Available online at www.sciencedirect.com

SCIENCE DIRECT®

Journal of Multivariate Analysis 96 (2005) 136–171

Journal of Multivariate Analysis

www.elsevier.com/locate/jmva

Second-order accurate inference on eigenvalues of covariance and correlation matrices

Robert J. Boik

Department of Mathematical Sciences, Montana State University, Bozeman, MT 59717-2400, USA

Received 12 September 2003
Available online 11 November 2004

Abstract

Edgeworth expansions and saddlepoint approximations for the distributions of estimators of certain eigenfunctions of covariance and correlation matrices are developed. These expansions depend on second-, third-, and fourth-order moments of the sample covariance matrix. Expressions for and estimators of these moments are obtained. The expansions and moment expressions are used to construct second-order accurate confidence intervals for the eigenfunctions. The expansions are illustrated and the results of a small simulation study that evaluates the finite-sample performance of the confidence intervals are reported.

AMS 1991 subject classification: 62H25; 62E20

Keywords: Confidence interval; Correlation matrix; Covariance matrix; Edgeworth expansion; Eigenvalue; Principal components analysis; Saddlepoint approximation

1. Introduction

A flexible spectral model for principal components of covariance matrices from several populations was proposed by Boik [6]. This model unifies and extends the common principal component model and related models of Flury [14] and others. The spectral model also is applicable to a covariance matrix from a single population. It allows arbitrary eigenvalue multiplicities and it allows the distinct eigenvalues to be modeled parametrically or

E-mail addresses: RJBoik@math.montana.edu.
URL: http://www.math.montana.edu/∼rjboik/

0047-259X/$ - see front matter
doi:10.1016/j.jmva.2004.09.009

nonparametrically. Procedures that provide the same flexibility for modeling principle components of correlation matrices were proposed by Boik [7]. In this article, second-order accurate confidence intervals for functions of eigenvalues of covariance and correlation matrices are developed. Asymptotic expansions of the distributions of model-based estimators of the eigenvalues also are constructed.

To be more specific, let $\mathbf{y}$ be a random p-vector with positive definite covariance matrix $\boldsymbol{\Sigma}$ and correlation matrix $\boldsymbol{\Psi}$. Denote the p-vector of eigenvalues of either $\boldsymbol{\Sigma}$ or $\boldsymbol{\Psi}$ by λ. Second-order accurate confidence intervals for

$$\psi_1 = \mathbf{h}'\lambda \quad \text{and} \quad \psi_2 = \mathbf{h}'\lambda/(\mathbf{1}_p'\lambda) \tag{1}$$

are developed without assuming normality, where $\mathbf{h}$ is a p-vector of known constants and $\mathbf{1}_p$ is a p-vector of ones. The eigenfunctions ψ_1 and ψ_2 could reflect partial sums or differences among eigenvalues. For example, if $\mathbf{h}$ is selected to be $\mathbf{h} = (\mathbf{1}_a' \ \mathbf{0}')'$ and eigenvalues are ordered from largest to smallest, then ψ_1 represents the variability associated with the first a principal components and ψ_2 represents the proportion of the total variability that is associated with the first a components. If λ is the vector of eigenvalues of a correlation matrix, then $\psi_1 = p\psi_2$ and the two functions yield equivalent information. If λ is the vector of eigenvalues of a covariance matrix, however, then ψ_1 and ψ_2 yield different information. Edgeworth and saddlepoint approximations for the distributions of

$$\widehat{\psi}_1 = \mathbf{h}'\widehat{\lambda} \quad \text{and} \quad \widehat{\psi}_2 = \mathbf{h}'\widehat{\lambda}/(\mathbf{1}_p'\widehat{\lambda}) \tag{2}$$

also are developed in this article, where $\widehat{\lambda}$ is a consistent model-based estimator of λ, possibly subject to constraints. The proposed methods can be extended to arbitrary differentiable functions of λ, but attention in this article is restricted to ψ_1 and ψ_2.

The expansions and confidence intervals depend on higher-order moments of the sample covariance matrix. In Section 2, matrix expressions for second-, third-, and fourth-order moments of the sample covariance are obtained. In addition, unbiased estimators of the second- and third-order moments and consistent estimators of the fourth-order moments are constructed.

Parameterizations for $\boldsymbol{\Sigma}$ and $\boldsymbol{\Psi}$ in terms of eigenvalues and eigenvectors are briefly reviewed in Section 3. Section 4 describes Edgeworth and saddlepoint approximations for the distributions of $\widehat{\psi}_1$ and $\widehat{\psi}_2$ when sampling from multivariate normal populations. Section 5 gives asymptotically distribution free (ADF) expansions of the distributions of $\widehat{\psi}_1$ and $\widehat{\psi}_2$. Normal theory and ADF confidence intervals that are based on the expansions are described in Section 6. The asymptotic expansions and confidence intervals are illustrated in Section 7. Section 8 reports the results of a simulation study that examines the accuracy of the Edgeworth and saddlepoint approximations under normality as well as the finite sample coverage of the confidence intervals under normality and under nonnormality. The proposed second-order accurate confidence intervals show a substantial improvement in coverage probability compared to first-order accurate intervals. Expressions for certain required derivatives are available in a supplement that can be down-loaded from <http://www.math.montana.edu/~rjboik/pca_eigen/>.

2. Moments of the sample covariance matrix

It is assumed that the observable data can be represented as

$$\mathbf{Y} = \mathbf{X}\mathbf{B} + \mathbf{E}, \tag{3}$$

where $\mathbf{Y}$ is an $N \times p$ observable random matrix, $\mathbf{X}$ is an $N \times q$ matrix of known constants, $\text{rank}(\mathbf{X}) = r \leqslant q$, and $\mathbf{E}$ is an $N \times p$ unobservable matrix of random deviations. The rows of $\mathbf{E}$ are assumed to be independently and identically distributed with mean zero and variance $\mathbf{\Sigma}$. The distribution of $\mathbf{E}$ is arbitrary except that the regularity conditions described in Section 5 are assumed to be satisfied. The usual unbiased estimator of $\mathbf{\Sigma}$ is

$$\mathbf{S} = \frac{1}{n}\mathbf{Y}'\mathbf{Q}\mathbf{Y}, \quad \text{where } \mathbf{Q} = \mathbf{I}_N - \mathbf{H}_X, \quad \mathbf{H}_X = \mathbf{X}\left(\mathbf{X}'\mathbf{X}\right)^{-}\mathbf{X}' \tag{4}$$

is the perpendicular projection operator that projects onto $\mathcal{R}(\mathbf{X})$, the vector space generated by the columns of $\mathbf{X}$, and $n = N - r$. It is assumed that $\mathcal{R}(\mathbf{X}')$ does not depend on N, the nonzero eigenvalues of $\mathbf{X}'\mathbf{X}$ diverge to infinity as $N \to \infty$, and $\lim_{N\to\infty} q_{ii} = 1$ for $i = 1, \ldots, N$, where q_{ii} is the ith diagonal element of $\mathbf{Q}$.

Expansions of $\widehat{\psi}_1$ and $\widehat{\psi}_2$ depend on the moments of $\sqrt{n}(\mathbf{s} - \boldsymbol{\sigma})$, where $\mathbf{s} = \text{vec}(\mathbf{S})$ and $\boldsymbol{\sigma} = \text{vec}(\mathbf{\Sigma})$. In particular, expansions of

$$W_i \stackrel{\text{def}}{=} \sqrt{n}(\widehat{\psi}_i - \psi_i) \quad \text{and} \quad Z_i \stackrel{\text{def}}{=} W_i/\widehat{\sigma}_{W_i} \quad \text{for } i = 1, 2 \tag{5}$$

require the following moments or consistent estimators thereof:

$$\mathbf{\Omega}_{22,n} \stackrel{\text{def}}{=} nE\left[(\mathbf{s} - \boldsymbol{\sigma})(\mathbf{s} - \boldsymbol{\sigma})'\right], \qquad \mathbf{\Omega}_{222,n} \stackrel{\text{def}}{=} n^{\frac{3}{2}}E\left[(\mathbf{s} - \boldsymbol{\sigma}) \otimes (\mathbf{s} - \boldsymbol{\sigma})(\mathbf{s} - \boldsymbol{\sigma})'\right],$$
$$\mathbf{\Omega}_{42,n} \stackrel{\text{def}}{=} n\,\text{Cov}\left[\text{vec}(\widehat{\mathbf{\Omega}}_{22,n}), \mathbf{s}\right]$$

and

$$\mathbf{\Omega}_{2222,n} \stackrel{\text{def}}{=} n^2E\left[(\mathbf{s} - \boldsymbol{\sigma})(\mathbf{s} - \boldsymbol{\sigma})' \otimes (\mathbf{s} - \boldsymbol{\sigma})(\mathbf{s} - \boldsymbol{\sigma})'\right], \tag{6}$$

where $\widehat{\sigma}^2_{W_i}$ is a consistent estimator of $\text{Var}(W_i)$ and $\widehat{\mathbf{\Omega}}_{22,n}$ is an estimator of $\mathbf{\Omega}_{22,n}$. Subscripts 22, 42, 222, and 2222 refer to the order of the moments. The matrix $\mathbf{\Omega}_{22,n}$, for example, is the expectation of the product of two second-order terms in $\mathbf{Y}$. The subscript n serves as a reminder that the moments depend on the model matrix, $\mathbf{X}$, which, in turn, depends on the sample size. The quantity $\mathbf{\Omega}_{222,n}$ is $O(n^{-1/2})$, whereas the remaining moments in (6) are $O(1)$. As sample size increases, the moments $\mathbf{\Omega}_{22,n}$, $\sqrt{n}\mathbf{\Omega}_{222,n}$, $\mathbf{\Omega}_{42,n}$, and $\mathbf{\Omega}_{2222,n}$ approach $\mathbf{\Omega}_{22,\infty}$, $\mathbf{\Omega}^*_{222,\infty}$, $\mathbf{\Omega}_{42,\infty}$, and $\mathbf{\Omega}_{2222,\infty}$, respectively, where

$$\mathbf{\Omega}_{22,\infty} \stackrel{\text{def}}{=} \lim_{n\to\infty} \mathbf{\Omega}_{22,n}, \qquad \mathbf{\Omega}^*_{222,\infty} \stackrel{\text{def}}{=} \lim_{n\to\infty} n^{\frac{1}{2}}\mathbf{\Omega}_{222,n},$$
$$\mathbf{\Omega}_{42,\infty} \stackrel{\text{def}}{=} \lim_{n\to\infty} \mathbf{\Omega}_{42,n} \quad \text{and} \quad \mathbf{\Omega}_{2222,\infty} \stackrel{\text{def}}{=} \lim_{n\to\infty} \mathbf{\Omega}_{2222,n}. \tag{7}$$

In this section, matrix expressions for the moments in (6) and (7) are obtained. Unbiased estimators of $\mathbf{\Omega}_{42,n}$ and $\mathbf{\Omega}_{222,n}$ and consistent estimators of the moments in (7) are derived. Expressions for and estimators of $\mathbf{\Omega}_{22,n}$ already are known, but for completeness and to illustrate the method of construction, these results also are given.

Let $\boldsymbol{\varepsilon}$ be any row of $\mathbf{E}$ in (3) and define γ_{21}, γ_{22}, γ_{42}, and γ_{44} as

$$\gamma_{21} \stackrel{\text{def}}{=} E\left(\boldsymbol{\varepsilon} \otimes \boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'\right), \qquad \gamma_{22} \stackrel{\text{def}}{=} E\left(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}' \otimes \boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'\right),$$
$$\gamma_{42} \stackrel{\text{def}}{=} E\left(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}' \otimes \boldsymbol{\varepsilon}\boldsymbol{\varepsilon}' \otimes \boldsymbol{\varepsilon} \otimes \boldsymbol{\varepsilon}\right) \quad \text{and} \quad \gamma_{44} \stackrel{\text{def}}{=} E\left(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}' \otimes \boldsymbol{\varepsilon}\boldsymbol{\varepsilon}' \otimes \boldsymbol{\varepsilon}\boldsymbol{\varepsilon}' \otimes \boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'\right), \tag{8}$$

respectively. Consistent estimators of these quantities are obtained by substituting the observable residual $\widetilde{\boldsymbol{\varepsilon}}$ for $\boldsymbol{\varepsilon}$ and averaging. That is,

$$\widetilde{\gamma}_{21} = \frac{1}{n}\sum_{i=1}^{N}\left(\widetilde{\boldsymbol{\varepsilon}}_i \otimes \widetilde{\boldsymbol{\varepsilon}}_i\widetilde{\boldsymbol{\varepsilon}}_i'\right), \qquad \widetilde{\gamma}_{22} = \frac{1}{n}\sum_{i=1}^{N}\left(\widetilde{\boldsymbol{\varepsilon}}_i\widetilde{\boldsymbol{\varepsilon}}_i' \otimes \widetilde{\boldsymbol{\varepsilon}}_i\widetilde{\boldsymbol{\varepsilon}}_i'\right),$$
$$\widetilde{\gamma}_{42} = \frac{1}{n}\sum_{i=1}^{N}\left(\widetilde{\boldsymbol{\varepsilon}}_i\widetilde{\boldsymbol{\varepsilon}}_i' \otimes \widetilde{\boldsymbol{\varepsilon}}_i\widetilde{\boldsymbol{\varepsilon}}_i' \otimes \widetilde{\boldsymbol{\varepsilon}}_i \otimes \widetilde{\boldsymbol{\varepsilon}}_i\right)$$

and

$$\widetilde{\gamma}_{44} = \frac{1}{n}\sum_{i=1}^{N}\left(\widetilde{\boldsymbol{\varepsilon}}_i\widetilde{\boldsymbol{\varepsilon}}_i' \otimes \widetilde{\boldsymbol{\varepsilon}}_i\widetilde{\boldsymbol{\varepsilon}}_i' \otimes \widetilde{\boldsymbol{\varepsilon}}_i\widetilde{\boldsymbol{\varepsilon}}_i' \otimes \widetilde{\boldsymbol{\varepsilon}}_i\widetilde{\boldsymbol{\varepsilon}}_i'\right), \tag{9}$$

where $\widetilde{\boldsymbol{\varepsilon}}_i$ is the ith residual vector. Specifically, $\widetilde{\boldsymbol{\varepsilon}}_i'$ is the ith row of $\mathbf{QY}$, where $\mathbf{Q}$ is defined in (4).

To obtain expressions for $\boldsymbol{\Omega}_{22,n}$ and $\boldsymbol{\Omega}_{22,\infty}$, first note that $\mathbf{QY} = \mathbf{QE}$. Accordingly,

$$\mathbf{S} = \frac{1}{n}\sum_{i=1}^{N}\sum_{j=1}^{N}\boldsymbol{\varepsilon}_i q_{ij}\boldsymbol{\varepsilon}_j' \quad \text{and} \quad \widetilde{\boldsymbol{\varepsilon}}_j = \sum_{i=1}^{N}\boldsymbol{\varepsilon}_i q_{ij}, \tag{10}$$

where $\boldsymbol{\varepsilon}_i'$ is the ith row of $\mathbf{E}$ and q_{ij} is the ijth component of $\mathbf{Q}$. Substituting the expressions for $\widetilde{\boldsymbol{\varepsilon}}_i$ and $\mathbf{S}$ in (10) into $\widetilde{\gamma}_{22}$, $\mathbf{ss}'$, and $(\mathbf{S} \otimes \mathbf{S})$, and then taking expectations reveals that

$$\begin{aligned} &E(\widetilde{\gamma}_{22}\ \mathbf{ss}'\ \mathbf{N}_p[\mathbf{S} \otimes \mathbf{S})])' \\ &= \left\{\frac{1}{n}\begin{pmatrix} c_2 & [c_1 - c_2] & [c_1 - c_2] \\ \frac{c_1}{n} & \left[n - \frac{c_1}{n}\right] & \left[1 - \frac{c_1}{n}\right] \\ \frac{c_1}{n} & \left[1 - \frac{c_1}{n}\right] & \left[n + 1 - \frac{2c_1}{n}\right] \end{pmatrix} \otimes \mathbf{I}_{p^2}\right\}\begin{pmatrix} \gamma_{22} \\ \boldsymbol{\sigma}\boldsymbol{\sigma}' \\ \mathbf{N}_p[\boldsymbol{\Sigma} \otimes \boldsymbol{\Sigma}] \end{pmatrix}, \end{aligned} \tag{11}$$

where $c_1 = \text{tr}(\mathbf{Q}^{\odot 2})$, $c_2 = \mathbf{1}_N'\mathbf{Q}^{\odot 4}\mathbf{1}_N$, $\mathbf{N}_p = (\mathbf{I}_{p^2} + \mathbf{I}_{(p,p)})/2$, $\mathbf{I}_{(a,b)}$ is the commutation matrix [27], and $\odot$ is the elementwise operator. For example, if $\mathbf{a}$ is a $q \times 1$ vector, then $e^{\odot \mathbf{a}'} = (e^{a_1} \cdots e^{a_q})$. The commutation matrix, $\mathbf{I}_{(a,b)}$ is denoted by $\mathbf{K}_{ba}$ in Magnus and Neudecker [28,29, Section 3.7]. By using (11) and the definitions in (6) and (7) it is readily shown that

$$\boldsymbol{\Omega}_{22,n} = \frac{c_1}{n}\left(\gamma_{22} - \boldsymbol{\sigma}\boldsymbol{\sigma}'\right) + \left(1 - \frac{c_1}{n}\right)2\mathbf{N}_p(\boldsymbol{\Sigma} \otimes \boldsymbol{\Sigma})$$

and

$$\boldsymbol{\Omega}_{22,\infty} = \gamma_{22} - \boldsymbol{\sigma}\boldsymbol{\sigma}'. \tag{12}$$

Solving (11) for the moment matrices yields

$$
\begin{aligned}
&(\gamma_{22}\ \boldsymbol{\sigma}\boldsymbol{\sigma}'\ 2\mathbf{N}_p[\boldsymbol{\Sigma}\otimes\boldsymbol{\Sigma}])' \\
&= \left[a \begin{pmatrix} n(n+2)-3c_1 & n(c_2-c_1) & n(c_2-c_1) \\ -c_1 & \dfrac{n(n+1)c_2-2c_1^2}{n-1} & \dfrac{c_1^2-nc_2}{n-1} \\ -2c_1 & \dfrac{2(c_1^2-nc_2)}{n-1} & \dfrac{n^2c_2-c_1^2}{n-1} \end{pmatrix} \otimes \mathbf{I}_{p^2} \right] \\
&\quad \times E \begin{pmatrix} \widetilde{\gamma}_{22} \\ \mathbf{s}\mathbf{s}' \\ 2\mathbf{N}_p[\mathbf{S}\otimes\mathbf{S})] \end{pmatrix},
\end{aligned} \tag{13}
$$

where $a = n/[n(n+2)c_2 - 3c_1^2]$, and c_1 and c_2 are defined in (11). It follows from (12) and (13) that

$$
\widehat{\boldsymbol{\Omega}}_{22,n} = a_1\widetilde{\gamma}_{22} + a_2\mathbf{s}\mathbf{s}' + a_3 2\mathbf{N}_p(\mathbf{S}\otimes\mathbf{S}) \tag{14}
$$

is an unbiased estimator of $\boldsymbol{\Omega}_{22,n}$ where

$$
a_1 = anc_1, \quad a_2 = \frac{-a[2nc_2+(n-3)c_1^2]}{(n-1)} \quad \text{and} \quad a_3 = \frac{-an(c_1^2-nc_2)}{(n-1)},
$$

a is defined in (13) and c_1 and c_2 are defined in (11). If $\mathbf{X} = \mathbf{1}_N$, then $n = N-1$ and the coefficients simplify to

$$
a_1 = \frac{n^2}{a_0}, \quad a_2 = -\frac{n(n^2-2)}{(n+1)a_0} \quad \text{and} \quad a_3 = -\frac{n^2}{(n+1)a_0},
$$

where $a_0 = (n-1)(n-2)$. Also, it is apparent that

$$
\widetilde{\boldsymbol{\Omega}}_{22,\infty} = \widetilde{\gamma}_{22} - \mathbf{s}\mathbf{s}' \tag{15}
$$

is a consistent estimator of $\boldsymbol{\Omega}_{22,\infty}$, where $\widetilde{\gamma}_{22}$ is defined in (9). Browne [10] and Koning et al. [22] derived the estimator in (14) for the special case when $\mathbf{X} = \mathbf{1}_N$. Boik [5, Theorem 5], derived the estimator in (14) for general $\mathbf{X}$ by a slightly different method than above.

The methods that were used to obtain the expressions in (12) and the estimators in (14) and (15) can, in principle, be extended to moments of any order. The derivations are rather tedious, however, so comparable results are obtained for third-order moments of $\mathbf{S}$ (sixth-order moments of $\boldsymbol{\varepsilon}$) only. These results are summarized in Theorem 1.

Theorem 1. *Matrix expressions for third-order moments of* $\mathbf{S}$ *are given by*

$$
\boldsymbol{\Omega}_{222,n} = \sum_{i=1}^{12} a_{222,i}\mathbf{M}_i \quad \textit{and} \quad \boldsymbol{\Omega}_{42,n} = \sum_{i=1}^{12} a_{42,i}\mathbf{M}_i,
$$

where the sixth-order moments, $\{\mathbf{M}_i\}_{i=1}^{12}$*, as well as the coefficients* c_i*,* $a_{222,i}$*, and* $a_{42,i}$ *for* $i = 1, \ldots, 12$ *are defined in Table* 1.

Table 1
Sixth-order moments and associated coefficients

i	$\mathbf{M}_i$	c_i	$n^{\frac{3}{2}} a_{222,i}$	$n^2 a_{42,i}$
1	γ_{42}	$\operatorname{tr}\left(\mathbf{Q}^{\odot 2}\right)$	c_8	na_1c_5 $+c_8(a_2+2a_3)$
2	$\left(\boldsymbol{\sigma}\otimes\boldsymbol{\sigma}\boldsymbol{\sigma}'\right)$	$\mathbf{1}_N'\mathbf{Q}^{\odot 4}\mathbf{1}_N$	$2c_8$	$n^2[a_1(c_1-c_2)+2a_3+c_1]$ $+n^3a_2+2na_1(c_5-c_9)$ $-nc_1(3a_2+2a_3)$ $+2[a_2c_8+2a_3(c_8-c_1)]$
3	$2\mathbf{N}_{p^2}\times$ $\left(\boldsymbol{\Sigma}\otimes\boldsymbol{\Sigma}\otimes\boldsymbol{\sigma}\right)2\mathbf{N}_p$	$\mathbf{q}'\mathbf{Q}\mathbf{q}$	$2c_8$ $-2c_1$	$+na_1[c_1-c_9+2(c_5-c_2)]$ $+n^2a_2+n(2a_3-a_2c_1)$ $+2[a_2(c_8-c_1)+a_3(2c_8-3c_1)]$
4	$\operatorname{vec}\left[(\boldsymbol{\Sigma}\otimes\boldsymbol{\Sigma})2\mathbf{N}_p\right]\boldsymbol{\sigma}'$	$\mathbf{1}_N'\mathbf{Q}^{\odot 3}\mathbf{1}_N$	$2c_8$ $-2c_1$	$n^3(a_3-1)$ $+n^2[a_1(c_1-c_2)+a_2+a_3+c_1]$ $+n[2a_1(c_5-c_9)-c_1(a_2+4a_3)]$ $+2[a_2(c_8-c_1)+a_3(2c_8-c_1)]$
5	$\left(2\mathbf{N}_p\otimes 2\mathbf{N}_p\right)\times$ $\left(\boldsymbol{\Sigma}\otimes\boldsymbol{\sigma}\otimes\boldsymbol{\Sigma}\right)2\mathbf{N}_p$	$\mathbf{q}'\mathbf{Q}^{\odot 4}\mathbf{1}_N$	n $-3c_1$ $+2c_8$	$na_1[c_1-c_9+2(c_5-c_2)]$ $+n^2a_3+n[a_2+a_3(1-c_1)]$ $+a_2(2c_8-3c_1)+a_3(4c_8-5c_1)$
6	$2\mathbf{N}_{p^2}(\gamma_{22}\otimes\boldsymbol{\sigma})$	$\mathbf{1}_N'\mathbf{Q}^{\odot 6}\mathbf{1}_N$	$-c_8$	$n[a_1(c_9-c_5)+a_2c_1]$ $+2a_3(c_1-c_8)-a_2c_8$
7	$\operatorname{vec}(\gamma_{22})\boldsymbol{\sigma}'$	$\mathbf{1}_N'\left(\mathbf{Q}^{\odot 3}\right)^2\mathbf{1}_N$	$-c_8$	$n^2(a_1c_2-c_1)$ $+n[c_1(2a_3+a_2)-a_1c_5]$ $-c_8(a_2+2a_3)$
8	$2\mathbf{N}_{p^2}(2\mathbf{N}_p\otimes\mathbf{I}_{p^2})\times$ $\left[\boldsymbol{\Sigma}\otimes\operatorname{dvec}(\gamma_{22},p^3,p)\right]$ $\times 2\mathbf{N}_p$	$\operatorname{tr}\left(\mathbf{Q}^{\odot 3}\right)$	c_1 $-c_8$	$na_1(c_2-c_5)$ $+(a_2+2a_3)(c_1-c_8)$
9	$(2\mathbf{N}_p\otimes 2\mathbf{N}_p)\times$ $(\mathbf{I}_p\otimes\mathbf{I}_{(p,p)}\otimes\mathbf{I}_p)\times$ $(\gamma_{22}\otimes\boldsymbol{\sigma})$	$\mathbf{q}'\mathbf{Q}^{\odot 2}\mathbf{q}$	c_1 $-c_8$	$n[a_1(c_9-c_5)+a_3c_1]$ $+a_2(c_1-c_8)+a_3(c_1-2c_8)$
10	$2\mathbf{N}_{p^2}(2\mathbf{N}_p\otimes\mathbf{I}_{p^2})\times$ $(\gamma_{21}'\otimes\operatorname{vec}\gamma_{21})$	$\mathbf{q}'\mathbf{Q}\mathbf{Q}^{\odot 3}\mathbf{1}_N$	c_3 $-c_8$	$na_1(c_{10}-c_5)$ $+(a_2+2a_3)(c_3-c_8)$
11	$(\gamma_{21}\otimes\gamma_{21})2\mathbf{N}_p$	$\operatorname{tr}\left[\left(\mathbf{Q}^{\odot 2}\right)^2\mathbf{Q}\right]$	c_3 $-c_8$	$na_1(c_{11}-c_5)$ $+a_2(c_3-c_8)+2a_3(c_4-c_8)$
12	$(2\mathbf{N}_p\otimes 2\mathbf{N}_p)\times$ $(\mathbf{I}_p\otimes\mathbf{I}_{(p,p)}\otimes\mathbf{I}_p)\times$ $(\gamma_{21}\otimes\gamma_{21})$	$\operatorname{tr}\left[\left(\mathbf{Q}\mathbf{Q}^{\odot 2}\right)^2\right]$	c_4 $-c_8$	$na_1(c_{11}-c_5)$ $+a_2(c_4-c_8)$ $+a_3(c_3+c_4-2c_8)$

$\mathbf{Q}$ is given in (4), $\mathbf{q}=(Q_{11}\ Q_{22}\cdots Q_{NN})'$, and $\{a_i\}_{i=1}^{3}$ are defined in (12).

Denote the sample version of the ith *sixth-order moment by* $\widetilde{\mathbf{M}}_i$ *and define* $\widetilde{\mathbf{M}}$ *as* $\widetilde{\mathbf{M}} \stackrel{\text{def}}{=} (\widetilde{\mathbf{M}}_1' \cdots \widetilde{\mathbf{M}}_{12}')'$. *Specifically,* $\widetilde{\mathbf{M}}_i$ *is obtained by replacing* $\boldsymbol{\Sigma}$, γ_{21}, γ_{22}, *and* γ_{42} *in* $\mathbf{M}_i$ **by** $\mathbf{S}$, $\widetilde{\gamma}_{21}$, $\widetilde{\gamma}_{22}$, *and* $\widetilde{\gamma}_{42}$, *respectively. The expectations of the sample sixth-order moments are*

$$E\left(\widetilde{\mathbf{M}}\right) = \left(\mathbf{W} \otimes \mathbf{I}_{p^4}\right)\mathbf{M}, \quad \textit{where } \mathbf{M} \stackrel{\text{def}}{=} (\mathbf{M}_1' \cdots \mathbf{M}_{12}')',$$

and the components of the 12×12 *coefficient matrix* $\mathbf{W} = \{w_{ij}\}$ ıare given in Appendix A. *An unbiased and consistent estimator of* $\mathbf{M}$ *is given by* $\widehat{\mathbf{M}} = \left(\mathbf{W}^{-1} \otimes \mathbf{I}_{p^4}\right)\widetilde{\mathbf{M}}$. *The estimator* $\widehat{\mathbf{M}}$ *can be used along with the coefficients in Table* 1 *to obtain unbiased estimators of* $\boldsymbol{\Omega}_{222,n}$ *and* $\boldsymbol{\Omega}_{42,n}$. *For example,*

$$\widehat{\boldsymbol{\Omega}}_{222,n} = n^{-\frac{3}{2}}\left(c_8\widehat{\mathbf{M}}_1 + 2c_8\widehat{\mathbf{M}}_2 + \cdots\right).$$

Simplifications of the coefficients for the unbiased estimators $\widehat{\boldsymbol{\Omega}}_{222,n}$ and $\widehat{\boldsymbol{\Omega}}_{42,n}$ exist under special conditions. For example, if $\mathbf{X} = \mathbf{1}_N$, where $\mathbf{X}$ is the model matrix in (3), then

$$\widehat{\boldsymbol{\Omega}}_{222,n} = \left(\mathbf{w}_{222}' \otimes \mathbf{I}_{p^4}\right)\widetilde{\mathbf{M}} \quad \text{and} \quad \widehat{\boldsymbol{\Omega}}_{42,n} = \left(\mathbf{w}_{42}' \otimes \mathbf{I}_{p^4}\right)\widetilde{\mathbf{M}}, \tag{16}$$

where the 12×1 coefficient vectors $\mathbf{w}_{222}$ and $\mathbf{w}_{42}$ are given in Table A.1 in Appendix A.

An unbiased estimator of $\boldsymbol{\Omega}_{2222,n}$ in (6) can be constructed by the methods employed in Theorem 1, but it is sufficient for present purposes to construct a consistent estimator of $\boldsymbol{\Omega}_{2222,\infty}$. Using $n^{\frac{1}{2}}(\mathbf{s} - \boldsymbol{\sigma}) \xrightarrow{\text{dist}} \mathrm{N}(\mathbf{0}, \boldsymbol{\Omega}_{22,\infty})$, it is readily shown that

$$\boldsymbol{\Omega}_{2222,n} = 2\mathbf{N}_{p^2}(\boldsymbol{\Omega}_{22,n} \otimes \boldsymbol{\Omega}_{22,n}) + \mathrm{vec}(\boldsymbol{\Omega}_{22,n})\left[\mathrm{vec}(\boldsymbol{\Omega}_{22,n})\right]' + O\left(n^{-1}\right)$$

and that

$$\boldsymbol{\Omega}_{2222,\infty} = 2\mathbf{N}_{p^2}(\boldsymbol{\Omega}_{22,\infty} \otimes \boldsymbol{\Omega}_{22,\infty}) + \mathrm{vec}(\boldsymbol{\Omega}_{22,\infty})\left[\mathrm{vec}(\boldsymbol{\Omega}_{22,\infty})\right]'.$$

Accordingly,

$$\widetilde{\boldsymbol{\Omega}}_{2222,\infty} = 2\mathbf{N}_{p^2}\left(\widetilde{\boldsymbol{\Omega}}_{22,\infty} \otimes \widetilde{\boldsymbol{\Omega}}_{22,\infty}\right) + \mathrm{vec}\left(\widetilde{\boldsymbol{\Omega}}_{22,\infty}\right)\left[\mathrm{vec}\left(\widetilde{\boldsymbol{\Omega}}_{22,\infty}\right)\right]' \tag{17}$$

is a consistent estimator of $\boldsymbol{\Omega}_{2222,\infty}$, where $\widetilde{\boldsymbol{\Omega}}_{22,\infty}$ is given in (15). Note, the estimator in (17) remains consistent if $\widetilde{\boldsymbol{\Omega}}_{22,\infty}$ is replaced by $\widehat{\boldsymbol{\Omega}}_{22,n}$.

The focus in this section is on the moments of $\mathbf{S}$ when sampling from nonnormal distributions. Nonetheless, it still is of interest to examine the moments under the assumption of multivariate normality of $\mathbf{Y}$. Boik [6, Theorem A.2] gave expressions for $\boldsymbol{\Omega}_{222,n}$ and $\boldsymbol{\Omega}_{2222,n}$ under normality. These results, along with known results for $\boldsymbol{\Omega}_{22,n}$ and a new result for $\boldsymbol{\Omega}_{42,n}$ are given below:

$$\begin{aligned}
\boldsymbol{\Omega}_{22,n} &= \boldsymbol{\Omega}_{22,\infty} = 2\mathbf{N}_p(\boldsymbol{\Sigma} \otimes \boldsymbol{\Sigma}),\\
\boldsymbol{\Omega}_{222,n} &= n^{-\frac{1}{2}}\boldsymbol{\Omega}^*_{222,\infty} = n^{-\frac{1}{2}}(2\mathbf{N}_p \otimes 2\mathbf{N}_p)(\boldsymbol{\Sigma} \otimes \boldsymbol{\sigma} \otimes \boldsymbol{\Sigma})2\mathbf{N}_p,\\
\boldsymbol{\Omega}_{42,n} &= \boldsymbol{\Omega}_{42,\infty} = n^{\frac{1}{2}}\boldsymbol{\Omega}_{222,n}
\end{aligned}$$

and

$$\begin{aligned}\mathbf{\Omega}_{2222,n} &= \mathbf{\Omega}_{2222,\infty} = 2\mathbf{N}_{p^2}(2\mathbf{N}_p \otimes 2\mathbf{N}_p)(\mathbf{\Sigma} \otimes \mathbf{\Sigma} \otimes \mathbf{\Sigma} \otimes \mathbf{\Sigma}) \\ &\quad + \operatorname{vec}\left[2\mathbf{N}_p(\mathbf{\Sigma} \otimes \mathbf{\Sigma})\right]\left\{\operatorname{vec}\left[2\mathbf{N}_p(\mathbf{\Sigma} \otimes \mathbf{\Sigma})\right]\right\}' + O\left(n^{-1}\right).\end{aligned}$$

See [6, Theorem A.2] for an explicit expression for the $O(n^{-1})$ term in $\mathbf{\Omega}_{2222,n}$. Consistent estimators under normality can be obtained by substituting $\mathbf{S}$ for $\mathbf{\Sigma}$ in the above equations.

3. Parameterizations of covariance matrices

3.1. Eigenvalues of the correlation matrix

In the remainder of this article it is assumed that the correlation matrix $\mathbf{\Psi}$ is irreducible. That is, it cannot be permuted into a nontrivial block diagonal matrix and, therefore, Theorem 2 in [7] is satisfied for $k = 1$. More generally, if a correlation matrix can be permuted into a nontrivial block diagonal matrix, then the following parameterization must be applied separately to each of the diagonal blocks.

Following Boik [7], the covariance matrix $\mathbf{\Sigma}$ is parameterized as a function of a $\dot{\nu}$-dimensional vector, $\boldsymbol{\theta}$, where

$$\mathbf{\Sigma} = \boldsymbol{\sigma}_D \mathbf{\Psi} \boldsymbol{\sigma}_D = \mathbf{\Sigma}(\boldsymbol{\theta}), \quad \mathbf{\Psi} = \mathbf{\Gamma}\mathbf{\Lambda}\mathbf{\Gamma}', \quad \boldsymbol{\theta} = \begin{pmatrix} \boldsymbol{\tau} \\ \boldsymbol{\mu} \\ \boldsymbol{\varphi} \end{pmatrix}, \quad \boldsymbol{\nu} = \begin{pmatrix} \dim(\boldsymbol{\tau}) \\ \dim(\boldsymbol{\mu}) \\ \dim(\boldsymbol{\varphi}) \end{pmatrix}, \tag{18}$$

$\dot{\nu} = \dim(\boldsymbol{\theta})$, $\boldsymbol{\sigma}_D = \operatorname{Diag}(\boldsymbol{\sigma}_d)$, $\boldsymbol{\sigma}_d = \boldsymbol{\sigma}_d(\boldsymbol{\tau})$ is the p-vector of standard deviations of the elements of $\mathbf{y}$, $\mathbf{\Gamma} = \mathbf{\Gamma}\mathbf{G}(\boldsymbol{\mu}, \boldsymbol{\varphi})\big|_{\boldsymbol{\mu}=\mathbf{0}}$, $\mathbf{\Lambda} = \operatorname{Diag}(\lambda)$, $\lambda = \lambda(\boldsymbol{\varphi})$, and $\mathbf{\Gamma}$ is a matrix of unit-norm orthogonal eigenvectors. The dimension of $\boldsymbol{\mu}$ is $\nu_2 = (p^2 - \mathbf{m}'\mathbf{m})/2 - (p-1)$, where $\mathbf{m}$ is a vector whose elements are the multiplicities of the distinct eigenvalues of $\mathbf{\Psi}$. The vector of standard deviations is parameterized as $\boldsymbol{\sigma}_d = \mathbf{T}_1 \exp\{\odot \mathbf{T}_2 \boldsymbol{\tau}\}$, where $\mathbf{T}_1 \colon p \times q_1$ and $\mathbf{T}_2 \colon q_1 \times \nu_1$ are full column-rank design matrices of known constants. Details on the parameterization of $\mathbf{G}$ in terms of $(\boldsymbol{\mu}, \boldsymbol{\varphi})$ can be found in [7, Section 2.3].

The vector of eigenvalues is parameterized as

$$\lambda = p \left(\frac{\mathbf{T}_3 \exp\{\odot \mathbf{T}_4 \boldsymbol{\varphi}\}}{\mathbf{1}_p' \mathbf{T}_3 \exp\{\odot \mathbf{T}_4 \boldsymbol{\varphi}\}} \right), \tag{19}$$

where $\mathbf{T}_3 \colon p \times q_3$ and $\mathbf{T}_4 \colon q_3 \times q_4$ are full column-rank design matrices of known constants. Without loss of generality, it can be assumed that $\mathbf{T}_4$ satisfies $\mathbf{1}_{q_3}' \mathbf{T}_4 = \mathbf{0}$. If this condition is not satisfied, then replace $\mathbf{T}_4$ by any matrix whose columns form a basis for $\mathcal{R}\left[(\mathbf{I}_{q_3} - \mathbf{1}_{q_3} q_3^{-1} \mathbf{1}_{q_3}')\mathbf{T}_4\right]$. If no restrictions are placed on the parameter vector $\boldsymbol{\varphi}$, then $\nu_3 = q_4$. More generally, λ can be represented by (19), subject to the constraint $\mathbf{C}_1'\lambda = \mathbf{c}_0$, where $\mathbf{C}_1$ is a known matrix of constants and $\mathbf{c}_0$ is a known vector of constants. In this case, $\boldsymbol{\varphi}$ in (19) is replaced by $\boldsymbol{\xi}$, where $\boldsymbol{\xi}$ is an implicit function of $\boldsymbol{\varphi}$ and $\nu_3 = \dim(\boldsymbol{\varphi}) < \dim(\boldsymbol{\xi})$.

For details, see [7, Section 2.4]. Derivatives of λ with respect to $\boldsymbol{\varphi}$ are denoted by

$$\mathbf{D}^{(1)}_{\lambda:\boldsymbol{\varphi}} \stackrel{\text{def}}{=} \frac{\partial\,\lambda}{\partial\,\boldsymbol{\varphi}'}, \quad \mathbf{D}^{(2)}_{\lambda:\boldsymbol{\varphi},\boldsymbol{\varphi}} \stackrel{\text{def}}{=} \frac{\partial^2\lambda}{\partial\,\boldsymbol{\varphi}'\otimes\partial\,\boldsymbol{\varphi}'},$$

and

$$\mathbf{D}^{(3)}_{\lambda:\boldsymbol{\varphi},\boldsymbol{\varphi},\boldsymbol{\varphi}} = \frac{\partial^3\lambda}{\partial\,\boldsymbol{\varphi}'\otimes\partial\,\boldsymbol{\varphi}'\otimes\partial\,\boldsymbol{\varphi}'}. \tag{20}$$

Expressions for these derivatives are given in the supplement.

3.2. Eigenvalues of the covariance matrix

If interest is in the eigenvalues of the covariance matrix, then $\boldsymbol{\Sigma}$ can be parameterized as a function of a $\dot{v}$-dimensional vector, $\boldsymbol{\theta}$, where

$$\boldsymbol{\Sigma} = \boldsymbol{\Gamma}\boldsymbol{\Lambda}\boldsymbol{\Gamma}' = \boldsymbol{\Sigma}(\boldsymbol{\theta}), \quad \boldsymbol{\theta} = \begin{pmatrix}\boldsymbol{\mu}\\ \boldsymbol{\varphi}\end{pmatrix}, \quad \boldsymbol{v} = \begin{pmatrix}\dim(\boldsymbol{\mu})\\ \dim(\boldsymbol{\varphi})\end{pmatrix}, \quad \boldsymbol{\Gamma} = \boldsymbol{\Gamma}\mathbf{G}(\boldsymbol{\mu})\big|_{\boldsymbol{\mu}=\mathbf{0}}, \tag{21}$$

$\dot{v} = \dim(\boldsymbol{\theta})$, and $\boldsymbol{\Lambda} = \boldsymbol{\Lambda}(\boldsymbol{\varphi})$. The dimension of $\boldsymbol{\mu}$ is $(p^2 - \mathbf{m}'\mathbf{m})/2$, where $\mathbf{m}$ is a vector whose elements are the multiplicities of the distinct eigenvalues. Details on the parameterization $\mathbf{G}$ in terms of $\boldsymbol{\mu}$ can be found in [6, Section 2.3].

If interest is in $\psi_1 = \mathbf{h}'\lambda$, then a suitable parameterization for λ is

$$\lambda = \mathbf{T}_1 \exp\{\odot\mathbf{T}_2\boldsymbol{\varphi}\}, \tag{22}$$

where $\mathbf{T}_1$: $p \times q_1$ and $\mathbf{T}_2$: $q_1 \times q_2$ are full column-rank design matrices of known constants. If no restrictions are placed on the parameter vector $\boldsymbol{\varphi}$, then $v_2 = q_2$. If $\boldsymbol{\varphi}$ must satisfy $\mathbf{C}_1'\lambda = \mathbf{c}_0$, then λ can be parameterized as (22) except that $\boldsymbol{\varphi}$ is replaced by $\boldsymbol{\xi}$, where $\boldsymbol{\xi}$ is an implicit function of $\boldsymbol{\varphi}$ and $v_2 = \dim(\boldsymbol{\varphi}) < \dim(\boldsymbol{\xi})$. Derivatives of λ with respect to $\boldsymbol{\varphi}$ are denoted as in (20). Details and expressions for these derivatives are given in the supplement.

If interest is in $\psi_2 = \mathbf{h}'\lambda/\operatorname{tr}(\boldsymbol{\Sigma})$, then a suitable parameterization for λ is

$$\lambda = \varphi_1\left(\frac{\mathbf{T}_1\exp\{\odot\mathbf{T}_2\boldsymbol{\varphi}_2\}}{\mathbf{1}_p'\mathbf{T}_1\exp\{\odot\mathbf{T}_2\boldsymbol{\varphi}_2\}}\right), \quad \text{where } \varphi_1 = \operatorname{tr}(\boldsymbol{\Sigma}) \text{ and } \boldsymbol{\varphi} = \begin{pmatrix}\varphi_1\\ \boldsymbol{\varphi}_2\end{pmatrix}. \tag{23}$$

It can be assumed, without loss of generality, that $\mathbf{T}_2$ satisfies $\mathbf{1}_{q_1}'\mathbf{T}_2 = \mathbf{0}$. If no restrictions are placed on the parameter vector $\boldsymbol{\varphi}_2$, then $v_2 = q_2 + 1$. If $\boldsymbol{\varphi}_2$ must satisfy $\mathbf{C}_1'\lambda\varphi_1^{-1} = \mathbf{c}_0$, then λ can be parameterized as (23) except that $\boldsymbol{\varphi}_2$ is replaced by $\boldsymbol{\xi}$, where $\boldsymbol{\xi}$ is an implicit function of $\boldsymbol{\varphi}_2$ and $v_2 - 1 = \dim(\boldsymbol{\varphi}_2) < \dim(\boldsymbol{\xi})$. Derivatives of λ with respect to $\boldsymbol{\varphi}_2$ are denoted by

$$\mathbf{D}^{(1)}_{\lambda:\boldsymbol{\varphi}_2} \stackrel{\text{def}}{=} \frac{\partial\,\lambda}{\partial\,\boldsymbol{\varphi}_2'}, \quad \mathbf{D}^{(2)}_{\lambda:\boldsymbol{\varphi}_2,\boldsymbol{\varphi}_2} \stackrel{\text{def}}{=} \frac{\partial^2\lambda}{\partial\,\boldsymbol{\varphi}_2'\otimes\partial\,\boldsymbol{\varphi}_2'},$$

and

$$\mathbf{D}^{(3)}_{\lambda:\boldsymbol{\varphi}_2,\boldsymbol{\varphi}_2,\boldsymbol{\varphi}_2} \stackrel{\text{def}}{=} \frac{\partial^3 \lambda}{\partial \boldsymbol{\varphi}_2' \otimes \partial \boldsymbol{\varphi}_2' \otimes \partial \boldsymbol{\varphi}_2'}. \tag{24}$$

Details and expressions for these derivatives are given in the supplement.

4. Edgeworth and saddlepoint expansions under normality

In this section, Edgeworth and saddlepoint expansions are constructed for the density of $\hat{\psi}_1 = \mathbf{h}'\widehat{\lambda}$, where $\widehat{\lambda}$ is the maximum likelihood estimator of the vector of eigenvalues of the correlation or covariance matrix based on a sample of size $N = n + r$ from a multivariate normal distribution and r is the rank of $\mathbf{X}$ in (3). The eigenvalues are parameterized as (19) if interest is in correlation matrices or as (22) if interest is in covariance matrices. Modifications for the expansion of the density of $\hat{\psi}_2 = \mathbf{h}'\widehat{\lambda}/(\mathbf{1}_p'\widehat{\lambda})$, where $\widehat{\lambda}$ is the maximum likelihood estimator of the vector of eigenvalues of the covariance matrix are described in Section 4.3. See [3, Chapter 4] and Reid [34] for descriptions of Edgeworth and saddlepoint (tilted Edgeworth) expansions. First-order asymptotic distributions of $\hat{\psi}_1$ and $\hat{\psi}_2$ in the case of covariance matrices and under multivariate normality were obtained by Anderson [1].

4.1. Edgeworth expansion under normality

Let $\mathbf{S}$ be a sample covariance matrix whose distribution is Wishart: $n\mathbf{S} \sim W_p(n, \boldsymbol{\Sigma})$, where $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}(\boldsymbol{\theta})$. Denote the corresponding log likelihood function as $\ell(\boldsymbol{\theta})$ and its ith derivative as $\boldsymbol{\ell}_i(\boldsymbol{\theta})$. Specifically,

$$\boldsymbol{\ell}_1 \stackrel{\text{def}}{=} \frac{\partial \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}, \quad \boldsymbol{\ell}_2 \stackrel{\text{def}}{=} \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}' \otimes \partial \boldsymbol{\theta}} \quad \text{and} \quad \boldsymbol{\ell}_3 \stackrel{\text{def}}{=} \frac{\partial^3 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}' \otimes \partial \boldsymbol{\theta}' \otimes \partial \boldsymbol{\theta}}. \tag{25}$$

These derivatives depend on $\boldsymbol{\theta}$ only through the derivatives of $\operatorname{vec} \boldsymbol{\Sigma}$ with respect to $\boldsymbol{\theta}$. The latter derivatives are denoted as

$$\mathbf{F}^{(1)} \stackrel{\text{def}}{=} \left.\frac{\partial \operatorname{vec} \boldsymbol{\Sigma}}{\partial \boldsymbol{\theta}'}\right|_{\boldsymbol{\mu}=0}, \quad \mathbf{F}^{(2)} \stackrel{\text{def}}{=} \left.\frac{\partial^2 \operatorname{vec} \boldsymbol{\Sigma}}{\partial \boldsymbol{\theta}' \otimes \partial \boldsymbol{\theta}'}\right|_{\boldsymbol{\mu}=0} \quad \text{and} \quad \mathbf{F}^{(3)} \stackrel{\text{def}}{=} \left.\frac{\partial^3 \operatorname{vec} \boldsymbol{\Sigma}}{\partial \boldsymbol{\theta}' \otimes \partial \boldsymbol{\theta}' \otimes \partial \boldsymbol{\theta}'}\right|_{\boldsymbol{\mu}=0}.$$

For notational convenience, the following definitions are used:

$$\ddot{\mathbf{F}}^{(1)} \stackrel{\text{def}}{=} \left(\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{\Sigma}^{-1}\right) \mathbf{F}^{(1)} \quad \text{and} \quad \ddot{\mathbf{F}}^{(2)} \stackrel{\text{def}}{=} \left(\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{\Sigma}^{-1}\right) \mathbf{F}^{(2)}. \tag{26}$$

Define $\mathbf{Z}_j$ as

$$\mathbf{Z}_j \stackrel{\text{def}}{=} \sqrt{n}\left(n^{-1}\boldsymbol{\ell}_j - \mathbf{K}_j\right), \quad \text{where } \mathbf{K}_j = n^{-1}E(\boldsymbol{\ell}_j). \tag{27}$$

For example,

$$\mathbf{Z}_1 = \tfrac{1}{2}\,\ddot{\mathbf{F}}^{(1)\prime}\sqrt{n}(\mathbf{s} - \boldsymbol{\sigma}), \quad \mathbf{K}_1 = \mathbf{0}, \quad \text{and} \quad \mathbf{K}_2 = -\tfrac{1}{2}\,\mathbf{F}^{(1)\prime}\ddot{\mathbf{F}}^{(1)}.$$

Explicit expressions for log likelihood derivatives, $\mathbf{Z}_2$, $\mathbf{Z}_3$, $\mathbf{K}_3$, and $\mathbf{K}_4$ are given in the supplement. Note that $\mathbf{Z}_j = O_p(1)$ for all j, $\sqrt{n}\mathbf{Z}_1$ is the score function and that $-\mathbf{K}_2 = \bar{\mathbf{I}}_{\boldsymbol{\theta}}$ is the average Fisher information. The MLE of $\boldsymbol{\theta}$ can be expanded in tensor notation as (7.10) in [31] or in vector notation as

$$\sqrt{n}(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}) = \widehat{\boldsymbol{\delta}}_0 + n^{-\frac{1}{2}}\widehat{\boldsymbol{\delta}}_1 + n^{-1}\widehat{\boldsymbol{\delta}}_2 + O_p\left(n^{-\frac{3}{2}}\right), \tag{28}$$

where

$$\widehat{\boldsymbol{\delta}}_0 = \bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1}\mathbf{Z}_1, \quad \widehat{\boldsymbol{\delta}}_1 = \bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1}[\mathbf{Z}_2\widehat{\boldsymbol{\delta}}_0 + \tfrac{1}{2}\mathbf{K}_3(\widehat{\boldsymbol{\delta}}_0 \otimes \widehat{\boldsymbol{\delta}}_0)]$$

and

$$\widehat{\boldsymbol{\delta}}_2 = \bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1}[\mathbf{Z}_2\widehat{\boldsymbol{\delta}}_1 + \tfrac{1}{2}\mathbf{Z}_3(\widehat{\boldsymbol{\delta}}_0 \otimes \widehat{\boldsymbol{\delta}}_0) + \mathbf{K}_3(\widehat{\boldsymbol{\delta}}_0 \otimes \widehat{\boldsymbol{\delta}}_1) + \tfrac{1}{6}\mathbf{K}_4(\widehat{\boldsymbol{\delta}}_0 \otimes \widehat{\boldsymbol{\delta}}_0 \otimes \widehat{\boldsymbol{\delta}}_0)].$$

An expansion of the density of $\widehat{\psi}_1$ can be obtained by inverting the characteristic function of W_1 in (5). First, the moment generating function of W_1 will be found. For convenience, denote W_1 by W, i.e., $W = \sqrt{n}(\hat{\psi}_1 - \psi_1)$. To obtain $M_W(t)$, first expand W around $\hat{\psi}_1 = \psi_1$. Let $\mathbf{E}_{\boldsymbol{\varphi}}$ be a matrix of ones and zeros that satisfies $\mathbf{E}'_{\boldsymbol{\varphi}}\boldsymbol{\theta} = \boldsymbol{\varphi}$. An explicit expression for $\mathbf{E}_{\boldsymbol{\varphi}}$ is obtained by writing $\mathbf{E}_{\boldsymbol{\varphi}}$ as $\mathbf{E}_{\boldsymbol{\nu},3}$ and then using Eq. (4) in [7]. The random variable W can be expanded as follows:

$$\begin{aligned} &W = Q_0 + \frac{1}{\sqrt{n}}Q_1 + \frac{1}{n}Q_2 + O_p\left(n^{-\frac{3}{2}}\right), \quad \text{where } Q_0 = \mathbf{h}'\mathbf{D}^{(1)}_{\lambda:\boldsymbol{\varphi}}\mathbf{E}'_{\boldsymbol{\varphi}}\widehat{\boldsymbol{\delta}}_0, \\ &Q_1 = \mathbf{h}'[\mathbf{D}^{(1)}_{\lambda:\boldsymbol{\varphi}}\mathbf{E}'_{\boldsymbol{\varphi}}\widehat{\boldsymbol{\delta}}_1 + \tfrac{1}{2}\mathbf{D}^{(2)}_{\lambda:\boldsymbol{\varphi},\boldsymbol{\varphi}}(\mathbf{E}'_{\boldsymbol{\varphi}}\widehat{\boldsymbol{\delta}}_0 \otimes \mathbf{E}'_{\boldsymbol{\varphi}}\widehat{\boldsymbol{\delta}}_0)] \end{aligned} \tag{29}$$

and

$$\begin{aligned} Q_2 = \; &\mathbf{h}'[\mathbf{D}^{(1)}_{\lambda:\boldsymbol{\varphi}}\mathbf{E}'_{\boldsymbol{\varphi}}\widehat{\boldsymbol{\delta}}_2 + \mathbf{D}^{(2)}_{\lambda:\boldsymbol{\varphi},\boldsymbol{\varphi}}(\mathbf{E}'_{\boldsymbol{\varphi}}\widehat{\boldsymbol{\delta}}_0 \otimes \mathbf{E}'_{\boldsymbol{\varphi}}\widehat{\boldsymbol{\delta}}_1) \\ &+ \tfrac{1}{6}\mathbf{D}^{(3)}_{\lambda:\boldsymbol{\varphi},\boldsymbol{\varphi},\boldsymbol{\varphi}}(\mathbf{E}'_{\boldsymbol{\varphi}}\widehat{\boldsymbol{\delta}}_0 \otimes \mathbf{E}'_{\boldsymbol{\varphi}}\widehat{\boldsymbol{\delta}}_0 \otimes \mathbf{E}'_{\boldsymbol{\varphi}}\widehat{\boldsymbol{\delta}}_0)]. \end{aligned}$$

The moment generating function of W can be obtained by expanding the exponential function e^{tW} and then taking expectations. That is,

$$\begin{aligned} M_W(t) &= E(e^{tW}) \\ &= E\left(\exp\{tQ_0\}\left[1 + \frac{t}{\sqrt{n}}Q_1 + \frac{t}{n}Q_2 + \frac{t^2}{2n}Q_1^2 + O_p\left(n^{-\frac{3}{2}}\right)\right]\right), \end{aligned} \tag{30}$$

where Q_0, Q_1, and Q_2 are defined in (29) and the expectation is taken with respect to the Wishart distribution $W_p(n, \boldsymbol{\Sigma})$. The exponential function on the right-hand side of (30) can be combined with the Wishart density function, $f_{\text{wish}}(n\mathbf{S}; n, \boldsymbol{\Sigma})$, to obtain

$$\begin{aligned} &\exp\{tQ_0\}\, f_{\text{wish}}(n\mathbf{S}; n, \boldsymbol{\Sigma}) \\ &\quad = \frac{|\boldsymbol{\Sigma}_t|^{\frac{n}{2}}}{|\boldsymbol{\Sigma}|^{\frac{n}{2}}} \exp\left\{-\frac{t}{2}\sqrt{n}\,\text{tr}(\mathbf{V}\boldsymbol{\Sigma})\right\} f_{\text{wish}}(n\mathbf{S}; n, \boldsymbol{\Sigma}_t), \end{aligned}$$

where

$$\mathbf{V} = \mathbf{V}' = \text{dvec}(\mathbf{v}, p, p), \qquad \mathbf{v} = \ddot{\mathbf{F}}^{(1)}\, \bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1} \mathbf{E}_{\varphi} \mathbf{D}^{(1)\prime}_{\lambda:\boldsymbol{\varphi}} \mathbf{h},$$
$$\boldsymbol{\Sigma}_t = \left(\boldsymbol{\Sigma}^{-1} - \frac{t}{\sqrt{n}}\mathbf{V}\right)^{-1} = \sum_{j=0}^{\infty} \left(\frac{t}{\sqrt{n}}\right)^{j} (\boldsymbol{\Sigma}\mathbf{V})^{j}\, \boldsymbol{\Sigma}, \tag{31}$$

$\ddot{\mathbf{F}}^{(1)}$ is defined in (26), and $\text{dvec}(\mathbf{M}, a, b)$ is an $a \times b$ matrix that satisfies $\text{vec}\,[\text{dvec}(\mathbf{M}, a, b)] = \text{vec}\,\mathbf{M}$. It is assumed that $|t|$ is sufficiently small so that $-\infty < r_j t < \sqrt{n}$ for $j = 1, \ldots, p$ is satisfied, where $r_1, \ldots, r_p$ are the eigenvalues of $\boldsymbol{\Sigma}\mathbf{V}$. The ratio of determinants in (31) can be expanded as

$$\frac{|\boldsymbol{\Sigma}_t|^{\frac{n}{2}}}{|\boldsymbol{\Sigma}|^{\frac{n}{2}}} = \exp\left\{\frac{n}{2}\sum_{j=1}^{\infty}\left(\frac{t}{\sqrt{n}}\right)^{j} \frac{\text{tr}(\boldsymbol{\Sigma}\mathbf{V})^{j}}{j}\right\}.$$

Accordingly, the moment generating function of W can be expressed as

$$\begin{aligned} M_W(t) &= \exp\left\{\frac{t^2}{4}\ \text{tr}(\boldsymbol{\Sigma}\mathbf{V})^2\right\} \\ &\quad \times \left\{1 + \frac{t^3}{3!\sqrt{n}}\text{tr}(\boldsymbol{\Sigma}\mathbf{V})^3 + \frac{3t^4}{4!n}\ \text{tr}(\boldsymbol{\Sigma}\mathbf{V})^4 + \frac{10t^6}{6!n}\left[\text{tr}(\boldsymbol{\Sigma}\mathbf{V})^3\right]^2 + O\left(n^{-\frac{3}{2}}\right)\right\} \\ &\quad \times E\left[1 + \frac{t}{\sqrt{n}}Q_1 + \frac{t}{n}\,Q_2 + \frac{t^2}{2n}\,Q_1^2 + O_p\left(n^{-\frac{3}{2}}\right)\right], \end{aligned}$$

where Q_1 and Q_2 are defined in (29) and the expectation is taken with respect to the Wishart distribution $W_p(n, \boldsymbol{\Sigma}_t)$ and $\boldsymbol{\Sigma}_t$ is given in (31). The results, after taking expectations, are summarized in Theorem 2.

Theorem 2. *The moment generating function of* $W = \sqrt{n}(\hat{\psi}_1 - \psi_1)$ *is*

$$\begin{aligned} M_W(t) &= \exp\left\{\frac{t^2}{2}\sigma_W^2\right\} \\ &\quad \times\left\{1 + \frac{t}{\sqrt{n}}\omega_1 + \frac{t^2}{2n}\omega_2 + \frac{t^3}{3!\sqrt{n}}\omega_3 + \frac{t^4}{4!n}\omega_4 + \frac{t^6}{6!n}\omega_6 + O\left(n^{-\frac{3}{2}}\right)\right\}, \end{aligned}$$

where

$$\sigma_W^2 = \tfrac{1}{2}\,\text{tr}(\boldsymbol{\Sigma}\mathbf{V})^2,$$

and expressions for ω_i, $i = 1, 2, 3, 4, 6$, *are given in Appendix* B.

In some cases, a more accurate approximation can be obtained from the moment generating function of $W_m = \sqrt{m}(\hat{\psi}_1 - \psi_1)$ rather than from the moment generating function of W, where $m = n + \varDelta$ and $\varDelta = O(1)$. It is readily shown that the moment generating function of W_m is identical to the moment generating function of W to order $O(n^{-\frac{3}{2}})$ except that n is replaced by m and

$$\frac{t^2}{2n}\omega_2 \text{ is replaced by } \frac{t^2}{2m}\left(\omega_2 + \Delta\sigma_W^2\right). \tag{32}$$

The Edgeworth expansion for the distribution of $\hat{\psi}_1$ is obtained by inverting the characteristic function $M_W(it)$ or $M_{W_m}(it)$. The results based on W_m are summarized in Theorem 3. Results based on W are obtained by equating Δ to zero and equating m to n.

Theorem 3. *The density and distribution functions for the random variable* $\hat{\psi}_1$ *are*

$$f_{\hat{\psi}_1}(x|\boldsymbol{\theta}) = \frac{\sqrt{m}}{\sigma_W}\,\phi(z)\Bigg\{1+\frac{1}{\sigma_W\sqrt{m}}\,H_1(z)\omega_1+\frac{1}{\sigma_W^2 2m}\,H_2(z)\left(\omega_2+\Delta\sigma_W^2\right)$$
$$+\frac{1}{\sigma_W^3 3!\sqrt{m}}\,\omega_3H_3(z)+\frac{1}{\sigma_W^4 4!m}\,\omega_4H_4(z)+\frac{1}{\sigma_W^6 6!m}\,\omega_6H_6(z)+O\left(m^{-\frac{3}{2}}\right)\Bigg\}$$

and

$$F_{\hat{\psi}_1}(x|\boldsymbol{\theta}) = P(\hat{\psi}_1\leqslant x)=\Phi(z)-\phi(z)\Bigg\{\frac{1}{\sigma_W\sqrt{m}}\omega_1+\frac{1}{\sigma_W^2 2m}H_1(z)\left(\omega_2+\Delta\sigma_W^2\right)$$
$$+\frac{1}{\sigma_W^3 3!\sqrt{m}}\omega_3H_2(z)+\frac{1}{\sigma_W^4 4!m}\omega_4H_3(z)+\frac{1}{\sigma_W^6 6!m}\omega_6H_5(z)+O\left(m^{-\frac{3}{2}}\right)\Bigg\},$$

where $z=\sqrt{m}(x-\psi_1)/\sigma_W$, $\phi(\cdot)$ *is the standard normal pdf*, $\Phi(\cdot)$ *is the standard normal cdf, and* $H_j(\cdot)$ *is the* jth *Hermite polynomial.*

Konishi [24,25] gave scalar expressions for the $O(n^{-\frac{1}{2}})$ terms in Theorem 3 applied to correlation matrices for the special case when (a) $\Delta=0$, (b) the eigenvalues of $\boldsymbol{\Psi}$ are not constrained as in (19), i.e., when λ is estimated by the eigenvalues of the sample correlation matrix and (c) either $\mathbf{h}$ contains a single nonzero entry or $\mathbf{h}$ has the form $\mathbf{h}=(\mathbf{1}_a'\ \mathbf{0}')'$ for $a\leqslant p$ and the eigenvalues are ordered from large to small.

Konishi [23] and Fujikoshi [15] derived scalar expressions for Edgeworth expansions of differentiable functions of the eigenvalues of sample covariance matrices. The error in these expansions is $O(n^{-3/2})$, the same as in Theorem 3. Fujikoshi's [15] expansion is more general than Konishi's in that the eigenvalues need not be simple. Unlike the expansion in Theorem 3, however, Fujikoshi's expansion does not allow the eigenvalues to be modeled in parametric form. The expansion in Theorem 3 agrees numerically with Fujikoshi's expansion in the special case when the distinct eigenvalues are unconstrained.

4.2. Saddlepoint expansions under normality

An approximation to the cumulant generating function of $\hat{\psi}_1$ is readily obtained from the moment generating function given in Theorem 2. The result, after using (32), is

$$K_{\hat{\psi}_1}(t)\approx t\kappa_1+\frac{t^2}{2m}\kappa_2+\frac{t^3}{3!m^2}\kappa_3+\frac{t^4}{4!m^3}\kappa_4, \tag{33}$$

where

$$\kappa_1=\psi_1+\frac{\omega_1}{m},\quad \kappa_2=\sigma_W^2+\frac{\omega_2-\omega_1^2+\Delta\sigma_W^2}{m},\quad \kappa_3=\omega_3,\quad \kappa_4=\omega_4-4\,\omega_1\omega_3,$$

and ω_i for $i = 1, \ldots, 4$ as well as σ^2_W are defined in Theorem 2. From Easton and Ronchetti [13], the general saddlepoint approximation to the density of $\hat{\psi}_1$ at x is

$$f_{\hat{\psi}_1}(x) = \widehat{f}_{\hat{\psi}_1}(x) + O\left(n^{-1}\right), \tag{34}$$

where

$$\begin{aligned}
\widehat{f}_{\hat{\psi}_1}(x) &= \left[\frac{m}{2\pi R''_m(t_0)}\right]^{\frac{1}{2}} \exp\left\{m\left[R_m(t_0) - xt_0\right]\right\}, \\
R_m(t) &= \frac{1}{m}\widehat{K}_{\hat{\psi}_1}(mt) = t\kappa_1 + \frac{t^2}{2}\kappa_2 + \frac{t^3}{3!}\kappa_3 + \frac{t^4}{4!}\kappa_4, \\
R''_m(t_0) &= \left.\frac{\partial^2 R_m(t)}{(\partial t)^2}\right|_{t=t_0},
\end{aligned}$$

t_0 is the solution to $R'_m(t_0) = x$, and

$$R'_m(t) = \frac{\partial R_m(t)}{\partial t}.$$

The renormalized saddlepoint approximation is $c_0 \widehat{f}_{\hat{\psi}_1}(x)$, where c_0 is chosen so that the approximation to the density integrates to one. The relative error in the renormalized saddlepoint approximation is only $O(n^{-3/2})$, at least in the normal deviation region $x - \psi_1 = O(n^{-1/2})$.

It is possible that $R'_m(t_0) = x$ has either no real solution or multiple real solutions. For this reason, Wang [38] suggested that $R_m(t)$ in (34) be replaced by

$$\widetilde{R}_m(t, b) = t\kappa_1 + \frac{t^2}{2}\kappa_2 + \left[\frac{t^3}{3!}\kappa_3 + \frac{t^4}{4!}\kappa_4\right] \exp\left\{-\frac{t^2}{2}\kappa_2 b^2\right\}, \tag{35}$$

where

$$b = \max[\tfrac{1}{2}, \inf\{b^*;\ \widetilde{R}''_m(t, b^*) \geqslant 0 \text{ for all } t\}].$$

Replacing $R_m(t)$ by $\widetilde{R}_m(t, b)$ does not change the order of the approximation.

The saddlepoint approximation to the CDF of $\hat{\psi}_1$ can be obtained by using the method of Lugannini and Rice [26]. The result is

$$\begin{aligned}
F_{\hat{\psi}_1}(x|\boldsymbol{\theta}) &= P(\hat{\psi}_1 \leqslant x) \\
&= \Phi(r\sqrt{m}) - \frac{\phi(r\sqrt{m})}{\sqrt{m}}\left(\frac{1}{t_0\sqrt{\widetilde{R}''_m(t_0, b)}} - \frac{1}{r}\right) + O\left(n^{-1}\right),
\end{aligned}$$

where

$$r = \mathrm{sign}(t_0)\sqrt{2\left[t_0 x - \widetilde{R}_m(t_0, b)\right]}, \quad t_0 \text{ is the solution to } \widetilde{R}'_m(t_0, b) = x, \tag{36}$$

$\phi(\cdot)$ is the standard normal pdf, and $\Phi(\cdot)$ is the standard normal cdf.

4.3. Modifications for $\widehat{\psi}_2$

The expansions in Sections 4.1 and 4.2 can be applied to the random variable $\widehat{\psi}_2 = \mathbf{h}'\widehat{\lambda}/(\mathbf{h}'\widehat{\lambda})$, where $\widehat{\lambda}$ is the maximum likelihood estimator of the vector of eigenvalues of the covariance matrix, by making the following modifications.

(1) Parameterize the eigenvalues as (23) rather than (22).
(2) Replace $\mathbf{D}^{(1)}_{\lambda:\boldsymbol{\varphi}}$, $\mathbf{D}^{(2)}_{\lambda:\boldsymbol{\varphi},\boldsymbol{\varphi}}$, and $\mathbf{D}^{(3)}_{\lambda:\boldsymbol{\varphi},\boldsymbol{\varphi},\boldsymbol{\varphi}}$ by $\varphi_1^{-1}\mathbf{D}^{(1)}_{\lambda:\boldsymbol{\varphi}_2}$, $\varphi_1^{-1}\mathbf{D}^{(2)}_{\lambda:\boldsymbol{\varphi}_2,\boldsymbol{\varphi}_2}$, and $\varphi_1^{-1}\mathbf{D}^{(3)}_{\lambda:\boldsymbol{\varphi}_2,\boldsymbol{\varphi}_2,\boldsymbol{\varphi}_2}$, respectively. Expressions for these derivatives are given in the supplement.
(3) Replace $\mathbf{E}_{\boldsymbol{\varphi}} = \mathbf{E}_{\boldsymbol{\nu},3}$ by $\mathbf{E}_{\boldsymbol{\varphi}_2} = \mathbf{E}_{\boldsymbol{\nu}^*,3}$, where $\boldsymbol{\nu}^* = [\dim(\boldsymbol{\mu})\ 1\ \dim(\boldsymbol{\varphi}_2)]'$.

The Edgeworth expansions of Konishi [23] and Fujikoshi [15] also can be used for $\widehat{\psi}_2$ in the special case where $\widehat{\lambda}$ is the vector of eigenvalues of the sample covariance matrix. Sugiyama and Tong [35] also gave an Edgeworth expansion for the density of $\widehat{\psi}_2$ in the same special case. Their expansion, however, is not invariant to scalar multiplication of $\boldsymbol{\Sigma}$, and, therefore, it is not correct.

5. Edgeworth and saddlepoint expansions without normality

In this section, expansions for the density of $\widehat{\psi}_1$ are constructed without assuming multivariate normality of $\boldsymbol{\varepsilon}$, an arbitrary row of $\mathbf{E}$ in (3). Expansions for $\widehat{\psi}_2$ are similar in structure and can be obtained by making the modifications described in Section 4.3.

5.1. Validity conditions

Sufficient conditions to ensure validity of Edgeworth expansions when sampling from nonnormal distributions have been described by Bhattacharya and Ghosh [4], Wallace [37], Hall [18, Section 2.4], and others. It follows from Theorem 2.2 in Hall that, with remainder $O(n^{-1})$, Edgeworth expansions of the distributions of W_k in (5) for $k = 1, 2$ are valid if the following conditions are satisfied.

(a) $E(||\boldsymbol{\varepsilon}||^6) < \infty$ and
(b) $\limsup_{||\mathbf{t}||\to\infty} |M_{\boldsymbol{\varepsilon}^*}(i\mathbf{t})| < 1$, where $\boldsymbol{\varepsilon}^* = \begin{pmatrix} \boldsymbol{\varepsilon} \\ \operatorname{vech}(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}') \end{pmatrix}$, the vech operator stacks the distinct elements of a symmetric matrix, and $M_{\boldsymbol{\varepsilon}^*}(i\mathbf{t})$ is the characteristic function of $\boldsymbol{\varepsilon}^*$.

Define $\widehat{\sigma}^2_{W_k}$ to be a consistent estimator of $\operatorname{Var}(W_k)$ based on observing $\left\{\operatorname{vech}(\widetilde{\boldsymbol{\varepsilon}}_j\widetilde{\boldsymbol{\varepsilon}}_j' \otimes \widetilde{\boldsymbol{\varepsilon}}_j\widetilde{\boldsymbol{\varepsilon}}_j')\right\}_{j=1}^n$ for $k = 1, 2$. Then, with remainder $O(n^{-1})$, Edgeworth expansions of the distributions of Z_k in (5) for $k = 1, 2$ are valid if the following conditions are satisfied.

(a*) $E(||\boldsymbol{\varepsilon}||^{12}) < \infty$ and

(b*) $\limsup_{||\mathbf{t}||\to\infty} |M_{\boldsymbol{\varepsilon}^{**}}(i\mathbf{t})| < 1$, where $\boldsymbol{\varepsilon}^{**} = \begin{pmatrix} \boldsymbol{\varepsilon}^{*} \\ \operatorname{vech}(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}' \otimes \boldsymbol{\varepsilon}\boldsymbol{\varepsilon}') \end{pmatrix}$.

Conditions (b) and (b*) are satisfied if the distribution of $\boldsymbol{\varepsilon}$ has an absolutely continuous component whose density is strictly positive on some nonempty open set. The above conditions are sufficient but not necessary. Babu [2] showed that Edgeworth expansions can be valid even if only one component of the random vector ($\boldsymbol{\varepsilon}$ in the present application) satisfies Cramér's condition, (b) or (b*). Booth, Hall, and Woods [8] and Kong and Levin [21] showed that Edgeworth expansions of the distribution of a statistic, T, can be valid even if one samples from a discrete distribution provided that the sampling distribution of T is not lattice and its support set is sufficiently dense. Booth, Hall, and Woods ensure that T is not lattice by requiring that the parent distribution be nonlattice and that its support set contain a minimal number of atoms. Kong and Levin ensure that T be nonlattice by requiring that covariates not cluster around too few points.

5.2. ADF edgeworth expansions

The first three moments of $\widehat{\psi}_1$ under general conditions can be obtained by expanding $\widehat{\psi}_1$, $\widehat{\psi}_1^2$, and $\widehat{\psi}_1^3$ around $\mathbf{s} = \boldsymbol{\sigma}$ and taking expectations. The results in Section 2 are useful for evaluating these expectations. The three moments, in turn, can be used to construct Edgeworth expansions. With remainder $O(n^{-1})$, the Edgeworth expansions for the density and distribution functions of $\hat{\psi}_1$ are given in Theorem 4.

Theorem 4. *If the data follow the model in* (3) *and* (a) *and* (b) *in Section* 5.1 *are satisfied, then the density and distribution functions of* $\widehat{\psi}_1$ *are as follows*:

$$f_{\hat{\psi}_1}(x|\boldsymbol{\theta}) = \frac{\sqrt{n}}{\sigma_W}\phi(z)\left\{1 + \frac{\omega_1 z}{\sigma_W\sqrt{n}} + \frac{\omega_3}{\sigma_W^3 3!\sqrt{n}}H_3(z) + O\left(n^{-1}\right)\right\}$$

and

$$F_{\hat{\psi}_1}(x|\boldsymbol{\theta}) = \Phi(z) - \phi(z)\left\{\frac{\omega_1}{\sigma_W\sqrt{n}} + \frac{\omega_3}{\sigma_W^3 3!\sqrt{n}}H_2(z) + O\left(n^{-1}\right)\right\},$$

where

$$z = \frac{\sqrt{n}(x - \psi_1)}{\sigma_W}; \quad W = \sqrt{n}(\hat{\psi}_1 - \psi_1); \quad \sigma_W^2 = \tfrac{1}{4}\mathbf{v}'\boldsymbol{\Omega}_{22,n}\mathbf{v};$$

$$\begin{aligned}\omega_1 = {} & \tfrac{1}{4}\operatorname{tr}\left\{\ddot{\mathbf{F}}^{(2)}\left[\mathbf{a}\otimes\bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1}\,\ddot{\mathbf{F}}^{(1)\prime}\,\boldsymbol{\Omega}_{22,n}\left(\mathbf{I}_{p^2} - \mathbf{P}_F\right)'\right]\right\} + \tfrac{1}{8}\operatorname{tr}\left(\mathbf{A}_2\,\ddot{\mathbf{F}}^{(1)\prime}\,\boldsymbol{\Omega}_{22,n}\,\ddot{\mathbf{F}}^{(1)}\right)\\ & - \tfrac{1}{16}\mathbf{v}'\,\mathbf{F}^{(2)}\operatorname{vec}\left(\bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1}\,\ddot{\mathbf{F}}^{(1)\prime}\,\boldsymbol{\Omega}_{22,n}\,\ddot{\mathbf{F}}^{(1)}\,\bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1}\right)\\ & - \operatorname{tr}\left[\left(\mathbf{V}\otimes\boldsymbol{\Sigma}^{-1}\right)\left(\mathbf{I}_{p^2} - \mathbf{P}_F\right)\boldsymbol{\Omega}_{22,n}\mathbf{P}_F{}'\right],\end{aligned}$$

$$\omega_3 = \frac{\sqrt{n}}{8}(\mathbf{v}\otimes\mathbf{v})'\boldsymbol{\Omega}_{222,n}\mathbf{v} + \frac{3}{16}\mathbf{b}'\bar{\mathbf{I}}_{\boldsymbol{\theta}}\mathbf{A}_2\bar{\mathbf{I}}_{\boldsymbol{\theta}}\mathbf{b} + \frac{3}{8}\mathbf{v}'\boldsymbol{\Omega}_{22,n}\left(\mathbf{I}_{p^2}-\mathbf{P}_F\right)'\ddot{\mathbf{F}}^{(2)}(\mathbf{a}\otimes\mathbf{b})$$
$$-\frac{3}{2}\mathbf{v}'\boldsymbol{\Omega}_{22,n}\left(\mathbf{I}_{p^2}-\mathbf{P}_F\right)'\left(\mathbf{V}\otimes\boldsymbol{\Sigma}^{-1}\right)\mathbf{P}_F\boldsymbol{\Omega}_{22,n}\mathbf{v} - \frac{3}{32}\mathbf{v}'\,\mathbf{F}^{(2)}(\mathbf{b}\otimes\mathbf{b}),$$

$$\mathbf{b} = \bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1}\,\ddot{\mathbf{F}}^{(1)\prime}\,\boldsymbol{\Omega}_{22,n}\mathbf{v},$$

$\mathbf{V}$ *and* $\mathbf{v}$ *are defined in* (31); $\mathbf{a}$, $\mathbf{A}_2$, *and* $\mathbf{P}_F$ *are defined in Appendix* B; *and the remaining terms are defined in Theorem* 3. *Under normality, the expressions for* ω_1, σ^2_W, *and* ω_3 *in simplify to those in* (30).

Waternaux [39] obtained the Cornish–Fisher expansion for the percentiles of the marginal distributions of the eigenvalues of a sample covariance matrix without assuming normality. The Cornish–Fisher expansion was based on an Edgeworth expansion with error $O(n^{-1})$. Fujikoshi [16] also obtained the Edgeworth expansion for the distribution of the eigenvalues of a sample covariance matrix without assuming normality. Fujikoshi's expansion has error only $O(n^{-\frac{3}{2}})$, but like Waternaux's expansion, it does not allow restrictions on the eigenvalues. In particular, both Waternaux's and Fujikoshi's expansions require that eigenvalue multiplicities be one.

5.3. ADF saddlepoint expansions

An ADF saddlepoint approximation to the distribution of $\widehat{\psi}_1$ can be obtained by using the method of Gatto and Ronchetti [17]. In the present application, their procedure consists first of keeping only the Q_0 and Q_1 terms of the expansion of W in (29). This truncated expansion, in turn, is approximated by a U statistic of degree 2. An ADF Edgeworth expansion for the U statistic is then obtained and is used to construct the saddlepoint approximation. The result is identical to the saddlepoint approximation in (34), except that κ_{ig} is substituted for κ_i for $i = 1, \ldots, 4$, where

$$\kappa_{1g} = \psi_1 + \frac{\omega_1}{m}, \quad \kappa_{2g} = \sigma^2_W + \frac{\omega_{2g} + \Delta\sigma^2_W}{m},$$
$$\kappa_{3g} = \omega_3\left(\frac{\kappa_{2g}}{\sigma^2_W}\right)^{\frac{3}{2}} \quad \text{and} \quad \kappa_{4g} = \omega_{4g}\left(\frac{\kappa_{2g}}{\sigma^2_W}\right)^{2}, \tag{37}$$

σ^2_W, ω_1, and ω_3 are given in Theorem 4; and expressions for ω_{2g} and ω_{4g} are given in Appendix C. Unlike the expansion in (34), however, the renormalized saddlepoint approximation of Gatto and Ronchetti has relative error $O(n^{-1})$ rather than $O(n^{-3/2})$. The loss of accuracy occurs because the error in the expansion of W is $O_p(n^{-1})$ rather than $O_p(n^{-3/2})$.

6. ADF confidence intervals

Confidence intervals for ψ_1 can be constructed from the density and distribution functions of the studentized statistic

$$Z \stackrel{\text{def}}{=} Z_1, \quad \text{where } \widehat{\sigma}_W^2 = \begin{cases} \frac{1}{2}\operatorname{tr}\left(\widehat{\mathbf{V}}\widehat{\mathbf{\Sigma}}\right)^2 & \text{if normality is assumed,} \\ \frac{1}{4}\widehat{\mathbf{v}}'\widehat{\mathbf{\Omega}}_{22,n}\widehat{\mathbf{v}} & \text{otherwise,} \end{cases}$$

Z_1 is defined in (5), $\widehat{\mathbf{V}}$ and $\widehat{\mathbf{v}}$ are $\mathbf{V}$ and $\mathbf{v}$ of (31) in which consistent estimators are substituted for parameters and $\widehat{\mathbf{\Omega}}_{22,n}$ is given in (14). Expanding the numerator and denominator of Z around $\widehat{\boldsymbol{\theta}} = \boldsymbol{\theta}$ and $\widehat{\mathbf{\Omega}}_{22,n} = \mathbf{\Omega}_{22,n}$ and then taking expectations reveals that

$$E(Z) = \frac{\omega_1^*}{\sqrt{n}} + O\left(n^{-\frac{3}{2}}\right), \quad \operatorname{Var}(Z) = 1 + O\left(n^{-1}\right)$$

and

$$E\left[Z - E(Z)\right]^3 = \frac{\omega_3^*}{\sqrt{n}} + O\left(n^{-\frac{3}{2}}\right),$$

where

$$\omega_1^* = \frac{\omega_1}{\sigma_W} - \frac{\omega_3}{3\sigma_W^3} - \frac{1}{6\sigma_W^3}\operatorname{tr}(\mathbf{V\Sigma})^3 + \frac{\delta}{\sigma_W^3},$$
$$\omega_3^* = -\frac{1}{\sigma_W^3}\left[\omega_3 + \operatorname{tr}(\mathbf{V\Sigma})^3\right] + 6\frac{\delta}{\sigma_W^3},$$

$$\begin{aligned}\delta = {} & \frac{1}{16}\mathbf{v}'\mathbf{\Omega}_{22,n}(\mathbf{I}_{p^2} - \mathbf{P}_F)'\ddot{\mathbf{F}}^{(2)}(\mathbf{a}\otimes\mathbf{b}) - \frac{1}{16}(\mathbf{v}\otimes\mathbf{v})'\mathbf{\Omega}_{42,n}\mathbf{v} + \frac{1}{6}\operatorname{tr}(\mathbf{V\Sigma})^3 \\ & + \frac{\sqrt{n}}{24}(\mathbf{v}\otimes\mathbf{v})'\mathbf{\Omega}_{222,n}\mathbf{v} - \frac{1}{4}\mathbf{v}'\mathbf{\Omega}_{22,n}(\mathbf{I}_{p^2} - \mathbf{P}_F)'\left(\mathbf{V}\otimes\mathbf{\Sigma}^{-1}\right)\mathbf{P}_F\mathbf{\Omega}_{22,n}\mathbf{v}, \end{aligned} \tag{38}$$

$\mathbf{\Omega}_{22,n}$, $\mathbf{\Omega}_{42,n}$, and $\mathbf{\Omega}_{222,n}$ are defined in (6); and the remaining terms are defined Theorem 4. If $\mathbf{Y}$ has a multivariate normal distribution, then ω_1 and ω_3 simplify to the expressions in Theorem 2 and δ simplifies to $\delta = 0$.

Define Z^* as

$$Z^* = Z - \frac{\widehat{\omega}_1^*}{\sqrt{n}}, \tag{39}$$

where Z is defined in (6) and $\widehat{\omega}_1^*$ is ω_1^* of (38) in which consistent estimators have been substituted for parameters. If conditions (a*) and (b*) in Section 5.1 are satisfied, then the Edgeworth expansions of the density and distribution of Z^* to $O(n^{-1})$ are

$$f_{Z^*}(z^*|\boldsymbol{\theta}) = \phi(z^*)\left\{1 + \frac{\omega_3^*}{3!\sqrt{n}}H_3(z^*) + O\left(n^{-1}\right)\right\}$$

and

$$F_{Z^*}(z^*|\boldsymbol{\theta}) = \Phi(z^*) - \phi(z^*)\left\{\frac{\omega_3^*}{3!\sqrt{n}}H_2(z^*) + O\left(n^{-1}\right)\right\}, \tag{40}$$

where ω_3^* is defined in (38). In practice, a consistent estimator of ω_3^* must be substituted for ω_3^*. This substitution does not affect the accuracy of the expansions. Several confidence interval procedures can be constructed directly from (40). These methods plus another method based on an empirical saddlepoint approximation are described in the following subsections. Confidence intervals for ψ_2 are obtained in the same manner after making the substitutions described in Section 4.3.

6.1. First-order method and second-order edgeworth method

To order $O(n^{-\frac{1}{2}})$, the random variable Z in (6) has distribution $Z \sim \mathrm{N}(0, 1)$. Accordingly,

$$\left(\widehat{\psi}_1 - z_{\alpha_2}\frac{\widehat{\sigma}_W}{\sqrt{n}}, \widehat{\psi}_1 - z_{\alpha_1}\frac{\widehat{\sigma}_W}{\sqrt{n}}\right) \tag{41}$$

is a $100(\alpha_2 - \alpha_1)\%$ confidence interval for ψ_1 with error $O(n^{-\frac{1}{2}})$, where z_α satisfies $\Phi(z_\alpha) = \alpha$.

A second-order accurate confidence interval can be based on the Edgeworth expansion of the distribution of Z^* in (39). Percentiles of this distribution to $O(n^{-1})$ can be computed by inverting the cumulative distribution function in (40). The desired inverse function is readily obtained from the Cornish–Fisher expansion [11]. The results are as follows:

$$P(Z^* \leqslant z_\alpha^*) = \alpha + O\left(n^{-1}\right), \quad \text{where } z_\alpha^* = z_\alpha + \frac{\widehat{\omega}_3^*(z_\alpha^2 - 1)}{6\sqrt{n}}.$$

Accordingly,

$$\left(\widehat{\psi}_1 - \frac{\widehat{\sigma}_W}{\sqrt{n}}\left[z_{\alpha_2}^* + \frac{\widehat{\omega}_1^*}{\sqrt{n}}\right], \widehat{\psi}_1 - \frac{\widehat{\sigma}_W}{\sqrt{n}}\left[z_{\alpha_1}^* + \frac{\widehat{\omega}_1^*}{\sqrt{n}}\right]\right) \tag{42}$$

is a $100(\alpha_2 - \alpha_1)\%$ confidence interval for ψ_1 with error $O(n^{-1})$. Note that the widths of the first-order interval in (41) and the second-order interval in (42) are identical. The second-order interval is merely shifted to correct for bias and skewness.

6.2. Hall's cubic transformation method

Hall [19] argued against using a confidence interval such as that in (42) when sample size is small. The problem with (42) is that $z_{\alpha_1}^*$ and $z_{\alpha_2}^*$ both diverge to ∞ or to $-\infty$, depending on the sign of ω_3^* as $\alpha_2 - \alpha_1 \to 1$. Instead of (42), Hall recommended that confidence intervals be based on the Cornish–Fisher quantity

$$T = Z^* - \frac{\widehat{\omega}_3^*(Z^{*^2} - 1)}{6\sqrt{n}}, \tag{43}$$

which to order $O\left(n^{-1}\right)$ has the $\mathrm{N}(0, 1)$ distribution. Hall added a term of size $O(n^{-1})$ to simplify the inversion from T to Z^*. Hall's quantity is $T_H = T + \widehat{\omega}_3^{*^2} Z^{*^3}/(108n)$ and this quantity also has the N(0, 1) distribution to $O(n^{-1})$. The inverse is

$$Z^* = \frac{6\sqrt{n}}{\widehat{\omega}_3^*}\left\{1 - \left[1 + \frac{\widehat{\omega}_3^*}{2\sqrt{n}}\left(\frac{\widehat{\omega}_3^*}{6\sqrt{n}} - T_H\right)\right]^{\frac{1}{3}}\right\}.$$

Applying the inverse transformation to $P(z_{\alpha_1} \leqslant T_H \leqslant z_{\alpha_2}) = \alpha_2 - \alpha_1 + O(n^{-1})$ reveals that (L_2, L_1) is a $100(\alpha_2 - \alpha_1)\%$ confidence interval for ψ_1 with error $O(n^{-1})$, where

$$L_i = \widehat{\psi}_1 - \widehat{\sigma}_W\left\{\frac{\widehat{\omega}_1^*}{n} + \frac{6}{\widehat{\omega}_3^*} - \frac{6}{\widehat{\omega}_3^*}\left[1 + \frac{\widehat{\omega}_3^*}{2\sqrt{n}}\left(\frac{\widehat{\omega}_3^*}{6\sqrt{n}} - z_{\alpha_i}\right)\right]^{\frac{1}{3}}\right\}. \tag{44}$$

6.3. DiCiccio and Monti's transformation method

DiCiccio and Monti [12] indirectly employed the Cornish–Fisher quantity in (43) to construct second-order correct confidence intervals. Their approach, applied to the current problem, is to work with Z_C rather than Z^*, where

$$Z_C = \frac{\sqrt{n}(\widehat{\phi} - \phi)}{\widehat{\sigma}_\phi}, \quad \widehat{\phi} = h(\widehat{\psi}_1),$$

and the function $h(\cdot)$ is chosen to remove skewness. If h is chosen to be a member of the Box–Cox [9] family of transformations, $h(\widehat{\psi}_1) = (\widehat{\psi}_1^\zeta - 1)/\zeta$, then Z_C can be expanded as

$$Z_C = Z^* - \frac{(\zeta - 1)\widehat{\sigma}_W}{2\widehat{\psi}_1\sqrt{n}}\left(Z^{*^2} - 1\right) + \frac{\widehat{\omega}_1^*}{\sqrt{n}} - \frac{\widehat{\omega}_3^*}{6\sqrt{n}} + O\left(n^{-1}\right),$$

Furthermore, if ζ is estimated as $\widehat{\zeta} = 1 + (\widehat{\psi}_1\widehat{\omega}_3^*)/(3\widehat{\sigma}_W)$, then

$$Z_C = T + \frac{\widehat{\omega}_1^*}{\sqrt{n}} - \frac{\widehat{\omega}_3^*}{6\sqrt{n}} + O\left(n^{-1}\right),$$

where T is defined in (43). Accordingly,

$$Z_C - \frac{\widehat{\omega}_1^*}{\sqrt{n}} + \frac{\widehat{\omega}_3^*}{6\sqrt{n}} \sim \mathrm{N}(0, 1)$$

to order $O(n^{-1})$. Applying the inverse Box–Cox transformation to

$$P\left(z_{\alpha_1} \leqslant Z_C - \frac{\widehat{\omega}_1^*}{\sqrt{n}} + \frac{\widehat{\omega}_3^*}{6\sqrt{n}} \leqslant z_{\alpha_2}\right) = \alpha_2 - \alpha_1 + O\left(n^{-1}\right)$$

reveals that (L_2, L_1) is a $100(\alpha_2 - \alpha_1)\%$ confidence interval for ψ_1 with error $O(n^{-1})$, where

$$L_i = \widehat{\psi}_1\left[1 - \left(\frac{\sigma_W}{\widehat{\psi}_1\sqrt{n}} + \frac{\widehat{\omega}_3^*}{3\sqrt{n}}\right)\left(z_{\alpha_i} + \frac{\widehat{\omega}_1^*}{\sqrt{n}} - \frac{\widehat{\omega}_3^*}{6\sqrt{n}}\right)\right]^{1/\hat{\zeta}}, \tag{45}$$

and $\hat{\zeta} = 1 + \widehat{\psi}_1\widehat{\omega}_3/(3\widehat{\sigma}_W)$.

6.4. Tingley and Field's saddlepoint method

Tingley and Field [36] proposed a technique for constructing confidence intervals for smooth scalar functions of a parameter vector, where the parameters are estimated by an *M*-estimator. For the problem under consideration, the technique begins by approximating the distribution of

$$\overline{G} = N^{-1}\sum_{i=1}^{N}\frac{1}{2}\widehat{\mathbf{v}}'\left[(\boldsymbol{\varepsilon}_i \otimes \boldsymbol{\varepsilon}_i) - \widehat{\boldsymbol{\sigma}}\right],$$

where $\widehat{\mathbf{v}}$ is $\mathbf{v}$ of (31) in which parameters are replaced by estimates, $\widehat{\mathbf{v}}$ and $\widehat{\boldsymbol{\sigma}}$ are treated as fixed vectors, and $\boldsymbol{\varepsilon}_i$ for $i = 1, \ldots, N$ have the same joint distribution as the rows of $\mathbf{E}$ in (31). If $\boldsymbol{\varepsilon}_i \sim \mathrm{N}(\mathbf{0}, \boldsymbol{\Sigma})$, then the cumulant generating function of $\overline{G}$ is

$$K_{\bar{G}}(t) = NK_G(N^{-1}t),$$

where

$$K_G(t) = -\frac{t}{2}\,\mathrm{tr}\left(\boldsymbol{\Sigma}\widehat{\mathbf{V}}\right) - \frac{1}{2}\ln|\mathbf{I}_p - t\boldsymbol{\Sigma}\widehat{\mathbf{V}}| = \frac{1}{2}\sum_{j=2}^{\infty}\left(\frac{t^j}{j}\right)\mathrm{tr}(\boldsymbol{\Sigma}\widehat{\mathbf{V}})^j. \tag{46}$$

More generally, the cumulant generating function of $\overline{G}$ can be approximated by the empirical cumulant generating function

$$\widehat{K}_{\bar{G}}(t) = N\widehat{K}_G(N^{-1}t), \quad \text{where } \exp\left\{\widehat{K}_G(t)\right\} = N^{-1}\sum_{i=1}^{N}\exp\{tg_i\}, \tag{47}$$

$\{g_i\}_{i=1}^N$ is the observed configuration, $g_i = \widehat{\mathbf{v}}'\left[(\widetilde{\boldsymbol{\varepsilon}}_i \otimes \widetilde{\boldsymbol{\varepsilon}}_i) - \widehat{\boldsymbol{\sigma}}\right]/2$, $\widetilde{\boldsymbol{\varepsilon}}_i$ is the ith residual, and $\widetilde{\boldsymbol{\varepsilon}}_i$ is defined in (9).

A confidence interval for ψ_1 is then constructed by inverting a test of $\mathrm{H}_0\colon \psi_1 = \psi_{10}$. The test itself is obtained by applying an exponential tilt to the distribution of $\overline{G}$:

$$f_{\bar{G}}(\bar{g}|\zeta) \approx \exp\left\{\zeta\bar{g} - K_{\bar{G}}(\zeta)\right\} f_{\bar{G}}(\bar{g}),$$

where $K_{\bar{G}}$ is the cumulant generating function of $\overline{G}$ and ζ is chosen to satisfy $E\left(\overline{G}|\zeta\right) = (\psi_{10} - \widehat{\psi}_1)(N-1)/N$. The test statistic is $\overline{G}$, the observed value is $\bar{g}_{\text{obs}} = 0$, and the confidence interval is the set of values

$$\left\{\psi_1;\ \alpha_1 < P(\overline{G} < \bar{g}_{\text{obs}}|\zeta) < \alpha_2\right\},$$

where the probability $P(\overline{G} < \bar{g}_{\text{obs}}|\zeta)$ is computed by using a saddlepoint approximation to the density $f_{\bar{G}}(\bar{g}|\zeta)$. Using the method of Lugannini and Rice [26], the resulting $100(\alpha_2 - \alpha_1)\%$ confidence interval with error $O(N^{-1})$ is (L_1, L_2), where

$$L_i = \left(\frac{N}{N-1}\right)K'_G(\zeta^*_{3-i}) + \widehat{\psi}_1,$$

$$\zeta^*_{3-i} \text{ satisfies } \Phi\left[-\,\mathrm{sign}(\zeta^*_{3-i})\sqrt{2(N-1)K_G(\zeta^*_{3-i})}\right]$$

$$+\frac{\exp\left\{-(N-1)K_G(\zeta^*_{3-i})\right\}}{\sqrt{2\pi(N-1)}}\left[\frac{1}{\zeta^*_{3-i}\sqrt{K''_G(0)}}-\frac{\operatorname{sign}(\zeta^*_{3-i})}{\sqrt{2K_G(\zeta^*_{3-i})}}\right]=\alpha_i,$$

$$K'_G(\zeta^*_{3-i})=\left.\frac{\partial K_G(\zeta)}{\partial\zeta}\right|_{\zeta=\zeta^*_{3-i}},\quad K''_G(\zeta^*_{3-i})=\left.\frac{\partial^2 K_G(\zeta)}{(\partial\zeta)^2}\right|_{\zeta=\zeta^*_{3-i}},\tag{48}$$

and K_G is given in (46) if normality is assumed or is replaced by $\widehat{K}_G$ in (47) otherwise.

7. Example

Data set #144 in [20] contains five measurements on skulls that had been collected in Tibet. The skulls were classified by region into two groups of size 17 and 15. For further details, see [32]. Mardia's [30] measures of multivariate skewness and kurtosis are $b_1 = 12.88, b_2 = 36.64$ for the first group and $b_1 = 13.99, b_2 = 32.61$ for the second group. These values suggest that the distribution of skull measurements is fairly symmetric and mildly meso-kurtic. The likelihood ratio test of equality of the two population covariance matrices is nonsignificant (Bartlett corrected $X^2 = 18.37$, $\mathrm{d}f = 15$, $p = 0.24$). The pooled sample covariance matrix together with the sample correlation matrix are displayed in the lower and upper triangular parts of the following matrix:

$$\mathbf{S}\backslash\mathbf{R}=\begin{pmatrix}59.01 & 0.17 & 0.37 & 0.61 & 0.40\\ 9.01 & 48.26 & 0.03 & 0.15 & 0.65\\ 17.22 & 1.08 & 36.20 & 0.19 & 0.10\\ 20.12 & 4.34 & 4.84 & 18.31 & 0.46\\ 20.11 & 30.05 & 4.11 & 12.99 & 43.70\end{pmatrix}.$$

A correlation model was fit to the skull measures. The model placed no restrictions on the standard deviations, but restricted the distinct eigenvalues of the correlation matrix to follow an exponential curve, $\lambda_2 = \lambda_3$, and $\lambda_4 = \lambda_5$. These restrictions can be imposed by choosing $\mathbf{T}_3$ and $\mathbf{T}_4$ in (19) as

$$\mathbf{T}_3 = (1 \oplus \mathbf{1}_2 \oplus \mathbf{1}_2) \quad \text{and} \quad \mathbf{T}_4 = \begin{pmatrix}1 & 1 & 1\\ 1 & 2 & 3\end{pmatrix}'.$$

The later matrix, however, does not satisfy $\mathbf{1}'_3\mathbf{T}_4 = \mathbf{0}$. As described below equation (19), $\mathbf{T}_4$ can be replaced by any matrix whose columns form a basis for $\mathcal{R}[(\mathbf{I}_3 - \mathbf{1}_3(\frac{1}{3})\mathbf{1}'_3)\mathbf{T}_4]$.

Table 2
Eigenvalues of $\mathbf{R}$ and $\widehat{\mathbf{\Psi}}$

Index	Eigenvalues		Standard error		Bias estimate	
	$\mathbf{R}$	$\widehat{\mathbf{\Psi}}$	Normal	ADF	Normal	ADF
1	2.331	2.531	0.249	0.243	0.090	0.131
2	1.236	0.908	0.054	0.053	−0.024	−0.033
3	0.804	0.908	0.054	0.053	−0.024	−0.033
4	0.366	0.326	0.071	0.069	−0.021	−0.033
5	0.263	0.326	0.071	0.069	−0.021	−0.033

One such matrix is $\mathbf{T}_4 = (1\ 0\ \ -1)'$. The estimate of φ, for this choice of model matrices, is $\widehat{\varphi} = 1.025$. The fitted covariance and correlation matrices are displayed in the lower and upper triangular parts of the following matrix:

$$\widehat{\mathbf{\Sigma}} \backslash \widehat{\mathbf{\Psi}} = \begin{pmatrix} 61.87 & 0.29 & 0.34 & 0.64 & 0.51 \\ 15.90 & 47.25 & 0.21 & 0.24 & 0.61 \\ 16.45 & 8.89 & 37.15 & 0.13 & 0.21 \\ 21.81 & 7.21 & 3.51 & 19.00 & 0.49 \\ 26.44 & 27.80 & 8.33 & 14.13 & 44.17 \end{pmatrix}. \tag{49}$$

The lack of fit X^2 statistic is 2.87 (2.65 after Bartlett correction) with 5 degrees of freedom. See Table 2 in [7] for details on the lack of fit test.

The eigenvalues of the sample correlation matrix, $\mathbf{R}$, and the fitted correlation matrix, $\widehat{\mathbf{\Psi}} = \widehat{\mathbf{\Gamma}}\widehat{\mathbf{\Lambda}}\widehat{\mathbf{\Gamma}}'$ are displayed in Table 2. Also displayed are standard errors from (6) and estimated bias of the eigenvalues of $\widehat{\mathbf{\Psi}}$, computed with and without the assumption of normality. See [7] for details on the estimator of bias. The estimated biases and standard errors do not depend too much on whether the normal theory or the ADF estimators are employed because the distribution of skull measurements does not strongly depart from normality.

Two-sided 90% confidence intervals for each of the distinct eigenvalues of the correlation matrix are displayed in Table 3. The limits of the intervals also serve as limits for one-sided 95% lower and upper confidence intervals. The intervals displayed in Table 3 are the first-order method from (41), the second-order Edgeworth method from (42), the second-order Hall method from (44), the second-order DiCiccio and Monti method from (45), and the second-order Tingley and Field method from (48). The intervals were computed using normal theory (Theorem 2, Eq. (46)) and ADF (Theorem 4, Eq. (47)) estimators of unknown quantities.

The three intervals based on the Edgeworth expansion (Edgeworth, Hall, DiCiccio) are similar to one another and differ from the first-order intervals primarily by a shift that adjusts for bias and skewness. The Tingley and Field intervals appear to be a compromise between the first-order intervals and the Edgeworth-based intervals, at least when ADF estimators are used.

Table 3
90% Confidence intervals for the eigenvalues of the correlation matrix of skull measurements

Method	Normal theory based					
	λ_1		λ_2 and λ_3		λ_4 and λ_5	
	Lower	Upper	Lower	Upper	Lower	Upper
First order	2.121	2.941	0.820	0.997	0.210	0.442
Edgeworth	2.028	2.848	0.827	1.004	0.249	0.482
Hall	2.028	2.848	0.825	1.006	0.248	0.483
DiCiccio	2.029	2.846	0.831	1.002	0.244	0.488
Tingley	2.096	3.078	0.790	1.002	0.171	0.450
Asymptotic distribution free						
First order	2.131	2.931	0.822	0.995	0.212	0.440
Edgeworth	1.987	2.788	0.841	1.014	0.265	0.493
Hall	1.987	2.788	0.840	1.014	0.263	0.496
DiCiccio	1.970	2.797	0.845	1.010	0.256	0.508
Tingley	2.114	2.875	0.834	0.999	0.229	0.445

8. Simulation study

8.1. Edgeworth and saddlepoint approximations under normality

A five-dimensional covariance matrix, $\mathbf{\Sigma}$, was set equal to the fitted covariance matrix from the skull data in (49) and 50, 000 samples were generated from the Wishart distribution $W_5(20, \mathbf{\Sigma})$. The restricted model was fit to each sample covariance matrix and the following functions of the eigenvalues of the correlation matrix were estimated:

$$\begin{pmatrix} \psi_{11} \\ \psi_{12} \\ \psi_{13} \\ \psi_{14} \end{pmatrix} = \begin{pmatrix} \lambda_1 \\ \lambda_2 + \lambda_3 \\ \lambda_4 + \lambda_5 \\ \lambda_1 + \lambda_2 + \lambda_3 \end{pmatrix} = \mathbf{H}'\lambda, \quad \text{where } \mathbf{H} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 \end{pmatrix}'. \tag{50}$$

Fig. 1 displays the kernel smoothed empirical pdf of $\widehat{\psi}_{11}, \ldots, \widehat{\psi}_{14}$. Also displayed are the Edgeworth and saddlepoint approximations from Theorem 3 and (34), respectively, and the first-order normal approximation

$$\widehat{\psi}_{1j} \dot{\sim} \mathrm{N}\left[\psi_{1j}, \operatorname{tr}(\mathbf{\Sigma}\mathbf{V}_j)^2/(2n)\right],$$

where $\mathbf{V}_j$ is $\mathbf{V}$ of (31) in which $\mathbf{h}$ is the jth column of $\mathbf{H}$ in (50). The value $\Delta = 0$ was used. Vertical bars are drawn at the means of the first-order normal approximations. The first-order normal approximations fail to account for the bias and skewness that are present in the distributions of the estimators. The saddlepoint approximations are substantially more accurate than the first-order normal approximations, but they are slightly less accurate than the Edgeworth approximations. The saddlepoint approximations of Gatto and Ronchetti in Section 5.3 also were computed but are not displayed in Fig. 1. They are more accurate

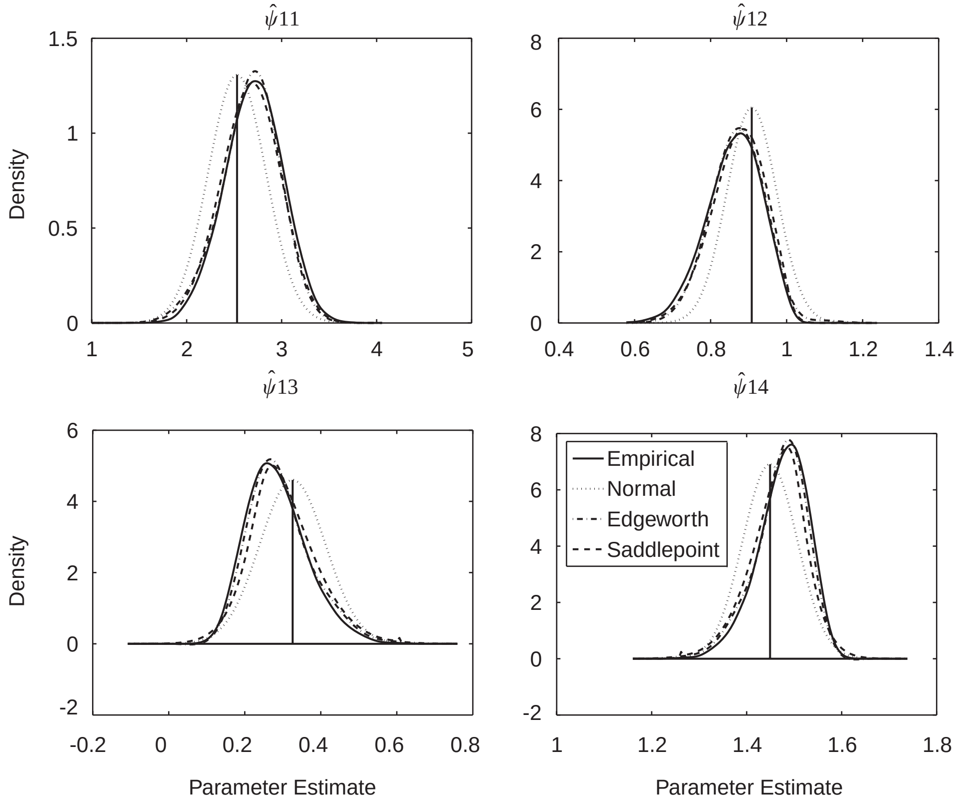

Fig. 1. Density of $\widehat{\psi}_{11}, \ldots, \widehat{\psi}_{14}$ for the skull model.

that the first-order normal approximations, but less accurate than the normal theory-based Edgeworth and saddlepoint approximations.

Fig. 2 displays the difference between the empirical CDF and the CDFs based on the normal theory Edgeworth approximation in Theorem 3, the normal theory saddlepoint approximation in (36), and ADF saddlepoint in Section 5.3. No single approximation is uniformly superior, but the Edgeworth expansion tends to have smaller errors followed in order by the normal theory saddlepoint and ADF saddlepoint approximations.

8.2. Finite-sample coverage of confidence intervals

Random samples were drawn from six multivariate distributions, each having covariance matrix equal to the fitted matrix in (49). The marginal standardized third and fourth cumulants of the five random variables within each distribution are listed in Table 4. The multivariate Bernoulli random variables were generated using the algorithm of Qaqish [33]. Marginally, each of the five random variables in this distribution has a Bernoulli distribution with probability of success 0.28. This is the smallest probability for which Qaqish's

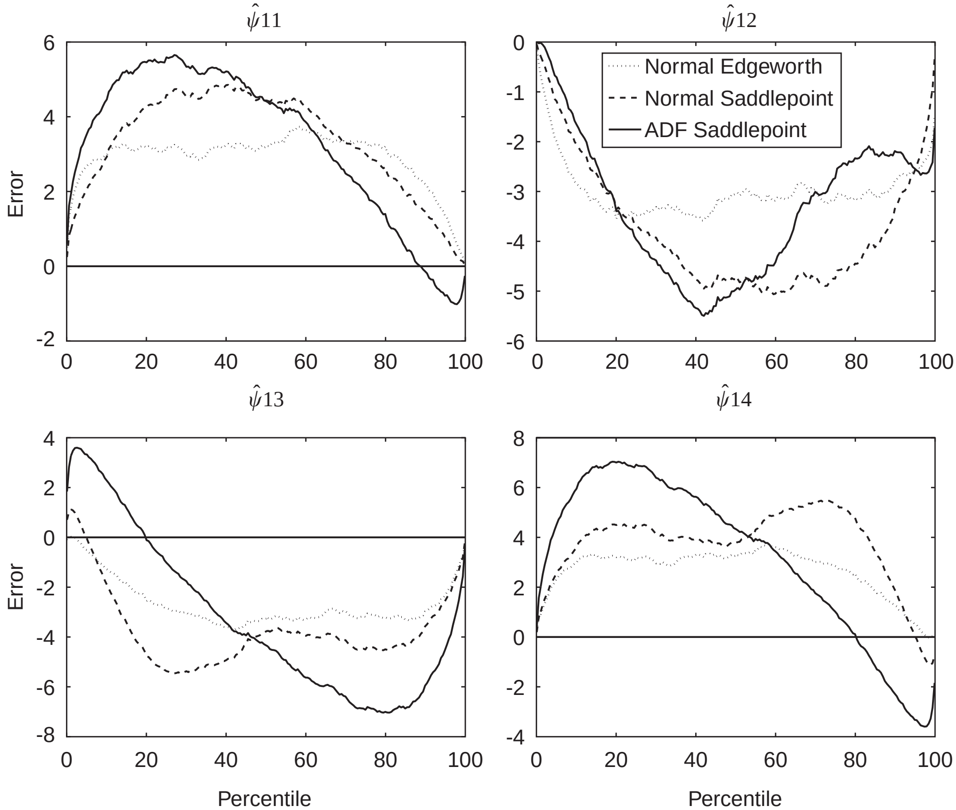

Fig. 2. Error in CDF of $\widehat{\psi}_{11}, \ldots, \widehat{\psi}_{14}$ for the skull model.

multivariate Bernoulli distribution exists, subject to $\mathbf{\Sigma} = \widehat{\mathbf{\Sigma}}$ of (49). The multivariate Bernoulli distribution does not satisfy regularity condition (b*) in Section 5.1. Therefore, there is no assurance that the associated Edgeworth expansions are valid. The parameters of the multivariate lognormal distributions were chosen to attain specific marginal standardized third cumulants and $\mathbf{\Sigma} = \widehat{\mathbf{\Sigma}}$ of (49). Random samples of size $N = 21$, 51, 101, 201, 501, and 1001 were drawn from each distribution except lognormal 4. Random samples of size $N = 201$, 501, 1001, and 5001 were drawn from the lognormal 4 distribution.

For each sample size, 5000 samples were drawn from the parent distribution in Table 4. If the covariance matrix based on multivariate Bernoulli sampling was singular, then the sample was discarded and a new sample was drawn. This occurred in less than 10% of the samples when $N = 21$ and did not occur when $N \geqslant 51$. For each sample, 95% one-sided lower and upper confidence intervals for the distinct eigenvalues of the correlation matrix were computed by the methods illustrated in Table 3. It is possible that the DiCiccio and Monti interval in 45 cannot be computed because the inverse Box–Cox transformation fails to exist for one or both endpoints. If an endpoint of a DiCiccio and Monti interval could not be computed, then the endpoint was equated to the corresponding first-order endpoint computed on the same data set.

Table 4
Marginal standardized cumulants of population distributions

Distribution	Variable				
	1	2	3	4	5
Normal					
κ_{3j}	0.00	0.00	0.00	0.00	0.00
κ_{4j}	0.00	0.00	0.00	0.00	0.00
Bernoulli					
κ_{3j}	0.98	0.98	0.98	0.98	0.98
κ_{4j}	−1.04	−1.04	−1.04	−1.04	−1.04
Lognormal 1					
κ_{3j}	0.25	0.50	0.75	1.00	1.25
κ_{4j}	0.11	0.45	1.02	1.83	2.90
Lognormal 2					
κ_{3j}	0.50	1.00	1.50	2.00	2.50
κ_{4j}	0.45	1.83	4.25	7.86	12.85
Lognormal 3					
κ_{3j}	1.00	2.00	3.00	4.00	5.00
κ_{4j}	1.83	7.86	19.40	38.00	65.26
Lognormal 4					
κ_{3j}	2.00	4.00	6.00	8.00	10.00
κ_{4j}	7.86	38.00	102.76	214.49	384.78

$\kappa_{3j} = (Y_j - \mu_j)^3/\sigma_j^3$, $\kappa_{4j} = (Y_j - \mu_j)^4/\sigma_j^4 - 3$.

Minimum and maximum coverage rates were estimated empirically for each (method, sample size) combination. For example, 5000 samples of size $N = 21$ were drawn from the multivariate normal distribution $\mathrm{N}(\mathbf{1}_n\boldsymbol{\mu}', \boldsymbol{\Sigma})$, where $\boldsymbol{\Sigma} = \widehat{\boldsymbol{\Sigma}}$ of (49). Empirical coverage rates of the first-order normal-theory intervals for $(\lambda_1, \lambda_2, \lambda_4)$ were (0.8258, 0.9688, 0.9930) for the lower intervals and (0.9856, 0.8762, 0.7778) for the upper intervals yielding minimum and maximum coverage rates of 0.7778 and 0.9930. The corresponding coverage rates for the first-order ADF method were (0.7892, 0.9572, 0.9836) for the lower intervals and (0.9762, 0.8298, 0.7480) for the upper intervals yielding minimum and maximum coverage rates of 0.7480 and 0.9836.

Fig. 3 displays the minimum and maximum coverage rates when sampling from either multivariate normal or multivariate Bernoulli distributions. If normality is satisfied, then coverage for all confidence intervals converges to $1 - \alpha = 0.95$ as $n \to \infty$. It is apparent in the left-hand panel of Figure 3 that the coverage rates for second-order Edgeworth-based methods (Edgeworth, Hall, and DiCiccio) converge to $1 - \alpha$ substantially faster than do the first-order methods. Using normal-theory estimators when normality actually is satisfied yields minimum coverage rates that are slightly closer to $1 - \alpha$ than are the minimum coverage rates of intervals based on ADF estimators. The penalty for using ADF estimators, however, decreases rapidly as n increases. The right-hand panel of Fig. 3 shows that ADF methods are superior to normal-theory methods and that second-order ADF methods are superior to the first-order ADF method when sampling from the multivariate Bernoulli

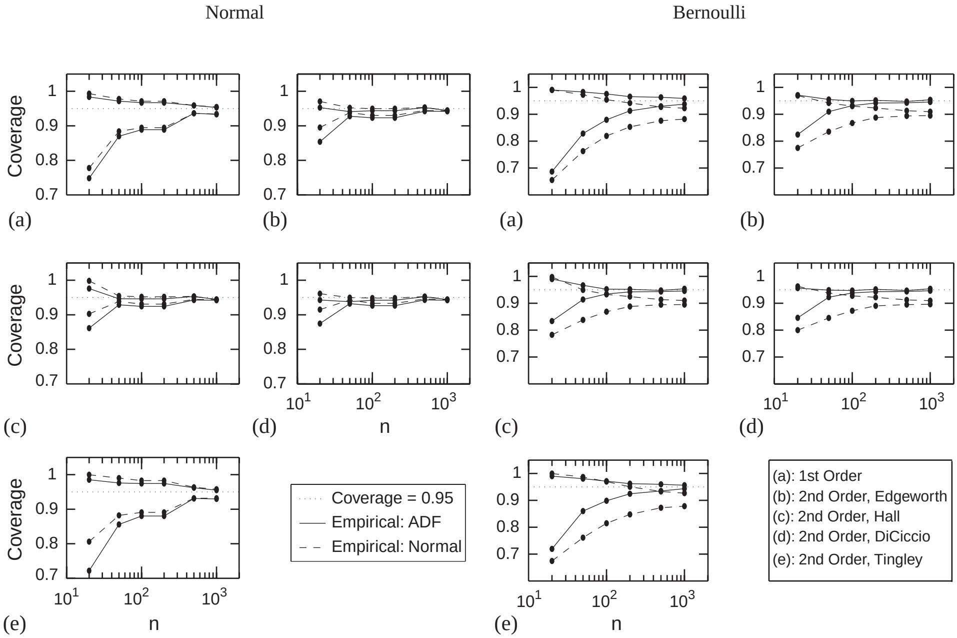

Fig. 3. Coverage when sampling from normal or Bernoulli distributions.

distribution. That is, employing skewness and bias corrections yields superior confidence intervals even though regularity condition (b*) in Section 5.1 is not satisfied. Under Bernoulli sampling, the distribution of Z in (6) is discrete with as many as $\binom{N+2^5-1}{N}$ atoms. Presumably, the probability mass is spread among these atoms in such a manner that the step sizes in the cdf of Z are no larger than $O(n^{-1})$. Fig. 3 also reveals that, under Bernoulli sampling, coverage rate of the normal-theory-based intervals converges to a value smaller that $1-\alpha$. Overall, the three Edgeworth-based methods perform best.

Fig. 4 displays minimum and maximum confidence interval coverage when sampling from multivariate lognormal distributions. Collectively, the four panels show that (a) if deviation from normality is slight (lognormal 1) and sample size is small, then normal-theory intervals are superior to ADF intervals; (b) as deviation from normality increases, performance of normal-theory intervals degrades and ADF intervals, both first and second order, are superior to the corresponding normal-theory intervals; (c) the performance of second-order ADF Edgeworth-based intervals (Edgeworth, Hall, DiCiccio) is superior to that of ADF Tingley intervals; (e) coverage of second-order ADF Edgeworth-based intervals is superior to that of first-order ADF intervals; and (d) if deviation from normality is large (lognormal 4), then sample size $N = 5001$ is too small to ensure coverage of $1-\alpha$, even for second-order ADF methods.

Of the 680, 000 DiCiccio and Monti confidence interval endpoints depicted in Figs. 3 and 4, 2037 endpoints could not be computed. Most of these failures (2004) were ADF endpoints and this occurred because the ADF estimators of bias and skewness are more variable than are the normal-theory estimators.

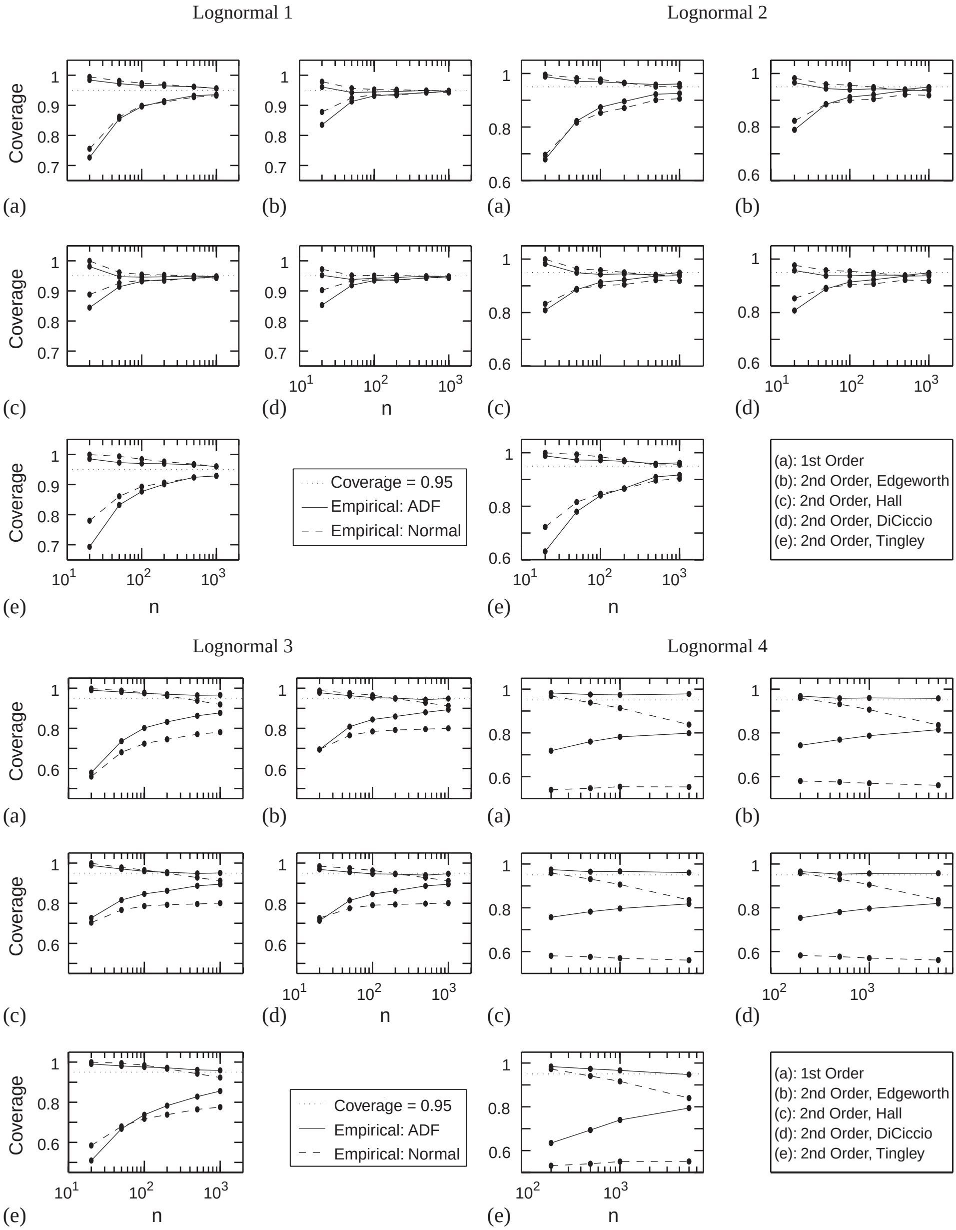

Fig. 4. Coverage when sampling from lognormal distributions.

Overall, the three Edgeworth-based ADF intervals in Eqs. (42), (44), (45) are recommended. These methods performed substantially better than first-order methods under all conditions; nearly as well as second-order normal-theory methods in cases where the parent distribution is normal or nearly so; and substantially better than normal theory methods when the parent distribution is not normal. Approximately $\frac{1}{2}$ of 1% of the ADF DiCiccio and Monti endpoints could not be computed, but in these instances, second-order Edgeworth or Hall endpoints can be substituted. If deviation from normality is large, then none of the methods perform well. In this case, intervals based on robust estimators of the eigenvalues likely would perform better.

Appendix A. Weights for sixth-order moments

The coefficient matrix $\mathbf{W}$ in Theorem 1 can be written as

$$\mathbf{W} = \mathbf{D}^{-1}\mathbf{W}^*, \quad \text{where } \mathbf{D} = n \oplus n^3\mathbf{I}_3 \oplus n^2\mathbf{I}_8, \quad \mathbf{W}^* = (\mathbf{w}_1^* \ \mathbf{w}_2^* \ \cdots \ \mathbf{w}_{12}^*),$$

$$(\mathbf{w}_1^* \ \mathbf{w}_2^* \ \mathbf{w}_3^*) = \begin{pmatrix} c_6 & c_8 - 3c_5 + 2c_6 & c_8 - 3c_5 + 2c_6 \\ c_8 & n^3 - 3nc_1 + 2c_8 & n^2 - (n+2)c_1 + 2c_8 \\ 4c_8 & 4[n^2 - (n+2)c_1 + 2c_8] & n(n^2+n+2) - 4(n+2)c_1 + 8c_8 \\ 2c_8 & 2[n^2 - (n+2)c_1 + 2c_8] & 2(n - 3c_1 + 2c_8) \\ 8c_8 & 8(n - 3c_1 + 2c_8) & 4[n(n+1) - (n+5)c_1 + 4c_8] \\ 2c_5 & 2(nc_1 + 2c_5 - nc_2 - 2c_9) & (n+1)c_1 + 4c_5 - (n+2)c_2 - 3c_9 \\ c_5 & nc_1 + 2c_5 - nc_2 - 2c_9 & c_1 + 2c_5 - 2c_2 - c_9 \\ 8c_5 & 8(c_1 + 2c_5 - 2c_2 - c_9) & 2[(n+3)c_1 + 8c_5 - (n+6)c_2 - 5c_9] \\ 4c_5 & 4(c_1 + 2c_5 - 2c_2 - c_9) & 4(c_1 + 2c_5 - 2c_2 - c_9) \\ 4c_7 & 4(c_3 + 2c_7 - 2c_{10} - c_{11}) & 2(c_3 + c_4 + 4c_7 - 2c_{10} - 4c_{11}) \\ 2c_7 & 2(c_3 + 2c_7 - 2c_{10} - c_{11}) & 2(c_3 + 2c_7 - 2c_{10} - c_{11}) \\ 4c_7 & 4(c_4 + 2c_7 - 3c_{11}) & 2(c_3 + c_4 + 4c_7 - 2c_{10} - 4c_{11}) \end{pmatrix},$$

$$(\mathbf{w}_4^* \ \mathbf{w}_5^*) = \begin{pmatrix} c_8 - 3c_5 + 2c_6 & c_8 - 3c_5 + 2c_6 \\ n^2 - (n+2)c_1 + 2c_8 & n - 3c_1 + 2c_8 \\ 4(n - 3c_1 + 2c_8) & 2[n(n+1) - (n+5)c_1 + 4c_8] \\ n^2(n+1) - 2(2n+1)c_1 + 4c_8 & n(n+1) - (n+5)c_1 + 4c_8 \\ 4[n(n+1) - (n+5)c_1 + 4c_8] & n(n^2+3n+4) - 6(n+3)c_1 + 16c_8 \\ 2(c_1 + 2c_5 - 2c_2 - c_9) & 2(c_1 + 2c_5 - 2c_2 - c_9) \\ nc_1 + 2c_5 - nc_2 - 2c_9 & c_1 + 2c_5 - 2c_2 - c_9 \\ 8(c_1 + 2c_5 - 2c_2 - c_9) & 2[(n+3)c_1 + 8c_5 - (n+6)c_2 - 5c_9] \\ 2[(n+1)c_1 + 4c_5 - (n+2)c_2 - 3c_9] & (n+3)c_1 + 8c_5 - (n+6)c_2 - 5c_9 \\ 4(c_3 + 2c_7 - 2c_{10} - c_{11}) & 2(c_3 + c_4 + 4c_7 - 2c_{10} - 4c_{11}) \\ 2(c_4 + 2c_7 - 3c_{11}) & c_3 + c_4 + 4c_7 - 2c_{10} - 4c_{11} \\ 2(c_3 + c_4 + 4c_7 - 2c_{10} - 4c_{11}) & 3c_3 + c_4 + 8c_7 - 6c_{10} - 6c_{11} \end{pmatrix},$$

$$
(\mathbf{w}_6^* \ \cdots \ \mathbf{w}_9^*) = \begin{pmatrix}
c_5 - c_6 & c_5 - c_6 & c_5 - c_6 & c_5 - c_6 \\
nc_1 - c_8 & nc_1 - c_8 & c_1 - c_8 & c_1 - c_8 \\
2[(n+1)c_1 - 2c_8] & 4(c_1 - c_8) & (n+3)c_1 - 4c_8 & 4(c_1 - c_8) \\
2(c_1 - c_8) & 2(nc_1 - c_8) & 2(c_1 - c_8) & (n+1)c_1 - 2c_8 \\
8(c_1 - c_8) & 8(c_1 - c_8) & 2[(n+3)c_1 - 4c_8] & 2[(n+3)c_1 - 4c_8] \\
nc_2 + c_9 - 2c_5 & 2(c_9 - c_5) & c_2 + c_9 - 2c_5 & 2(c_2 - c_5) \\
c_9 - c_5 & nc_2 - c_5 & c_2 - c_5 & c_9 - c_5 \\
4(c_2 + c_9 - 2c_5) & 8(c_2 - c_5) & (n+4)c_2 + 3c_9 - 8c_5 & 4(c_2 + c_9 - 2c_5) \\
4(c_2 - c_5) & 4(c_9 - c_5) & 2(c_2 + c_9 - 2c_5) & (n+2)c_2 + c_9 - 4c_5 \\
2(c_{10} + c_{11} - 2c_7) & 4(c_{10} - c_7) & c_{10} + 3c_{11} - 4c_7 & 2(c_{10} + c_{11} - 2c_7) \\
2(c_{10} - c_7) & 2(c_{11} - c_7) & c_{10} + c_{11} - 2c_7 & 2(c_{11} - c_7) \\
4(c_{11} - c_7) & 4(c_{11} - c_7) & 2(c_{10} + c_{11} - 2c_7) & 2(c_{10} + c_{11} - 2c_7)
\end{pmatrix},
$$

$$
(\mathbf{w}_{10}^* \ \cdots \ \mathbf{w}_{12}^*) = \begin{pmatrix}
c_7 - c_6 & c_7 - c_6 & c_7 - c_6 \\
c_3 - c_8 & c_3 - c_8 & c_4 - c_8 \\
2(c_3 + c_4 - 2c_8) & 4(c_3 - c_8) & 2(c_3 + c_4 - 2c_8) \\
2(c_3 - c_8) & 2(c_4 - c_8) & c_3 + c_4 - 2c_8 \\
4(c_3 + c_4 - 2c_8) & 4(c_3 + c_4 - 2c_8) & 2(3c_3 + c_4 - 4c_8) \\
c_{10} + c_{11} - 2c_5 & 2(c_{10} - c_5) & 2(c_{11} - c_5) \\
c_{10} - c_5 & c_{11} - c_5 & c_{11} - c_5 \\
2(c_{10} + 3c_{11} - 4c_5) & 4(c_{10} + c_{11} - 2c_5) & 4(c_{10} + c_{11} - 2c_5) \\
2(c_{10} + c_{11} - 2c_5) & 4(c_{11} - c_5) & 2(c_{10} + c_{11} - 2c_5) \\
3c_{12} + c_4^2 - 4c_7 & 4(c_{12} - c_7) & 4(c_{12} - c_7) \\
2(c_{12} - c_7) & c_{12} + c_4^2 - 2c_7 & 2(c_{12} - c_7) \\
4(c_{12} - c_7) & 4(c_{12} - c_7) & 3c_{12} + c_4^2 - 4c_7
\end{pmatrix},
$$

and c_i for $i = 1, \ldots, 12$ are given in Table 1.

The coefficient vectors $\mathbf{w}_{222}$ and $\mathbf{w}_{42}$ in (16) are displayed in Table A.1.

Appendix B. Expressions for ω_i in Theorem 2

The quantities ω_i for $i = 1, 2, 3, 4, 6$ can be written as follows:

$$
\omega_1 = \tfrac{1}{2}\mathbf{a}_2' \operatorname{vec}(\bar{\mathbf{I}}_{\boldsymbol{\theta}}) - \tfrac{1}{4}\mathbf{v}' \mathbf{F}^{(2)} \operatorname{vec}\left(\bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1}\right),
$$

$$
\begin{aligned}
\omega_2 = {} & \omega_1^2 - 2\operatorname{tr}\left[\mathbf{F}^{(1)\prime}\left(\mathbf{V}\otimes\boldsymbol{\Sigma}^{-1}\right)\left(\mathbf{I}_{p^2} - \mathbf{P}_F\right)\mathbf{F}^{(2)}(\mathbf{a}\otimes\bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1})\right] + \mathbf{a}'\bar{\mathbf{I}}_{\boldsymbol{\theta}}\mathbf{A}_3\operatorname{vec}\left(\bar{\mathbf{I}}_{\boldsymbol{\theta}}\right) \\
& + 4\operatorname{tr}\left[\mathbf{P}_F\left(\boldsymbol{\Sigma}\mathbf{V}\otimes\mathbf{I}_p\right)\left(\mathbf{I}_{p^2} - \mathbf{P}_F\right)\mathbf{N}_p\left(\boldsymbol{\Sigma}\mathbf{V}\otimes\mathbf{I}_p\right)\right] - \tfrac{1}{2}\mathbf{a}'\bar{\mathbf{I}}_{\boldsymbol{\theta}}\mathbf{A}_2\,\mathbf{F}^{(1)\prime}\,\ddot{\mathbf{F}}^{(2)}\operatorname{vec}\left(\bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1}\right) \\
& + \tfrac{1}{2}\operatorname{tr}\left[\ddot{\mathbf{F}}^{(2)\prime}\left(\mathbf{I}_{p^2} - \mathbf{P}_F\right)\mathbf{F}^{(2)}(\mathbf{a}\mathbf{a}'\otimes\bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1})\right] + \operatorname{tr}\left[\mathbf{A}_2\mathbf{F}^{(1)\prime}\left(\mathbf{V}\otimes\boldsymbol{\Sigma}^{-1}\right)\mathbf{F}^{(1)}\right] \\
& - \tfrac{1}{2}\mathbf{v}'\,\mathbf{F}^{(2)}\,\mathbf{a}_2 - \frac{1}{2}\mathbf{v}'\,\mathbf{F}^{(3)}\left[\mathbf{a}\otimes\operatorname{vec}\left(\bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1}\right)\right] + \tfrac{1}{8}\mathbf{v}'\,\mathbf{F}^{(2)}\left(\bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1}\otimes\bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1}\right)\mathbf{F}^{(2)\prime}\,\mathbf{v}
\end{aligned}
$$

Table A.1
Coefficients for unbiased estimators of $\mathbf{\Omega}_{222,n}$ and $\mathbf{\Omega}_{42,n}$ when $\mathbf{X} = \mathbf{1}_N$

i	$w_{222,i}/n^{3/2}$	$w_{42,i}/n^2$
1	$\frac{\left(n^3+3n^2+2n+12\right)}{(n-4)_4\,(n+1)}$	$\frac{(n+2)\left(n^2-n+2\right)}{(n-4)_4\,(n+1)}$
2	$2\frac{\left(n^4-n^3-12n^2+8n+12\right)}{(n-4)_4\,(n+1)^2}$	$2\frac{\left(n^4-2n^3-9n^2+8n+12\right)}{(n-4)_4\,(n+1)^2}$
3	$2\frac{n\left(2n^2-5n-1\right)}{(n-4)_4\,(n+1)^2}$	$2\frac{n\left(n^2-2n-1\right)}{(n-4)_4\,(n+1)^2}$
4	$2\frac{n\left(2n^2-5n-1\right)}{(n-4)_4\,(n+1)^2}$	$2\frac{n\left(2n^2-5n-5\right)}{(n-4)_4\,(n+1)^2}$
5	$2\frac{n\left(n^2-2n+3\right)}{(n-4)_4\,(n+1)^2}$	$\frac{n\left(n^2-n+2\right)}{(n-4)_4\,(n+1)^2}$
6	$-\frac{\left(n^4+n^3-8n^2-16n+4\right)}{(n-4)_4\,(n+1)^2}$	$-\frac{\left(n^4-n^3-6n^2-4n+4\right)}{(n-4)_4\,(n+1)^2}$
7	$-\frac{\left(n^4+n^3-8n^2-16n+4\right)}{(n-4)_4\,(n+1)^2}$	$-\frac{(n+2)\left(n^3-n^2-8n-2\right)}{(n-4)_4\,(n+1)^2}$
8	$-\frac{\left(2n^3+n^2-5n+8\right)}{(n-4)_4\,(n+1)^2}$	$-\frac{(n+2)\left(n^2-n+2\right)}{(n-4)_4\,(n+1)^2}$
9	$-\frac{\left(2n^3+n^2-5n+8\right)}{(n-4)_4\,(n+1)^2}$	$-\frac{n\left(2n^2-n-7\right)}{(n-4)_4\,(n+1)^2}$
10	$-\frac{\left(n^3-n^2-2n+4\right)}{(n-1)\,(n-4)_2\,(n+1)^2}$	$-\frac{n\left(n^2-2\right)}{(n-2)_2\,(n-4)\,(n+1)^2}$
11	$-\frac{\left(n^3-n^2-2n+4\right)}{(n-1)\,(n-4)_2\,(n+1)^2}$	$-\frac{\left(n^3+n^2-4n-8\right)}{(n-4)_4\,(n+1)^2}$
12	$-\frac{\left(5n^2+5n-12\right)}{(n-4)_4\,(n+1)^2}$	$-\frac{\left(3n^2+3n-4\right)}{(n-4)_4\,(n+1)^2}$

$(n-a)_b$ is Pochhammer's symbol, $(n-a)_b = (n-a)(n-a+1)\cdots(n-a+b-1)$.

$$
\begin{aligned}
&-\tfrac{1}{2}\mathbf{v}'\,\mathbf{F}^{(2)}\,\mathrm{vec}\left[\bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1}\mathbf{F}^{(1)\prime}\left(\mathbf{V}\otimes\mathbf{\Sigma}^{-1}\right)\mathbf{F}^{(1)}\bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1}-\bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1}\,\mathbf{F}^{(1)\prime}\,\ddot{\mathbf{F}}^{(2)}\left(\mathbf{a}\otimes\bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1}\right)\right]\\
&-\mathrm{tr}\left[\mathbf{F}^{(1)\prime}\,\ddot{\mathbf{F}}^{(2)}\,(\mathbf{a}\otimes\mathbf{A}_2)\right]+\left[\mathrm{vec}(\mathbf{V\Sigma V})\right]'\left(\mathbf{I}_{p^2}-\mathbf{P}_F\right)\mathbf{F}^{(2)}\,\mathrm{vec}\left(\bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1}\right)\\
&+\tfrac{1}{4}\mathbf{v}'\,\mathbf{F}^{(2)}\left[\mathbf{a}\otimes\bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1}\,\mathbf{F}^{(1)\prime}\,\ddot{\mathbf{F}}^{(2)}\,\mathrm{vec}\left(\bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1}\right)\right]+\tfrac{1}{2}\,\mathrm{tr}\left[\mathbf{A}_2\bar{\mathbf{I}}_{\boldsymbol{\theta}}\mathbf{A}_2\bar{\mathbf{I}}_{\boldsymbol{\theta}}\right]
\end{aligned}
$$

$$-\tfrac{1}{2}\left[\operatorname{vec}\left(\bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1}\right)\right]'\ddot{\mathbf{F}}^{(2)\prime}\left(\mathbf{I}_{p^2}-\mathbf{P}_F\right)\mathbf{F}^{(2)}(\mathbf{a}\otimes\mathbf{a}),$$

$$\omega_3 = -\tfrac{3}{2}\mathbf{v}'\,\mathbf{F}^{(2)}(\mathbf{a}\otimes\mathbf{a}) + 3\mathbf{a}'\bar{\mathbf{I}}_{\boldsymbol{\theta}}\mathbf{A}_2\bar{\mathbf{I}}_{\boldsymbol{\theta}}\mathbf{a} + \operatorname{tr}(\boldsymbol{\Sigma}\mathbf{V})^3,$$

$$\begin{aligned}
\omega_4 = {} & 4\omega_1\omega_3 - 6\mathbf{v}'\,\mathbf{F}^{(2)}\left[\mathbf{a}\otimes\bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1}\,\mathbf{F}^{(1)\prime}\operatorname{vec}(\mathbf{V}\boldsymbol{\Sigma}\mathbf{V})\right] + 12\left[\operatorname{vec}(\mathbf{V}\boldsymbol{\Sigma}\mathbf{V})\right]'\mathbf{F}^{(1)}\mathbf{A}_2\bar{\mathbf{I}}_{\boldsymbol{\theta}}\mathbf{a} \\
& +12\mathbf{a}'\bar{\mathbf{I}}_{\boldsymbol{\theta}}\mathbf{A}_2\bar{\mathbf{I}}_{\boldsymbol{\theta}}\mathbf{A}_2\bar{\mathbf{I}}_{\boldsymbol{\theta}}\mathbf{a} - 12\mathbf{v}'\,\mathbf{F}^{(2)}\left[\mathbf{a}\otimes\mathbf{A}_2\bar{\mathbf{I}}_{\boldsymbol{\theta}}\mathbf{a}\right] + 3\operatorname{tr}(\boldsymbol{\Sigma}\mathbf{V})^4 \\
& -2\mathbf{v}'\,\mathbf{F}^{(3)}(\mathbf{a}\otimes\mathbf{a}\otimes\mathbf{a}) - 6\mathbf{a}'\bar{\mathbf{I}}_{\boldsymbol{\theta}}\mathbf{A}_2\,\mathbf{F}^{(1)\prime}\,\ddot{\mathbf{F}}^{(2)}(\mathbf{a}\otimes\mathbf{a}) + 4\mathbf{a}'\bar{\mathbf{I}}_{\boldsymbol{\theta}}\mathbf{A}_3\left[\bar{\mathbf{I}}_{\boldsymbol{\theta}}\mathbf{a}\otimes\bar{\mathbf{I}}_{\boldsymbol{\theta}}\mathbf{a}\right] \\
& +3\mathbf{v}'\,\mathbf{F}^{(2)}\left[\mathbf{a}\otimes\bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1}\,\mathbf{F}^{(1)\prime}\,\ddot{\mathbf{F}}^{(2)}(\mathbf{a}\otimes\mathbf{a})\right] + 3\mathbf{v}'\,\mathbf{F}^{(2)}\left(\mathbf{a}\mathbf{a}'\otimes\bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1}\right)\mathbf{F}^{(2)\prime}\,\mathbf{v}, \\
& \omega_6 = 10\omega_3^2, \\
& \mathbf{P}_F = \tfrac{1}{2}\mathbf{F}^{(1)}\bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1}\,\ddot{\mathbf{F}}^{(1)\prime}, \quad \mathbf{a} = \bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1}\mathbf{E}_{\varphi}\mathbf{D}_{\lambda:\boldsymbol{\varphi}}^{(1)\prime}\mathbf{h}, \\
& \mathbf{a}_2 = \left(\bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1}\mathbf{E}_{\boldsymbol{\varphi}}\otimes\bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1}\mathbf{E}_{\boldsymbol{\varphi}}\right)\mathbf{D}_{\lambda:\boldsymbol{\varphi},\boldsymbol{\varphi}}^{(2)\prime}\mathbf{h}, \quad \mathbf{A}_2 = \operatorname{dvec}(\mathbf{a}_2, \dot{\nu}, \dot{\nu}), \\
& \mathbf{a}_3 = \left(\bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1}\mathbf{E}_{\boldsymbol{\varphi}}\otimes\bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1}\mathbf{E}_{\boldsymbol{\varphi}}\otimes\bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1}\mathbf{E}_{\boldsymbol{\varphi}}\right)\mathbf{D}_{\lambda:\boldsymbol{\varphi},\boldsymbol{\varphi},\boldsymbol{\varphi}}^{(3)\prime}\mathbf{h}, \text{ and } \mathbf{A}_3 = \operatorname{dvec}(\mathbf{a}_3, \dot{\nu}, \dot{\nu}^2).
\end{aligned}$$

The quantity $\mathbf{P}_F$ is the projection operator that projects onto $\mathcal{R}(\mathbf{F}^{(1)})$ along $\mathcal{N}(\ddot{\mathbf{F}}^{(1)\prime})$.

Appendix C. Expressions for ω_{2g} and ω_{4g} in (37)

The quantities ω_{2g} and ω_{4g} can be written as follows:

$$\begin{aligned}
\omega_{2g} = {} & \left(\frac{2m}{m-1}\right)\Bigg[\frac{1}{64}\operatorname{tr}\left(\mathbf{V}_2\boldsymbol{\Omega}_{22}\mathbf{V}_2\boldsymbol{\Omega}_{22}\right) \\
& +\frac{1}{16}\operatorname{tr}\left\{\ddot{\mathbf{F}}^{(2)\prime}\left(\mathbf{I}_{p^2}-\mathbf{P}_F\right)\boldsymbol{\Omega}_{22}\left(\mathbf{I}_{p^2}-\mathbf{P}_F\right)'\ddot{\mathbf{F}}^{(2)}\left(\mathbf{a}\mathbf{a}'\otimes\bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1}\,\ddot{\mathbf{F}}^{(1)\prime}\,\boldsymbol{\Omega}_{22}\,\ddot{\mathbf{F}}^{(1)}\,\bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1}\right)\right\} \\
& +\operatorname{tr}\left\{\mathbf{P}_F{}'\left(\boldsymbol{\Sigma}^{-1}\otimes\mathbf{V}\right)\left(\mathbf{I}_{p^2}-\mathbf{P}_F\right)\boldsymbol{\Omega}_{22}\left(\mathbf{I}_{p^2}-\mathbf{P}_F\right)'\left(\boldsymbol{\Sigma}^{-1}\otimes\mathbf{V}\right)\mathbf{P}_F\boldsymbol{\Omega}_{22}\right\} \\
& +\frac{1}{16^2}\mathbf{v}'\,\mathbf{F}^{(2)}\left(\bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1}\,\ddot{\mathbf{F}}^{(1)\prime}\,\boldsymbol{\Omega}_{22}\,\ddot{\mathbf{F}}^{(1)}\,\bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1}\otimes\bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1}\,\ddot{\mathbf{F}}^{(1)\prime}\,\boldsymbol{\Omega}_{22}\,\ddot{\mathbf{F}}^{(1)}\,\bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1}\right)\mathbf{F}^{(2)\prime}\,\mathbf{v} \\
& -\frac{1}{2}\operatorname{tr}\left\{(\mathbf{a}'\otimes\ddot{\mathbf{F}}^{(1)}\,\bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1})\,\ddot{\mathbf{F}}^{(2)\prime}\left(\mathbf{I}_{p^2}-\mathbf{P}_F\right)\boldsymbol{\Omega}_{22}\left(\mathbf{I}_{p^2}-\mathbf{P}_F\right)'\left(\boldsymbol{\Sigma}^{-1}\otimes\mathbf{V}\right)\mathbf{P}_F\boldsymbol{\Omega}_{22}\right\} \\
& -\frac{1}{32}\mathbf{v}'\,\mathbf{F}^{(2)}\operatorname{vec}\left\{\bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1}\,\ddot{\mathbf{F}}^{(1)\prime}\,\boldsymbol{\Omega}_{22}\left(\mathbf{I}_{p^2}-\mathbf{P}_F\right)'\ddot{\mathbf{F}}^{(2)}\left(\mathbf{a}\otimes\bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1}\,\ddot{\mathbf{F}}^{(1)\prime}\,\boldsymbol{\Omega}_{22}\,\ddot{\mathbf{F}}^{(1)}\,\bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1}\right)\right\} \\
& +\frac{1}{16}\operatorname{tr}\left\{\left(\mathbf{a}'\otimes\ddot{\mathbf{F}}^{(1)}\,\bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1}\right)\ddot{\mathbf{F}}^{(2)\prime}\left(\mathbf{I}_{p^2}-\mathbf{P}_F\right)\boldsymbol{\Omega}_{22}\mathbf{V}_2\boldsymbol{\Omega}_{22}\right\} \\
& +\frac{1}{8}\mathbf{v}'\,\mathbf{F}^{(2)}\operatorname{vec}\left\{\bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1}\,\ddot{\mathbf{F}}^{(1)\prime}\,\boldsymbol{\Omega}_{22}\left(\mathbf{I}_{p^2}-\mathbf{P}_F\right)'\left(\boldsymbol{\Sigma}^{-1}\otimes\mathbf{V}\right)\mathbf{P}_F\boldsymbol{\Omega}_{22}\,\ddot{\mathbf{F}}^{(1)}\,\bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1}\right\} \\
& -\frac{1}{4}\operatorname{tr}\left\{\mathbf{P}_F{}'\left(\boldsymbol{\Sigma}^{-1}\otimes\mathbf{V}\right)\left(\mathbf{I}_{p^2}-\mathbf{P}_F\right)\boldsymbol{\Omega}_{22}\mathbf{V}_2\boldsymbol{\Omega}_{22}\right\} \\
& -\frac{1}{64}\mathbf{v}'\,\mathbf{F}^{(2)}\operatorname{vec}\left\{\bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1}\,\ddot{\mathbf{F}}^{(1)\prime}\,\boldsymbol{\Omega}_{22}\mathbf{V}_2\boldsymbol{\Omega}_{22}\,\ddot{\mathbf{F}}^{(1)}\,\bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1}\right\}\Bigg]
\end{aligned}$$

and

$$
\begin{aligned}
\omega_{4g} = {} & \tfrac{1}{16}(\mathbf{v}\otimes\mathbf{v})'\gamma_{44}(\mathbf{v}\otimes\mathbf{v}) + \tfrac{3}{2}\left[\operatorname{tr}(\mathbf{V\Sigma})\right]^2\sigma_W^2 \\
& + \tfrac{3}{16}\left[\operatorname{tr}(\mathbf{V\Sigma})\right]^4 - \tfrac{1}{4}\operatorname{tr}(\mathbf{V\Sigma})(\mathbf{v}\otimes\mathbf{v})'\gamma_{42}\mathbf{v} - 3\sigma_W^4 \\
& + 3(\mathbf{v}\otimes\mathbf{v})'\Big\{\left[\gamma_{42} - 2(\boldsymbol{\sigma}\otimes\boldsymbol{\Omega}_{22}) - \operatorname{vec}(\boldsymbol{\Omega}_{22})\boldsymbol{\sigma}' - (\boldsymbol{\sigma}\otimes\boldsymbol{\sigma}\boldsymbol{\sigma}')\right]\otimes\mathbf{v}'\boldsymbol{\Omega}_{22}\Big\} \\
& \times\Big\{\tfrac{1}{4}\operatorname{vec}\Big[(\mathbf{I}_{p^2}-\mathbf{P}_F)'\,\ddot{\mathbf{F}}^{(2)}\Big(\mathbf{a}\otimes\bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1}\,\ddot{\mathbf{F}}^{(1)\prime}\Big)\Big] + \tfrac{1}{8}\mathbf{v}_2 \\
& - \operatorname{vec}\Big[(\mathbf{I}_{p^2}-\mathbf{P}_F)'(\boldsymbol{\Sigma}^{-1}\otimes\mathbf{V})\mathbf{P}_F\Big] - \tfrac{1}{16}\Big(\ddot{\mathbf{F}}^{(1)}\,\bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1}\otimes\ddot{\mathbf{F}}^{(1)}\,\bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1}\Big)\mathbf{F}^{(2)\prime}\,\mathbf{v}\Big\} \\
& + \tfrac{3}{4}(\mathbf{a}\otimes\mathbf{b})'\,\ddot{\mathbf{F}}^{(2)\prime}\left(\mathbf{I}_{p^2}-\mathbf{P}_F\right)\boldsymbol{\Omega}_{22}\left(\mathbf{I}_{p^2}-\mathbf{P}_F\right)'\ddot{\mathbf{F}}^{(2)}(\mathbf{a}\otimes\mathbf{b}) \\
& + 12\mathbf{v}'\boldsymbol{\Omega}_{22}\mathbf{P}_F{}'\Big(\boldsymbol{\Sigma}^{-1}\otimes\mathbf{V}\Big)\left(\mathbf{I}_{p^2}-\mathbf{P}_F\right)\boldsymbol{\Omega}_{22}\left(\mathbf{I}_{p^2}-\mathbf{P}_F\right)'\Big(\boldsymbol{\Sigma}^{-1}\otimes\mathbf{V}\Big)\mathbf{P}_F\boldsymbol{\Omega}_{22}\mathbf{v} \\
& + \tfrac{3}{64}\mathbf{v}'\,\mathbf{F}^{(2)}\Big(\mathbf{b}\mathbf{b}'\otimes\bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1}\,\ddot{\mathbf{F}}^{(1)\prime}\,\boldsymbol{\Omega}_{22}\,\ddot{\mathbf{F}}^{(1)}\,\bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1}\Big)\mathbf{F}^{(2)\prime}\,\mathbf{v} \\
& + \tfrac{3}{16}\mathbf{v}'\boldsymbol{\Omega}_{22}\mathbf{V}_2\boldsymbol{\Omega}_{22}\mathbf{V}_2\boldsymbol{\Omega}_{22}\mathbf{v} \\
& - 6(\mathbf{a}\otimes\mathbf{b})'\,\ddot{\mathbf{F}}^{(2)\prime}\left(\mathbf{I}_{p^2}-\mathbf{P}_F\right)\boldsymbol{\Omega}_{22}\left(\mathbf{I}_{p^2}-\mathbf{P}_F\right)'\Big(\boldsymbol{\Sigma}^{-1}\otimes\mathbf{V}\Big)\mathbf{P}_F\boldsymbol{\Omega}_{22}\mathbf{v} \\
& - \tfrac{3}{8}\mathbf{v}'\,\mathbf{F}^{(2)}\Big\{\mathbf{b}\otimes\bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1}\,\ddot{\mathbf{F}}^{(1)\prime}\,\boldsymbol{\Omega}_{22}\left(\mathbf{I}_{p^2}-\mathbf{P}_F\right)'\ddot{\mathbf{F}}^{(2)}\,(\mathbf{a}\otimes\mathbf{b})\Big\} \\
& + \tfrac{3}{4}(\mathbf{a}\otimes\mathbf{b})'\,\ddot{\mathbf{F}}^{(2)\prime}\left(\mathbf{I}_{p^2}-\mathbf{P}_F\right)\boldsymbol{\Omega}_{22}\mathbf{V}_2\boldsymbol{\Omega}_{22}\mathbf{v} \\
& + \tfrac{3}{2}\mathbf{v}'\,\mathbf{F}^{(2)}\Big\{\mathbf{b}\otimes\bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1}\,\ddot{\mathbf{F}}^{(1)\prime}\,\boldsymbol{\Omega}_{22}\left(\mathbf{I}_{p^2}-\mathbf{P}_F\right)'\Big(\boldsymbol{\Sigma}^{-1}\otimes\mathbf{V}\Big)\mathbf{P}_F\boldsymbol{\Omega}_{22}\mathbf{v}\Big\} \\
& - 3\mathbf{v}'\boldsymbol{\Omega}_{22}\mathbf{P}_F{}'\Big(\boldsymbol{\Sigma}^{-1}\otimes\mathbf{V}\Big)\left(\mathbf{I}_{p^2}-\mathbf{P}_F\right)\boldsymbol{\Omega}_{22}\mathbf{V}_2\boldsymbol{\Omega}_{22}\mathbf{v} \\
& - \tfrac{3}{16}\mathbf{v}'\,\mathbf{F}^{(2)}\Big(\mathbf{b}\otimes\bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1}\,\ddot{\mathbf{F}}^{(1)\prime}\,\boldsymbol{\Omega}_{22}\mathbf{V}_2\boldsymbol{\Omega}_{22}\mathbf{v}\Big)\Big],
\end{aligned}
$$

where

$$
\begin{aligned}
\mathbf{v}_2 &= \Big(\ddot{\mathbf{F}}^{(1)}\,\bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1}\mathbf{E}_{\boldsymbol{\varphi}}\otimes\ddot{\mathbf{F}}^{(1)}\,\bar{\mathbf{I}}_{\boldsymbol{\theta}}^{-1}\mathbf{E}_{\boldsymbol{\varphi}}\Big)\mathbf{D}^{(2)\prime}_{\lambda:\boldsymbol{\varphi},\boldsymbol{\varphi}}\mathbf{h}, \\
\mathbf{V}_2 &= \operatorname{dvec}\Big(\mathbf{v}_2,\,p^2,\,p^2\Big),
\end{aligned}
$$

and the remaining terms are defined in (8) and in Appendix B.

References

[1] T.W. Anderson, Asymptotic theory for principal component analysis, Ann. Math. Statist. 34 (1963) 122–148.

[2] G.J. Babu, Z.D. Bai, Edgeworth expansions of a function of sample means under minimal moment conditions and partial Cramér's condition, Sankhyā 55 (1993) 244–258.

[3] O.E. Barndorff-Nielsen, D.R. Cox, Asymptotic Techniques for use in Statistics, Chapman & Hall, London, 1989.

[4] R.N. Bhattacharya, J.K. Ghosh, On the validity of the formal Edgeworth expansion, Ann. Statist. 6 (1978) 434–451.
[5] R.J. Boik, A local parameterization of orthogonal and semi-orthogonal matrices with applications, J. Multivariate Anal. 67 (1998) 244–276.
[6] R.J. Boik, Spectral models for covariance matrices, Biometrika 89 (2002) 159–182.
[7] R.J. Boik, Principal Components models for correlation matrices, Biometrika 90 (2003) 679–701.
[8] J.G. Booth, P. Hall, A.T. Wood, On the validity of Edgeworth and saddlepoint approximations, J. Multivariate Anal. 51 (1994) 121–138.
[9] G.E.P. Box, D.R. Cox, An analysis of transformations (with discussion), J. Roy. Statist. Soc. Ser. B 26 (1964) 211–252.
[10] M.W. Browne, Asymptotically distribution-free methods for the analysis of covariance structures, Brit. J. Math. Statist. Psychol. 37 (1984) 62–83.
[11] E.A. Cornish, R.A. Fisher, Moments and cumulants in the specification of distributions, Internat. Statist. Rev. 5 (1937) 307–322.
[12] T.J. DiCiccio, A.C. Monti, Accurate confidence limits for scalar functions of vector M-estimands, Biometrika 89 (2002) 437–450.
[13] G.S. Easton, E. Ronchetti, General saddlepoint approximations with applications to L statistics, J. Amer. Statist. Assoc. 81 (1986) 420–430.
[14] B. Flury, Common Principal Components and Related Models, Wiley, New York, 1988.
[15] Y. Fujikoshi, Asymptotic expansions for the distributions of some function of the latent roots of matrices in three situations, J. Multivariate Anal. 8 (1978) 63–72.
[16] Y. Fujikoshi, Y. Asymptotic expansions for the distributions of the sample roots under nonnormality, Biometrika 67 (1980) 45–51.
[17] R. Gatto, E. Ronchetti, General saddlepoint approximations of marginal densities and tail probabilities, J. Amer. Statist. Assoc. 91 (1996) 666–673.
[18] P. Hall, The Bootstrap and Edgeworth Expansion, Springer, New York, 1992.
[19] P. Hall, On the removal of skewness by transformation, J. Roy. Statistic. Soc. Ser. B 54 (1992) 221–228.
[20] D.J. Hand, F. Daly, A.D. Lunn, K.J. McConway, E. Ostrowski, A Handbook of Small Data Sets, Chapman & Hall, London, 1994.
[21] F. Kong, B. Levin, Edgeworth expansions for the conditional distributions in logistic regression models, J. Statist. Plan. Inference 52 (1996) 109–129.
[22] R.H. Koning, H. Neudecker, T. Wansbeek, Unbiased estimation of fourth-order matrix moments, Linear Algebra Appl. 160 (1992) 163–174.
[23] S. Konishi, Asymptotic expansions for the distribution of a function of the latent roots of the covariance matrix, Ann. Inst. Statist. Math. 29 (1977) 89–96.
[24] S. Konishi, Asymptotic expansions for the distribution of statistics based on a correlation matrix, Canad. J. Statist 6 (1978) 49–56.
[25] S. Konishi, Asymptotic expansions for the distributions of statistics based on the sample correlation matrix in principal component analysis, Hiroshima Math. J. 9 (1979) 647–700.
[26] R. Lugannini, S. Rice, Saddlepoint approximation for the distribution of the sum of independent random variables, Advan. Appl. Probab. 12 (1980) 475–490.
[27] E.C. MacRae, Matrix derivatives with an application to an adaptive linear decision problem, Ann. Statist. 2 (1974) 337–346.
[28] J.R. Magnus, H. Neudecker, The commutation matrix: some properties and applications, Ann. Statist. 7 (1979) 381–394.
[29] J.R. Magnus, H. Neudecker, H. Matrix, Differential Calculus with Applications in Statistics and Econometrics, revised ed, Wiley, Chichester, 1999.
[30] K.V. Mardia, Measures of multivariate skewness and kurtosis with applications, Biometrika 57 (1970) 519–530.
[31] P. McCullagh, Tensor Methods in Statistics, Chapman & Hall, London, 1987.
[32] G.M. Morant, A first study of the Tibetan skull, Biometrika 14 (1923) 193–260.
[33] B.F. Qaqish, A family of binary distributions for simulating correlated binary variables with specified marginal means and correlations, Biometrika 90 (2003) 455–463.
[34] N. Reid, Saddlepoint methods and statistical inference, Statist. Sci. 3 (1988) 213–238.

[35] T. Sugiyama, H. Tong, On a statistic useful in dimensionality reduction in multivariable linear stochastic system, Comm. Statist.—Theory and Methods A5 (1976) 711–721.
[36] M. Tingley, C. Field, Small-sample confidence intervals, J. Amer. Statist. Assoc. 85 (1990) 427–434.
[37] D.L. Wallace, Asymptotic approximations to distributions, Ann. Math. Statist. 29 (1958) 635–654.
[38] S. Wang, General saddlepoint approximations in the bootstrap, Statist. Probab. Lett. 13 (1992) 61–66.
[39] C.M. Waternaux, Asymptotic distribution of the sample roots for a nonnormal population, Biometrika 63 (1976) 639–645.