Cain, M., Zhang, Z., & Yuan, K. (2017). Univariate and Multivariate Skewness and Kurtosis for Measuring Nonnormality: Prevalence, Influence and Estimation. Behavior Research Methods, 49(5), 1716–1735.

Univariate and Multivariate Skewness and Kurtosis for Measuring Nonnormality: Prevalence, Influence and Estimation

Meghan K. Cain, Zhiyong Zhang, and Ke-Hai Yuan

University of Notre Dame

Author Note

This research is supported by a grant from the U.S. Department of Education (R305D140037). However, the contents of the paper do not necessarily represent the policy of the Department of Education, and you should not assume endorsement by the Federal Government.

Correspondence concerning this article can be addressed to Meghan Cain (mcain3@nd.edu), Ke-Hai Yuan (kyuan@nd.edu), or Zhiyong Zhang (zzhang4@nd.edu), Department of Psychology, University of Notre Dame, 118 Haggar Hall, Notre Dame, IN 46556.

Abstract

Nonnormality of univariate data has been extensively examined previously (Blanca et al., 2013; Micceri, 1989). However, less is known of the potential nonnormality of multivariate data although multivariate analysis is commonly used in psychological and educational research. Using univariate and multivariate skewness and kurtosis as measures of nonnormality, this study examined 1,567 univariate distriubtions and 254 multivariate distributions collected from authors of articles published in Psychological Science and the American Education Research Journal. We found that 74% of univariate distributions and 68% multivariate distributions deviated from normal distributions. In a simulation study using typical values of skewness and kurtosis that we collected, we found that the resulting type I error rates were 17% in a *t*-test and 30% in a factor analysis under some conditions. Hence, we argue that it is time to routinely report skewness and kurtosis along with other summary statistics such as means and variances. To facilitate future report of skewness and kurtosis, we provide a tutorial on how to compute univariate and multivariate skewness and kurtosis by SAS, SPSS, R and a newly developed Web application.

Keywords: nonnormality, skewness, kurtosis, software

Univariate and Multivariate Skewness and Kurtosis for Measuring Nonnormality: Prevalence, Influence and Estimation

Almost all commonly used statistical methods in psychology and other social sciences are based on the assumption that the collected data are normally distributed. For example, *t*- and *F*-distributions for mean comparison, Fisher Z-transformation for inferring correlation, and standard errors and confidence intervals in multivariate statistics are all based on the normality assumption (Tabachnick & Fidell, 2012). Researchers rely on these methods to accurately portray the effects under investigation, but may not be aware that their data do not meet the normality assumption behind these tests or what repercussions they face when the assumption is violated. From a methodological perspective, if quantitative researchers know the type and severity of nonnormality that researchers are facing, they can examine the robustness of normal-based methods as well as develop new methods that are better suited for the analysis of nonnormal data. It is thus critical to understand whether practical data satisfy the normality assumption and if not, how severe the nonnormality is, what type of nonnormality it is, what the consequences are, and what can be done about it.

To understand normality or nonnormality, we need to first define a measure of it. Micceri (1989) evaluated deviations from normality based on arbitrary cut-offs of various measures of nonnormality, including asymmetry, tail weight, outliers, and modality. He found that all 440 large-sample achievement and psychometric measures distributions were nonnormal, 90% of which had sample sizes larger than 450. More recently, Blanca et al. (2013) evaluated nonnormality using the skewness and kurtosis[1] of 693 small samples, with sample size ranging from 10 to 30. The study includes many psychological variables, and the authors found that 94.5% of distributions were outside the range of [-0.25, 0.25] on either skewness or kurtosis and therefore violated the normality assumption. However, neither Micceri nor Blanca et al. discuss the distribution of skewness or kurtosis, how to test violations of normality, or how much effect

[1]Without specific mention, the skewness and kurtosis refer to the sample skewness and kurtosis throughout the paper.

they can have on the typically used methods such as *t*-test and factor analysis.

Scheffe (1959, p.333) has commented that kurtosis and skewness are “the most important indicators of the extent to which nonnormality affects the usual inferences made in the analysis of variance.” Skewness and kurtosis are also an intuitive means to understand normality. If skewness is different from 0, the distribution deviates from symmetry. If kurtosis is different from 0, the distribution deviates from normality in tail mass and shoulder for univariate data (DeCarlo, 1997b).[2]

In practice, normality measures such as skewness and kurtosis are rarely reported. In order to study nonnormality, we have contacted and obtained responses from 124 researchers, among whom only three reported skewness and kurtosis in their papers. The under-report of normality measures can be due to several reasons. First, many researchers are still not aware of the prevalence and influence of nonnormality. Second, not every researcher is familiar with skewness and kurtosis or their interpretation. Third, extra work is needed to compute skewness and kurtosis than the commonly used summary statistics such as means and standard deviations.

This paper provides a simple and practical response to the continuing under-report of nonnormality measures in published literature by elucidating the problem of nonnormality and offering feasible recommendations. We begin with an easy-to-follow introduction to univariate and multivariate skewness and kurtosis, their calculations, and interpretations. We then report on a review we conducted assessing the prevalence and severity of skewness and kurtosis in recent psychology and education publications. We show the influence of skewness and kurtosis on commonly used statistical tests in our field using data of typical skewness, kurtosis, and sample size found in our review. We offer a tutorial on how to compute the skewness and kurtosis measures we report here through commonly used software including SAS, SPSS, R, and a Web application. Finally, we offer practical recommendations for our readers that they can follow in their own research, including a guideline on how to report sample statistics in empirical research

[2]Kurtosis measures can be centered at either 0 or 3, the former is usually referred to as “excess kurtosis”. This is because the normal distribution has a kurtosis of 3, and therefore an excess kurtosis of 0.

and some possible solutions for nonnormality.

Univariate and Multivariate Skewness and Kurtosis

Different formulations for skewness and kurtosis exist in the literature. Joanes & Gill (1998) summarize three common formulations for univariate skewness and kurtosis that they refer to as g_1 and g_2, G_1 and G_2, and b_1 and b_2. The R package moments (Komsta & Novomestky, 2015), SAS proc means with vardef=n, Mplus, and STATA report g_1 and g_2. Excel, SPSS, SAS proc means with vardef=df, and SAS proc univariate report G_1 and G_2. Minitab reports b_1 and b_2, and the R package e1071 (Meyer et al., 2015) can report all three. There are also several measures of multivariate skewness and kurtosis, though Mardia's measures (Mardia, 1970) are by far the most common. These are currently only available in STATA, or as add-on macros multnorm in SAS or mardia in SPSS (DeCarlo, 1997a).

Univariate Skewness and Kurtosis

For the univariate case, we adopt Fisher's skewness (G_1) and kurtosis (G_2). Specifically, the skewness, G_1, is calculated as

$$G_1 = \frac{\sqrt{n(n-1)}}{n-2} \cdot \frac{m_3}{m_2^{3/2}}, \tag{1}$$

and the kurtosis, G_2, as

$$G_2 = \frac{n-1}{(n-2)(n-3)} \cdot \left[(n+1)\left(\frac{m_4}{m_2^2} - 3\right) + 6\right], \tag{2}$$

where $m_r = \sum_{i=1}^{n}(x_i - \bar{x})^r/n$ is the rth central moment with $\bar{x}$ being the sample mean and n the sample size. The sample skewness G_1 can take any value between negative infinity and positive infinity. For a symmetric distribution such as a normal distribution, the expectation of skewness is 0. A non-zero skewness indicates that a distribution "leans" one way or the other and has an asymmetric tail. Distributions with positive skewness have a longer right tail in the positive direction, and those with negative skewness have a longer left tail in the negative direction.

Figure 1 portrays three distributions with different values of skewness. The one in the middle is a normal distribution and its skewness is 0. The one on the left is a lognormal distribution with a positive skewness = 1.41. A commonly used example of a distribution with a long positive tail is the distribution of income where most households make around $53,000 a year[3] and fewer and fewer make more. In psychology, typical response time data often show positive skewness because much longer response time is less common (Palmer et al., 2011). The distribution on the right in Figure 1 is a skew-normal distribution with a negative skewness = -0.3. For example, high school GPA of students who apply for colleges often shows such a distribution because students with lower GPA are less likely to seek a college degree. In psychological research, scores on easy cognitive tasks tend to be negatively skewed because the majority of participants can complete most tasks successfully (Wang et al., 2008).

Kurtosis is associated with the tail, shoulder and peakedness of a distribution. Generally, kurtosis increases with peakedness and decreases with flatness. However, as DeCarlo (1997b) explains, it has as much to do with the shoulder and tails of a distribution as it does with the peakedness. This is because peakedness can be masked by variance. Figures 2a and 2b illustrate this relationship clearly. Figure 2a shows the densities of three normal distributions each with kurtosis of 0 but different variances, and Figure 2b shows three distributions with different kurtosis but the same variance. Normal distributions with low variance have high peaks and light tails as in Figure 2a, while distributions with high kurtosis have high peaks and heavy tails as in Figure 2b. Hence, peakedness alone is not indicative of kurtosis, but rather it is the overall shape that is important. Skewness cannot increase without kurtosis also increasing because of the relationship: kurtosis $\geq$ skewness$^2 - 2$ (Shohat, 1929).

Kurtosis has a range of $[-2(n-1)/(n-3), \infty)$ in a sample of size *n* and a range of [-2,∞] in the population.[4] The expectation of kurtosis of a normal distribution is 0. If a distribution is leptokurtic, meaning it has positive kurtosis, the distribution has a fatter tail than the normal distribution with the same variance. Generally speaking, if a data set is contaminated or contains

[3]The inflation adjusted medium household income is $53,657 in 2014 based on census.

[4]Note that if $g_2 = m_4/m_2^2 - 3$ is used to estimate kurtosis it also has a minimum value of -2.

extreme values, its kurtosis is positive. If a distribution is platykurtic, meaning it has negative kurtosis, the distribution has a relatively flat shoulder and short tails (e.g., see Figure 2b). For example, the distribution of age of the US population has negative kurtosis because there are generally the same number of people at each age.

Because for a normal distribution both skewness and kurtosis are equal to 0 in the population, we can conduct hypothesis testing to evaluate whether a given sample deviates from a normal population. Specifically, the hypothesis testing can be conducted in the following way.[5] We first calculate the standard errors of skewness (SES) and kurtosis (SEK) under the normality assumption (Bliss, 1967, p.144-145),

$$SES = \sqrt{\frac{6n(n-1)}{(n-2)(n+1)(n+3)}}, \quad (3)$$

$$SEK = 2(SES)\sqrt{\frac{n^2-1}{(n-3)(n+5)}}. \quad (4)$$

Note that the standard errors are functions of sample size. In particular, standard error decreases as sample size increases, and the strictness with which we call a distribution "normal" becomes more and more rigid. This is a natural consequence of statistical inference. With these standard errors, two statistics,

$$Z_{G1} = G_1/SES$$

and

$$Z_{G2} = G_2/SEK,$$

can be formed for skewness and kurtosis, respectively. Both of these statistics can be compared against the standard normal distribution, $N(0,1)$, to obtain a p-value to test a distribution's departure from normality (Bliss, 1967). If there is a significant departure, the p-value is smaller

[5]Other hypothesis testing methods available for skewness and kurtosis are available (Anscombe & Glynn, 1983; D'Agostino, 1970). The reason for adopting the method discussed here is that the standard errors of skewness and kurtosis are reported in popular statistical software such as SPSS and SAS, and, therefore, it is a feasible method for evaluating skewness and kurtosis through existing software.

than .05 and we can infer that the underlying population is nonnormal. If neither test is significant, there is not enough evidence to reject normality based on skewness or kurtosis although it may still be nonnormal in other characteristics.

Multivariate Skewness and Kurtosis

The univariate skewness and kurtosis have been extended to multivariate data. Multivariate skewness and kurtosis measure the same shape characteristics as in the univariate case. However, instead of making the comparison of the distribution of one variable against a univariate normal distribution, they are comparing the joint distribution of several variables against a multivariate normal distribution.

In this study, we use Mardia's measures (Mardia, 1970) of multivariate skewness and kurtosis, because they are most often included in software packages. Mardia defined multivariate skewness and kurtosis, respectively, as

$$b_{1,p} = \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} \left[(\mathbf{x}_i - \bar{\mathbf{x}})' \mathbf{S}^{-1} (\mathbf{x}_j - \bar{\mathbf{x}}) \right]^3, \tag{5}$$

$$b_{2,p} = \frac{1}{n} \sum_{i=1}^{n} \left[(\mathbf{x}_i - \bar{\mathbf{x}})' \mathbf{S}^{-1} (\mathbf{x}_i - \bar{\mathbf{x}}) \right]^2, \tag{6}$$

where $\mathbf{x}$ is a $p \times 1$ vector of random variables and $\mathbf{S}$ is the biased sample covariance matrix of $\mathbf{x}$ defined as

$$\mathbf{S} = \frac{1}{n} \sum_{i=1}^{n} \left[(\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})' \right]. \tag{7}$$

Both measures have a p subscript, so they are specific to a set of p variables. The expected Mardia's skewness is 0 for a multivariate normal distribution and higher values indicate a more severe departure from normality. The expected Mardia's kurtosis is $p(p+2)$ for a multivarite normal distribution of p variables. As in the univariate case, values under this expectation indicate platykurtism and higher values indicate leptokurtism.

Standardized measures can be formed for Mardia's skewness and kurtosis using the following formulations:

$$z_{1,p} = \frac{n}{6} b_{1,p}, \tag{8}$$

and

$$z_{2,p} = \frac{b_{2,p} - [p(p+2)(n-1)]/(n+1)}{\sqrt{[8p(p+2)]/n}}. \tag{9}$$

Standardized multivariate skewness $z_{1,p}$ can be compared against the chi-squared distribution $\chi^2_{p(p+1)(p+2)/6}$, and standardized multivariate kurtosis $z_{2,p}$ can be compared against the standard normal distribution $N(0,1)$. If the test statistic $z_{1,p}$ is significant, e.g. the p-value is smaller than .05, the joint distribution of the set of p variables has significant skewness; if the test statistic $z_{2,p}$ is significant, the joint distribution has significant kurtosis. If at least one of these tests is significant, it is inferred that the underlying joint population is nonnormal. As in the univariate case, non-significance does not necessarily imply normality.

Review of Skewness and Kurtosis in Practical Data

Although Micceri (1989) and Blanca et al. (2013) have studied univariate nonnormality, we are not aware of any study that has investigated multivariate skewness and kurtosis with empirical data or has tested the significance of nonnormality. Therefore, we conducted a study to further evaluate the severity of nonnormality in our field, especially in the multivariate case. Focusing on published research, we contacted 339 researchers with publications that appeared in Psychological Science from January 2013 to June 2014 and 164 more researchers with publications that appeared in the American Education Research Journal from January 2010 to June 2014. The two journals were chosen due to their prestige in their corresponding fields. We asked the researchers to provide the univariate and multivariate skewness and kurtosis of continuous variables used in their papers. Binary, categorical, and nominal variables were excluded, though likert items were included because they are often treated as normal in the literature. To help the researchers compute the skewness and kurtosis, we provided a tutorial for different software as we will present later in this paper. Our data collection ended in November, 2014, by which point we had obtained 1,567 univariate measures and 254 multivariate measures of skewness and kurtosis from 194 studies. Some authors submitted univariate results without multivariate results so not all 1,567 univariate measures are included as part of a multivariate measure. The median sample size for these studies was 106, and the sample size ranged from 10 to 200,000. The median number of variables included in a multivariate measure was 3, and ranged from 1 to 36. Since researchers had the option to submit skewness and kurtosis anonymously, it is unclear how many authors responded to our request or what their study characteristics may be.

Univariate Skewness and Kurtosis

As shown in Table 1a, univariate skewness ranged from -10.87 to 25.54 and univariate kurtosis from -2.20 to 1,093.48, far wider than previously reported or tested. Because these most extreme values may be outliers, we also report 1st through 99th percentiles of univariate skewness and kurtosis. Percentiles can be interpreted as the percent of samples with lower skewness or kurtosis than that value. There is clearly a large range from the 1st to the 99th percentile, especially for kurtosis. The correlation between sample size and skewness is $r = -0.005$, and with kurtosis is $r = 0.025$. These are comparable to what Blanca et al. (2013) have reported in which correlations between sample size and skewness and kurtosis were .03 and -.02, respectively. The results in Table 1a include skewness and kurtosis when the sample size is smaller and larger than 106, the median sample size of all collected data. As shown in this table, negative skewness and kurtosis are much more common than previously reported: 38% of distributions have negative skewness and 47% have negative kurtosis. This could be due to the number of likert measures provided, but because of the anonymous submission option there is no way to confirm this. Means and sample size-weighted means are also provided in Table 1. Sample size-weighted means are helpful because we expect sample measures to better-reflect that of the population as sample size increases. Therefore, measures from large samples are given higher weight than those from smaller samples. The mean univariate skewness is 0.51, and the sample size-weighted mean is 0.47. The mean univariate kurtosis is 4.29, and the sample size-weighted mean is 8.41. Therefore, on average, the skewness and kurtosis are larger than that of a normal distribution. To further visualize what these distributions look like, Figure 3 shows histograms of 20 randomly selected distributions from our review. Note that there is no common shape that explains skewness or kurtosis.

Percentages of univariate distributions with significant skewness or kurtosis by sample size are presented in Table 1b. About 66% of univariate distributions had significant skewness and 54% had significant kurtosis. Almost 74% of distributions had either significant skewness or kurtosis and were therefore classified as nonnormal. As expected, it becomes easier for tests to become significant with larger sample sizes. Over 95% of distributions with sample sizes greater than the median sample size, 106, were tested as nonnormal. Conversely, when the sample size was less than 106 only 56% of distributions were significantly nonnormal.

Multivariate Skewness and Kurtosis

The 254 collected Mardia's multivariate skewness ranged from 0 to 1,332 and multivariate kurtosis from 1.80 to 1,476. Percentiles of Mardia's skewness and kurtosis split by median sample size and median number of variables used in their calculation are presented in Table 2. The correlation between sample size and Mardia's skewness is $r = -0.01$ and with Mardia's kurtosis is $r = 0.02$. The correlation between the number of variables and Mardia's skewness is $r = 0.58$ and with Mardia's kurtosis is $r = 0.73$. After centering Mardia's kurtosis on $p(p + 2)$, the expected value under normality, the correlation between kurtosis and the number of variables becomes $r = 0.05$. The mean multivariate skewness is 32.94, and the sample size-weighted mean is 28.26. The mean multivariate kurtosis is 78.70, and the sample size-weighted mean is 92.03. Therefore, the average skewness and kurtosis are greater than that of a multivariate normal distribution. This has important ramifications especially for SEM, for which multiple outcome measures are often used and for which multivariate kurtosis can asymptotically affect standard errors.

Percentages of multivariate distributions with significant Mardia's skewness and kurtosis are presented in Table 3. About 58% of multivariate skewness measures and 57% of multivariate kurtosis measures reached significance. Combining these, 68% of multivariate distributions were significantly nonnormal. In particular, 94% of Mardia's measures were tested significant when the sample size was larger than 106. Similarly, more Mardia's measures became significant with more variables.

To summarize, based on the test of 1,567 univariate and 254 multivariate skewness and kurtosis from real data, we conclude that 74% of univariate data and 68% of multivariate data significantly deviated from a univariate or multivariate normal distribution. In examining only those univariate measures included in a multivariate measure, 68% have significant nonnormality. Therefore, nonnormality is a severe problem in real data, though multivariate nonnormality does not appear to be a severe problem above and beyond that of univariate normality. However, this relationship requires further study to evaluate.

Influences of Skewness and Kurtosis

In order to clearly show the influence of skewness and kurtosis, we conducted simulations one on the one-sample *t*-test, simple regression, one-way ANOVA, and confirmatory factor analysis (CFA). Simulation studies are helpful because we know what results the statistical tests should show, and so we can evaluate how nonnormality affects those results. Note that for all of these models, the normality of the dependent variable is what is of interest. There are no normality assumptions put on the independent variable.

Influence of Univariate Skewness and Kurtosis

Yuan et al. (2005) show that the properties of mean estimates are not affected by either skewness or kurtosis *asymptotically*, but that the standard error of sample variance is a function of kurtosis. If normality is assumed (kurtosis = 0), the standard error of variance will be underestimated when kurtosis is positive and overestimated when kurtosis is negative. In other words, kurtosis will still have an effect on variance estimates at very large sample sizes while mean estimates are only affected in small samples. For example, Yanagihara & Yuan (2005) found that the expectation and variance of the *t*-statistic depends on skewness, but that the effect lessens as sample size increases.

To concretely demonstrate the influence of univariate skewness and kurtosis, we conducted a simulation study on a one-sample *t*-test. In the simulation, we set the skewness to the 1st, 5th, 25th, 50th, 75th, 95th, and 99th percentiles of univariate skewness found in our review of practical data. These were tested in sample sizes of the 5th, 25th, 50th, 75th, and 95th percentiles of sample size found in our review. Therefore, these conditions should represent typical results found in our field. Because kurtosis does not influence the *t*-test, it was kept at the 99th percentile, 95.75, throughout all conditions. In total, we considered 35 conditions for each test. Under each condition, we generated 10,000 sets of data with mean 0, variance 1, and the specified skewness and kurtosis from a Pearson distribution in R (R Core Team, 2016) using the package PearsonDS (Becker & Klöß ner, 2016).[6] Then, we obtained the empirical type I error rate to reject the null hypothesis that the population mean is equal to 0 using the significance level 0.05 in a two-tailed, a lower-tail, and an upper-tail one-sample t-test.

Table 4 displays the empirical type I error rate for each condition. For brevity, type I error rates of just the lowest sample size are presented for conditions with skewness between -1.17 and 0.94 because these conditions did not present any problems. To better understand the empirical type I error rate, we bold those that are outside of the range [0.025, 0.075]. When the skewness and kurtosis are 0, the generated data are from a normal distribution and the empirical type I error rate is close to 0.05 even when the sample size is as small as 18 for all three tests. When data deviate from normality, the results show that a two-sided test is more robust than a one-sided test. The two-sided test only has increased type 1 error rate for a skewness of 6.32, for which a sample size of 554 is necessary to dissipate the effect. A lower tail t-test has even higher type 1 error rates at this skewness, and an upper tail t-test has an increased type 1 error rate with negative skewness and very low rates with high positive skewness.

[6]Pearson distribution includes a class of distributions. It is used here because it allows us to keep the mean and variance fixed but at the same time change the skewness and kurtosis.

A simple regression and a one-way ANOVA with three groups were also tested at all of these conditions. The regression was robust to all conditions, even at the lowest sample size. Type 1 errors in the ANOVA were also robust to all conditions examined here, though it is known that power can suffer when the population is platykurtic (Glass et al., 1972).

Influence of Multivariate Skewness and Kurtosis

In order to show the influence of multivariate skewness and kurtosis, we conducted simulation studies on CFAs. First, we focus on a one-factor model with four manifest variables. For each manifest variable, the factor loading is fixed at 0.8 and the uniqueness factor variance is 0.36. The variance of the factor is set to 1. Note that when kurtosis = 24 data are from a multivariate normal distribution and so the centered kurtosis is 0. Although in our review of practical data about half of the data sets had centered Mardia's kurtosis less than 0, 21 is the only multivariate kurtosis less than 24 we were able to successfully simulate. Hence, we used these two values of Mardia's kurtosis (21 and 24) along with the 75th, 95th, and 99th percentiles of Mardia's kurtosis found in our review of practical data of four manifest variables (30, 60, and 100). The same sample sizes from our review were used as in the previous simulation, with the exception of 18. A sample size of 18 was excluded because it is not a sufficient sample size for this analysis. Because skewness does not influence SEM, it was kept at 0 throughout all conditions. In total, 20 conditions were considered. 1,000 data sets were used to evaluate each condition.The authors are currently unaware of any method to simulate data with a particular multivariate skewness or kurtosis, so instead we used the R package lavaan (Rosseel, 2012) to simulate data from a model with certain univariate skewness and kurtosis. Appropriate univariate values were found to simulate multivariate values of a population by trial and error.

The influence of skewness and kurtosis is evaluated through the empirical type I error rate of rejecting the factor model using the normal-distribution-based chi-squared goodness-of-fit test. This test is significant when the model does not fit the data. Because the true one-factor model was fit to the simulated data, one would expect the empirical type I error rate to be close to the nominal level 0.05. Deviation from it indicates the influence of skewness and kurtosis. The empirical type I error rates at different levels of Mardia's kurtosis are summarized in Table 5.

The results show that when the data are from a multivariate normal distribution (kurtosis = 24), the empirical type I error rates were close to the nominal level 0.05. However, when the data deviate from a multivariate normal distribution to a Mardia's kurtosis of 60, the empirical type I error rates are all greater than 0.05. Unsurprisingly, the problem becomes worse with an increase in sample size. For example, when the multivariate kurtosis is 100 and the sample size is 1489, the normal-distribution-based chi-squared test rejects the correct one-factor model 29.8% of the time.

Type 1 error rates were also compared in a one-factor model with eight manifest variables and a two-factor model with four manifest variables each to investigate the effects of an increase in the number of manifest variables or number of factors. Factor loadings were adjusted to maintain uniqueness factor variance at 0.36 and total variance at 1. The same conditions were tested as in the simulation study above, with the exception of those with a sample size of 48. This sample size is not sufficient for an analysis of eight manifest variables. The same univariate kurtoses were used to simulate the data, though they result in different multivariate measure for eight variables than they do for four. The resulting empirical type I error rates of these multivariate kurtoses for both of these models can be found in Table 6.

Once again, type I error is maintained when the distribution is multivariate normal (kurtosis = 80), but once kurtosis reaches 150 all type I errors are above 0.05. As sample size increases, the problem worsense. In comparison to the results shown in Table 5, type I errors are worse with an increase in the number of manifest variables. However, holding the number of manifest variables constant, an increase in the number of factors lowers type I error rate.

In summary, if either univariate or multivariate nonnormal data are analyzed using normal-distribution-based methods, it will lead to incorrect statistical inference. Given the prevalence of nonnormality as we have shown in the previous section, it is very important to quantify the nonnormality. We suggest using skewness and kurtosis to measure nonnormality and we will show how to obtain both univariate and multivariate skewness and kurtosis in the next section.

Computing Univariate and Multivariate Skewness and Kurtosis

In this section, we illustrate how to compute univariate and multivariate skewness and kurtosis in popular statistical software including SAS, SPSS, and R as well as a newly developed Web application. As previously mentioned, different softwares produce different types of univariate skewness and kurtosis. Furthermore, most don't report tests or multivariate measures. Using our software and macros for SAS, SPSS, and R produces consistent and full results across software. Some software requires macros that can be downloaded from our website at `http://psychstat.org/nonnormal`. Our Web application can be found at `https://webpower.psychstat.org/models/kurtosis`. As an example, we use a subset of data from the Early Childhood Longitudinal Study, Kindergarten Class of 1998-99 (ECLS-K) to show the use of different software. The ECLS-K is a longitudinal study with data collected in kindergarten in the fall and spring of 1998-99, in 1st grade in the fall and spring of 1999-2000, in 3rd grade in the spring of 2002, in 5th grade in the spring of 2004, and in 8th grade in the spring of 2007. The data used here consist of four consecutive mathematical ability measures of 563 children from kindergarten to 1st grade. To simplify our discussion, we assume that all files to be used are in the folder of "C:\nonnormal", which needs to be changed accordingly.

SAS

To use SAS for computing the univariate and multivariate skewness and kurtosis, first download the mardia.sas macro file from our website. Our macro was modified from a SAS macro MULTNORM provided by the SAS company. After saving the sas macro file, the following code can be used to get the skewness and kurtosis for the ECLS-K data.[7]

SAS input

[7] The number on the right is used to identify the code only and is not part of the SAS code.

```
1 DATA eclsk;
2   INFILE "eclsk563.txt";
3   INPUT y1 y2 y3 y4;
4 RUN;
5 %INCLUDE "mardia.sas";
6 %mardia(data=eclsk, var=y1 y2 y3 y4)
```

In the SAS input, Line 1 through Line 4 read the ECLS-K data in the file "`eclsk563.txt`" into SAS. Line 5 includes the SAS macro file downloaded from our website for use within SAS. The sixth line uses the function `mardia` in the macro to calculate skewness and kurtosis. The argument "`data=`" specifies the SAS database to use and "`var=`" specifies the variables to use in calculating the skewness and kurtosis.

The SAS output from the analysis of the ECLS-K data is given below. The first part of the output, from Line 1 to Line 8, displays the univariate skewness and kurtosis as well as their corresponding standard error. For example, the skewness for the ECLS-K data at time 1 is 0.69 with a standard error 0.10 (Line 5). Based on a *z*-test, one would conclude that the skewness is significantly large than 0. For another example, the kurtosis for the data at time 4 is 1.29 with a standard error 0.21 (Line 8), indicating the kurtosis is significantly larger than 0.

The second part of the output, from Line 10 to Line 23 includes the information on multivariate skewness and kurtosis. First, the multivariate skewness is 2.26 (Line 16) with a standardized measure of 212.24 (Line 17). The p-value for a chi-squared test is approximately 0 (Line 18). Therefore, the multivariate skewness is significantly larger than 0. Second, the multivariate kurtosis is 25.47 (Line 21) with the standardized measure of 2.51 (Line 22). The *p*-value for a *z*-test is approximately 0.01 (Line 23). Therefore, the multivariate kurtosis is significantly different from that of a multivariate normal distribution with 4 variables (24). Consequently, the data do not follow a multivariate normal distribution and therefore violate the normality assumption if used in multivariate analysis.

SAS output

```
### Univariate Skewness and Kurtosis ###

    Skewness   SE_skew  Kurtosis   SE_kurt

y1 0.6932137 0.1029601  0.229546 0.2055599
y2 0.0368512 0.1029601  -0.41783 0.2055599
y3 -0.225271 0.1029601 -0.252103 0.2055599
y4 -1.000066 0.1029601 1.2898344 0.2055599

### Mardia's multivariate skewness and kurtosis ###

Sample size =        563
Number of variables =          4

Multivariate skewness
b1p =  2.2618775
z1 =  212.23951
p-value =          0

Multivariate kurtosis
b2p =  25.468192
z2 =   2.514123
p-value =  0.0119329
```

SPSS

DeCarlo (1997b) has developed an SPSS macro to calculate multivariate skewness and kurtosis.[8] We slightly modified the macro to make the output of univariate skewness and kurtosis consistent to other software. To use the SPSS macro, first download the macro file mardia.sps to your computer from our website. Then, open a script editor (File->New->Syntax) within SPSS and include the following SPSS script.

The code on the first eight lines in the input is used to read the ECLS-K data into SPSS. These lines are not necessary if your data are already imported into SPSS. Line 10 gets the SPSS macro into SPSS for use. The function `mardia` calculates univariate and multivariate skewness and kurtosis for the variables specified by the `vars` option on Line 11. Note that the folder to the data file and the SPSS macro file needs to be modified to reflect the actual location of them.

SPSS input

```
1  get data
2   /type = txt
3   /file = "C:\nonnormal\eclsk563.txt"
4   /delimiters = " "
5   /firstcase = 1
6   /variables =  y1 f2.0 y2 f2.0 y3 f2.0 y4 f2.0.
7
8  execute.
9
10 INCLUDE file="C:\nonnormal\mardia.sps".
11 mardia vars=y1 y2 y3 y4 /.
12 execute.
```

The SPSS output from the analysis of the ECLS-K data is given below. Similar to the SAS

[8]The original macro can be downloaded at `http://www.columbia.edu/~ld208/Mardia.sps`.

output, the first part of the output includes univariate skewness and kurtosis and the second part is for the multivariate skewness and kurtosis. SPSS obtained the same skewness and kurtosis as SAS because the same definition for skewness and kurtosis was used.

SPSS output

```
Sample size:
   563

Number of variables:
     4

Univariate Skewness
         y1         y2         y3         y4     SE_skew
      .6932      .0369     -.2253    -1.0001       .1030

Univariate Kurtosis
         y1         y2         y3         y4     SE_kurt
      .2295     -.4178     -.2521     1.2898       .2056

Mardia's multivariate skewness
        b1p         z1    p-value
     2.2619   212.2395      .0000

Mardia's multivariate kurtosis
        b2p         z2    p-value
    25.4682     2.5141      .0119
```

R

To use R, first download the R code file mardia.r to your computer from our website. Then, in the editor of R, type the following code. The code on Line 1 gets the ECLS-K data into R and Line 2 provides names for the variables in the data. The third line loads the R function `mardia` into R. Finally, the last line uses the function `mardia` to carry out the analysis on Line 4.

R input

```
1 eclsk <- read.table('eclsk563.txt')
2 names(eclsk)<-c("y1", "y2", "y3", "y4")
3 source("mardia.r")
4 mardia(eclsk)
```

The output from the R analysis is presented below. Clearly, it obtains the same univariate and multivariate skewness and kurtosis as SAS and SPSS.

R output

```
Sample size:  563
Number of variables:  4

Univariate skewness and kurtosis
      Skewness   SE_skew   Kurtosis   SE_kurt
y1  0.69321372 0.1029601  0.2295460 0.2055599
y2  0.03685117 0.1029601 -0.4178298 0.2055599
y3 -0.22527112 0.1029601 -0.2521029 0.2055599
y4 -1.00006618 0.1029601  1.2898344 0.2055599

Mardia's multivariate skewness and kurtosis
                 b          z    p-value
Skewness  2.261878 212.239506 0.00000000
```

```
Kurtosis 25.468192   2.514123 0.01193288
```

Web Application for Skewness and Kurtosis

To further ease the calculation of univariate and multivariate skewness and kurtosis, we also developed a Web application that can work within a Web browser and does not require knowledge of any specific software. The Web application utilizes the R function discussed in the previous section to obtain skewness and kurtosis on a Web server and produces the same results as SAS, SPSS, and R.

To access the Web application, type the URL `http://psychstat.org/kurtosis` in a Web browser and a user will see an interface as shown in Figure 4. To use the Web application, the following information needs to be provided on the interface.

Data. The data file can be chosen by clicking the "Choose File" button[9] and locating the data set of interest on the local computer.

Type of Data. The Web application allows commonly used data types such as SPSS, SAS, Excel, and text data. To distinguish the data used, it recognizes the extension names of the data file. For example, a SPSS data file ends with the extension name `.sav`, a SAS data file with the extension name `.sas7bdat`, and an Excel data file with the extension name `.xls` or `.xlsx`. In addition, a CSV file (comma separated value data file) with the extension name `.csv` and a TXT file (text file) with the extension name `.txt` can also be used. If a `.csv` or `.txt` file is used, the user needs to specify whether variable names are included in the file. For Excel data, it requires the first row of the data file to be the variable names.

Select Variables to Be Used. Skewness and kurtosis can be calculated on either all the variables or a subset of variables in the data. To use all the variables, leave this field blank. To select a subset of variables, provide the column numbers separated by comma "`,`". Consecutive

[9]Note that different operating systems and/or browsers might show the button differently. For example, for Internet Explorer, the button reads "Browse...".

variables can be specified using "–". For example, `1, 2-5, 7-9, 11` will select variables 1, 2, 3, 4, 5, 7, 8, 9, 11.

Missing Data. Missing data are allowed in the data although they will be removed before the calculation of skewness and kurtosis. This field should be left blank if the data file has no missing values. If multiple values are used to denote missing data, they can be specified all together separated by a comma (,). For example, `-999, -888, NA` will specify all three values as missing data.

After providing the required information, clicking the "Calculate" button will start the calculation of skewness and kurtosis. The output of the analysis is provided below. The output is identical to the R output except for the variable names for univariate skewness and kurtosis. This is because by default the variable names are constructed using "V" and an integer in R.

Web application output

```
Sample size:  563
Number of variables:  4

Univariate skewness and kurtosis
      Skewness   SE_skew   Kurtosis   SE_kurt
V1  0.69321372 0.1029601  0.2295460 0.2055599
V2  0.03685117 0.1029601 -0.4178298 0.2055599
V3 -0.22527112 0.1029601 -0.2521029 0.2055599
V4 -1.00006618 0.1029601  1.2898344 0.2055599

Mardia's multivariate skewness and kurtosis
                 b          z    p-value
Skewness  2.261878 212.239506 0.00000000
Kurtosis 25.468192   2.514123 0.01193288
```

Discussion and Recommendations

The primary goals of this study were to assess the prevalence of nonnormality in recent psychology and education publications and its influence on statistical inference, as well as to provide a software tutorial on how to compute univariate and multivariate skewness and kurtosis. First, nonnormality clearly exists in real data. Based on the test of skewness and kurtosis of data from 1,567 univariate variables, we found that 74% of either skewness or kurtosis were significantly different from that of a normal distribution. Furthermore, 68% of 254 multivariate data sets had significant Mardia's multivariate skewness or kurtosis. Our results together with those of Micceri (1989) and Blanca et al. (2013) strongly suggest the prevalence of nonnormality in real data.

Our investigation on the influence of skewness and kurtosis involved simulation studies on the one-sample *t*-test and factor analysis. Through simulation, we concretely showed that nonnormality, as measured by skewness and kurtosis, exerted great influence on statistical tests that bear the normality assumption. For example, the use of the *t*-test incorrectly rejected a null hypothesis 17% of the time and the chi-squared test incorrectly rejected a correct factor model 30% of the time under some conditions. Therefore, nonnormality can cause severe problems. For example, a significant result might be simply an artificial effect caused by nonnormality.

Given the prevalence of nonnormality and its influence on statistical inference, it is critical to report statistics such as skewness and kurtosis to understand the violation of normality. In Table 7, we list the summary statistics that are critical to different statistical methods in empirical data analysis. For example, mean comparisons would be influenced by skewness while factor analysis is more influenced by kurtosis. To facilitate the report of univariate and multivariate skewness and kurtosis, we have provided SAS, SPSS, and R code as well as a Web application to compute them.

Once nonnormality has been identified as a problem, the main options for handling it in a statistical analysis include transformation, nonparametric methods, and robust analysis. Transforming data so that it becomes normal is an easy option, because after transformation the researcher can proceed with whichever normality-based method they desire. In psychology, log transformation is a common way to get rid of positive skewness, for example. The Box-Cox transformation method (Box & Cox, 1964) is also very popular because it's easy to use and can accomodate many types of nonnormality. However, it has been suggested that Box-Cox and other transformations seldom maintain linearity, normality, and homoscedasticity simultaneously (Sakia, 1992, for example), and even if transformation is successful the resulting parameter estimates often have little substantive meaning.

Corder & Foreman (2014) offer an easy-to-follow review of nonparametric techniques, including the Mann–Whitney U-test, Kruskal-Wallis test, and Spearman rank order correlation, among others. The basic premise of most of these methods is to perform analysis on ranks rather than the raw data. This is, of course, a more robust procedure than assuming normality of raw data, but can be less powerful in some circumstances and the results can be less meaningful. However, for data that is already ordinal or ranked these methods are certainly the best option, and can still be a good option in other circumstances, as well.

Robust analysis is often the best method, though historically it has also been the most difficult to conduct. Robust analysis generally addresses three points of concern: parameter estimates, standard errors of those estimates, and test statistics. Within the context of SEM the three most common methods with the best performance in dealing with each of these issues are robust estimation using Huber-type weights (Huber, 1967), sandwich-type standard errors, and the Satorra-Bentler scaled chi-squared statistic (Satorra & Bentler, 1988), respectively. See Fouladi (2000) for a review of other adjusted test statistics and Yuan & Schuster (2013) for a review of other estimation procedures. We focus on SEM at this time because those models are asymptotically affected by nonnormality, and so provide the largest opportunity for improvement.

Recently, some software packages have begun to include these procedures, making robust analysis a much easier option than it has ever been before. Table 8 shows which software packages include which robust procedures. Currently, EQS (Multivariate Software, Inc.), WebSEM (Zhang & Yuan, 2012), and the R package rsem (Yuan & Zhang, 2012) are the only softwares to offer all three of the aforementioned methods, and WebSEM and rsem offer them for free. Additionally, WebSEM has a user-friendly interface in which researchers can draw the path diagram they wish to fit.

As shown in Figure 3, there is no common distribution of practical data in psychology and education. With such diversity in data shapes and research goals, it is impossible to create one universal solution. However, we hope that through this paper we were able to elucidate the problem through our review of practical data and simulation and offer some feasible recommendations to researchers in our field. It is our hope that researchers begin to take nonnormality seriously and start to report them along with means and variances that have already been established in data analysis. We believe that reporting skewness and kurtosis in conjunction with moving toward robust SEM analysis offer two high-impact changes that can be made in the literature at this time. These actions will not only increase the transparency of data analysis but also encourage quantitative methodologists to develop better techniques to deal with nonnormality, improve statistical practices and conclusions in empirical analysis, and increase awareness and knowledge of the nonormality problem for all researchers in our field.

References

Anscombe, F. J., & Glynn, W. J. (1983, April). Distribution of the Kurtosis Statistic b2 for Normal Samples. *Biometrika*, *70*(1), 227. doi: 10.2307/2335960

Becker, M., & Klöß ner, S. (2016). *PearsonDS: Pearson Distribution System.* Retrieved from `https://CRAN.R-project.org/package=PearsonDS` (R package version 0.98)

Blanca, M. J., Arnau, J., L$\tilde{A}^3 pez - Montiel, D., Bono, R., \&Bendayan, R.$ (2013, January). Skewness and Kurtosis in Real Data Samples. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences*, *9*(2), 78–84. doi: 10.1027/1614-2241/a000057

Bliss, C. (1967). Statistical Tests of Skewness and Kurtosis. In *Statistics in Biology: Statistical Methods for Research in the Natural Sciences* (Vol. 1, pp. 140–146). New York: McGraw-Hill Book Company.

Box, G. E., & Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society. Series B (Methodological)*, 211–252. Retrieved 2016-07-26, from `http://www.jstor.org/stable/2984418`

Corder, G. W., & Foreman, D. I. (2014). *Nonparametric statistics: A step-by-step approach*. John Wiley & Sons.

D'Agostino, R. B. (1970, December). Transformation to Normality of the Null Distribution of g 1. *Biometrika*, *57*(3), 679. doi: 10.2307/2334794

DeCarlo, L. (1997a). *Mardia's multivariate skew (b1p) and multivariate kurtosis (b2p).* Retrieved from `http://www.columbia.edu/~ld208/Mardia.sps`

DeCarlo, L. (1997b). On the meaning and use of kurtosis. *Psychological methods*, *2*(3), 292. doi: 10.1037/1082-989X.2.3.292

Fouladi, R. T. (2000, July). Performance of Modified Test Statistics in Covariance and Correlation Structure Analysis Under Conditions of Multivariate Nonnormality. *Structural Equation Modeling: A Multidisciplinary Journal*, *7*(3), 356–410. Retrieved 2016-07-26, from `http://www.tandfonline.com/doi/abs/10.1207/S15328007SEM0703_2` doi: 10.1207/S15328007SEM0703_2

Glass, G. V., Peckham, P. D., & Sanders, J. R. (1972, January). Consequences of Failure to Meet Assumptions Underlying the Fixed Effects Analyses of Variance and Covariance. *Review of Educational Research*, *42*(3), 237–288. doi: 10.3102/00346543042003237

Huber, P. J. (1967). The behavior of maximum likelihood estimates under nonstandard conditions. In *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability* (Vol. 1, pp. 221–233).

Joanes, D., & Gill, C. (1998). Comparing measures of sample skewness and kurtosis. *Journal of the Royal Statistical Society: Series D (The Statistician)*, *47*(1), 183–189. doi: 10.1111/1467-9884.00122

Komsta, L., & Novomestky, F. (2015). *moments: Moments, cumulants, skewness, kurtosis and related tests*. Retrieved from `https://CRAN.R-project.org/package=moments` (R package version 0.14)

Mardia, K. V. (1970, December). Measures of Multivariate Skewness and Kurtosis with Applications. *Biometrika*, *57*(3), 519. doi: 10.2307/2334770

Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A., & Leisch, F. (2015). *e1071: Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071), TU Wien*. Retrieved from `https://CRAN.R-project.org/package=e1071` (R package version 1.6-7)

Micceri, T. (1989). The unicorn, the normal curve, and other improbable creatures. *Psychological Bulletin*, *105*(1), 156. doi: 10.1037/0033-2909.105.1.156

Palmer, E. M., Horowitz, T. S., Torralba, A., & Wolfe, J. M. (2011). What are the shapes of response time distributions in visual search? *Journal of Experimental Psychology: Human Perception and Performance*, *37*(1), 58–71. Retrieved 2016-07-26, from `http://doi.apa.org/getdoi.cfm?doi=10.1037/a0020747` doi: 10.1037/a0020747

R Core Team. (2016). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from `https://www.R-project.org/`

Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software*, *48*(2), 1–36. Retrieved from `http://www.jstatsoft.org/v48/i02/`

Sakia, R. M. (1992). The Box-Cox Transformation Technique: A Review. *The Statistician*, *41*(2), 169. Retrieved 2016-07-26, from `http://www.jstor.org/stable/10.2307/2348250?origin=crossref` doi: 10.2307/2348250

Satorra, A., & Bentler, P. (1988). Scaling corrections for statistics in covariance structure analysis (UCLA Statistics Series 2). *Los Angeles: University of California at Los Angeles, Department of Psychology*.

Scheffe, H. (1959). *The analysis of variance.* New York: Wiley.

Shohat, J. (1929, December). Inequalities for Moments of Frequency Functions and for Various Statistical Constants. *Biometrika*, *21*(1/4), 361. doi: 10.2307/2332566

Tabachnick, B. G., & Fidell, L. S. (2012). *Using Multivariate Statistics* (6th ed.). Pearson.

Wang, L., Zhang, Z., McArdle, J. J., & Salthouse, T. A. (2008, September). Investigating Ceiling Effects in Longitudinal Data Analysis. *Multivariate Behavioral Research*, *43*(3), 476–496.

Retrieved 2016-07-26, from http://www.tandfonline.com/doi/abs/10.1080/00273170802285941 doi: 10.1080/00273170802285941

Yanagihara, H., & Yuan, K.-H. (2005, November). Four improved statistics for contrasting means by correcting skewness and kurtosis. *British Journal of Mathematical and Statistical Psychology*, *58*(2), 209–237. doi: 10.1348/000711005X64060

Yuan, K.-H., Bentler, P. M., & Zhang, W. (2005, November). The Effect of Skewness and Kurtosis on Mean and Covariance Structure Analysis: The Univariate Case and Its Multivariate Implication. *Sociological Methods & Research*, *34*(2), 240–258. doi: 10.1177/0049124105280200

Yuan, K.-H., & Schuster, C. (2013). Overview of statistical estimation methods. *Oxford library of psychology. The Oxford handbook of quantitative methods*, *1*, 361–387.

Yuan, K.-H., & Zhang, Z. (2012). Robust structural equation modeling with missing data and auxiliary variables. *Psychometrika*, *77*(4), 803–826. doi: 10.1007/s11336-012-9282-4

Zhang, Z., & Yuan, K.-H. (2012). *WebSEM: Conducting structual equation modelling online.* Notre Dame, IN. Retrieved from https://websem.psychstat.org

Table 1

Univariate skewness and kurtosis

(a) *Skewness and kurtosis by sample size*

	$n \leq 106$		$n > 106$		Overall	
Percentile	Skewness	Kurtosis	Skewness	Kurtosis	Skewness	Kurtosis
Minimum	-4.35	-2.20	-10.87	-1.99	-10.87	-2.20
1st	-1.68	-1.79	-2.68	-1.56	-2.08	-1.70
5th	-1.10	-1.28	-1.27	-1.28	-1.17	-1.28
25th	-0.33	-0.60	-0.33	-0.52	-0.33	-0.57
Median	0.27	0.02	0.15	0.12	0.20	0.07
75th	0.91	1.35	1.00	2.12	0.94	1.62
95th	2.25	5.89	3.56	19.39	2.77	9.48
99th	4.90	30.47	10.81	154.60	6.32	95.75
Maximum	6.32	40.00	25.54	1,093.48	25.54	1,093.48

(b) *Percent of significant skewness and kurtosis by sample size*

	$n \leq 106$	$n > 106$	Overall
Skewness	51	82	66
Kurtosis	33	77	54
Either	56	95	74

Note. There were 805 distributions with $n \leq 106$ and 762 with $n > 106$. Nonnormality is defined by significant statistics Z_{G_1} or Z_{G_2}, $p < .05$.

Table 2

Mardia's measures by sample size and number of variables

(a) *Mardia's Skewness*

	By Sample Size		By # of Variables		Overall
Percentile	$n \leq 106$	$n > 106$	$p \leq 3$	$p > 3$	
Minimum	0.01	0.00	0.00	0.02	0.00
1st	0.03	0.00	0.00	0.43	0.00
5th	0.23	0.02	0.03	1.08	0.035
25th	1.15	0.35	0.33	5.72	0.76
Median	3.04	3.26	1.14	1.40	3.08
75th	13.91	14.92	2.95	44.43	14.32
95th	124.97	107.54	23.97	211.31	112.82
99th	635.90	496.77	343.60	786.84	610.66
Maximum	1,263.60	796.92	496.77	1,263.60	1,263.60

(b) *Mardia's Kurtosis.*

	By Sample Size				By # of Variables				Overall	
	$n \leq 106$		$n > 106$		$p \leq 3$		$p > 3$			
	$b_{2,p}$	$b^*_{2,p}$	$b_{2,p}$	$b^*_{2,p}$	$b_{2,p}$	$b^*_{2,p}$	$b_{2,p}$	$b^*_{2,p}$	$b_{2,p}$	$b^*_{2,p}$
Min	2.19	-90.50	1.99	-18.57	2.00	-7.77	15.09	-90.50	1.99	-90.50
1st	2.23	-61.02	2.34	-15.43	2.20	-7.72	18.90	-63.61	2.23	-54.55
5th	3.35	-23.59	2.79	-7.51	2.39	-3.74	22.26	-30.83	2.92	-17.01
25th	8.08	-2.33	8.81	0.26	7.02	-0.82	37.76	-2.38	8.26	-1.35
Median	14.24	-0.70	31.69	5.37	8.71	0.26	60.86	5.55	18.90	0.59
Mean	61.40	0.01	98.63	48.34	22.45	12.87	152.3	35.02	78.70	22.46
Mean*	72.11	2.17	92.31	50.36	16.49	9.36	272.5	146.1	92.03	49.71
75th	43.00	2.22	90.89	29.32	14.84	2.34	153.3	27.36	56.69	7.47
95th	190.1	28.18	419.4	179.25	52.69	44.54	614.4	119.3	323.1	98.17
99th	942.6	87.45	755.4	732.9	384	369	1,356	719.4	914.9	541
Max	1,476	108.1	1,392	1,368	556	541	1,476	1,368	1,476	1,368

Note. There were 136 multivariate distributions with $n \leq 106$, 118 with $n > 106$, 144 with $p \leq 3$, and 110 with $p > 3$. $b^*_{2,p}$ is $b_{2,p}$ centered on $p(p+2)$.

Table 3

Percent significant Mardia's skewness and kurtosis at significance level 0.05.

	By Sample Size		By # of Variables		Overall
	$n \leq 106$	$n > 106$	$p \leq 3$	$p > 3$	
Skewness	34	86	53	65	58
Kurtosis	35	82	47	70	57
Either	46	94	60	79	68

Note. There were 136 multivariate distributions with $n \leq 106$, 118 with $n > 106$, 144 with $p \leq 3$, and 110 with $p > 3$. Nonnormality is defined by significant statistics $z_{1,p}$ or $z_{2,p}$, $p < .05$.

Table 4

Type I error rates of the one-sample t-test

		Tail Tested		
Sample Size	Skewness	Two-tailed	Lower-tail	Upper-tail
18	-2.08	0.057	0.029	**0.079**
48	-2.08	0.055	0.033	0.072
105	-2.08	0.052	0.037	0.065
555	-2.08	0.05	0.043	0.058
1488	-2.08	0.05	0.046	0.057
18	-1.17	0.048	0.035	0.064
18	-0.33	0.046	0.045	0.053
18	0.2	0.045	0.051	0.046
18	0.94	0.049	0.061	0.038
18	2.77	0.064	**0.092**	**0.023**
48	2.77	0.06	**0.082**	0.027
105	2.77	0.056	0.075	0.031
555	2.77	0.05	0.062	0.039
1488	2.77	0.052	0.059	0.045
18	6.32	**0.177**	**0.216**	**0.005**
48	6.32	**0.123**	**0.157**	**0.011**
105	6.32	**0.09**	**0.12**	**0.016**
555	6.32	0.062	**0.081**	0.028
1488	6.32	0.055	0.069	0.035

Note. Bolded entries are those outside of the range [0.025,0.075] and are therefore considered different from the nominal 0.05.

Table 5

Type I error rates of the χ^2 test for factor analysis.

		Sample Size			
Kurtosis	Centered Kurtosis	48	106	554	1489
21	-3	0.061	0.058	0.060	0.060
24	0	0.053	0.046	0.048	0.050
30	6	0.055	0.052	0.055	0.056
60	36	**0.108**	**0.121**	**0.149**	**0.152**
100	76	**0.161**	**0.215**	**0.287**	**0.298**

Note. Bolded entries are those outside of the range [0.025,0.075] and are therefore considered different from the nominal 0.05.

Table 6

Empirical Type I error rates of the χ^2 test for factor analysis with 8 manifest variables.

			Sample Size		
# of Factors	Kurtosis	Centered Kurtosis	106	554	1489
1	75	-5	0.0654	0.0695	0.0688
	80	0	0.0533	0.0528	0.0502
	90	10	0.0546	0.0591	0.0574
	150	70	**0.191**	**0.2449**	**0.2603**
	250	170	**0.4159**	**0.5847**	**0.6373**
2	75	-5	**0.0861**	0.0675	0.0609
	80	0	0.0729	0.0549	0.0522
	90	10	**0.0781**	0.061	0.0597
	150	70	**0.1664**	**0.1695**	**0.1652**
	250	170	**0.3134**	**0.3746**	**0.4126**

Note. Bolded entries are those outside of the range [0.025,0.075] and are therefore considered different from the nominal 0.05.

Table 7

Critical summary statistics for different methods in empirical research

Methods	Recommendations
Mean comparison (t-test, F-test)	means, variances, univariate skewness
Repeated measures ANOVA, MANOVA	means, variances, univariate & multivariate skewness
Regression and correlation	mean, variance-covariances, univariate & multivariate kurtosis
Factor and principal and component analysis	variances-covariances, univariate & multivariate kurtosis
SEM without mean structures	variances-covariances, univariate & multivariate kurtosis
SEM with mean structures	means, variances-covariances, univariate & multivariate skewness & kurtosis
Growth curve analysis	means, variances-covariances, univariate & multivariate skewness & kurtosis

Table 8

Robust procedures available in current software

	Robust estimation	Sandwich-type SE	Satorra-Bentler	Free
WebSEM	X	X	X	X
rsem	X	X	X	X
EQS	X	X	X	
Mplus		X	X	
Amos				

Note. This table shows which software packages currently offer robust estimation using Huber-type weighting, sandwich-type standard errors, the Satorra-Bentler scaled chi-square statistic, and are available for free to their users.

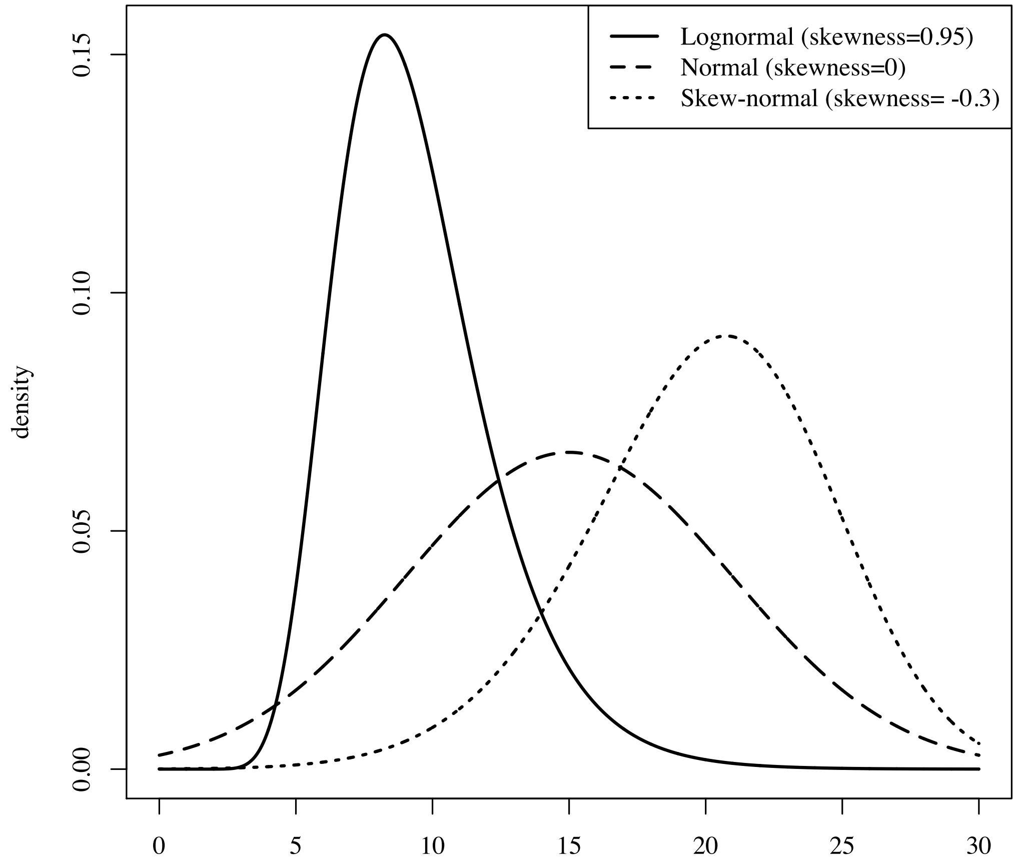
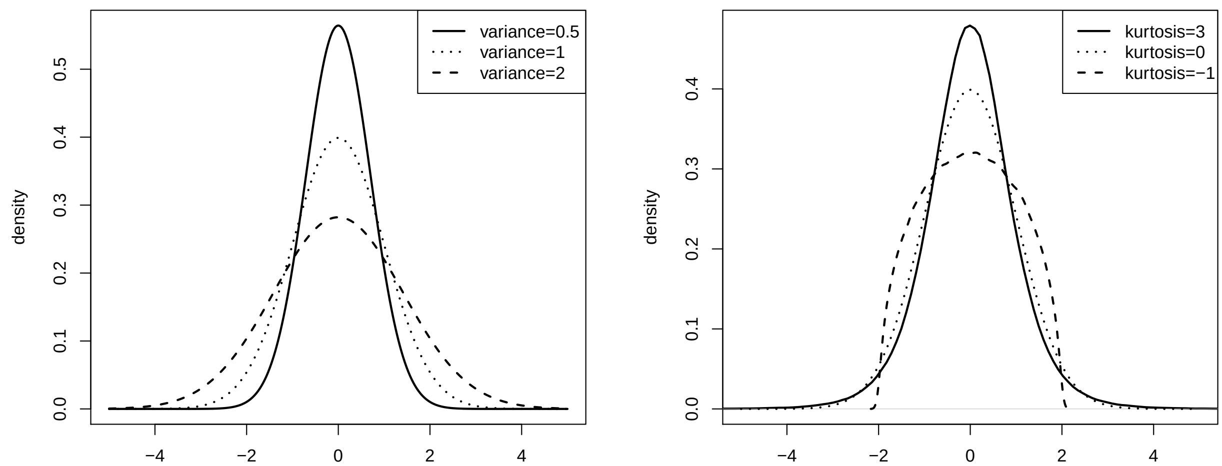


Figure 1. Illustration of positive and negative skewness.



(a) Normal distributions with the same kurtosis = 0, different variance.

(b) Distributions with same variance = 1, different kurtosis.

Figure 2. Illustration of the relationship between kurtosis and variance. In Figure 2(a) each population has a kurtosis of 0, and variance varies from 0.5 to 2.0. In Figure 2(b) each population has a variance of 1, and kurtosis varies from -1 to 3.

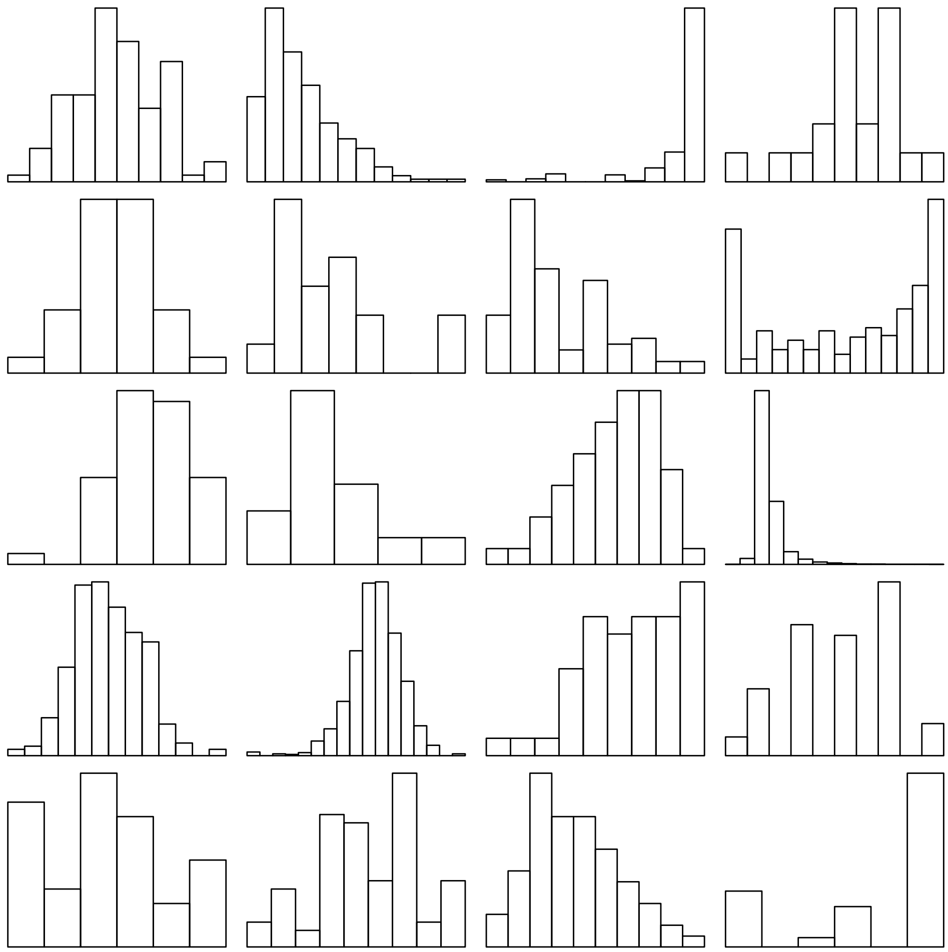


Figure 3. Histograms of 20 randomly selected distributions collected for review

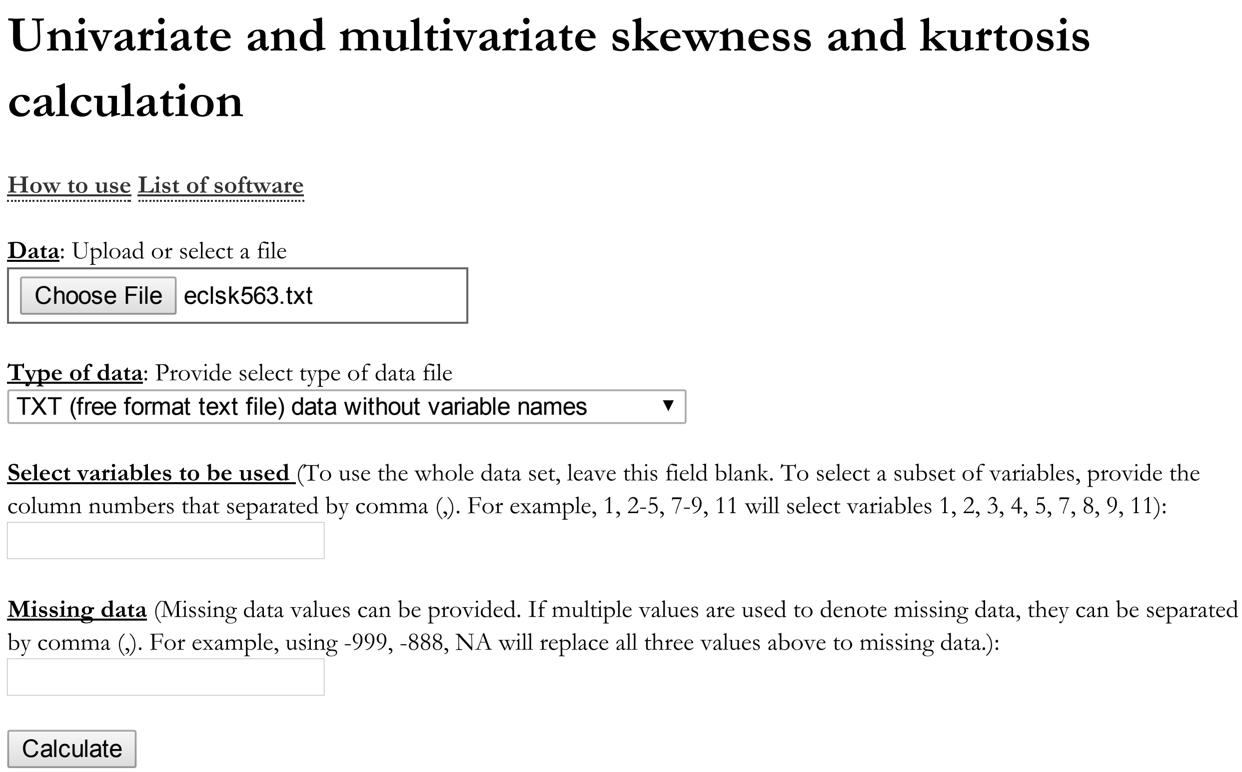


Figure 4. Interface of the Web application