COMMUN. STATIST.-THEOR. METH., 14(6), 1405-1411 (1985)

ELLIPSOIDAL CONFIDENCE REGIONS FOR A NORMAL COVARIANCE MATRIX

Doug Wiens
Department of Mathematics, Statistics
and Computing Science
Dalhousie University
Halifax, Nova Scotia
Canada B3H 4H8

Keywords and phrases: confidence ellipsoids; covariance matrices; simultaneous confidence intervals.

ABSTRACT

We obtain an asymptotic expansion of the confidence coefficient for an ellipsoidal confidence region on the elements of a normal covariance matrix. This leads to simultaneous confidence intervals on all linear functions of the elements of this matrix, which are compared with those of Roy (1954).

1. INTRODUCTION

Using results of Tyler (1982), Wiens (1983) exhibited confidence ellipsoids for covariance matrices of elliptically symmetric populations, and for their inverses. For a normally distributed population, the ellipsoid assumes a particularly simple form.

Theorem 1: Let $\underline{y}_1,\ldots,\underline{y}_n$ be i.i.d. $N_p(\underline{\mu},\Sigma)$ random vectors. Put $N = n-1$, $q = p(p+1)/2$, $c_{p,N} = (p+1)(N-p)(N-p+3)/((N+1)(N(p+1) - (p+1)^2 + 2))$, $V_n = \sum_{i=1}^{n} (\underline{y}_i-\overline{\underline{y}})(\underline{y}_i-\overline{\underline{y}})'/(N-p-1)$. Define

 0361-0926/85/1406-1405$3.50/0

$$Y_n = nc_{p,N}(\mathrm{vec}(\Sigma-V_n))'(V_n^{-1}\otimes V_n^{-1})\mathrm{vec}(\Sigma-V_n),\ G_n(y) = P(Y_n\le y).$$

Then $Y_n \sim \frac{N+1}{2} c_{p,N}\ \mathrm{tr}((\frac{U_N}{N-p-1})^{-1} - I_p)^2$, where $U_N \sim W_p(I,N)$. A $100(1-\alpha)\%$ confidence region for Σ is the q-dimensional ellipsoid $\{\Sigma | Y_n \le G_n^{-1}(1-\alpha)\}$. Furthermore, $E[Y_n] = q$ and $Y_n \xrightarrow{L} \chi_q^2$.

Proof: See Wiens (1983).

In this paper, an asymptotic expansion of the distribution of Y_n is given up to $O(n^{-2})$. In Section 2, we will prove

Theorem 2: With $P_q = P(\chi_q^2 \le y)$, $G_n(y)$ is given by

$$P_q + \frac{p}{n}\{\frac{4p^2+9p+7}{3}P_{q+6} - \frac{22p^2+47p+31}{8}P_{q+4} + \frac{6p^2+17p+9}{4}P_{q+2}$$
$$- \frac{2p^2+33p+17}{24}P_q\} + O(n^{-2}) .$$

The convergence of G_n to the χ_q^2 d.f. is quite slow, and the χ^2-approximation alone is inadequate for practical purposes. Using the methods of Hill and Davis (1968), we find

$$G_n^{-1}(1-\alpha) = \chi^2_{q;1-\alpha} + k_{p;\alpha}/n + O(n^{-2}) ,$$

where, with $\chi^2 = \chi^2_{q;1-\alpha}$,

$$k_{p;\alpha} = 2p\chi^2[\frac{(4p^2+9p+7)}{3q(q+2)(q+4)}\{(\chi^2)^2 + (q+4)\chi^2 + (q+2)(q+4)\}$$
$$- \frac{22p^2+47p+31)}{8q(q+2)}(\chi^2+q+2) + \frac{(6p+17p+9)}{4q}] .$$

Some values of $k_{p;\alpha}$ are given in the tables of Section 3, where simultaneous confidence intervals on all linear functions of vec Σ are exhibited, and compared with those of Roy (1954).

Nagao (1973) proposed

$$T_1 = N(\mathrm{vec}(\Sigma_0^{-1}-V_n^{-1}))'(V_n\otimes V_n)\mathrm{vec}(\Sigma_0^{-1} - V_n^{-1}) = \frac{N}{2}\mathrm{tr}(V_n\Sigma_0^{-1}-I_p)^2$$
$$\sim \frac{N}{2}\ \mathrm{tr}(\frac{U_N}{N} - I_p)^2$$

as a test statistic for the hypothesis that $\Sigma = \Sigma_0$, and obtained an expansion similar to that above for the d.f. of T_1. Since the methods used here are similar to those used by Nagao, the proof of Theorem 2 is only outlined.

2. PROOF OF THEOREM 2

Put $k = N-p-1$, $Z = (\frac{k}{2})^{1/2} \log \frac{U_N}{k}$, so that $(\frac{U_N}{N-p-1})^{-1} = \exp(-(\frac{2}{k})^{1/2}Z)$

and

$$Y_n = \frac{(p+1)(k+1)(k+4)}{2[(p+1)k+2]} \operatorname{tr} ((\frac{U_N}{k})^{-1} - I_p)^2$$
$$= \operatorname{tr}Z^2 - (\frac{2}{k})^{\frac{1}{2}} \operatorname{tr}Z^3 + \frac{1}{k}[\frac{7}{6} \operatorname{tr}Z^4 + \frac{5p+3}{p+1} \operatorname{tr}Z^2] + O(k^{-3/2}) \quad . \tag{1}$$

As at (2.4) of Nagao (1973), the density of Z has the asymptotic expansion

$$g(Z) = c_1 \operatorname{etr}\{(\frac{k}{2}+1)(\frac{2}{k})^{1/2}Z - (\frac{k}{2})e^{(\frac{2}{k})^{1/2}Z}\} \cdot$$
$$[1+\frac{p-1}{2}(\frac{2}{k})^{1/2} \operatorname{tr}Z+\frac{1}{12k}\{(3p^2-6p+2)\operatorname{tr}^2Z+p\operatorname{tr}Z^2\} + O(k^{-3/2})] , \tag{2}$$

where $c_1 = (\frac{k}{2})^{p(2k+p+1)/4}/\,\Gamma_p(\frac{N}{2})$. Combining (1) and (2) gives an expression for $e^{isY_n}g(Z)$. Then expanding e^{isY_n} and $(\frac{k}{2})e^{(k/2)^{1/2}Z}$ gives

$$e^{isY_n}g(Z) = c_1e^{-kp/2 - \frac{1}{2}(1-2is)\operatorname{tr}Z^2} \cdot\exp\{(\frac{2}{k})^{1/2}a_1+\frac{b_1}{k}+O(k^{-3/2})\} \cdot$$
$$\cdot[1+(\frac{2}{k})^{1/2}a_2+\frac{b_2}{k} + O(k^{-3/2})] , \tag{3}$$

where $a_1 = -is\operatorname{tr}Z^3 + \operatorname{tr}Z - \frac{1}{6}\operatorname{tr}Z^3$, $b_1 = is[\frac{7}{6}\operatorname{tr}Z^4+\frac{5p+3}{p+1} \operatorname{tr}Z^3] - \frac{1}{12}\operatorname{tr}Z^4$, $a_2=\frac{p-1}{2}\operatorname{tr}Z$, $b_2= \frac{1}{12} [(3p^2-6p+2)\operatorname{tr}^2Z+p\operatorname{tr}Z^2]$. Expanding $\exp\{\cdot\}$ in (3) then gives, as the characteristic function $\psi_n(s)$ of Y_n,

$$\psi_n(s)=c_1e^{-kp/2} \int_{Z=Z'} e^{-\frac{1}{2}(1-2is)\operatorname{tr}Z^2} [1+A(\frac{2}{k})^{1/2}+\frac{B}{k}+O(k^{-3/2})]dZ, \tag{4}$$

where

$$A = A_1 trZ + A_2 trZ^3 \ , \ B = B_1 trZtrZ^3 + B_2 tr^2 Z + B_3 trZ^2 + B_4 trZ^4 + B_5 tr^2 Z^3 ,$$

$$A_1 = \frac{p+1}{2}, \ A_2 = -(is+\frac{1}{6}), \ B_1 = -(p+1)(is+\frac{1}{6}), \ B_2 = \frac{3p^2+6p+2}{12} ,$$

$$B_3 = \frac{5p+3}{p+1} is + \frac{p}{12} \ , \ B_4 = \frac{7}{6} is - \frac{1}{12}, \ B_5 = (is+\frac{1}{6})^2 .$$

Now let $\mathrm{vec}_s Z$ be the $q\times 1$ vector formed from those elements of Z on and below the main diagonal, ordered anti-lexicographically.

Define $D : q\times q$ and $\underline{y} = q\times 1$ by

$$D = \mathrm{diag}(1,2,\ldots,2;\ldots;1,2,;1); \quad \underline{y} = (1-2is)^{1/2} D^{1/2} \mathrm{vec}_s Z .$$

Then (4) becomes

$$\psi_n(s) = c \int_{R^q} \frac{e^{-(\underline{y}'\underline{y})/2}}{(2\pi)^{q/2}} [1 + A(\frac{2}{k})^{1/2} + \frac{B}{k} + O(k^{-3/2})] d\underline{y} , \tag{5}$$

where

$$c = (2\pi)^{p/2}(1-2is)^{-q/2} e^{-kp/2} (\frac{k}{2})^{p(2k+p+1)/4} / \prod_{\alpha=1}^{p} \Gamma(\frac{k+\alpha+1}{2})$$

$$= (1-2is)^{-q/2}[1 - \frac{p}{24k}(2p^2+3p-1) + O(k^{-2})] . \tag{6}$$

We may thus treat $\underline{y}$ as a $N_q(\underline{0},I)$ vector. With respect to this distribution we have, by symmetry, $E[A] = 0$. Also

$$E[trZtrZ^3] = \frac{3}{2}p(p+1)(1-2is)^{-2} \ , \ E[tr^2 Z] = p(1-2is)^{-1} ,$$

$$E[trZ^2] = \frac{p(p+1)}{2}(1-2is)^{-2} ,$$

$$E[trZ^4] = \frac{p}{4}(2p^2+5p+5)(1-2is)^{-2} \ , \ E[tr^2 Z^3] = \frac{3p}{4}(4p^2+9p+7) \cdot$$

$$\cdot (1-2is)^{-3} .$$

Substituting these expectations, and (6), into (5) and inverting $\psi_n(s)$ then completes the proof.

3. SIMULTANEOUS CONFIDENCE INTERVALS

Put $a = N/(N-p-1)$, $b = (G_n^{-1}(1-\alpha)/nc_{p,N})^{1/2}$, so that the unbiased sample covariance matrix is $S = a^{-1}V_n$, and the level $1-\alpha$ confidence region of Theorem 1 becomes

$$\{\Sigma \mid (\mathrm{vec}(\Sigma - aS))'(S^{-1} \otimes S^{-1})\mathrm{vec}(\Sigma - aS) \le (ab)^2\} .$$

Applying Scheffe's (1959) method, we find that simultaneous confidence intervals on all linear functions of $\mathrm{vec}\Sigma$ are given by

$$1-\alpha = P\{a(\mathrm{tr}MS - b(\mathrm{tr}(MS)^2)^{1/2}) \le \mathrm{tr}M\Sigma \le a(\mathrm{tr}MS + b(\mathrm{tr}(MS)^2)^{1/2}$$

for all symmetric M } .

Putting $M = \underline{m}\,\underline{m}'$ gives the intervals

$$a(1-b)\underline{m}'S\underline{m} \le \underline{m}'\Sigma\underline{m} \le a(1+b)\underline{m}'S\underline{m} \ . \tag{7}$$

Choosing M to have 1's in the $(i,j)^{th}$ and $(j,i)^{th}$ positions, zereos elsewhere, gives

$$1-\alpha \ge P\{a(s_{ij} - b(\frac{s_{ii}s_{jj}(1+r_{ij}^2)}{2})^{1/2}) \le \sigma_{ij} \le a(s_{ij} + b(\frac{s_{ii}s_{jj}(1+r_{ij}^2)}{2})^{1/2}) \ ; \ \text{all} \ i,j\} \ , \tag{8}$$

where r_{ij} is the sample correlation coefficient. In (7), choose $\underline{m}$ to have $\sigma_{ii}^{-1/2}$ and $\sigma_{jj}^{-1/2}$ in the i^{th} and j^{th} positions, zeroes elsewhere; combine with (8), and assume that n is large enough that $b < 1$. Then simultaneous confidence intervals on the population correlation coefficients, still at a combined level exceeding $1-\alpha$, are

$$\frac{-2b}{1+b}(1+r_{ij}^{+}) + r_{ij} \le \rho_{ij} \le \frac{2b}{1-b}(1+r_{ij}^{+}) + r_{ij} \ ,$$

where

$$r_{ij}^{+} = \max(0, r_{ij}) \ .$$

Corresponding to (7), Roy (1954) gave the intervals

$$1-\alpha = P\{\underline{m}'S\underline{m}/u \le \underline{m}'\Sigma\underline{m} \le \underline{m}'S\underline{m}/\ell \ ; \ \text{all} \ \underline{m}\} \tag{9}$$

where $\ell < u$ are such that $[\ell,u]$ contains all roots of $\Sigma^{-1}S$ with probability $(1-\alpha)$. Using (9), Anderson (1965) obtained

$$1-\alpha \ge P\{\frac{(\ell^{-1}+u^{-1})s_{ij} - (\ell^{-1}-u^{-1})(s_{ii}s_{jj})^{1/2}}{2} \le \sigma_{ij} \le \frac{(\ell^{-1}+u^{-1})s_{ij} + (\ell^{-1}-u^{-1})(s_{ii}s_{jj})^{1/2}}{2} \quad \text{all} \ i,j\} \ . \tag{10}$$

Put $R(N,p,\alpha) = 2ab/(\ell^{-1}-u^{-1})$. Then the intervals in (7) are shorter than those in (9) if $R < 1$; those in (8) are shorter than those in (10) if $R < ((1+r_{ij}^2)/2)^{-1/2} \in [1,\sqrt{2}]$. Tables I-III below give some comparative values. We have approximated $G_n^{-1}(1-\alpha)$ by $\chi^2_{q;1-\alpha} + k_{p;\alpha}/n$. For $p=2$, the values of ℓ and u were obtained from Thompson (1962), for $p=4$ and 6 they were obtained from Pearson and Hartley (1976).

For some pairs (p,N) , the intervals in (7) and (8) are uniformly shorter. For others, $1 < R < \sqrt{2}$, so that there will be values r_{ij} for which each method gives shorter intervals. An asterisk (*) indicates a combination for which $b > 1$.

TABLE I: $p=2$, $\alpha=.05$, $\chi^2_{3;.95}=7.815$, $k_{2;.05}=76.9603$

N	2ab	$\ell^{-1}-u^{-1}$	R(N,2,.05)
20	1.723	2.285	.754
40	1.047	1.360	.770
60	.810	1.050	.771
80	.682	.885	.770
100	.600	.779	.770

TABLE II: $p=4$, $\alpha=.10$, $\chi^2_{10;.9}=15.9872$, $k_{4;.1}=182.011$

N	2ab	$\ell^{-1}-u^{-1}$	R(N,4,.1)
*20	2.980	3.479	.857
40	1.640	1.836	.894
60	1.231	1.369	.899
80	1.021	1.135	.900
100	.890	.990	.899

TABLE III: $p=6$, $\alpha=.10$, $\chi^2_{21;.9}=29.615$, $k_{6;.1}=445.458$

N	2ab	$\ell^{-1}-u^{-1}$	R(N,6,.1)
*20	5.183	5.646	.918
*40	2.509	2.527	.993
60	1.810	1.785	1.014
80	1.473	1.448	1.017
100	1.269	1.247	1.017

ACKNOWLEDGEMENTS

This research was supported by the Natural Sciences and Engineering Research Council of Canada, under Grant No. A8603. We also thank a referee for his helpful criticism of an earlier version of this paper.

BIBLIOGRAPHY

Anderson, T.W. (1965). Some optimum confidence bounds for roots of determinantal equations. Ann. Math. Statist. (36), 468-488.

Hill, G.W., and Davis, A.W. (1968). Generalized asymptotic expansions of Cornish-Fisher type. Ann. Math. Statist. (39), 1264-1273.

Nagao, H. (1973). On some test criteria for covariance matrix. Ann. Statist. (1), 700-709.

Pearson, E.S. and Hartley, H.O. (1976). Biometrika Tables for Statisticians II, Biometrika Trust.

Roy, S.N. (1954). Some further results in simultaneous confidence interval estimation. Ann. Math. Statist. (25), 752-761.

Scheffé, H. (1959). The Analysis of Variance. J. Wiley and Sons, New York.

Thompson, W.A., Jr., (1962). Estimation of dispersion parameters. Journ. Res. Nat. Bur. Standards. (66B), 161-164.

Tyler, D.E. (1982). Radial estimates and the test for sphericity. Biometrika (69), 429-436.

Wiens, D. (1983). On some pattern reduction matrices which appear in statistics. To appear in Linear Algebra and its Applications.

Received by Editorial Board member December, 1984; Revised June and November, 1984 and Retyped March, 1985.

Recommended by R. P. Gupta, Dalhousie University, Halifax, Canada.

Refereed Anonymously.