A NOTE ON THE SPHERICITY TEST[1]

By Leon J. Gleser[2]

Columbia University

1. Introduction and summary. Let x be a random $p \times 1$ column vector having a multivariate normal distribution with unknown mean vector μ and unknown covariance matrix Σ. We wish to test the hypothesis of "sphericity," namely $H: \Sigma = \sigma^2 I_p$, where $\sigma^2 > 0$ is an unknown positive constant. Alternatives to H which are considered are $H_A: \Sigma$ positive definite, but $\Sigma \neq \sigma^2 I$.

Given N observation vectors $x^{(1)}, x^{(2)}, \cdots, x^{(N)}$, independently distributed, each with the distribution of x, we can reduce consideration to the sufficient statistic $(\bar{x}, S)$, where

$$\bar{x} = N^{-1} \textstyle\sum_{i=1}^{N} x^{(i)}, \qquad S = \textstyle\sum_{i=1}^{N} (x^{(i)} - \bar{x})(x^{(i)} - \bar{x})'.$$

Then $\bar{x}$ has a multivariate normal distribution with mean vector μ and covariance matrix Σ/N, and S has the Wishart distribution, i.e., has density

$$p(S) = C_{p,n} |S|^{(n-p-1)/2} |\Sigma|^{-n/2} \exp[-\tfrac{1}{2} \operatorname{tr} \Sigma^{-1} S], \qquad S > 0; \tag{1.1}$$

where

$$C_{p,n}^{-1} = \pi^{p(p-1)/4} 2^{np/2} \textstyle\prod_{i=1}^{p} \Gamma((n - i + 1)/2), \qquad p \leqq n,$$

and $n = N - 1$. Henceforth we shall denote the fact that a random matrix Z has the density (1.1) by writing $\mathcal{L}(Z) = \mathcal{W}(\Sigma, p, n)$; thus, $\mathcal{L}(S) = \mathcal{W}(\Sigma, p, n)$.

Mauchly [4] has found the likelihood ratio test for H v.s. H_A. The rejection region of this test can be written in the form:

$$T(S) \equiv (\operatorname{tr} S)^p / |S| > K, \tag{1.2}$$

where $T(S)/p^p$ is the $-2/N$th power of the likelihood ratio statistic λ.

The moments of the likelihood ratio statistic λ under H were obtained by Mauchly [4]. Anderson [1] uses these moments to give the exact distribution of λ under H and to obtain an asymptotic expansion of this null distribution. The distribution of λ under H_A has been obtained for the case $p = 2$ by Girshick [3], but the distribution of λ under H_A for $p > 2$ appears to be highly untractable.

In this note, we show that the distribution of $T(S)$ is related to the distribution of Bartlett's statistic for testing homogeneity of variances (viz., Anderson [1]). From this relation, we derive that Mauchly's test (1.2) is unbiased. A derivation of the asymptotic distribution of $T(S)$ under H_A completes the note.

It should be mentioned here that a direct relationship between the likelihood ratio *statistic* λ and the Bartlett *statistic* for testing the homogeneity of variances

Received 28 July 1965; revised 15 November 1965.

[1] Research supported in part by National Science Foundation Grant NSF-GP-3694 at Columbia University.

[2] Now at Johns Hopkins University.

for the elements of x is given by Anderson [1]. He shows that $H: \Sigma = \sigma^2 I$ is a combination of two hypotheses $H_1: \Sigma$ is diagonal, and $H_2: \Sigma = \sigma^2 I$ *given* that Σ is diagonal. Hypothesis H_2 is the hypothesis of the homogeneity of the variances of the elements of the vector x given that these random elements are stochastically independent. The likelihood ratio statistic λ_2 for testing this hypothesis is a monotone function of Bartlett's statistic. Further, the likelihood ratio statistic λ for H is the product $\lambda = \lambda_1\lambda_2$ of λ_2 and the likelihood ratio statistic λ_1 for testing H_1 (Anderson [1], pp. 260–2). Unfortunately, both the distribution of λ_1 and the distribution of λ_2 depend upon the unknown Σ, and, unless Σ is diagonal, λ_1 and λ_2 are dependent. As a result, this relationship between λ and Bartlett's statistic is difficult to exploit in finding the distribution of λ. In this note, we use the invariance of λ under orthogonal transformations to enable us to change to new variables having a diagonal covariance matrix. Homogeneity of variances for these new variables is shown to be equivalent to $H: \Sigma = \sigma^2 I$ under the old variables. Using Anderson's representation for H, but now expressed in terms of the new variables, we have $\lambda = \lambda_1'\lambda_2'$, where λ_2' tests homogeneity of variances for the new variables and λ_1' tests the diagonality of the new covariance matrix. Since the new covariance matrix is diagonal, λ_1' and λ_2' are independent and λ_1' has a distribution independent of the parameters. Such a representation is (hopefully) convenient for determining the properties of the likelihood ratio test based on λ. This representation, however, only connects the *distribution* of λ and the *distribution* of Bartlett's statistic, for the "new" variables used in our representation are not observable, but rather are functions of the unknown covariance matrix Σ.

2. A canonical form for the distribution of $T(S)$ and its implications. Note that $T(S)$ is invariant under the transformation $S \to \Gamma S \Gamma'$ for any $p \times p$ orthogonal matrix Γ. In particular let Γ_0 be a matrix of eigenvectors for Σ; that is, Γ_0 is orthogonal and

$$\Gamma_0 \Sigma \Gamma_0' = \text{Diag}(\lambda_1, \cdots, \lambda_p) \equiv D_\lambda$$

where λ_i, $i = 1, \cdots, p$ are the eigenvalues of $\Sigma > 0$. Then letting $V = \Gamma_0 S \Gamma_0'$, we see by invariance that $T(V) = T(S)$. Further, $\mathcal{L}(V) = \mathcal{W}(D_\lambda, p, n)$. The distribution of $T(S)$ can now be found from a consideration of the distribution of $T(V)$.

Letting $R = D_v^{-\frac{1}{2}} V D_v^{-\frac{1}{2}}$ where $D_v \equiv \text{Diag}(v_{11}, v_{22}, \cdots, v_{pp})$, $V = (v_{ij})$ we find that

$$T(V) = \left[\frac{(\sum_{i=1}^p v_{ii})^p}{\prod_{i=1}^p v_{ii}}\right]\frac{1}{|R|} \equiv \frac{L(V)}{|R|}. \tag{2.1}$$

Since $\mathcal{L}(V) = \mathcal{W}(D_\lambda, p, n)$ and since D_λ is diagonal, $|R|$ and $L(V)$ are independent (Anderson [1], p. 174). Further under both H and H_A the distribution of $|R|$ is the same as that of $\prod_{i=1}^{p-1} b_i$, where the b_i's are independent and b_i has the Beta distribution with $\frac{1}{2}(n - i)$ and $\frac{1}{2}i$ degrees of freedom, $i = 1, \cdots, p - 1$ (viz. Anderson [1], p. 237). Thus, $T(V)$ is the ratio of two independent random variables, one of which has a known distribution independent of the parameters (μ, Σ). Since $\mathcal{L}(V) = \mathcal{W}(D_\lambda, p, n)$, the random variables v_{ii} are independent, and v_{ii} has the distribution of λ_i times a chi-square distribution with n degrees of freedom. The statistic $L(V)$ thus has the distribution of Bartlett's statistic for testing homogeneity of variances (i.e., $\lambda_1 = \lambda_2 = \cdots = \lambda_p$).

The hypothesis $H: \Sigma = \sigma^2 I$ is, however, equivalent to the hypothesis $\lambda_1 = \lambda_2 = \cdots = \lambda_p$. Since $|R|$ and $L(V)$ are independent, and since the distribution of R is independent of Σ, we have:

$$\begin{aligned} P[T(V) > K \mid H] &= \int_R P[L(V) > K|R| \mid \lambda_1 = \lambda_2 = \cdots = \lambda_p]\, dP(R) \\ &\leqq \int_R P[L(V) > K|R| \mid \lambda_i \neq \lambda_j \quad \text{some} \quad i \neq j]\, dP(R) \\ &= P[T(V) > K \mid H_A], \end{aligned}$$

the inequality above following from the result of Brown [2] that for all L_0,

$$(2.2) \quad P[L(V) > L_0 \mid \lambda_1 = \lambda_2 = \cdots = \lambda_p] \leqq P[L(V) > L_0 \mid \lambda_i \neq \lambda_j \quad \text{some} \quad i \neq j].$$

We thus conclude that:

THEOREM 1. *The Mauchly test of sphericity is unbiased.*

It might be hoped that the relation (2.1) between Mauchly's statistic and Bartlett's statistic could be exploited to find the distribution of $T(S)$ under H_A. Certainly $Q(S) = \log T(S)$ is distributed as the convolution of $\log L(V)$ and $-\log|R|$. Since we know the distribution of $-\log|R|$, we need only find the distribution of $\log L(V)$ under H_A.

Unfortunately, little work has been done on this distribution, and the results known to the author that have been obtained are not suited for the extra step of forming a convolution. We shall content ourselves for the present with the asymptotic distribution of $T(S)$ as $n \to \infty$.

THEOREM 2. *Under* H_A

$$\lim_{n\to\infty} \mathcal{L}(n^{\frac{1}{2}}[\log T(S) - \log L(D_\lambda)]) = \mathfrak{N}(0, [2 \textstyle\sum_{i=1}^{p} ((\lambda_i/\bar{\lambda}) - 1)^2 + 2(p-1)])$$

where $\bar{\lambda} = p^{-1} \sum_{i=1}^{p} \lambda_i$.

PROOF. Apply Theorem 4.2.5 in Anderson [1] first to $\log L(V)$ and then to $-\log|R|$. A convolution of the limiting distributions of $\log L(V)$ and $-\log|R|$ gives us the desired result.

REFERENCES

[1] ANDERSON, T. W. (1958). *Introduction to Multivariate Statistical Analysis*. Wiley, New York.

[2] BROWN, G. W. (1939). On the power of the L_1 test for equality of several variances. *Ann. Math. Statist.* **10** 119–128.

[3] GIRSHICK, M. A. (1940). The distribution of the ellipticity statistic L_e when the hypothesis is false. *Terr. Mag.* **37** 455–457.

[4] MAUCHLY, J. W. (1940). Significance test for sphericity of a normal n-variate distribution. *Ann. Math. Statist.* **11** 204–209.

[5] NEYMAN, J. and PEARSON, E. S. (1931). On the problem of k samples. *Bull. Acad. Polonaise Sci. Cracovie* **6** 460.

[6] THOMPSON, C. M. and MERRINGTON, M. (1946). Tables for testing the homogeneity of a set of estimated variances. *Biometrika* **33** 296–304.