Ann. Inst. Statist. Math.
38 (1986), Part A, 277-283

LIKELIHOOD RATIO TESTS FOR COMPARING k POPULATIONS —THE TWO-PARAMETER NONREGULAR MODELS

SHAUL K. BAR-LEV* AND BENZION BOUKAI**

(Received July 2, 1984; revised July 30, 1985)

Summary

The null and nonnull distributions of the likelihood ratio statistics for testing the homogeneity of k given populations, each associated with a nonregular density depending on two truncation parameters, are investigated. This generalizes to the two-parameter case the work of Hogg (1956, *Ann. Math. Statist.*, **27**, 529–532), Barr (1966, *J. Amer. Statist. Assoc.*, **61**, 856–864) and Khatri and Jaiswal (1969, *Aust. J. Statist.*, **11**, 79–84; 1969, 1971, *Ann. Inst. Statist. Math.*, **21**, 127–136; **23**, 199–210).

1. Introduction

Let (c, d) be a given (finite or infinite) interval, $h(x)$ a positive integrable function over every closed interval contained in (c, d), and $\Theta = \{(\theta_1, \theta_2): c<\theta_1<\theta_2<d\}$. Let $f(x: \theta_1, \theta_2)$, $(\theta_1, \theta_2) \in \Theta$, be a two-parameter density defined as

$$(1.1) \qquad f(x: \theta_1, \theta_2)=\begin{cases} h(x)/g(\theta_1, \theta_2) & \theta_1 \leq x \leq \theta_2 \\ 0 & \text{elsewhere}, \end{cases}$$

where $g(\theta_1, \theta_2)=\int_{\theta_1}^{\theta_2} h(t)dt$.

Let k populations be given with $f(x: \theta_1^i, \theta_2^i)$ as the parent density associated with the i-th population, $i=1,\cdots,k$. Let X_i and Y_i be the minima and maxima, respectively, of a random sample of size n_i (≥ 2) drawn from $f(x: \theta_1^i, \theta_2^i)$, $i=1,\cdots,k$, and assume that the k samples are independent. Based on these data and the likelihood ratio test (LRT),

Key words: Nonregular distributions, likelihood ratio tests, power functions.

* Now visiting in the Department of Statistics, State University of New York at Buffalo.
** Now in the Department of Mathematical Sciences, State University of New York at Binghamton.

we test for the homogeneity of the k populations, as given by the following hypotheses:

$$\begin{cases} H_1\colon (\theta_1^i, \theta_2^i)=(\theta_1^0, \theta_2^0) & \text{for every } i=1,\cdots,k, \\ & \qquad \text{where } (\theta_1^0, \theta_2^0)\in\Theta \text{ is specified} \\ K_1\colon (\theta_1^i, \theta_2^i)\neq(\theta_1^0, \theta_2^0) & \text{for some } i=1,\cdots,k. \end{cases}$$

and

$$\begin{cases} H_2\colon (\theta_1^i, \theta_2^i)=(\theta_1, \theta_2) & \text{for every } i=1,\cdots,k, \\ & \qquad \text{where } (\theta_1, \theta_2) \text{ is unspecified} \\ K_2\colon (\theta_1^i, \theta_2^i)\neq(\theta_1^j, \theta_2^j) & \text{for some } i \text{ and } j,\ i\neq j,\ i,j=1,\cdots,k. \end{cases}$$

Hogg [5] analyzed the one-parameter problem for testing the homogeneity of the k populations. For the case in which the θ_1^i's are known to be equal to a given constant, Hogg showed that -2 times the logarithm of the LRT statistic has $\chi^2(2k)$ as its null distribution if the common value of the θ_2^i's is specified (under the null hypothesis), and $\chi^2(2k-2)$ otherwise. Barr [2] and Khatri and Jaiswal ([6]-[8]) derived the nonnull distribution of the LRT statistics.

For the two-parameter problem, let Λ_i be the LRT statistic for H_i vs K_i, and set $l_i=-2\log\Lambda_i$, $i=1, 2$. In contrast to the one-parameter results obtained by Hogg [5], it is shown in Section 2 that chi-square fails to be the exact null distribution of the l_i's. Nevertheless it serves as a limiting distribution since as $n_i\rightarrow\infty$, $i=1,\cdots,k$, the limiting null distributions of l_1 and l_2 are $\chi^2(4k)$ and $\chi^2(4k-4)$, respectively. The exact nonnull distribution of l_1 is discussed in Section 3. The corresponding distribution of l_2 is quite difficult to derive (even for the two and three population case). Accordingly, it will not be treated here and is left as an open question.

2. The limiting null distributions of the LRT statistics

To simplify the presentation of the results we first present a lemma. In what follows we use $\psi_Q(\cdot)$ and $f_Q(\cdot)$ to denote the characteristic and density functions, of a r.v. Q.

LEMMA 2.1. *Let X and Y be the minima and maxima, respectively, of a random sample of size n drawn from* (1.1), *and define* $W_1=g(X, Y)/g(\theta_1, Y)$, $W_2=g(X, Y)/g(\theta_1, \theta_2)$, $R=g(\theta_1, X)/g(\theta_1, \theta_2)$ *and* $S=g(\theta_1, Y)/g(\theta_1, \theta_2)$. *Then,* i) $-2\log W_1^{n-1}\sim\chi^2(2)$, ii) $-2\log W_2^n\overset{\mathcal{D}}{=}Z_1+(n/(n-1))Z_2$, *where Z_1 and Z_2 are* i.i.d. r.v.'s *with a common $\chi^2(2)$ distribution, and* iii) *the distribution of (R, S) is free of (θ_1, θ_2).*

PROOF. By transforming from (X, Y) to (W_1, W_2), we obtain $f_{W_1,W_2}(w_1, w_2)=n(n-1)w_2^{n-1}/w_1^2$ for $0<w_2<w_1<1$. The marginal densities of W_1 and W_2 yield (i) and (ii). (iii) is obtained by transforming from (X, Y) to (R, S). One obtains that $f_{R,S}(r, s)=n(n-1)(s-r)^{n-2}$ for $0<r<s<1$, which is the desired result.

We now consider the LRT's for H_i vs K_i, $i=1, 2$. Using the monotonicity properties of $g(\cdot, \cdot)$, we can express Λ_1 and Λ_2 as

$$\Lambda_1=\begin{cases} \prod_{j=1}^{k} \{g(X_j, Y_j)/g(\theta_1^0, \theta_2^0)\}^{n_j}, & \theta_1^0<X_j<Y_j<\theta_2^0,\ j=1,\cdots,k \\ 0, & \text{elsewhere}, \end{cases} \tag{2.1}$$

where $\{\Lambda_1=0\}$ is a null event under H_1, and

$$\Lambda_2=\prod_{j=1}^{k} \{g(X_j, Y_j)/g(X^*, Y^*)\}^{n_j}, \tag{2.2}$$

where $X^*=\min_{1\leq i\leq k} X_i$ and $Y^*=\max_{1\leq i\leq k} Y_i$. The limiting null distributions of l_1 and l_2 are given by the following theorem.

THEOREM 2.1. *Let $n_j\to\infty$ for $j=1,\cdots,k$, then* a) $l_1\xrightarrow{\mathcal{D}}\chi^2(4k)$ *under H_1, and* b) $l_2\xrightarrow{\mathcal{D}}\chi^2(4k-4)$ *under H_2.*

PROOF. Assume that H_1 holds. Λ_1 can be written as $\Lambda_1=\prod_{j=1}^{k} \nu_j$, where $\nu_j=\{g(X_j, Y_j)/g(\theta_1^0, \theta_2^0)\}^{n_j}$, $j=1,\cdots,k$. Because of the independence of the k samples, application of Lemma 2.1 (ii) yields for every t

$$\phi_{l_1}(t)=\prod_{j=1}^{k} \phi_{-2\log\nu_j}(t)=(1-2it)^{-k} \prod_{j=1}^{k} (1-2itn_j/(n_j-1))^{-1}\to(1-2it)^{-2k} \quad \text{as } n_j\to\infty,$$

$j=1,\cdots,k$, and this implies a).

Now assume that H_2 holds. We show that the distribution of Λ_2 is free of (θ_1, θ_2), and that Λ_2 is independent of (X^*, Y^*). We then use Lemma 2.1 (ii) to complete the proof. Define $R_j=g(\theta_1, X_j)/g(\theta_1, \theta_2)$, $S_j=g(\theta_1, Y_j)/g(\theta_1, \theta_2)$, $j=1,\cdots,k$, $R^*=\min_{1\leq j\leq k} R_j$, and $S^*=\max_{1\leq j\leq k} S_j$. Then rewriting (2.2) in terms of these quantities, we obtain $\Lambda_2=\prod_{j=1}^{k} \{(S_j-R_j)/(S^*-R^*)\}^{n_j}$. By Lemma 2.1 (iii), the distribution of (R_j, S_j) is free of (θ_1, θ_2) for all $j=1,\cdots,k$. Since independence of the k samples implies independence of the (R_j, S_j)'s, it follows that the distribution of the random vector $(R_j, S_j, j=1,\cdots,k)$ is free of (θ_1, θ_2) and hence the distribution of Λ_2 is also free of (θ_1, θ_2).

Under H_2, (X^*, Y^*) is complete and sufficient for (θ_1, θ_2). This, the fact that the distribution of Λ_2 is free of (θ_1, θ_2), and Basu's Theorem (Basu [3], Theorem 2) imply that Λ_2 is independent of (X^*, Y^*).

Finally, rewrite Λ_2 in (2.2) as $\Lambda_2=\prod_{j=1}^{k} \xi_j/\xi^*$, where $\xi_j=\{g(X_j, Y_j)/g(\theta_1, \theta_2)\}^{n_j}$, $j=1,\cdots,k$ and $\xi^*=\{g(X^*, Y^*)/g(\theta_1, \theta_2)\}^N$, $N=\sum_{j=1}^{k} n_j$. Since Λ_2 is independent of (X^*, Y^*), the r.v.'s $-2\log\Lambda_2$ and $-2\log\xi^*$ are independent, so that we have for every t, $\psi_{-2\log\Lambda_2}(t)\psi_{-2\log\xi^*}(t)=\prod_{j=1}^{k}\psi_{-2\log\xi_j}(t)$. Application of Lemma 2.1 (ii) for the r.v.'s $-2\log\xi^*$ and $-2\log\xi_j$, $j=1, \cdots, k$, yields

$$\psi_{-2\log\Lambda_2}(t)=(1-2it)^{-(k-1)}(1-2itN/(N-1))\prod_{j=1}^{k}(1-2itn_j/(n_j-1))^{-1}\,.$$

Letting $n_j\to\infty$, $j=1,\cdots,k$, we obtain b).

Remark. If the θ_1^i's and the θ_2^i's are considered as the structural and incidental parameters, respectively, one may be interested in testing

$$\begin{cases} H_3:\ \theta_1^i=\theta_1^0 & \text{for every } i=1,\cdots,k, \text{ where } \theta_1^0 \text{ is specified} \\ K_3:\ \theta_1^i\neq\theta_1^0 & \text{for some } i=1,\cdots,k\,. \end{cases}$$

For testing these hypotheses one can use the LRT as well as the conditional likelihood ratio test (CLRT). The CLRT, which is based on the conditional likelihood of the structural parameters, is defined in a manner analogous to the ordinary LRT. Such a conditional test was introduced by Andersen [1], who derived its asymptotic behaviour for a certain regular model.

Let Λ_3 and Λ_3^c be the LRT and CLRT statistics, respectively, for testing H_3 vs K_3. These statistics have the forms

$$\Lambda_3=\begin{cases} \prod_{j=1}^{k}\{g(X_j, Y_j)/g(\theta_1^0, Y_j)\}^{n_j}\,, & \theta_1^0<X_j<Y_j,\ j=1,\cdots,k \\ 0\,, & \text{elsewhere}\,, \end{cases}$$

and

$$\Lambda_3^c=\begin{cases} \prod_{j=1}^{k}\{g(X_j, Y_j)/g(\theta_1^0, Y_j)\}^{n_j-1}\,, & \theta_1^0<X_j<Y_j,\ j=1,\cdots,k \\ 0\,, & \text{elsewhere}\,. \end{cases}$$

If $l_3=-2\log\Lambda_3$ and $l_3^c=-2\log\Lambda_3^c$, then under H_3, $l_3\xrightarrow{\mathscr{D}}\chi^2(2k)$ as $n_j\to\infty$, $j=1,\cdots,k$, whereas $l_3^c\sim\chi^2(2k)$. The derivation of these results is omitted since it is similar to that of Hogg [5] for the one-parameter case.

3. The nonnull distribution of l_1

Here, we derive the nonnull distribution of l_1 for the k population case. Calculations of standard nature are omitted for the sake of brevity.

We shall find $P(\Lambda_1 \leqq \lambda)$ for all values of the parameters and from this compute the power function of the test. Define the events $B_j = \{\theta_1^0 \leqq X_j \leqq Y_j \leqq \theta_2^0\}$, $\bar{B}_j$ the complement of B_j, $j=1,\cdots,k$, and the indexing sets $K=\{1,\cdots,k\}$, $A_1=\{j \in K\colon \theta_2^j \leqq \theta_1^0\}$, $A_2=\{j \in K\colon \theta_1^j < \theta_1^0 < \theta_2^j \leqq \theta_2^0\}$, $A_3=\{j \in K\colon \theta_1^j < \theta_1^0 < \theta_2^0 < \theta_2^j\}$, $A_4=\{j \in K\colon \theta_1^0 \leqq \theta_1^j < \theta_2^j \leqq \theta_2^0\}$, $A_5=\{j \in K\colon \theta_1^0 \leqq \theta_1^j < \theta_2^0 < \theta_2^j\}$, and $A_6=\{j \in K\colon \theta_2^0 \leqq \theta_1^j\}$.

If $\lambda<0$, then $P(\Lambda_1 \leqq \lambda)=0$, while if $\lambda \geqq 0$,

$$P(\Lambda_1 \leqq \lambda)=P(\Lambda_1 \leqq \lambda,\ \bar{B}_j \text{ occurs for at least one } j=1,\cdots,k) \\ +(P(\Lambda_1 \leqq \lambda,\ B_j \text{ occurs for all } j=1,\cdots,k)\,. \tag{3.1}$$

Let c_1 and c_2 denote the first and second terms, respectively, on the right hand side of (3.1). Since the occurance of $\bigcup_{j=1}^{k} \bar{B}_j$ implies $\Lambda_1 =0 \leqq \lambda$, we have $c_1=P\left(\bigcup_{j=1}^{k} \bar{B}_j\right)=1-\prod_{i=1}^{6} \prod_{j_i \in A_i} P(B_{j_i})$, where $P(B_{j_i})=0$ for $j_i \in A_i$, $i=1, 6$, $P(B_{j_2})=\{g(\theta_1^0, \theta_2^{j_2})/g(\theta_1^{j_2}, \theta_2^{j_2})\}^{n_{j_2}}$, $j_2 \in A_2$, $P(B_{j_3})=\{g(\theta_1^0, \theta_2^0)/g(\theta_1^{j_3}, \theta_2^{j_3})\}^{n_{j_3}}$, $j_3 \in A_3$, $P(B_{j_4})=1$, $j_4 \in A_4$, and $P(B_{j_5})=\{g(\theta_1^{j_5}, \theta_2^0)/g(\theta_1^{j_5}, \theta_2^{j_5})\}^{n_{j_5}}$, $j_5 \in A_5$. If $A_1 \cup A_6=\phi$ then $\bigcup_{i=2}^{5} A_i=K$ and $\bigcap_{i=1}^{k} B_i=\bigcap_{i=2}^{5} \bigcap_{j_i \in A_i} B_{j_i}$, and hence we have

$$c_1=\begin{cases} 1\,, & \text{if } A_1 \cup A_6 \neq \phi \\ 1-\prod_{i=2}^{5} \prod_{j_i \in A_i} P(B_{j_i})\,, & \text{if } A_1 \cup A_6 = \phi\,. \end{cases} \tag{3.2}$$

Let

$$D_1=-2\log\left\{\prod_{j_2\in A_2}\left[\frac{g(X_{j_2}, Y_{j_2})}{g(\theta_1^0, \theta_2^{j_2})}\right]^{n_{j_2}} \prod_{j_3\in A_3}\left[\frac{g(X_{j_3}, Y_{j_3})}{g(\theta_1^0, \theta_2^0)}\right]^{n_{j_3}} \right. \\ \left. \times \prod_{j_4\in A_4}\left[\frac{g(X_{j_4}, Y_{j_4})}{g(\theta_1^{j_4}, \theta_2^{j_4})}\right]^{n_{j_4}} \prod_{j_5\in A_5}\left[\frac{g(X_{j_5}, Y_{j_5})}{g(\theta_1^{j_5}, \theta_2^0)}\right]^{n_{j_5}}\right\},$$

and $d_1=-2\log(\lambda b_1)$, where

$$b_1=\prod_{j_2\in A_2}\left[\frac{g(\theta_1^0, \theta_2^0)}{g(\theta_1^0, \theta_2^{j_2})}\right]^{n_{j_2}} \prod_{j_4\in A_4}\left[\frac{g(\theta_1^0, \theta_2^0)}{g(\theta_1^{j_4}, \theta_2^{j_4})}\right]^{n_{j_4}} \prod_{j_5\in A_5}\left[\frac{g(\theta_1^0, \theta_2^0)}{g(\theta_1^{j_5}, \theta_2^0)}\right]^{n_{j_5}}.$$

Then, we have

$$c_2=\begin{cases} 0\,, & \text{if } A_1 \cup A_6 \neq \phi \\ P\left(D_1 \geqq d_1 \,\middle|\, \bigcap_{i=2}^{5} \bigcap_{j_i \in A_i} B_{j_i}\right) \prod_{i=2}^{5} \prod_{j_i \in A_i} P(B_{j_i})\,, & \text{if } A_1 \cup A_6 = \phi\,. \end{cases} \tag{3.3}$$

Derivation of the distribution of (X_{j_i}, Y_{j_i}) conditional on B_{j_i} for $j_i \in A_i$, $i=2,\cdots,5$, and application of Lemma 2.1 (ii) show that the distribution of D_1 conditional on $\bigcap_{i=2}^{5} \bigcap_{j_i \in A_i} B_{j_i}$ equals the distribution of $\sum_{j=1}^{k} (Z_{1j}+(n_j/(n_j-1))Z_{2j})$, where the Z_{1j}'s and the Z_{2j}'s are i.i.d. r.v.'s having a common $\chi^2(2)$ distribution. Letting $G_{k,n_1,\cdots,n_k}$ denote the distribution of the latter summation, combining (3.2) with (3.3), and noting that $d_1 \leqq 0$ is equivalent to $\lambda \geqq b_1^{-1}$, we obtain for $\lambda \geqq 0$

$$(3.4)\qquad P(\Lambda_1 \leqq \lambda) = \begin{cases} 1, & \text{if } A_1 \cup A_6 \neq \phi \text{ or } A_1 \cup A_6 = \phi \text{ and } \lambda \geqq b_1^{-1} \\ 1 - \int_0^{d_1} dG_{k,n_1,\cdots,n_k}(x) \cdot \prod_{j_2 \in A_2} \left[\frac{g(\theta_1^0, \theta_2^{j_2})}{g(\theta_1^{j_2}, \theta_2^{j_2})} \right]^{n_{j_2}} & \\ \quad \times \prod_{j_3 \in A_3} \left[\frac{g(\theta_1^0, \theta_2^0)}{g(\theta_1^{j_3}, \theta_2^{j_3})} \right]^{n_{j_3}} \prod_{j_5 \in A_5} \left[\frac{g(\theta_1^{j_5}, \theta_2^0)}{g(\theta_1^{j_5}, \theta_2^{j_5})} \right]^{n_{j_5}}, & \\ & \text{if } A_1 \cup A_6 = \phi \text{ and } \lambda < b_1^{-1}. \end{cases}$$

If $\alpha \in (0, 1)$ is the significance level, then H_1 is rejected if the given sample value of Λ_1 is less than λ_0, the critical value determined by $1-\alpha = \int_0^{-2\log \lambda_0} dG_{k,n_1,\cdots,n_k}(x)$. The power function of the corresponding LRT is obtained by replacing in (3.4), λ by λ_0. The technical difficulty connected with the derivation of the power function lies with the fact that no simple expression exists for $G_{k,n_1,\cdots,n_k}$—the distribution of a linear combination of chi-square variates. However, in some special cases it can be given a simpler form. For example, in the case $n_1 = \cdots = n_k = n$, we have for $t \geqq 0$

$$G_{k,n,\cdots,n}(t) = \frac{(n-1)^k}{2^{2k} n^k ((k-1)!)^2} \sum_{i=0}^{k-1} \binom{k-1}{i} (-1)^i \int_0^t w^{k-1-i} e^{-w/2} \cdot \left[\int_0^w y^{k-1-i} e^{y/(2n)} dy \right] dw ,$$

which can be evaluated by use of Gradshtein and Ryzhik ([4], p. 92 (2.321)). Hence for $k=2, 3$, we obtain, respectively,

$$G_{2,n,n}(t) = (1/2)e^{-t/2} \{ n e^{t/(2n)} [-(n-1)t + 2n(2n-3)] - (n-1)^2 (t+4n+2) + 2e^{t/2} \} ,$$

and

$$G_{3,n,n,n}(t) = (1/8)e^{-t/2} \{ n e^{t/(2n)} [-(n-1)^2 t^2 + 4n(n-1)(3n-4)t - 8n^2(6n^2-15n+10)] + (n-1)^3 [t^2 + 4(3n+1)t + 8(6n^2+3n+1)] + 8e^{t/2} \} .$$

Acknowledgement

The authors are grateful to a referee whose suggestions resulted in improvement both in content and style.

UNIVERSITY OF HAIFA, ISRAEL

REFERENCES

[1] Andersen, E. B. (1971). The asymptotic distribution of conditional likelihood ratio tests, *J. Amer. Statist. Assoc.*, **66**, 630-633.

[2] Barr, D. R. (1966). On testing the equality of uniform and related distributions, *J. Amer. Statist. Assoc.*, **61**, 856-864.

[3] Basu, D. (1955). On statistics independent of a complete sufficient statistic, *Sankhya*, **15**, 377-380.

[4] Gradshtein, I. S. and Ryzhik, I. M. (1965). *Tables of Integrals, Series, and Products* (4th ed.), Academic Press, New York.

[5] Hogg, R. V. (1956). On the distribution of the likelihood ratio, *Ann. Math. Statist.*, **27**, 529-532.

[6] Khatri, C. G. and Jaiswal, M. C. (1969). Testing the equality of distributions when the range depends upon the parameter, *Aust. J. Statist.*, **11**, 79-84.

[7] Jaiswal, M. C. and Khatri, C. G. (1969). The power function of the likelihood ratio test when range depends upon the parameter, *Ann. Inst. Statist. Math.*, **21**, 127-136.

[8] Jaiswal, M. C. and Khatri, C. G. (1971). On certain tests and monotonicity of their power for the parameters involved in the nonregular density functions, *Ann. Inst. Statist. Math.*, **23**, 199-210.