View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Elsevier - Publisher Connector

JOURNAL OF MULTIVARIATE ANALYSIS 4, 114–122 (1974)

Distribution of the Likelihood Ratio Criterion for Testing $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_0$, $\boldsymbol{\mu} = \boldsymbol{\mu}_0$

B. N. NAGARSENKER* AND K. C. S. PILLAI†

Purdue University

The exact null distribution of the likelihood ratio criterion for testing H_0: $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_0$ and $\boldsymbol{\mu} = \boldsymbol{\mu}_0$ against alternatives H_1: $\boldsymbol{\Sigma} \neq \boldsymbol{\Sigma}_0$ or $\boldsymbol{\mu} \neq \boldsymbol{\mu}_0$ in $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ has been obtained as (a) a chi-square series and (b) a beta series. Percentage points have been tabulated for $p = 2(1)\,6$, $\alpha = .005, .01, .025, .05, .1$, and $.25$ and various values of sample size N.

1. INTRODUCTION

Let a p-variate random sample of size N from the normal distribution with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$ be denoted by $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N$. The likelihood ratio criterion for testing the hypothesis H_0: $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_0$ and $\boldsymbol{\mu} = \boldsymbol{\mu}_0$ against alternatives H_1: $\boldsymbol{\Sigma} \neq \boldsymbol{\Sigma}_0$ or $\boldsymbol{\mu} \neq \boldsymbol{\mu}_0$, where $\boldsymbol{\Sigma}_0$ is a given positive definite matrix and $\boldsymbol{\mu}_0$ a given vector, is expressed as (Anderson [1])

$$L = (e/N)^{Np/2} \,|\, \mathbf{S}\boldsymbol{\Sigma}_0^{-1} \,|^{N/2} \exp[-\tfrac{1}{2} \operatorname{Tr} \boldsymbol{\Sigma}_0^{-1}\{\mathbf{S} + N(\bar{\mathbf{x}} - \boldsymbol{\mu}_0)(\bar{\mathbf{x}} - \boldsymbol{\mu}_0)'\}], \quad (1.1)$$

where $\mathbf{S} = \sum_{\alpha=1}^{N} (\mathbf{x}_\alpha - \bar{\mathbf{x}})(\mathbf{x}_\alpha - \bar{\mathbf{x}})'$ and $\bar{\mathbf{x}} = \sum_{\alpha=1}^{N} \mathbf{x}_\alpha / N$. Although the exact distribution of L is unknown, it has been shown that the asymptotic distribution of $-2 \log L$ is a chi-square with $\frac{1}{2}p(p+1) + p$ degrees of freedom. No further information on the distribution of L appears to be available.

Received April 23, 1973.

AMS 1970 subject classifications: Primary 62H10; Secondary 62E15.

Key words and phrases: Exact null distribution; likelihood ratio criterion; chi-square series; beta series; test specifying mean vector and covariance matrix; percentage points.

* Work of this author was supported in part by the Aerospace Research Laboratories, Air Force Systems Command, U.S. Air Force, Contract F33615-72-C-1400, and in part by David Ross Grant from Purdue Research Foundation.

† Work of this author was sponsored by Aerospace Research Laboratories, Air Force Systems Command, U.S. Air Force, Contract F33615-72-C-1400. Reproduction in whole or part permitted for any purpose of the United States Government.

Copyright © 1974 by Academic Press, Inc.
All rights of reproduction in any form reserved.

In this paper, the null distribution of L is obtained first as a chi-square series and then as a beta series by methods similar to those used in Nagarsenker and Pillai [3]. From these, percentage points of L can be computed to any degree of accuracy even for small sample sizes. Tabulations of percentage points for $p = 2(1)6$ for various significance levels are given. In this paper are presented only tabulations of upper 5% and 1% points of $-2 \log L$ for different values of N.

2. Distribution of L as a Chi-Square Series

The hth moment of L under the null hypothesis is known to be (Anderson [1])

$$E(L^h) = (2e/N)^{Nhp/2}[\Gamma_p\{(n + Nh)/2\}/\Gamma_p(n/2)] \cdot (1 + h)^{-(Np/2)(1+h)}, \quad (2.1)$$

where $n = N - 1$. Let

$$\lambda = -2 \log L. \quad (2.2)$$

If $\phi(t)$ is the characteristic function of λ, then

$$\phi(t) = \frac{(2e/N)^{-Npit}}{(1 - 2it)^{-Np(1-2it)/2}} \frac{\prod_{\alpha=1}^{p} \Gamma\left\{\frac{N}{2}(1 - 2it) - \frac{\alpha}{2}\right\}}{\prod_{\alpha=1}^{p} \Gamma\left(\frac{N - \alpha}{2}\right)}, \quad (2.3)$$

and therefore

$$\begin{aligned} \log \phi(t) = &-Npit \log(2e/N) + \sum_{\alpha=1}^{p} \log \Gamma\left\{\frac{N}{2}(1 - 2it) - \frac{\alpha}{2}\right\} \\ &- \sum_{\alpha=1}^{p} \log \Gamma\left(\frac{N - \alpha}{2}\right) - \frac{Np}{2}(1 - 2it) \log(1 - 2it). \end{aligned} \quad (2.4)$$

Using the following expansion for the gamma function in (2.4),

$$\begin{aligned} \log \Gamma(x + h) = &\tfrac{1}{2} \log(2\pi) + (x + h - \tfrac{1}{2}) \log x - x \\ &- \sum_{r=1}^{m} \frac{(-1)^r B_{r+1}(h)}{r(r + 1)\, x^r} + R_{m+1}(x), \end{aligned} \quad (2.5)$$

where $R_m(x)$ is the remainder such that $| R_m(x)| \leqslant \theta/| x^m |$, θ a constant independent of x, and $B_r(h)$ the Bernoulli polynomial of degree r and order one, we obtain

$$\log \phi(t) = \frac{p}{2}\log(2\pi) + \sum_{\alpha=1}^{p} \log \Gamma\left(\frac{N-\alpha}{2}\right) + \frac{pN}{2}\log(N/2e)$$
$$- \left(\frac{p^3+3p}{4}\right)\log \Gamma\left\{\frac{N}{2}(1-2it)\right\} + \sum_{\gamma=1}^{m} (Q_\gamma/N^\gamma)\left(\frac{1-2it}{2}\right)^{-\gamma}$$
$$+ R'_{m+1}(N, t), \tag{2.6}$$

where the coefficients Q_r's are given by

$$Q_r = (-1)^{r-1} \sum_{\alpha=1}^{p} B_{r+1}(-\alpha/2)/r(r+1). \tag{2.7}$$

The characteristic function of L can then be obtained from (2.6) as

$$\phi(t) = K(p, N)[N(1-2it)/2]^{-v}\left(\sum_{j=0}^{\infty} (B_j/N^j)\left(\frac{1-2it}{2}\right)^{-j}\right) + R''_{m+1}(N, t), \tag{2.8}$$

where

$$K(p, N) = (2\pi)^{p/2}(N/2e)^{pN/2}\left[\prod_{\alpha=1}^{p} \Gamma\left(\frac{N-\alpha}{2}\right)\right]^{-1} \tag{2.9}$$

and

$$v = (p^2+3p)/4. \tag{2.10}$$

The coefficients B_j's can be expressed in terms of Q_j's as in Nagarsenker and Pillai [3]. Since $(1-2it)^{-n/2}$ is the characteristic function of a chi-square density with n degrees of freedom, say $g_n(\chi^2)$, the density of λ can be derived from (2.8) in the form

$$f(\lambda) = K(p, N)\sum_{j=0}^{\infty} (B_j)(2/N)^{j+v} g_{2(j+v)}(\chi^2) + R'''_{m+1}(N). \tag{2.11}$$

The probability that λ is larger than any value, say λ_0, is

$$P(\lambda \geqslant \lambda_0) = K(p, N)\sum_{j=0}^{\infty} (B_j)(2/N)^{j+v} G_{2(j+v)}(\chi^2) + R_{m+1}(N), \tag{2.12}$$

where

$$G_{2(j+v)}(\chi^2) = \int_{\lambda_0}^{\infty} g_{2(j+v)}(\chi^2)\, d\chi^2$$

and

$$R_{m+1}(N) = \frac{1}{2\pi} K(p, N) \int_{\lambda_0}^{\infty} \int_{-\infty}^{\infty} e^{-it\lambda} \sum_{j=0}^{\infty} (B_j)(2/N)^{j+v}(1 - 2it)^{-(j+v)}$$

$$\cdot [\exp\{R'_{m+1}(N, t)\} - 1]\, dt\, d\lambda.$$

Thus from (2.12), we see that the distribution of λ is a series of chi-square distributions.

3. Distribution of L as a Beta Series

Let

$$\lambda_1 = L^{2/N}. \tag{3.1}$$

Then from (2.1), we have

$$E(\lambda_1{}^h) = \frac{(2e/N)^{ph}}{\prod_{\alpha=1}^{p} \Gamma\left(\frac{N-\alpha}{2}\right)} \frac{\prod_{\alpha=1}^{p} \Gamma\left\{\frac{N}{2} + h - \frac{\alpha}{2}\right\}}{(1 + 2h/N)^{Np(1+2h/N)/2}}. \tag{3.2}$$

Using inverse Mellin's transform, the density of λ_1 is given by

$$f(\lambda_1) = \frac{1}{\prod_{\alpha=1}^{p} \Gamma\left(\frac{N-\alpha}{2}\right)} \cdot \frac{1}{2\pi i} \int_{-i\infty}^{i\infty} \frac{\lambda_1^{-h-1}(2e/N)^{ph} \prod_{=1}^{p} \Gamma\left\{\frac{N}{2} + h - \frac{\alpha}{2}\right\}}{(1 + 2h/N)^{Np(1+2h/N)/2}}\, dh. \tag{3.3}$$

Putting $(N/2) + h = t$ in (3.3), we have

$$f(\lambda_1) = K_1(p, N)\, \lambda_1^{N/2-1} \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \lambda_1^{-t} C(t)\, dt, \tag{3.4}$$

where $c = N/2$,

$$C(t) = (e/t)^{pt} \prod_{\alpha=1}^{p} \Gamma\left(t - \frac{\alpha}{2}\right) \tag{3.5}$$

and

$$K_1(p, N) = (2e/N)^{-pN/2} \left[\prod_{\alpha=1}^{p} \Gamma\left(\frac{N-\alpha}{2}\right)\right]^{-1}. \tag{3.6}$$

Using the expansion (2.5) to each gamma function in (3.6), we have

$$\log C(t) = (p/2)\log(2\pi) - \frac{(p^2+3p)}{4}\log t + [A_1/t + A_2/t^2 + \cdots + A_r/t^r + \cdots], \quad (3.7)$$

where the coefficients A_r's are given by

$$A_r = (-1)^{r-1}\Big[\sum_{\alpha=1}^{p} B_{r+1}(-\alpha/2)\Big]\Big/ r(r+1). \quad (3.8)$$

Thus from (3.7), we have

$$C(t) = (2\pi)^{p/2} t^{-(p^2+3p)/4}[1 + B_1/t + B_2/t^2 + \cdots + B_r/t^r + \cdots], \quad (3.9)$$

where the coefficients B_r's can be computed as in Nagarsenker and Pillai [3].

Using (3.9) in (3.4), we have the density of λ_1 as

$$f(\lambda_1) = K_1(p, N)\,\lambda_1^{N/2-1}(2\pi)^{p/2}\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty}\lambda_1^{-t}t^{-v}\Big[1 + \sum_{r=1}^{\infty}(B_r/t^r)\Big]\,dt, \quad (3.10)$$

where $K_1(p, N)$ is as given in (3.6) and $v = (p^2 + 3p)/4$.

The integral on the right-hand side of (3.10) can be easily computed if v is an integer and its value is by Cauchy's theorem of residues, the residue of $\lambda_1^{-t}t^{-v}[1 + \sum_{r=1}^{\infty} B_r/t^v]$ at $t = 0$. This is easily seen to be

$$\sum_{r=0}^{\infty}[(-\log\lambda_1)^{v+r-1}B_r/\Gamma(v+r)], \qquad B_0 = 1, \quad (3.11)$$

and thus from (3.10) the density of λ_1 is

$$f(\lambda_1) = K_1(p, N)(2\pi)^{p/2}\sum_{r=0}^{\infty}(B_r)(-\log\lambda_1)^{v+r-1}/\Gamma(v+r). \quad (3.12)$$

The probability that λ_1 is less than any value, say λ_0, is

$$P(\lambda_1 \leqslant \lambda_0) = K_1(p, N)(2\pi)^{p/2}\sum_{r=0}^{\infty}(B_r)\int_0^{\lambda_0}\lambda_1^{N/2-1}(-\log\lambda_1)^{v+r-1}\,d\lambda_1/\Gamma(v+r). \quad (3.13)$$

For computational purposes, we let

$$I_{v+r-1,u} = \int_0^{\lambda_0}\lambda_1{}^u(-\log\lambda_1)^{v+r-1}\,d\lambda_1/\Gamma(v+r), \quad (3.14)$$

where $u = (N/2) - 1$. Then integrating by parts the right-hand side of (3.14), we have the following recurrence relation:

$$I_{v+r-1,u}(\lambda_0) = [\lambda_0^{u+1}(-\log \lambda_0)^{v+r-1}/\Gamma(v+r) + I_{v+r-2,u}(\lambda_0)]/(u+1) \quad (3.15)$$

and

$$I_{0,u}(\lambda_0) = \lambda_0^{(u+1)}/(u+1). \quad (3.16)$$

With this notation, (3.14) can be written as

$$P(\lambda_1 \leqslant \lambda_0) = K_1(p, N)(2\pi)^{p/2} \sum_{r=0}^{\infty} (B_r)\, I_{v+r-1,u}(\lambda_0), \quad (3.17)$$

where $u = (N/2) - 1$ and $I_{v+r-1,u}(\lambda_0)$ satisfies the recurrence relation (3.15) and (3.16). It is to be noted that (3.17) holds only if $v = (p^2 + 3p)/4$ is an integer. Otherwise, we can appeal to the theorem stated in Nair [5], since in this case

$$\phi(t) = t^{-v}\left[1 + \sum_{r=1}^{\infty} B_r/t^r\right] = 0(t^{-r}). \quad (3.18)$$

Thus according to the theorem, we can expand $\phi(t)$ in the factorial series as

$$\begin{aligned} \phi(t) &= t^{-v}\left[1 + \sum_{r=1}^{\infty} B_r/t^r\right] \\ &= \sum_{i=0}^{\infty} R_i \Gamma(t+a)/\Gamma(t+v+i+a), \end{aligned} \quad (3.19)$$

where a is an arbitrary positive constant and can be chosen to govern the rate of convergence of the resulting series. The coefficients R_i's can be determined explicitly as in Nagarsenker and Pillai [3]. Now using (3.19) in (3.10) and noting that term by term integration is valid since a factorial series is uniformly convergent in a half-plane (see Doetsch [2, pp. 135–36]), we have the density of λ_1 in the case that v is not an integer in the form

$$\begin{aligned} f(\lambda_1) &= K_1(p, N)\, \lambda_1^{N/2-1}(2\pi)^{p/2} \sum_{i=0}^{\infty} R_i \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \lambda_1^{-t}[\Gamma(t+a+r+i)]\, dt \\ &= K_1(p, N)(2\pi)^{p/2} \sum_{i=0}^{\infty} R_i \lambda_1^{N/2+a-1}(1-\lambda_1)^{v+i-1}/\Gamma(v+i). \end{aligned} \quad (3.20)$$

TABLE I

Upper 5% Points of $-2 \log L$ [a]

N \ p	2	3	4	5	6
4	17.381				
5	15.352	27.706			
6	14.318	24.431	39.990		
7	13.689	22.713	35.307	54.261	
8	13.265	21.646	32.787	48.039	70.475
9	12.960	20.915	31.190	44.610	62.660
10	12.729	20.382	30.080	42.400	58.222
11	12.549	19.975	29.261	40.843	55.321
12	12.404	19.655	28.631	39.683	53.254
13	12.285	19.396	28.131	38.782	51.698
14	12.186	19.181	27.723	38.061	50.480
15	12.101	19.002	27.384	37.470	49.499
16	12.029	18.848	27.098	36.977	48.691
17	11.966	18.716	26.854	36.559	48.013
18	11.911	18.601	26.642	36.200	47.436
19	11.862	18.499	26.457	35.888	46.938
20	11.819	18.410	26.294	35.614	46.504
22	11.745	18.258	26.019	35.157	45.785
24	11.684	18.134	25.797	34.790	45.212
26	11.633	18.031	25.614	34.489	44.745
28	11.591	17.944	25.460	34.237	44.357
30	11.554	17.870	25.329	34.023	44.029
32	11.522	17.806	25.215	33.840	43.748
34	11.494	17.750	25.117	33.681	43.505
36	11.469	17.701	25.030	33.541	43.292
38	11.447	17.657	24.954	33.417	43.105
40	11.427	17.618	24.885	33.307	42.938
45	11.386	17.536	24.742	33.079	42.594
50	11.353	17.471	24.630	32.900	42.324
55	11.327	17.419	24.539	32.755	42.107
60	11.305	17.375	24.465	32.636	41.929
65	11.286	17.339	24.402	32.537	41.780
70	11.271	17.308	24.348	32.452	41.654
75	11.257	17.281	24.302	32.379	41.546
80	11.245	17.258	24.262	32.316	41.451
85	11.235	17.237	24.227	32.261	
90	11.225	17.219	24.196	32.211	
95	11.217	17.203	24.168		
100	11.210	17.188	24.143		

[a] p = number of variates; N = number of observations; $-2 \log L = \text{Tr}\, \boldsymbol{\Sigma}_0^{-1} \{\mathbf{S} + N(\bar{\mathbf{x}} - \boldsymbol{\mu}_0)(\bar{\mathbf{x}} - \boldsymbol{\mu}_0)'\} - N \log |\mathbf{S}| + N \log |\boldsymbol{\Sigma}_0| + Np \log N - Np$, where $\bar{\mathbf{x}}$ = sample mean vector and $\mathbf{S}$ = sample sum of product matrix.

TABLE II

Upper 1 % Points of $-2 \log L$ [a]

N \ p	2	3	4	5	6
4	24.087				
5	21.114	36.308			
6	19.625	31.682	50.512		
7	18.729	29.318	44.073	66.728	
8	18.129	27.871	40.713	58.348	84.937
9	17.700	26.890	38.621	53.885	74.530
10	17.377	26.180	37.184	51.063	68.874
11	17.125	25.642	36.133	49.100	65.244
12	16.923	25.219	35.328	47.650	62.690
13	16.758	24.878	34.692	46.531	60.784
14	16.620	24.597	34.176	45.639	59.302
15	16.503	24.361	33.748	44.911	58.114
16	16.403	24.161	33.388	44.305	57.139
17	16.316	23.988	33.080	43.793	56.324
18	16.239	23.838	32.814	43.353	55.631
19	16.172	23.706	32.582	43.973	55.035
20	16.112	23.589	32.378	42.639	54.517
22	16.010	23.392	32.035	42.083	53.658
24	15.927	23.231	31.758	41.637	52.977
26	15.857	23.098	31.529	41.272	52.422
28	15.798	22.986	31.337	40.967	51.961
30	15.747	22.890	31.174	40.708	51.573
32	15.703	22.807	31.033	40.486	51.240
34	15.665	22.734	30.911	40.294	50.953
36	15.631	22.671	30.803	40.125	50.701
38	15.601	22.614	30.708	39.976	50.480
40	15.574	22.564	30.623	39.844	50.284
45	15.517	22.458	30.447	39.568	49.877
50	15.473	22.375	30.308	39.353	49.559
55	15.436	22.307	30.196	39.179	49.303
60	15.406	22.252	30.103	39.036	49.094
65	15.381	22.205	30.025	38.916	48.919
70	15.359	22.165	29.959	38.815	48.770
75	15.341	22.131	29.903	38.727	48.643
80	15.324	22.101	29.853	38.651	48.532
85	15.310	22.074	29.810	38.585	
90	15.297	22.051	29.771	38.526	
95	15.286	22.030	29.737		
100	15.276	22.011	29.706		

[a] p = number of variates; N = number of observations; $-2 \log L = \mathrm{Tr}\, \boldsymbol{\Sigma}_0^{-1} \{\mathbf{S} + N(\bar{\mathbf{x}} - \boldsymbol{\mu}_0)(\bar{\mathbf{x}} - \boldsymbol{\mu}_0)'\} - N \log |\mathbf{S}| + N \log |\boldsymbol{\Sigma}_0| + Np \log N - Np$, where $\bar{\mathbf{x}}$ = sample mean vector and $\mathbf{S}$ = sample sum of product matrix.

The distribution of λ_1 is then given by

$$P(\lambda_1 \leqslant \lambda_0) = K_1(p, N)(2\pi)^{p/2} \sum_{i=0}^{\infty} R_i I_{\lambda_0}\left(\frac{N}{2} + a, v + i\right) \Big/ \Gamma(v + i), \quad (3.21)$$

where $I_{\lambda_0}(p, q)$ is the incomplete beta function $\int_0^{\lambda_0} x^{p-1}(1 - x)^{q-1}\, dx$.

4. Computations of Percentage Points

Percentage points of $\lambda_1 = L^{2/N}$ for $L = .005$, .01, .025, .05, .1, and .25 were computed for $p = 2(1)6$ and various values of N using (2.12), (3.13), and (3.21), and these are available in a technical report [4]. Table I gives the upper 5% and Table II 1% points of $-2 \log L$ for the same values of p and selected values of N.

References

[1] Anderson, T. W. (1958). *An Introduction to Multivariate Analysis.* John Wiley, New York.
[2] Doetsch, G. (1971). *Guide to the Applications of the Laplace and Z-transformations.* Van Nostrand–Reinhold, New York.
[3] Nagarsenker, B. N. and Pillai, K. C. S. (1973). The distribution of the sphericity test criterion. *J. Multivariate Analysis* **3**, 226–235.
[4] Nagarsenker, B. N. and Pillai, K. C. S. (1973). Distribution of the likelihood ratio criterion for testing $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_0$, $\boldsymbol{\mu} = \boldsymbol{\mu}_0$. ARL Technical Report 73–0014, Aerospace Research Laboratories, Wright–Patterson AFB, Ohio.
[5] Nair, U. S. (1940). Application of factorial series in the study of distribution laws in statistics. *Sankhyā* **5** 175.