TRANSFORMATIONS OF THE PEARSON TYPE III DISTRIBUTION

By A. C. Olshen

I. Introductory

Transformations of the normal curve have been used as a basis for the representation of skew frequency distributions by Edgeworth, Kapteyn, Van Uven, Bernstein, and others. Various studies have been made of the distributions obtained by replacing each of a set of normally distributed variates by a logarithmic function of the variates. Among the earlier investigators along this line were Galton, and McAllister; later, works by Jorgensen, Fisher, Wicksell, Davies, and a more recent study by Pae-Tsi-Yuan, were added.

Rietz[1] restated and treated, in a general fashion, the question as to the properties of the distribution of powers of a set of variates which are known to be normally distributed. By a suitable choice for the origin of the normal curve, he obtained results which are applicable in answering questions which frequently arise in the applied field concerning the properties of families of interrelated distributions, one strain of which is known to be normally distributed. For example, in the family made up of the diameters, surface areas, volumes, etc. of some physical quantity, if it were known that one set, the surface areas for instance, were distributed normally, then from his results we have the properties of the distributions of any of the other sets.

Likewise it has seemed of interest to investigate, in a similar fashion, the properties of the transformed Type III Pearson distribution. We shall treat both the power and logarithmic transformations. For instance, if we knew that any one of the physical measurements, velocity, kinetic energy, momentum, or centrifugal force (all of which are functions of the velocity) were distributed according to a Type III curve, then we raise the question as to the properties of the distributions of any of the others. Similarly, if the intensity of certain light, I, were known to be distributed according to a Type III law, we will discuss the properties of the distribution of the brightness, B, of the light as seen by the eye, since the two are known to be related by the law $B = K \log I$. The same analysis applies to the relationship between L, the loudness of a sound, and E, the energy in the sound wave, since $L = K \log E$.

Two forms of the Type III distribution will be considered. In the first form, all the variates are taken positive; in the second form, the origin is at the mean and the variates are measured in units of standard deviation.

---

[1] H. L. Rietz, Frequency Distributions Obtained By Certain Transformations of Normally Distributed Variates, Annals of Math., Vol. 23, (1922) pp. 291–300.

In the last section, a transformation is developed which will transform the ordinates of a given probability function into the ordinates of the normal curve, $y = Ce^{\frac{-t^2}{2}}$, to within certain approximations. This transformation is applied to the Type III distribution and to the distribution obtained under power transformations of variates of the Type III distribution.

II. Power Transformations

a. *Type III curve with all variates positive.*

Given the Pearson Type III law,

$$(1) \qquad y = y_0 x^{\gamma\bar{x}-1} e^{-\gamma x}, \qquad 0 \leq x < \infty,$$

where

$$(2) \quad \gamma = \frac{2\mu_2}{\mu_3} > 0, \qquad y_0 = \frac{(\gamma)^{\gamma\bar{x}}}{\Gamma(\gamma\bar{x})}, \qquad \bar{x} = \mu_1' > \frac{1}{\gamma}, \qquad x_{\text{mo.}} = \bar{x} - \frac{1}{\gamma}.$$[2]

The probability function (1) is a single-valued, real-valued, non-negative, continuous function of x with $\int_0^\infty y\,dx = 1$. The probability that a variate chosen at random will fall into the interval α_1 to α_2 is given by

$$(3) \qquad P = \int_{\alpha_1}^{\alpha_2} y\,dx.$$

Let us make a transformation by replacing each variate x by x', where $x' = x^n$, and n is a real number on which restrictions will be placed as we proceed. When n is such that x' may have more than one value corresponding to an assigned value of x we shall consider only the principal value of x'. Then $dx' = nx^{n-1}dx$, except at $x = 0$ when $n < 1$, and $dx = \dfrac{dx'}{nx'^{\frac{n-1}{n}}}$ except at $x' = 0$ when $n > 1$, or $n < 0$.

The frequency function of the x' variates is given by

$$(4) \qquad f(x') = \frac{y_0}{n} x'^{\frac{\gamma\bar{x}}{n}-1} e^{-\gamma x'^{\frac{1}{n}}},$$

which does not represent a Type III curve when $n \neq 1$. The function (4) is discontinuous at $x' = 0$ if $\dfrac{\gamma\bar{x}}{n} < 1$. Likewise, corresponding to (3) we have

$$(5) \qquad P = \frac{y_0}{n} \int_{\alpha_1^n}^{\alpha_2^n} x'^{\frac{\gamma\bar{x}}{n}-1} e^{-\gamma x'^{\frac{1}{n}}} dx'.$$

---

[2] The expression $x_{\text{mo.}}$ represents the mode, and $x_{\text{md.}}$ represents the median.

In order to study the maxima and minima points of (4) we take the derivative

$$\frac{df(x')}{dx'} = f(x')(nx')^{-1}\left\{-\gamma x'^{\frac{1}{n}} + \gamma\bar{x} - n\right\}. \tag{6}$$

The derivative changes signs at

$$x' = \left(\bar{x} - \frac{n}{\gamma}\right)^n. \tag{7}$$

Thus, variates in an interval, dx', at the mode of the new distribution (4) came by transformation, from an interval in the neighborhood of $x = \bar{x} - \frac{n}{\gamma}$, which is to the left of the mode of (1) when $n > 1$. The function (4) will be a monotone decreasing or a unimodal continuous distribution with mode given by (7) according as $\bar{x}$ is equal to or is greater than $\frac{n}{\gamma}$.

It will prove convenient to discuss the properties of (4) under three headings, according as $n > 1$, $0 < n < 1$, and $n < 0$, where n or its reciprocal is an integer.

Case I. $n > 1$.

When $\bar{x} < \frac{n}{\gamma}$, (4) is a monotone decreasing function, infinite at the origin and asymptotic to the x' axis; in this case the distribution of x' is similar to the distribution arising in the corresponding transformation of a set of normally distributed[3] variates, when $\bar{x} \leq 4(n - 1)$, where $\bar{x}$ is the arithmetic mean of the x's of the normal curve. However, we are primarily interested in the case when $\bar{x} \geq \frac{n}{\gamma}$, under which condition a mode exists on the frequency curve $f(x')$ and is given by $x'_{mo.} = \left(\bar{x} - \frac{n}{\gamma}\right)^n$. Henceforth in discussing the comparative values of the measures of central tendency, it will be assumed that the condition $\bar{x} \geq \frac{n}{\gamma}$ is satisfied. We have,

$$\left(\bar{x} - \frac{n}{\gamma}\right)^n < \left(\bar{x} - \frac{1}{\gamma}\right)^n, \quad \text{where } x_{mo.} = \bar{x} - \frac{1}{\gamma}.$$

Thus, while variates at the modal value of x in the Type III distribution transform into $x' = \left(\bar{x} - \frac{1}{\gamma}\right)^n$, the mode of the new distribution is at $x'_{mo.} = \left(\bar{x} - \frac{n}{\gamma}\right)^n$ which, when $n > 1$, is to the left of the positions to which variates at the mode of the Type III distribution were transformed. Furthermore, as n increases, $x'_{mo.}$ approaches the origin.

[3] Cf. Rietz, loc. cit. p. 296.

The arithmetic mean of the x''s distributed in accord with the function (4) is given by

$$\text{(8)}\qquad \begin{aligned} \bar{x}' = {}_{x'}\mu_1' &= \frac{y_0}{n}\int_0^\infty x'^{\frac{\gamma\bar{x}}{n}} e^{-\gamma x'^{\frac{1}{n}}}\,dx' \\ &= \frac{\Gamma(\gamma\bar{x}+n)}{\gamma^n\Gamma(\gamma\bar{x})}. \end{aligned}$$

Similarly the sth moment about the origin is

$$\text{(9)}\qquad {}_{x'}\mu_s' = \frac{\Gamma(\gamma\bar{x}+sn)}{\gamma^{sn}\Gamma(\gamma\bar{x})} = \frac{(\gamma\bar{x}+\overline{sn-1})^{(sn)}}{\gamma^{sn}}.$$

But,

$$\frac{\Gamma(\gamma\bar{x}+n)}{\gamma^n\Gamma(\gamma\bar{x})} = \left(\bar{x}+\frac{n-1}{\gamma}\right)\left(\bar{x}+\frac{n-2}{\gamma}\right)\cdots(\bar{x}),$$

which is greater than $(\bar{x})^n$, hence $\bar{x}' > \bar{x}^n$. Thus, while variates at the mean value of x in (1) transform into $(\bar{x})^n$, the mean of (4) is at $\frac{\Gamma(\gamma\bar{x}+n)}{\gamma^n\Gamma(\gamma\bar{x})}$ which is to the right of the positions to which variates at the mean of (1) were transformed. We have

$$\text{(10)}\qquad \left(\bar{x}-\frac{n}{\gamma}\right)^n < (\bar{x})^n < \frac{\Gamma(\gamma\bar{x}+n)}{\gamma^n\Gamma(\gamma\bar{x})},$$

hence

$$\text{(11)}\qquad x'_{\text{mo.}} < \bar{x}'.$$

In 1895, Karl Pearson[4] showed that the median of the Type III curve was approximately two-thirds of the distance from the mode to the mean, and later Doodson[5] gave similar results. The analysis of (1) along this line is given in Section IV. However, since $x_{\text{mo.}} = \bar{x} - \frac{1}{\gamma}$, we may take $x_{\text{md.}} = \bar{x} - \frac{1}{c\gamma}$ where $c > 1$, (approximately equal to 3). Then $x'_{\text{md.}} = \left(\bar{x} - \frac{1}{c\gamma}\right)^n$. We have

$$\text{(12)}\qquad \left(\bar{x}-\frac{n}{\gamma}\right)^n < \left(\bar{x}-\frac{1}{c\gamma}\right)^n < (\bar{x})^n,$$

hence from (10)

$$\text{(13)}\qquad x'_{\text{mo.}} < x'_{\text{md.}} < \bar{x}'.$$

---

[4] Karl Pearson, Skew Variation in Homogeneous Material, Philosophical Transactions, Vol. 186A, part 1, (1895) pp. 343–414.

[5] Arthur T. Doodson, Relation of Mode, Median and Mean in Frequency Curves, Biometrika, Vol. 11, (1917) p. 425.

Considering the case when $n = 2$, we have with the aid of (9),

$$_{x'}\beta_1 = \frac{8(5\gamma^2\bar{x}^2 + 17\gamma\bar{x} + 15)^2}{(\gamma\bar{x})(\gamma\bar{x} + 1)(2\gamma\bar{x} + 3)^3},$$

and

$$_{x'}\beta_2 = \frac{3(4\gamma^4\bar{x}^4 + 72\gamma^3\bar{x}^3 + 337\gamma^2\bar{x}^2 + 629\gamma\bar{x} + 420)}{(\gamma\bar{x})(\gamma\bar{x} + 1)(2\gamma\bar{x} + 3)^2}.$$

From the moments of (1), one readily gets

$$_{x}\beta_1 = \frac{4}{\gamma\bar{x}},$$

and

$$_{x}\beta_2 = \frac{3(\gamma\bar{x} + 2)}{\gamma\bar{x}}.$$

It can be shown easily that $_{x'}\beta_1 > {}_{x}\beta_1$ and $_{x'}\beta_2 > {}_{x}\beta_2$; hence the distribution of the squares of variates is more leptokurtic and more skew than the original distribution.

From (10) and (12) it is evident that the mode approaches neither the median nor the mean as n increases, subject to the condition $\gamma\bar{x} \geq n$. Each of the ratios of the mode to the median and to the mean approaches the limit 1 as $\bar{x}$ is increased indefinitely, the rapidity of approach to the limiting value depending on the size of n.

Taking the second derivative to find the points of inflection of the function (4), we have

$$\frac{d^2f(x')}{dx'^2} = f(x')(nx')^{-2}\left\{\gamma^2 x'^{\frac{2}{n}} + \gamma x'^{\frac{1}{n}}(3n - 2\gamma\bar{x} - 1) + (\gamma^2\bar{x}^2 - 3n\gamma\bar{x} + 2n^2)\right\}.$$

When the points of inflection exist they are given by

$$x' = \frac{(2\gamma\bar{x} - 3n + 1) \pm\sqrt{n^2 - 6n + 1 + 4\gamma\bar{x}}}{2\gamma}. \tag{14}$$

Under the restriction that $\gamma\bar{x} \geq n$, the expression under the radical in (14) cannot vanish, and will always be positive.

Case II. $0 < n < 1$.

We now consider the distribution obtained by taking positive integral roots of a set of variates distributed in accord with (1). The mode of $f(x')$, as given by (7), will always exist since from (2), $\gamma\bar{x} > 1 > n$. We have

$$\begin{aligned} \left(\bar{x} - \frac{n}{\gamma}\right)^n &< \left(\bar{x} - \frac{1}{c\gamma}\right)^n \qquad \text{if } n > \frac{1}{c}, \\ \left(\bar{x} - \frac{n}{\gamma}\right)^n &> \left(\bar{x} - \frac{1}{c\gamma}\right)^n \qquad \text{if } n < \frac{1}{c}, \\ \left(\bar{x} - \frac{n}{\gamma}\right)^n &= \left(\bar{x} - \frac{1}{c\gamma}\right)^n \qquad \text{if } n = \frac{1}{c}. \end{aligned} \tag{15}$$

Hence it is evident that $x'_{mo.}$ is less than, greater than, or equal to $x'_{md.}$ according as n is greater than, less than, or equal to $1/c$. In power transformations of symmetrical distributions[6] the results differ in that the modal value is always greater or less than the median according as the value of n lies between 0 and 1 or outside of these bounds.

Here, $\left(\bar{x} - \frac{n}{\gamma}\right)^n > \left(\bar{x} - \frac{1}{\gamma}\right)^n$, hence in contrast to Case I, the mode of the new distribution is to the right of the position to which variates at the mode of the Type III distribution were transformed.

It has been shown[7] for every set of positive values that $\mu'_n > (\bar{x})^n$ when n lies outside of the interval 0 to 1, and that $\mu'_n < (\bar{x})^n$ when n lies between 0 and 1. We have then

$$\frac{\Gamma(\gamma\bar{x} + n)}{(\gamma)^n\Gamma(\gamma\bar{x})} = {}_{x'}\mu'_1 = {}_x\mu'_n < (\bar{x})^n. \tag{16}$$

The mean of the new distribution is to the left of the position to which the variates at the mean of the Type III distribution were transformed.

In Section IV, with the aid of certain approximating assumptions, it will be shown that $\bar{x}' > x'_{md.}$ when $x'_{md.} > x'_{mo.}$ and conversely, $\bar{x}' < x'_{md.}$ when $x'_{md.} < x'_{mo.}$

Case III. $n < 0$.

Let $n = -m$, where m is a positive integer. Then we have

$$f(x') = \frac{y_0}{m} x'^{-\left(\frac{\gamma\bar{x}}{m}+1\right)} e^{-\gamma x'^{-\frac{1}{m}}}, \qquad f(x') = 0 \quad \text{at} \quad x' = 0. \tag{17}$$

In place of (6) we have

$$\frac{df(x')}{dx'} = f(x')\left(mx'^{\frac{m+1}{m}}\right)^{-1}\left\{(\gamma\bar{x} + m)\, x'^{\frac{1}{m}} - \gamma\right\}, \tag{18}$$

and (7) becomes

$$x' = \left(\frac{1}{\bar{x} + \frac{m}{\gamma}}\right)^m. \tag{19}$$

But

$$\left(\frac{1}{\bar{x} + \frac{m}{\gamma}}\right)^m < \left(\frac{1}{\bar{x} - \frac{1}{\gamma}}\right)^m,$$

and

$$\bar{x}' = \frac{1}{\left(\bar{x} - \frac{m}{\gamma}\right)\left(\bar{x} - \frac{m-1}{\gamma}\right)\cdots\left(\bar{x} - \frac{1}{\gamma}\right)} > \left(\frac{1}{\bar{x}}\right)^m.$$

---

[6] H. L. Rietz, On Certain Properties of Frequency Distributions, Proc. National Academy of Science, Vol. 13, (1927) p. 820.

[7] J. L. W. V. Jensen, On Convex Functions and the Inequalities between the Means, Acta Mathematica, Vol. 30, pp. 175–193.

Hence, as in Case I, the mode of (17) is to the left of the position to which variates at the mode of (1) were transformed while the mean of (17) is to the right of the position to which the variates at the mean of (1) were transformed. Since

$$\left(\frac{1}{\bar{x} + \frac{m}{\gamma}}\right)^m < \left(\frac{1}{\bar{x} - \frac{1}{c\gamma}}\right)^m,$$

we have $x'_{\text{mo.}} < x'_{\text{md.}}$. Also

$$\frac{1}{\left(\bar{x} - \frac{m}{\gamma}\right)\left(\bar{x} - \frac{m-1}{\gamma}\right)\cdots\left(\bar{x} - \frac{1}{\gamma}\right)} > \left(\frac{1}{\bar{x} - \frac{1}{c\gamma}}\right)^m,$$

hence $\bar{x}' > x'_{\text{md.}}$. Therefore

$$x'_{\text{mo.}} < x'_{\text{md.}} < \bar{x}'. \tag{20}$$

As a special case, when $n = -1$, (17) reduces to

$$f(x') = y_0 x'^{-(\gamma\bar{x}+1)} e^{-\frac{\gamma}{x'}}, \tag{21}$$

which is a Pearson Type V distribution.

b. *Type III curve with mean zero and unit variance.*

Even though the form of the Type III distribution with which we have been dealing, wherein all the variates are positive, is more closely akin to actual distributions that may arise in applied problems, nevertheless it will be of interest to examine the properties of the transformed curve when the mean is taken as zero, with unit variance.

The second and third moments about the mean of the distribution (1) are ${}_x\mu_2 = \frac{\bar{x}}{\gamma}$, ${}_x\mu_3 = \frac{2\bar{x}}{\gamma^2}$. If we write α_3 for the third standard moment $\frac{\mu_3}{\mu_2^{3/2}}$, then

$$\gamma = \frac{4}{\bar{x}\alpha_3^2}. \tag{22}$$

By replacing the variable, x, in (1) by the expression

$$x = \bar{x}\left(1 + \frac{\alpha_3 t}{2}\right), \tag{23}$$

we obtain the Type III distribution

$$\eta = \eta_0\left(1 + \frac{\alpha_3 t}{2}\right)^{\frac{4}{\alpha_3^2}-1} e^{-\frac{2}{\alpha_3}t}, \tag{24}$$

where

$$\eta_0 = \frac{\left(\frac{4}{\alpha_3^2}\right)^{\frac{4}{\alpha_3^2}-\frac{1}{2}}}{e^{\frac{4}{\alpha_3^2}}\Gamma\left(\frac{4}{\alpha_3^2}\right)}, \qquad \alpha_3 < 2, \qquad -\frac{2}{\alpha_3} \le t < \infty, \qquad \bar{t} = 0, \qquad {}_t\mu_2 = 1.$$

Equation (22) lends itself to a simple interpretation of the restriction made in Section IIa, Case I, that $\gamma\bar{x} \geq n$ in order that the mode of (4) exist and be given by (7). The upper bound in the values of α_3 considered by Salvosa[8] in the computation of his tables was $\alpha_3 = 1.1$. Upon examination of the tables it is obvious that in most cases the skewness of the Type III distribution, as measured by α_3, will be less than 1.1. Hence in most cases we will have $3.22 \leq \gamma\bar{x} < \infty$. The effect of the limitations imposed by the condition $\gamma\bar{x} \geq n$ may be inferred to some extent from the following table.

TABLE I

The upper bound of α_3 for the existence of a mode in Case I, Section IIa

n	2	3	4	5	6	7	8	9	10	25	50	100
α_3	1.41	1.15	1.00	.89	.82	.76	.71	.67	.63	.40	.28	.20

When we make a transformation by replacing each variate t in (24) by t' where $t' = t^n$ $(n \neq 0)$ and n is an integer (positive or negative) or the reciprocal of an integer, then $dt' = nt^{n-1}\,dt$, except at $t = 0$ when $n < 1$, and $dt = \dfrac{dt'}{nt'^{\frac{n-1}{n}}}$, except at $t' = 0$ when $n > 1$, or $n < 0$. The function (24) becomes

$$(25) \qquad f(t') = \frac{\eta_0}{n}\,\frac{\left(1 + \frac{\alpha_3}{2}t'^{\frac{1}{n}}\right)^{\frac{4}{\alpha_3^2}-1} e^{-\frac{2}{\alpha_3}t'^{\frac{1}{n}}}}{t'^{\frac{n-1}{n}}}.$$

The distribution function, $f(t')$, is infinite at $t' = 0$ when $n > 1$. In place of (5), we have

$$(26) \qquad P = \frac{\eta_0}{n}\int_{a_1^n}^{a_2^n} \frac{\left(1 + \frac{\alpha_3}{2}t'^{\frac{1}{n}}\right)^{\frac{4}{\alpha_3^2}-1} e^{-\frac{2}{\alpha_3}t'^{\frac{1}{n}}}}{t'^{\frac{n-1}{n}}}\,dt'.$$

Here a_1 and a_2 are taken to be positive or zero when n is even. When n is odd a_1 and a_2 may be taken negative as (25) will give the frequency curve for negative values of t' that arise from setting $t' = t^n$ when t is negative. Examining for maxima and minima points, we have

$$(27) \qquad \frac{df(t')}{dt'} = -f(t')\left\{nt'\left(1 + \frac{\alpha_3}{2}t'^{\frac{1}{n}}\right)\right\}^{-1}\left\{t'^{\frac{2}{n}} + \frac{n\alpha_3 t'^{\frac{1}{n}}}{2} + (n-1)\right\}.$$

[8] Luis R. Salvosa, Tables of Pearson's Type III Function, Annals of Mathematical Statistics, Vol. 1, (1930) pp. 191–198.

The derivative changes signs at

$$t' = \frac{1}{2^n}\left(\frac{-n\alpha_3}{2} \pm \sqrt{\left(\frac{n\alpha_3}{2}\right)^2 - 4(n-1)}\right)^n \tag{28}$$

when $\left(\frac{n\alpha_3}{2}\right)^2 > 4(n-1)$, and at $t' = 0$ for certain values of n. When $n > 1$, $\left(\frac{n\alpha_3}{2}\right)^2 > 4(n-1)$ only for very large values of n since $\alpha_3 < 2$. When n is odd and positive, the derivative changes signs at $t' = 0$. If n is the reciprocal of an odd positive integer greater than one, there is a minimum at $t' = 0$, and the function $f(t')$ is zero at this point. Further properties of the frequency curves given by (25) will be discussed under the three cases treated in Section IIa.

Case I. $n > 1$.

When $\left(\frac{n\alpha_3}{2}\right)^2 > 4(n-1)$, it can be shown that (28) gives neither a maximum nor minimum point of $f(t')$ since $t'^{\frac{1}{n}}$ will always be less than t_l.[9] Similarly, when $\left(\frac{n\alpha_3}{2}\right)^2 < 4(n-1)$, there is neither a maximum nor minimum since (28) is imaginary. When n is odd, $f(t')$ is infinite at the origin and is a monotone increasing function of t' from the lower bound to the origin, and a monotone decreasing function of t' from the origin. When n is even, $f(t')$ is a monotone decreasing function of t', infinite at the origin. The forms of the distributions in this case are similar to those arising in power transformations of normally distributed variates[10] when $n > 1$ and $\bar{x}^2 \leq 4(n-1)$ and also to the forms arising in Section IIa, Case I, when $\bar{x} < \frac{n}{\gamma}$.

Even though we have a discontinuity at the origin, the total area under the curve is one, which is evident since we can integrate function (25) over the entire range of t' when n is odd and positive.

Case II. $0 < n < 1$.

This case includes the distribution obtained by taking positive integral roots of a set of variates. As in the study of the normal distribution,[11] we limit our considerations to the principal real values of the functions. When n is odd, there is a minimum at $t' = 0$ and a maximum given by each of the two signs before the radical in (28). Hence in this case, we have one minimum and two maxima.

With the values for n and α_3 in (24), $t_{\text{md.}} = -.164$ and $t_{\text{mo.}} = -.500$. The transformed distribution gives $t'_{\text{md.}} = -.547$ and two modes, the primary mode $t'_{\text{mo.}} = -.967$, and the secondary mode $t'_{\text{mo.}} = .903$. In contrast to the corresponding transformation of normally distributed variates, the primary mode is less than the median.

---

[9] The expression t_l represents the lower bound of t in distribution (24).

[10] H. L. Rietz, Cf. loc. cit. p. 296.

[11] Cf. Rietz. loc. cit. p. 297.

TABLE II

Comparison of the Type III Distribution and the Transformed Distribution when $\alpha_3 = .6$, $n = 3$

t	η	t'	$f(t')$
−3.00	.000001	−27.000000	0
−2.50	.001347	−15.625000	.000072
−2.00	.029467	− 8.000000	.002456
−1.75	.072787	− 5.359375	.007922
−1.50	.139285	− 3.375000	.020635
−1.25	.220462	− 1.953125	.047032
−1.00	.301350	− 1.000000	.100450
− .50	.405345	− .125000	.540460
− .25	.414211	− .015625	2.209125
− .10	.406131	− .001000	13.537700
− .05	.401485	− .000125	53.531333
− .02	.398272	− .000008	331.893333
0	.395962	0	∞
.02	.393522	.000008	327.935000
.05	.389628	.000125	51.950400
.10	.382549	.001000	12.751633
.15	.374795	.003375	5.552519
.25	.357533	.015625	1.906843
.50	.307293	.125000	.409724
.75	.252971	.421875	.149909
1.00	.200493	1.000000	.066831
1.50	.114233	3.375000	.016923
2.00	.058376	8.000000	.004865
2.50	.027285	15.625000	.001455
3.00	.011836	27.000000	.000438
3.50	.004820	42.875000	.000131
4.00	.001859	64.000000	.000039
5.00	.000242	125.000000	.000003
6.00	.000027	216.000000	0

Case III. $n < 0$.

Let $n = -m$, where m is a positive integer. Then (25) becomes

$$f(t') = \frac{\eta_0}{m} \frac{\left(1 + \frac{\alpha_3}{2} t'^{-\frac{1}{m}}\right)^{\frac{4}{\alpha_3^2}-1} e^{-\frac{2}{\alpha_3} t'^{-\frac{1}{m}}}}{t'^{\frac{m+1}{m}}}. \tag{29}$$

TABLE III

Comparison of the Type III Distribution and the Transformed Distribution When $\alpha_3 = 1$, $n = 1/3$

t	η	t'	$f(t')$
−2.00	0	−1.259921	0
−1.75	.025272	−1.205071	.11010
−1.50	.122626	−1.144714	.48206
−1.25	.251021	−1.077217	.87385
−1.00	.360894	−1.000000	1.08268
− .90	.393277	− .965489	1.09998
− .75	.427526	− .908560	1.05874
− .50	.448084	− .793701	.84683
− .27	.433958	− .646330	.54385
− .08	.405678	− .430887	.22596
0	.390734	0	0
.08	.374536	.430887	.20861
.27	.332927	.646330	.41723
.50	.280748	.793701	.53058
.64	.249865	.861774	.55669
.74	.228711	.904504	.56134
.90	.196904	.965489	.55064
1.00	.178470	1.000000	.53541
1.50	.104259	1.144714	.40985
2.00	.057252	1.259921	.27265
2.50	.029989	1.357209	.16572
3.00	.015133	1.442250	.09443
3.50	.007410	1.518295	.05125
4.00	.003539	1.587401	.02675
4.50	.001655	1.650964	.01353
5.00	.000761	1.709976	.00368
5.50	.000344	1.765174	.00322

Taking the derivative,

$$(30)\quad \frac{df(t')}{dt'} = -f(t')\left\{mt'^{\frac{m+2}{m}}\left(1 + \frac{\alpha_3}{2}t'^{-\frac{1}{m}}\right)\right\}^{-1}\left\{(m+1)t'^{\frac{2}{m}} + \frac{m}{2}\alpha_3 t'^{\frac{1}{m}} - 1\right\},$$

and in place of (28) we have

$$(31)\qquad t' = \left\{\frac{\dfrac{-m\alpha_3}{2} \pm \sqrt{\left(\dfrac{m\alpha_3}{2}\right)^2 + 4(m+1)}}{2(m+1)}\right\}^m.$$

The transformed distribution has little statistical significance for odd values of m, since $f(t')$ is a disjoined distribution. There are no values for $f(t')$ in the interval, $\left(\frac{-\alpha_3}{2}\right)^m < t' < 0$, since $\frac{-2}{\alpha_3} \leq t < \infty$. The transformed distribution is thus composed of two sections, each with its own mode. The section for negative values of t', with range $-\infty < t' \leq \left(\frac{-\alpha_3}{2}\right)^m$, has a mode given by (31) with the negative sign before the radical. The section for positive values of t', with range $0 \leq t' < \infty$, has a mode given by (31) with the positive sign before the radical.

When m is an even integer, if we assign to $f(t')$ the value 0 when $t' = 0$, $f(t')$ becomes a continuous unimodal distribution in the interval $0 \leq t' < \infty$, with the mode given by (31), with the positive sign before the radical.

III. Logarithmic Transformations

As indicated in the introduction, numerous studies have been made of the distributions obtained by replacing normally distributed variates by exponential functions of the variates. If a variate x, with range $-\infty < x < \infty$, is distributed normally with mean zero and unit variance, then by replacing x by x', where $x' = c + e^{kx}$ the range of x' becomes, $c \leq x' < \infty$. Likewise if a variate x is distributed in accord with a Type III law, with range $0 \leq x < \infty$, then in making the above transformation, the range of x' becomes $(c + 1) \leq x' < \infty$. We shall now study the properties of the distribution of x' obtained by the above transformation applied to distribution (1). Because of the similarity of the properties of the transformed frequency distributions, we shall take $k = 1$ and $c = 0$.

Letting $x' = e^x$, (1) is transformed into

$$f(x') = y_0 (\log x')^{\gamma\bar{x}-1} x'^{-(\gamma+1)}, \qquad 1 \leq x' < \infty. \tag{32}$$

Then,

$$\frac{df(x')}{dx'} = f(x')\{x' \log x'\}^{-1}\{(\gamma\bar{x} - 1) - (\gamma + 1) \log x'\}. \tag{33}$$

The derivative changes signs at

$$x' = e^{\frac{\gamma\bar{x}-1}{\gamma+1}}. \tag{34}$$

The arithmetic mean of the x''s distributed in accord with (32) is given by

$$\begin{aligned} \bar{x}' &= y_0 \int_1^\infty (\log x')^{\gamma\bar{x}-1} x'^{-\gamma}\, dx' \\ &= y_0 \int_0^\infty x^{\gamma\bar{x}-1} e^{-x(\gamma-1)}\, dx \\ &= \left(\frac{\gamma}{\gamma - 1}\right)^{\gamma\bar{x}} \quad \text{if} \quad \gamma > 1. \end{aligned} \tag{35}$$

The integral is divergent when $0 < \gamma \leq 1$, hence for these cases we take $0 < k < \gamma$, then

$$\bar{x}' = y_0 \int_0^\infty x^{\gamma\bar{x}-1} e^{-x(\gamma-k)}\, dx$$

$$= \left(\frac{\gamma}{\gamma - k}\right)^{\gamma\bar{x}}. \tag{36}$$

Likewise in order that the first s moments about the origin be finite when $0 < \gamma \leq s$, we must have $0 < k < \dfrac{\gamma}{s}$.

The median of the distribution of x''s corresponds to

$$x = \log x' = \bar{x} - \frac{1}{c\gamma},$$

hence

$$x'_{\text{md.}} = e^{\bar{x} - \frac{1}{c\gamma}}. \tag{37}$$

1. **The relative positions of the averages.** We have $e^{\frac{\gamma\bar{x}-1}{\gamma+1}} < e^{\bar{x}-\frac{1}{c\gamma}}$ since $\dfrac{\gamma\bar{x} - 1}{\gamma + 1} < \bar{x} - \dfrac{1}{c\gamma}$. Hence

$$x'_{\text{mo.}} < x'_{\text{md.}}. \tag{38}$$

Also,

$$\gamma\bar{x} \log\left(\frac{\gamma}{\gamma - 1}\right) = \bar{x}\left[1 + \frac{1}{2\gamma} + \frac{1}{3\gamma^2} + \cdots\right] > \bar{x} - \frac{1}{c\gamma}.$$

Therefore

$$e^{\bar{x}-\frac{1}{c\gamma}} < \left(\frac{\gamma}{\gamma - 1}\right)^{\gamma\bar{x}}$$

and hence

$$x'_{\text{md.}} < \bar{x}'. \tag{39}$$

From (38) and (39), we have

$$x'_{\text{mo.}} < x'_{\text{md.}} < \bar{x}'. \tag{40}$$

We shall now investigate the locations of the various averages as related to the upper and lower points of inflection whose abscissas will be denoted by I'_u and I'_l respectively. Taking the second derivative,

$$\frac{d^2 f(x')}{dx'^2} = f(x')\,\{x' \log x'\}^{-2}\,\{(\gamma\bar{x} - 1)(\gamma\bar{x} - 2) - (2\gamma^2\bar{x} + 3\gamma\bar{x} - 2\gamma - 3)\log x'$$

$$+ (\gamma + 1)(\gamma + 2) \ldots$$

The points of inflection are given by

$$x' = e^{\varphi(\gamma, \bar{x})}, \tag{41}$$

where

$$\varphi(\gamma, \bar{x}) = \frac{(\gamma\bar{x} - 1)(2\gamma + 3) \pm \sqrt{(\gamma\bar{x} - 1)\{(\gamma\bar{x} - 1) + 4(\gamma + 1)(\gamma + 2)\}}}{2(\gamma + 1)(\gamma + 2)}.$$

(a). To show $x'_{mo.} > I'_l$.
We have,

$$\frac{\gamma\bar{x} - 1}{\gamma + 1} > \frac{(\gamma\bar{x} - 1)(2\gamma + 3) - (\gamma\bar{x} - 1)(p + 1)}{2(\gamma + 1)(\gamma + 2)},$$

where

$$p + 1 = \sqrt{1 + \frac{4(\gamma + 1)(\gamma + 2)}{\gamma\bar{x} - 1}}, \quad \text{since } 2\gamma + 4 > 2\gamma + 2 - p.$$

Therefore

$$e^{\frac{\gamma\bar{x} - 1}{\gamma + 1}} > e^{\varphi_l(\gamma, \bar{x})}.$$

(b). To show $x'_{mo.} < I'_u$.
We have,

$$\frac{\gamma\bar{x} - 1}{\gamma + 1} < \frac{(\gamma\bar{x} - 1)(2\gamma + 3) + (\gamma\bar{x} - 1)(p + 1)}{2(\gamma + 1)(\gamma + 2)},$$

since $2\gamma + 4 < 2\gamma + 4 + p$.
Therefore,

$$e^{\frac{\gamma\bar{x} - 1}{\gamma + 1}} < e^{\varphi_u(\gamma, \bar{x})}.$$

From (a) and (b), we have

$$I'_l < x'_{mo.} < I'_u. \tag{42}$$

(c). To show $x'_{md.} > I'_l$.
We have,

$$x'_{md.} > x'_{mo.} \text{ and } x'_{mo.} > I'_l.$$

Consequently

$$x'_{md.} > I'_l.$$

(d). To show the conditions under which $x'_{md.}$ is less than or greater than I'_u.
Upon simplifying the inequality we find that $e^{\varphi_u(\gamma, \bar{x})}$ will be greater or less than $e^{\bar{x} - \frac{1}{c\gamma}}$ according as the expression

$$-2\bar{x}^2 + \bar{x}\left\{\frac{1}{c}\left(3 + \frac{4}{\gamma}\right) + (\gamma - 3)\right\} + \frac{1}{c}\left(2 + \frac{3}{\gamma}\right) - \frac{1}{c^2\gamma}\left(3 + \frac{2}{\gamma}\right) - 2 - \frac{1}{c^2} \tag{43}$$

is positive or negative. But (43) will be negative for all values of $\bar{x}$ if its discriminant is negative or zero. Upon further examination it can be seen that the discriminant will be negative or zero according as

$$\gamma^2 - 6\gamma\left(1 - \frac{1}{c}\right) + \frac{1}{c}\left(6 + \frac{1}{c}\right) - 7 \leq 0. \tag{44}$$

The quadratic equation in γ, given by (44), factors into

$$\left(\gamma - \frac{A - B}{2}\right)\left(\gamma - \frac{A + B}{2}\right),$$

where

$$A = 6\left(1 - \frac{1}{c}\right), \qquad B = \sqrt{36\left(1 - \frac{1}{c}\right)^2 - \frac{4}{c}\left(6 + \frac{1}{c}\right) + 28}, \qquad B > A.$$

Hence in order that $x'_{\text{md.}}$ be greater than I'_u for all values of $\bar{x}$, we must have

$$\gamma \leq \frac{1}{2}\left\{6\left(1 - \frac{1}{c}\right) + \sqrt{36\left(1 - \frac{1}{c}\right)^2 - \frac{4}{c}\left(6 + \frac{1}{c}\right) + 28}\right\}.$$

When c lies in the neighborhood of 3, γ must be less than 5. Proceeding further, we can divide (43) by negative 2, reverse the inequalities, and factor the expression into

$$\left(\bar{x} - \frac{A' - B'}{4}\right)\left(\bar{x} - \frac{A' + B'}{4}\right),$$

where

$$A' = \left\{\frac{1}{c}\left(3 + \frac{4}{\gamma}\right) + (\gamma - 3)\right\}$$

$$B' = \sqrt{A'^2 + 8\left\{\frac{1}{c}\left(2 + \frac{3}{\gamma}\right) - \frac{1}{c^2\gamma}\left(3 + \frac{2}{\gamma}\right) - 2 - \frac{1}{c^2}\right\}}$$

$$B' < A'.$$

Then

$$x'_{\text{md.}} > I'_u \quad \text{if} \quad \bar{x} < \frac{A' - B'}{4} \quad \text{or} \quad \bar{x} > \frac{A' + B'}{4}, \tag{45}$$

and

$$x'_{\text{md.}} < I'_u \quad \text{if} \quad \frac{A' - B'}{4} < \bar{x} < \frac{A' + B'}{4}. \tag{46}$$

(e). To show $I'_u < \bar{x}'$.

We have, $I_l' < \bar{x}'$, since $\bar{x}' > x'_{\text{md.}} > x'_{\text{mo.}}$ and $x'_{\text{mo.}} > I_l'$. Also, $I_u' < \bar{x}'$ for those values of γ and $\bar{x}$ for which $I_u' < x'_{\text{md.}}$. It remains to be shown that $\bar{x}'$ is always greater than I_u' for all values of γ and $\bar{x}$. To show that

$$e^{\varphi_u(\gamma,\bar{x})} < \left(\frac{\gamma}{\gamma - 1}\right)^{\gamma\bar{x}},$$

it will be sufficient to show that $\varphi_u(\gamma, \bar{x}) < \bar{x}$, since

$$\gamma\bar{x} \log\left(\frac{\gamma}{\gamma - 1}\right) = \bar{x}\left(1 + \frac{1}{2\gamma} + \frac{1}{3\gamma^2} + \cdots\right) > \bar{x}.$$

The inequality is satisfied if

$$(\gamma\bar{x} - 1)\{(\gamma\bar{x} - 1) + 4(\gamma + 1)(\gamma + 2)\} < \{2\bar{x}(\gamma + 1)(\gamma + 2) - (\gamma\bar{x} - 1)(2\gamma + 3)\}^2.$$

This expression, however, reduces to the condition that we must have

$$-2 - \gamma\bar{x} - 6\bar{x} - 6\gamma\bar{x}^2 - \gamma^2\bar{x} - 8\bar{x}^2 < 0,$$

which is always true. Hence

$$e^{\varphi_u(\gamma,\bar{x})} < \left(\frac{\gamma}{\gamma - 1}\right)^{\gamma\bar{x}},$$

and we have

$$I_u' < \bar{x}'. \tag{47}$$

2. **Contact at the ends of the range.** We shall now investigate whether the frequency function, $f(x')$, has high contact with the x' axis. The function, $f(x')$, vanishes at both limits, and thus it will be sufficient to test the derivatives. The nth derivative can be expressed in the form of $f(x')(x' \log x')^{-n}$ multiplied by an nth degree polynomial in $\log x'$. It can easily be shown that each derivative will vanish at the upper limit, while the nth derivative will vanish at the lower limit provided $\gamma\bar{x} > n$. Therefore, $f(x')$ does not have high contact at the lower end of the range.

3. **Moments.** The sth moment about the origin is given by

$$_{x'}\mu_s' = y_0 \int_1^\infty (\log x')^{\gamma\bar{x}-1} x'^{s-\gamma-1}\, dx'$$

$$= y_0 \int_0^\infty x^{\gamma\bar{x}-1} e^{-x(\gamma-s)}\, dx$$

$$= \left(\frac{\gamma}{\gamma - s}\right)^{\gamma\bar{x}}, \quad \text{if} \quad \gamma > s. \tag{48}$$

If $\gamma \leq s$, then taking k such that $\gamma > sk > 0$, we get in place of (48),

$$_{x'}\mu_s' = \left(\frac{\gamma}{\gamma - sk}\right)^{\gamma\bar{x}}. \tag{49}$$

We easily obtain the recurring relationship

$$(50)\qquad {}_{x'}\mu_s' = \left(\frac{\gamma - \overline{s-1}}{\gamma - s}\right)^{\gamma\bar{x}} {}_{x'}\mu_{s-1}'.$$

The sth moment of x' about the mean is

$$\begin{aligned} {}_{x'}\mu_s &= y_0 \int_1^\infty \left\{x' - \left(\frac{\gamma}{\gamma - 1}\right)^{\gamma\bar{x}}\right\}^s (\log x')^{\gamma\bar{x}-1} x'^{-(\gamma+1)}\, dx' \\ &= y_0 \int_0^\infty \left\{e^x - \left(\frac{\gamma}{\gamma - 1}\right)^{\gamma\bar{x}}\right\}^s x^{\gamma\bar{x}-1} e^{-\gamma x}\, dx \end{aligned}$$

$$(51)\qquad = \left(\frac{\gamma}{\gamma - s}\right)^{\gamma\bar{x}} \sum_{j=0}^{s} (-1)^j \binom{s}{j} \left(\frac{\gamma}{\gamma - 1}\right)^{j\gamma\bar{x}} \left(\frac{\gamma - s}{\gamma - \overline{s - j}}\right)^{\gamma\bar{x}}.$$

If we do not take the value of k to be 1, then

$$(52)\qquad {}_{x'}\mu_s = \left(\frac{\gamma}{\gamma - sk}\right)^{\gamma\bar{x}} \sum_{j=0}^{s} (-1)^j \binom{s}{j} \left(\frac{\gamma}{\gamma - k}\right)^{j\gamma\bar{x}} \left(\frac{\gamma - sk}{\gamma - \overline{s - j}k}\right)^{\gamma\bar{x}}.$$

IV. Transformation into a Normal Distribution

We shall now consider a unimodal probability function $y = f(x)$ with range, $a \leq x \leq b$, and shall seek to express x as such a function of t as will transform $y = f(x)$, into $y = Ce^{-\frac{t^2}{2}}$. For simplicity, we assume that $y = f(x)$ has its modal value at $x = 0$, and thus each of the curves $y = f(x)$ and $y = Ce^{-\frac{t^2}{2}}$ has its one maximum value at the origin.

In $y = f(x)$ let $\log y = V$, then equating densities,[12] we have,

$$V - \log C + \frac{t^2}{2} = 0.$$

Then

$$(53)\qquad \begin{aligned} \frac{dV}{dt} + t &= 0, \\ \frac{d^2V}{dt^2} + 1 &= 0, \\ \frac{d^nV}{dt^n} &= 0, \qquad n \geq 3. \end{aligned}$$

[12] If $f(x)$ is a probability function or density of a distribution, then $f(x)\, dx$ is, to within infinitesimals of higher order, the probability that a value taken at random will fall into the interval dx at x.

Under the assumption that x is a function which can be expanded in a Maclaurin series in powers of t, we shall use equations (53) to determine the values of $\left.\frac{d^n x}{dt^n}\right]_{t=0}$ in the series

$$x = A_0 + A_1 t + A_2 \frac{t^2}{2!} + \cdots \tag{54}$$

Let v_n represent $\left.\frac{d^n V}{dx^n}\right]_{t=0}$, $n = 0, 1, 2 \cdots$. Then $v_0 = \log C$, and $v_1 = 0$, since $\frac{dV}{dx}$ and $\frac{dy}{dt}$ vanish when $x = 0$, that is when $t = 0$, since A_0 is taken to be zero. Taking the second derivative,

$$\frac{d^2 V}{dt^2} = \frac{d^2 V}{dx^2}\left(\frac{dx}{dt}\right)^2 + \frac{d^2 x}{dt^2}\left(\frac{dV}{dx}\right) = -1,$$

and

$$\left.\frac{d^2 V}{dt^2}\right]_{t=0} = v_2 A_1^2 = -1.$$

Therefore we have

$$A_1 = (-v_2)^{-\frac{1}{2}}, \text{ when } v_2 < 0. \tag{55}$$

Also,

$$\left.\frac{d^3 V}{dt^3}\right]_{t=0} = v_3 A_1^3 + 3v_2 A_1 A_2 = 0,$$

and we have

$$A_2 = \frac{v_3}{3v_2^2}. \tag{56}$$

Similarly,

$$A_3 = \frac{5v_3^2 - 3v_4 v_2}{12 v_2^4}(-v_2)^{\frac{1}{2}},$$

$$A_4 = \frac{-(40v_3^3 - 45v_2 v_3 v_4 + 9v_2^2 v_5)}{45 v_2^5}, \tag{57}$$

$$A_5 = \frac{-\{385 v_3^4 - 630 v_2 v_3^2 v_4 + 21 v_2^2 (8 v_3 v_5 + 5 v_4^2) - 24 v_2^3 v_6\}(-v_2)^{\frac{1}{2}}}{144 v_2^7},$$

$$\vdots$$

Though the procedure is straightforward, the work becomes somewhat involved in determing A_n as n gets larger. For this reason, we proceed in the following manner. By the use of Burmann's[13] theorem we can write (54) in the form

$$x = \left\{\frac{x}{t}\right\}_{x=0} t + \left\{\frac{d}{dx}\left(\frac{x}{t}\right)^2\right\}_{x=0} \frac{t^2}{2!} + \left\{\frac{d^2}{dx^2}\left(\frac{x}{t}\right)^3\right\}_{x=0} \frac{t^3}{3!} + \cdots \tag{58}$$

But

$$\frac{x}{t} = \frac{x}{\sqrt{2}} (\log C - V)^{-\frac{1}{2}},$$

where V is a function of x. We have,

$$\log C = V\Big]_{x=0}, \qquad \frac{dV}{dx} = 0 \quad \text{at} \quad x = 0,$$

and we may write

$$V = \log C + \frac{d^2V}{dx^2}\Big]_{x=0} \frac{x^2}{2!} + \frac{d^3V}{dx^3}\Big]_{x=0} \frac{x^3}{3!} + \cdots$$

Hence,

$$\frac{x}{t} = \frac{x}{\sqrt{2}} (a_2 x^2 + a_3 x^3 + a_4 x^4 + \cdots)^{-\frac{1}{2}}$$

$$= \frac{1}{\sqrt{2}} (a_2 + a_3 x + a_4 x^2 + \cdots)^{-\frac{1}{2}}, \quad \text{where} \quad a_n = -\frac{v_n}{n!}.$$

We can now write,

$$\left(\frac{x}{t}\right)^n = \frac{1}{2^{\frac{n}{2}}} \left\{\sum_{s=0}^{\infty} a_{s+2} x^s\right\}^{-\frac{n}{2}}. \tag{59}$$

But

$$A_n = \frac{d^{n-1}}{dx^{n-1}}\left(\frac{x}{t}\right)^n\Big]_{x=0} = (n-1)! \text{ multiplied by the coefficient of } x^{n-1} \text{ in (59)}.$$

Hence,

$$A_n = (n-1)!\,(2a_2)^{-\frac{n}{2}} \sum \frac{\left(-\frac{n}{2}\right)\left(-\frac{n}{2}-1\right)\left(-\frac{n}{2}-2\right)\cdots\left(-\frac{n}{2}-p+1\right)}{\lambda_1!\lambda_2!\lambda_3!\lambda_n!} \left(\frac{a_3}{a_2}\right)^{\lambda_1}\left(\frac{a_4}{a_2}\right)^{\lambda_2}\left(\frac{a_5}{a_2}\right)^{\lambda_3}\cdots\left(\frac{a_{n+2}}{a_2}\right)^{\lambda_n} \tag{60}$$

where the summation is over all values of λ_s such that

$$\sum_{s=1}^{n} s\lambda_s = 0, \quad \text{and} \quad p = \sum_{s=1}^{n} \lambda_s.$$

[13] A. De Morgan, Differential and Integral Calculus, (1842) page 305.

This expression can be written in the form,

$$A_n = (n-1)!\,(2a_2)^{-\frac{n}{2}} \sum_{s=0}^{n-2} (-1)^{s+1} \frac{n}{2} \cdot \frac{n+2}{4} \cdot \frac{n+4}{6} \cdots \frac{n+2s}{2(s+1)} \frac{D^{n-s-2} a_3^{s+1}}{a_2^{s+1}}, \qquad n > 1, \tag{61}$$

where D is the derivative operator of Arbogast.[14]

If we take expression (1) as our function of x, we obtain,

$$v_n = (-1)^{n-1} \frac{(n-1)!\,\gamma^n}{(\gamma\bar{x}-1)^{n-1}},$$

which gives

$$x = \frac{(\gamma\bar{x}-1)}{\gamma} + \frac{(\gamma\bar{x}-1)^{\frac{1}{2}}}{\gamma} t + \frac{2}{3\gamma}\frac{t^2}{2!} + \frac{1}{6\gamma(\gamma\bar{x}-1)^{\frac{1}{2}}}\frac{t^3}{3!} + \frac{4}{45\gamma(\gamma\bar{x}-1)}\frac{t^4}{4!} + \frac{1}{36\gamma(\gamma\bar{x}-1)^{\frac{3}{2}}}\frac{t^5}{5!} + \cdots \tag{62}$$

where $A_n = (n-1)!\,\dfrac{(\gamma\bar{x}-1)^{\frac{n}{2}}}{2^{\frac{n}{2}}\gamma^n}$ multiplied by the coefficient of

$$\left[x - \left(\bar{x} - \frac{1}{\gamma}\right)\right]^{n-1} \quad \text{in} \quad \left\{\sum_{s=0}^{\infty} \frac{(-1)^s \gamma^s [x - (\bar{x} - 1/\gamma)]^s}{(s+2)(\gamma\bar{x}-1)^s}\right\}^{-\frac{n}{2}}$$

This series is known to diverge for large values of t. However, the series is defined for those values of t that correspond to x for the interval $0 < x < 2\left(\bar{x} - \frac{1}{\gamma}\right)$. With the aid of (22) and Salvosa's[15] tables we give in Table IV the percentage of the total population which is included in this interval.

TABLE IV

The percentage of the population, characterized by (1), which is included in the interval $0 < x < 2(\bar{x} - 1/\gamma)$, for different degrees of skewness

α_3	1.1	1.0	.9	.8	.7	.6	.5	.4	.3	.2	.1
Percent of Population	79.386	84.880	89.781	93.908	97.021	98.959	99.805	99.990	100.000	100.000	100.000

Thus, in dealing with samples as large as 10,000, with moderate degrees of skewness, the probability of getting a value that falls beyond this interval

[14] Augustus De Morgan, On Arbogast's Formulae of Expansion, Cambridge and Dublin Mathematical Journal, Vol. 1, (1848) pp. 238-255.

[15] Cf. Salvosa, loc. cit. page 2.

becomes negligible. Hence it may be expected that with the use of a comparatively few terms of series (62) we may transform the ordinates of a moderately skew Type III function to within close approximations of the ordinates of the normal function.

Baker[16] considered the transformations of a non-normal frequency distribution represented by $f(t)dt$, where the origin is taken at some central point and the scale is the standard deviation of the distribution. By equating probabilities he found a function φ, such that by setting $t = \varphi(u)$, he obtained

$$f[\varphi(u)]. \quad \varphi'(u)\, du = e^{-\frac{1}{2}u^2}\, du.$$

It seemed of interest to compare the results obtained by applying transformation (62), which is found by equating densities, to the illustration treated by Baker,[17] where the transformation giving equality of probabilities was used. The example treated was

$$f(t) = .9929\left(1 + \frac{t}{10}\right)^{99} e^{-10t}. \tag{63}$$

This is a Type III distribution of form (24), with $\alpha_3 = .2$. From (22), $\gamma\bar{x} = \frac{4}{\alpha_3^2} = 100$, and from (62) we obtain the series

$$x = \frac{99}{\gamma}(1 + .1005038u + .0033670u^2 + .0000282u^3 - .0000004u^4 + \cdots). \tag{64}$$

We shall utilize only the first four terms and rewrite (64) in the form

$$\gamma x = 99(1 + .1005038u + .0033670u^2 + .0000282u^3).$$

However, from (23),

$$x = \bar{x}\left(1 + \frac{\alpha_3 t}{2}\right),$$

which gives

$$\gamma x = \gamma\bar{x}\left(1 + \frac{\alpha_3 t}{2}\right) = 100(1 + .1t).$$

Therefore,

$$t = .1(\gamma x - 100).$$

---

[16] G. A. Baker, Transformation of Non-normal Frequency Distributions into Normal Distributions, Annals of Mathematical Statistics, Vol. 5, (1934) pp. 113–123.

[17] Baker, loc. cit. page 117.

With the aid of Salvosa's[18] tables, we obtain the following results.

TABLE V

Comparison of the ordinates of the normal function, the function with skewness .2, and the skewed function transformed by

$t = 9.9\ (-.0101010 + .1005038\ u + .0033670\ u^2 + .0000282\ u^3)$

u	Normal Curve	Function with skewness .2	Transformed skew curve	Transformed skew curve minus normal
−2.0	.053991	.049243	.054226	.000235
−1.8	.078950	.076810	.079291	.000341
−1.6	.110921	.112956	.111393	.000472
−1.4	.149727	.157043	.150359	.000632
−1.2	.194186	.206951	.195003	.000817
−1.0	.241971	.259120	.242986	.001015
− .8	.289692	.308958	.290905	.001213
− .6	.333225	.351538	.334618	.001393
− .4	.368270	.382453	.369811	.001541
− .2	.391043	.398583	.392678	.001635
0	.398942	.398610	.400615	.001673
.2	.391043	.383157	.392682	.001639
.4	.368270	.354545	.369811	.001541
.6	.333225	.316273	.334621	.001396
.8	.289692	.272360	.290905	.001213
1.0	.241971	.226714	.242984	.001013
1.2	.194186	.182641	.194999	.000823
1.4	.149727	.142563	.150353	.000626
1.6	.110921	.107939	.111383	.000462
1.8	.078950	.079354	.079277	.000327
2.0	.053991	.056702	.054214	.000223

The ordinates of the transformed distribution are more symmetrical and approximate the ordinates of the normal curve more closely than the values obtained by Baker even though we have used only four terms in the transforming series.

Returning to the general case, we may write

$$\begin{aligned}\int_a^b y\,dx &= \int_{-\infty}^{\infty} Ce^{-\frac{t^2}{2}} \frac{dx}{dt}\cdot dt \\ &= C\int_{-\infty}^{\infty} e^{-\frac{t^2}{2}}\left(A_1 + A_2 t + A_3 \frac{t^2}{2!} + A_4 \frac{t^3}{3!} + \cdots\right) dt,\end{aligned} \tag{65}$$

[18] Salvosa, loc. cit. pp. 64 et seq.

provided the series converges for all values of t. Under the assumption that the integrand satisfies conditions for the term by term integration of the series, we get

$$(66)\quad 1 = \int_a^b y\,dx = C\sqrt{2\pi}\left\{A_1 + \frac{A_3}{2!} + \frac{3A_5}{4!} + \cdots + \frac{A_{2n+1}}{2^n n!} + \cdots\right\}.$$

The area to the right of the modal ordinate is

$$\int_{x_{mo.}}^b y\,dx = C\int_0^\infty e^{-\frac{t^2}{2}}\left(A_1 + A_2 t + \frac{A_3 t^2}{2!} + \cdots\right)dt$$

$$= \tfrac{1}{2} + C\int_0^\infty e^{-\frac{t^2}{2}}\left(A_2 t + A_4\frac{t^3}{3!} + \cdots \frac{A_{2n}\,t^{2n-1}}{(2n-1)!} + \cdots\right)dt$$

$$(67)\quad = \tfrac{1}{2} + C\left(A_2 + \frac{A_4}{3} + \cdots\right).$$

Hence the area from the mode to the median is

$$(68)\quad C\left(A_2 + \frac{A_4}{3} + \cdots\right).$$

Let us consider distribution (1) again. The coefficients in series (62) are functions of the skewness, and become smaller with smaller degrees of skewness. Indications are that with moderate skewness, the series converges sufficiently to be used for certain formal purposes. If we assume this and proceed in a formal manner we obtain some interesting results that are consistent with approximations that have been obtained elsewhere.

Thus, it is interesting to note that using the coefficients of series (62) in equation (66), we obtain

$$\Gamma(\gamma\bar{x}) = \sqrt{2\pi(\gamma\bar{x}-1)}\,(\gamma\bar{x}-1)^{\gamma\bar{x}-1}e^{-(\gamma\bar{x}-1)}\left\{1 + \frac{1}{12(\gamma\bar{x}-1)} + \frac{1}{288(\gamma\bar{x}-1)^2} + \cdots\right\},$$

which is Stirling's asymptotic form for $\Gamma(\gamma\bar{x})$.[19]

From (68), the area from the mode to the median in the Type III distribution characterized by (1) is approximately

$$(69)\quad C\left(\frac{2}{3\gamma} - \frac{4}{135\gamma(\gamma\bar{x}-1)} + \cdots\right),$$

where

$$C = \frac{\gamma(\gamma\bar{x}-1)^{\gamma\bar{x}-1}e^{-(\gamma\bar{x}-1)}}{\Gamma(\gamma\bar{x})}.$$

[19] E. Czuber, Wahrscheinlichkeitsrechnung, Volume 1, (1908) pp. 23-24.

Since $(\gamma\bar{x} - 1)$ is large when the skewness is moderate, and since the terms of (69) are rapidly decreasing, the area from the mode to the median is approximately equal to $\frac{2C}{3\gamma}$. But $\frac{1}{\gamma}$ is the distance from the mode to the mean and C is the ordinate at the mode, hence the area from the mode to the median is approximately equal to the ordinate at the mode multiplied by 2/3 of the distance between the mode and mean. Therefore with moderate skewness the median is approximately 2/3 of the distance between the mode and mean, which conforms to the approximate result first obtained by Karl Pearson[20] for the Type III distribution. We may, for all cases resulting in (68), take A_2 as being approximately equal to the distance from the mode to the median. This becomes somewhat more apparent by finding the arithmetic mean of distribution y. Thus,

$$\mu_1' = \frac{\int_a^b xy\,dx}{\int_a^b y\,dx} = \frac{C\int_{-\infty}^{\infty} e^{-\frac{t^2}{2}}\left(A_0 + A_1 t + \frac{A_2 t^2}{2!} + A_3\frac{t^3}{3!} + \cdots\right)\left(A_1 + A_2 t + A_3\frac{t^2}{2!} + A_4\frac{t^3}{3!} + \cdots\right)dt}{C\sqrt{2\pi}\left(A_1 + \frac{A_3}{2!} + \frac{3A_5}{4!} + \cdots\right)}$$

$$= \frac{A_0A_1 + \left(\frac{3A_1A_2}{2!} + \frac{A_0A_3}{2!}\right) + 3\left(\frac{A_0A_5}{4!} + \frac{5A_1A_4}{4!} + \frac{5A_2A_3}{2!\,3!}\right) + \cdots}{A_1 + \frac{A_3}{2!} + \frac{3A_5}{4!} + \cdots}$$

$$= A_0 + \frac{3A_2}{2!} + \frac{15A_4}{4!} + \cdots. \tag{70}$$

Remembering that A_0 is the abscissa of the mode, it becomes apparent that the mean is, in general, approximately equal to the mode plus 3/2 of the distance from the mode to the median.

Though series (62) is known not to converge for large values of t, it is interesting to note that if we use distribution (1) for y, we have from (70)

$$\mu_1' = \left(\bar{x} - \frac{1}{\gamma}\right) + \frac{3}{2}\left(\frac{2}{3\gamma}\right) - \frac{1}{3\,\gamma(\gamma\bar{x} - 1)\,3!} + \cdots, \tag{71}$$

the first two terms of which give $\bar{x}$, which is μ_1', and hence if (71) were an exact formula, the sum of the terms beyond the second would be zero.

For example (63), it can be seen from the following that A_2 furnishes a close approximation to the distance from the mode to the median. Here, $t = .1\,(\gamma\bar{x} - 100)$; putting $x = \bar{x} - \frac{1}{\gamma}$ we have $t_{\text{mo.}} = .1(\gamma\bar{x} - 101) = -.1$. Putting

---

[20] Karl Pearson, loc. cit.

$x = \left(\bar{x} - \frac{1}{\gamma}\right) + \frac{2}{3\gamma}$, where $A_2 = \frac{2}{3\gamma}$, we have as our approximation to the median

$$t_{\text{md.}} = .1\left(\gamma\bar{x} - \frac{301}{3}\right)$$
$$= -.03333.$$

Interpolating in the Salvosa tables, we find for $\alpha_3 = .2$, $t_{\text{md.}} = -.03331$ approximately. Hence it is seen that the interpolated values checks very closely with that obtained by using the A_2 criterion.

We shall now consider briefly the transforming series, when for y, we take distribution (4). Then, corresponding to (54), we obtain the series

$$(72) \qquad x' = \frac{(\gamma\bar{x} - n)^n}{\gamma^n} + \frac{n(\gamma\bar{x} - n)^{n-\frac{1}{2}}}{\gamma^n} t + \frac{n(3n-1)(\gamma\bar{x} - n)^{n-1}}{3\gamma^n}\frac{t^2}{2!} + \frac{n(6n^2 - 6n + 1)(\gamma\bar{x} - n)^{n-\frac{3}{2}}}{6\gamma^n}\frac{t^3}{3!} + \frac{n(45n^3 - 90n^2 + 45n - 4)(\gamma\bar{x} - n)^{n-2}}{45\gamma^n}\frac{t^4}{4!} + \cdots.$$

When $n = 1$, (72) reduces to the series given by (62). Suppose we are primarily interested in the cases for which $0 < n < 1$. For these cases the coefficients of (72) decrease more rapidly than do those of series (62). Under the same assumptions as to convergence which were made in working with the latter series we have, from (68) and (72), the area from the mode to the median given approximately by

$$(73) \quad C\left\{\frac{n(3n-1)(\gamma\bar{x} - n)^{n-1}}{3\gamma^n} + \frac{n(45n^3 - 90n^2 + 45n - 4)(\gamma\bar{x} - n)^{n-2}}{135\gamma^n} + \cdots\right\}.$$

When $0 < n < 1$ we always have $\gamma\bar{x} > n$; then $A_2 > 0$ if $n > 1/3$, and $A_2 < 0$ if $n < 1/3$. Therefore, if A_2 is taken to be approximately equal to the distance from the mode to the median, we have $x'_{\text{mo.}} > x'_{\text{md.}}$ if $n < 1/3$, and $x'_{\text{mo.}} < x'_{\text{md.}}$ if $n > 1/3$, since A_2 is positive or negative according as n is greater or less than $1/3$. Combining these results with (70), we have $\bar{x}' > x'_{\text{md.}}$ if $x'_{\text{mo.}} < x'_{\text{md.}}$, and $\bar{x}' < x'_{\text{md.}}$ if $x'_{\text{mo.}} > x'_{\text{md.}}$, which are the relations given in Section IIa, for case II.