Two Simple Approximations to the Distributions of Quadratic Forms*

Ke-Hai Yuan and
University of Notre Dame

Peter M. Bentler
University of California, Los Angeles

Abstract

Many test statistics are asymptotically equivalent to quadratic forms of normal variables, which are further equivalent to $T = \sum_{i=1}^{d} \lambda_i z_i^2$ with z_i being independent and following $N(0, 1)$. Two approximations to the distribution of T have been implemented in popular software and are widely used in evaluating various models. It is important to know how accurate these approximations are when compared to each other and to the exact distribution of T. The paper systematically studies the quality of the two approximations and examines the effect of λ_i's and the degrees of freedom d by analysis and Monte Carlo. The results imply that the adjusted distribution for T can be as good as knowing its exact distribution. When the coefficient of variation of the λ_i's is small, the rescaled statistic $T_R = dT / \left(\sum_{i=1}^{d} \lambda_i \right)$ is also adequate for practical model inference. But comparing T_R against χ_d^2 will inflate type I errors when substantial differences exist among the λ_i's, especially, when d is also large.

Keywords

Adjusted chi-square; rescaled statistic; coefficient of variation; Kolmogorov-Smirnov statistic; type I error

1. Introduction

In many statistical problems, the statistics for testing null hypotheses are asymptotically equivalent to quadratic forms of normal variables, which may not follow a chi-square distribution. Examples include the general likelihood ratio (LR) statistic when the distribution is misspecified (Foutz & Srivastava, 1977; Vuong, 1989); the Pearson chi-square statistic for contingency tables when the true covariance matrix of the estimated cells cannot be consistently estimated (Rao & Scott, 1984); test statistics in covariance structure analysis when the discrepancy function is specified using the normality assumption but the true underlying population distribution of the sample is unknown (Shapiro, 1983); test statistics for dimension reduction in inverse regression when the underlying distribution of the predictors is unknown (Li, 1991, 1992; Bura & Cook, 2001; Cook & Ni, 2005); the likelihood ratio statistic in testing the number of components in a normal mixture model when the null hypothesis holds (Lo, Mendell & Rubin, 2001). The quadratic forms are also

the building blocks for the commonly used *F*-statistics in ANOVA and regression. The distribution of a quadratic form of normal variables can be characterized by a linear combination of independent chi-square variates, each with one degree of freedom. Because the exact distribution of a linear combination of independent chi-square variates is difficult to obtain in general, various approximations to its distribution have been proposed (Solomon & Stephens, 1977). Two relatively simple ones are widely used in practice, one is to rescale the involved statistic, the other is to adjust the chi-square distribution. The purpose of this paper is to study these two distribution approximations using analysis and Monte Carlo. In section 2 we review the two approximations, their use in practice, and existing studies. The necessity and framework for the current study are also made clear after the review. In section 3 we study the effect of the coefficients on the approximations. Section 4 presents Monte Carlo results. Conclusion and discussion are provided in section 5.

2. Two Approximations to the Distribution of Quadratic Forms

Let $\mathbf{x} \sim N_p(\mathbf{0}, \mathbf{\Gamma})$ and $T = \mathbf{x}'\mathbf{W}\mathbf{x}$ be a quadratic form in $\mathbf{x}$. The matrix $\mathbf{\Gamma}$ is typically of full rank while $\mathbf{W}$ is nonnegative definite. Let the rank of $\mathbf{W}$ be d and the nonzero eigenvalues of $\mathbf{W}\mathbf{\Gamma}$ be $\lambda_1, \lambda_2, \ldots, \lambda_d$. There exists

$$T = \mathbf{x}'\mathbf{W}\mathbf{x} = \sum_{i=1}^{d} \lambda_i z_i^2, \tag{1}$$

where $z_i \sim N(0, 1)$ and are independent. Let $c = \sum_{i=1}^{d} \lambda_i / d$, the first approximation to the distribution of T is to rescale T by $T_R = c^{-1}T$ and compare T_R against χ_d^2 for inference. We will use the notation

$$T_R \overset{.}{\sim} \chi_d^2 \text{ or } T \overset{.}{\sim} c\chi_d^2 \tag{2}$$

to imply approximating the distribution of T_R by χ_d^2 or that of T by $c\chi_d^2$. It is obvious that $E(T_R) = d$, so that the rescaling is actually a mean correction. A more sophisticated correction is to also adjust the degrees of freedom of the chi-square distribution as in

$$T \overset{.}{\sim} a\chi_b^2, \tag{3}$$

where a and b are determined by matching the first two moments of T with those of $a\chi_b^2$. Straightforward calculation leads to

$$a = \frac{\sum_{i=1}^{d} \lambda_i^2}{\sum_{i=1}^{d} \lambda_i} \text{ and } b = \frac{\left(\sum_{i=1}^{d} \lambda_i\right)^2}{\sum_{i=1}^{d} \lambda_i^2}.$$

These approximations were originally proposed by Welch (1938) and further studied by Satterthwaite (1941) and Box (1954). When both $\mathbf{\Gamma}$ and $\mathbf{W}$ can be consistently estimated, c, a and b will be estimated as

$$\widehat{c} = \text{tr}\left(\widehat{\mathbf{W}\mathbf{\Gamma}}\right) / d, \quad \widehat{a} = \text{tr}\left[\left(\widehat{\mathbf{W}\mathbf{\Gamma}}\right)^2\right] / \text{tr}\left(\widehat{\mathbf{W}\mathbf{\Gamma}}\right), \quad \widehat{b} = \left[\text{tr}\left(\widehat{\mathbf{W}\mathbf{\Gamma}}\right)\right]^2 / \text{tr}\left[\left(\widehat{\mathbf{W}\mathbf{\Gamma}}\right)^2\right].$$

In dealing with the effect of survey design on analyzing multiway contingency tables, Rao and Scott (1984) noted that the approximations in (2) and (3) are practically adequate. In the context of covariance structure analysis, Satorra and Bentler (1988, 1994) proposed using the two approximations when T is the normal distribution based LR statistic. Monte Carlo results in Hu, Bentler and Kano (1992) showed that the approximation in (2) performed very well. The rescaled statistic T_R in (2) has been in standard software (EQS, LISREL, MPLUS) for many years and used in numerous publications by researchers in psychology, education, sociology, medicine, business, etc. The adjusted distribution for T in (3) has also been in popular software (e.g., EQS, MPLUS) and widely used in analyzing data with violation of distribution conditions.

Although these two approximations have been used for inference on a variety of models, their relative merits are not well-understood. In the context of covariance structure analysis, Fouladi (1997, 2000) reported that (3) performs better than (2). In testing the dimensionality of the space of the effective predictors using inverse regression, Bura and Cook (2003) also found that (3) performs better than (2). However, Bentler and Xie (2000) found that (2) performs much better than (3). These conclusions are based on examples and simulated type I errors, not the overall distribution approximation. Satorra and Bentler (1994) reported a few percentiles of T and T_R using a small simulation, they did not contrast the two approximations. As we shall see, the performance of the two approximations depends on the values of the coefficients λ_i's in (1). None of the above studies have controlled these coefficients. Actually, in any of these contexts, it is rather difficult to control the λ_i's when $\mathbf{\Gamma}$ and $\mathbf{W}$ are derived from models. Even when all the λ_i's can be specified, their effect on (2) and (3) will be confounded with sampling errors due to finite sample sizes.

In practice, the significance of a statistic is reported either using its p-value or indicating whether the null hypothesis is rejected at a certain level. For the statistic T_R, the p-value is the probability

$$P(T_R \geq t_R),$$

where t_R is the observed value of T_R and the probability is evaluated according to the true distribution of T_R. Because the true distribution of T_R is unknown, the reported p-value in the output of a software is calculated using χ^2_d according to (2). Obviously, for the p-value to make sense, the true distribution of T_R needs to be well approximated by χ^2_d in the interval $[t_R, \infty)$. Because $t_R \in [0, \infty)$, the overall distribution of T_R needs to be well described by χ^2_d in order to trust the reported p-value. Similarly, the overall distribution of T needs to be well described by $a\chi^2_b$ in order for the p-value based on (3) to make sense. When the statistic is used for purely hypothesis testing, the reference distribution needs to describe the tail behavior of the statistic well in order to properly control type I errors. In this paper, we will contrast the approximations (2) and (3) in both tail behavior and the overall distribution. We will also study how these approximations perform when compared to a statistic that exactly follows a chi-square distribution. When studying them through a LR or Pearson chi-square statistic, we will not be able to separate the approximation of the distribution of the statistic with that of a quadratic form from those in (2) and (3). So we will only study (2) and (3) when $\mathbf{W}$ and $\mathbf{\Gamma}$ or the λ_i's in (1) are given. Using known λ_i's also allows us to easily design conditions on their relative size, which has the strongest effect on the two approximations. We will discuss the implication and limitation of the obtained results in the concluding section.

3. Effect of the Coefficients λ_i's on the Approximating Distributions

In this section we study the effect of the λ_i's on the approximations in (2) and (3) by analysis, and relate the a and b to the coefficient of variations of the λ_i's. We will also introduce the Kolmogorov-Smirnov distance and a related measure of mean distance between two distributions, which will be used to study the performance of the two approximations in the next section by Monte Carlo.

Consider when $\lambda_1 = \lambda_2 = \ldots = \lambda_d = \lambda$, then $c = \lambda$, $a = \lambda$ and $b = d$, and the approximations in (2) and (3) are perfect. When all the λ_i's in (1) change proportionally, i.e., λ_i becoming $\tau\lambda_i$, then T changes to τT; c changes to τc; a changes to τa and b remains the same. In such a case, the qualities of the approximations in (2) and (3) do not change. So it is the relative sizes of the λ_i's that affect the two approximations.

When $\Sigma_{i=1}^{d}\lambda_i$ is a constant while the λ_i's change, the distribution of $\Sigma_{i=1}^{d}\lambda_i z_i^2$ will change. But the scaling factor c remains the same. So the quality of the approximation in (2) is affected as variations occur among the λ_i's. It is obvious that the relative sizes of the λ_i's also affect the approximation in (3). To see how a and b change when the λ_i's change, we rewrite b as

$$b = \left(\sum_{i=1}^{d} \tau_i^2 \right)^{-1},$$

where $\tau_i = \lambda_i / \Sigma_{i=1}^{d}\lambda_i$. Because $\Sigma_{i=1}^{d}\tau_i = 1$, $\Sigma_{i=1}^{d}\tau_i^2$ reaches its minimum when $\tau_1 = \tau_2 = \ldots = \tau_d = 1/d$. This implies that b reaches its maximum value at d when all the λ_i's are equal; b decreases as the λ_i's depart from each other. Because $ab = \Sigma_{i=1}^{d}\lambda_i$, when holding $\Sigma_{i=1}^{d}\lambda_i$ constant, a will increase when the λ_i's depart from each other. Of course, when $\Sigma_{i=1}^{d}\lambda_i$ decreases, it is very likely that both a and b decrease.

We may use the coefficient of variation of the λ_i's,

$$\mathrm{CV}(\lambda) = \frac{\mathrm{SD}(\lambda)}{\bar{\lambda}} = \frac{\left\{ \Sigma_{i=1}^{d} \left(\lambda_i - \bar{\lambda} \right)^2 / d \right\}^{1/2}}{\bar{\lambda}},$$

to measure the relative variations among the λ_i's, where $\bar{\lambda} = \Sigma_{i=1}^{d}\lambda_i / d$. When $\mathrm{CV}(\lambda) = 0$, both the approximations in (2) and (3) are perfect. They become poorer as $\mathrm{CV}(\lambda)$ increases. Actually, both a and b are closely related to $\mathrm{CV}(\lambda)$. It follows from

$$\mathrm{CV}^2(\lambda) = \frac{\Sigma_{i=1}^{d}\lambda_i^2 / d - \bar{\lambda}^2}{\bar{\lambda}^2} = d \frac{\Sigma_{i=1}^{d}\lambda_i^2}{\left(\Sigma_{i=1}^{d}\lambda_i \right)^2} - 1$$

and $ab = \Sigma_{i=1}^{d}\lambda_i$ that

$$b = \frac{d}{\mathrm{CV}^2(\lambda) + 1} \text{ and } a = \bar{\lambda} \left[\mathrm{CV}^2(\lambda) + 1 \right].$$

So the approximations in (2) and (3) are equivalent only when $CV(\lambda) = 0$. The distribution approximation in (2) can be regarded as approximating (3) by treating $CV(\lambda) = 0$ even when it is not. So we would expect that the difference between (2) and (3) becomes obvious when $CV(\lambda)$ increases.

The approximation will also depend on the degrees of freedom. As d increases, according to the central limit theorem, the distribution of T may even be approximately described by a normal distribution, and so may $a\chi_b^2$, T_R and χ_d^2. Thus, we may expect that the approximations in (2) and (3) will improve as d increases, which will be examined by Monte Carlo in the next section.

We will use the well-known Kolmogorov-Smirnov (KS) statistic to evaluate the overall distribution approximations in (2) and (3). The KS statistic measures the distance between the empirical distribution function (EDF) $\hat{F}(t)$ and the proposed target distribution function $G(t)$; $F(t)$ will be reserved for the true cumulative distribution function (CDF) of T. Suppose we have N independent observations on T. Let the ordered statistics be $t_{(1)} \leq t_{(2)} \leq \ldots t_{(N)}$, the KS is calculated by

$$\text{KS} = \max_{1 \leq i \leq N} \text{KS}_i \text{ with KS}_i = \max\left\{ \left| \frac{i-1}{N} - G\left(t_{(i)}\right) \right|, \left| \frac{i}{N} - G\left(t_{(i)}\right) \right| \right\}.$$

Because KS is determined by one point on the real line, it does not tell us the whole picture of the approximation. For example, with one statistic and a given distribution we may have $\text{KS}_1 = .2$ and $\text{KS}_i < .001$ for $i > 1$; with another statistic and a given distribution we may have $\text{KS}_i = .1$ for all i. Then, $\text{KS} = .2$ for the first statistic and $\text{KS} = .1$ for the second statistic. However, the distribution description for the first statistic is better than that for the second statistic except at the left tail. Another measure that better characterizes the overall discrepancy between $\hat{F}(t)$ and $G(t)$ is the average or the mean of the KS_i,

$$\text{MKS} = \sum_{i=1}^{N} \text{KS}_i / N.$$

This statistic was proposed in Yuan, Hayashi and Bentler (2007). We will use it in the next section to study the distribution approximations in (2) and (3). The maximum value of the KS is 1.0, which implies that $\hat{F}(t)$ and $G(t)$ do not have any overlap. To see the maximum value of MKS, we may assume that $\hat{F}(t)$ is above $G(t)$ or $G(t_{(1)}) = 1.0$, then $\text{KS}_i = 1 - (i - 1)/N$ and

$$\begin{aligned}
\text{MKS} \quad &= \frac{1}{N} \sum_{i=1}^{N} \left[1 - \frac{(i-1)}{N} \right] \\
&= 1 - \frac{1}{N^2} \left[\frac{N(N+1)}{2} - N \right] \approx \frac{1}{2}.
\end{aligned}$$

. The KS and MKS will be used to measure the distance between the EDF of T_R and the CDF of χ_d^2 as well as that between the EDF of T/a and the CDF of χ_b^2 in the next section.

4. Monte Carlo Results

In this section we use Monte Carlo to study the effect of $CV(\lambda)$ and d on the two approximations in (2) and (3). First, we let $CV(\lambda)$ change with a fixed d; next, we let d

change; then we let both d and CV(λ) change. We will start with the overall distribution and then turn to the tail behavior as reflected by type I errors.

4.1 Overall distribution

Let $d = 2$ and the vector of the λ_i's = $(1, k)'$. The CV(λ), KS and MKS for $k = 2$ to 10, with N = 2000 replications, are reported in Table 1(a). The KS and MKS under χ^2_d are obtained by comparing a simulated chi-square variate to the chi-square distribution χ^2_d, they represent what KS and MKS are like under a perfect situation. Because KS does not depend on $F(t)$ (Serfling, 1980, p. 62), we can also regard that KS and MKS under χ^2_d correspond to the discrepancy between the EDF of T and the CDF of T. The KS and MKS under $c\chi^2_d$ correspond to the approximation in (2), those under $a\chi^2_d$ correspond to the approximation in (3). Because applied researchers commonly use the nominal χ^2_d as the reference distribution for the LR statistic without checking the distribution of the sample, we also include

$$T \overset{\cdot}{\sim} \chi^2_d \tag{4}$$

in the study. The KS and MKS corresponding to (4) are reported under $L\chi^2_d$, where L is for "linear combination" of chi-square variates.

When $k = 2$, CV(λ) = .333, the KS under $c\chi^2_d$ is the smallest, the MKS under $c\chi^2_d$ is also comparable to that for the ideal case. As k or CV(λ) increases, the KS and MKS under χ^2_d fluctuate; the KS and MKS under $c\chi^2_d$ also fluctuate, but they tend to increase; those under $a\chi^2_b$ also tend to increase, but the speed is a lot smaller; those under $L\chi^2_d$ are always the greatest. The biggest number in each column is marked in boldface, which indicates how large KS and MKS can be under the worst case. Because the greatest number under χ^2_d is just by chance and the great numbers under $L\chi^2_d$, $c\chi^2_d$ and $a\chi^2_b$ are due to systematic errors in addition to chance, comparing these numbers gives us the information on the quality of the approximation in (2), (3) and (4). The KS corresponding to (4) is about 15 times of the perfect case; that corresponding to $c\chi^2_d$ is 4 times of the perfect case; that corresponding to $a\chi^2_b$ is about twice of the perfect case. Comparisons of the largest MKS's are similar to those of the KS's. Each number in the last row of Table 1(a) is the average of the previous rows, according to which the approximation in (3) is a lot better than that in (2). Actually, only when $k = 2$ does the approximation in (2) enjoy smaller KS and MKS than those for (3); we put this condition in boldface in the first column of the table.

To see the effect of the degrees of freedom on the approximations in (2) and (3), we next study the conditions of $d = 6$ with $\lambda_k = \left(1'_3, k1'_3\right)'$ and $d = 10$ with $\lambda_k = \left(1'_5, k1'_5\right)'$, where 1_j represents a vector of j 1's. The KS and MKS are reported in Tables 1(b) and (c), respectively. Although the degrees of freedom increased, the CV(λ) for a given k is the same due to the same two distinct λ_i's. The patterns of KS and MKS under $L\chi^2_d$ and $c\chi^2_d$ in Tables 1(b) and (c) are about the same as in (a), they tend to increase as CV(λ) increases. However, the KS and MKS under $a\chi^2_b$ may just fluctuate. Actually, the greatest KS or MKS under $a\chi^2_b$ in either Table 1(b) or (c) is smaller than that under χ^2_d. Comparing the averaged KS or MKS at the bottom of Tables 1(a), (b) and (c), we notice that those corresponding to χ^2_d tend

to be stable as d changes, since the distribution of KS does not depend on F; the KS and MKS corresponding to $c\chi_d^2$ also appear not affected when d changes; the KS and MKS corresponding to $L\chi_d^2$ obviously increase when d increases; those corresponding to $a\chi_b^2$ tend to decrease as d increases. At $d = 10$, with three decimals, the average KS under $a\chi_b^2$ is identical to that under χ_d^2, and so is the average MKS.

Mean and covariance structure analysis typically involves many variables, the degrees of freedom can be much larger than those studied in Table 1; there can be many predictors in regression and the degrees of freedom can also be very large in testing the number of principal Hessian directions when using inverse regression. It is most likely that, as the dimension increases, the corresponding CV(λ) also change. To further compare the two approximations in (2) and (3) under these conditions, we choose (a) $d = 10$ with ten conditions on λ_i's: $\lambda_1 = (1, 1.1, 1.2, \ldots, 1.9)'$, $\lambda_2 = (1, 1.2, 1.4, \ldots, 2.8)'$, $\ldots$, $\lambda_{10} = (1, 2, 3, \ldots, 10)'$; (b) $d = 30$ with ten conditions on λ_i's: $\lambda_1 = (1, 1.1, 1.2, \ldots, 3.9)'$, $\lambda_2 = (1, 1.2, 1.4, \ldots, 6.8)'$, $\ldots$, $\lambda_{10} = (1, 2, 3, \ldots, 30)'$; and (c) $d = 50$ with ten conditions on λ_i's: $\lambda_1 = (1, 1.1, 1.2, \ldots, 5.9)'$, $\lambda_2 = (1, 1.2, 1.4, \ldots, 10.8)'$, $\ldots$, $\lambda_{10} = (1, 2, 3, \ldots, 50)'$. The KS and MKS using $N = 2000$ as well as the associated CV(λ) are reported in Tables 2(a), (b) and (c), respectively. Except when $d = 10$ and $k = 5$ where KS and MKS under $c\chi_d^2$ are smaller than those under both χ_d^2 and $a\chi_b^2$, all the other KS and MKS corresponding to the approximation in (3) are smaller than those corresponding to the approximation in (2). The KS and MKS under $a\chi_b^2$ in Table 2(a) are almost as small as those under χ_d^2; the average KS and MKS under $a\chi_b^2$ in Table 2(b) are even smaller than those under χ_d^2, due to sampling errors. The average KS and MKS under $a\chi_b^2$ are identical to those under χ_d^2 in Table 2(c). As d and CV(λ) increase, the KS and MKS under $L\chi_d^2$ reach their maximum; then it is meaningless to approximate the linear combination of chi-square variate by the nominal chi-square distribution.

In the practice of principal components and factor analysis, when ordering the eigenvalues of a sample covariance matrix from large to small, it often happens that the first few drop dramatically in size, the remaining ones slowly decrease. The phenomenon that most smaller eigenvalues sit on a line is called the scree test in factor analysis (see Gorsuch, 1983, pp. 165–169). We also include the following conditions to mimic such a phenomenon: $d = 10$ and $\lambda_{10} = (1, 1.1, 1.2, \ldots, 1.7, 1.8, 10)'$, 9 eigenvalues are evenly spaced except the largest one; $d = 20$ and $\lambda_{20} = (1, 1.1, 1.2, \ldots, 2.6, 2.7, 10, 20)'$, 18 eigenvalues are evenly spaced except the largest two; $\ldots$, $d = 100$ and $\lambda_{100} = (1, 1.1, 1.2, \ldots, 9.7, 9.8, 9.9, 10, 20, 30, \ldots, 100)'$, 90 eigenvalues are evenly spaced except the largest ten. Table 3 contains the CV(λ) as well as the KS and MKS for these conditions. The CV(λ) increases as d increases, the KS and MKS under χ_d^2 remain stable as they should be; those under $L\chi_d^2$ reach their maximum values after $d = 50$ or 40; the KS and MKS under $c\chi_d^2$ tend to increase due to the increase of CV(λ); but the KS and MKS under $a\chi_b^2$ tend to decrease due to the increase of d although CV(λ) also increases.

We may conclude from Tables 1 to 3 that, when controlling CV(λ), the approximation in (2) is almost not affected by the degrees of freedom while the approximation in (3) improves as the degrees of freedom increase. For a given d, when CV(λ) increases, the approximation in (2) tends to become worse; the approximation in (3) also tends to become worse when d is small. At a large d, the approximation in (3) is almost not affected by the change of CV(λ).

4.2 Type I errors

Type I errors for each distribution description are obtained under the same conditions as those for the overall distribution approximation. We also use the same notation as introduced in the previous subsection. For each condition we report type I errors corresponding to nominal level $\alpha = .01$, .025, .05 and .10, which are most widely used in the applied literature.

Tables 4(a) to (c) contain Monte Carlo type I errors corresponding to the conditions in Tables 1(a) to (c), respectively. Each table also contains the average of the absolute differences (AAD) between the Monte Carlo type I errors and the nominal level α across all the conditions of CV(λ)'s for each distribution description.

Because type I errors under χ_d^2 are obtained by referring a chi-square random variable to the critical value from its true distribution, we may regard the differences between the Monte Carlo type I errors and the nominal level as due to sampling error. As expected, the Monte Carlo type I errors under $L\chi_d^2$ are much greater than the nominal level, due to half of the λ_j's being greater than 1.0. Not expected is that all the Monte Carlo type I errors under $c\chi_d^2$ are substantially greater than the nominal level, indicating poor approximation for (2) at the tail. The AAD under $c\chi_d^2$ ranges from 3 to 8 times of that under χ_d^2. The Monte Carlo type I errors under $a\chi_b^2$ fluctuate around the nominal levels, and are very comparable to those under $a\chi_b^2$. For the twelve AADs under $a\chi_b^2$, four of them equal those under $a\chi_b^2$; three are slightly greater than those under $a\chi_b^2$; five are slightly smaller than those under $a\chi_b^2$. Comparing the AADs in Tables 4(a) to (c), we find that the AAD under $c\chi_d^2$ increases with d and so does the AAD under $L\chi_d^2$. But the AADs under $a\chi_b^2$ are stable as are those under χ_d^2.

We may notice that, under χ_d^2, the AAD tends to increase as increases. This is because the variance of the Monte Carlo type I error, given by $\alpha(1 - \alpha)/N$, is an increasing function of α on the interval of [0, .5]. The smaller variance at a smaller leads to the smaller AADs. Similarly, the super behavior of the Monte Carlo type I errors in Tables 4(b) and (c) for the approximation in (3) might be explained by Var(T/a) = $2b$, and $b < d$ unless CV(λ) = 0.

Tables 5(a) to (c) contain the Monte Carlo type I errors corresponding to the conditions in Tables 2(a) to (c), respectively. Similar to those in Table 4, the Monte Carlo type I errors under $a\chi_b^2$ fluctuate around the nominal levels and are comparable to those under χ_d^2. As expected, all the Monte Carlo type I errors under $L\chi_d^2$ are way above the nominal levels. Except at $\alpha = .10$ for the condition of $k = 3$ in Table 5(a), all the Monte Carlo type I errors under $c\chi_d^2$ in Table 5 are above the nominal levels. The AADs under $c\chi_d^2$ are 2 to 9 times of those under χ_d^2. Comparing the AADs in Tables 5(a) to (c), we notice that those under $c\chi_d^2$ increase with d, while those under $a\chi_b^2$ are stable as are those under χ_d^2.

Table 6 contains the Monte Carlo type I errors corresponding to the conditions in Table 3. Those under $c\chi_d^2$ obviously depart more from the nominal levels than in the previous tables, those corresponding to the greater CV(λ)'s are more than 10 times of the nominal level when $\alpha = .01$, more than 5 times of the nominal level when $\alpha = .025$, more than triple of the nominal level when $\alpha = .05$, and more than twice of the nominal level when $\alpha = .10$. The AAD under $c\chi_d^2$ is from 14 to 38 times of the AAD under χ_d^2. At $\alpha = .01$, .025, the Monte Carlo type I errors under $a\chi_b^2$ are also systematically greater than the nominal levels,

indicating that the statistic T has a heavier tail than that of $a\chi_b^2$ in the extreme right tail. At α = .05, .10, the Monte Carlo type I errors under $a\chi_b^2$ are very comparable to those under χ_d^2. At α = .10, the AAD under $a\chi_b^2$ is even smaller than that under χ_d^2.

We may conclude from Tables 4 to 6 that, the tail approximation in (2) gets worse as the degrees of freedom increase or CV(λ) increases; both lead to greater type I errors. At the nominal levels α = .01, .025, .05 and .10, with CV(λ) ≤ .8, the approximation in (3) controls type I errors as well as knowing the true distribution of T . The approximation in (3) also controls type I errors as well as knowing the true distribution of T at α = .05 and .10 for greater CV(λ)'s. However, at α = .01 or .025 and with CV(λ) > 1, the approximation in (3) may lead to greater type I errors than the nominal levels.

5. Discussion and Conclusion

In this paper, we quantified the conditions that may affect the two widely used approximations. The quality of the two approximations was studied by varying and controlling the conditions. Because the true CDF, $F(t)$, of a quadratic form is hard to evaluate, we used the EDF $\hat{F}(t)$ to estimate it. In addition to using Monte Carlo, one may use a numerical method to approximate $F(t)$, which can be defined through an integral with an infinite upper limit. The procedure involves replacing the infinity limit by a finite number and followed by a numerical integration (see Farebrother, 1990). Errors will occur when replacing the infinite limit by a finite limit and when using a numerical method to calculate the area under a continuous curve. The amount of error depends on the chosen upper limit and the number of rectangles or trapezoids used in the numerical integration, it also depends on the value of x. The amount of computation in the numerical method can be huge although the error can be made arbitrarily small. Comparing to using numerical method to evaluate $F(t)$, the Monte Carlo EDF $\hat{\Phi}(t)$ approximates $F(t)$ with $E[\hat{\Phi}(t)] = F(t)$. The mean square error (MSE) in $\hat{\Phi}(t)$ can be characterized by

$$\text{MSE} = E\left[\widehat{F}(x) - F(x)\right]^2 = \text{Var}\left[\widehat{F}(x)\right] = F(x)\left[1 - F(x)\right]/N \leq 1/(4N).$$

With N = 2000 in the study, there exists MSE ≤ 1/(8000) = .000125. The MSE can be made smaller if we choose a larger N. But N = 2000 is enough for our purpose, that is, we can clearly tell the pros and cons of each of the two approximations under varied conditions.

The overall distribution approximations in (2) and (3) are comparable when both CV(λ) and the degrees of freedom are small. The approximation in (3) generally performs better, especially when d is large. When CV(λ) is not large, say less than 0.5 and d is greater than 10, the overall distribution approximation in (3) can be as good as knowing the exact distribution of T. The approximation in (3) also describes the right tail of the T in (1) as well as knowing the exact distribution of T. The approximation in (2) does not describe the right tail of the T in (1) well. In particular, the right tail of T_R is heavier than that of χ_d^2. Either a larger d or a larger CV(λ) makes the approximation of the tail behavior in (2) worse.

The results in section 4 suggest that, unless all the λ_i's are equal, (3) should be used to describe the distribution of T instead of (2). In practice, we will not have the λ_i's and thus do not know whether CV(λ) = 0 or not. Analytical result (see Muirhead, 1982, p. 388) implies that larger sample eigenvalues $\hat{\lambda}_i$'s tend to over-estimate their population counterparts and smaller ones tend to under-estimate their population counterparts. Thus, even when CV(λ) = 0, we still have a positive CV($\hat{\lambda}$). Further study on how to test CV(λ) = 0 will be valuable for

properly choosing the two approximations. With real data, it is unlikely for CV(λ) = 0. The adjusted statistic is preferred before a reliable procedure for testing CV(λ) = 0 is available.

In Monte Carlo studies with LR or other statistics, it may happen that the λ_i's are equal (see Yuan & Bentler, 1998). Then the approximation in (2) uses the correct assumption about the λ_i's and thus it will perform better than that in (3). Conflicting results on controlling type I errors by the two approximations (e.g., Fouladi, 1997, 2000; Bentler & Xie, 2000; Bura & Cook, 2003) would most likely be resolved if the population λ_i's were known.

Although many test statistics are asymptotically equivalent to quadratic forms, the two approximations are most widely used in mean and covariance structure analysis. As mentioned in sections 1 and 2, the rescaled statistic is available in EQS, LISREL and MPLUS; the adjusted statistic is available in MPLUS and in recent builds of EQS. The command "Method=xx, Robust;" in model specification for EQS computes both the rescaled and adjusted statistics for any estimator "xx" such as xx=ML or xx=GLS (see Bentler, in press, p. 8, 289). The same command is applicable to nonnormal continuous data, categorical data, and nonnormal missing data as well as various model types such as multilevel models and correlation structures. In MPLUS, the command for computing the rescaled statistic is "ESTIMATOR=MLR;", "ESTIMATOR=MLM;" or "ESTIMATOR=WLSM;", depending on data type and estimation method used; the command for generating the adjusted statistic is "ESTIMATOR=MLMV;" or "ESTIMATOR=WLSMV;", depending on data type and estimation method used (see Muthén & Muthén, 2006, p. 426). In LISREL the sample estimate of the asymptotic covariance matrix is computed with PRELIS and saved in a file with suffix "ACC." This is then read into LISREL, where the scaled statistic "C3" is computed as a correction to "C2," which is asymptotically equivalent to a quadratic form. The rescaled statistic is available for methods ULS, GLS, ML, and DWLS, as defined in LISREL (Jöreskog et al., 2000, Ch. 4).

A test statistic is asymptotically equivalent to a quadratic form only under the hypothesis of a correctly specified model[1]. So the results in section 4 are on type I errors. They are also related to type II errors or power. For example, the results in Tables 4 to 6 suggest that the right tail of T_R is heavier than that of χ_d^2. If one uses $T_R \tilde{\chi}_d^2$ for inference when CV(λ) is large, then the power will be artifically inflated. With misspecified models, a test statistic, including the LR statistic, cannot be asymptotically described by quadratic forms in general (see Yuan et al., 2007).

The paper has focused on the quadratic form in (1) with given W and Γ or λ_i's. In practice, a statistic is only approximated by a quadratic form rather than equals a quadratic form. The discrepancy between the quadratic form and the corresponding statistic depends on the sample size and the underlying population distribution; and it approaches zero as the sample size goes to infinity. Similarly, W and Γ or λ_i's will have to be estimated when applying either of the approximations in (2) or (3) for real data analysis. The discrepancy between the estimates $\hat{\lambda}_i$'s and λ_i's will also depend on the sample size and the underlying population distribution. When the underlying population distribution is unknown, we will not be able to quantify either of the discrepancies at a finite sample size. Such a difficulty is associated with almost all statistical inferences beyond the regression model with normally distributed errors.

---

[1]Model misspecification and distribution misspecification are different in this characterization. For example, regardless of the distributions of the factors and errors/uniquenesses in the common factor model, a three-factor model may be misspecified as a two-factor model; the distribution for the observed variables may be correctly or incorrectly specified as multivariate normal.

Despite the discrepancies, the obtained results in this paper agree well with many simulated results at finite sample sizes. For example, for the normal distribution based LR statistic and 340 conditions studied by Fouladi (2000), at $\alpha = .05$, the mean rejection rate of the approximation in (2) is .143, almost triple of the nominal level; while the mean rejection rate of the approximation in (3) is .067, only slight greater than the nominal level. Although Fouladi (2000) did not report $CV(\lambda)$, the results in the previous section and the substantial difference between the two approximations imply that many of the conditions must have $CV(\lambda)$ substantially different from 0. Yuan and Bentler (1998) studied the approximation (2) for the normal distribution based LR statistic and reported the population $CV(\lambda)$. When $CV(\lambda) = 0$ or .089, type I errors of the approximation (2) is very close to the nominal level when the sample size is greater than 500. However, at $CV(\lambda) = 2.38$, type I errors of the approximation in (2) move away/departure from the nominal level as the sample size increases.

Although we are unable to remove the discrepancy between the LR statistic and the quadratic form, the obtained results does provide us a clear picture on which statistic to choose. While recommending the approximation in (3), we need to emphasize that, even when $CV(\lambda = 0$, close distribution description of T by (3) as reported in Tables 1 to 3 or type I errors as reported in Table 4 to 6 may not be obtainable unless the sample size is huge. Actually, at smaller sample sizes, the LR statistic over-rejects the correct model substantially even when the population distribution is correctly specified (e.g., Bentler & Yuan, 1999). With using the approximation in (3), what we can expect is that the test controls type I errors as well as when the distribution is correctly specified when referring the LR test to the nominal chi-square distribution. Without studying the approximation in (2), many simulation studies in the literature endorsed the approximation in (2) when the sample size is moderate (e.g., Hu et al., 1992). The results in this paper together with those reported in Fouladi (1997,2000) imply that the approximation in (3) will perform better than that in (2) at both moderate sample sizes and large sample sizes. The approximation in (3) may even describe the behavior of T better than that in (2) at smaller sample sizes, but it needs to be studied on a cases-by-case basis.

The results in this section allows the separation of sampling error from systematic errors for inference based on (2) or (3). At a smaller n and with a $CV(\lambda)$ substantially greater than 0, the performance of the statistic in (2) suffers from both sampling error and the systematic error. If the sampling error is positive, then (2) will lead to serious over-rejection of the correct model. If sampling error is negative, then (2) may lead to slight-over rejection or slight under-rejection (see Table 10 and 11 of Yuan & Bentler, 1998). At a relative large n, inference based on (2) will lead to over-rejection unless $CV(\lambda)$ is small. The over-rejection or under-rejection of (3) will be mainly at smaller sample sizes.

Information can be obtained by working with a specific statistic at a given model using simulation. But it is very likely that simulated conditions may not reflect those in real data analysis. When the chosen statistic is the LR statistic, say based on the normal distribution, it is relatively easy to simulate conditions that all the λ_i's are equal. It is rather difficult to simulate conditions with varying $CV(\lambda)$'s, as given in the previous section. If the sample size n is relatively large, both the discrepancies will be small, then the results obtained here will be applicable to real data analysis.

Acknowledgments

References

Bentler, PM. EQS 6 structural equations program manual. Multivariate Software; Encino, CA: in press

Bentler PM, Xie J. Corrections to test statistics in principal Hessian directions. Statistics & Probability Letters 2000;47:381–89.

Bentler PM, Yuan K-H. Structural equation modeling with small samples: Test statistics. Multivariate Behavioral Research 1999;34:181–197.

Box GEP. Some theorems on quadratic forms applied in the study of analysis of variance problems: I. Effect of inequality of variance in one-way classification. Annals of Mathematical Statistics 1954;25:290–302.

Bura E, Cook RD. Extending sliced inverse regression: The weighted chi-squared test. Journal of the American Statistical Association 2001;96:996–1003.

Bura E, Cook RD. Assessing corrections to the weighted chi-squared test for dimension. Communication in Statistics: Simulation and Computation 2003;32:127–146.

Cook RD, Ni L. Sufficient dimension reduction via inverse regression: a minimum discrepancy approach. Journal of the American Statistical Association 2005;100:410–428.

Farebrother RW. Algorithm AS 256: The distribution of a quadratic form in normal variables. Applied Statistics 1990;39:294–309.

Fouladi RT. Type I error control of some covariance structure analysis techniques under conditions of multivariate nonnormality. Computing Science and Statistics 1997;29:526–532.

Fouladi RT. Performance of modified test statistics in covariance and correlation structure analysis under conditions of multivariate nonnormality. Structural Equation Modeling 2000;7:356–410.

Foutz RV, Srivastava RC. The performance of the likelihood ratio test when the model is incorrect. Annals of Statistics 1977;5:1183–1194.

Gorsuch, RL. Factor analysis. 2nd ed.. Lawrence Erlbaum Associates; Hillsdale, NJ: 1983.

Hu L, Bentler PM, Kano Y. Can test statistics in covariance structure analysis be trusted? Psychological Bulletin 1992;112:351–362. [PubMed: 1454899]

Jöreskog, KG.; Sörbom, D.; du Toit, S.; du Toit, M. LISREL 8: New statistical features. Scientific Software International; Lincolnwood, IL: 2000.

Li KC. Sliced inverse regression for dimension reduction. Journal of the American Statistical Association 1991;86:316–342.

Li KC. On principal Hessian directions for data visualization and dimension reduction: Another application of Stein's lemma. Journal of the American Statistical Association 1992;87:1025–1039.

Lo Y, Mendell NR, Rubin DB. Testing the number of components in a normal mixture. Biometrika 2001;88:767–778.

Muirhead, RJ. Aspects of multivariate statistical theory. Wiley; New York: 1982.

Muthén, LK.; Muthén, BO. Mplus users guide. 4th ed.. Muthén & Muthén; Los Angeles, CA: 2006. http://www.statmodel.com/download/usersguide/Mplus%20Users%20Guide%20v41.pdf

Rao JNK, Scott AJ. On chi-squared tests for multi-way contingency tables with cell proportions estimated from survey data. Annals of Statistics 1984;12:46–60.

Satorra, A.; Bentler, PM. 1988 Proceedings of Business and Economics Sections. American Statistical Association; Alexandria, VA: 1988. Scaling corrections for chi-square statistics in covariance structure analysis; p. 308-313.

Satorra, A.; Bentler, PM. Corrections to test statistics and standard errors in covariance structure analysis. In: von Eye, A.; Clogg, CC., editors. Latent variables analysis: Applications for developmental research. Sage; Newbury Park, CA: 1994. p. 399-419.

Satterthwaite FE. Synthesis of variance. Psychometrika 1941;6:309–316.

Serfling, RJ. Approximation theorems of mathematical statistics. Wiley; New York: 1980.

Shapiro A. Asymptotic distribution theory in the analysis of covariance structures (a unified approach). South African Statistical Journal 1983;17:33–81.

Solomon H, Stephens MA. Distribution of a sum of weighted chi-square variables. Journal of the American Statistical Association 1977;72:881–885.

Vuong QH. Likelihood ratio tests for model selection and nonnested hypotheses. Econometrica 1989;57:307–333.

Welch BL. The significance of the difference between two means when the population variances are unequal. Biometrika 1938;29:350–361.

Yuan K-H, Bentler PM. Normal theory based test statistics in structural equation modeling. British Journal of Mathematical and Statistical Psychology 1998;51:289–309. [PubMed: 9854947]

Yuan K-H, Hayashi K, Bentler PM. Normal theory likelihood ratio statistic for mean and covariance structure analysis under alternative hypotheses. Journal of Multivariate Analysis 2007;98:1262–1282.

Table 1(a)

Kolmogorov-Smirnov (KS) distances and the mean KS (MKS) between the empirical distributions of the statistics and their proposed distributions, $d = 2$, $\lambda_k = (1, k)'$, 2000 replications.

k	CV(λ)	KS				MKS			
		χ_d^2	$L\,\chi_d^2$	$c\chi_d^2$	$a\chi_b^2$	χ_d^2	$L\,\chi_d^2$	$c\chi_d^2$	$a\chi_b^2$
2	0.333	0.018	0.150	0.017	0.029	0.006	0.098	0.008	0.011
3	0.500	0.012	0.235	0.037	0.033	0.004	0.149	0.016	0.014
4	0.600	0.011	0.278	0.054	0.050	0.004	0.183	0.030	0.020
5	0.667	0.028	0.330	0.074	0.046	0.011	0.207	0.038	0.018
6	0.714	0.018	0.376	0.064	0.045	0.007	0.233	0.038	0.020
7	0.750	0.018	0.408	0.077	0.060	0.010	0.256	0.039	0.022
8	0.778	0.012	0.430	0.084	0.056	0.003	0.267	0.047	0.021
9	0.800	**0.030**	0.421	**0.120**	0.062	**0.014**	0.269	**0.063**	**0.026**
10	0.818	0.015	**0.462**	0.109	**0.066**	0.004	**0.289**	0.053	0.021
Ave	0.662	0.018	0.343	0.071	0.050	0.007	0.217	0.037	0.019

Table 1(b)

Kolmogorov-Smirnov (KS) distances and the mean KS (MKS) between the empirical distributions of the statistics and their proposed distributions, $d = 6$, $\lambda_k=\left(1'_3, k1'_3\right)'$, 2000 replications.

k	CV(λ)	KS				MKS			
		χ^2_d	$L\chi^2_d$	$c\chi^2_d$	$a\chi^2_b$	χ^2_d	$L\chi^2_d$	$c\chi^2_d$	$a\chi^2_b$
2	0.333	0.020	0.269	0.020	0.024	0.008	0.178	0.010	0.008
3	0.500	0.014	0.409	0.040	0.018	0.007	0.270	0.021	0.007
4	0.600	0.021	0.510	0.057	0.025	0.007	0.324	0.030	0.010
5	0.667	0.018	0.572	0.079	0.029	0.006	0.361	0.041	0.013
6	0.714	0.016	0.623	0.086	**0.033**	0.006	0.382	0.046	0.014
7	0.750	**0.042**	0.666	0.098	0.025	**0.018**	0.400	**0.051**	0.011
8	0.778	0.015	0.709	**0.103**	0.022	0.004	0.421	0.048	0.009
9	0.800	0.031	0.754	0.071	0.026	0.014	0.433	0.040	**0.016**
10	0.818	0.018	**0.760**	0.101	0.020	0.006	**0.441**	0.049	0.006
Ave	0.662	0.022	0.586	0.073	0.025	0.009	0.357	0.037	0.010

Table 1(c)

Kolmogorov-Smirnov (KS) distances and the mean KS (MKS) between the empirical distributions of the statistics and their proposed distributions, $d = 10$, $\lambda_k = \left(1'_5, k1'_5\right)'$, 2000 replications.

k	CV(λ)	KS				MKS			
		χ^2_d	$L\chi^2_d$	$c\chi^2_d$	$a\chi^2_b$	χ^2_d	$L\chi^2_d$	$c\chi^2_d$	$a\chi^2_b$
2	0.333	0.022	0.332	0.025	0.012	0.010	0.220	0.012	0.004
3	0.500	0.015	0.523	0.047	0.016	0.004	0.336	0.019	0.005
4	0.600	0.016	0.630	0.071	0.025	0.007	0.386	0.032	0.012
5	0.667	0.029	0.698	0.083	0.031	0.015	0.422	0.043	0.015
6	0.714	**0.034**	0.786	0.054	**0.033**	**0.018**	0.453	0.032	**0.016**
7	0.750	0.016	0.815	0.069	0.014	0.005	0.461	0.037	0.005
8	0.778	0.010	0.838	0.085	0.022	0.004	0.467	0.043	0.007
9	0.800	0.019	0.859	**0.094**	0.012	0.005	0.476	**0.047**	0.005
10	0.818	0.019	**0.882**	0.082	0.015	0.006	**0.483**	0.044	0.006
Ave	0.662	0.020	0.707	0.068	0.020	0.008	0.412	0.034	0.008

Table 2(a)

Kolmogorov-Smirnov (KS) distances and the mean KS (MKS) between the empirical distributions of the statistics and their proposed distributions, $d = 10$, $\lambda_k = \mathbf{1}_{10} + k(0, 0.1, 0.2, \ldots, 0.9)'$, 2000 replications.

k	CV(λ)	KS				MKS			
		χ_d^2	$L\chi_d^2$	$c\chi_d^2$	$a\chi_b^2$	χ_d^2	$L\chi_d^2$	$c\chi_d^2$	$a\chi_b^2$
1	0.198	0.022	0.304	0.022	0.019	0.010	0.208	0.010	0.007
2	0.302	0.015	0.505	0.020	0.015	0.004	0.331	0.007	0.003
3	0.367	0.016	0.632	0.041	0.029	0.007	0.388	0.016	0.011
4	0.410	0.029	0.701	**0.053**	**0.035**	0.015	0.424	**0.025**	0.015
5	0.442	**0.034**	0.799	0.032	0.032	**0.018**	0.460	0.015	**0.018**
6	0.466	0.016	0.824	0.029	0.012	0.005	0.466	0.016	0.005
7	0.484	0.010	0.853	0.036	0.018	0.004	0.475	0.018	0.006
8	0.500	0.019	0.875	0.040	0.013	0.005	0.482	0.018	0.004
9	0.512	0.019	0.903	0.032	0.023	0.006	0.488	0.018	0.007
10	0.522	0.022	**0.924**	0.036	0.024	0.010	**0.491**	0.017	0.006
Ave	0.420	0.020	0.732	0.034	0.022	0.008	0.421	0.016	0.008

Table 2(b)

Kolmogorov-Smirnov (KS) distances and the mean KS (MKS) between the empirical distributions of the statistics and their proposed distributions, $d = 30$, $\lambda_k = 1_{30} + k(0, 0.1, 0.2, …, 2.9)'$, 2000 replications.

k	CV(λ)	KS				MKS			
		χ_d^2	$L\chi_d^2$	$c\chi_d^2$	$a\chi_b^2$	χ_d^2	$L\chi_d^2$	$c\chi_d^2$	$a\chi_b^2$
1	0.353	0.020	0.899	0.027	0.017	0.005	0.488	0.012	0.005
2	0.444	0.022	0.987	0.047	0.021	0.010	0.500	0.021	0.007
3	0.485	0.013	0.996	0.039	0.011	0.004	0.500	0.020	0.004
4	0.509	0.027	0.999	0.032	0.016	0.009	0.500	0.018	0.006
5	0.525	0.027	1.000	0.036	0.022	**0.016**	0.500	0.019	0.011
6	0.535	0.020	1.000	0.042	0.022	0.007	0.500	0.021	0.008
7	0.543	0.019	1.000	0.043	0.019	0.007	0.500	0.023	0.007
8	0.550	0.021	1.000	0.039	0.013	0.005	0.500	0.020	0.004
9	0.554	**0.028**	1.000	0.034	0.017	0.008	0.500	0.017	0.007
10	0.558	0.015	**1.000**	**0.050**	**0.026**	0.006	**0.500**	**0.024**	**0.013**
Ave	0.506	0.021	0.988	0.039	0.018	0.008	0.499	0.019	0.007

Table 2(c)

Kolmogorov-Smirnov (KS) distances and the mean KS (MKS) between the empirical distributions of the statistics and their proposed distributions, $d = 50$, $\lambda_k = \mathbf{1}_{50} + k(0, 0.1, 0.2, \ldots, 4.9)'$, 2000 replications.

k	CV(λ)	KS				MKS			
		χ_d^2	$L\,\chi_d^2$	$c\chi_d^2$	$a\chi_b^2$	χ_d^2	$L\,\chi_d^2$	$c\chi_d^2$	$a\chi_b^2$
1	0.418	0.015	0.997	0.031	0.010	0.005	0.500	0.015	0.003
2	0.489	0.025	1.000	0.026	0.020	0.009	0.500	0.015	0.005
3	0.518	0.029	1.000	0.043	0.026	0.014	0.500	0.021	0.014
4	0.534	0.018	1.000	0.033	0.017	0.006	0.500	0.017	0.007
5	0.545	0.017	1.000	0.032	0.018	0.006	0.500	0.016	0.006
6	0.551	0.013	1.000	0.040	0.022	0.004	0.500	0.022	0.006
7	0.557	0.010	1.000	0.054	0.016	0.004	0.500	0.024	0.007
8	0.560	0.028	1.000	**0.057**	**0.037**	0.012	0.500	**0.030**	0.013
9	0.563	**0.040**	1.000	0.054	0.028	**0.018**	0.500	0.028	**0.014**
10	0.566	0.015	**1.000**	0.049	0.016	0.005	**0.500**	0.027	0.006
Ave	0.530	0.021	1.000	0.042	0.021	0.008	0.500	0.022	0.008

Table 3

Kolmogorov-Smirnov (KS) distances and the mean KS (MKS) between the empirical distributions of the statistics and their proposed distributions, λ_d = (1, 1.1, 1.2, …, 1 + .1[d − d/10 − 1], 10, 20, …, d)′, 2000 replications.

d	CV(λ)	KS				MKS			
		χ^2_d	$L\chi^2_d$	$c\chi^2_d$	$a\chi^2_b$	χ^2_d	$L\chi^2_d$	$c\chi^2_d$	$a\chi^2_b$
10	1.147	0.022	0.519	0.116	0.062	0.010	0.328	0.067	0.030
20	1.352	0.022	0.838	0.154	0.054	0.008	0.472	0.083	0.029
30	1.462	0.022	0.960	0.157	0.042	0.010	0.498	0.088	0.024
40	1.531	0.017	0.993	0.181	0.051	0.007	0.500	0.098	0.023
50	1.579	0.029	1.000	0.172	0.039	0.014	0.500	0.096	0.018
60	1.613	0.013	1.000	0.171	0.033	0.005	0.500	0.093	0.016
70	1.640	0.018	1.000	0.185	0.034	0.008	0.500	0.100	0.018
80	1.660	0.012	1.000	0.190	0.041	0.004	0.500	0.103	0.024
90	1.677	0.028	1.000	0.191	0.029	0.008	0.500	0.103	0.014
100	1.691	0.019	1.000	0.191	0.029	0.005	0.500	0.107	0.013
Ave	1.535	0.020	0.931	0.171	0.041	0.008	0.480	0.094	0.021

λ_{10} = (1, 1.1, 1.2, …, 1.7, 1.8, 10)′; λ_{20} = (1, 1.1, 1.2, …, 2.6, 2.7, 10, 20)′; …, λ_{100} = (1, 1.1, 1.2, …, 9.7, 9.8, 9.9, 10, 20, 30, …, 100)′.

Table 4(a)

Type I errors of exact chi-square and three approximations to the distributions of quadratic forms when CV(λ) changes, $d = 2$, $\lambda_k = (1, k)'$, 2000 replications.

k	CV(λ)	nominal level $\alpha = .01$				nominal level $\alpha = .025$			
		χ_d^2	$L\chi_d^2$	$c\chi_d^2$	$a\chi_b^2$	χ_d^2	$L\chi_d^2$	$c\chi_d^2$	$a\chi_b^2$
2	0.333	0.016	0.056	0.017	0.014	0.025	0.097	0.028	0.025
3	0.500	0.012	0.107	0.013	0.009	0.027	0.162	0.030	0.021
4	0.600	0.010	0.153	0.018	0.009	0.024	0.211	0.035	0.024
5	0.667	0.010	0.202	0.023	0.012	0.023	0.259	0.039	0.026
6	0.714	0.011	0.250	0.022	0.012	0.026	0.321	0.047	0.027
7	0.750	0.013	0.297	0.029	0.013	0.034	0.368	0.052	0.032
8	0.778	0.014	0.318	0.027	0.015	0.027	0.375	0.050	0.030
9	0.800	0.010	0.320	0.021	0.010	0.027	0.390	0.043	0.022
10	0.818	0.013	0.366	0.021	0.010	0.026	0.432	0.044	0.021
AAD		0.002	0.220	0.011	0.002	0.002	0.265	0.016	0.003

k	CV(λ)	nominal level $\alpha = .05$				nominal level $\alpha = .10$			
		χ_d^2	$L\chi_d^2$	$c\chi_d^2$	$a\chi_b^2$	χ_d^2	$L\chi_d^2$	$c\chi_d^2$	$a\chi_b^2$
2	0.333	0.054	0.149	0.059	0.054	0.108	0.225	0.111	0.107
3	0.500	0.052	0.215	0.059	0.047	0.100	0.303	0.107	0.098
4	0.600	0.048	0.272	0.064	0.048	0.103	0.355	0.111	0.096
5	0.667	0.051	0.324	0.064	0.045	0.095	0.417	0.113	0.095
6	0.714	0.050	0.372	0.073	0.055	0.101	0.457	0.117	0.103
7	0.750	0.060	0.425	0.084	0.061	0.115	0.506	0.133	0.110
8	0.778	0.053	0.441	0.076	0.052	0.091	0.522	0.119	0.101
9	0.800	0.050	0.448	0.074	0.048	0.096	0.521	0.111	0.095
10	0.818	0.043	0.495	0.069	0.047	0.097	0.561	0.116	0.091
AAD		0.003	0.299	0.019	0.004	0.005	0.329	0.015	0.005

Table 4(b)

Type I errors of exact chi-square and three approximations to the distributions of quadratic forms when CV(λ) changes, $d = 6$, $\lambda_k=(1'_3, k1'_3)'$, 2000 replications.

k	CV(λ)	nominal level α =.01				nominal level α =.025			
		χ^2_d	$L\chi^2_d$	$c\chi^2_d$	$a\chi^2_b$	χ^2_d	$L\chi^2_d$	$c\chi^2_d$	$a\chi^2_b$
2	0.333	0.013	0.094	0.019	0.014	0.035	0.155	0.036	0.033
3	0.500	0.018	0.229	0.027	0.018	0.037	0.311	0.043	0.034
4	0.600	0.008	0.339	0.024	0.011	0.024	0.417	0.038	0.026
5	0.667	0.010	0.423	0.019	0.011	0.023	0.507	0.038	0.020
6	0.714	0.010	0.505	0.025	0.013	0.026	0.586	0.047	0.025
7	0.750	0.008	0.568	0.029	0.012	0.020	0.634	0.046	0.028
8	0.778	0.010	0.624	0.022	0.008	0.027	0.686	0.048	0.020
9	0.800	0.010	0.697	0.031	0.010	0.021	0.753	0.054	0.024
10	0.818	0.015	0.708	0.034	0.011	0.029	0.761	0.054	0.028
AAD		0.002	0.455	0.015	0.002	0.005	0.509	0.020	0.004

k	CV(λ)	nominal level α =.05				nominal level α =.10			
		χ^2_d	$L\chi^2_d$	$c\chi^2_d$	$a\chi^2_b$	χ^2_d	$L\chi^2_d$	$c\chi^2_d$	$a\chi^2_b$
2	0.333	0.061	0.226	0.065	0.057	0.114	0.322	0.115	0.106
3	0.500	0.059	0.387	0.068	0.056	0.105	0.480	0.118	0.106
4	0.600	0.047	0.485	0.067	0.047	0.095	0.587	0.120	0.093
5	0.667	0.045	0.581	0.077	0.044	0.095	0.666	0.128	0.101
6	0.714	0.051	0.651	0.078	0.049	0.093	0.721	0.122	0.098
7	0.750	0.044	0.701	0.080	0.047	0.092	0.765	0.136	0.099
8	0.778	0.051	0.741	0.083	0.046	0.105	0.807	0.133	0.098
9	0.800	0.045	0.802	0.081	0.051	0.100	0.843	0.137	0.098
10	0.818	0.049	0.806	0.088	0.051	0.103	0.851	0.136	0.097
AAD		0.005	0.548	0.026	0.004	0.006	0.571	0.027	0.003

Table 4(c)

Type I errors of exact chi-square and three approximations to the distributions of quadratic forms when CV(λ) changes, $d = 10$, $\lambda_k = \left(1'_5, k1'_5\right)'$, 2000 replications.

k	CV(λ)	nominal level $\alpha = .01$				nominal level $\alpha = .025$			
		χ^2_d	$L\,\chi^2_d$	$c\chi^2_d$	$a\chi^2_b$	χ^2_d	$L\,\chi^2_d$	$c\chi^2_d$	$a\chi^2_b$
2	0.333	0.014	0.126	0.020	0.017	0.033	0.194	0.039	0.033
3	0.500	0.011	0.323	0.018	0.012	0.024	0.424	0.037	0.023
4	0.600	0.009	0.468	0.024	0.010	0.026	0.557	0.039	0.028
5	0.667	0.006	0.583	0.018	0.009	0.022	0.660	0.042	0.020
6	0.714	0.011	0.710	0.028	0.009	0.027	0.784	0.051	0.027
7	0.750	0.011	0.765	0.030	0.014	0.028	0.821	0.052	0.028
8	0.778	0.010	0.804	0.028	0.013	0.022	0.851	0.053	0.024
9	0.800	0.011	0.837	0.035	0.014	0.028	0.874	0.066	0.031
10	0.818	0.008	0.876	0.031	0.012	0.020	0.901	0.062	0.026
AAD		0.002	0.600	0.016	0.003	0.003	0.649	0.024	0.003

k	CV(λ)	nominal level $\alpha = .05$				nominal level $\alpha = .10$			
		χ^2_d	$L\,\chi^2_d$	$c\chi^2_d$	$a\chi^2_b$	χ^2_d	$L\,\chi^2_d$	$c\chi^2_d$	$a\chi^2_b$
2	0.333	0.057	0.272	0.064	0.056	0.110	0.388	0.114	0.105
3	0.500	0.050	0.506	0.068	0.052	0.097	0.609	0.118	0.100
4	0.600	0.048	0.639	0.062	0.045	0.094	0.715	0.119	0.087
5	0.667	0.038	0.728	0.074	0.046	0.090	0.797	0.131	0.096
6	0.714	0.056	0.834	0.084	0.052	0.109	0.874	0.140	0.109
7	0.750	0.054	0.865	0.083	0.052	0.108	0.900	0.141	0.099
8	0.778	0.053	0.885	0.085	0.048	0.102	0.916	0.141	0.096
9	0.800	0.055	0.908	0.095	0.059	0.110	0.940	0.150	0.106
10	0.818	0.045	0.930	0.097	0.055	0.097	0.956	0.149	0.108
AAD		0.005	0.679	0.029	0.004	0.007	0.688	0.033	0.006

Table 5(a)

Type I errors of exact chi-square and three approximations to the distributions of quadratic forms when CV(λ) changes, $d = 10$, $\lambda_k = \mathbf{1}_{10} + k(0, 0.1, 0.2, \ldots, 0.9)'$, 2000 replications.

k	CV(λ)	nominal level $\alpha = .01$				nominal level $\alpha = .025$			
		χ^2_d	$L\,\chi^2_d$	$c\chi^2_d$	$a\chi^2_b$	χ^2_d	$L\,\chi^2_d$	$c\chi^2_d$	$a\chi^2_b$
1	0.198	0.014	0.108	0.016	0.014	0.033	0.174	0.036	0.034
2	0.302	0.011	0.279	0.013	0.010	0.024	0.376	0.027	0.023
3	0.367	0.009	0.431	0.015	0.009	0.026	0.520	0.032	0.028
4	0.410	0.006	0.553	0.012	0.009	0.022	0.655	0.027	0.019
5	0.442	0.011	0.711	0.019	0.013	0.027	0.782	0.035	0.026
6	0.466	0.011	0.766	0.019	0.012	0.028	0.824	0.040	0.030
7	0.484	0.010	0.817	0.021	0.011	0.022	0.862	0.037	0.028
8	0.500	0.011	0.849	0.019	0.012	0.028	0.895	0.043	0.030
9	0.512	0.008	0.894	0.017	0.011	0.020	0.924	0.040	0.026
10	0.522	0.008	0.928	0.019	0.009	0.024	0.948	0.037	0.023
AAD		0.002	0.623	0.007	0.002	0.003	0.671	0.010	0.004

k	CV(λ)	nominal level $\alpha = .05$				nominal level $\alpha = .10$			
		χ^2_d	$L\,\chi^2_d$	$c\chi^2_d$	$a\chi^2_b$	χ^2_d	$L\,\chi^2_d$	$c\chi^2_d$	$a\chi^2_b$
1	0.198	0.057	0.249	0.060	0.057	0.110	0.361	0.108	0.107
2	0.302	0.050	0.465	0.054	0.051	0.097	0.590	0.105	0.099
3	0.367	0.048	0.627	0.054	0.047	0.094	0.717	0.097	0.089
4	0.410	0.038	0.729	0.052	0.042	0.090	0.797	0.106	0.092
5	0.442	0.056	0.840	0.067	0.053	0.109	0.893	0.120	0.104
6	0.466	0.054	0.871	0.065	0.051	0.108	0.912	0.121	0.097
7	0.484	0.053	0.903	0.063	0.052	0.102	0.938	0.112	0.095
8	0.500	0.055	0.924	0.072	0.055	0.110	0.950	0.126	0.112
9	0.512	0.045	0.945	0.071	0.057	0.097	0.971	0.124	0.105
10	0.522	0.048	0.962	0.063	0.049	0.108	0.978	0.122	0.100

k	CV(λ)	nominal level $\alpha = .01$				nominal level $\alpha = .025$			
		χ_d^2	$L\chi_d^2$	$c\chi_d^2$	$a\chi_b^2$	χ_d^2	$L\chi_d^2$	$c\chi_d^2$	$a\chi_b^2$
AAD		0.005	0.701	0.012	0.004	0.007	0.710	0.014	0.006

Table 5(b)

Type I errors of exact chi-square and three approximations to the distributions of quadratic forms when CV(λ) changes, $d = 30$, $\lambda_k = \mathbf{1}_{30} + k(0, 0.1, 0.2, \dots, 2.9)'$, 2000 replications.

k	CV(λ)	nominal level $\alpha = .01$				nominal level $\alpha = .025$			
		χ^2_d	$L\,\chi^2_d$	$c\chi^2_d$	$a\chi^2_b$	χ^2_d	$L\,\chi^2_d$	$c\chi^2_d$	$a\chi^2_b$
1	0.353	0.011	0.869	0.014	0.012	0.024	0.915	0.034	0.028
2	0.444	0.009	0.995	0.021	0.015	0.025	0.998	0.043	0.030
3	0.485	0.011	0.999	0.017	0.009	0.024	1.000	0.041	0.025
4	0.509	0.008	1.000	0.017	0.011	0.024	1.000	0.036	0.025
5	0.525	0.012	1.000	0.024	0.013	0.031	1.000	0.050	0.034
6	0.535	0.013	1.000	0.020	0.012	0.030	1.000	0.039	0.026
7	0.543	0.014	1.000	0.022	0.016	0.027	1.000	0.042	0.026
8	0.550	0.011	1.000	0.024	0.013	0.027	1.000	0.042	0.026
9	0.554	0.010	1.000	0.018	0.007	0.023	1.000	0.041	0.025
10	0.558	0.011	1.000	0.017	0.010	0.023	1.000	0.043	0.024
AAD		0.001	0.976	0.009	0.002	0.002	0.966	0.016	0.002

k	CV(λ)	nominal level $\alpha = .05$				nominal level $\alpha = .10$			
		χ^2_d	$L\,\chi^2_d$	$c\chi^2_d$	$a\chi^2_b$	χ^2_d	$L\,\chi^2_d$	$c\chi^2_d$	$a\chi^2_b$
1	0.353	0.051	0.945	0.058	0.052	0.099	0.970	0.109	0.098
2	0.444	0.055	1.000	0.073	0.060	0.111	1.000	0.122	0.107
3	0.485	0.047	1.000	0.067	0.052	0.098	1.000	0.119	0.095
4	0.509	0.046	1.000	0.064	0.043	0.098	1.000	0.123	0.099
5	0.525	0.061	1.000	0.077	0.060	0.116	1.000	0.134	0.108
6	0.535	0.055	1.000	0.071	0.052	0.100	1.000	0.119	0.097
7	0.543	0.050	1.000	0.075	0.053	0.101	1.000	0.136	0.110
8	0.550	0.052	1.000	0.065	0.049	0.099	1.000	0.121	0.097
9	0.554	0.047	1.000	0.071	0.051	0.101	1.000	0.116	0.093
10	0.558	0.045	1.000	0.066	0.048	0.095	1.000	0.109	0.088

k	CV(λ)	nominal level $\alpha = .01$				nominal level $\alpha = .025$			
		χ^2_d	$L\,\chi^2_d$	$c\chi^2_d$	$a\chi^2_b$	χ^2_d	$L\,\chi^2_d$	$c\chi^2_d$	$a\chi^2_b$
AAD		0.004	0.944	0.019	0.004	0.004	0.897	0.021	0.006

Table 5(c)

Type I errors of exact chi-square and three approximations to the distributions of quadratic forms when CV(λ) changes, $d = 50$, $\lambda_k = \mathbf{1}_{50} + k(0, 0.1, 0.2, \ldots, 4.9)'$, 2000 replications.

k	CV(λ)	nominal level $\alpha = .01$				CV(λ)	nominal level $\alpha = .025$			
		χ^2_d	$L\,\chi^2_d$	$c\chi^2_d$	$a\chi^2_b$		χ^2_d	$L\,\chi^2_d$	$c\chi^2_d$	$a\chi^2_b$
1	0.418	0.010	0.999	0.019	0.014	0.418	0.023	0.999	0.037	0.025
2	0.489	0.010	1.000	0.020	0.011		0.028	1.000	0.041	0.029
3	0.518	0.010	1.000	0.027	0.013		0.031	1.000	0.045	0.031
4	0.534	0.016	1.000	0.022	0.015		0.032	1.000	0.043	0.029
5	0.545	0.012	1.000	0.021	0.013		0.027	1.000	0.045	0.027
6	0.551	0.009	1.000	0.017	0.011		0.024	1.000	0.039	0.020
7	0.557	0.012	1.000	0.024	0.013		0.032	1.000	0.043	0.027
8	0.560	0.008	1.000	0.018	0.009		0.019	1.000	0.035	0.022
9	0.563	0.013	1.000	0.027	0.013		0.032	1.000	0.054	0.031
10	0.566	0.010	1.000	0.019	0.010		0.024	1.000	0.040	0.022
AAD		0.002	0.990	0.011	0.002		0.004	0.975	0.017	0.004

k	CV(λ)	nominal level $\alpha = .05$				CV(λ)	nominal level $\alpha = .10$			
		χ^2_d	$L\,\chi^2_d$	$c\chi^2_d$	$a\chi^2_b$		χ^2_d	$L\,\chi^2_d$	$c\chi^2_d$	$a\chi^2_b$
1	0.418	0.047	1.000	0.063	0.046	0.418	0.097	1.000	0.120	0.106
2	0.489	0.050	1.000	0.071	0.055		0.106	1.000	0.120	0.098
3	0.518	0.061	1.000	0.078	0.057		0.109	1.000	0.135	0.111
4	0.534	0.054	1.000	0.074	0.053		0.102	1.000	0.122	0.100
5	0.545	0.052	1.000	0.070	0.050		0.098	1.000	0.129	0.101
6	0.551	0.051	1.000	0.073	0.049		0.097	1.000	0.121	0.097
7	0.557	0.056	1.000	0.074	0.048		0.104	1.000	0.124	0.100
8	0.560	0.044	1.000	0.077	0.043		0.092	1.000	0.125	0.105
9	0.563	0.062	1.000	0.088	0.062		0.113	1.000	0.143	0.116
10	0.566	0.053	1.000	0.073	0.045		0.107	1.000	0.130	0.100

k	CV(λ)	nominal level $\alpha = .01$				nominal level $\alpha = .025$			
		χ_d^2	$L\chi_d^2$	$c\chi_d^2$	$a\chi_b^2$	χ_d^2	$L\chi_d^2$	$c\chi_d^2$	$a\chi_b^2$
AAD		0.005	0.950	0.024	0.005	0.006	0.900	0.027	0.004

Table 6

Type I errors of exact chi-square and three approximations to the distributions of quadratic forms when d and $CV(\lambda)$ changes, $\lambda_d = (1, 1.1, 1.2, \ldots, 1 + .1[d - d/10 - 1], 10, 20, \ldots, d)'$, 2000 replications.

d	$CV(\lambda)$	nominal level $\alpha = .01$				nominal level $\alpha = .025$			
		χ^2_d	$L\chi^2_d$	$c\chi^2_d$	$a\chi^2_b$	χ^2_d	$L\chi^2_d$	$c\chi^2_d$	$a\chi^2_b$
10	1.147	0.014	0.358	0.048	0.014	0.033	0.442	0.073	0.030
20	1.352	0.012	0.797	0.066	0.019	0.026	0.850	0.091	0.032
30	1.462	0.009	0.967	0.072	0.020	0.025	0.981	0.098	0.035
40	1.531	0.010	0.998	0.091	0.019	0.030	1.000	0.124	0.038
50	1.579	0.010	1.000	0.092	0.015	0.031	1.000	0.121	0.031
60	1.613	0.009	1.000	0.101	0.019	0.027	1.000	0.137	0.033
70	1.640	0.006	1.000	0.101	0.018	0.024	1.000	0.135	0.037
80	1.660	0.008	1.000	0.090	0.021	0.019	1.000	0.122	0.036
90	1.677	0.011	1.000	0.109	0.018	0.027	1.000	0.145	0.036
100	1.691	0.017	1.000	0.103	0.018	0.029	1.000	0.140	0.035
AAD		0.002	0.902	0.077	0.008	0.003	0.902	0.094	0.009

d	$CV(\lambda)$	nominal level $\alpha = .05$				nominal level $\alpha = .10$			
		χ^2_d	$L\chi^2_d$	$c\chi^2_d$	$a\chi^2_b$	χ^2_d	$L\chi^2_d$	$c\chi^2_d$	$a\chi^2_b$
10	1.147	0.057	0.512	0.103	0.048	0.110	0.609	0.143	0.094
20	1.352	0.052	0.886	0.121	0.048	0.098	0.924	0.162	0.090
30	1.462	0.055	0.987	0.136	0.052	0.111	0.995	0.181	0.090
40	1.531	0.055	1.000	0.154	0.060	0.102	1.000	0.201	0.100
50	1.579	0.061	1.000	0.154	0.052	0.109	1.000	0.203	0.097
60	1.613	0.046	1.000	0.166	0.057	0.106	1.000	0.207	0.103
70	1.640	0.049	1.000	0.164	0.057	0.105	1.000	0.212	0.102
80	1.660	0.041	1.000	0.164	0.058	0.092	1.000	0.209	0.088
90	1.677	0.057	1.000	0.174	0.062	0.106	1.000	0.223	0.107
100	1.691	0.053	1.000	0.171	0.057	0.108	1.000	0.224	0.099

| d | CV(λ) | nominal level $\alpha = .01$ | | | | nominal level $\alpha = .025$ | | | |
		χ_d^2	$L\chi_d^2$	$c\chi_d^2$	$a\chi_b^2$	χ_d^2	$L\chi_d^2$	$c\chi_d^2$	$a\chi_b^2$
AAD		0.005	0.888	0.100	0.006	0.007	0.853	0.096	0.005

$\lambda_{10} = (1, 1.1, 1.2, \ldots, 1.7, 1.8, 10)'; \lambda_{20} = (1, 1.1, 1.2, \ldots, 2.6, 2.7, 10, 20)'; \ldots, \lambda_{100} = (1, 1.1, 1.2, \ldots, 9.7, 9.8, 9.9, 10, 20, 30, \ldots, 100)'.$