Statistics and Computing **15**: 281–287, 2005

Approximation of power in multivariate analysis

RONALD W. BUTLER* and ANDREW T.A. WOOD†

**Department of Statistics, Colorado State University, Fort Collins, CO 80523 USA*
walrus@stat.colostate.edu
†*School of Mathematical Sciences, University of Nottingham, Nottingham NG7 2RD, UK*
atw@maths.nott.ac.uk

Received December 2002 and accepted May 2005

We consider the calculation of power functions in classical multivariate analysis. In this context, power can be expressed in terms of tail probabilities of certain noncentral distributions. The necessary noncentral distribution theory was developed between the 1940s and 1970s by a number of authors. However, tractable methods for calculating the relevant probabilities have been lacking. In this paper we present simple yet extremely accurate saddlepoint approximations to power functions associated with the following classical test statistics: the likelihood ratio statistic for testing the general linear hypothesis in MANOVA; the likelihood ratio statistic for testing block independence; and Bartlett's modified likelihood ratio statistic for testing equality of covariance matrices.

Keywords: block independence, covariance equality, general linear hypothesis, hypergeometric functions, Laplace approximation, power function, saddlepoint approximation

1. Introduction

In this paper we present simple yet extremely accurate saddlepoint approximations to the power functions of three well-known test statistics in classical multivariate analysis: the likelihood ratio statistic, denoted Λ_{GLM}, for testing the general linear model; the likelihood ratio statistic, denoted Λ_{BI}, for testing block independence; and Bartlett's M statistic, denoted Λ_{ECM}, for testing equality of covariance matrices across populations. See Section 2.1 for further details.

The noncentral distribution theory for these statistics, which is mathematically sophisticated, has been developed in a sequence of papers by Anderson, James, Constantine, Pillai, Sugiura, Muirhead and many others from the 1940s onwards. These developments made use of new tools, especially zonal polynomials (James 1960, 1961) and hypergeometric functions of matrix argument (Herz 1955). The monograph by Muirhead (1982) is an excellent starting point; see also Anderson (1984).

Despite impressive achievements, the multivariate noncentral distribution theory has unfortunately had relatively little practical impact due to the practical difficulty in computing the theoretical expressions that specify the relevant noncentral distributions. We shall use saddlepoint methods to determine these distributions. Saddlepoint methods have already been successfully employed to calculate null distributions of statistics which fall in the Box class (first introduced by Box (1949)). The Box class is characterized as consisting of statistics whose null moment generating function (MGF) can be expressed as a ratio of products of gamma functions. A prominent member of this class is the likelihood ratio test statistic for the general linear hypothesis, Λ_{GLM}, whose null distribution has been approximated in Srivastava and Yao (1989) and Butler, Huzurbazar and Booth (1992). For further details of these and other applications in the Box class see Butler, Booth and Huzurbazar (1993) and Booth *et al.* (1995) for saddlepoint approximations based on the normal distribution, and see Jensen (1991, 1995) for saddlepoint approximations based on the gamma distribution. The statistics Λ_{BI} and Λ_{ECM} are also in the Box class.

Our approach in the non-null case involves two steps: replacing the relevant hypergeometric function of matrix argument in the noncentral MGF (see Section 2.2) with a Laplace approximation (see Section 2.3), followed by the application of saddlepoint methods to this approximate MGF as in the null case (see Sections 2.4 and 2.5). This two step approach is referred to as a *sequential* saddlepoint approximation. The Laplace approximations to the matrix–argument hypergeometric functions ${}_1F_1$

Table 1. One-way MANOVA. *Approximations for the noncentral CDF of* $\ln \Lambda_{GLM}$ *in the general linear MANOVA hypothesis. The various methods are sequential saddlepoint approximations (Seq. Sad.), the* $O(n^{-3})$ *expansions of Sugiura and Fujikoshi (1969), the noncentral F approximations of Muller and Peterson (1984), the* $O(n^{-3/2})$ *expansion of Sugiura (1973a), and Monte Carlo estimation with 10,000 repetitions. The table entries are the CDF approximations evaluated at the simulated empirical percentiles and should be compared with the true percentiles in the top row*

	1	5	10	30	50	70	90	95	99
(p, n, m)		(I, J)				$\Omega = J\Omega_1$			
(3, 24, 7)		(8, 4)				$\Omega_1 = \text{diag}(1/2, 1/2, 3/4)$			
Seq. Sad.	1.011	4.992	9.979	29.94	49.93	69.98	89.95	94.96	98.99
$O(n^{-3})\chi^2$	1.385	5.525	10.48	30.16	49.95	69.91	89.90	94.94	98.98
Noncent. F	.9920	4.878	9.740	29.31	49.07	69.12	89.43	94.65	98.90
$O(n^{-3/2})$	.0000	.0000	.0000	.0134	.591	7.290	79.60	153.7	296.0
10^4 Simulations	1.310	3.940	10.46	31.61	48.30	70.83	89.74	94.45	99.36
(7, 56, 7)		(8, 8)				$\Omega_1 = \text{diag}\{1/4(1/4)7/4\}$			
Seq. Sad.	1.010	5.003	10.00	29.98	49.99	70.01	89.99	95.01	99.01
$O(n^{-3})\chi^2$	4.434	7.679	11.35	28.76	48.50	69.13	89.88	95.02	99.03
Noncent. F	.6482	3.539	7.436	24.55	43.44	64.00	86.70	93.00	98.45
$O(n^{-3/2})$	.0000	.0020	.1867	.7944	5.612	24.90	103.8	158.0	229.1
(7, 24, 7)		(8, 4)				$\Omega_1 = \text{diag}\{1(1)7\}$			
Seq. Sad.	.9633	4.883	9.762	29.62	49.67	69.78	89.89	94.95	99.00
$O(n^{-3})\chi^2$	−60.4	−102	−82.8	−.896	49.30	79.36	96.04	98.46	99.80
Noncent. F	.1655	1.213	2.919	12.29	27.24	46.84	75.11	85.14	95.79
(16, 84, 11)		(12, 8)				$\Omega_1 = \{5(0), 3(1/4), 3(1/2), 3(3/4), 2(1)\}$			
Seq. Sad	.9851	4.982	9.957	29.94	49.94	69.97	90.00	95.01	99.00
$O(n^{-3})\chi^2$	2.087	6.561	11.39	30.18	49.42	69.25	89.62	94.80	98.96
Noncent. F	.5616	3.195	6.788	23.06	41.52	62.13	85.64	92.33	98.24
(50, 150, 49)		(50, 4)				$\Omega_1 = \{7(1/16, 1/8, 3/16, 1/4, 3/8, 1/2, 5/8)\}$			
Seq. Sad.	.9954	5.001	9.998	29.96	50.03	70.02	89.96	94.97	98.98
$O(n^{-3})\chi^2$	.7444	4.194	8.819	28.64	49.54	70.66	91.25	96.05	99.45
Noncent. F	.7276	3.919	8.117	26.07	45.42	65.87	87.76	93.64	98.62

and ${}_2F_1$, and the sequential saddlepoint approximation to the noncentral distribution of Λ_{GLM}, were first presented in Butler and Wood (2002).

We now summarize the main contributions of this paper: numerical results, presented in Tables 1–3, demonstrating the high accuracy of the sequential saddlepoint approximation for the three statistics considered; comparison with alternative approximations due to Sugiura and Fujikoshi (1969), Sugiura (1973a) and Muller *et al.* (1992); and details of how to implement the sequential saddlepoint approach, which are given in the Appendix.

The accuracy of the approximations in all three instances is exceptional. For example, in practical settings of the general linear hypothesis with small and moderately large noncentrality eigenvalues, the relative errors of the noncentral saddlepoint probabilities are typically less that 1% for all percentiles of the distribution. This accuracy is retained in high dimensions; see Butler and Wood (2004) for central limit results in which dimension increases.

2. Computational details of the approximations

We now explain how to calculate our approximations to the non-null distributions of the three test statistics considered in this paper.

2.1. *The test statistics*

2.1.1. The likelihood ratio statistic in MANOVA

Consider the general linear model $Y = XB + E$ where the rows of E are i.i.d. $N_p(0, \Sigma)$, $X\,(N \times q)$ is a known full rank matrix of covariates, $B\,(q \times p)$ is a matrix of unknown parameters, $E\,(N \times p)$ is the "error" matrix and $Y\,(N \times p)$ is the observation matrix. We assume that Σ is an unknown covariance matrix of full rank. The general linear hypothesis H_{GLM} can be expressed as $H_{GLM} : CB = 0_{m,p}$ where $C\,(m \times q)$ is a suitable rank m matrix of known constants, and $0_{m,p}$ is the $m \times p$ matrix of zeros.

Table 2. *Sequential saddlepoint approximation to the CDF of the noncentral distribution for* $\ln \Lambda_{BI}$. *The table entries are interpreted in the same manner as those in Table 1*

(p_1, p_2, n)	1	5	10	30	50	70	90	95	99
(2, 3, 10)					$P = \text{diag}\{.1, .2\}$				
Seq. Sad.	.9877	4.995	10.03	30.06	50.04	70.02	90.04	95.02	98.99
$O(n^{-3/2})$	.0006	1.637	9.332	57.07	91.91	112.7	121.1	120.0	115.4
$O(n^{-3})$	1.453	9.449	18.62	47.42	68.56	84.52	96.12	98.28	99.70
(5, 7, 20)					$P = \text{diag}\{.05(.05).25\}$				
Seq. Sad.	1.003	4.992	9.972	30.04	50.03	70.08	89.98	95.01	99.00
$O(n^{-3/2})$	.0000	.0000	.0000	.0005	1.065	10.28	93.19	186.1	402.6
$O(n^{-3})$	3.455	15.64	27.52	59.68	79.29	91.72	98.63	99.54	99.96
(10, 13, 40)					$P = \text{diag}\{.1(.1).9, .95\}$				
Seq. Sad.	.9790	4.953	9.957	29.91	49.95	69.99	90.03	95.01	99.00
(25, 28, 70)					$P = \text{diag}\{2(.1), 2(.2), 3(.3), \ldots, 3(.7), 2(.8), 2(.9), 2(.95)\}$				
Seq. Sad.	.9903	5.004	9.990	29.96	49.97	69.99	90.05	95.01	98.99

Table 3. *Sequential saddlepoint approximation to the CDF of the noncentral distribution for* $\ln \Lambda_{ECM}$. *The table entries are interpreted in the same manner as those in Table 1*

(p, n_1, n_2)	1	5	10	30	50	70	90	95	99
(3, 5, 8)					$\Delta = \text{diag}\{.9, .95, .975\}$				
Lug. & Rice	1.002	4.964	9.950	29.92	49.93	69.98	90.01	94.97	98.99
χ^2 to $O(n^{-3/2})$	.8234	4.435	9.169	28.78	48.86	69.22	89.73	94.83	98.96
(3, 5, 8)					$\Delta = \text{diag}\{.2, .5, .7\}$				
Lug. & Rice	1.022	5.033	10.04	29.81	49.95	69.60	89.86	94.95	99.00
χ^2 to $O(n^{-3/2})$	.4049	3.113	6.921	23.47	41.75	61.90	85.41	92.26	98.31
(8, 15, 20)					$\Delta = \text{diag}\{.8(.025).975\}$				
Lug. & Rice	.9793	4.981	9.978	29.91	49.87	69.89	89.98	94.99	99.02
χ^2 to $O(n^{-3/2})$	.6594	3.833	8.121	26.63	46.39	67.10	88.73	94.29	98.86
(16, 22, 27)					$\Delta = \text{diag}\{.80(.0125).9875\}$				
Lug & Rice	.9753	4.964	9.970	29.96	50.05	70.03	89.98	95.00	99.00
χ^2 to $O(n^{-3/2})$	.1731	1.436	3.605	15.97	33.00	54.39	81.59	89.96	97.67
(16, 50, 60)					$\Delta = \text{diag}\{.05(.05).80\}$				
Lug. & Rice	1.117	5.400	10.66	31.08	51.08	70.82	90.34	95.18	99.06
(32, 120, 125)					$\Delta = I_2 \otimes \text{diag}\{.05(.05).80\}$				
Lug. & Rice	1.170	5.654	11.00	31.81	51.89	71.47	90.69	95.41	99.08

The likelihood ratio statistic Λ_{GLM} for testing the null hypothesis H_{GLM}, may be written $\Lambda_{GLM} = |SS_{err}|/|SS_{hypo} + SS_{err}|$, where $|\cdot|$ denotes determinant. Under the model described above, the following results hold: SS_{hypo} and SS_{err} are independent; SS_{err} has a p-dimensional central Wishart distribution with $n = N - q$ degrees of freedom and shape matrix Σ, denoted as $\text{Wishart}_p(n, \Sigma)$; and SS_{hypo} has a p-dimensional non-central Wishart distribution with m degrees of freedom, shape matrix Σ, and non-centrality matrix $\Omega = \Sigma^{-1} M_1^T M_1$, where $m \times p$ matrix M_1 is described in Section 10.2 of Muirhead (1982). The hypothesis H_{GLM} holds if and only if $M_1 = 0_{m,p}$.

2.1.2. *The likelihood ratio statistic for block independence*

Consider a random sample $x_1, \ldots, x_N$ of vectors from a multivariate normal population $N_p(\mu, \Sigma)$. Suppose that the observation vector splits into two components of dimensions p_1 and p_2, with $p_1 \le p_2$ and $p_1 + p_2 = p$. Suppose that the two

components correspond to the block representation $\mu = (\mu_1^T, \mu_2^T)^T$ and $\Sigma = [\Sigma_{ij}]_{i,j=1,2}$ where μ_i is of dimension p_i, and Σ_{ij} is of dimension $p_i \times p_j$. The null hypothesis of block independence (BI) may be specified as

H_{BI}: $\Sigma_{12}\,(=\Sigma_{21}^T) = 0_{p_1,p_2}$ with Σ_{11}, Σ_{22} and μ unrestricted, where $0_{p_1,p_2}$ is the $p_1 \times p_2$ matrix of zeros.

Let A denote the sample covariance matrix based on $x_1, \ldots, x_N$ with $n = N - 1$ degrees of freedom. Specify A in block form as $A = [A_{ij}]_{i,j=1,2}$ where A_{ij} is $p_i \times p_j$. Then it can be shown (see Muirhead 1982, Section 11.2) that the likelihood ratio test for block independence rejects the null hypothesis H_{BI} for small values of $\Lambda_{\mathrm{BI}} = |A|/(|A_{11}||A_{22}|)$. It turns out that the noncentral distribution of Λ_{BI} (i.e. the distribution of Λ_{BI} when $\Sigma_{12} \neq 0_{p_1,p_2}$) is determined by the quantities n, p_1, p_2 and $P = \mathrm{diag}\{\rho_1, \ldots, \rho_{p_1}\}$ where $\rho_1, \ldots, \rho_{p_1}$ are the eigenvalues of $\Sigma_{11}^{-1}\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$.

2.1.3. *Bartlett's M statistic for testing equality of covariances*

Let $x_1, \ldots, x_{N1}$ be a random sample from a $N_p(\mu_1, \Sigma_1)$ population, and let $y_1, \ldots, y_{N2}$ be a random sample from a $N_p(\mu_2, \Sigma_2)$ population. The null hypothesis of equal covariance matrices (ECM) may be written

H_{ECM}: $\Sigma_1 = \Sigma_2 = \Sigma$, with Σ, μ_1 and μ_2 unrestricted.

The Bartlett M test (or the modified likelihood ratio test) rejects H_{ECM} for small values of $\Lambda_{\mathrm{ECM}} = |A_1|^{n_1/n}|A_2|^{n_2/n}/|A_1+A_2|$ where $n_i = N_i - 1$, $n = n_1 + n_2$,

$$A_1 = \sum_{i=1}^{N_1}(x_i - \bar{x})(x_i - \bar{x})^T, \quad A_2 = \sum_{j=1}^{N_2}(y_j - \bar{y})(y_j - \bar{y})^T,$$

and $\bar{x}$ and $\bar{y}$ are the respective sample means. See Muirhead (1982).

The non-central distribution of Λ_{ECM} (i.e. the distribution of Λ_{ECM} when $\Sigma_1 \neq \Sigma_2$) is determined by the quantities p, n_1, n_2, and $\Delta = \mathrm{diag}\{\delta_1, \ldots, \delta_p\}$ where $\delta_1, \ldots, \delta_p$ are the eigenvalues of $\Sigma_1\Sigma_2^{-1}$.

2.2. *Moment generating functions*

In what follows,

$$\Gamma_p(a) = \pi^{p(p-1)/4}\prod_{i=1}^{p}\Gamma\left\{a - \frac{1}{2}(i-1)\right\}$$

is the multivariate gamma function, and $\Gamma_1(a) \equiv \Gamma(a) = \int_0^\infty x^{a-1}e^{-x}dx$ is the usual (scalar) gamma function; see e.g. Muirhead (1982, Section 2.1.2). The quantities ${}_1F_1(a; b; X)$ and ${}_2F_1(a, b; c; X)$ below are hypergeometric functions of matrix argument.

(i) A compact expression for the non-null MGF of $\ln \Lambda_{\mathrm{GLM}}$, due originally to Constantine (1963), is given in Theorem 10.5.1 of Muirhead (1982) as

$$\begin{aligned} M_{\mathrm{GLM}}(s) &= E(e^{s \ln \Lambda_{\mathrm{GLM}}}) \\ &= \frac{\Gamma_p\left(\frac{n}{2}+s\right)\Gamma_p\left\{\frac{1}{2}(n+m)\right\}}{\Gamma_p\left(\frac{n}{2}\right)\Gamma_p\left\{\frac{1}{2}(n+m)+s\right\}} \\ &\quad \times {}_1F_1\left\{s; \frac{1}{2}(n+m)+s; -\frac{1}{2}\Omega\right\}. \end{aligned}$$

(ii) The log likelihood ratio statistic for testing block independence has non-null MGF (see Sugiura and Fujikoshi 1969; Muirhead 1982, Theorem 11.2.6)

$$\begin{aligned} M_{\mathrm{BI}}(s) &= E(e^{s \ln \Lambda_{\mathrm{BI}}}) \\ &= \frac{\Gamma_{p_1}\left(\frac{n}{2}\right)\Gamma_{p_1}\left\{\frac{1}{2}(n-p_2)+s\right\}}{\Gamma_{p_1}\left(\frac{n}{2}+s\right)\Gamma_{p_1}\left\{\frac{1}{2}(n-p_2)\right\}}|I_{p_1} - P^2|^{n/2} \\ &\quad \times {}_2F_1\left(\frac{n}{2}, \frac{n}{2}; \frac{n}{2}+s; P^2\right). \end{aligned}$$

(iii) The log Bartlett M-statistic, $\ln \Lambda_{\mathrm{ECM}}$, for testing equality of covariance matrices has non-null MGF (see Sugiura 1969, Muirhead 1982, Theorem 8.2.11)

$$\begin{aligned} &M_{\mathrm{ECM}}(s) \\ &= E(e^{s \ln \Lambda_{\mathrm{ECM}}}) \\ &= \frac{\Gamma_p\left(\frac{n}{2}\right)\Gamma_p\left\{\frac{n_1}{2}\left(1+\frac{2s}{n}\right)\right\}\Gamma_p\left\{\frac{n_2}{2}\left(1+\frac{2s}{n}\right)\right\}}{\Gamma_p\left(\frac{n_1}{2}\right)\Gamma_p\left(\frac{n_2}{2}\right)\Gamma_p\left\{\frac{n}{2}\left(1+\frac{2s}{n}\right)\right\}} \\ &\quad \times |\Delta|^{n_1 s/n}\,{}_2F_1\left\{s, \frac{n_1}{2}\left(1+\frac{2s}{n}\right); \frac{n}{2}\left(1+\frac{2s}{n}\right); I_p - \Delta\right\}. \end{aligned}$$

2.3. *Laplace approximations to ${}_1F_1$ and ${}_2F_1$*

The following Laplace approximations to ${}_1F_1$ and ${}_2F_1$ were given in Butler and Wood (2002), to which we refer the reader for a more detailed account.

The calibrated approximations ${}_1\hat{F}_1(a; b; X)$ and ${}_2\hat{F}_1(a, b; c; X)$ both have explicit form. The approximation ${}_1\hat{F}_1(a; b; X)$ is given by

$$\begin{aligned} {}_1\hat{F}_1(a; b; X) &= b^{pb - p(p+1)/4}R_{1,1}^{-1/2} \\ &\quad \times \prod_{i=1}^{p}\left\{\left(\frac{\hat{y}_i}{a}\right)^a\left(\frac{1-\hat{y}_i}{b-a}\right)^{b-a}e^{x_i\hat{y}_i}\right\} \end{aligned}$$

where $X = \mathrm{diag}(x_1, \ldots, x_p)$,

$$R_{1,1} = \prod_{i=1}^{p}\prod_{j=i}^{p}\left\{\frac{\hat{y}_i\hat{y}_j}{a} + \frac{(1-\hat{y}_i)(1-\hat{y}_j)}{b-a}\right\} \text{ and}$$

$$\hat{y}_i = \frac{2a}{b - x_i + \sqrt{(x_i - b)^2 + 4ax_i}}$$

for $i = 1, \ldots, p$. The approximation ${}_2\hat{F}_1(a, b; c; X)$ is given by

$$ {}_2\hat{F}_1(a, b; c; X) = c^{pc - p(p+1)/4} R_{2,1}^{-1/2} \times \prod_{i=1}^{p} \left\{ \left(\frac{\hat{y}_i}{a} \right)^a \left(\frac{1 - \hat{y}_i}{c - a} \right)^{c-a} (1 - x_i \hat{y}_i)^{-b} \right\} $$

where $X = \text{diag}(x_1, \ldots, x_p)$, $S_i = x_i \hat{y}_i (1 - \hat{y}_i)/(1 - x_i \hat{y}_i)$,

$$ R_{2,1} = \prod_{i=1}^{p} \prod_{j=i}^{p} \left\{ \frac{\hat{y}_i \hat{y}_j}{a} + \frac{(1 - \hat{y}_i)(1 - \hat{y}_j)}{c - a} - \frac{b}{a(c-a)} S_i S_j \right\}; $$

and, putting $\tau_i = x_i(b - a) - c$,

$$ \hat{y}_i = 2a / \left(\sqrt{\tau_i^2 - 4a x_i (c - b)} - \tau_i \right). $$

These approximations are calibrated so that they achieve equality at $X = 0_p$, i.e. ${}_1\hat{F}_1(a; b; 0_p) = {}_1F_1(a; b; 0_p) = 1$ and ${}_2\hat{F}_1(a; b; c; 0_p) = {}_2\hat{F}_1(a; b; c; 0_p) = 1$; for this reason we often refer to ${}_1\hat{F}_1$ and ${}_2\hat{F}_1$ as calibrated Laplace approximations.

2.4. *The Lugananni and Rice approximation*

Given a random variable X with MGF $M(s)$ and cumulant generating function $K(s) = \ln M(s)$, the Lugananni and Rice (1980) approximation to the tail probability $\Pr(X > y)$ is given by

$$ \hat{\Pr}(X > y) = 1 - \Phi(r) + \phi(r)(u^{-1} - r^{-1}) $$

where Φ and ϕ are the cumulative distribution function and density, respectively, of a standard normal random variable; $r = \text{sgn}(\hat{s}) \sqrt{2\{\hat{s}y - K(\hat{s})\}}$ and, $u = \hat{s} \sqrt{K''(\hat{s})}$, where $\text{sgn}(\hat{s}) = \pm 1$ or 0 depending on whether $\hat{s}$ is positive, negative or zero; and $\hat{s}$ is the solution to the saddlepoint equation $K'(\hat{s}) = y$. In the above, K' and K'' are the first and second derivative, respectively, of the cumulant generating function K. The Lugananni and Rice approximation is known to be extremely accurate in a wide variety of settings, and this is shown to be the case for the statistics considered in this paper.

2.5. *Summary of the approximations*

Our approximation to the non-null distribution of Λ_{GLM} consists of the following two-stage approach.

Stage I: Use the calibrated Laplace approximation ${}_1\hat{F}_1$ given in Section 2.3 to approximate ${}_1F_1$ in $M_{\text{GLM}}(s)$ and call the approximate MGF $\hat{M}_{\text{GLM}}(s)$.

Stage II: Use the Lugananni and Rice tail probability approximation, but replace K by $\hat{K}(s) = \ln \hat{M}_{\text{GLM}}(s)$.

All that changes for the two statistics Λ_{BI} and Λ_{ECM} is that we approximate ${}_2F_1$ with the calibrated Laplace approximation ${}_2\hat{F}_1$ in their respective MGFs.

In each of the three cases, $\hat{K}(s)$ is an explicit function of s which results in nearly all of the required calculations as explicit computations. The hardest and only non-explicit computation is solving the (approximate) saddlepoint equation $\hat{K}'(\hat{s}) = y$. It is straightforward to solve this equation numerically using analytical differentiation to calculate $\hat{K}'$; see the appendix. It is also convenient to use one numerical derivative to calculate $\hat{K}''$ from $\hat{K}'$ thus simplifying the programming considerably but also maintaining sufficient computational accuracy.

3. Numerical results

Consider a balanced 1-way MANOVA design in p dimensions with I levels and J repetitions. We wish to test a hypothesis of the form $CB = 0_{m,p}$ corresponding to the standard null hypothesis in a one-way MANOVA, e.g. the level effects $\{\alpha_i : i = 1, \ldots, I\}$, constrained to add to the zero vector, are all zero. This test leads to $m = I - 1$ degrees of freedom for hypothesis and $n = I(J - 1)$ degrees of freedom for error. The noncentrality matrix in Wilks' test is given by

$$ \Omega = J\Sigma^{-1} \sum_{i=1}^{I} \alpha_i \alpha_i^T := J\Omega_1 $$

and increases with J where Ω_1 is that portion which is fixed.

The accuracy of the sequential saddlepoint is shown in Table 1 for various settings of the balanced one-way MANOVA. It displays percentages that measure saddlepoint accuracy in terms of how close the entries are to the true percentages listed in the top row. The first, second and fourth examples consider power approximation for local alternatives. In these three examples the largest percentage absolute relative error is 1.5%. The third example is a repeat of the second example with the eigenvalues of Ω_1 multiplied by 4. The largest relative error is now 3.7%. In those examples shown as well as others, the greatest accuracy was achieved with more local power computations dealing with smaller entries in Ω_1. In power computations with even larger eigenvalues than those shown for Example 3 and also considered in higher dimensions, saddlepoint accuracy was consistently maintained under 6% relative error. Table entries for the saddlepoint approximation were determined in the following way. Empirical quantiles for the noncentral distribution of $\ln \Lambda_{\text{GLM}}$ associated with the probabilities in the top row were determined by simulating 10^6 independent values of $\ln \Lambda_{\text{GLM}}$. Then sequential saddlepoint approximations were evaluated at these empirical percentiles so that Table 1 shows the accuracy of the saddlepoint approximations were the empirical percentiles regarded as exact.

Table 1 also shows the same sort of computations to determine the comparable accuracy for other power approximations discussed in the literature. These include the noncentral χ^2 expansions of Sugiura and Fujikoshi (1969) with error $O(n^{-3})$ for local power computations, the noncentral F approximation introduced in Muller and Peterson (1984) and numerically evaluated in Muller *et al.* (1992), and the expansions of Sugiura (1973a) with error $O(n^{-3/2})$. These approximations provide modestly good approximations with local power computations but deteriorate quickly as the noncentrality eigenvalues increase. Each of these computations, and particularly those of Sugiura and Fujikoshi (1969), were exceedingly difficult to program and ran much slower than the sequential saddlepoint method. None of

these methods can replicate the consistent accuracy of the sequential saddlepoint method across the range of examples considered. The expansions of Kulp and Nagarsenkar (1984) are not shown but were found to be very inaccurate.

In order to compare these analytical methods with Monte Carlo estimation, rejection sampling was used with 10,000 repetitions to estimate the power as shown for the first example. If $\hat{F}(x)$ is the empirical CDF, then

$$\sqrt{10^4}\{(\hat{F}oF^{-1})(p) - p\} \approx B(p),$$

a Brownian bridge for $p \in [0, 1]$. The error structure of the simulation is essentially determined by the Brownian bridge which is independent of F; thus the other four examples in Table 1, and those in Tables 2 and 3, lead to comparable results.

Table 2 displays sequential saddlepoint approximations for the noncentral CDF of $\ln \Lambda_{BI}$ and has a similar structure and interpretation as Table 1. The entries are the sequential saddlepoint approximation evaluated at empirical quantiles for the listed percentage levels computed from simulations of 10^6 values of Λ_{BI}. We see remarkable accuracy with the largest percentage relative error as 2.1%.

Table 2 compares these saddlepoint approximations to two other approximations for the cases in which $p_1 = 2, 5$. Sugiura and Fujikoshi (1969) suggest a $O(n^{-3/2})$ expansion based on the central limit tendency of $\ln\Lambda_{BI}$ which is denoted as $O(n^{-3/2})$. Lee (1971), Muirhead (1972), and Sugiura (1973b) proposed a $O(n^{-3})$ expansion under local alternatives in which $P \to 0$ which is denoted as $O(n^{-3})$. None of these alternative approximations show the capability of delivering the consistent accuracy seen with the saddlepoint approximation.

In Table 3 several examples involving the statistic $\ln \Lambda_{ECM}$ are provided. The first, third and fourth examples are more local power computations in which it maintains extremely high accuracy; this accuracy becomes only slightly worse in the remaining examples where power is no longer local.

For the local power examples, Table 3 also shows the results of an expansion denoted as "χ^2 to $O(n^{-3/2})$." This approximation was suggested by Sugiura (1974 equation (3.5)) as a local expansion about the null hypothesis in which $\Delta - I = O(n^{-1/2})$ and which has a leading term that is noncentral χ^2. Another $O(n^{-3/2})$ expansion for fixed Δ that was also suggested by Sugiura (1974, equation (3.9)) is not shown since it was not competitive with the previous one.

Appendix

The expression needed to solve the saddlepoint equation in the case of Wilks' statistic is

$$\hat{K}'(s) = \sum_{i=1}^{p}\left\{\psi\left[\frac{n}{2} + s - \frac{1}{2}(i-1)\right] - \psi\left[\frac{n+m}{2} + s - \frac{1}{2}(i-1)\right]\right\} + \nabla(s)$$

where $\psi(z) = d \ln \Gamma(z)/dz$ and

$$\nabla(s) = \frac{\partial}{\partial s}\ln {}_1\hat{F}_1\left(a, b; -\frac{1}{2}\Omega\right).$$

Here and also below we use $a = s$ and $b = (n + m)/2 + s$ to simplify the expressions. Derivative ∇ is computed using implicit differentiation and, after substantial simplification, yields

$$\begin{aligned}\nabla(s) = {} & \left\{b - \frac{1}{4}(p+1)\right\}\frac{p}{b} + p\ln b \\ & - \frac{1}{2}\sum_{i=1}^{p}\sum_{j=i}^{p}\left[\left\{\frac{\hat{y}_i\hat{y}_j}{a} + \frac{(1-\hat{y}_i)(1-\hat{y}_j)}{b-a}\right\}^{-1}\right. \\ & \times\left\{\left(\frac{\hat{y}_i}{a} - \frac{1-\hat{y}_i}{b-a}\right)\frac{\partial\hat{y}_j}{\partial s}\right. \\ & \left.\left. + \left(\frac{\hat{y}_j}{a} - \frac{1-\hat{y}_j}{b-a}\right)\frac{\partial\hat{y}_i}{\partial s} - \frac{\hat{y}_i\hat{y}_j}{a^2}\right\}\right] \\ & + \sum_{i=1}^{p}\left\{\ln\left(\frac{\hat{y}_i}{a}\right) - 1 + \left(\frac{a}{\hat{y}_i} - \frac{b-a}{1-\hat{y}_i} + x_i\right)\frac{\partial\hat{y}_i}{\partial s}\right\},\end{aligned}$$

with

$$\begin{aligned}\partial\hat{y}_i/\partial s = {} & 2\{b - x_i + \sqrt{q_i} \\ & - a - a(x_i + b)/\sqrt{q_i}\}/(b - x_i + \sqrt{q_i})^2\end{aligned}$$

and $q_i = (x - b)^2 + 4ax_i$. The computation of $\hat{K}''(s)$ is best performed using a numerical derivative of $\hat{K}'(s)$.

References

Anderson T.W. 1984. An introduction to multivariate analysis. Second Edition. Wiley, New York.

Booth J.G., Butler R.W., Huzurbazar S., and Wood A.T.A. 1995. Saddlepoint approximations for p-values of some tests of covariance matrices. J. Statist. Comput. Simul. 53: 165–180.

Box G.E.P. 1949. A general distribution theory for a class of likelihood ratio criteria. Biometrika 36: 317–346.

Butler R.W., Booth J.G., and Huzurbazar S. 1993. Saddlepoint approximations for tests of block independence, sphericity, and equal variances and covariances. J.R. Statist. Soc. B 55: 171–184.

Butler R.W., Huzurbazar S., and Booth, J.G. 1992. Saddlepoint approximations for the generalized variance and Wilks' statistic. Biometrika 79: 157–169.

Butler R.W. and Wood A.T.A. 2002. Laplace approximations for hypergeometric functions of matrix argument. Ann. Statist 30: 1155–1177.

Butler R.W. and Wood A.T.A. 2004. A dimensional CLT for Wilks' lambda in multivariate analysis. Scandinavian Journal of Statistics 31: 585–601.

Constantine A.G. 1963. Some noncentral distribution problems in multivariate analysis. Ann. Math. Statist. 34: 1270–1285.

Herz C.S. 1955. Bessel functions of matrix argument. Ann. Math. 61: 474–523.

Jensen J.L. 1991. A large deviation type approximation for the 'Box class' of likelihood ratio criteria. J. Amer. Statist. Assoc. 86: 437–440.

Jensen J.L. 1995. Saddlepoint Approximations. Clarendon Press, Oxford.

James A.T. 1960. The distribution of the latent roots of the covariance matrix. Ann. Math. Statist. 31: 151–158.

James A.T. 1961. Zonal polynomials of the real positive definite symmetric matrices. Ann. Math. 74: 456–469.

Kulp R.W. and Nagarsenkar B.N. 1984. An asymptotic expansion of the nonnull distribution of Wilks criterion for testing the multivariate linear hypothesis Ann. Statist. 12: 1576–1583.

Lee Y.-S. 1971. Distribution of the canonical correlations and asymptotic expansions for distributions of certain independence test statistics. Ann. Math. Statist. 42: 526–537.

Lugananni R. and Rice, S. 1980. Saddlepoint approximations for the distribution of the sum of independent random variables. Adv. Appl. Probab 12: 475–490.

Muirhead R.J. 1972. On the test of independence between two sets of variates. Ann. Math. Statist. 43: 1491–1497.

Muirhead R.J. 1982. Aspects of Multivariate Statistical Theory. John Wiley, New York.

Muller K.E., LaVange L.M., Ramey S.L., and Ramey C.T. 1992. Power calculations for general linear multivariate models including repeated measures applications. J. Amer. Statist. Assoc. 87: 1209–1226.

Muller K.E. and Peterson B.L. 1984. Practical methods for computing power in testing the multivariate general linear hypothesis. Computational Statistics and Data Analysis 2: 143–158.

Srivastava M.S. and Yao W.K. 1989. Saddlepoint method for obtaining tail probability of Wilks' likelihood ratio test. J. Mult. Anal. 31: 117–126.

Sugiura N. 1969. Asymptotic non-null distribution of the likelihood ratio criteria for covariance matrix under local alternatives. Mimeo Ser. 609, Institute of Statistics, University of North Carolina, Chapel Hill.

Sugiura N. 1973a. Further asymptotic formulas for the non-null distributions of three statistics for multivariate linear hypothesis. Ann. Inst. Statist. Math. 25: 153–163.

Sugiura N. 1973b. Asymptotic non-null distributions of the likelihood ratio criteria for covariance matrix under local alternatives. Ann. Statist. 1: 718–728.

Sugiura N. 1974. Asymptotic formulas for the hypergeometric function ${}_2F_1$ of matrix argument, useful in multivariate analysis. Ann. Inst. Statist. Math. 26: 117–125.

Sugiura N. and Fujikoshi Y. 1969. Asymptotic expansion of the non-null distributions of the likelihood ratio criteria for multivariate linear hypothesis and independence. Ann. Math. Statist. 40: 942–952.