k_1^2}{k_n^2} & \dfrac{-k_2^2}{k_n^2} & \cdot\cdot & \cdot\cdot & \cdot\cdot & \dfrac{-k_{n-2}^2}{k_n^2} & \dfrac{-k_{n-1}^2}{k_n^2}
\end{vmatrix},$$

les dérivées de $z + \alpha_n$ qui forment la dernière ligne étant déduites de la relation (37).

Le déterminant est égal à

$$\frac{k_1^2 + k_2^2 + \ldots + k_n^2}{k_n^2}.$$

La probabilité (36) peut donc être remplacée par

$$(38)\quad \left\{ \begin{aligned} & \frac{k_1 k_2 \ldots k_n}{\pi^{\frac{n}{2}}}\, e^{-[z^2(k_1^2+k_2^2+\ldots+k_n^2)+k_1^2\alpha_1^2+k_2^2\alpha_2^2+\ldots+k_n^2\alpha_n^2]} \\ & \times \frac{dz\, d\alpha_1\, d\alpha_2 \ldots d\alpha_n}{k_n^2}(k_1^2 + k_2^2 + \ldots + k_n^2). \end{aligned} \right.$$

Le terme du premier degré en z, dans l'exposant de e, disparaît de l'expression (38), à cause de la relation (37).

Lorsque $\alpha_1, \alpha_2, \ldots, \alpha_{n-1}$ sont donnés, la probabilité est de la forme

$$G e^{-(k_1^2+k_2^2+\ldots+k_n^2)z^2}\,dz, \tag{39}$$

et comme il faut bien que, dans cette hypothèse, z ait une valeur comprise entre $-\infty$ et $+\infty$, l'équation

$$G\int_{-\infty}^{\infty} e^{-(k_1^2+k_2^2+\ldots+k_n^2)z^2}\,dz = 1$$

donne

$$G = \frac{\sqrt{k_1^2+k_2^2+\ldots+k_n^2}}{\sqrt{\pi}}.$$

La valeur probable du carré de l'erreur commise sur z, déduite de la formule (39), est

$$\frac{1}{2(k_1^2+k_2^2+\ldots+k_n^2)},$$

indépendante des valeurs supposées connues de $\alpha_1, \alpha_2, \ldots, \alpha_n$.

Dans le plus grand nombre des cas, quand on entreprend une série de mesures, l'habileté de l'observateur n'est ni parfaitement connue, comme nous venons de le supposer, ni complètement inconnue, comme on l'a admis (148); ce sont deux cas extrêmes. Il arrivera presque toujours que, toutes les valeurs de k étant possibles, elles seront, *a priori*, inégalement vraisemblables. Si la loi de leurs probabilités avant l'épreuve est inconnue, le problème est insoluble.

CHAPITRE IX.

ERREURS DE SITUATION D'UN POINT.

Sit p locus objecti alicujus ex observatione prima definitus, q, r, s, . ejusdem objecti loca in observationibus sequentibus, sint insuper P, Q, R, S pondera reciproce proportionalia spatiis evagationum, per quæ se diffundere possint errores ex observationibus singulis prodeuntes, quæ dantur ex datis errorum limitibus, et ad puncta p, q, r, s, .. posita intelligantur pondera P, Q, R, S, .., et inveniatur eorum gravitatis centrum Z dico punctum Z fore locum objecti, qui pro vero ejus loco tutissime haberi potest

Roger Cotes.

175. Confiance de Bravais dans la loi élémentaire de la probabilité des erreurs proposée et abandonnée par Gauss. — 176. Conséquence dans le cas où l'erreur sur chaque coordonnée est la résultante de deux erreurs élémentaires. — 177. Formule de Bravais déduite d'un postulatum. — 178. Détermination du facteur que la démonstration laisse indéterminé. — 179. Probabilité des écarts dans le tir à la cible. — 180. Ellipse dans l'intérieur de laquelle il y a probabilité donnée de voir la balle se placer — 181. Moyenne probable des valeurs de la constante. — 182, 183. Quelle est la mesure de l'habileté d'un tireur. Difficulté d'une réponse précise. — 184. Vérifications de la formule. — 185. Erreur à craindre sur la moyenne des valeurs de la constante. — 186. Valeur probable du carré de l'erreur commise dans l'évaluation du nombre des balles intérieures à une certaine ellipse. — 187. Calculs relatifs à 1000 balles tirées dans une même cible.

175. Lorsque les coordonnées d'un point sont déterminées indirectement par l'observation de grandeurs dont elles dépendent, l'erreur commise sur sa véritable place et la probabilité d'un écart donné dépendent de la loi de probabilité des erreurs commises dans les diverses mesures.

Les études faites sur cette question, particulièrement par Bravais dans un Mémoire remarqué, supposent une confiance absolue dans la loi proposée par Gauss sur la probabilité des erreurs élémentaires ; la probabilité d'une erreur z est proportionnelle à $e^{-k^2z^2}$, la constante k variant seule avec la précision des mesures.

Les formules déduites de cette hypothèse sont confirmées par les faits connus.

Bravais exprime les coordonnées du point observé en fonction linéaire des grandeurs mesurées ; cela revient à prendre pour variables les différences entre les grandeurs mesurées et les valeurs observées, en supposant ces différences assez petites pour qu'il soit permis d'en négliger le carré.

176. Supposons, par exemple, que, x et y étant les coordonnées d'un point inconnu dans un plan, on ait

$$x = X + au + bv,$$
$$y = Y + a'u + b'v,$$

X et Y étant les valeurs approchées obtenues en acceptant les mesures comme exactes et u et v les erreurs commises sur ces mesures.

La probabilité pour que l'erreur u soit comprise entre α et $\alpha + d\alpha$ étant

$$\frac{k}{\sqrt{\pi}} e^{-k^2\alpha^2}\, d\alpha,$$

et celle pour que l'erreur v soit comprise entre β et $\beta + d\beta$,

$$\frac{k'}{\sqrt{\pi}} e^{-k'^2\beta^2}\, d\beta,$$

la probabilité pour que le point dont les coordonnées sont x et y se trouve dans la partie du plan qui correspond aux limites indiquées est

$$\frac{kk'}{\pi} e^{-k^2\alpha^2 - k'^2\beta^2}\, d\alpha\, d\beta.$$

On peut, dans cet élément d'intégrale double, substituer aux variables α et β les différences

$$x - X = x_1,$$
$$y - Y = y_1.$$

Le produit $d\alpha\, d\beta$, d'après la théorie de la transformation des intégrales, sera remplacé par

$$\frac{dx_1\, dy_1}{\frac{dx_1}{d\alpha}\frac{dy_1}{d\beta} - \frac{dx_1}{d\beta}\frac{dy_1}{d\alpha}} = \frac{dx_1\, dy_1}{ab' - ba'},$$

et la probabilité pour que le point soit compris dans le rectangle infiniment petit $dx_1\, dy_1$ prendra la forme

$$\mathrm{G}\, e^{-k_1^2 x_1^2 - 2\lambda x y_1 - k_2^2 y_1^2}\, dx_1\, dy_1,$$

l'exposant de e étant la somme $-k^2\alpha^2 - k'^2\beta^2$ exprimée en fonction de x_1 et de y_1.

177. La méthode peut être étendue au cas d'un nombre quelconque de variables élémentaires. On transformera, dans tous les cas, la probabilité d'un système d'erreurs mis sous la forme d'un élément d'intégrale multiple, en introduisant au nombre des variables les coordonnées du point étudié auxquelles il faudra, dans le cas général, associer d'autres variables. Ces variables, arbitrairement choisies, disparaissent à la fin du calcul; mais les transformations intermédiaires sont fort compliquées.

On obtient très simplement le même résultat en acceptant un postulatum équivalent à celui de Gauss (138), énoncé par Cotes, en 1709, et employé ingénieusement par M. Schols pour démontrer le théorème de Bravais.

Si plusieurs positions d'un même point ont été successivement obtenues et méritent la même confiance, la position la plus probable est le centre de gravité du système des différents points considérés comme de même masse.

Soit $\varphi(\alpha, \beta)\, d\alpha\, d\beta$ la probabilité pour que l'erreur commise sur l'abscisse du point inconnu soit comprise entre α et $\alpha + d\alpha$ et l'erreur sur l'ordonnée entre β et $\beta + d\beta$, le point se trouvant compris, par conséquent, dans un rectangle infiniment petit de surface $d\alpha\, d\beta$.

Si $x_1 y_1$, $x_2 y_2$, ..., $x_n y_n$ sont les coordonnées des positions présumées d'un même point et x, y les coordonnées véritables, les erreurs successivement commises dans les diverses déterminations des coordonnées inconnues sont représentées par $x - x_1$,

$y-y_1$, ..., et la probabilité de leur concours est proportionnelle au produit

$$(1)\qquad \varphi(x-x_1, y-y_1)\varphi(x-x_2, y-y_2)\ldots\varphi(x-x_n, y-y_n).$$

La position la plus probable a pour coordonnées, d'après le postulatum,

$$(2)\qquad \begin{cases} x = \dfrac{x_1+x_2+\ldots+x_n}{n}, \\ y = \dfrac{y_1+y_2+\ldots+y_n}{n}. \end{cases}$$

Ces valeurs de x et de y doivent rendre le produit (1) maximum et, par conséquent, annuler les dérivées de ce produit par rapport aux variables x et y.

Si l'on pose

$$(3)\qquad \begin{cases} \dfrac{dl\,\varphi(x, y)}{dx} = F_1(x, y), \\ \dfrac{dl\,\varphi(x, y)}{dy} = F_2(x, y), \end{cases}$$

les coordonnées x et y définies par les équations (2) doivent satisfaire aux équations

$$F_1(x-x_1, y-y_1)+F_1(x-x_2, y-y_2)+\ldots+F_1(x-x_n, y-y_n)=0,$$
$$F_2(x-x_1, y-y_1)+F_2(x-x_2, y-y_2)+\ldots+F_2(x-x_n, y-y_n)=0.$$

La somme des variables $x-x_1$, $x-x_2$, ..., $x-x_n$ est, en vertu des équations (2), égale à zéro, ainsi que celle des variables $y-y_1$, $y-y_2$, ..., $y-y_n$; aucune autre condition ne leur est imposée. Les fonctions F_1 et F_2 sont donc définies par les conditions que les équations

$$F_1(u_1, v_1)+F_1(u_2, v_2)+\ldots+F_1(u_n, v_n)=0,$$
$$F_2(u_1, v_1)+F_2(u_2, v_2)+\ldots+F_2(u_n, v_n)=0$$

soient les conséquences nécessaires des relations

$$u_1+u_2+\ldots+u_n=0,$$
$$v_1+v_2+\ldots+v_n=0$$

entre les variables.

En réduisant à deux le nombre des variables, la condition se réduit à

$$F_1(u_1, v_1) + F_1(-u_1, -v_1) = 0;$$

la fonction F_1 change de signe quand les deux variables changent de signe.

Si l'on suppose trois variables, on obtiendra, en tenant compte du résultat précédent,

$$(4) \qquad F_1(u_1, v_1) + F_1(u_2, v_2) = F_1(u_1 + u_2, v_1 + v_2),$$

quelles que soient les variables u_1, v_1, u_2, v_2. On en déduit, en posant $\frac{dF_1(x, y)}{dx} = \varphi_1(x, y)$,

$$\varphi_1(u_1, v_1) = \varphi_1(u_1 + u_2, v_1 + v_2),$$
$$\varphi_1(u_2, v_2) = \varphi_1(u_1 + u_2, v_1 + v_2);$$

par conséquent,

$$\varphi_1(u_1, v_1) = \varphi_1(u_2, v_2).$$

La dérivée de F_1 par rapport à x est donc une constante. En différentiant l'équation (4) par rapport à v_1 et à v_2, on verra qu'il en est de même de la dérivée par rapport à y. Les dérivées de la fonction F_1 étant constantes, cette fonction est linéaire par rapport aux deux variables; il en sera de même de F_2, et nous pouvons poser, en remplaçant F_1 et F_2 par les dérivées qui les définissent (3),

$$\frac{dl\,\varphi(x, y)}{dx} = ax + by,$$
$$\frac{dl\,\varphi(x, y)}{dy} = a'x + b'y;$$

$ax + by$ et $a'x + b'y$ étant les dérivées d'une même fonction, on doit avoir

$$\frac{d(ax + by)}{dy} = \frac{d(a'x + b'y)}{dx};$$

par conséquent, $b = a'$. La fonction $l\varphi(x, y)$ définie par ses deux dérivées est nécessairement de la forme

$$l\,\varphi(x, y) = \frac{ax^2}{2} + bxy + \frac{b'y^2}{2} + C,$$

a, b, b' et C étant des constantes.

Les erreurs infinies étant impossibles, $\varphi(x, y)$ doit s'annuler quand x et y sont infinis, les constantes a et b' seront négatives.

La probabilité d'une erreur comprise entre u et $u+du$ pour x et v et $v+dv$ pour y est donc de la forme

$$G e^{-k^2u^2-2\lambda uv-k'^2v^2}\,du\,dv,$$

G étant une constante et λ, constant aussi, pouvant être positif ou négatif.

Les points d'égale probabilité sont sur une même ellipse ayant pour équation

$$k^2u^2+2\lambda uv+k'^2v^2=\mathrm{H},$$

u et v désignant les différences entre les coordonnées du point considéré et la position véritable, centre commun de toutes les ellipses semblables entre elles dont les dimensions sont proportionnelles à $\sqrt{\mathrm{H}}$.

178. Les quatre constantes G, λ, k, k' sont liées dans tous les cas par la relation nécessaire

$$\mathrm{G}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-k^2u^2-2\lambda uv-k'^2v^2}\,du\,dv=1. \tag{5}$$

Il faut bien, en effet, que les erreurs aient des valeurs comprises entre $-\infty$ et $+\infty$, et la somme des probabilités relatives à toutes les erreurs possibles est égale à l'unité.

Posons

$$k^2u^2+2\lambda uv+k'^2v^2=\mathrm{H};$$

à chaque valeur de H correspond une ellipse dont la surface est

$$\frac{\pi\mathrm{H}}{\sqrt{k^2k'^2-\lambda^2}}.$$

Tous les points de la couronne comprise entre les ellipses correspondant aux valeurs H et $\mathrm{H}+d\mathrm{H}$ du paramètre ont même probabilité, et la somme des probabilités relatives aux éléments de cette couronne est

$$G e^{-H} \frac{\pi \, dH}{\sqrt{k^2 k'^2 - \lambda^2}},$$

le facteur $\frac{\pi \, dH}{\sqrt{k^2 k'^2 - \lambda^2}}$ représentant la surface de la couronne.

La condition (5) devient

$$\frac{G\pi}{\sqrt{k^2 k'^2 - \lambda^2}} \int_0^\infty e^{-H} \, dH = 1.$$

et, comme

$$\int_0^\infty e^{-H} \, dH = 1,$$

on déduit de (5)

$$G = \frac{1}{\pi} \sqrt{k^2 k'^2 - \lambda^2}.$$

La probabilité pour que le point se trouve entre les deux ellipses qui correspondent aux paramètres H et $H + dH$ est

$$e^{-H} \, dH.$$

179. La formule précédente s'applique à la probabilité des écarts dans le tir à la cible.

La première question à résoudre dans l'étude des questions relatives à une arme donnée et à un tireur qui en fait usage est de déterminer pour cette arme et pour ce tireur les constantes caractéristiques k, k' et λ. Nous chercherons pour cela, en considérant ces constantes comme connues, les valeurs probables de u^2, v^2 et uv, u et v désignant les coordonnées du point où frappe la balle par rapport à deux axes passant par le centre de gravité de tous les points frappés, supposés, bien entendu, très nombreux.

Les valeurs probables, en vertu du théorème de Bernoulli, doivent différer très peu des valeurs données aux moyennes par le hasard.

Les coordonnées du point où frappe la balle par rapport aux axes dont l'origine est au centre de gravité des points frappés étant désignées par u et v, la probabilité pour que ce point se trouve dans le rectangle $du\,dv$ étant

$$\frac{\sqrt{k^2k'^2-\lambda^2}}{\pi}e^{-k^2u^2-2\lambda uv-k'^2v^2}\,du\,dv,$$

la valeur probable de u^2 est

$$(6)\qquad \frac{\sqrt{k^2k'^2-\lambda^2}}{\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}u^2e^{-k^2u^2-2\lambda uv-k'^2v^2}\,du\,dv.$$

On peut l'écrire

$$(7)\qquad \frac{\sqrt{k^2k'^2-\lambda^2}}{\pi}\int_{-\infty}^{\infty}u^2e^{-k^2u^2}\,du\int_{-\infty}^{\infty}e^{-k'^2\left(v+\frac{\lambda u}{k'^2}\right)^2}\,dv\,e^{\frac{\lambda^2u^2}{k'^2}};$$

on a

$$\int_{-\infty}^{\infty}e^{-k'^2\left(v+\frac{\lambda u}{k'^2}\right)^2}\,dv=\frac{\sqrt{\pi}}{k'},$$

et l'expression (7) devient

$$(8)\qquad \frac{1}{k'\sqrt{\pi}}\sqrt{k^2k'^2-\lambda^2}\int_{-\infty}^{\infty}u^2e^{-\left(k^2-\frac{\lambda^2}{k'^2}\right)u^2}\,du.$$

La formule

$$\int_{-\infty}^{\infty}z^2e^{-\mu^2z^2}\,dz=\frac{\sqrt{\pi}}{2\mu^3}$$

donne pour (8) la valeur

$$(9)\qquad \frac{2k'^2}{2(k^2k'^2-\lambda^2)}.$$

Telle est la valeur probable de u^2.

La valeur probable de uv est exprimée par

$$\frac{\sqrt{k^2k'^2-k^2}}{\pi}\int_{-\infty}^{\infty}ue^{-k^2u^2}\,du\int_{-\infty}^{\infty}ve^{-k'^2v^2-2\lambda uv}\,dv;$$

on a

$$e^{-k'^2v^2-2\lambda uv}=e^{-k'^2\left(v+\frac{\lambda u}{k'^2}\right)^2}e^{\frac{\lambda^2u^2}{k'^2}}.$$

Posons

$$v+\frac{\lambda u}{k'^2}=\frac{t}{k'},$$

on aura

$$\int_{-\infty}^{\infty} v e^{-k'^2\left(v+\frac{\lambda u}{k'^2}\right)^2} dv = \int_{-\infty}^{\infty} e^{-t^2} \frac{k't - \lambda u}{k'^3} dt = -\frac{\lambda u}{k'^3}\sqrt{\pi}.$$

La valeur probable de uv se réduit donc à

$$-\frac{\sqrt{k^2 k'^2 - \lambda^2}}{k'^3\sqrt{\pi}} \lambda \int_{-\infty}^{\infty} u^2\, du\, e^{-u^2\left(k^2 - \frac{\lambda^2}{k'^2}\right)} = -\frac{\lambda}{2(k^2 k'^2 - \lambda^2)}.$$

La valeur probable de v^2, calculée comme celle de u^2, est

$$\frac{k^2}{2(k^2 k'^2 - \lambda^2)}. \tag{10}$$

En égalant les valeurs probables aux valeurs moyennes données par l'ensemble des résultats obtenus, on obtiendra des équations dont les deux membres, si les observations sont nombreuses, différeront probablement très peu. Posons donc

$$\frac{u_1^2 + u_2^2 + \ldots + u_n^2}{n} = \mathrm{A},$$

$$\frac{v_1^2 + v_2^2 + \ldots + v_n^2}{n} = \mathrm{B},$$

$$\frac{u_1 v_1 + u_2 v_2 + \ldots + u_n v_n}{n} = \mathrm{C}.$$

Nous aurons, pour déterminer k, k' et λ, les équations

$$\frac{k'^2}{2(k^2 k'^2 - \lambda^2)} = \mathrm{A},$$

$$\frac{k^2}{2(k^2 k'^2 - \lambda^2)} = \mathrm{B},$$

$$\frac{-\lambda}{2(k^2 k'^2 - \lambda^2)} = \mathrm{C}:$$

on en déduit

$$\frac{1}{k^2 k'^2 - \lambda^2} = 4(\mathrm{AB} - \mathrm{C}^2),$$

$$k^2 = \frac{\mathrm{B}}{2(\mathrm{AB} - \mathrm{C}^2)},$$

$$k'^2 = \frac{\mathrm{A}}{2(\mathrm{AB} - \mathrm{C}^2)},$$

$$\lambda = \frac{-\mathrm{C}}{2(\mathrm{AB} - \mathrm{C}^2)}.$$

180. Les constantes k, k' et λ étant connues pour une certaine arme et pour un certain tireur, si l'on pose

$$k^2 u^2 + 2\lambda uv + k'^2 v^2 = \mathrm{H},$$

la probabilité pour que la balle soit placée entre les deux ellipses qui correspondent aux valeurs H et H + dH sera (178)

$$e^{-\mathrm{H}}\, d\mathrm{H}.$$

La probabilité pour que la balle frappe en dehors de l'ellipse correspondant à une valeur H_1 du paramètre est

$$\int_{\mathrm{H}_1}^{\infty} e^{-\mathrm{H}}\, d\mathrm{H} = e^{-\mathrm{H}_1}.$$

Si l'on pose

$$\mathrm{H}_1 = 0,69315,$$

la probabilité sera $\frac{1}{2}$. L'ellipse correspondant à cette valeur de H_1 contiendra, si les coups sont nombreux, la moitié des points frappés à très peu près.

Si l'on donne successivement à H les valeurs

0,10536,
0,22315,
0,35669,
0,51082,
0,69315,
0,91329,
1,20677,
1,60944,
2,30359,

les neuf ellipses semblables correspondant à ces valeurs du paramètre partageront le plan en dix régions contenant chacune, vraisemblablement sur un très grand nombre de coups, la dixième partie à peu près du nombre des balles. La première de ces régions est la plus petite des ellipses; la dernière est la portion du plan située au delà de la plus grande.

181. A chaque point frappé par la balle correspond une valeur

de H. La valeur probable de H est

$$\int_0^\infty \mathrm{H}e^{-\mathrm{H}}\,d\mathrm{H} = 1.$$

Telle doit donc être, avec une probabilité d'autant plus grande que les épreuves seront plus nombreuses, la moyenne des valeurs de H correspondant aux points frappés.

182. Une question fort importante doit rester indécise. Quelle est, dans un concours de tir, la règle à conseiller pour juger les tireurs? Le problème ne me semble pas comporter de solution absolue.

Si les balles tirées par deux concurrents peuvent être classées de telle sorte que chaque point frappé par le premier soit plus près du centre que le point correspondant frappé par le second, la décision semble facile. Dans ce cas-là même, si les différences sont petites, le tireur dont les balles sont moins approchées du but peut quelquefois prétendre au premier rang, en alléguant l'importance d'un écart horizontal plus grande que celle d'un écart qui laisse la balle dans l'alignement vertical.

Il semble naturel, à première vue, de mesurer le mérite d'un tireur par la surface de l'ellipse à l'intérieur de laquelle il y a probabilité donnée de voir la balle se placer.

La valeur de la différence

$$k^2 k'^2 - \lambda^2$$

représenterait alors le mérite de chacun.

Cette appréciation peut, dans des cas extrêmes, qui, très probablement, ne se sont jamais présentés et ne se présenteront jamais, donner des conséquences inacceptables.

Si l'un des tireurs plaçait toutes ses balles sur une ligne droite passant par l'origine des coordonnées qui sert de but, la règle proposée lui assignerait le premier rang, quelle que fût la distance de ses balles au centre de la cible. La surface de l'ellipse caractéristique serait nulle, en effet, dans ce cas qui, véritablement, n'est pas à craindre. Il suffit qu'il soit possible pour condamner la règle. La théorie n'en reçoit aucune atteinte.

183. Si l'on regardait *a priori* toutes les directions comme indifférentes, il faudrait supposer $\lambda = 0$, $k = k'$; la probabilité pour frapper la cible à une distance comprise entre R et $R + dR$ serait

$$2k^2 e^{-k^2 R^2} R\, dR.$$

La valeur probable de R et celle de R^2 seraient $\frac{\sqrt{\pi}}{2k}$ et $\frac{1}{2k^2}$; elles représentent, pour un grand nombre d'épreuves, la valeur moyenne des distances au centre de la cible et celle de leurs carrés.

Si dans un concours de tir on accorde le premier rang au tireur pour lequel une des deux moyennes désignée à l'avance a la plus petite valeur, les deux règles, si les épreuves sont nombreuses, donneront le même résultat; celui des tireurs pour lequel k a la plus grande valeur sera *certainement* vainqueur.

Si les épreuves sont peu nombreuses, le concours devient un jeu de hasard dans lequel, comme il est juste, un avantage est fait au plus adroit. La valeur de k, déduite de la moyenne des distances au centre de la cible, mérite moins de confiance que celle qui résulte des carrés. Le rapport des carrés des erreurs à craindre sur k, dans les deux hypothèses, se calculera comme (152); il est égal à $16 - 4\pi$.

184. La difficulté, on peut dire même l'impossibilité de donner une règle précise de préférence entre deux séries de coups résulte d'une autre circonstance encore. Les formules précédentes supposent que, sur un grand nombre de coups, le centre de gravité des points frappés coïncide avec le centre de la cible. Nous avons écarté, en un mot, les erreurs constantes. Elles ne sont jamais nulles cependant, et, quand un grand nombre d'épreuves auront été faites, l'examen du mérite de leur ensemble devra commencer par la détermination du centre de gravité du système des points frappés. La distance de ce centre à celui de la cible représentera avec une grande probabilité la partie constante de l'écart. Pour déterminer les constantes caractéristiques k, k' et λ, on devra rapporter les coordonnées des points frappés à deux axes ayant pour origine le centre de gravité. Le mérite d'un tireur se trouvera ainsi défini par cinq paramètres : les deux coordonnées a et b

du centre de gravité G de l'ensemble des points frappés et les constantes k, k', λ, déterminées au moyen des coordonnées prises par rapport à des axes passant par le point G. La somme a^2+b^2 pourrait mesurer l'imperfection systématique de l'arme employée ou le défaut constant du tireur dans sa manière de viser, et la différence $k^2k'^2-\lambda^2$ représenterait la précision. Le tireur est d'autant plus habile que ces quantités sont plus petites, mais il est impossible de leur assigner une importance relative.

Supposons deux tireurs, Pierre et Paul, ayant tiré chacun 100 balles. Les balles de Pierre sont toutes à $0^m,1$ du centre de la cible dans un cercle de $0^m,05$ de rayon. Il y a dans sa manière de tirer une cause d'erreur commune à tous les coups; à cela près, son tir approche de la perfection.

Paul, au contraire, n'a pas dans son tir d'erreur systématique. Les points frappés par ses balles entourent le centre de la cible; ils sont tous dans un cercle de $0^m,15$ de rayon.

Entre ces deux tireurs, dont l'un tire avec plus de précision, mais dont l'autre est exempt de toute erreur systématique, quel est le plus habile? La question ne peut être résolue. On pourrait la comparer à la suivante : Deux chronomètres ont été étudiés. Lorsque la température est constante, la marche du premier est plus régulière; mais, si l'on échauffe ou refroidit l'enceinte, il subit une influence plus considérable. Comment décider la préférence méritée par l'un d'eux?

Le problème est évidemment impossible à résoudre.

185. J'ai appliqué les résultats précédents à l'examen de 1000 coups tirés par des tireurs habiles à 200^m de distance avec dix armes de même modèle, chaque tireur tirant dix coups avec chaque arme.

Le centre de gravité des 1000 points frappés avait pour coordonnées

$$X = 0^m,08,$$
$$Y = 0^m,21;$$

on a trouvé, en nommant u et v les coordonnées par rapport à des axes passant par ce centre de gravité, pour l'ensemble des points frappés,

$$\frac{\Sigma u^2}{n} = 568,255 = A,$$

$$\frac{\Sigma v^2}{n} = 648,666 = B,$$

$$\frac{\Sigma uv}{n} = 33,221 = C.$$

Les équations

$$A = \frac{k'^2}{2(k^2k'^2 - \lambda^2)},$$

$$B = \frac{k^2}{2(k^2k'^2 - \lambda^2)},$$

$$C = \frac{-\lambda}{2(k^2k'^2 - \lambda^2)}$$

donnent

$$k^2 = 0,0008825,$$

$$k'^2 = 0,0007732,$$

$$\lambda = -0,0000452.$$

L'équation d'une ellipse d'égale probabilité est

$$k^2u^2 + 2\lambda uv + k'^2v^2 = H;$$

à chaque point du plan correspond une valeur de H. En faisant le calcul pour les 1000 points frappés par les balles, on a formé le Tableau suivant :

Tir de 1000 balles. — *Valeurs de* H.

	<0,105	<0,223	<0,356	<0,511	<0,693	<0,913	<1,206	<1,609	<2,303	>2,303
1.......	0,001	0,108	0,224	0,357	0,514	0,695	0,914	1,208	1,620	2,306
2.......	0,002	0,110	0,224	0,357	0,519	0,695	0,915	1,208	1,648	2,325
3.......	0,003	0,110	0,225	0,357	0,519	0,697	0,916	1,215	1,655	2,329
4	0,003	0,111	0,225	0,358	0,519	0,704	0,924	1,218	1,664	2,341
5.......	0,004	0,112	0,226	0,358	0,522	0,705	0,935	1,219	1,664	2,353
6.. ...	0,004	0,112	0,226	0,359	0,522	0,706	0,937	1,219	1,672	2,368
7..	0,008	0,112	0,227	0,359	0,523	0,706	0,939	1,223	1,675	2,379
8.... ..	0,011	0,112	0,227	0,359	0,524	0,707	0,940	1,232	1,680	2,387
9.......	0,012	0,112	0,229	0,359	0,525	0,707	0,941	1,232	1,683	2,387
10	0,013	0,112	0,231	0,360	0,525	0,709	0,942	1,232	1,687	2,400
11	0,013	0,115	0,232	0,360	0,525	0,711	0,945	1,236	1,688	2,413
12.......	0,015	0,118	0,232	0,360	0,526	0,711	0,945	1,241	1,689	2,415
13	0,016	0,118	0,232	0,361	0,527	0,712	0,947	1,241	1,704	2,445
14	0,016	0,118	0,235	0,362	0,528	0,713	0,947	1,241	1,704	2,446
15	0,016	0,119	0,235	0,363	0,529	0,714	0,950	1,241	1,707	2,460
16.......	0,016	0,122	0,236	0,363	0,531	0,717	0,950	1,244	1,710	2,468
17	0,017	0,123	0,242	0,363	0,543	0,718	0,950	1,254	1,713	2,510
18	0,020	0,124	0,249	0,366	0,544	0,719	0,950	1,264	1,713	2,516
19.......	0,021	0,124	0,251	0,369	0,545	0,719	0,954	1,270	1,714	2,543
20.. ...	0,021	0,129	0,251	0,372	0,545	0,719	0,958	1,274	1,720	2,601
21	0,022	0,130	0,252	0,374	0,548	0,719	0,961	1,277	1,724	2,608
22.......	0,023	0,130	0,252	0,375	0,549	0,722	0,965	1,280	1,724	2,666
23	0,023	0,131	0,252	0,375	0,549	0,731	0,967	1,282	1,737	2,672
24	0,025	0,131	0,253	0,376	0,551	0,733	0,969	1,285	1,737	2,673
25	0,026	0,131	0,255	0,377	0,551	0,735	0,969	1,300	1,738	2,675
26.	0,027	0,131	0,256	0,377	0,552	0,739	0,972	1,302	1,742	2,691
27.......	0,027	0,132	0,257	0,380	0,553	0,742	0,973	1,306	1,745	2,695
28.... .	0,028	0,133	0,259	0,382	0,557	0,743	0,976	1,310	1,747	2,706
29.... .	0,029	0,133	0,264	0,385	0,558	0,744	0,976	1,315	1,747	2,707
30.. ...	0,031	0,134	0,267	0,386	0,565	0,745	0,977	1,315	1,759	2,710
31.....	0,032	0,134	0,268	0,386	0,566	0,747	0,978	1,320	1,761	2,717
32	0,032	0,134	0,274	0,388	0,566	0,748	0,990	1,324	1,766	2,739
33.. ..	0,034	0,136	0,274	0,388	0,571	0,748	0,993	1,324	1,799	2,742
34......	0,035	0,137	0,277	0,389	0,572	0,751	0,995	1,329	1,801	2,762
35..... .	0,037	0,137	0,278	0,390	0,572	0,753	0,998	1,330	1,803	2,772
36	0,037	0,138	0,279	0,394	0,573	0,753	1,002	1,341	1,804	2,806
37.. ...	0,038	0,138	0,279	0,395	0,574	0,757	1,005	1,342	1,810	2,827
38.	0,039	0,140	0,281	0,396	0,575	0,759	1,009	1,343	1,813	2,891
39.. ...	0,040	0,141	0,283	0,397	0,578	0,759	1,016	1,343	1,821	2,905
40	0,042	0,141	0,284	0,404	0,580	0,764	1,020	1,343	1,833	2,923
41......	0,042	0,141	0,285	0,404	0,583	0,765	1,024	1,349	1,889	2,959
42.......	0,043	0,141	0,286	0,405	0,584	0,767	1,024	1,355	1,857	2,970
43.... .	0,045	0,143	0,286	0,406	0,584	0,768	1,028	1,357	1,862	2,976
44.......	0,045	0,147	0,286	0,407	0,586	0,768	1,037	1,361	1,883	2,980
45.......	0,045	0,149	0,287	0,409	0,587	0,770	1,041	1,367	1,884	3,029

Tir de 1000 balles. — *Valeurs de* H. (Suite.)

	<0,105	<0,223	<0,356	<0,511	<0,693	<0,913	<1,206	<1,609	<2,303	>2,303
46.......	0,046	0,150	0,287	0,410	0,589	0,770	1,044	1,370	1,885	3,114
47	0,047	0,150	0,287	0,410	0,589	0,771	1,044	1,375	1,885	3,125
48.......	0,047	0,153	0,287	0,410	0,596	0,777	1,053	1,376	1,888	3,145
49.......	0,050	0,154	0,288	0,411	0,597	0,779	1,055	1,377	1,891	3,168
50	0,050	0,154	8,288	0,411	0,598	0,779	1,056	1,384	1,895	3,225
51... ..	0,054	0,154	0,288	0,412	0,602	0,780	1,057	1,389	1,901	3,270
52	0,056	0,157	0,290	0,414	0,603	0,782	1,059	1,393	1,908	3,348
53.......	0,056	0,158	0,292	0,421	0,604	0,785	1,059	1,397	1,934	3,357
54.......	0,056	0,158	0,293	0,422	0,604	0,792	1,059	1,406	1,953	3,382
55.......	0,057	0,159	0,295	0,422	0,607	0,794	1,060	1,406	1,960	3,382
56.......	0,061	0,159	0,295	0,424	0.608	0,794	1,068	1,407	1,969	3,386
57.......	0,062	0,159	0,296	0,425	0,609	0,800	1,069	1,415	1,970	3,392
58.	0,064	0,160	0,296	0,427	0,610	0,801	1,069	1,419	1,971	3,402
59.... .	0,064	0,161	0,296	0,427	0,610	0,804	1,069	1,426	1,981	3,402
60..... .	0,067	0,161	0,297	0,427	0,611	0,806	1,079	1,427	1,988	3,413
61..... .	0,068	0,161	0,298	0,428	0,614	0,806	1,081	1,430	1,999	3,502
62.......	0,068	0,164	0,299	0,429	0,618	0,806	1,082	1,430	2,001	3,611
63......	0,068	0,168	0,301	0,431	0,623	0,807	1,093	1,430	2,006	3,611
64.......	0,069	0,168	0,303	0,437	0,623	0.812	1,094	1,435	2,018	3,619
65.......	0,069	0,170	0,306	0,437	0,633	0,814	1,098	1,437	2,023	3,622
66.... ..	0,069	0,171	0,309	0,441	0,634	0,820	1,100	1,437	2,029	3,680
67..... .	0.069	0,172	0,309	0,442	0,636	0,821	1,100	1,450	2,039	3,704
68..... .	0,072	0,173	0,310	0,445	0,637	0.821	1,109	1,454	2,048	3,716
69	0,072	0,174	0,310	0,447	0,640	0,823	1,116	1,459	2,055	3,740
70.......	0.073	0,174	0,310	0,448	0,641	0,823	1,118	1,462	2,081	3,787
71.......	0,073	0,174	0,310	0,448	0,643	0,825	1,122	1,475	2,085	3,795
72.......	0,074	0,176	0,311	0,449	0,643	0,826	1,124	1,482	2,086	3,806
73.......	0,075	0,176	0,311	0,451	0,645	0,827	1,130	1,486	2,107	3,924
74.......	0,075	0,177	0,315	0,451	0,646	0,831	1,134	1,486	2,114	3,925
75	0,077	0,178	0.318	0,458	0,646	0,836	1,145	1,490	2,115	3,946
76..... .	0,077	0,178	0,319	0,461	0,649	0,836	1,147	1,500	2,117	3,949
77.......	0,078	0,179	0,320	0,463	0,649	0,840	1,148	1,503	2,127	4,135
78... ...	0,079	0,180	0,320	0,464	0,651	0,840	1,153	1,509	2,131	4,155
79.......	0,080	0,183	0,321	0,464	0,652	0,840	1,155	1,513	2,142	4,161
80.......	0,080	0,183	0,322	0,464	0,652	0,842	1,159	1,513	2,193	4,252
81.......	0,080	0,185	0,322	0,464	0,652	0,843	1,163	1,519	2,195	4,361
82.......	0,082	0,188	0,328	0,467	0,654	0,846	1,168	1,528	2,212	4,372
83.	0,083	0,191	0,328	0,468	0,656	0,847	1,179	1,534	2,238	4,490
84.......	0,089	0,191	0,330	0,470	0,657	0,847	1,180	1,561	2,240	4,637
85...... .	0,089	0,194	0,331	0,470	0,657	0,850	1,181	1,561	2,242	4,710
86.......	0,090	0,194	0,336	0,472	0,657	0,852	1,186	1,562	2,249	4.896
87	0,092	0,195	0,337	0,473	0,660	0,852	1,192	1,566	2,265	4,943
88.......	0,092	0,195	0,337	0,474	0,660	0,855	1,197	1,568	2,273	5.613
89.......	0,093	0,197	0,340	0,475	0,662	0,858	1,197	1.579	2,281	5,868
90.......	0,095	0,197	0,340	0,480	0,664	0,859	″	1,582	2,300	5,892

TIR DE 1000 BALLES. — *Valeurs de* H. (Fin.)

	$<0,105$	$<0,223$	$<0,356$	$<0,511$	$<0,693$	$<0,913$	$<1,206$	$<1,609$	$<2,303$	$>2,303$
91......	0,096	0,198	0,342	0,485	0,667	0,859	″	1,586	″	5,955
92......	0,096	0,199	0,343	0,485	0,668	0,859	″	1,588	″	5,980
93......	0,097	0,200	0,346	0,486	0,670	0,865	″	1,603	″	7,609
94......	0,097	0,202	0,346	0,486	0,673	0,865	″	1,609	″	8,125
95......	0,097	0,203	0,351	0,489	0,676	0,866	″	″	″	8,181
96......	0,097	0,204	0,355	0,489	0,677	0,869	″	″	″	9,645
97	0,099	0,206	0,356	0,491	0,682	0,869	″	″	″	10,657
98......	0,102	0.208	0,356	0,491	0,687	0,870	″	″	″	(1)
99......	0,103	0,209	0,356	0,494	0,687	0,872	″	″	″	″
100......	″	0,214	0,356	0,494	0,690	0,881	″	″	″	″
101......	″	0,214	″	0,495	″	0,886	″	″	″	″
102......	″	0,215	″	0,500	″	0,887	″	″	″	″
103.. ...	″	0,215	″	0,501	″	0,888	″	″	″	″
104......	″	0,215	″	0,503	″	0,893	″	″	″	″
105......	″	0,217	″	0,507	″	0,894	″	″	″	″
106......	″	0,225	″	0,509	″	0,897	″	″	″	″
107......	″	″	″	0,510	″	0,897	″	″	″	″
108......	″	″	″	0,510	″	0,899	″	″	″	″
109......	″	″	″	″	″	0,900	″	″	″	″
110......	″	″	″	″	″	0,901	″	″	″	″
111......	″	″	″	″	″	0,903	″	″	″	″
112......	″	″	″	″	″	0,904	″	″	″	″
113......	″	″	″	″	″	0,906	″	″	″	″
114......	″	″	″	″	″	0,909	″	″	″	″
115.. . .	″	″	″	″	″	0,913	″	″	″	″

Les nombres de balles placées dans les divers intervalles auxquels correspondent des probabilités égales à $\frac{1}{10}$ sont, d'après ce Tableau, 99, 106, 100, 108, 100, 115, 89, 94, 90, 97.

Aucun d'eux ne s'écarte assez du nombre le plus probable 100 pour démentir la théorie.

A chacune des limites adoptées pour H correspond, nous l'avons dit, une probabilité $\frac{1}{10}$, et 100 balles par intervalle représenteraient l'événement le plus probable. L'événement le plus probable se présente rarement, il y a toujours un écart. La valeur probable du carré de l'écart calculée plus loin (187) est

$$\frac{9000}{100} = 90.$$

(1) Plus deux balles mises hors la cible.

La somme des carrés des écarts observés est

$$1 + 36 + 64 + 225 + 121 + 36 + 100 + 1 = 624;$$

la moyenne est 62,4.

Le carré de l'écart moyen est donc inférieur à sa valeur probable, et l'accord de la théorie avec les faits est aussi satisfaisant que possible.

La valeur probable de H est, nous l'avons démontré, égale à l'unité. La valeur moyenne des 998 valeurs données par le hasard est 0,981. L'accord, on le voit, est de nouveau très satisfaisant.

186. Nous avons trouvé (181) la valeur probable du paramètre H, caractéristique des ellipses de probabilité donnée, égale à l'unité. Il est intéressant de chercher la valeur probable du carré de l'erreur commise en égalant à l'unité la moyenne des valeurs de H, c'est-à-dire de calculer la valeur probable avant l'épreuve de

$$\left(\frac{H_1 + H_2 + \ldots + H_n}{n} - 1\right)^2,$$

H_1, H_2, ..., H_n étant les valeurs de H relatives aux diverses épreuves.

On a

$$\left(\frac{H_1 + H_2 + \ldots + H_n}{n} - 1\right)^2 = \frac{1}{n^2}\Sigma H_i^2 + \frac{2}{n^2}\Sigma H_i H_{i'} - \frac{2}{n}\Sigma H_i + 1. \tag{11}$$

La valeur probable de H_i^2 est

$$\int_0^\infty e^{-H} H^2\, dH = 2;$$

celle de H_i est l'unité, et la somme (11), en ayant égard au nombre de termes compris dans chaque somme Σ, est

$$\frac{2}{n} + \frac{n-1}{n} - 2 + 1 = \frac{1}{n}.$$

Telle est la valeur probable du carré de l'erreur commise en égalant la moyenne des valeurs de H à l'unité.

187. Nous avons trouvé la valeur de H pour laquelle la probabilité de voir la balle se placer à l'intérieur de l'ellipse correspondante est $\frac{1}{10}$.

Soient N le nombre des balles qui frapperont dans l'intérieur de l'ellipse, n celui des balles tirées; la différence

$$N - \frac{n}{10}$$

sera petite si n est grand. Cherchons la valeur probable du carré

$$\left(N - \frac{n}{10}\right)^2; \tag{12}$$

on a

$$\left(N - \frac{n}{10}\right)^2 = N^2 - \frac{2nN}{10} + \frac{n^2}{100}.$$

La valeur probable de N est $\frac{n}{10}$: par conséquent, celle de $\frac{2nN}{10}$ est $\frac{2n^2}{100}$, et l'expression (12) peut être remplacée par

$$N^2 - \frac{n^2}{100}.$$

$\frac{n^2}{100}$ est donné, puisque n est le nombre des balles qui ont été tirées; nous devons chercher seulement la valeur probable de N^2.

N est le nombre de balles placées dans l'intérieur de l'ellipse. Le problème est donc celui-ci :

Un événement a pour probabilité $\frac{1}{10}$; quelle est la valeur probable du carré du nombre de fois qu'il se présentera sur n épreuves?

Soient p la probabilité de l'événement (p est ici égal à $\frac{1}{10}$) et q la probabilité de l'événement contraire.

Le développement

$$(p+q)^n = p^n + np^{n-1}q + \frac{n.n-1}{1.2}p^{n-2}q^2 + \ldots + q^n$$

donne, par ses différents termes, les probabilités des combinaisons qui peuvent se produire.

Si donc on représente cette somme par

$$\Sigma A_m p^m,$$

la valeur probable du carré du nombre m d'arrivées de l'événement dont la probabilité est p est

$$\Sigma m^2 A_m p^m;$$

on a

$$(p+q)^n = \Sigma A_m p^m.$$

En prenant la dérivée par rapport à p et multipliant par p,

$$n(p+q)^{n-1}p = \Sigma m A_m p^m;$$

prenant de nouveau la dérivée par rapport à p et multipliant par p,

$$n(p+q)^{n-1}p + n(n-1)(p+q)^{n-2}p^2 = \Sigma m^2 A_m p^m$$

et, puisque $p+q=1$,

$$\Sigma m^2 A_m p^m = np + n(n-1)p^2 = n^2p^2 + npq.$$

p est $\frac{1}{10}$, q est $\frac{9}{10}$; la valeur probable de N^2 est donc

$$\frac{n^2}{100} + \frac{9n}{100}$$

et, par conséquent, celle de $N^2 - \frac{n^2}{100}$ est

$$\frac{9n}{100}.$$

La différence $N - \frac{n}{10}$ doit très probablement augmenter indéfiniment, comme le font toujours les valeurs des différences absolues entre les grandeurs dont les valeurs probables sont égales; mais la différence des valeurs relatives

$$\left(\frac{N}{n} - \frac{1}{10}\right)$$

tendra vers zéro, car la valeur probable de

$$\left(\frac{N}{n} - \frac{1}{10}\right)^2$$

est

$$\frac{1}{n^2}\left(\frac{9n}{100}\right) = \frac{9}{100n}.$$

Le rapport du nombre N des balles qui se placeront dans l'intérieur de la petite ellipse, au nombre total n des balles tirées, tend vers $\frac{1}{10}$ quand n augmente, puisque la valeur probable du carré de la différence avec $\frac{1}{10}$ est $\frac{9}{100n}$.

CHAPITRE X.

LA THÉORIE DES MOYENNES.

Errorum regularium consideratio proprie ab instituto nostro excluditur

GAUSS.

188. Abandon nécessaire de la loi de Gauss. — 189. Conditions imposées à la loi inconnue qui devrait la remplacer. — 190. Détermination expérimentale de la partie constante de l'erreur — Évaluation de l'erreur à craindre. — 191. La moyenne des mesures converge vers la valeur véritable augmentée de l'erreur constante. — 192. Valeur probable de la constante caractéristique désignée par m^2. L'évaluation de l'erreur à craindre dépend d'une constante nouvelle. — 193. La constante m^2 diminue quand on retranche l'erreur constante. — 194. Importance de la valeur de m^2; insuffisance de la formule la plus simple. Correction proposée sans preuve bien satisfaisante. — 195. Observations de mérite inégal. — Poids d'une observation. — 196 Objection de Poisson à la théorie des moyennes. — Cause de l'exception.

188. Ni le succès près des observateurs de sa loi de probabilité des erreurs, ni la simplicité des conséquences, ni leur accord constant avec les faits n'ont décidé son illustre inventeur à y voir une vérité démontrée. Nous avons indiqué les graves objections que laisse subsister (138) la démonstration. Jamais Gauss ne les a proposées, mais l'abandon de sa première théorie permet de croire qu'elles s'étaient présentées à son esprit.

Sans renoncer aux méthodes déduites de cette théorie et devenues indispensables, Gauss a voulu les établir sur des principes plus certains.

La recherche d'une loi rigoureuse pour représenter la probabilité des erreurs ne semble laisser aucun espoir de succès : les plus illustres y ont échoué et les données du problème ne semblent donner prise à aucune recherche théorique.

Gauss, sans chercher cette loi inaccessible, variable sans aucun doute d'un cas à l'autre, a su, tout en laissant la fonction indéterminée, résoudre rigoureusement le problème.

La fonction inconnue, d'après l'ingénieuse manière dont il pose la question, figure seulement dans des intégrales définies dont les valeurs numériques deviennent les constantes caractéristiques d'un système d'observations.

189. Supposons qu'en mesurant une grandeur la probabilité d'une erreur comprise entre z et $z + dz$ soit représentée par $\varphi(z)\,dz$. La fonction inconnue $\varphi(z)$ doit satisfaire à quelques conditions qu'il faut dire :

On a rigoureusement

$$\int_{-\infty}^{\infty} \varphi(z)\,dz = 1. \tag{1}$$

Il faut bien, en effet, que l'erreur ait une valeur, et la somme des probabilités pour tous les cas possibles, entre $-\infty$ et $+\infty$, représentant la certitude, doit être égale à l'unité.

Si les mesures n'ont pas d'erreur systématique et que l'instrument rende les erreurs positives aussi probables, exactement, que les erreurs négatives, on aura

$$\varphi(z) = \varphi(-z).$$

C'est ce que nous avons supposé jusqu'ici, admettant qu'avant l'étude d'un cas particulier on ait déterminé l'erreur constante de l'instrument, pour la faire disparaître ou pour en corriger les résultats.

Ne faisons pas d'abord cette hypothèse, et posons

$$\int_{-\infty}^{\infty} z\,\varphi(z)\,dz = a; \tag{2}$$

a sera une constante que l'on peut appeler l'*erreur probable*. Gauss la nomme la *partie constante de l'erreur*. Quand on l'aura déterminée pour un instrument donné et un observateur désigné, on la retranchera de chaque mesure; l'erreur, qui était z, deviendra $z - a$. En posant

$$z - a = y,$$

la probabilité de l'erreur y sera toujours $\varphi(z)\,dz$; si on la nomme $f(y)\,dy$, on aura

$$\int_{-\infty}^{\infty} y f(y)\,dy = \int_{-\infty}^{\infty} (z - a)\varphi(z)\,dz$$

et, à cause des conditions (1) et (2),

$$\int_{-\infty}^{\infty} y f(y)\,dy = 0.$$

La valeur probable des erreurs corrigées est donc égale à zéro et, ce qui revient au même, leur partie constante est nulle.

Nous parlons de la différence entre la valeur exacte et la valeur observée, qui peut être positive ou négative, et non de l'erreur absolue, toujours positive, dont la valeur probable, évidemment, ne saurait être nulle.

190. La détermination de la constante a sera facile, en général : on mesurera un grand nombre de fois une grandeur bien connue.

Soient $e_1, e_2, \ldots, e_n$ les erreurs successivement commises, on prendra

$$(3) \qquad \frac{e_1 + e_2 + \ldots + e_n}{n} = a.$$

La valeur probable de l'erreur doit, en effet, d'après le théorème de Bernoulli, différer peu de la moyenne des erreurs.

On peut aller plus loin et donner une appréciation de l'erreur à craindre, en acceptant l'équation (3).

Posons

$$(4) \qquad \int_{-\infty}^{\infty} z^2 \varphi(z)\,dz = m^2.$$

La constante m^2 est déterminée pour chaque système d'expérience. On en trouvera la valeur approchée, expérimentalement comme on a trouvé celle de a.

Cherchons la valeur probable de

$$\left(\frac{e_1+e_2+\ldots+e_n}{n}-a\right)^2; \tag{5}$$

elle donnera, évidemment, une indication de l'erreur à craindre quand on accepte comme nulle la grandeur positive qu'elle représente.

Cette expression (5) peut s'écrire

$$\frac{\Sigma e_i^2}{n^2}+\frac{2\Sigma e_i e_{i'}}{n^2}-\frac{2a}{n}\Sigma e_i+a^2.$$

La valeur probable de e_i^2 est, quel que soit i,

$$\int_{-\infty}^{\infty} z^2\varphi(z)\,dz=m^2;$$

celle de e_i,

$$\int_{-\infty}^{\infty} z\,\varphi(z)\,dz=a$$

et, par conséquent, celle de $e_i e_{i'}$ est a^2.

Ces valeurs sont les mêmes pour toutes les valeurs de i, car l'appréciation de l'erreur *à craindre* est supposée faite à l'avance : elle est relative aux instruments dont on dispose, aux méthodes employées et à l'habileté connue de l'observateur. Sa valeur n'a rien de fortuit.

L'expression (5) devient, en ayant égard au nombre des termes de chaque somme,

$$\frac{m^2}{n}+\frac{n(n-1)a^2}{n^2}-\frac{2na^2}{n}+a^2,$$

c'est-à-dire

$$\frac{m^2-a^2}{n}. \tag{6}$$

La valeur probable du carré de l'erreur commise en prenant pour a la moyenne des erreurs tend donc vers zéro lorsque n augmente.

191. La moyenne d'un nombre de mesures de plus en plus grand convergera vers la valeur véritable augmentée de l'erreur constante a.

On aura, en effet, en nommant $x_1, x_2, \ldots, x_n$ les évaluations successives d'une même grandeur z et $e_1, e_2, \ldots, e_n$ les erreurs correspondantes,

$$x_1 = z + e_1,$$
$$x_2 = z + e_2,$$
$$\ldots\ldots\ldots\ldots,$$
$$x_n = z + e_n,$$
$$\frac{x_1 + x_2 + \ldots + x_n}{n} = z + \frac{e_1 + e_2 + \ldots + e_n}{n},$$

et, puisque la moyenne des erreurs diffère peu de a quand n est grand, la moyenne des valeurs de x différera peu de $z + a$; et, si a a été donné par l'étude préalable de l'instrument et de la méthode, en le retranchant de la moyenne, on aura une évaluation de la grandeur mesurée d'autant plus certaine que les mesures seront plus nombreuses.

L'erreur commise sera exactement

$$\frac{e_1 + e_2 + \ldots + e_n}{n} - a;$$

la valeur probable de son carré est inversement (190) proportionnelle à n : on peut donc la regarder elle-même comme de l'ordre $\frac{1}{\sqrt{n}}$.

La confiance méritée par la moyenne d'une série de mesures s'accroît comme la racine carrée de leur nombre.

192. Pour obtenir, dans un système donné d'observations, la valeur probable de la constante m^2, on fera une série de mesures $x_1, x_2, \ldots, x_n$ d'une grandeur bien connue à l'avance. $e_1, e_2, \ldots, e_n$ étant les erreurs successivement commises, on pourra prendre

$$\frac{e_1^2 + e_2^2 + \ldots + e_n^2}{n} = m^2.$$

L'erreur à craindre en adoptant cette équation sera d'autant moindre que le nombre n des mesures sera plus grand. Il faut, pour l'évaluer, chercher la valeur probable de

$$\left(\frac{e_1^2 + e_2^2 + \ldots + e_n^2}{n} - m^2\right)^2.$$

Posons

$$\int_{-\infty}^{\infty} z^4 \varphi(z)\, dz = h^4;$$

h sera une nouvelle constante liée à la perfection du système d'observation et, comme m, d'autant plus petite que le système de mesures sera meilleur.

h^4 est la valeur probable de la quatrième puissance de l'erreur commise dans une observation.

On a

$$\left(\frac{e_1^2 + e_2^2 + \ldots + e_n^2}{n} - m^2\right)^2 = \frac{1}{n^2}\Sigma e_i^4 + \frac{2\Sigma e_i^2 e_{i'}^2}{n^2} - \frac{2m^2}{n}\Sigma e_i^2 + m^4. \tag{7}$$

La valeur probable de e_i^4 est h^4, quel que soit i; celle de e_i^2 est m^2 et celle de $e_i^2 e_{i'}^2$, par conséquent, m^4. La valeur probable de (7) est, par conséquent,

$$\frac{h^4}{n} + \frac{n(n-1)}{n^2} m^4 - \frac{2m^4 n}{n} + m^4,$$

c'est-à-dire

$$\frac{h^4 - m^4}{n}; \tag{8}$$

elle tend vers zéro lorsque n augmente.

193. Lorsque l'on étudie un instrument, si l'on ne peut pas faire disparaître les erreurs constantes, le premier soin doit être de déterminer l'erreur probable a; elle sera retranchée de chaque résultat donné par l'instrument, la différence devenant l'évaluation acceptée.

La valeur de m^2, relative aux mesures ainsi prises, est toujours plus petite qu'avant la correction; plus petite même que si, au lieu de retrancher a, on faisait une autre correction constante, quelle qu'elle fût.

Lorsque z est remplacé par $(z - a)$, la valeur probable du carré de l'erreur devient

$$\int_{-\infty}^{\infty}(z-a)^2\varphi(z)\,dz=\int_{-\infty}^{\infty}z^2\varphi(z)\,dz-2a\int_{-\infty}^{\infty}z\,\varphi(z)\,dz+a^2\int_{-\infty}^{\infty}\varphi(z)\,dz$$

et, à cause des équations

$$\int_{-\infty}^{\infty}\varphi(z)\,dz=1,\qquad \int_{-\infty}^{\infty}z\varphi(z)\,dz=a,$$

$$\int_{-\infty}^{\infty}(z-a)^2\varphi(z)\,dz=m^2-a^2,$$

plus petite que m^2. Il n'y a pas à craindre que a^2 soit plus grand que m^2, car le premier membre est essentiellement positif.

Si, au lieu de a, on retranchait de z une autre constante α, on aurait

$$\begin{aligned}\int_{-\infty}^{\infty}&(z-\alpha)^2\varphi(z)\,dz\\ &=\int_{-\infty}^{\infty}z^2\varphi(z)\,dz-2\alpha\int_{-\infty}^{\infty}z\varphi(z)\,dz+\alpha^2\int_{-\infty}^{\infty}\varphi(z)\,dz\\ &=m^2-2a\alpha+\alpha^2=m^2-a^2+(a-\alpha)^2,\end{aligned}$$

plus grande que m^2-a^2.

194. La valeur de m^2, quand on la calcule, comme on doit le faire, après avoir corrigé chaque observation de sa partie constante, est la mesure du degré de confiance à accorder au système considéré.

Si m^2 est petit, toute erreur qui n'est pas très petite en valeur absolue a une probabilité très petite. La valeur probable du carré de l'erreur ne pourrait pas évidemment, sans cela, être très petite.

Si m^2 est grand, on peut craindre de grandes erreurs; leur probabilité ne peut pas être petite.

La détermination de m^2 est donc importante. La règle donnée la fait dépendre de l'équation

$$\frac{e_1^2+e_2^2+\ldots+e_n^2}{n}=m^2,$$

et, pour connaître m, il faut, par conséquent, connaître d'abord les erreurs commises dans une série de mesures.

Si cette condition n'est pas remplie, on procédera comme on a fait (160) pour un problème semblable, ou, plutôt, pour résoudre le même problème que nous avons déjà rencontré. Nous avions trouvé, en étudiant une loi de probabilité d'erreurs,

$$\frac{e_1^2 + e_2^2 + \ldots + e_n^2}{n} = \frac{1}{2k^2},$$

$e_1, e_2, \ldots, e_n$ étant les erreurs successivement commises. C'est précisément la même formule, démontrée de la même manière, dans laquelle $\frac{1}{2k^2}$ représente l'intégrale

$$\int_{-\infty}^{\infty} \frac{k}{\sqrt{\pi}} z^2 e^{-k^2 z^2} dz,$$

qui remplace

$$\int_{-\infty}^{\infty} z^2 \varphi(z)\, dz,$$

lorsque la probabilité d'une erreur z, au lieu d'être $\varphi(z)$, est $\frac{k}{\sqrt{\pi}} e^{-k^2 z^2}$.

Nous pourrons, comme nous l'avons fait (160), remplacer les erreurs $e_1, e_2, \ldots, e_n$, si elles ne sont pas connues, par leurs valeurs approchées, qui seront les différences entre chaque mesure et la moyenne, et l'on devra, comme il a été expliqué, remplacer après cette substitution le dénominateur n par $n - 1$.

La démonstration, nous l'avons vu, suppose que l'on néglige un terme dont la petitesse nécessaire est très imparfaitement démontrée.

195. Lorsqu'une même grandeur X a été mesurée par des procédés différents, ou par divers observateurs avec des instruments de mérite inégal, on ne doit pas prendre la moyenne. Les observations les plus dignes de confiance doivent garder une influence plus grande.

Soient $x_1, x_2, \ldots, x_n$ n évaluations d'une même grandeur.

Supposons que chacune des évaluations soit corrigée de la partie constante, de telle sorte que la valeur probable de l'erreur ait été rendue nulle.

Soient m_1^2, m_2^2, ..., m_n^2 les valeurs supposées connues de la constante m pour chacun des systèmes de mesures qui ont fourni ces n valeurs. Si l'on fait entrer dans la détermination plusieurs mesures prises dans les mêmes circonstances, on supposera les valeurs de m égales.

Cherchons parmi les expressions de la forme

$$X = \lambda_1 x_1 + \lambda_2 x_2 + \ldots + \lambda_n x_n \tag{9}$$

celle qui doit inspirer le plus de confiance. La valeur à adopter est celle, évidemment, qui rendra minima la valeur probable du carré de l'erreur commise.

On doit avoir nécessairement

$$\lambda_1 + \lambda_2 + \ldots + \lambda_n = 1; \tag{10}$$

car, sans cela, toutes les mesures étant supposées exactes, la valeur qu'on en déduit ne le serait pas.

L'erreur commise dans l'évaluation (9) sera

$$\lambda_1 e_1 + \lambda_2 e_2 + \ldots + \lambda_n e_n;$$

elle a pour carré

$$\Sigma \lambda_i^2 e_i^2 + 2\Sigma \lambda_i \lambda_{i'} e_i e_{i'}.$$

La valeur probable de e_i est nulle; par conséquent aussi, celle de $e_i e_{i'}$; celle de e_i^2 est m_i^2. La valeur probable du carré de l'erreur commise est, par conséquent,

$$\lambda_1^2 m_1^2 + \lambda_2^2 m_2^2 + \ldots + \lambda_n^2 m_n^2.$$

Il faut choisir les valeurs de λ_4, λ_2, ..., λ_n qui rendent cette somme minima en satisfaisant à l'équation (10).

Il faut égaler à zéro les deux différentielles

$$d\lambda_1 + d\lambda_2 + \ldots + d\lambda_n = 0,$$
$$2m_1^2 \lambda_1\, d\lambda_1 + 2m_2 \lambda_2\, d\lambda_2 + \ldots + 2m_n \lambda_n\, d\lambda_n = 0.$$

La seconde de ces équations doit être la conséquence nécessaire de la première. Il faut pour cela que les coefficients des différentielles soient égaux et que l'on ait

$$m_1^2\lambda_1 = m_2^2\lambda_2 = \ldots = m_n^2\lambda_n;$$

on en déduit, à cause de (10),

$$\lambda_i = \frac{\frac{1}{m_i^2}}{\frac{1}{m_1^2}+\frac{1}{m_2^2}+\ldots+\frac{1}{m_n^2}},$$

et la valeur de X qu'il faut adopter est

$$\frac{\frac{x_1}{m_1^2}+\frac{x_2}{m_2^2}+\ldots+\frac{x_n}{m_n^2}}{\frac{1}{m_1^2}+\frac{1}{m_2^2}+\ldots+\frac{1}{m_n^2}}.$$

C'est la moyenne des valeurs successivement obtenues après que chacune a été multipliée par un facteur égal à l'inverse de la valeur correspondante m^2.

Ce facteur $\frac{1}{m^2}$, d'autant plus grand que m^2 est plus petit et que, par conséquent, la mesure mérite plus de confiance, se nomme le *poids de l'observation*.

r observations semblables équivalent, d'après la formule, à une seule qui aurait un poids r fois plus grand.

Si la probabilité d'une erreur z est

$$\frac{k}{\sqrt{\pi}}e^{-k^2z^2},$$

on a

$$m^2 = \frac{1}{2k^2}.$$

Le poids d'une observation est proportionnel à k^2.

La définition nouvelle se trouve d'accord avec celle qui a été donnée.

Le mot *précision* ne peut pas (165), dans le cas général, être défini avec la même rigueur.

196. La règle relative aux moyennes, et la sécurité qui en résulte, est démontrée indépendamment de toute hypothèse sur la loi de probabilité des erreurs.

Poisson a signalé comme une objection le cas où la fonction $\varphi(z)$ serait proportionnelle à

$$\frac{1}{k^2+z^2}.$$

En la représentant par

$$\frac{G\,dz}{k^2+z^2},$$

on doit avoir

$$\int_{-\infty}^{\infty}\frac{G\,dz}{k^2+z^2}=1;$$

on en déduit

$$\frac{\pi G}{k}=1,$$

$$G=\frac{k}{\pi}.$$

La probabilité d'une erreur comprise entre z et $z+dz$ est alors

$$(11)\qquad \frac{k}{\pi}\frac{dz}{k^2+z^2}.$$

Dans ce cas, l'intégrale

$$\frac{k}{\pi}\int_{-\infty}^{\infty}\frac{z^2\,dz}{k^2+z^2},$$

désignée dans la démonstration par m^2, est infinie. Le poids $\frac{1}{m^2}$ d'une observation est nul.

Il n'y a pas lieu de s'étonner si la conclusion est en défaut.

L'hypothèse est réalisée par une girouette qui tourne librement sans que rien la dirige. Considérée comme un instrument destiné à montrer une direction, celle du nord par exemple, elle donnera précisément la probabilité d'erreur exprimée par la formule (11).

Nommons z, en effet, la distance à laquelle la direction prolongée de la girouette coupera une perpendiculaire à la ligne que l'on prétend déterminer, si la girouette fait avec cette ligne un angle φ, on aura

$$(12)\qquad \varphi=\operatorname{arc\,tang}\frac{z}{k}.$$

Toutes les valeurs de φ comprises entre $-\frac{\pi}{2}$ et $+\frac{\pi}{2}$ sont également probables.

La probabilité pour que l'angle désigné par le hasard tombe entre φ et $\varphi + d\varphi$ est

$$\frac{d\varphi}{\pi}.$$

L'équation (12) donne

$$d\varphi = \frac{k\,dz}{k^2 + k^2}.$$

La probabilité d'une erreur, sur z, comprise entre z et $z + dz$ est donc

$$\frac{k}{\pi}\,\frac{dz}{k^2 + z^2}.$$

C'est précisément la formule (11).

CHAPITRE XI.

COMBINAISONS DES OBSERVATIONS.

Nachdem der Observator das Seinige gethan hat, ist es an dem Geometer die Unsicherheit der Beobachtungen und der Rechnung daraus abgeleiteten Grössen, nach streng mathematischen Principien zu wurdigen.

GAUSS.

197. La théorie des moyennes n'est pas applicable, en général, à la détermination simultanée de plusieurs grandeurs. — 198. Lorsque plusieurs valeurs d'une même inconnue sont indépendantes, on peut prendre la moyenne en ayant égard à leur poids; premier exemple. — 199. Deuxième exemple. 200. Troisième exemple. — 201, 202. Problème dans lequel les valeurs d'une même inconnue ne sont pas indépendantes, resolu en suivant le principe de la démonstration, dont il faut changer le détail. — 203. Problème général; première solution de Gauss. — 204. En ne faisant, en apparence, aucune hypothèse sur la loi de probabilité, on ne change pas essentiellement les conditions de l'énoncé. — 205 Substitution de la plus petite valeur probable du carré de l'erreur à l'erreur la plus probable. — 206 Lorsque le nombre des équations surpasse celui des inconnues, il existe entre les erreurs des relations nécessaires qui ne sont pas satisfaites. — 207. Expression adoptée pour l'une des inconnues; on rend le carré de la valeur probable de l'erreur minimum. — 208. Les erreurs étant très petites, la solution est la plus générale. — 209 Valeur probable du carré de l'erreur à craindre — 210. Premier exemple. — 211. Second exemple. — 212. Les valeurs probables des carrés des erreurs commises sont indépendantes de la concordance des résultats; explication de ce paradoxe. — 213. Les formules sont démontrées pour des observations qui ne sont pas encore faites. — 214. On peut se placer à un point de vue très différent; le problème devient insoluble. — Développement sur un exemple. La valeur probable *a priori* de l'inconnue que l'on veut calculer *a posteriori* est un élément nécessaire de la solution. — 215. La question appartient à la théorie de la probabilité des causes; faute de l'une des données indispensables, la solution est impossible. — 216. Discussion d'un problème analogue. — 217. Étude du problème général; les solutions sont en nombre infini. — 218. Premier exemple. — 219. Second

197. Lorsqu'une même grandeur a été mesurée plusieurs fois et que les résultats ne s'accordent pas, s'ils inspirent une égale confiance, il faut en prendre la moyenne; si leurs *poids* sont inégaux, on tient compte (195) dans le calcul de leurs valeurs relatives.

Lorsque plusieurs grandeurs ont été mesurées et qu'elles doivent servir à déterminer des inconnues par des équations plus nombreuses qu'il n'est nécessaire, le problème semble de même sorte. On possède, en effet, autant d'appréciations différentes de chaque grandeur que de groupes d'équations pouvant les déterminer; mais ces appréciations ne sont pas indépendantes : cela exige un changement de méthode.

Si l'on a, par exemple, mesuré les trois angles A, B, C et les trois côtes a, b, c d'un même triangle, on pourra adopter comme valeur de l'angle A, soit la mesure A directement obtenue, soit le supplément de la somme $B+C$, soit celle que l'on obtient en associant B ou C à deux quelconques des côtés, soit enfin prendre pour données les trois côtés.

Lors même que l'on aurait évalué les poids relatifs de ces neuf valeurs de l'angle A, la théorie des moyennes ne serait pas applicable. La combinaison qu'elle prescrit vaudrait mieux, peut-être, que l'une des mesures adoptée sans correction, mais elle n'est pas la plus plausible entre toutes. La théorie des moyennes suppose, en effet, l'indépendance des mesures associées. La valeur probable de chaque erreur est supposée nulle, ainsi que celle des produits de deux erreurs. Les erreurs positives, en d'autres termes, ont, par hypothèse, même probabilité que les erreurs négatives : cette condition n'est pas remplie dans le cas qui nous occupe. Si l'on s'est trompé, par exemple, en mesurant un angle, et cela est inévitable, les calculs par lesquels on fera servir cet angle à deux déterminations d'un autre angle donneront, vraisemblablement, des erreurs de signes contraires à celles du premier. La valeur probable du produit de ces deux erreurs sera positive.

198. Lorsque la dépendance des erreurs n'existe pas, on peut appliquer la théorie des moyennes. Nous en donnerons quelques exemples.

On veut déterminer la direction d'une ligne droite partant d'un point pris pour origine des coordonnées. On mesure pour cela les ordonnées $y_1, y_2, \ldots, y_n$ de n points de cette droite correspondant à n abscisses connues $x_1, x_2, \ldots, x_n$. Quelle valeur faut-il adopter pour le coefficient angulaire de la droite? Les mesures prises donnent, en désignant ce coefficient par a,

$$a = \frac{y_1}{x_1}, \qquad a = \frac{y_2}{x_2}, \qquad \ldots, \qquad a = \frac{y_n}{x_n}.$$

Ces déterminations indépendantes ont des poids inégaux qu'il faut calculer.

Si l'on nomme m_i^2 la valeur probable du carré de l'erreur commise sur y_i, le carré de l'erreur commise sur $\frac{y_i}{x_i}$, x_i étant exactement connu, a pour valeur probable $\frac{m_i^2}{x_i^2}$; le poids de la valeur correspondante de a est, par conséquent, $\frac{x_i^2}{m_i^2}$. On doit, avant de prendre la moyenne, multiplier chaque valeur de a par son poids. On aura

$$a = \frac{\dfrac{y_1 x_1}{m_1^2} + \dfrac{y_1 x_2}{m_2^2} + \ldots + \dfrac{y_n x_n}{m_n^2}}{\dfrac{x_1^2}{m_1^2} = \dfrac{x_2^2}{m_2^2} + \ldots + \dfrac{m_n^2}{x_n^2}}.$$

Si les mesures des ordonnées inspirent toutes la même confiance, l'expression se réduit à

$$a = \frac{x_1 y_1 + x_2 y_2 + \ldots + x_n y_n}{x_1^2 + x_2^2 + \ldots + x_n^2}.$$

199. Prenons pour inconnues, dans un second exemple, les trois angles directement mesurés d'un triangle. La somme des trois mesures ne se trouvant pas égale à deux angles droits, quelles corrections faut-il adopter?

En nommant les angles A, B, C, on a deux mesures de l'angle A :

$$A \quad \text{et} \quad 180^\circ - B - C.$$

Ces mesures sont indépendantes; on peut donc en prendre la moyenne, mais il faut calculer leur poids.

Soient m^2 le carré de l'erreur à craindre sur chacune des trois mesures; e_1, e_2, e_3 les erreurs réellement commises. L'erreur sur $180^\circ - B - C$ est $e_2 + e_3$; elle a pour carré

$$e_2^2 + e_3^2 + 2e_2e_3,$$

dont la valeur probable est $2m^2$. Les poids des deux déterminations de l'angle A sont donc $\frac{1}{m^2}$ et $\frac{1}{2m^2}$, et l'on adoptera la valeur

$$\frac{A + \frac{1}{2}(180^\circ - B - C)}{1 + \frac{1}{2}} = A + \frac{1}{3}(180^\circ - A - B - C).$$

L'erreur commise est

$$\frac{2e_1}{3} - \frac{e_2}{3} - \frac{e_3}{3}.$$

La valeur probable du carré de cette erreur est

$$m^2\left(\frac{4}{9} + \frac{1}{9} + \frac{1}{9}\right) + \frac{2m^2}{3}.$$

200. Le calcul précédent suppose les chances d'erreur dans la mesure d'un angle indépendantes de la grandeur de l'angle. Quand les mesures sont prises dans les mêmes conditions, cela est, en effet, presque absolument vrai.

Si, au lieu de mesurer les trois angles d'un triangle, on mesurait les trois parties d'une ligne très bien connue par des mesures antérieures, le problème, en apparence identique, serait, en réalité, très différent.

Soit l la longueur, supposée parfaitement connue, d'une ligne dont les trois parties a, b, c sont directement mesurées; on trouve

$$a + b + c = l + \alpha.$$

Comment doit-on répartir l'erreur α entre les trois mesures? On a deux évaluations de a : a et $l - b - c$.

Si e_1, e_2, e_3 sont les erreurs commises sur a, b, c, les erreurs commises sur les deux mesures sont e_1 et $e_2 + e_3$; mais les valeurs probables de e_1^2, e_2^2, e_3^2 sont inégales : c'est ce qui distingue ce problème du précédent. Soient m_1^2, m_2^2, m_3^2 les valeurs probables de e_1^2, e_2^2, e_3^2; on a

$$(e_2 + e_3)^2 = e_2^2 + e_3^2 + 2e_2e_3.$$

La valeur probable de $e_2\,e_3$ est nulle s'il n'y a pas d'erreur constante; les carrés des erreurs commises sur les valeurs de a ont donc pour valeurs probables m_1^2 et $m_2^2 + m_3^2$ et les poids des deux déterminations sont $\frac{1}{m_1^2}$ et $\frac{1}{m_2^2 + m_3^2}$.

On devra prendre pour valeur de a

$$\frac{\dfrac{a}{m_1^2} + (l - b - c)\dfrac{1}{m_2^2 + m_3^2}}{\dfrac{1}{m_1^2} + \dfrac{1}{m_2^2 + m_3^2}} = a + (l - a - b - c)\frac{m_1^2}{m_1^2 + m_2^2 + m_3^2}$$

La difficulté est d'évaluer les erreurs probables.

Si les mesures ont été prises en portant sur chaque ligne une unité de longueur, la valeur probable du carré de l'erreur commise, dont chaque partie peut être positive ou négative, est proportionnelle (171) au nombre des unités; on prendra donc $m_1^2 = a$, $m_2^2 = b$, $m_3^2 = c$, et la valeur la plus plausible de a est

$$a + (l - a - b - c)\frac{a}{a + b + c}.$$

201. Résolvons un problème très simple auquel la théorie des moyennes n'est pas applicable.

Supposons que, d'une station O, on ait observé quatre points A_1, A_2, A_3, A_4. On a mesuré, en les réduisant à l'horizon, les angles sous lesquels sont vues les cinq distances A_1A_2, A_1a_3, A_1A_4, A_2A_4, A_3A_4. Soient l_1, l_2, l_3, l_4, l_5 les cinq valeurs obtenues ; elles donnent évidemment trois mesures de l'angle A_1OA_2

$$l_1, \quad l_3 - l_4, \quad l_2 + l_5 - l_4.$$

Ces valeurs ne sont pas indépendantes. Les erreurs commises sur $l_3 - l_4$ et sur $l_2 + l_5 - l_4$ sont liées l'une et l'autre à l'exactitude de l_4. La probabilité pour qu'elles soient de mêmes signes est plus grande que pour qu'elles soient de signes contraires. La théorie des moyennes n'est pas applicable.

Cherchons sans changer de méthode, en modifiant seulement la démonstration, la meilleure combinaison à adopter.

Nous résoudrons deux problèmes :

Quelle est la meilleure combinaison des deux mesures $l_3 - l_4$ et $l_2 + l_5 - l_4$ qui ne sont pas indépendantes ?

Quelle est la meilleure combinaison de l_1 avec la valeur déduite des deux autres mesures ?

Pour déduire de $l_3 - l_4$ et de $l_2 + l_5 - l_4$ la valeur la plus plausible de l'angle dont ces expressions représentent deux valeurs approchées, nous prendrons pour cet angle

$$\lambda_1 (l_3 - l_4) + \lambda_2 (l_2 + l_5 - l_4), \tag{1}$$

avec la condition nécessaire

$$\lambda_1 + \lambda_2 = 1;$$

car il faut bien que, dans le cas où les deux évaluations s'accorderaient, leur moyenne soit égale à leur valeur commune.

En nommant e_1, e_2, e_3, e_4, e_5 les erreurs commises sur les cinq mesures, l'erreur commise en adoptant (1) est

$$\lambda_1 (e_3 - e_4) + \lambda_2 (e_2 + e_5 - e_4) = \lambda_2 e_2 + \lambda_1 e_3 - (\lambda_1 + \lambda_2) e_4 + \lambda_2 e_5.$$

La valeur probable du carré de l'erreur commise, en nommant m^2 la valeur probable des carrés e_1^2, e_2^2, ..., e_5^2, est

$$m^2 [\lambda_2^2 + \lambda_1^2 + (\lambda_1 + \lambda_2)^2 + \lambda_2^2] = m^2 (2\lambda_1^2 + 3\lambda_2^2 + 2\lambda_1\lambda_2).$$

Il faut déterminer λ_1 et λ_2 de telle sorte que, en supposant

$$\lambda_1 + \lambda_2 = 1,$$

la somme

$$2\lambda_1^2 + 3\lambda_2^2 + 2\lambda_1\lambda_2$$

soit minima.

On trouve

$$\lambda_1 = \frac{2}{3}, \qquad \lambda_2 = \frac{1}{3},$$

et la valeur la plus plausible est

$$(2) \qquad \frac{2}{3}(l_3 - l_4) + \frac{1}{3}(l_2 + l_5 - l_4) = \frac{1}{3} l_2 + \frac{2}{3} l_3 - l_4 + \frac{1}{3} l_5.$$

Le carré de l'erreur commise sur cette détermination a pour valeur probable

$$m^2 (2\lambda_1^2 + 3\lambda_2^2 + 2\lambda_1\lambda_2) = \frac{5}{3} m^2.$$

La valeur probable du carré de l'erreur commise sur l_1 étant m^2, le poids de cette détermination sera $\frac{1}{m^2}$; celui de l'expression (2), $\frac{3}{5m^2}$; ces deux valeurs du même angle sont indépendantes. On prendra donc, enfin, pour valeur la plus plausible déduite de l'ensemble des mesures,

$$\frac{l_1 + \frac{3}{5}\left(\frac{1}{3} l_2 + \frac{2}{3} l_3 - l_4) + \frac{1}{3} l^5\right)}{1 + \frac{3}{5}} = \frac{5 l_1 + l_2 + 2 l_3 - 3 l_4 + l_5}{8}.$$

L'erreur commise sera

$$\frac{5 e_1 + e_2 + 2 e_3 - 3 e_4 + e_5}{8},$$

dont le carré a pour valeur probable

$$\frac{m^2}{64}(25 + 1 + 4 + 9 + 1) = \frac{5 m^2}{8}.$$

202. Si, oubliant que les évaluations ne sont pas indépendantes, on avait cherché les poids des trois mesures

$$l_1, \quad l_3 - l_4, \quad l_2 + l_5 - l_\nu,$$

ils sont

$$\frac{1}{m^2}, \quad \frac{1}{2m^2}, \quad \frac{1}{3m^2},$$

on aurait adopté pour valeur de l'angle

$$\frac{l_1 + \frac{1}{2}(l_3 - l_4) + \frac{1}{3}(l_2 + l_5 - l_4)}{1 + \frac{1}{2} + \frac{1}{3}} = \frac{6l_1 + 2l_2 + 3l_3 - 5l_4 + 2l_5}{11}.$$

Le carré de l'erreur aurait pour valeur probable $\frac{78}{121}m^2$, c'est-à-dire $0{,}644\,m^2$ au lieu de $0{,}625\,m^2$.

203. Le problème général qu'il faut résoudre est le suivant :

On a fait, pour déterminer n grandeurs inconnues, $n+p$ mesures, dont les résultats s'y rattachent par des équations nécessaires. Les équations se trouvent incompatibles; quel est le meilleur système de valeurs à adopter?

Si l'on accepte pour loi de probabilité des erreurs la formule $\frac{k}{\sqrt{\pi}}e^{-k^2z^2}dz$, en supposant à la constante k une même valeur pour toutes les grandeurs directement mesurées, la théorie devient fort simple.

Les mesures obtenues étant $l_1, l_2, \ldots, l_{n+p}$, on devra, pour rendre les équations compatibles, leur faire subir des corrections $e_1, e_2, \ldots, e_{n+p}$. La probabilité pour que ces erreurs supposées aient été réellement commises est proportionnelle au produit

$$e^{-k^2(e_1^2+e_2^2+\ldots+e_{n+p}^2)};$$

elle sera maxima quand la somme des carrés des corrections sera la plus petite possible.

Le meilleur système de corrections est celui pour lequel la somme des carrés des erreurs supposées commises est un minimum.

204. Après avoir proposé le théorème précédent, dont il a suivi les conséquences avec une merveilleuse habileté, Gauss, dans ses derniers Mémoires sur la combinaison des observations, a voulu s'affranchir de toute hypothèse sur la loi de probabilité des erreurs.

Les règles prescrites n'ont pas changé pour cela.

Il doit sembler étrange que la loi de probabilité des erreurs soit sans influence sur les conclusions d'une théorie dans laquelle elle joue un si grand rôle.

L'explication est simple : une hypothèse, compatible en apparence avec toutes les lois, est introduite dans la démonstration; elle impose en réalité la même forme à toutes.

Nous supposons, dit Gauss, les observations assez exactes pour que les carrés et les produits des erreurs soient négligeables.

Toutes les équations se trouvent par là réduites au premier degré, et toutes les lois sont équivalentes.

Si l'on nomme $\varphi(z)\,dz$ la probabilité pour qu'une erreur d'observation soit comprise entre z et $z+dz$, on peut, z étant très petit, remplacer $\varphi(z)$ par le développement

$$(3)\qquad \varphi(z) = \varphi(0) + z\varphi'(0) + \frac{z^2}{1.1}\varphi''(0) + \frac{z^3}{1.2.3}\varphi'''(0) + \frac{z^4}{1.2.3.4}\varphi^{\text{IV}}(0).$$

Les erreurs constantes étant écartées, on doit avoir

$$\varphi(z) = \varphi(-z);$$

$\varphi'(0)$ et $\varphi'''(0)$ sont donc nuls.

Le droit de négliger z^2 devant z donne, *a fortiori*, celui de négliger z^4 devant z^2 et de réduire l'équation (3) à

$$\varphi(z) = \varphi(0) + \frac{1}{2}z^2\varphi''(0);$$

mais on a, en négligeant toujours z^4 devant z^2,

$$a + bz^2 = ae^{\frac{bz^2}{a}}.$$

La fonction $\varphi(z)$ est donc, en réalité, assimilée à une exponentielle de la forme

$$Ge^{-k^2z^2}.$$

205. Gauss, on le voit, aurait pu, dans les conditions où il se place, conserver sa théorie primitive. Il a fait beaucoup mieux. Cette théorie conseille, en effet, l'adoption du système de corrections dont la probabilité est maxima. La théorie nouvelle semble préférable. La valeur probable du carré de l'erreur commise est rendue minima. Les deux conditions s'accordent, mais les principes sont très différents. Il aurait pu arriver que les corrections les plus probables eussent accru, dans le cas où elles ne sont pas les véritables, les chances de commettre de très grandes erreurs. Toutes les probabilités doivent intervenir pour décider le meilleur choix à faire.

206. Soient

$$(4)\qquad \begin{cases} \varphi_1(x, y, z, \ldots) = l_1, \\ \varphi_2(x, y, z, \ldots) = l_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots \end{cases}$$

les équations qui rattachent n inconnues x, y, z, ... à $n+p$ grandeurs mesurées l_1, l_2, ..., l_{n+p}. Elles sont incompatibles et donneront seulement une valeur approchée de chaque inconnue. Nous supposerons les carrés des erreurs commises dans cette première approximation négligeables. La théorie, sans cette simplification, serait inextricable.

Le théorème de Taylor permettra d'exprimer φ_1, φ_2, ..., φ_{n+p} en fonction linéaire des accroissements dont on néglige les carrés. En éliminant entre les $n+p$ équations ainsi transformées les erreurs commises sur x, y, z, ... dans la première approximation, on obtiendra, entre les erreurs e_1, e_2, ..., e_{n+p} commises sur les grandeurs directement mesurées l_1, l_2, ..., l_{n+p}, p équations du premier degré de la forme

$$(5)\qquad \begin{cases} P_1 e_1 + P_2 e_2 + \ldots + P_{n+p} e_{n+p} = h_1, \\ Q_1 e_1 + Q_2 e_2 + \ldots + Q_{n+p} e_{n+p} = h_2, \\ R_1 e_1 + R_2 e_2 + \ldots + R_{n+p} e_{n+p} = h_3, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \end{cases}$$

207. Pour ne pas compliquer les calculs, nous supposerons six grandeurs observées et trois inconnues, dont elles sont des fonctions déterminées; les équations (5) seront alors au nombre de trois et $n+p$ sera égal à six.

Nous adopterons, pour représenter la grandeur dont la mesure a été trouvée égale à l_1, la somme

$$(6)\qquad l_1 + \lambda_1 h_1 + \lambda_2 h_2 + \lambda_3 h_3.$$

L'expression (6) se réduirait à l_1 si les mesures étaient parfaites; mais de petites erreurs ayant été commises, en les nommant e_1, e_2, e_3, e_4, e_5, e_6, celle qui en résulte pour (6) est

$$(7)\quad \left\{\begin{aligned} &e_1+\lambda_1(P_1e_1+P_2e_2+\ldots+P_6e_6)\\ &\quad+\lambda_2(Q_1e_1+Q_2e_2+\ldots+Q_6e_6)+\lambda_3(R_1e_1+R_2e_2+\ldots+R_6e_6).\end{aligned}\right.$$

Il faut choisir λ_1, λ_2, λ_3 de manière à rendre minima la valeur probable du carré de cette erreur.

En supposant les mesures dignes d'une égale confiance, et nommant m^2 la valeur probable du carré e_i^2 de l'erreur commise sur l'une quelconque d'entre elles, les valeurs probables de e_i et de $e_ie_{i'}$ étant nulles par hypothèse, la valeur probable du carré de (7) est

$$(8)\quad \left\{\begin{aligned} &m^2(1+\lambda_1^2\Sigma P^2+\lambda_2^2\Sigma Q^2+\lambda_3^2\Sigma R^2+2\lambda_1\lambda_2\Sigma PQ\\ &\qquad+2\lambda_2\lambda_3\Sigma QR+2\lambda_3\lambda_1\Sigma RP+2\lambda_1P_1+2\lambda_2Q_1-2\lambda_3R_1).\end{aligned}\right.$$

Pour rendre cette expression minima, il faut égaler à zéro les dérivées par rapport aux facteurs arbitraires λ_1, λ_2, λ_3; on écrira donc les équations

$$(9)\quad \left\{\begin{aligned} &\lambda_1\Sigma P^2+\lambda_2\Sigma PQ+\lambda_3\Sigma PR+P_1=0,\\ &\lambda_1\Sigma PQ+\lambda_2\Sigma Q^2+\lambda_3\Sigma QR+Q_1=0,\\ &\lambda_1\Sigma PR+\lambda_2\Sigma QR+\lambda_3\Sigma R^2+R_1=0.\end{aligned}\right.$$

Les facteurs λ_1, λ_2, λ_3 étant déterminés par ces équations, l'expression

$$(10)\qquad l_1+\lambda_1h_1+\lambda_2h_2+\lambda_3h_3$$

sera la meilleure valeur à adopter pour l_1.

208. Une objection se présente. L'expression (7), dans laquelle les indéterminées λ_1, λ_2, λ_3 ont été choisies le plus avantageusement possible, n'est pas la plus générale parmi celles qui, si les mesures étaient exactes, se réduiraient à l_1. On pourrait obtenir d'autres valeurs approchées en nombre infini. Pourquoi ne pas chercher entre toutes celle qui donne la plus petite erreur probable?

Si les erreurs n'étaient pas très petites, l'objection serait fondée mais, les carrés étant négligeables par hypothèse, toute fonction qui se réduit à zéro quand h_1, h_2, ... sont nuls peut être supposée du premier degré par rapport à ces quantités.

209. Les équations (9) font connaître les coefficients les meilleurs à adopter pour la formule (10). La valeur probable du carré de l'erreur est rendue minima ; elle est représentée par la somme (8), dans laquelle λ_1, λ_2, λ_3 seront déduits des équations (9).

L'expression (8) peut s'exprimer plus simplement. Si l'on ajoute les équations (9) après avoir multiplié la première par λ_1, la deuxième par λ_2 et la troisième par λ_3, en retranchant du coefficient de m^2 dans (8) la somme qui est égale à zéro, la valeur probable du carré de l'erreur commise sur l_1 prend la forme

$$m^2(1 + P_1\lambda_1 + Q_1\lambda_2 + R_1\lambda_3);$$

elle est proportionnelle à m^2. Les valeurs de P_1, Q_1, R_1, λ_1, λ_2, λ_3 ne dépendent nullement de la concordance plus ou moins parfaite des observations, révélée par les valeurs h_1, h_2, h_3 des fonctions qui devraient être nulles.

210. Supposons, pour donner un exemple très simple, que x, y, z, u soient les quatre angles d'un quadrilatère. On a trouvé pour ces angles, directement mesurés,

$$\begin{aligned} x &= l_1, \\ y &= l_2, \\ z &= l_3, \\ u &= l_4. \end{aligned}$$

Si les mesures étaient exactes, on aurait

$$l_1 + l_2 + l_3 + l_4 = 360^\circ.$$

Cette condition n'est pas remplie; on a

$$l_1 + l_2 + l_3 + l_4 - 360^\circ = h,$$

h étant supposé très petit.

On adoptera alors pour l'angle x la valeur

$$X = l_1 + \lambda h.$$

Si e_1, e_2, e_3, e_4 sont les erreurs respectivement commises sur les quatre mesures, l'erreur E commise sur X sera

$$E = e_1 + \lambda(e_1 + e_2 + e_3 + e_4).$$

En nommant m^2 la valeur probable du carré de chacune des erreurs de mesure, la valeur probable du carré de E est

$$m^2(1 + 4\lambda^2 + 2\lambda). \tag{11}$$

Elle est minima pour la valeur

$$\lambda = -\frac{1}{4},$$

et l'on doit prendre

$$X = l_1 - \frac{1}{4}h.$$

La valeur probable du carré de l'erreur est donnée par l'expression (11) quand on y suppose $\lambda = -\frac{1}{4}$. Elle est donc

$$\frac{3m^2}{4},$$

indépendante, on doit le remarquer, de la valeur de h.

211. Supposons, pour donner un second exemple, que, d'une même station O, on ait observé quatre points A_1, A_2, A_3, A_4. On a mesuré, en les réduisant à l'horizon, les angles sous lesquels les distances A_1A_2, A_1A_3, A_1A_4, A_2A_4 et A_3A_4 sont vues du point O, en désignant par l_1, l_2, l_3, l_4, l_5 les valeurs trouvées. On veut en déduire les valeurs les plus plausibles des angles x, y, z formés par OA_1 avec les trois autres directions OA_2, OA_3, OA_4.

On aurait, si les mesures étaient parfaites,

$$x = l_1, \qquad y = l_2, \qquad z = l_3, \qquad z - x = l_4, \qquad z - y = l_5 \tag{12}$$

et, par conséquent,

$$\begin{cases} l_4 + l_1 - l_3 = 0, \\ l_5 + l_2 - l_3 = 0. \end{cases} \tag{13}$$

Les équations (13) ne sont pas satisfaites et, à cause des erreurs d'observation, on a

$$(14)\qquad \begin{cases} l_4 + l_1 - l_3 = h_1, \\ l_5 + l_2 - l_3 = h_2. \end{cases}$$

On prendra pour x la valeur

$$(15)\qquad X = l_1 + \lambda_1 h_1 + \lambda_2 h_2.$$

Si e_1, e_2, e_3, e_4, e_5 sont les erreurs commises sur les quatre mesures, l'erreur commise sur X sera

$$(16)\qquad E = e_1 + \lambda_1(e_4 + e_1 - e_3) + \lambda_2(e_5 + e_2 - e_3).$$

Si m^2 désigne la valeur probable du carré de chaque erreur de mesure e_1, e_2, e_3, e_4, e_5, la valeur probable de E^2 est

$$(17)\qquad m^2(1 + 3\lambda_1^2 + 3\lambda_2^2 + 2\lambda_1 - 2\lambda_1\lambda_2).$$

Le minimum de cette expression correspond à

$$\lambda_1 = -\frac{3}{8}, \qquad \lambda_2 = \frac{1}{8}.$$

Nous adopterons donc, pour l'angle x, la valeur

$$(18)\qquad X = l_1 - \frac{3}{8} h_1 + \frac{1}{8} h_2.$$

Elle s'accorde, on le voit en remplaçant h_1 et h_2 par leurs valeurs (14), avec la solution obtenue (201) par une voie différente.

Le carré de l'erreur commise en adoptant l'expression (18) est $\frac{5}{8} m^2$; elle est indépendante de h_1 et de h_2, par conséquent de l'exactitude des mesures.

On trouverait, par des calculs semblables,

$$Y = l_2 + \frac{1}{8} h_1 - \frac{3}{8} h_2,$$

$$Z = l_3 + \frac{1}{4} h_1 + \frac{1}{4} h_2.$$

Les carrés des erreurs commises ayant pour valeurs probables $\frac{5m^2}{8}$

et $\frac{7m^2}{8}$, les valeurs les plus plausibles des angles l_4 et l_5 sont

$$L_4 = l_4 - \frac{3}{8}h_1 + \frac{1}{8}h_2,$$

$$L_5 = l_5 + \frac{1}{8}h_1 - \frac{3}{8}h_2.$$

212. Les valeurs probables des carrés des erreurs commises sont, dans tous les cas, indépendantes de l'accord plus ou moins parfait des observations. Les quantités désignées par h_1, h_2, ..., qui seraient nulles si les observations étaient parfaites, ne figurent pas dans l'évaluation de l'erreur à craindre.

Nous avons déjà (173) rencontré et expliqué ce paradoxe. Il n'est pas inutile d'y revenir.

Dans la détermination de l'erreur probable, la précision des observations a été supposée connue. Le facteur m^2 représente la valeur probable du carré de l'erreur commise sur chaque mesure. En supposant ainsi l'habileté de l'observateur évaluée à l'avance, sans qu'il soit tenu compte dans cette appréciation des discordances révélées par la comparaison des mesures, il ne faut pas s'étonner de ne pas voir figurer ces discordances dans le calcul de l'erreur à craindre.

On mesure, par exemple, les trois angles d'un triangle; on a vu déjà l'observateur à l'œuvre, il fait usage d'excellents instruments; on apprécie en conséquence la constante m^2 : l'erreur probable est $0'',50$. La somme des angles obtenus surpasse cependant $180°$ de $12''$. L'observateur ne méritait pas la confiance accordée. La valeur de m^2 était très probablement mal choisie. Mais, pour en adopter une autre, on manque de données suffisantes.

213. Les formules démontrées sont applicables à des observations qui ne sont pas encore faites; elles indiquent les calculs par lesquels les inconnues devront se déduire des grandeurs directement mesurées. Les valeurs probables des carrés des erreurs à craindre dépendent de la précision espérée pour les mesures qu'on va prendre. Le cas où cette précision est assez bien connue, *a priori,* pour que les résultats obtenus n'y puissent rien changer, quoi qu'il arrive, est tout à fait exceptionnel.

C'est à lui que se rapportent les formules.

214. On s'est placé quelquefois à un point de vue absolument opposé. La précision des mesures est supposée inconnue; la concordance plus ou moins parfaite des observations est le seul renseignement d'après lequel on puisse l'apprécier.

Ce problème est le contraire du précédent. Nous supposions la précision connue *a priori;* le résultat plus ou moins heureux des observations n'y pouvait rien changer; pour rendre cette hypothèse acceptable, nous supposions même que les observations ne fussent pas faites encore.

On suppose, au contraire, dans le nouveau problème, la précision complètement inconnue. Il faut la calculer d'après les résultats, qui sont, cette fois, la seule donnée.

Gauss a fait reposer la solution de ce problème sur une formule très élégante, qui sera démontrée à la fin de ce Chapitre.

La formule est irréprochable; mais l'application est rarement permise.

Le problème n'est pas nettement posé.

Quelques exemples rendront la difficulté très claire.

On a mesuré les trois angles d'un triangle; l'excès de leur somme sur deux angles droits peut-il faire connaître, indépendamment de tout autre renseignement, la valeur probable du carré de l'erreur commise dans la mesure de chacun des angles?

Si l'on admet, comme il est vrai, qu'à chaque instrument manié par un observateur désigné correspond, *objectivement,* une valeur probable déterminée du carré de l'erreur et que, en donnant la somme des trois angles obtenus, on demande la valeur vraie de cette constante caractéristique de la précision, le problème est insoluble.

Une valeur vraisemblable est, évidemment, tout ce qu'il est permis d'espérer.

En réduisant le problème à ces termes très vagues, une lacune subsiste dans l'énoncé; elle doit enlever toute confiance dans le résultat.

La connaissance de la valeur probable *a priori* de l'inconnue dont on veut déterminer *a posteriori* la valeur vraisemblable est un élément essentiel de la question; on ne donne sur lui aucune indication.

Les trois angles d'un triangle ont été mesurés par un observateur très habile; il a fait usage d'un excellent instrument; chaque mesure a été prise trois fois; les résultats proposés sont les moyennes des trois observations. La somme des angles, après toutes ces précautions, surpasse 180° de $0'',25$.

Les mêmes angles sont mesurés par un débutant qui s'exerce; l'instrument qu'on lui a confié est médiocre. Chaque angle n'est mesuré qu'une fois; les angles obtenus diffèrent des précédents de plusieurs secondes chacun. La somme des angles, pour ces secondes mesures, est exactement 180°.

Quels sont les résultats les plus dignes de confiance?

Les premiers, évidemment.

Les formules qui déduiront la précision des mesures de l'accord plus ou moins parfait des résultats ne peuvent manquer de donner l'avantage aux seconds.

Le cas, pourrait-on répondre, n'est pas celui qu'on a supposé. Les deux séries de mesures sont faites dans des conditions telles qu'avant d'en connaître le résultat, sans donner la mesure numérique des deux précisions, on les propose comme très inégales. L'énoncé du problème résolu par Gauss suppose, au contraire, que l'on ne sache rien sur la précision des mesures.

Est-il possible, quand on combine des mesures, de ne rien savoir sur leur précision? Savoir, comme on l'admet, que toutes les valeurs de l'erreur probable sont *a priori* également vraisemblables serait un renseignement très précis qui, vraisemblablement, n'a dans aucun cas représenté la vérité.

Sans avoir étudié un instrument, le nom du fabricant, le prix dont on l'a payé, la situation de l'observateur qui s'en sert, font que certaines évaluations de sa précision, sans être tenues pour impossibles, seraient accueillies avec étonnement. Cela suffit pour changer les conditions de l'énoncé.

215. La question appartient à la théorie de la probabilité des causes. Le désaccord entre les mesures prises est un fait observé. Les causes possibles, en nombre infini, sont la précision de chaque mesure. Quelle que soit cette précision, l'événement observé est possible. Le plus adroit peut avoir une défaillance; le plus maladroit peut, par un heureux hasard, obtenir de bons résultats : les résultats très inexacts peuvent se compenser fortuitement.

Plus la concordance est grande, assurément, plus il est probable qu'elle a pour cause l'exactitude des mesures et que celle-ci est due à l'habileté de l'observateur.

La probabilité assignée à chaque cause dépend (115) de deux facteurs : la probabilité que la cause donne à l'événement observé et la probabilité, *a priori*, pour que la cause ait agi.

On veut, d'après l'énoncé, comme on l'a fait trop souvent en d'autres circonstances (**123**, **124**, **130**), se passer complètement de la seconde donnée. C'est une faute contre les principes. La lacune laissée dans l'énoncé sera forcément remplacée dans chaque solution obtenue par une condition introduite arbitrairement.

216. Supposons, pour faire connaître le principe de la méthode adoptée, qu'on veuille déterminer la *chance* pour qu'une pièce de monnaie désignée retombe sur le côté face quand on la jette en l'air.

La pièce est jetée μ fois : elle a montré m fois face et n fois pile.

Soient p la probabilité inconnue qu'elle donne à l'arrivée de face, q celle qu'elle donne à l'arrivée de pile.

L'événement le plus probable sur μ épreuves est μp fois face. En égalant la valeur probable à la valeur vraie, nous aurons

$$\mu p = m.$$

La valeur *vraisemblable* de p serait donc $\frac{m}{\mu}$.

Le principe, appliqué à un petit nombres d'épreuves, donnerait des résultats inacceptables. Si, sur trois épreuves, la pièce a montré deux fois face, oserait-on proposer $\frac{2}{3}$ comme valeur vraisemblable de la probabilité qu'elle donne à l'arrivée de face?

Lorsque le nombre des épreuves est grand, le résultat cesse d'être choquant. La précision de la valeur proposée pour la probabilité p n'est pas pour cela mieux justifiée.

Une même pièce a été jetée 1 000 000 de fois : on a obtenu 500 391 fois face; on en conclut que la pièce donne, *vraisemblablement*, à la sortie de face, la probabilité

$$p = 0{,}500\,391. \tag{19}$$

Aucune de ces six décimales ne mérite confiance.

La probabilité p a une valeur objective. A chaque pièce, d'après sa structure, correspond une valeur déterminée de p. Personne n'a pu croire, bien entendu, que cette valeur soit égale à (19); mais il n'est pas même très probable qu'elle soit plus grande que 0,50.

Supposons que, la pièce étant bien connue, la valeur exacte de p soit

$$p = 0,499609, \tag{20}$$

c'est-à-dire qu'elle s'écarte de $\frac{1}{2}$ précisément autant que la valeur (19) indiquée par le calcul, mais en sens inverse.

Cherchons quelle serait, dans cette hypothèse, la probabilité de l'événement observé.

Le nombre le plus probable des arrivées de face est

$$499609.$$

L'événement observé est l'arrivée de 500391 fois face; l'écart est 782.

La probabilité d'un écart égal à h est

$$\frac{1}{\sqrt{2\mu\pi pq}} e^{-\frac{h^2}{2\mu pq}};$$

on a

$$\frac{h^2}{2\mu pq} = \frac{(782)^2}{500000} = 1,22305.$$

La probabilité donnée par l'hypothèse à l'événement observé est

$$\frac{1}{\sqrt{2\pi\mu pq}} e^{-1,22305};$$

celle que donne au même événement l'hypothèse la plus plausible pour laquelle il correspond à un écart nul est

$$\frac{1}{\sqrt{2\mu\pi pq}}.$$

Le rapport est

$$e^{-1,22305} = 0,294332.$$

Le rapport des probabilités de deux causes également probables *a priori* est exactement celui des probabilités observées. La valeur (19) de p n'est donc pas quatre fois plus probable que la valeur (20), que nous avons choisie, on peut le dire, absolument contraire.

217. Reprenons la question générale.

On a mesuré $n+p$ grandeurs l_1, l_2, l_3, ..., l_{n+p}. Elles sont liées par des équations rigoureuses à n inconnues x, y, z, ... En éliminant les inconnues, on obtient p équations nécessaires entre les $n+p$ grandeurs mesurées :

$$\begin{gathered}
F_1(l_1, l_2, \ldots, l_{n+p}) = 0, \\
F_2(l_1, l_2, \ldots, l_{n+p}) = 0, \\
\ldots\ldots\ldots\ldots\ldots\ldots, \\
F_p(l_1, l_2, \ldots, l_{n+p}) = 0.
\end{gathered}$$

Les mesures n'étant pas parfaites, ces équations ne seront pas satisfaites : les premiers membres auront de petites valeurs $-h_1$, $-h_2$, ..., $-h_n$ qui seront exactement connues, puisque les mesures l_1, l_2, ..., l_{n+p} le sont.

Si l'on nomme e_1, e_2, ..., e_{n+p} les erreurs commises sur ces mesures et qu'on en néglige les carrés, la fonction F_1, qui est égale à $-h_1$ et qui doit être nulle après les corrections, recevra un accroissement h_1 ; on obtiendra ainsi p équations du premier degré

$$\left\{\begin{aligned}
P_1 e_1 + P_2 e_2 + \ldots + P_{n+p} e_{n+p} &= h_1, \\
Q_1 e_1 + Q_2 e_2 + \ldots + Q_{n+p} e_{n+p} &= h_2, \\
R_1 e_1 + R_2 e_2 + \ldots + R_{n+p} e_{n+p} &= h_3, \\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots&\ldots,
\end{aligned}\right. \tag{21}$$

h_1, h_2, ..., h_n étant des nombres connus dans chaque problème.

Le principe admis est celui-ci :

Il est permis, à titre d'approximation, d'égaler une fonction des premiers membres de (21), dont la valeur est exactement connue, à la valeur probable calculée avant l'épreuve.

Cette égalité, bien entendu, n'a jamais été proposée comme certaine : *Valor verus*, dit Gauss, *prout fors errores obtulit, major minorve medii fieri potest.*

Ces deux grandeurs, la valeur véritable et la valeur probable, que le hasard peut faire plus grandes ou plus petites l'une que l'autre dans une proportion inconnue, sont égalées cependant pour former l'équation dont la solution est déduite.

La critique est rendue difficile. Quand on a dit : le second membre de l'équation proposée peut être, suivant la décision du hasard, plus grand ou plus petit que le premier, on s'est mis en règle avec la rigueur : le lecteur averti sait qu'on n'y prétend pas. De quel droit reprocher une cause d'erreur nettement signalée?

Le résultat est donné comme une approximation; il est tout naturel, au contraire, de chercher quelle confiance mérite le principe sur lequel elle repose. Nous montrerons que, le principe étant admis, on peut en déduire, pour la précision, des valeurs très différentes et dont aucune, par conséquent, ne mérite confiance.

Formons un polynome

$$\lambda_1 h_1^2 + \lambda_2 h_2^2 + \ldots + \lambda_p h_p^2 + \lambda_{p+1} h_1 h_2 + \ldots, \tag{22}$$

homogène et du second degré par rapport aux seconds membres, numériquement connus, des équations (21). Quels que soient les coefficients choisis, $\lambda_1, \lambda_2, \ldots$, l'expression (22) est connue.

On peut en déterminer la valeur probable, avant les épreuves faites, en fonction de la valeur probable m^2 du carré de l'erreur commise sur chaque mesure.

$h_1, h_2, \ldots, h_p$ sont exprimés par les équations (21) en fonction des erreurs $e_1, e_2, \ldots, e_{n+p}$. Le polynome (22) est donc une fonction connue des erreurs commises dans les mesures; on en peut former la valeur probable en remarquant que celle de e_i^2 est m^2 et que celle de $e_i e_{i'}$ est nulle, quels que soient i et i'. La valeur probable du polynome (22) sera donc de la forme Gm^2, G étant connu. En l'égalant à sa valeur vraie, puisque tel est le principe accepté, on obtiendra une valeur de m^2 dans laquelle figureront les facteurs arbitraires désignés par λ.

218. Il ne sera pas inutile de donner une application.

On a mesuré les trois angles d'un triangle; les erreurs commises e_1, e_2, e_3 sont inconnues, mais leur somme est exactement connue; on a

h_1 étant l'excès de la somme des angles mesurés sur deux angles droits.

On en déduit

$$h_1^2 = e_1^2 + e_2^2 + e_3^2 + 2e_1e_2 + 2e_2e_3 + 2e_3e_1.$$

Si m^2 est la valeur probable du carré de chacune des erreurs e_1, e_2, e_3, la valeur probable du second membre est $3m^2$, et l'on écrira, *en égalant cette valeur probable à la valeur vraie,*

$$h^2 = 3m^2;$$

par conséquent,

$$m^2 = \frac{h^2}{3}.$$

Comme il n'y a qu'une seule équation entre les erreurs, il n'y a pas, dans ce cas, de choix à faire entre les combinaisons.

219. Reprenons le problème résolu (**203**). On a mesuré cinq angles l_1, l_2, l_3, l_4, l_5, entre lesquels les conditions géométriques du problème donnent les relations nécessaires

$$l_4 + l_1 - l_3 = 0,$$
$$l_5 + l_2 - l_3 = 0$$

Les mesures étant imparfaites, on a trouvé

$$l_4 + l_1 - l_3 = h_1,$$
$$l_5 + l_2 - l_3 = h_2.$$

En désignant les erreurs réellement commises par e_1, e_2, e_3, e_4, e_5, on a donc

$$e_4 + e_1 - e_3 = h_1,$$
$$e_5 + e_2 - e_3 = h_2.$$

Quels que soient les facteurs λ_1, λ_2, λ_3, le trinome

$$\lambda_1 h_1^2 + \lambda_2 h_2^2 + \lambda_3 h_1 h_2 \tag{23}$$

est connu.

Ce trinome est une fonction homogène du second degré des erreurs e_1, e_2, e_3, e_4, e_5, et, si l'on nomme m^2 la valeur probable du carré de l'une de ces erreurs, celle du produit de deux d'entre elles étant nulle, on trouvera pour valeur probable de l'expression (23), calculée avant les mesures prises,

$$m^2(3\lambda_1^2 + 3\lambda_2^2 + \lambda_3).$$

En égalant cette valeur probable à la valeur vraie, on aura

$$(24) \qquad m^2 = \frac{\lambda_1 h_1^2 + \lambda_2 h_2^2 + \lambda_3 h_1 h_2}{3\lambda_1^2 + 3\lambda_2 + \lambda_3};$$

λ_1, λ_2, λ_3 sont arbitraires.

Si l'on voulait choisir entre les valeurs en nombre infini représentées par la formule (24), il faudrait chercher la valeur probable du carré de la différence des deux membres et disposer de λ_1, λ_2, λ_3 de manière à la rendre minima.

Mais la formule (24), qui peut évidemment donner des valeurs de m^2 très inégales, reste, quels que soient les facteurs λ_1, λ_2, λ_3, la conséquence de l'égalité admise entre la valeur vraie d'une grandeur et la valeur probable.

220. Il est aisé de prouver par un exemple combien sont grandes les erreurs à craindre en égalant les valeurs vraies aux valeurs probables.

Un observateur mesure les trois angles d'un triangle; la probabilité d'une erreur comprise entre z et $z + dz$ est, pour lui, $\frac{k}{\sqrt{\pi}} e^{-k^2 z^2} dz$. Quelle est la probabilité d'une erreur y sur la somme des trois angles?

On peut dire :

La valeur probable du carré de l'erreur commise sur chaque angle est $\frac{1}{2k^2}$;

La valeur probable du carré de l'erreur commise sur la somme des trois angles est $\frac{3}{2k^2}$.

Si donc on représente la probabilité d'une erreur y sur la somme des trois angles par $\frac{k'}{\sqrt{\pi}} e^{-k'^2 y^2}$, il faudra supposer

$$\frac{1}{2k'^2} = \frac{3}{2k^2},$$

$$k' = \frac{k}{\sqrt{3}}.$$

La probabilité d'une erreur plus petite que α sur la somme des trois angles est donc

$$\frac{2k'}{\sqrt{\pi}}\int_0^\alpha e^{-k'^2z^2}\,dz = \Theta(k'\alpha).$$

La démonstration peut laisser un doute. En acceptant la loi représentée par $e^{-k^2z^2}$ pour les probabilités d'erreurs partielles sur chaque angle, est-il permis d'en conclure une expression de même forme pour l'erreur totale? L'assertion n'est pas évidente; on peut la démontrer.

Soient x, y, z les erreurs commises sur les trois angles d'un triangle; posons

$$x + y + z = \alpha,$$
$$x^2 + y^2 + z^2 = \rho^2.$$

La probabilité du concours des trois erreurs était, *a priori*,

$$\frac{k^3}{\pi\sqrt{\pi}}e^{-k^2\rho^2}\,dx\,dy\,dz,$$

elle est le produit de $\frac{k^3}{\pi\sqrt{\pi}}e^{-k^2\rho^2}$ par l'élément de volume $dx\,dy\,dz$, en considérant x, y, z comme des coordonnées rectangulaires.

La probabilité pour que, $x+y+z$ étant compris entre α et $\alpha+d\alpha$, ρ le soit entre ρ et $\rho+d\rho$ sera le produit de $\frac{k^3}{\pi\sqrt{\pi}}e^{-k^2\rho^2}$ par le volume compris entre les deux sphères qui correspondent aux rayons ρ et $\rho+d\rho$ et les plans dont les équations sont

$$x + y + z = \alpha,$$
$$x + y + z = \alpha + d\alpha.$$

Ce volume est

$$\frac{2\pi\rho\,d\rho\,d\alpha}{\sqrt{3}}. \tag{25}$$

La probabilité *a priori* pour que, la somme des erreurs étant comprise entre α et $\alpha+d\alpha$, celle de leurs carrés le soit entre ρ^2 et $(\rho+d\rho)^2$ est égale à

$$\frac{2k^3}{\sqrt{\pi}}\,\frac{e^{-k^2\rho^2}\rho\,d\rho\,d\alpha}{\sqrt{3}}. \tag{26}$$

La probabilité pour que la somme des trois erreurs soit comprise entre α et $\alpha + d\alpha$ est la somme des valeurs de (26) pour toutes les valeurs possibles de ρ qui sont comprises entre $\frac{\alpha}{\sqrt{3}}$ et ∞.

La probabilité pour que la somme des trois erreurs soit comprise entre α et $\alpha + d\alpha$ est donc

$$(27) \qquad \frac{2k^3}{\sqrt{\pi}} \frac{d\alpha}{\sqrt{3}} \int_{\frac{\alpha}{\sqrt{3}}}^{\infty} e^{-k^2\rho^2} \rho \, d\rho;$$

on a

$$\int_{\frac{\alpha}{\sqrt{3}}}^{\infty} e^{-k^2\rho^2} \rho \, d\rho = \frac{1}{2k^2} e^{-\frac{k^2\alpha^2}{3}}.$$

La formule (27) se réduit à

$$\frac{k}{\sqrt{\pi}} \frac{d\alpha}{\sqrt{3}} e^{-\frac{k^2\alpha^2}{3}},$$

précisément celle que nous avions admise.

221. Supposons, pour entrer au détail, qu'en mesurant les trois angles d'un triangle l'erreur probable sur chaque mesure soit égale à $2''$, la valeur correspondante de k est $\frac{1}{2\sqrt{\pi}}$.

La probabilité d'une erreur plus petite que $1''$ sur la somme des trois angles est, dans ce cas,

$$\Theta\left(\frac{1}{2\sqrt{3\pi}}\right) = 0,15.$$

L'événement n'a rien d'invraisemblable.

S'il arrive cependant qu'en mesurant les trois angles d'un triangle, on trouve une somme d'erreurs égale à $1''$, la théorie adoptée donnera, pour déterminer la valeur vraisemblable du carré de l'erreur sur chaque mesure, la relation acceptée

$$3m^2 = 1;$$

par conséquent,

$$\frac{1}{2k^2} = m^2 = \frac{1}{3},$$

$$k^2 = \frac{2}{3},$$

ce qui correspond à une erreur probable de chaque mesure

$$\frac{1}{k\sqrt{\pi}} = \frac{\sqrt{3}}{\sqrt{2\pi}} = 0'',7.$$

Lorsque, dans une seule épreuve, l'erreur commise sur la somme des angles d'un triangle est égale à $1''$, est-il possible de proposer avec confiance $0'',7$ comme la valeur vraisemblable de l'erreur probable sur chaque observation?

La valeur probable, pour l'observateur qui a mesuré les angles, est un nombre parfaitement déterminé; en la supposant égale à $2''$, c'est-à-dire triple de la solution proposée, la probabilité d'une erreur moindre que celle qui s'est produite serait $0,15$.

La valeur $0'',7$ adoptée pour l'erreur probable donnerait à l'événement observé, c'est-à-dire à une somme d'erreurs moindre que $1''$, une probabilité plus petite que $\frac{1}{2}$ égale à

$$\Theta\left(\frac{10}{7\sqrt{3\pi}}\right) = 0,447.$$

La probabilité pour qu'en adoptant $0'',7$ pour valeur probable de l'erreur on s'écarte peu de la vérité est, on le voit, fort éloignée de la certitude.

222. Pour expliquer plus clairement les principes théoriques de la méthode des moindres carrés, nous n'avons tenu aucun compte de la longueur des calculs. On peut, dans presque tous les cas, les abréger considérablement.

Les corrections prescrites par la méthode exposée (**210**) pour les grandeurs directement mesurées l_1, l_2, ..., l_{n+p} satisfont à une condition remarquable qui peut servir de base à une détermination plus rapide des inconnues.

La somme des carrés des corrections faites aux mesures est un minimum.

Reprenons pour le démontrer la question déjà résolue (**210**), en supposant toujours, pour éviter la longueur des formules, trois inconnues seulement, entre lesquelles on a six équations.

En nommant l_1, l_2, ..., l_6 les grandeurs mesurées et e_1, e_2, ..., e_6 les erreurs très petites commises dans leur évaluation, les conditions du problème donneront trois équations du premier degré entre les six erreurs

$$(28)\qquad \left\{\begin{aligned} P_1 e_1 + P_2 e_2 + P_3 e_3 + P_4 e_4 + P_5 e_5 + P_6 e_6 &= h_1,\\ Q_1 e_1 + Q_2 e_2 + Q_3 e_3 + Q_4 e_4 + Q_5 e_5 + Q_6 e_6 &= h_2,\\ R_1 e_1 + R_2 e_2 + R_3 e_3 + R_4 e_4 + R_5 e_5 + R_6 e_6 &= h_3. \end{aligned}\right.$$

Nous allons chercher à déterminer entre les solutions de ces équations le système pour lequel la somme

$$(29)\qquad e_1^2 + e_2^2 + e_3^2 + e_4^2 + e_5^2 + e_6^2$$

est minima. Nous constaterons l'identité des corrections ainsi déterminées avec celles qui résultent des conditions adoptées (210).

Pour rendre la somme (29) minima, il faut égaler sa différentielle à zéro en même temps que celle des expressions (28).

Nous aurons ainsi les équations

$$(30)\qquad e_1\,de_1 + e_2\,de_2 + e_3\,de_3 + e_4\,de_4 + e_5\,de_5 + e_6\,de_6 = 0;$$

$$(31)\qquad \left\{\begin{aligned} P_1\,de_1 + P_2\,de_2 + P_3\,de_3 + P_4\,de_4 + P_5\,de_5 + P_6\,de_6 &= 0,\\ Q_1\,de_1 + Q_2\,de_2 + Q_3\,de_3 + Q_4\,de_4 + Q_5\,de_5 + Q_6\,de_6 &= 0,\\ R_1\,de_1 + R_2\,de_2 + R_3\,de_3 + R_4\,de_4 + R_5\,de_5 + R_6\,de_6 &= 0. \end{aligned}\right.$$

On ajoutera, conformément à la méthode de Lagrange, ces quatre équations, après avoir multiplié trois d'entre elles par des facteurs μ_1, μ_2, μ_3, et l'on égalera à zéro les coefficients des six différentielles. Les équations

$$(32)\qquad \left\{\begin{aligned} e_1 + \mu_1 P_1 + \mu_2 Q_1 + \mu_3 R_1 &= 0,\\ e_2 + \mu_1 P_2 + \mu_2 Q_2 + \mu_3 R_2 &= 0,\\ e_3 + \mu_1 P_3 + \mu_2 Q_3 + \mu_3 R_3 &= 0,\\ \cdots\cdots\cdots\cdots&\cdots,\\ e_6 + \mu_1 P_6 + \mu_2 Q_6 + \mu_3 R_6 &= 0, \end{aligned}\right.$$

jointes aux trois équations (28), détermineront les corrections e_1, e_2, e_3, e_4, e_5, e_6 et les multiplicateurs μ_1, μ_2, μ_3.

En ajoutant les équations (32) après les avoir multipliées par les coefficients de chacune des inconnues μ_1, μ_2, μ_3, et en ayant égard aux relations (28), on formera les équations

$$(33)\qquad \left\{\begin{aligned} \mu_1(P_1^2) + \mu_2(P_1 Q_1) + \mu_3(P_1 R_1) &= h_1,\\ \mu_1(P_1 Q_1) + \mu_2(Q_1^2) + \mu_3(R_1 Q_1) &= h_2,\\ \mu_1(P_1 R_1) + \mu_2(Q_1 R_1) + \mu_3(R_1^2) &= h_3, \end{aligned}\right.$$

où (P_1^2), $(P_1 Q_1)$, ... représentent les sommes

$$\begin{array}{l} P_1^2 \quad + P_2^2 \quad + P_3^2, \\ P_1 Q_1 + P_2 Q_2 + P_3 Q_3, \\ \dots\dots\dots\dots\dots\dots \end{array}$$

Les équations (33) feront connaître μ_1, μ_2, μ_3. On déduira ensuite de (32) e_1, e_2, e_3, ...

Les corrections ainsi obtenues sont identiques à celles qui ont été déduites (**209**) du Calcul des probabilités. Reprenons, en effet, les équations (7) qui font connaître λ_1, λ_2, λ_3; elles ne diffèrent de (33) que par le changement des termes indépendants des inconnues.

Cherchons par les deux méthodes la correction désignée par e_1.

Les valeurs de λ_1, λ_2, λ_3 étant données par les équations

$$(34) \qquad \left\{ \begin{array}{l} \lambda_1(P_1^2) \quad + \lambda_2(P_1 Q_1) + \lambda_3(P_1 R_1) = - P_1, \\ \lambda_1(P_1 Q_1) + \lambda_2(Q_1^2) \quad + \lambda_3(Q_1 R_1) = - Q_1, \\ \lambda_1(R_1 P_1) + \lambda_2(R_1 Q_1) + \lambda_3(R_1^2) \quad = - R_1, \end{array} \right.$$

la correction prescrite est

$$e_1 = \lambda_1 h_1 + \lambda_2 h_2 + \lambda_3 h_3.$$

La condition du minimum donne

$$e_1 = - \mu_1 P_1 - \mu_2 Q_1 - \mu_3 R_1.$$

En ajoutant les équations (33) multipliées par λ_1, λ_2, λ_3 et les équations (9) multipliées par μ_1, μ_2, μ_3, on constate l'identité des deux sommes.

223. Le principe des moindres carrés étant admis, on peut s'en servir pour obtenir les meilleures valeurs des inconnues, sans s'astreindre au calcul préalable des erreurs commises sur les grandeurs directement déterminées.

Les équations imposées aux inconnues étant

$$(35) \qquad \left\{ \begin{array}{l} \varphi_1(x, y, z, \dots) = l_1, \\ \varphi_2(x, y, z, \dots) = l_2, \\ \dots\dots\dots\dots\dots\dots, \\ \varphi_{n+p}(x, y, z, \dots) = l_{n+p}, \end{array} \right.$$

et leur nombre $n+p$ surpassant le nombre n des inconnues, on devra d'abord déterminer des valeurs approchées X, Y, Z, ... de x, y, z, ..., pour lesquelles ils subsistera entre les deux membres des équations (35) des différences très petites $\alpha_1, \alpha_2, \ldots, \alpha_{n+p}$. Si l'on nomme $x_1, y_1, z_1, \ldots$ les corrections qu'il faut faire à X, Y, Z, ..., et $e_1, e_2, \ldots, e_{n+p}$ celles de $l_1, l_2, \ldots, l_{n+p}$, on aura, en négligeant les carrés de ces corrections, des équations nécessaires de la forme

$$(36)\quad \left\{\begin{array}{l} P_1 x_1 + P_2 y_1 + P_3 z_1 + \ldots + \alpha_1 = e_1, \\ Q_1 x_1 + Q_2 y_1 + Q_3 z_1 + \ldots + \alpha_2 = e_2, \\ R_1 x_1 + R_2 y_1 + R_3 z_1 + \ldots + \alpha_3 = e_3, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \end{array}\right.$$

Pour rendre minima la somme

$$e_1^2 + e_2^2 + e_3^2 + \ldots,$$

il faut, après l'avoir formée, égaler à zéro les dérivées par rapport à x_1, y_1, z_1, ..., et l'on aura ainsi, entre les inconnues véritables de la question,

$$\begin{array}{l} P_1(P_1 x_1) + Q_1(Q_1 x_1) + R_1(R_1 x_1) + \ldots + P_1\alpha_1 + Q_1\alpha_2 + \ldots = 0, \\ P_2(P_1 x_1) + Q_2(Q_1 x_1) + R_2(R_1 x_1) + \ldots + P_2\alpha_1 + Q_2\alpha_2 + \ldots = 0, \\ P_3(P_1 x_1) + Q_3(Q_1 x_1) + R_3(R_1 x_1) + \ldots + P_3\alpha_1 + Q_3\alpha_2 + \ldots = 0, \end{array}$$

$(P_1 x_1)$, $(Q_1 x_1)$, ... désignant l'ensemble des termes qui, dans les équations (36), précèdent α_1, α_2, ... On est conduit à la règle suivante :

On ajoutera les premiers membres des équations du premier degré, dont les seconds membres sont les corrections à faire subir aux grandeurs directement mesurées, après les avoir multipliées successivement par le coefficient de chaque inconnue.

Les équations ainsi formées, en nombre égal à celui des inconnues, donneront la solution du problème.

224. Supposons que cinq points A_1, A_2, A_3, A_4, A_5 aient été visés d'une même station O. Sur les dix angles que forment deux à deux les lignes de visée, on en a mesuré huit. En nommant x_1, x_2, x_3, x_4 les angles de la direction OA avec les quatre autres, ces angles déterminant tous les autres, les mesures choisies sont telles qu'en nommant l_1, l_2, l_3, l_4, l_5, l_6, l_7, l_8 leurs valeurs exactes, on a

$$(37)\quad \left\{\begin{array}{llll} x_3 = l_1, & x_4 = l_2, & x_4 - x_3 = l_3, & x_4 - x_2 = l_4, \\ x_2 = l_5, & x_3 - x_1 = l_6, & x_1 = l_7, & x_2 - x_1 = l_8. \end{array}\right.$$

Les relations entre les angles mesurés doivent être

$$(38)\quad \left\{\begin{array}{l} l_2 - l_1 - l_3 = 0, \\ l_2 - l_5 - l_4 = 0, \\ l_1 - l_7 - l_6 = 0, \\ l_5 - l_7 - l_8 = 0. \end{array}\right.$$

Ces expressions ne seront pas nulles, mais auront des valeurs très petites. Nommons ces quatre valeurs α_1, α_2, α_3, α_4; elles donneront, indépendamment de tout calcul, une première idée des erreurs commises.

Pour appliquer la méthode des moindres carrés, il est inutile de former les équations (5); nous ajouterons les équations (37) après avoir multiplié chacune par le coefficient de l'une des inconnues successivement choisies. Nous aurons ainsi

$$\begin{aligned} 3x_1 - x_2 - x_3 &= l_7 - l_6 - l_8, \\ 3x_2 - x_4 - x_1 &= l_8 + l_5 - l_4, \\ 3x_3 - x_4 - x_1 &= l_1 - l_3 + l_6, \\ 3x_4 - x_3 + x_2 &= l_1 + l_3 + l_4. \end{aligned}$$

On en déduira les valeurs des inconnues

$$\begin{aligned} x_1 &= \frac{1}{5}l_1 + \frac{2}{15}l_2 - \frac{1}{15}l_3 - \frac{1}{15}l_4 + \frac{1}{5}l_5 - \frac{4}{15}l_6 + \frac{7}{15}l_7 - \frac{4}{15}l_8, \\ x_2 &= \frac{2}{15}l_1 + \frac{1}{5}l_2 + \frac{1}{15}l_3 - \frac{1}{3}l_4 + \frac{7}{15}l_5 - \frac{1}{15}l_6 + \frac{1}{5}l_7 + \frac{4}{15}l_8, \\ x_3 &= \frac{7}{15}l_1 + \frac{1}{5}l_2 - \frac{4}{15}l_3 + \frac{1}{15}l_4 + \frac{2}{15}l_5 + \frac{4}{15}l_6 + \frac{1}{5}l_7 - \frac{1}{15}l_8, \\ x_4 &= \frac{1}{5}l_1 + \frac{7}{15}l_2 + \frac{4}{15}l_3 + \frac{4}{15}l_4 + \frac{1}{5}l_5 + \frac{1}{15}l_6 + \frac{2}{15}l_7 + \frac{1}{15}l_8. \end{aligned}$$

225. Lorsque les corrections seront calculées, on devra chercher, en tenant compte des réserves faites, la valeur probable du carré de l'erreur à craindre pour chaque inconnue.

Gauss a donné pour ce calcul, comme pour tous les détails de cette théorie, une méthode devenue classique de laquelle résulte une démonstration du principe des moindres carrés très différente de celle que nous avons adoptée. Nous la reproduisons textuellement.

Problème. — *Désignons par v, v', v'', ... les fonctions linéaires suivantes des indéterminées x, y, z, ... :*

$$(G_1)\quad \begin{cases} v = ax + by + cz + \ldots + l, \\ v' = a'x + b'y + c'z + \ldots + l', \\ v'' = a''x + b''y + c''z + \ldots + l'', \\ \ldots\ldots\ldots\ldots\ldots\ldots \end{cases}$$

Parmi tous les systèmes des coefficients $\varkappa$, $\varkappa'$, $\varkappa''$, ... qui donnent identiquement

$$\varkappa v + \varkappa' v' + \varkappa'' v'' + \ldots = x - k,$$

k étant indépendant de x, y, z, ..., trouver celui pour lequel $\varkappa^2 + \varkappa'^2 + \varkappa''^2 + \ldots$ est minimum.

Posons

$$(G_2)\quad \begin{cases} av + a'v' + a''v'' + \ldots = \xi, \\ bv + b'v' + b''v'' + \ldots = \eta, \\ cv + c'v' + c''v'' + \ldots = \zeta, \\ \ldots\ldots\ldots\ldots\ldots\ldots; \end{cases}$$

ξ, η, ζ seront des fonctions linéaires de x, y, z, et l'on aura

$$(G_3)\quad \begin{cases} \xi = x\Sigma a^2 + y\Sigma ab + z\Sigma ac + \ldots + \Sigma al, \\ \eta = x\Sigma ab + y\Sigma b^2 + z\Sigma bc + \ldots + \Sigma bl, \\ \zeta = x\Sigma ac + y\Sigma bc + z\Sigma c^2 + \ldots + \Sigma cl, \\ \ldots\ldots\ldots\ldots\ldots\ldots, \end{cases}$$

où

$$\Sigma a^2 = a^2 + a'^2 + a''^2 + \ldots,$$

et de même pour les autres Σ.

Le nombre des quantités ξ, η, ζ, ... est égal au nombre ϖ des inconnues x, y, z, ...; on pourra donc obtenir, par élimination, une équation de la forme suivante [1]

$$x = A + (\alpha\alpha)\xi + (\alpha\beta)\eta + (\alpha\gamma)\zeta + \ldots,$$

qui sera satisfaite identiquement lorsqu'on remplacera ξ, η, ζ par leurs valeurs (G_3). Par conséquent, si l'on pose

$$(G_4) \qquad \left\{ \begin{aligned} a\,(\alpha\alpha) + b\,(\alpha\beta) + c\,(\alpha\gamma) + \ldots &= \alpha, \\ a'(\alpha\alpha) + b'(\alpha\beta) + c'(\alpha\gamma) + \ldots &= \alpha', \\ a''(\alpha\alpha) + b''(\alpha\beta) + c''(\alpha\gamma) + \ldots &= \alpha'', \end{aligned} \right.$$

on aura identiquement

$$(G_5) \qquad \alpha v + \alpha' v' + \alpha'' v'' + \ldots = x - A.$$

Cette équation montre que, parmi les différents systèmes de coefficients $\varkappa$, $\varkappa'$, $\varkappa''$, ..., on doit compter le système

$$\varkappa = \alpha, \qquad \varkappa' = \alpha', \qquad \varkappa'' = \alpha'', \qquad \ldots.$$

On aura d'ailleurs, pour un système quelconque,

$$(\varkappa - \alpha)v + (\varkappa' - \alpha')v' + (\varkappa'' - \alpha'')v'' + \ldots = A - k,$$

et cette équation, étant identique, entraîne les suivantes :

$$\begin{aligned} (\varkappa - \alpha)a + (\varkappa' - \alpha')a' + (\varkappa'' - \alpha'')a'' + \ldots &= 0, \\ (\varkappa - \alpha)b + (\varkappa' - \alpha')b' + (\varkappa'' - \alpha'')b'' + \ldots &= 0, \\ (\varkappa - \alpha)c + (\varkappa' - \alpha')c' + (\varkappa'' - \alpha'')c'' + \ldots &= 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots& \end{aligned}$$

Ajoutons ces équations après les avoir multipliées, respectivement, par $(\alpha\alpha)$, $(\alpha\beta)$, $(\alpha\gamma)$, ..., nous aurons, en vertu du système (G_4),

$$(\varkappa - \alpha)\alpha + (\varkappa' - \alpha')\alpha' + (\varkappa'' - \alpha'')\alpha'' + \ldots = 0,$$

c'est-à-dire

$$\varkappa^2 + \varkappa'^2 + \ldots = \alpha^2 + \alpha'^2 + \ldots + (\varkappa - \alpha)^2 + (\varkappa' - \alpha')^2 + \ldots;$$

---

[1] On verra plus loin la raison qui a conduit à désigner les coefficients de cette formule par la notation $(\alpha\alpha)$, $(\alpha\beta)$.

par conséquent, la somme

$$\varkappa^2 + \varkappa'^2 + \varkappa''^2 + \ldots$$

aura une valeur minimum, lorsque l'on aura

$$\varkappa = \alpha, \qquad \varkappa' = \alpha', \qquad \varkappa'' = \alpha'', \qquad \ldots.$$

D'ailleurs, cette valeur minimum s'obtiendra de la manière suivante.

L'équation (G_5) montre que l'on a

$$\begin{aligned} a\alpha + a'\alpha' + a''\alpha'' + \ldots &= 0, \\ b\alpha + b'\alpha' + b''\alpha'' + \ldots &= 0, \\ c\alpha + c'\alpha' + c''\alpha'' + \ldots &= 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \end{aligned}$$

Multiplions ces équations, respectivement, par $(\alpha\alpha)$, $(\alpha\beta)$, $(\alpha\gamma)$, ..., et ajoutons; en ayant égard aux relations (G_4), on trouvera

$$\alpha^2 + \alpha'^2 + \alpha''^2 + \ldots = (\alpha\alpha).$$

Lorsque les observations auront donné des équations approximatives

$$v = 0, \qquad v' = 0, \qquad v'' = 0, \qquad \ldots,$$

il faudra, pour déterminer l'inconnue x, choisir une combinaison de la forme suivante

$$\varkappa v + \varkappa' v' + \varkappa'' v'' + \ldots = 0,$$

telle que l'inconnue x acquière un coefficient égal à 1 et que les autres inconnues se trouvent éliminées.

Le poids de cette détermination sera

$$\frac{1}{\varkappa^2 + \varkappa'^2 + \varkappa''^2 + \ldots}.$$

On obtiendra la détermination la plus convenable en prenant

$$\varkappa = \alpha, \qquad \varkappa' = \alpha', \qquad \varkappa'' = \alpha'', \qquad \ldots;$$

alors x aura la valeur A. On obtiendrait évidemment la même valeur sans connaître les multiplicateurs α, α', α'', ..., en effectuant l'élimination sur les équations

$$\xi = 0, \qquad \eta = 0, \qquad \zeta = 0, \qquad \ldots;$$

le poids de cette détermination sera

$$\frac{1}{(\alpha\alpha)},$$

et l'erreur moyenne à craindre

$$m\sqrt{p(\alpha\alpha)} = m'\sqrt{p'(\alpha\alpha)} = m''\sqrt{p''(\alpha\alpha)} = \ldots.$$

Une marche analogue conduirait aux valeurs les plus convenables des autres inconnues y, z, ..., qui seront celles que l'on obtiendrait en effectuant l'élimination sur les équations

$$\xi = 0, \qquad \eta = 0, \qquad \zeta = 0, \qquad \ldots.$$

Si nous désignons par Ω la somme

$$v^2 + v'^2 + v''^2 + \ldots$$

ou, ce qui revient au même,

$$p(V - L)^2 + p'(V' - L')^2 + p''(V'' - L'')^2 + \ldots,$$

on aura évidemment

$$2\xi = \frac{d\Omega}{dx}, \qquad 2\eta = \frac{d\Omega}{dy}, \qquad 2\zeta = \frac{d\Omega}{dz}, \qquad \ldots;$$

par conséquent, les valeurs des inconnues, déduites de la combinaison la plus convenable, et que nous pouvons appeler les valeurs *les plus plausibles,* sont précisément celles qui donnent à Ω une valeur minimum. Or $V - L$ représente la différence entre la valeur observée et la valeur calculée; donc les valeurs les plus plausibles des inconnues sont celles qui rendent minimum la somme des carrés des différences entre les valeurs calculées et observées des quantités V, V', V'', ..., ces carrés étant respectivement multipliés par le poids des observations. Ces principes ont été depuis longtemps établis par d'autres considérations (*Theoria motus corporum cœlestium*).

Si l'on veut assigner la précision relative de chacune des déterminations, il faut déduire des équations (G_3) les valeurs de x, y, z, ..., qui se présenteront sous la forme suivante :

$$(G_7)\quad \begin{cases} x = A + (\alpha\alpha)\xi + (\alpha\beta)\eta + (\alpha\gamma)\zeta + \ldots, \\ y = B + (\beta\alpha)\xi + (\beta\beta)\eta + (\beta\gamma)\zeta + \ldots, \\ z = C + (\gamma\alpha)\xi + (\gamma\beta)\eta + (\gamma\gamma)\zeta + \ldots. \end{cases}$$

Les valeurs les plus plausibles des inconnues x, y, z, ... seront A, B, C, ... Les poids de ces déterminations seront

$$\frac{1}{(\alpha\alpha)},\quad \frac{1}{(\beta\beta)},\quad \frac{1}{(\gamma\gamma)},\quad \ldots,$$

et les erreurs moyennes à craindre

$$\begin{array}{ll} \text{Pour } x \ldots\ldots\ldots\ldots\ldots & m\sqrt{p}(\alpha\alpha) = m'\sqrt{p'}(\alpha\alpha) = \ldots, \\ \text{Pour } y \ldots\ldots\ldots\ldots\ldots & m\sqrt{p}(\beta\beta) = m'\sqrt{p'}(\beta\beta) = \ldots, \\ \text{Pour } z \ldots\ldots\ldots\ldots\ldots & m\sqrt{p}(\gamma\gamma) = m'\sqrt{p'}(\gamma\gamma) = \ldots, \end{array}$$

ce qui s'accorde avec les résultats obtenus antérieurement (*Theoria motus corporum cœlestium*).

Le cas où il n'y a qu'une seule inconnue est le plus fréquent et le plus simple de tous. On a alors

$$V = x,\qquad V' = x,\qquad V'' = x,\qquad \ldots;$$

il sera utile d'en dire quelques mots.

On aura

$$\begin{array}{llll} a = \sqrt{p}, & a' = \sqrt{p'}, & a'' = \sqrt{p''}, & \ldots, \\ l = -L\sqrt{p}, & l' = -L'\sqrt{p'}, & l'' = -L''\sqrt{p''}, & \ldots \end{array}$$

et, par conséquent,

$$\xi = (p + p' + p'' + \ldots)x - (pL + p'L' + p''L'' + \ldots);$$

d'où

$$(\alpha\alpha) = \frac{1}{p + p' + p'' + \ldots},$$
$$(A) = \frac{pL + p'L' + p''L'' + \ldots}{p + p' + p'' + \ldots}.$$

Ainsi, si, par plusieurs observations qui n'ont pas la même précision et dont les poids respectifs sont p, p', p'', ..., on a trouvé, pour une même quantité, une première valeur L′, une deuxième L′, une troisième L″, ..., la valeur la plus plausible sera

$$\frac{p\,\mathrm{L} + p'\,\mathrm{L}' + p''\,\mathrm{L}'' + \ldots}{p + p' + p'' + \ldots},$$

et le poids de cette détermination sera

$$p + p' + p'' + \ldots.$$

Si toutes les observations sont également plausibles, la valeur la plus probable sera

$$\frac{\mathrm{L} + \mathrm{L}' + \mathrm{L}'' + \ldots}{\varpi},$$

c'est-à-dire la moyenne arithmétique entre les valeurs observées ; en prenant pour unité le poids d'une observation isolée, le poids de la moyenne sera ϖ.

226. Quoique les formules proposées pour exprimer la valeur probable du carré de l'erreur à craindre, à quelque point de vue qu'on se place pour les obtenir, méritent peu de confiance, il n'est pas inutile de les défendre contre un reproche injustement adressé.

Bienaymé, auteur de l'objection, a proposé « une modification profonde » ; il parle de la « défectuosité du calcul ordinaire ». Le défaut qu'il signale lui paraît si simple, « qu'aux premiers mots tout le monde en reconnaîtra l'existence ». — « L'erreur consiste, dit-il, à calculer la probabilité d'une erreur commise comme si elle était la seule. Un des premiers principes de la théorie des probabilités est que, quand plusieurs événements arrivent simultanément, la probabilité de leur concours est le produit des probabilités de chacun, de sorte que la probabilité de ce concours est inférieure à la probabilité de chaque événement pris à part ; elle est d'autant plus petite qu'il y a plus d'événements. »

« Évidemment, ajoute Bienaymé, il en est de même des erreurs de plusieurs inconnues. La probabilité que ces erreurs resteront toutes à la fois dans certaines limites ne peut être que le produit des probabilités séparées pour que chacune ne s'écarte pas de ses limites propres et, par conséquent, cette probabilité du concours des erreurs de grandeur limitée doit être notablement inférieure à la probabilité des limites de chaque erreur considérée isolément, quelles que puissent être les autres. »

L'assertion est évidente; mais le tort est d'accuser les auteurs de la théorie et des applications qu'on en a faites de l'avoir ignorée ou oubliée.

Quand on a calculé une inconnue, il importe de savoir quelle confiance mérite le résultat. Les formules de Gauss répondent plus ou moins rigoureusement à cette question. Si une seconde inconnue est calculée, le même problème sera résolu pour elle.

Si l'on connaît les probabilités pour que les erreurs commises sur deux angles soient plus petites que $0'',10$, on pourra, les deux résultats n'étant pas contestés, chercher la probabilité pour que les deux épreuves soient toutes deux plus petites que $0'',10$; l'intérêt de cet autre problème sera plus ou moins grand, mais c'est une étrange prétention d'accuser d'*erreur* ceux qui n'ont pas désiré le résoudre.

Prenons un exemple.

On veut connaître un angle A. Cet angle fait partie d'un triangle ABC. On mesure les trois angles A, B, C et l'on prend pour A la valeur

$$A + \frac{1}{3}(180^\circ - A - B - C). \tag{38}$$

Si m^2 est la valeur probable du carré de l'erreur à craindre dans la mesure d'un angle, le carré de l'erreur à craindre en adoptant l'expression (38) est (199) $\frac{2}{3}m^2$.

L'objection consiste à dire : Sur les deux angles B et C vous avez des erreurs à craindre; elles doivent entrer en compte; elles ont leur part nécessaire dans l'évaluation du mérite de la solution. Cela est vrai si le problème est de résoudre le triangle; mais, si le calcul est entrepris pour déterminer l'angle A, on n'aura nul souci des deux autres.

Le triangle a trois côtés; on peut y inscrire un cercle, ou le circonscrire, déterminer la surface, calculer les bissectrices, etc., et résoudre cent problèmes différents pour chacun desquels, puisque le triangle est imparfaitement connu, une erreur sera à craindre. Cherchera-t-on la probabilité pour que toutes ces erreurs soient inférieures à des limites données? Si persuadé qu'on soit qu'il faut le faire, le nombre des grandeurs qui dépendent du triangle étant infini, il faudra s'arrêter; où commencera la *faute* commise?

Un exemple réduira l'objection à sa véritable valeur.

On construit une carte géographique. Les villes, villages et bourgades y sont inscrits par milliers. On étudie l'un des points principaux et l'on cherche l'erreur probable à craindre sur chacune de ses coordonnées géographiques. Les calculs sont irréprochables; l'auteur de l'objection, sans y contredire, signale une faute très grave. Votre carte, dit-il, contient mille déterminations; il *fallait,* en vertu d'*un principe dont la vérité frappera tout le monde,* faire pour les mille points marqués le calcul exécuté pour un seul et multiplier les mille probabilités.

Les erreurs n'étant pas indépendantes, la solution aurait le mérite d'une difficulté vaincue, mais elle condamnerait la carte la plus admirée d'autant plus sévèrement qu'elle serait plus riche de détails. Comment espérer que le produit de mille probabilités ne soit pas très petit?

Le produit étant supposé connu, on accueillerait certainement comme un grand progrès la recherche isolée de chaque facteur. C'est elle seulement qui peut intéresser.

227. Nous terminerons ce Chapitre par la démonstration d'un élégant théorème de Gauss, annoncé (**214**) et dont les conséquences relatives à la détermination de la précision d'un système d'observations ne me paraissent pas acceptables.

On a mesuré directement $n+p$ grandeurs. Les mesures inspirent la même confiance; mais la valeur probable m^2 du carré de l'erreur commise sur l'une d'elles est *a priori* complètement inconnue. Ces $n+p$ grandeurs mesurées sont liées par des équations, que l'énoncé du problème fait connaître, à n grandeurs inconnues. La méthode des moindres carrés détermine ces inconnues par la condition que les corrections sur les grandeurs directement mesurées, qui rendent les équations compatibles, aient une somme de carrés minima.

Les calculs font connaître exactement cette somme de carrés, plus petite par l'énoncé même de la condition imposée, que la somme des carrés des erreurs réellement commises.

Le théorème de Gauss se compose de deux parties :

La somme des carrés des corrections prescrites par la méthode des moindres carrés est une fonction homogène du second degré, parfaitement déterminée, des erreurs réellement commises.

En désignant par m^2 la valeur probable, *a priori*, du carré de l'erreur commise sur une observation, la valeur probable de la fonction qui représente la somme des carrés des corrections prescrites par la méthode, et par conséquent la valeur probable de cette somme de carrés dont la valeur numérique est connue, est égale à pm^2.

Nous allons démontrer ces deux théorèmes, sans pouvoir accepter qu'il soit permis ensuite d'égaler la valeur vraie de la somme des carrés des corrections à la valeur probable et d'en conclure pour valeur probable du carré de l'une des erreurs d'observations

$$\frac{e_1^2 + e_2^2 + \ldots + e_{n+p}^2}{p},$$

$e_1, e_2, \ldots, e_{n+p}$ étant les corrections faites aux grandeurs mesurées et le dénominateur p étant l'excès du nombre $n+p$ des mesures prises sur le nombre n des inconnues qu'on en a déduites.

Non seulement la somme des carrés des corrections, mais chaque correction en particulier peut s'exprimer en fonction des erreurs réellement commises. Si ces erreurs, en effet, sont connues, les équations auxquelles les inconnues satisfont, et qui deviennent incompatibles, sont parfaitement déterminées ; elles s'accorderaient si l'on ajoutait à chaque grandeur l'erreur commise en la mesurant, mais elles peuvent s'accorder d'une infinité de manières. Les corrections véritables étant inconnues, on rend la somme des carrés minima. Le calcul à faire pour cela est parfaitement déterminé, et le résultat ne peut contenir, outre les données de la question et les grandeurs mesurées, que les erreurs réellement commises.

La méthode exposée (207) permet de calculer cette fonction, dont l'existence est évidente *a priori*.

Soient

$$(39)\qquad \begin{cases} P_1 e_1 + P_2 e_2 + \ldots + P_{n+p} e_{n+p} = \alpha_1, \\ Q_1 e_1 + Q_2 e_2 + \ldots + Q_{n+p} e_{n+p} = \alpha_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots, \\ T_1 e_1 + T_2 e_2 + \ldots + T_{n+p} e_{n+p} = \alpha_p \end{cases}$$

les équations qui lient les erreurs possibles $e_1, e_2, \ldots, e_{n+p}$ aux évaluations des grandeurs $l_1, l_2, \ldots, l_{n+p}$, dont $\alpha_1, \alpha_2, \ldots, \alpha_p$ sont des fonctions numériquement connues et très petites, puisqu'elles seraient rigoureusement nulles si les mesures étaient exactes. Les corrections $e_1, e_2, \ldots, e_{n+p}$, prescrites par la méthode des moindres carrés, sont des fonctions de $\alpha_1, \alpha_2, \ldots, \alpha_p$ que la méthode fait connaître. La somme de leurs carrés est une fonction homogène du second degré de $\alpha_1, \alpha_2, \ldots, \alpha_p$. Mais les équations (39) sont satisfaites par tous les systèmes de corrections compatibles avec les données du problème; elles le sont donc par les corrections égales aux erreurs véritablement commises, et, en remplaçant $\alpha_1, \alpha_2, \ldots, \alpha_p$ par leurs valeurs données par (39) en fonction des erreurs véritables $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_{n+p}$, on aura exprimé la somme des carrés des corrections en fonction de ces erreurs véritablement commises qui restent inconnues.

Pour obtenir les corrections prescrites par la méthode des moindres carrés, il faut rendre minima la somme

$$e_1^2 + e_2^2 + \ldots + e_{n+p}^2;$$

adjoindre, par conséquent, à l'équation

$$e_1 de_1 + e_2 de_2 + \ldots + e_{n+p} de_{n+p} = 0$$

les p équations obtenues en différentiant le système (39).

On devra, conformément à la théorie des maxima et minima, ajouter ces équations après les avoir multipliées par des facteurs indéterminés, et écrire le système

$$(40)\qquad \begin{cases} e_1 + \mu_1 P_1 + \mu_2 Q_1 + \ldots + \mu_p T_1 = 0, \\ e_2 + \mu_1 P_2 + \mu_2 Q_2 + \ldots + \mu_p T_2 = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots, \\ e_{n+p} + \mu_1 P_{n+p} + \mu_2 Q_{n+p} + \ldots + \mu_p T_{n+p} = 0; \end{cases}$$

les facteurs $\mu_1, \mu_2, \ldots, \mu_p$ et les inconnues $e_1, e_2, \ldots, e_{n+p}$ se déduiront de ces $n+p$ équations adjointes aux p équations (39).

Pour obtenir les facteurs $\mu_1, \mu_2, \ldots, \mu_p$, ajoutons les équations (40), après les avoir multipliées par $P_1, P_2, \ldots$, puis par $Q_1, Q_2, \ldots$, et ainsi de suite, nous aurons

$$
(41) \qquad \left\{
\begin{array}{l}
\alpha_1 + \mu_1(P^2) \;\; + \mu_2(PQ) + \ldots + \mu_p(P_1 T) = 0, \\
\alpha_2 + \mu_1(Q_1 P) + \mu_2(Q^2) \;\; + \ldots + \mu_p(Q_1 T) = 0, \\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots, \\
\alpha_p + \mu_1(TP) \;\; + \mu_2(TQ) + \ldots + \mu_p(T^2) \;\; = 0,
\end{array}
\right.
$$

(P^2), (PQ), ... représentant la somme des carrés des valeurs de P_1^2, de $P_1 Q_1$, ...

Ajoutons les équations (40), après les avoir multipliées par e_1, $e_2, \ldots, e_{n+p}$, nous aurons

$$
\Sigma e_i^2 + \mu_1 \alpha_1 + \mu_2 \alpha_2 + \ldots + \mu_p \alpha_p = 0,
$$

et cette équation fera connaître Σe_i^2, exprimé, comme nous l'avions annoncé, en fonction de $\alpha_1, \alpha_2, \ldots, a_n$ qui sont eux-mêmes des fonctions connues d'un système quelconque de corrections possibles et par conséquent, en particulier, des erreurs réellement commises.

228. Il nous reste à chercher la valeur probable de l'expression

$$
(42) \qquad \mu_1 \alpha_1 + \mu_2 \alpha_2 + \ldots + \mu_p \alpha_p.
$$

Si $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_{n+p}$ sont les erreurs réellement commises, la valeur probable de e_i^2 est, quel que soit i, la quantité inconnue m^2 qui représente la précision des observations, et, les mesures étant indépendantes, la valeur probable de $\varepsilon_i, \varepsilon_{i'}$, quels que soient i et i', est égale à zéro.

La valeur probable de (42) est donc le produit par m^2 de la somme des coefficients des carrés de $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_{n+p}$.

En substituant aux α dans la somme (42) leurs valeurs en fonction des ε, données par le système (39), l'expression devient, en ayant égard aux équations (40),

$$
(43) \qquad \varepsilon_1 e_1 + \varepsilon_2 e_2 + \ldots + \varepsilon_n e_n;
$$

et, comme les corrections e sont des fonctions linéaires des μ et des α, par conséquent des ε, on voit que la somme des coefficients des carrés $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_{n+p}$ dans (42) est égale à

$$\frac{de_1}{d\varepsilon_1} + \frac{de_2}{d\varepsilon_2} + \ldots + \frac{de_{n+p}}{d\varepsilon_{n+p}},$$

c'est-à-dire, d'après le système (39), égale, au signe près, à

$$\begin{aligned}
&P_1 \frac{d\mu_1}{d\varepsilon_1} + Q_1 \frac{d\mu_2}{d\varepsilon_1} + \ldots + T_1 \frac{d\mu_p}{d\varepsilon_1} \\
&+ P_2 \frac{d\mu_1}{d\varepsilon_2} + Q_2 \frac{d\mu_2}{d\varepsilon_2} + \ldots + T_2 \frac{d\mu_p}{d\varepsilon_2} \\
&+ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
&+ P_{n+p} \frac{d\mu_1}{d\varepsilon_{n+p}} + Q_{n+p} \frac{d\mu_2}{d\varepsilon_{n+p}} + \ldots + T_{n+p} \frac{d\mu_p}{d\varepsilon_{n+p}}.
\end{aligned}$$

On a, en remarquant que les ε satisfont au système (39) et que les μ sont des fonctions linéaires des α,

$$\frac{d\mu_1}{d\varepsilon_1} = \frac{d\mu_1}{d\alpha_1} P_1 + \frac{d\mu_1}{d\alpha_2} Q_1 + \ldots + \frac{d\mu_1}{d\alpha_p} T_1;$$

on en déduit les équations

$$\begin{aligned}
P_1 \frac{d\mu_1}{d\varepsilon_1} &= \frac{d\mu_1}{d\alpha_1} P_1^2 + \frac{d\mu_1}{d\alpha_2} P_1 Q_1 + \ldots + \frac{d\mu_1}{d\alpha_p} P_1 T_1, \\
P_2 \frac{d\mu_1}{d\varepsilon_2} &= \frac{d\mu_1}{d\alpha_1} P_2^2 + \frac{d\mu_1}{d\alpha_2} P_2 Q_2 + \ldots + \frac{d\mu_1}{d\alpha_p} P_2 T_2, \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots, \\
P_{n+p} \frac{d\mu_1}{d\varepsilon_{n+p}} &= \frac{d\mu_1}{d\alpha_1} P_{n+p}^2 + \frac{d\mu_1}{d\alpha_2} P_{n+p} Q_{n+p} + \ldots + \frac{d\mu_1}{d\alpha_p} P_{n+p} T_{n+p}.
\end{aligned}$$

La somme des premiers membres est donc égale à

$$\left(\frac{d\mu_1}{d\alpha_1}\right)(P^2) + \frac{d\mu_1}{d\alpha_2}(PQ) + \ldots + \left(\frac{d\mu_1}{d\alpha_p}\right)(PT), \tag{44}$$

c'est-à-dire à l'unité; car cette expression (44) est précisément le résultat de la substitution dans la valeur de μ, déduite de (40), des coefficients de μ_1 aux termes tout connus $\alpha_1, \alpha_2, \ldots, \alpha_p$.

La première colonne du Tableau est égale à l'unité.

Il en est de même de toutes les autres, et la somme est égale à p.

Cette élégante démonstration a été donnée par M. Guyou.

Lors donc que l'on résout un problème par la méthode des moindres carrés, $n+p$ mesures ayant été prises et le carré de l'erreur à craindre sur chacune d'elles étant m^2, la valeur probable de la somme des carrés des erreurs réellement commises est

$$(n+p)m^2;$$

mais la valeur probable, nécessairement plus petite, de la somme des carrés des corrections indiquées par la méthode comme les plus plausibles est seulement pm^2.

229. Reprenons, pour donner un exemple, le problème résolu (214). A_1, A_2, A_3, A_4 étant quatre points observés du point O, et les angles sous lesquels A_1A_2, A_1A_3, A_1A_4, A_4A_3 et A_4A_2 sont vus du point O étant l_1, l_2, l_3, l_4, l_5, en posant

$$l_4+l_1-l_3=\alpha_1,$$
$$l_5+l_2-l_3=\alpha_2,$$

α_1 et α_2 étant très petits, les corrections à faire aux angles mesurés sont

$$-\frac{3}{8}\alpha_1+\frac{1}{8}\alpha_2,$$
$$\frac{1}{8}\alpha_1-\frac{3}{8}\alpha_2,$$
$$\frac{1}{4}\alpha_1+\frac{1}{4}\alpha_2,$$
$$\frac{1}{8}\alpha_1-\frac{3}{8}\alpha_2.$$

La somme de leurs carrés est

$$(45)\qquad \frac{1}{8}(3\alpha_1^2+3\alpha_2^2-2\alpha_1\alpha_2).$$

On a, en nommant ε_1, ε_2, ε_3, ε_4, ε_5 les erreurs réellement commises,

$$\alpha_1=\varepsilon_4+\varepsilon_1-\varepsilon_3,$$
$$\alpha_2=\varepsilon_5+\varepsilon_2-\varepsilon_3.$$

La valeur probable de α_1^2 est $3m^2$; celle de α_2^2, $3m^2$ et celle de

$\alpha_1 \alpha_2$, m^2; la valeur probable de (45) est donc

$$\frac{1}{8}(9m^2 + 9m^2 - 2m^2) = 2m^2,$$

c'est-à-dire, conformément au théorème de Gauss, le produit de m^2 par l'excès $5 - 3$ du nombre des angles mesurés sur celui des angles inconnus réellement distincts.

En égalant cette somme $2m^2$ à la somme des carrés des corrections prescrites par les formules pour les huit angles directement mesurés, la valeur ainsi obtenue pour m^2 ne peut nullement être acceptée pour mesure certaine ou vraisemblable de l'erreur à craindre dans les observations.

On peut affirmer seulement, et c'est là le point important, que, si la somme des carrés des corrections est petite, la probabilité est grande pour que les observations aient été bien faites.

230. Indépendamment de l'incertitude du principe sur lequel la démonstration repose, je veux dire le droit d'égaler la somme des carrés des corrections à leur valeur probable, une autre cause, non moins grave que la première, suffirait, dans le plus grand nombre des cas, pour enlever toute confiance dans l'évaluation précise des chances d'erreur proposées après chaque application de la méthode.

On suppose, *a priori,* toutes les mesures également précises; il est impossible, dans la plupart des cas, de croire à cette égalité : c'est faute de connaître aucune raison de préférence qu'on accepte l'équivalence des résultats. Mais, connues ou inconnues, ces raisons, si elles existent, doivent exercer une influence sur l'erreur *réellement* commise, et c'est celle-là dont on prétend donner les chances.

Il ne faudrait pas dire : On obtient une précision moyenne. Des mesures dont la précision est inégale ne donneront nullement le même résultat qu'un nombre égal de mesures prises avec une précision uniforme de quelque manière qu'on la choisisse.

Les formules, enfin, supposent pour toutes les observations les erreurs constantes absolument écartées; c'est une condition difficilement remplie quand on combine des observations d'origine différente.

Le calcul de la précision d'un système d'observations et l'évaluation qu'on en déduit pour la confiance méritée par le résultat ont compromis plus d'une fois la méthode des moindres carrés.

Après avoir discuté par d'immenses calculs les observations du passage de Vénus sur le Soleil en 1761, Encke a trouvé pour la parallaxe du Soleil $8'',49$ et pour erreur probable $0'',06$. Il y avait, en conséquence, plus de 300000 à parier contre 1 que l'erreur n'atteindrait pas $0'',42$, représentant sept fois l'erreur probable.

Les astronomes, cependant, acceptent aujourd'hui pour parallaxe $8'',91$, qui correspond précisément à l'erreur $0'',42$.

Sur un nombre total de 149 observations, Encke, par des raisons dont il serait difficile de donner le détail, en avait considéré 90 comme meilleures que les 59 autres. Les premières avaient même poids dans les calculs et les secondes un poids moitié moindre. Il n'en faut pas davantage, indépendamment de toute objection théorique, pour expliquer, j'oserai dire pour prévoir, des erreurs plus grandes encore que celles qu'on a commises.

Si le résultat final est exact, l'un des observateurs, Short, s'est trompé de $25''$ sur l'instant du second contact et Justander de $40''$ sur celui du premier.

L'un et l'autre, cependant, sont admis dans la première classe, comme Lacaille, qui se serait trompé de $2''$ seulement, et Lalande de $1'',7$.

L'erreur probable sur l'instant du premier contact, pour tous les observateurs de première classe, étant $7''$, il y aurait, si l'on s'en rapporte aux formules, 19000 à parier contre 1 qu'une erreur de $40''$ ne sera pas commise; n'était-ce pas une raison suffisante pour faire passer Justander dans la seconde classe, peut-être même pour supprimer ses chiffres, en voyant qu'à la sortie il s'est trompé de $20''$?

Encke, en prenant ce parti, aurait manqué, je le sais, à un principe que je n'accepte pas : les observations sont des témoins; si elles sont, avant l'épreuve, jugées dignes de confiance, leur déclaration, quelle qu'elle soit, doit être recueillie et conservée.

Laplace a évalué la masse de Jupiter $\frac{1}{1060}$ de celle du Soleil.

L'erreur commise, affirmait-il, est plus petite que $\frac{1}{50}$ du nombre proposé, et le Calcul des probabilités démontre qu'il y a 100000 contre 1 à parier pour qu'elle n'atteigne pas cette limite. La limite cependant a été dépassée : aucun astronome n'en doute aujourd'hui.

Il serait intéressant de refaire et de discuter de tels calculs. Je veux parler ici des principes seulement. Lorsque des inconnues sont déterminées par un grand nombre de mesures, les équations étant plus nombreuses qu'il ne faut, le calcul fait connaître les corrections pour lesquelles la somme des carrés est minima.

Quelle est la confiance méritée par les résultats?

Nous avons résolu deux problèmes différents :

Les observations n'étant pas faites encore, ou, ce qui revient au même, leur résultat étant encore inconnu, mais leur précision étant appréciée d'après l'habileté de l'observateur, quelle est la précision du résultat? La solution est irréprochable, mais sans intérêt dans la plupart des cas. Lorsque les observateurs sont différents et les observations nombreuses, il est impossible, évidemment, d'exprimer *a priori* par un nombre la confiance méritée par chacun, en écartant les circonstances particulières qui ont pu le troubler, comme, par exemple, dans les observations du passage de Vénus, ce phénomène imprévu de la goutte qui rendait les contacts incertains.

C'est après les avoir obtenus qu'il faut juger les résultats, et le véritable problème est celui-ci : Les observations sont faites, les calculs terminés, la somme des carrés des corrections est connue en chiffres; en déduire la précision supposée égale des observations combinées.

Nous devons répéter ce qui a été dit (**174**) :

Quand on entreprend une série de mesures, l'habileté des observateurs n'est ni parfaitement connue ni complètement inconnue. Ce sont des cas extrêmes. Il arrivera presque toujours que, toutes les valeurs de la précision étant possibles, elles seront, *a priori*, inégalement vraisemblables. La loi de leurs probabilités avant l'épreuve étant inconnue, le problème est insoluble.

Si les observations étaient *mal* faites, les équations seraient discordantes. La probabilité pour que le hasard, et non la perfection des mesures, les rende compatibles après de petites corrections, peut être considérée comme une impossibilité. On peut, en conséquence, quand la somme des carrés des erreurs est petite, accepter sans crainte le résultat, mais il est téméraire d'évaluer en chiffres la confiance qu'il doit inspirer.

CHAPITRE XII.

LES LOIS DE LA STATISTIQUE.

Were this calculus founded on the experience of a very great number of years, it would very well be worth the while to think of methods to facilite the computation of two, three or more lives.

HALLEY.

231. Il existe plus d'une manière de consulter le sort; quand la probabilité est la même, la moyenne est la même sur un grand nombre d'épreuves, mais les chances d'écart peuvent être différentes. — 232 Expression algébrique du problème à résoudre. — 233. Laplace et Poisson, dans leurs études sur la statistique des naissances, ont négligé cette remarque. — 234. Le tirage dans plusieurs urnes donne, pour une même probabilité moyenne, une valeur plus petite au carré de l'écart. — 235. Influence de l'importance des sommes assurées sur les chances d'écart de la moyenne. La formule obtenue en supposant les tirages faits dans la même urne n'est pas acceptable. — 236. Loi de mortalité de Gompertz.

231. Les géomètres ont tacitement assimilé les événements fortuits à une série de tirages au sort faits dans une urne de composition inconnue. Rien n'autorise, *a priori,* une telle hypothèse. Toutes les manières de consulter le hasard ne sont pas équivalentes. Sans vouloir le contester, on s'est montré souvent trop peu sévère dans le choix à faire entre elles. Une première condition est évidente, c'est l'invariabilité approchée du rapport entre le nombre des événements et celui des épreuves.

Quand cette première vérification réussit, on s'en contente presque toujours; le rapport constant fait connaître la composition d'une urne dans laquelle doivent se faire les tirages fictifs; on en déduit, pour un grand nombre d'épreuves, les conséquences estimées de plus en plus probables.

Si l'on observe, par exemple, dans un pays dont la population est stationnaire, le nombre des décès, celui des naissances, le rapport du nombre des filles à celui des garçons, le nombre des incendies, celui des jours où le vent souffle dans une direction désignée, etc., on trouvera, avec une approximation inégale, mais toujours grande à la longue, un rapport invariable entre le nombre d'événements d'un genre désigné et le nombre des épreuves. On comprend dans quel sens est pris le mot *épreuve*. S'il s'agit, par exemple, des incendies, chaque maison sert d'épreuve, et le rapport dont nous parlons est celui de leur nombre total à celui des sinistres.

Chaque événement fortuit acquiert par ces relevés, quand ils portent sur de grands nombres, une probabilité déterminée sur laquelle, lorsque les rapports restent constants, ne peut s'élever aucun doute.

Si, sur 10 000 individus âgés de 30 ans, 5000 atteignent l'âge de 65 ans, on conclura, très légitimement, que pour un homme de 30 ans choisi au hasard la probabilité de vivre 35 ans est $\frac{1}{2}$.

La conclusion étant acceptée, elle n'autorise pas l'assimilation des chances de décès des hommes de 30 ans, dans une période de 35 ans, à celles du tirage au sort dans une urne contenant 1 boule blanche et 1 boule noire.

Si dans une telle urne on fait 10000 tirages, le nombre des boules blanches obtenues sera 5000 environ, un peu plus ou un peu moins, selon les caprices du hasard.

Si, sur 10000 individus âgés de trente ans, on compte les survivants 35 ans après, ce nombre, d'après les Tables qui sont très exactes, sera 5000 environ, un peu plus, un peu moins, suivant des circonstances que nul ne peut prévoir.

Les deux cas sous ce rapport sont identiques.

Est-ce là tout ce qu'on doit demander?

Les nombres comparés diffèrent peu de 5000.

Mais l'écart dans un cas, celui des décès, est complètement inconnu; nous n'en pouvons rien dire, moins encore affirmer. Dans le cas des tirages au sort dans une urne, il est soumis à des lois précises.

La moyenne de ses valeurs absolues, celles des valeurs de son carré peuvent être, avec confiance, calculées à l'avance.

On peut affirmer que, dans le cas pris pour exemple, sur 10000 tirages renouvelés un grand nombre de fois, la valeur moyenne de l'écart sera 40 ; celle de son carré, 2500.

Si, en considérant un grand nombre de groupes de 10000 hommes de 30 ans, la moyenne générale des décès en 35 ans étant égale à 5000, la moyenne des écarts, au lieu d'être 40, se trouve égale à 100, on pourra, sans en conclure l'existence d'une cause perturbatrice, accuser de la discordance la prétention d'assimiler deux problèmes très différents.

Il y a, nous l'avons dit, bien des moyens de consulter le hasard ; quand ils donnent le même résultat moyen, ils ne donnent pas pour cela les mêmes probabilités d'écart. Au lieu de tirer des boules dans une urne de composition donnée, on peut associer plusieurs urnes de composition différente et puiser alternativement dans chacune d'elles : les résultats moyens sont les mêmes que pour des tirages faits dans une urne de composition moyenne, les chances d'écart ne le sont pas.

Si, pour prendre un cas extrême, au lieu de puiser 10000 fois dans une urne contenant 1 boule noire et 1 boule blanche, on puisait alternativement dans deux urnes contenant, l'une la boule noire, l'autre la boule blanche, on obtiendrait avec certitude 5000 fois la boule blanche, et l'écart deviendrait nul.

La substitution de plusieurs urnes à une seule pour représenter les Tables de mortalité paraît, *a priori*, très plausible. Parmi les individus du même âge, il est impossible de ne pas faire des catégories pour lesquelles les chances de vie sont inégales. Elles ne sont pas les mêmes pour la ville et pour la campagne ; on doit tenir compte des habitudes d'oisiveté ou de travail, de la profession exercée, de l'intempérance ou de la sobriété, de la longévité des parents, etc. La statistique confond tous les cas et donne une moyenne ; on approcherait davantage de la vérité en formant une Table pour chaque catégorie : chaque Table alors serait remplacée par une urne et les compositions seraient différentes.

Si l'événement étudié est la chute de la pluie et que, dans un lieu déterminé, on observe, en moyenne, sur un grand nombre de siècles, 92 jours de pluie par an, la probabilité sera, pour qu'il pleuve un jour donné, $\frac{92}{365}$, cela n'est pas contesté; mais, chaque année, le nombre des jours de pluie s'écartera plus ou moins de 92 : la moyenne des écarts n'a rien de commun avec celle qui se produirait si, chaque année, on faisait 365 tirages dans une urne contenant 92 boules blanches et 273 boules noires.

La cause de la différence, très probablement, est autre dans ce cas que dans le précédent. La probabilité pour qu'il pleuve un jour désigné longtemps à l'avance est $\frac{92}{365}$; mais, pour qu'il pleuve deux jours de suite, elle est très différente de $(\frac{92}{365})^2$. Quand il fait mauvais temps, ce n'est pas d'habitude pour un jour seulement; la probabilité d'un tirage est influencée par celle du tirage précédent. Cela ne change rien aux moyennes, puisque l'urne a été composée précisément pour les rendre égales; mais il n'y a plus entre les écarts, indépendamment de toute cause perturbatrice, aucune relation nécessaire.

232. La question générale semble devoir être posée de la manière suivante :

Un événement fortuit peut, sur μ épreuves, arriver un nombre inconnu de fois : la probabilité pour qu'il arrive n fois est p_n; le nombre probable des arrivées sera

$$(1) \qquad p_1 + 2p_2 + 3p_3 + \ldots + \mu p_\mu;$$

en le désignant par μp, la statistique indiquera le nombre p comme probabilité de l'événement à chaque épreuve.

Si l'on renouvelle n fois les μ épreuves, le nombre d'arrivées s'écartera, pour chaque série, de la valeur moyenne μp. En nommant N_1, N_2, ..., N_n les nombres successifs sur μ épreuves, la moyenne

$$(2) \qquad \frac{N_1 + N_2 + N_3 + \ldots + N_n}{n}$$

différera peu de μp. Nous pouvons même la regarder comme égale à μp, car c'est ce rapport seul qui nous fait connaître la probabilité moyenne désignée par p.

L'écart dans les μ épreuves formant la série du rang i sera

$$(3)\qquad \frac{N_1 + N_2 + N_3 + \ldots + N_n}{n} - N_i.$$

La moyenne des carrés des valeurs de cette différence aurait (62) une expression très simple, $\mu p(1 - p)$, si l'événement était le tirage au sort dans une urne ; elle pourra, sans qu'on s'en étonne, prendre dans le cas général une valeur très différente.

L'étude de ces valeurs dans tous les cas possibles serait intéressante.

La moyenne des carrés de l'expression (3) est identiquement (163)

$$(4)\qquad \frac{N_1^2 + N_2^2 + \ldots + N_n^2}{n} - \left(\frac{N_1 + N_2 + \ldots + N_n}{n}\right)^2.$$

Le second terme différera peu de $\mu^2 p^2$; cela résulte, nous l'avons dit, de la définition même de p.

Le premier terme

$$\frac{N_1^2 + N_2^2 + \ldots + N_n^2}{n}$$

n'est pas une fonction déterminée de p.

Si les nombres N_1, N_2, ..., N_n résultent de tirages au sort dans une urne donnant à la sortie d'une boule blanche la probabilité p, on aura approximativement, pour un grand nombre d'épreuves,

$$(5)\qquad \frac{N_1 + N_2 + \ldots + N_n}{n} = \mu p.$$

$$(6)\qquad \frac{N_1^2 + N_2^2 + \ldots + N_n^2}{n} = \mu^2 p^2 + \mu p(1 - p).$$

L'équation (5) est évidente.

Le premier membre de (6) peut, si n est grand, être remplacé par la valeur probable de N^2, c'est-à-dire par le carré du nombre d'arrivées sur μ épreuves de l'événement dont la probabilité est la somme des termes du développement de $(p + q)^\mu$ multipliés chacun par le carré de l'exposant de p. Cette somme a été calculée (183).

Si l'on ne connaît sur les probabilités $p_1, p_2, \ldots, p_\mu$ que la seule équation

$$p_1 + 2p_2 + 3p_3 + \ldots + \mu p_\mu = \mu p,$$

à laquelle il faut adjoindre la condition identique

$$p_1 + p_2 + p_3 + \ldots + p_\mu = 1,$$

l'équation (6) n'est plus démontrée. Il est impossible de connaître, d'après les données, la valeur probable de

$$\frac{N_1^2 + N_2^2 + \ldots + N_\mu^2}{n}$$

égale à

$$p_1 + 4p_2 + 9p_3 + \ldots + \mu^2 p_\mu;$$

et par conséquent aussi la valeur probable de l'écart représentée par l'expression (4) doit rester inconnue.

233. Lorsque Laplace et Poisson ont cherché les probabilités de certaines anomalies locales dans le rapport du nombre de naissances masculines et féminines, ils n'ont pas tenu compte des différences très grandes que nous venons de signaler. Leurs calculs sont faits comme si, la naissance d'un garçon ayant une certaine probabilité, les résultats possibles d'un nombre quelconque de naissances avaient, à moins de causes perturbatrices, les mêmes chances que si l'on tirait des boules d'une même urne convenablement préparée.

234. Lorsque la probabilité d'un événement est p, la valeur probable du nombre d'arrivées sur μ épreuves est μp, et celle du carré de l'écart entre le nombre véritable et le nombre probable μp est (62) $\mu p(1-p)$.

Si, à une urne donnant une probabilité p à l'événement, on substitue n urnes différentes donnant les probabilités $p_1, p_2, \ldots, p_n$, dans lesquelles on puisera alternativement, la probabilité moyenne étant égale à p, la valeur moyenne du carré de l'écart sera diminuée.

Si l'on tire, en effet, successivement dans les diverses urnes et que, le nombre total des tirages étant μn, on ait puisé μ fois dans chacune, le nombre probable des boules blanches sorties sera

$$\mu(p_1 + p_2 + p_3 + \ldots + p_n).$$

Sur un même nombre μn de tirages dans la première urne, le nombre probable des boules blanches serait

$$\mu n p.$$

Ces nombres sont égaux, puisque, par hypothèse, la moyenne

$$\frac{p_1 + p_2 + p_3 + \ldots + p_n}{n}$$

est égale à p.

Les chances d'écart sont très différentes.

Dans le cas des μn tirages faits dans l'urne qui donne à la sortie d'une boule blanche la probabilité p, si l'on désigne le nombre des boules blanches sorties par

$$\mu n p + z,$$

la valeur probable de z^2 est (62)

$$(7) \qquad \mu n p(1 - p).$$

Dans le cas des n séries de μ tirages qui forment la seconde épreuve, les nombres de boules blanches pourront être représentés par

$$\mu p_1 + z_1, \quad \mu p_2 + z_2, \quad \ldots, \quad \mu p_n + z_n;$$

l'écart entre leur nombre total et le nombre le plus probable sera

$$z_1 + z_2 + \ldots + z_n,$$

dont le carré peut être représenté par

$$(8) \qquad \Sigma z_i^2 + 2\Sigma z_i z_{i'}.$$

La valeur probable de $z_i z_{i'}$ est nulle, quels que soient i et i'; celle de z_i^2 est

$$\mu p_1(1 - p_1).$$

La valeur probable du carré de l'écart dans l'ensemble des μn

épreuves faites dans n urnes différentes est

$$(9) \qquad \mu[p_1(1-p_1)+p_2(1-p_2)+\ldots+p_n(1-p_n)].$$

La différence des expressions (7) et (9) est, on le voit aisément,

$$\frac{\mu}{n}\Sigma(p_i-p_{i'})^2;$$

elle est essentiellement positive, et la valeur probable du carré de l'écart dans le cas d'une seule urne est, pour une même probabilité moyenne, plus grande que pour les urnes associées.

235. Les remarques précédentes peuvent s'appliquer à la théorie des assurances. Le bénéfice d'une Compagnie d'assurances sur la vie, par exemple, dépend du nombre des décès qui surviendront dans l'année parmi les assurés. Ce nombre se compose de deux parties : un terme fixe, proportionnel au nombre des assurés et donné par les Tables, et un terme aléatoire, inconnu de grandeur et de signe, que nous nommerons l'*écart*. Le premier terme fait connaître la valeur équitable de la prime à payer, le second représente les variations du bénéfice annuel; il est très probablement petit par rapport au premier, si le nombre des affaires est considérable.

L'appréciation réduite à ces termes vagues n'est pas contestable; mais il n'est pas permis de la réduire en formule en assimilant les écarts à ceux que peuvent produire des tirages au sort dans une urne de composition fixe.

Si l'on considère, par exemple, une Compagnie d'assurances mutuelles contre l'incendie, la part de chacun dans la répartition des sinistres variera d'autant moins, toutes choses égales d'ailleurs, d'une année à l'autre, que le nombre des assurés sera plus grand et que les sommes assurées à chacun différeront moins de l'égalité.

Soient :

μ_1 le nombre des assurés à qui la prime à payer en cas de sinistre est α_1;

μ_2 le nombre de ceux pour qui la prime est α_2;

. . . . ;

μ_n le nombre de ceux pour qui elle est α_n.

Soient :

$e_1, e_2, \ldots, e_n$ les écarts relatifs à chacune de ces catégories;
p la probabilité d'un sinistre.

La somme probable à payer sera

$$p(\mu_1\alpha_1 + \mu_2\alpha_2 + \ldots + \mu_n\alpha_n),$$

et la part proportionnelle de celui qui doit recevoir α_1 est

$$\frac{p\alpha_1(\mu_1\alpha_1 + \mu_2\alpha_2 + \ldots + \mu_n\alpha_n)}{\alpha_1 + \alpha_2 + \ldots + \alpha_n}.$$

L'*écart* de la somme à payer, c'est-à-dire la différence entre la somme prévue et celle qui sera réellement due, est pour la Compagnie

$$\alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_n e_n.$$

Les écarts $e_1, e_2, \ldots, e_n$ ayant des valeurs probables égales à zéro, ainsi que leurs produits deux à deux, le carré de cette expression a même valeur probable que

$$\alpha_1^2 e_1^2 + \alpha_2^2 e_2^2 + \ldots + \alpha_n^2 e_n^2. \tag{10}$$

Si le sinistre dont la probabilité est p était assimilé à un tirage au sort dans une urne de composition invariable, la valeur probable de e_i^2 serait $\mu_i p(1-p)$, et celle de la somme (10) aurait pour valeur

$$p(1-p)(\mu_1\alpha_1^2 + \mu_2\alpha_2^2 + \ldots + \mu_n\alpha_n^2). \tag{11}$$

La part correspondante de l'assuré qui doit recevoir α_i serait

$$\frac{\alpha_i p(1-p)(\mu_1\alpha_1^2 + \mu_2\alpha_2^2 + \ldots + \mu_n\alpha_n^2)}{\alpha_1 + \alpha_2 + \ldots + \alpha_n}.$$

Cette évaluation, déduite d'une assimilation que rien n'autorise, ne doit inspirer aucune confiance.

236. Je terminerai ce Chapitre en indiquant une loi remarquable de probabilité proposée par Gompertz, et qui, dans des limites assez écartées, paraît s'approcher de la vérité.

La condition arbitrairement imposée à la fonction inconnue est que la probabilité pour que deux individus d'âges connus vivent l'un et l'autre après un nombre donné d'années soit proportionnelle à celle pour qu'un troisième individu, d'âge convenablement choisi, vive lui-même après ce même nombre d'années.

Si $\varphi(z)$ désigne, pour un nombre donné de naissances, le nombre des survivants à l'âge z, la condition demandée est exprimée par l'équation

$$\frac{\varphi(a+x)}{\varphi(a)}\frac{\varphi(b+x)}{\varphi(b)} = \mathrm{G}\frac{\varphi(c+x)}{\varphi(c)}. \tag{12}$$

Cette équation doit avoir lieu quel que soit x, quand on choisit pour G une fonction convenable de a et de b.

$\frac{\varphi(a+x)}{\varphi(a)}$ est, en effet, la probabilité pour qu'un individu dont l'âge est a vive encore dans n années.

En prenant les logarithmes des deux membres de (12) et posant $l\,\varphi(u) = \mathrm{F}(u)$, la condition devient

$$\mathrm{F}(a+x)+\mathrm{F}(b+x) = \mathrm{F}(c+x)+\mathrm{H},$$

H étant une fonction de a et de b indépendante de x. En prenant la dérivée par rapport à x et posant

$$\mathrm{F}'(u) = \psi(u),$$

on a

$$\psi(a+x)+\psi(b+x) = \psi(c+x); \tag{13}$$

par conséquent, en faisant $x=0$,

$$\psi(a)+\psi(b) = \psi(c); \tag{14}$$

on doit avoir aussi

$$\psi(a+dx)+\psi(b+dx) = \psi(c+dx) \tag{15}$$

et, en retranchant,

$$\psi'(a)+\psi'(b) = \psi'(c). \tag{16}$$

Les seconds membres des équations (14) et (15) sont fonctions de c; ils dépendent donc l'un de l'autre. Il doit en être de même des premiers membres, et une relation doit exister entre les deux fonctions

$$\psi(a)+\psi(b),$$
$$\psi'(a)+\psi'(b).$$

La condition nécessaire et suffisante pour qu'il en soit ainsi est que le déterminant fonctionnel soit nul. Écrivons donc

$$\psi'(a)\,\psi''(b)-\psi''(a)\,\psi'(b)=0.$$

Telle est l'équation qui détermine ψ. Si l'on suppose que ni ψ' ni ψ'' ne soient nuls, on aura

$$\frac{\psi''(a)}{\psi'(a)}=\frac{\psi''(b)}{\psi'(b)};$$

on en déduit aisément

$$\psi'(z)=G e^{kz}$$

et, successivement,

$$\psi(z)=H e^{kz}+C_1,$$
$$F(z)=H_1 e^{kz}+C_1 z+C_2,$$
$$\varphi(z)=e^{H_1 e^{kz}+C_1 z+C_2};$$

H_1, C_1, C_2 et k sont des constantes.

Cette formule, plus générale que celle de Gompertz, a été proposée par Makeham.

Gompertz suppose C_1 et C_2 égaux à zéro et prend

$$\varphi(z)=G e^{He^{kz}}.$$

J'ai déterminé les coefficients par la condition d'accorder autant que possible les coefficients avec les meilleures Tables connues.

Les résultats sont donnés par le Tableau suivant :

$$\varphi(z)=G e^{He^{kz}},$$

$$\varphi(30)=890,\qquad \varphi(60)=584,\qquad \varphi(90)=16,$$

$K=$	$0{,}071485,$	$\log K=8{,}8542146,$
$H=-$	$0{,}0065461,$	$\log H=7{,}8159811,$
$G=$	$941{,}160.$	$\log G=2{,}9736634.$

z.	$\varphi(z)$ Calcul.	Tables des 20 C^{ies} anglaises	de Duvillard.	de Deparcieux.	de Northampton.
30....	890,00	890	890,00	890,00	890,00
35....	868,88	854	820,59	841,50	813,89
40....	839,56	813	750,30	796,63	737,78
45....	799,35	769	678,54	754,20	659,23
50....	745,18	718	603,38	704,48	579,87
55....	674,06	657	522,39	637,79	496,86
60....	584,00	584	433,78	561,40	413,64
65....	475,74	491	337,93	478,95	331,24
70....	354,87	382	238,97	375,88	250,05
75....	233,39	259	145,72	255,84	168,87
80....	124,85	142	70,49	143,08	95,19
85....	54,44	59	19,44	58,20	37,75
90....	16,00	16	7,78	12,33	9,33

Les différences entre les valeurs de $\varphi(z)$ et celles que donnent les Tables des vingt Compagnies anglaises sont moindres que les différences des diverses Tables entre elles.

CHAPITRE XIII.

PROBABILITÉS DES DÉCISIONS.

Cela fait, comment sentenciez-vous, mon amy?
Comme vous autres, messieurs, répondit Bridoye, pour celuy je donne sentence duquel la chance livrée par le sort des dez judiciaires premier advient.

RABELAIS.

237. Résumé critique des tentatives faites pour appliquer le Calcul des probabilités aux décisions judiciaires.

237. L'assimilation la plus téméraire d'un tirage au sort aux effets de causes inconnues et variables a été proposée par Condorcet.

Le Livre trop longtemps admiré sur la probabilité des décisions prises à la majorité repose tout entier sur cette confusion. Aucun de ses principes n'est acceptable, aucune de ses conclusions n'approche de la vérité.

La théorie de Condorcet a été commentée, refaite même en entier par des savants illustres ou célèbres; aucun progrès n'a pu en corriger l'impuissance.

Les successeurs de Condorcet, tout en le louant d'avoir porté le flambeau de la Science dans ces mystérieuses questions, ont reconnu l'insuffisance de ses formules : ils n'en ont pas proposé de meilleures.

Laplace a rejeté les résultats de Condorcet, Poisson n'a pas accepté ceux de Laplace; ni l'un ni l'autre n'a pu soumettre au calcul ce qui y échappe essentiellement : les chances d'erreur d'un esprit plus ou moins éclairé, devant des faits mal connus et des droits imparfaitement définis.

Dans la discussion d'une loi sur le jury, Arago allégua l'autorité de Laplace.

On pouvait, disait-il, diminuer les erreurs judiciaires dans le rapport de 5 à 7. La théorie le démontre. Ces chiffres sont aussi certains que la parallaxe du Soleil.

Un député osa exprimer un doute, Arago le traita fort mal. Quand il parlait au nom de la Science, il n'appartenait pas à des ignorants de le contredire.

On a changé la parallaxe du Soleil pour une autre plus exacte. Les chiffres de Laplace n'ont pas à être changés, ils ne méritent que l'oubli.

L'analyse du Livre de Condorcet est difficile à faire. Les erreurs y sont tellement évidentes, la confiance qu'elles inspirent tellement naïve, que l'approbation connue de juges très justement illustres rend les citations invraisemblables.

Comment croire qu'à côté des aberrations singulières, textuellement rapportées, ne se trouve pas quelque idée de génie qu'il serait juste de produire?

Le Livre n'est pas rare, chacun peut chercher.

Tout se passe, suivant Condorcet, comme si les magistrats, imitant Bridoye, juge de Myrelingues, sentenciaient par le sort des dés. L'assimilation de l'opinion d'un juge au tirage d'une boule dans une urne de composition déterminée est pour lui une identité. Si la boule est blanche, la décision sera bonne. Le juge se trompera s'il tire une boule noire. L'urne dans laquelle puise un juge éclairé et honnête contient beaucoup de boules blanches; les boules noires abondent dans celle d'un juge sans conscience.

Le difficile est de trouver la composition de l'urne. Telle n'est pas l'opinion de Condorcet. Il suppose qu'une même urne serve à tous les juges, pour toutes les causes et pour tous les tribunaux d'un même pays. Le problème n'a plus qu'une seule inconnue. Dans cette hypothèse, favorable au calcul, Condorcet est en droit de rassurer les innocents en menaçant les coupables d'un inévitable châtiment; on doit supposer, dans l'urne dont tout dépend, les boules blanches en majorité. En douter serait faire injure à la magistrature. Si les juges se trompaient plus d'une fois sur deux, il faudrait supprimer les procès.

On doit les conserver, mais assurer de bons jugements. Rien n'est plus facile; il faut accroître le nombre des juges. Quand les boules blanches sont en plus grand nombre que les noires, leur sortie en plus grand nombre est probable; elle devient certaine si les tirages sont nombreux : la probabilité d'une décision prise par la majorité peut approcher ainsi de la certitude.

Nous supposerons, dit Condorcet, les assemblées composées de votants ayant une égale justesse d'esprit et des lumières égales; nous supposerons qu'aucun votant n'ait d'influence sur les voix des autres et que tous opèrent de bonne foi.

Plus le nombre des votants augmentera, plus la probabilité de la décision sera grande : la limite de cette probabilité est la certitude.

Les illusions de Condorcet ne s'étendent pas à toutes les assemblées.

Une assemblée nombreuse ne peut pas, dit-il, être composée d'hommes très éclairés : il y aura un grand nombre de questions sur lesquelles la probabilité de la voix de chaque votant sera au-dessous de $\frac{1}{2}$. Alors, plus l'assemblée sera nombreuse, plus elle sera exposée à rendre des décisions fausses.

On peut dire plus, elle en sera certaine.

Une assemblée nombreuse, dont chaque membre se trompe plus d'une fois sur deux, se prononcera certainement contre la vérité : elle donnera un moyen sûr de la connaître. Condorcet ne l'a pas proposé, mais il résulte de ses formules; il serait injuste de lui en refuser l'honneur.

Tous les calculs ont pour base la probabilité pour qu'un juge se trompe; on ne dit ni quel juge ni dans quel procès : c'est une constante qu'il faut déterminer. Condorcet donne plusieurs solutions.

La plus assurée, malheureusement d'une exécution difficile, consisterait à réunir pour former un tribunal d'examen un assez grand nombre d'hommes *véritablement éclairés* pour que leurs décisions fussent considérées comme certaines. On saurait alors combien de fois les juges se seront trompés dans leurs décisions prises à la majorité; en admettant pour tous la même chance d'erreur on dégagera aisément des formules qui ne contiendront pas d'autre inconnue la valeur exacte de cette chance.

Cette méthode, dit Condorcet, ne peut avoir qu'un inconvénient. Il en énumère trois cependant :

La difficulté de composer un tribunal d'examen, le long temps qui serait nécessaire pour examiner un grand nombre de décisions, les embarras qui peuvent rendre l'examen difficile.

Condorcet, on le voit, ne dissimule pas les difficultés. Mais, quand on les aura surmontées, quel dédommagement !

La certitude d'un bon jugement pourra croître sans limite, il n'y aura qu'à choisir.

Si le risque de l'erreur, dit Condorcet, est tel qu'on néglige un risque semblable quand il s'agit de sa propre vie, les plus exigeants devront s'en contenter.

Beaucoup de gens réputés sages prennent à Lyon le bateau pour se rendre à Avignon. Le pont Saint-Esprit cependant est sur la route.

Que faudrait-il penser d'un tribunal qui donnerait aux innocents autant de chances d'être pendus qu'un voyageur en a de se noyer au pont Saint-Esprit ?

Cette idée ne lui plaît pas complètement.

Supposons, dit-il, que l'on sache combien il périt de paquebots sur le nombre de ceux qui vont de Douvres à Calais ou qui reviennent de Calais à Douvres, et qu'on n'ait égard qu'à ceux qui sont partis par un temps regardé comme bon et sûr par les hommes instruits dans la navigation. Il est clair qu'on aura ainsi la valeur d'un risque qu'on peut négliger sans imprudence.

Après de longues et consciencieuses recherches, Condorcet se décide à accepter la fraction $\frac{1}{144768}$: C'est la dernière concession qu'il puisse faire. C'est là, dit-il, le risque le plus considérable qu'il soit permis de négliger. C'est la probabilité d'erreur qu'une nation bien gouvernée peut laisser subsister dans les jugements et dans les décisions des assemblées délibérantes. Une erreur sur 144768 jugements est le dernier mot de Condorcet.

Laplace promet moins, mais ne tient pas mieux sa promesse ; il assimile, comme Condorcet, l'opinion d'un juge à un tirage fait dans une urne, mais il repousse l'hypothèse d'une probabilité invariable.

Laplace suppose toutefois que, dans une même cause, tous les juges ont chance égale de se tromper; il admet aussi, supposition non moins étrange, que cette chance, à l'ouverture des débats, soit complètement inconnue.

Qu'il s'agisse d'un jury d'expropriation, d'un tribunal de première instance, d'une Cour d'appel ou de la Cour de cassation, d'une question de droit ou d'une question de fait, d'un crime contre les personnes ou contre les propriétés, ses formules et ses chiffres n'en reçoivent aucun changement. Un seul renseignement figure dans ses formules : le nombre des voix émises en faveur de chaque opinion. Deux jugements portés à la majorité de cinq contre trois se valent, quels que soient les juges. Si le partage se fait dans la proportion de sept contre un, celui des juges qui s'est séparé des sept autres puise, comme eux, dans la presque unanimité la même garantie de sagacité. La chance d'erreur est la même pour tous, telle a été la base du calcul. Ces huit juges puisent dans la même urne, les boules blanches y sont en grande majorité. Si le hasard a mis une boule noire dans les mains du huitième juge, c'est un pur accident : il n'en faut rien conclure contre lui.

Les conséquences de ces hypothèses sont moins assurées, quoi qu'en ait dit Arago, que la théorie du Soleil.

Dans les tribunaux où cinq voix sont nécessaires pour une condamnation, la probabilité d'une erreur est $\frac{65}{256}$, et cela quels que soient les juges.

Si le tribunal est réduit à six membres qui ne pourraient condamner qu'à la pluralité de quatre voix, la probabilité d'une erreur à craindre serait au-dessous de $\frac{1}{4}$.

Dans le jury de douze membres, si la pluralité exigée pour la condamnation est de huit voix sur douze, la probabilité de l'erreur à craindre est $\frac{1093}{8192}$; elle est à peu près $\frac{1}{22}$ si la pluralité est de neuf voix. Dans le cas de l'unanimité, la probabilité d'une erreur est réduite à $\frac{1}{8192}$.

Telle serait, suivant le calcul, la mesure de la sécurité assurée aux innocents par la loi anglaise.

Poisson n'accepte pas la solution de Laplace, il le déclare timidement.

Laplace, dit-il, fait une hypothèse qui n'est point incontestable.

L'insuccès de son maître ne le décourage pas ; il fait reposer à son tour des calculs exacts sur des hypothèses sans fondement et propose le résultat avec la même confiance qu'un théorème de Géométrie.

Avant 1831, dit-il, et pour la France entière, la probabilité qu'un juré ne se tromperait pas dans son vote était un peu supérieure à $\frac{2}{3}$ dans le cas des crimes contre les personnes, et à peu près égale à $\frac{13}{17}$ dans le cas des crimes contre les propriétés. Sans distinction de l'espèce de crimes, cette chance était très peu inférieure à $\frac{3}{4}$ pour toute la France et un peu supérieure à cette fraction pour le département de la Seine.

La probabilité de la culpabilité de l'accusé se trouverait, pour la France entière, comprise entre 0,53 et 0,54 : elle surpasse $\frac{2}{3}$ dans le cas des crimes contre les propriétés.

Dans les années qui ont précédé 1831 et pour la France entière, la probabilité de l'erreur d'une condamnation prononcée à la majorité minima de sept voix contre cinq était, à peu près, 0,16 ou 0,04, selon qu'il s'agissait d'un crime contre les personnes ou d'un crime contre les propriétés. Sans distinction de l'espèce de crime, elle avait pour valeur 0,06.

Que faut-il croire de tout cela?

Absolument rien.

Poisson, comme Condorcet et comme Laplace, assimile les jurés à des urnes. Comme Laplace, il suppose la probabilité la même pour tous ceux qui jugent une même cause; comme Condorcet, il la suppose égale pour toutes les causes.

Il déclare formellement, il est vrai, ces hypothèses inacceptables; elles n'en sont pas moins la base de ses calculs : il croit tout concilier en substituant dans les énoncés ce qu'il appelle une *probabilité moyenne* à la constante introduite dans les démonstrations, erreur de principe moins excusable peut-être que les hypothèses les plus hasardées.

Si, sur douze jurés, sept ne se trompent jamais et cinq se trompent toujours, la probabilité moyenne d'erreur sera $\frac{5}{12}$. Elle le sera aussi si chaque juré tire sa décision bonne ou mauvaise dans une urne contenant cinq boules noires ou sept blanches. Le jury cependant, dans le premier cas, ne se trompe jamais; les boules noires, dans le second cas, seront souvent en majorité. Poisson dans ses calculs ne distingue pas les deux hypothèses.

L'une des formes les plus étranges de l'illusion, dont on fait honneur à Condorcet, a été proposée par Cournot. Il déduit des formules par un calcul très exact le mérite des trois juges qui composent un même tribunal, non seulement le mérite relatif, mais le mérite absolu, la probabilité pour chacun d'eux de ne pas se tromper dans une cause qui leur est soumise. On a peine à comprendre qu'un tel résultat n'ait pas mis en défiance un esprit rigoureux et subtil.

Le tribunal se compose de trois juges. Trois inconnues seulement sont à déterminer. Cournot, qui fait un pas vers la réalité, en supposant aux juges une sagacité inégale, leur attribue la même chance d'erreur dans toutes les causes qui leur sont soumises, croyant, comme Poisson, obtenir par cette singulière hypothèse ce qu'il nomme une *probabilité moyenne;* il suppose en outre, et c'est là la moins acceptable de ses erreurs, la chance de bien juger indépendante, pour chaque juge, de celle des deux autres. Si chaque juge se trompe une fois sur quatre, ils se tromperont tous les trois ensemble une fois sur soixante-quatre.

C'est se placer trop loin de la vérité pour que l'application des formules puisse donner même une grossière approximation. Que l'on veuille faire ou non la fiction contraire, quand un juge se trompe il y a pour cela des raisons : il n'a pas réellement mis la main dans une urne où le hasard l'a mal servi. Il a ajouté foi à un faux témoignage, le concours fortuit de plusieurs circonstances a éveillé à tort sa défiance, un avocat trop habile l'a ému, de hautes influences peut-être l'ont ébranlé. Ses collègues ont entendu les mêmes témoins, on les a instruits des mêmes circonstances, le même avocat a plaidé devant eux, on a tenté sur eux la même pression : la chance d'opiner dans le même sens n'est aucunement comparable à celle de tirer trois boules de même couleur dans trois tirages indépendants les uns des autres.

Si, comme le demande très sérieusement Cournot, on invitait le greffier à noter, après chaque jugement, l'opinion de chacun des juges, pour appliquer, quand les chiffres sont nombreux, la formule qui donne leur mérite, la perspicacité de chacun étant contrôlée par celle de ses deux collègues, le juge le mieux noté de France serait celui qui, sans discuter ni réfléchir, voterait toujours comme son président : s'il faut en croire la formule, un tel juge ne se trompe jamais.

Ni Cournot ni Poisson n'ont commis la plus petite faute comme géomètres; ils traduisent rigoureusement leurs hypothèses. Mais les hypothèses n'ont pas le moindre rapport avec la situation d'un accusé devant les juges.

Ils ont aperçu les différences et croient, en les signalant, acquérir le droit de ne pas en tenir compte.

Poisson, qui, comme Condorcet, a consacré à la théorie des jugements un volume entier rempli des plus savants calculs, croit atténuer les objections qu'il ne pouvait manquer d'apercevoir, en altérant, dans ses énoncés, la signification du mot *coupable*. On rendrait, dit-il, le langage plus exact en substituant le mot *condamnable*, qui est toute la vérité, au mot *coupable* qui avait besoin d'explications et que nous continuerons d'employer pour nous conformer à l'usage.

L'innocent, accablé sous des indices trompeurs ou victime de machinations trop habiles pour qu'aucun juge puisse les soupçonner, est un accusé *condamnable*. Poisson, *pour se conformer à l'usage*, le classe parmi les coupables. L'erreur unanime des juges devient alors une preuve de sagacité dont l'algèbre leur tient compte en évaluant leur mérite avec son infaillible précision. Dans cette suite de calculs stériles, qui resteront, comme l'a dit justement Stuart Mill, le scandale des Mathématiques, Condorcet seul a donné un sage conseil : celui de choisir pour composer les assemblées des hommes *véritablement éclairés*.

TABLE

DES

VALEURS DE L'INTÉGRALE

$$\Theta(t) = \frac{\sqrt{\pi}}{2}\int_0^t e^{-t^2}\,dt.$$

t	Θ.	t.	Θ.	t.	Θ.
0,00....	0,0000000	0,27....	0,2974182	0,54....	0,5549392
0,01....	0,0112833	0,28....	0,3078800	0,55....	0,5633233
0,02....	0,0225644	0,29....	0,3182834	0,56....	0,5716157
0,03....	0,0338410	0,30....	0,3286267	0,57....	0,5798158
0,04....	0,0451109	0,31....	0,3389081	0,58....	0,5879229
0,05....	0,0563718	0,32....	0,3491259	0,59....	0,5959365
0,06....	0,0676215	0,33 ...	0,3592785	0,60....	0,6038561
0,07....	0,0788577	0,34....	0,3693644	0,61.. .	0,6116812
0,08....	0,0900781	0,35....	0,3793819	0,62 ...	0,6194114
0,09....	0,1012806	0,36....	0,3893296	0,63....	0,6270463
0,10....	0,1124630	0,37....	0,3992059	0,64....	0,6345857
0,11....	0,1236230	0,38....	0,4090093	0,65....	0,6420292
0,12....	0,1347584	0,39....	0,4187385	0,66....	0,6493765
0,13....	0,1458671	0,40....	0,4283922	0,67. ..	0,6566275
0,14....	0,1569470	0,41. ..	0,4379690	0,68....	0,6637820
0,15....	0,1679959	0,42....	0,4474676	0,69....	0,6708399
0,16....	0,1790117	0,43....	0,4568867	0,70....	0,6778010
0,17....	0,1899923	0,44....	0,4662251	0,71....	0,6846654
0,18....	0,2009357	0,45....	0,4754818	0,72....	0,6914330
0,19....	0,2118398	0,46....	0,4846555	0,73....	0,6981038
0,20....	0,2227025	0,47....	0,4937452	0,74....	0,7046780
0,21....	0,2335218	0,48....	0,5027498	0,75....	0,7111556
0,22....	0,2442958	0,49....	0,5116683	0,76 ...	0,7175367
0,23....	0,2550225	0,50....	0,5204999	0,77....	0,7238216
0,24....	0,2657000	0,51....	0,5292437	0,78 ...	0,7300104
0,25....	0,2763263	0,52....	0,5378987	0,79....	0,7361035
0,26....	0,2868997	0,53....	0,5464641	0,80....	0,7421010

t.	Θ	t.	Θ	t.	Θ.
0,81....	0,7480033	1,25....	0,9229001	1,69....	0,9831526
0,82....	0,7538108	1,26...	0,9252359	1,70....	0,9837904
0,83....	0,7595238	1,27 ...	0,9275136	1,71....	0,9844070
0,84....	0,7651427	1,28....	0,9297342	1,72....	0,9850028
0,85....	0,7706680	1,29 ...	0,9318987	1,73 ...	0,9855785
0,86....	0,7761002	1,30...	0,9340080	1,74. ..	0,9861346
0,87....	0,7814398	1,31....	0,9360632	1,75....	0,9866717
0,88....	0,7866873	1,32....	0,9380652	1,76....	0,9871903
0,89....	0,7918432	1,33....	0,9400150	1,77. ..	0,9876910
0,90....	0,7969082	1,34....	0,9419137	1,78.. .	0,9881742
0,91....	0,8018828	1,35...	0,9437622	1,79 ...	0,9886406
0,92....	0,8067677	1,36....	0,9455614	1,80....	0,9890905
0,93....	0,8115635	1,37....	0,9473124	1,81. ..	0,9895245
0,94....	0,8162710	1,38....	0,9490160	1,82...	0,9899431
0,95....	0,8208908	1,39....	0,9506733	1,83. ..	0,9903467
0,96....	0,8254236	1,40....	0,9522851	1,84....	0,9907359
0,97....	0,8298703	1,41....	0,9538524	1,85. ..	0,9911110
0,98....	0,8342315	1,42....	0,9553762	1,86....	0,9914725
0,99....	0,8385081	1,43....	0,9568573	1,87....	0,9918207
1,00....	0,8427008	1,44....	0,9582966	1,88.. .	0,9921562
1,01....	0,8468105	1,45....	0,9596950	1,89 ...	0,9924793
1,02....	0,8508380	1,46....	0,9610535	1,90....	0,9927904
1,03....	0,8547842	1,47....	0,9623729	1,91....	0,9930899
1,04....	0,8586499	1,48....	0,9636541	1,92 ...	0,9933782
1,05....	0,8624360	1,49....	0,9648979	1,93....	0,9936557
1,06....	0,8661435	1,50....	0,9661052	1,94....	0,9939226
1,07....	0,8697732	1,51....	0,9672768	1,95....	0,9941794
1,08....	0,8733261	1,52....	0,9684135	1,96 ...	0,9944263
1,09....	0,8768030	1,53....	0,9695162	1,97....	0,9946637
1,10....	0,8802050	1,54....	0,9705857	1,98.. .	0,9948920
1,11....	0,8835330	1,55....	0,9716227	1,99....	0,9951114
1,12....	0,8867879	1,56....	0,9726281	2,00. ..	0,9953223
1,13....	0,8899707	1,57....	0,9736026	2,01. ..	0,9955248
1,14. ..	0,8930823	1,58....	0,9745470	2,02....	0,9957195
1,15....	0,8961238	1,59....	0,9754620	2,03....	0,9959063
1,16....	0,8990962	1,60....	0,9763484	2,04....	0,9960858
1,17....	0,9020004	1,61. ..	0,9772069	2,05.. .	0,9962581
1,18....	0,9048374	1,62. ..	0,9780381	2,06....	0,9964235
1,19....	0,9076083	1,63....	0,9788429	2,07....	0,9965822
1,20....	0,9103140	1,64....	0,9796218	2,08.. .	0,9967344
1,21....	0,9129555	1,65....	0,9803756	2,09....	0,9968805
1,22....	0,9155339	1,66....	0,9811049	2,10....	0,9970205
1,23....	0,9180501	1,67....	0,9818104	2,11....	0,9971548
1,24....	0,9205052	1,68....	0,9824928	2,12....	0,9972836

t.	Θ.	t.	Θ.	t.	Θ.
2,13....	0,9974070	2,57....	0,9997215	3,01....	0,9999793
2,14....	0,9975253	2,58....	0.9997364	3,02....	0,9999805
2,15....	0,9976386	2,59....	0,9997505	3,03....	0,9999817
2,16....	0,9977472	2,60....	0,9997640	3,04....	0,9999829
2,17...	0,9978511	2,61....	0,9997767	3,05....	0,9999839
2,18....	0,9979505	2,62....	0,9997888	3,06....	0,9999849
2,19....	0,9980459	2,63....	0,9998003	3,07....	0,9999859
2,20....	0,9981372	2,64....	0,9998112	3,08....	0,9999867
2,21....	0,9982244	2,65....	0,9998215	3,09....	0,9999876
2,22....	0,9983079	2,66....	0,9998313	3,10....	0,9999884
2,23....	0,9983878	2,67....	0,9998406	3,11....	0,9999891
2,24....	0,9984642	2,68....	0,9998494	3,12....	0,9999898
2,25....	0,9985373	2,69....	0,9998578	3,13....	0,9999904
2,26....	0,9986071	2,70....	0,9998657	3,14 ...	0,9999910
2,27....	0,9986739	2,71....	0,9998732	3,15....	0,9999916
2,28....	0,9987377	2,72....	0,9998803	3,16....	0,9999921
2,29....	0,9987986	2,73....	0,9998870	3,17....	0,9999926
2,30....	0,9988568	2,74....	0,9998933	3,18....	0,9999931
2,31....	0,9989124	2,75....	0,9998994	3,19....	0,9999936
2,32....	0.9989655	2,76....	0;9999051	3,20....	0,9999940
2,33....	0,9990162	2,77....	0,9999105	3,21....	0,9999944
2,34....	0,9990646	2,78....	0,9999156	3,22....	0,9999947
2,35....	0,9991107	2,79....	0,9999204	3,23....	0,9999951
2,36....	0,9991548	2,80....	0,9999250	3,24....	0,9999954
2,37....	0,9991968	2,81....	0,9999293	3,25 ...	0,9999957
2,38....	0,9992369	2,82....	0,9999334	3,26....	0,9999960
2,39....	0,9992751	2,83....	0,9999372	3,27....	0,9999962
2,40....	0,9993115	2,84....	0,9999409	3,28....	0,9999965
2,41....	0,9993462	2,85....	0,9999443	3,29....	0,9999967
2,42....	0,9993793	2,86....	0,9999476	3,30....	0,9999969
2,43....	0,9994108	2,87....	0,9999507	3,31....	0,9999971
2,44....	0,9994408	2,88....	0,9999536	3,32....	0,9999973
2,45....	0,9994694	2,89....	0,9999563	3,33 ...	0,9999975
2,46....	0,9994966	2,90....	0,9999589	3,34....	0,9999977
2,47....	0,9995226	2,91....	0,9999613	3,35....	0,9999978
2,48....	0,9995472	2,92....	0,9999636	3,36....	0,9999980
2,49....	0,9995707	2,93....	0,9999658	3,37....	0,9999981
2,50....	0,9995930	2,94....	0,9999679	3,38....	0,9999982
2,51....	0,9996143	2,95....	0,9999698	3,39 ..	0,9999984
2,52....	0,9996345	2,96...	0,9999716	3,40....	0,9999985
2,53....	0,9996537	2,97....	0,9999733	3,41....	0,9999986
2,54....	0,9996720	2,98....	0,9999750	3,42 ...	0,9999987
2,55....	0,9996893	2,99....	0,9999765	3,43....	0,9999988
2,56....	0,9997058	3,00....	0,9999779	3,44....	0,9999989

t.	Θ.	t.	Θ.	t.	Θ.
3,45...	0,9999989	3,67...	0,99999978990	3,88...	0,99999995915
3,46..	0,99999900780	3,68...	0,99999980528	3,89 ..	0,99999996230
3,47...	0,99999907672	3,69...	0,99999981957	3,90...	0,99999996522
3,48...	0,99999914101	3,70...	0,99999983285	3,91...	0,99999996790
3,49...	0,99999920097	3,71...	0,99999984517	3,92...	0,99999997039
3,50...	0,99999925691	3,72...	0,99999985663	3,93...	0,99999997260
3,51...	0,99999930905	3,73...	0,99999986726	3,94...	0,99999997482
3,52...	0,99999935766	3,74..	0,99999987712	3,95...	0,99999997678
3,53...	0,99999940296	3,75...	0,99999988629	3,96...	0,99999997860
3,54...	0,99999944519	3,76...	0,99999989477	3,97...	0,99999998028
3,55...	0,99999948452	3,77...	0,99999990265	3,98...	0,99999998183
3,56..	0,99999952115	3,78...	0,99999990995	3,99...	0,99999998327
3,57...	0,99999955527	3,79...	0,99999991672	4,00...	0,99999998459
3,58...	0,99999958703	3,80...	0,99999992200	4,10...	0,99999999330
3,59...	0,99999961661	3,81...	0,99999992881	4,20...	0,99999999714
3,60 ..	0,99999964414	3,82...	0,99999993421	4,30...	0,99999999880
3,61...	0,99999966975	3,83...	0,99999993921	4,40...	0,99999999951
3,62...	0,99999969358	3,84...	0,99999994383	4,50...	0,99999999981
3,63...	0,99999971574	3,85...	0,99999994812	4,60...	0,99999999992
3,64...	0,99999973636	3,86...	0,99999995208	4,70...	0,99999999997
3,65...	0,99999975551	3,87...	0,99999995575	4,80...	0,99999999999
3,66...	0,99999977333				

FIN.

38649 Paris. — Imprimerie GAUTHIER-VILLARS, quai des Grands-Augustins, 55.