The Distribution of a Statistic Used for Testing Sphericity of Normal Distributions

Author(s): S. John

Source: *Biometrika*, Apr., 1972, Vol. 59, No. 1 (Apr., 1972), pp. 169-173
Published by: Oxford University Press on behalf of Biometrika Trust
Stable URL: https://www.jstor.org/stable/2334628

Biometrika (1972), **59**, 1, *p.* 169
Printed in Great Britain

The distribution of a statistic used for testing sphericity of normal distributions

By S. JOHN
Australian National University, Canberra

SUMMARY

The joint distribution of the sum of the rth powers ($r = 1, \ldots, p-1$) and the product of all the latent roots of a $p \times p$ Wishart matrix is obtained and used to derive the null distributions of the likelihood ratio test criterion and the locally most powerful invariant test criterion for detecting deviations of the variances and covariances of a p-variate normal distribution from proportionality to specified numbers. The sphericity of the distribution is a special case. Explicit expressions are given for the null distributions in the trivariate case. In the bivariate case the two test criteria coincide and their null distribution has been known. The distribution of the locally most powerful test criterion being complicated for values of p larger than three, some approximations are fitted by the method of moments and compared.

Some key words: Analysis of covariance matrices; Likelihood ratio test; Locally most powerful test; Latent roots; Asymptotic distribution; Beta distribution; Chi-squared approximation; Method of moments; Sphericity of multivariate normal distribution.

1. INTRODUCTION

Mauchly (1940) obtained the likelihood ratio test for testing the hypothesis H_0 that the covariance matrix of a p-variate normal distribution equals $\sigma^2\mathbf{\Sigma}_0$, where $\mathbf{\Sigma}_0$ is a given positive definite matrix and σ^2 is some positive number. The exact null distribution of the likelihood ratio in the case $p = 2$ and an asymptotic expansion of the distribution function for general p are given by Anderson (1958, p. 263).

Only the canonical form of H_0, where $\mathbf{\Sigma}_0 = \mathbf{I}$, needs consideration, for, if $\mathbf{\Sigma}_0 \neq \mathbf{I}$, we may work with $\mathbf{y} = \mathbf{Mx}$, where $\mathbf{M}$ is any $p \times p$ matrix such that $\mathbf{M\Sigma}_0\mathbf{M}' = \mathbf{I}$. If $\mathbf{X}$ denotes the $p \times N$ matrix whose jth column is the jth of N independent observations from the p-variate normal population, the null hypothesis is invariant under the following transformations: (1) the addition of the same vector to each column of $\mathbf{X}$; (2) the premultiplication of $\mathbf{X}$ by a scalar multiple of an orthogonal matrix; (3) the postmultiplication of $\mathbf{X}$ by an orthogonal matrix. Among tests that are invariant under these transformations the locally most powerful one is that which rejects H_0 when $\Sigma T_i^2/(\Sigma T_i)^2 > c$, where $T_1, \ldots, T_p$ are the latent roots of $\mathbf{S} = N\mathbf{V}$ in ascending order, $\mathbf{V}$ being the sample covariance matrix, and where c is a constant such that the test has the required size; see John (1971, § 3). We shall use U to denote $\Sigma T_i^2/(\Sigma T_i)^2$. Alternatively, $U = \operatorname{tr}(\mathbf{S}^2)/\{\operatorname{tr}(\mathbf{S})\}^2$. If $\mathbf{\Sigma}_0 \neq \mathbf{I}$, the T's should be taken to be the latent roots of $\mathbf{\Sigma}_0^{-1}\mathbf{S}$ and U should be taken to be $\operatorname{tr}\{(\mathbf{\Sigma}_0^{-1}\mathbf{S})^2\}/\{\operatorname{tr}(\mathbf{\Sigma}_0^{-1}\mathbf{S})\}^2$. Note that the likelihood ratio

$$R = (|\mathbf{V}|/|\mathbf{\Sigma}_0|)^{\frac{1}{2}N}\{\operatorname{tr}(\mathbf{\Sigma}_0^{-1}\mathbf{V}/p)\}^{\frac{1}{2}Np}.$$

If $p = 2$, R and U are functions of each other and the two tests coincide. The exact null distribution of $(n-1)[\{2(1-U)\}^{-\frac{1}{2}}-1]$ derived from that of R is the F distribution with 2 and $2n-2$ degrees of freedom, where $n = N-1$; see also § 3. If $p > 2$, the two tests differ. Our main objective will be to obtain the null distribution of U for values of p greater than two.

2. The exact distribution

The density of $(T_1, \ldots, T_p)$ was found by Fisher (1939), Hsu (1939) and Roy (1939); at $(t_1, \ldots, t_p)$ it is

$$\frac{\pi^{\frac{1}{2}p}\left\{\prod_{j=1}^{p} t_i^{\frac{1}{2}(n-p-1)} \exp\left(-\frac{1}{2}\sum_{j=1}^{p} t_i\right)\right\} \prod_{j>k}(t_j-t_k)}{2^{\frac{1}{2}pn}\prod_{j=1}^{p}\{\Gamma(\frac{1}{2}n+\frac{1}{2}-\frac{1}{2}j)\,\Gamma(\frac{1}{2}p+\frac{1}{2}-\frac{1}{2}j)\}} \tag{2·1}$$

if $0 \leqslant t_1 \leqslant \ldots \leqslant t_p$, and zero otherwise. Let

$$Z_r = \sum_{j=1}^{p} T_j^r, \quad z_r = \sum_{j=1}^{p} t_j^r \quad (r = 1, \ldots, p-1), \quad Z_p = \prod_{j=1}^{p} T_j, \quad z_p = \prod_{j=1}^{p} t_j.$$

The Jacobian of the transformation from the t's to the z's is

$$\{\Gamma(p)\prod_{j>k}(t_j-t_k)\}^{-1}. \tag{2·2}$$

Multiplying (2·1) by (2·2) and expressing the product in terms of the z's, we get the density of $(Z_1, \ldots, Z_p)$ at $(z_1, \ldots, z_p)$ as

$$\frac{\pi^{\frac{1}{2}p}\,2^{-\frac{1}{2}pn} z_p^{\frac{1}{2}(n-p-1)}\,e^{-\frac{1}{2}z_1}}{\Gamma(p)\prod_{j=1}^{p}\{\Gamma(\frac{1}{2}n+\frac{1}{2}-\frac{1}{2}j)\,\Gamma(\frac{1}{2}p+\frac{1}{2}-\frac{1}{2}j)\}}. \tag{2·3}$$

Integration of (2·3) with respect to z_i $(i = 3, \ldots, p)$ will yield the density of (Z_1, Z_2) at (z_1, z_2), from which the density of (Z_1, U) at (z_1, u) may be obtained by replacing z_2 by uz_1^2 and multiplying the result by the Jacobian of this transformation, i.e. z_1^2. The marginal density of U at u is now obtained by integrating the density of (Z_1, U) at (z_1, u) with respect to z_1. We go from U to $T = (p-1)^{-1}(pU-1)$ for reasons that will become clear in § 3. When $p = 3$, the density of T at t is

$$\frac{2\pi^{\frac{1}{2}}\,\Gamma(\frac{3}{2}n)\{(1-t^{\frac{1}{2}})^{n-2}(1+2t^{\frac{1}{2}})^{\frac{1}{2}n-1}-(1+t^{\frac{1}{2}})^{n-2}(1-2t^{\frac{1}{2}})^{\frac{1}{2}n-1}\}}{3^{\frac{3}{2}n-2}\{\Gamma(\frac{1}{2}n)\}^2\,\Gamma(\frac{1}{2}n-\frac{1}{2})} \tag{2·4}$$

if $t < \frac{1}{4}$ and

$$\frac{2\pi^{\frac{1}{2}}\,\Gamma(\frac{3}{2}n)\,(1-t^{\frac{1}{2}})^{n-2}(1+2t^{\frac{1}{2}})^{\frac{1}{2}n-1}}{3^{\frac{3}{2}n-2}\{\Gamma(\frac{1}{2}n)\}^2\,\Gamma(\frac{1}{2}n-\frac{1}{2})} \tag{2·5}$$

if $t \geqslant \frac{1}{4}$.

It is of some interest to note that the exact density function of $W = p^p Z_1^{-p} Z_p$, which is essentially the likelihood ratio criterion, also can be obtained from (2·3) by an analogous method. If $p = 3$, the density of W at w is

$$\frac{3^{-3n/2+2}\,\pi^{\frac{1}{2}}\Gamma(\frac{3}{2}n)}{2\,\Gamma(\frac{1}{2}n)\,\Gamma(\frac{1}{2}n-\frac{1}{2})\,\Gamma(\frac{1}{2}n-1)}\,w^{\frac{1}{2}(n-3)}\{\psi_2(w)-\psi_1(w)\}\,[w^{\frac{1}{3}}\{\psi_1(w)+\psi_2(w)\}-4],$$

if

$$\psi_1(w) = [\cos\{\tfrac{1}{3}\arccos(-27^{-\frac{1}{2}}w)\}]^{-1}, \quad \psi_2(w) = [\cos\{\tfrac{2}{3}\pi-\tfrac{1}{3}\arccos(-27^{-\frac{1}{2}}w)\}]^{-1}$$

and $0 \leqslant w \leqslant 1$; see also Consul (1967).

3. Some approximations to the distribution of U

The Taylor expansions of $\frac{1}{2}Np(p-1)T$ and $-2\log R$ in powers of the deviations of the elements of **S** from their expected values agree up to terms of order two. Hence the asymptotic distribution of $\frac{1}{2}Np(p-1)T$ must be the same as that of $-2\log R$, namely the chi-squared distribution with $\frac{1}{2}p(p+1)-1$ degrees of freedom. This is the first approximation.

The multiplication of a statistic by a quantity that tends to one as $N \to \infty$ will not alter its asymptotic distribution. By using a multiplier μ that makes the expected value of $\frac{1}{2}\mu Np(p-1)T$ equal to the expected value of a chi-squared with $\frac{1}{2}p(p+1)-1$ degrees of freedom we may expect to get an improved approximation. Such a value for μ is

$$1-(1-2p^{-1})N^{-1}.$$

The expected value of U is obtained in the appendix. Thus the second approximation is to treat $(\frac{1}{2}np+1)(p-1)T$ as a chi-squared with $\frac{1}{2}p(p+1)-1$ degrees of freedom.

Table 1. *Percentage points of T for $p = 2$*

	Upper 5 %		Upper 1 %	
n	Chi-squared approximation	Exact	Chi-squared approximation	Exact
2	1·9972	0·9975	3·0701	0·9999
4	1·1983	0·8643	1·8421	0·9536
6	0·8559	0·6983	1·3158	0·8415
10	0·5447	0·4861	0·8373	0·6406
16	0·3524	0·3293	0·5418	0·4588
25	0·2304	0·2209	0·3542	0·3187

We have seen that the asymptotic distribution of $\frac{1}{2}Np(p-1)T$ is that of a chi-squared distribution. We also note that

$$T = (p-1)^{-1}p\Sigma(T_i-\bar{T})^2/(\Sigma T_i)^2 \geqslant 0,$$
$$T = (p-1)^{-1}\{p-1-p\sum_{i\neq j}\Sigma T_iT_j/(\Sigma T_i)^2\} \leqslant 1.$$

It is thus suggested that the distribution of T may be approximated by the distribution having the beta density function

$$\frac{\Gamma(\frac{1}{2}\nu_1+\frac{1}{2}\nu_2)}{\Gamma(\frac{1}{2}\nu_1)\Gamma(\frac{1}{2}\nu_2)}t^{\frac{1}{2}\nu_1-1}(1-t)^{\frac{1}{2}\nu_2-1} \quad (0 \leqslant t \leqslant 1), \tag{3·1}$$

where ν_1 and ν_2 are so determined that the first two moments of this distribution agree with the corresponding moments of T, easily obtained from the moments of U given in the appendix. The values of ν_1 and ν_2 so determined are respectively $m(p+2)$ and $mp(n-1)$, where $m = \frac{1}{2}(p-1)-\{(n+2)p\}^{-1}(p-3)(p-2)$. Note that approximating the distribution of T by the distribution having density function (3·1) is equivalent to taking the distribution of $\{(1-U)\nu_1p\}^{-1}(pU-1)\nu_2$ to be approximately the F distribution with degrees of freedom ν_1 and ν_2.

If $p = 2$ the two chi-squared approximations coincide and the beta approximation coincides with the exact distribution. Some idea of the goodness of the approximations may be obtained from Tables 1 and 2. The true levels given in Table 1 were calculated by integrating (2·5) and the first of the two terms of (2·4) with the help of Pearson's (1948) table of the incomplete beta function after the substitution $x = 2(1-t^{\frac{1}{2}})/3$; the integration of the second term of (2·4) was done with the help of the same table after the substitution $y = 2(1+t^{\frac{1}{2}})/3$. In violation of expectations, for $p = 3$ and levels 0·05 and 0·01, the approximation obtained by regarding $\frac{1}{2}Np(p-1)\,T$ as distributed as a chi-squared with $\frac{1}{2}p(p+1)-1$ degrees of freedom performed better than the approximation obtained by regarding $(\frac{1}{2}np+1)\,(p-1)\,T$ as distributed as a chi-squared with $\frac{1}{2}p(p+1)-1$ degrees of freedom for all values of n in Table 2. Comparisons have not yet been made for values of p greater than three.

Table 2. *True levels when* $p = 3$

	Nominal level			
	0·05		0·01	
n	$3NT$ regarded as chi-squared with 5 degrees of freedom	Beta approximation	$3NT$ regarded as chi-squared with 5 degrees of freedom	Beta approximation
6	0·038	0·055	0·003	0·015
10	0·042	0·051	0·006	0·013
21	0·045	0·049	0·008	0·011

Appendix. Expected values of U and U^2

In this appendix we shall assume that $\mathbf{\Sigma}_0 = \mathbf{I}$. No generality is lost by this.

Lemma 1. *The equation* $E(U^\theta) = E\{(\Sigma T_i^2)^\theta\}/E\{(\Sigma T_i)^{2\theta}\}$ *is true, provided the expectations involved exist.*

Proof. We have

$$E\{(\Sigma T_i^2)^\theta\} = E\{(\Sigma T_i)^{2\theta}\, U^\theta\},$$
$$= E\{(\Sigma T_i)^{2\theta}\}\, E(U^\theta),$$

since under H_0, ΣT_i is distributed independently of U, $T_i/(\Sigma T_i)$ $(i = 1, \ldots, p-1)$, and ΣT_i being distributed independently. This independence may be verified by obtaining the joint density of $T_i/\Sigma T_i$ $(i = 1, \ldots, p-1)$ and ΣT_i from the joint density of the T_i's. It will be found that it factorizes into two factors, one involving only ΣT_i and the other involving only $T_i/\Sigma T_i$ $(i = 1, \ldots, p-1)$.

Lemma 2. *The following equations are true*:

$$E\{(\Sigma T_i)^2\} = (np+2)\,np\sigma^4, \quad E\{(\Sigma T_i)^4\} = (np+6)\,(np+4)\,(np+2)\,np\sigma^8,$$
$$E(\Sigma T_i^2) = np(n+p+1)\,\sigma^4,$$
$$E\{(\Sigma T_i^2)^2\} = \{pn^3 + (2p^2+2p+8)\,n^2 + (p^3+2p^2+21p+20)\,n + 8p^2 + 20p + 20\}\,np\sigma^8.$$

Proof. The first two equations follow from the fact that ΣT_i being equal to tr $(\mathbf{S})$ has under H_0 the distribution of $\sigma^2\chi^2_{np}$. The last two equations follow from the fact that

$$\Sigma T_i^2 = \operatorname{tr}(\mathbf{S}^2) = \sum_{i=1}^{p}\sum_{j=1}^{p} S_{ij}^2,$$

S_{ij} being the element in the position (i,j) in $\mathbf{S}$. The values of $E(\Sigma T_i^2)$ and $E\{(\Sigma T_i^2)^2\}$ can therefore be written down using the following equations:

$$E(S_{ii}^2) = n(n+2)\,\sigma^4, \quad E(S_{ij}^2) = n\sigma^4, \quad E(S_{ii}^4) = 16(\tfrac{1}{2}n)_4\,\sigma^8, \quad E(S_{ii}^2 S_{ij}^2) = 8(\tfrac{1}{2}n)_3\,\sigma^8,$$
$$E(S_{ii}^2 S_{jj}^2) = n^2(n+2)^2\,\sigma^8, \quad E(S_{ii}^2 S_{jk}^2) = n^2(n+2)\,\sigma^8,$$
$$E(S_{ij}^4) = 3n(n+2)\,\sigma^8, \quad E(S_{ij}^2 S_{ik}^2) = n(n+2)\,\sigma^8, \quad E(S_{ij}^2 S_{kl}^2) = n^2\sigma^8,$$

where we use the notation $(x)_r = x(x+1)\ldots(x+r-1)$. In these equations $i \neq j \neq k \neq l$. The equations follow from the fact that, if H_0 is true, given the ith row of $\mathbf{X}$, S_{ij} has the normal distribution with mean zero and variance $\sigma^2 S_{ii}$ and the fact that S_{ii} has the distribution of $\sigma^2\chi_n^2$.

References

Anderson, T. W. (1958). *An Introduction to Multivariate Statistical Analysis*. New York: Wiley.

Consul, P. C. (1967). On the exact distribution of the criterion W for testing sphericity in a p-variate normal distribution. *Ann. Math. Statist.* **38**, 1170–4.

Fisher, R. A. (1939). The sampling distribution of some statistics obtained from non-linear equations. *Ann. Eugen.* **9**, 238–49.

Hsu, P. L. (1939). On the distribution of the roots of certain determinantal equations. *Ann. Eugen.* **9**, 250–8.

John, S. (1971). Some optimal multivariate tests. *Biometrika* **58**, 123–7.

Mauchly, J. W. (1940). Significance test for sphericity of a normal n-variate distribution. *Ann. Math. Statist.* **11**, 204–9.

Pearson, K. (1948). *Tables of the Incomplete Beta-Function*, 1st edition. Cambridge University Press.

Roy, S. N. (1939). p-statistics or some generalizations in the analysis of variance appropriate to multivariate problems. *Sankhyā* **4**, 381–96.

[*Received June* 1971. *Revised August* 1971]