Behaviormetrika
Vol.33, No.1, 2006, 3–26

APPROXIMATIONS TO THE DISTRIBUTION OF THE SAMPLE COEFFICIENT ALPHA UNDER NONNORMALITY

Haruhiko Ogasawara*

Approximate distributions of the sample coefficient alpha under nonnormality as well as normality are derived by using the single- and two-term Edgeworth expansions up to the term of order $1/n$. The case of the standardized coefficient alpha including the weights for the components of a test is also considered. From the numerical illustration with simulation using the normal and typical nonnormal distributions with different types/degrees of nonnormality, it is shown that the variances of the sample coefficient alpha under nonnormality can be grossly different from those under normality. The corresponding biases and skewnesses are shown to be negative under various conditions. The method of developing confidence intervals of the population coefficient alpha using the Cornish-Fisher expansion with sample cumulants is presented.

1. Introduction

Cronbach's (1951) alpha is one of the coefficients for reliability or internal consistency of a test consisting of multiple components (e.g., subtests). It is known that coefficient alpha gives a lower bound for the reliability of the test, where reliability is defined by the proportion of the variance of the true score of the test with uncorrelated errors to the total variance. The lower bound becomes the true reliability if and only if the components are essentially tau-equivalent or equivalently the off-diagonal elements of the covariance matrix of the component scores are the same. Although this pattern may not be obtained in practice, the simplicity of the definition of coefficient alpha has attracted many researchers in test theory or psychometrics (for the actual differences of coefficient alpha and its corresponding true reliability in typical situations, see Osburn, 2000).

In practice, coefficient alpha is only available as its estimate. The distribution of the sample coefficient alpha was given by Kristof (1963) and Feldt (1965) with the assumption of the pattern of compound symmetry for the population covariance matrix of the component scores under normality. The term compound symmetry indicates that the set of the off-diagonal elements of a covariance matrix and that of the diagonal elements have the same values in each set. Van Zyl, Neudecker and Nel (2000) relaxed the assumption of the compound symmetry. Their result, however, was based on the asymptotic normality of the sample coefficient alpha under normality. That is, they derived the asymptotic variance of order $O(n^{-1})$ for the sample coefficient alpha under the normality for component scores, where $n + 1 = N$ is the number of independent observations or subjects. Under the same condition as in van Zyl et al. (2000), Kistner and Muller (2004) gave the exact

Key Words and Phrases: Cronbach's alpha, tau-equivalence, Edgeworth expansion, Cornish-Fisher expansion, nonnormality, cumulant

* Department of Information and Management Science, Otaru University of Commerce, 3–5–21, Midori, Otaru 047–8501 Japan. E-mail: hogasa@res.otaru-uc.ac.jp

Partially supported by Grant-in-Aid for Scientific Research from the Japanese Ministry of Education, Culture, Sports, Science and Technology.

and approximate distributions of the sample coefficient alpha.

In the case of compound symmetry under normality, Kristof (1963, Equation (43)) derived the exact bias of the sample coefficient alpha. Using the equivalent result van Zyl et al. (2000, Equation (11)) concluded that the sample coefficient alpha is asymptotically unbiased. This is mathematically correct since the asymptotic standard error is of order $O(n^{-1/2})$ while the asymptotic bias is of order $O(n^{-1})$, which can be neglected when n is large. However, in a later section, we will show that the bias is substantial especially when the sample size is relatively small. Yuan and Bentler (2002) showed that the normal theory asymptotic standard error derived by van Zyl et al. (2000) holds for some family of nonnormal distributions irrespective of the violation of normality. It is often pointed out that in practice data are seldom, if ever, normally distributed, which gives an advantage to Yuan and Bentler (2002). However, it is not clear whether the family of nonnormal distributions with such robustness shown by Yuan and Bentler (2002) can approximate typical nonnormal distributions observed in practice or used in simulation. Further, we have a problem for the asymptotic standard error in finite samples, even when normality is satisfied, since the asymptotic approximation is not necessarily tolerable in relatively small samples (see van Zyl et al., 2000, Table 1).

In this article, we will derive the asymptotic approximations to the distributions of the sample coefficient alpha under arbitrary distributions with relatively mild regularity conditions using Edgeworth expansions up to order $O(n^{-1})$, We will show that the approximation by the usual asymptotic standard error given by van Zyl et al. (2000) will be substantially improved by the asymptotic expansions under normality and nonnormality. Further, it will be shown that the distributions of the sample coefficient alpha under typical nonnormal distributions are considerably different from that under normality, when nonnormality is not mild, by using theoretical and simulated distributions.

2. Asymptotic expansion of the distribution of the sample coefficient alpha for unstandardized variables

Let α be the population coefficient alpha for a test with p component scores whose population covariance matrix is denoted by $\mathbf{\Sigma}$. Then,

$$\alpha = \frac{p}{p-1}\left\{1 - \frac{\mathrm{tr}(\mathbf{\Sigma})}{\mathbf{1}_p{}'\mathbf{\Sigma}\,\mathbf{1}_p}\right\},\quad (p \geq 2), \tag{1}$$

where $\mathbf{1}_p$ is the $p \times 1$ vector whose elements are 1's (in the remainder of this article $p \geq 2$ is assumed). The sample counterpart of α is given by

$$\hat{\alpha} = \frac{p}{p-1}\left\{1 - \frac{\mathrm{tr}(\mathbf{S})}{\mathbf{1}_p{}'\mathbf{S}\,\mathbf{1}_p}\right\}, \tag{2}$$

where $\mathbf{S}$ is the usual unbiased sample covariance matrix for p components based on N observations. Coefficient alpha can be generalized to the case with fixed weights $\mathbf{w} = (w_1, ..., w_p)'$ for the component scores (see ten Berge & Hofstee, 1999 for the cases with weights being functions of $\mathbf{\Sigma}$ or $\mathbf{S}$). However, the case with given weights can be treated without weights when $\mathbf{\Sigma}$ and $\mathbf{S}$ are replaced, and redefined, by $\mathrm{diag}(\mathbf{w})\,\mathbf{\Sigma}\,\mathrm{diag}(\mathbf{w})$ and $\mathrm{diag}(\mathbf{w})\,\mathbf{S}\,\mathrm{diag}(\mathbf{w})$, respectively, where $\mathrm{diag}(\mathbf{w})$ is the diagonal matrix whose diagonal elements are the elements of $\mathbf{w}$. Consequently, only the case with (1) and (2) is dealt with in this section without loss of generality. Let

$$v = n^{1/2}(\hat{\alpha} - \alpha), \tag{3}$$

where v depends on n, but the subscript n in v is omitted for simplicity of notation. Suppose that the following general results for the cumulants of the functions of $\mathbf{S}$ are available with the assumption of the existence of the finite moments up to the required order (see e.g., Hall, 1992, pp.46–47):

$$\begin{aligned} \kappa_1(v) &= \mathrm{E}(v) = n^{-1/2}a_1 + O(n^{-3/2}), \\ \kappa_2(v) &= \mathrm{E}[\{v - \mathrm{E}(v)\}^2] = a_2 + n^{-1}\Delta a_2 + O(n^{-2}), \\ \kappa_3(v) &= \mathrm{E}[\{v - \mathrm{E}(v)\}^3] = n^{-1/2}a_3 + O(n^{-3/2}), \\ \kappa_4(v) &= \mathrm{E}[\{v - \mathrm{E}(v)\}^4] - 3\{\kappa_2(v)\}^2 = n^{-1}a_4 + O(n^{-2}), \end{aligned} \tag{4}$$

where a_1/n, a_2/n, $\Delta a_2/n^2$, a_3/n^2 and a_4/n^3 are the asymptotic bias, asymptotic variance, higher-order asymptotic (added) variance with (first-order) bias-correction, third asymptotic cumulant and fourth asymptotic cumulant of $\hat{\alpha}$, respectively.

Then, it is known that under its validity the asymptotic distribution function of the transformed v or $\hat{\alpha}$ is given by Edgeworth expansion as:

$$\begin{aligned} \Pr\left(\frac{v}{a_2^{1/2}} \le x\right) &= \Phi(x) - n^{-1/2}\left\{\frac{a_1}{a_2^{1/2}} + \frac{a_3}{6a_2^{3/2}}(x^2 - 1)\right\}\phi(x) \\ -n^{-1}&\left\{\frac{1}{2}(\Delta a_2 + a_1^2)\frac{x}{a_2} + \left(\frac{a_4}{24} + \frac{a_1 a_3}{6}\right)\frac{x^3 - 3x}{a_2^2} + \frac{a_3^2(x^5 - 10x^3 + 15x)}{72a_2^3}\right\}\phi(x) + o(n^{-1}) \end{aligned} \tag{5}$$

(see e.g., Fujikoshi, 1980, p.46; Hall, 1992, Theorem 2.2; Ogasawara, 2006), where $\Phi(\cdot)$ and $\phi(\cdot)$ are the distribution and density functions of the standard normal distribution, respectively. The asymptotic density function corresponding to (5) is given as

$$\begin{aligned} f\left(\frac{v}{a_2^{1/2}} = x\right) &= \left[1 + n^{-1/2}\left\{\frac{a_1 x}{a_2^{1/2}} + \frac{a_3}{6a_2^{3/2}}(x^3 - 3x)\right\} + n^{-1}\left\{\frac{1}{2}(\Delta a_2 + a_1^2)\frac{x^2 - 1}{a_2}\right.\right. \\ &\left.\left. + \left(\frac{a_4}{24} + \frac{a_1 a_3}{6}\right)\frac{x^4 - 6x^2 + 3}{a_2^2} + \frac{a_3^2(x^6 - 15x^4 + 45x^2 - 15)}{72a_2^3}\right\}\right]\phi(x) + o(n^{-1}). \end{aligned} \tag{6}$$

Note that in (5) and (6), we have three approximations i.e., those up to $O(1)$, $O(n^{-1/2})$ and $O(n^{-1})$. The first approximations in (5) and (6) are simply $\Phi(x)$ and $\phi(x)$, respectively, which correspond to the usual results using the asymptotic normality of $\hat{\alpha}$ and its (first-order) asymptotic standard error $(a_2/n)^{1/2}$. The second and third approximations up to $O(n^{-1/2})$ and $O(n^{-1})$ are those by the single- and two-term Edgeworth expansions, respectively.

3. Asymptotic cumulants

The constants a_1 to a_4 and Δa_2 associated with the asymptotic cumulants in (5) and (6) are given from the Taylor expansion of v about its true value with the assumption of the existence of the moments of observable variables up to the required order:

$$v = \frac{\partial \hat{\alpha}}{\partial \mathbf{s}'}|_{\mathbf{s}=\boldsymbol{\sigma}}\mathbf{u} + \frac{n^{-1/2}}{2}\left\{\left(\frac{\partial}{\partial \mathbf{s}'}\right)^{\langle 2\rangle}|_{\mathbf{s}=\boldsymbol{\sigma}}\hat{\alpha}\right\}\mathbf{u}^{\langle 2\rangle} + \frac{n^{-1}}{6}\left\{\left(\frac{\partial}{\partial \mathbf{s}'}\right)^{\langle 3\rangle}|_{\mathbf{s}=\boldsymbol{\sigma}}\hat{\alpha}\right\}\mathbf{u}^{\langle 3\rangle} + O_p(n^{-3/2}), \quad (7)$$

where $\mathbf{s} = \mathrm{ve}(\mathbf{S})$; $\boldsymbol{\sigma} = \mathrm{ve}(\boldsymbol{\Sigma})$; $\mathrm{ve}(\cdot)$ is the vectorizing operator taking nonduplicated elements of a symmetric matrix; $\mathbf{u} = n^{1/2}(\mathbf{s} - \boldsymbol{\sigma})$; $\mathbf{s}$ and $\hat{\alpha}$ are also used as variables, which were estimates or estimators, for simplicity of notation; and $\mathbf{x}^{\langle k\rangle}$ denotes the k-fold Kronecker product of $\mathbf{x}$ i.e., $\mathbf{x}^{\langle k\rangle} = \mathbf{x}\otimes\cdots\otimes\mathbf{x}$ (k times of $\mathbf{x}$).

From (7), the following are easily obtained:

$$a_1 = \frac{1}{2}\mathrm{tr}\left(\frac{\partial^2 \alpha}{\partial \boldsymbol{\sigma}\,\partial \boldsymbol{\sigma}'}\boldsymbol{\Omega}\right), \; a_2 = \frac{\partial \alpha}{\partial \boldsymbol{\sigma}'}\boldsymbol{\Omega}\frac{\partial \alpha}{\partial \boldsymbol{\sigma}}, \quad (8)$$

where $\partial\alpha/\partial\boldsymbol{\sigma}$ denotes $\partial\hat{\alpha}/\partial\mathbf{s}$ evaluated at $\mathbf{s} = \boldsymbol{\sigma}$ for simplicity of notation with other derivatives defined similarly;

$$n\left\{\mathrm{acov}(n^{-1/2}\mathbf{u})\right\}_{ab,cd} = (\boldsymbol{\Omega})_{ab,cd} = \sigma_{abcd} - \sigma_{ab}\sigma_{cd}, \quad (9)$$

$$(p \geq a \geq b \geq 1; p \geq c \geq d \geq 1)$$

using double subscript notation; $\mathrm{acov}(\cdot)$ denotes the asymptotic covariance matrix of order $O(n^{-1})$ for an argument vector; $\sigma_{a\ldots d}$ is the multivariate central moment of the variables $X_a, ..., X_d$, whose special case is $\sigma_{ab} = (\boldsymbol{\Sigma})_{ab}$; and $(\cdot)_{ab}$ is the element of the a-th row and b-th column of the argument matrix.

For the other asymptotic cumulants, from Ogasawara (2006), we have

$$\begin{aligned}\Delta a_2 = &-\sum_{a\geq b}\sum_{c\geq d}\frac{\partial\alpha}{\partial\sigma_{ab}}\frac{\partial\alpha}{\partial\sigma_{cd}}(\sigma_{abcd} - \sigma_{ab}\sigma_{cd} - \sigma_{ac}\sigma_{bd} - \sigma_{ad}\sigma_{bc})\\ &+\sum_{a\geq b}\sum_{c\geq d}\sum_{e\geq f}\left\{\frac{\partial\alpha}{\partial\sigma_{ab}}\frac{\partial^2\alpha}{\partial\sigma_{cd}\partial\sigma_{ef}}(\sigma_{abcdef} - \sigma_{ab}\sigma_{cdef} - 2\sigma_{cd}\sigma_{abef}\right.\\ &\qquad - 2\sigma_{acd}\sigma_{bef} - 2\sigma_{abc}\sigma_{def} - 2\sigma_{abd}\sigma_{cef} + 2\sigma_{ab}\sigma_{cd}\sigma_{ef})\\ &\left.+\sum_{g\geq h}\left(\frac{1}{2}\frac{\partial^2\alpha}{\partial\sigma_{ab}\partial\sigma_{ef}}\frac{\partial^2\alpha}{\partial\sigma_{cd}\partial\sigma_{gh}} + \frac{\partial\alpha}{\partial\sigma_{gh}}\frac{\partial^3\alpha}{\partial\sigma_{ab}\partial\sigma_{cd}\partial\sigma_{ef}}\right)(\boldsymbol{\Omega})_{ab,cd}(\boldsymbol{\Omega})_{ef,gh}\right\},\end{aligned}$$

$$a_3 = \sum_{a\geq b}\sum_{c\geq d}\sum_{e\geq f}\frac{\partial\alpha}{\partial\sigma_{ab}}\frac{\partial\alpha}{\partial\sigma_{cd}}\frac{\partial\alpha}{\partial\sigma_{ef}}(\sigma_{abcdef} - 3\sigma_{ab}\sigma_{cdef}$$

$$- 6\sigma_{abc}\sigma_{def} + 2\sigma_{ab}\sigma_{cd}\sigma_{ef}) + 3\frac{\partial\alpha}{\partial\boldsymbol{\sigma}'}\boldsymbol{\Omega}\frac{\partial^2\alpha}{\partial\boldsymbol{\sigma}\,\partial\boldsymbol{\sigma}'}\boldsymbol{\Omega}\frac{\partial\alpha}{\partial\boldsymbol{\sigma}}, \tag{10}$$

$$\begin{aligned} a_4 = &\sum_{a\ge b}\sum_{c\ge d}\sum_{e\ge f}\sum_{g\ge h}\Bigg[\frac{\partial\alpha}{\partial\sigma_{ab}}\frac{\partial\alpha}{\partial\sigma_{cd}}\frac{\partial\alpha}{\partial\sigma_{ef}}\frac{\partial\alpha}{\partial\sigma_{gh}}\Bigg(\kappa_{abcdefgh} + \sum^{24}\kappa_{ac}\kappa_{bdefgh} + \sum^{32}\kappa_{ace}\kappa_{bdfgh} \\ &+ \sum^{8}\kappa_{aceg}\kappa_{bdfh} + \sum^{24}\kappa_{abeg}\kappa_{cdfh} + \sum^{96}\kappa_{ac}\kappa_{be}\kappa_{dfgh} + \sum^{48}\kappa_{ac}\kappa_{eg}\kappa_{bdfh} \\ &+ \sum^{96}\kappa_{ac}\kappa_{beg}\kappa_{dfh} + \sum^{48}\kappa_{bc}\kappa_{de}\kappa_{fg}\kappa_{ha} - \sum^{6}\kappa_{abcd}M(ef,gh)\Bigg) \\ &+ \sum_{j\ge k} 2\frac{\partial^2\alpha}{\partial\sigma_{ab}\partial\sigma_{cd}}\frac{\partial\alpha}{\partial\sigma_{ef}}\frac{\partial\alpha}{\partial\sigma_{gh}}\frac{\partial\alpha}{\partial\sigma_{jk}}\sum^{10}M(ab,cd)M(ef,gh,jk) \\ &+ \sum_{j\ge k}\sum_{l\ge m}\left(\frac{3}{2}\frac{\partial^2\alpha}{\partial\sigma_{ab}\partial\sigma_{cd}}\frac{\partial^2\alpha}{\partial\sigma_{ef}\partial\sigma_{gh}} + \frac{2}{3}\frac{\partial^3\alpha}{\partial\sigma_{ab}\partial\sigma_{cd}\partial\sigma_{ef}}\frac{\partial\alpha}{\partial\sigma_{gh}}\right) \\ &\times\frac{\partial\alpha}{\partial\sigma_{jk}}\frac{\partial\alpha}{\partial\sigma_{lm}}\sum^{15}M(ab,cd)M(ef,gh)M(jk,lm)\Bigg] \\ &- (4a_1a_3 + 6a_2\Delta a_2 + 6a_2a_1^2), \end{aligned}$$

where $\sum_{a\ge b}$ is the summation over the range $p \ge a \ge b \ge 1$; $\kappa_{a\ldots h}$ is the multivariate cumulant of the variables $X_a, ..., X_h$;

$$\begin{aligned} &M(ab,cd) = \kappa_{abcd} + \kappa_{ac}\kappa_{bd} + \kappa_{ad}\kappa_{bc} = (\boldsymbol{\Omega})_{ab,cd}, \\ &M(ab,cd,ef) = \kappa_{abcdef} + \sum^{12}\kappa_{abce}\kappa_{df} + \sum^{4}\kappa_{ace}\kappa_{bdf} + \sum^{8}\kappa_{ac}\kappa_{be}\kappa_{df}, \\ &(p \ge a \ge b \ge 1 : p \ge c \ge d \ge 1; p \ge e \ge f \ge 1); \end{aligned} \tag{11}$$

$\sum^{k}$ is the sum of the k products with similar patterns.

Though (10) with (11) looks involved, we should note that for the first approximation by the usual asymptotic standard error, only a_2 in (8) is necessary, for the second one by the single-term Edgeworth expansion a_1, a_2 and a_3 are required, and for the third one by the two-term Edgeworth expansion, all of a_1 to a_4 and Δa_2 are needed. As will be shown later, the single-term Edgeworth expansion has substantial improvement of approximation over the usual one by the asymptotic standard error. So, the computation of a_1 and a_3 in addition to a_2 is of practical value.

In (8) and (10), the partial derivatives of $\hat{\alpha}$ with respect to $s_{ab} = (\mathbf{S})_{ab}$ up to the third order are involved. Recall that for the above first, second and third approximations, the partial derivatives up to the first, second and third orders are required, respectively. The partial derivatives are derived straightforwardly since $\hat{\alpha}$ is an elementary function of $\mathbf{s}$.

Let

$$\sigma_{\mathrm{D}} = \mathrm{tr}(\boldsymbol{\Sigma}) \quad \text{and} \quad \sigma_{\mathrm{T}} = \mathbf{1}_p{}'\boldsymbol{\Sigma}\,\mathbf{1}_p. \tag{12}$$

Then, noting $\alpha = \{p/(p-1)\}\{1-(\sigma_{\mathrm{D}}/\sigma_{\mathrm{T}})\}$, we have

$$\begin{aligned}\frac{p-1}{p}\frac{\partial\alpha}{\partial\boldsymbol{\sigma}'} &= \frac{\sigma_{\mathrm{D}}}{\sigma_{\mathrm{T}}^2}\mathbf{1}_{p^2}{}'\frac{\mathrm{vec}(\boldsymbol{\Sigma})}{\partial\boldsymbol{\sigma}'} - \frac{1}{\sigma_{\mathrm{T}}}\mathbf{1}_{p^2}{}'\frac{\mathrm{vec}\{\mathrm{Diag}(\boldsymbol{\Sigma})\}}{\partial\boldsymbol{\sigma}'}\\ &= \frac{\sigma_{\mathrm{D}}}{\sigma_{\mathrm{T}}^2}\mathbf{1}_{p^2}{}'\mathbf{D}_p - \frac{1}{\sigma_{\mathrm{T}}}\mathbf{1}_{p^2}{}'\,\mathbf{C}_p, \end{aligned} \tag{13}$$

where vec(·) is the vectorizing operator stacking the columns of a matrix sequentially whose top is the first column; $\mathbf{D}_p$ and $\mathbf{C}_p$ are the matrices with $\mathrm{vec}(\boldsymbol{\Sigma}) = \mathbf{D}_p\boldsymbol{\sigma}$ and $\mathrm{vec}\{\mathrm{Diag}(\boldsymbol{\Sigma})\} = \mathbf{C}_p\boldsymbol{\sigma}$; and Diag(·) is the diagonal matrix whose diagonal elements are given by those of the argument matrix. The elementwise expression corresponding to (13) is

$$\frac{p-1}{p}\frac{\partial\alpha}{\partial\sigma_{ab}} = \frac{\sigma_{\mathrm{D}}}{\sigma_{\mathrm{T}}^2}(2-\delta_{ab}) - \frac{\delta_{ab}}{\sigma_{\mathrm{T}}},\ (p \geq a \geq b \geq 1), \tag{14}$$

where δ_{ab} is the Kronecker delta.

The second and third partial derivatives with matrix and elementwise expression are given from (13) and (14) as follows:

$$\begin{aligned}
&\frac{p-1}{p}\frac{\partial^2\alpha}{\partial\boldsymbol{\sigma}\,\partial\boldsymbol{\sigma}'} = -\frac{2\sigma_{\mathrm{D}}}{\sigma_{\mathrm{T}}^3}\mathbf{D}_p{}'\mathbf{1}_{p^2}\mathbf{1}_{p^2}{}'\mathbf{D}_p + \frac{1}{\sigma_{\mathrm{T}}^2}(\mathbf{C}_p{}'\mathbf{1}_{p^2}\mathbf{1}_{p^2}{}'\mathbf{D}_p + \mathbf{D}_p{}'\mathbf{1}_{p^2}\mathbf{1}_{p^2}{}'\mathbf{C}_p),\\
&\frac{p-1}{p}\frac{\partial^2\alpha}{\partial\sigma_{ab}\partial\sigma_{cd}} = -\frac{2\sigma_{\mathrm{D}}}{\sigma_{\mathrm{T}}^3}(2-\delta_{ab})(2-\delta_{cd}) + \frac{1}{\sigma_{\mathrm{T}}^2}\{(2-\delta_{ab})\delta_{cd} + \delta_{ab}(2-\delta_{cd})\},\\
&\frac{p-1}{p}\left(\frac{\partial}{\partial\boldsymbol{\sigma}'}\right)^{\langle 3\rangle}\alpha = \frac{6\sigma_{\mathrm{D}}}{\sigma_{\mathrm{T}}^4}(\mathbf{1}_{p^2}{}'\mathbf{D}_p)^{\langle 3\rangle} - \frac{2}{\sigma_{\mathrm{T}}^3}\{(\mathbf{1}_{p^2}{}'\mathbf{D}_p)^{\langle 2\rangle}\otimes\mathbf{1}_{p^2}{}'\mathbf{C}_p\\
&\qquad + (\mathbf{1}_{p^2}{}'\mathbf{D}_p)\otimes(\mathbf{1}_{p^2}{}'\mathbf{C}_p)\otimes(\mathbf{1}_{p^2}{}'\mathbf{D}_p) + (\mathbf{1}_{p^2}{}'\mathbf{C}_p)\otimes(\mathbf{1}_{p^2}{}'\mathbf{D}_p)^{\langle 2\rangle}\},\\
&\frac{p-1}{p}\frac{\partial^3\alpha}{\partial\sigma_{ab}\partial\sigma_{cd}\partial\sigma_{ef}} = \frac{6\sigma_{\mathrm{D}}}{\sigma_{\mathrm{T}}^4}(2-\delta_{ab})(2-\delta_{cd})(2-\delta_{ef})\\
&-\frac{2}{\sigma_{\mathrm{T}}^3}\{(2-\delta_{ab})(2-\delta_{cd})\delta_{ef} + (2-\delta_{ab})\delta_{cd}(2-\delta_{ef}) + \delta_{ab}(2-\delta_{cd})(2-\delta_{ef})\},\\
&(p \geq a \geq b \geq 1; p \geq c \geq d \geq 1; p \geq e \geq f \geq 1).
\end{aligned} \tag{15}$$

From (15) we see that the matrix expressions do not necessarily give simpler results than the corresponding elementwise ones. The above results are summarized as:

Theorem 1. The asymptotic approximation of the distribution of $\hat{\alpha}$ under nonnormality with the assumptions of the existence of the finite moments of the component scores of a test up to the eighth order and the validity of the two-term Edgeworth expansion is given by (5) and (6) with the asymptotic cumulants of (8) and (10), and the partial derivatives of (12) through (15).

When the moments higher than the eighth order do not exist, the order of the remainder term for $\kappa_4(v)$ in (4) should be replaced by $o(n^{-1})$.

4. Asymptotic variance and bias under normality

When the component scores of a test are normally distributed, simplified results are obtained. Let $\mathrm{avar}(\hat{\alpha})_{\mathrm{NT}}$ be the asymptotic variance of $\hat{\alpha}$ of order $O(n^{-1})$ under normality. Then, we have

Corollary 1 (van Zyl et al., 2000).

$$n\,\mathrm{avar}(\hat{\alpha})_{\mathrm{NT}} = \frac{2p^2}{(p-1)^2\sigma_{\mathrm{T}}^3}[\sigma_{\mathrm{T}}\{\sigma_{\mathrm{D}}^2 + \mathrm{tr}(\boldsymbol{\Sigma}^2)\} - 2\sigma_{\mathrm{D}}\mathbf{1}_p{}'\boldsymbol{\Sigma}^2\mathbf{1}_p]. \tag{16}$$

Proof. The proof is given in a way somewhat different from van Zyl et al.'s (2000) method since the intermediate results will be used later. Under normality, it is known that $\boldsymbol{\Omega} = 2\mathbf{D}_p^+(\boldsymbol{\Sigma}\otimes\boldsymbol{\Sigma})\mathbf{D}_p^{+}{}'$, where $\mathbf{D}_p^+$ is the Moore-Penrose g-inverse of $\mathbf{D}_p$ (see e.g., Browne, 1974). Then, from (8) and (13), we have

$$n\,\mathrm{avar}(\hat{\alpha})_{\mathrm{NT}} = \frac{2p^2}{(p-1)^2\sigma_{\mathrm{T}}^4}\mathbf{1}_{p^2}{}'(\sigma_{\mathrm{D}}\mathbf{D}_p - \sigma_{\mathrm{T}}\mathbf{C}_p)\mathbf{D}_p^+(\boldsymbol{\Sigma}\otimes\boldsymbol{\Sigma})\mathbf{D}_p^{+}{}'(\sigma_{\mathrm{D}}\mathbf{D}_p{}' - \sigma_{\mathrm{T}}\mathbf{C}_p{}')\mathbf{1}_{p^2}. \tag{17}$$

Let $\mathbf{e}_p = \mathrm{vec}(\mathbf{I}_p)$ with $\mathbf{I}_p$ being the $p\times p$ identity matrix, then noting $\mathbf{1}_{p^2}{}'\mathbf{D}_p\mathbf{D}_p^+ = \mathbf{1}_{p^2}{}'$ and $\mathbf{1}_{p^2}{}'\mathbf{C}_p\mathbf{D}_p^+ = \mathbf{e}_p{}'$ (17) becomes

$$= \frac{2p^2}{(p-1)^2\sigma_{\mathrm{T}}^4}(\sigma_{\mathrm{D}}\mathbf{1}_{p^2}{}' - \sigma_{\mathrm{T}}\mathbf{e}_p{}')(\boldsymbol{\Sigma}\otimes\boldsymbol{\Sigma})(\sigma_{\mathrm{D}}\mathbf{1}_{p^2} - \sigma_{\mathrm{T}}\mathbf{e}_p). \tag{18}$$

In (18), noting

$$\begin{aligned}
&\mathbf{1}_{p^2}{}'(\boldsymbol{\Sigma}\otimes\boldsymbol{\Sigma})\mathbf{1}_{p^2} = \mathrm{tr}\{(\boldsymbol{\Sigma}\mathbf{1}_p\mathbf{1}_p{}')^2\} = \sigma_{\mathrm{T}}^2,\\
&\mathbf{1}_{p^2}{}'(\boldsymbol{\Sigma}\otimes\boldsymbol{\Sigma})\mathbf{e}_p = \mathrm{tr}(\boldsymbol{\Sigma}\,\mathbf{I}_p\boldsymbol{\Sigma}\mathbf{1}_p\mathbf{1}_p{}') = \mathbf{1}_p{}'\boldsymbol{\Sigma}^2\mathbf{1}_p,\\
&\mathbf{e}_p{}'(\boldsymbol{\Sigma}\otimes\boldsymbol{\Sigma})\mathbf{e}_p = \mathrm{tr}(\boldsymbol{\Sigma}^2),
\end{aligned} \tag{19}$$

(18) becomes

$$= \frac{2p^2}{(p-1)^2\sigma_{\mathrm{T}}^4}\{\sigma_{\mathrm{D}}^2\sigma_{\mathrm{T}}^2 - 2\sigma_{\mathrm{D}}\sigma_{\mathrm{T}}\mathbf{1}_p{}'\boldsymbol{\Sigma}^2\mathbf{1}_p + \sigma_{\mathrm{T}}^2\mathrm{tr}(\boldsymbol{\Sigma}^2)\}, \tag{20}$$

which gives (16). Q. E. D.

Let $\mathrm{abis}(\hat{\alpha})_{\mathrm{NT}}$ be the asymptotic bias of $\hat{\alpha}$ of order $O(n^{-1})$ under normality. Then, we have

Corollary 2.

$$n\,\mathrm{abis}(\hat{\alpha})_{\mathrm{NT}} = \frac{2p}{(p-1)\sigma_{\mathrm{T}}^2}(\mathbf{1}_p{}'\boldsymbol{\Sigma}^2\mathbf{1}_p - \sigma_{\mathrm{T}}\sigma_{\mathrm{D}}). \tag{21}$$

Proof. From (15) and (19),

$$\begin{aligned}
n\,\mathrm{abis}(\hat{\alpha})_{\mathrm{NT}} &= \frac{2p}{(p-1)\sigma_{\mathrm{T}}^3}\mathbf{1}_{p^2}{}'\{\sigma_{\mathrm{T}}\mathbf{D}_p\mathbf{D}_p^+(\boldsymbol{\Sigma}\otimes\boldsymbol{\Sigma})\mathbf{D}_p^{+}{}'\mathbf{C}_p{}'\\
&\qquad - \sigma_{\mathrm{D}}\mathbf{D}_p\mathbf{D}_p^+(\boldsymbol{\Sigma}\otimes\boldsymbol{\Sigma})\mathbf{D}_p^{+}{}'\mathbf{D}_p{}'\}\mathbf{1}_{p^2}\\
&= \frac{2p}{(p-1)\sigma_{\mathrm{T}}^3}(\sigma_{\mathrm{T}}\mathbf{1}_p{}'\boldsymbol{\Sigma}^2\mathbf{1}_p - \sigma_{\mathrm{D}}\sigma_{\mathrm{T}}^2),
\end{aligned} \tag{22}$$

which yields (21). Q. E. D.

When the components of a test are congeneric, the covariance structure becomes that of the one-factor model:

$$\boldsymbol{\Sigma} = \boldsymbol{\lambda}\boldsymbol{\lambda}' + \boldsymbol{\Psi}, \tag{23}$$

where $\boldsymbol{\lambda} = (\lambda_1, ..., \lambda_p)'$ is the vector of factor loadings and $\boldsymbol{\Psi} = \mathrm{diag}(\psi_1, ..., \psi_p)$ is the covariance matrix of the uncorrelated unique factors. When (23) holds, (21) becomes

$$\begin{aligned} n\,\mathrm{abis}(\hat{\alpha})_{\mathrm{NT}} &= \frac{2p}{(p-1)\sigma_{\mathrm{T}}^2}[(\mathbf{1}_p{}'\boldsymbol{\lambda})^2\boldsymbol{\lambda}'\boldsymbol{\lambda} + 2\mathbf{1}_p{}'\boldsymbol{\lambda}\boldsymbol{\lambda}'\boldsymbol{\Psi}\mathbf{1}_p + \mathrm{tr}(\boldsymbol{\Psi}^2) \\ &\qquad - \{\boldsymbol{\lambda}'\boldsymbol{\lambda} + \mathrm{tr}(\boldsymbol{\Psi})\}\{(\mathbf{1}_p{}'\boldsymbol{\lambda})^2 + \mathrm{tr}(\boldsymbol{\Psi})\}] \\ &= \frac{2p}{(p-1)\sigma_{\mathrm{T}}^2}[2\mathbf{1}_p{}'\boldsymbol{\lambda}\boldsymbol{\lambda}'\boldsymbol{\Psi}\mathbf{1}_p + \mathrm{tr}(\boldsymbol{\Psi}^2) \\ &\quad - \mathrm{tr}(\boldsymbol{\Psi})\{(\mathbf{1}_p{}'\boldsymbol{\lambda})^2 + \boldsymbol{\lambda}'\boldsymbol{\lambda}\} - \{\mathrm{tr}(\boldsymbol{\Psi})\}^2]. \end{aligned} \tag{24}$$

When the model is spherical or $\boldsymbol{\Psi} = \psi\mathbf{I}_p$ with $\psi > 0$, (24) becomes

$$\begin{aligned} n\,\mathrm{abis}(\hat{\alpha})_{\mathrm{NT}} &= \frac{2p}{(p-1)\sigma_{\mathrm{T}}^2}[2\psi(\mathbf{1}_p{}'\boldsymbol{\lambda})^2 + p\,\psi^2 - p\,\psi\{(\mathbf{1}_p{}'\boldsymbol{\lambda})^2 + \boldsymbol{\lambda}'\boldsymbol{\lambda}\} - p^2\psi^2] \\ &= \frac{-2p\psi}{(p-1)\sigma_{\mathrm{T}}^2}\{p\boldsymbol{\lambda}'\boldsymbol{\lambda} + (p-2)(\mathbf{1}_p{}'\boldsymbol{\lambda})^2 + p(p-1)\psi\} < 0. \end{aligned} \tag{25}$$

That is, under the assumption, $\mathrm{abis}(\hat{\alpha})_{\mathrm{NT}}$ is negative.

Similarly, when the components are tau-equivalent, or $\boldsymbol{\lambda}\boldsymbol{\lambda}' = \lambda\,\mathbf{1}_p\mathbf{1}_p{}'$ with $\lambda > 0$ and $\psi_i > 0,\ (i = 1, ..., p)$, we have from (24)

$$\begin{aligned} n\,\mathrm{abis}(\hat{\alpha})_{\mathrm{NT}} &= \frac{2p}{(p-1)\sigma_{\mathrm{T}}^2}[2p\lambda\mathrm{tr}(\boldsymbol{\Psi}) - \mathrm{tr}(\boldsymbol{\Psi})(p^2\lambda + p\lambda) + \mathrm{tr}(\boldsymbol{\Psi}^2) - \{\mathrm{tr}(\boldsymbol{\Psi})\}^2] \\ &= \frac{-2p}{(p-1)\sigma_{\mathrm{T}}^2}\{(p-1)p\lambda\mathrm{tr}(\boldsymbol{\Psi}) + \sum_{i \neq j}\psi_i\psi_j\} < 0. \end{aligned} \tag{26}$$

Under the assumption, $\mathrm{abis}(\hat{\alpha})_{\mathrm{NT}}$ is again negative. In case of compound symmetry with $\boldsymbol{\Sigma} = \lambda\,\mathbf{1}_p\mathbf{1}_p{}' + \psi\mathbf{I}_p,\ \lambda > 0$ and $\psi > 0$, it is obvious from above results that $\mathrm{abis}(\hat{\alpha})_{\mathrm{NT}}$ is negative. In this case, we have from (25) or (26),

$$\begin{aligned} n\,\mathrm{abis}(\hat{\alpha})_{\mathrm{NT}} &= \frac{-2p\psi}{(p-1)\sigma_{\mathrm{T}}^2}\{p^2\lambda + (p-2)p^2\lambda + p(p-1)\psi\} \\ &= -\frac{2p\,\psi(p-1)(p^2\lambda + p\psi)}{(p-1)(p^2\lambda + p\psi)^2} = -\frac{2\,\psi}{p\lambda + \psi} \\ &= -2(1-\alpha) < 0. \end{aligned} \tag{27}$$

Equation (27) holds exactly when n is replaced by $n-2$, which was first obtained by Kristof (1963) as was addressed earlier. From the negative values of $\mathrm{abis}(\hat{\alpha})_{\mathrm{NT}}$ in (25), (26) and (27), we expect that $\mathrm{abis}(\hat{\alpha})_{\mathrm{NT}}$ tends to be negative. Actually under various conditions with different types of nonnormal distributions $\mathrm{abis}(\hat{\alpha})$'s are negative, which will be shown later.

When compound symmetry is satisfied, $\mathrm{avar}(\hat{\alpha})_{\mathrm{NT}}$ becomes

$$\mathrm{avar}(\hat{\alpha})_{\mathrm{NT}} = \frac{2n^{-1}p}{p-1}(1-\alpha)^2, \tag{28}$$

Table 1: The ratio of the absolute value of the bias of $\hat{\alpha}$ to its standard error under compound symmetry and normality.

	p						
n	2	4	8	16	32	64	∞
Exact values							
25	.187	.227	.244	.252	.255	.257	.259
50	.137	.167	.180	.186	.189	.190	.192
100	.098	.120	.130	.134	.136	.137	.139
200	.070	.086	.093	.096	.097	.098	.099
400	.050	.061	.066	.068	.069	.070	.070
800	.035	.043	.047	.048	.049	.049	.050
Asymptotic values							
25	.200	.245	.265	.274	.278	.281	.283
50	.141	.173	.187	.194	.197	.198	.200
100	.100	.122	.132	.137	.139	.140	.141
200	.071	.087	.094	.097	.098	.099	.100
400	.050	.061	.066	.068	.070	.070	.071
800	.035	.043	.047	.048	.049	.050	.050

which is given from (16) and was given by van Zyl et al. (2000, Equation (22)). The corresponding exact result was first given by Kristof (1963, Equations (58) and (63)) when $2n^{-1}p$ in (28) is replaced by $2n(np-2)/\{(n-2)^2(n-4)\}$ (see also, van Zyl et al., 2000, Equation (12)). From (27) and (28), we have the relationship:

$$\mathrm{abis}(\hat{\alpha})_{\mathrm{NT}} = -n^{-1/2}\sqrt{\frac{2(p-1)}{p}}\,\mathrm{ase}(\hat{\alpha})_{\mathrm{NT}}, \tag{29}$$

where $\mathrm{ase}(\cdot) = \sqrt{\mathrm{avar}(\cdot)}$. Let $\mathrm{bis}(\cdot)$, $\mathrm{var}(\cdot)$ and $\mathrm{se}(\cdot)$ be the exact values corresponding to the asymptotic values. Then, the exact version of (29) is given by

$$\mathrm{bis}(\hat{\alpha})_{\mathrm{NT}} = -\sqrt{\frac{2(n-4)(p-1)}{n(np-2)}}\,\mathrm{se}(\hat{\alpha})_{\mathrm{NT}}. \tag{30}$$

Table 1 shows $-\mathrm{abis}(\hat{\alpha})_{\mathrm{NT}}/\mathrm{ase}(\hat{\alpha})_{\mathrm{NT}}$, the ratio of the absolute value of the asymptotic bias to the corresponding asymptotic standard error, and its exact counterpart for various values of p and n. Note that the relative bias is of order $O(n^{-1/2})$ and becomes negligible when n becomes large. However, from the table we see that when $n = 100$ the relative biases are almost more than 10%, and when $n = 200$ they are about 10%, which should not be neglected. We also find that the asymptotic values in the table are close to the corresponding exact values in practical sense even in small samples.

From the exact version of (27) and (28), we have the following expression of mean square error under compound symmetry and normality:

$$\mathrm{E}\{(\hat{\alpha}-\alpha)^2\}_{\mathrm{NT}} = \mathrm{var}(\hat{\alpha})_{\mathrm{NT}} + \{\mathrm{bis}(\hat{\alpha})_{\mathrm{NT}}\}^2 = \tfrac{(1-\alpha)^2}{(n-2)^2}\left\{\tfrac{2n(np-2)}{(n-4)(p-1)}+4\right\}, \tag{31}$$

whose asymptotic version is $n^{-1}\{(1-\alpha)^2/(p-1)\}\{2p+n^{-1}(20p-8)\}$, which decreases as n, p and/or α increase.

5. Coefficient alpha for standardized component scores

When the component scores of a test are standardized, we have the so-called standardized coefficient alpha:

$$\alpha_\rho = \frac{p}{p-1}\left(1 - \frac{p}{\mathbf{1}_p{}'\mathbf{P}\ \mathbf{1}_p}\right), \tag{32}$$

where $\mathbf{P} = \{\mathrm{Diag}(\mathbf{\Sigma})\}^{-1/2}\mathbf{\Sigma}\,\{\mathrm{Diag}(\mathbf{\Sigma})\}^{-1/2}$. The ase of $\hat{\alpha}_\rho$ and its relationships with p and α_ρ under normality were given by Hayashi and Kamata (2005). In this section, we generalize (32) using fixed weights $\mathbf{w}$ for standardized scores as follows:

$$\eta = \frac{p}{p-1}\left(1 - \frac{\mathbf{w}'\mathbf{w}}{\mathbf{w}'\mathbf{P}\ \mathbf{w}}\right) = \frac{p}{p-1}\left(1 - \frac{\mathbf{w}'\mathbf{w}}{\rho_{\mathrm{W}}}\right), \tag{33}$$

where $\rho_{\mathrm{W}} = \mathbf{w}'\mathbf{P}\ \mathbf{w}$ and the different notation η is used for clarity. The sample counterpart of (33) is

$$\hat{\eta} = \frac{p}{p-1}\left(1 - \frac{\mathbf{w}'\mathbf{w}}{\mathbf{w}'\mathbf{R}\ \mathbf{w}}\right) = \frac{p}{p-1}\left(1 - \frac{\mathbf{w}'\mathbf{w}}{r_{\mathrm{W}}}\right), \tag{34}$$

where $\mathbf{R}$ is the $p \times p$ usual sample correlation matrix and $r_{\mathrm{W}} = \mathbf{w}'\mathbf{R}\ \mathbf{w}$.

It is obvious that the basic results of (1) through (11) hold for $\hat{\eta}$ when α and $\hat{\alpha}$ are replaced by η and $\hat{\eta}$, respectively with some minor adjustment for the weights. The substantial difference is found for the partial derivatives of $\hat{\eta}$ with respect to s_{ab}'s evaluated at the true values. They are given in Subsection 1 of the appendix.

6. Simulation

To confirm the accuracy of the formulas in finite samples, simulations were performed under normality and nonnormality with the assumptions of congeneric component scores or parallel tests (subtests). The nonnormal data were given by the uniform, t ($df = 9$) and chi-square ($df = 10$, 3 and 1) distributions for the common factor or true score and for the unique factors or errors. The values $df = 9$ for the t-distribution and $df = 10$ for the chi-square distribution were employed for ease of comparison. That is, the third (fourth) standardized cumulants for the uniform, t ($df = 9$) and chi-square ($df = 10$) distributions are $0\,(-6/5)$, $0\,(6/5)$ and $2/\sqrt{5}\,(6/5)$, respectively (see e.g., Stuart & Ort, 1994, Sections 16.3 and 16.11) with the existence of the moments up to the eighth order for the t-distribution with $df = 9$. The chi-square distributions with $df = 10$, 3 and 1 were also used to represent mild, moderate and strong nonnormality respectively (the standardized r-th cumulant is given by $\nu 2^{r-1}(r-1)!/(2\nu)^{r/2} = 2^{(r/2)-1}(r-1)!/\nu^{(r/2)-1}$ when $df = \nu$).

The population covariance matrices for $\hat{\alpha}$ were constructed in six ways:

$$
\begin{aligned}
&[1]\ \mathbf{\Sigma} = \lambda \mathbf{1}_3\,\mathbf{1}_3{}' + \psi \mathbf{I}_3,\ \lambda = 0.3,\ \psi = 0.7,\ \ \alpha \doteq 0.56,\\
&[2]\ \mathbf{\Sigma} = \begin{bmatrix} 1.0 & \text{sym.} & \\ 0.2 & 1.0 & \\ 0.2 & 0.5 & 1.0 \end{bmatrix},\ \ \alpha \doteq 0.56,\\
&[3]\ \mathbf{\Sigma} = \begin{bmatrix} 1.0 & \text{sym.} & \\ 0.6 & 4.0 & \\ 0.9 & 1.8 & 9.0 \end{bmatrix},\ \ \alpha \doteq 0.48,\\
&[4]\ \mathbf{\Sigma} = \lambda \mathbf{1}_6\,\mathbf{1}_6{}' + \psi \mathbf{I}_6,\ \lambda = 0.1,\ \psi = 0.9,\ \ \alpha = 0.40,\\
&[5]\ \mathbf{\Sigma} = \lambda \mathbf{1}_6\,\mathbf{1}_6{}' + \psi \mathbf{I}_6,\ \lambda = 0.3,\ \psi = 0.7,\ \ \alpha = 0.72,\\
&[6]\ \mathbf{\Sigma} = \lambda \mathbf{1}_6\,\mathbf{1}_6{}' + \psi \mathbf{I}_6,\ \lambda = 0.8,\ \psi = 0.2,\ \ \alpha = 0.96.
\end{aligned}
\tag{35}
$$

The covariance matrices [1], [4], [5] and [6] show compound symmetry while [2] is the matrix of non-compound symmetry with the same α as that in [1]. The number of components in [1] to [3] is three while that in [4] to [6] is six. The population covariance matrices for $\hat{\eta}$ are set to be the same as above when $\mathbf{w} = \mathbf{1}_p$. The case of $\hat{\eta}$ with $\mathbf{w} = (1, 2, 3)'$ is given only by the population covariance matrix of [1]. That is, in this case the population covariance matrix [3] for $\hat{\alpha}$ becomes equal to that of the weighted standardized variables in [1] with $\mathbf{w} = (1, 2, 3)'$. So, the matrix numbers [1] to [6] common to $\hat{\alpha}$ and $\hat{\eta}$ are used (note that when $\mathbf{w} = (1, 2, 3)'$ for $\hat{\eta}$, number [3] is used instead of [1] for ease of comparison to $\hat{\alpha}$ with [3]).

Observations were randomly generated using the independently distributed common factor and unique factors under normality and nonnormality with $n = N - 1 = 25$, 50, 100 and 200. From the generated observations, the sample covariance matrix was obtained to have $\hat{\alpha}$ while for $\hat{\eta}$ the sample correlation matrix was computed. This procedure was replicated 1,000,000 times under each condition. From the set of 1,000,000 estimates for α or η, the simulated cumulants were obtained using k-statistics (unbiased estimators of cumulants; Stuart & Ort, 1994, Equation (12.28)) up to the fourth order. To derive the theoretical or asymptotic cumulants of $\hat{\alpha}$ and $\hat{\eta}$, we require the multivariate cumulants (moments) of the component scores (see (8) to (11)), which can be obtained by the cumulative property of the cumulant of the sum of independent variables (the single common factor and a unique factor):

$$
\begin{aligned}
\kappa_r(X_a, X_b, ..., X_h) = {} & \lambda_a \lambda_b \cdots \lambda_h \kappa_r(f) \\
& + \delta_{ab}\delta_{bc} \cdots \delta_{gh} (\psi_a \psi_b \cdots \psi_h)^{1/2} \kappa_r(e_a),
\end{aligned}
\tag{36}
$$

where $X_a = \lambda_a f + \psi_a^{1/2} e_a$; and f and e_a are the common and unique factors, respectively with standardization to have unit variances.

Tables 2 and 3 show the simulated and theoretical cumulants for $\hat{\alpha}$. The simulated values of a_1, a_2, a_3 and a_4 were given from the corresponding k-statistics multiplied by n, n, n^2 and n^3, respectively. The simulated value of $\Delta\alpha_2$ is given by $n^2\{\mathrm{SD}^2 - (\alpha_1/n)\}$, where SD is the square root of the usual unbiased sample variance from 1,000,000 estimates. From Table 2, we see that when n is increased, the simulated values become similar

Table 2: Simulated and theoretical cumulants of $\hat{\alpha}$ ($p = 3$).

	[1] Cmp. sym. (.3), $\alpha \doteq .56$				[2] Non-cmp. sym., $\alpha \doteq .56$				[3] SD(1, 2, 3), $\alpha \doteq .48$			
n	Normal	U	T9	C10	Normal	U	T9	C10	Normal	U	T9	C10
a_2: variance												
25	.787	.642	.895	.888	.794	.668	.889	.879	.730	.607	.820	.814
50	.670	.551	.765	.765	.680	.576	.762	.761	.637	.528	.722	.724
100	.620	.512	.713	.715	.628	.536	.711	.712	.595	.494	.684	.687
200	.596	.495	.692	.693	.602	.516	.689	.687	.574	.478	.669	.669
Th.	.574	.477	.671	.671	.583	.500	.666	.666	.558	.464	.652	.652
Δa_2: higher-order variance												
25	5.31	4.13	5.59	5.43	5.28	4.21	5.58	5.33	4.30	3.58	4.21	4.05
50	4.80	3.70	4.69	4.70	4.86	3.79	4.78	4.78	3.98	3.21	3.53	3.64
100	4.53	3.48	4.22	4.37	4.54	3.59	4.48	4.63	3.69	3.06	3.20	3.54
200	4.29	3.54	4.21	4.47	3.89	3.28	4.64	4.30	3.32	2.86	3.38	3.44
Th.	4.21	3.26	3.89	4.20	4.23	3.35	3.99	4.17	3.48	2.86	2.85	3.13
a_1: bias												
25	−.950	−.826	−1.053	−1.048	−.940	−.823	−1.038	−1.032	−.798	−.728	−.866	−.855
50	−.913	−.783	−1.017	−1.012	−.904	−.783	−1.008	−1.019	−.769	−.693	−.830	−.845
100	−.899	−.763	−1.007	−1.004	−.879	−.768	−.993	−1.005	−.748	−.677	−.819	−.839
200	−.869	−.758	−1.026	−.999	−.865	−.758	−1.002	−.998	−.731	−.678	−.827	−.827
Th.	−.875	−.752	−.998	−.998	−.867	−.745	−.990	−.990	−.737	−.659	−.816	−.816
a_3: third cumulant												
25	−4.86	−3.27	−6.00	−5.73	−4.88	−3.48	−6.03	−5.65	−3.96	−2.81	−4.76	−4.51
50	−3.43	−2.36	−4.25	−4.17	−3.56	−2.54	−4.28	−4.16	−2.94	−2.07	−3.54	−3.43
100	−2.95	−2.03	−3.65	−3.57	−2.99	−2.18	−3.69	−3.61	−2.53	−1.77	−3.09	−3.06
200	−2.72	−1.86	−3.42	−3.37	−2.74	−2.02	−3.53	−3.35	−2.32	−1.65	−2.96	−2.88
Th.	−2.51	−1.73	−3.13	−3.15	−2.58	−1.87	−3.18	−3.14	−2.20	−1.55	−2.73	−2.70
a_4: fourth cumulant												
25	61.7	34.8	82.1	73.0	59.8	37.2	84.3	72.0	44.6	27.3	58.1	52.6
50	33.6	19.7	45.6	43.4	36.0	21.7	46.1	43.8	27.1	16.1	33.9	31.3
100	26.6	15.6	36.3	32.4	26.4	16.8	36.6	33.8	20.8	12.3	27.3	25.8
200	23.4	13.0	33.0	30.5	23.1	14.9	33.7	29.3	17.6	11.1	26.4	23.4
Th.	20.4	11.8	26.5	27.4	21.2	13.2	27.6	27.3	16.5	10.0	21.0	20.9

Note. $n = N - 1$ in the simulation; Cmp. sym. (.3) = Compound symmetric data with $\lambda = .3$ and $\psi = .7$; SD(1, 2, 3) = Data with SDs being 1, 2 and 3; U, T9, C10=Uniform, t ($df = 9$) and chi-square ($df = 10$) distributions; Th.=Theoretical or asymptotic values.

to the corresponding theoretical or asymptotic values. The simulated values of a_1 are close to the asymptotic values even when n is relatively small. The simulated/asymptotic absolute values of a_2 and a_1 under the uniform distribution are smaller than those under normality while those under the t- and chi-square distributions are greater than those under normality. This is an expected result since the uniform distribution has negative fourth cumulant while t- and chi-square distributions have positive one. The tendency found in a_2 and a_1 is also observed for the results of a_3 and a_4. We find that the biases and third cumulants are all negative and that the added higher-order variances and fourth cumulants are positive.

Table 3 shows the similar results for $\hat{\alpha}$ when $p = 6$ with compound symmetry (the

Table 3: Simulated and theoretical cumulants of $\hat{\alpha}$ ($p = 6$).

	[4] Cmp. sym. (.1), $\alpha = .40$				[5] Cmp. sym. (.3), $\alpha = .72$				[6] Cmp. sym. (.8), $\alpha = .96$			
n	Normal	U	T9	C10	Normal	U	T9	C10	Normal	U	T9	C10
a_2: variance												
200	.898	.814	.982	.976	.195	1.36	.252	.253	.0040	.0018	.0060	.0061
Th.	.864	.783	.945	.945	.188	1.31	.245	.245	.0038	.0018	.0059	.0059
Δa_2: higher-order variance												
200	6.88	6.03	7.43	6.20	1.44	.92	1.39	1.57	.030	.011	.029	.043
Th.	6.62	5.93	6.42	6.46	1.44	.86	1.16	1.59	.029	.011	.017	.039
a_1: bias												
200	−1.21	−1.12	−1.31	−1.29	−.567	−.389	−.722	−.715	−.081	−.036	−.124	−.122
Th.	−1.20	−1.11	−1.29	−1.29	−.560	−.397	−.723	−.723	−.080	−.036	−.124	−.124
a_3: third cumulant												
200	−4.95	−4.14	−5.58	−5.44	−.495	−.252	−.704	−.736	−.0015	−.0003	−.0026	−.0029
Th.	−4.56	−3.81	−5.09	−5.08	−.464	−.235	−.621	−.679	−.0014	−.0003	−.0022	−.0026
a_4: fourth cumulant												
200	54.1	40.3	61.6	56.0	2.35	.88	4.04	4.22	.0010	.0001	.0024	.0026
Th.	45.1	35.0	51.2	51.5	2.14	.81	2.65	3.55	.0009	.0001	.0010	.0022

Note. $n = N - 1$ in the simulation; Cmp. sym. (x) = Compound symmetric data with $\lambda = x$ and $\psi = 1 - x$; U, T9, C10=Uniform, t ($df = 9$) and chi-square ($df = 10$) distributions; Th.=Theoretical or asymptotic values.

results for $n = 25$ to 100 are omitted to save space). The simulated values with $n = 200$ are again similar to the corresponding theoretical values except for the results of Δa_2 and a_4 in [6] under the t-distribution. Note that the absolute cumulants in [6] with $\alpha = 0.96$ are all small, which can be partially explained by a ceiling effect with the upper bound 1.0 for α. Most of the absolute cumulants by the uniform distribution are again smaller than those by the t- and chi-square distributions.

Table 4 shows the results for $\hat{\eta}$ corresponding to those in Table 2 for $\hat{\alpha}$. The variances of $\hat{\eta}$ are the same or slightly smaller than those of $\hat{\alpha}$. It is of interest to find that the absolute biases of $\hat{\eta}$ are consistently smaller than those of $\hat{\alpha}$. Table 5 shows the results for $\hat{\eta}$ under normality and varied strength of nonnormality when $p = 6$ with compound symmetry. From the table, we immediately find that when nonnormality is strong, the distribution of $\hat{\eta}$ can be quite different from that under normality. That is, the variance of $\hat{\eta}$ in [6] by the chi-square ($df = 1$) distribution is six times as large as that under normality. We also find that some of the added higher-order variances are negative.

Table 6 gives the simulated and asymptotic ratios of the higher-order standard error to the first-order one, where the simulated value is defined as SD/ase with ase $= \sqrt{a_2/n}$ while the asymptotic value is hase/ase with hase $= \{a_2/n + (\Delta a_2/n^2)\}^{1/2}$. Except for some of the results of the chi-square distribution with $df = 1$, the asymptotic values are reasonably similar to the corresponding simulated ones. The values tend to be larger when n is smaller, which indicates that the higher-order added variances are not negligible when n is relatively small. The same values as those for $\alpha = .40$ denoted by asterisks under

Table 4: Simulated and theoretical cumulants of $\hat{\eta}$ ($p = 3$).

	[1] Cmp. sym. (.3), $\eta \doteq .56$				[2] Non-cmp. sym., $\eta \doteq .56$				[3] W(1, 2, 3), $\eta \doteq .48$			
n	Normal	U	T9	C10	Normal	U	T9	C10	Normal	U	T9	C10
a_2: variance												
200	.597	.497	.692	.694	.589	.517	.661	.659	.566	.479	.651	.651
Th.	.574	.477	.671	.671	.570	.500	.639	.639	.548	.463	.633	.633
Δa_2: higher-order variance												
200	4.56	3.88	4.20	4.50	3.86	3.49	4.43	4.02	3.72	3.23	3.66	3.78
Th.	4.63	3.69	4.07	4.46	4.39	3.64	4.09	4.21	3.96	3.26	3.38	3.65
a_1: bias												
200	−.769	−.744	−.844	−.815	−.755	−.724	−.817	−.813	−.681	−.685	−.724	−.723
Th.	−.772	−.736	−.809	−.809	−.756	−.710	−.802	−.802	−.686	−.665	−.707	−.707
a_3: third cumulant												
200	−2.73	−1.87	−3.42	−3.38	−2.56	−2.02	−3.18	−2.99	−2.40	−1.73	−3.00	−2.94
Th.	−2.51	−1.73	−3.13	−3.15	−2.41	−1.87	−2.87	−2.80	−2.26	−1.61	−2.76	−2.74
a_4: fourth cumulant												
200	23.6	13.2	33.0	30.7	20.4	14.8	28.4	24.3	18.6	11.8	27.3	24.5
Th.	20.5	11.9	26.5	27.4	18.9	13.1	23.5	22.8	17.2	10.6	21.8	21.7

Note. $n = N - 1$ in the simulation; Cmp. sym. (.3) = Compound symmetric data with $\lambda = .3$ and $\psi = .7$; W(1, 2, 3) = Data with weights being 1, 2 and 3; U, T9, C10=Uniform, t ($df = 9$) and chi-square ($df = 10$) distributions; Th.=Theoretical or asymptotic values.

Table 5: Simulated and theoretical cumulants of $\hat{\eta}$ ($p = 6$).

	[4] Cmp. sym. (.1), $\eta = .40$				[5] Cmp. sym. (.3), $\eta = .72$				[6] Cmp. sym. (.8), $\eta = .96$			
n	Normal	C10	C3	C1	Normal	C10	C3	C1	Normal	C10	C3	C1
a_2: variance												
200	.90	.98	1.15	1.63	.196	.252	.382	.737	.0040	.0061	.0110	.0262
Th.	.86	.94	1.13	1.67	.188	.245	.378	.757	.0038	.0059	.0107	.0245
Δa_2: higher-order variance												
200	7.4	6.2	3.7	−7.5	1.54	1.45	.85	−4.02	.026	.034	.064	.353
Th.	7.2	6.5	4.0	−11.3	1.64	1.63	.97	−5.81	.050	.084	.221	1.051
a_1: bias												
200	−1.22	−1.16	−1.01	−.61	−.53	−.58	−.73	−1.12	−.052	−.087	−.178	−.423
Th.	−1.21	−1.15	−1.00	−.56	−.52	−.59	−.74	−1.16	−.051	−.089	−.179	−.436
a_3: third cumulant												
200	−4.97	−5.44	−6.44	−8.57	−.50	−.73	−1.38	−4.33	−.0014	−.0029	−.0085	−.0476
Th.	−4.56	−5.08	−6.10	−7.67	−.46	−.68	−1.32	−4.18	−.0014	−.0026	−.0078	−.0395
a_4: fourth cumulant												
200	53.9	55.9	70.9	89.8	2.4	4.2	9.0	41.9	.0009	.0026	.0121	.1609
Th.	45.1	51.5	65.0	100.6	2.1	3.6	9.0	47.0	.0011	.0030	.0156	.1810

Note. $n = N - 1$ in the simulation; Cmp. sym. (x) = Compound symmetric data with $\lambda = x$ and $\psi = 1 - x$; C10, C3, C1=Chi-square distributions with df = 10, 3, 1; Th.=Theoretical or asymptotic values.

Table 6: Simulated and theoretical ratios of the bias-corrected higher-order asymptotic standard error to the first-order asymptotic one for $\hat{\alpha}$ and $\hat{\eta}$ ($p = 6$; compound symmetric data).

				Chi-square					
		Normal		$df = 10$		$df = 3$		$df = 1$	
	n	Sim.	Th.	Sim.	Th.	Sim.	Th.	Sim.	Th.
[4] $\alpha = .40$	25	1.178	1.143	1.161	1.129	1.126	1.086	1.028	.917
	50	1.083	1.074	1.075	1.066	1.055	1.044	.998	.960
	100	1.039	1.038	1.035	1.034	1.025	1.022	.993	.980
	200	1.020	1.019	1.016	1.017	1.010	1.011	.994	.990
[5] $\alpha = .72$	25	1.178	*	1.159	1.122	1.136	1.081	1.103	.976
	50	1.082	*	1.073	1.063	1.061	1.042	1.041	.988
	100	1.040	*	1.033	1.032	1.028	1.021	1.014	.994
	200	1.019	*	1.016	1.016	1.012	1.011	1.005	.997
[6] $\alpha = .96$	25	1.179	*	1.166	1.125	1.227	1.146	1.535	1.277
	50	1.081	*	1.077	1.064	1.100	1.076	1.244	1.147
	100	1.041	*	1.036	1.033	1.048	1.039	1.111	1.076
	200	1.019	*	1.018	1.016	1.021	1.019	1.051	1.039
[4] $\eta = .40$	25	1.193	1.155	1.168	1.130	1.120	1.068	1.016	.853
	50	1.090	1.080	1.076	1.067	1.049	1.034	.985	.930
	100	1.042	1.041	1.035	1.034	1.021	1.017	.985	.965
	200	1.021	1.021	1.016	1.017	1.008	1.009	.989	.983
[5] $\eta = .72$	25	1.193	1.161	1.155	1.125	1.098	1.050	1.018	.833
	50	1.089	1.084	1.069	1.064	1.039	1.025	.985	.920
	100	1.043	1.043	1.031	1.033	1.015	1.013	.981	.961
	200	1.020	1.022	1.015	1.016	1.006	1.006	.987	.981
[6] $\eta = .96$	25	1.162	1.235	1.139	1.254	1.171	1.351	1.367	1.648
	50	1.072	1.124	1.063	1.134	1.074	1.188	1.170	1.363
	100	1.036	1.064	1.029	1.069	1.035	1.098	1.077	1.196
	200	1.017	1.032	1.014	1.035	1.015	1.050	1.035	1.102

Note. Sim.=SD/ase and Th.=hase/ase, where SD=Standard deviation from simulation, ase = $\sqrt{a_2/n}$ and hase = $\{a_2/n + (\Delta a_2/n^2)\}^{1/2}$; The asterisks indicate the same values as those for $\alpha = .40$.

normality in the table stem from the property of the asymptotic (higher-order) variances proportional to $(1-\alpha)^2$ (recall (28) and its exact counterpart).

Figure 1 presents the theoretical (curved lines) and simulated (histograms) distributions of the transformed estimates of α, i.e. $(\hat{\alpha}-\alpha)/\mathrm{ase}(\hat{\alpha})$ in the case of [1] with compound symmetry ($\lambda = 0.3,\ \psi = 0.7$), $p = 3$ and $n = 50$ under the six different distributions. The symmetric dashed lines show the standard normal distributions and correspond to the approximations by the usual asymptotic standard errors based on the asymptotic normality of $\hat{\alpha}$. We easily find that with $n = 50$ the actual distributions are somewhat different from the normal approximations. On the other hand, the theoretical distributions given by the single-term Edgeworth expansions (solid lines) and two-term one (dotted lines) well approximate the simulated ones. Although it is not easy to differentiate the two types of Edgeworth expanded distributions when the component scores are normally or weakly nonnormally distributed, we see that the two-term Edgeworth expansions give improved

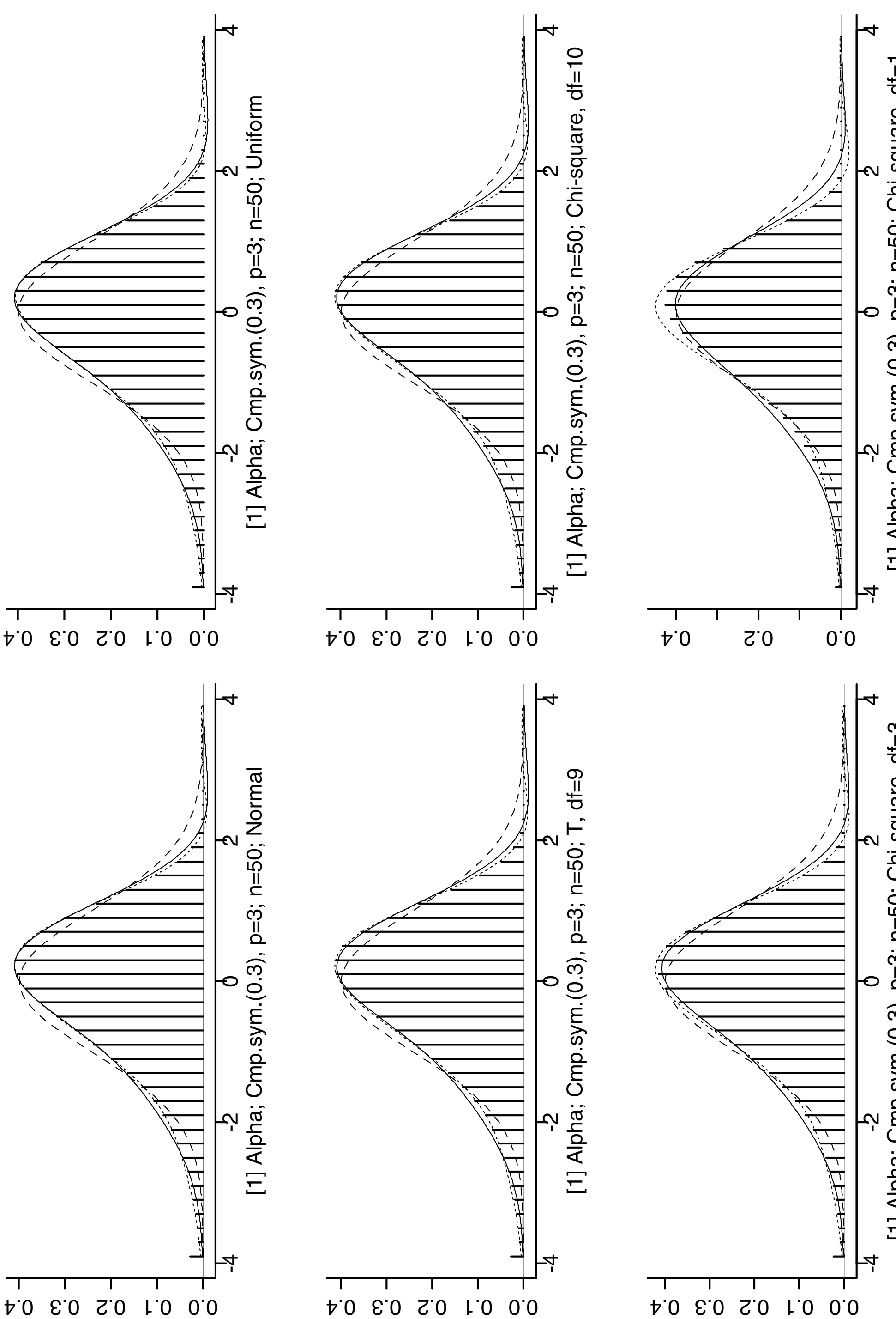

Figure 1: Theoretical (curved lines) and simulated (histograms) distributions of the transformed sample alpha coefficients (dashed lines=standard normal distribution, solid lines=single-term Edgeworth expansion, dotted lines=two-term Edgeworth expansion).

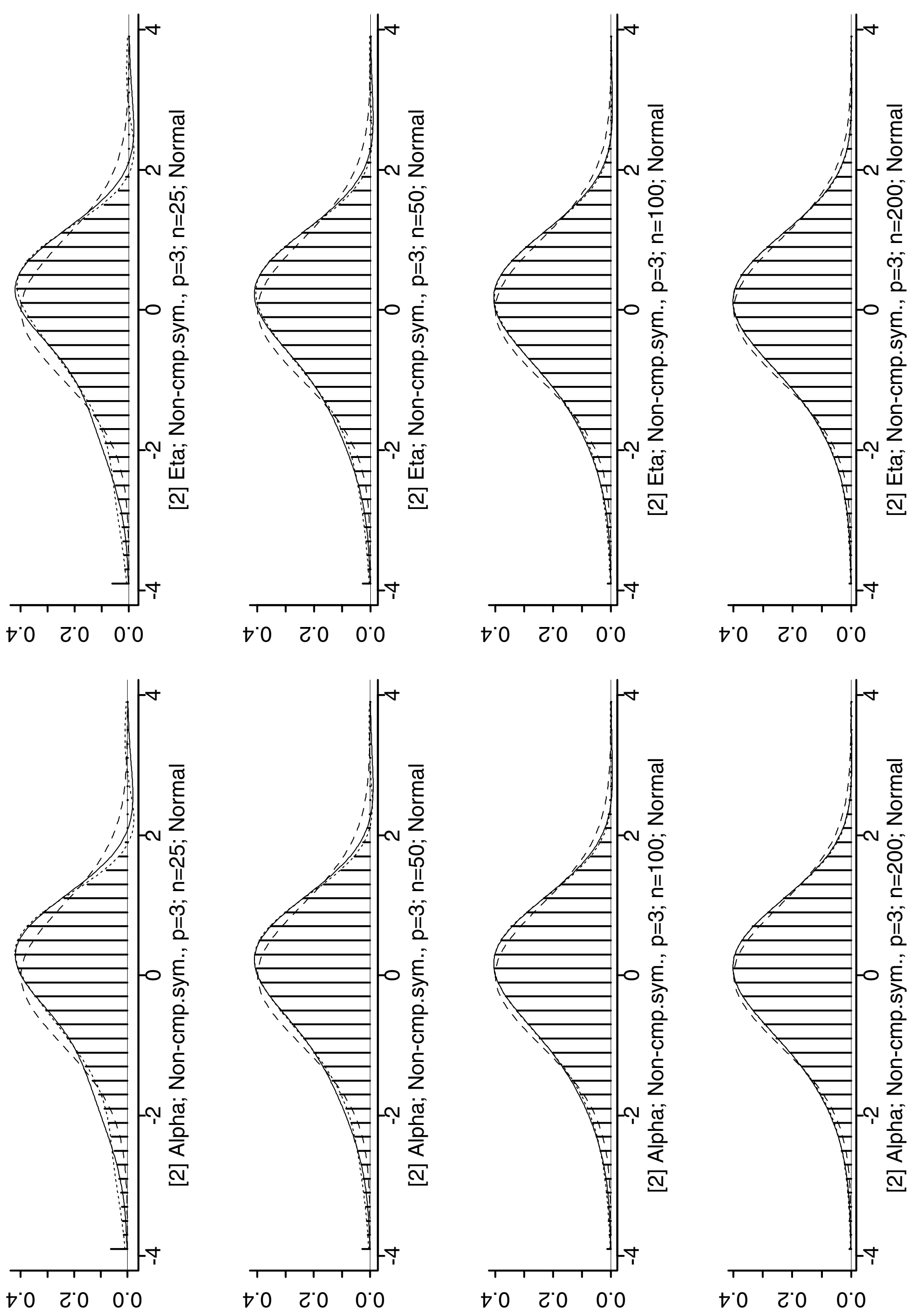

Figure 2: Theoretical (curved lines) and simulated (histograms) distributions of the transformed sample alpha/eta coefficients (dashed lines=standard normal distribution, solid lines=single-term Edgeworth expansion, dotted lines=two-term Edgeworth expansion).

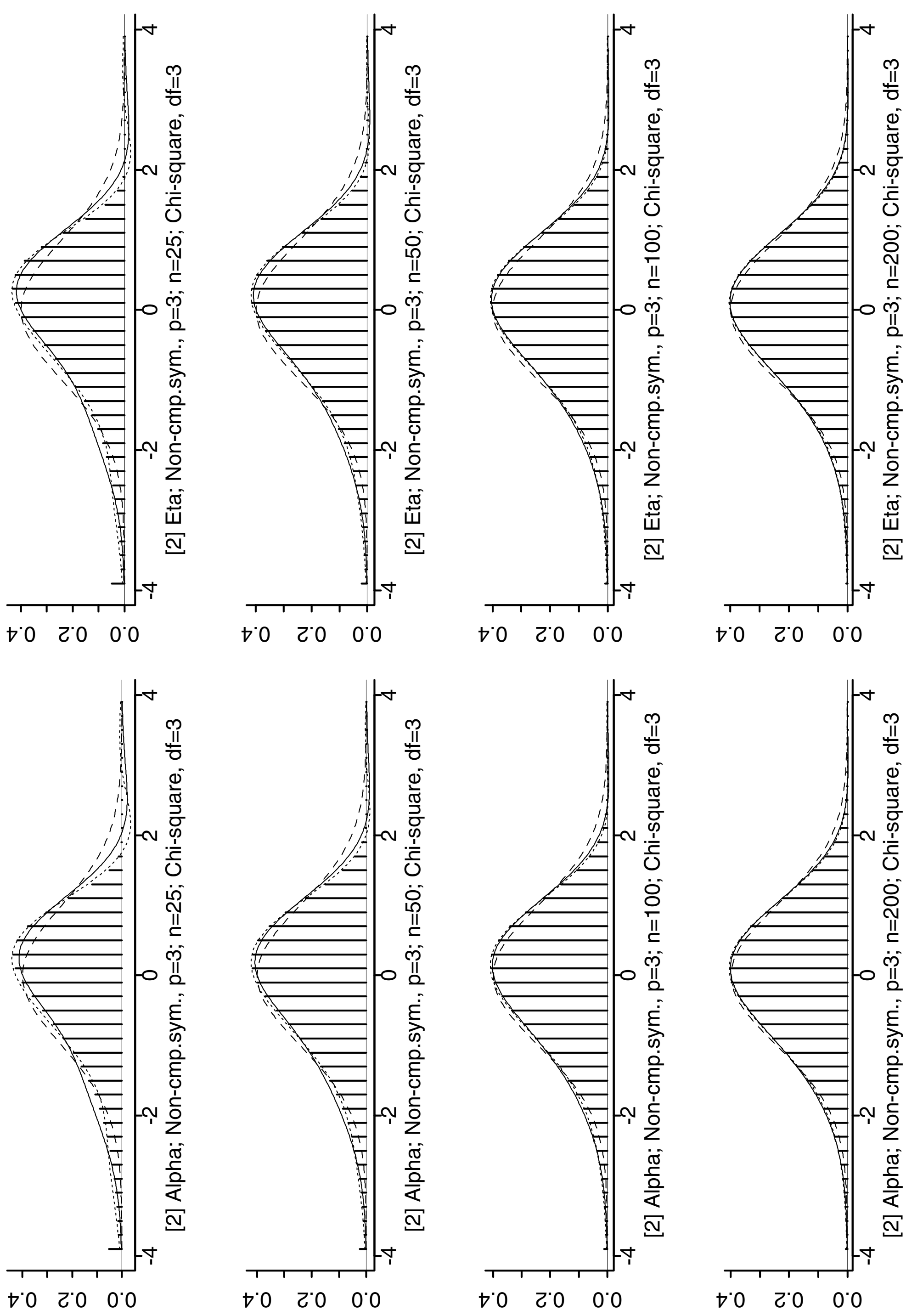

Figure 3: Theoretical (curved lines) and simulated (histograms) distributions of the transformed sample alpha/eta coefficients (dashed lines=standard normal distribution, solid lines=single-term Edgeworth expansion, dotted lines=two-term Edgeworth expansion).

approximation over the single-term ones in the case of the chi-square distributions with $df = 3$ and 1. However, the two-term expansions seem to excessively adjust the single-term ones in some area of the density functions e.g., the central area in the case of the chi-square distribution with $df = 1$. The theoretical densities in some upper tail area are negative, which generally happens especially when n is small. The negative densities are to be set to zero when necessary.

Figures 2 and 3 give the similar plots for $\hat{\alpha}$ and $\hat{\eta}$ in the case of [2] with $n = 25$ to 200, non-compound symmetry and $p = 3$ under the normal and chi-square distributions ($df = 3$), respectively. Again, we find the improved approximations by the Edgeworth expansions especially when n is small. We also find that the asymptotic normality gradually appears as the sample size becomes larger.

7. Discussion

From the simulation study with numerical illustration, we find that the distributions of $\hat{\alpha}$ and $\hat{\eta}$ under nonnormality are different from those under normality, which was well shown by the large difference of the variances in Table 2 especially when nonnormality is strong. Even under normality, we find that the distributions of $\hat{\alpha}$ and $\hat{\eta}$ are negatively biased and negatively skewed when n is relatively small. These findings support the use of the Edgeworth expansions. The two-term Edgeworth expansions are useful to obtain analytically the approximations to the distributions of $\hat{\alpha}$ and $\hat{\eta}$ when population α and η are available as in simulations. However, in practice α and η are unknown. In such a case, we have difficulty in making statistical inference using the two-term Edgeworth expansion since the estimates of a_1 to a_4 with Δa_2 are accurate only up to $n^{-1/2}$ i.e., $\hat{a}_i - a_i = O_p(n^{-1/2})$ (recall (5) and (6)). So, in practice only the single-term expansion using $\hat{a}_1$, $\hat{a}_2$ and $\hat{a}_3$ is recommended to use (with some adaptation to the corresponding Studentized $\hat{\alpha}$). In many cases, as shown in the simulation study, the improvement given by the two-term Edgeworth expansion over the single-term one is slight in practical sense. In addition, for $\hat{a}_1$, $\hat{a}_2$ and $\hat{a}_3$ we do not require the third partial derivatives of $\hat{\alpha}$ and $\hat{\eta}$ with respect to $\mathbf{s}$, and the moments of the component scores higher-than the sixth order (see (8) to (10)). When $\hat{a}_2$ is used in place of a_2, we have to consider the results for the Studentized statistic, which is given in Subsection 2 of the appendix:

As addressed in the introductory section, α does not give the true reliability, when a test is multidimensional or a test has multiple common factors. In this case, α generally gives values smaller than the corresponding true values (see e.g., Zinbarg, Revelle, Yovel, & Li, 2005). When we have a multidimensional test which is composed of subtests with each subtest having several components, stratified α (Cronbach, Schönemann, & McKie, 1965) can be used, which is expressed by Feldt and Brennan's (1989, Equation (28)) formulation as:

$$\text{Stratified } \alpha = 1 - \frac{1}{\mathbf{1}_p{}'\boldsymbol{\Sigma}\,\mathbf{1}_p} \sum_{i=1}^{K} \mathbf{1}_{p_i}{}'\boldsymbol{\Sigma}_{ii}\mathbf{1}_{p_i}(1 - \alpha_i), \tag{37}$$

where $\boldsymbol{\Sigma}_{ii}$ is the $p_i \times p_i$ covariance matrix for the i-th subtest; p_i is the number of components in the i-th subtest; $\sum_{i=1}^{K} p_i = p$; K is the number of subtests; and α_i is coefficient alpha for the i-th subtest.

It is known that in multidimensional cases, stratified α gives values closer to the true ones than the usual α (see Osburn, 2000). The sample counterpart of stratified α is given by using $\mathbf{S}$, $\mathbf{S}_{ii}$(the (i,i)th submatrix of $\mathbf{S}$) and $\hat{\alpha}_i$ in place of $\boldsymbol{\Sigma}$, $\boldsymbol{\Sigma}_{ii}$ and α_i in (37). Asymptotic expansions for $\hat{\alpha}_i$'s are obtained in the same manner as for $\hat{\alpha}$. The asymptotic expansion of the distribution of the estimator of stratified α can also be obtained with some added algebra for the partial derivatives of the estimator with respect to $\mathbf{s}$.

So far, simulated methods have not been dealt with except for the Monte Carlo method for comparison with the asymptotic ones in the numerical examples. It is well known that resampling methods such as bootstrapping often have various advantages over the asymptotic methods. In the case of α, the resampling methods may give economical results in computation. Further, since the estimators of α and its stratified version are elementary functions of sample variances and covariances, they are free from the anomalous cases such as non-convergence, (rotational indeterminacy) and Heywood cases in factor analysis. So, advantages of the asymptotic methods may be found as general ones. That is, asymptotic results do not have randomness found in the results of resampling which depend on e.g., methods of generating pseudo-random numbers, their seeds and the numbers of resamples used. Further, the various functional properties discussed in the asymptotic variance and bias under normality in Section 4 are due to the transparent property of the asymptotic cumulants. It is also useful to use the asymptotic results indirectly or use them with resampling methods (see Abramovitch & Singh, 1985; Hall, 1992, Ch.3; Zhou & Gao, 2000).

Appendix

1. The partial derivatives of $\hat{\eta}$ with respect to s evaluated at the true values

From (34), the matrix expression of the first partial derivative is

$$\frac{p-1}{p\,\mathbf{w}'\mathbf{w}}\frac{\partial \eta}{\partial \boldsymbol{\sigma}'} = \frac{1}{\rho_{\mathrm{W}}^2}\frac{\partial \rho_{\mathrm{W}}}{\partial \boldsymbol{\rho}'}\frac{\partial \boldsymbol{\rho}}{\partial \boldsymbol{\sigma}'}, \tag{A1}$$

where $\boldsymbol{\rho} = \mathrm{vb}(\mathbf{P})$ and $\mathrm{vb}(\cdot)$ denotes the vectorizing operator taking the elements below the main diagonal of a square matrix; and $\partial\rho_{\mathrm{W}}/\partial\rho_{ij} = 2w_i w_j$, $(p \geq i > j \geq 1)$ with $\rho_{ij} = (\mathbf{P})_{ij}$. The elementwise version of (A1) is

$$\begin{aligned}\frac{p-1}{p\,\mathbf{w}'\mathbf{w}}\frac{\partial \eta}{\partial \sigma_{ab}} &= \frac{1}{\rho_{\mathrm{W}}^2}\frac{\partial \rho_{\mathrm{W}}}{\partial \sigma_{ab}} \\ &= \frac{1}{\rho_{\mathrm{W}}^2}\{2w_a w_b(1-\delta_{ab})\sigma_{aa}^{-1/2}\sigma_{bb}^{-1/2} - \delta_{ab}\sigma_{aa}^{-1} w_a \sum_{g\neq a}^{p} \rho_{ag} w_g\},\ (p \geq a \geq b \geq 1).\end{aligned} \tag{A2}$$

The second partial derivatives are given as

$$\frac{p-1}{p\,\mathbf{w}'\mathbf{w}}\left(\frac{\partial}{\partial\boldsymbol{\sigma}'}\right)^{\langle 2\rangle}\eta=\frac{1}{\rho_{\mathrm{W}}^2}\frac{\partial\rho_{\mathrm{W}}}{\partial\boldsymbol{\rho}'}\left(\frac{\partial}{\partial\boldsymbol{\sigma}'}\right)^{\langle 2\rangle}\boldsymbol{\rho}-\frac{2}{\rho_{\mathrm{W}}^3}\left(\frac{\partial\rho_{\mathrm{W}}}{\partial\boldsymbol{\rho}'}\frac{\partial\boldsymbol{\rho}}{\partial\boldsymbol{\sigma}'}\right)^{\langle 2\rangle}\tag{A3}$$

or

$$\frac{p-1}{p\,\mathbf{w}'\mathbf{w}}\frac{\partial^2\eta}{\partial\sigma_{ab}\partial\sigma_{cd}}=\frac{1}{\rho_{\mathrm{W}}^2}\frac{\partial^2\rho_{\mathrm{W}}}{\partial\sigma_{ab}\partial\sigma_{cd}}-\frac{2}{\rho_{\mathrm{W}}^3}\frac{\partial\rho_{\mathrm{W}}}{\partial\sigma_{ab}}\frac{\partial\rho_{\mathrm{W}}}{\partial\sigma_{cd}},\tag{A4}$$

where

$$\begin{aligned}\frac{\partial^2\rho_{\mathrm{W}}}{\partial\sigma_{ab}\partial\sigma_{cd}}&=-w_aw_b(1-\delta_{ab})\delta_{cd}\sigma_{cc}^{-1}(\delta_{ca}+\delta_{cb})\sigma_{aa}^{-1/2}\sigma_{bb}^{-1/2}\\&\quad-\delta_{ab}\sum_{g\neq a}^{p}[\{(\delta_{ac}\delta_{gd}+\delta_{ad}\delta_{gc})\sigma_{aa}^{-1/2}\sigma_{gg}^{-1/2}-\frac{1}{2}\delta_{cd}\rho_{ag}\sigma_{cc}^{-1}(\delta_{ca}+\delta_{cg})\}\sigma_{aa}^{-1}\\&\quad-\rho_{ag}\delta_{cd}\delta_{ca}\sigma_{aa}^{-2}]w_aw_g,\quad(p\geq a\geq b\geq 1;p\geq c\geq d\geq 1).\end{aligned}\tag{A5}$$

The third partial derivatives become

$$\begin{aligned}&\frac{p-1}{p\,\mathbf{w}'\mathbf{w}}\left(\frac{\partial}{\partial\boldsymbol{\sigma}'}\right)^{\langle 3\rangle}\eta=\frac{p-1}{p\,\mathbf{w}'\mathbf{w}}\mathrm{vec}'\left\{\frac{\partial}{\partial\boldsymbol{\sigma}}\left(\frac{\partial}{\partial\boldsymbol{\sigma}'}\right)^{\langle 2\rangle}\eta\right\}\\&=\mathrm{vec}'\left[\frac{1}{\rho_{\mathrm{W}}^2}\frac{\partial}{\partial\boldsymbol{\sigma}}\left\{\frac{\partial\rho_{\mathrm{W}}}{\partial\boldsymbol{\rho}'}\left(\frac{\partial}{\partial\boldsymbol{\sigma}'}\right)^{\langle 2\rangle}\boldsymbol{\rho}\right\}-\frac{2}{\rho_{\mathrm{W}}^3}\left\{\left(\frac{\partial}{\partial\boldsymbol{\sigma}}\left(\frac{\partial\rho_{\mathrm{W}}}{\partial\boldsymbol{\rho}'}\frac{\partial\boldsymbol{\rho}}{\partial\boldsymbol{\sigma}'}\right)\right)\otimes\left(\frac{\partial\rho_{\mathrm{W}}}{\partial\boldsymbol{\rho}'}\frac{\partial\boldsymbol{\rho}}{\partial\boldsymbol{\sigma}'}\right)\right.\right.\\&\quad\left.+\left(\frac{\partial\rho_{\mathrm{W}}}{\partial\boldsymbol{\rho}'}\frac{\partial\boldsymbol{\rho}}{\partial\boldsymbol{\sigma}'}\right)\otimes\left(\frac{\partial}{\partial\boldsymbol{\sigma}}\left(\frac{\partial\rho_{\mathrm{W}}}{\partial\boldsymbol{\rho}'}\frac{\partial\boldsymbol{\rho}}{\partial\boldsymbol{\sigma}'}\right)\right)+\left(\frac{\partial\rho_{\mathrm{W}}}{\partial\boldsymbol{\rho}'}\frac{\partial\boldsymbol{\rho}}{\partial\boldsymbol{\sigma}'}\right)'\frac{\partial\rho_{\mathrm{W}}}{\partial\boldsymbol{\rho}'}\left(\frac{\partial}{\partial\boldsymbol{\sigma}'}\right)^{\langle 2\rangle}\boldsymbol{\rho}\right\}\\&\quad\left.+\frac{6}{\rho_{\mathrm{W}}^4}\left(\frac{\partial\rho_{\mathrm{W}}}{\partial\boldsymbol{\rho}'}\frac{\partial\boldsymbol{\rho}}{\partial\boldsymbol{\sigma}'}\right)'\left(\frac{\partial\rho_{\mathrm{W}}}{\partial\boldsymbol{\rho}'}\frac{\partial\boldsymbol{\rho}}{\partial\boldsymbol{\sigma}'}\right)^{\langle 2\rangle}\right]\end{aligned}\tag{A6}$$

or

$$\begin{aligned}&\frac{p-1}{p\,\mathbf{w}'\mathbf{w}}\frac{\partial^3\eta}{\partial\sigma_{ab}\partial\sigma_{cd}\partial\sigma_{ef}}\\&=\frac{1}{\rho_{\mathrm{W}}^2}\frac{\partial^3\rho_{\mathrm{W}}}{\partial\sigma_{ab}\partial\sigma_{cd}\partial\sigma_{ef}}-\frac{2}{\rho_{\mathrm{W}}^3}\left(\frac{\partial^2\rho_{\mathrm{W}}}{\partial\sigma_{ab}\partial\sigma_{cd}}\frac{\partial\rho_{\mathrm{W}}}{\partial\sigma_{ef}}+\frac{\partial^2\rho_{\mathrm{W}}}{\partial\sigma_{ab}\partial\sigma_{ef}}\frac{\partial\rho_{\mathrm{W}}}{\partial\sigma_{cd}}+\frac{\partial^2\rho_{\mathrm{W}}}{\partial\sigma_{cd}\partial\sigma_{ef}}\frac{\partial\rho_{\mathrm{W}}}{\partial\sigma_{ab}}\right)\\&\quad+\frac{6}{\rho_{\mathrm{W}}^4}\frac{\partial\rho_{\mathrm{W}}}{\partial\sigma_{ab}}\frac{\partial\rho_{\mathrm{W}}}{\partial\sigma_{cd}}\frac{\partial\rho_{\mathrm{W}}}{\partial\sigma_{ef}},\end{aligned}\tag{A7}$$

where $\mathrm{vec}'(\cdot)=\{\mathrm{vec}(\cdot)\}'$ and

$$\begin{aligned}\frac{\partial^3\rho_{\mathrm{W}}}{\partial\sigma_{ab}\partial\sigma_{cd}\partial\sigma_{ef}}&=w_aw_b(1-\delta_{ab})\delta_{cd}(\delta_{ca}+\delta_{cb})\delta_{ef}\\&\quad\times\{\delta_{ce}\sigma_{cc}^{-2}\sigma_{aa}^{-1/2}\sigma_{bb}^{-1/2}+(1/2)\delta_{ae}\sigma_{cc}^{-1}\sigma_{aa}^{-3/2}\sigma_{bb}^{-1/2}+(1/2)\delta_{be}\sigma_{cc}^{-1}\sigma_{aa}^{-1/2}\sigma_{bb}^{-3/2}\}\\&\quad+\delta_{ab}\sum_{g\neq a}^{p}\Big[(\delta_{ac}\delta_{gd}+\delta_{ad}\delta_{gc})\delta_{ef}\{(3/2)\delta_{ae}\sigma_{aa}^{-5/2}\sigma_{gg}^{-1/2}+(1/2)\delta_{ge}\sigma_{aa}^{-3/2}\sigma_{gg}^{-3/2}\}\\&\qquad+\{(1/2)\delta_{cd}\sigma_{cc}^{-1}(\delta_{ca}+\delta_{cg})\sigma_{aa}^{-1}+\delta_{cd}\delta_{ca}\sigma_{aa}^{-2}\}\end{aligned}$$

$$\times \{(\delta_{ae}\delta_{gf} + \delta_{af}\delta_{ge})\sigma_{aa}^{-1/2}\sigma_{gg}^{-1/2} - (1/2)\delta_{ef}\rho_{ag}\sigma_{ee}^{-1}(\delta_{ea} + \delta_{eg})\}$$
$$- (1/2)\delta_{cd}\rho_{ag}(\delta_{ca} + \delta_{cg})\delta_{ef}(\delta_{ce}\sigma_{cc}^{-2}\sigma_{aa}^{-1} + \delta_{ae}\sigma_{cc}^{-1}\sigma_{aa}^{-2})$$
$$- 2\rho_{ag}\delta_{cd}\delta_{ca}\delta_{ef}\delta_{ae}\sigma_{aa}^{-3}\Big]\ w_a w_g,$$
$$(p \geq a \geq b \geq 1; p \geq c \geq d \geq 1;\ p \geq e \geq f \geq 1). \qquad \text{(A8)}$$

In (A1), (A3) and (A6), the elementwise expressions of $\frac{\partial \boldsymbol{\rho}}{\partial \boldsymbol{\sigma}'}$, $\left(\frac{\partial}{\partial \boldsymbol{\sigma}'}\right)^{\langle 2\rangle} \boldsymbol{\rho}$ and $\left(\frac{\partial}{\partial \sigma'}\right)^{\langle 3\rangle} \boldsymbol{\rho}$ are available, but omitted since the final elementwise expressions (A2), (A4) with (A5), and (A7) with (A8) are given above. The values of $\sigma_{ii}, (i = 1, ..., p)$ in this appendix can be set to 1 without loss of generality after taking the partial derivatives.

2. Asymptotic expansion for the Studentized statistic

Let t be a Studentized statistic of $\hat{\alpha}$ or $\hat{\eta}$ as

$$t = v/\hat{a}_2 = n^{-1/2}(\hat{\theta} - \theta)/\hat{a}_2, \qquad \text{(A9)}$$

where θ stands for α or η. Let $a_i{}'$, $(i = 1, 2, 3)$ be the asymptotic cumulants of t whose estimates are $\hat{a}_i{}'$, $(i = 1, 2, 3)$. Then from Ogasawara (2005, Theorem 2):

$$a_1{}' = a_2^{-1/2}a_1 - \frac{1}{2}a_2^{-3/2}\left\{\frac{\partial \theta}{\partial \boldsymbol{\sigma}'}\boldsymbol{\Omega}\frac{\partial a_2}{\partial \boldsymbol{\sigma}} + \frac{\partial \theta}{\partial \boldsymbol{\sigma}'} n\ \mathrm{acov}(\mathbf{s}, \mathbf{s}_{(4)}{}')\frac{\partial a_2}{\partial \boldsymbol{\sigma}_{(4)}}\right\},$$
$$a_2{}' = 1, \qquad \text{(A10)}$$
$$a_3{}' = a_2^{-3/2}a_3 - 3a_2^{-3/2}\left\{\frac{\partial \theta}{\partial \boldsymbol{\sigma}'}\boldsymbol{\Omega}\frac{\partial a_2}{\partial \boldsymbol{\sigma}} + \frac{\partial \theta}{\partial \boldsymbol{\sigma}'} n\ \mathrm{acov}(\mathbf{s}, \mathbf{s}_{(4)}{}')\frac{\partial a_2}{\partial \boldsymbol{\sigma}_{(4)}}\right\},$$

where $\boldsymbol{\sigma}_{(4)}$ is the ${}_{p+3}C_4 \times 1$ vector:

$$\boldsymbol{\sigma}_{(4)} = (\sigma_{1111}, \sigma_{2111}, ..., \sigma_{ijkl}, ..., \sigma_{pppp})', (p \geq i \geq j \geq k \geq l \geq 1), \qquad \text{(A11)}$$

and $\mathbf{s}_{(4)}$ is its sample counterpart; $\mathrm{acov}(\mathbf{s}, \mathbf{s}_{(4)}{}')$ is the $p(p+1)/2 \times {}_{p+3}C_4$ asymptotic covariance matrix of $\mathbf{s}$ and $\mathbf{s}_{(4)}$. From (5), using (A10) (note $a_2{}' = 1$), we have the single-term Edgeworth expansion of t:

$$\Pr(t \leq x) = \Phi(x) - n^{-1/2}\left\{a_1{}' + \frac{a_3{}'}{6}(x^2 - 1)\right\}\phi(x) + O(n^{-1}). \qquad \text{(A12)}$$

In practice, the Cornish-Fisher expansion (see e.g., Hall, 1992) corresponding to the single-term Edgeworth expansion is convenient for developing confidence intervals of α and η with fixed asymptotic confidence coefficients e.g., $1 - \tilde{\alpha} = 0.95$. That is,

$$\Pr\left(t \leq t_{\tilde{\alpha}/2}\right) = \Pr\left(\frac{n^{1/2}(\hat{\theta} - \theta)}{\hat{a}_1^{1/2}} \leq t_{\tilde{\alpha}/2}\right) = 1 - \frac{\tilde{\alpha}}{2} + O(n^{-1}), \qquad \text{(A13)}$$

where

$$t_{\tilde{\alpha}/2} = z_{\tilde{\alpha}/2} + n^{-1/2}\left\{a_1{}' + \frac{a_3{}'}{6}(z_{\tilde{\alpha}/2}^2 - 1)\right\} \quad \text{with} \quad \Phi(z_{\tilde{\alpha}/2}) = 1 - (\tilde{\alpha}/2) \qquad \text{(A14)}$$

From (A13), the confidence interval of θ with asymptotic confidence coefficient $1-\tilde{\alpha}$ using $\hat{a}_1{}'$, $\hat{a}_3{}'$ with $\hat{a}_2$ is

$$\begin{aligned}(\hat{\theta} &+ [-z_{\tilde{\alpha}/2} - n^{-1/2}\{\hat{a}_1{}' + (\hat{a}_3{}'/6)(z_{\tilde{\alpha}/2}^2 - 1)\}]n^{-1/2}\hat{a}_2^{1/2}, \\ \hat{\theta} &+ [z_{\tilde{\alpha}/2} - n^{-1/2}\{\hat{a}_1{}' + (\hat{a}_3{}'/6)(z_{\tilde{\alpha}/2}^2 - 1)\}]n^{-1/2}\hat{a}_2^{1/2}).\end{aligned} \tag{A15}$$

Note that the confidence interval using only the estimated usual standard error is

$$(\hat{\theta} - z_{\tilde{\alpha}/2}n^{-1/2}\hat{a}_2^{1/2}, \hat{\theta} + z_{\tilde{\alpha}/2}n^{-1/2}\hat{a}_2^{1/2}). \tag{A16}$$

For (A15) we require $\hat{a}_1'$ and $\hat{a}_3{}'$ with $\hat{a}_2$ which depend on the sample moments of the component scores up to the sixth order while for (A16) those up to the fourth order are required.

References

Abramovitch, L., & Singh, K. (1985). Edgeworth corrected pivotal statistics and the bootstrap. *Annals of Statistics*, **13**, 116–132.

Browne, M.W. (1974). Generalized least-squares estimators in the analysis of covariance structures. *South African Statistical Journal*, **8**, 1–24. Reprinted in D.J. Aigner & A.S. Goldberger (Eds.), *Latent variables in socioeconomic models* (pp.205–226). Amsterdam: North Holland, 1977.

Cronbach, L.J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, **16**, 297–334.

Cronbach, L.J., Schönemann, P., & McKie, D. (1965). Alpha coefficients for stratified-parallel tests. *Educational and Psychological Measurement*, **25**, 291–312.

Feldt, L.S. (1965). The approximate sampling distribution of Kuder-Richardson reliability coefficient twenty. *Psychometrika*, **30**, 357–370.

Feldt, L.S., & Brennan, R.L. (1989). Reliability. In R.L. Linn (Ed.), *Educational measurement* (3rd ed., pp.105–146). New York: Macmillan.

Fujikoshi, Y. (1980). Asymptotic expansions for the distributions of the sample roots under nonnormality. *Biometrika*, **67**, 45–51.

Hall, P. (1992). *The bootstrap and Edgeworth expansion.* New York: Springer. Corrected printing, 1997.

Hayashi, K., & Kamata, A. (2005). A note on the estimator of the alpha coefficient for standardized variables under normality. *Psychometrika*, **70**, 579–589.

Kistner, E.O., & Muller, K.E. (2004). Exact distributions of intraclass correlation and Cronbach's alpha with Gaussian data and general covariance. *Psychometrika*, **69**, 459–474.

Kristof, W. (1963). The statistical theory of speeded-up reliability coefficients when a test has been divided into several equivalent parts. *Psychometrika*, **28**, 221–238.

Ogasawara, H. (2005). *Asymptotic expansion of the distributions of the estimators in factor analysis under nonnormality.* Paper submitted for publication.

Ogasawara, H. (2006). Asymptotic expansion of the sample correlation coefficient under nonnormality. *Computational Statistics and Data Analysis*, **50**, 891–910.

Osburn, H.G. (2000). Coefficient alpha and related internal consistency reliability coefficients. *Psychological Methods*, **5**, 343–355.

Stuart, A., & Ord, J.K. (1994). *Kendall's advanced theory of statistics: Distribution theory* (6th ed., Vol.1). London: Arnold.

ten Berge, J.M.F., & Hofstee, W.K.B. (1999). Coefficient alpha and reliabilities of unrotated and rotated components. *Psychometrika*, **64**, 83–90.

van Zyl, J.M., Neudecker, H., & Nel, D.G. (2000). On the distribution of the maximum likelihood estimator of Cronbach's alpha. *Psychometrika*, **65**, 271–280.

Yuan, K.-H., & Bentler, P. (2002). On robustness of the normal-theory based asymptotic distributions of three reliability coefficient estimates. *Psychometrika*, **67**, 251–259.

Zhou, X.-H., & Gao, S. (2000). One-sided confidence intervals for means of positively skewed distributions. *American Statistician*, **54**, 100–104.

Zinbarg, R.E., Revelle, W., Yovel, I., & Li, W. (2005). Cronbach's α, Revelle's β, and McDonald's ω_H: Their relations with each other and two alternative conceptualizations of reliability. *Psychometrika*, **70**, 123–133.

(Received May 21 2005, Revised December 17 2005)