Uniform Saddlepoint Approximations for Ratios of Quadratic Forms*

Ronald W. Butler[†]
Department of Statistics
Colorado State University
Fort Collins, CO 80523
walrus@stat.colostate.edu

Marc S. Paolella
Swiss Banking Institute
University of Zurich
CH - 8032 Zurich, Switzerland
paolella@isb.unizh.ch

February 12, 2004

Abstract

Ratios of quadratic forms in correlated normal variables which introduce noncentrality into the quadratic forms are considered. The denominator is assumed to be positive (w.p.1). Various serial correlation estimates such as least squares, Yule-Walker, and Burg as well as Durbin-Watson statistics provide important examples of such ratios. The cumulative distribution function (cdf) and density for such ratios admit saddlepoint approximations. These approximations are shown to preserve uniformity of relative error over the entire range of support. Furthermore, explicit values for the limiting relative errors at the extreme edges of support are derived.

1 Introduction

Consider the ratio of quadratic forms

$$R = \frac{\epsilon' \mathbf{A} \epsilon}{\epsilon' \mathbf{B} \epsilon} \tag{1}$$

where, without loss in generality, $\mathbf{A}$ and $\mathbf{B}$ are assumed to be $n \times n$ symmetric. Let $\epsilon \sim \mathbf{N}(\mu, \mathbf{I}_n)$ and suppose $\mathbf{B}$ is also positive semidefinite thereby assuring that the denominator is positive with probability one. There is no loss in generality in having the covariance of ϵ as the identity. This is because, if the distribution of ϵ were $\mathbf{N}(\mu, \boldsymbol{\Sigma})$, then (1) describes the model with $\boldsymbol{\Sigma}^{\frac{1}{2}} \mathbf{A} \boldsymbol{\Sigma}^{\frac{1}{2}}$ and $\boldsymbol{\Sigma}^{\frac{1}{2}} \mathbf{B} \boldsymbol{\Sigma}^{\frac{1}{2}}$ replacing $\mathbf{A}$ and $\mathbf{B}$ respectively, and $\boldsymbol{\Sigma}^{-\frac{1}{2}} \mu$ replacing

[*]Abbreviated Title: Saddlepoint Uniformity

AMS 1991 subject classifications. Primary 62E17, 62M10; secondary 62F11

Key words and phrases. Ratios of quadratic forms, saddlepoint approximations, serial correlations.

[†]Research supported by NSF grant DMS-9970785.

μ in the distributional assumption on ϵ. Thus model (1) incorporates all dependence among the components of ϵ as well as noncentrality that occurs when $\mu \neq \mathbf{0}$.

Various sorts of saddlepoint approximations for the distribution (cdf) and density of R have been proposed beginning with the seminal work on serial correlations in Daniels (1956). Further marginal distributional approximations are given in McGregor (1960), Phillips (1978), Jensen (1988), Wang (1992), Lieberman (1994a,b), Butler and Paolella (1998a), and Marsh (1998). Joint distributional approximations for the set of serial correlations comprising the correlogram were initiated by Daniels (1956) and continued in Durbin (1980) and Butler and Paolella (1998b).

The main contributions of the current paper are in establishing the uniformity of relative errors for the saddlepoint cdf and density approximations in the right tail when used with univariate ratios R in a class $\mathcal{C}_{\mathcal{R}}$. This class encompasses all the examples in the aforementioned papers. Expressions for the limiting relative error are given as the right edge of support for R is approached. These expressions are explicit in the more elementary settings in which a certain defining eigenvalue is simple and mostly implicitly defined when multiple. The noncentral beta distribution provides an important example in which the limiting error is explicit but the defining eigenvalue is multiple.

The left tail of R is dealt with by changing $\mathbf{A}$ to $-\mathbf{A}$ thus switching the left tail of R to the right tail of $-R$. The results for the right tail of $-R$ can now provide similar uniformity results for the left tail when applicable. If $-R$ is a member of class $\mathcal{C}_{\mathcal{R}}$ then we say that R is in $\mathcal{C}_{\mathcal{L}}$. However, for the most part, the paper concentrates on class $\mathcal{C}_{\mathcal{R}}$.

The class of ratios $\mathcal{C}_{\mathcal{R}}$ is characterized technically in terms of a sequence of largest eigenvalues. Let (l, r) be the support of R, with r perhaps infinite, and define $\lambda_n(r)$ as the largest eigenvalue of $\mathbf{A} - r\mathbf{B}$ for $r \in (\mathsf{l}, \mathsf{r})$. The class $\mathcal{C}_{\mathcal{R}}$ is characterized as those ratios R whose matrices $\mathbf{A}$ and $\mathbf{B}$ admit the limit

$$0 = \lim_{r \to \mathsf{r}} \lambda_n(r). \tag{2}$$

Class $\mathcal{C}_{\mathcal{R}}$ contains the subset $\mathcal{B}$ that consists of all ratios with bounded support (l, r); this is a property of R that is guaranteed when $\mathbf{B} > \mathbf{0}$, or positive definite. Such ratios include the

Durbin-Watson statistics as well as the Yule-Walker and Burg estimators of serial correlation with arbitrary lag computed from least squares residuals. If $\mathbf{B} \geq \mathbf{0}$ has at least one zero eigenvalue, then r may be finite or infinite. The portion of $\mathcal{C}_{\mathcal{R}}$ with $\mathsf{r} = \infty$ includes least squares estimators of serial correlation in various sorts of models with arbitrary lag and computed from residuals with trend or covariates removed. Such models include those with autoregressive lag in the dependent variable and those with lag in the additive noise.

Large *sample space* asymptotics have not been previously considered for the class $\mathcal{B}$. The only previous consideration for a member of the class $\mathcal{C}_{\mathcal{R}} - \mathcal{B}$ is in Jensen (1988, 1995 §9.4). He obtained results in agreement with those below for the least squares estimator of lag one serial correlation, when the time series is a mean zero $AR(1)$ model.

The class $\mathcal{C}_{\mathcal{R}}$ excludes F-statistic and Satterthwaite-type ratios as have been considered in Butler and Paolella (2002). In this work $\lambda_n(r) > 0$ does not depend on r. For this setting, saddlepoint uniformity is also maintained, however, a different asymptotic saddlepoint behavior results from these different assumptions.

Some alternative large *sample size* asymptotics for the lag one least squares estimator, showing that the error is $O(n^{-1})$ and $O(n^{-3/2})$ on compact sets, are given in Lieberman (1994b) and Jensen (1995 §9.4) respectively when $\mu = \mathbf{0}$.

2 Saddlepoint Approximations

2.1 Distribution Theory

The cdf for R in the most general setting with noncentrality is

$$\begin{aligned} \Pr(R \leq r) &= \Pr\left(\frac{\epsilon'\mathbf{A}\epsilon}{\epsilon'\mathbf{B}\epsilon} \leq r\right) = \Pr\left(\epsilon'(\mathbf{A} - r\mathbf{B})\,\epsilon \leq 0\right) \\ &= \Pr(X_r \leq 0) \end{aligned} \tag{3}$$

where X_r has been so defined. Assume the spectral decomposition

$$\mathbf{A} - r\mathbf{B} = \mathbf{P}_r'\mathbf{\Lambda}_r\mathbf{P}_r \tag{4}$$

where $\mathbf{P}_r$ is orthogonal and $\mathbf{\Lambda}_r = \text{diag}(\lambda_1, \ldots, \lambda_n)$, with

$$\lambda_1 = \lambda_1(r) \leq \cdots \leq \lambda_n = \lambda_n(r)$$

consisting of the ordered eigenvalues of (4). Whenever convenient, we suppress the dependence of the various quantities on r. The distribution of X_r is therefore

$$X_r = \sum_{i=1}^{n} \lambda_i \chi^2\left(1, \nu_i^2\right), \tag{5}$$

where $\left\{\nu_i^2\right\}$ are determined as $(\nu_1, ..., \nu_n)' = \nu_r = \mathbf{P}_r \mu$ and represent the noncentrality parameters of the independent noncentral χ_1^2 variables specified in (5). The ordered values of $\{\lambda_i\}$ are in 1-1 correspondence with the components of ν_r specified through the particular choice of $\mathbf{P}_r$. Notationally we use χ_k^2 for the central chi-square instead of $\chi^2(k, 0)$.

Before proceeding with the development of a saddlepoint approximation for the distribution of R, we must first characterize the support of R, its relationship to the eigenvalues $\lambda_1(r)$ and $\lambda_n(r)$ and the convergence strip for the moment generating function of X_r.

Lemma 1 *All the eigenvalues of $\mathbf{\Lambda}_r$ are strictly decreasing in r for $\mathbf{B} > \mathbf{0}$ and decreasing when $\mathbf{B} \geq \mathbf{0}$.*

Proof. This is a direct consequence of theorem 9 in Magnus and Neudecker (1988, §10.9). ■

Lemma 2 *The distribution of R is degenerate at a single point if and only if $\mathbf{A} = c\mathbf{B}$ for some scalar constant c.*

A description for the support of R requires consideration of the various cases involved which depend on eigenvalue decompositions for $\mathbf{A}$ and $\mathbf{B}$. Suppose that $\mathbf{B}$ has $p \geq 0$ zero eigenvalues and let $\mathbf{O}_\mathbf{B}'$ be the orthogonal matrix of eigenvectors for $\mathbf{B}$ such that

$$\mathbf{O_B B O_B'} = \begin{pmatrix} \mathbf{\Lambda_B} & \mathbf{0} \\ \mathbf{0} & \mathbf{0}_{p \times p} \end{pmatrix}.$$

Denote

$$\mathbf{O_B A O_B'} = \begin{pmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} \\ \mathbf{C}_{21} & \mathbf{C}_{22} \end{pmatrix}$$

where $\mathbf{C}_{11}$ is $(n-p)\times(n-p)$ and $\mathbf{C}_{22}$ is $p\times p$. Let $N(\mathbf{C}_{12})$ denote the null space in $\Re^p$ for matrix $\mathbf{C}_{12}$.

Lemma 3 *The support of R is specified in the following set of cases.*

1. Suppose $\mathbf{B}>0$, hence $p=0$, and $\mathbf{A}$ has rank of at least one. Then the support of R is the finite interval (l,r) with l and r as the smallest and largest eigenvalues of $\mathbf{B}^{-1}\mathbf{A}$.

2. If $p\geq 1$, so $\mathbf{B}$ has at least one zero eigenvalue, then the right edge r is given as follows.

   (a) If $\mathbf{C}_{22}$ has a positive eigenvalue, then $\mathsf{r}=\infty$.

   (b) If $\mathbf{C}_{22}<\mathbf{0}$ then $\mathsf{r}<\infty$ and r is the largest eigenvalue of

   $$\mathbf{\Lambda}_{\mathbf{B}}^{-1}(\mathbf{C}_{11}-\mathbf{C}_{12}\mathbf{C}_{22}^{-1}\mathbf{C}_{21}). \tag{6}$$

   (c) If $\mathbf{C}_{22}\leq\mathbf{0}$ and $\mathbf{C}_{22}$ has at least one zero eigenvalue, then $\mathsf{r}=\infty$ if $N(\mathbf{C}_{22})\nsubseteq N(\mathbf{C}_{12})$; otherwise $\mathsf{r}<\infty$ and is the largest eigenvalue of

   $$\mathbf{\Lambda}_{\mathbf{B}}^{-1}(\mathbf{C}_{11}-\mathbf{C}_{12}\mathbf{O}_{\mathbf{C}1}\mathbf{\Lambda}_{\mathbf{C}}^{-1}\mathbf{O}_{\mathbf{C}1}'\mathbf{C}_{21}).$$

   Here, $\mathbf{O}_{\mathbf{C}}'=(\mathbf{O}_{\mathbf{C}1},\mathbf{O}_{\mathbf{C}2})$ consists of the eigenvectors of $\mathbf{C}_{22}$,

   $$\mathbf{O}_{\mathbf{C}}\mathbf{C}_{22}\mathbf{O}_{\mathbf{C}}'=\begin{pmatrix}\mathbf{\Lambda}_{\mathbf{C}} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}_{m\times m}\end{pmatrix},$$

   $\mathbf{\Lambda}_{\mathbf{C}}<0$, m is the multiplicity of the zero eigenvalue, and $\mathbf{O}_{\mathbf{C}1}$ are eigenvectors with nonzero eigenvalues.

Proof. For case 1, set $\mathbf{z}_1=\mathbf{B}^{1/2}\epsilon$ so that

$$R=\frac{\epsilon'\mathbf{A}\epsilon}{\epsilon'\mathbf{B}\epsilon}=\frac{\mathbf{z}_1'\mathbf{B}^{-1/2}\mathbf{A}\mathbf{B}^{-1/2}\mathbf{z}_1}{\mathbf{z}_1'\mathbf{z}_1}. \tag{7}$$

The ratio R in (7) is bounded between the smallest and largest eigenvalues of $\mathbf{B}^{-1/2}\mathbf{A}\mathbf{B}^{-1/2}$ or equivalently $\mathbf{B}^{-1}\mathbf{A}$. Proof for case 2 is relegated to Appendix 8.1. ■

Not all of the settings of Lemma 3 pertain to the class $\mathcal{C}_{\mathcal{R}}$.

Lemma 4 *When considering the right tail, matrices* $\mathbf{A}$ *and* $\mathbf{B}$ *admit a ratio* R *in class* $\mathcal{C}_{\mathcal{R}}$ *only for cases 1 or 2(b) or 2(c). When considering both the left and right tails, then class* $\mathcal{C}_{\mathcal{R}} \cap \mathcal{C}_{\mathcal{L}}$ *encompasses case 1 or the special setting of case 2 in which* $\mathbf{C}_{22} = \mathbf{0}$.

Proof. See Appendix 8.1. ■

Lemmas 3 and 4 are most easily understood by using some simple examples. Consider an $F_{1,1}$ distribution for R. Then

$$\mathbf{A} - r\mathbf{B} = \begin{pmatrix} 1 & 0 \\ 0 & -r \end{pmatrix}$$

and $\lambda_1(r) = -r$ with $\lambda_2(r) \equiv 1$. Clearly this is not in the class $\mathcal{C}_{\mathcal{R}}$ nor in $\mathcal{C}_{\mathcal{L}}$.

Next consider $n = 2$ and the least squares estimate of a lag 1 serial correlation in the simplest setting with $R = \epsilon_1\epsilon_2/\epsilon_1^2 = \epsilon_2/\epsilon_1$. Note this has the Cauchy distribution when $\mu = 0$ and the support is $(-\infty, \infty)$. To see that this ratio is in the classes $\mathcal{C}_{\mathcal{R}}$ and $\mathcal{C}_{\mathcal{L}}$, note that

$$\mathbf{A} - r\mathbf{B} = \begin{pmatrix} -r & 1/2 \\ 1/2 & 0 \end{pmatrix} \tag{8}$$

and the limiting eigenvalues are

$$\lim_{r\to \mathsf{l}=-\infty} (-r - \sqrt{r^2+1})/2 = 0 = \lim_{r\to \mathsf{r}=\infty} (-r + \sqrt{r^2+1})/2. \tag{9}$$

The same results hold more generally with least squares estimates of serial correlation from regression residuals.

Lemma 5 *Suppose* R *has a non-degenerate distribution in class* $\mathcal{C}_{\mathcal{R}}$ *as described in Lemma 4,* $\mathbf{B} \geq 0$, *and* $\mathbf{A}$ *has rank of at least one. The upper range of support* $\mathsf{r} \leq \infty$ *for* R, *as given in cases 1,2(b), and 2(c) of Lemma 3, solves*

$$\lambda_n(\mathsf{r}) = 0. \tag{10}$$

If r *is an interior point for the support of* R, *then the mgf of* X_r *is*

$$M_{X_r}(s) = \left(\prod_{i=1}^{n} (1 - 2s\lambda_i)^{-1/2}\right) \exp\left\{ s \sum_{i=1}^{n} \frac{\lambda_i \nu_i^2}{1 - 2s\lambda_i} \right\} \tag{11}$$

and convergent on the neighborhood of zero given as

$$\frac{1}{2\lambda_1(r)} < s < \frac{1}{2\lambda_n(r)}. \tag{12}$$

Proof. Value r is interior to the support of R if and only if $\lambda_1(r) < 0 < \lambda_n(r)$. Using the continuity of $\lambda_n(\cdot)$ and the restriction to class $\mathcal{C}_{\mathcal{R}}$, then (10) must hold. ■

2.2 Cdf Saddlepoint Approximation

The saddlepoint approximation is based on the cumulant generating function (cgf) for X_r given as $K_{X_r}(s) = \ln M_{X_r}(s)$. The saddlepoint $\hat{s}$ is the unique root of

$$0 = K'_{X_r}(\hat{s}) = \sum_{i=1}^{n}\left(\frac{\lambda_i}{1-2\hat{s}\lambda_i} + \frac{\lambda_i\nu_i^2}{(1-2\hat{s}\lambda_i)^2}\right) \tag{13}$$

in the range (12). The Lugannani and Rice (1980) approximation to first-order is

$$\widehat{\Pr}_1(R \le r) = \begin{cases} \Phi(\hat{w}) + \phi(\hat{w})\left\{\hat{w}^{-1} - \hat{u}^{-1}\right\} & \text{if} \quad 0 \ne \mathcal{E}[X_r] \\ \frac{1}{2} + \frac{K'''_{X_r}(0)}{6\sqrt{2\pi}K''_{X_r}(0)^{3/2}} & \text{if} \quad 0 = \mathcal{E}[X_r] \end{cases} \tag{14}$$

where $\Phi(\cdot)$ and $\phi(\cdot)$ denote the distribution and density function of a standard normal random variable, respectively, and

$$\hat{w} = \operatorname{sgn}(\hat{s})\sqrt{-2K_{X_r}(\hat{s})} \quad \hat{u} = \hat{s}\sqrt{K''_{X_r}(\hat{s})}. \tag{15}$$

Higher derivatives of K_{X_r} are given as

$$K_X^{(j)}(\hat{s}) = 2^{j-1}(j-1)!\sum_{i=1}^{n}\lambda_i^j(1-2\hat{s}\lambda_i)^{-j}\left(1 + \frac{j\nu_i^2}{1-2\hat{s}\lambda_i}\right). \tag{16}$$

Third and fourth derivatives allow computation of the second-order approximation which includes further terms in the saddlepoint expansion. This has been given in Daniels (1987) as

$$\widehat{\Pr}_2(R \le r) = \widehat{\Pr}_1(R \le r) - \phi(\hat{w})\left\{\hat{u}^{-1}\left(\frac{\hat{\kappa}_4}{8} - \frac{5}{24}\hat{\kappa}_3^2\right) - \hat{u}^{-3} - \frac{\hat{\kappa}_3}{2\hat{u}^2} + \hat{w}^{-3}\right\} \tag{17}$$

for r values such that $0 \ne \mathcal{E}[X_r]$ where $\hat{\kappa}_j = K_{X_r}^{(j)}(\hat{s})/K''_{X_r}(\hat{s})^{j/2}$.

2.3 Density Saddlepoint Approximation

The saddlepoint density approximation for $f_R(r)$, the density of R at r, is derived in the Appendix as

$$\hat{f}_R(r) = \frac{J_r(\hat{s})}{\sqrt{2\pi K''_{X_r}(\hat{s})}} M_{X_r}(\hat{s}), \tag{18}$$

where $\hat{s}$ is the same saddlepoint used in the cdf approximation and which solves (13). Factor $J_r(\hat{s})$ is computed from

$$J_r(s) = \operatorname{tr}(\mathbf{I} - 2s\mathbf{\Lambda}_r)^{-1}\mathbf{H}_r + \nu_r'(\mathbf{I} - 2s\mathbf{\Lambda}_r)^{-1}\mathbf{H}_r(\mathbf{I} - 2s\mathbf{\Lambda}_r)^{-1}\nu_r \tag{19}$$

with

$$\mathbf{H}_r = \mathbf{P}_r\mathbf{B}\mathbf{P}_r'.$$

The second-order saddlepoint density in this context is

$$\hat{f}_{R2}(r) = \hat{f}_R(r)(1 + O) \tag{20}$$

where

$$O = \left(\frac{\hat{\kappa}_4}{8} - \frac{5}{24}\hat{\kappa}_3^2\right) + \frac{J_r'(\hat{s})\hat{\kappa}_3}{2J_r(\hat{s})\sqrt{K''_{X_r}(\hat{s})}} - \frac{J_r''(\hat{s})}{2J_r(\hat{s})K''_{X_r}(\hat{s})}. \tag{21}$$

Example 6 *For matrices* **A** *and* **B** *in which* $R \sim Beta\left(\frac{m}{2}, \frac{n-m}{2}\right)$, *the saddlepoint density in (18) is*

$$\hat{f}_R(r) = \frac{B\left(\frac{m}{2}, \frac{n-m}{2}\right)}{\hat{B}\left(\frac{m}{2}, \frac{n-m}{2}\right)} f_R(r)$$

where $\hat{B}$ *is Stirling's approximation for the Beta function* B.

3 Uniformity of the Approximations in r

The relative errors of the Lugannani and Rice approximation in (14) and the density approximation in (18) are shown to be uniform over $[0, \mathsf{r})$ when in class $\mathcal{C}_{\mathcal{R}}$. These results follow as a consequence of deriving their finite limiting ratios as $r \to \mathsf{r}$. The limiting ratios are derived in Theorems 10, 11, 16, and 17 below. Our approach for computing these limiting ratios follows that also used in Jensen (1988, 1995 §9.4) and generalizes these results to accommodate both noncentrality and the special concerns of multiple eigenvalues.

The nature of these asymptotics is dependent on the multiplicity of eigenvalue $\lambda_n(\mathsf{r}) = 0$, denoted as m. As a simple eigenvalue with $m = 1$, the limiting ratios are derived in Theorems 10 and 11. This is a common setting encountered while dealing with serial correlations. With $m \geq 2$ however, the asymptotics are more difficult and such results are deferred to Theorems 16 and 17. Examples of the multiple eigenvalue setting are also common and include least squares estimates and Yule-Walker estimates for lag l serial correlation with $l \geq 2$. One important multiple eigenvalue example is the noncentral beta distribution discussed in §4.

The Case 2(a) setting is not in $\mathcal{C}_{\mathcal{R}}$, however the relative error can still be shown to be uniform over $[0, \infty)$. Section 8.5 of the Appendix discusses this setting and shows that the asymptotics are those found with Satterthwaite F-type ratios as detailed in Butler and Paolella (2002).

The following result is used extensively in computing these limiting ratios.

Lemma 7 *Suppose* $Z_0 \sim \chi^2_{n_0}$ *independently of* $Z_1 \sim \chi^2(n_1, 2\omega)$. *If* $a_0, a_1 > 0$, *then the density of* $a_1 Z_1 - a_0 Z_0$ *at zero is*

$$f_{a_1 Z_1 - a_0 Z_0}(0) = \frac{a_0^{\frac{1}{2}n_1 - 1} a_1^{\frac{1}{2}n_0 - 1} 2^{-1} e^{-\omega}}{(a_0 + a_1)^{\bar{n}-1} B\left(\frac{n_1}{2}, \frac{n_0}{2}\right)(\bar{n} - 1)} \, {}_1F_1\left(\bar{n} - 1; \frac{n_1}{2}; \frac{\omega a_0}{a_0 + a_1}\right)$$

where $\bar{n} = (n_0 + n_1)/2$.

Proof. The density is the convolution

$$\int_0^\infty \frac{1}{a_0 a_1} f_{Z_0}\left(\frac{t}{a_0}\right) f_{Z_1}\left(\frac{t}{a_1}\right) dt$$

computed by expressing the noncentral χ^2 as a weighted sum of central χ^2 densities as in Johnson and Kotz (1970, p. 132). Tedious calculation lead to

$$\frac{a_0^{\frac{1}{2}n_1 - 1} a_1^{\frac{1}{2}n_0 - 1} 2^{-1} e^{-\omega}}{(a_0 + a_1)^{\bar{n}-1} \Gamma\left(\frac{n_0}{2}\right)} \left\{ \sum_{k=0}^{\infty} \frac{\Gamma(\bar{n} + k - 1)}{\Gamma\left(\frac{n_1}{2} + k\right) k!} \left(\frac{\omega a_0}{a_0 + a_1}\right)^k \right\}.$$

The lemma results upon recognizing that the summation in curly braces is the Taylor expansion for the confluent hypergeometric function ${}_1F_1$ in Abramowitz and Stegun (1972, 13.1.2). ■

3.1 Simple Eigenvalue $\lambda_n(\mathsf{r}) = 0$

Suppose $\mathsf{r} < \infty$, under the circumstances of Cases 1, 2(b), or 2(c). We assume here that $\mathbf{A} - \mathsf{r}\mathbf{B}$ has a simple zero eigenvalue with multiplicity $m = 1$. For general $\mathbf{A}$ and $\mathbf{B}$, this multiplicity is often difficult to anticipate. Define

$$\nu_0 = \nu_n(\mathsf{r}) := \mathbf{p}_n(\mathsf{r})'\mu, \tag{22}$$

where $\mathbf{p}_n(\mathsf{r})$ is the eigenvector associated with the zero eigenvalue of $\mathbf{A} - \mathsf{r}\mathbf{B}$.

The situation with $\mathsf{r} = \infty$ is more complicated.

Lemma 8 *Suppose Case 2(c) with* $\mathsf{r} = \infty$. *Then* m, *the multiplicity of zero eigenvalues in* $\{\lambda_i(\infty)\}$ *is the number of zero eigenvalues for* $\mathbf{C}_{22}$. *If* $m = 1$ *then*

$$\nu_0 = \nu_n(\infty) := \mathbf{o}_n'\mathbf{O}_{\mathbf{B}2}'\mu \tag{23}$$

where $\mathbf{o}_n$ *is the* $p\times 1$ *eigenvector associated with the zero eigenvalue of* $\mathbf{C}_{22}$, $\mathbf{O}_{\mathbf{B}}' = (\mathbf{O}_{\mathbf{B}1}, \mathbf{O}_{\mathbf{B}2})$ *and* $\mathbf{O}_{\mathbf{B}2}$ *is* $n \times p$ *and the orthonormal basis for the null space of* $\mathbf{B}$ *used to determine* $\mathbf{C}_{22}$.

Proof. See Appendix 8.1. ■

The $AR(1)$ example in (8) with $n = 2$ provides a simple example. Here, $\mathbf{C}_{22}$ is the scalar 0 so that $\mathbf{o}_n' = 1$, and $\mathbf{O}_{\mathbf{B}2}' = (0, 1)$; hence $\nu_0 = \mu_2$.

Lemma 9 *Suppose the conditions of Lemma 5 and let* $m = 1$. *Then, as* $r \to \mathsf{r} \leq \infty$,

$$\epsilon = \lambda_n(r) \to \lambda_n(\mathsf{r}) = 0$$

and

$$\hat{s} = \frac{t_0}{\epsilon} + O(1) \to \infty, \tag{24}$$

where

$$t_0 = \frac{1}{4n}\left\{2n - 1 + \nu_0^2 - \sqrt{\left(\nu_0^2 + 2n - 1\right)^2 - (2n-1)^2 + 1}\right\} \tag{25}$$

and ν_0 *is defined in (22) or (23). In addition,*

$$\hat{u} \to u_0 = \sqrt{\frac{n-1}{2} + \frac{2t_0^2}{(1-2t_0)^2} + \frac{4\nu_0^2 t_0^2}{(1-2t_0)^3}}. \tag{26}$$

Proof. The largest eigenvalue $\lambda_n(r)$ is continuous at $r = \mathsf{r}$ so that $\epsilon = \lambda_n(r) \to 0$ as $r \to \mathsf{r}$. Define $\hat{t} = \epsilon\hat{s}$ and replace $\hat{s}$ with $\hat{t}$ in saddlepoint equation (13) to get

$$\begin{aligned} 0 &= \sum_{i=1}^{n} \left\{ \frac{\lambda_i}{1 - 2\hat{t}\lambda_i/\epsilon} + \frac{\lambda_i \nu_i^2}{\left(1 - 2\hat{t}\lambda_i/\epsilon\right)^2} \right\} \\ &= \epsilon \left[\sum_{i=1}^{n-1} \left\{ \frac{\lambda_i}{\epsilon - 2\hat{t}\lambda_i} + O(\epsilon) \right\} + \left\{ \frac{1}{1 - 2\hat{t}} + \frac{\nu_n^2}{\left(1 - 2\hat{t}\right)^2} \right\} \right]. \end{aligned} \tag{27}$$

Remove the factor ϵ and take the limit of this equation as $r \to \mathsf{r}$ to get

$$0 = \frac{n-1}{-2t_0} + \frac{1}{1 - 2t_0} + \frac{\nu_0^2}{\left(1 - 2t_0\right)^2}, \tag{28}$$

where $\lim_{r\to\mathsf{r}} \nu_n = \nu_0$ and t_0 is defined to be the appropriate root of (28). There are two roots to this quadratic equation and the smaller root in (25) is the correct limiting value of $\hat{t}$. The reason for this is that the larger root exceeds $1/2$ and $1/2$ is an upper bound on the range of values for $\hat{t}$ as deduced from (12). To show that $\hat{t} \to t_0$ as given in (25), note that the estimating equation in (27), without factor ϵ, converges uniformly to (28) in a compact neighborhood of t_0; therefore $\hat{t}$ must converge to t_0 by a uniform continuity argument. The expansion for $\hat{s}$ in ϵ in (24) now follows. Similar arguments provide the expansion

$$\begin{aligned} K''_{X_r}(\hat{s}) &\sim 2 \sum_{i=1}^{n-1} \left\{ \frac{\lambda_i^2}{\left(1 - 2t_0\lambda_i/\epsilon\right)^2} + \frac{2\lambda_i^2\nu_i^2}{\left(1 - 2t_0\lambda_i/\epsilon\right)^3} \right\} + 2\left\{ \frac{\epsilon^2}{\left(1 - 2t_0\right)^2} + \frac{2\epsilon^2\nu_n^2}{\left(1 - 2t_0\right)^3} \right\} \\ &\sim \epsilon^2 \left\{ \frac{n-1}{2t_0^2} + \frac{2}{\left(1 - 2t_0\right)^2} + \frac{4\nu_0^2}{\left(1 - 2t_0\right)^3} \right\}. \end{aligned}$$

Thus $\hat{\sigma} = \sqrt{K''_{X_r}(\hat{s})} = O(\epsilon)$ and $\hat{s} = O\left(\epsilon^{-1}\right)$ and the product $\hat{u}$ converges to u_0 as given in (26). ■

Theorem 10 *Suppose $n \geq 2$, R has a non-degenerate distribution in $\mathcal{C}_{\mathcal{R}}$, $\mathbf{B} > 0$, and $\mathbf{A}$ has rank of at least one. If $m = 1$, then the limiting ratio of the true tail probability to its first order Lugannani and Rice approximation in (14) is*

$$\lim_{r\to\mathsf{r}} \frac{\Pr(R > r)}{\widehat{\Pr}_1(R > r)} = \frac{\sqrt{2\pi\left(1 - 2t_0\right)}\left(2t_0\right)^{\frac{n-1}{2}} u_0 e^{-\eta_2}}{B\left(\frac{1}{2}, \frac{n+1}{2}\right)\frac{n}{2}} \, {}_1F_1\left(\frac{n}{2}; \frac{1}{2}; \frac{\nu_0^2}{2}\right). \tag{29}$$

where parameters t_0, u_0, ν_0, and η_2 are specified in (25), (26), (22,23), and (36) respectively. All of these parameters are determined by ν_0 so the right side of (29) is a function of ν_0 alone.

Proof. The general method of proof entails writing the tail probability ratio for finite r as an inversion integral. The limiting ratio is then determined by passing the inversion integral to its limit as $r \to \mathsf{r}$.

The right tail probability of R is determined in terms of X_r, using the inversion formula as

$$\begin{aligned}
\Pr\left(X_r > 0\right) &= \frac{1}{2\pi i}\int_{\hat{s}-i\infty}^{\hat{s}+i\infty} z^{-1} M_{X_r}\left(z\right) dz \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty} (\hat{s}+it)^{-1} M_{X_r}(\hat{s}+it) dt \\
&= \frac{M_{X_r}(\hat{s})}{\hat{s}\hat{\sigma}}\frac{1}{2\pi}\int_{-\infty}^{\infty}\left(1+\frac{it}{\hat{s}\hat{\sigma}}\right)^{-1}\frac{M_{X_r}(\hat{s}+it/\hat{\sigma})}{M_{X_r}(\hat{s})} dt \qquad (30)
\end{aligned}$$

where the last integral results from a scale change to the variable of integration with $\hat{\sigma} = \sqrt{K''_{X_r}(\hat{s})}$.

The asymptotic behavior of the first order saddlepoint approximation in (14) is related to the leading term in (30) according to

$$\widehat{\Pr}_1\left(X_r > 0\right) \sim \phi\left(\hat{w}\right)/\hat{u}^{-1} = \frac{M_{X_r}(\hat{s})}{\hat{s}\hat{\sigma}\sqrt{2\pi}} \qquad (31)$$

as $r \to \mathsf{r}$ if it can be shown that $\hat{w} \to \infty$ under such limits. From Lemma 9, $\hat{s} \to \infty$ and $\lambda_1(r) < 0$ for sufficiently large r so that the term $\{1 - 2\hat{s}\lambda_1(r)\}^{-1/2} \to 0$ in $M_{X_r}(\hat{s})$ as do some other terms; thus $\hat{w} = \sqrt{-2K_{X_r}(\hat{s})} \to \infty$.

Using (31) in (30) then

$$\lim_{r\to\mathsf{r}}\frac{\Pr\left(X_r > 0\right)}{\widehat{\Pr}_1\left(X_r > 0\right)} = \sqrt{2\pi}\lim_{r\to\mathsf{r}}\frac{1}{2\pi}\int_{-\infty}^{\infty}\left(1+\frac{it}{\hat{s}\hat{\sigma}}\right)^{-1}\frac{M_{X_r}(\hat{s}+it/\hat{\sigma})}{M_{X_r}(\hat{s})} dt. \qquad (32)$$

The limiting inversion is now determined by finding the pointwise limit of its integrand and applying the dominated convergence theorem. A dominating function that is integrable is provided in the Appendix.

The limiting values of the integrand in (32) are now determined. For the first term,

$$\left(1+\frac{it}{\hat{s}\hat{\sigma}}\right)^{-1} \sim \left(1+\frac{it}{u_0}\right)^{-1}.$$

The centrality factors of the M_{X_r}-ratio are

$$\prod_{i=1}^{n}\left\{\frac{1-2\left(\hat{s}+it/\hat{\sigma}\right)\lambda_i}{1-2\hat{s}\lambda_i}\right\}^{-\frac{1}{2}} = \prod_{i=1}^{n}\left\{1-\frac{2it\lambda_i}{\hat{\sigma}\left(1-2\hat{s}\lambda_i\right)}\right\}^{-\frac{1}{2}}. \qquad (33)$$

As $\epsilon \to 0$, $\hat{\sigma} = \hat{u}/\hat{s} \sim \epsilon u_0/t_0$ so that

$$\hat{\sigma}\left(1 - 2\hat{s}\lambda_i\right) \sim \left(\frac{\epsilon u_0}{t_0}\right)\left(1 - 2t_0\lambda_i/\epsilon\right) \tag{34}$$
$$\sim \begin{cases} -2u_0\lambda_i > 0 & \text{if} \quad i = 1, \ldots, n-1 \\ \left(\epsilon u_0/t_0\right)\left(1 - 2t_0\right) & \text{if} \quad i = n\,. \end{cases}$$

The limit of (33) is therefore

$$\left(1 + \frac{it}{u_0}\right)^{-\frac{n-1}{2}} \left(1 - 2it\eta_1\right)^{-\frac{1}{2}}$$

where

$$\eta_1 = \frac{t_0}{u_0\left(1 - 2t_0\right)}.$$

For the limit of the noncentral portions of the M_{X_f}-ratio, the exponent of the i^{th} term of $\exp\left(\cdot\right)$ is

$$\frac{\left(\hat{s} + it/\hat{\sigma}\right)\lambda_i\nu_i^2}{1 - 2\left(\hat{s} + it/\hat{\sigma}\right)\lambda_i} - \frac{\hat{s}\lambda_i\nu_i^2}{1 - 2\hat{s}\lambda_i} = \frac{\nu_i^2}{2}\left(\frac{1}{1 - 2\left(\hat{s} + it/\hat{\sigma}\right)\lambda_i} - \frac{1}{1 - 2\hat{s}\lambda_i}\right). \tag{35}$$

As $\epsilon \to 0$, then $\hat{s} \to \infty$ and $1 - 2\hat{s}\lambda_i \to +\infty$ since $\lambda_i\left(\mathsf{r}\right) < 0$ for $i = 1, \ldots, n-1$. Simple computations show the real portion of (35) is $O\left(\hat{s}^{-1}\right) \to 0$ and the imaginary portion is $O\left(\hat{\sigma}\right) \to 0$. Thus the limit of the M_{X_f}-ratio contribution is 1 for $i = 1, \ldots, n-1$. For $i = n$, the exponent from the right side of (35) is

$$\begin{aligned}
&\sim \frac{\nu_n^2}{2}\left(\frac{1}{1 - 2\epsilon\left(\frac{t_0}{\epsilon} + it\frac{t_0}{\epsilon u_0}\right)} - \frac{1}{1 - 2\frac{t_0}{\epsilon}\epsilon}\right) \\
&\sim \frac{\nu_0^2}{2}\left(\frac{1}{1 - 2t_0\left(1 + it/u_0\right)} - \frac{1}{1 - 2t_0}\right) \\
&= \frac{\eta_2}{1 - 2it\eta_1} - \eta_2
\end{aligned}$$

where

$$\eta_2 = \frac{\nu_0^2}{2\left(1 - 2t_0\right)}. \tag{36}$$

Using the dominating convergence theorem, the limiting ratio in (32) is

$$\sqrt{2\pi}\frac{1}{2\pi}\int_{-\infty}^{\infty}\left(1 + \frac{it}{u_0}\right)^{-\frac{n+1}{2}}\left(1 - 2it\eta_1\right)^{-\frac{1}{2}}\exp\left(\frac{\eta_2}{1 - 2it\eta_1} - \eta_2\right)dt. \tag{37}$$

In (37), the first exponential term in t is expanded in a Taylor series to give an integrand of

$$\left(1+\frac{it}{u_0}\right)^{-\frac{n+1}{2}} \quad \times \quad \sum_{k=0}^{\infty}(1-2it\eta_1)^{-\frac{1}{2}(1+2k)}\left(e^{-\eta_2}\frac{\eta_2^k}{k!}\right).$$

This may be recognized as the product of two characteristic functions representing the convolution of $-\frac{1}{2u_0}\chi^2_{n+1} = -a_0 Z_0$ and $\eta_1\chi^2(1, 2\eta_2) = a_1 Z_1$ variables as in Lemma 7. The inversion integral in (37) is the density of $a_1 Z_1 - a_0 Z_0$ at 0 as given by Lemma 7. Reducing this to an expression in t_0 and using $\eta_2(1+2u_0\eta_1)^{-1} = \nu_0^2/2$ leads to

$$f_{a_1 Z_1 - a_0 Z_0}(0) = \frac{\sqrt{1-2t_0}\,(2t_0)^{\frac{n-1}{2}}\, u_0 e^{-\eta_2}}{B\left(\frac{1}{2},\frac{n+1}{2}\right)\frac{n}{2}}\ {}_1F_1\left(\frac{n}{2};\frac{1}{2};\frac{\nu_0^2}{2}\right)$$

which gives (29). ■

In the central case with $\nu = \mathbf{0}$, the limiting ratio of tail probabilities in Theorem 10 works out to be

$$\hat{B}\left(\frac{1}{2},\frac{n-1}{2}\right)/B\left(\frac{1}{2},\frac{n-1}{2}\right), \tag{38}$$

where $\hat{B}$ is Stirling's approximation. This same limiting error was derived in Jensen (1995, §9.4) which considered the tail ratio for the distribution of the least squares estimate in a mean zero $AR(1)$ model. Jensen's (9.4.7) is this value when the difference in notation is accounted for (our n is his $n+1$).

As $\nu_0^2 \to \infty$, then the limiting ratio in Theorem 10 is

$$\hat{\Gamma}\left(\frac{1}{2},\frac{n-1}{2}\right)/\Gamma\left(\frac{1}{2},\frac{n-1}{2}\right)\{1+O(\nu_0^{-2})\}$$

where $\hat{\Gamma}$ is Stirling's approximation. This follows from the large argument asymptotics for ${}_1F_1$ as given in 13.1.4 of Abramowitz and Stegun (1972).

Theorem 11 *Under the conditions of Theorem 10, the first order saddlepoint density has the same relative limit given in (29). The intermediate limit*

$$\lim_{r\to \mathrm{r}}\frac{f_R(r)}{\hat{f}_R(r)} = \sqrt{2\pi}\,(1-2t_0)^{\frac{3}{2}}(2t_0)^{\frac{n-3}{2}}\,u_0 e^{-\eta_2}\left\{p_0\frac{{}_1F_1\left(\frac{n}{2};\frac{3}{2};\frac{\nu_0^2}{2}\right)}{B\left(\frac{3}{2},\frac{n-1}{2}\right)\frac{n}{2}} + (1-p_0)(1-2t_0)\frac{{}_1F_1\left(\frac{n+2}{2};\frac{5}{2};\frac{\nu_0^2}{2}\right)}{B\left(\frac{5}{2},\frac{n-1}{2}\right)\frac{n+2}{2}}\right\} \tag{39}$$

is derived where

$$p_0 = \frac{1-2t_0}{1-2t_0+\nu_0^2}.$$

This expression is shown to be analytically the same as the right side of (29).

Proof. Follow the approach used for the cdf so that

$$\begin{aligned} f_R(r) &= \mathcal{E}(W) f_{Y_r}(0) = \mathcal{E}(W) \frac{1}{2\pi} \int_{-\infty}^{\infty} M_{Y_r}(\hat{s}+it)\, dt \\ &= \mathcal{E}(W) \frac{M_{Y_r}(\hat{s})}{\hat{\sigma}} \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{M_{Y_r}(\hat{s}+it/\hat{\sigma})}{M_{Y_r}(\hat{s})} dt \\ &= \hat{f}_R(r) \sqrt{2\pi} \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{M_{X_r}(\hat{s}+it/\hat{\sigma}) J_r(\hat{s}+it/\hat{\sigma})}{M_{X_r}(\hat{s}) J_r(\hat{s})} dt \end{aligned}$$

from (67). Then,

$$\lim_{r\to \mathrm{r}} \frac{f_R(r)}{\hat{f}_R(r)} = \sqrt{2\pi} \lim_{r\to \mathrm{r}} \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{M_{X_r}(\hat{s}+it/\hat{\sigma}) J_r(\hat{s}+it/\hat{\sigma})}{M_{X_r}(\hat{s}) J_r(\hat{s})} dt. \tag{40}$$

The inversion is much the same as with the cdf except as concerns the different factor $J(\hat{s}+it/\hat{\sigma})/J(\hat{s})$ in place of $(1+it/\hat{s}\hat{\sigma})^{-1}$. The dominating convergence theorem can still be applied since, in the Appendix, the norm of this new factor is shown to be uniformly bounded in t for sufficiently large r. In addition it is also shown that

$$\frac{J_r(\hat{s}+it/\hat{\sigma})}{J_r(\hat{s})} \sim p_0 (1-2it\eta_1)^{-1} + (1-p_0)(1-2it\eta_1)^{-2}. \tag{41}$$

The limit of the M_{X_r}-ratio from Theorem 10 leads to a limiting inversion in (40) as

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} \{p_0 \phi_{a_1 Z_1}(t) \phi_{-a_0 Z_0}(t) + (1-p_0) \phi_{a_2 Z_2}(t) \phi_{-a_0 Z_0}(t)\}\, dt$$

where $\phi(\cdot)$ denotes a characteristic function, $a_0 Z_0 \sim \frac{1}{2u_0}\chi^2_{n-1}$, $a_1 Z_1 \sim \eta_1 \chi^2(3, 2\eta_2)$, and $a_2 Z_2 \sim \eta_1 \chi^2(5, 2\eta_2)$. The limiting ratio is therefore

$$\sqrt{2\pi} \{p_0 f_{a_1 Z_1 - a_0 Z_0}(0) + (1-p_0) f_{a_2 Z_2 - a_0 Z_0}(0)\},$$

which is (39) upon using Lemma 7.

The proof that (39) agrees analytically with the right side of (29) is deferred to the Appendix. ■

Relative errors for the second order approximations, with the cdf given in (17) and the density given in (20), are also uniform in the right tail. The additional correction terms for second order cdf remain bounded as $r \to \mathsf{r}$. This occurs because $\hat{\kappa}_3$, $\hat{\kappa}_4$ and $\hat{u}$ have finite nonzero limits and because the last term $\hat{w}^{-3} \to 0$. Some additional work is needed to deal with the last two terms in (21), the correction term of (20). From the asymptotics in §8.3, one can show that

$$J_r^{(j)}(\hat{s}) = O\left(\epsilon^j\right)$$

as $\hat{s} \to \infty$ or $\epsilon \to 0$. Using $\sqrt{K''_{X_r}(\hat{s})} = O(\epsilon)$, then these last two terms converge to nonzero values as $\epsilon \to 0$.

3.2 Multiple Eigenvalue $\lambda_n(\mathsf{r}) = 0$

In this setting, the asymptotics depend on the relative rates of convergence to zero for the multiple eigenvalues of $\mathbf{A} - r\mathbf{B}$ that approach 0 as $r \to \mathsf{r}$. If m denotes its multiplicity, then $m \geq 1$ by definition of r. The allowable values for m are $1 \leq m \leq n-1$ but not $m = n$. This latter value would make $\lim_{r \to \mathsf{r}}(\mathbf{A} - r\mathbf{B}) \equiv \mathbf{0}$ in which case the distribution of R approaches a degenerate distribution by Lemma 2. For unbounded ratios in Case 2(c), the value of m is the dimension of the null space for $\mathbf{C}_{22}$, whereas for ratios in $\mathcal{B}$ the value of m is less transparent.

We must first determine the relative rates at which the m largest eigenvalues of $\mathbf{A} - r\mathbf{B}$ vanish as $r \to \mathsf{r}$ in the two separate settings, $\mathsf{r} < \infty$ and $\mathsf{r} = \infty$. For the former setting, general formulae for these relative rates are given in the next lemma. When $\mathsf{r} = \infty$, the relative rates must be determined on a case by case basis.

Lemma 12 *Suppose $\mathsf{r} < \infty$ and 0 is an eigenvalue of multiplicity m for $\mathbf{A} - \mathsf{r}\mathbf{B}$. Let the columns of $n \times m$ matrix $\mathbf{U}_0$ be an orthonormal basis for the null space of $\mathbf{A} - \mathsf{r}\mathbf{B}$. Furthermore, denote the ordered eigenvalues of $\mathbf{U}_0'\mathbf{B}\mathbf{U}_0$ as*

$$0 \leq \tau_{n-m+1} \leq \cdots \leq \tau_n.$$

If $\tau_n > 0$, then the limiting relative rates of convergence to zero for the largest m eigenvalues *of* $\mathbf{A} - r\mathbf{B}$ *are*

$$\lim_{r \to \mathsf{r}} \frac{\lambda_i(r)}{\lambda_n(r)} = \frac{\tau_i}{\tau_n} = \omega_i \tag{42}$$

for $i = n - m + 1, \ldots, n$ *where* $0 \leq \omega_{n-m+1} \leq \cdots \leq \omega_n = 1$.

For the most common case in which $\mathbf{B} > \mathbf{0}$, then $\tau_{n-m+1} > 0$ so that $\omega_{n-m+1} > 0$.

Proof. The results follow by using theorem 13 in Magnus and Neudecker (1988, p. 167) that has been taken from Lancaster (1964). A nontrivial requirement for using this result is that all the eigenvalues of $\mathbf{A} - \mathsf{r}\mathbf{B}$ have only linear elementary divisors; this holds in our setting since the matrix is symmetric.

Theorem 13 states that the m derivatives $\{\partial\lambda_i(r)/\partial r|_{r=\mathsf{r}}\}$ assume values as the eigenvalues of $-\mathbf{U}_0'\mathbf{B}\mathbf{U}_0$. Using L'Hospital's rule,

$$\lim_{r \to \mathsf{r}} \frac{\lambda_i(r)}{\lambda_n(r)} = \lim_{r \to \mathsf{r}} \frac{\partial\lambda_i(r)/\partial r}{\partial\lambda_n(r)/\partial r}$$

which are ratios of $\{-\tau_i\}$. The larger the eigenvalue, the greater its rate of decrease to 0 has to be to catch up with the others as $r \to \mathsf{r}$. Therefore these rates are the ratios of τ_i-values specified in (42). ■

In case 2(c) for which $\mathsf{r} = \infty$, the results of Lancaster (1964) cannot be applied directly since the components of the matrix $\mathbf{A} - r\mathbf{B}$ are not analytic at $r = \infty$. A partial result may be obtained by reparametrizing

$$\mathbf{D}(\varepsilon) = (\mathbf{A} - r\mathbf{B})/r = \varepsilon\mathbf{A} - \mathbf{B} \tag{43}$$

and letting $\varepsilon \to 0$. If $\{\lambda_i(\varepsilon)\}$ denote the ordered eigenvalues of $\mathbf{A} - \varepsilon^{-1}\mathbf{B}$ then

$$\psi_i(\varepsilon) = \varepsilon\lambda_i(\varepsilon) \tag{44}$$

are the ordered eigenvalues of (43).

Lemma 13 *Consider case 2(c) in which* $\mathsf{r} = \infty$ *and assume that the zero eigenvalue has multiplicity* m. *Then* $\lambda_{n-m+1}(r), \ldots, \lambda_n(r)$ *are analytic at* $r = \infty$. *If* $\lambda_n'(\infty) > 0$ *then the relative rates of convergence are*

$$\omega_i = \frac{\partial\lambda_i(\varepsilon)/\partial\varepsilon|_{\varepsilon=0}}{\partial\lambda_n(\varepsilon)/\partial\varepsilon|_{\varepsilon=0}} = \frac{\partial^2\psi_i(\varepsilon)/\partial\varepsilon^2|_{\varepsilon=0}}{\partial^2\psi_n(\varepsilon)/\partial\varepsilon^2|_{\varepsilon=0}} \tag{45}$$

for $i = n - m + 1, \ldots, n$.

Proof. Whereas m eigenvalues among $\{\lambda_i(\varepsilon)\}$ converge to zero as $\varepsilon \to 0$, now $p \geq m$ eigenvalues of $\mathbf{D}(\varepsilon)$ converge to zero as $\varepsilon \to 0$. Lancaster's (1964) results may be applied to determine $\{\psi_i'(0)\}$ for the p zero roots of $\mathbf{D}(0)$. This leads to $\psi_i'(0) = 0$ for $i = n - m + 1, \ldots, n$. To see this, note that Lancaster's results specify these derivatives as the eigenvalues of

$$\mathbf{O}_{\mathbf{B}2}'\mathbf{A}\mathbf{O}_{\mathbf{B}2} = \mathbf{C}_{22}$$

or the elements of $\mathbf{\Lambda_C}$ for the $p - m$ smallest and zero for the m largest.

Lancaster's results also assure that $\{\psi_i(\varepsilon)\}$ are analytic functions at $\varepsilon = 0$. Since $\psi_i'(0) = 0$ for $i \geq n - m + 1$ then their Taylor expansions are

$$\psi_i(\varepsilon) = \varepsilon^2 \psi_i''(0)/2 + O(\varepsilon^3)$$

which yields Taylor expansion for $\lambda_i(\varepsilon)$ as

$$\lambda_i(\varepsilon) = \varepsilon \psi_i''(0)/2 + O(\varepsilon^2). \tag{46}$$

Hence $\{\lambda_i(\varepsilon)\}$ are analytic at $\varepsilon = 0$ and $\{\lambda_i(r)\}$ are analytic at $r = \infty$. The relative rates of convergence are now determined from the leading terms in (46) or alternatively by using L'Hospital's rule applied to (44). Both of these arguments lead to the expressions in (45).
■

The limiting noncentrality parameters $\{\nu_{0i} : i = n - m + 1, \ldots, n\}$ are more difficult to determine for $m \geq 2$ because they are expressed in terms of the limiting eigenvectors associated with the eigenvalues that vanish. In the case $\mathsf{r} < \infty$, it is intuitively clear, and Lancaster (1964) has shown formally that these eigenvectors are smoothly defined as $r \to \mathsf{r}$. Let $\mathbf{P}_{2r}$ be $n \times m$ and denote the last m columns of $\mathbf{P}_r'$ which are the eigenvectors for the m largest eigenvalues of $(\mathbf{A} - r\mathbf{B})$ (which increase in size with column number). Then $\mathbf{P}_{2r}$ is continuous at $r = \mathsf{r}$ and the limiting noncentrality parameters are

$$(\nu_{0,n-m+1}, \ldots, \nu_{0,n})' = \mathbf{P}_{2\mathsf{r}}'\mu. \tag{47}$$

In the unbounded setting with $\mathsf{r} = \infty$, let $n \times m$ matrix $\mathbf{P}_{2\varepsilon}$ consist of the eigenvectors corresponding to the largest m eigenvalues of $\mathbf{D}(\varepsilon)$ in (43). Then the limiting noncentrality parameters are given in (47) with $\mathbf{P}_{20} = \lim_{\varepsilon\to 0}\mathbf{P}_{2\varepsilon}$ replacing $\mathbf{P}_{2\mathsf{r}}$. In complicated practical examples in which these computations are not explicit, these limiting eigenvectors are best computed numerically by taking $\varepsilon = 10^{-7}$ or $r = \mathsf{r} - 10^{-7}$.

Example 14 *The least squares estimate of a lag 2 serial correlation with $n = 3$ and zero mean has the form $R = \epsilon_1\epsilon_3/\epsilon_1^2$ and leads to the matrix*

$$\mathbf{D}(\varepsilon) = \varepsilon\mathbf{A} - \mathbf{B} = \begin{pmatrix} -1 & 0 & \frac{1}{2}\varepsilon \\ 0 & 0 & 0 \\ \frac{1}{2}\varepsilon & 0 & 0 \end{pmatrix} = \mathbf{Q}'_\varepsilon \begin{pmatrix} \psi_-(\varepsilon) & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \psi_+(\varepsilon) \end{pmatrix} \mathbf{Q}_\varepsilon$$

where

$$\mathbf{Q}'_\varepsilon = \begin{pmatrix} 2\psi_-(\varepsilon)/\varepsilon & 0 & 2\psi_+(\varepsilon)/\varepsilon \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \qquad \psi_\pm(\varepsilon) = -\tfrac{1}{2} \pm \tfrac{1}{2}\sqrt{(1+\epsilon^2)}.$$

The eigenvectors in matrix $\mathbf{Q}'_\varepsilon$ have not been normalized as would be needed to use the notation $\mathbf{P}'_\varepsilon$. The limits of the eigenvalues are

$$\lim_{\varepsilon\to 0}\{\psi_-(\varepsilon), 0, \psi_+(\varepsilon)\} = (-1, 0, 0)$$

and the limiting normed eigenvectors have $\mathbf{P}'_\varepsilon \to \mathbf{I}_3$ as $\varepsilon \to 0$. Note that $\partial\psi_+(\varepsilon)/\partial\varepsilon|_{\varepsilon=0} = 0$. Also the eigenvalues of

$$\mathbf{O}'_{\mathbf{B}2}\mathbf{A}\mathbf{O}_{\mathbf{B}2} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 \\ \frac{1}{2} & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

are both zero as discussed in Lemma 13. The limiting rate

$$\omega_2 = \lim_{r\to\infty}\frac{\lambda_2(r)}{\lambda_3(r)} = \lim_{\varepsilon\to 0}\frac{0}{\partial^2\psi_+(\varepsilon)/\partial\varepsilon^2} = \frac{0}{1/2} = 0.$$

The limiting noncentrality parameters are

$$\begin{pmatrix} \nu_{02} \\ \nu_{03} \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\mu = \begin{pmatrix} \mu_2 \\ \mu_3 \end{pmatrix}.$$

In this context the derivations for the relative errors become rather messy due to their dependence on $\{\omega_i, \nu_{0i}\}$. The derivations, however, are identical to those in the simple eigenvalue setting, except for the modifications needed to account for these differential rates of convergence. We shall briefly demonstrate these modifications in the next lemma and then simply state all the results without details since the proofs lend little further insight. All summations in the remainder of this subsection are over $\mathcal{S} = \{n-m+1, \ldots, n\}$.

Lemma 15 *Suppose R is in class $\mathcal{C}_{\mathcal{R}}$ and let m be the multiplicity of the zero eigenvalue of $\mathbf{A} - \mathsf{r}\mathbf{B}$. Then, as $r \to \mathsf{r}$,*

$$\epsilon = \lambda_n(r) \to \lambda_n(\mathsf{r}) = 0$$

and

$$\hat{s} = \frac{t_0}{\epsilon} + O(1) \to \infty, \tag{48}$$

where t_0 is the unique solution to

$$0 = -\frac{n-m}{2t_0} + \sum_{i \in \mathcal{S}} \omega_i \left\{ \frac{1}{1-2t_0\omega_i} + \frac{\nu_{0i}^2}{(1-2t_0\omega_i)^2} \right\} \tag{49}$$

in $(0, 1/2)$ with $\mathcal{S} = \{n-m+1, \ldots, n\}$. In addition,

$$\hat{u} \to u_0 = \sqrt{\frac{n-m}{2} + 2t_0^2 \sum_{i \in \mathcal{S}} \omega_i^2 \left\{ \frac{1}{(1-2t_0\omega_i)^2} + \frac{2\nu_{0i}^2}{(1-2t_0\omega_i)^3} \right\}}. \tag{50}$$

Proof. The saddlepoint equation in this instance, with $\hat{t}/\epsilon$ replacing $\hat{s}$, is

$$0 = \epsilon \left[\sum_{i=1}^{n-m} \left\{ \frac{\lambda_i}{\epsilon - 2\hat{t}\lambda_i} + O(\epsilon) \right\} + \sum_{i \in \mathcal{S}} \left\{ \frac{\lambda_i/\epsilon}{1-2\hat{t}\lambda_i/\epsilon} + \frac{\lambda_i \nu_i^2/\epsilon}{(1-2\hat{t}\lambda_i/\epsilon)^2} \right\} \right]. \tag{51}$$

Passing to the limit using $\lambda_i/\epsilon \to \omega_i$ for $i \in \mathcal{S}$, then it may be seen that the right side of (51) in the square brackets converges to the equation of (49). There is a unique solution t_0 to this equation because the right side of (49) is strictly increasing in t_0 and maps $(0, 1/2)$ onto $(-\infty, \infty)$. The remainder of the proof is straightforward. ■

Theorem 16 *Suppose $n \geq 2$, and the conditions of Lemma 15. Define operator $D_0(X)$ to be the density of random variable X evaluated at zero. Then the limiting ratio of the true* *tail probability for R to its first order Lugannani and Rice approximation in (14) is*

$$\lim_{r \to \mathrm{r}} \frac{\Pr(R > r)}{\widehat{\Pr}_1(R > r)} = \sqrt{2\pi} D_0 \left\{ \sum_{i \in \mathcal{S}} \eta_{1i} \chi^2(1, 2\eta_{2i}) - \frac{1}{2u_0} \chi^2_{n-m+2} \right\} \tag{52}$$

where the χ^2 *terms are independent random variables. Parameters* η_{1i} *and* η_{2i} *for* $i \in \mathcal{S}$ *are*

$$\eta_{1i} = \frac{t_0 \omega_i}{u_0(1 - 2t_0\omega_i)} \qquad \eta_{2i} = \frac{\nu_{0i}^2}{2(1 - 2t_0\omega_i)}.$$

In the case $m = 1$, this result reduces to that in Theorem 10.

The comparable result for the density requires some further notation. Let $\mathbf{H}_{\mathsf{r}} = (h_{ij})$ and define

$$W_J = \sum_{i \in \mathcal{S}} h_{ii}\eta_{3i} + \sum_{i \in \mathcal{S}} \sum_{j \in \mathcal{S}} \nu_{0i}\nu_{0j}\eta_{3i}\eta_{3j}h_{ij} > 0$$

where $\eta_{3i} = (1 - 2t_0\omega_i)^{-1}$.

Theorem 17 *Under the conditions of Theorem 16, the limiting ratio for the density approximation is*

$$\lim_{r \to \mathrm{r}} \frac{f_R(r)}{\hat{f}_R(r)} = \frac{\sqrt{2\pi}}{W_J} \left[\sum_{i \in \mathcal{S}} h_{ii}\eta_{3i} D_0 \left\{ \sum_{j \in \mathcal{S}} \eta_{1j} \chi^2(1, 2\eta_{2j}) + \eta_{1i}\chi_2^2 - \frac{1}{2u_0}\chi^2_{n-m} \right\} + \right.$$
$$\left. \sum_{i \in \mathcal{S}} \sum_{j \in \mathcal{S}} \nu_{0i}\nu_{0j}\eta_{3i}\eta_{3j}h_{ij} D_0 \left\{ \sum_{k \in \mathcal{S}} \eta_{1k}\chi^2(1, 2\eta_{2k}) + \eta_{1i}\chi_2^2 + \eta_{1j}\chi_2^2 - \frac{1}{2u_0}\chi^2_{n-m} \right\} \right],$$

where all χ^2 *variates are assumed to be independent.*

In the case $m = 1$, this can be shown equivalent to the result of Theorem 11. In showing this, we must use the fact that

$$\chi^2(l, 2\eta_2) + \chi_2^2 = \chi^2(l + 2, 2\eta_2)$$

as given in Johnson and Kotz (1970, p. 132)

4 Noncentral Beta $\left(\frac{m}{2}, \frac{n-m}{2}\right)$ Distribution

This distribution has

$$\mathbf{A} = \begin{pmatrix} \mathbf{I}_m & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}$$

and $\mathbf{B} = \mathbf{I}_n$ so that $\mathsf{r} = 1$ and $\omega_i \equiv 1$. This leads to the explicit expression

$$t_0 = \frac{1}{2} + \frac{1}{4n}\left\{(\theta - m) - \sqrt{(\theta - m)^2 + 4\theta n}\right\},$$

where

$$\theta = \sum_{i \in \mathcal{S}} \nu_{0i}^2 = \sum_{i=1}^{m} \mu_i^2.$$

Furthermore,

$$u_0 = \sqrt{\frac{n-m}{2} + 2t_0^2\left\{\frac{m}{(1-2t_0)^2} + \frac{2\theta}{(1-2t_0)^3}\right\}}.$$

Theorem 18 *For a noncentral Beta* $\left(\frac{m}{2}, \frac{n-m}{2}\right)$ *with* $\min(m, n-m) \geq 1$, *the limiting ratio of the true tail probability to its first order Lugannani and Rice approximation is*

$$RE_{cdf} = \frac{\sqrt{2\pi}\,(1-2t_0)^{\frac{m}{2}}\,(2t_0)^{\frac{n-m}{2}}\,u_0 e^{-\eta_2}}{B\left(\frac{m}{2}, \frac{n-m}{2}\right)\frac{n-m}{2}}\;{}_1F_1\left(\frac{n}{2}; \frac{m}{2}; \frac{\theta}{2}\right), \tag{53}$$

where

$$\eta_2 = \frac{\theta}{2(1-2t_0)}.$$

The first order saddlepoint density has the same relative error limit.

Proof. The proof follows the general flow of the single eigenvalue proof. For the density approximation, an intermediate result is that

$$\begin{aligned}\mathrm{RE}_{\mathrm{den}} =& \sqrt{2\pi}\,(1-2t_0)^{\frac{m}{2}+1}\,(2t_0)^{\frac{n-m}{2}-1}\,u_0 e^{-\eta_2}\left\{p_0\frac{{}_1F_1\left(\frac{n}{2}; \frac{m}{2}+1; \frac{\theta}{2}\right)}{B\left(\frac{m}{2}, \frac{n-m}{2}\right)\frac{m}{2}}\right.\\ &\left. + (1-p_0)\frac{(1-2t_0)\,{}_1F_1\left(\frac{n}{2}+1; \frac{m}{2}+2; \frac{\theta}{2}\right)}{B\left(\frac{m}{2}+1, \frac{n-m}{2}\right)\left(\frac{m}{2}+1\right)}\right\},\end{aligned}$$

where

$$p_0 = \frac{(1-2t_0)\sum_{i\in\mathcal{S}} h_{ii}}{(1-2t_0)\sum_{i\in\mathcal{S}} h_{ii} + \sum_{i\in\mathcal{S}}\sum_{j\in\mathcal{S}} \nu_{0i}\nu_{0j}h_{ij}} = \frac{(1-2t_0)\,m}{(1-2t_0)\,m + \theta}.$$

The same recursions for the confluent hypergeometric function, used in Theorem 11, lead to a simplification of this result so that $\mathrm{RE}_{\mathrm{den}}$ is as given in (53). ■

In the central setting with $\theta = 0$, the value in (53) reduces to

$$\mathrm{RE}_{\mathrm{den}} = \mathrm{RE}_{\mathrm{cdf}} = \hat{B}\left(\frac{m}{2}, \frac{n-m}{2}\right)/B\left(\frac{m}{2}, \frac{n-m}{2}\right). \tag{54}$$

This is consistent with the computation of the central Beta$\left(\frac{m}{2}, \frac{n-m}{2}\right)$ density in Example 5.

As $\theta \to \infty$, the limiting ratio for (53) is

$$\hat{\Gamma}\left(\frac{1}{2}, \frac{n-m}{2}\right) / \Gamma\left(\frac{1}{2}, \frac{n-m}{2}\right) \{1 + O(\theta^{-1})\}$$

which follows from the asymptotics for ${}_1F_1$ given in 13.1.4 of Abramowitz and Stegun (1972).

5 Serial Correlations

Least squares, Yule-Walker, and Burg estimates for lag l correlations without the sample mean have the respective forms

$$R_{ls} = \frac{\sum_{i=1}^{n-l} \epsilon_i \epsilon_{i+l}}{\sum_{i=1}^{n-l} \epsilon_i^2} \qquad R_{yw} = \frac{\sum_{i=1}^{n-l} \epsilon_i \epsilon_{i+l}}{\sum_{i=1}^{n} \epsilon_i^2} \qquad R_b = \frac{\sum_{i=1}^{n-l} \epsilon_i \epsilon_{i+l}}{\frac{1}{2} \sum_{i \in \mathcal{I}} \epsilon_i^2 + \sum_{i=l+1}^{n-l} \epsilon_i^2}, \tag{55}$$

where $\mathcal{I} = \{1, \ldots, l,\ n-l+1, \ldots, n\}$. Estimator R_{ls} is in class $\mathcal{C}_{\mathcal{R}} - \mathcal{B}$ whereas the other two are in $\mathcal{B}$. All are represented in the large sample space asymptotics.

5.1 Least Squares Estimates

For estimator R_{ls}, $\mathbf{O_B} = \mathbf{I}_n$ and $\mathbf{C}_{22} = \mathbf{0}$ so R_{ls} is in both $\mathcal{C}_{\mathcal{R}}$ and $\mathcal{C}_{\mathcal{L}}$. The multiplicity of the zero eigenvalue at $\mathsf{r} = \infty$ is $m = l$.

Example 19 *Consider R_{ls} with $n = 7$ and $l = 3 = m$. The matrix $\mathbf{D}(\varepsilon)$ admits 3 eigenvalues*

$$0 < \psi_5(\varepsilon) = \psi_6(\varepsilon) < \psi_7(\varepsilon)$$

which converge to 0 as $\varepsilon \to 0$. For each $\psi_i''(0) = 1/2$ so the relative rates are $\omega_5 = \omega_6 = \omega_7 = 1$. The limiting noncentrality parameters are $(\mu_5, \mu_6, \mu_7)'$.

Example 20 *Consider the previous example but suppose ϵ_4^2 is excluded from the denominator so the estimator is $R = \sum_{i=1}^{4} \epsilon_i \epsilon_{i+3} / \sum_{i=1}^{3} \epsilon_i^2$. In this instance $l = 3$ but $m = 2$. The matrix*

$$\mathbf{C}_{22} = \begin{pmatrix} 0 & 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & 0 & 0 \end{pmatrix}$$

has eigenvalues $-1/2, 0, 0,$ *and* $1/2$*. This* R *is in case 2(a) which is addressed in §8.5 of the Appendix.*

5.2 Yule-Walker and Burg Estimates

Some examples and results help to distinguish the simple and multiple eigenvalue settings that may arise using the Yule-Walker estimates. Suppose that $\mathbf{A}_l$ is the numerator matrix for lag l with the examples in (55).

Example 21 $l = 1$. *In such settings, the largest eigenvalue of* $\mathbf{A}_1$ *has multiplicity* 1 *and the simple setting applies. This result follows from the next lemma and the fact that* $\mathbf{A}_1$ *is an irreducible matrix.*

Lemma 22 *Consider a Yule-Walker ratio with general lag* l *and suppose that the* $n \times n$ *matrix* $\mathbf{A}_l$ *is irreducible. Then the simpler settings of Theorems 10 and 11 apply for large sample space relative errors.*

Proof. Consider the n components as states of a Markov chain with transition matrix as $\mathbf{A}_l$. All states communicate and the period of the chain must be 2 since step l followed by step $-l$ returns to any state. By the Perron-Frobenius theorem (Seneta, 1981, thm. 1.7), the 2 eigenvalues of $\mathbf{A}_l$ attaining the largest magnitude must be opposites that differ by the factor -1. ■

Example 23 $n = 5$ *and* $l = 2$. *The* 5×5 *matrix* $\mathbf{A}_2$ *is not irreducible and consists of two irreducible subchains with states* $\{1, 3, 5\}$ *and* $\{2, 4\}$. *The 5 eigenvalues are* 0, $\pm 1/2$, $\pm 1/\sqrt{2}$. *The support of* R *is* $\left(-1/\sqrt{2}, 1/\sqrt{2}\right)$ *and the largest eigenvalue is simple.*

Example 24 $n = 8$ *and* $l = 3$. *The* 8×8 *matrix* $\mathbf{A}_3$ *is not irreducible and consists of three irreducible subchains with states* $\{1, 4, 7\}, \{2, 5, 8\}$ *and* $\{3, 6\}$. *The 8 eigenvalues, with their multiplicities in parentheses, are* $0(2)$, $\pm 1/2(2)$, *and* $\pm 1/\sqrt{2}(2)$. *The support of* R *is* $(-1/\sqrt{2}, 1/\sqrt{2})$ *and the largest eigenvalue has multiplicity 2 so Theorems 16 and 17 are applicable.*

The irreducibility assumption of Lemma 22 guarantees the simple situation for the Yule-Walker estimate. When $\mathbf{A}_l$ is not irreducible, both simple and multiple eigenvalue settings are common.

Example 25 *The Burg estimator with* $n = 6$ *and* $l = 2$ *has support* $(-1, 1)$. *The largest eigenvalue of* $\mathbf{A}-r\mathbf{B}$ *has multiplicity 2 and is*

$$\lambda_5(r) = \lambda_6(r) = -\tfrac{3}{4}r + \tfrac{1}{4}\sqrt{r^2+8} \downarrow 0 \qquad as \quad r \to 1.$$

Thus $\omega_5 = \omega_6 = 1$. *The null space of* $\mathbf{A}-\mathbf{B}$ *has orthonormal basis given by the rows of*

$$\mathbf{U}_0' = \frac{1}{\sqrt{3}} \begin{pmatrix} 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 \end{pmatrix}$$

which delineates the communicating states with lag 2.

6 Numerical Example

In practice, serial correlations with arbitrary lag l are computed from least squares residuals and this often assures in practice that the largest eigenvalue of $\mathbf{A} - \mathrm{r}\mathbf{B}$ has algebraic multiplicity one. Thus the simpler situation for the large deviation errors occurs most often in practical data analysis.

Numerical confirmation of the large deviation errors in Theorems 10 and 11 is possible by considering the simplest model of §2.1. This is the least squares estimate of lag one with $n = 2$ in a model without a location effect. Then $R = \epsilon_2/\epsilon_1$ with $\epsilon_i \overset{\text{indep}}{\sim} \mathrm{N}(\mu_i, 1)$. The exact density can be expressed as

$$\begin{aligned} f_R(r) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} |x| \exp\left\{-\frac{1}{2}(x-\mu_1)^2\right\} \exp\left\{-\frac{1}{2}(rx-\mu_2)^2\right\} dx \\ &= (\pi\delta)^{-1} \exp\left\{-\frac{1}{2}\left(\mu_1^2+\mu_2^2\right)\right\} + \frac{\lambda\theta(\mu_1 + r\mu_2)}{\delta\sqrt{2\pi\delta}} \end{aligned} \tag{56}$$

where

$$\delta = 1 + r^2, \quad \theta = \operatorname{erf}\left(\frac{\mu_1 + r\mu_2}{\sqrt{2\delta}}\right), \quad \lambda = \exp\left(-\frac{1}{2}\frac{(\mu_1 r - \mu_2)^2}{\delta}\right).$$

From Theorem 10, the limiting relative errors in the left and right tails are dependent on ν_0 alone; in both tails this value is $\nu_0 = \mu_2$ so that the limiting relative errors are the same in both tails regardless of the values of μ_1 and μ_2.

Density (56) is both heavy tailed and bimodal for $\mu_1 = 0.2$ and $\mu_2 = 2$. Figure 1 plots the exact density, the normalized version of $\hat{f}_R$ in (18) denoted as $\bar{f}_R$, and the second order saddlepoint $\hat{f}_{R2}$ in (20) for this case. While both appear highly accurate in the tails, only the latter captures the bimodality. Figure 2 plots the ratio of the exact to the three approximate densities including $\hat{f}_R$, $\bar{f}_R$ and $\hat{f}_{R2}$. As $|r|$ increases, we have numerically confirmed that $f_R(r) \,/\, \hat{f}_R(r) \to 0.8222$, in agreement with the value computed from Theorems 10 and 11. This value is virtually achieved at $|r| = 10$. Both $\bar{f}_R$ and $\hat{f}_{R2}$ perform better than $\hat{f}_R$ in the tails, the latter most notably so.

The true cdf of R, or $F_R(r)$, must be computed from (56) using numerical integration. In this case, $|r|$ must be substantially larger before the same limiting ratio, as specified in Theorem 10, is reached. Figure 3 plots

$$\frac{F_R(r)}{\hat{F}_R(r)} 1_{\{\hat{s}<0\}} + \frac{1-F_R(r)}{1-\hat{F}_R(r)} 1_{\{\hat{s}>0\}} \quad \text{vs.} \quad r \tag{57}$$

with $\hat{F}_R(r)$ as $\widehat{\Pr}_2$ in (17) and as $\widehat{\Pr}_1$ in (14). For these values of μ_i, $\widehat{\Pr}_1$ is more accurate than $\widehat{\Pr}_2$ only in the range $-1.8 < r < 1.2$. At $r = -25,000$, $F_R(r) \,/\, \hat{F}_R(r) = 0.8226$ for $\widehat{\Pr}_1$, as given by Theorem 10, while for $\widehat{\Pr}_2$ the ratio is 1.015. This latter ratio necessarily includes the factor $(1+O_F)$ where O_F approximates the limit of the second-order correction term.

If $\mu_1 = \mu_2 = 0$, R is Cauchy, and the saddlepoint density reduces to

$$\hat{f}_R(r) = \frac{1}{\sqrt{2\pi}\,(1+r^2)} = \sqrt{\frac{\pi}{2}} f_R(r).$$

Thus $\bar{f}_R$ is exact and the saddlepoint solution to $0 = K'_X(\hat{s})$ is given by $\hat{s} = r$. The relative error is,

$$f_R(r) \,/\, \hat{f}_R(r) = \sqrt{\frac{2}{\pi}} = \frac{\hat{B}(1/2, 1/2)}{B(1/2, 1/2)},$$

in agreement with the large sample space theory. In this case, the values for $\hat{w}$ and $\hat{u}$ in the Lugannani-Rice formula reduce to

$$\hat{w} = \text{sgn}(r)\sqrt{\ln\left(1+r^2\right)}, \quad \hat{u} = r\left(1+r^2\right)^{-1/2}.$$

7 Acknowledgments

The authors gratefully acknowledge a referee whose comments lead to the correct versions of Lemmas 3 and 4. We are also grateful to Hari Iyer who clarified some algebraic issues arising in Section 3.2 concerning rational forms.

8 Appendix

8.1 Proofs for Lemmas

8.1.1 Lemma 3

Consider case 2(a) of Lemma 3. Transform $(\mathbf{z}_1', \mathbf{z}_2') = \epsilon'\mathbf{O}_\mathbf{B}'$ and rewrite

$$R = \frac{\mathbf{z}_1'\mathbf{C}_{11}\mathbf{z}_1 + 2\mathbf{z}_1'\mathbf{C}_{12}\mathbf{z}_2 + \mathbf{z}_2'\mathbf{C}_{22}\mathbf{z}_2}{\mathbf{z}_1'\mathbf{\Lambda}_\mathbf{B}\mathbf{z}_1}. \tag{58}$$

Let the the angle between $\mathbf{z}_2$ and the eigenvector for the positive eigenvalue of $\mathbf{C}_{22}$ be sufficiently small. As $||\mathbf{z}_2|| \to \infty$, then $R \to \infty$ and establishes part (a).

If $\mathbf{C}_{22} < \mathbf{0}$, R in (58) is concave in $\mathbf{z}_2$ for fixed $\mathbf{z}_1$ and attains a maximum at $\hat{\mathbf{z}}_2 = -\mathbf{C}_{22}^{-1}\mathbf{C}_{21}\mathbf{z}_1$ with R value

$$R = \frac{\mathbf{z}_1'(\mathbf{C}_{11} - \mathbf{C}_{12}\mathbf{C}_{22}^{-1}\mathbf{C}_{21})\mathbf{z}_1}{\mathbf{z}_1'\mathbf{\Lambda}_\mathbf{B}\mathbf{z}_1}. \tag{59}$$

Now the maximum value of R in (59) over $\mathbf{z}_1$ is specified in (6) of case 2(b).

If $\mathbf{C}_{22} \leq \mathbf{0}$ as in case (c), then further transform $\mathbf{z}_2$ to $(\mathbf{z}_3', \mathbf{z}_4') = \mathbf{z}_2'\mathbf{O}_\mathbf{C}'$ and rewrite

$$R = \frac{\mathbf{z}_1'\mathbf{C}_{11}\mathbf{z}_1 + 2\mathbf{z}_1'\mathbf{C}_{12}(\mathbf{O}_{\mathbf{C}1}\mathbf{z}_3 + \mathbf{O}_{\mathbf{C}2}\mathbf{z}_4) + \mathbf{z}_3'\mathbf{\Lambda}_\mathbf{C}\mathbf{z}_3}{\mathbf{z}_1'\mathbf{\Lambda}_\mathbf{B}\mathbf{z}_1}.$$

If $N(\mathbf{C}_{22}) \subseteq N(\mathbf{C}_{12})$ then $\mathbf{C}_{12}\mathbf{O}_{\mathbf{C}2} = \mathbf{0}$. Now R is concave in $\mathbf{z}_3$ due to $\mathbf{\Lambda}_\mathbf{C} < \mathbf{0}$ and the successive maximization of R over $\mathbf{z}_3$ and $\mathbf{z}_1$ leads to the maximum in part (c). In the event that $\mathbf{C}_{12}\mathbf{O}_{\mathbf{C}2} \neq \mathbf{0}$, then sequences of $\mathbf{z}_4$ values which increase $\mathbf{z}_1'\mathbf{C}_{12}\mathbf{O}_{\mathbf{C}2}\mathbf{z}_4$ lead to unbounded support for R.

8.1.2 Lemma 4

Explanations are needed only for cases 2(a) and 2(c) in instances with $\mathsf{r} = \infty$. For case (a) in which $\mathbf{C}_{22}$ has a positive eigenvalue, then $\lambda_n(r)$ cannot converge to 0 as $r \to \infty$. To show this, note that the eigenvalues of $(\mathbf{A} - r\mathbf{B})$ are also the eigenvalues of

$$\begin{aligned} \mathbf{Q}_r &= \begin{pmatrix} \mathbf{I}_{n-p} & \mathbf{0} \\ \mathbf{0} & \mathbf{O_C} \end{pmatrix} \mathbf{O_B} \left(\mathbf{A} - r\mathbf{B}\right) \mathbf{O}'_{\mathbf{B}} \begin{pmatrix} \mathbf{I}_{n-p} & \mathbf{0} \\ \mathbf{0} & \mathbf{O}'_{\mathbf{C}} \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{C}_{11} - r\mathbf{\Lambda_B} & \mathbf{C}_{12}\mathbf{O}'_{\mathbf{C}} \\ \mathbf{O_C}\mathbf{C}_{21} & \mathbf{\Lambda_C} \end{pmatrix}. \end{aligned} \tag{60}$$

Suppose the positive eigenvalue of $\mathbf{C}_{22}$ is the last diagonal element of $\mathbf{\Lambda_C}$ or the (n, n) element of $\mathbf{Q}_r$. If ξ_n is the n-vector with a 1 in its last component and zeros elsewhere, then

$$\lambda_n(r) \geq \xi'_n \mathbf{Q}_r \xi_n = (\mathbf{\Lambda_C})_{pp} > 0$$

for all r. Thus $\lambda_n(r) \nrightarrow 0$ if $\mathbf{\Lambda_C}$ has a positive eigenvalue.

Case 2(c) requires further decomposition of $\mathbf{C}_{22}$. The eigenvalues of $\mathbf{A} - r\mathbf{B}$ are now the eigenvalues of

$$\mathbf{Q}_r = \begin{pmatrix} \mathbf{C}_{11} - r\mathbf{\Lambda_B} & \mathbf{C}_{12}\mathbf{O_{C1}} & \mathbf{C}_{12}\mathbf{O_{C2}} \\ \mathbf{O}'_{\mathbf{C1}}\mathbf{C}_{21} & \mathbf{\Lambda_C} & \mathbf{0} \\ \mathbf{O}'_{\mathbf{C2}}\mathbf{C}_{21} & \mathbf{0} & \mathbf{0}_{m\times m} \end{pmatrix}. \tag{61}$$

With the condition $\mathbf{C}_{12}\mathbf{O_{C2}} = \mathbf{0}$, then 0 is an eigenvalue of multiplicity m for all r, and for sufficiently large r, all the eigenvalues in (61) are non-positive; this confirms that $\mathsf{r} < \infty$. However, suppose $\mathbf{C}_{12}\mathbf{O_{C2}} \neq \mathbf{0}$ and, for example, the $(1, n)$ element of $\mathbf{Q}_r$ is the value $(\mathbf{Q}_r)_{1n} > 0$. Then for any r, a positive quadratic form in $\mathbf{Q}_r$ can be constructed that assures $\lambda_n(r) > 0$. To construct this quadratic form, let ξ_1 be the indicator of the 1st component. For $\varepsilon = \varepsilon(r) > 0$ sufficiently small and $\theta = \pi/2 - \varepsilon$, then

$$\begin{aligned} \lambda_n(r) &\geq (\xi_1 \cos\theta + \xi_n \sin\theta)' \mathbf{Q}_r (\xi_1 \cos\theta + \xi_n \sin\theta) \\ &= (\mathbf{Q}_r)_{11} \cos^2\theta + 2(\mathbf{Q}_r)_{1n} \cos\theta \sin\theta > 0. \end{aligned}$$

If $(\mathbf{Q}_r)_{1n} < 0$ then take $\theta = -\pi/2 + \varepsilon$. Thus, for finite r, $\lambda_n(r) > 0$ but without $r < \infty$, this argument fails.

For any $r < \infty$, the multiplicity of the zero eigenvalue is at least the number of columns in $\mathbf{C}_{12}\mathbf{O}_{\mathbf{C}2}$ that are filled entirely with zeros. For each column having a nonzero entry, the above argument may be used to show that an additional eigenvalue converges to zero. Furthermore, with respect to the usual basis $\{\xi_1, ..., \xi_n\}$, the first $n - m$ coordinates of the eigenvectors for the largest m eigenvalues must all converge to zero as $r \to \infty$; thus the largest m eigenvalues $\lambda_{n-m+1}(r), ..., \lambda_n(r) \to 0$ as $r \to \infty$ and their associated eigenvectors eventually span $\{\xi_{n-m+1}, \ldots, \xi_n\}$.

8.1.3 Lemma 8

The structure of $\mathbf{Q}_r$ in (61) assures the following eigenstructure as $r \to \infty$: (i) the $n-p$ smallest eigenvalues are $O(-r)$ with associated eigenvectors that eventually span $\{\xi_1, \ldots, \xi_{n-p}\}$; (ii) the $p - m$ finite intermediate eigenvalues are the diagonal elements of $\mathbf{\Lambda_C}$ whose eigenvectors are eventually $\{\xi_{n-p+1}, \ldots, \xi_{n-m}\}$ if the diagonal entries of $\mathbf{\Lambda_C}$ increase with indices; and (iii) the largest eigenvalue is zero and has multiplicity m with vector space eventually spanning $\{\xi_{n-m+1}, \ldots, \xi_n\}$. Clearly (i) must hold. The argument at the end of Lemma 4 proves that (iii) holds with at least multiplicity m. Now by exclusion the remaining eigenvectors must span $\{\xi_{n-p+1}, \ldots, \xi_{n-m}\}$ resulting in (ii).

If $m = 1$ then, as $r \to \infty$,

$$\mathbf{0} \longleftarrow (\mathbf{A} - r\mathbf{B})\,\mathbf{O}'_{\mathbf{B}} \begin{pmatrix} \mathbf{I}_{n-p} & \mathbf{0} \\ \mathbf{0} & \mathbf{O}'_{\mathbf{C}} \end{pmatrix} \xi_n = (\mathbf{A} - r\mathbf{B})\,\mathbf{O}_{\mathbf{B}2}\mathbf{o}_n,$$

where $\mathbf{o}_n$ is the last column of $\mathbf{O}'_{\mathbf{C}}$ and $\mathbf{O}_{\mathbf{B}2}\mathbf{o}_n$ is the eigenvector for $\lambda_n(\infty) = 0$.

8.2 Derivation of (18)

Ratio R has the form $R = V/W$ so that its density at r may be expressed as the density of a "constructed" random variable Y_r using the Geary (1944) representation for the density of a ratio of random variables. This is the essential aspect to the approaches used by Daniels (1954, Sec. 9) and later Lieberman (1994b). The density of R at r, or $f_R(r)$ for fixed value r, can be expressed in terms of the density of random variable Y_r at 0, or $f_{Y_r}(0)$, where Y_r

is the "constructed" random variable associated with mgf

$$M_{Y_r}(s) = \frac{1}{\mathcal{E}(W)} \frac{\partial}{\partial t} M_{V,W}(s,t)|_{t=-rs}. \tag{62}$$

The relationship is

$$f_R(r) = \mathcal{E}(W)\, f_{Y_r}(0) \tag{63}$$

and is developed in Stuart and Ord (1994, Sec. 11.10). Density $f_{Y_r}(0)$ is easily approximated with $\hat{f}_{Y_r}(0)$, its saddlepoint approximation as in Daniels' (1954); a density approximation for $f_R(r)$ is therefore

$$\hat{f}_R(r) = \mathcal{E}(W)\, \hat{f}_{Y_r}(0) \tag{64}$$

which is Daniels (1954, §9) approximation.

The joint mgf of (V, W) required in (62) is easily computed as

$$M(s,t) = M_{V,W}(s,t) = |\mathbf{\Omega}|^{-\frac{1}{2}} \exp\left\{-\tfrac{1}{2}\mu'\left(\mathbf{I}_n - \mathbf{\Omega}^{-1}\right)\mu\right\}$$

where $\mathbf{\Omega} = \mathbf{I}_n - 2(s\mathbf{A} + t\mathbf{B})$. Take $\partial/\partial t$ by using the chain rule and the rules of matrix differentiation and evaluate this at $t = -rs$ to get

$$\frac{\partial M(s,t)}{\partial t}|_{t=-rs} = |\mathbf{\Xi}_s|^{-\frac{1}{2}} \exp\left\{s\mu'\mathbf{\Xi}_s^{-1}\mathbf{D}_r\mu\right\}\left\{\mu'\mathbf{\Xi}_s^{-1}\mathbf{B}\mathbf{\Xi}_s^{-1}\mu + \operatorname{tr}\mathbf{\Xi}_s^{-1}\mathbf{B}\right\} \tag{65}$$

where $\mathbf{\Xi}_s = \mathbf{I} - 2s\mathbf{D}_r$ and $\mathbf{D}_r = \mathbf{A} - r\mathbf{B}$. The canonical reduction $\mathbf{P}_r\mathbf{D}_r\mathbf{P}_r' = \mathbf{\Lambda}_r$ also applies to $\mathbf{\Xi}_s$ and is used to rewrite (65) in the simpler form

$$\begin{aligned}\frac{\partial M(s,t)}{\partial t}|_{t=-rs} &= |\mathbf{I} - 2s\mathbf{\Lambda}_r|^{-\frac{1}{2}} \exp\left\{s\nu_r'\left(\mathbf{I} - 2s\mathbf{\Lambda}_r\right)^{-1}\mathbf{\Lambda}_r\nu_r\right\} J_r(s) \\ &= M_{X_r}(s)\, J_r(s) \end{aligned} \tag{66}$$

with $J_r(s)$ given in (19) and $M_{X_r}(s)$ in (11).

The expression in (66) provides a saddlepoint approximation for the density of Y_r at 0 as required in (64). First note that

$$\mathcal{E}(W) = \frac{\partial M(s,t)}{\partial t}|_{t=-rs=0} = J_r(0),$$

so we see that the mgf of Y_r is

$$M_{Y_r}(s) = M_{X_r}(s)\frac{J_r(s)}{J_r(0)} \tag{67}$$

and the cgf is therefore $K_{X_r}(s) + \ln[J_r(s)/J_r(0)]$. We may ignore the second term in this expression when determining the saddlepoint since $K_{X_r}(s)$ is strictly convex and clearly the dominant term in the expression. By doing so we also assure that the saddlepoint at r for the density approximation is the same as the saddlepoint used for the cdf approximation at r. Following this approach, then

$$\begin{aligned}\hat{f}_R(r) &= \mathcal{E}(W)\hat{f}_{Y_r}(0) \\ &= \frac{1}{\sqrt{2\pi K''_{X_r}(\hat{s})}} M_{X_r}(\hat{s}) J_r(\hat{s})\end{aligned}$$

as specified in (18).

8.3 Dominating function for (32)

The norm of the first factor of the integrand in (32) is

$$\left(1 + \frac{t^2}{\hat{s}^2\hat{\sigma}^2}\right)^{-\frac{1}{2}} \le \left(1 + \frac{t^2}{c_1^2}\right)^{-\frac{1}{2}} \tag{68}$$

for some $c_1^2 > u_0^2$ and sufficiently large r. The centrality factors inside the M_{X_r}-ratio norm are bounded as

$$\begin{aligned}\prod_{i=1}^{n}\left\|\frac{1-2(\hat{s}+it/\hat{\sigma})\lambda_i}{1-2\hat{s}\lambda_i}\right\|^{-\frac{1}{2}} &= \prod_{i=1}^{n}\left\{1 + \frac{4t^2\lambda_i^2}{\hat{\sigma}^2(1-2\hat{s}\lambda_i)^2}\right\}^{-\frac{1}{2}} \\ &\le \left(1+\frac{t^2}{c_1^2}\right)^{-\frac{n-1}{2}}\left(1+\frac{t^2}{c_2^2}\right)^{-\frac{1}{2}} \end{aligned} \tag{69}$$

for sufficiently large r where $c_2^2 > 1/\left(4\eta_1^2\right)$. This holds because of the convergence shown in (34). The norm for the i^{th} noncentrality term in the M_{X_r}-ratio involves only the real portion of the exponent in (35) or

$$\frac{\nu_i^2}{2}\left(\frac{1-2\hat{s}\lambda_i}{(1-2\hat{s}\lambda_i)^2 + 4t^2\lambda_i^2/\hat{\sigma}^2} - \frac{1}{1-2\hat{s}\lambda_i}\right) \le 0$$

for all i. Thus all the noncentrality terms have norms bounded above by 1. The dominating function is the product of expressions (69) and (68) which is integrable.

8.4 Results for Theorem 11

Denote the components of $\mathbf{H}_r$ as $\mathbf{H}_r = (h_{ij})$. Consider $J_r(s)$ as a function of complex variable at the two values $s = \hat{s} + it/\hat{\sigma}$ or $\hat{s}$. At either point we have

$$J_r(s) = \sum_{i=1}^{n} \frac{h_{ii}}{1-2s\lambda_i} + \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{h_{ij}\nu_i\nu_j}{(1-2s\lambda_i)(1-2s\lambda_j)} \tag{70}$$

$$= \frac{h_{nn}}{1-2s\lambda_n} + \frac{h_{nn}\nu_n^2}{(1-2s\lambda_n)^2} + O(\epsilon) \tag{71}$$

as $\epsilon \to 0$. At $s = \hat{s} + it/\hat{\sigma}$ and for $i = 1, \ldots, n-1$, each term $(1-2s\lambda_i)^{-1}$ is uniformly bounded in t, for sufficiently large r, and each term also converges to zero. The $i = j = n$ terms are retained in (71) and are bounded. Thus $\|J_r(\hat{s} + it/\hat{\sigma})\|$ is bounded and $J_r(\hat{s})$ stays bounded away from zero so the dominating converge theorem can be used. A simple computation shows

$$J_r(\hat{s} + it/\hat{\sigma}) \sim h_{nn}\left\{\frac{1}{(1-2t_0)(1-2it\eta_1)} + \frac{\nu_0^2}{(1-2t_0)^2(1-2it\eta_1)^2}\right\}, \tag{72}$$

and the limiting behavior of $J_r(\hat{s})$ is determined by setting $t = 0$ in (72). The limiting ratio in (41) now follows.

To show that (39) agrees analytically with the right side of (29), two different recursions for the confluent hypergeometric function are used. Expression 13.4.7 of Abramowitz and Stegun (1972), taken with $a = n/2$ and $b = 3/2$, leads to the recursion

$$\frac{n\nu_0^2}{3}\,{}_1F_1\left(\frac{n}{2}+1;\frac{5}{2};\frac{\nu_0^2}{2}\right) = \left(\nu_0^2 - 1\right){}_1F_1\left(\frac{n}{2};\frac{3}{2};\frac{\nu_0^2}{2}\right) + {}_1F_1\left(\frac{n}{2}-1;\frac{1}{2};\frac{\nu_0^2}{2}\right). \tag{73}$$

The expression in the curly braces of (39) may be reduced to

$$\frac{2(1-2t_0)\Gamma\left(\frac{n}{2}\right)}{\sqrt{\pi}\Gamma\left(\frac{n-1}{2}\right)\left(1-2t_0+\nu_0^2\right)}\left\{{}_1F_1\left(\frac{n}{2};\frac{3}{2};\frac{\nu_0^2}{2}\right) + \frac{n\nu_0^2}{3}\,{}_1F_1\left(\frac{n}{2}+1;\frac{5}{2};\frac{\nu_0^2}{2}\right)\right\}.$$

After substitution of (73), this becomes

$$\frac{2(1-2t_0)\Gamma\left(\frac{n}{2}\right)}{\sqrt{\pi}\Gamma\left(\frac{n-1}{2}\right)\left(1-2t_0+\nu_0^2\right)}\left\{\nu_0^2\,{}_1F_1\left(\frac{n}{2};\frac{3}{2};\frac{\nu_0^2}{2}\right) + {}_1F_1\left(\frac{n}{2}-1;\frac{1}{2};\frac{\nu_0^2}{2}\right)\right\}. \tag{74}$$

Recursion 13.4.4 of Abramowitz and Stegun (1972) is now used with $a = n/2$ and $b = 1/2$ so that

$$\nu_0^2\,{}_1F_1\left(\frac{n}{2};\frac{3}{2};\frac{\nu_0^2}{2}\right) = {}_1F_1\left(\frac{n}{2};\frac{1}{2};\frac{\nu_0^2}{2}\right) - {}_1F_1\left(\frac{n}{2}-1;\frac{1}{2};\frac{\nu_0^2}{2}\right).$$

Upon substitution, (74) becomes

$$\frac{2\left(1-2t_0\right)\Gamma\left(\frac{n}{2}\right)}{\sqrt{\pi}\Gamma\left(\frac{n-1}{2}\right)\left(1-2t_0+\nu_0^2\right)}\,{}_1F_1\left(\frac{n}{2};\frac{1}{2};\frac{\nu_0^2}{2}\right). \tag{75}$$

With (75) used in the curly braces of (39), then (39) simplifies to (29) using simple algebra.

8.5 Relative Errors for Case 2(a) of Lemma 3

The $F_{1,1}$ example of §2 provides a simple example that illustrates the asymptotics involved in this case. For finite r, the cgf admits saddlepoint $\hat{s}$ solving

$$0=K_r'(\hat{s})=\frac{1}{1-2\hat{s}}-\frac{r}{1+2r\hat{s}}\rightarrow\frac{1}{1-2\hat{s}}-\frac{1}{2\hat{s}} \tag{76}$$

as $r\to\infty$. The solution to (76) at $r=\infty$ is $\hat{s}=1/4$ which is well inside the edge of the convergence strip at 1/2. This is the same asymptotics as occurs with Satterthwaite F-type ratios as detailed in Butler and Paolella (2002). Because $\hat{s}$ stops short of the edge of the convergence strip, $\hat{u}$ also converges to a finite limit as $r\to\infty$. However, the value $\hat{w}\to\infty$.

More generally, suppose $\lambda_n(r)\nrightarrow 0$ as $r\to\infty$. Here, $\lambda_{n-p+1}(r),\ldots,\lambda_n(r)$ converge to the ordered entries in $\mathbf{\Lambda_C}$ denoted as $\lambda_{n-p+1}\leq\cdots\leq\lambda_n>0$ and $\lambda_1(r),\ldots,\lambda_{n-p}(r)\to-\infty$. Then as $r\to\infty$, the saddlepoint solution to (13) approaches $\hat{s}_\infty$, the solution to

$$0=-\frac{n-p}{2\hat{s}_\infty}+\sum_{i=n-p+1}^{n}\frac{\lambda_i}{1-2\hat{s}_\infty\lambda_i}\left(1+\nu_{i0}^2+\frac{2\hat{s}_\infty\lambda_i\nu_{i0}^2}{1-2\hat{s}_\infty\lambda_i}\right), \tag{77}$$

where $\{\nu_{i0}\}$ are the limiting noncentrality parameters. The right side of (77) is a strictly increasing function of $\hat{s}_\infty$ that maps $(0,1/(2\lambda_n))$ into $(-\infty,\infty)$ and its root is the upper bound to the saddlepoint which is well below $1/(2\lambda_n)$ the right edge of the convergence strip. For the $F_{1,1}$ example, this root is 1/4.

The remaining asymptotics for the right tail in this case are given in Butler and Paolella (2002) where an explicit value for the limiting relative error has also been derived.

References

[1] Abramowitz, M. and Stegun, I.A. (1972). *Handbook of Mathematical Functions.* Dover, New York.

[2] Butler, R. and Paolella, M.S. (1998a). Computing the density and distribution of ratios of quadratic forms in normal variables. *Working Paper, Institut für Statistik und Ökonometrie der Universität Kiel.*

[3] Butler, R. and Paolella, M.S. (1998b). Approximate distributions for the various serial correlograms. *Bernoulli* **4**:497-518.

[4] Butler, R.W and Paolella, M.S. (2002) Saddlepoint approximation and bootstrap inference for the Satterthwaite class of ratios. *Journal of the American Statistical Association* **97** 836-846.

[5] Daniels, H.E. (1954). Saddlepoint approximation in statistics. *Ann. Math. Statist.* **25**:631-650.

[6] Daniels, H.E. (1956). The approximate distribution of serial correlation coefficients. *Biometrika* **43**:169-185.

[7] Daniels, H.E. (1987). Tail probability approximation. *International Statistical Review* **55**:37-48.

[8] Durbin, J. (1980). The approximate distribution of serial correlation coefficients calculated from residuals on Fourier series. *Biometrika* **67**:335-350.

[9] Geary, R. C. (1944). Extension of a theorem by Harald Cramér on the frequency distribution of the quotient of two variables. *Journal of the Royal Statistical Society.* **17**:56-7.

[10] Imhof, J.P. (1961). Computing the distribution of quadratic forms in normal variables. *Biometrika* **48**:417-26.

[11] Johnson, N. L. and Kotz, S. (1970). *Distributions in Statistics, Continuous Univariate Distributions, Vol. 2.* Houghton Mifflin, Boston.

[12] Jensen, J.L. (1988). Uniform saddlepoint approximations. *Adv. Appl. Prob.* **20**:622-634.

[13] Jensen, J.L. (1995). *Saddlepoint Approximations.* Oxford University Press, Oxford.

[14] Lancaster, P. (1964). On eigenvalues of matrices dependent on a parameter. *Numerische Mathematik* **6**:377-387.

[15] Lieberman, O. (1994a). Saddlepoint approximation for the least squares estimator in first-order autoregression. *Biometrika* **81**:807-11. Amendment 1996, **83**:246.

[16] Lieberman, O. (1994b). Saddlepoint approximation for the distribution of a ratio of quadratic forms in normal variables. *Journal of the American Statistical Association* **89**:924-928.

[17] Lugannani, R. and Rice, S. O. (1980). Saddlepoint approximations for the distribution of sums of independent random variables. *Adv. Appl. Probab.* **12**:475–490.

[18] Magnus, R. and Neudecker, H. (1988). *Matrix Differential Calculus with Applications in Statistics and Econometrics.* Wiley, Chichester.

[19] Marsh, P. W. R. (1998). Saddlepoint approximations and non-central quadratic forms. *Econometric Theory* **14**:539-559.

[20] McGregor, J.R. (1960). An approximate test for serial correlation in polynomial regression. *Biometrika* **47**:111-9.

[21] Phillips, P.C.B. (1978). Edgeworth and saddlepoint approximations in the first-order non-circular autoregression. *Biometrika* **65**:91-8.

[22] Seneta, J.L. (1981). *Non-negative Matrices and Markov Chains.* Springer, NewYork.

[23] Stuart, A. and Ord, J. (1994). *Kendall's Advanced Theory of Statistics*, Volume 1. Edward Arnold, London, Sixth Edition.

[24] Wang, S. (1992). Tail probability approximations in the first-order non-circular autoregression. *Biometrika* **79**:431-4.

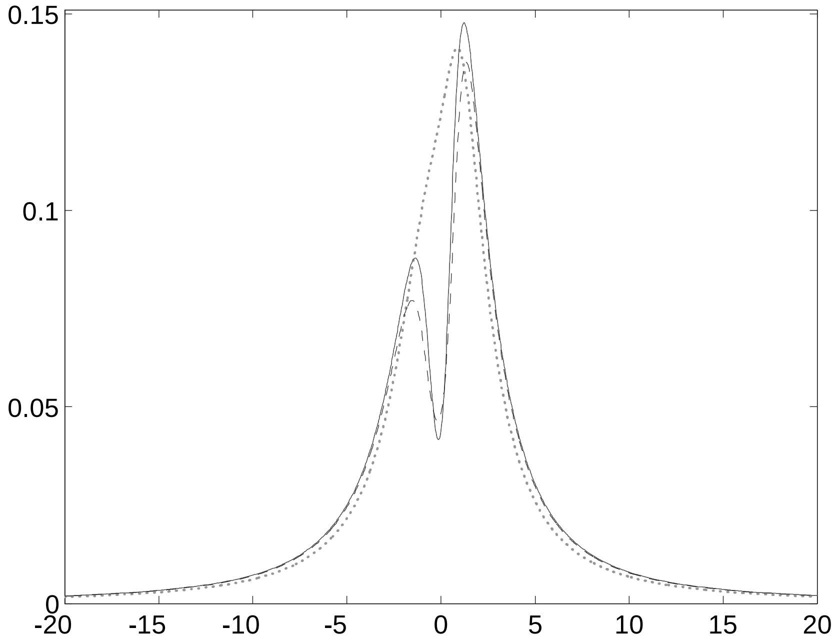

Figure 1: Exact density f_R (solid), second order $\hat{f}_{R2}$ (dashed), and normalized $\bar{f}_R$ (dotted) approximations.

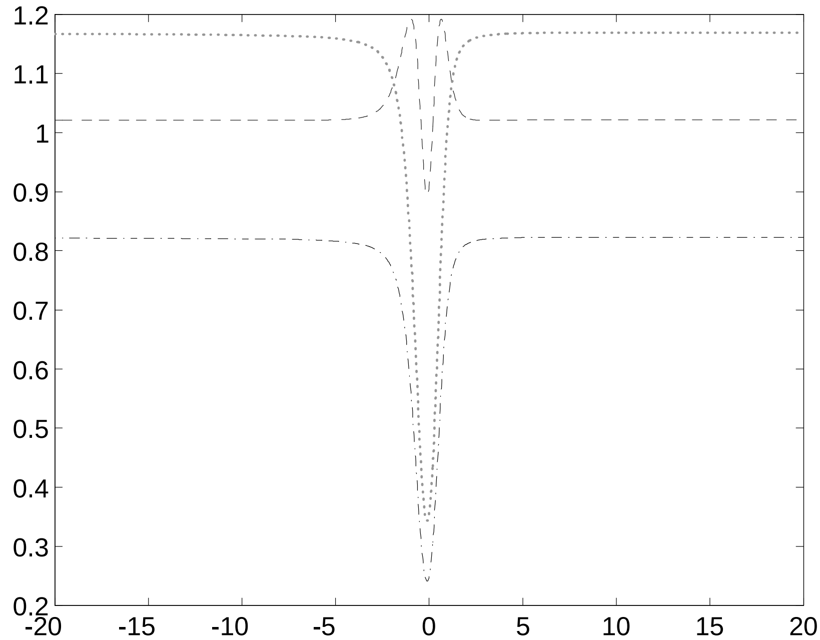

Figure 2: Error ratios $f_R/\hat{f}_{R2}$ (dashed), $f_R/\hat{f}_R$ (dashed-dot) and $f_R/\bar{f}_R$ (dotted).

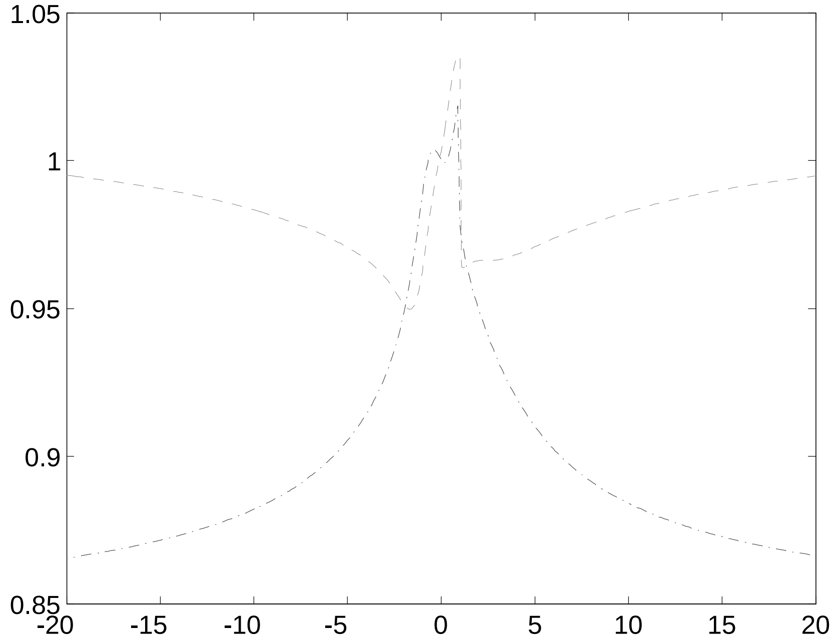

Figure 3: The tail error ratios described in (57) for $\hat{F}_R = \widehat{\mathrm{Pr}}_2$ (dashed) and $\hat{F}_R = \widehat{\mathrm{Pr}}_1$ (dash-dot).