Adv. & Appl. in Stat. 3(3) (2003), 285-292

TESTING MULTISAMPLE COMPOUND SYMMETRY

DAYA K. NAGAR and MARÍA EUGENIA CASTAÑEDA

Departamento de Matemáticas, Universidad de Antioquia
Medellín, A. A. 1226, Colombia

Abstract

This paper is concerned with testing multisample compound symmetry of k multivariate Gaussian models. The likelihood ratio test statistic and its null moments for testing multisample compound symmetry have been derived. The asymptotic null distribution of the test statistic has been obtained using Box's method. This test can be viewed as an extension of Wilks' test for testing intraclass correlation structure of a covariance matrix.

1. Introduction

Let $\prod$ be a normal m-variate population with mean vector $\boldsymbol{\mu}$ and positive definite covariance matrix $\Sigma = (\sigma_{ij})$. Let H_{vc} be the hypothesis that the variances are equal and the covariances are equal, i.e., $H_{vc} : \sigma_{11} = \cdots = \sigma_{mm} = \sigma_a;\ \sigma_{12} = \cdots = \sigma_{1m} = \cdots = \sigma_{m-1,m} = \sigma_b$, where σ_a and σ_b are unknown constants with suitable restrictions. A covariance matrix for which H_{vc} is true is said to have *compound symmetry*. Compound symmetry is often found in repeated measures designs which are not time dependent (Winer [8]). If the order of the responses does not matter, then the covariances are exchangeable. The

---

2000 Mathematics Subject Classification: Primary 62H40; Secondary 62H10.

Key words and phrases: compound symmetry, distribution, likelihood ratio test, intraclass correlation, null moments, Box's approximation.

Received June 15, 2003

term exchangeable is synonymous with compound symmetry. The problem of testing H_{vc} was first considered by Wilks [7] who derived the likelihood ratio test statistic Λ_{vc}, its null moments and distributional results for small values of m. The problem of testing H_{vc} plays a very useful role in areas like medical research and psychometrics. Such models also arise in the study of familial data.

Let $\Pi_1, \ldots, \Pi_k$ be k independent m-variate normal populations with mean vectors $\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_k$ and positive definite covariance matrices $\Sigma_1, \ldots, \Sigma_k$, respectively. The hypothesis of multisample compound symmetry can be stated as

$$H_{vc(k)} : \Sigma_1 = \cdots = \Sigma_k = \sigma^2[(1-\rho)I_m + \rho J], \tag{1.1}$$

where $\sigma^2 > 0$, $\rho(-1/(m-1) < \rho < 1)$ are unknown constants, I_m is an identity matrix of order m and J is an $m \times m$ matrix having all its elements unity. Such a hypothesis arises in repeated measures designs with two or more repeated factors where if the assumption of homogeneity of group covariance matrices cannot be made a priori it needs to be tested.

In this article, we will derive the modified likelihood ratio test statistic Λ^* for testing $H_{vc(k)}$, its null moments and asymptotic expansion of the distribution function of a constant multiple of $-2\ln\Lambda^*$.

2. The Likelihood Ratio Test Statistic

Let $\boldsymbol{X}_{gj}$, $j = 1, \ldots, N_g$ be a random sample from $N_m(\boldsymbol{\mu}_g, \Sigma_g)$, where $\boldsymbol{\mu}_g$ and Σ_g are unknown, $g = 1, \ldots, k$. Let $\overline{\boldsymbol{X}}_g$ and A_g be, respectively, the mean vector and the matrix of sum of squares and products from the gth sample, i.e., $N_g\overline{\boldsymbol{X}}_g = \sum_{j=1}^{N_g}\boldsymbol{X}_{gj}$, $A_g = \sum_{j=1}^{N_g}(\boldsymbol{X}_{gj} - \overline{\boldsymbol{X}}_g)(\boldsymbol{X}_{gj} - \overline{\boldsymbol{X}}_g)'$, $g = 1, \ldots, k$ and put $A = \sum_{g=1}^{k} A_g$ and $N_0 = \sum_{g=1}^{k} N_g$. The likelihood ratio test statistic Λ for testing $H_{vc(k)}$ can be derived as

$$\Lambda = \frac{(m-1)^{N_0(m-1)/2}(mN_0)^{N_0 m/2}}{\prod_{g=1}^{k} N_g^{N_g m/2}} \frac{\prod_{g=1}^{k} \det(A_g)^{N_g/2}}{\{\mathrm{tr}(JA)\}^{N_0/2}[\mathrm{tr}\{(mI_m - J)A\}]^{N_0(m-1)/2}} \tag{2.1}$$

The null hypothesis $H_{vc(k)}$ is rejected if $\Lambda \leq \Lambda_0$, where Λ_0 is determined by the null distribution and the level of significance. Using modification suggested by Bartlett, the modified likelihood ratio test statistic Λ^* for testing $H_{vc(k)}$ is obtained as

$$\Lambda^* = \frac{(m-1)^{n_0(m-1)/2}(mn_0)^{n_0 m/2}}{\prod_{g=1}^{k} n_g^{n_g m/2}} \frac{\prod_{g=1}^{k} \det(A_g)^{n_g/2}}{\{\mathrm{tr}(JA)\}^{n_0/2}[\mathrm{tr}\{(mI_m - J)A\}]^{n_0(m-1)/2}}, \tag{2.2}$$

where $n_g = N_g - 1$ and $n_0 = \sum_{g=1}^{k} n_g = N_0 - k$. Note that for $k = 1$, the hypothesis of multisample symmetry reduces to the Wilks' H_{vc} hypothesis for testing intraclass-correlation structure of a covariance matrix.

3. The Null Moments

Since $A_1, \ldots, A_k$ are independent Wishart matrices (Anderson [1], Gupta and Nagar [3]) and under $H_{vc(k)}$, $A_g \sim W_m(n_g, \Sigma)$, $\Sigma = \sigma^2[(1-\rho)I_m + \rho J]$, we obtain the hth null moment of Λ^* by integrating over Wishart densities. That is,

$$E(\Lambda^{*h}) = \frac{(m-1)^{n_0(m-1)h/2}(mn_0)^{n_0 mh/2}}{\prod_{g=1}^{k} n_g^{n_g mh/2}} \frac{1}{\det(2\Sigma)^{n_0/2}\prod_{g=1}^{k}\Gamma_m(n_g/2)}$$

$$\times \int_{A_1>0} \cdots \int_{A_k>0} \frac{\prod_{g=1}^{k} \det(A_g)^{n_g h/2}}{\{\mathrm{tr}(JA)\}^{n_0 h/2}[\mathrm{tr}\{(mI_m - J)A\}]^{n_0(m-1)h/2}}$$

$$\times \prod_{g=1}^{k} \left\{ \det(A_g)^{(n_g-m-1)/2} \operatorname{etr}\left(-\frac{1}{2}\Sigma^{-1} A_g\right)\right\} dA_1 \cdots dA_k$$

$$= \frac{(m-1)^{n_0(m-1)h/2}(mn_0)^{n_0mh/2}}{\prod_{g=1}^{k} n_g^{n_gmh/2}} \det(2\,\Sigma)^{n_0h/2} \prod_{g=1}^{k} \frac{\Gamma_m[n_g(1+h)/2]}{\Gamma_m(n_g/2)}$$

$$\times E[\{\operatorname{tr}(JA)\}^{-n_0h/2}\{\operatorname{tr}(mI_m - J)A\}^{-n_0(m-1)h/2}], \tag{3.1}$$

where now A_g have independent $W_m(n_g(1+h), \Sigma)$, $g = 1, \ldots, k$. Hence the random matrix $A = \sum_{g=1}^{k} A_g$ has $W_m(n_0(1+h), \Sigma)$ and

$$E[\{\operatorname{tr}(JA)\}^{-n_0h/2}\{\operatorname{tr}(mI_m - J)A\}^{-n_0(m-1)h/2}]$$

$$= \frac{1}{\det(2\,\Sigma)^{n_0(1+h)/2}\Gamma_m[n_0(1+h)/2]}$$

$$\times \int_{A>0} \{\operatorname{tr}(JA)\}^{-n_0h/2}\{\operatorname{tr}(mI_m - J)A\}^{-n_0(m-1)h/2}$$

$$\times \det(A)^{(n_0+n_0h-m-1)/2} \operatorname{etr}\left(-\frac{1}{2}\Sigma^{-1} A\right) dA. \tag{3.2}$$

Replacing $\{\operatorname{tr}(JA)\}^{-n_0h/2}$ and $[\{\operatorname{tr}(mI_m - J)A\}]^{-n_0(m-1)h/2}$ by their equivalent gamma integrals, namely

$$\{\operatorname{tr}(JA)\}^{-n_0h/2} = \frac{1}{2^{n_0h/2}\Gamma(n_0h/2)} \int_0^{\infty} x^{n_0h/2-1} \exp\left\{-\frac{1}{2} x \operatorname{tr}(JA)\right\} dx,$$

$$\operatorname{Re}(h) > 0,$$

and

$$[\operatorname{tr}\{(mI_m - J)A\}]^{-n_0(m-1)h/2} = \frac{1}{2^{n_0(m-1)h/2}\Gamma[n_0(m-1)h/2]}$$

$$\times \int_0^{\infty} y^{n_0(m-1)h/2-1}$$

$$\times \exp\left[-\frac{1}{2} y \operatorname{tr}\{(mI_m - J)A\}\right] dy, \ \operatorname{Re}(h) > 0,$$

respectively, where Re(·) denotes the real part of (·) in (3.2) and integrating out A using multivariate gamma integral (Gupta and Nagar [3, p. 18]) we have

$$E[\{\text{tr}(JA)\}^{-n_0h/2}\{\text{tr}(mI_m - J)A\}^{-n_0(m-1)h/2}]$$

$$= \left\{2^{n_0mh/2}\Gamma\left(\frac{1}{2}n_0h\right)\Gamma\left[\frac{1}{2}n_0(m-1)h\right]\right\}^{-1}\left\{\det(2\Sigma)^{n_0(1+h)/2}\Gamma_m\left[\frac{1}{2}n_0(1+h)\right]\right\}^{-1}$$

$$\times \int_0^\infty x^{n_0h/2-1}\int_0^\infty y^{n_0(m-1)h/2-1}\int_{A>0}\det(A)^{(n_0+n_0h-m-1)/2}$$

$$\times \text{etr}\left[-\frac{1}{2}\{\Sigma^{-1} + xJ + y(mI_m - J)\}A\right]dA\,dx\,dy$$

$$= \left\{2^{n_0mh/2}\det(\Sigma)^{n_0(1+h)/2}\Gamma\left(\frac{1}{2}n_0h\right)\Gamma\left[\frac{1}{2}n_0(m-1)h\right]\right\}^{-1}\int_0^\infty x^{n_0h/2-1}$$

$$\times \int_0^\infty y^{n_0(m-1)h/2-1}\det(\Sigma^{-1} + xJ + y(mI_m - J))^{-n_0(1+h)/2}dx\,dy. \tag{3.3}$$

Now, using the results $\det(\Sigma^{-1} + xJ + y(mI_m - J)) = (a^{-1} + xm)(b^{-1} + ym)^{m-1}$, where $a = \sigma^2[1 + (m-1)\rho]$ and $b = \sigma^2(1-\rho)$,

$$\int_0^\infty x^{n_0h/2-1}(a^{-1} + xm)^{-n_0(1+h)/2}dx = \frac{a^{n_0/2}}{m^{n_0h/2}}\frac{\Gamma(n_0/2)\Gamma(n_0h/2)}{\Gamma[n_0(1+h)/2]}$$

and

$$\int_0^\infty y^{n_0(m-1)h/2-1}(b^{-1} + ym)^{-n_0(m-1)(1+h)/2}dy$$

$$= \frac{b^{n_0(m-1)/2}}{m^{n_0(m-1)h/2}}\frac{\Gamma[n_0(m-1)/2]\Gamma[n_0(m-1)h/2]}{\Gamma[n_0(m-1)(1+h)/2]}$$

the expression (3.3) is further simplified as

$$E[\{\text{tr}(JA)\}^{-n_0h/2}\{\text{tr}(mI_m - J)A\}^{-n_0(m-1)h/2}]$$

$$= \det(2m\,\Sigma)^{-n_0h/2}\frac{\Gamma(n_0/2)\Gamma[n_0(m-1)/2]}{\Gamma[n_0(1+h)/2]\Gamma[n_0(m-1)(1+h)/2]}. \tag{3.4}$$

Finally, substituting (3.4) in (3.1) and simplifying the resulting expression, we obtain

$$E(\Lambda^{*h}) = \frac{(m-1)^{n_0(m-1)h/2} n_0^{n_0 mh/2}}{\prod_{g=1}^{k} n_g^{n_g ph/2}}$$

$$\times \frac{\Gamma(n_0/2)\Gamma[n_0(m-1)/2]}{\Gamma[n_0(1+h)/2]\Gamma[n_0(m-1)(1+h)/2]} \prod_{g=1}^{k} \frac{\Gamma_m[n_g(1+h)/2]}{\Gamma_m(n_g/2)}, \quad (3.5)$$

where $\operatorname{Re}[n_g(1+h)] > m-1$, $g = 1, \ldots, k$. Further simplification of $E(\Lambda^{*h})$ can be achieved by using

$$\Gamma_m(a) = \pi^{m(m-1)/4} \prod_{j=1}^{m} \Gamma\left(a - \frac{j-1}{2}\right), \ \operatorname{Re}(a) > \frac{m-1}{2}.$$

Thus we get

$$E(\Lambda^{*h}) = \frac{(m-1)^{n_0(m-1)h/2} n_0^{n_0 mh/2}}{\prod_{g=1}^{k} n_g^{n_g mh/2}} \frac{\Gamma(n_0/2)\Gamma[n_0(m-1)/2]}{\Gamma[n_0(1+h)/2]\Gamma[n_0(m-1)(1+h)/2]}$$

$$\times \prod_{g=1}^{k} \prod_{j=1}^{m} \frac{\Gamma[n_g(1+h)/2 - (j-1)/2]}{\Gamma[(n_g - j + 1)/2]}, \quad (3.6)$$

where $\operatorname{Re}[n_g(1+h)] > m-1$, $g = 1, \ldots, k$. Substituting $k = 1$, $n_1 = n = N-1$ in (3.6) and simplifying, we get the hth moment of modified Wilks' Λ_{vc}^{*} statistic as

$$E(\Lambda_{vc}^{*h}) = \frac{(m-1)^{n(m-1)h/2} \rho[(m-1)n/2]}{\rho[(m-1)n(1+h)/2]} \prod_{j=1}^{m-1} \frac{\Gamma[n(1+h)/2 - j/2]}{\Gamma[(n-j)/2]} \quad (3.7)$$

4. Asymptotic Null Distribution

By using (3.6) and Box's method of expansion (see Anderson [1]), we find the correction factor for $L = -2\ln\Lambda^{*}$ to be

$$c = 1 - \frac{2m}{3(m-1)[km(m+1)-4]n_0}\left[\frac{1}{4}(m-1)(2m^2+3m-1)\sum_{g=1}^{k}\frac{n_0}{n_g} - 1\right]$$

and the asymptotic expansion of the null distribution of $M = cL$ as

$$P[M \leq x]$$

$$= P(\chi_f^2 \leq x) + \frac{\gamma_2}{(cn_0)^2}\{P(\chi_{f+4}^2 \leq x) - P(\chi_f^2 \leq x)\} + O(cn_0)^{-3}, \qquad (4.1)$$

where $f = km(m+1)/2 - 2$ and

$$\gamma_2 = \frac{1}{48}m(m^2-1)(m+2)\sum_{g=1}^{k}\left(\frac{n_0}{n_g}\right)^2$$

$$- \frac{m^2}{18(m-1)^2[km(m+1)-4]}\left[\frac{1}{4}(m-1)(2m^2+3m-1)\sum_{g=1}^{k}\frac{n_0}{n_g} - 1\right]^2.$$

Applying the general inversion expansion of Hill and Davis [4] (also see Sugiura and Nagao [6]) to (4.1), we get the asymptotic formula for the percentage point of M as

$$u + \frac{2\gamma_2}{(cn_0)^2}\sum_{j=1}^{2}\frac{u^j}{f_{(j)}} + O(cn_0)^{-3},$$

where $f_{(j)} = f(f+2)\cdots(f+2j-2)$ and u is the upper percentage point of chi-square distribution with f degrees of freedom. For the exact distribution of Λ^* one may resort to methods given in Gupta and Tang [2] and Nagar and Sánchez [5].

Acknowledgement

This research work was supported by the SIU, Universidad de Antioquia research grant no. IN0011CE.

References

[1] T. W. Anderson, An Introduction to Multivariate Statistical Analysis, 2nd ed., John Wiley & Sons, New York, 1984.

[2] A. K. Gupta and J. Tang, Distribution of likelihood ratio statistic for testing equality of covariance matrices of multivariate Gaussian models, Biometrika 71(3) (1984), 555-559.

[3] A. K. Gupta and D. K. Nagar, Matrix Variate Distributions, Chapman & Hall/CRC, Boca Raton, 2000.

[4] G. W. Hill and A. W. Davis, Generalized asymptotic expansion of Cornish-Fisher type, Ann. Math. Stat. 39 (1968), 150-163.

[5] D. K. Nagar and C. C. Sánchez, Exact percentage points for testing multisample sphericity in the bivariate case, Comm. Stat.-Simul. Comp. (to appear).

[6] N. Sugiura and H. Nagao, On Bartlett's test and Lehmann's test for homogeneity of variances, Ann. Math. Stat. 40(6) (1969), 2018-2032.

[7] S. S. Wilks, Sample criteria for testing equality of means, equality of variances and equality of covariances in a normal multivariate distribution, Ann. Math. Stat. 17 (1946), 257-281.

[8] B. J. Winer, Statistical Principles in Experimental Design, McGraw-Hill, New York, 1971.

■