TRANSFORMATIONS OF BIMODAL DISTRIBUTIONS

By

G. A. BAKER

I. INTRODUCTION

Several men have concerned themselves extensively with the transformation of frequency distributions, for instance, Edgeworth, Kapteyn, Arne Fisher, and H. L. Reitz (see 1, bibliography). The first three of these men have been concerned with transformations as a means of extending the scope of the normal distribution and Gram-Charlier system as a method of description. Rietz has been more interested in the properties of the transformed distributions.

There are three types of transformations that are of particular importance:

(1) $u = x^n$ because it has a physical interpretation.

(2) $u = \log x$ because Arne Fisher and others find it useful.

(3) $u = e^x$ because it is the inverse of (2).

These three transformations will be discussed in some detail for bimodal frequency distributions. It is interesting to note that it is possible to transform a bimodal distribution into a unimodal distribution and vice versa by means of these transformations. The general scheme of the first part of the following is that of H. L. Heitz (see 1, bibliography).

The latter part of this paper consists of a few remarks on transformations in general.

Institute of Mathematical Statistics is collaborating with JSTOR to digitize, preserve, and extend access to *The Annals of Mathematical Statistics.*
www.jstor.org

II. THE TRANSFORMATION $u = x^n$

In the following theorems it will be understood that *one* means *at least one* and that a frequency function is to have a total area of unity.

The transformation $u = x^n$ has a very clear physical interpretation, for if the diameters of oranges are distributed as $f(x)$ then the distribution of the volumes of these oranges would be obtained by making the transformation $u = kx^3$.

Theorem I.

Given a continuous bimodal frequency function of positive variates $y = f(x)$ with a range $0 < a \leq x \leq e$ with modes at $x = b$, $x = d$ and antimode at $x = c$, $(a < b < c < d < e)$ $f(a) = f(e) = 0$ and with a continuous derivative, then the frequency distribution $v = \phi(u)$, $\left[\phi(u) \equiv \frac{1}{n} u^{\frac{1-n}{n}} f\left(u^{\frac{1}{n}}\right)\right]$ of positive variates, $u = x^n$ has modes as follows:

Case I. $n > 1$

(1) one mode $a^n < u \leq b^n$ always, and (2) one mode and one antimode $c^n < u \leq d^n$ if $\left|(1-n) f(u^{\frac{1}{n}})\right| < u^{\frac{1}{n}} f'(u^{\frac{1}{n}})$ somewhere in this interval.

Case II: $0 < n < 1$

(1) always one mode $b^n \leq u < c^n$, (2) a mode and antimode $d^n \leq u < e^n$ if $\left| u^{\frac{1}{n}} f'(u^{\frac{1}{n}}) \right| > (1-n) f(u^{\frac{1}{n}})$ somewhere in this interval.

Case III. $n < 0$

(1) One mode $d^n \leq u < e^n$ (2) one mode and one antimode $b^n \leq u < c^n$ if $\left| u^{\frac{1}{n}} f'(u^{\frac{1}{n}}) \right| > (1-n) f(u^{\frac{1}{n}})$ somewhere in this interval.

Proof:

Since $u^{\frac{1}{n}}$ is taken to be positive, then if $\frac{dv}{du}$ is to be zero we must have

$$(1) \qquad (1-n) f(u^{\frac{1}{n}}) + u^{\frac{1}{n}} f'(u^{\frac{1}{n}}) = 0$$

Also we have by hypothesis

$$(2) \qquad f(a) = f(e) = 0$$

$$(3) \qquad f'(b) = f'(c) = f'(d) = 0$$

and that $f'(x)$ is continuous.

From these considerations the proof of the theorem follows quite simply; for instance:

Case I. $n > 1$

In the interval $e^n > u \geqq d^n$ (1) is negative. At $u - c^n$ (1) is negative. In the interval $c^n < u < d^n$, $u^{1/n} f'(u^{1/n})$ is positive and hence from continuity there is a maximum and minimum or not according as $u^{\frac{1}{n}} f'(u^{\frac{1}{n}}) < |(1-n) f(u^{\frac{1}{n}})$ or not for every u in this interval. At a^n (1) is $u^{\frac{1}{n}} f'(a)$ which is zero or positive, while at b^n (1) is negative. If $f'(a)$ is positive there is clearly a maximum at the point where the sign of the continuous derivative changes from positive to negative in the interval $a^n < u \leq b^n$ If $f'(a)$ is zero it follows that there is also a maximum, since $v = 0$ at $u = a^n$ and then increases before decreasing at $u = b^n$

The other cases follow from exactly similar reasoning

Theorem II.

In case the bimodal continuous frequency function $y = f(x)$ (of Theorem I) is symmetrical about the antimodal line $x = c$, then the mean value of u in the frequency distribution $v = \phi(u)$ of $u = x^n (n \neq 0 \text{ nor } 1)$ is less or greater than its median value according as the value of n lies between 0 and 1 or outside of these bounds.

The first moment of the transformed distribution is given by $u = \frac{1}{n}\int_{a^n}^{e^n} u^{\frac{1}{n}} f(u^{\frac{1}{n}})\,du = \int_a^e x^n f(x)\,dx$ i. e., we have $\bar{\mu} = \mu'_n$ where μ'_n is the n th moment about the origin of the original frequency distribution $y = f(x)$. Denoting the mean value of x by $\bar{x}$, it is known[1] for every set of positive values that $u'_n < \bar{x}^n$, when u lies between 0 and 1, and that $\mu'_n > \bar{x}^n$ when u lies outside this interval.

Since $x = c$ when $y = f(x)$ is symmetrical about this line, the theorem follows. This follows Rietz exactly.

Theorem III.

In case the continuous bimodal frequency function $y = f(x)$ (of Theorem I) is symmetrical about the antimodal line $x = c$, the frequency distribution $v = \phi(u)$ of $u = x^n$ $(n \neq 0 \text{ nor } 1)$ has the following relations between its modes and its median.

Case I. $n > 1$

One mode $\leq$ median, in any case, and one greater if $|(1-n)f(u^{\frac{1}{n}})| < u^{\frac{1}{n}}f'(u^{\frac{1}{n}})$ somewhere in the interval $c^n < u \leq d^n$

Case II. $0 < n < 1$

One mode $\leq$ median, in any case, and one greater if

$$|u^{\frac{1}{n}}f'(u^{\frac{1}{n}}) > (1-n)f(u^{\frac{1}{n}})$$

at some point $d^n \leq u \leq e^n$

Case III. $n < 0$

One mode $\geq$ median, in any case, and one less if

$$|u^{\frac{1}{n}}f'(u^{\frac{1}{n}})| > (1-n)f(u^{\frac{1}{n}})$$

at some point $b^n \leq u \leq c^n$.

As an example of a transformation $u = x^n$ which transforms a bimodal frequency function satisfying the conditions of the previous theorems into a unimodal distribution consider the following.

Take $n = 37$ and $f(x) = -x^4 + 12x^3 - 50x^2 + 84x - 44$,
$0 < \alpha \leq x \leq \beta < 6$
If $\frac{dv}{du} = 0$, then

$$(1-n)f(u^{\frac{1}{n}}) + u^{\frac{1}{n}}f'(u^{\frac{1}{n}}) = 0 \tag{1}$$

---

[1] See J. L. W. V. Jensen, Acta Mathematica, Vol. 30. (1906), pp. 180-187.

Instead of the variable $u^{\frac{1}{n}}$ we may just as well write an x. Hence (1) becomes

$$(2) \qquad F(x) = 32x^4 - 396x^3 + 1700x^2 - 2940x + 1984$$

Calculating Sturm's functions for (2), it is easily seen that the transformed distribution has only one mode and that in the interval $0 < u < 2^{37}$

III. TRANSFORMATION $u = \log x$

Theorem IV.

Given a continuous bimodal frequency function (of Theorem 1) with a range $1 < a \leq x \leq e$, then the frequency distribution $v = \phi(u)$ $[\phi(u) \equiv e^u f(e^u)]$ of positive variates $u = \log x$ has one mode, in any case, $\log d \leq u < \log e$ and has a mode and antimode in the interval $\log b \leq u < \log c$ if $|e^u f'(e^u)| > f(e^u)$ somewhere in this interval.

This follows very simply from considering $\frac{dv}{du}$, which is

$$e^{2u} f'(e^u) + e^u f(e^u)$$

Theorems similar to those stated under the transformation $u = x^n$ concerning the relative position of the modes and median of the transformed distribution may be stated here.

As an example of a bimodal frequency distribution satisfying our hypothesis and which is transformed into a unimodal distribution by the transformation $u = \log x$ consider

$$f(x) = -x^4 + 16x^3 - 92x^2 + 224x - 148 , \quad 1 < \alpha \leq x \leq \beta < 7$$

The condition for the vanishing of the derivative of the transformed distribution takes the form

$$F(x) = -5x^4 + 64x^3 - 276x^2 + 448x - 148$$

By calculating Sturm's functions for $F(x)$ it is easily seen that the transformed distribution has but one mode and that in the interval $\log 6 < u < \log 7$.

IV. TRANSFORMATION $u = e^x$

Theorem V.

Given a continuous bimodal frequency function (of Theorem 1), then the frequency distribution $v = \phi(u)$, $\left[\phi(u) = \frac{1}{u} f(\log u)\right]$ of positive variates $u = e^x$ has one mode $e^a < u \le e^b$, and has a mode and antimode in the interval $e^c < u \le e^d$ if $f'(\log u) = f(\log u)$ at some point in this interval.

For $$\frac{dv}{du} = \frac{1}{u^2}\left[f'(\log u) - f(\log u)\right]$$

from this the theorem follows.

Theorems similar to those stated concerning the relative positions of the median and modes of the transformed distribution in the case of the transformation $u = x^n$ may be stated here also.

As an example of a bimodal frequency distribution that satisfies our hypothesis and is transformed into a unimodal distribution by the transformation $u = e^x$ consider

$$f(x) = -x^4 + 16x^3 - 92x^2 + 224x - 148, \quad 1 < \alpha \le x \le \beta < 7$$

The condition that the derivative of the transformed distribution vanish takes the form

$$F(x) = x^4 - 20x^3 + 140x^2 - 408x + 372$$

By calculating Sturm's functions for $F(x)$ it is easily seen that the transformed distribution has only one mode and that in the interval $e^1 < u < e^3$.

V. TRANSFORMATIONS IN GENERAL

Suppose that we have a frequency distribution the distribution of whose parameters due to random sampling we know. If we transform this distribution what will happen to the distributions of the estimates of the parameters? It appears that, in view of the fact that bimodal and possibly multimodal distributions may be transformed by fairly simple transformations into unimodal distributions, there will be no simple relation between the change in the frequency distribution and the corresponding changes in the distributions of the estimates of the parameters by means of random samples. As a specific example of these general remarks consider the following.

If the normal curve

(1) $$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$$

is transformed by the transformation

(2) $$u = x^2$$

giving

(3) $$F(u) = \frac{1}{2\sqrt{2u}} \quad \frac{e^{-\frac{1}{2}u}}{u}$$

Then, applying a general method for finding the distribution of the means of samples, first developed by J. O. Irwin (2, bibliography), the mean values of the u's are found to be distributed as proportional to

(4) $$\frac{\left(\frac{n}{2}\right)^{\frac{n}{2}} x^{\frac{n-2}{2}} e^{-\frac{1}{2}nx}}{\Gamma\left(\frac{n}{2}\right)}$$

Then

(5) $$u'_m = \frac{(n+2) \text{ ---- } (n+2m-2)}{n^{m-1}}$$

If $m \geq 1$

Whence $\mu_0 = 1$ $\qquad \beta_1 = \frac{8}{n}$

$\mu_1 = 1$ $\qquad \beta_2 = 3 + \frac{12}{n}$

$\mu_2 = \frac{2}{n}$ $\qquad \underset{n \to \infty}{L} \beta_1 = 0$

$\mu_3 = \frac{8}{n^2}$ $\qquad \underset{n \to \infty}{L} \beta_2 = 3$

$\mu_4 = \frac{12n+48}{n^3}$

Thus we see that, although the sampled population is J-shaped, the distribution of the estimates of the means ultimately approaches the normal distribution but that this approach is rather slow.

It has been shown (see 3) that (4) is also the distribution of the estimates of the second moment by means of samples of n drawn from (1), the second moment of the sample being taken about the mean of (1). This is a special example of a general consideration that is of considerable interest in this connection.

It has been shown (see 3) that, formally, the distribution of the estimates by means of samples of n of the mth moment of a population represented by $f(x), a \leq x \leq b$, the m th moment of the sample being taken about the mean of $f(x)$ is given by the solution of the integral equation

$$(6) \qquad F(z) = \int_{\alpha}^{\beta} \psi(x)\, e^{zx}\, dx$$

where $\psi(x)$ is the unknown distribution of n times the estimates of the m th moment of the population about the mean of the population and

$$F(z) = \left(\int_{a}^{b} f(x)\, e^{zx^m}\, dx\right)^n$$

and if m is even

$$\alpha = 0$$

$$\beta = \text{larger of } na^m,\ nb^m$$

if m is odd

$$\alpha = na^m$$
$$\beta = nb^m$$

Now, the formal development for finding the distribution of the means of samples of n drawn from a population represented by $f(x)$ transformed by the transformation $u = x^m$ leads to a relation equivalent to (6) (see 2). This result may be stated as

Theorem VI.[1]

If the distribution of the estimates of n times the m th moment of a population represented by $f(x)$, $a \leq x \leq b$ about the mean of $f(x)$ exists as a solution of (6) it is identical with the distribution of the estimates by means of samples of n of n times the mean, measured from the mean of $f(x)$, of the population represented by $f(x)$ transformed by the transformation $u = x^m$.

---

[1] This theorem permits of an obvious generalization to the case of the k th moment of the transformed distribution.

This enables us to formally identify these two problems so that anything that is true of one distribution is also true of the other.

With other transformations the relation between the distribution of the means of random samples from the transformed distribution and the distribution of the estimates of the parameters of the original distribution become much more complicated.

Further, we might say a few words with regard to the possibility of transforming various types of distributions into various other types.

Suppose that $f(x)$ is a continuous frequency function of positive variates, $a \leq x \leq b$, and that $f'(x)$ is continuous in this closed interval. Now make the transformation

(7) $u = \phi(x)$

and suppose that $\phi(x)$ is such that (7) can be solved explicitly for x, i. e.,

(8) $x = \psi(u)$

Then $f(x)\,dx$ becomes, assuming $\psi'(u)$ is continuous,

$$f[\psi(u)]\,\psi'(u)\,du$$

(9) , $U(u) = f[\psi(u)]\,\psi'(u)$

Supposing that f is known, what can we do towards fixing the form of U by a suitable choice of ψ ?

Now, the simplest of all possible frequency distributions, from the standpoint of description by means of a continuous function, is one in which the probabilities of all values of the variate are equal. Hence we will suppose for illustration that

$$(10)\qquad U(u) = f[\psi(u)]\,\psi'(u) = k$$

whence, putting $\psi(u) = y$

we have

$$(11)\qquad \int f(y)\,dy = ku + c$$

Suppose that

$$(12)\qquad f(x) = \alpha x + \beta$$

Then

$$(13)\qquad \psi(u) = \frac{-\beta \pm \sqrt{\beta^2 + 2\alpha(ku + c)}}{\alpha}$$

From this it is apparent that if $f(x)$ is any polynomial whose degree is less than four and which is positive $a \leq x \leq b$ may, conceivably, be transformed into a rectangular distribution.

If in place of k we were to put a specified function, say the normal function, we would run into considerable difficulty.

In (9) we may regard ψ as known and then ask what forms of f may be transformed into certain specified forms. For instance, let us take

$$u = \log x$$
$$x = e^u$$

Then $U(u) = f(e^u)\,e^u$

$$(14)\qquad U'(u) = e^u\left[f(e^u) + e^u f'(e^u)\right]$$

Now, since $u>0$, it is apparent that if $f(x)=c$ that (14) has no zero.

Let us put, for illustration, $U'(u)=0$ or $U(u)=k$

Then $f(x)=\frac{k}{x}$

However, if we were to suppose that (14) vanished at only one point. at exactly two points, etc., instead of identically it would be very difficult to express this in terms of the form of

VI. SUMMARY

It has been shown that unimodal distributions may be transformed into bimodal distributions by means of rather simple transformations. This suggests that bimodal distributions are not necessarily the result of heterogeneity.

The fact that a badly misshapen distribution may be transformed into something that is approximately normal does not seem to be of much aid in determining the distribution of the estimates of the constants of the original distribution.

The problem of transforming a specified distribution into another specified distribution is very difficult in gneral but could, perhaps, be handled to an adequate degree of approximation in special cases.

G. A. Baker

BIBLIOGRAPHY

1. Rietz, H. L. On Certain Properties of Frequency Distributions of the Powers and Roots of the Variates of a Given Distribution. *Proceedings of the National Academy of Science,* vol. 13, No. 12 (Dec. 1927), pp. 817-820.

2. Irwin, J. O., M.A., M.S. On the Frequency Distribution of the Means of Samples from a Population Having Any Law of Frequency with Finite Moments, with Special Reference to Pearson's Type II. *Biometrika,* vol. 19 (1927), pp. 225.

3. Baker, G. A. Random Sampling from Non-Homogeneous Populations. *Metron,* vol. VIII, Part 3 (Feb. 28, 1930), page 1.