Psychological Methods
1997, Vol. 2, No. 3, 292–307

1082-989X/97/$3.00

On the Meaning and Use of Kurtosis

Lawrence T. DeCarlo
Fordham University

For symmetric unimodal distributions, positive kurtosis indicates heavy tails and peakedness relative to the normal distribution, whereas negative kurtosis indicates light tails and flatness. Many textbooks, however, describe or illustrate kurtosis incompletely or incorrectly. In this article, kurtosis is illustrated with well-known distributions, and aspects of its interpretation and misinterpretation are discussed. The role of kurtosis in testing univariate and multivariate normality; as a measure of departures from normality; in issues of robustness, outliers, and bimodality; in generalized tests and estimators, as well as limitations of and alternatives to the kurtosis measure β_2, are discussed.

It is typically noted in introductory statistics courses that distributions can be characterized in terms of central tendency, variability, and shape. With respect to shape, virtually every textbook defines and illustrates skewness. On the other hand, another aspect of shape, which is kurtosis, is either not discussed or, worse yet, is often described or illustrated incorrectly. Kurtosis is also frequently not reported in research articles, in spite of the fact that virtually every statistical package provides a measure of kurtosis. This occurs most likely because kurtosis is not well understood and because the role of kurtosis in various aspects of statistical analysis is not widely recognized. The purpose of this article is to clarify the meaning of kurtosis and to show why and how it is useful.

On the Meaning of Kurtosis

Kurtosis can be formally defined as the standardized fourth population moment about the mean,

$$\beta_2 = \frac{E\,(X-\mu)^4}{(E\,(X-\mu)^2)^2} = \frac{\mu_4}{\sigma^4},$$

where E is the expectation operator, μ is the mean, μ_4 is the fourth moment about the mean, and σ is the standard deviation. The normal distribution has a kurtosis of 3, and $\beta_2 - 3$ is often used so that the reference normal distribution has a kurtosis of zero ($\beta_2 - 3$ is sometimes denoted as γ_2). A sample counterpart to β_2 can be obtained by replacing the population moments with the sample moments, which gives

$$b_2 = \frac{\Sigma(X_i - \bar{X})^4/n}{(\Sigma(X_i - \bar{X})^2/n)^2},$$

where b_2 is the sample kurtosis, X bar is the sample mean, and n is the number of observations.

Given a definition of kurtosis, what information does it give about the shape of a distribution? The left and right panels of Figure 1 illustrate distributions with positive kurtosis (leptokurtic), $\beta_2 - 3 > 0$, and negative kurtosis (platykurtic), $\beta_2 - 3 < 0$. The left panel shows that a distribution with positive kurtosis has heavier tails and a higher peak than the normal, whereas the right panel shows that a distribution with negative kurtosis has lighter tails and is flatter.

Kurtosis and Well-Known Distributions

Although a stylized figure such as Figure 1 is useful for illustrating kurtosis, a comparison of well-known distributions to the normal is also informative. The t distribution, which is discussed in introductory textbooks, provides a useful example. Figure 2 shows the t distribution with 5 *df*, which has a positive kurtosis of $\beta_2 - 3 = 6$, and the normal distribution, for which $\beta_2 - 3 = 0$. Note that the t distribution with 5 *df* has a variance of 5/3, and the normal distribution shown in the figure is scaled to also have a variance of 5/3.

The figure shows that the t_5 distribution has heavier

I thank Barry H. Cohen for motivating me to write the article and Richard B. Darlington and Donald T. Searls for helpful comments.

Correspondence concerning this article should be addressed to Lawrence T. DeCarlo, Department of Psychology, Fordham University, Bronx, New York 10458. Electronic mail may be sent via Internet to decarlo@murray.fordham.edu.

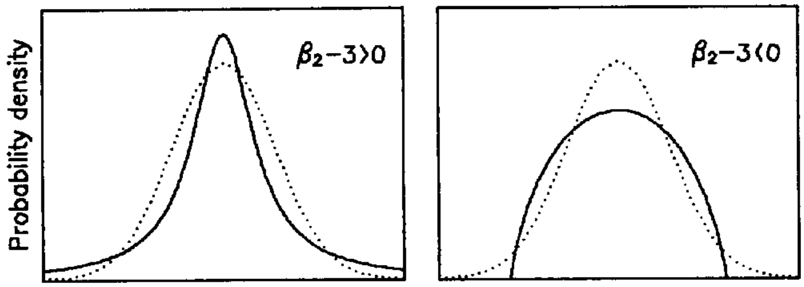

Figure 1. An illustration of kurtosis. The dotted lines show normal distributions, whereas the solid lines show distributions with positive kurtosis (left panel) and negative kurtosis (right panel).

tails and a higher peak than the normal. It is informative to note in introductory courses that, because of the heavier tails of the t distribution, the critical values for the t test are larger than those for the z test and approach those of the z as the sample size increases (and the t approaches the normal). Also note that the t_5 distribution crosses the normal twice on each side of the mean, that is, the density shows a pattern of higher-lower-higher on each side, which is a common characteristic of distributions with excess kurtosis.

With respect to negative kurtosis, a simple example is the continuous uniform (rectangular) distribution, for which $\beta_2 - 3 = -1.2$. Figure 3 shows the uniform distribution and the normal distribution, both with a variance of unity (the range for the uniform distribution is $\pm \sqrt{3}$). The figure shows that, relative to the normal, the uniform distribution has light tails, a flat center, and heavy shoulders. Also note that the uniform density, like that for the t, crosses the normal twice on each side of the mean.

Other examples of symmetric distributions with positive kurtosis are the logistic distribution, for which $\beta_2 - 3 = 1.2$, and the Laplace (double exponential) distribution, for which $\beta_2 - 3 = 3$; the logistic distribution has been used in psychology in signal detection theory and in item response theory, for example, whereas the Laplace has been used in vision research and in mathematical psychology. The symmetric binomial distribution with $p = .5$ offers an interesting example of a distribution with negative kurtosis: $\beta_2 - 3$ is negative, with a maximum of -2 for the two-point binomial ($n = 1$), and approaches zero as the index n increases (and the distribution approaches the normal).

Kurtosis and Density Crossings

Figures 2 and 3 show a basic characteristic of distributions with excess kurtosis: The densities cross the normal twice on each side of the mean. Balanda and MacGillivray (1988) referred to standardized densities that cross twice as satisfying the Dyson–Finucan condition, after Dyson (1943) and Finucan (1964), who showed that the pattern of density crossings is often associated with excess kurtosis.

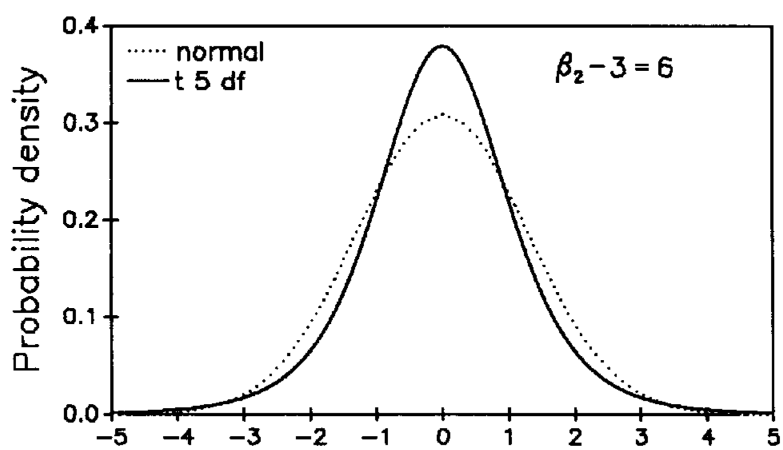

Figure 2. The t distribution with 5 df (solid curve) and the normal distribution (dotted curve), both with a variance of 5/3.

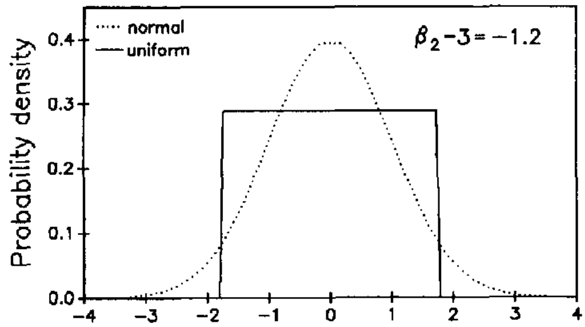

Figure 3. The uniform distribution and the normal distribution, both with a variance of unity.

A Simplified Explanation of Kurtosis

Why are tailedness and peakedness both components of kurtosis? It is basically because kurtosis represents a movement of mass that does not affect the variance. Consider the case of positive kurtosis, where heavier tails are often accompanied by a higher peak. Note that if mass is simply moved from the shoulders of a distribution to its tails, then the variance will also be larger. To leave the variance unchanged, one must also move mass from the shoulders to the center, which gives a compensating decrease in the variance and a peak. For negative kurtosis, the variance will be unchanged if mass is moved from the tails and center of the distribution to its shoulders, thus resulting in light tails and flatness. A similar explanation of kurtosis has been given by several authors (e.g., Balanda & MacGillivray, 1988; Ruppert, 1987). Balanda and MacGillivray noted that the definition of kurtosis is "necessarily vague" because the movement of mass can be formalized in more than one way (such as where the shoulders are located, p. 116).

The above explanation divides the distribution into tails, shoulders, and center, where for β_2 the shoulders are located at $\mu \pm \sigma$, as noted by Darlington (1970) and Moors (1986). It should be recognized that although tailedness and peakedness are often both components of kurtosis, kurtosis can also reflect the effect of primarily one of these components, such as heavy tails (which gives rise to some of the limitations discussed below). Thus, for symmetric distributions, positive kurtosis indicates an excess in either the tails, the center, or both, whereas negative kurtosis indicates a lightness in the tails or center or both (an excess in the shoulders). An approach by means of influence functions, discussed below, shows that kurtosis primarily reflects the tails, with the center having a smaller influence.

On Some Common Misconceptions Concerning Kurtosis

Further insight into kurtosis can be gained by examining some misconceptions about it that appear in a number of textbooks, ranging from those used in introductory courses to those used in advanced graduate courses. Three common errors are that (a) kurtosis is defined solely in terms of peakedness, with no mention of the importance of the tails; (b) the relation between the peak and tails of a distribution with excess kurtosis is described or illustrated incorrectly; and (c) descriptions and illustrations of kurtosis fail to distinguish between kurtosis and the variance.

An Old Error Revisited: Kurtosis As Simply Peakedness

Many textbooks describe kurtosis as simply indicating peakedness (positive kurtosis) or flatness (negative kurtosis), with no mention of the importance of the tails. Kaplansky (1945) referred to the tendency to describe kurtosis in terms of peakedness alone as a "common error," apparently made in statistics textbooks of the 1940s. As counterexamples to this notion, Kaplansky gave density functions for a distribution with positive kurtosis but a lower peak than the normal, and a distribution with negative kurtosis but a higher peak than the normal. The counterexamples illustrate why the definition of kurtosis solely in terms of peakedness or flatness can be misleading. Unfortunately, the error noted by Kaplansky (1945) and others still appears in a number of textbooks.

It is interesting to note that Kaplansky's (1945) two counterexamples to kurtosis as peakedness alone do not satisfy the Dyson–Finucan condition, because the distributions cross the normal more than twice on each side of the mean. As noted by Balanda and MacGillivray (1988), "If distributions cross more than the required minimum number of times, the value of β_2 cannot be predicted without more information. It is the failure to recognize this that causes most of the mistakes and problems in interpreting β_2" (p. 113).

A Recent Error: On Tailedness and Peakedness

Although the above error persists, many textbooks now correctly recognize that tailedness and peakedness are both components of kurtosis. However, and somewhat surprisingly, the description of the tails is often incorrect. In particular, a number of textbooks, ranging from introductory to advanced graduate texts, describe positive kurtosis as indicating peakedness and light (rather than heavy) tails and negative kurtosis as indicating flatness and heavy (rather than light) tails (e.g., Bollen, 1989; Howell, 1992; Kirk, 1990; Tabachnick & Fidell, 1996). This is a serious error, because it leads to conclusions about the tails that are exactly the opposite of what they should be.

Kurtosis and the Variance

Another difficulty is that a number of textbooks do not distinguish between kurtosis and the variance. For example, positive and negative kurtosis are sometimes described as indicating large or small variance, respectively. Note, however, that the kurtosis measure β_2 is scaled with respect to the variance, so it is not affected by it (it is scale free). Kurtosis reflects the shape of a distribution apart from the variance.

A related problem is that many textbooks use distributions with considerably different variances to illustrate kurtosis. This is apparently another old yet persistent problem; Finucan (1964), for example, noted that it appeared in statistics textbooks over 30 years ago:

> But a falsely simplified version of this as "peakedness" has unfortunately gained some currency and has even misled some elementary texts into presenting two curves of markedly unequal variances (e.g., intersecting only once on each side of the mean) as their example of a difference in kurtosis. (p. 112)

This is exactly the error made in more recent textbooks.

For the purpose of illustrating the shape of a distribution relative to the normal, as measured by β_2, the distributions should be scaled to have equal variances, as in Figures 2 and 3. Otherwise, any difference in the appearance of the distributions will not simply reflect a difference in kurtosis, but will also reflect the difference in variance. This is illustrated by Figure 4, which shows three normal distributions with variances of 0.5, 1, and 2. Relative to the standard normal ($\sigma^2 = 1$), the distribution with smaller variance ($\sigma^2 = 0.5$) appears to have a higher peak and lighter tails, whereas the distribution with larger variance ($\sigma^2 = 2$) appears flatter with heavier tails, which matches the incorrect descriptions and illustrations of kurtosis that are commonly given (note that the distributions also cross only once on each side of the mean). However, the three distributions are all normal, so they have exactly the same shape and kurtoses of $\beta_2 - 3 = 0$; the differences shown in Figure 4 simply reflect the difference in variance, not kurtosis.

Figure 4 is also relevant to the illustration of the t distribution with varying degrees of freedom that is given in many textbooks. The typical illustration shows that, as the degrees of freedom decrease, the t distribution appears flatter with heavier tails than the normal, and it is often described in this way. However, the t is actually more peaked than the normal, as Figure 2 shows (for $df = 5$); the apparent flatness in textbook illustrations arises because of the larger variance of the t as the degrees of freedom decrease (the variance is $df/[df - 2]$ for $df > 2$). In fact, Horn's (1983) measure of peakedness, noted below, suggests that, contrary to becoming flatter, the t distribution becomes more peaked as the degrees of freedom decrease (and the tails are heavier).

On the Use of Kurtosis

As taught in introductory courses, a basic goal of statistics is to organize and summarize data, and the mean and variance are introduced as summary mea-

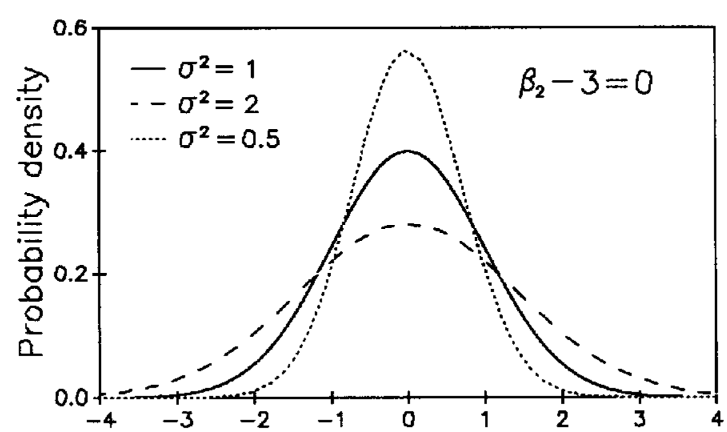

Figure 4. Three normal distributions with variances (σ^2) of 0.5, 1, and 2.

sures of the location and variability of a distribution. Similarly, skew and kurtosis provide summary information about the shape of a distribution. Although there are limitations to β_2 as a measure of kurtosis, as discussed below, the concepts of kurtosis, tail weight, and peakedness of a distribution "nevertheless play an important role in both descriptive and inferential statistics" (Balanda & MacGillivray, 1988, p. 114). Some of these roles are examined in this section.

First considered are uses of kurtosis that are practical and research oriented, followed by uses that are more conceptual and teaching oriented. From a practical perspective, the kurtosis and skewness statistics provided by virtually every statistical package provide information about shape that researchers should consider (and report), and tests based on these (or related) statistics have been shown to have excellent properties. Kurtosis is also relevant to issues of robustness, outliers, and modified tests and estimators, each of which is discussed in turn (also see Hopkins & Weeks, 1990; Jobson, 1991). At a more conceptual level, the simplified view of kurtosis given above serves to introduce the concepts of tails, center, and shoulders of a distribution, and these in turn are useful for a discussion of limitations of the measure β_2, alternatives to β_2, nonnormal distributions, and different approaches to formalizing the concept of kurtosis, as addressed in the last sections of this article.

Kurtosis and Normality

Part of a complete statistical analysis is an assessment of assumptions, including any distributional assumptions. When using normal theory methods, therefore, the assumption of normality should be checked. Other reasons for assessing normality are because departures from normality can affect tests and confidence intervals based on normal theory methods, and because the reduction of multivariate data to covariance matrices may overlook important aspects of the data; robustness and multivariate normality are discussed below.

Univariate normality. The use of the kurtosis statistic b_2, together with the skewness statistic $\sqrt{b_1}$, to assess normality has a long history in statistics, which was reviewed by D'Agostino (1986). The skewness statistic $\sqrt{b_1}$, like the kurtosis statistic b_2, is obtained from the sample moments as

$$\sqrt{b_1} = \frac{\Sigma(X_i - \overline{X})^3/n}{(\Sigma(X_i - \overline{X})^2/n)^{3/2}}$$

and indicates departures from symmetry.

Although tests based on the kurtosis and skewness statistics have been shown to have excellent properties for detecting departures from normality, D'Agostino, Belanger, and D'Agostino Jr. (1990) have noted that the results "have not been disseminated very well" (p. 316). In particular, studies have shown that tests based on the kurtosis and skewness statistics b_2 and $\sqrt{b_1}$ have good power properties, and D'Agostino et al. (1990) recommended their use for sample sizes as small as nine. In addition, the statistics provide information about the type and magnitude of departures from normality. Thus, the recommended strategy for assessing normality is to use tests and measures of skew and kurtosis in conjunction with omnibus tests, such as the Shapiro–Wilk test and the D'Agostino and Pearson (1973) K^2 test, both of which have good power properties, and graphical checks, such as normal probability plots.

With respect to software, the Shapiro–Wilk test and various plots (e.g., normal probability, quantile–quantile) are provided by many packages, such as SAS (SAS Institute, 1989) and SPSS (SPSS Inc., 1994). The skewness and kurtosis tests and the omnibus test K^2 can be obtained in SAS by using the macro given in D'Agostino et al. (1990) or in SPSS by using the macro given in the appendix of this article.

Multivariate normality. The multivariate normal distribution has several simplifying properties, one of which is that it is completely defined by the first two moments, which is important to the many multivariate methods that use a covariance matrix as input. The multivariate normal distribution plays a central role in multivariate methods because of this property and others (e.g., the marginals of multivariate normal random variables are also normal). Multivariate normality is a stronger assumption than univariate normality, and, just as for univariate normality, it should be checked; Cox and Wermuth (1994) noted a number of reasons for testing multivariate normality.

A first step in assessing multivariate normality is to separately test each variable for univariate normality, because univariate normality is a necessary condition for multivariate normality. This can be done by using plots and the skew, kurtosis, and omnibus tests noted above. Because several tests are performed, a Bonferroni correction can be used to control the Type I error rate (using α/p, where α is the desired experimentwise error rate and p is the number of variables).

Although univariate normality is a necessary condition for multivariate normality, it is not sufficient, which means that a nonnormal multivariate distribution can have normal marginals. So if univariate normality is not rejected, then the next step is to check for multivariate normality. Looney (1995) discussed an interesting example (Royston's hematology data) where, after transforming some of the variables, univariate tests do not indicate departures from normality for the marginal distributions, but multivariate tests indicate departures from multivariate normality.

Multivariate normality can be assessed by again using both formal and informal tools. For example, many textbooks discuss a graphical check of multivariate normality that is based on the squared Mahalanobis distances (e.g., Gnanadesikan, 1977; Jobson, 1992; Johnson & Wichern, 1988; Stevens, 1996; Tabachnick & Fidell, 1996). With respect to tests, Looney (1995) has recently discussed several multivariate tests that are based on the skew, kurtosis, and Shapiro–Wilk statistics; several tests are used because the different tests are sensitive to different types of departures from multivariate normality. SAS macros and FORTRAN routines are available to perform the tests (see Looney, 1995), and the SPSS macro in the appendix gives several multivariate tests of normality as well as a plot of the squared Mahalanobis distances. In addition, Cox and Wermuth (1994) discussed testing multivariate normality by using regression tests of linearity, which can easily be performed using standard software.

Robustness. This section notes that distributional shape is relevant to issues of robustness (e.g., see E. S. Pearson & Please, 1975) and to decisions about robust alternatives.

The frequent finding of departures from normality for many psychological variables (see Micceri, 1989) has led to interest in the robustness of various tests and estimators. Robustness has been investigated in sampling (Monte Carlo) studies by examining the effects of specific types of departures from normality; for example, symmetric distributions with heavy tails, such as the Laplace, Cauchy, and contaminated normal, are commonly used. Sampling studies and theoretical considerations (such as approximations of moments) have together shown that shape has a different effect on different tests and estimators. For example, a general finding for univariate and multivariate data is that tests of means appear to be affected by skew more than kurtosis, whereas tests of variances and covariances are affected by kurtosis more than skew (e.g., Jobson, 1991, p. 55; Mardia, Kent, & Bibby, 1979, p. 149). So, knowledge about expected types of departures from normality for a variable is relevant; this knowledge may come from prior research experience with the variable(s), from theoretical considerations, or both. For example, reaction time distributions tend to be positively skewed with a heavy tail, as shown by extensive research, and examples of nonnormal distributions arising from theory can be found in mathematical psychology (e.g., Luce, 1986).

A specific example is that it has long been recognized that kurtosis can have a large effect on tests of equality of variances (see Box, 1953; E. S. Pearson, 1931; Rivest, 1986). In response, robust tests, such as Levene's (1960) test (one version of which is given by the EXAMINE procedure of SPSS), have been developed. Note that shape is also relevant to Levene's test (and others; see Algina, Olejnik, & Ocanto, 1989) in that Brown and Forsythe (1974) found that different versions of the test appear to perform better, depending on whether the distribution is symmetric with heavy tails, in which case the trimmed mean can be used in place of the mean, or asymmetric, in which case the median can be used in place of the mean.

Kurtosis can also affect tests of equality of covariance matrices (e.g., Layard, 1974). More generally, the effect of kurtosis on analyses based on covariance matrices has received extensive attention in structural equation modeling, which is widely used in psychology (see Tremblay & Gardner, 1996). Browne (1982, 1984), for example, noted that kurtosis can have a large effect on significance tests and standard errors of parameter estimates, and measures of univariate and multivariate kurtosis are now included in most software for structural equation modeling. If kurtosis is thought to be a problem, a number of alternative (and possibly more robust) tests and estimators are available (see Hu, Bentler, & Kano, 1992).

Note that, even if a test is generally robust, shape can still be relevant in some situations, such as for small sample sizes (as are often obtained in applied research) or for models with random effects. For example, it is well known that tests of means, such as t tests and analyses of variance, are robust to moderate departures from normality (see Harwell, Rubinstein, Hayes, & Olds, 1992; Lindman, 1992; Stevens, 1996). However, Tiku, Tan, and Balakrishnan (1986) noted that, for small sample sizes, the power and Type I errors of the t test can be heavily affected by skewness and kurtosis; they also noted the relevance of shape to the choice of robust alternatives (p. 113).

Outliers. The topic of robustness is closely related to that of outliers. As noted above, kurtosis largely reflects tail behavior, and so its use for detecting outliers has been considered. In fact, kurtosis can be quite useful for detecting outliers in some situations (location slippage); discussions of approaches to detecting outliers can be found in Barnett and Lewis (1996), Jobson (1991), Tietjen (1986), and Tiku et al. (1986). Note that positive kurtosis can arise either because outliers are present, yet the distribution is normal, or because the underlying distribution is nonnormal, in which case heavy tailed nonnormal distributions can be considered as alternatives to the normal.

For multivariate data, a classical approach to detecting multivariate outliers that is discussed in many textbooks (e.g., Gnanadesikan, 1977; Jobson, 1992; Johnson & Wichern, 1988; Seber, 1984) is to examine the squared Mahalanobis distance for each case; a large value for a case relative to other cases can indicate a multivariate outlier. Note that the Mahalanobis distances are also related to Mardia's measure of multivariate kurtosis (see Mardia, 1970, 1980), in that the average of the sum of the Mahalanobis distances raised to the fourth power gives Mardia's measure (see Mardia et al., 1979). In fact, Mardia's test of multivariate kurtosis has been shown to have good properties for detecting multivariate outliers in some situations (Schwager & Margolin, 1982).

The relation of Mardia's (1970) measure to the Mahalanobis distances is also helpful for understanding the measure: A large value of Mardia's measure (relative to the expected value under multivariate normality) suggests the presence of one or more cases with large Mahalanobis distances, which are cases that are far from the centroid of all cases (potential outliers). So, Mardia's multivariate kurtosis in part indicates if the tails are heavy or light relative to those of the multivariate normal distribution; of course, a possible effect of the center also has to be kept in mind.

Generalized tests and estimators. Kurtosis also appears in a number of tests and estimators. For example, Searls and Intarapanich (1990) showed that an estimator of the variance that uses (a known value of) kurtosis in the divisor has a smaller mean squared error. Other examples are Box's (1953; Box & Andersen, 1955) modification of Bartlett's test for equal variances, which corrects the degrees of freedom using an estimate of kurtosis, and Layard's (1973) modification of a chi-square test for equality of covariance matrices, which also utilizes kurtosis.

Kurtosis also plays a role in structural equation modeling, in that it appears in one form or another in generalizations of normal theory methods, such as elliptical theory (see Bentler, 1989), which uses Mardia's (1970) multivariate kurtosis, and heterogeneous kurtosis theory (Kano, Berkane, & Bentler, 1990), which uses estimates of univariate kurtosis. These theories are more general in that they allow for multivariate distributions with heavier or lighter tails than the multivariate normal. Of course, an understanding of kurtosis is a requisite for understanding how the above theories generalize normal theory methods, so teaching about kurtosis lays the groundwork for later courses.

It should be noted that this section simply points out the role of kurtosis in various tests and estimators, and is not meant to imply that the use of kurtosis is necessarily a plus. In fact, some difficulties with elliptical theory estimators and tests appear to be due to problems with estimating multivariate kurtosis, and current research is examining other approaches.

Kurtosis and Nonnormal Distributions

The use of the standardized moments β_2 and $\sqrt{\beta_1}$ to describe shape goes back to Karl Pearson (1895), who also introduced a system of frequency curves to model departures from normality often found for real-world data; these and other systems of distributions, such as the Johnson system (Johnson, 1949), approach shape through the standardized moments, and plots of the (β_1, β_2) plane (Pearson diagrams) are often used for illustrations. The next sections examine several nonnormal distributions and note limitations of and alternatives to the kurtosis measure β_2.

Bimodality. The relation of kurtosis to bimodality illustrates both advantages and limitations of the measure β_2. In particular, Finucan (1964) noted that, because bimodal distributions can be viewed as having "heavy shoulders," they should tend to have negative kurtosis, that is, "a bimodal curve in general has also a strong negative kurtosis" (p. 112). Darlington (1970) took this view a step further and argued that kurtosis can be interpreted as a measure of unimodality versus bimodality, with large negative kurtosis indicating a tendency toward bimodality (the uniform distribution, with $\beta_2 - 3 = -1.2$, provides a dividing point). The symmetric binomial distribution with $n = 1$ and $p = .5$ offers a simple example: The mean is $np = .5$, the standard deviation is $\sqrt{np(1-p)} = 0.5$, and the (Bernoulli) distribution consists of mass at 0 and 1, so all the probability mass is concentrated at $\mu \pm \sigma$, the shoulders, and $\beta_2 - 3$ is -2, which is the lowest possible value.

Hildebrand (1971) noted that the family of symmetric beta distributions provides an example of a continuous distribution that nicely illustrates Darlington's (1970) point. The family has a shape parameter $\nu > 0$, and as ν varies, $\beta_2 - 3$ varies between −2 and 0. For $\nu = 1$, the distribution is uniform and $\beta_2 - 3 = -1.2$. For $\nu > 1$, the distribution is unimodal and approaches the normal as ν increases, and $\beta_2 - 3$ approaches zero (from the left). For $\nu < 1$, the distribution is bimodal and $\beta_2 - 3 < -1.2$, and as ν approaches zero, $\beta_2 - 3$ approaches −2 (and the modes approach ±1, as for the symmetric binomial). Figure 5 presents an illustration. The top left panel shows a standardized symmetric beta distribution with $\nu = 3$, for which $\beta_2 - 3 = -0.67$. Note that, relative to the standard normal, the distribution is flat with light tails, and also satisfies the Dyson–Finucan condition (it crosses the normal twice on each side of the mean). As ν approaches 1, the distribution approaches the uniform, as shown by the top right and bottom left panels of Figure 5 with $\nu = 1.2$ and $\nu = 1$. For $\nu < 1$, the distribution is bimodal and $\beta_2 - 3 < -1.2$, as shown by the bottom right panel for $\nu = 0.9$. Thus, the symmetric beta is a family of light-tailed distributions that range from unimodal to no mode (uniform) to bimodal, and this is reflected by $\beta_2 - 3$ going from near zero (close to normal) to −1.2 (uniform) to less than −1.2 (bimodal).

A limitation, however, is that kurtosis for bimodal distributions is not necessarily negative. Hildebrand (1971) noted, for example, that the double gamma family of distributions (also known as the reflected gamma; see Johnson, Kotz, & Balakrishnan, 1994) can have values of $\beta_2 - 3$ ranging from −2 to 3 when the distributions are bimodal (for an illustration, see Balanda & MacGillivray, 1988). This means that $\beta_2 - 3$ can be zero or positive for bimodal distributions; it depends on where the modes are located and on the heaviness of the tails. Adding contamination to the tails, for example, can result in zero or positive kurtosis for a bimodal or flat distribution (the double gamma has heavy tails). Moors (1986) noted that, as a consequence of Hildebrand's counterexample, "Darlington's result did not receive the attention it deserves" (p. 284).

Recognizing the above limitation, it should nevertheless be kept in mind that large negative kurtosis may indicate bimodality. In a similar vein, Bajgier and Aggarwal (1991) noted that a one-tailed test of negative kurtosis can be useful for detecting balanced mixtures of normal distributions in some situations. Other tools for detecting bimodality and mixtures are plots, such as a normal probability plot (see D'Agostino et al., 1990, for an illustration), and possibly Horn's (1983) measure of peakedness, discussed below. The CLUSTER procedure of SAS (SAS Insti-

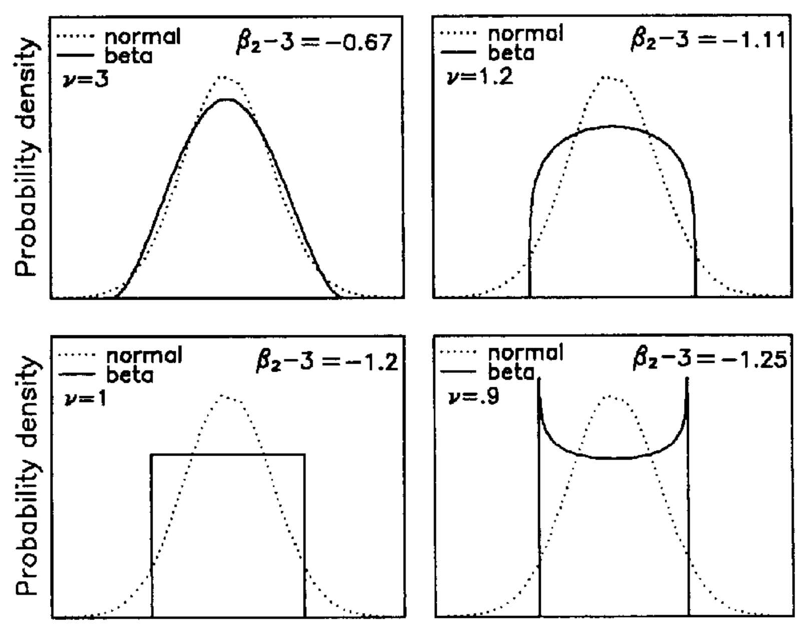

Figure 5. The solid lines show examples from the family of standardized symmetric beta distributions, which vary from unimodal to bimodal with shape parameter ν. The dotted lines show the standard normal distribution.

tute, 1989) also gives a "bimodality coefficient" (p. 561) based on kurtosis and skew that might be useful for skewed distributions (it remains to be investigated).

Limitations of β_2. A discussion of kurtosis would not be complete without noting some further limitations of the measure β_2. One problem is that more than one distributional shape can correspond to a single value of β_2. The family of symmetric Tukey lambda distributions provides an example: There are two distributions (two values of lambda) with different shapes for each value of β_2 (for an illustration, see Balanda & MacGillivray, 1988; Johnson et al., 1994; Joiner & Rosenblatt, 1971). For example, for $\lambda = 0.135$, the distribution approximates the normal and $\beta_2 - 3 = 0$. However, for $\lambda = 5.2$, again $\beta_2 - 3 = 0$, as for the normal, but the distribution is considerably more peaked than the normal. So, for $\lambda = 5.2$, β_2 does not reflect the peakedness of the distribution. This might occur because the distribution is peaked yet the tails are truncated; Chissom (1970), for example, used discrete distributions to show that truncation of the tails of a peaked distribution decreases its positive kurtosis, and similarly, symmetrically truncating the tails of a normal distribution can lower $\beta_2 - 3$ from 0 to -1.2 (see Figure 1 of Sugiura & Gomi, 1985). So, just as adding mass to the tails can eliminate the negative kurtosis of a flat or bimodal distribution, as noted above, removing mass from the tails can eliminate the positive kurtosis of a peaked distribution.

Another limitation of β_2 is that it cannot be used when the moments are not finite. The familiar t distribution provides an example, in that β_2 is only defined for $df > 4$ (the variance is also not finite for $df = 1$ or 2). Similarly, the Cauchy distribution does not have finite moments, so β_2 and the variance do not exist (note that a t with $df = 1$ is the Cauchy). In these and other examples, alternative measures are more useful. For example, Hogg's (1974) measure of tail heaviness can be used for the t (for all df) and Cauchy distributions, and Hogg noted that the normal, logistic, t, and Cauchy distributions can be ordered as given with respect to increasing tail heaviness.

Another difficulty is that β_2 does not necessarily allow a comparison of nonnormal distributions with respect to each other, but only with respect to the normal. For example, the normal, Laplace, and t (with 5 df) have values of $\beta_2 - 3$ of 0, 3, and 6, respectively, which reflects that the Laplace and t distributions are both more peaked with heavier tails than the normal. However, the Laplace is more peaked than the t_5, but its value of $\beta_2 - 3$ is smaller (3 vs. 6), so β_2 fails to reflect the greater peakedness of the Laplace. In this case, Horn's (1983) measure of peakedness is more useful, in that it indicates the peakedness of the Laplace relative to the t. Further discussion of comparative kurtosis and approaches based on partial orderings of distributions is provided by Balanda and MacGillivray (1988).

In sum, by their very nature, there are always limitations to summary measures, and this applies to the mean, variance, and skew, as well as to kurtosis. The purpose here is not to dictate the best or only approach to describing shape, but rather to clarify the meaning and relevance of kurtosis and to motivate researchers to look at the kurtosis statistic already included in their output. As is the case for any measure, informed use of kurtosis requires knowledge of both its advantages and limitations.

Alternatives to β_2. Because of limitations of β_2, a number of alternative measures of kurtosis have been proposed. Balanda and MacGillivray (1988) provide a review and note that the measures basically differ with respect to how they are scaled and where they position the shoulders of the distribution. For example, β_2 locates the shoulders at $\mu \pm \sigma$ and the measure is scaled with respect to the standard deviation, but other possibilities are to locate the shoulders at the quartiles, as done for a measure proposed by Groeneveld and Meeden (1984), or to scale the measure using the interquartile range, as done for some quantile based measures. Balanda and MacGillivray (1988) noted that the alternative measures "together form a haphazardly constructed collection of alternatives rather than a coherent alternative approach to the standardized fourth central moment" (p. 114), but the measures can nevertheless be useful (this is also a rapidly developing area in statistics).

In some applications, interest centers on part of the distribution rather than on the entire distribution. For example, in some situations (e.g., the study of floods or pollution levels), the tails (the extremes) of the distribution are of primary interest. Thus, some of the alternative measures attempt to measure tailedness alone or peakedness alone. For example, Hogg (1974) proposed a measure of tailedness, whereas Horn (1983) proposed a measure of peakedness. If interest is primarily on tail weight, then Hogg's measure of tailedness is less affected by outliers than β_2, whereas if the center (or shoulders) of the distribution is of greater interest, then Horn's measure better reflects peakedness than β_2 (in that the influence of the tails is bounded).

It should be recognized, however, that the above are not "pure measures of peakedness or of tail weight" (Ruppert, 1987, p. 5), because they are affected by both the center and tails. Ruppert offered a "simple intuitive reason" (p. 5) for this, which is the same as the simplified explanation of kurtosis given above. More important, he showed that the influence function is useful for understanding β_2 and other measures. For example, plots of the (symmetric) influence function show that β_2 is largely affected by tail weight and to a lesser extent by peakedness. The plots also show that Hogg's (1974) measure of tailedness indeed reflects tail weight and is less affected by outliers than β_2 but is still affected by the center (though to a lesser extent than β_2), and Horn's (1983) measure indeed reflects peakedness but is also affected by the tails (though to a lesser extent than β_2). Balanda and MacGillivray (1988) argue that a simultaneous consideration of tailedness and peakedness provides a better understanding of distributional shape (through an ordering based approach) than a separation of the concepts.

A Note on Kurtosis and Skewness Statistics

As shown above, the kurtosis statistic b_2 and the skewness statistic $\sqrt{b_1}$ are obtained by substituting the sample moments for the population moments. These statistics form the basis for univariate tests of kurtosis and skew, as discussed by D'Agostino et al. (1990), and for the multivariate tests of kurtosis and skew discussed by Looney (1995). They are also used in structural equation modeling in, for example, the mean scaled univariate kurtosis estimates used in elliptical estimators (e.g., Bentler, 1989, p. 214).

Many readers will recognize that the estimator of the second population moment (the variance) used in the kurtosis statistic b_2 is biased, because it uses n in the denominator instead of $n - 1$; similarly, the third and fourth sample moments are biased estimators of the third and fourth population moments. Another approach is to use unbiased estimators of the population moments, which gives the Fisher g statistics (see Fisher, 1970, p. 75), g_1 for skewness and g_2 for kurtosis,

$$g_1 = \frac{n\Sigma(X_i - \overline{X})^3}{(n-1)(n-2)[\Sigma(X_i - \overline{X})^2/(n-1)]^{3/2}}$$

and

$$g_2 = \frac{n(n+1)\Sigma(X_i - \overline{X})^4}{(n-1)(n-2)(n-3)[\Sigma(X_i - \overline{X})^2/(n-1)]^2} - \frac{3(n-1)^2}{(n-2)(n-3)}.$$

The Fisher g statistics are related to $\sqrt{b_1}$ and b_2, and D'Agostino et al. (1990) used this relation to compute $\sqrt{b_1}$ and b_2 from the g statistics given by SAS (SAS Institute, 1989) and SPSS (SPSS Inc., 1994; note that using the BIASKUR option in the CALIS procedure of SAS gives $\sqrt{b_1}$ and b_2).

A Macro for Measures and Tests of Skew and Kurtosis

The appendix provides a macro for measures and tests of univariate and multivariate skewness and kurtosis based on $\sqrt{b_1}$ and b_2. The macro can be used to supplement the graphs and statistics provided by many statistical packages. For example, normal probability plots and the Shapiro–Wilk statistic are provided by SAS (SAS Institute, 1989), SPSS (SPSS Inc., 1994) and other software.

For univariate data, the macro gives $\sqrt{b_1}$ and b_2 and tests based on them, as discussed by D'Agostino et al. (1990). It also provides two omnibus tests: K^2 (D'Agostino & Pearson, 1973), which simply sums the two chi-squares for skewness and kurtosis, and a score (Lagrange multiplier) test (see Jarque & Bera, 1987), which is a function of $\sqrt{b_1}$ and b_2. D'Agostino (1986) noted that K^2 might be less affected by ties than the Shapiro–Wilk statistic (Looney, 1995, recommends the use of a correction for ties for the Shapiro–Wilk).

In addition to the univariate statistics, the macro gives for multivariate data (a) Mardia's (1970) multivariate kurtosis; (b) Srivistava's (1984) and Small's (1980) measures and tests of multivariate kurtosis and skew, both of which are discussed by Looney (1995); (c) an omnibus test of multivariate normality based on Small's statistics (see Looney, 1995); (d) a list of the five cases with the largest squared Mahalanobis distances; (e) a plot of the squared Mahalanobis distances, which is useful for checking multivariate normality and for detecting multivariate outliers; and (f) Bonferroni adjusted critical values for testing for a single multivariate outlier by using the Mahalanobis distance, as discussed by Penny (1996), who also noted that the test gives results equivalent to those obtained by using jackknifed Mahalanobis distances.

To gain experience with the macro, one can use it to replicate the results of D'Agostino et al. (1990), who provided data in the form of a stem and leaf plot (62 participants from the Framingham heart study) and illustrated the use of the univariate statistics and normal probability plot. For multivariate analysis, the macro can be used with Fisher's iris data, which is readily available (e.g., in examples of the procedures CLUSTER and DISCRIM in the *SAS/STAT User's Guide* [SAS Institute, 1989]; also by anonymous ftp from Statlib: ftp to lib.stat.cmu.edu, it's in the datasets directory as part of visualizing.data) and the analysis reported by Looney (1995) can be (partially) replicated; note that the p value for Small's Q_2 in Looney's Table 2 should be .072 and not .074 (S. W. Looney, personal communication, September 1995). Gnanadesikan (1977, pp. 161–195) gave examples of the use of the plot of ordered squared Mahalanobis distances, which provides a visual check of multivariate normality (the points should lie along the diagonal) and can also help in the detection of multivariate outliers.

To use the macro, one needs two lines, one to include the macro in the program and the other to execute it. In SPSS 6.1 (SPSS Inc., 1994), the commands can be typed directly in the syntax window, for example, as the following:

```
include 'c:\spsswin\normtest.sps'.
normtest vars=x1,x2,x3,x4.
```

The first line includes the macro, which in this case is named normtest.sps and is in the spsswin directory, and the second line invokes the macro for variables x1 to x4, for example.

Conclusions

At the level of an introductory course, kurtosis can be illustrated with a stylized figure, such as Figure 1. Well-known distributions, such as the t and uniform, are also useful as examples. It is informative to note the relevance of density crossings to the kurtosis measure, and to distinguish kurtosis from the variance. In second and higher level courses, the role of kurtosis for assessing normality; for describing the type and magnitude of departures from normality; for detecting outliers, bimodality, and mixtures; in issues of robustness; in generalized tests and estimators, as well as limitations and alternatives to kurtosis, can be discussed.

References

Algina, J., Olejnik, S., & Ocanto, R. (1989). Type I error rates and power estimates for selected two-sample tests of scale. *Journal of Educational Statistics, 14,* 373–384.

Bajgier, S. M., & Aggarwal, L. K. (1991). Powers of goodness-of-fit statistics in detecting balanced mixed normal distributions. *Educational and Psychological Measurement, 51,* 253–269.

Balanda, K. P., & MacGillivray, H. L. (1988). Kurtosis: A critical review. *American Statistician, 42,* 111–119.

Barnett, V., & Lewis, T. (1996). *Outliers in statistical data* (3rd ed.). New York: Wiley.

Bentler, P. A. (1989). *EQS: Structural equations program manual.* Los Angeles: BMDP Statistical Software.

Bollen, K. A. (1989). *Structural equations with latent variables.* New York: Wiley.

Box, G. E. P. (1953). Non-normality and tests on variances. *Biometrika, 40,* 318–335.

Box, G. E. P., & Andersen, S. L. (1955). Permutation theory in the derivation of robust criteria and the study of departures from assumption. *Journal of the Royal Statistical Society,* Series B, *17,* 1–34.

Brown, M. B., & Forsythe, A. B. (1974). Robust tests for the equality of variances. *Journal of the American Statistical Association, 69,* 364–367.

Browne, M. W. (1982). Covariance structures. In D. M. Hawkins (Ed.), *Topics in applied multivariate analysis* (pp. 72–141). New York: Cambridge University Press.

Browne, M. W. (1984). Asymptotically distribution-free methods for the analysis of covariance structures. *British Journal of Mathematical and Statistical Society, 37,* 62–83.

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

Cox, D. R., & Wermuth, N. (1994). Tests of linearity, multivariate normality, and the adequacy of linear scores. *Applied Statistics, 43,* 347–355.

D'Agostino, R. B. (1986). Tests for the normal distribution. In R. B. D'Agostino & M. A. Stephens (Eds.). *Goodness-of-fit techniques* (pp. 367–419). New York: Marcel Dekker.

D'Agostino, R. B., Belanger, A., & D'Agostino, R. B., Jr. (1990). A suggestion for using powerful and informative tests of normality. *American Statistician, 44,* 316–321.

D'Agostino, R. B., & Pearson, E. S. (1973). Tests for departure from normality. Empirical results for the distributions of b_2 and $\sqrt{b_1}$. *Biometrika, 60,* 613–622.

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

Dyson, F. J. (1943). A note on kurtosis. *Journal of the Royal Statistical Society, 106,* 360–361.

Finucan, H. M. (1964). A note on kurtosis. *Journal of the Royal Statistical Society,* Series B, *26,* 111–112.

Fisher, R. A. (1970). *Statistical methods for research workers* (14th ed.). Edinburgh, Scotland: Oliver & Boyd.

Gnanadesikan, R. (1977). *Methods for statistical data analysis of multivariate observations.* New York: Wiley.

Groeneveld, R. A., & Meeden, G. (1984). Measuring skewness and kurtosis. *Statistician, 33,* 391–399.

Harwell, M. R., Rubinstein, E. N., Hayes, W. S., & Olds, C. C. (1992). Summarizing Monte Carlo results in methodological research: The one- and two-factor fixed effects ANOVA cases. *Journal of Educational Statistics, 17,* 315–339.

Hildebrand, D. K. (1971). Kurtosis measures bimodality? *American Statistician, 25,* 42–43.

Hogg, R. V. (1974). Adaptive robust procedures. *Journal of the American Statistical Association, 69,* 909–921.

Hopkins, K. D., & Weeks, D. L. (1990). Tests for normality and measures of skewness and kurtosis: Their place in research reporting. *Educational and Psychological Measurement, 50,* 717–729.

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

Howell, D. C. (1992). *Statistical methods for psychology* (3rd ed.). Boston: PWS-Kent.

Hu, L., Bentler, P. M., & Kano, Y. (1992). Can test statistics in covariance structure analysis be trusted? *Psychological Bulletin, 112,* 351–362.

Jarque, C. M., & Bera, A. K. (1987). A test for normality of observations and regression residuals. *International Statistical Review, 55,* 163–172.

Jobson, J. D. (1991). *Applied multivariate data analysis: Vol. 1. Regression and experimental design.* New York: Springer-Verlag.

Jobson, J. D. (1992). *Applied multivariate data analysis: Vol. 2. Categorical and multivariate methods.* New York: Springer-Verlag.

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

Johnson, N. L., Kotz, S., & Balakrishnan, N. (1994). *Continuous univariate distributions* (Vol. 1, 2nd ed.). New York: Wiley.

Johnson, R. A., & Wichern, D. W. (1988). *Applied multivariate statistical analysis* (2nd ed.). Englewood Cliffs, NJ: Prentice Hall.

Joiner, B. L., & Rosenblatt, J. R. (1971). Some properties of the range in samples from Tukey's symmetric lambda distributions. *Journal of the American Statistical Association, 66,* 394–399.

Kano, Y., Berkane, M., & Bentler, P. M. (1990). Covariance structure analysis with heterogeneous kurtosis parameters. *Biometrika, 77,* 575–585.

Kaplansky, I. (1945). A common error concerning kurtosis. *Journal of the American Statistical Association, 40,* 259.

Kirk, R. E. (1990). *Statistics: An introduction* (3rd ed.). Philadelphia: Holt, Rinehart & Winston.

Layard, M. W. J. (1973). Robust large-sample tests for homogeneity of variances. *Journal of the American Statistical Association, 68,* 195–198.

Layard, M. W. J. (1974). A Monte Carlo comparison of tests for equality of covariance matrices. *Biometrika, 16,* 461–465.

Levene, H. (1960). Robust tests for equality of variances. In I. Olkin, S. G. Churye, W. Hoeffding, W. G. Madow, & H. B. Mann (Eds.), *Contributions to probability and statistics* (p. 278–292). Stanford, CA: Stanford University Press.

Lindman, H. R. (1992). *Analysis of variance in experimental design.* New York: Springer-Verlag.

Looney, S. W. (1995). How to use tests for univariate normality to assess multivariate normality. *American Statistician, 49,* 64–70.

Luce, R. D. (1986). *Response times: Their role in inferring elementary mental organization.* New York: Oxford University Press.

Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika, 57,* 519–530.

Mardia, K. V. (1980). Tests of univariate and multivariate normality. In P. R. Krishnaiah (Ed.), *Handbook of statistics* (Vol. 1, pp. 279–320). Amsterdam: North-Holland.

Mardia, K. V., Kent, J. T., & Bibby, J. M. (1979). *Multivariate analysis.* New York: Academic Press.

Micceri, T. (1989). The unicorn, the normal curve, and other improbable creatures. *Psychological Bulletin, 105,* 156–166.

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

Pearson, E. S. (1931). The analysis of variance in cases of non-normal variation. *Biometrika, 23,* 114–133.

Pearson, E. S., & Please, N. W. (1975). Relation between the shape of population distribution and the robustness of four simple test statistics. *Biometrika, 62,* 223–241.

Pearson, K. (1895). Contributions to the mathematical theory of evolution II: Skew variation in homogeneous material. *Philosophical Transactions of the Royal Society of London, 186,* 343–412.

Penny, K. I. (1996). Appropriate critical values when testing for a single multivariate outlier by using the Mahalanobis distance. *Applied Statistics, 45,* 73–81.

Rivest, L. P. (1986). Bartlett's, Cochran's, and Hartley's tests on variances are liberal when the underlying distribution is long-tailed. *Journal of the American Statistical Association, 81,* 124–128.

Ruppert, D. (1987). What is kurtosis? *American Statistician, 41,* 1–5.

SAS Institute. (1989). *SAS/STAT user's guide* (Version 6, 4th ed.). Cary, NC: Author.

Schwager, S. J., & Margolin, B. H. (1982). Detection of multivariate normal outliers. *Annals of Statistics, 10,* 943–954.

Searls, D. T., & Intarapanich, P. (1990). A note on an estimator for the variance that utilizes the kurtosis. *American Statistician, 44,* 295–296.

Seber, G. A. F. (1984). *Multivariate observations.* New York: Wiley.

Small, N. J. H. (1980). Marginal skewness and kurtosis in testing multivariate normality. *Applied Statistics, 29,* 85–87.

SPSS Inc. (1994). *SPSS Advanced Statistics 6.1.* Chicago, IL: Author.

Srivistava, M. S. (1984). A measure of skewness and kurtosis and a graphical method for assessing multivariate normality. *Statistics and Probability Letters, 2,* 263–267.

Stevens, J. (1996). *Applied multivariate statistics for the social sciences* (3rd ed.). Hillsdale, NJ: Erlbaum.

Sugiura, N., & Gomi, A. (1985). Pearson diagrams for truncated normal and truncated Weibull distributions. *Biometrika, 72,* 219–222.

Tabachnick, B. G., & Fidell, L. S. (1996). *Using multivariate statistics* (3rd ed.). New York: Harper & Row.

Tietjen, G. L. (1986). The analysis and detection of outliers. In R. B. D'Agostino & M. A. Stephens (Eds.), *Goodness-of-fit techniques* (pp. 497–522). New York: Marcel Dekker.

Tiku, M. L., Tan, W. Y., & Balakrishnan, N. (1986). *Robust inference.* New York: Marcel Dekker.

Tremblay, P. F., & Gardner, R. C. (1996). On the growth of structural equation modeling in psychological journals. *Structural Equation Modeling, 3,* 93–104.

Appendix

An SPSS Macro for Univariate and Multivariate Skew and Kurtosis

```
preserve
set printback = none
define normtest (vars = !charend('/'))
matrix
get x /variables = !vars /names = varnames /missing = omit
compute n = nrow(x)
compute p = ncol(x)
compute s1 = csum(x)
compute s2 = csum(x&**2)
compute s3 = csum(x&**3)
compute s4 = csum(x&**4)
compute xbar = s1/n
compute j = make(n,1,1)
compute xdev = x-j * xbar
release x
compute m2 = (s2-(s1&**2/n))/n
compute m3 = (s3-(3/n*s1&*s2)+(2/(n**2)*(s1&**3)))/n
compute m4 = (s4-(4/n*s1&*s3)+(6/(n**2)*(s2&*(s1&**2)))-(3/(n**3)*(s1&**4)))/n
compute sqrtb1 = t(m3/(m2&**1.5))
compute b2 = t(m4/(m2&**2))
******** quantities needed for multivariate statistics ********
computes s = sscp(xdev)/(n-1)
compute sinv = inv(s)
compute d = diag(s)
compute dmat = make(p,p,0)
call setdiag(dmat,d)
compute sqrtdinv = inv(sqrt(dmat))
compute corr = sqrtdinv*s*sqrtdinv
*** principal components for Srivastava's tests ***
```

```
call svd(s,u,q,v)
compute pc = xdev*v
release xdev
*** Mahalanobis distances ***
compute sqrtqinv = inv(sqrt(q))
compute stdpc = pc*sqrtqinv
compute dsq = rssq(stdpc)
release stdpc
**************** univariate skew and kurtosis *****************
*** approximate Johnson's SU transformation for skew ***
compute y = sqrtb1*sqrt((n+1)*(n+3)/(6*(n-2)))
compute beta2 = 3*(n**2+27*n-70)*(n+1)*(n+3)/((n-2)*(n+5)*(n+7)*(n+9))
compute w = sqrt(-1+sqrt(2*(beta2-1)))
compute delta = 1/sqrt(ln(w))
compute alpha = sqrt(2/(w*w-1))
compute sub1 = delta*ln(y/alpha+sqrt((y/alpha)&**2+1))
compute psub1 = 2*(1-cdfnorm(abs(sub1)))
print {n}/title''Number of observations:'' /format = f5
print {p}/title''Number of variables:'' /format = f5
print {sqrtb1,sub1,psub1}
  /title''Measures and tests of skew:''
  /clabels = ''sqrt(b1)'',''z(b1)'',''p-value''
  /rnames = varnames /format = f10.4
*** Anscombe and Glynn's transformation for kurtosis
compute eb2 = 3*(n-1)/(n+1)
compute vb2 = 24*n*(n-2)*(n-3)/(((n+1)**2)*(n+3)*(n+5))
compute stm3b2 = (b2-eb2)/sqrt(vb2)
compute beta1 = 6*(n*n-5*n+2)/((n+7)*(n+9))*sqrt(6*(n+3)*(n+5)/(n*(n-2)*(n-3)))
compute a = 6+(8/beta1)*(2/beta1+sqrt(1+4/(beta1**2)))
compute zb2 = (1-2/(9*a)-((1-2/a)/(1+stm3b2*sqrt(2/(a-4))))&**(1/3))/sqrt(2/(9*a))
compute pzb2 = 2*(1-cdfnorm(abs(zb2)))
compute b2minus3 = b2-3
print {b2minus3,zb2,pzb2}
  /title''Measures and tests of kurtosis:''
  /clabels = ''b2-3'',''z(b2)'',''p-value''
  /rnames = varnames /format = f10.4
compute ksq = sub1&**2+zb2&**2
compute pksq = 1-chicdf(ksq,2)
compute lm = n*((sqrtb1&**2/6)+(b2minus3&**2/24))
compute plm = 1-chicdf(lm,2)
print
  /title''Omnibus tests of normality (both chisq, 2 df):''
print {ksq,pksq,lm,plm}
  /title'' D'Agostino & Pearson K sq          Jarque & Bera LM test''
  /clabels = ''K sq'',''p-value'',''LM'',''p-value''
  /rnames = varnames /format = f10.4
do if p>1
print
/title''*************** Multivariate Statistics ***************''
*** Small's multivariate tests ***
compute uinv = inv(corr&**3)
compute uinv2 = inv (corr&**4)
compute q1 = t(sub1)*uinv*sub1
* note: the variant of Small's kurtosis uses Anscombe & Glynn's
* transformation in lieu of SU (A & G is simpler to program)
```

```
compute q2 = t(zb2)*uinv2*zb2
compute pq1 = 1-chicdf(q1,p)
compute pq2 = 1-chicdf(q2,p)
print /title''Tests of multivariate skew:''
print {q1,p,pq1}/title'' Small's test (chisq)''
  /clabels = ''Q1'',''df'',''p-value''/format = f10.4
*** Srivastava's multivariate tests ***
compute pcs1 = csum(pc)
compute pcs2 = csum(pc&**2)
compute pcs3 = csum(pc&**3)
compute pcs4 = csum(pc&**4)
release pc
compute mpc2 = (pcs2-(pcs1&**2/n))/n
compute mpc3 = (pcs3-(3/n*pcs1&*pcs2) +(2/(n**2)* (pcs1&**3)))/n
compute mpc4 = (pcs4-(4/n*pcs1&*pcs3) +(6/(n**2)* (pcs2&*(pcs1&**2)))-(3/(n**3)* (pcs1&**4)))/n
compute pcb1 = mpc3/(mpc2&**1.5)
compute pbc2 = mpc4/(mpc2&**2)
compute sqb1p = rsum(pcb1&**2)/p
compute b2p = rsum(pcb2)/p
compute chib1 = sqb1p*n*p/6
compute normb2 = (b2p-3)*sqrt(n*p/24)
compute pchib1 = 1-chicdf(chib1,p)
compute pnormb2 = 2*(1-cdfnorm(abs(normb2)))
print {chib1,p,pchib1}
  /title'' Srivastava's test''
  /clabels = ''chi(b1p)'',''df'',''p-value''/format = f10.4
print /title''Tests of multivariate kurtosis:''
print {q2,p,pq2}
  /title'' A variant of Small's test (chisq)''
  /clabels = ''VQ2'',''df'',''p-value''/format = f10.4
print {b2p,normb2,pnormb2}
  /title'' Srivastava's test''
  /clabels = ''b2p'',''N(b2p)'',''p-value''/format = f10.4
*** Mardia's multivariate kurtosis ***
compute b2pm = csum(dsq&**2)/n
compute nb2pm = (b2pm-p*(p+2))/sqrt(8*p*(p+2)/n)
compute pnb2pm = 1-cdfnorm(abs(nb2pm))
print {b2pm,nb2pm,pnb2pm}
  /title'' Mardia's test''
  /clabels = ''b2p'',''N(b2p)'',''p-value''/format = f10.4
compute q3 = q1+q2
compute q3df = 2*p
compute pq3 = 1-chicdf(q3,q3df)
print /title''Omnibus test of multivariate normality:''
print {q3,q3df,pq3}
  /title'' (based on Small's test, chisq)''
  /clabels = ''VQ3'',''df'',''p-value''/format = f10.4
end if
compute cse = {1:n}
compute case = t(cse)
compute rnk = rnkorder(dsq)
compute top = (n+1)-rnk
compute pvar = make(n,1,p)
compute ddf = make(n,1,(n-p-1))
compute ncase = make(n,1,n)
```

```
compute a01 = make(n,1,(1-.01/n))
compute a05 = make(n,1,(1-.05/n))
compute mahal = {case,rnk,top,dsq,pvar,ddf,ncase,a01,a05}
save mahal /outfile = temp
  /variables = case,rnk,top,dsq,pvar,ddf,ncase,a01,a05
end matrix
get file = temp
sort cases by top (a)
do if case = 1
compute f01 = idf.f(a01,pvar,ddf)
compute f05 = idf.f(a05,pvar,ddf)
compute fc01 = (f01*pvar*(ncase-1)**2)/(ncase*(ddf+pvar*f01))
compute fc05 = (f05*pvar*(ncase-1)**2)/(ncase*(ddf+pvar*f05))
print space
print
  /'Critical values (Bonferroni) for a single multivar. outlier:'
print space
print
  /' critical F(.05/n) = 'fc05 (f5.2)' df = 'pvar (f3)','ddf (f4)
print
  /' critical F(.01/n) = 'fc01 (f5.2)' df = 'pvar (f3)', 'ddf (f4)
print space
print /'5 observations with largest Mahalanobis distances:'
end if
execute
do if top < 6
print
  /' rank = 'top (f2)' case# = 'case (f4)' Mahal D sq = 'dsq
end if
execute
compute chisq = idf.chisq((rnk-.5)/ncase,pvar)
plot
  /title = ''Plot of ordered squared distances''
  /symbols = '*'
  /horizontal = ' Chi-square p,(i-.5/n)' min(0)
  /vertical = ' Mahalanobis D squared' min(0)
  /plot = dsq with chisq
execute
!enddefine
restore
```

Received January 24, 1996
Revision received August 28, 1996
Accepted December 5, 1996 ■