Int. Stat. Rev., Vol. 41, No. 2, 1973, pp. 203–223/*Longman Group Ltd/Printed in Great Britain*

Transformations – An Introduction and a Bibliography

M. H. Hoyle
Peat, Marwick, Mitchell & Co., Management Consultants

Summary

This paper reviews the reasons for transforming data, the ways of developing a suitable transformation and the various transformations that are to be found in the literature. It also discusses the various problems involved in presenting results in the *original* variable when the analysis has been performed on the transformed variable.

1. Reasons for using Transformations

The standard statistical techniques associated with the linear model have been developed with the following basic assumptions:

(*a*) additivity – that is to say the main effects combine linearly to "explain" the observations,

(*b*) constant variance – that is the observations are assumed to have a constant variance about their varying means. Explicitly this means that the variance is independent of both the expected value of the observations and the sample size,

(*c*) normality – that is to say the observations are assumed to have a normal distribution.

The first assumption is important in the interpretation of the data; it is not an assumption which is always necessary for estimation or testing of hypotheses though in many circumstances it will be necessary to ensure identification of the parameters. The assumption of constant variance is usually made because it simplifies the estimation techniques. With it, Least Squares Estimators are also Minimum Variance Unbiased Linear Estimators (LSE = MVULE). Without it, a weighted least squares analysis gives the MVULE's (Kendall and Stuart, 1961). The third assumption is critically important in the testing of hypotheses, for the normality of the observations leads to comparatively simple and standard testing procedures which have been thoroughly investigated and, more importantly, leads to distributions which have been tabulated.

There are many cases in practice where these assumptions are known, or believed, not to hold. The statistician then has the option of either developing new theory or of bending the situation so that one or more of the assumptions are met, or nearly met. Transforming the data in one way or another can result in one or more of the assumptions being tenable.

Kruskal (1968) ranks the order of importance of the assumptions as:

1. Additivity.
2. Constant variance (homogeneity).
3. Normality.

Tukey (1958) has suggested that the same data may be analysed several times over using different transformations depending on the purpose of the particular stage in the analysis of the data. Other authors, Box and Cox (1964) and Draper and Hunter (1969), have indicated methods of finding transformations which simultaneously satisfy, or nearly satisfy, the three assumptions. This can usually be done so long as one does not press too hard for any one of them (Kendall and Stuart, 1966). The robustness of the usual linear model procedures means that only a little is lost if the assumptions are not exactly met (Kendall and Stuart, 1966).

2. Notation

It is convenient at this early stage to give the notation that will be used throughout the paper. The data or variable as it arises naturally will be denoted by x, with $E(x) = \theta$ and $V(x) = \phi^2$. ($E(.)$ is the expected value and $V(.)$ the variance.) The transformed data will be represented by ξ, with $E(\xi) = \mu$ and $V(\xi) = \sigma^2$. It will usually, but not always, be assumed that ξ has a normal distribution.

The relationship between x and ξ can be expressed in two ways:

the transformation function $g(.)$, with $\xi = g(x)$,
the inverse function, $f(.)$, with $x = f(\xi)$,

where $g(.) = f^{-1}(.)$. In some cases it will be necessary to make f or g depend explicitly on a parameter, w. For example, the transformation, $g_w(x) = x^w$. If more than one parameter is required, subscripts will be used. In the context of the linear model, either x or ξ may be the dependent variable, but in either case z, if necessary subscripted, will denote the independent variable(s).

It must be pointed out that for some $g(.)$ the corresponding $f(.)$ may not be known exactly or even up to a scalar multiple. Consider, for instance, the practice of transforming the original data into normal scores. In this case whilst the transformation carrying x into ξ is known, it is impossible to recover x given ξ.

Sample estimators of μ and σ^2 will be based on a sample of n independent observations, $\xi_1, \xi_2, \ldots, \xi_n$, typically identically and normally distributed. In this case

$$E^{-1}(\mu) = \hat{\mu} \sim N(\mu, \lambda^2\sigma^2)$$

and

$$E^{-1}(\sigma^2) = \hat{\sigma}^2 = S/v,$$

where $S/\sigma^2 \sim \chi^2(v)$. Here $E^{-1}(.)$ is the linear operator "the minimum variance unbiased estimator of". Let $\boldsymbol{\theta} \in \Theta$ be a parameter in the general parametric space Θ and let $\boldsymbol{t}$ be the MVUE of $\boldsymbol{\theta}$, then the linear operator E^{-1} is defined by $\boldsymbol{t} = E^{-1}(\boldsymbol{\theta})$. λ^2 defines the variance of μ as a multiple of the variance of the random variable ξ, i.e. $V(\hat{\mu}) = \lambda^2\sigma^2$.

3. Developing the Transformation

This section is concerned with the choice of the function $g(.)$, or equivalently $f(.)$. What methods are available, given a sample of data or a theoretical situation, to lead the statistician to the most suitable transformation? In other words, given some knowledge about x (its distribution, a sample of values $x_1, \ldots, x_n$ etc.) how can the statistician find $\xi = g(x)$ such that one or more of the three basic assumptions is more tenable?

3.1 *Constant variance*

Suppose x is such that its variance, σ^2, is a function of its mean, μ. Specifically let $\phi^2 = D^2(\theta)$ and $g(x)$ is to be chosen so that $V(\xi) = c$, say. In other words, the dependence of the variance on the mean is to be removed. In general, however, this is impossible and one is satisfied with

$$V(\xi) = V(g(x)) = c\{1 + O(R^{-1})\},$$

where R is some known constant which is large enough for R^{-1} to be negligible. By a well-known approximate argument, e.g. Kendall and Stuart (1966), it can be shown that $g(x)$ should be chosen so that

$$g(x) \propto \left\{\int \frac{d\theta}{D(\theta)}\right\}_{\theta = x}.$$

Although this result is approximate, its validity or otherwise can be tested if the theoretical distribution of x is known. If $V(\xi)$ is not quite constant then it may be possible to improve on $g(x)$ by further analytical work, e.g. Anscombe (1948). Alternatively, if an analytical expression is available to describe the non-constancy of $V(\xi)$ then Hotelling (1953) points out that the variance stabilizing technique may be applied a second time or more.

Where, on the other hand, only a sample is available and there is no knowledge of the distribution of x, then even the function $D^2(\theta)$ is not known. In such cases, the mean and variance of x in separate groups of observations can be calculated and an empirical relationship derived. This is discussed in Fisher and Mather (1943) and Quenouille (1950). Using this method, the approximation is more hazardous than that derived analytically but nevertheless often gives satisfactory results. In either case the real justification for the transformation derived is that, as Tippett (1934) says, it works.

Section 4 will illustrate the use that has been made of this approach.

3.2. *Additivity*

Tukey (1949) and Moore and Tukey (1954) derived a simple test, based on the residuals calculated from a standard analysis of the original data, showing whether or not the hypothesis of additivity is tenable. The latter paper, and this work has subsequently been developed by Anscombe (1955, 1961), Elston (1961), and Anscombe and Tukey (1963), showed how an estimate could be made of the transformation necessary to restore additivity to the data. The class of transformations considered was restricted to the class $\xi = x^w$ so that these papers were concerned with estimating w.

3.3. *Cornish-Fisher Expansions*

Several important distributions occurring in statistics depend on some variable n in such a way that as n tends to infinity the distribution tends to normality. For large n this is often an acceptable approximation, but for smaller n, Cornish and Fisher (1937, 1960) have developed the transformation

$$\xi = b_0 + b_1 x + b_2 x^2 + b_3 x^3 + \dots,$$

where the b's are of order $n^{-\frac{1}{2}}$ or smaller and are functions of the cumulants of the original distribution.

The non-normality of x usually shows itself most clearly in x having a skewed distribution. If $\xi = g_w(x)$ is the class of transformations under consideration, w is often chosen to minimize a measure of skewness, e.g. the third standardized moment of ξ.

Curtis (1943) gives a careful mathematical discussion of the limiting normality of many of the standard transformations to be described in section 4.

3.4. *Johnson's Transformations*

Johnson (1949) considered the general family of transformations

$$\xi = \gamma + \delta g\left(\frac{x-\beta}{\eta}\right),$$

where γ, δ, β, η are parameters at choice and g is some convenient function. For practical purposes it is desirable that $g(.)$ should not depend on any other parameters and that it should be a monotonic function. The parameters and the function g must be chosen so that ξ is normal or approximately so.

Johnson examines three types of system:

(i) The lognormal system: $g(x) = \log(x)$

(ii) The S_B system: $g(x) = \log\left\{\frac{x}{1-x}\right\}$

(iii) The S_U system: $g(x) = \sinh^{-1}(x)$.

Other systems are of course possible but the variety of shapes given by these three types is very great and does, for example, cover the shapes given by the Pearson system of curves.

3.5. *Box and Cox*

If the statistician is prepared to restrict his attention to the power class of transformation functions

$$g_w(x) = x^w,$$

then Box and Cox (1964) have developed a powerful method of estimating w in order that ξ is simultaneously normal, of constant variance and is explained additively by the appropriate treatment effects. Moreover, in their paper they develop a method of testing the hypotheses of normality, homogeneity and additivity.

The Box-Cox method does, however, require sophisticated computer calculations and Draper and Hunter (1969) have suggested a simpler approach. It is based on plotting against w simple statistics, e.g. Levene (1960), calculated from the standard analysis of the transformed data for differing values of w. Turning points in the graphs indicate appropriate values of w.

3.6. *Other Approaches*

Blom (1954) derives a transformation of the binomial distribution that gives the Cornish-Fisher expansion the property that the term measuring "skewness" is as small as possible. The application of this principle leads to a differential equation which contains the desired transformation as a solution. He subsequently shows that all the transformations used hitherto in the literature for stabilizing the binomial and related variables can be developed from this common viewpoint. He thus confirms theoretically the observed result that transformations inducing constant variance also commonly produce normality. Kendall and Stuart (1966) point out that they do not produce optimum normalization. There is always a need to compromise on the conflicting requirements of the three assumptions.

4. The Transformations

The literature indicates 19 transformations in general use. Some are special cases of each other but they are quoted individually because they have a separate and important identity. All are used to make one or more of the three assumptions more tenable.

4.1. *Power Transformation*

$$\begin{aligned} \xi &= (x+w_1)^{w_2} & \quad w_2 \neq 0 \\ &= \log(x+w_1) & \quad w_2 = 0. \end{aligned}$$

To avoid the discontinuity at $w_2 = 0$, this may equivalently be written as

$$\begin{aligned} \xi &= \{(x+w_1)^{w_2} - 1\}/w_2 & \quad w_2 \neq 0 \\ &= \log(x+w_1) & \quad w_2 = 0. \end{aligned}$$

Tukey (1957) and Dolby (1963) have studied the characteristics of the family of transformations generated by this equation. Moore (1957) has considered the necessary values of w_1 and w_2

required to induce normality in ξ and the transformation is the starting-point of the methods of Box and Cox (1964). Turner, Monroe and Lucas (1961) discuss the maximum likelihood estimation of the parameters and Healy and Taylor (1962) give a table to facilitate fractional power transformations when $w_1 = 0$ and w_2 is a multiple of 0·2.

Transformations 4.2 to 4.5 are special cases of this family of transformations.

4.2 *Reciprocal Transformation*

$$\xi = x^{-1}.$$

This transformation is often an entirely empirical one based upon reducing extreme skewness. It sometimes turns out that x^{-1}, and not x itself, more nearly satisfies the three basic assumptions. It has almost always been used where $\xi = x^{-1}$ has a definite physical meaning and where the probability of the random variable being less than or equal to zero is negligible. For example, data on the failures of a machine may be collected as either "intervals between failures" or "the number of failures per unit time". One of these may not be approximately normal but the other, its reciprocal, may well be.

It can be seen from the results given in section 3.1 that if

$$V(x) = D^2(\theta) \propto \theta^4$$

then

$$g(x) \propto \left\{\int \frac{d\theta}{\theta^2}\right\}_{\theta = x} \propto \frac{1}{x}$$

is the appropriate variance stabilising transformation.

4.3. *Logarithmic Transformation*

$$\xi = \log_e(x+w).$$

This transformation arises in three ways. In the first place there is a specific distribution, called the lognormal distribution, which represents remarkably well, for instance, the distribution of income over the population of a country. If ξ is normal, then x is said to be lognormally distributed. Aitchison and Brown (1957) present a full account of its uses and of methods available to estimate the parameters of the lognormal distribution.

Other authors have empirically suggested using the logarithmic transformation as a means of making the data at hand conform more nearly to the three assumptions, for example Bartlett and Kendall (1946), Anscombe (1948), Kleczowski (1949), Moore (1958) and Healy (1967). It is one way of reducing the skewness of a distribution but it is less severe than that of the reciprocal transformation.

The third way in which the logarithmic transformation arises is in stabilising the variance. If

$$V(x) = D^2(\theta) \propto \theta^2$$

or equivalently the standard deviation of x is proportional to θ, then $g(x) = \log(x)$ is the appropriate variance stabilizing transformation.

4.4 *Square Root Transformation*

$$\xi = (x+w)^{\frac{1}{2}}$$

This transformation is used where x has a Poisson distribution, with mean θ, and it arises as a result of finding, in the standard way, a transformation which makes $V(\xi)$ independent of θ. It appears to have been used in the early 1930s, see for example Tippett (1934), with $w = 0$. Bartlett (1936) considered the problem in more detail and suggested putting $w = \frac{1}{2}$. However, Anscombe (1948), using results by Johnson, showed that $w = \frac{3}{8}$ was optimal for variance-stabilization. In this case

$$V(\xi) \sim \frac{1}{4}\left(1+\frac{1}{16\theta^2}\right)$$

$$\gamma_1 \sim -\frac{1}{2\sqrt{\theta}}\left(1+\frac{7}{16\theta}\right)$$

$$\gamma_2 \sim \frac{1}{\theta}\left(1+\frac{369}{256\theta}\right)$$

where γ_1 and γ_2 are the standard third and fourth moments of ξ, measuring respectively skewness and kurtosis. These compare with $\gamma_1 = \sqrt{\theta}$ and $\gamma_2 = \theta^{-1}$ for the original Poisson variable, x.

In more recent work, Kihlberg, Herson and Schotz (1967) have concluded from an extensive computer study that $w = 0{\cdot}386$ is optimal for most values of θ, though for small θ, $\theta < 2$, $w = \frac{1}{4}$ is better. Indeed for small θ it is difficult to choose a single value of w to stabilize the variance of ξ at $\frac{1}{4}$ and Freeman and Tukey (1950) were led to the related compound transformation

$$\xi = \sqrt{x}+\sqrt{x+1}.$$

This transformation, sometimes known as the *chordal* transformation, more nearly stabilizes the variance ξ for all values of θ.

The square root transformation, $\xi = \sqrt{2x}$, has also been used where x has a χ^2 distribution with v degrees of freedom. Fisher (1925) showed that ξ was approximately distributed as $N(\sqrt{2v-1}, 1)$ and that this provided one way of extending the published χ^2 tables.

In another application of this transformation, Smith, Adderley and Bethwaite (1965) found it could be usefully employed in analysing rainfall data.

4.5. *Cube Root Transformation*

$$\xi = x^{\frac{1}{3}}.$$

Fisher's approximation to the χ^2 distribution using the square root transformation is not entirely satisfactory and Wilson and Hilferty (1931) developed the result that $\xi = \left(\frac{x}{v}\right)^{\frac{1}{3}}$ is approximately $N\left(1-\frac{2}{9v}, \frac{2}{9v}\right)$, where $x \sim \chi^2(v)$. Haldane (1937) refined this approximation but Garwood (1936) and Merrington (1941) showed empirically that the Wilson-Hilferty approximation was adequate for almost all practical purposes and better than Fisher's square root transformation.

The square root and cube root transformations have a further use in common as Howell (1965) found that the cube root transformation can also be usefully employed in analysing rainfall data.

4.6. *The ar tanh Transformation*

$$\xi = \text{ar tanh}\,(x).$$

The distribution of the product moment correlation coefficient, r, from a bivariate normal distribution is not easily tabulated and Fisher (1920) suggested, and in Fisher (1921) demonstrated, that $\xi = \text{ar tanh}\,(r)$ is approximately $N\left(\text{ar tanh}\,(\rho), \frac{1}{n-3}\right)$, where ρ is the population correlation coefficient. It can be derived by applying the variance stabilizing technique of section 3.1 to the approximate variance of r, $V(r) \doteq \frac{(1-\rho^2)^2}{n}$. Hotelling (1953) discusses this transformation in detail and derives a more refined transformation using the techniques of section 3.1 a second time.

Quenouille (1948) has suggested using the ar tanh transformation when r is a serial correlation, but on investigation found it unsatisfactory. Kendall, in the discussion of Williams (1967), raised the possibility of applying the ar tanh transformation to canonical correlations.

4.7. *Arc sin Transformation*

$$\xi = \arcsin(x).$$

This transformation must not be confused with what in section 4.8 is called the angular transformation, $\xi = \arcsin(\sqrt{x})$. The names "arc sin" and "angular" have been used inconsistently in the literature.

The arc sin transformation was first suggested by Fisher (1922) and again in Fisher (1930), to stabilize the variance of a binomial variate. Unfortunately, whilst the variance of ξ, being approximately equal to $\frac{1}{2n}$, is almost independent of π, the probability of success, it still depends on the sample size and for binomial variates the angular transformation is more suitable.

Two further uses of the arc sin transformation have, however, been suggested. Jenkins (1954) showed that if r is a serial correlation coefficient then the transformation to stabilize the variance is $\xi = \arcsin(r)$. It also has a limiting normal distribution and although $E(\xi)$ is still sensitive to ρ as $\rho \to 1$, this sensitivity is not so great as that of the ar tanh transformation suggested by Quenouille (1948). Secondly, some theoretical work by Hotelling (1953), Harley (1956, 1957) and Daniels and Kendall (1958) has considered the problem of the existence of a function $g(r)$ such that $E(g(r)) = g(\rho)$, where r and ρ are respectively the ordinary (i.e. not serial) estimated and population correlation coefficients. It turns out that $g(r) = \arcsin(r)$ is the unique function satisfying this condition. Sankaran (1958) has suggested a more complicated version of the arc sin transformation for use with the ordinary correlation coefficient.

4.8. *Angular Transformation*

$$\xi = \arcsin \sqrt{\left(\frac{x+w_1}{n+w_2}\right)}.$$

The name of this transformation has often been confused with what has been defined in section 4.7 as the arc sin transformation.

The angular transformation arises out of attempts to stabilize the variance of a binomial variate, x, the number of successes. In this case, the standard method of section 3.1 gives $w_1 = w_2 = 0$. Fisher (1954) has questioned this approach, preferring additivity as a criterion.

The history, nature and effectiveness of the transformation is fully discussed in Eisenhart, Hastay and Wallis (1947) who consider refinements of it. Early users of the transformation had put $w_1 = w_2 = 0$, but Bartlett (1936) suggested that putting $w_1 = \frac{1}{2}$ and $w_2 = 0$ gave ξ a more stable variance. Anscombe (1948) subsequently suggested the slightly different transformation

$$\xi = \arcsin \sqrt{\left(\frac{x+\frac{3}{8}}{n+\frac{3}{4}}\right)}$$

In this case

$$V(\xi) = \tfrac{1}{4} + O\left(\frac{1}{n^2}\right)$$

$$\gamma_1 \sim \frac{2\theta - n}{2\{n\theta(n-\theta)\}^{\frac{1}{2}}}$$

$$\gamma_2 \sim \frac{n^2 - 2\theta(n-\theta)}{n\theta(n-\theta)}.$$

A minor point requiring attention is when $x = 0$ or n. Bartlett (1947) and Eisenhart *et al.* (1947) suggested, and Ghurye (1949) confirmed, the use of

$$\begin{aligned} \xi &= \arcsin \sqrt{\frac{1}{4n}} & x &= 0 \\ &= 90° - \arcsin \sqrt{\frac{1}{4n}} & x &= n. \end{aligned}$$

Mosteller and Tukey (1949), following Fisher and Mather (1943), have suggested graphical techniques to facilitate the use of the angular transformation.

4.9. *Hyperbolic Transformation*

$$\xi = \operatorname{ar\,sinh} \sqrt{\left(\frac{x+w_1}{k+w_2}\right)}.$$

This transformation is used to stabilize the variance when x follows the negative binomial distribution

$$\text{Prob}\,(X = x) = \frac{\Gamma(x+k)}{x!\,\Gamma(k)}\left(\frac{\theta}{\theta+k}\right)^x \qquad x = 0, 1, 2, \ldots.$$

Beall (1942) appears to have been the first to have considered this transformation with $w_1 = w_2 = 0$. It does in fact follow directly from the use of the variance stabilizing techniques. Anscombe (1948) considered the transformation again and suggested putting $w_1 = \frac{3}{8}$ and $w_2 = -\frac{3}{4}$. In this case, for large θ and constant k/θ, $V(\xi) = \frac{1}{4} + O\left(\frac{1}{\theta^2}\right)$. He also pointed out that $\xi = \log(x + \frac{1}{2}k)$ is a simpler but none the less good transformation.

4.10. *Log log Transformation*

$$\xi = \log(-\log x).$$

In many physical and biological phenomena the probability of a change can be expressed as

$$p = 1 - e^{-\beta x}.$$

In such cases Finney (1951) and Fisher and Yates (1957) have shown that the log log transformation is of value. It has also been considered by Yates (1955) and Naylor (1964) who show, as does Kruskal (1968), that it is very similar in effect to other transformations, particularly the angular and probit transformations. This transformation has been used in the statistics of extremes where the asymptotic cumulative distribution of the smallest value is

$$F(x) = 1 - \exp(-e^{-y})$$

with

$$y = \alpha_n (x - x_{(n)})$$

Here $x_{(n)}$ is the nth largest value and α_n is a factor with dimension x^{-1}. A very full account is given in Gumbel (1958).

4.11. *Probability Integral Transformation*

$$\xi = \int_{-\infty}^{x} p(t)\, dt,$$

where $p(x)$ is the probability density function of x. In this case ξ is uniformly distributed on the interval (0, 1) so that this transformation does not induce normality although it does stabilize the variance perfectly at $V(\xi) = \frac{1}{12}$. David and Johnson (1948, 1950) and David (1950) consider this transformation in some detail for the case where $p(x)$ depends on some unknown parameters. The transformation appears to have been first used by Fisher (1932) and Karl Pearson (see Pearson, 1938), in combining independent significance tests.

4.12. *Co-ordinate Transformation*

$$\xi = \Phi^{-1}\{P(x)\},$$

where $P(x)$ is the cumulative distribution function of x and Φ^{-1} is the inverse of the distribution function of the standard normal distribution.

This transformation is an application of the preceding probability integral transformation, $\eta_1 = P(x)$, followed by what might be called the "inverse probability integral transformation", $\xi = \Phi^{-1}(\eta_1)$. For this transformation ξ is exactly a standard normal variate but of course, P is in general unknown and Φ^{-1} cannot be expressed in closed form. However, Kowalski and Tarter (1969) show how, using approximations to Φ^{-1}, for example that of Tarter (1968), and using the Fourier estimator of the distribution function discussed in Tarter and Kronmal (1968), practical use can be made of the transformation. They use it to transform a variety of non-normal distributions to investigate the power of the normal tests for independence.

4.13. *Equivalent Deviate Transformation*

$$x = \int_{-\infty}^{\xi} h(t)\, dt,$$

where $h(t)$ is any specified probability distribution function. This transformation was introduced by Finney (1949), and is further discussed in Finney (1964*a*), in connection with biological assay work. Other transformations, e.g. the probability integral and the probit, are, ignoring additive constants, special cases of this transformation.

4.14. *Probit Transformation*

$$x = \int_{-\infty}^{\xi-5} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt.$$

This transformation was first used by Gaddum (1933) as a normal equivalent deviate, but Bliss (1935) decreased the normal deviate by 5 with the object of making the occurrence of negative values very rare. This latter transformation is known as the probit transformation. It is used where data may be interpreted on the supposition that a normal deviate is linearly dependent on some observable concomitant variable, and that an observable frequency is that with which this deviate is exceeded in a normal distribution. For example, the frequency with which a high jumper clears the bar decreases with the height at which it is placed.

4.15. *Logit Transformation*

$$\xi = \log\left(\frac{x}{1-x}\right).$$

In some cases, particularly in biological assay and population growth, the probability of an event can be expressed as

$$p(x) = \frac{1}{1+e^{-(\alpha+\beta x)}}.$$

This is the logistic function and the logit transformation can, since the inverse transformation is $x = 1/(1+e^{-\xi})$, usefully be used to analyse data generated by it. It is a special case of the equivalent deviate transformation when the frequency function, i.e. $\frac{dp(x)}{dx}$, in this notation, is $h(t) = \frac{1}{2}\,\mathrm{sech}^2\, t$. Since $\mathrm{ar\,tanh}\,(r) = \frac{1}{2}\log_e\left(\frac{1+r}{1-r}\right)$, this transformation is also equivalent to the ar tanh (r) transformation with $r = 2x-1$. The transformation is considered in Cox (1970), Finney (1964, 1964*a*), Fisher and Yates (1957), and Yates (1955, 1961).

4.16. *Legit Transformation*

Intermediate in character between the distributions of probit and logit is the deviate, ξ, appropriate to gene frequencies, x, determined by selection and diffusion, where x and $y\,(= 1-x)$ are related to the standard deviate, ξ, by the equation

$$\frac{d^2y}{d\xi^2} = 4xy\xi.$$

Fisher (1950) gave the name "legit" to this transformation. In this paper he also gives the necessary tables.

4.17 *Normal Scores*

It is often necessary to draw statistical conclusions from data giving the rank order of a number of magnitudes without knowledge of their quantitative values. Thus in tests of psychological preferences, subjects can often express preferences without being able to assign numerical values to the force with which the preference is felt. Sometimes also, an experimenter who possesses quantitative values may suspect that the metric used is unsuitable for the comparisons he wishes to make, so that he may prefer to draw conclusions only from the order of the magnitudes observed.

In either case, after ranking the observations, they are replaced by the corresponding expected values of the order statistics in a sample of n from a standardized normal distribution. These are the normal scores introduced by Fisher and Yates in the first edition of their *Tables* published in 1938. They are tabulated in Fisher and Yates (1957).

Fisher and Yates suggested that the normal scores should be used for significance tests involving comparisons with ranked data, assumed to be derived from an underlying normal distribution. For example, to compare the locations of two samples, the ranks are replaced by the corresponding normal scores and, as an approximation, the standard t test is then applied.

The intuitive basis of this procedure is, of course, that the normal scores provide a good reconstruction of the underlying variate values on a standardized scale, the null hypothesis being assumed true. This intuitive thinking has been put on a more formal basis by Terry (1952) and Hoeffding (1951, 1953) who, amongst other things, show that asymptotically there is perfect correlation between the expected values of the order statistics and the variate-values they replace. In addition, Chernoff and Savage (1958) showed that for detecting shifts in location of an arbitrary distribution the test based on normal scores has asymptotically a power greater or equal to that of the usual large-sample normal test using variate values. In particular, if the populations are normal there is asymptotically no loss of power from using normal scores rather than variate values.

4.18. *Exponential Scores*

The use of normal scores is only a special case of a more general treatment discussed for example by Lehmann (1959, pp. 232–40). Another special case was introduced by Cox (1964), who suggested the following procedure for significance tests involving the comparison of exponential distributions: rank the observations, replace the ranks by the corresponding expected order statistics in sampling from the unit exponential distribution and then calculate the usual exponential theory test statistic. The special case of this procedure for the comparison of two samples was given by Savage (1956) and Cox (1964) investigates some tests based on exponential scores.

4.19. *Monotone Transformation*

In two papers, Kruskal (1964, 1965) gives a computer based method of finding the monotone transformation, which after the transformation, minimizes the residual sum of squares, suitably scaled. The sum of squares is calculated from an assumed linear model. No parametric family and no normality assumption is required.

5. Transformation of Independent Variables

So far we have written $\xi = g(x)$ with $E(x) = \theta$, but there will be many cases where the statistician will know that

$$\theta = a_1 z_1 + a_2 z_2 + \ldots + a_k z_k$$

say, in which the $a_1, \ldots, a_k$ are unknown regression coefficients and the $z_1, \ldots, z_k$ concomitant variables. This situation arises particularly in the construction of models of chemical reactions and leads very rapidly to the problem of non-linear optimization discussed, for example, in Box and Tidwell (1962). Hill (1966) pointed out that if in addition to x the z's were also to be transformed, then *all* the transformations should be carried out simultaneously and not in stages. If the family of transformations is restricted to the class of power transformations, then Draper and Hunter (1969) have indicated comparatively simple ways of finding the appropriate powers of x and of $z_1, z_2, \ldots, z_k$.

6. Removal of Transformation Bias

Whatever the purpose of the transformation, it often raises problems when the analysis of the transformed data is complete. For example, in weather modification experiments, the analysis of the data may be best carried out in terms of the cube root of the rainfall and will indicate whether or not an effect due to "seeding" is present. Suppose the evidence suggests that it is. Clearly the estimated magnitude of the effect cannot be presented in terms of "cube-rooted" inches of rain. Is it sufficient to merely cube the effect as measured in the transformed variables and present this as the MVUE of the effect of cloud seeding?

Put more generally, we have the transformations $g(.)$, the estimators $\hat{\mu}$ and $\hat{\sigma}^2$, and we know their distributions. The problem is to find the MVUE of θ, $\hat{\theta}$. In practice we want more than this for it is desirable to put a standard error on this estimate, so that the variance $V(\hat{\theta})$ and an estimate of it, $\hat{V}(\hat{\theta})$, are also required. We should also ideally wish to indicate a confidence region for θ. Further work in cloud seeding by Howell (1966) and Smith, Adderley and Bethwaite (1965) has indicated that the variance may be a more important parameter than the mean. This implies that there are situations when we wish to estimate ϕ^2 as well as θ.

Neyman and Scott (1960) considered in general terms the problem of finding the MVUE of θ. They show that provided $f(\xi)$ is any entire function of second order or less, θ and $E^{-1}(\theta)$ exist. A function of a complex variable ζ will be called regular in a region if it is analytic and single valued there. An entire function (sometimes known as an integral function) is one which is regular for all finite ζ. The order of an entire function measures the rate of growth of the function as ζ increases. All polynomials and exponential function are entire functions. A standard reference is Boas (1954).

A function $f(z)$ is entire of second order if the radii of convergence of the two series

$$\sum_{n=0}^{\infty} \frac{1}{n!} f_0^{(2n)} z^n \quad \text{and} \quad \sum_{n=0}^{\infty} \frac{1}{n!} f_0^{(2n+1)} z^n \tag{A}$$

are infinite where $f_a^{(r)}$ stands for the rth derivative of $f(z)$ evaluated at $z = a$. Neyman and Scott (1960) point out this is stronger than that the Taylor Series expansion of f,

$$f(z) = \sum_{n=0}^{\infty} \frac{1}{n!} f_0^{(n)} z^n \tag{B}$$

is convergent for all real z. If the radii of convergence of (A) are infinite, then the radius of (B) will also be infinite, but the converse is not necessarily true.

Neyman and Scott go on to show that, when f(z) is entire of second order,

$$\theta = f(0) + \sum_{r=1}^{\infty} \frac{1}{r!} f_0^{(r)} E(\xi^r)$$

$$\hat{\theta} = \hat{\theta}(\hat{\mu}, S) = f(0) + \sum_{r=1}^{\infty} \frac{1}{r!} f_0^{(r)} T_r,$$

where T_r is such that

$$E(T_r) = E(\xi^r)$$

Specifically,

$$T_{2r} = \sum_{k=0}^{r} \frac{(2r)!}{(2k)!(r-k)!} \hat{\mu}^{2k} [\tfrac{1}{4}S(1-\lambda^2)]^{r-k} \frac{\Gamma(v/2)}{\Gamma\left(\frac{v}{2}+r-k\right)}$$

and

$$T_{2r+1} = \sum_{k=0}^{r} \frac{(2r+1)!}{(2k+1)!(r-k)!} \hat{\mu}^{2k+1} [\tfrac{1}{4}S(1-\lambda^2)]^{r-k} \frac{\Gamma(v/2)}{\Gamma\left(\frac{v}{2}+r-k\right)}.$$

That $\hat{\theta}$ derived in this way is the MVUE estimator of θ follows from the definition of the T_r's, the sufficiency of $\hat{\mu}$ and $\hat{\sigma}^2$ for μ and σ^2, and the result of Lehmann and Scheffé (1950) that any function of the sufficient statistics is a MVUE of its expectation. Moreover, since $(\hat{\mu}, \hat{\sigma}^2)$ are complete sufficient statistics, $\hat{\theta}$ is unique (Lehmann and Scheffé, 1950).

Schmetterer (1960) has interpreted the results of Neyman and Scott in terms of the solution $h(\hat{\mu}, \hat{\sigma}^2)$ of the integral equation

$$E[h(\hat{\mu}, \hat{\sigma}^2)] = E(x).$$

Kolmogorov (1950) had earlier considered the problem of finding unbiased estimators in terms of the solutions of integral equations but he relied heavily upon the results of Blackwell (1947) in using sufficient statistics for (μ, σ^2) to turn unbiased but inefficient estimators into MVUE's.

The problem of estimating the variance ϕ^2 can also be formulated as an integral equation, viz, find $h^*(\hat{\mu}, \hat{\sigma}^2)$ such that

$$E[h^*(\hat{\mu}, \hat{\sigma}^2)] = E[x - E(x)]^2.$$

This does not seem a promising way of proceeding because of the non-linearity of the integral equation.

The formula given by Neyman and Scott for $\hat{\theta}$ is complicated but they point out that four commonly used transformations – the square root, logarithm, angular and hyperbolic – are linked by a simple differential equation concerning their inverse functions. These functions are also entire functions. The equation is

$$f_{\xi}^{(2)} = A + Bf(\xi).$$

Table I gives the values of A and B for these four transformations.

If the transformation is taken after a linear function has been made of the natural or original variable, no additional problem arises. Let y be the natural variable with mean, ψ, and let $x = cy+d\,(c \neq 0)$. Then $\theta = c\psi + d$ and

$$\hat{\psi} = \frac{1}{c}(\hat{\theta}-d).$$

By the results of Lehmann and Scheffé (1950), $\hat{\psi}$ is the unique MVUE of ψ. By a similar argument no problems arise when considering $\hat{V}(\hat{\theta})$, $\hat{V}(\hat{\theta})$ and $\hat{\phi}^2$.

Table I. *The recursive transformations*

$\xi = g(x)$	$x = f(\xi)$	A	B
$\sqrt{x}$	ξ^2	2	0
$\log_M(x)$	$e^{\log_e M} = e^m$	0	m^2
$\arcsin\sqrt{x}$	$\sin^2\xi$	2	-4
$\operatorname{arc\,sinh}\sqrt{x}$	$\sinh^2\xi$	2	4

Neyman and Scott go on to derive the following results for θ and $\hat{\theta}$.

$$\theta = \begin{cases} f(\mu)\,e^{B\sigma^2/2} + \dfrac{A}{B}(e^{B\sigma^2/2}-1) & B \neq 0 \\ f(\mu) + A\sigma^2/2 & B = 0 \end{cases}$$

and

$$\hat{\theta} = \begin{cases} \Phi\,[B(1-\lambda^2)\,S,\, v]\left[f(\hat{\mu}) + \dfrac{A}{B}\right] - \dfrac{A}{B} & B \neq 0 \\ f(\hat{\mu}) + A(1-\lambda^2)\,\hat{\sigma}^2/2 & B = 0 \end{cases}$$

where

$$\Phi(aS, v) = \sum_{r=0}^{\infty} \frac{1}{r!} \frac{\Gamma(v/2)}{\Gamma\left(\dfrac{v}{2}+r\right)} \left(\frac{aS}{4}\right)^r$$

$$= \left(\frac{2}{S\sqrt{a}}\right)^{v/2-1} \Gamma\left(\frac{v}{2}\right) I_{v/2-1}(S\sqrt{a})$$

and $I_\varepsilon(u)$ is the Bessel function of imaginary argument. This series converges very rapidly, only a few terms usually being required for adequate accuracy.

It follows from these results that the bias of the crude estimator, $f(\hat{\mu})$, of θ is

$$E\{f(\hat{\mu})-\theta\} = \begin{cases} \left[\theta + \dfrac{A}{B}\right][\exp\{-B(1-\lambda^2)\,\sigma^2/2\}-1] & B \neq 0 \\ -A(1-\lambda^2)\,\sigma^2/2 & B = 0. \end{cases}$$

It can be seen that the absolute value of this bias is always a monotonic decreasing function of λ^2. This implies, by recalling that λ^2 measures the precision (in terms of σ^2) with which μ is estimated, that the larger the sample size $\left(\lambda^2 = \frac{1}{n}\right)$ or the better the experimental design (λ^2 is smaller than for a less efficient design) the *worse* the bias of the crude estimated $f(\hat{\mu})$. Thus if the analysis of the data is incorrect and no correction is made to allow for the bias in the inverse transformation, some of the advantages of a large sample or an efficient design will be lost.

There are of course many transformations whose inverse function is an entire function but which is not a solution to the recursive condition of Neyman and Scott. In these cases the more general results of Neyman and Scott must generally be used. For some simple transformations, e.g. the cube-root transformation, an expression for $\hat{\theta}$ can be easily obtained by simple manipulation of the known expressions for the moments of the normal and chi-square distributions.

It is interesting to note that the results of the Neyman and Scott indicate a theoretical difficulty with the reciprocal transformation, $\xi = 1/x$. In this case $\theta = E(x) = E(1/\xi)$ does not exist so that it cannot be estimated.

However, Box (1971) has suggested defining the pseudo expectation, θ', of x as

$$\theta' = PE(x) = \lim_{\varepsilon \to 0} \int_{\varepsilon}^{\infty} \frac{1}{\xi} \cdot \frac{e^{-\frac{(\xi-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi}\sigma}\, d\xi$$

and following Ghurye and Olkin (1969) this can be estimated unbiasedly as

$$\widehat{\theta'} = \widehat{PE}(x) = \lim_{\varepsilon \to 0} \int_{\varepsilon}^{\infty} \frac{1}{\xi} E^{-1} \left\{ \frac{e^{-\frac{(\xi-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi}\sigma} \right\} d\xi.$$

Barton (1961) has given the necessary expression for the unbiased estimator of the normal density so that $\hat{\theta}'$ can be calculated.

In Hoyle (1968) consideration is given to ϕ^2, $\hat{\phi}^2$ and $\hat{V}(\hat{\theta})$ and it is shown that, provided $f(.)$ is restricted to the class of entire functions of second order or less, ϕ^2 exists and that $\hat{\phi}^2$ and $\hat{V}(\hat{\theta})$ are MVUE's. Expressions for these functions are given for the square root, cube root, logarithmic, angular and hyperbolic transformations.

The results for the logarithmic transformation have also been considered by Finney (1941), Sichel (1951–2) and Meulenberg (1965).

Goldberger (1968), Heien (1968) and Bradu and Mundlak (1970) have considered a more general problem concerning the logarithmic transformation and the log normal distribution. They consider the case of $x = e^{a_1 z_1 + \dots + a_k z_k + u}$ where $u \sim N(0, \sigma^2)$ is the error term. Bradu and Mundlak (1970) also discuss $\hat{V}(\hat{\theta})$ and consider confidence intervals for $V(\hat{\theta})$.

Yates (1961) considers a related problem involving the use of the logit (or indeed any other generally used transformation) in the analysis of quantal data in multiway tables. It is pointed out that if the constants obtained in the analysis are retransformed directly to percentages considerable distortions may occur. It is shown that the ordinary maximum likelihood procedure of fitting, using provisional and working values and successive approximation, removes this distortion. The result follows from the fact that the weighted mean of a set of percentages, with weights proportional to the frequencies, is the maximum likelihood estimate of the mean percentage.

7. Confidence Intervals for θ

For the class of recursive transformations Neyman and Scott (1960) have obtained the MVUE of θ but it is clearly desirable to have in addition some indication of the variability of $\hat{\theta}$ and in Hoyle (1968) I have discussed estimating the variance of $\hat{\theta}$. However, it is by no means self-evident how this estimate should be used in order to derive a satisfactory confidence interval for θ.

In a series of papers, Land (1969, 1970, 1971) has given results and extensive tables enabling the exact confidence interval for θ to be calculated when the logarithmic transformation is used. Given the estimated $\hat{\mu}$ and π^2, he shows how to test hypotheses of the form

$$H_p: \mu+\beta\sigma^2 = p,$$

for arbitrary β. The confidence sets defined in terms of these tests are intervals, in the one-sided case at least and it also appears, but it has not yet been proved, to be the case for the two-sided test. Land goes on to show that exact confidence intervals for functions of the form $h(\mu, \sigma^2) = k(\mu+\beta\sigma^2)$, such as the mean of the lognormal distribution,

$$h(\mu, \sigma^2) = \exp(\mu+\sigma^2/2),$$

and hence for θ when the logarithmic transformation is used, can be obtained by the above method.

For the single sample model, Land's method is based on the conditional distribution of

$$T = \frac{(\hat{\mu}-\log\theta)}{\sqrt{\hat{\sigma}^2/n}} \quad \text{given} \quad Z = \left[\left(1-\frac{1}{n}\right)\hat{\sigma}^2+(\hat{\mu}-\log\theta)^2\right]^{\frac{1}{2}}$$

under the null hypothesis, $E(x) = \theta$. The conditional distribution of T given $Z = z$, when $E(x) = \theta$, has the density

$$p_v(t \mid w) \propto (v+t^2)^{-\left(\frac{v+1}{2}\right)} \exp\left[(v+1)\, wt\, (v+t^2)^{-\frac{1}{2}}\right],$$

where $v = n-1$ and $w = -\frac{1}{2}z$. If $t(v, w, \alpha)$ denotes the αth quantile of this distribution, the uniformly most powerful unbiased (UMPU) test of level α of the null hypothesis against the alternative $E(x)<\theta$ is given by the rule "reject if $T<t(v, w, \alpha)$". The optimal two-sided tests require different tables of critical values. A level α upper confidence interval for $E(x)$ is defined as the smallest value of θ such that the UMPU level test of $H: E(x) = \theta$ does not reject in favour of the alternative $E(x)<0$.

Land (1969, 1970) goes on to suggest that approximate intervals for more general functions, such as $\theta = \theta(\mu, \sigma^2)$ in the non-logarithmic transformation case, can be obtained by constructing a confidence interval for a linear function of μ and σ^2 that approximates to $\theta(\mu, \sigma^2)$ in the region of interest.

Kanofsky (1968) has proposed a method of simultaneous confidence intervals for all functions of μ and σ. He constructs a trapezoidal-shaped confidence region of level $1-\alpha$ for μ and σ, and for an arbitrary function, $h(\mu, \sigma)$, defines a confidence set for this function as the set of values m such that the curve $h(\mu, \sigma) = m$ intersects this confidence region. If one is interested in a single function, the procedure in general is conservative, but, according to Land (1970), it is the only method based on exact distribution theory that has been proposed for a general function $h(\mu, \sigma)$.

8. Approximate Confidence Intervals for θ

The usual method has been to rely on approximate confidence intervals for θ and Land (1970, 1971) distinguishes between two general approaches, the transformation methods and the direct methods. In the transformation methods one essentially transforms the α-level confidence interval for μ, say $\hat{\mu}_L \leqq \mu \leqq \hat{\mu}_U$, where $P\{\hat{\mu}_L \leqq \mu \leqq \hat{\mu}_U\} = \alpha$, into a confidence interval for θ and the various methods are concerned with different ways of making the transformation. The direct methods on the other hand go directly from an estimate $\hat{\theta}$ and some measure of its variability to an approximate confidence interval for θ. The results in Land (1971) and Hoyle (1971) indicate that neither approach is satisfactory for the square root, cube root, logarithmic, angular and hyperbolic transformations.

The best approximate method seems to be that developed by Professor D. R. Cox in which an estimate is made of $\beta = h(\theta)$, where $h(\theta)$ is a monotonic function. An approximate confidence interval for β, $\hat{\beta}_L \leqq \beta \leqq \hat{\beta}_U$, based on approximate or large sample consideration is transformed into a confidence interval for θ by using the inverse function $h^{-1}(.)$. The problem comes in identifying a suitable function $h(\theta)$. Land (1971) has shown that for the logarithmic transformation

$$\begin{aligned}\beta = h(\theta) &= \log(\theta) \\ &= \mu+\sigma^2/2\end{aligned}$$

is a good choice of $h(\theta)$. The interval $(\hat{\beta}_L, \hat{\beta}_U)$ is based on the fact that

$$\hat{\beta} = \hat{\mu}+\hat{\sigma}^2/2,$$

$$\operatorname{var}(\hat{\beta}) = \sigma^2\left\{\frac{1}{n}+\frac{1}{2(n+1)}\right\},$$

and that β has approximately a normal distribution.

In another approach, Mantel and Parwary (1961) have suggested treating the ratio of the unrestricted maximum of the likelihood to the maximum of the likelihood under the constraint that $\theta(\mu, \sigma^2)$ is equal to some constant, as a chi-square variate with one degree of freedom. By using a series of such constants, the range, or ranges, of values consistent with the data can be found, thus leading to approximate confidence intervals for $\theta(\mu, \sigma^2)$. Their method was developed to apply to functions $h(\delta_1, \delta_2, \ldots, \delta_k)$ where the δ_i are parameters of interest but their method can clearly be applied to the problem at hand. Their original problem has been considered further by Halperin and Mantel (1963), Halperin (1964, 1965) for the case where δ_i are means of normal variates. Likelihood theory has also been used by Box and Cox (1964) to determine approximate confidence intervals for the parameter of a power transformation.

A still further approximate approach has been developed in Sichel (1966, 1967) for the problem of setting confidence intervals for the logarithmic transformation. In Sichel (1967) the exact sampling distribution of $\hat{\theta}$ is given, but it depends on the values of the unknown parameters μ and σ^2. However, in Sichel (1966) an excellent approximation to it is given from which confidence intervals can be computed. It does require, however, the use of a further set of tables (for the T-distribution).

The approximation of Sichel (1966) is that $\hat{\theta}$ has the approximate density

$$\frac{1}{\sqrt{2\pi}\rho\hat{\theta}}\exp\left\{-\tfrac{1}{2}\left(\frac{\log_e\hat{\theta}-\eta}{\rho}\right)^2\right\},$$

where

$$\rho^2 = \lambda^2\sigma^2+\log_e\Phi\{(1-\lambda^2)^2\sigma^4, v\}$$

and

$$\eta = \mu+\tfrac{1}{2}(\sigma^2-\rho^2).$$

This lognormal approximation has mean and variance identical to the mean and variance of the exact sampling distribution. ρ^2 is estimated as

$$\hat{\rho}^2 = \hat{\sigma}^2+\log_e\Phi(S^2, v)$$

leading to the lower confidence limit for the population mean given by Sichel (1966) as

$$\hat{\theta}_L = \hat{\theta}\exp\{\tfrac{1}{2}\hat{\rho}^2-T_\alpha\hat{\rho}\},$$

where T_α is a deviate cutting of a proportion α in the tail of the T-distribution and

$$T = \frac{\log_e \hat{\theta} - \log_e \theta}{\hat{\rho}} + \tfrac{1}{2}\hat{\rho}.$$

Sichel (1966) has given the exact distribution of T which unfortunately depends on the unknown population parameter σ^2. However, the T-distribution is robust against changes in v^2 for $0{\cdot}3 \leqq \sigma^2 \leqq 1{\cdot}5$, a domain covering many practical applications of the lognormal theory, particularly in mining.

Sichel (1966) gives simple tables of factors by which the estimates $\hat{\theta}$ must be multiplied to obtain approximate confidence limits for the lognormal population mean θ at nominal confidence levels $\alpha = 0{\cdot}05$, $0{\cdot}10$ and $0{\cdot}95$. For these tables the nuisance parameter σ^2 was set equal to $0{\cdot}7$.

Acknowledgement

This review formed part of the author's Ph.D. thesis at the University of London. I am grateful to Professors Stuart and Plackett for their comments. I am also grateful to Miss Kearsey, the Librarian of the Royal Statistical Society, and her staff, for their help in compiling this bibliography.

9. Bibliography

1 Aitchison, J. and Brown, J. A. C. (1957). *The Lognormal Distribution.* Cambridge University Press.

2 Anscombe, F. J. (1948). The transformation of poisson, binomial and negative binomial data. *Biometrika*, **35**, 246–54.

3 Anscombe, F. J. (1950). Table of the hyperbolic transformation $\sinh^{-1}\sqrt{x}$. *J. Roy. Stat. Soc.* A, **113**, 228–9.

4 Anscombe, F. J. (1955). Contribution to the discussion of paper by Box, G. E. P. and Anderson, S. E. L. *J. Roy. Stat. Soc.* B, **17**, 1–26.

5 Anscombe, F. J. (1956). On estimating binomial response relations. *Biometrika*, **43**, 461–4.

6 Anscombe, F. J. (1961). Examination of residuals. *Fourth Berkeley Symposium of Mathematical Statistics and Probability.* University of California Press, Berkeley, **1**, 1–36.

7 Anscombe, F. J. and Tukey, J. W. (1955). The criticism of transformations. *J. Amer. Stat. Assoc.* **50**, 566 (abstract).

8 Anscombe, F. J. and Tukey, J. W. (1963). The examination and analysis of residuals. *Technometrics*, **5**, 151–60.

9 Bartlett, M. S. (1936). The square-root transformation in analysis of variance. *J. Roy. Stat. Soc. Suppl.* **3**, 68–78.

10 Bartlett, M. S. (1937). Subsampling for attributes. *J. Roy. Stat. Soc. Suppl.* **4**, 131–5.

11 Bartlett, M. S. (1947). The use of transformations. *Biometrics*, **3**, 39–52.

12 Bartlett, M. S. and Kendall, D. G. (1946). The statistical analysis of variance-heterogeneity and the logarithmic transformation. *J. Roy. Stat. Soc. Suppl.* **8**, 128–38.

13 Beall, G. (1942). The transformation of data from entomological field experiments so that the analysis of variance becomes applicable. *Biometrika*, **29**, 243–62.

14 Bennett, B. M. (1956). On the variance stabilizing properties of certain logarithmic transformations I. *Trabajos de Estadistica*, **7**, 295–7.

15 Bennett, B. M. (1957). On the variance stabilizing properties of certain logarithmic transformations II. *Trabajos de Estadistica*, **8**, 69–74.

16 Berkson, J. (1953). A statistically precise and relatively simple method of estimating the bio-assay and quantal response. *J. Amer. Stat. Assoc.* **48**, 565–99.

17 Biggers, J. D. and Claringfold, P. J. (1954). The angular transformation in quantal analysis. *Biometrics*, **10**, 570 (abstract).

18 Blackwell, D. (1947). Conditional expectation and unbiased estimation. *Ann. Math. Statist.* **18**, 105–10.

19 Blom, G. (1954). Transformation of the binomial, negative binomial, poisson and chi-square distribution. *Biometrika*, **41**, 302–16. (Correction **43**, 235.)

20 Bliss, C. I. (1935). The calculation of the dosage mortality curve. *Ann. Appl. Biol.* **22**, 134–67.

21 Bliss, C. I. (1937). The analysis of field experimental data expressed in percentages. *Pl. Prot. Leningrad*, **12**, 70–2.

22 Bliss, C. I. (1938). The transformation of percentages for use in the analysis of variance. *Ohio. J. Science*, **38**, 9–12.

23 Boas, R. P. Jr. (1954). *Entire Functions.* Academic Press, New York.

24 Box, G. E. P. and Anderson, S. E. L. (1955). Permutation theory in the derivation of robust criteria and the study of departures from assumption. *J. Roy. Stat. Soc.* B, **17**, 1–26.

25 Box, G. E. P. and Cox, D. R. (1964). An analysis of transformations. *J. Roy. Stat. Soc.* B, **26**, 211–52.

26 Box, G. E. P. and Tidwell, P. W. (1962). Transformation of the independent variable. *Technometrics*, **4**, 5321–50.

27 Box, M. J. (1971). Bias in nonlinear estimation. *J. Roy. Stat. Soc.* B, **33**, 171–201.

28 Bradu, D. and Mundlak, Y. (1970). Estimation in lognormal linear models. *J. Amer. Stat. Assoc.* **65**, 198–211.

29 Chandra, Sekar C. (1942). A note on the inverse sine transformation. *Sankhyā*, **6**, 195–8.

30 Chernoff, H. and Savage, I. R. (1958). Asymptotic normality and efficiency of certain nonparametric test statistics. *Ann. Math. Statist.* **29**, 972–94.

31 Claringfold, P. J., Biggers, J. B. and Emmens, C. W. (1953). The angular transformation in quantal analysis. *Biometrics*, **9**, 467–83.

32 Clark, A. and Leonard, W. H. (1939). The analysis of variance with special reference to data expressed as percentages. *J. Amer. Soc. Agron.* **31**, 55–6.

33 Cochran, W. G. (1938). Some difficulties in the statistical analysis of replicated experiments. *Emp. J. Expt. Agric.* **6**, 157–75.

34 Cochran, W. G. (1940). Analysis of variance when experimental errors follow poisson or binomial laws. *Ann. Math. Statist.* **9**, 335–47.

35 Cornish, E. A. and Fisher, R. A. (1937). Moments and cumulants in the specification of distributions. *Rev. Inst. Int. Statist.* **5**, 307–20.

36 Cornish, E. A. and Fisher, R. A. (1960). The percentile points of distributions having known cumulants. *Technometrics*, **2**, 209–25.

37 Cox, D. R. (1964). Some applications of exponential ordered scores. *J. Roy. Stat. Soc.* B, **26**, 103–10.

38 Cox, D. R. (1970). *Analysis of Binary Data.* Methuen, London.

39 Curtis, J. H. (1943). On transformations used in the analysis of variance. *Ann. Math. Statist.* **14**, 107–32.

40 Daniels, H. E. and Kendall, M. G. (1958). Short proof of Miss Harley's theorem on the correlation coefficient. *Biometrika*, **45**, 571–2.

41 David, F. N. (1955). The transformation of discrete variables. *Ann. Human Genet.* **19**, 174–82.

42 David, F. N. and Johnson, N. L. (1948). The probability integral transformation when parameters are to be estimated. *Biometrika*, **35**, 182–90.

43 David, F. N. and Johnson, N. L. (1950). The probability integral transformation when the variable is discontinuous. *Biometrika*, **37**, 42–9.

44 Dolby, J. L. (1963). A quick method for choosing a transformation. *Technometrics*, **5**, 317–26.

45 Draper, J. (1952). Properties of distributions resulting from certain simple transformations of the normal distribution. *Biometrika*, **39**, 290–301. (Corrections, *Biometrika*, **40**, 236.)

46 Draper, N. R. and Cox, D. R. (1969). On distributions and their transformations to normality. *J. Roy. Stat. Soc.* B, **31**, 472–6.

47 Draper, N. R. and Hunter, W. G. (1969). Transformations: some examples revisited. *Technometrics*, **11**, 23–40.

48 Dyke, G. V. and Patterson, H. D. (1952). Analysis of factorial arrangements when the data are proportions. *Biometrics*, **8**, 1–12.

49 Eisenhart, C., Hastay, H. W. and Wallis, W. A. (eds) (1947). Inverse sine transformation of proportions. *Techniques of Statistical Analysis.* McGraw-Hill, New York.

50 Elston, R. E. (1961). On additivity in the analysis of variance. *Biometrics*, **17**, 209–19.

51 Finney, D. J. (1941). On the distribution of a variate whose logarithmic is normally distributed. *J. Roy. Stat. Soc. Suppl.* **7**, 155–61.

52 Finney, D. J. (1949). The estimation of the parameters of tolerance distributions. *Biometrika*, **36**, 239–56.

53 Finney, D. J. (1951). Estimation of bacterial densities from dilution series. *J. Hygiene*, **49**, 26–35.

54 Finney, D. J. (1964). *Probit Analysis.* Cambridge University Press, Cambridge.

55 Finney, D. J. (1964). *Statistical Method in Biological Assay* (2nd edition). Griffins, London.

56 Fisher, A. (1936). *The Mathematical Theory of Probabilities* **1** (2nd edition). Macmillan, New York.

57 Fisher, R. A. (1920). Studies in crop variation. 1. An examination of the yield of dressed grain from Broadwalk. *J. Agric. Sci.* **11**, 107–35.

58 Fisher, R. A. (1921). On the "probable error" of a correlation coefficient deduced from a small sample. *Metron*, **4**, 1–32.

59 Fisher, R. A. (1922). On the mathematical foundations of theoretical statistics. *Phil. Trans. Roy. Soc. London*, **222**, 309–68.

60 Fisher, R. A. (1925). *Statistical Methods for Research Workers* (1st edition). Griffin, London.

61 Fisher, R. A. (1930). The distribution of gene ratio for rare mutations. *Proc. Roy. Soc. Edinburgh*, **50**, 205–20.

62 Fisher, R. A. (1950). Gene frequencies in a cline determined by selection and diffusion. *Biometrics*, **6**, 353–61.

63 Fisher, R. A. (1954). The analysis of variance with various binomial transformations. *Biometrics*, **10**, 130–51.

64 Fisher, R. A. and Mather, K. (1943). The inheritance of style length "Lythrum Salicano". *Ann. Eug.* (1943), **12**, 1–23.

65 Fisher, R. A. and Yates, F. (1957). *Statistical Tables for Biological, Agricultural and Medical Research* (5th edition). Oliver and Boyd, Edinburgh.

66 Freeman, M. F. and Tukey, J. W. (1950). Transformations related to the angular and the square root. *Ann. Math. Statist.* **21**, 607–11.

67 Gaddum, J. H. (1945). Lognormal distributions. *Nature*, **156**, 463–6.

68 Garwood, F. (1936). Fiducial limits for the Poisson distribution. *Biometrika*, **28**, 437–42.

69 Ghurye, S. G. (1949). Transformations of a binomial variate for the analysis of variance. *J. Ind. Soc. Agric. Stat.* **2**, 94–109.

70 GOLDBERGER, A. S. (1968). The interpretation and estimation of Cobb-Douglas functions. *Econometrica*, **36**, 464–72.

71 GRIZZLE, L. E. (1962). Asymptotic power of tests of linear hypotheses using the probit and logit transformations. *J. Amer. Stat. Assoc.* **57**, 877–94.

72 GUMBEL, E. J. (1958). *Statistics of Extremes*. Columbia University Press, New York.

73 HALDANE, J. B. S. (1937). The approximate normalization of a class of frequency distributions. *Biometrika*, **29**, 392–404.

74 HALDANE, J. B. S. (1942). Moments of the distributions of powers and products of normal variates. *Biometrika*, **32**, 226–42.

75 HALPERIN, M. (1964). Interval estimation of non-linear parametric functions II. *J. Amer. Stat. Assoc.* **59**, 168–81.

76 HALPERIN, M. (1965). Interval estimation of non-linear parametric functions III. *J. Amer. Stat. Assoc.* **60**, 1191–9.

77 HALPERIN, M. and MANTEL, N. (1963). Interval estimation of non-linear parametric functions I. *J. Amer. Stat. Assoc.* **58**, 611–27.

78 HARLEY, B. I. (1956). Some properties of an angular transformation for the correlation coefficient. *Biometrika*, **43**, 219–24.

79 HARLEY, B. I. (1957). Further properties of the angular transformation of the correlation coefficient. *Biometrika*, **44**, 273–5.

80 HEALY, M. J. R. (1967). A critical table of two-decimal logarithms. *Applied Stats.* **16**, 86.

81 HEALY, M. J. R. and TAYLOR, L. R. (1962). Tables for power-law transformations. *Biometrika*, **49**, 557–9.

82 HEIEN, D. M. (1968). A note on log-linear regression. *J. Amer. Stat. Assoc.* **63**, 1034–8.

83 HILL, W. J. (1966). *Statistical Techniques for Model Building*. Unpublished Ph.D. thesis, University of Wisconsin.

84 HOEFFDING, W. (1951). Optimum nonparametric tests. *Proceedings Second Berkeley Symposium on Mathematical Statistics*. University of California Press, Berkeley, **1**, 83–92.

85 HOEFFDING, W. (1953). On the distribution of the expected values of the order statistics. *Ann. Math. Statist.* **24**, 93–100.

86 HOTELLING, H. (1953). New Light on the correlation coefficient and its transform. *J. Roy. Stat. Soc.* B, **15**, 193–232.

87 HOWELL, W. E. (1965). Cloud seeding against the 1964 drought in the North-East. *J. Appl. Meteorology*, **4**, 553–9.

88 HOWELL, W. E. (1966). Effect on mean rainfall of artificially increased variance. *J. Appl. Meteorology*, **5**, 128–9.

89 HOYLE, M. H. (1968). The estimation of variances after using a gaussianating transformation. *Ann. Math. Statist.* **39**, 1125–43.

90 HOYLE, M. H. (1971). *Transformations and Spoilt Data in Linear Models*. Unpublished Ph.D. thesis, University of London.

91 JENKINS, G. M. (1954). An angular transformation for the serial correlation coefficient. *Biometrika*, **41**, 261–5.

92 JOHNSON, N. L. (1949). Systems of frequency curves generated by methods of translation. *Biometrika*, **36**, 149–176.

93 KANOFSKY, P. (1968). Derivation of simultaneous confidence intervals for parametric functions from a parametric confidence region. *Sankhyā* A, **30**, 379–86.

94 KEMPTHORNE, O. (1968). Aims, logic and methods of transformation. *Biometrics*, **24**, 1034 (abstract).

95 KENDALL, M. G. and STUART, A. (1963). *The Advanced Theory of Statistics*, **1**. Griffin, London.

96 KENDALL, M. G. and STUART, A. (1961). *The Advanced Theory of Statistics*, **2**, Griffin, London.

97 KENDALL, M. G. and STUART, A. (1966). *The Advanced Theory of Statistics*, **3**. Griffin, London.

98 KIHLBERG, J. K., HERSON, J. H. and SCHOTZ, D. E. (1967). Square root transformations revisited. Presented to American Statistical Association Annual Meeting.

99 KLECZOWSKI, A. (1949). The transformation of local lesion counts for statistical analysis. *Ann. Appl. Biol.* **36**, 139–52.

100 KOLMOGOROFF, A. N. (1950). Unbiased estimates. Izvestya Akademii Nauk SSSR. *Seriya Mat.* **14**, 303-26.

101 KOWALSKI, C. J. (1970). The performance of some rough tests for bivariate normality before and after co-ordinate transformations to normality. *Technometrics*, **12**, 517–44.

102 KOWALSKI, C. J. and TARTER, M. E. (1969). Co-ordinate transformations to normality and the power of normal tests for independence. *Biometrika*, **56**, 139–48.

103 KRUSKAL, J. B. (1964). Nonmetric multidimensional scaling – a numerical method. *Psychometrika*, **29**, 115–29.

104 KRUSKAL, J. B. (1965). Analysis of factorial experiments by estimating monotone transformations of the data. *J. Roy. Stat. Soc.* B, **27**, 251–63.

105 KRUSKAL, J. B. (1968). Statistical analysis: transformations of data. *International Encyclopaedia of the Social Science*. The Macmillan Co. and the Free Press, Chicago.

106 LAND, C. E. (1969). Confidence interval estimation of functions of the normal mean and variance. *Ann. Math. Statist.* **40**, 1860 (abstract).

107 LAND, C. E. (1970). Confidence intervals for linear functions of the normal mean and variance. Submitted to *Ann. Math. Statist.*

108 LAND, C. E. (1971). An evaluation of approximate confidence intervals estimation methods for lognormal means. Submitted to *Technometrics*.

109 LAURENT, A. G. (1963). The lognormal distribution and the translation method: description and estimation problems. *J. Amer. Stat. Assoc.* **58**, 231–5.

110 LEHMANN, E. L. (1959). *Testing of Statistical Hypotheses*. Wiley, New York.

111 LEHMANN, E. L. and SCHEFFÉ, H. (1950). Completeness, similar regions and unbiased estimation – part 1. *Sankhyā*, **10**, 305–40.

112 LEVENE, H. (1960). Robust tests for equality of variances. *Contributions to Probability and Statistics*, ed. I. Olkin. Stanford University Press.

113 MANTEL, N. and PATWARY, K. M. (1961). Interval estimation of single parameter functions. *Bull. Inst. Int. Stat.* **38** (4), 227–39.

114 MERRINGTON, M. (1941). Numerical approximations to percentage points of the χ^2 distribution. *Biometrika*, **32**, 200–2.

115 MEULENBERG, M. T. G. (1965). On the estimation of an exponential function. *Econometrica*, **33**, 863–8.

116 MOORE, P. G. (1957). Transformations to normality using fractional powers of the variables. *J. Amer. Stat. Assoc.* **52**, 237–46.

117 MOORE, P. G. (1958). Interval analysis and the logarithmic transformation. *J. Roy. Stat. Soc.* B, **20**, 189–92.

118 MOORE, P. G. and TUKEY, J. W. (1954). Answer to query 112. *Biometrics*, **10**, 562–8.

119 MOSHMAN, J. (1952). Critical values of the lognormal distribution. *J. Amer. Stat. Assoc.* **48**, 600–9.

120 MOSTELLER, F. and TUKEY, J. W. (1949). The uses and usefulness of binomial probability paper. *J. Amer. Stat. Assoc.* **44**, 174–212.

121 NAYLOR, A. F. (1964). Comparisons of regression constants fitted by maximum likelihood to four common transformations of binomial data. *Ann. Hum. Genet. Lond.* **27**, 241–6.

122 NEYMAN, J. and SCOTT, E. L. (1960). Correction for bias introduced by a transformation of variables. *Ann. Math. Statist.* **31**, 643–55.

123 O'CARROLL, F. M. (1967). Precision evaluation and monitoring. *The Statistician*, **17**, 127–40.

124 PACHARES, J. (1968). Letter to the Editor. *American Statistician*, **22**, (4), 50.

125 PATTERSON, R. L. (1966). Difficulties involved in the estimation of a population mean using transformed sample data. *Technometrics*, **8**, 535–7.

126 PEARCE, S. C. (1945). Lognormal distributions. *Nature*, **156**, 747.

127 PEARSON, E. S. (1938). The probability integral transformation for testing goodness of fit and combining independent tests of significance. *Biometrika*, **30**, 134–48.

128 QUENOUILLE, M. H. (1948). Some results in the testing of serial coefficients. *Biometrika*, **35**, 261–7.

129 QUENOUILLE, M. H. (1949). Approximate tests of correlation-time series: 3. *Proc. Camb. Phil. Soc.* **45**, 483–4.

130 QUENOUILLE, M. H. (1950). *Introductory Statistics*. Butterworth, London.

131 QUENOUILLE, M. H. (1953). *The Design and Analysis of Experiments*. Griffin, London.

132 QUENOUILLE, M. H. (1956). Notes on bias estimation. *Biometrika*, **43**, 353–60.

133 QUENOUILLE, M. H. (1957). The effect of transformations of variables upon their correlation coefficient. *Biometrika*, **42**, 272–3.

134 SANKARAN, MUNUSWAMY (1958). On Nair's transformation of the correlation coefficient. *Biometrika*, **45**, 567–71.

135 SAVAGE, I. E. (1956). Contributions to the theory of rank order statistics: two sample case. I. *Ann. Math. Statist.* **27**, 590–615.

136 SCHMETTERER, L. (1960). On a problem of J. Neyman and E. Scott. *Ann. Math. Statist.* **31**, 656–61.

137 SICHEL, H. S. (1951–2). New methods in the statistical evaluation of mine sampling data. *Trans. Inst. Mining and Metallurgy*, **61**, 261–88.

138 SICHEL, H. S. (1966). The estimation of means and associated confidence limits for small samples from lognormal populations. *Symposium on Math. Stats. and Comp. Appl. in Ore Valuation. S.A. Inst. Min. Met. Johannesburg* (March), **106**, 1–17.

139 SICHEL, H. S. (1967). Application of statistical techniques to the evaluation of mineral deposits. *Bull. Int. Stats. Inst.* **42**, 245–65.

140 SLUTSKY, E. E. (1950). Tables for the calculation of the incomplete gamma function and the χ^2 probability function. *Moscow-Leningrad Isdat Akad. Nauk. SSR.*

141 SMITH, E. J., ADDERLEY, E. E. and BETHWAITE, F. D. (1965). A cloud seeding experiment in New England, Australia. *J. Appl. Meteorology*, **4**, 433–41.

142 SNEDECOR, G. W. (1946). *Statistical Methods.* Iowa State College Press.

143 STEVENS, W. L. (1953). Tables of the angular transformation. *Biometrika*, **40**, 70–3.

144 TARTER, M. E. (1968). Approximation of the inverse cumulative distribution function and applications. *Biometrika*, **55**, 29–41.

145 TARTER, M. E. and KRONMAL, R. A. (1968). Estimation of the cumulative by Fourier series methods and application to the insertion problem. *Proc. Ass. Comp. Mach.* **23**, 491–8.

146 TATE, R. F. (1959). Unbiased estimation: functions of location and scale parameters. *Ann. Math. Statist.* **30**, 341–66.

147 TERRY, M. E. (1952). Some rank order tests which are most powerful against specific parametric alternatives. *Ann. Math. Statist.* **23**, 346–66.

148 THONI, H. (1969). A table for estimating the mean of a lognormal distribution. *J. Amer. Stat. Assoc.* **64**, 632–6.

149 TIPPETT, L. H. C. (1934). Statistical methods in textile research. Part 2. Uses of the binomial and poisson distributions. *Shirley Inst. Mem.* **13**, 35–72.

150 TUKEY, J. W. (1949). One degree of freedom for non-additivity. *Biometrics*, **5**, 232–42.

151 TUKEY, J. W. (1957). On the comparative anatomy of transformations. *Ann. Math. Statist.* **28**, 602–32.

152 Tukey, J. W. (1958). The true purpose of transformations. *Biometrics*, **24**, 1041 (abstract).
153 Turner, M. E., Monroe, M. J. and Lucas, H. L., Jr. (1961). Generalised asymptotic regression and non-linear path analysis. *Biometrics*, **17**, 120–43.
154 Van Belle, G. (1968). Contingency table analysis by means of transformations. *Biometrics*, **24**, 1041 (abstract).
155 Williams, C. B. (1937). The use of logarithms in the interpretation of certain entymological problems. *Ann. Appl. Biol.* **24**, 404–14.
156 Williams, E. J. (1967). The analysis of association among many variates. *J. Roy. Stat. Soc.* **29**, 199–242.
157 Wilson, E. B. and Hilferty, M. M. (1931). The distribution of chi-square. *Proc. Nat. Acad. Sci.* **17**, 684–8.
158 Yates, F. (1955). The use of transformations and maximum likelihood in the analysis of quantal-experiments involving two treatments. *Biometrika*, **42**, 382–403.
159 Yates, F. (1961). Marginal percentages in multiway tables of quantal data with disproportionate frequencies. *Biometrics*, **17**, 1–9.
160 Zacks, S. (1966). Sequential estimation of the mean of a lognormal distribution having a prescribed proportional closeness. *Ann. Math. Statist.* **37**, 1688–96.
161 Zubin, J. (1936). Note on a transformation function for proportions and percentages. *J. Appl. Psychology*, **19**, 213–220.

Résumé

Cet article passe en revue les raisons qui motivent la transformation des données, les moyens d'obtenir une transformation adéquate, et les différentes transformations que l'on rencontre dons la littérature. Il considère aussi les différents problèmes qui se posent lorsqu'il s'agit de présenter sur la variable originale les résultats d'une analyse effectuée sur la variable transformée.

Les techniques statistiques standard associées au modèle linéaire reposent sur trois hypothèses fondamentales: additivité, variance constante, et normalité. Il arrive souvent que la totalité ou certaines de ces hypothèses ne sont pas valables, alors qu'elles peuvent être admissibles sur des données transformées. Théorie et méthodes concernant les techniques qui permettent de déterminer la transformation appropriée sont décrites en détail.

L'article décrit 19 transformations d'emploi général et on donne les références. Dans la pratique, toutes permettent de mieux satisfaire à l'une au moins des trois hypothèses.

L'article considère aussi les méthodes utilisables pour la transposition sur la variable originale non transformée des résultats obtenus sur la variable transformée. Il discute en particulier le problème de la recherche de l'estimateur à variance minimale, de l'erreur type qui lui est associée, et des intervalles de confiance.

La bibliographie comporte 160 références.

LINKED CITATIONS

You have printed the following article:

Transformations: An Introduction and a Bibliography
M. H. Hoyle
International Statistical Review / Revue Internationale de Statistique, Vol. 41, No. 2. (Aug., 1973), pp. 203-223.
Stable URL:
http://links.jstor.org/sici?sici=0306-7734%28197308%2941%3A2%3C203%3ATAIAAB%3E2.0.CO%3B2-J

[Footnotes]

[2] **The Transformation of Poisson, Binomial and Negative-Binomial Data**
F. J. Anscombe
Biometrika, Vol. 35, No. 3/4. (Dec., 1948), pp. 246-254.
Stable URL:
http://links.jstor.org/sici?sici=0006-3444%28194812%2935%3A3%2F4%3C246%3ATTOPBA%3E2.0.CO%3B2-W

[3] **Table of the Hyperbolic Transformation sinh -1 #x**
F. J. Anscombe
Journal of the Royal Statistical Society. Series A (General), Vol. 113, No. 2. (1950), pp. 228-229.
Stable URL:
http://links.jstor.org/sici?sici=0035-9238%281950%29113%3A2%3C228%3ATOTHT%3E2.0.CO%3B2-H

[5] **On Estimating Binomial Response Relations**
F. J. Anscombe
Biometrika, Vol. 43, No. 3/4. (Dec., 1956), pp. 461-464.
Stable URL:
http://links.jstor.org/sici?sici=0006-3444%28195612%2943%3A3%2F4%3C461%3AOEBRR%3E2.0.CO%3B2-2

[8] **The Examination and Analysis of Residuals**
F. J. Anscombe; John W. Tukey
Technometrics, Vol. 5, No. 2. (May, 1963), pp. 141-160.
Stable URL:
http://links.jstor.org/sici?sici=0040-1706%28196305%295%3A2%3C141%3ATEAAOR%3E2.0.CO%3B2-0

NOTE: *The reference numbering from the original has been maintained in this citation list.*

LINKED CITATIONS

[9] **The Square Root Transformation in Analysis of Variance**
M. S. Bartlett
Supplement to the Journal of the Royal Statistical Society, Vol. 3, No. 1. (1936), pp. 68-78.
Stable URL:
http://links.jstor.org/sici?sici=1466-6162%281936%293%3A1%3C68%3ATSRTIA%3E2.0.CO%3B2-V

[10] **Sub-Sampling for Attributes**
M. S. Bartlett
Supplement to the Journal of the Royal Statistical Society, Vol. 4, No. 1. (1937), pp. 131-135.
Stable URL:
http://links.jstor.org/sici?sici=1466-6162%281937%294%3A1%3C131%3ASFA%3E2.0.CO%3B2-1

[11] **The Use of Transformations**
M. S. Bartlett
Biometrics, Vol. 3, No. 1. (Mar., 1947), pp. 39-52.
Stable URL:
http://links.jstor.org/sici?sici=0006-341X%28194703%293%3A1%3C39%3ATUOT%3E2.0.CO%3B2-6

[12] **The Statistical Analysis of Variance-Heterogeneity and the Logarithmic Transformation**
M. S. Bartlett; D. G. Kendall
Supplement to the Journal of the Royal Statistical Society, Vol. 8, No. 1. (1946), pp. 128-138.
Stable URL:
http://links.jstor.org/sici?sici=1466-6162%281946%298%3A1%3C128%3ATSAOVA%3E2.0.CO%3B2-2

[13] **The Transformation of Data from Entomological Field Experiments so that the Analysis of Variance Becomes Applicable**
Geoffrey Beall
Biometrika, Vol. 32, No. 3/4. (Apr., 1942), pp. 243-262.
Stable URL:
http://links.jstor.org/sici?sici=0006-3444%28194204%2932%3A3%2F4%3C243%3ATTODFE%3E2.0.CO%3B2-Y

[16] **A Statistically Precise and Relatively Simple Method of Estimating the Bioassay with Quantal Response, Based on the Logistic Function**
Joseph Berkson
Journal of the American Statistical Association, Vol. 48, No. 263. (Sep., 1953), pp. 565-599.
Stable URL:
http://links.jstor.org/sici?sici=0162-1459%28195309%2948%3A263%3C565%3AASPARS%3E2.0.CO%3B2-R

NOTE: *The reference numbering from the original has been maintained in this citation list.*

LINKED CITATIONS

[18] **Conditional Expectation and Unbiased Sequential Estimation**
David Blackwell
The Annals of Mathematical Statistics, Vol. 18, No. 1. (Mar., 1947), pp. 105-110.
Stable URL:
http://links.jstor.org/sici?sici=0003-4851%28194703%2918%3A1%3C105%3ACEAUSE%3E2.0.CO%3B2-D

[19] **Transformations of the Binomial, Negative Binomial, Poisson and #2 Distributions**
Gunnar Blom
Biometrika, Vol. 41, No. 3/4. (Dec., 1954), pp. 302-316.
Stable URL:
http://links.jstor.org/sici?sici=0006-3444%28195412%2941%3A3%2F4%3C302%3ATOTBNB%3E2.0.CO%3B2-M

[24] **Permutation Theory in the Derivation of Robust Criteria and the Study of Departures from Assumption**
G. E. P. Box; S. L. Andersen
Journal of the Royal Statistical Society. Series B (Methodological), Vol. 17, No. 1. (1955), pp. 1-34.
Stable URL:
http://links.jstor.org/sici?sici=0035-9246%281955%2917%3A1%3C1%3APTITDO%3E2.0.CO%3B2-9

[25] **An Analysis of Transformations**
G. E. P. Box; D. R. Cox
Journal of the Royal Statistical Society. Series B (Methodological), Vol. 26, No. 2. (1964), pp. 211-252.
Stable URL:
http://links.jstor.org/sici?sici=0035-9246%281964%2926%3A2%3C211%3AAAOT%3E2.0.CO%3B2-6

[26] **Transformation of the Independent Variables**
G. E. P. Box; Paul W. Tidwell
Technometrics, Vol. 4, No. 4. (Nov., 1962), pp. 531-550.
Stable URL:
http://links.jstor.org/sici?sici=0040-1706%28196211%294%3A4%3C531%3ATOTIV%3E2.0.CO%3B2-C

[27] **Bias in Nonlinear Estimation**
M. J. Box
Journal of the Royal Statistical Society. Series B (Methodological), Vol. 33, No. 2. (1971), pp. 171-201.
Stable URL:
http://links.jstor.org/sici?sici=0035-9246%281971%2933%3A2%3C171%3ABINE%3E2.0.CO%3B2-I

NOTE: *The reference numbering from the original has been maintained in this citation list.*

LINKED CITATIONS

[28] **Estimation in Lognormal Linear Models**
Dan Bradu; Yair Mundlak
Journal of the American Statistical Association, Vol. 65, No. 329. (Mar., 1970), pp. 198-211.
Stable URL:
http://links.jstor.org/sici?sici=0162-1459%28197003%2965%3A329%3C198%3AEILLM%3E2.0.CO%3B2-P

[30] **Asymptotic Normality and Efficiency of Certain Nonparametric Test Statistics**
Herman Chernoff; I. Richard Savage
The Annals of Mathematical Statistics, Vol. 29, No. 4. (Dec., 1958), pp. 972-994.
Stable URL:
http://links.jstor.org/sici?sici=0003-4851%28195812%2929%3A4%3C972%3AANAEOC%3E2.0.CO%3B2-B

[31] **The Angular Transformation in Quantal Analysis**
P. J. Claringbold; J. D. Biggers; C. W. Emmens
Biometrics, Vol. 9, No. 4. (Dec., 1953), pp. 467-484.
Stable URL:
http://links.jstor.org/sici?sici=0006-341X%28195312%299%3A4%3C467%3ATATIQA%3E2.0.CO%3B2-5

[34] **The Analysis of Variance when Experimental Errors Follow the Poisson or Binomial Laws**
W. G. Cochran
The Annals of Mathematical Statistics, Vol. 11, No. 3. (Sep., 1940), pp. 335-347.
Stable URL:
http://links.jstor.org/sici?sici=0003-4851%28194009%2911%3A3%3C335%3ATAOVWE%3E2.0.CO%3B2-L

[35] **Moments and Cumulants in the Specification of Distributions**
E. A. Cornish; R. A. Fisher
Revue de l'Institut International de Statistique / Review of the International Statistical Institute, Vol. 5, No. 4. (Jan., 1938), pp. 307-320.
Stable URL:
http://links.jstor.org/sici?sici=0373-1138%28193801%295%3A4%3C307%3AMACITS%3E2.0.CO%3B2-Q

[36] **The Percentile Points of Distributions Having Known Cumulants**
Ronald A. Fisher; E. A. Cornish
Technometrics, Vol. 2, No. 2. (May, 1960), pp. 209-225.
Stable URL:
http://links.jstor.org/sici?sici=0040-1706%28196005%292%3A2%3C209%3ATPPODH%3E2.0.CO%3B2-Y

NOTE: *The reference numbering from the original has been maintained in this citation list.*

LINKED CITATIONS

[37] **Some Applications of Exponential Ordered Scores**
D. R. Cox
Journal of the Royal Statistical Society. Series B (Methodological), Vol. 26, No. 1. (1964), pp. 103-110.
Stable URL:
http://links.jstor.org/sici?sici=0035-9246%281964%2926%3A1%3C103%3ASAOEOS%3E2.0.CO%3B2-2

[39] **On Transformations Used in the Analysis of Variance**
J. H. Curtiss
The Annals of Mathematical Statistics, Vol. 14, No. 2. (Jun., 1943), pp. 107-122.
Stable URL:
http://links.jstor.org/sici?sici=0003-4851%28194306%2914%3A2%3C107%3AOTUITA%3E2.0.CO%3B2-O

[40] **Short Proof of Miss Harley's Theorem on the Correlation Coefficient**
H. E. Daniels; M. G. Kendall
Biometrika, Vol. 45, No. 3/4. (Dec., 1958), pp. 571-572.
Stable URL:
http://links.jstor.org/sici?sici=0006-3444%28195812%2945%3A3%2F4%3C571%3ASPOMHT%3E2.0.CO%3B2-H

[42] **The Probability Integral Transformation When Parameters are Estimated from the Sample**
F. N. David; N. L. Johnson
Biometrika, Vol. 35, No. 1/2. (May, 1948), pp. 182-190.
Stable URL:
http://links.jstor.org/sici?sici=0006-3444%28194805%2935%3A1%2F2%3C182%3ATPITWP%3E2.0.CO%3B2-Z

[43] **The Probability Integral Transformation when the Variable is Discontinuous**
F. N. David; N. L. Johnson
Biometrika, Vol. 37, No. 1/2. (Jun., 1950), pp. 42-49.
Stable URL:
http://links.jstor.org/sici?sici=0006-3444%28195006%2937%3A1%2F2%3C42%3ATPITWT%3E2.0.CO%3B2-R

[44] **A Quick Method for Choosing a Transformation**
James L. Dolby
Technometrics, Vol. 5, No. 3. (Aug., 1963), pp. 317-325.
Stable URL:
http://links.jstor.org/sici?sici=0040-1706%28196308%295%3A3%3C317%3AAQMFCA%3E2.0.CO%3B2-G

NOTE: *The reference numbering from the original has been maintained in this citation list.*

LINKED CITATIONS

[45] **Properties of Distributions Resulting from Certain Simple Transformations of the Normal Distribution**

J. Draper

Biometrika, Vol. 39, No. 3/4. (Dec., 1952), pp. 290-301.

Stable URL:

http://links.jstor.org/sici?sici=0006-3444%28195212%2939%3A3%2F4%3C290%3APODRFC%3E2.0.CO%3B2-D

[46] **On Distributions and Their Transformation to Normality**

N. R. Draper; D. R. Cox

Journal of the Royal Statistical Society. Series B (Methodological), Vol. 31, No. 3. (1969), pp. 472-476.

Stable URL:

http://links.jstor.org/sici?sici=0035-9246%281969%2931%3A3%3C472%3AODATTT%3E2.0.CO%3B2-V

[47] **Transformations: Some Examples Revisited**

Norman R. Draper; William G. Hunter

Technometrics, Vol. 11, No. 1. (Feb., 1969), pp. 23-40.

Stable URL:

http://links.jstor.org/sici?sici=0040-1706%28196902%2911%3A1%3C23%3ATSER%3E2.0.CO%3B2-6

[48] **Analysis of Factorial Arrangements when the Data are Proportions**

G. V. Dyke; H. D. Patterson

Biometrics, Vol. 8, No. 1. (Mar., 1952), pp. 1-12.

Stable URL:

http://links.jstor.org/sici?sici=0006-341X%28195203%298%3A1%3C1%3AAOFAWT%3E2.0.CO%3B2-T

[50] **On Additivity in the Analysis of Variance**

R. C. Elston

Biometrics, Vol. 17, No. 2. (Jun., 1961), pp. 209-219.

Stable URL:

http://links.jstor.org/sici?sici=0006-341X%28196106%2917%3A2%3C209%3AOAITAO%3E2.0.CO%3B2-D

[51] **On the Distribution of a Variate Whose Logarithm is Normally Distributed**

D. J. Finney

Supplement to the Journal of the Royal Statistical Society, Vol. 7, No. 2. (1941), pp. 155-161.

Stable URL:

http://links.jstor.org/sici?sici=1466-6162%281941%297%3A2%3C155%3AOTDOAV%3E2.0.CO%3B2-W

NOTE: *The reference numbering from the original has been maintained in this citation list.*

[52] **The Estimation of the Parameters of Tolerance Distributions**
D. J. Finney
Biometrika, Vol. 36, No. 3/4. (Dec., 1949), pp. 239-256.
Stable URL:
http://links.jstor.org/sici?sici=0006-3444%28194912%2936%3A3%2F4%3C239%3ATEOTPO%3E2.0.CO%3B2-5

[59] **On the Mathematical Foundations of Theoretical Statistics**
R. A. Fisher
Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character, Vol. 222. (1922), pp. 309-368.
Stable URL:
http://links.jstor.org/sici?sici=0264-3952%281922%29222%3C309%3AOTMFOT%3E2.0.CO%3B2-2

[62] **Gene Frequencies in a Cline Determined by Selection and Diffusion**
R. A. Fisher
Biometrics, Vol. 6, No. 4. (Dec., 1950), pp. 353-361.
Stable URL:
http://links.jstor.org/sici?sici=0006-341X%28195012%296%3A4%3C353%3AGFIACD%3E2.0.CO%3B2-K

[63] **The Analysis of Variance with Various Binomial Transformations**
Ronald Fisher
Biometrics, Vol. 10, No. 1. (Mar., 1954), pp. 130-139.
Stable URL:
http://links.jstor.org/sici?sici=0006-341X%28195403%2910%3A1%3C130%3ATAOVWV%3E2.0.CO%3B2-0

[66] **Transformations Related to the Angular and the Square Root**
Murray F. Freeman; John W. Tukey
The Annals of Mathematical Statistics, Vol. 21, No. 4. (Dec., 1950), pp. 607-611.
Stable URL:
http://links.jstor.org/sici?sici=0003-4851%28195012%2921%3A4%3C607%3ATRTTAA%3E2.0.CO%3B2-O

[68] **Fiducial Limits for the Poisson Distribution**
F. Garwood
Biometrika, Vol. 28, No. 3/4. (Dec., 1936), pp. 437-442.
Stable URL:
http://links.jstor.org/sici?sici=0006-3444%28193612%2928%3A3%2F4%3C437%3AFLFTPD%3E2.0.CO%3B2-E

LINKED CITATIONS

[70] **The Interpretation and Estimation of Cobb-Douglas Functions**
Arthur S. Goldberger
Econometrica, Vol. 36, No. 3/4. (Jul. - Oct., 1968), pp. 464-472.
Stable URL:
http://links.jstor.org/sici?sici=0012-9682%28196807%2F10%2936%3A3%2F4%3C464%3ATIAEOC%3E2.0.CO%3B2-X

[71] **Asymptotic Power of Tests of Linear Hypotheses Using the Probit and Logit Transformations**
James E. Grizzle
Journal of the American Statistical Association, Vol. 57, No. 300. (Dec., 1962), pp. 877-894.
Stable URL:
http://links.jstor.org/sici?sici=0162-1459%28196212%2957%3A300%3C877%3AAPOTOL%3E2.0.CO%3B2-%23

[73] **The Approximate Normalization of a Class of Frequency Distributions**
J. B. S. Haldane
Biometrika, Vol. 29, No. 3/4. (Feb., 1938), pp. 392-404.
Stable URL:
http://links.jstor.org/sici?sici=0006-3444%28193802%2929%3A3%2F4%3C392%3ATANOAC%3E2.0.CO%3B2-3

[74] **Moments of the Distributions of Powers and Products of Normal Variates**
J. B. S. Haldane
Biometrika, Vol. 32, No. 3/4. (Apr., 1942), pp. 226-242.
Stable URL:
http://links.jstor.org/sici?sici=0006-3444%28194204%2932%3A3%2F4%3C226%3AMOTDOP%3E2.0.CO%3B2-C

[75] **Interval Estimation of Non-Linear Parametric Functions, II**
Max Halperin
Journal of the American Statistical Association, Vol. 59, No. 305. (Mar., 1964), pp. 168-181.
Stable URL:
http://links.jstor.org/sici?sici=0162-1459%28196403%2959%3A305%3C168%3AIEONPF%3E2.0.CO%3B2-G

[76] **Interval Estimation of Non-Linear Parametric Functions, III**
Max Halperin
Journal of the American Statistical Association, Vol. 60, No. 312. (Dec., 1965), pp. 1191-1199.
Stable URL:
http://links.jstor.org/sici?sici=0162-1459%28196512%2960%3A312%3C1191%3AIEONPF%3E2.0.CO%3B2-E

NOTE: *The reference numbering from the original has been maintained in this citation list.*

LINKED CITATIONS

[77] **Interval Estimation of Non-Linear Parametric Functions**
Max Halperin; Nathan Mantel
Journal of the American Statistical Association, Vol. 58, No. 303. (Sep., 1963), pp. 611-627.
Stable URL:
http://links.jstor.org/sici?sici=0162-1459%28196309%2958%3A303%3C611%3AIEONPF%3E2.0.CO%3B2-L

[78] **Some Properties of an Angular Transformation for the Correlation Coefficient**
B. I. Harley
Biometrika, Vol. 43, No. 1/2. (Jun., 1956), pp. 219-224.
Stable URL:
http://links.jstor.org/sici?sici=0006-3444%28195606%2943%3A1%2F2%3C219%3ASPOAAT%3E2.0.CO%3B2-0

[79] **Further Properties of an Angular Transformation of the Correlation Coefficient**
B. I. Harley
Biometrika, Vol. 44, No. 1/2. (Jun., 1957), pp. 273-275.
Stable URL:
http://links.jstor.org/sici?sici=0006-3444%28195706%2944%3A1%2F2%3C273%3AFPOAAT%3E2.0.CO%3B2-Y

[80] **A Critical Table of Two-Decimal Logarithms**
M. J. R. Healy
Applied Statistics, Vol. 16, No. 1. (1967), p. 86.
Stable URL:
http://links.jstor.org/sici?sici=0035-9254%281967%2916%3A1%3C86%3AACTOTL%3E2.0.CO%3B2-N

[81] **Tables for Power-Law Transformations**
M. J. R. Healy; L. R. Taylor
Biometrika, Vol. 49, No. 3/4. (Dec., 1962), pp. 557-559.
Stable URL:
http://links.jstor.org/sici?sici=0006-3444%28196212%2949%3A3%2F4%3C557%3ATFPT%3E2.0.CO%3B2-U

[82] **A Note on Log-Linear Regression**
Dale M. Heien
Journal of the American Statistical Association, Vol. 63, No. 323. (Sep., 1968), pp. 1034-1038.
Stable URL:
http://links.jstor.org/sici?sici=0162-1459%28196809%2963%3A323%3C1034%3AANOLR%3E2.0.CO%3B2-R

NOTE: *The reference numbering from the original has been maintained in this citation list.*

LINKED CITATIONS

[85] **On the Distribution of the Expected Values of the Order Statistics**
Wassily Hoeffding
The Annals of Mathematical Statistics, Vol. 24, No. 1. (Mar., 1953), pp. 93-100.
Stable URL:
http://links.jstor.org/sici?sici=0003-4851%28195303%2924%3A1%3C93%3AOTDOTE%3E2.0.CO%3B2-B

[86] **New Light on the Correlation Coefficient and its Transforms**
Harold Hotelling
Journal of the Royal Statistical Society. Series B (Methodological), Vol. 15, No. 2. (1953), pp. 193-232.
Stable URL:
http://links.jstor.org/sici?sici=0035-9246%281953%2915%3A2%3C193%3ANLOTCC%3E2.0.CO%3B2-6

[89] **The Estimation of Variances After Using a Gaussianating Transformation**
M. H. Hoyle
The Annals of Mathematical Statistics, Vol. 39, No. 4. (Aug., 1968), pp. 1125-1143.
Stable URL:
http://links.jstor.org/sici?sici=0003-4851%28196808%2939%3A4%3C1125%3ATEOVAU%3E2.0.CO%3B2-4

[91] **An Angular Transformation for the Serial Correlation Coefficient**
G. M. Jenkins
Biometrika, Vol. 41, No. 1/2. (Jun., 1954), pp. 261-265.
Stable URL:
http://links.jstor.org/sici?sici=0006-3444%28195406%2941%3A1%2F2%3C261%3AAATFTS%3E2.0.CO%3B2-S

[92] **Systems of Frequency Curves Generated by Methods of Translation**
N. L. Johnson
Biometrika, Vol. 36, No. 1/2. (Jun., 1949), pp. 149-176.
Stable URL:
http://links.jstor.org/sici?sici=0006-3444%28194906%2936%3A1%2F2%3C149%3ASOFCGB%3E2.0.CO%3B2-1

[101] **The Performance of Some Rough Tests for Bivariate Normality before and after Coordinate Transformations to Normality**
Charles J. Kowalski
Technometrics, Vol. 12, No. 3. (Aug., 1970), pp. 517-544.
Stable URL:
http://links.jstor.org/sici?sici=0040-1706%28197008%2912%3A3%3C517%3ATPOSRT%3E2.0.CO%3B2-8

NOTE: *The reference numbering from the original has been maintained in this citation list.*

LINKED CITATIONS

[102] **Co-Ordinate Transformations to Normality and the Power of Normal Tests for Independence**
Charles J. Kowalski; Michael E. Tarter
Biometrika, Vol. 56, No. 1. (Mar., 1969), pp. 139-148.
Stable URL:
http://links.jstor.org/sici?sici=0006-3444%28196903%2956%3A1%3C139%3ACTTNAT%3E2.0.CO%3B2-L

[104] **Analysis of Factorial Experiments by Estimating Monotone Transformations of the Data**
J. B. Kruskal
Journal of the Royal Statistical Society. Series B (Methodological), Vol. 27, No. 2. (1965), pp. 251-263.
Stable URL:
http://links.jstor.org/sici?sici=0035-9246%281965%2927%3A2%3C251%3AAOFEBE%3E2.0.CO%3B2-N

[108] **Fitting Time Series Models for Prediction**
William S. Cleveland
Technometrics, Vol. 13, No. 4. (Nov., 1971), pp. 713-723.
Stable URL:
http://links.jstor.org/sici?sici=0040-1706%28197111%2913%3A4%3C713%3AFTSMFP%3E2.0.CO%3B2-S

[109] **The Lognormal Distribution and the Translation Method: Description and Estimation Problems**
Andre G. Laurent
Journal of the American Statistical Association, Vol. 58, No. 301. (Mar., 1963), pp. 231-235.
Stable URL:
http://links.jstor.org/sici?sici=0162-1459%28196303%2958%3A301%3C231%3ATLDATT%3E2.0.CO%3B2-B

[114] **Numerical Approximations to the Percentage Points of the #2 Distribution**
Maxine Merrington
Biometrika, Vol. 32, No. 2. (Oct., 1941), pp. 200-202.
Stable URL:
http://links.jstor.org/sici?sici=0006-3444%28194110%2932%3A2%3C200%3ANATTPP%3E2.0.CO%3B2-W

[115] **On the Estimation of an Exponential Function**
M. T. G. Meulenberg
Econometrica, Vol. 33, No. 4. (Oct., 1965), pp. 863-868.
Stable URL:
http://links.jstor.org/sici?sici=0012-9682%28196510%2933%3A4%3C863%3AOTEOAE%3E2.0.CO%3B2-Q

NOTE: *The reference numbering from the original has been maintained in this citation list.*

LINKED CITATIONS

[116] **Transformations to Normality Using Fractional Powers of the Variable**
P. G. Moore
Journal of the American Statistical Association, Vol. 52, No. 278. (Jun., 1957), pp. 237-246.
Stable URL:
http://links.jstor.org/sici?sici=0162-1459%28195706%2952%3A278%3C237%3ATTNUFP%3E2.0.CO%3B2-7

[117] **Interval Analysis and the Logarithmic Transformation**
P. G. Moore
Journal of the Royal Statistical Society. Series B (Methodological), Vol. 20, No. 1. (1958), pp. 187-192.
Stable URL:
http://links.jstor.org/sici?sici=0035-9246%281958%2920%3A1%3C187%3AIAATLT%3E2.0.CO%3B2-M

[118] **112. Query**
Peter G. Moore; John W. Tukey
Biometrics, Vol. 10, No. 4. (Dec., 1954), pp. 562-568.
Stable URL:
http://links.jstor.org/sici?sici=0006-341X%28195412%2910%3A4%3C562%3A1Q%3E2.0.CO%3B2-L

[119] **Critical Values of the Log-Normal Distribution**
Jack Moshman
Journal of the American Statistical Association, Vol. 48, No. 263. (Sep., 1953), pp. 600-609.
Stable URL:
http://links.jstor.org/sici?sici=0162-1459%28195309%2948%3A263%3C600%3ACVOTLD%3E2.0.CO%3B2-X

[120] **The Uses and Usefulness of Binomial Probability Paper**
Frederick Mosteller; John W. Tukey
Journal of the American Statistical Association, Vol. 44, No. 246. (Jun., 1949), pp. 174-212.
Stable URL:
http://links.jstor.org/sici?sici=0162-1459%28194906%2944%3A246%3C174%3ATUAUOB%3E2.0.CO%3B2-C

[122] **Correction for Bias Introduced by a Transformation of Variables**
Jerzy Neyman; Elizabeth L. Scott
The Annals of Mathematical Statistics, Vol. 31, No. 3. (Sep., 1960), pp. 643-655.
Stable URL:
http://links.jstor.org/sici?sici=0003-4851%28196009%2931%3A3%3C643%3ACFBIBA%3E2.0.CO%3B2-M

NOTE: *The reference numbering from the original has been maintained in this citation list.*

[123] **Precision Evaluation and Monitoring**
F. M. O'Carroll
The Statistician, Vol. 17, No. 2. (1967), pp. 127-140.
Stable URL:
http://links.jstor.org/sici?sici=0039-0526%281967%2917%3A2%3C127%3APEAM%3E2.0.CO%3B2-H

[124] **Letters to the Editor**
James Pachares; Peter J. Huber
The American Statistician, Vol. 22, No. 4. (Oct., 1968), p. 50.
Stable URL:
http://links.jstor.org/sici?sici=0003-1305%28196810%2922%3A4%3C50%3ALTTE%3E2.0.CO%3B2-3

[125] **Difficulties Involved in the Estimation of a Population Mean Using Transformed Sample Data**
R. L. Patterson
Technometrics, Vol. 8, No. 3. (Aug., 1966), pp. 535-537.
Stable URL:
http://links.jstor.org/sici?sici=0040-1706%28196608%298%3A3%3C535%3ADIITEO%3E2.0.CO%3B2-H

[127] **The Probability Integral Transformation for Testing Goodness of Fit and Combining Independent Tests of Significance**
E. S. Pearson
Biometrika, Vol. 30, No. 1/2. (Jun., 1938), pp. 134-148.
Stable URL:
http://links.jstor.org/sici?sici=0006-3444%28193806%2930%3A1%2F2%3C134%3ATPITFT%3E2.0.CO%3B2-J

[128] **Some Results in the Testing of Serial Correlation Coefficients**
M. H. Quenouille
Biometrika, Vol. 35, No. 3/4. (Dec., 1948), pp. 261-267.
Stable URL:
http://links.jstor.org/sici?sici=0006-3444%28194812%2935%3A3%2F4%3C261%3ASRITTO%3E2.0.CO%3B2-9

[132] **Notes on Bias in Estimation**
M. H. Quenouille
Biometrika, Vol. 43, No. 3/4. (Dec., 1956), pp. 353-360.
Stable URL:
http://links.jstor.org/sici?sici=0006-3444%28195612%2943%3A3%2F4%3C353%3ANOBIE%3E2.0.CO%3B2-4

NOTE: *The reference numbering from the original has been maintained in this citation list.*

LINKED CITATIONS

[133] **The Effect of Transformations of Variables Upon Their Correlation Coefficients**
M. H. Quenouille
Biometrika, Vol. 44, No. 1/2. (Jun., 1957), pp. 272-273.
Stable URL:
http://links.jstor.org/sici?sici=0006-3444%28195706%2944%3A1%2F2%3C272%3ATEOTOV%3E2.0.CO%3B2-E

[134] **On Nair's Transformation of the Correlation Coefficient**
Munuswamy Sankaran
Biometrika, Vol. 45, No. 3/4. (Dec., 1958), pp. 567-571.
Stable URL:
http://links.jstor.org/sici?sici=0006-3444%28195812%2945%3A3%2F4%3C567%3AONTOTC%3E2.0.CO%3B2-F

[135] **Contributions to the Theory of Rank Order Statistics-the Two-Sample Case**
I. Richard Savage
The Annals of Mathematical Statistics, Vol. 27, No. 3. (Sep., 1956), pp. 590-615.
Stable URL:
http://links.jstor.org/sici?sici=0003-4851%28195609%2927%3A3%3C590%3ACTTTOR%3E2.0.CO%3B2-J

[136] **On a Problem of J. Neyman and E. Scott**
L. Schmetterer
The Annals of Mathematical Statistics, Vol. 31, No. 3. (Sep., 1960), pp. 656-661.
Stable URL:
http://links.jstor.org/sici?sici=0003-4851%28196009%2931%3A3%3C656%3AOAPOJN%3E2.0.CO%3B2-E

[143] **Tables of the Angular Transformation**
W. L. Stevens
Biometrika, Vol. 40, No. 1/2. (Jun., 1953), pp. 70-73.
Stable URL:
http://links.jstor.org/sici?sici=0006-3444%28195306%2940%3A1%2F2%3C70%3ATOTAT%3E2.0.CO%3B2-D

[144] **Inverse Cumulative Approximation and Applications**
M. E. Tarter
Biometrika, Vol. 55, No. 1. (Mar., 1968), pp. 29-41.
Stable URL:
http://links.jstor.org/sici?sici=0006-3444%28196803%2955%3A1%3C29%3AICAAA%3E2.0.CO%3B2-N

NOTE: *The reference numbering from the original has been maintained in this citation list.*

LINKED CITATIONS

[146] **Unbiased Estimation: Functions of Location and Scale Parameters**
R. F. Tate
The Annals of Mathematical Statistics, Vol. 30, No. 2. (Jun., 1959), pp. 341-366.
Stable URL:
http://links.jstor.org/sici?sici=0003-4851%28195906%2930%3A2%3C341%3AUEFOLA%3E2.0.CO%3B2-2

[147] **Some Rank Order Tests which are most Powerful Against Specific Parametric Alternatives**
Milton E. Terry
The Annals of Mathematical Statistics, Vol. 23, No. 3. (Sep., 1952), pp. 346-366.
Stable URL:
http://links.jstor.org/sici?sici=0003-4851%28195209%2923%3A3%3C346%3ASROTWA%3E2.0.CO%3B2-A

[148] **A Table for Estimating the Mean of a Lognormal Distribution**
Hanspeter Thoni
Journal of the American Statistical Association, Vol. 64, No. 326. (Jun., 1969), pp. 632-636.
Stable URL:
http://links.jstor.org/sici?sici=0162-1459%28196906%2964%3A326%3C632%3AATFETM%3E2.0.CO%3B2-8

[150] **One Degree of Freedom for Non-Additivity**
John W. Tukey
Biometrics, Vol. 5, No. 3. (Sep., 1949), pp. 232-242.
Stable URL:
http://links.jstor.org/sici?sici=0006-341X%28194909%295%3A3%3C232%3AODOFFN%3E2.0.CO%3B2-1

[151] **On the Comparative Anatomy of Transformations**
John W. Tukey
The Annals of Mathematical Statistics, Vol. 28, No. 3. (Sep., 1957), pp. 602-632.
Stable URL:
http://links.jstor.org/sici?sici=0003-4851%28195709%2928%3A3%3C602%3AOTCAOT%3E2.0.CO%3B2-L

[153] **Generalized Asymptotic Regression and Non-Linear Path Analysis**
Malcolm E. Turner; Robert J. Monroe; Henry L. Lucas, Jr.
Biometrics, Vol. 17, No. 1. (Mar., 1961), pp. 120-143.
Stable URL:
http://links.jstor.org/sici?sici=0006-341X%28196103%2917%3A1%3C120%3AGARANP%3E2.0.CO%3B2-U

NOTE: *The reference numbering from the original has been maintained in this citation list.*

LINKED CITATIONS

[156] **The Analysis of Association Among Many Variates**

E. J. Williams

Journal of the Royal Statistical Society. Series B (Methodological), Vol. 29, No. 2. (1967), pp. 199-242.

Stable URL:

http://links.jstor.org/sici?sici=0035-9246%281967%2929%3A2%3C199%3ATAOAAM%3E2.0.CO%3B2-%23

[157] **The Distribution of Chi-square**

Edwin B. Wilson; Margaret M. Hilferty

Proceedings of the National Academy of Sciences of the United States of America, Vol. 17, No. 12. (Dec. 15, 1931), pp. 684-688.

Stable URL:

http://links.jstor.org/sici?sici=0027-8424%2819311215%2917%3A12%3C684%3ATDOC%3E2.0.CO%3B2-H

[158] **The Use of Transformations and Maximum Likelihood in the Analysis of Quantal Experiments Involving Two Treatments**

F. Yates

Biometrika, Vol. 42, No. 3/4. (Dec., 1955), pp. 382-403.

Stable URL:

http://links.jstor.org/sici?sici=0006-3444%28195512%2942%3A3%2F4%3C382%3ATUOTAM%3E2.0.CO%3B2-I

[159] **Marginal Percentages in Multiway Tables of Quantal Data with Disproportionate Frequencies**

F. Yates

Biometrics, Vol. 17, No. 1. (Mar., 1961), pp. 1-9.

Stable URL:

http://links.jstor.org/sici?sici=0006-341X%28196103%2917%3A1%3C1%3AMPIMTO%3E2.0.CO%3B2-T

[160] **Sequential Estimation of the Mean of a Log-Normal Distribution Having a Prescribed Proportional Closeness**

S. Zacks

The Annals of Mathematical Statistics, Vol. 37, No. 6. (Dec., 1966), pp. 1688-1696.

Stable URL:

http://links.jstor.org/sici?sici=0003-4851%28196612%2937%3A6%3C1688%3ASEOTMO%3E2.0.CO%3B2-X