THE FREQUENCY DISTRIBUTION OF SOME MEASURED PARALLAXES AND OF THE PARALLAXES THEMSELVES

By Edwin B. Wilson and Willem J. Luyten

Harvard School of Public Health and Harvard College Observatory

Read before the Academy April 28, 1925

It is well known that the frequency functions of a set of quantities and of their measures are different;[1] if the error of measurement is comparable in magnitude to the quantity determined, they may be decidedly dissimilar. In the case of good photographic parallaxes the reported probable errors average about eight one-thousandths of a second of arc, making a dispersion of twelve. Moreover, as we have shown,[2] there appears to be a Lexian ratio of 1.25 here involved which would increase the true dispersion to $\sigma = 15$. This might well seriously disturb the frequency distribution of the large number of parallaxes which lie between —0.″15 and +0.″050.

To obtain a fairly comparable set of stars for discussion we restricted our list to a spread of one magnitude in apparent brightness and found 313 determinations[3] of parallax by the Allegheny, McCormick and Mt. Wilson observatories of stars between magnitudes 5 and 6. For some stars there were several determinations; it was decided to discuss the frequency distribution of the measures, and thus each measure was counted separately. To have combined by averaging, the plural measurements would have introduced a variety of different weights; to have chosen one of the measures and discarded the rest would have involved a more or less arbitrary choice.[4]

The frequency curve obtained is very skew, but looks like a logarithmic transform, and is, in fact, normal in $x = \log_{10}(\tilde{\omega} + 20)$ the parallax $\tilde{\omega}$ being taken in thousandths of a second of arc. Plotted on probability paper, the distribution nowhere departs from a straight line more than 2 or 3 per cent, which is about the magnitude of the errrors of sampling. The fit appears better than would be expected on chance (figure 1, line with points).

It may therefore be assumed that the frequency distribution of these parallax-measurements is:

$$f(x) = \frac{1}{\sqrt{2\pi}\tau} e^{-(x-x_0)^2/2\tau^2}, \quad x_0 = 1.625 \quad \tau = 0.275.$$

The method of passing from this to the frequency function of the parallaxes themselves, given by Whittaker and Robinson[1] is not very satisfactory on account of the slow convergence of the series

$$f(\tilde{\omega}) = e^{-D^2/2\sigma^2} \frac{\log_{10} e}{\sqrt{2\pi} + (\tilde{\omega} + 20)} e^{-[\log_{10}(\tilde{\omega}+20) - 1.62]^2/2\tau^2}, \; D = d/d\tilde{\omega}$$

The term D^4 cannot be neglected and the expressions become much involved. It is, however, possible to calculate the moments[5] of $f(x)$ on the basis of the variable $\tilde{\omega}$ and to express the moments of the true distribution (φ) in terms of them and of $\sigma = 15$.

Let $\log_{10} G = x_0 = 1.625, \qquad \log_{10} H = \tau^2/_2 . \log_{10} e = 0.087,$

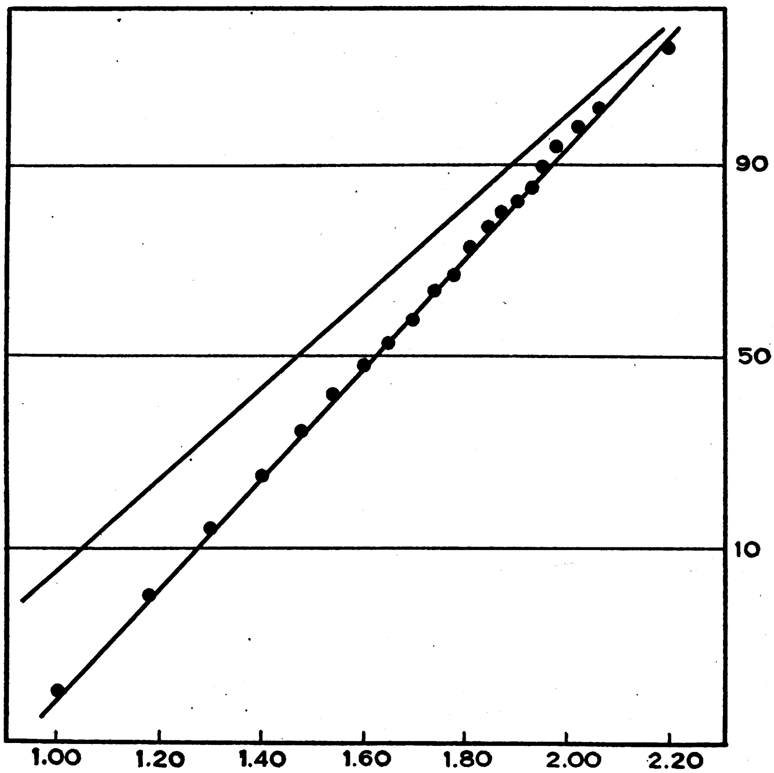

FIGURE 1

Plot on logarithmic-probability paper of the distribution of observed parallaxes (pointed line) with scale giving $\log_{10}(\tilde{\omega} + 20)$ and of the fitted logarithmic transform (above) with scale giving $\log_{10}(\tilde{\omega} + 6.3)$.

then G is the median of the distribution f and GH is its arithmetic mean measured from the origin $\tilde{\omega} = -20$. The moments about the mean are:

$$\mu_2 = G^2H^2(H^2 - 1), \qquad \mu_3 = G^3H^3(H^2 - 1)^2 (H^2 + 2)$$

$$\mu_4 = G^4H^4(H^2 - 1)^2 (H^8 + 2H^6 + 3H^4 - 3)$$

The observational errors distributed normally with $\sigma = 15$ do not disturb the mean and the third moment but change the second and fourth moment of the true frequency function into:[6]

$$\mu_2' = \mu_2 - \sigma^2, \qquad \mu_4' = \mu_4 - 6\sigma^2\mu_2 + 3\sigma^4.$$

We have $\log_{10} GH = 1.712$, $GH = 51.5$, and the mean of $\bar{\omega}$ is 31.5. Furthermore $\mu_2 = 1308$, $\mu_3 = 116{,}000$, $\mu_4 = 26{,}400{,}000$; and if the value of σ be 15 we have $\mu_2' = 1083$, $\mu_4' = 24{,}600{,}000$. This gives for the new curve the Pearsonian constants:

$$\beta_1 = 10.6 \qquad \beta_1^{1/2} = 3.26, \qquad \beta_2 = 21.0 \quad k_1 = 4, \qquad k_2 = 7.4.$$

With these values a Pearson curve of type VI may be fitted, but the origin thus found has apparently nothing to do with the astronomical problem, the numerical formula is involved, and the solution is generally unsatisfactory, as is often the case.

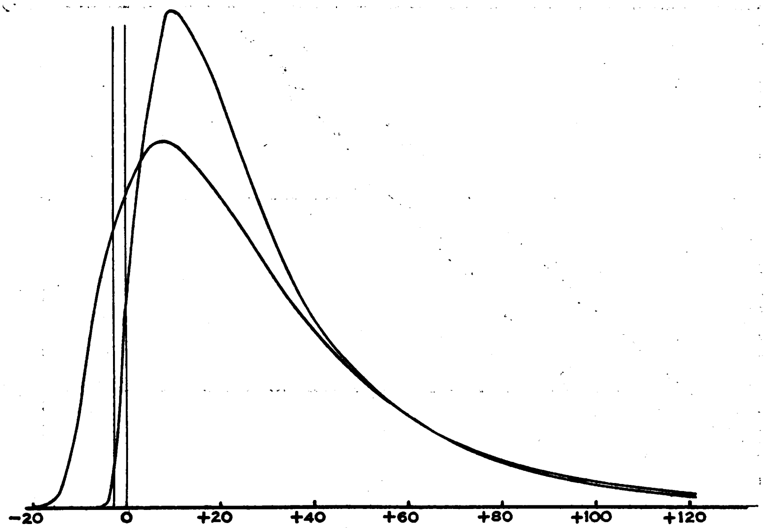

FIGURE 2

Graphs on a parallax base of the frequency distributions given in figure 1. The vertical line to the left of the origin is located at $\bar{\omega} = -2.7$ and represents the true origin of parallaxes when allowance is made for the average parallax of the comparison stars. The lower curve gives the distributions of observations and the higher curve of the parallaxes themselves.

There is, however, available another method which may be followed. The transformed curve must in this case be a good deal like a logarithmic transform with the origin drawn in from $\bar{\omega} = -20$ toward $\bar{\omega} = 0$. As a first approximation we may determine that logarithmic transform which has its first three moments identical with the above computed for the transformed true frequency curve.

Let the new origin be at a, on the scale of $\bar{\omega} + 20$. Then

$$GH = 51.5 = G'H' + a; \qquad (G'H')^2 (H'^2 - 1) = u_2' = u_2 - \sigma^2 = 1083$$

$$\mu_3 = \mu_3' = 116{,}000 = (G'H')^3 (H'^2 - 1)^2 (H'^2 + 2).$$

These equations may be solved for G'; H'; and a, to give

$$a = 13.7, \quad G' = 28.7, \quad H' = 1.32, \quad H'^2 = 1.75, \text{ and}$$

$$f(\bar{\omega}) = \frac{\log_{10} e}{\sqrt{2\pi}\,\rho(\bar{\omega} + 6.3)} e^{-[\log_{10}(\bar{\omega} + 6.3) - 1.46]^2/2\rho^2}, \quad \rho = 0.324.$$

The observed median or geometric mean is $41.7 - 20 = 21.7$; the median or geometric mean ex observational errors is $28.7 - 6.3 = 22.4$. The observed mode is $G/H^2 = 27.6$ measured from the origin $\bar{\omega} = -20$, or at parallax $\bar{\omega} = 7.6$; the corrected mode is at $G'/H'^2 = 16.4$, or at $\bar{\omega} = 16.4 - 6.3 = 10.1$. There are no parallaxes under —6.3; there are only seven parallaxes out of 311 that are negative, instead of 41 observed. As there is a systematic correction[2] of 2.7 which should be added to the parallaxes to reduce them to absolute values some of the parallaxes might be negative. In this distribution $\log(-2.7 + 6.3) = \log 3.6 = 0.56$ which departs from 1.46 by $0.90 = 2.8\rho$ and occurs only once in 350 times, or once in our series instead of 25 times as before. The new distribution is therefore a great improvement.

The fourth moments of the new and old curves do not check as they should. But the maximum contribution to these moments on either curve lies out in the region of parallaxes 200 to 250 where observations are so few that no confidence can be placed in the values of the moments. Indeed the maximum contribution to the third moments, though not nearly so far out, is still far enough to render somewhat doubtful the identification of $M_3 = M_3' = 116{,}000$ as a device to determine the new curve. Consequently the location of the mode on the new curve and the minutiae of its configuration cannot be regarded as well established; but we believe that its general characteristics are as portrayed and that it gives a far better picture of the frequency distribution of these parallaxes than does the original curve. The example serves to point out how seriously two such frequency distributions may differ (fig. 2).

[1] Whittaker and Robinson, *Calculus of Observations*, Art. 105, p. 206.

[2] These PROCEEDINGS, **10,** 129 (1924).

[3] Two parallaxes, at —25 and —31, were rejected.

[4] On the whole, the procedure followed seems as good as any, though it makes the frequency distribution of the parallax-measurements slightly different from that in which no star was represented by more than one parallax. The effect is most pronounced in the part of the distribution determined by the larger parallaxes, which is the part of less interest in the present discussion.

[5] See, e.g., Arne Fisher, *Mathematical Theory of Probabilities*, 2nd Ed., Chap. XVI, p. 237.

[6] If a frequency curve $\varphi(x)$ has each of its elements $\varphi\, dx$ subject to dispersion by a normal error function, the resulting observed frequency function $f(\xi)$ is

$$f(\xi) = \int_x \frac{\varphi(x)dx}{\sqrt{2\pi}\sigma} e^{-(\xi - x)^2 : 2\sigma^2}.$$

Then $$\int_\xi \xi^k f(\xi)d\xi = \int_\xi \int_x [(\xi - x) + x]^k \frac{\varphi(x)}{\sqrt{2\pi}\sigma} e^{-(\xi-x)^2:2\sigma^2} d\xi\, dx.$$

As $d(\xi - x) = d\xi$, the integration in ξ may be performed as

$$\int_\xi \xi^k f(\xi)d\xi = \int_x (M_k + kM_{k-1}\, x + \tfrac{1}{2}\, k(k-1)M_{k-2}\, x^2 + \ldots + x^k)\, \varphi(x)dx$$

where M_k are the moments of order k for the normal frequency function and are simply expressible in terms of σ (the moments of odd order vanish). Next an integration with respect to x gives

$$\nu_k = \int_\xi \xi^k f(\xi)d\xi = M_k + kM_{k-1}\, \nu_1' + \tfrac{1}{2}\, k(k-1)M_{k-2}\, \nu_2' + \ldots + \nu_k'$$

where ν are the moments about any origin the observed frequency function f and ν' those about the same origin for the true frequency function φ.

THE EFFECT OF VARYING MASS ON A BINARY SYSTEM

By Ernest W. Brown

Yale University, New Haven

Read before the Academy April 28, 1925

1. In the *Monthly Notices* for 1924, November, Dr. Jeans considers the effect on a variable mass M of a central acceleration M/r^2. He writes down the equations in polar coördinates, namely,

$$\frac{d^2r}{dt^2} - r\left(\frac{d\theta}{dt}\right)^2 = -\frac{M}{r^2}, \qquad r^2\left(\frac{d\theta}{dt}\right) = \text{const.,} \tag{1.1}$$

and the equation deducible from these, namely,

$$\frac{d}{dt}\left\{\frac{1}{2}\left(\frac{dr}{dt}\right)^2 + \frac{1}{2}\, r^2\left(\frac{d\theta}{dt}\right)^2 - \frac{M}{r}\right\} = -\frac{1}{r}\frac{dM}{dt}. \tag{1.2}$$

He puts $-M/2a$ for the terms in brackets where $2a$ is the major axis in an elliptic orbit, so as to obtain

$$\frac{d}{dt}\left(\frac{M}{2a}\right) = \frac{1}{r}\frac{dM}{dt}. \tag{1.3}$$

Then, pointing out that the mean value of $1/r$ is $1/a$, he deduces

$$Ma = \text{const.} \tag{1.4}$$

2. Suppose, however, we should write (1.2) in the form

$$\frac{d}{dt}\left\{\frac{1}{2}\left(\frac{dr}{dt}\right)^2 + \frac{1}{2}\, r^2\left(\frac{d\theta}{dt}\right)^2 - \frac{M_0}{r}\right\} = (M - M_0)\,\frac{d}{dt}\left(\frac{1}{r}\right) \tag{2.5}$$

and put $-M_0/2a$ for the portion in parentheses in the left hand member, we should get