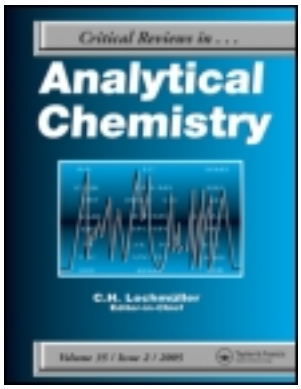

Critical Reviews in Analytical Chemistry
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/batc20

Fitting Straight Lines with Replicated Observations by Linear Regression. IV. Transforming Data

Agustin G. Asuero [a] & Julia Martin Bueno [b]

[a] Department of Analytical Chemistry, Faculty of Pharmacy, The University of Seville, Seville, Spain

[b] Department of Analytical Chemistry, Higher Polytechnic School, The University of Seville, Seville, Spain

Available online: 13 Feb 2011

To cite this article: Agustin G. Asuero & Julia Martin Bueno (2011): Fitting Straight Lines with Replicated Observations by Linear Regression. IV. Transforming Data, Critical Reviews in Analytical Chemistry, 41:1, 36-69

To link to this article: http://dx.doi.org/10.1080/10408347.2010.523589

Critical Reviews in Analytical Chemistry, 41:36–69, 2011

ISSN: 1040-8347 print / 1547-6510 online
DOI: 10.1080/10408347.2010.523589

Fitting Straight Lines with Replicated Observations by Linear Regression. IV. Transforming Data

Agustin G. Asuero[1] and Julia Martin Bueno[2]

[1]*Department of Analytical Chemistry, Faculty of Pharmacy, The University of Seville, Seville, Spain*
[2]*Department of Analytical Chemistry, Higher Polytechnic School, The University of Seville, Seville, Spain*

By transforming variables it is possible to introduce non-linear terms to the mathematical framework of linear regression. The purpose of this article is to stress the importance of transforming data in the context of linear regression analysis. However, often problems arise when people unfamiliar with mathematical statistics attempt to put this theory into practice for a certain application. Reasons for making transformations, probability plots and normality, transformations to simplify relationships, and weighting transformation data are covered in this paper. Special attention has also been paid to the Box-Cox method, i.e., transformation based on sample data observations, which is very easy to apply in practice despite its mathematical background. Statistical measurements are also re-expressed after data transformation, and a number of applications concerning the use of transformation and variance stabilization in analytical chemistry are given in tabular form. The analytical, pharmaceutical, biochemical, and clinical literature has been thoroughly revised. Variance analysis and applications on fitting straight lines with replicated observations by linear regression will be the subject of the later paper of the series.

INTRODUCTION

The importance of least squares methods (Meites, 1979; de Levie, 2000; Sayago et al., 2004; Lavagnini and Magno, 2007; Asuero and Gonzalez, 1989), nonhomogeneous variance (Sayago and Asuero, 2004; Herrador et al., 1987), weights (Tellinghuisen, 2007; Asuero and Gonzalez, 2007), correlation coefficient (Asuero et al., 2006), random error propagation law (Asuero et al., 1988), accuracy and recovery (Gonzalez et al., 1999), uncertainty (Gonzalez et al., 2005), and other topics in chemometrics (Esteban et al., 2006; Norman and Maeder, 2006) has been stressed properly in the literature. To extend the range of problems where linear models may be applied, it is useful to restore the linear structure which does not seem to apply to the raw data.

However much the analyst may wish that nature's mold be linear, the simple truth is often found in curves (Acton, 1959). Real world systems may fail to satisfy the necessary conditions required (Natrella, 1963) for the strict or even approximate validity of otherwise appropriate methods of analysis. In such situations, a transformation (change of scale) may be applied to the raw data (Mandel, 1976) in order to properly perform a conventional analysis. Any mathematical function that is applied to every point in a data set is called a transformation. Nevertheless, at a time when the number of statistical books on most subjects seems to be grown rapidly (Meier and Zünd, 2000; Graham, 1993; Miller and Miller, 2005; Einax et al., 1997; Mullins, 2003; Haswell, 1992), there are few statistical and chemometric analytical books dealing with transformations (Atkinson, 1985; Carroll and Rupper, 1988; Meloun et al., 1992; Draper and Smith, 1998; Rawlings et al., 1998) in enough detail.

Although it might be expected that the best way to estimate coefficients of a non-linear equation is to employ directly a non-linear-regression program, the statistical basis is not altogether solid and the mathematical computation is tedious (Mager, 1991): i) different final solutions may be obtained depending on the quality of the starting values and the type of internal structure of the data; ii) it is difficult to discriminate between rival models; iii) while linear regression is robust to departures from homoscedasticity, non-linear regression is relatively sensitive; iv) non-linear regression may itself produce a substantially functionally generated multi-collinearity, which is the major reason that estimates are not robust. Most approaches to non-linear regression assume that asymptotic theory can be applied. Unfortunately, the number of observations of real experiments is too small, so that the use of asymptotic estimates may cause some trouble (Mager, 1991).

Address correspondence to Agustin G. Asuero, Department of Analytical Chemistry, Faculty of Pharmacy, The University of Seville, C/Prof. García Gonzáles, No. 2. 41012–Seville. E-mail: asuero@us.es

There are some advantages in employing mathematical transformations on empirical data. They can successfully be applied to stabilize variances (to achieve homocedasticity), to get an approximate normality, a more robust model, additivity, and test approximately the type of model. (Meloun et al., 1992; Draper and Smith, 1998; Weisberg, 2005). A first stage to consider in any particular problem is whether data transformations might be necessary and useful; we could carry out a number of graphical examinations of the data (Lavagnini and Magno, 2007; Herrader et al., 1987; Barnet, 2004) or numerical-data screening activities to investigate whether the key assumptions (linearity of relationship, independence, constancy of residual variance, and normally distributed errors or residuals) are separately or jointly reasonable for the study at hand, and whether there might be anomalous observations in the form of outliers (Weisberg, 2005; Belloto and Sokolovski, 1985).

Often, informal plots of the data will clearly reveal the need of transformation of an obvious kind (such as $\ln x$ or $1/y$). In those situations, the more formal analysis may be viewed as a useful check procedure to hold in reserve (Draper and Smith, 1998). Two empirical rules are often helpful (Weisberg, 2005) in this context: the log rule and the range rule. If the value of a variable range over more than one order of magnitude and the variable is strictly positive, then replacing the variable by its logarithm is likely to be helpful. If, on the other hand, the range of a variable is considerably less than one order of magnitude, then any transformation of that variable is unlikely to be helpful. Thus, the effect of the transformation is greater the greater the ratio of $y_{\max}/y_{\min}$ considered; the effect of a power transformation $Y = y^{\alpha}$ is greater the more α differs from unity (Box and Draper, 1987).

Although linear regression is, by definition, a process of linear modelling, it is possible to introduce non-linear terms to the linear mathematical framework by transforming variables (Asuero and Gonzalez, 1989; Sayago and Asuero, 2004; Draper and Smith, 1998). When a non-linear function may be transformed into a linear one, it is said to be intrinsically linear (Sayago et al., 2004; Asuero and Gonzalez, 1987; Bates and Watts, 2007). Functions which cannot be so transformed are called intrinsically non-linear instead, i.e. a single model of the form $y = \alpha + e^{\beta x}$. Many non-linear models in the parameters can be linearized, re-expressed as a linear function of the parameters by appropriate transformations (Rawlings et al., 1998; Wang et al., 1992; Tomassone et al., 1983). For example, the relationship

$$y = \alpha x^{b} \qquad [1]$$

is linearized by taking the logarithm of both sides of the equality giving

$$\ln y = \ln \alpha + \beta \ \ln x \qquad [2]$$

Scientists who study the relationships between attributes of individuals or species call it an allometric model (Tomassone et al., 1983; Warton et al., 2006), and the value of β plays an important roll in allometric studies. The non-linear relationship between y and x is represented by the linear relationship between $y' = \ln y$ and $x' = \ln x$. In the previous example, transformations of both the response and the predictor are required to get a linear mean function. In other problems, transformation of only one variable may be desirable.

A linear transformation of y, say $Y = (y - k_0)/k_1$ where the k's are constants, is in fact a rescaling and relocation of the graduating function, and has no effect on usual normality and distribution assumptions (Box and Draper, 1987). When referring to transformation, we mean actually changing the shape of the distribution by stretching the scale in some places and compressing it in others. This cannot be accomplished simply by adding or subtracting a constant from each data value. This would simply shift the distribution without changing its shape. Transformation also cannot be accomplished by simply multiplying or dividing each data value by a constant; this would only change the scale of the data, again without changing the actual shape of the distribution.

The most common transformations involve the use of logarithms (i.e., plotting y against $\ln x$, or $\ln y$ against $\ln x$) and exponentials (Tomassone et al., 1983; Daniel and Wood, 1980; EPA, 2000). Less commonly used transformations include reciprocals, square roots, and trigonometric functions. Two or more of these functions are sometimes used in combination, especially in calibration programs supplied with commercial analytical instruments (Bysouth and Tyson, 1986).

Non-linear transformations such as the square root, log, and reciprocal of some necessarily positive response y have the effect of expanding the scale at one part of the range and contracting it at another (Box and Draper, 1987). These are all examples of simple powder transformations characterized by $Y = y^{\alpha}$, as indicated above. If α is less than unity they have the effect of contracting the range at high values and may be called contractive transformations. Power transformations with α greater than unity have the reverse effect and may be called expansive. Contractive transformation of the kind $\sqrt{y}$, $\log y$, $1/\sqrt{y}$, $1/y$ are most often needed in practical problems (Box and Draper, 1987).

If we have no prior information as to which transformation should be used, we have to resort to trial and error. The transformation that yields the best straight line when the accumulative frequency of the transformed variable is plotted on the normal probability paper is considered most suitable. The statistical tests, comparisons, confidence limits, etc., may be applied to the transformed data, and the results thus obtained may be transformed, if desired, back to the original curvilinear scale (Acton, 1959; Meloun et al., 1992). Finally, transformations may not prove to be adequate for this purpose and we may have to reject all thought of a linear model, and proceed instead to consider a range of alternative non-linear models.

REASONS FOR MAKING TRANSFORMATIONS

Two approaches for handling non-uniform variance are transfor1mation of the dependent variable and use of weighted least

squares (Asuero and Gonzalez, 1989; Sayago and Asuero, 2004; Gad, 1999); the former is probably the most common. In summary, non-normality, non-uniform variance, and to simplify the relationship between the dependent variables (response) and the independent variable(s) are the three basic reasons for transforming variables in regression (Meloun et al., 2000; Hoyle, 1973; Zorn et al., 1997; Mosteller and Youtz, 1961). However, it is unreasonable to expect that, for any type of data, there will always be a transformation so that normality of distribution, constancy of error variable, and simplicity (linearity) of the model structure may be achieved. The starting point of the study is, of course, the linearization of the model; this is a simple assumption that shows advantages (Tomassone et al., 1983). As a matter of fact, if the nature of the subject under study is well known is possible to establish a linear relationship by making a convenient transformation on either one of the x or y variables, or even both of them. We emphasize, however, that not all useful transformations will correspond to interpretable physical models. The important question posed is whether a fitted curve is an artifact or is based on an underlying mechanism for biological or physicochemical reasons.

A second assumption is the homogeneity (uniformity or homocedasticity) of y variances. When there is a theoretical or probabilistic reason to determine an optimal transformation, the examination of residuals of a previous model may help guide the choice of transformation (Sayago and Asuero, 2004; Draper and Smith, 1998; Tomassone, 1983). The variance may be related to the mean value of y (or to the value of the independent variable x) and then will not be a constant. In this case it is non-homogeneous (non-uniform or heterocedastic) and a transformation may be useful in order to be stabilized (Sayago and Asuero, 2004; Gad, 1999). The log-transformation strongly changes the distributions of errors and the originally homocedastic observation vector according to, i.e., the Arrhenius equation becomes strongly heterocedastic (Klicka and Kubácek, 1997). However, Sundberg (1998) argues theoretically for the appropriateness of using ordinary least squares on log-transformed data on parameter estimation in the Arrhenius equation. Usually, the same transformation which stabilizes the variance led to a normalization of the variable distribution. In short according to Brownlee (1984) "this is one case where usually we can both have our cake and eat it."

A transformation on the dependent variable to simplify a nonlinear relationship will destroy both homogeneous variances and normality if these assumptions were met with the original dependent variable (Natrella, 1963). Or, a transformation to stabilize variance may cause non-normality. Fortunately, as indicated above, transformation for homogeneity of variance and normality tend to go hand-in-hand so that often both assumptions are more nearly satisfied after an appropriate transformation. If one must make a choice, stabilizing variance is usually given precedence over improving normality. In any case, "the gods who favor statisticians have frequently ordained that the world be well behaved, and so we often find that a transformation to obtain one of these *desiderata* in fact achieves them all (well, almost achieves them!)"(Acton, 1959).

PROBABILITY PLOTS AND NORMALITY

An easy way to graphically check for normality would be to construct a plot with all data points lying on a straight line if the sample were drawn from a normal distribution $N(\mu,\ \sigma^2)$: i.e., a scatter plot of the ordered sample values $y_1 \leq y_2 \leq \ldots \leq y_n$ (arranged in ascending order) versus $z_1 \leq z_2 \leq \ldots \leq z_n$ where z_i is given by $z_i = F^{-1}(p_i), i = 1, 2, \ldots, n$ (i denotes the ranking or order number), and p_i is a reasonable empirical estimate of the cumulative distribution function $F[(y_i - \mu)/\sigma]$.

If a set of data is normally distributed with $N(\mu,\ \sigma^2)$, the area from $-\infty$ to some point y is given by the corresponding cumulative density function (Brownlee, 1984)

$$F(y) = \int_{-\infty}^{y} f(y)\, dy = \int_{-\infty}^{y} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2\right) dy \qquad [3]$$

A plot of 100 $F(y)$ versus $y_i (i = 1, 2, \ldots, n)$ called the cumulative probability curve is approximately S-shaped (Fig. 1). It begins at $(-\infty,\ 0)$ and passes through points such as $(\mu - 1.96\,\sigma,\ 2.5)$, $(\mu - \sigma,\ 15.87)$, $(\mu,\ 50)$, $(\mu + \sigma,\ 84.13)$, $(\mu + 1.96\,\sigma,\ 97.5)$, and $(\infty,\ 100)$ and in general through the points $(\mu + z\sigma,\ F(\mu + z\sigma))$ where z is called the normal score (Box and Draper, 1987). Now, to determine whether the

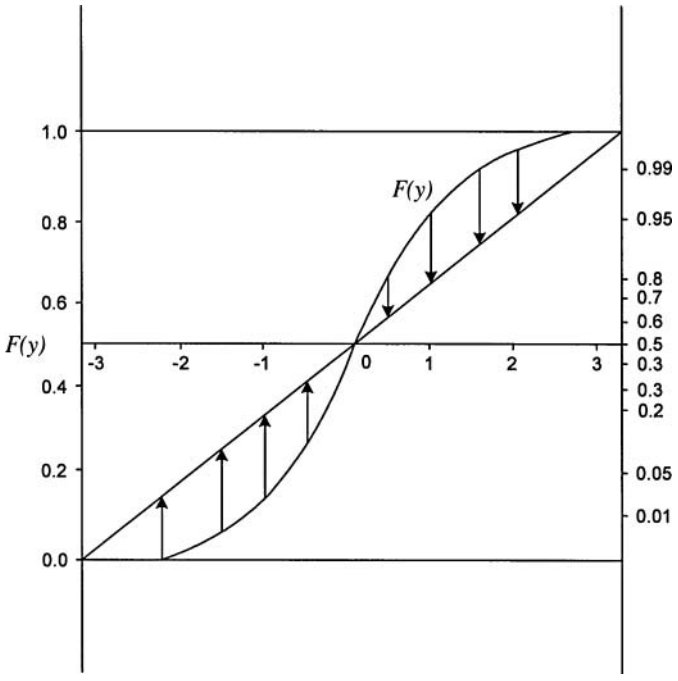

FIG. 1. Graph of the standard normal cumulative distribution function on cartesian and probability scales, showing how $F(y)$ is transformed to a straight line via scaling.

observations are normally distributed, we can check whether the empirical cumulative distribution function resembles the S-shaped curve that we expect from a normal distribution. A course of such shape is not convenient to use and it is preferable to rectify it. This is achieved by the use of probability paper. The probability graphics was described for the first time by P. Henry in 1894 (Phillippe, 1967), and rediscovered later in 1914 by Hazen (1914).

It is usually difficult for us to recognize non-linear patterns and a curve of such, shape is not convenient to use (Mage, 1982). The human eye is much better at recognizing linear tendencies. However, plotting 100 $F(y)$ versus y on normal probability paper straightens out (du Toit et al., 1986) the S-shaped curve. So, normal probability paper is specially prepared paper in which one of the scales (the vertical) represents a linearization of the cumulative normal distribution (Fig. 1). The resulting scale is non-linear with graduations bunched in the middle and increasing in each direction. Since the normal distribution is symmetrical, the spacing is symmetrical about the ordinate of 50%.

A variety of probability papers are available from private companies such as Kenffel & Esser Co. (reference 46-8003), Technical and Engineering Aids for Management (TEAM) Co. (reference 3211), Codex 3227, or Dietzgen 340-PS-90. Personal computers can make these plots using commercial software statistical programs. The plotting is facilitated if probability paper corresponding to $F[(y_i - \mu)/\sigma]$ is available, but such paper is by no means essential. The use of probability paper is unnecessary when the expected values of the order statistics are directly available.

Cumulative frequencies are obtained by adding for each class frequencies up to that point, divided by the total number of data in the data tests. For mathematical reasons (Barnett, 1975), the data quantile (standardized normal scores) order z_i of each ranked data value is defined slightly differently from the data cumulative frequencyi/n. If normal probability paper is available, z_i values can be plotted against the y_i where z_i is usually chosen from (Mage, 1982; du Toit et al., 1986; Barnett, 1975) the following possibility

$$i/(n+1)$$
$$(i-0.5)/n$$
$$(i-0.3)/(n+0.4)$$
$$(i-0.375)/(n+0.25)$$
$$(i-0.33)/(n+0.33)$$

If the unit area under the normal curve is divided into n equal sections it can be expected that if normality holds, one observation from the total of n observations will fall in each section. For $z_i = (i-0.5)/n$ (Graham, 1993; Taylor, 1990) this consequently means that y_i is plotted against the midpoint $(i-0.5)/n$ of the accumulative area of the ith section. The factor 100 adapts this to the vertical scale given in normal probability paper. It will be noted in particular that the simple choice $(i-0.5)/n$ which results in some bias performs quite well, whereas $p_i = (i-3/8)/(n+1/4)$ (Ryan et al., 1985; Mandel, 1964) leads to a practically unbiased estimator of σ. The ISO rules select this later (International Organization of Standards, 1997). BMDP uses $p_i = (i-0.33)/(n+0.33)$ in the 5-D program (Chatterjee and Hadi, 1988). Interestingly, each of the different definitions can be justified on different theoretical grounds. The differences between these different systems are typically unimportant in practical use.

If the normality assumption holds, this scatter plot should also be an approximate straight line. The (rough) estimate of the mean μ is simply the abscissa value associated with $F(y) = 50$. The difference $F(y)_{84.13} - F(y)_{50}$ is an (rough) estimate of σ, which also may be estimated as $2/5[F(y)_{90} - F(y)_{10}]$. The graphical representation is sensitive in detecting such features as differences in symmetry (or asymmetry), length of tails, and the existence of outliers.

The ordered values are random variables and are not independent since they have been ranked and satisfy the constraints$y_1 \leq y_2 \leq \ldots \leq y_n$; the mere act of ordering the data destroys their statistical independence (Feinberg, 1996). Therefore, if one point is above the line there is a good chance that the next one will be above the line. Hence, the plotted points will not be randomly scattered about the line. There could be significant runs above and below the line even though the data came from a normal population. On the other hand the variances of the extreme points are much higher than the points in the middle of the plot and when drawing a line through the data the center points should be given more weight than the extremes. Thus, one major drawback of the probability plotting technique is its lack of objectivity. It is relatively easy to reject an assumed model which is quite disparate from the data, but it is more difficult to make a decision for borderline cases. So, the plot may be supplemented with an objective procedure which allows one to specify a Type I error and to assess the risk involved in rejecting the model erroneously (Bayne and Rubin, 1986; Shapiro, 1990; Kateman and Buidens, 1993).

The normal probability plot of the data will work reasonably well if the sets are not too small (Taylor, 1990), that is to say for more than 10 points, and ideally for a much larger number of points. The advantages inherent on the use of probability plots have been outlined by Filliben (1975). Probability plots, however, are incapable of distinguishing (Filliben, 1977) likely alternative error distributions unless $n \geq 50$ (Filliben, 1977); moderate outliers may easily be missed if $n \leq 20$. The problem is compounded because the most commonly used methods for testing assumptions and estimating parameters (and their uncertainty) rely in turn on normality and absence of bad data.

Examining the residuals may also check the assumption of normality and it will be more likely to be appropriate (Bates and Watts, 2007) if the order of experiments is randomized. Probability plots are usually constructed in computers by plotting the ordered observations against the expected values of the order statistics for a random sample of the same size from the specified

distribution of interest (often normal). The vertical axis is often converted to a normal score, that is, the normal deviate value that would correspond to the plotted probability level. Such linear plots are sometimes called theoretical quantile-quantile plots as they essentially plot the observed quantiles (the ranked data) against theoretical quantiles.

The *pth* quantile (or the percentile) is defined to be the value of x below which $p\,\%$ of the sample value lies. The *pth* quantile separates the order statistics into two parts so that each contains the required percent of the sample elements, $p\,\%$ and $(100 - p)\%$. If possible, graphical displays such as box plots and normal quantile plots should be supplemented with more objective statistical procedures when an assumption of normality is being investigated (Kateman and Buidens, 1993).

MOMENTS OF A RANDOM VARIABLE: SKEWNESS AND KURTOSIS

The most commonly used measures of skewness and kurtosis are based upon moments about the mean. The term "moment" originates in mechanics and is used to denote a measure of the tendency of a force to cause rotation of an object about a center point. For samples less than 20 in size ($n \leq 20$), the hypothesis about the form of distribution may be tested (Aknazarova and Kafarov, 1982) by simple tests based on a comparison of population distribution parameters with their estimates formed from sample values. We shall use (Table 1) the notation μ_j to denote the jth moment about the origin (raw moment) and m_j to denote the jth moment about the mean, μ, in the following for probability models.

Important measures of skewness and kurtosis (from the Greek word kyrtos, meaning curved) in a population are the third and fourth moments about the mean (central moments), i.e., the average values of $(y - \mu)^3$ and $(y - \mu)^4$over the whole population. To render these measures scale invariant they are divided by σ^3 and σ^4, respectively (σ is the standard deviation). The coefficient of skewness, γ_1, is a measure (Meloun and Militki, 1995) of the symmetry of the distribution of a random variable with respect to the mean

$$\gamma_1 = \frac{E\,(y-\mu)^3}{\sigma^3} = \frac{m_3}{\sigma^3} = \frac{m_3}{m_2^{3/2}} = \sqrt{\beta_1} \qquad [4]$$

It is a pure number that characterizes only the shape of the distribution. For a unimodal (those with a single hump) symmetric distribution, i.e., the normal one, γ_1 is given equal to zero. A negative value is due to skewness toward lower (left-skewed) values and mode > median > mean, while a positive value (right-skewed) indicates excess higher values (Groeneveld and Meeden, 1997) and mean > median > mode (Fig. 2). In a symmetrical population, mean, median, and mode coincide.

It is, thus, natural to take distance from mean to mode or mean to median as measuring the skewness of the distribution. K. Pearson (Havilcek and Crain, 1988) proposed for measure: (mean – mode)/σ. However, the skewness value may be zero for non-symmetrical distributions, so care must be exercised in interpreting its value. If skewness is to be changed into symmetry, then it will be necessary to distort the scale differently in the two tails of the distribution. A transformation will be required to be capable of doing this (Natrella, 1963; Mandel, 1976; Meloun et al., 1992; Box and Draper, 1987; Gad, 1999).

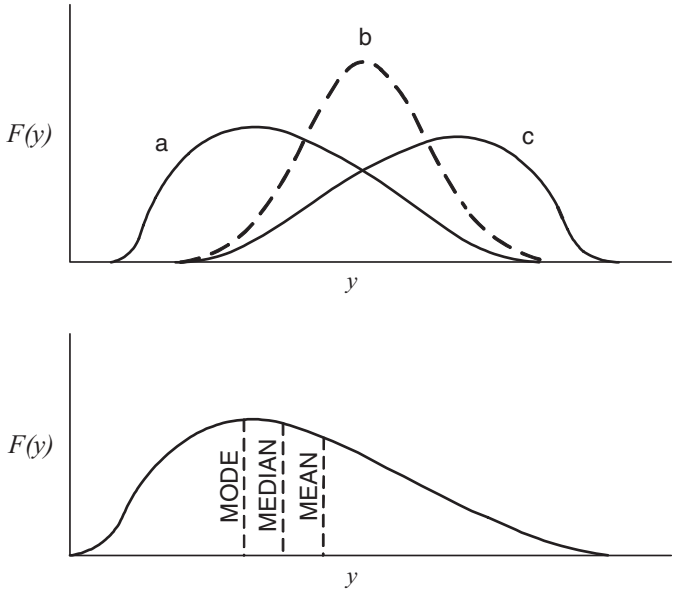

FIG. 2. Top: the coefficient of skewness indicates the direction of the preponderant tail of a distribution. A) right-skewed $\sqrt{\beta_1} > 0$; b) symmetrical $\sqrt{\beta_1} = 0$; c) left-skewed $\sqrt{\beta_1} < 0$. Bottom: typical locations of the mode, median, and mean for a right skewed probability distribution.

TABLE 1
Moments of a Random Variable for Probability Models

Discrete with values $\{y_i\}$ and probability p_i	Continuous with probability function $f(x)$
$\mu_j = \sum (y_i - \mu)^j p_i$	$\mu_j = \int_{-\infty}^{\infty} (y-a)^j f(x)dx$
Raw moment, $\mu = 0$	Central moments (a = mean): in terms of the raw moments
$\mu_1 = E(y)$: arithmetic mean $E(y) = \sum_{i=1}^{n} y_i p_i$	$E(y) = \int_{-\infty}^{\infty} y\, f(y)\, dy$
	$\mu_1 = 0$
	$\mu_2 = \alpha_2 - \alpha_1^2$
	μ_3, skewness: $\mu_3 = \alpha_3 - 3\alpha_1\alpha_2 + 2\alpha_1^2$
	μ_4, peakedness or kurtosis: $\mu_4 = \alpha_4 - 4\alpha_1\alpha_3 + 6\alpha_1^2\alpha_2 - 3\alpha_1^4$

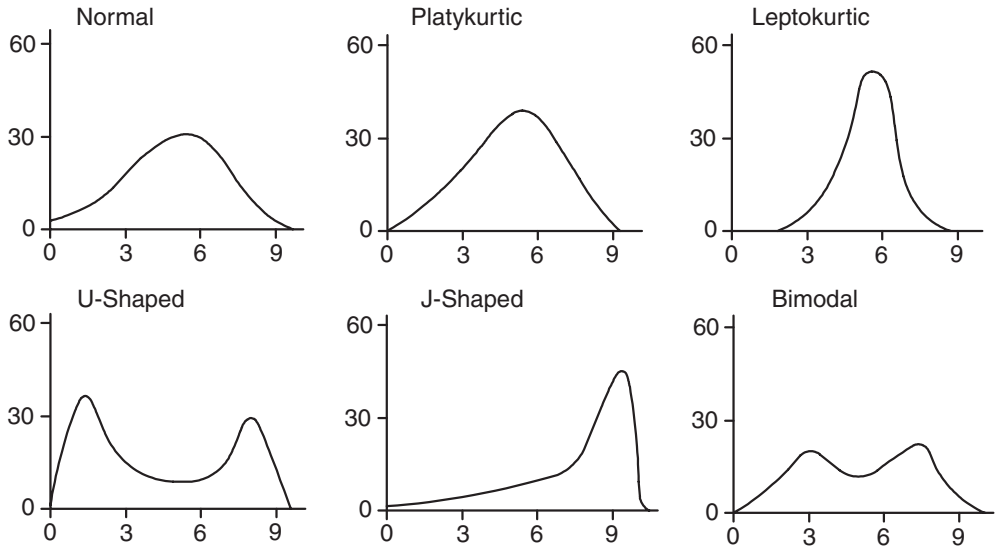

FIG. 3. Frequency distribution of different shapes.

The coefficient of kurtosis is the standardized fourth moment

$$\gamma_2 = \frac{E\,(y-\mu)^4}{\sigma^4} = \frac{m_4}{\sigma^4} = \frac{m_4}{m_2^2} = \beta_2 \qquad [5]$$

and is more difficult (Havilcek and Crain, 1988; Moors, 1986; Chissan, 1970) to interpret. It is a statistical measure of the concentration of the distribution around the mean. The resulting distribution is leptokurtic (peak at the center) if the majority of the values cluster around the center, i.e., the Laplace (double exponentially) and logistic distributions, which are termed super gaussian, and platykurtic (platy means flat in Greek) if the values lend to scatter away from the center (Fig. 3), i.e., continuous or discrete uniform distributions and Bernoulli distributions, termed subgaussians. The normal distribution ($\gamma_2 = 3$) is the reference point for determining kurtosis and is referred to as being mesokurtic. A few other well known distributions can be mesokurtic, depending on parameter values, i.e., the binomial distribution is mesokurtic for $p = 1/2 \pm \sqrt{1/12}$ (where p is proportion).

Although kurtosis has been used as a measure (Darlington, 1970; Horn, 1983) of peakedness (the relation of its height to its width), due to its vague definition, it may be as though another quantity that characterizes a distribution without necessarily giving it a geometric interpretation. Peakedness is more difficult to determine from inspection because this characteristic may result from the choice of dimensions for the vertical and horizontal scales (Havilcek and Crain, 1988). Ruppert has quoted (Ruppert, 1987) that "there is no agreement on what kurtosis measures."

For observed data we denote the moments about the origin and about the mean by the following general relationships, respectively

$$\bar{y}_j = \frac{1}{n}\sum_{i=1}^{n} y_i^j \qquad [6]$$

$$s_j = \frac{1}{n}\sum_{i=1}^{n} (y_i - \bar{y})^j \qquad [7]$$

and then taking into account the properties of expectation (Brownlee, 1984); the estimates g_1 and g_2 of the coefficients of skewness and kurtosis obtained from Eqs. [4] and [5], respectively, which are dimensionless, are given by

$$g_1 = \frac{\sqrt{n}\sum (y_i - \bar{y})^3}{\left[\sum (y_i - \bar{y})^3\right]^{3/2}} = \sqrt{b_1} = \frac{s_3}{s_2^{1.5}} = \frac{\bar{y}_3 - 3\bar{y}\ \bar{y}_2 + 2\bar{y}^3}{\left(\bar{y}_2 - \bar{y}^2\right)^{3/2}} \qquad [8]$$

$$g_2 = \frac{n\sum (y_i - \bar{y})^4}{\left[\sum (y_i - \bar{y})^2\right]^2} = b_2 = \frac{s_4}{s_2^2} = \frac{\bar{y}_4 - 4\bar{y}\ \bar{y}_3 + 6\,\bar{y}^2\bar{y}_2 - 3\bar{y}^4}{\left(\bar{y}_2 - \bar{y}^2\right)^2} \qquad [9]$$

(note that s, the estimated standard deviation, is here defined with the denominator n and not with "$n-1$"). The right hand expressions of Eqs. [8] and [9] express the result in terms of the moments about the origin and have an important computational advantage: one can proceed through a list of observations without first calculating the mean. Thus, these expressions often are called "one pass" formulas (Ott, 1995) and often are useful in certain computer programs and in hand calculations.

TABLE 2
Limits for the Skewness Factor, g_1, in the Case of a Normal Distribution (Taylor, 1990)

	Probability level	
Size of simple (n)	5 %	1 %
5	–1.058 to 1.058	–1.397 to 1.397
10	– 0.950 to 0.959	– 1.342 to 1.342
15	–0.802 to 0.802	– 1.275 to 1.275
20	–0.777 to 0.777	– .152 to 1.152
25	–0.711 to 0.711	– 1.061 to 1.061

Since s_3 and s_4 are developed (Eqs. [8] and [9]) from the cube and the fourth power, respectively, of the difference between each observation and the mean, a single unusual observation (i.e., an outlier) can exert considerable influence on both estimates. Consequently, the statistics g_1 and g_2 are very sensitive to occasional extreme values found among the observations; they are not considered (Press et al., 1999; Rousseew, 1991) as robust statistics.

For small data sets, one often gets values of g_1 and g_2 different from 0 and 3, respectively. Statistical tests are available to help judge the significance of any observed departure. Transgression of these bounds would indicate a non-normal distribution at the probability level of the test. Tables 2 and 3 contain percentiles for the tails of the distribution of the standardized third and fourth moments for tests for departure from normality (Sachs, 1982) for several sizes of samples. Although a passed test for skewness and kurtosis doesn't necessarily prove that the distribution is normal, a failed test obviously indicates that normal statistics should be applied with care (Rice, 1988).

The distributions of the above estimates are very complicated and little investigated. The respective asymptotic variances, however, are well known (Aknazarova and Kafarov, 1982; Meloun and Militki, 1995)

$$s_{g_1}^2 \approx \frac{6\,(n-2)}{(n+1)\,(n+3)} \qquad [10]$$

$$s_{g_2}^2 \approx \frac{24n\,(n-2)\,(n-3)}{(n+1)^2\,(n+3)\,(n+5)} \qquad [11]$$

If we know $s_{g_1}^2$ and $s_{g_2}^2$, we can readily check if the estimated coefficients of skewness and excess differ significantly from zero and three, respectively. If

$$|g_1| \leq 3\ s_{g_1} \qquad [12]$$

$$|g_2| \leq 5\ s_{g_2} \qquad [13]$$

the empirical distribution may be taken (Aknazarova and Kafarov, 1982) as being normal.

For great samples ($n > 50$), g_1 and g_2 are approximately normally distributed with standard errors of $\sqrt{6/n}$ and $\sqrt{24/n}$, as we may easily derive by taking limits in Eqs. [10] and [11], respectively.

Measures of skewness and kurtosis provide quantitative descriptions of non-normal variations and are helpful in determining whether a distribution departs too much from normality to be analyzed by standard parametric methods. Other more reliable tests for normality are summarized in Table 4 (Bayne and Rubin, 1986; Kateman and Buildens, 1993; Aknazarova and Kafarov, 1982; Sachs, 1982). All of these methods (Kateman and Buidens, 1993) are based on cumulative distribution functions except the Shapiro-Wilk method (Shapiro and Wilk, 1965), which is recommended by several authors since this test shows the best (Royston, 1995; Royston, 1982) overall power. A test for combined sample skewness and kurtosis is available:

$$C_1 = \frac{g_1^2}{s_{g_1}^2} + \frac{(g_2-3)^2}{s_{g_2}^2} \qquad [14]$$

For a normal distribution, the test criterion C_1 has approximately the $\chi^2_{1-\alpha}$; the null hypothesis about normality of the sample may test against disturb statistical testing (Chambers et al., 1983; Jarque and Bera, 1987). In this test, the super-normality effect

TABLE 3
Lower and Upper Percentiles of the Standardized 3rd and 4th Moments, $\sqrt{\beta_1}$ and β_2, for Test for Departure from Normality (Darlington, 1970)

	Skewness $\sqrt{\beta_1}$			Kurtosis, β_2					
	Upper percentiles			Lower percentiles			Upper percentiles		
Size of sample n	10 %	5 %	1 %	1 %	5 %	10 %	10 %	5 %	1 %
7	0.787	1.008	1.432	1.25	1.41	1.53	3.20	3.55	4.23
10	0.722	0.950	1.397	1.39	1.56	1.68	3.53	3.95	5.00
15	0.648	0.862	1.275	1.55	1.72	1.84	3.62	4.13	5.20

Since the sampling distribution of $\sqrt{\beta_1}$ is symmetrical about zero, the same values, with negative sign, correspond to the lower percentiles

TABLE 4
Principal Methods for Normality Testing

Skewness (index)
Chi-square
Lilliefords
Kolmogorov D^+ (Kolmogorov one-sample statistics)
Kolmogorov D (modified Kolmogorov, E_n)
Kuiper
Cramer-von Mises (Cramer-Smirnov)
Watson
Anderson-Darling
Shapiro-Wilk

(symmetric distribution with positive kurtosis) of small samples may disturb against statistical testing.

Compared with the problem of non-constant variance, chemists have paid less attention (Phillips and Eyring, 1983) to the normality assumption underlying the least squares method.

TRANSFORMATIONS TO SIMPLIFY RELATIONSHIPS

The majority of the published work on transformation is concerned with transforming the y's to achieve simplicity when the necessary assumptions could not otherwise be realistically made (Box and Tidwell, 1962; Draper and Hunter, 1969). Many recommend that simplifying the relationship should take precedence over all. This in fact depends both on the intrinsic value and the general acceptance of the non-linear relationship being considered. If a transformed model is meaningful and is readily interpreted, a transformation to linearize the model would not be wise if it creates heterogeneous variance or non-normality (Natrella, 1963). A basic rule of science says that, all the other things being equal, the simplest model that describes the observed behavior of the system should be adopted (Sayago and Asuero, 2004; Rawlings et al., 1998).

Two situations dealing with transformations to simplify relationships may be mainly envisaged (Draper and Smith, 1998; Rawlings et al., 1998). When there is no prior idea of the model fitting to the data, the objective is to empirically establish mathematical forms of the dependence between the variables in order that the observed relationship may be represented in the simplest form, i.e., a straight line. Note that the model is to be linear in the parameters (Lavagnini and Magno, 2007); only the form in which the variables are expressed is being considered.

If a previous knowledge of the system, however, suggests a non-linear mathematical function, non-linear in the parameters, for relating the dependent variable to the independent one (s), the purpose of the transformation is to re-express the non-linear model in a form that is linear in the parameters and for which ordinary least squares can be used. Such linearization of non-linear models is not always possible but when it is possible the transformation to be used is dictated by the functional form of the model.

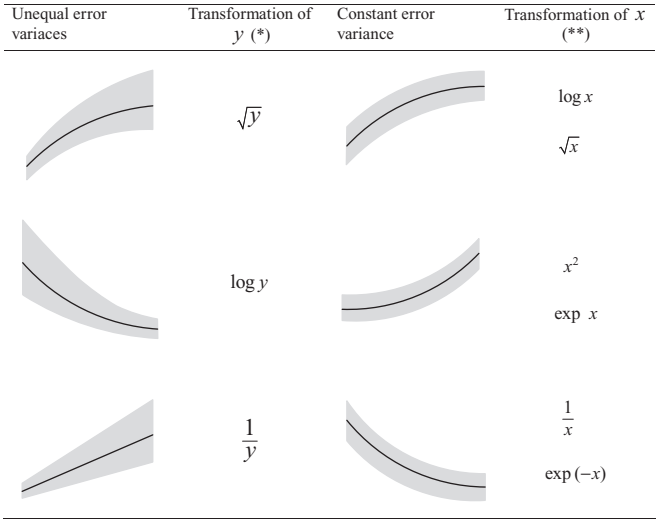

FIG. 4. Prototype regression patterns with unequal error variances and simple transformations of y and prototype non-linear regression patterns with constant variance and simple transformations of x; (*) a simultaneous transformation on x may also be helpful or necessary. (**) Linearize the regression relation without affecting the distribution of y.

Transformations for symmetry is carried out by a simple power transformation (which does not retain the scale, is not always continuous, and is suitable only for positive values) (Gad, 1999). The power family of transformations $x^* = x^k$ or $y^* = y^k$, referred to as the "one bend" transformations (Rawlings et al., 1998), provides a useful set of transformations for "straightening" a single bend in the relationship between two variables. Note that it can be used on either x or y (Fig. 4) (Neter et al., 1996). By ordering the transformations according to the exponent k we obtain a sequence of power transformations, named by Mosteller and Tukey (Tukey, 1977; Mosteler and Tukey, 1977) ladder *of reexpressions* (Table 5). The power $k = 1$ implies no transformation whereas the power transformation $k = 0$ is reinterpreted as the logarithmic transformation as we will show later.

Figure 4-left contains some relations where the skewness and the error variance increase with the mean response, with some simple transformations on y that may be helpful (Neter et al., 1996). When the distribution of the error is close to the normal one with constant variance, transformation of x should be attempted as exemplified in Fig. 4-right (Neter et al., 1996). A transformation of y in this case would change the normal distribution of the error term leading also to heterocedasticity (differing error term variances). Scatter plots and residual plots based on each transformation should then be prepared and analyzed to decide which transformation is most effective.

The rule for straightening a one bend relationship is to move up or down the ladder of transformations according to the direction in which the bulge of the curve of y versus $x$points. How

TABLE 5
Tukey's Ladder of Transformations (for choosing a function to change a distributions's shape)

Need to correct	Strength of transformation	Mathematical function	k exponent value $y* = y^k$
Positive skew	Stronger	$-\frac{1}{y^2}$	-2
	Mild	$-\frac{1}{y}$	-1
	"	$\ln y$	0
	"	$\sqrt{y}$	$1/2$
No shape change	—	y	1
Negative skew	Mild	y^2	2
	"	y^3	3
	Stronger	$\exp y$	

far one moves on the ladder of transformations depends on the sharpness of the curvature (Rawlings et al., 1998). When only one independent variable is involved, this is easily determined by trying several transformations on a few observations covering the range of the data and then choosing that transformation which makes the points most nearly collinear.

The transformations y^3, y^2, $\sqrt{y}$, $\ln y$, $-1/\sqrt{y}$, and $-1/y$ are among those often used, where y denotes an observed value of the response variable (Tukey, 1957). The first two shorten the tail of a left-skewed distribution and the last four shorten the tail of a right-skewed distribution. Using -1 in the numerators of the last two transformations preserves the order of the y values. The $\ln y$ and $\sqrt{y}$ transformations are often used to correct non-normality (Fig. 5). If any of the sample values are negative, a positive constant of sufficient size may be added to each sample value before the transformation is applied. If some of the original values are zero, it is customary to add a small quantity to make the data value non-zero as the logarithm of zero does not exist. The size of the small quantity depends on the magnitude of non-zero data and the consequences of potentially erroneous inference from the resulting transformed data. As a working point, a value of one tenth the smallest non-zero value could be selected (EPA, 2000).

Logarithms to any base can be used, but common logarithms (to the base 10) are generally the easiest. These logarithms differ only by a constant factor r, and so only the scale of the numbers involved is affected, not the basic nature of the subsequent analysis. It should be noted that the square root transformation overcorrects when very small values and zero appear in the original data. In these cases, $\sqrt{x+1}$ is often used as a transformation (Gad, 1999).

Very large data values are made much smaller, in applying the square root transformation, condensing and shrinking the scale, whereas small data values are made larger, expanding the scale for smaller values. In the case of logarithm transformation very large data values are made much smaller, moving them close together, while very small data values are spread a part from each other. This is a lot like the effect of the square root, but much stronger (Gad, 1999).

When data exhibit a few elevated values such that the frequency distribution is "skewed" with a long right tail, estimates based on the assumption of normality do not apply. In practice, transformations of data are used to stabilize variance and bring about normality. Perhaps the most commonly used transformation in this situation is (Gibbons, 1994; Aroian, 1941) the natural log transformation, where x is a lognormal random variable, such that $y = \ln x \sim N(\mu,\ \sigma^2)$. Then, the normal-theory procedures could be applied to the transformed data.

The log transformation is likely to work well if the ratio of the standard deviation to the mean is similar among several groups of observations (Altman, 1991) (constant coefficient of variation over all possible data values). It is evident that y cannot be negative, since the logarithm of a negative number is not defined. This fact has contributed, albeit probably in a small measure, to the use of the lognormal distribution when a finite probability of a negative value is physically absurd. For non-normal distributions, it is common in water quality work to assume that the log-normal distribution applies so that the logarithms of the raw data can be used in a Gaussian likelihood (Gibbons, 1994). In particular, it is important not to apply a transformation such as the logarithmic form when one is not needed.

The square root transformation is less dramatic than taking logs. It is particularly used when the variables is a count (frequency) and thus would be expected to follow a Poisson distribution (Shumway, 2002). The reciprocal transformation has a much more drastic effect than taking logs (note that it reverses the order of observations), and may be useful if the observed data have an extremely skewed distribution. Their use is not common, however, and there are certain reasons for using the log transformation in preference to any other as long as it yields satisfactory results.

TRANSFORMATIONS TO LINEARIZE THE MODEL

Various non-linear models can be written in a linear form through appropriate transformations. Table 6 gives a number of non-linear models corresponding to the most usual cases found together with the transformations for converting them into linear ones (Tomassone et al., 1983; Daniel and Wood, 1980, du Toit et al., 1986; Bayne and Rubin, 1986; de Levie, 2004). It should be noted that the use of these transformations is certain to accomplish one thing only, i.e., to yield a straight-line form. The transformed data, however, will not necessarily satisfy certain assumptions which are theoretically necessary in order to apply the least squares method (Belloto and Sokolovski, 1985; Bates and Watts, 2007; Seber, 1977). In these examples, the placement of the error in the original model was such that the transformed model had an additive error. If there were reasons

TABLE 6
Non-linear Function That can be Written in a Linear Form by Means of a Transformation

Function	Formula	Transformation	Linear form
Power function	$y = \alpha x^b$	$y' = \log y \quad x' = \log x$	$y' = \log \alpha + \beta x'$
Exponential grown model	$y = \alpha e^{\beta x}$	$y' = \log y$	$y' = \log \alpha + \beta x$
Logarithmic	$y = \alpha + \beta \log x$	$x' = \log x$	$y = \alpha + \beta x'$
Hyperbolic	$y = \dfrac{x}{\alpha x - \beta}$	$y' = 1/y \quad x' = 1/x$	$y' = \alpha - \beta x'$
Logit	$y = \dfrac{e^{\alpha+\beta x}}{1 + e^{\alpha+\beta x}}$	$y' = \log\left(\dfrac{y}{1-y}\right) = 2\tanh^{-1}(2y-1)$	$y' = \alpha + \beta x$

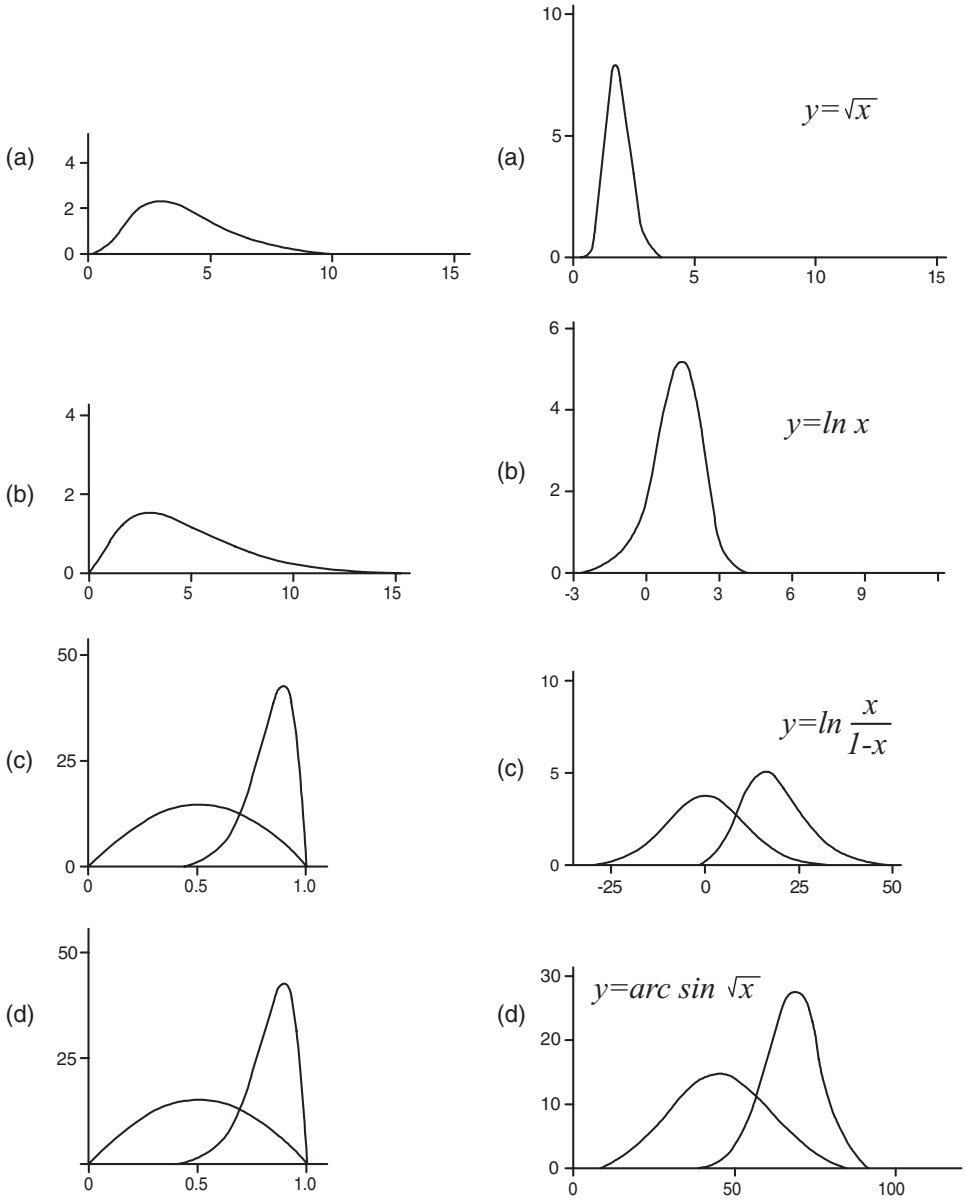

FIG. 5. Normalizing effect of some frequently used transformations. Left: original distributions. Right: transformed distributions.

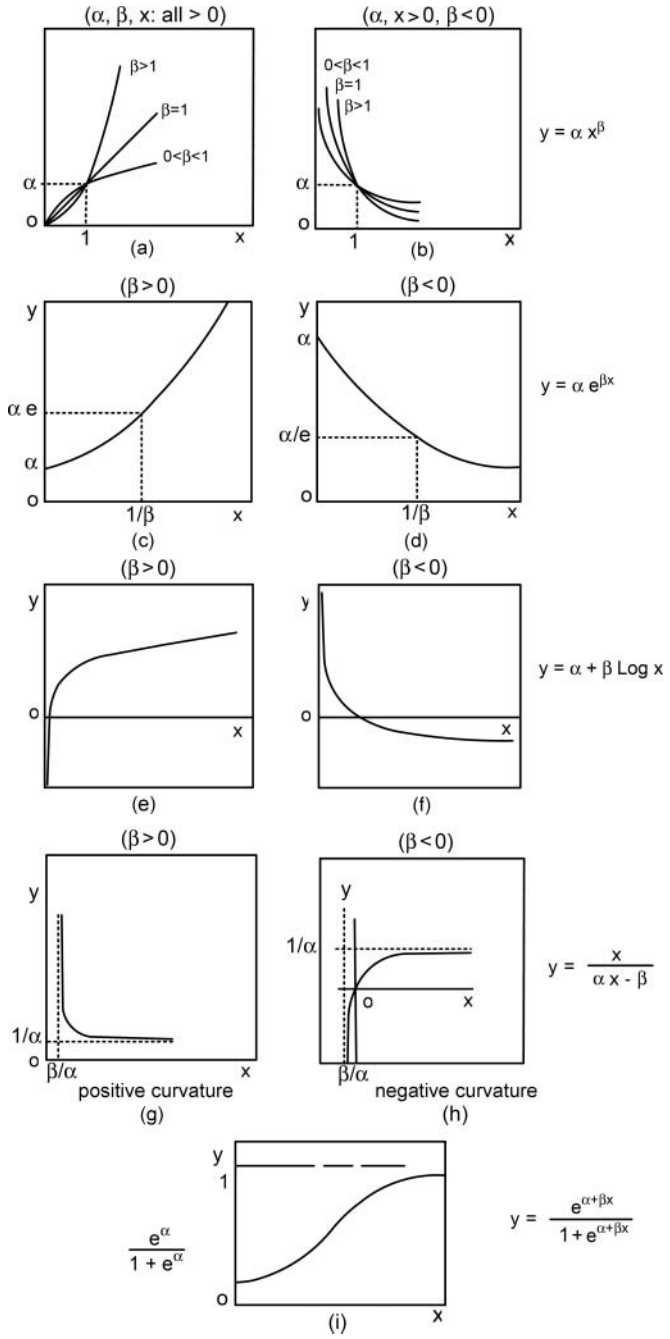

FIG. 6. Plots of linearizable curves.

to believe that the error were additive in the original models, all would have become intrinsically non-linear (Draper and Smith, 1998). Figure 6 shows examples of linearizable functions. Quite a few commonly occurring shapes are not linearizable; three are shown in Fig. 7. The *S*-shaped curve of Fig. 7 (bottom left) may be fitted by the so-called logistic equations (Daniel and Wood, 1980). The concentration of an intermediate compound, where x represents time (or duration) of a chemical reaction, often follows Fig. 7 (bottom right).

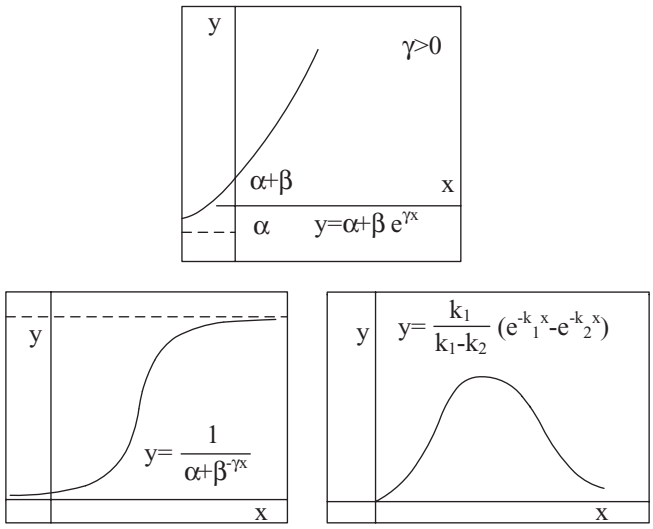

FIG. 7. Plots of non-linearizable curves.

It is important to understand, however, that a transformation of the data involves a transformation of the disturbance term, too, which affects the assumptions on it (Box and Draper, 1987; Bates and Watts, 2007; Seber, 1977). Thus, if we assume the model function with an additive and normal disturbance term is an appropriate representation of the experimental situation, then the same assumptions will not be appropriate for the transformed data. Hence, we should use non-linear regression on the original data, or else weighted least squares on the transformed data. The transformed model could then be fitted to give some initial estimates (Meloun et al., 1992; Mager, 1991).

So we must give some thought to what is being assumed about the error structure when we transform a response. In practice, it is sometimes simpler to decide on a transformation, fit it, and then examine the residuals in the metric of the transformed variable to see if they are reasonably well-behaved (Draper and Smith, 1998; Seber, 1977). If they are, the error specifications in the transformed response space are assumed to be all right.

We may return to the original y-space after the model has been fitted and evaluate the former function and make predictions, but the confidence statement will not be symmetric about the predicted value $\hat{y}$, and the residuals are not, however, checked; these are not the residuals that should satisfy the residual checks for normality and so on (Draper and Smith, 1998). The least squares assumptions on the behavior of the errors apply to the errors after transformations. Decisions as to how the errors should be incorporated into the models will depend on one's best judgment as to how the system operates and the analysis of the behavior of the residuals before and after transformation (Draper and Smith, 1998; Rowlings et al., 1998; Weisberg, 2005; Seber, 1977).

Note that relationships that show more than one bend, such as the classical *S*-shaped growth curve, cannot be straightened with the power family of transformations. A few commonly

used two-bend transformations are logit, arcsin (or angular), and probit. The effect of the transformation in all three cases is to "stretch" the upper and lower tails, the values of p(proportion) near one and zero making the relationship more linear (Natrella, 1963). The logit is sometimes preferred as a mean of simplifying a model that involves the product of probabilities. It has the desirable property that it is increases from $-\infty$ to ∞ as y increases from 0 to 1.

The integrated normal or probit transformation is related to the percentage area under the normal distribution curve from negative infinity (probability units) and has application in exploratory graphics analysis and specialized regression modelling of binary response variables (Collect, 2004; Petersen et al., 2000; Petersen, 1999). The term probit was coined in the 1930s by Bliss (1934) and was defined as $5 +$ the $1-p$ quantile from the standard normal distribution where p is a proportion (Armitage et al., 2002). The number 5 was added to ensure that the resulting scores were positive; the value can be taken from statistical compilations (Finney, 1971). Nevertheless, in the major statistical software packages probit are defined without the addition of 5. As probit methodology, including numerical optimization for fitting of probit functions, was introduced before widespread availability of electronic computing, it was convenient to have probits uniformly positive when using tables. Common areas of application do not require positive probits.

The tail-stretching effect of probit is lower than that of the angular transformation. The values for $p =$ (0% effect) and $p = 1$ (100% effect) are not defined theoretically; as a good approximation, we use

$$p = \frac{1}{2v} \qquad \text{(for } p = 1\text{)} \qquad [15]$$

$$p = \frac{2\,v - 1}{2\,v} \qquad \text{(for } p = 0\text{)} \qquad [16]$$

or the so-called working probits obtained by visual fitting (Mager, 1991); $p = u/v$. The procedure introduces a significant heterocedasticity in the upper and lower asymptotic curve regions. The probit transformation arises as the logical transformation when, for example, the chance of survival of an organism to a toxic substance is related to the dose, or ln(dose), of the toxin through a normal probability distribution of sensitivities (Rowlings et al., 1998). Such a so-called probit model is still important in toxicology, as well as other fields.

The logit transformation has a similar interpretation but here the threshold distribution is the logistic distribution. Logit models are usually more popular than probit models because on the one hand the exponentiated logistic coefficients may be interpreted as odd ratios, and on the other hand there are more diagnostic tools available in logistic regression. This later argument, however, is circular; i.e., a chicken-egg issue, that is, there might be more diagnostic tools because it is being used more often (Ender).

Many of the most common transformations (such as the arcsin, the square root, the logarithmic, and the logistic transformations) were developed for situations in which the random variables were expected a priori to have specific nonnormal distributions (Draper and Smith, 1998).

If the hazard rate is double-exponentially distributed, the extremal-value of complementary ln / ln-transformation may be used (Mager, 1991)

$$y = \ln\left(-\ln\left(1 - p\right)\right) \qquad [17]$$

It is applicable also to bimodal distributions. In general, it gives the best estimates in lndose-response assays.

Most bioassays and immunoassays have sigmoidal dose-response relationships. The logistic function may be used to describe such relationships if either the response range is naturally limited to 0–100% or a sufficiently wide range of doses is used so that asymptotic responses are obtained (Das and Tydeman, 1980). This function was proposed as early as 1940 to describe a sigmoid dose-response relationship for bioassays (Emmens, 1940) and has more recently been used to describe the dose-response relationship for immunoassays (Rodbard, 1975; Rodbard, 1974; Dudley et al., 1985; Gottschalk and Dunn, 2005a; Gottschalk and Dunn, 2005b; Findlay and Dillard, 2007).

However, because linear relationships are graphically and computationally less complex, analysis of data after a linearizing transformation has often been preferred. Two basic transformations (Mateu, 1997) can be used here: logits

$$\text{Logit}(y) = \log \frac{y}{1 - y} \qquad [18]$$

and foldest (square roots)

$$\text{foldest root}(y) = \sqrt{y} - \sqrt{1 - y} \qquad [19]$$

Logits are the more important transformation for proportions and percentages. Both of these transformation methods have basically similar properties. They stretch the scale near the ends (0 and 1) in a symmetric way around the middle value of $1/2$. The logit scale stretches farther and farther with no limits as a data value approaches 0 or 1, whereas the foldest roots are weaker and do no stretch indefinitely. Depending on the characteristics of each variable, different transformations are proposed.

Data sets that come to us a group of percentages or proportions are special because they are often bounded on both sides. Percentages are often forced to be between 0 and 100%, and proportions must be between 0 and 1. Because of this, the methods for transforming them will be somewhat different from the methods above (Mateu, 1997). Nonetheless, the immediate objective will be to make the transformed data distribution as symmetric as possible. When there are restrictive bounds, data values often trend to concentrate near them. In order to promote symmetry in the data, a transformation should stretch the distribution near these areas of concentration.

The literature on radio immunoassay (RIA) statistics and data processing is large (Gottschalk and Dunn, 2005a; Gottschalk and Dunn, 2005b; Findlay and Dillard, 2007; Mateu, 1997; Healy, 1972). Immunoassays, unlike the traditional bioassays,

is dependent on specific chemical reactions that obey the law of mass action and that are not subject to errors introduced by the biological variability of test systems (Buncher and Tsay, 1994). Nevertheless, they are so similar in structure that they need consideration from the viewpoint of bioassay. RIA gives a reliable measure of drug concentration in the low nanogram/milliliter range or below and has the ability to process a large number of samples in a relatively short time. The technique has been applied in virtually all areas of pre-clinical and clinical pharmacology, including pharmacokinetics and pharmacodinamics.

The four-parameter logistic function (Gottschalk and Dunn, 2005b; Findlay and Dillard, 2007; Healy, 1972; Buncher and Tsay, 1994) is a very flexible model for data following a sigmoidal shaped curve

$$E(y) = \beta_2 + \frac{\beta_1 - \beta_2}{1 + \left(\frac{c}{\beta_3}\right)^{\beta_4}} \qquad [20]$$

where y is the bound count (B) and c is the dose. The parameter β_1 is the asymptote as the concentration $c \to 0$ for $\beta_4 > 0$, β_2 is the asymptote as $c \to \infty$, β_3 is the predicted concentration at the response halfway between the two asymptotes (also denoted ED_{50}), and β_4 is related to the slope at the center of the curve (O'Connell et al., 1993). If the values of and β_1 and β_2 are assumed to be known, then Eq. [15] may be rearranged in a linearized form (O'Connell et al., 1993; Hatch et al., 1976; Rodbard and McClean, 1977; Cernosek, Jr., and Gutierrez-Cernosek, 1978; McDonald, 1981; Fischer and Rodbard, 1983; Plikaytis et al., 1991; Miller, 1991; Ritz and Streibis, 2005; Connors, 1984)

$$\text{logit}\left(\frac{y - \beta_2}{\beta_1 - \beta_2}\right) = -\beta_4 \ \log c + \beta_3^* \qquad [21]$$

where β_3^* is a function of β_3; β_3^* and $-\beta_4$ are the intercept and slope of a linear regression between the logit of the normalized or transformed response variable y and the logarithm (using either natural or common logarithms) of dose. This is commonly referred to as the "logit-log" method. Thus, if the values of β_1 and β_2, and accordingly the difference $(\beta_1 - \beta_2)$ are known to be very stable from assay to assay, then one may utilize a linearized form of the dose/response curve (Rodbart et al., 1987; Davidian, 2002). Note that the transformation to linearity may induce heterocedasticity in the new response (Ruppert et al., 1989), so that unweighted regression analysis cannot be applied. Thus, analysis in the logit scale may be inefficient. The importance of the choice of the upper and lower limits and the effect of these limits on the linearity of the transformed responses has been described (Das and Tydeman, 1980; Hatch et al., 1976).

The model is applied to RIA, immunoradiometric assay (IRMA or labelled anti-body assay), two-side IRMA, enzyme linked immunosorbent assay (ELISA), and enzyme multiplied immunological techniques (EMIT) (Rodbard and Frazier, 1975; Gottschalk and Dunn, 2005a; Buncher and Tsay, 1994).

Another model for sigmoidal dose response curve is the Weibul model (Plikaytis et al., 1991) given by the formula

$$f\,(x, (b, c, d, e)) = c + (d - c)\exp\{-\exp\{b\,(\log(x) - e)\}\} \qquad [22]$$

The parameters c, and d, are the lower and upper limits; as for the four-parameter logistic model, b is the relative slope around e and the e parameter is the logarithm of the inflexion point. The Weibull model is not symmetric around any point. The three-parameter Weibull model with the lower limit equal to 0 is then

$$f\,(x, (b, d, e)) = d\exp\{-\exp\{b\log(x) - e\}\} \qquad [23]$$

The frequently used logistic growth model is (Rowlings et al., 1998)

$$y = \frac{\alpha}{1 + \gamma e^{-\beta x_i}} \qquad [24]$$

This function gives the characteristic growth curve starting at $y = \alpha/(1 + \gamma)$ at $x = 0$ and asymptoting to $y = \alpha$ as x gets large. The function is intrinsically linear only if the value of α is known, as is the case, for example, when the dependent variable is the proportion of individuals showing reaction to a treatment. Ifα is known, the model is linearized by defining

$$y^* = \ln\left(\frac{\alpha}{y} - 1\right) \qquad [25]$$

and the model becomes

$$y_i^* = \gamma^* - \beta\ x_i \qquad [26]$$

where $\gamma^* = \ln\gamma$.

Proper non-linear fitting data is often difficult, and the error due to the bias by rectification is negligible with the random error of the experiment, at least in most cases. The proper non-linear and quasi-linear models are different, but in many cases, the consequences are less dangerous than the problems from employing non-linear regression models (Mager, 1991).

TRANSFORMATIONS TO STABILIZE VARIANCES

The first of variance heterogeneity is usually associated with the data whose distribution is non-normal (Rios, 1977; Canavos, 1984) viz, negative binomial, Poisson, binomial, etc. (Table 7). Data transformation is the most appropriate remedial measure in such situations. The second kind of variance heterogeneity usually occurs in experiments where, due to the nature of the treatment tested, some treatments have errors that are substantially higher (or lower) than others. The variance and the mean are independent in the normal probability distribution. All other common distributions (Table 7) have a direct link between the mean and the variance. In practice, the nature of the dependence between σ_y and $\bar{y}$ may be suggested by theoretical considerations or by preliminary empirical analysis, or by both (Box and Draper, 1987). If the functional relationship between the variance and the mean is known, a transformation exists that will

TABLE 7
Probability Distribution Functions (Natrella, 1963; Havilcek and Crain, 1988; Collect, 2004; Petersen et al., 2000)

Law	Function	Parameters	Mean	Variance	Skewness	Kurtosis
Uniform $U(a,b)$	$f(x)=\dfrac{1}{b-a}$	$-\infty<a<b<\infty$	$\dfrac{a+b}{2}$	$\dfrac{(b-a)^2}{12}$	0	1.8
Normal	$\dfrac{1}{\sigma\sqrt{2\pi}}\ e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$	$-\infty<\mu<\infty$ $\sigma>0$	μ	σ^2	0	3
Exponential	$\lambda\, e^{-\lambda x}$ $x>0$	$\lambda>0$	$\dfrac{1}{\lambda}$	$\dfrac{1}{\lambda^2}$	2	9
Binomial	$\dfrac{n!}{(n-r)!\,r!}\,p^r q^{n-r}$	$n=1,2,\ldots;\ 0<p<1$ $q=1-p$	$n\,p$	$n\,p\,q$	$\dfrac{1-2p}{\sqrt{n\,p\,q}}$	$3+\dfrac{1-6p\,q}{n\,p\,q}$
Poisson	$\dfrac{\lambda^r}{r!}\ e^{-\lambda}$	$\lambda>0$ $r=0,1,2,\ldots$	λ	λ	$\dfrac{1}{\sqrt{\lambda}}$	$3+\dfrac{1}{\lambda}$
Distribution of s^2	$\dfrac{1}{\Gamma\left[\dfrac{n-1}{2}\right]2^{\frac{n-1}{2}}}\ y^{\frac{n-1}{2}-1}\ e^{-\frac{y}{2}}$	$y>0$	σ^2	$\dfrac{2\,\sigma^4}{n-1}$		
Beta	$\dfrac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\,\Gamma(\beta)}x^{\alpha-1}(1-x)^{\beta-1}$	$0<x<1$ $\alpha,\ \beta>0$	$\dfrac{\alpha}{\alpha+\beta}$	$\dfrac{\alpha\ \beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$	$\dfrac{2(\beta-\alpha)\sqrt{\alpha+\beta+1}}{\sqrt{\alpha+\beta}\,(\alpha+\beta+2)}$	*
Gamma	$\dfrac{1}{\Gamma(\alpha)\,\theta^{\alpha}}\ x^{\alpha-1}\ e^{-\frac{x}{\theta}}$	$x>0$ $\alpha,\theta>0$	$\alpha\theta$	$\alpha\,\theta^2$	$\dfrac{2}{\sqrt{\alpha}}$	$3\left(1+\dfrac{2}{\alpha}\right)$
Chi-square	$\dfrac{1}{\Gamma\left(\dfrac{\upsilon}{2}\right)2^{\frac{\upsilon}{2}}}\ x^{\frac{\upsilon}{2}-1}\ e^{-\frac{x}{2}}$	$x>0$	υ	$2\,\upsilon$	$\dfrac{4}{\sqrt{2\,\upsilon}}$	$3\left(1+\dfrac{4}{\upsilon}\right)$
Bernoulli	$P(\xi=1)=p$ $P(\xi=0)=q=1-p$		p	$p\,q$	$p\,q\,(q-p)$	
Geometric	$(1-p)\,p^{k-1}=q\,p^{k-1}$	$k=1,2,\ldots$ 0 $0<p<1$	$\dfrac{1}{1-p}=\dfrac{1}{q}$	$\dfrac{p}{q^2}$	$\dfrac{1+p}{\sqrt{p}}$	$\dfrac{(1-p)^2+9p}{p}$
Hypergeometric	$\dfrac{\binom{Np}{r}\binom{Nq}{n-r}}{\binom{N}{r}}$	$N=1,2,\ldots$ $n=1,2,\ldots,N$ $p=0,\frac{1}{N},\ldots,1$	np	$npq\left(\dfrac{N-n}{N-1}\right)$	$\dfrac{npq\,(q-p)\,(N-n)\,(N-2n)}{(N-1)\,(N-2)}$	
Negative Binomial $\binom{n+r-1}{r}p^n q^r=\binom{-n}{r}p^n(-q)^r$		$r=0,1,\ldots,n$ $n>0$ $0\le p\le 1$	$\dfrac{nq}{p}$	$\dfrac{nq}{p^2}$	$\dfrac{n\,q}{p^2}\ \left(1+2\dfrac{q}{p}\right)$	
Weibull	$\dfrac{n}{\theta}\left(\dfrac{x}{\theta}\right)^{\eta-1}e^{-\left(\frac{x}{\theta}\right)^{\eta}}$ 0	$x\ge 0,\ \ \eta>0,\ \ \theta>0$elsewhere	$\theta\,a$	$\theta^2\ \left(b-a^2\right)$	$\dfrac{c-3ab+2a}{\left(b-a^2\right)^{3/2}}$	$\dfrac{d-4ac+6ba^2-3a^4}{\left(b-a^2\right)^2}$

*$\dfrac{3(\alpha+\beta+1)\left[2(\alpha+\beta)^2+\alpha\,\beta(\alpha+\beta-6)\right]}{\alpha\,\beta(\alpha+\beta+2)(\alpha+\beta+3)}$

**$a=\Gamma\left(1+\dfrac{1}{\eta}\right),\ b=\Gamma\left(1+\frac{2}{\eta}\right),\ c=\Gamma\left(1+\frac{3}{\eta}\right),\ d=\Gamma\left(1+\frac{4}{\eta}\right)$

***$\Gamma(a)=\int_0^{\infty}u^{a-1}e^{u}du$ or, for integers, $\Gamma(a)=(a-1)!$

make the variance (approximately) constant. This kind of transformation is usually applied when it is known that the variable follows a probability distribution whose variance is a function of the mean value of the distribution (Draper and Smith, 1998). Non-uniform variance (heterogeneous variance) as well as non-normality, are expected a priori with certain kinds of data. The same situations that give non-normal distributions will usually give heterogeneous variances since the variance in most non-normal distributions is related to the mean of the distribution (Natrella, 1963; Brownlee, 1984). This knowledge is in general alien to the statistical analysis, i.e., the variance is equal to the mean in the Poisson distribution. The plot of the Poisson variance against the mean would be a straight line of one slope. A priori, one should expect variances to be heterogeneous when the random variable is not normally distributed. Nevertheless, it is necessary to be cautious and select a simple transformation.

Let z represent y expressed in a transformed scale defined by

$$z = f(y) \tag{27}$$

Then, according to the random error propagation law (Asuero et al., 2006; Asuero et al., 1989; Gonzalez et al., 2005; Mandel, 1964; Tellinghuisen, 2000), the standard deviation of z is given approximately by

$$s_z = \left(\frac{df(y)}{dy}\right) s_y \tag{28}$$

If we choose the function $f(y)$ in such a way that s_z is a constant (homocedasticity in the $z$scale), then (Mandel, 1964; Bartlett, 1947; Hoaglin et al., 1983)

$$\left(\frac{df(y)}{dy}\right) s_y = k \tag{29}$$

where k is independent of z. This is because a constant variance is necessarily independent of the mean (and anything else for that matter), and this independence property is fundamental to the normal density (Bates and Watts, 2007), as we have indicated above. More generally, if the relation between measurement and standard deviation is

$$s_y = \phi_y \tag{30}$$

the proper transformation of scale to achieve homocedasticity is given by the differential Eq. [5] which becomes

$$df(y) = K\frac{dy}{\phi_y} \tag{31}$$

which upon integration gives:

$$z = f(y) = k\int \frac{dy}{\phi_y} + C \tag{32}$$

where C is an arbitrary constant. Some well-known transformations that arise in this way are shown in Table 8. The transformations selected have an advantage: in stabilizing variance, the transformed variable is more Gaussian. Note that some of these are members of the power family.

For example, for "a binomial proportion H" with mean valueθ, the variance (Brownlee, 1984) is given by

$$s_H^2 = \frac{\theta(1-\theta)}{n} \tag{33}$$

where n is the number of observations.

Inserting this in Eq. [31]

$$\phi(\theta) = \int \frac{d\theta}{\sqrt{\theta(1-\theta)/n}} = \sqrt{n}\int \frac{d\theta}{\sqrt{\theta(1-\theta)}} \tag{34}$$

This integral can be evaluated by changing the variable to $t = \sqrt{\theta}$, so that $t^2 = \theta$, so $2t\,dt = d\theta$, and that

$$\int \frac{d\theta}{\sqrt{\theta(1-\theta)}} = \int \frac{2tdt}{\sqrt{t^2\left(1-t^2\right)}} = 2\int \frac{dt}{\sqrt{1-t^2}}$$
$$= 2\ \arcsin t = 2\ \arcsin\sqrt{\theta} \tag{35}$$

TABLE 8
Transformation to Correct for Homogeneity and Approximate Normality

Data type	$s_y = f(y)$	Variance stabilizing transformation
Poisson (Count)* $y \geq 0$	$\sqrt{y}$	$\sqrt{y}$
Small counts (y)**		$\sqrt{y+1}$ or $\sqrt{y} + \sqrt{y+1}$
Binomial $(0 < y < 1)$	$\sqrt{y(1-y)}$	$a \sin\sqrt{y}$
Negative binomial $0 \leq y \leq 1$	$\dfrac{\sqrt{1-y}}{y}$	$\left(\dfrac{2+y}{3}\right)\sqrt{1-y} = (1-y)^{1/2} - \dfrac{(1-y)^{3/2}}{3}$
Variance = (mean)^2 $y \geq 0$	y	$\ln y$
Correlation coefficient $-1 \leq y \leq 1$	$\dfrac{1}{\sqrt{1-y^2}}$	$0.5\,[\ln(1+y) - \ln(1-y)]$

* Modifications for the Poisson and binomial cases have been suggested by Natrella, 1963.

** It should be noted that the square root transformation overcorrects when very small values and zero appears in the original data. In these cases, $\sqrt{y+1}$ is often used as a transformation.

Thus, if we use the transformation $\phi(H) = 2\sqrt{n}\ \arcsin\sqrt{H}$, the variance of $\theta(H)$ will be 1. If we omit the factor $\sqrt{n}$ and instead use $\phi(H) = 2\arcsin\sqrt{H}$, then the variance of this transformed variable will be $1/n$.

The problem that arises when analyzing proportions (where the data consists of proportions of widely different magnitudes) is the lack of homogeneity of variance; the variance depends on the magnitude of the proportion. Though methods are available to analyze such data taking into account the disparate variance, the calculations are relatively complicated. If the proportions to be analyzed are approximately normally distributed ($n\ p$ and $n\ q \geq 5; q = 1 - p$), thearcsin transformation will equalize the variances (Acton, 1959; Bisgaard and Fuller, 1995).

Figure 8 presents a comparison of the actual values of the variance of the transformed values and the corresponding approximate values by giving Eq. [28], for four of the transformations listed in Table 7.

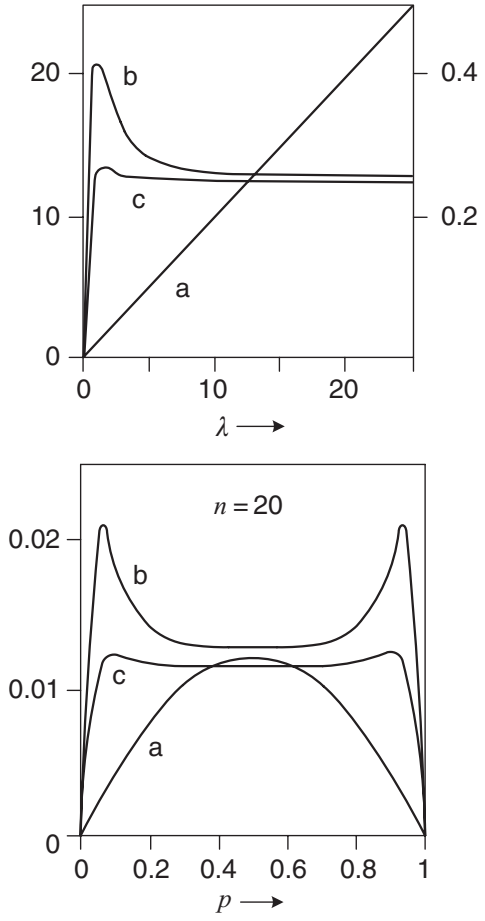

FIG. 8. Top: The variance functions for Poisson distributed counts when using (a) no transformations, (b) the square root of counts, and (c) Freeman and Tukey's modifications. The y-axes scales on the left applies to (a) and the scale on the right applies to (b) and (c). Bottom: The variance functions for sample size of $n = 20$ for binomial distributed proportions when using (a) no transformation, (b) the arsin square root of proportions and (c) Tukey's modification.

When using thearcsin transformation, each proportion should be based on the same number of observations, n. If the number of observations is similar for each proportion, the analysis using arcsines will be close to correct. However, if the number of observations is very different for the different observed proportions, the use of the transformation is not appropriate. Also, for very small or very large proportions (less than 0.3 or greater than 0.97), a more accurate transformation is given by Mosteller and Youtz (1961).

The logit and the probit, although they are generally applied to binomial data, will not stabilize the variance

If the variance is proportional to the mean, so that (Brownlee, 1984)

$$s_y^2 = k\ y \qquad [36]$$

then

$$\phi\,(y) = \int \frac{dy}{\sqrt{k\ y}} = \frac{2}{\sqrt{k}}\sqrt{y} \qquad [37]$$

Thus, a square root transformation will be a constant variance. In particular, if $k = 1$, as occurs with the Poisson distribution, $2\sqrt{y}$ will have variance 1, and $\sqrt{y}$ will have variance 1/4 (Bayne and Rubin, 1986). Poisson data arise if we are plotting radiation counts for a time interval, indeed, almost any number of random events occurring in a prescribed time or space. The $\sqrt{y}$ transformation was improved by Anscombe (Tellinghuisen, 2000; Bartlett, 1947; Hoaglin et al., 1983; Bisgaard and Fuller, 1995: Anscanbe, 1948; Kihlberg et al., 1972) to $\sqrt{y + 3/8}$, and a somewhat better one was developed by Freeman and Tukey (1950) $\sqrt{y} + \sqrt{y+1}$. This last transformation has a variance equal to one (plus or minus 6%) for an expected value of y greater than one, and the two earlier transformations (Acton, 1959) stabilize to a variance of approximately 0.25.

In some instrumental method of analytical and physical chemistry, the relative standard deviation of the measured variable is constant (Meloun et al., 1992). This means that the standard deviation is proportional to the mean; then

$$s_y^2 = k^2 y^2 \qquad [38]$$

$$\phi(y) = \int \frac{dy}{\sqrt{k^2 y^2}} = \frac{1}{k}\ \ln y \qquad [39]$$

So that $\ln y$ will have a constant variance k^2 (and $\log y$ a constant variance $\log^2 e/k^2$). Probably the most common transformation used in scientific research is the log transformation (Bolton, 2004). Data that are skewed to the right can be often shown to have close to a log normal distribution (a log normal distribution is a distribution that would be normal following a log transformation). The log transformation is applicable in many cases, especially when the range of observations covers several order of magnitude (Rodriguez and Asuero, 1980; Asuero, 1978). The measurement of acidity by pH is an example of

such a transformation in ordinary scientific work (Kennedy and Neville, 1986).

Data involving radioactive decay, first order chemical kinetics (Connors, 1990; de Levie, 2002), all follow an exponential decay of the type $y = a\,e^{-bt}$ where t denotes time. In the pharmaceutical sciences, the logarithmic transformation has particular applications in kinetic studies, when ascertaining stability and pharmacokinetic parameters (Bolton, 2004). The log transformation may be used when data are presented in the form of ratios. Ratios are often used to express the comparative absorption of drug from two formulations based on the area under the plasma level versus time curve from the bioavailability study. However, it is important not to apply a transformation such as the logarithmic form when one is not needed (Shumway et al., 2002).

If the standard deviation is proportional to the square of the mean value, so that

$$s_y^2 = k^2 y^4 \qquad [40]$$

then

$$\phi(y) = \int \frac{dy}{\sqrt{k^2 y^4}} = \frac{1}{k}\int \frac{dy}{y^2} = -\frac{1}{k\,y} \qquad [41]$$

To summarize these three results, to obtain constant variance, if $\sigma^2 \propto y$ we use $\sqrt{y}$, if $\sigma^2 \propto y^2$ we use $\ln x$, and if $\sigma^2 \propto y^4$ we use $1/y$. A further transformation, which has a strong cedastis effect (variance stabilizing effect) and a normalizing property, $Y = \tanh^{-1} x$, was originally suggested by Fisher and applied to the sample correlation coefficient to employ normal theory testing (Herrador et al., 1987; González et al., 2006; Sonnergaard, 2006).

Transforming to constant variance often has the additional effect of making the disturbances behave more normally. This is because a constant variance is necessarily independent of the mean (and anything else, for that matter), and this independent property is fundamental to the normal density (Bates and Watts, 2007).

An experiment which includes replications allows further tests to be made on the appropriateness of assumptions (Lavagnini and Magno, 2007; Asuero and Gonzalez, 1989; Sayago and Asuero, 2004; Bates and Watts, 2007). For example, even before an expectation function has been proposed $\mu = E(y)$, it is possible to check the assumption of constant variance by using an analysis of variance to get averages and variances for each set of replications and plotting the variances and standard deviations versus the averages. If the plots show systematic relationships, then one can use a variance-stabilizing procedure to transform to constant variance (Asuero and Gonzalez, 2007).

In particular (Draper and Smith, 1998; Bates and Watts, 2007; Atkinson, 2003; Montgomery and Peck, 1982), if it is assumed that σ_y is proportional to some power of μ so that

$$\sigma_y^2 \propto \mu^{2(1-\lambda)} \qquad [42]$$

if the power law holds with $\lambda < 1$, large observations will have large standard deviations rather than small observations. Taking logarithms of the square root of both sides of this relationship yields

$$\log\ \sigma_y = c + (1-\lambda)\ \log\ \mu \qquad [43]$$

If replicate observations are available, a plot of log-standard deviation against log-mean will indicate whether the power law holds (y^λ for $\lambda \neq 0$ and $\log y$ for $\lambda = 0$, as can also be deduced by Taylor series expansion). Kleczkowski (1949) established that a linear relation of the standard deviation of y to the mean of y, with intercept $-c$ at standard deviation $y = 0$, led to a shifted log transformation $\log(y + c)$.

Experience has shown that normality is a reasonable assumption in many cases. However, in some situations it is not appropriate to assume normality; count data will frequently behave more like Poisson distributed random variables (Schwartz, 1978). Transformation of the dependent variable to a form that is more nearly normally distributed is the usual recourse to non-normality. Statistical theory says that such transformation exists if the distribution of the original dependent variable is known. Many of the common transformations (such as the arcsin, the square root, the logarithmic, and the logistic transformation) were developed for situations in which the random variables were expected a priori to have specific non-normal distributions (Rawlings et al., 1998).

In many cases, the sample data provide the only information available for determining the appropriate normalizing transformation. The plots of the residuals may suggest transformations, or several transformations might be tried and the one adopted that most nearly satisfies the normality criteria. A picture is worth a thousand words. This old saying emphasizes the importance of graphs at all stages of a statistical analysis (Chambers et al., 1983; Buja and Tukey, 1991; Cook and Weisberg, 1994; Cook and Weisberg, 1994).

The logit and the probit, although they are generally applied to binomial data, will not stabilize the variance.

WEIGHTING TRANSFORMATION DATA

Noisy information is crucial to do any quantitative interpretation of data (de Brauwere et al., 2007). Many examples in analytical chemistry indicate that the variability of the response often increases with the response level (Sayago and Asuero, 2004; Asuero and Gonzalez, 2007; Meier and Zünd, 2000; Miller and Mille, 2005; Mullins, 2003; Baumann, 1997). In such a situation, some remedial actions like transformation or using weighted least squares regression must be done in order to stabilize the variance of the response and, thus, heterocedasticity can be accounted for (Asuero and Gonzalez, 2007; Zorn et al., 1997). Experimental errors often are heterocedastics (Asuero and Gonzalez, 2007; Steliopolous and Sticke, 2007); sometimes we know that an instrument is noisier in one region of its range than in another. In that case we can put more weight

on some of the data obtained in a different region (de Levie, 2001). In spite of this, up to now, the acceptance of weighted least squares regression is quite low in routine laboratories.

Assigning weights to data requires that we know how much (relative) weight to allow to each measurement (Sayago and Asuero, 2004) and how to handle such individual weights or weighting factors w_i in an analysis. If we have sufficient replicates of each observation, we might, e.g., assign each measurement its proper individual weight, equal to the reciprocal of its variance, $w_i = 1/s_i^2$ (Asuero and Gonzalez, 1989; Asuero and Gonzalez 1988; Miller, 1991). Unfortunately such information is seldom available. Replicate measurements are often not performed because they are too expensive (Sayago and Asuero, 2004; Asuero and Gonzalez, 2007; Mullins, 2003). Although the use of m-replicates is considered in many experimental studies, a careful analysis shows that the benefits of such a practice in term of improving the precision of estimates may not be worthwhile without an analysis of its benefits (Singer et al., 2007).

Iterative methods based on the residuals are proven to be more reliable, especially when the number of replicates is small (Seber, 1977). Various weight schemes have been used to remove the heterogeneity of response variability: different type of variance functions and methods for estimation of them have been introduced (Asuero and Gonzalez, 2007; Tellinghuisen, 2008; Shumway et al., 1989; Wilson et al., 2004; Rocke and Lorenzato, 1995; Rocke et al., 2003). An alternative method uses an algorithm which starts from residuals and so it necessitates a model to describe the data (de Brauwere et al., 2007). Replicate data are desirable, but these should be used to derive the data variance function, not to provide sample-variance based weights. The weights are then calculated from the variance function (de Levie, 2000; Asuero and Gonzalez, 2007; Tellinghuisen, 2008; Davidan and Haaland, 1990).

Least-squares parameters are normally correlated, and in the calculation of statistical errors in functions of the parameters, this correlation must be taken into account (Tellinghuisen, 2007; Tellinghuisen, 2000).

There are many situations in experimental physical science where data are transformed to facilitate analysis by linear regression (Asuero and Gonzalez, 2007; Tellinghuisen, 2000). Logarithmic conversion is often used to render first order kinetics data and thermodynamic temperature dependences into linear functions of the independent variables (Tellinghuisen, 2000). Reciprocation is employed in a variety of situations including analysis of equilibrium and binding constant data, enzyme kinetics, adsorption isotherms, and fluorescence quenching (Tellinghuisen, 2000). In the pharmaceutical sciences, the logarithmic transformation has particular applications in kinetic studies, when ascertaining stability and pharmacokinetics parameters (Bolton, 2004).

Frequently data are transformed before fitting, and if the original data are normal, the transformed data may not be. In particular, inversion and logarithm conversion yield biased, non-Gaussian distributions, so least squares analysis of such data yields biased, non-normally distributed parameters, even when the transformed data are properly weighted in accord with the transformation (Tellinghuisen, 2000). If we fit the transformed data using least squares, we minimize the sum of the residuals in Y rather than those in y. In some cases it may well be correct, namely when the errors in y are relative ones, proportional to the magnitude of the signal y. But when the experimental errors are absolute, the resulting fit will overemphasize the tail end of the data set (de Levie, 2001).

Rearranging a non-linear equation into a linear form and then performing linear regression to fit the data is very common, i.e., Scatchard plot of binding data and the Lineweaver-Burk plot of enzyme-kinetics. It is necessary to take into account that the rearrangements involved also rearrange the error distribution and so invalidate the assumptions behind the linear regression technique (Ruppert et al., 1989). It has been recommended to apply non-linear models to non-normal data rather than linear models to non-normal (transformed) data (Tellinghuisen, 2000).

In this particular example, the transformation runs into additional trouble when the signal decays to the baseline, because the experimental noise will then make a number of observations negative, in which case we will not take the corresponding logarithms (de Levie, 2000; de Levie, 2001).

Data involving radioactive decay, first order chemical kinetics (Connors, 1990), or the electrical current following a stepwise voltage change in a resistor-capacitor circuit all follow an exponential decay of the type $y = a\,e^{-bt}$ where t denotes time. It is usual to rectify such an expression by taking (natural) logarithms, so that $\ln y = \ln a - bt$, which is the expression for a straight line of $\ln y$ versus t. In general, upon transforming the dependent parameter y into Y when the original indications have normally distributed errors (Asuero and Gonzalez, 1989; de Levie, 2004; de Levie, 2001; de Levie, 1986; de Levie, 2004)

$$w_i = \frac{1}{\left(\frac{dY}{dy}\right)^2} \qquad [44]$$

If the system is both transformed and heterocedastic then the products of the weights are applied; i.e., the global weight will be (Asuero and Gonzalez, 2007; de Levie, 1986; Leatherbarrow, 1990; Hibbert, 2006) the product of the individual and the global weights (Table 9)

$$w_i = \frac{1}{s_i^2\left(\frac{dY}{dy}\right)^2} \qquad [45]$$

There may be additional problems with a transformation (Tellinghuisen, 2006). For example, when an exponential decays to zero, the experimental data will be scattered around zero, so that the same of them may be negative (we would bias the result if we merely left out the logarithms of the negative data).

Most of these methods, because of the nature of the relevant equations, do not fulfil one or more of the pre-requisites (for instance normal error distribution, uncorrelated variables) to

TABLE 9
Weighting Factors Associated with a Given Transformation

Transformation	Weighting factor (*)
$1/y$	$y^4\sigma_0^2/\sigma_y^2$
$\ln y$	$y^2\sigma_0^2/\sigma_y^2$
y^2	$\sigma_0^4/\left(4y^2\sigma_y^2\right)$
e^y	$\sigma_0^2/\left(e^{2y}\sigma_y^2\right)$
Logit y	$y^2(1-y)^2\sigma_0^2/\sigma_y^2$

* σ_0^2, proportionality factor; i.e., the variance of a function of unit weight (Asuero and Gonzalez, 2007).

be treated with ordinary least squares methods (Tellinghuisen, 2000; Tellinghuisen, 2006; Maccá, 1990). However, they can still be very useful for preliminary screening of data and also yield reliable results, if appropriately used (Maccá and Merkoci, 1994; Maccá, 1990).

TRANSFORMATION BASED ON SAMPLE DATA OBSERVATIONS: BOX-COX METHOD

In many cases, the sample data provides the only information available for determining the appropriate normalization transformation. The plot of residuals may suggest transformation; several transformations might be tried and the one adopted that most nearly satisfies the normality criteria. Alternatively, an empirical method of estimating the appropriate power transformation might be used (Box and Cox, 1964). A transformation family is a collection of transformations that are indexed by one or a few parameters that the analyst can select (Weisberg, 2005; Chinn, 1996; Sakia, 1992). A useful family of transformation is the following (Table 10)

$$T = \begin{cases} y^{\lambda}, & \text{for } \lambda \neq 0 \\ \ln y, & \text{for } \lambda = 0 \end{cases} \qquad [46]$$

This particular family, which was studied in detail by Tukey (1957) for $\lambda \leq 1$, contain the well-known log, square root, and inverse transformations. There is a particular difficulty when dealing with powers y^{λ}, because as λ approaches zero, y^{λ} approaches 1, thus lacking of sense the transformation (Draper and Smith, 1998). In order to choose the best λ value to run smoothly as λ approaches zero, Box and Cox (Box and Cox, 1964; Chinn, 1996; Sakia, 1992; Peace, 1988; Schlesselman, 1971) perform the calculations using no y^{λ}, which is discontinuous at $\lambda = 0$, but with either (Table 10)

$$W = \begin{cases} \left(y^{\lambda} - 1\right)/\lambda, & \text{for } \lambda \neq 0 \\ \ln y, & \text{for } \lambda = 0 \end{cases} \qquad [47]$$

which is essentially identical to Eq. [46] when the regression model contains a constant term b_0(Peace, 1988).

However, as early as 1903, Kapteyn (Kapteyn, 1903; Kapteyn and van Uwen, 1916) had suggested essentially the same transformation in his work on growth. A test of $\lambda = 0$($\log y$ versus x) is given by Sclove (1972). Eqs. [46] and [47] are equivalent since the Fstatistics in the analysis of variance are invariant under linear transformations (Malaeb, 1997).

It is better, however, to use (Table 10)

$$V = \begin{cases} \left(y^{\lambda} - 1\right)/\left(\lambda\, \dot{y}^{\lambda-1}\right), & \text{for } \lambda \neq 0 \\ \dot{y}\ \ln y, & \text{for } \lambda = 0 \end{cases} \qquad [48]$$

being $\dot{y}$ the geometric mean of the y_i in the data set

$$\dot{y} = (y_1\, y_2 \cdots y_k)^{1/k} = \left(\prod y_i\right)^{1/k} \qquad [49]$$

$\dot{y}$ is a constant and it would be evaluated at the beginning of the calculation procedure, usually by anti-loggin (exponentiating) the formula

$$\ln \dot{y} = \frac{\sum_{i=1}^{k} \ln \dot{y}_i}{k} \qquad [50]$$

A disadvantage of Eq. [47] is that, as λ varies, the sizes of the W's can change enormously, leading to minor problems in the analysis and requiring a special program to get the best λ value. For that reason, it is preferable to use the alternative form Eq. [48]. The additional divisor ($\dot{y}^{\lambda-1}$) in Eq. [48], compared with Eq. [47], is the *nth* power of the appropriate Jacobian of the transformation, which converts the set of y_i into the set of W_i (Mateu, 1997)

$$J(\lambda, y) = \prod_{i=1}^{k} \left| \frac{dy_i^{(\lambda)}}{dy_i} \right| = \prod_{i=1}^{k} y_i^{\lambda-1} \qquad [51]$$

This ensures that unit volume is preserved in moving from the set of y_i into the set of V_i in Eq. [47].

In the same paper, Box and Cox (1964) also proposed a more flexible family, an extender power or shifted power transformation, which could accommodate negative y's

$$y(\lambda) = \begin{cases} \dfrac{(y+\lambda_2)^{\lambda_1} - 1}{\lambda_1} & \text{if} \lambda_1 \neq 0 \\ \log(y+\lambda_2) & \text{if} \lambda_1 = 0 \end{cases} \qquad [52]$$

In practice, we would choose λ_2 such that $y + \lambda_2 > 0$ for any y. So, we could only view λ_1 as a model parameter. Note that if a shifted Box-Cox transformation is the variance stabilizing transformation, then the variance of the untransformed scale must be proportional to $(mean + \lambda_2)^{(1-\lambda_1)}$ (Chinn, 1996).

TABLE 10
Values of Certain Power Functions for Transformations to Stabilize Variances

λ	y^{λ}	$W = (y^{\lambda} - 1)/\lambda$	$V = (y^{\lambda} - 1)/(\lambda \dot{y}^{\lambda-1})$
1	y	$y - 1$	$y - 1$
0.5	$\sqrt{y}$	$2(\sqrt{y} - 1)$	$2\sqrt{\dot{y}}(\sqrt{y} - 1)$
0	1(?)	$\ln y$	$\dot{y}\ \ln y$
−0.5	$1/\sqrt{y}$	$2(1 - 1/\sqrt{y})$	$2\dot{y}\sqrt{\dot{y}}(1 - 1/\sqrt{y})$
−1	y^{-1}	$2(1 - 1/\sqrt{y})$	$\dot{y}^2(1 - 1/\sqrt{y})$

It is important to note that the range of $y^{(\lambda)}$ in Eqs. [47], [48], and [51] is restricted according to whether λ is positive or negative. This implies that the transformed value does not cover the entire range $(-\infty,\ +\infty)$ and, hence, their distributions are of bounded support. Consequently, only approximate normality is to be expected (Sakia, 1992).

The Box-Cox transformation has following properties (Meloun et al., 1992):

1) The curves of transformation are monotonic and continuous with respect to parameter λ because

$$y^{\lambda} = e^{\lambda \ln y} \approx 1 + \lambda \ \ln y + \frac{1}{2}\lambda^2 (\ln y)^2 + \cdots \qquad [53]$$

and according to the l'Hospital's rule (actually proved by Johann Bernouilli)

$$\lim_{\lambda \to 0} \left(\frac{y^{\lambda} - 1}{\lambda} \right) = \ln y \qquad [54]$$

Logarithms to any base (e.g., 10) can be used in this transformation. The difference amounts only to a constant factor. Also, scaled power transformations preserve the direction of association, in the sense that if (y, z) are positively related, then $(V,\ z)$ are positively related for all values of λ. With basic power transformations, the direction of association changes when $\lambda < 0$.

2) All transformation curves shape one point $[V = 0, y = 1]$ for all values of λ. The curves nearly coincide at points close to [0, 1]; that is, they share a common tangent line at this point (Fig. 9).
3) The power transformations with exponent -2, -3/2, -1, -1/2 o 1/2, 3/2, 2 have equal spacing between curves in the family of Box-Cox transformation graphs.

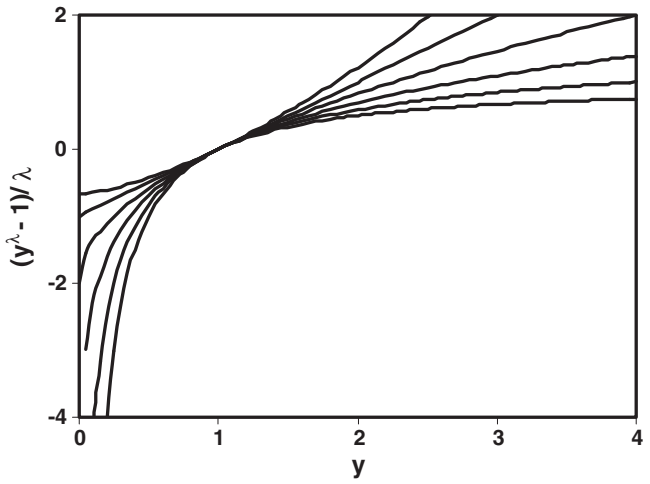

FIG. 9. Box-Cox transformations for various valores of λ. Graphs from stop to down corresponds to λ values of 1.5, 1, 0.5, 0, -0.5, and -1.

The idea underlying the Box and Cox method (Draper and Smith, 1998) is that, if an appropriate λ could be found, an additive model with normally distributed, independent, and homogeneous error structure could be fitted by the maximum likelihood method. However, in order to perform the procedure we do not have to understand maximum likelihood, Bayesian statistics, or the Jacobian. The necessary steps are as follows (Draper and Smith, 1998).

Choose a value of λ from a selected range. Usually we look at λ's in the range (-1, 1) or perhaps even (-2, 2), at first, and expand the range later if necessary. The selected range is usually covered with about 11–21 values of λ. The interval may divided up in additional portions if required later, which proves often to be unnecessary.

Now fit and record the residual sum of squares for the regression. Any ordinary least squares regression program can be used for this calculation.

Plot either the residual sum of squares for the regression $S(\lambda, U)$ versus λ or its logarithm $\ln S(\lambda, U)$ versus λ, depending on how big the numbers are (Draper and Smith, 1998; Rawlings et al., 1998; Mager, 1991). Draw a smooth curve through the plotted points, and find at what value of λ the lowest point of the curve lies $(\hat{\lambda})$. Maximizing the likelihood is equivalent to minimizing the residual sum of squares (Draper and Smith, 1998). We do not need to know λ very precisely in subsequent calculations the nearest convenient value in the sequence, $-2, -1\frac{1}{2}, -1, \frac{1}{2}, 0, \frac{1}{2}, 1\frac{1}{2}, 2, \ldots$, is usually adequate, after first checking that such a value lies within a selected confidence interval. We then analyze the transformed data—transformed via whatever value of λ was finally selected—and report the results.

A simple Box-Cox transformation reduces heterogeneity sufficiently. Carroll and Ruppert (1988) suggested that an excess of maximum over minimum standard deviation of 50% was unimportant. Precise estimation requires robust methods because the variance function will be significantly affected by outliers, which can be detected by applying graphical methods as discussed by Atkinson (1985). The Box and Cox method will also produce a confidence interval for the transformation parameter. If α is a level of significance, a 100(1- α)% confidence interval can be found for λ by calculating a critical sum of squares SS_c from

$$SS_c = S_{\min}(\lambda, z) \left(1 + \frac{t_{\upsilon}^2 (\alpha/2)}{\upsilon} \right) \qquad [55]$$

where $S_{\min}(\lambda, z)$ is the minimal residual sum of squares with respect to λ, with the associated υdegrees of freedom and t_{υ} is the corresponding value from the t table (Rawlings et al., 1998) with $\alpha/2$ probability in each tail.

This Box-Cox transformation is less suitable if the confidence interval for λ is too wide, and if the sample size is small then the confidence interval for the parameter will be wide. When the value $\lambda = 1$ is also covered by this confidence interval, the transformation is not efficient.

TABLE 11
Transformations

Transformations	Comments	Ref.
$y = \begin{cases} (\exp(\lambda y) - 1)\lambda & \lambda \neq 0 \\ y & \lambda = 0 \end{cases}$	Succesful in transform unimodal skewed distribution into normal distribution, but is not quite useful for bimodal or U-shaped distribution.	Manly, 1976
$y = \begin{cases} sign(y)\left\{(\lvert y\rvert + 1)^{\lambda} - 1\right\}/\lambda & \lambda \neq 0 \\ sign(y)\{\log(\lvert y\rvert + 1)\} & \lambda = 0 \end{cases}$ $sign(y) = \begin{cases} 1 & \text{if} \quad y \geq 0 \\ -1 & \text{if} \quad y < 0 \end{cases}$	It works well at those distributions that are somewhat symmetric. A power transformation on aymmetric distribution is likely going to introduce some degree of skewness.	John and Draper, 1980
$y = \left\{\lvert y\rvert^{\lambda} sign(y) - 1\right\}/\lambda \qquad \lambda > 0$		Bickel and Doksum, 1981
$y = \begin{cases} \frac{(y+1)^{\lambda}-1}{\lambda} & \text{if } \lambda \neq 0 \quad y \geq 0 \\ \log(y+1) & \text{if } \lambda = 0 \quad y \geq 0 \\ \frac{(1-y)^{2-\lambda}-1}{\lambda-2} & \text{if } \lambda \neq 2 \quad y < 0 \\ -\log(1-y) & \text{if } \lambda = 2 \quad y < 0 \end{cases}$	When estimating the transformation parameter, they found the value of λ that minimizes the Kullback-Leibler distance between the normal and the transformed distribution.	Yeo and Jonson, 2000

The Box-Cox transformation linearizes a regression relationship directly via a power transformation to the dependent variable y as shown (Xu and Hee, 2006)

$$z = \frac{y^{\lambda} - 1}{\lambda} = b_0 + b_1 x \quad [56]$$

Rearranging Eq. [56]

$$z' = y^{\lambda} = b'_0 + b'_1 x \quad [57]$$

where

$$b'_0 = \lambda b_0 + 1 \quad [58]$$

$$b' = \lambda b \quad [59]$$

and

$$z' = z\lambda + 1 \quad [60]$$

Once the optimal λ was found, Eq. [60] was used to plot $z'(= y^{\lambda})$ versus x.

The Box-Cox method is not transforming for linearity, but rather it is transforming for normality; λ is chosen to make the residuals from the regression of $W(y, \lambda)$ on y as close to normally distributed as possible. Hinkley (1975) stated that the distribution of the transformed variable is normal only if λ is zero or if λ^{-1} is an integer. Hernandez and Johnson (1980) point out that "as close as normal possible" need not be very close to normal, and so graphical checks are desirable after selecting a transformation; the distribution of the transformed variable is actually normal only in the log normal case. Draper and Cox (1969) showed that "this family of transformations can be useful even in situations where no power transformation can produce normality exactly."

The Box-Cox transformation is only appropriate for positive data (Cressie, 1978). Various modifications to the Box-Cox transformation have been proposed (Hinkley, 1975; Manly, 1976; John and Draper, 1980; Yeo and Johnson, 2000; Bickel and Doksum, 1981), some of which are designed to achieve heterogeneity simultaneously with normality. Hinkley (1975), Manly (1976), John and Draper (1980), and Yeo and Johnson (2000) proposed alternative families of transformations (Table 11) that can be used to compensate the restrictions on y, to obtain an approximate symmetry or to make the distribution normal (Chen et al., 2004). However, these modifications are at expense of simplicity, and may be unnecessary in practical applications. A simple variance-stabilizing transformation should be sought first (Rocke et al., 2003).

The parameter of a transformation, e.g., λ, can be selected through a trial and error approach until good normal probability plots are obtained, through optimization based on maximum likelihood estimation or Bayesian estimation (Draper and Smith, 1998; Rawlings et al., 1998), likelihood ratio test, or the use of M-estimators, etc. Draper and Hunter (1969) discussed other criteria for transformation selection and, in particular, proposed the minimization of a statistic F_1, as the appropriate criterion for the choice of a variance-stabilizing transformation. The F_1 statistics are the Fratio of the between subject (or between groups) to within subject (or within group) mean squares of the absolute residuals.

It is often more useful to apply transformations of predictor (model input) variables, along with the transformations of the dependent (model output) variables. Box and Tidwell (1962) provides an iterative procedure to estimate appropriate transformations of the original model inputs. Atkinson and Riari (1997) discussed different models and reasons for what transformations of predictor variables can be applied.

It should also be noted that applying the existing transformation techniques may have little effect if the values of the response are far from zero and the scatter in the observation is relatively small; in other words, the ratio of the largest to smallest observation should not be close to one (Atkinson, 1973).

Computer programs have been devised to apply the Box-Cox transformation technique (Huang et al., 1978; Chang, 1977). Some statistical packages include graphical tools that can help on select power transformations of both the predictor and the response. A program has been provided by Rode and Chinchilli (MULTBXX) which conducts a multivariate Box-Cox transformation (Rode and Chinchilli, 1988), of which the uni-variate version is a special case. Some statistical packages include graphical tools that can help you to select power transformations of both the predictor and the response (Weisberg, 2005; Malaeb, 1997).

RE-EXPRESSION OF THE STATISTICAL MEASUREMENTS AFTER DATA TRANSFORMATIONS

After an appropriate transformation of the original data y has been found, such that the transformed data give an approximately normal symmetrical distribution with constant variance, the statistical measurements of location and spread for the transformed data z are calculated. These include the sample mean $\bar{z}$, the sample variance s_z^2, and the confidence interval of the mean

$$\bar{z} \pm t_{1-\alpha/2}\,(n-1)\,s_z/\sqrt{k} \quad [61]$$

These estimates must then be re-calculated for the original data y. Two different approaches to the re-expression of the statistics for transformed data can be used without difficulty (Meloun et al., 1992; Meloun et al., 2000).

Rough re-expression is represented by a single reverse transformation $\bar{y}_R = g^{-1}(z)$. This re-expression for a simple powder transformation leads to the general re-expressed mean

$$\bar{y}_R = \bar{y}_{\lambda} = \left(\frac{\sum_{i=1}^{k} y_i^{\lambda}}{k} \right)^{1/\lambda} \quad [62]$$

where for $\lambda = 0$, $\ln y$ is used instead of $y^{1/\lambda}$.

More correct re-expressions are based on the Taylor series expansions of the function $z = g(y)$ in a neighborhood of the value $\bar{z}$. The re-expressed mean $\bar{y}_R$ is then given by

(Meloun et al., 1992; Meloun et al., 2000)

$$\bar{y}_R \approx g^{-1}\left\{\bar{z} - \frac{1}{2}\frac{d^2g(y)}{dy^2}\left(\frac{dg(y)}{dy}\right)^{-2} s_z^2\right\} \quad [63]$$

The variance is then expressed as

$$s_{\bar{y}_R}^2 = \left(\frac{dg(y)}{dy}\right)^{-2} s_z^2 \quad [64]$$

where individual derivatives are calculated at the point $y = \bar{y}_R$. The $100\,(1-\alpha)\,\%$ confidence interval of the re-expressed mean for the original data may be defined as (Meloun et al., 1992; Meloun et al., 2000)

$$\bar{y}_R - I_L \leq \mu \leq \bar{y}_R + I_U \quad [65]$$

where

$$I_L = g^{-1}\left[\bar{z} + G - t_{1-\alpha/2}\,(n-1)\,\frac{s_z}{\sqrt{n}}\right] \quad [66]$$

$$I_U = g^{-1}\left[\bar{z} + G + t_{1-\alpha/2}\,(n-1)\,\frac{s_z}{\sqrt{n}}\right] \quad [67]$$

and

$$G = -\frac{1}{2}\frac{d^2g(y)}{dy^2}\left(\frac{dg(y)}{dy}\right)^{-2} s_z^2 \quad [68]$$

On the basis of the (known) actual transformation $z = g(y)$ and the estimates $\bar{z}$, s_z^2, it is easy to calculate the re-expressed estimates $\bar{y}_R$ and $s_{\bar{y}_R}^2$:

For a logarithmic transformation (when $\lambda = 0$) and $g(y) = \ln y$ the re-expressed mean and variance are calculated by (Meloun et al., 1992; Meloun et al., 2000)

$$\bar{y}_R \approx \exp\left(\bar{z} + 0.5\, s_z^2\right) \quad [69]$$

and

$$s_{\bar{y}_R}^2 = \bar{y}_R^2\, s_z^2 \quad [70]$$

-For $\lambda \neq 0$ and the Box-Cox transformation, the re-expressed mean $\bar{y}_R$ will be represented by one of the two roots of the quadratic equation (Meloun et al., 1992; Meloun et al., 2000)

$$\bar{y}_{R,1,2} = \frac{1}{2}\left(1 + \lambda_{\bar{z}} \pm \sqrt{1 + 2\lambda\left(\bar{z} + s_z^2\right) + \lambda^2\left(\bar{z}^2 - 2s_z^2\right)}\right)^{1/\lambda} \quad [71]$$

which is closest to the median $\tilde{y}_{0.5} = g^{-1}(\tilde{z}_{0.5})$. If $\bar{y}_R$ is known, the corresponding variance may be calculated from

$$s_y^2 = \bar{y}_R^{2(1-\lambda)}\, s_z^2 \quad [72]$$

APPLICATIONS

A survey of the analytical applications of the use of transformations in linear regression is given in Table 12. Hyperbolic, sigmoidal, and exponential decay curves may be transformed into a quasi-linear form of a simple regression equation. Binding and complexation processes are of great interest in many areas of research and are studied by a variety of experimental methods. Examples of logistic transformations as well as Box-Cox transformations are also included in the table. The analytical, pharmaceutical, biochemical, and clinical literature has been thoroughly revised.

CONCLUDING REMARKS

Topics in chemometrics have been treated in this journal, i.e., introductory references and references (Brown and Bear, 1993; Rusling, 1989; Lochmüller and Reese, 1998; Kubista et al., 1999). The use of transformations, however, has been considered in the present paper in detail. The aim of the transformation of response and or stimulus variables might be to restore linearity of the relationship (Meloun and Militky, 1994), but it may be also intended to stabilize the variance or to yield normally distributed residual errors. An added benefit of the transformation approach is that sometimes it can remove heterogeneous variance (Natrella, 1963). In practice, however, we usually cannot obtain one simple transformation which satisfies a number of different criteria (Draper and Hunter, 1969; Bartlett, 1947). Sometimes such simple transformations (a logarithmic a power, or a root) are variously supported by a general scientific theory or by some plausible distributional assumptions, or may be merely indicated empirically by plotting the data (Weisberg, 2005; AMC, 1994). Of course the transformed data should be checked for model inadequacies before proceeding.

Most authors have been concerned about transformations of the response. In contrast, transformation diagnostics for explanatory variables have been studied to a lesser degree. Not every data set should be transformed (Box and Draper, 1987). We may require that the distribution be strongly skewed (i.e., clearly in need of transformation). In addition, transformation will generally be successful only when the data set spans at least two orders of magnitude (Weisberg, 2005). One of the problems of transforming the response, and thereby of the model, is that the results from statistical analysis will be on a scale or metric different from the scale of original measurements (Andersen et al., 1999). In the words of Tukey (1977), "We now regard re-expression as a tool, something to let us do a better job of grasping data."

The ratio of skewness to its standard error can be read roughly as a standardized score from a normal distribution (i.e., absolute values exceeding 3 are unusual). The ratio of kurtosis to its standard error can be used as a normal score. A ratio less than 3 indicates shorter tails than a normal distribution (Aknazarova and Kafarov, 1982).

Box and Cox (1964) provided another general method for selecting transformations of the response that is applicable both in simple and multiple regression (Lee et al., 1999; Li and Moor, 2002). Since the seminal paper by Box and Cox in 1964, the Box-Cox type of power transformations has inspired a large amount of research on its applicability as well as on the drawback arising

TABLE 12
Some Selected Applications of Transformations

Comments	Ref.
A theoretical model to fit the four-parameter equation in logit-log applied in RIA analysis is presented	Diez Montoro et al., 2007
Bilogarithmic hyperbolic cosine method applied to mole ratio data	Boccio et al., 2007
A logistic regression model based on dose, solubility, and permeability of compounds is developed to predict the food effect on AUC (area under the curve of time-plasma concentration profile)	Gu et al., 2007
The thyroid gland decision-making system developed through the logistic regression is an excellent system demonstrated 98.7% accurate by the classification table	Kim and Yoon, 2007
Logistic regression was used to determine whether the probabilities of cell survival and electroporation depend on experimental conditions and cell properties	Agarwal et al., 2007
A global logistic model was used to study the effects of both quantitative variables (NaCl, acid, and potassium sulfate concentrations) and dummy variables (laboratory medium or brine, and citric, lactic, or acetic acids) on growth of Saccharomyces cerevisiae	López et al., 2007
Alternative method to the Arrhenius equation for termogravimetric analysis based on a logistic mixture model	Naya et al., 2006
Kinetics modeling applied to predict the stability of cholecystokinin fragment CCK-4 in aqueous solution	Oliva et al., 2006
Decimal logistic transformation to the sigmoidal calibration curve of ion-selective bulk optodes, for the determination of cations based on ionophore-chromoionophore chemistry	Capitan-Vallvey et al., 2006
Short-term exposure data measurements of food micotoxin ochratoxin A (OTA) transformed through two-parameter Box-Cox transformation	Counil et al., 2006
Log-log transformation without weighting is the simplest model to fit the calibration data for the determination of piperaquine (PC) in urine	Singtoroj et al., 2006
Decimal logistic transformation to the sigmoidal calibration curve of ion-selective bulk optodes, for the determination of anions based on hydrophobic membranes containing neutral ionophore and chromoionofore	Capitan-Vallvey et al., 2006
A bi-logarithmic transformation to evaluate stability constants from continuous variation data	Sayago and Asuero, 2006
A bi-logarithmic transformation method to evaluate stability constants from mole ratio data	Boccio et al., 2006
Linear model of $y^{0.8}$ (from Box-Cox power transformation) versus di-n-octildisulphide (DOD) concentration showed negligible bias from 1.51 to 75.5 mg/L	Xu and Hee, 2006
A bi-logarithmic transformation to evaluate overlapping acidity constants via spectrophotometric measurements	Fernández-Recamales et al., 2006
The Box-Cox transformation applied to soil data improves sample symmetry and stabilizes spread; the logarithmic plot of a profile likelihood function enables the optimum transformation parameter to be found	Meloun et al., 2005
Different methods are used, i.e., Box-Cox transformation, which are known to be able to deal with non-linearities present in data	Dieterle et al., 2004
The minimum detectable value (MDV) of polychlorinated biphenyls (PBCs) is assessed by linearizing the calibration graph using the transformation $y^* = y^p$ and $y^* = a + bx$, the p parameter being determined iteratively	Van Loco et al., 2003
Models generated when dealing with first order kinetic to profile the degradation of pest-control compounds in soil does not support the broad use of logarithmic transformation to stabilize the variance	Herman and Scherer, 2003
It is necessary to transform both sides of the model, because of the nature of the relationship between the response and the mechanistic model, in order to achieve symmetry and constant variance	Atkinson, 2003
Box-Cox transformation improves a sample symmetry in the mean value of 17-hydroxy pregnenolone in the umbilical blood of newborns	Meloun et al., 2003

(Continued on next page)

TABLE 12
Some Selected Applications of Transformations *(Continued)*

Comments	Ref.
A new approach to choosing the righ calibration model in introduced based on the well known DoE (Design of Experiments) methodology, including as one of its variables the Box-Cox transformation	Flaten and Walmsley, 2003
Opacing data, coupled with literature reports on phlebitis occurrence, were used to generate a logistic regression that indicates the probability of phlebitis given an opacity value measured by the apparatus	Johnson et al., 2003
In estimating mean water quality concentrations with detection limits, robustness is improved by searching a class of power transformations (Box-Cox) for the best approximating normal distribution	Shumway et al., 2002
In order to describe the relationship between the responses y and the concentration x, a transformation of data $\sqrt{x} - \sqrt{y}$was carried out and a weighted regression model $1/x^{\lambda}$ was chosen ($\lambda = 1.278$) for the determination of sotalol in human plasma	Chiap et al., 2001
Bias in fitting reciprocal data to a linear model and linear fitting of logarithmic transformed data: Montecarlo method	Tellinghuisen, 2001
Determination of 17-hydroxypregnenolone in umbilical blood of newborns	Meloun et al., 2001
The logistic analysis showed that the nicotine concentration was a significant predictor of reduction in craving during the free-smoking period	Gomeni et al., 2001
Heckel model provides a method for transforming a parametric view of the force and displacement signal to a linear relationship for purely plastic materials	Krumme et al., 2000
Influence of the calculated indices on the observed clinical efficacy of ciprofloxacin; the results prove that log-transformation of the independent variable improves the data fitting to linear model	Sánchez-Recio et al., 2000
Literature data, reporting the kinetic reaction rate order of the termal denaturation of β-lactoglobulin (β-lg), were reviewed using two statistical approaches including R^2 and the Box-Cox transformation	Jaskulka et al., 2000
The retention times in ion-interaction chromatography for the simultaneous separation of aromatic sulphonates; results shown not to be homogeneously distributed in the variable domain, the problem being solved by using the Box-Cox transformation	Marengo et al., 2000
An easy generalization to general linear model prediction is another main advantage of the Box-Cox transformation for prediction	Yang, 1999
Determination of 3,5,6,-trichloro-2-pyridinol (3,5,6-TCP) in human urine, by using a standard curve constructed by linear regression after performing a ln/logit data transformation of the concentration and the absorbance readings, respectively	Shackelford et al. 1999
Box-Cox transformation and power-of-the-mean modeling of the variance in analysis of standardized toxicity studies	Andersen et al., 1999
Appropriateness of using ordinary least squares on log-transformed data when one discuss statistical parameter estimation in the Arrhenius equation, which relates kinetics reaction rate to temperature	Sundberg, 1998
Investigation of the variability of optimal power models in contrast to common regression models within and between analytical methods	Kimanani and Lavigne, 1998
Box-Cox-type power transformation method is proposed as an objective criteria to fit and cross-validate bioanalytical calibration curves	Kimanani, 1998
Box-Cox and link function transformation can assist in the determination of reaction orders in pharmaceutical studies	Mälkki-Laine and Valkeila, 1998
A method for transforming data (Box-Cox) so that the assumption of **ANOVA** are met (or violated to a lesser degree) and apply it in analysis of data from a physiology experiment	Peltier et al., 1998
The log-transformation strongly changes the distribution of errors and the originally homocedastic observation vector according to the Arrhenius equation becomes strongly heterocedastic	Klicka and Kubácek, 1997

(Continued on next page)

TABLE 12
Some Selected Applications of Transformations *(Continued)*

Comments	Ref.
A computer program written in SAS code for the Box-Cox family of power transformations is presented. The purpose of the program is to suggest a power transformation for the positive continuous response variables in only regression and analysis of variance (ANOVA) models	Malaeb, 1997
Basic and Box-Cox transformations are the suggested methods to achieve normal variables. An application related to environmental data with atmospheric parameters and SO_2 and particle concentrations is developed	Mateu, 1997
First-order kinetic data for the persulfate-iodide reaction	Hemalatha and Noorbatcha, 1997
A first-order model was selected to represent quality deterioration kinetics from a variance-stabilizing characteristics of a logarithmic transformation (Arrhenius equation)	Vankerschaver et al., 1996
Retention data in micellar-reverse-phase liquid chromatography fitted after linearization by using sensitivity weights	García-Alvarez-Coque et al., 1996
HPLC method bilogarithmic transformation to define the calibration model and deal with the heterocedasticity of methotrexate peak area	Cociglio et al., 1995
Exponential decay with random errors added to the points	Logan, 1995.
Variance-stabilizing transformation applied to the penetration data in vitro skin penetration studies followed by outlier testing	Kasting et al., 1994
Calibration graphs constructed by applying linear regression on the logarithmic transformed data of peak-heights ratio analyte/internal standard and the amount of analyte; HPLC method for the quantification of cocuronium and its putative metabolites in biological samples	Kleef et al., 1993
Isotope dilution mass spectrometry hyperbolic calibration graph obtained by plotting peak-area ratio against the ratio of the amount of unlabelled compound to that of labelled compound, part of which may be linearized	Sabot and Pinatel, 1993
Logistic function is proposed to model the capacity factor (k') as a function of pH in HPLC	de Aguilar et al., 1993
The non-linear four-parameter logistic model may be linearized if β_1 and β_2 are assumed known; the transformation to linearity may induce heterocedasticity in the error response	O'Connell et al., 1993
Margules and Van Laar equation parameters are calculated, using several different linearization methods	Shacham and Wisniak, 1993
Methods to linearize atomic absorption spectrometry (AAS) calibration curves, based on a mathematical transformation of the absorbance and a correction for stray light are evaluated	Wang et al., 1992
A bilogarithmic transformation to evaluate overlapping acidity constants from potentiometric data	Asuero, 1992
Linear regression after suitable non-linear transformation applied to Arrhenius equation and other approaches	Ebel et al., 1991
Transformation involving the normalization of the dependent variable peak height or peak area ratio, y, and the plasma drug content	McLean et al., 1990
Linearized models, such as the Scatchard plot, of y/x versus y and the double reciprocal linearization, are applied to the Langmuir isotherm comparing results obtained with the NLR method	Harrison and katti, 1990
Bi-logarithmic hyperbolic cosine method for determining acidity constants of amphoteric substances from solubility measurements	Asuero, 1989
A flexible model that can be used to determine the proper transformation and choice of weight, illustrated by examples, is proposed	Ruppert et al., 1989
Recommendation as plotting procedure $\log \lvert S - a/(bc)\rvert$ vs. $\log c$, where an analytical signal S is expected to vary linearly with concentration c ($S = a + b\ c$)	Johnson, 1988; Welch et al., 1988

(Continued on next page)

TABLE 12
Some Selected Applications of Transformations *(Continued)*

Comments	Ref.
Equations expressing the mass ratio of natural to labeled material in the sample in terms of the isotope peak intensities, the calibration graph being described as either a straight line or a curve	Sabot et al., 1988
Comparison of six commercially available curve-fitting algorithms for calibration in flame atomic absorption spectrometry (FAAS)	Bysouth and Tyson, 1986
Bilogarithmic method for the spectrophotometric determination of overlapping acidity constants	Asuero et al., 1986 Asuero et al., 1986
Representation of most extended calibration data by using log-log plot of absorbance concentration for compactness and ease in reviewing data	Miller-Ihli et al., 1984
Weighted vs. unweighted logit-log radioimmunoassay data reduction	Fischer and Rodbard, 1983
Three linearization methods are presented for determining the parameters of a first-order kinetic decay superposed on a finite sum of power terms (three general line-lag linearization procedures)	Schwartz, 1981
Calculator program for weighted logit-log Radioimmunoassay data	McDonald, 1981
A program WRANL for the analysis of immunoassays which have a logistic dose-response relationship is described, iteratively regression analysis being used to obtain log dose-logit response lines for all preparations compared in an assay	Das and Tydeman, 1980
Statistical quality control system for linearized dose-response curve modeled after the suggestions of Rodbard	Cernosek., and Gutierrez-Cernosek 1978
Fortran IV program written to perform the transformation of variables and to determine, by the method of maximizing the likelihood function suggested by Box and Cox, the transformation parameters, and thus the appropriate functional form	Huang et al., 1978; Chang, 1977.
The four-parameter logistic function provides an empirical description of the dose/response curve by use of an objective lest-squares regression analysis (with the advantages of computerization and automation)	Rodbard and McClean, 1977
Arrhenius linearized equation: statistical analysis of the enthalpy-entropy compensation effect	Krug et al., 1976
Improved method of logit-log curve fitting, by adjusting end-point parameters in radioimmunoassay studies	Hatch et al., 1976
The relative merits of a number of methods available for the graphical display of radioimmunoassay dose-response curves (for curve fitting and dose interpolation) are discussed; e.g., log transformation of the dose axis, and logit (B/B_0) vs. log dose	Rodbard, 1974
Reviews of the reason for transforming data, the ways of developing a suitable transformation, and the various transformations found in the literature	Hoyle, 1973
A logarithmic method is described for the calculation of the transport parameters, K_m and $V_{\max}$, of a biological system obeying Michaelis-Menten kinetics	Barber et al., 1967
Study of linear transformations derived from Michaelis-Menten	Dowd and Riggs, 1965
Chymotripin-catalysed hydrolysis of methyl hippurate	Elmore et al., 1963

from it use; a great deal of interest was generated, both in theoretical work and in practical applications (Mateu, 1997; Chinn, 1996; Sakia, 1992; Andersen et al., 1999).

The Box-Cox data transformation is a simple method that can enable analysis of heterocedastic data sets so that the assumptions of the analysis of variance may be satisfied better, especially when other transformation procedures fail. Although the Box-Cox procedure could be less robust in the presence of outliers and could, after all, induce a mean bias (Logothetis, 1990), this technique has many advantages that should always kept in mind: i) it makes full use of the information provided in the available data; ii) it assures that the simplicity and normality of the model are valid assumptions; iii) it can provide reliable and shorter confidence intervals for λ; and iv) it can deal with

cases with no replication in each cell (Box and Meyer, 1986). Zarembka (1974), however, has indicated that the procedure is not robust with respect to heterocedasticity.

Transforming the data is sometimes felt to be a trick used by statisticians, a belief that is based on the idea that the natural scale of measurement is in some way sacrosanct (Altman, 1991). This is not the case, and indeed some measurements, i.e., pH, are effectively already log transformed values. Much of the statistical literature on transformations concentrates on transformation to normality rather than on variance-stabilizing transformation, despite the greater importance of the assumption of the homogeneity of variance than that of normality for analysis of variance.

The distribution of the random errors, e_i, is changed by any non-linear transformation of the y (Bates and Watts, 2007; Seber, 1977). We will often inspect the "empirical cumulative distribution of the residuals" to see whether y or its transform appears more nearly normal. If the transformation of y chosen for physical or algebraic reasons produces wide changes in the residuals when they are plotted against the fitted-y values, we may want to make some guess about how the variance of y is changing with μ or with some guess (Asuero and Gonzalez, 2007; Baumann, 1997; Tellinghuisen, 2008; Davidian and Haaland, 1990). If regression with equal weight is applied to transformed data as is often done, the results may be different from the intention, but this does not deny the intention.

The importance of transformation stems from the various limitations of non-linear approaches (Mager, 1991). Fortunately, the resulting bias after re-transformation (rectification) of the estimators is often negligible, compared to the experimental uncertainty of the data and in many cases, the consequences are less dangerous than the problems from employing non-linear regression models. As many theoretical questions of non-linear regression are unsolved up to now, the old technique of linearizing transformations should again be made the center of interest (Mager, 1991). In most cases, the transformations should be considered as a first approach in analyzing the data, while non-linear regression should be applied to the final stage of data analysis. Fitting the Michaelis-Menten equation (and extensions thereof) provides an interesting example of the limitations of both graphical methods and analytical methods based on data transformation (Nelder, 1991; Ruppert, 1991; Jurs, 1986).

The analysis of best transformation helps to avoid an unthinking automatic use of computer packages, according to the golden rule of data analysis, "keep in touch with your own data" (Mager, 1991).

ACKNOWLEDGMENTS

Work supported by the Junta de Andalucia (Spain) through grant excellence research project P06-FQM-02029, to who author is grateful.

REFERENCES

Acton, F. S. *Analysis of Straight Line Data* (Wiley, New York, 1959).

Agarwal, A.; Zudans, I.; Weber, E. A.; Olofsson, J.; Orwar, O.; Weber, S. G. Effect of cell size and shape on single-cell electroporation. *Analytical Chemistry* **2007**, *79*(10), 3589–3596.

Aknazarova, S.; Kafarov, V. *Experiment Optimization in Chemistry and Chemical Engineering* (Mir Publishers, Moscow, 1982).

Altman, D. G. *Practical Statistics for Medical Research* (Chapman and Hall, Boca Raton, FL, 1991), 143–146.

AMC, Is my calibration linear? *Analyst* **1994**, *119*(11), 2363–2366.

Andersen, H.; Spliid, H.; Larsen, S.; Christensen, N. D. Consistency in statistical analysis of standardized toxicity studies. *Toxicology Methods* **1999**, *9*(2), 71–89.

Anscombe, F. J. The transformation of Poisson, binomial and negative-binomial data. *Biometrika* **1948**, *35*(3–4), 246–254.

Armitage, P.; Berry, G.; Matthews, J. N. S. *Statistical Methods in Medical Research* (Blackwell Science, Oxford, 2002).

Aroian, L. A. A study of Fisher's z distribution and the related F distribution. *The Annals of Mathematical Statistics* **1941**, *12*(4), 429–448.

Asuero, A. G. A bilogarithmic method for the evaluation of amphoteric substances from solubility measurements. *International Journal of Pharmaceutics* **1989**, *52*(2), 129–137.

Asuero, A. G. A bilogarithmic method for the potentiometric evaluation of stability constants of two-step overlapping equilibria. *Analytical Letters* 25 (6) (1992):1143–1155.

Asuero, A. G. Buffer capacity of a polyprotic acid: first derivative of the buffer capacity and pKa values of single and overlapping equilibria. *Critical Reviews in Analytical Chemistry* **2007**, *37*(4), 269–301.

Asuero, A. G. Preliminary evaluation of bis(2-pyridyl)hydrazone as an analytical reagent. *Microchemical Journal* **1978**, *23*(3), 390–399.

Asuero, A. G.; Gonzalez, A. G. Some observations of fitting a straight line to data. *Microchemical Journal* **1989**, *40*(2), 216–225.

Asuero, A. G.; Gonzalez, G. Fitting straight lines with replicated observations by linear regression: Part III Weighting data. *Critical Reviews in Analytical Chemistry* **2007**, *37*(3), 143–172.

Asuero, A. G.; Gonzalez, G.; de Pablos, F.; Gomez Ariza, J. L. Determination of the optimum working range in spectrophotometric procedures. *Talanta* **1988**, *35*(7), 531–537.

Asuero, A. G.; Jiménez-Trillo, J. L.; Navas, M. J. Mathematical treatment of absorbance versus pH graphs of polybasic acids. *Talanta* **1986**, *33*(11), 929–934.

Asuero, A. G.; Jiménez-Trillo, J. L.; Navas, M. J. Spectrophotometric methods for the evaluation of acidity constants. II. Numerical methods for two-step overlapping equilibria. *Talanta* **1986**, *33*(6), 531–536.

Asuero, A. G.; Sayago, A.; A. G. Gonzalez. The correlation coefficient: An overview. *Critical Reviews in Analytical Chemistry* **2006**, *36*(1), 41–59.

Atkinson, A. C. Testing transformation to normality. *Journal of the Royal Statistical Society* B **1973**, *35*(3), 473–479.

Atkinson, A. C. *Plots, Transformations and Regression: An Introduction to Graphical Methods of Diagnostic Regression Analysis* (Clarendon Press, Oxford, 1985).

Atkinson, A. C.; Horwitz's rule, transforming both sides and the design of experiments for mechanistic models. *Applied Statistics* **2003**, *52*(3), 261–278.

Atkinson, A. C.; Riari, M.; Bivariate boxplots, multiple outliers, multivariate transformations and discriminant analysis: The 1997 Hunter Lecture. *Environmetrics* **1997**, *8*(6), 538–562.

Barber, H. E.; Welch, B. L.; Mackay, D. The use of the logarithmic transformation in the calculation of the transport parameters of a system that obeys Michaelis-Menten kinetics. *Biochemical Journal* **1967**, *103*(1), 251–255.

Barnet, V. *Environmental Statistical Methods and Applications* (New York, Wiley, 2004), 161–173.

Barnett, V. Probability plotting methods and order statistics. *Applied Statistics* **1975**, *24*(1), 95–108.

Bartlett, M. S.; The use of transformations. *Biometrics* **1947**, *3*(1), 39–52.

Bates, D. M.; Watts, D. G. *Nonlinear Regression Analysis and its Applications* (Wiley, New York, 2007).

Baumann, K. Regression and calibration techniques. Part II. Validation, weighted and robust regression. *Process Control Quality* **1997**, *10*(1–2), 75–112.

Bayne, C. K.; Rubin, I. B. *Practical Experimental Designs and Optimization Methods for Chemistry* (VCR Publishers, Deerfield Beach, FL, USA, 1986).

Belloto, J. R. T; Sokolovski, T. D. Residual analysis in regression. *American Journal of Pharmaceutical Education* **1985**, *49*(3), 295–303.

Bickel, P. J.; Doksum, K. A. An analysis of transformations revisited. *Journal of the American Statistical Association* **1981**, *76*(374), 296–311.

Bisgaard, S.; Fuller, H.; Report No. 119: Analysis of factorial experiments with defects or defectives as the response, Center for Quality and Productivity Improvement, in *Report Series in Quality and Productivity*, (University of Wisconsin, Madison, Wisconsin, 1994), *Quality Engineering* 7(2) (1994–95):429–443.

Bliss, C. I. The method of probits. *Science* **1934**, *79*(2037), 38–39.

Boccio, M.; Asuero, A. G.; Sayago, A. Spectrophotometric evaluation of stability constants of 1:1 weak complexes from mole ratio data using the bilogarithmic hyperbolic cosine method. *Journal of Analytical Chemistry* **2007**, *62*(9), 840–844.

Boccio, M.; Sayago, A.; Asuero, A. G. A bilogarithmic method for the spectrophotometric evaluation of stability constants of 1:1 weak complexes from mole ratio data. *International Journal of Pharmaceutics* **2006**, *318*(1–2), 70–77.

Bolton, S. *Pharmaceutical Statistics, Practical and Clinical Applications*, 4th ed. (Marcel Dekker, New York, 2004).

Box, G. E. P.; Cox, D. R. An analysis of transformations. *Journal of the Royal Statistical Society Ser. B* **1964**, *26*(2), 211–252.

Box, G. E. P.; Tidwell, P. W. Transformation of the independent variables. *Technometrics* **1962**, *4*(4), 531–549.

Box, G. E. P.; Tidwell, P. W. Transformation of the independent variables. *Technometrics* **1962**, *4*(4), 531–550.

Box, G. E. P; Draper, N. R. *Empirical-Model-Building and Response Surfaces* (Wiley, New York, 1987).

Box, G. E. P; Meyer, D. Dispersion effects from fractional designs. *Technometrics* **1986**, *28*(1), 19–28.

Brown, S. D.; Bear, R. S. Jr. Chemometrics techniques in electrochemistry: A critical review. *Critical Reviews in Analytical Chemistry* **1993**, *24*(2), 99–131.

Brownlee, K. A. *Statistical Theory and Methodology in Science and Engineering*, 2nd ed. (Robert E. Krieger, Malabar, FL, 1984).

Buja, A.; Tukey, P. A. *Computing and Graphics in Statistics* (Springer-Verlag, New York, 1991).

Buncher, C. R.; Tsay, J-Y. (Eds.) *Statistics in the Pharmaceutical Industry*, 2nd ed. (Marcel Dekker, New York, 1994), 85–92.

Bysouth, S. R.; Tyson, J. F. A comparison of curve fitting algorithms for flame absorption spectrometry. *Journal of Analytical Atomic Spectroscopy* **1986**, *1*(1), 85–87.

C.-H. Gu, Li, H.; Levons, J.; Lentz, K.; Gandhi, R. B.; Raghavan, K.; Smith, R. L. Predicting effect of food on extent of drug absorption on physicochemical properties. *Pharmaceutical Research* **2007**, *24*(6), 1118–1130.

Canavos, G. C. *Applied Probability and Statistical Methods* (Little, Brown and Company, Toronto, 1984).

Capitan-Vallvey, L. F.; Lapresta-Fernández, A.; Fernández-Ramos, M. D.; Cuadros-Rodríguez, L. *Sensors and Actuators B* **2006**, *117*(1), 27–34.

Capitan-Vallvey, L. F.; Arroyo-Guerrero, E.; Fernández-Ramos, M. D.; Cuadros-Rodríguez, L. Logic linearization of analytical response curves in optical disposable sensors based on coextraction for monovalent anions. *Analytica Chimica Acta* **2006**, *561*(1–2), 156–163.

Carroll, R. J.; Rupper, D. *Transformation and Weighting in Regression* (Chapman and Hall, New York, 1988).

Cernosek, S. F. Jr.; Gutierrez-Cernosek, R. M. Use of a programmable desk-top calculator for the statistical quality control of radioimmunoassay. *Clinical Chemistry* **1978**, *24*(7), 1121–1125.

Chambers, J.; Cleveland, W.; Kleiner, W.; Tukey, P. *Graphical Methods for Data Analysis* (Duxbury Press, Boston, MA, 1983).

Chang, H. S. A computer program for Box-Cox transformation and estimation technique. *Econometrica* **1977**, *45*(7), 1741.

Chatterjee, S.; Hadi, A. S. *Sensitivity Analysis in Linear Regression* (Wiley, New York, 1988).

Chiap, P.; Ceccato, A.; Buraglia, B. M.; Boulander, B.; Hubert, P.; Crommen, J. Development and validation of an automated method for the liquid chromatographic determination of sotalol in plasma using dialysis and trace enrichment on a cation-exchange pre-column as on-line sample preparation. *Journal of Pharmaceutical and Biomedical Analysis* **2001**, *24*(5–6), 801–814.

Chen, W.; Baghdasaryan, L.; Buranathiti, T.; Cao, J. Model validation via uncertainty propagation and data transformation. *AIAA Journal* **2004**, *42*(7), 1406–1415.

Chinn, S. Choosing a transformation. *Journal of Applied Statistics* **1996**, *23*(4), 395–404.

Chissom, B. S. Interpretation of the kurtosis statistic. *The American Statistician* **1970**, *24*(4), 19–22.

Cociglio, M.; Hillaire-Buys, D.; Alric, C. Determination of methotrexate and 7-hydroxymethotrexate by liquid chromatography for routine monitoring of plasma levels. *Journal of Chromatography B* **1995**, *674*(1), 101–110.

Collect, D. *Modelling Binary Data* (Chapman and Hall, Boca Raton, FL, 2004).

Connors, K. A. *A Textbook of Pharmaceutical Analysis*, 3rd, ed., (New York, Wiley, 1984)

Connors, K. A. *Chemical Kinetics: The Study of Reaction Rates in Solution* (Wiley-VCH, New York, 1990).

Cook, D. R.; Weisberg, S. *Applied Regression including Computing and Graphics* (Wiley, Chichester, 1999).

Cook, R. D.; Weisberg, S. *An Introduction to Regression Statistics* (Wiley, New York, 1994).

Counil, E.; Verger, P.; Volatier, J.-L. Fitness-for-purpose of dietary survey duration: A case study with the assessment of exposure to ochratoxin A. *Food and Chemical Toxicology* **2006**, *44*(4), 499–509.

Cressie, N. The exponential and power data transformations. *The Statistician* **1978**, *27*(1), 57–60.

Daniel, C.; Wood, F. S. *Fitting Equations to Data: Computer Analysis of Multifactor Data*, 2nd ed. (New York, Wiley, 1980).

Darlington, R. B. Is kurtosis really "peakedness?". *The American Statistician* **1970**, *24*(2), 19–22.

Das, R. E. G.; Tydeman, M. S. Iterative weighted regression analysis of logit responses. A computer program for analysis of bioassays and immunoassays. *Computer Programs in Biomedicine* **1980**, *15*(1), 13–22.

Davidian, M.; Dose calibration, in *Encyclopedia of Environmmetrics* eds. A. El-Shaarwari and W. Piegorsh (Wiley, New York, 2002).

Davidian, M.; Haaland, P. D. Regression and calibration with non constant error variance. *Chemometrics and Intelligent Laboratory Systems* **1990**, *9*(3), 231–248.

de Aguilar, P. F.; Bourguignon, B.; Khots, M. S.; Penninckx, W.; Massart, D. L. The use of logistic transformation in HPLC optimization. *Quimica Analitica* 12 (1993):177–182.

de Brauwere, A.; Pintelon, R.; De Ridder, D.; Schoukens, J.; Baeyens, W. Estimation of heterocedastic measurement noise variances. *Chemometrics and Intelligent Laboratory Systems* **2007**, *86*(1), 130–138.

de Levie, R. Curve fitting with least squares. *Critical Reviews in Analytical Chemistry* **2000**, *30*(1), 59–74.

de Levie, R. Estimating precision in derived quantities. *Chemical Educator* **2004**, *9*(2), 80–88.

de Levie, R. Spreadsheet simulation of chemical kinetics. *Critical Reviews in Analytical Chemistry* **2002**, *32*(1), 97–107.

de Levie, R. When, why, and how to use weighted least squares. *Journal of Chemical Education* **1986**, *63*(1), 10–15.

de Levie, R. *Advanced Excel for Scientific Data Analysis* (Oxford University Press, Oxford, 2004), 142–149.

de Levie, R. *How to use Excel in Analytical Chemistry and in General Scientific Data Analysis* (Cambridge University Press, Cambridge, 2001).

Dieterle, F.; Busche, S.; Gauglitz, G. Different approaches to multivariate calibration of nonlinear sensor data. *Analytical and Bioanalytical Chemistry* **2004**, *380*(3), 383–396.

Diez Montoro, R.; Salaver Salvador, M. T.; Moreno Frigols, J. L. General model and four equations for the fit calibration curves in RIA. Comparison of results. *Anales de la Real Academia Nacional de Farmacia* **2007**, *73*(1), 65–86.

Dowd, J. E.; Riggs, D. S. A comparison of estimates of Michaelis-Menten kinetic constants from various linear transformations. *Journal of Biological Chemistry* **1965**, *240*(2), 863–869.

Draper, N. R.; Cox, D. R. On distributions and their transformation to normality. *Journal of the Royal Statistical Society B* **1969**, *31*(3), 472–476.

Draper, N. R.; Hunter, W. H. Transformations: Some examples revisited. *Technometrics* **1969**, *11*(1), 23–40.

Draper, N. R.; Smith, H. *Applied Regression Analysis*, 3rd ed. (New York: Wiley, 1998).

du Toit, S. H. C.; Steyn, A. G. W; Stumf, R. H. *Graphical Exploratory Data Analysis* (Springer Verlag, New York, 1986).

Dudley, R. A.; Edwards, P.; Ekins, R. P.; Finney, D. J.; McKenzie, I. G. M; Raab, G. M.; Rodbard, D.; Rodgers, R. P. C; Guidelines for immunoassay data processing. *Clinical Chemistry* **1985**, *31*(8), 1264–1271.

Ebel, S.; Ledermann, M.; Reyer, B. Prediction of dosage form stability from degradation kinetics of the drug substance. 2. Evaluation of Arrhenius equation. *European Journal of Pharmaceutics and Biopharmaceutics* 37(2) (1991):80–87 (in German).

Einax, J. W.; Zwanziger, H. W.; Geiβ, S. *Chemometrics in Environmental Analysis* (VCH, Weiheim, 1997).

Elmore, D. T.; Kingston, A. E.; Shields, D. B. The computation of velocities and kinetic constants of reactions with particular reference to enzime-catalysed processes. *Journal of Chemical Society* (1963):2070–2078.

Emmens, C. W.; The dose/response relation for certain principles of the pituitary gland, and of the serum and urine of pregnancy. *Journal of Endocrinology* **1940**, *2*(2), 194–225.

Ender, P. <http://www.philender.com/courses/categorical/notes3/probit1.html>

EPA Guidance for Data Quality Assessment. *Practical Methods for Data Analysis*. EPA QA/G-9 QA00 Update, United States Environmental Protection Agency, EPA/600/R-96/084, July 2000, 4–42.

Esteban, M.; Ariño, C.; Diaz-Cruz, J. M. Chemometrics in electroanalytical chemistry. *Critical Reviews in Analytical Chemistry* **2006**, *36*(3–4), 295–313.

Feinberg, M. *La Validation des Méthodes d'Analyse. Une approache chimiometrique de l'assurance qualité au laboratoire* (Masson, Paris, 1996).

Fernandez-Recamales, A.; Vega-Perez, J. M.; Asuero, A. G.; Sayago, A. Acid-base equilibria of biacetylmonoxime-isonicotinoylhydrazone. *Journal of Analytical Chemistry* **2006**, *61*(4), 393–395.

Filliben, J. J.; Testing basic assumptions in the measurement process, in *Validation of the Measurement Process*, ACS Symposium Series 63 ed. J. R. Devoe (ACS, Washington, DC, 1977).

Filliben, J. J.; The probability plot correlation coefficient test for normality. *Technometrics* 17(1) (1975) 111–117.

Findlay, J. W. A.; Dillard, R. F. Appropriate calibration curve fitting in ligand binding assays. *The AAPS Journal* 9(2) (2007):E260–E267.

Finney, J. *Probit Analysis*, 3rd ed. (Cambridge University Press, Cambridge, 1971).

Fischer, L.; Rodbard, D. Logit-log radioimmunoassay data reduction: Weighted vs. unweighted. *Clinical Chemistry* **1983**, *29*(2), 391–392.

Flaten, G. R.; Walmsley, A. D. Usign design of experiments to select optimum calibration model parameters. *Analyst* **2003**, *128*(7), 935–943.

Freeman, M. F.; Tukey, J. W. Transformations related to the angular and the square root. *The Annals of Mathematical Statistics* **1950**, *21*(4), 607–611.

Gad, S. C. *Statistics and Experimental Design for Toxicologists*, 3th ed. (CRC Press, FL, Boca Raton, 1999), 49–51.

García-Alvarez-Coque, M. C.; Torres-Lapasió, J. R.; Baeza-Baeza, J. J. Description of the partitioning behaviour of solutes and data treatment in micellar liquid chromatography with modifiers. *Analytica Chimica Acta* **1996**, *324*(2–3), 163–173.

Gibbons, R. D. *Statistical Methods for Groundwater Monitoring* (Wiley, New York, 1994), 77–79.

Gomeni, R.; Teneggi, V.; Iavarone, L.; Squassante, L.; Bye, A. Population pharmacokinetic-pharmacodynamic model of craving in an enforced smoking cessation populations: Indirect response and probabilistic modeling. *Pharmaceutical Research* **2001**, *18*(4), 537–543.

Gonzalez, A. G.; Herrador, A. M.; Asuero, A. G. Practical digest for evaluating the uncertainty of analytical assays from validation data according to the LGVAM protocol. *Talanta* **2005**, *65*(4), 1022–1030.

Gonzalez, A. G.; Herrador, M. A.; Asuero, A. G. Intra-laboratory testing of method accuracy from recovery assays. *Talanta* **1999**, *48*(3), 729–736.

González, A. G.; Herrador, M. A.; Asuero, A. G.; Sayago, A. The correlation coefficient attacks again. *Accreditation and Quality Assurance* **2006**, *11*(5), 256–258.

Gottschalk, P. G.; Dunn, J. R. Determining the error of dose estimates and minimum and maximum acceptable concentrations from assay with nonlinear dose-response curves. *Computer Methods and Programs in Biomedicine* **2005a**, *80*(3), 204–215.

Gottschalk, P. G.; Dunn, J. R. The five-parameter logistics: A characterization and comparison with the four-parameter logistic. *Analytical Biochemistry* **2005b**, *343*(1), 54–65.

Graham, R. C. *Data Analysis for the Chemical Sciences. A Guide to Statistical Techniques* (VCH Publishers, New York, 1993).

Groeneveld, R. A.; Meeden, G.; The mode, median, and mean inequality. *The American Statistician* **1997**, *31*(3), 120–121.

Harrison, F.; Katti, S. K. Hazards of linearization of Langnuir's model. *Chemometrics and Intelligent Laboratory Systems* **1990**, *9*(3), 249–255.

Haswell, S. J. (Ed.), *Practical Guide to Chemometrics* (Marcel Dekker, New York, 1992).

Hatch, K. F.; Coles, E.; Busey, H.; Goldman, S. C. End-point parameter adjustment on a small desk-top programmable calculator for logit-log analysis of radioimmunoassay data. *Clinical Chemistry* **1976**, *22*(8), 1383–1389.

Havilcek, L. L.; Crain, R. C. *Practical Statistics for the Physical Sciences* (ACS, Washington D.C., 1988).

Hazen, A. Storage to be provided in impounding reservoirs for municipal water supply. *Transactions of the American Society of Civil Engineering* 77 (1914):1539–1659.

Healy, M. J. R; Statistical analysis of radioimmunoassay data. *Biochemical Journal* **1972**, *130*(1), 207–210.

Hemalatha, M. R. K.; I. Noorbatcha. An undergraduate physical chemistry experiment on the analysis of first-order kinetic data. *Journal of Chemical Education* **1997**, *74*(8), 972–974.

Herman, R. A.; Scherer, P. N. Comparison of linear and nonlinear regression for modelling the first-order degradation of pest-control substances in soil. *Journal of Food and Agricultural Food Chemistry* **2003**, *51*(16), 4722–4726.

Hernandez, F.; Johnson, R. A. The large-sample behavior of transformations to normality. *Journal of the American Statistical Association* **1980**, *75*(372), 855–861.

Herrador, M. A.; Jimenez, A. M.; Asuero, A. G. Spectrophotometric determination of zinc in potable waters and insulin with methylglyoxal bis(4-phenyl-3-thiosemicarbazone). *Analyst* **1987**, *112*(9), 1237–1246.

Hibbert, D. B. The uncertainty of a result from a linear calibration. *Analyst* **2006**, *131*(12), 1273–1278.

Hinkley, D. V. On power transformations to symmetry. *Biometrika* **1975**, *62*(1), 101–111.

Hoaglin, D. C.; Mosteller, F.; Tukey, J. *Understanding Robust and Exploratory Data Analysis* (Wiley, New York, 1983).

Horn, P. S. A measure for peakedness. *The American Statistician* **1983**, *37*(1), 55–56.

Hoyle, M. H. Transformations—An introduction and a bibliography. *International Statistical Reviews* **1973**, *41*(2), 203–223.

Huang, C.-L.; Moon, L. C.; Chang, H. S. A computer program using the Box-Cox transformation technique for the specification of functional form. *The American Statistician* **1978**, *32*(4), 144.

International Organization of Standards. Statistical interpretation of data. Tests for departure from the normal distribution ISO 5479, (International Organization of Standards, Geneva, 1997).

Jarque, C. M.; Bera, A. K. A test for normality of observations and regression residuals. *Internacional Statistical Reviews* **1987**, *55*(2), 163–172.

Jaskulka, F. J.; Smith, D. E.; Larntz, K. Determining the kinetic reaction rate order for the thermal denaturation of β-lactoglobulin using two statistical approaches. *International Dairy Journal* 10 (9) (2000):589–595.

John, J. A.; N.R. Draper, An alternative family of transformations. *Applied Statistics* **1980**, *29*(2), 190–197.

Johnson, D. C. Guest editorial. *Analytica Chimica Acta* **1988**, *204*(1–2), 1–5.

Johnson, J. L. H; He, Y.; Yalkowsky, S. H. Prediction of precipitation-induced phlebitis: A statistical validation of an *in vitro* model. *Journal of Pharmaceutical Sciences* **2003**, *92*(8), 1574–1581.

Jurs, P. C. *Computer Software Applications in Chemistry* (Wiley, New York, 1986).

Kapteyn, J. C. *Skew Frequency Curves in Biology and Statistics* (P. Noordhoff, Croningen, The Netherlands, 1903).

Kapteyn, J. C.; van Uwen, M. J. *Skew Frequency Curves in Biology and Statistics* (Hoitrema Brothers, Groningen, 1916).

Kasting, G. B.; Filloon, T. G.; Francis, W. R.; Meredith, M. P. Improving the sensitivity of in vitro skin penetration experiments. *Pharmaceutical Reseach* **1994**, *11*(12), 1747–1754.

Kateman, G.; Buidens, L. *Quality Control in Analytical Chemistry* (Wiley, New York, 1993).

Kennedy, J. B.; Neville, A. M. *Basic Statistical Methods for Engineers and Scientists*, 3rd ed. (New York, Harper, 1986).

Kihlberg, J. K.; Herson, J. H.; Schotz, W. E. Square root transformation revisited. *Applied Statistics* **1972**, *21*(1), 76–81.

Kim, Y. S.; Yoon, C. N. Methodology of the thyroid gland disease decision-making using profiling in steroid hormone pathway. *Journal of Pharmaceutical and Biomedical Analysis* **2007**, *43*(3), 1100–1105.

Kimanani, E. K. Bioanalytical calibration curves: Proposal for statistical criteria. *Journal of Pharmaceutical and Biomedical Analysis* **1998**, *16*(6), 1117–1124.

Kimanani, E. K.; Lavigne, J. Bioanalytical calibration curves: Variability of optimal powers between and within analytical methods. *Journal of Pharmaceutical and Biomedical Analysis* **1998**, *16*(6), 1107–1115.

Kleczkowski, A. The transformation of local lesion counts for statistical analysis. *Annals of Applied Biology* **1949**, *36*(1), 139–152.

Kleef, U. W.; Proost, J. H.; Roggeveld, J.; Wierda, J. M. K. H. Determination of rocuronium and its putative metabolites in body fluids and tissue homogenates. *Journal of Chromatography* **1993**, *621*(1), 65–76.

Klicka, R.; Kubácek, L. Statistical properties of linearization of the Arrhenius equation via the logarithmic transformation. *Chemometrics and Intelligent Laboratory Systems* **1997**, *39*(1), 69–75.

Krug, R. R.; Hunter, W. G.; Grieger, R. A. Enthalpy-entropy compensation! Some fundamental statistical problems associated with the

analysis of Vant't Hoff and Arrhenius data. *The Journal of Physical Chemistry* **1976**, *80*(21), 2335–2341.

Krumme, M.; Schwabe, L.; Frömming, K.-H. Development of computarized procedures for the characterisation of the tableting properties with eccentric machines: Extended Heckel analysis. *European Journal of Pharmaceutics and Biopharmaceutics* **2000**, *49*(3), 275–286.

Kubista, M.; Nygrem, J.; Elbergali, A.; Sjöback, R. Making reference samples redundant. *Critical Revies in Analytical Chemistry* **1999**, *29*(1), 1–28.

Lavagnini, I.; Magno, F. A statistical overview of univariate calibration, inverse regression, and detection limits: Application to gas chromatography/mass spectrometry technique. *Mass Spectrometry Reviews* **2007**, *26*(1), 1–18.

Leatherbarrow, R. J. Using linear and nonlinear regression to fit biochemical data. *Trends in Biochemical Sciences* **1990**, *15*(12), 455–458.

Lee, J. C.; Chen, D-T.; Hung, H-N.; Chen, J. J. Analysis of drug dissolution data. *Statistics in Medicine* **1999**, *18*(7), 799–814.

Li, B. B.; Moor, B. The general Box-Cox transformation in multiple regression analysis. *Communications in Statistics-Simulation and Computation* **2002**, *31*(4), 673–687.

Lochmüller, C. H.; Reese, C. E. Introduction to factor analysis. *Critical Reviews in Analytical Chemistry* **1998**, *28*(1), 21–49.

Logan, S. R. How to determine the best straight line. *Journal of Chemical Education* **1995**, *72*(10), 896–898.

Logothetis, N. Box-Cox transformations and the Taguchi methods. *Applied Statistics* **1990**, *39*(1), 31–48.

López, F. N. A.; Quintana, M. C. D; Fernández, A. G. Use of logistic regression with dummy variables for modelling the growth-non growth limits of *Saccharomyces cerevisiae* IGAL01 as a function of sodium chloride, acid type, and potassium sorbate concentration according to growth media. *Journal of Food Protection* **2007**, *70*(2), 456–465.

M. M. Sánchez-Recio, Colino, C. I.; Sánchez-Navarro, A. A retrospective analysis of pharmacokinetic/pharmacodynamic indices as indicators of the clinical efficacy of ciprofloxacin. *Journal of Antimicrobial Chemotherapy* **2000**, *45*(3), 321–328.

Macca, C.; Merkoci, A. Potentiometric characterization of weak acids by multiple sample addition –I. Linear equations and intrinsic performance of the method. *Talanta* **1994**, *41*(12), 2033–2042.

Maccá, C. Gran plots and rigorous linear plots for weak acids titration: the "chemical" rationale. *Fresenius Journal of Analytical Chemistry* **1990**, *336*(1), 29–35.

Mage, D. T.; An objective graphical method for testing normal distributions using probability plots. *The American Statistician* **1982**, *36*(2), 116–120.

Mager, P. P. *Design Statistics in Pharmacochemistry* (Wiley, New York, 1991), 20–44.

Malaeb, Z. A. A SAS code to correct for non-normality and non-constant variance in regression and ANOVA models using the Box-Cox method of power transformation. *Environmental Monitoring and Assessment* **1997**, *47*(3), 255–273.

Mandel, J. *The Statistical Analysis of Experimental Data* (Wiley-Interscience, New York, 1964).

Mandel, J.; Models, transformation of scale, and weighting. *Journal of Quality Technology* **1976**, *8*(2), 86–97.

Manly, B. F. J. Exponential data transformations. *The Statistician* **1976**, *25*(1), 37–42.

Marengo, E.; Gennaro, M. C.; Gianotti, V. Chemometrically assisted simultaneous separation of 21 aromatic sulfonates in ion-interaction RP-HPLC. *Chemometrics and Intelligent Laboratory Systems* **2000**, *53*(1–2), 57–67.

Mateu, J. Methods of assessing and achieving normality applied to environmental data. *Environmental Management* **1997**, *21*(5), 766–777.

McDonald, M. Calculator program for weighted logit-log radioimmunoassay data reduction. *Clinical Chemistry* **1981**, *27*(11), 1946.

McLean, A. M.; Ruggirello, D. A.; Banfield, C.; Gonzalez, M. A.; Bialer, M. Application of a variance-stabilizing transformation approach to linear regression of calibration lines. *Journal of Pharmaceutical Sciences* **1990**, *79*(11), 1005–1008.

Meier, P. C.; Zünd, R. E. *Statistical Methods in Analytical Chemistry* (Wiley, New York, 2000), 129–131.

Meites, L. Some new techniques for the analysis and interpretation of chemical data. *Critical Reviews in Analytical Chemistry* **1979**, *8*(1), 1–53.

Meloun, M.; Hill, M.; Militky, J.; Kupka, K. Analysis of large and small samples of biochemical and clinical data. *Clinical Chemistry and Laboratory Medicine* **2001**, *39*(3), 53–61.

Meloun, M.; Hill, M.; Militky, J.; Kupka, K. Assessment of the mean-value of 17-hydroxypregnenolone in the umbilical blood of newborns by the exploratory analysis of biochemical data. *Computer Methods and Programs in Biomedicine* **2003**, *70*(3), 187–197.

Meloun, M.; Hill, M.; Militky, J.; Kupka, K. Transformation in the PC-aided biochemical data analysis. *Clinical Chemistry and Laboratory Medicine* **2000**, *38*(6), 553–559.

Meloun, M.; Militki, J.; Computer-assisted data treatment in analytical chemometrics.IV. Classical estimates of parameters of location, scale, and shape. *Chemical Papers* **1995**, *49*(1), 68–73.

Meloun, M.; Militky, J. Computer assisted data treatment in analytical chemometrics. 3. Data transformation. *Chemical Papers* 48 (1994):164–169.

Meloun, M.; Militky, J.; Forina, M. *Chemometrics for Analytical Chemistry, Vol. 1: PC-Aided Statistical Data Analysis* (Ellis Horwood, New York, 1992), 71–77.

Meloun, M.; Sanka, M.; Nemec, P.; Kritkova, S.; Kupka, K. The analysis of soil cores polluted with certain metals using the Box-Cox transformation. *Environmental Pollution* **2005**, *137*(2), 273–280.

Miller, J. N. Basic statistical methods for analytical chemistry. Part 2. Calibration and regression methods. A review. *Analyst* **1991**, *116*(1), 3–14.

Miller, J. N.; Miller, J. C. *Statistics and Chemometrics for Analytical Chemistry*, 5th ed. (Ellis Horwood, Essex, 2005).

Miller-Ihli, N. J.; O'Haver, T. C.; Harnly, J. M. Calibration and curve fitting for extended range AAS. *Spectrochimica Acta* 39B (2–3) (1984):1603–1614.

Montgomery, D. C.; Peck, E. A. *Introduction to Linear Regression Analysis* (Wiley, New York, 1982).

Moors, J. J. A. The meaning of kurtosis: Darlington re-examined. *The American Statistician* **1986**, *40*(4), 283–284.

Mosteler, R.; Tukey, J. W. *Data Analysis and Regression: A second course in statistics* (Addison-Wesley, Reading, MA, 1977).

Mosteller, F.; Youtz, C. Tables of the Freeman-Tukey transformations for the binomial and Poisson distribution. *Biometrika* **1961**, *48*(3–4), 433–440.

Mullins, E. *Statistics for the Quality Control Laboratory* (RSC, Cambridge, 2003).

Mälkki-Laine, L.; Valkeila, E. Application of regression transformations to the determination of reaction orders in stability studies. *International Journal of Pharmaceutics* **1998**, *161*(1), 29–35.

Natrella, M. G.; The use of transformations, *Experimental Statistics*, National Bureau of Standards Handbook 91 (NBS, Washington, DC, 1963), Chapter 20, 20-1 to 20-13.

Naya, S.; Cao, R.; I. L. de Ullibarri, Artiaga, R.; Barbadillo, F.; García, A. Logistic mixture model versus Arrhenius for kinetic study of material degradation by dynamic thermogravimetric analysis. *Journal of Chemometrics* **2006**, *20*(3–4), 158–163.

Nelder, J. A. Generalized linear models for enzyme-kinetic data. *Biometrics* **1991**, *47*(4), 1605–1615.

Neter, J.; Kutner, M.; Nachtshiau, C.; Wasserman, W. *Applied Linear Statistical Models* (Richard Irwing, D.; Chicago, IL, 1996).

Norman, S.; Maeder, M.; Model-based analysis for kinetic and equilibrium investigations. *Critical Reviews in Analytical Chemistry* **2006**, *36*(3–4), 199–209.

O'Connell, M. A.; Belanger, B. A.; Haaland, P. D. Calibration and assay development using the four parameter logistic model. *Chemometrics and Intelligent Laboratory Systems* **1993**, *20*(2), 97–114.

Oliva, A.; Llabrés, M.; Fariña, J. B. Data analysis of kinetic modeling used in drug stability studies: Isothermal versus nonisothermal assay. *Pharmaceutical Research* **2006**, *23*(11), 2595–2602.

Ott, W. R. *Environmental Statistics and Data Analysis* (Lewis Publishers CRC Press, Boca Raton, FL, 1995).

Peace, K. E. *Biopharmaceutical Statistics for Drug Development* (Marcel Dekker, New York, 1988), 357–359.

Peltier, M. R.; Wilcos, C. J.; Sharp, D. C. Technical note: Application of the Box-Cox data transformation to animal science experiments. *Journal of Animal Science* **1998**, *76*(3), 847–849.

Petersen, P. H. Analytical quality specification for measurements reported on an ordinal scale. *Accreditation and Quality Assurance* **1999**, *4*(9–10), 406–409.

Petersen, P. H.; Sandberg, S.; Fraser, C. G.; Goldschmidt, H. A model for setting analytical quality specifications and design of control for measurements on the ordinate scale. *Clinical Chemistry and Laboratory Medicine* **2000**, *38*(6), 545–551.

Phillippe, J. *Les Méthodes Statistiques en Pharmacie et en Chimie (applications à la recherche, à la production et au controle)* (Masson, Paris, 1967), 17.

Phillips, G. R.; Eyring, E. M. Comparison of conventional and robust regression in analysis of chemical data. *Analytical Chemistry* **1983**, *55*(7), 1134–1138.

Plikaytis, B. D.; Turner, S. H.; Gheesling, L. L.; Carlone, G. M. Comparison of standard curve-fitting methods to quantitate *Neisseria meningitidis* group. A polysaccharide antibody levels by enzyme-linked immunosorbent assay. *Journal of Clinical Microbiology* **1991**, *29*(7), 1439–1446.

Press, W. H.; Teukolsky, S. A.; Vetterling, W. T.; Flannery, B. P. *Numerical Recipes inc. The Art of Scientific Computing* (Cambridge University Press, Cambridge, 1999).

Rawlings, J. O.; Pantula, S. G.; Dickey, D. A. *Applied Regression Analysis. A Research Tool*, 2nd ed. (Springer-Verlag, New York, 1998).

Rice, J. A. *Mathematical Statistics and Data Analysis* (Wadsworth & Brooks/Cole, Belmont, CA, 1988).

Rios, S. *Métodos Estadísticos*, 2nd ed. (Ediciones del Castillo, Madrid, 1977).

Ritz, C.; Streibis, J. C. Bioassay analysis using R. *Journal of Statistical Software* **2005**, *12*(5), 1–12.

Rocke, D. M.; Durbin, B.; Wilson, M.; Kahn, H. D. Modeling uncertainty in the measurement of low-level analytes in environmental analysis. *Ecotoxicology and Environmental Safety* **2003**, *56*(1), 78–92.

Rocke, D. M.; Lorenzato, S.; A two-component model for measurement error in analytical chemistry. *Technometrics* **1995**, *37*(2), 176–184.

Rodbard, D. Statistical quality control and routine data processing for radioimmunoassay and immunoradiometric assays. *Clinical Chemistry* **1974**, *20*(10), 1255–1270.

Rodbard, D.; Frazier, G. R. Statistical analysis of radioligand assay data. *Methods in Enzymology* 37 (1975):3–22.

Rodbard, D.; McClean, S. W. Automated computer analysis for enzyme-multiplied immunological techniques. *Clinical Chemistry* **1977**, *23*(1), 112–115.

Rodbart, D.; Guardabasso, V.; Munson, P. J. Statistical aspects of radioimmunoassay, handbook of experimental pharmacology, in *Radioimmunoassay in Basic and Clinical Pharmacology* ch. 8, pp. 193–212, eds. C. Patrono and B. A. Peskar, (Springer-Verlag, New York, 1987).

Rode, R. A.; Chinchilli, V. M. The use of Box-Cox transformations in the development of multivariate tolerance regions with applications to clinical chemistry. *The American Statistician* **1988**, *42*(1), 23–30.

Rodríguez, M. M.; Asuero, A. G. Studies on pyridylhydrazones derived from biacetyl as analytical reagents. *Microchemical Journal* **1980**, *25*(3), 309–322.

Rousseew, P. J.; Tutorial to robust statistics. *Journal of Chemometrics* **1991**, *5*(1), 1–20.

Royston, J. P.; A remark on algorithm AS-181 - The W test for normality (Algorithm R94). *Applied Statistics* **1995**, *44*(4), 547–551.

Royston, J. P.; The W test for normality (Algorithm AS-181). *Applied Statistics* **1982**, *31*(2), 176–180.

Ruppert, D. What is kurtosis, an influence function approach. *The American Statistician* **1987**, *41*(1), 1–5.

Ruppert, D.; Cressie, N.; Carroll, R. J. A transformation/weigthing model for estimating Michaelis-Menten parameters. *Biometrics* **1989**, *45*(2), 637–656.

Ruppert, D.; Cressie, N.; Carroll, R. J. Response to "Generalized linear models for enzyme-kinetic data" by J. A. Nelder. *Biometrics* **1991**, *47*(4), 1610–1614.

Rusling, J. F. Analysis of chemical data by computer modeling. *Critical Reviews in Analytical Chemistry* **1989**, *21*(1), 49–81.

Ryan, B. F.; Joiner, B. L.; Ryan, T. A. *Minitab Handbook*, 2nd ed. (Duxbury Press, Boston, 1985).

Sabot, J.-F.; Pinatel, H. Calculation of the confidence range in order to obtain a linear calibration graph in stable isotope dilution mass spectrometry: Application to reference methods and pharmacological studies. *Analyst* **1993**, *118*(7), 831–834.

Sabot, J.-F.; Ribon, B.; Kouadio-Kouakou, L.-P.; Pinatel, H. Comparison of two calculation procedures for gas chromatography mass spectrometry associated with stable isotope dilution. *Analyst* **1988**, *113*(12), 1843–1847.

Sachs, L. *Applied Statistics. A handbook of Technique*, 2nd ed. (Springer-Verlag New York, 1982).

Sakia, R. M. The Box-Cox transformation technique—A review. *The Statistician* **1992**, *41*(2), 169–178.

Sayago, A.; Asuero, A. G. Fitting straight lines with replicated observations by linear regression: Part II. Testing for homogeneity of variances. *Critical Reviews in Analytical Chemistry* **2004**, *34*(3–4), 133–146.

Sayago, A.; Asuero, A. G. Spectrophotometric evaluation of stability constants of 1:1 weak complexes from continuous variation data. *International Journal of Pharmaceutics* **2006**, *321*(1–2), 94–100.

Sayago, A.; Boccio, M.; Asuero, A. G. Fitting straight lines with replicated observations by linear regression: the least squares postulates. *Critical Reviews in Analytical Chemistry* **2004**, *34*(1), 39–50.

Schlesselman, J. Power families: A note on the Box and Cox transformation. *Journal of the Royal Statistical Society B* **1971**, *33*(2), 307–311.

Schwartz, L. M. Linearization methods for first-order kinetic analysis. *Analytical Chemistry* **1981**, *53*(2), 206–213.

Schwartz, L. M.; Statistical uncertainties of analyses by calibration of counting measurements. *Analytical Chemistry* **1978**, *50*(7), 980–985.

Sclove, S. L. (Y vs x) or (Log y vs x), *Technometrics* 14 (1972): 391.

Seber, G. A. F *Linear Regression Analysis* (Wiley, New York, 1977).

Shacham, M.; Wisniak, J. Error analysis of linearization methods in regression data for the Van Laar and Margules equations. *Industrial Engineering Chemical Research* **1993**, *32*(11), 2820–2825.

Shackelford, D. D.; Young, D. L.; Mihaliak, C. A.; Shurdut, B. A.; Itak, J. A. Practical immunochemical method for determination of 3,5,6-trichloro-2-pyridinol in human urine: Applications and considerations for exposure assessment. *Journal of Agricultural and Food Chemistry* **1999**, *47*(1), 177–182.

Shapiro, S. S. *How to Test Normality and Other Distributional Assumptions* (American Society for Quality Control, ASCQ, Milwaukee, Winsconsin, 1990).

Shapiro, S. S.; Wilk, M. B. An analysis of variance test for normality (complete samples). *Biometrika* **1965**, *52*(3–4), 591–611.

Shumway, R. H.; Azari, A. S.; Johnson, P. Estimating mean concentration under transformation for environmental data with detection limits. *Technometrics* **1989**, *31*(3), 347–356.

Shumway, R. H.; Azari, R. S.; Kayhanian, M. Statistical approaches to estimating mean water quality concentrations with detection limits. *Environmental Science and Tecnology* 36(15) (2002):3345–3353

Singer, J. M.; Pedroso-de Lima, A. C.; Tanaka, N. I.; Gonzalez-Lopez, V. A. To triplicate or not to triplicate. *Chemometrics and Intelligent Laboratory Systems* **2007**, *86*(1), 82–85.

Singtoroj, T.; Tarning, J.; Annerberg, A.; Ashton, M.; Berqvist, Y.; White, N. J.; Lindegardh, N.; Day, N. P. J; A new approach to evaluate regression models during validation of bioanalytical assays. *Journal of Pharmaceutical and Biomedical Analysis* **2006**, *41*(1), 219–227.

Sonnergaard, J. M. On the misinterpretation of the correlation coefficient in pharmaceutical sciences. *International Journal of Pharmaceutics* **2006**, *321*(1–2), 12–17.

Steliopoulos, P.; Sticke, E. Estimation of performance characteristics of a confirmation method for thyreostats in plasma by means of a weighted least-squares approach. *Analytica Chimica Acta* **2007**, *592*(2), 181–186.

Sundberg, R. Statistical aspects on fitting the Arrhenius equation. *Chemometrics and Intelligent Laboratory Systems* **1998**, *41*(2), 249–252.

Taylor, J. K. *Statistical Techniques for Data Analysis* (Lewis Publishers, Inc., Chelsea, MI, 1990).

Tellinghuisen, J. Bias and inconsistency in linear regression. *Journal of Physical Chemistry* A **2000**, *104*(50), 11829–11835.

Tellinghuisen, J. Least squares with non-normal data: Estimating experimental variance functions. *Analyst* **2008**, *133*(2), 161–166.

Tellinghuisen, J. Statistical error propagation. *Journal of Physical Chemistry A* **2001**, *105*(15), 3917–3921.

Tellinghuisen, J. Weighted least-squares in calibration: What difference does it make? *Analyst* **2007**, *132*(6), 536–543.

Tellinghuisen, J.; Van't Hoff analysis of K° (T). How good or bad? *Biophysical Chemistry* **2006**, *120*(2), 114–120.

Tomassone, R.; Lesquoy, E.; Millier, C. *La Régression, nouveaux regards sur une ancienne méthode statistique* (Masson, Paris, 1983).

Tukey, J. W. *Exploratory Data Analysis* (Addison-Wesley, Reading, MA, 1977).

Tukey, J. W.; On the comparative anatomy of transformations. *The Annals of Mathematical Statistics* **1957**, *28*(3), 602–632.

van Loco, J.; Hanot, V.; Huysmans, G.; Elskens, M.; Degroodt, J. M.; Beemaert, H. Estimation of the minimum detectable value for the determination of PCBs in fatty food samples by GC-ECD: A curvilinear calibration case. *Analytica Chimica Acta* **2003**, *483*(1–2), 413–418.

Vankerschaver, K.; Willock, F.; Smouth, C.; Hendrick, M.; Tobback, P. Mathematical modelling of temperature and gas composition effects on visual quality changes of cut endive. *Journal of Food Science* **1996**, *61*(3), 613–620.

Wang, X. N.; Smeyers-Verbeke, J.; Massart, D. L. Linearization of atomic absorption calibration curves. *Analusis* **1992**, *20*(4), 209–215.

Warton, D. I.; Wright, I. J.; Falster, D. S.; Westoby, M. Bivariate line-fitting methods for allometry. *Biological Reviews* 81 (2006):259–291.

Weisberg, S. *Applied Linear Regression*, 3rd ed. Wiley, New York, 2005).

Welch, L. E.; Mead, D. A.; Johnson, D. C. A comparison of pulsed amperometric detection and conductivity detection for carbohydrates. *Analytica Chimica Acta* **1988**, *204*(1–2), 323–327.

Wilson, M. D.; Rocke, D. M.; Durbin, B.; Kahn, H. D. Detection limits and goodness-of-fits measures for the two-component model of chemical analytical error. *Analytica Chimica Acta* **2004**, *509*(2), 197–208.

Xu, W.; Hee, S. S. Q. Gas-chromatography-mass spectrometry analysis of di-n-octyl disulfide in a straight oil metal working fluid. Application of differential permeation and Box-Cox transformation. *Journal of Chromatography* A **2006**, *1101*(1–2), 25–31.

Yang, Z. Predicting a future lifetime through Box-Cox transformation. *Lifetime Data Analysis* **1999**, *5*(3), 265–279.

Yeo, I. K.; Johnson, R. A. A new family of power transformations to improve normality or symmetry. *Biometrika* **2000**, *87*(4), 954–959.

Zarembka, P. Transformation of variables in econometrics, in *Frontiers of Econometrics* ed. P. Zarembka (Academic Press: New York), **1974**, 81–104.

Zorn, M. E.; Gibbons, R. D.; Sonzogni, W. C. Weighted least squares approach to calculating limits of detection and quantification by modeling variability as a function of concentration. *Analytical Chemistry* **1997**, *69*(15), 3089–3075.