The Annals of Mathematical Statistics
1969, Vol. 40, No. 2, 371-380

DISTRIBUTION OF LIKELIHOOD RATIO IN TESTING AGAINST TREND

BY M. T. BOSWELL AND H. D. BRUNK[1]

Pennsylvania State University and University of Missouri

Summary. The phrase "testing against trend" in the title refers to a situation in which observations are made with equal sample sizes on several populations belonging to a common univariate exponential family. Order relations among the parameters associated with the various populations are assumed known, and it is desired to test the null hypothesis that the parameters are all equal. The likelihood ratio test is described in Section 3. Slight extensions, developed in Section 1, of known theorems suffice to determine, in a certain sense, the asymptotic distribution of an appropriate function of the likelihood ratio. This asymptotic distribution is that of Bartholomew's combination of Chi-squares.

1. Preliminary theorems. We give first an extension of known theorems on limiting distributions of functions of random variables which converge in distribution (Mann and Wald, 1943; Sverdrup, 1952; Chernoff, 1956; Prokhorov, 1956).[2]

THEOREM 1.1. *Let m, q be positive integers. For each m − tuple $\alpha = (\alpha_1, \cdots, \alpha_m)$ of positive integers, let $X^{(\alpha)} = (X_1^{(\alpha)}, \cdots, X_q^{(\alpha)})$ be a q-dimensional random vector. Let $X^{(\alpha)}$ converge in distribution as $\min(\alpha_1, \cdots, \alpha_m) \to \infty$ to a q-dimensional random vector X. For each α let $g^{(\alpha)}$ be a real-valued function on R^q, and let the random variables $g^{(\alpha)}(X)$ be identically distributed. Suppose there exists a closed set S in R^q*

(1.1) *such that the functions $g^{(\alpha)}$ are equicontinuous in S;*

$$P[X \varepsilon S] = 1. \tag{1.2}$$

Then $g^{(\alpha)}(X^{(\alpha)})$ converges in distribution as $\min(\alpha_1, \cdots, \alpha_m) \to \infty$ to $g(X)$' where $g = g^{(\alpha_0)}$ for any fixed α_0.

The method which appears to lend itself most readily to this extension is Sverdrup's (1952). His theorem is the case in which $m = 1$ and the functions $g^{(\alpha)}$ are all identical. (Hypothesis B(iii) on page 3 of (Sverdrup, 1952) can be shown to be implied by the others.) To allow different $g^{(\alpha)}$ requires only minor modifications in the proof; and as D. L. Hanson pointed out to one of the authors, an argument by contradiction reduces the case of m-way arrays to the case of one-way arrays.

Received 26 February 1968.

[1] The research of this author was partially supported by the Air Force Office of Scientific Research, Office of Aerospace Research, United States Air Force, under AFOSR Grant Nr. 746-65A.

[2] Note added in proof: Theorem 1.1 is also a consequence of Theorem 2 in "Preservation of weak convergence under mappings" by Flemming Topsoe, *Ann. Math. Statist.* **38** (1967) 1661–1665.

We shall require in Section 3 also an extension of Theorem 2.1 in (Brunk, 1960). (The conclusion of Theorem 2.1 in (Brunk, 1960) is there incorrectly stated, for discrete random variables. It should be stated as follows:

$$P[f_M(Z, X) < q] = \sum_{m=1}^{n} \sum_{k \varepsilon \mathcal{K}_m} P\{f_m[y(\alpha, X), X] < q\} / \prod_{i=1}^{n} k_i! i^{k_i}.)$$

Let random variables Ξ_i, $i = 1, 2, \cdots, n$, have as common range an interval $I \subset R$ (the reals) so that $\Xi = (\Xi_1, \cdots, \Xi_n) \varepsilon I^n = \times_{i=1}^{n} I \subset R^n$. For $m = 1, 2, \cdots, n$, define $\mathcal{K}_m = \{k = (k_1, \cdots, k_n): k_1, \cdots, k_n$ are nonnegative integers, $\sum_{i=1}^{n} ik_i = n, \sum_{i=1}^{n} k_i = m\}$. For $k \varepsilon \mathcal{K}_m$, define $\mathfrak{A}^k = \{\alpha = (\alpha_1, \cdots, \alpha_m)$: $\alpha_1, \cdots, \alpha_m$ are positive integers; for $i = 1, 2, \cdots, n$, exactly k_i of the components of α are equal to $i\}$. Define $\mathfrak{A}_m = \bigcup_{k \varepsilon \mathcal{K}_m} \mathfrak{A}^k$, $\mathfrak{A} = \bigcup_{m=1}^{n} \mathfrak{A}_m$, and

$$\mathcal{V}_m = \{v = (v_1, \cdots, v_m): v_j = (\alpha_j, w_j), j = 1, 2, \cdots, m,$$
$$\alpha = (\alpha_1, \cdots, \alpha_m) \varepsilon \mathfrak{A}_m, w = (w_1, \cdots, w_m) \varepsilon I^m\}.$$

Set $\beta_0 = 0$, $\beta_j = \beta_j(\alpha) = \sum_{\nu=1}^{j} \alpha_\nu$, $j = 1, 2, \cdots, m$. For $\alpha \varepsilon \mathfrak{A}_m$ and $\xi \varepsilon I^n$, define

$$u_j = u_j(\alpha, \xi) = \sum_{r=\beta_{j-1}+1}^{j} \xi_r / \alpha_j, \qquad j = 1, 2, \cdots, m,$$
$$u = u(\alpha, \xi) = (u_1, \cdots, u_m), \qquad y_j = y_j(\alpha, \xi) = (\alpha_j, u_j),$$
$$j = 1, 2, \cdots, m, y = y(\alpha, \xi) = v(\alpha, u) = (y_1, \cdots, y_m).$$

For $\xi \varepsilon I^n$, consider the least concave majorant of the set of points (k, Σ_k), $k = 0, 1, 2, \cdots, n$, where $\Sigma_k = \sum_{i=1}^{k} \xi_i$. Let its vertices have abscissas $b_0(\xi) = 0$, $b_1(\xi), \cdots, b_{m(\xi)}(\xi)$, and set $\alpha_j(\xi) = b_j(\xi) - b_{j-1}(\xi)$, $j = 1, 2, \cdots, m(\xi)$. Set $\alpha(\xi) = (\alpha_1(\xi), \cdots, \alpha_m(\xi))$, where $m = m(\xi)$, and $w(\xi) = u(\alpha(\xi), \xi) = (w_1(\xi), \cdots, w_m(\xi))$, where $w_j(\xi) = u_j(\alpha(\xi), \xi)$. Define $z(\xi) = y[\alpha(\xi), \xi)] = v(\alpha(\xi), u(\xi))$, $Z = y[\alpha(\Xi), \Xi]$, $M = m(\Xi)$, and $W_j = w_j(\Xi) = u_j(\alpha(\Xi), \Xi)$, $j = 1, 2, \cdots, M$.

THEOREM 1.2. *Let $\Xi_1, \cdots, \Xi_n$ be exchangeable random variables. For $m = 1, 2, \cdots, n$, let $f_m(v, \xi)$ be symmetric in the components of $v \varepsilon \mathcal{V}_m$ and in the components of $\xi \varepsilon I^n$. If*

(1.3) *the joint distribution function of $\Xi_1, \cdots, \Xi_n$ is continuous,*

or if

(1.4) *$f_{m(\xi)}[z(\xi), \xi]$, and for $\alpha \varepsilon \mathfrak{A}_m$, the functions $f_m[y(\alpha, \xi), \xi]$ are continuous in ξ on I^n, $m = 1, 2, \cdots, n$*

then

$$P[f_M(Z, \Xi) \leqq f] = \sum_{m=1}^{n} \sum_{k \varepsilon \mathcal{K}_m} P\{f_m[y(\alpha, \Xi), \Xi] \leqq f\} / \prod_{i=1}^{n} k_i! i^{k_i} \tag{1.5}$$

where α is selected arbitrarily from $\mathfrak{A}^k$ for each $k \varepsilon \mathcal{K}_m$.

PROOF. The proof under hypothesis (1.3) is just the proof in (Brunk, 1960). To carry out the proof under hypothesis (1.4), for fixed $\nu = 1, 2, \cdots$, let Ξ_1^ν,

$\Xi_2^\nu, \cdots, \Xi_n^\nu$ be exchangeable and have a continuous joint distribution function such that Ξ^ν converges in distribution to Ξ as $\nu \to \infty$. The conclusion follows from the first part of Theorem 1.2 and from the standard case of Theorem 1.1 on setting $g(\xi)$ first equal to $f_m[y(\alpha, \xi), \xi]$ for fixed $m = 1, 2, \cdots, n$ and $\alpha \varepsilon \mathcal{A}^k(k \varepsilon \mathcal{K}_m)$, and then $g(\xi)$ equal to $f_{m(\xi)}(z(\xi), \xi)$.

Theorem 1.2 is oriented toward application in the present paper by describing $\alpha(\xi)$ in terms of a least concave majorant. It remains valid, however, if the more general "mean value" functions E of (Brunk, 1960) replace the arithmetic averages used here in determining the least concave majorant.

The product $\prod_{i=1}^n 1/k_i!^{k_i}$ occurring in equation (1.5) is also the probability that a randomly chosen permutation of $(1, 2, \cdots, n)$ will have k_i cycles of length i, $i = 1, 2, \cdots, n$ (cf. Brunk, 1960, e.g. (2.9), p. 322; p. 313). A permutation of $(1, 2, \cdots, n)$ may be described in terms of cycles in a way easily illustrated by example. If the permutation carries $(1, 2, 3, 4, 5, 6, 7)$ into $(5, 1, 4, 7, 6, 2, 3)$, then $1 \to 5, 5 \to 6, 6 \to 2, 2 \to 1; 3 \to 4, 4 \to 7, 7 \to 3$. This permutation has two cycles, of lengths 4 and 3 respectively, which may be represented by $(5, 6, 2, 1), (4, 7, 3)$. With this convention, the first cycle stops when 1 is reached. The second stops when the next smaller number not previously encountered is reached, etc. With each of the $n!$ permutations equally likely, one verifies that the probability $p(\alpha)$ that the number M of cycles will be m and that the length A_1 of the first cycle will be α_1, the length A_2 of the second cycle α_2, etc., is given by

$$\begin{aligned} p(\alpha) &= P[M = m, A_1 = \alpha_1, A_2 = \alpha_2, \cdots, A_m = \alpha_m] \\ &= \textstyle\prod_{j=1}^m (1/\sum_{\nu=j}^m \alpha_\nu), \end{aligned}$$

where $\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_m)$. For $k \varepsilon \mathcal{K}_m$, set $P_k = \prod_{i=1}^n 1/k_i!i^{k_i}$. We then have also for $m = 1, 2, \cdots, n$, and for $k \varepsilon \mathcal{K}_m$,

$$P_k = \textstyle\sum_{\alpha \varepsilon A^k} p(\alpha),$$

since each represents the probability that a randomly chosen permutation will have k_i cycles of length i, $i = 1, 2, \cdots, n$.

The probability function of the number of cycles is given by

$$P[M = m] = |S_n^m|/n!,$$

where $|S_n^m|$ is the coefficient of z^m in $\prod_{i=1}^n (z + i - 1)$; S_n^m are Stirling's Numbers of the First Kind (cf. Miles, 1959).

LEMMA 1.1. *For $m = 1, 2, \cdots$, and for each positive integer r,* $\lim_{n\to\infty} P[\min (A_1, A_2, \cdots, A_m) \leqq r \mid M = m] = 0$.

This lemma is proved in (Boswell, 1966). For $m = 1, 2, \cdots, n$ and for positive integers r, set

$$D(m, r) = \{\alpha: \alpha \varepsilon \mathcal{A}_m, \min (\alpha_1, \cdots, \alpha_m) \leqq r\}.$$

Then Lemma 1.1 may be restated as follows: for fixed m and r,

$$\lim_{n\to\infty} n! \textstyle\sum_{\alpha \varepsilon D(m,r)} p(\alpha)/|S_n^m| = 0.$$

2. Testing against trend in populations belonging to an exponential family. For the case in which the populations are normal, Bartholomew (1959a, 1959b, 1961) found the distribution of the likelihood ratio. For the case in which they belong to an arbitrary exponential family, a simplification of the distribution problem was pointed out in (Brunk, 1960). Using this simplification and a result of one of the authors (Lemma 1.1) we are now able to discuss the asymptotic distribution.

Let F be a non-degenerate distribution function which admits a moment generating function

$$\exp [\Theta(\tau)] = \int \exp (x\tau)\, dF(x) \tag{2.1}$$

converging in a neighborhood of the origin. The functions $\exp \{x\tau - \Theta(\tau)\}$ are densities with respect to F of an exponential family $F_\tau(x)$ of distributions. Set $\theta(\tau) = \Theta'(\tau)$, and let X be a random variable having distribution function $F_\tau(x)$. Then

$$E(X_\tau) = \theta(\tau), \qquad \operatorname{Var} X_\tau = \theta'(\tau). \tag{2.2}$$

The function Θ is convex. Let T denote its convex conjugate:

$$T(\theta) = \sup_\tau [\theta\tau - \Theta(\tau)]. \tag{2.3}$$

The increasing function $\tau(\theta) = T'(\theta)$ is the inverse of $\theta(\tau)$, and with the change of parameter $\tau = \tau(\theta)$, the exponential densities can be written

$$f(x; \theta) = \exp [T(\theta) + (x - \theta)\tau(\theta)], \tag{2.4}$$

where θ is the mean.

Now suppose that $\Xi_1, \Xi_2, \cdots, \Xi_n$ are independent, and that Ξ_i has density $f(\xi, \theta_i)$ with respect to F, given by (2.4). We consider the likelihood ratio test of the null hypothesis

$$H_0 : \theta_1 = \theta_2 = \cdots = \theta_n \tag{2.5}$$

within the class of alternatives

$$H_1 : \theta_1 \geqq \theta_2 \geqq \cdots \geqq \theta_n . \tag{2.6}$$

The logarithm of the joint density of $\Xi_1, \cdots, \Xi_n$ at an observed point $\xi = (\xi_1, \cdots, \xi_n)$ is given by

$$\textstyle\sum_{i=1}^n [T(\theta_i) + (\xi_i - \theta_i)\tau(\theta_i)],$$

whose maximum under H_0 is $nT(\bar{\xi})$, where $\bar{\xi} = \sum_{i=1}^n \xi_i/n$, obtained by setting $\theta_i = \bar{\xi}$, $i = 1, 2, \cdots, n$. It is shown in (Brunk, 1955) (cf. also Brunk, 1958) that the maximum likelihood under H_1 is obtained by setting $\theta_i = w_j(\xi)$ for $[i \,\varepsilon\, (b_{j-1}(\xi), b_j(\xi)]$. Thus the likelihood ratio is

$$\lambda = \max_{\theta \varepsilon H_0} \textstyle\prod_{i=1}^n f(\xi_i ; \theta_i)/\max_{\theta \varepsilon H_1} \prod_{i=1}^n f(\xi_i ; \theta_i), \tag{2.7}$$

$$\lambda = \exp (-\{ \textstyle\sum_{j=1}^{m(\xi)} a_j(\xi) T[w_j(\xi)] - nT(\bar{\xi})\}).$$

As a random variable, a function of Ξ,

$$-2 \log \Lambda = 2 \textstyle\sum_{j=1}^M \alpha_j(\Xi) T(W_j) - nT(\bar{\Xi}).$$

We shall require the following theorem on the asymptotic distribution of a quadratic form. It is in the spirit of Madow's results (1940) but is not contained in them. Let $\Xi_1, \Xi_2, \cdots, \Xi_n$ be random variables. For positive integers $\alpha_1, \alpha_2, \cdots, \alpha_m$ whose sum is n, define

$$\beta_j = \sum_{i=1}^{j} \alpha_i, \quad j = 1, 2, \cdots, m; \quad \beta_0 = 0; \quad n = \beta_m = \sum_{j=1}^{m} \alpha_j;$$

$$U_j = U_j(\alpha) = \sum_{\nu=\beta_{j-1}}^{\beta_j} \Xi_\nu/\alpha_j, \quad j = 1, 2, \cdots, m, \quad \text{where}$$

$$\alpha = (\alpha_1, \cdots, \alpha_m); \quad \bar{\Xi} = \sum_{\nu=1}^{n} \Xi_\nu/n.$$

THEOREM 2.1. *Let $\Xi_1, \Xi_2, \cdots$ be independent, identically distributed random variables, each with mean μ, variance σ^2. Let m be a positive integer. Set*

$$Q^{(\alpha_1,\cdots,\alpha_m)} = \sum_{j=1}^{m} \alpha_j U_j^2 - n\bar{\Xi}^2/\sigma^2,$$

for positive integers $\alpha_1, \cdots, \alpha_m$. Then the limiting distribution of $Q^{(\alpha_1,\cdots,\alpha_m)}$ as $\alpha_1, \cdots, \alpha_m \to \infty$ is χ^2 with $m - 1$ degrees of freedom if $m > 1$; $Q = 0$ if $m = 1$.

PROOF. Set $X_j = (\alpha_j)^{\frac{1}{2}}(U_j - \mu)/\sigma$, $j = 1, 2, \cdots, m$, $X^T = (X_1, \cdots, X_m)$ (i.e., X^T is the transpose of the column vector of which X_j is the jth component). Then $Q = X^T(I - A/n)X$, where I is the $m \times m$ identity matrix, and A is the $m \times m$ symmetric matrix with $(\alpha_i\alpha_j)^{\frac{1}{2}}$ in the ith row and jth column. One verifies that the matrix $I - A/n$ is idempotent as well as symmetric. The norm of $I - A/n$ (as a linear operator on R^m: $\|I - A/n\| = \max_{\|x\|=1}\|(I - A/n)x\|$) is then 1. It follows that the functions $Q^{(\alpha_1,\cdots,\alpha_m)}(x) = x^T(I - A/n)x$ are equicontinuous. Further, from the Central Limit Theorem, the random variables $X_1, \cdots, X_m$ have a limiting distribution as $\min(\alpha_1, \cdots, \alpha_m) \to \infty$ which is that of independent normal (0, 1) random variables $Y_1, \cdots, Y_m$. Since $I - A/n$ is symmetric and idempotent, its rank is equal to its trace, which is $m - 1$. It follows that the distribution of $Q^{(\alpha_1,\cdots,\alpha_m)}(Y) = Y^T(I - A/n)Y$ is χ^2 with $m - 1$ degrees of freedom. Applying Theorem 1.1 we have the desired conclusion.

THEOREM 2.2. *With notation as in Theorem 2.1, let T be a real function on an open interval containing μ, having an integrable second derivative, T'', continuous at μ. Set*

$$Z = Z^{(\alpha_1,\cdots,\alpha_m)} = \sum_{j=1}^{m} \alpha_j T(U_j) - nT(\Xi).$$

Then the limiting distribution of $2Z^{(\alpha_1,\cdots,\alpha_m)}/\sigma^2 T''(\mu)$ as $\alpha_1, \cdots, \alpha_m \to \infty$ is χ^2 with $m - 1$ degrees of freedom if $m > 1$; $Z = 0$ if $m = 1$.

PROOF. Set $\epsilon(x) = 2\int_\mu^x (x - t)[T''(t) - T''(\mu)]\,dt/(x - \mu)^2 = 2\{T(x) - T(\mu) - (x - \mu)T'(\mu) - (x - \mu)^2T''(\mu)/2\}/(x - \mu)^2$. We have $\epsilon(x) \to 0$ as $x \to \mu$. Then

$$\begin{aligned} Z^{(\alpha_1,\cdots,\alpha_m)} &= \sum_{j=1}^{m} \alpha_j[T(U_j) - T(\mu)] - n[T(\bar{\Xi}) - T(\mu)] \\ &= T'(\mu)\{\sum_{j=1}^{m} \alpha_j(U_j - \mu) - n(\bar{\Xi} - \mu)\} \\ &\quad + T''(\mu)\{\sum_{j=1}^{m} \alpha_j(U_j - \mu)^2 - n(\bar{\Xi} - \mu)^2\}/2 \\ &\quad + \{\sum_{j=1}^{m} \alpha_j(U_j - \mu)^2\epsilon(U_j) - n(\bar{\Xi} - \mu)^2\epsilon(\bar{\Xi})\}. \end{aligned}$$

The first term on the right is zero. Each term in the last set of braces converges in probability to zero, as $\min(\alpha_1, \cdots, \alpha_m) \to \infty$, since U_j and $\bar{\Xi}$ converge in probability to μ, while $\alpha_j(U_j - \mu)^2$ and $n(\bar{\Xi} - \mu)^2$ converge in distribution to χ^2 with one degree of freedom. By Theorem 2.1, the random variable in the second pair of braces, multiplied by $1/\sigma^2$, has as limiting distribution the χ^2 with $m - 1$ degrees of freedom (is 0 if $m = 1$). This completes the proof of the theorem.

In discussing the asymptotic distribution of $-2 \log \Lambda$, we apply Theorem 1.2, with

$$f_m(v, \xi) = 2[\sum_{j=1}^m \alpha_j T(w_j) - nT(\sum_{j=1}^m \alpha_j w_j/n))],$$

where $v_j = (\alpha_j, w_j)$, $j = 1, 2, \cdots, m$, $v = (v_1, \cdots, v_m) \in \mathcal{V}_m$. (Thus $f_m(v, \xi)$ is constant as a function of ξ for fixed $v \in \mathcal{V}_m$.) We observe that $f_m(v, \xi)$ is symmetric in the components of v (and, of course, in the components of ξ). We have, for $\alpha \in \mathcal{A}_m$,

$$f_m[y(\alpha, \xi), \xi] = 2\{\sum_{j=1}^m \alpha_j T[u_j(\alpha, \xi)] - nT(\sum_{j=1}^m \alpha_j u_j(\alpha, \xi)/n)\},$$

and

$$f_{m(\xi)}[z(\xi), \xi] = 2\{\sum_{j=1}^{m(\xi)} \alpha_j(\xi) T[w_j(\xi)] - nT[\sum_{j=1}^{m(\xi)} \alpha_j(\xi) w_j(\xi)/n]\}.$$

The integral (2.1) giving $\exp\{\Theta(\tau)\}$ is assumed to converge in an open interval J containing the origin; in J, Θ has derivatives of all orders, and is convex. Recalling that $\theta = \Theta'$, we set $K = \theta(J)$; in K, T has derivatives of all orders, and is convex. It follows that $f_m[y(\alpha, \xi), \xi]$ is continuous in ξ for $\alpha \in \mathcal{A}_m$. Also, while $\alpha_j(\xi)$ $(j = 1, 2, \cdots)$, is not continuous in ξ, nor is $m(\xi)$, the sums $\sum_{j=1}^{m(\xi)} \alpha_j(\xi) T[u_j(\alpha(\xi), \xi)]$ and $\sum_{j=1}^{m(\xi)} \alpha_j(\xi) u_j(\alpha(\xi), \xi)$ are continuous in ξ. Thus the hypotheses of Theorem 1.2 are satisfied. For $\alpha \in \mathcal{A}_m$, we set

$$-2 \log \Lambda(\alpha) = 2 \sum_{j=1}^m \alpha_j T(u_j(\alpha, \xi) - nT(\bar{\xi})).$$

We conclude that

$$P[-2 \log \Lambda \leqq \delta] = \sum_{m=1}^n \sum_{k \in \mathcal{K}_m} P[-2 \log \Lambda(\alpha_k) \leqq \delta] \prod_{i=1}^n 1/k_i! i^{k_i},$$

where for $k \in \mathcal{K}_m$, α_k is chosen arbitrarily from $\mathcal{A}^k$. Since $P[-2 \log \Lambda(\alpha) \leqq \delta]$ is the same for all $\alpha \in \mathcal{A}^k$ and since for $k \in \mathcal{K}_m$ we have $\sum_{\alpha \in \mathcal{A}^k} p(\alpha) = \prod_{i=1}^n 1/k_i! i^{k_i}$ (cf. Section 1), we may write also

$$(2.8) \quad P[-2 \log \Lambda \leqq \delta] = \sum_{m=1}^n \sum_{\alpha \in \mathcal{A}_m} P[-2 \log \Lambda(\alpha) \leqq \delta] p(\alpha).$$

THEOREM 2.3. *For* $m = 1, 2, \cdots$,

$$\lim_{n \to \infty} n! \sum_{\alpha \in \mathcal{A}_m} P[-2 \log \Lambda(\alpha) \leqq \delta] p(\alpha)/|S_n^m| = P[\chi_{m-1}^2 \leqq \delta].$$

PROOF. From (2.2) it follows that if Ξ_1 has density $f(\xi, \theta)$ and if $E\Xi_1 = \mu$, $\operatorname{Var} \Xi_1 = \sigma^2$, then in applying Theorem 2.2 we have $\mu = \theta$ and $\sigma^2 T''(\mu) = \sigma^2 \tau'(\theta) = 1$. We conclude that the limiting distribution of $-2 \log \Lambda(\alpha)$ as $\min(\alpha_1, \cdots, \alpha_m) \to \infty$ is χ^2 with $m - 1$ degrees of freedom. Thus if $\epsilon > 0$, there is a positive integer r_0 such that $r \geqq r_0$, $\alpha \in \mathcal{A}_m - D(m, r)$ imply $|P[-2$

$\log \Lambda(\alpha) \leqq \delta] - P[\chi^2_{m-1} \leqq \delta]| < \epsilon$. We have

$$|n!\sum_{\alpha\epsilon\mathfrak{A}_m} P[-2 \log \Lambda(\alpha) \leqq \delta]p(\alpha)/|S_n^m| - P[\chi^2_{m-1} \leqq \delta]|$$
$$= |n!\sum_{\alpha\epsilon\mathfrak{A}_m} \{P[-2 \log \Lambda(\alpha) \leqq \delta] - P[\chi^2_{m-1} \leqq \delta]\}p(\alpha)/|S_n^m|$$
$$\leqq 2n!\sum_{\alpha\epsilon D(m,r_0)} p(\alpha)/|S_n^m| + \epsilon.$$

The conclusion of the theorem now follows from Lemma 1.1.

Formula (2.8) and Theorem 2.3 suggest the following approximation:

$$P[-2 \log \Lambda \leqq \delta] \doteq \sum_{m=1}^n P[\chi^2_{m-1} \leqq \delta]|S_n^m|/n!. = P[\bar{\chi}_n^2 \leqq \delta].$$

The right member is Bartholomew's combination of chi-squares (1959a): Bartholomew's case is the case of sampling from normal distributions, when the formula is exact for all n.

3. Example: sampling from exponential distributions. For $k = 1, 2, \cdots, n$, let Ξ_k be an observation on the distribution having density $f(\xi, k) = (1/\theta_k) \exp(-\xi/\theta_k)$, $\xi > 0$. For example, Ξ_k may be the length of life of an item of a certain kind. It may be known that the mean life θ_k is nonincreasing in k, $k = 1, 2, \cdots, n$. It may then be desired to test the null hypothesis

$$H_0: \theta_1 = \theta_2 = \cdots = \theta_n$$

within the class of alternatives

$$H_1: \theta_1 \geqq \theta_2 \geqq \cdots \geqq \theta_n .$$

In the notations of Sections 1 and 2 we have $T(x) = x - 1 - \ln x$ for $x > 0$, and

$$\lambda = \prod_{j=1}^m w_j^{\alpha_j}/\xi^n,$$

where $m = m(\xi)$, $w_j = w_j(\xi)$, $\alpha_j = \alpha_j(\xi)$ or

$$\lambda = (n^n \prod_{j=1}^m \alpha_j^{\alpha_j}) \prod_{j=1}^m (s_j/s)^{\alpha_j} \tag{3.1}$$

where $s_j = s_j(\xi) = \alpha_j w_j(\xi) = \sum_{i=b_{j-1}+1}^{b_j} \xi_i$, $s = \sum_{i=1}^n \xi_i$.

To use Λ as a test statistic, one must know, at least approximately, λ such that $P[\Lambda \leqq \lambda \mid H_0] = \alpha$, where α is the desired level of significance. The approximation suggested in Section 2 for the left member is the corresponding probability for Bartholomew's combination of χ^2 (Bartholomew, 1959a) :$P[\bar{\chi}_n^2 > -2 \log \lambda] = \sum_{m=1}^n P[\chi^2_{m-1} > -2 \log \lambda]|S_n^m|/n!$. Table 1 compares these probabilities for various combinations of n and λ, the exact probability for $P[\Lambda \leqq \lambda \mid H_0]$ being computed as indicated below. Table 2 lists values of λ corresponding to various significance levels α for the exact distribution of Λ for the present example; and Table 3 lists values of λ for Bartholomew's distribution. A table corresponding to Table 3 but giving $-2 \log \lambda$ instead of λ for $n = 3, 4, \cdots, 12, \alpha = .005, .01, .025, .05, .10$ may be found in (Bartholomew, 1959b).

The exact distribution of Λ was found using (1.5). For fixed m ($m = 1, 2, \cdots, n$) and fixed $k \epsilon \mathcal{K}_m$, let α be chosen arbitrarily from $\mathfrak{A}^k$. Let $\lambda(\alpha)$ be given

by (3.1) in which, however, m now does not depend on ξ, nor does α, the α_j bring the components of α;
and

$$s_j = \sum_{r=\beta_{j-1}+1}^{\beta_j} \xi_r = \alpha_j u_j(\alpha, \xi), \quad s = \sum_{i=r}^{n} \xi_i .$$

If $\Lambda(\alpha)$ is the random variable obtained by replacing ξ_i by Ξ_i, $i = 1, 2, \cdots, r$, then applying (1.5) we have

$$P[\Lambda \leqq f \mid H_0] = \sum_{m=1}^{n} \sum_{k\varepsilon\mathcal{K}_m} P[\Lambda(\alpha) \leqq f \mid H_0] / \prod_{i=1}^{n} k_i! i^{k_i}, \tag{3.2}$$

where α is selected arbitrarily from $\mathcal{A}^k$ for each $k \varepsilon \mathcal{K}_m$. The distribution function of $\Lambda(\alpha)$ in turn was approximated by the first few terms of its expansion in Legendre polynomials $\{P_k(x)\}$ over [0, 1]:

$$F(x) = \sum_{k=0}^{\infty} a_k P_k(x) \tag{3.3}$$

where

$$a_k = (-\tfrac{1}{2}) E\{P_{k+1}[\Lambda(\alpha)] - P_{k-1}[\Lambda(\alpha)]\}. \tag{3.4}$$

The computation of $EP_k[\Lambda(\alpha)]$ required the moments $E[\Lambda(\alpha)]^t$, computed as follows. The joint distribution of $\Xi_1, \cdots, \Xi_n$ under H_0 is that of the first n interarrival times of a Poisson process; and the joint distribution of $\Xi_1/s, \cdots, \Xi_n/s$ is that of $U_1, U_2 - U_1, \cdots, U_{n-1} - U_{n-2}, 1 - U_{n-1}$ where $U_1, \cdots, U_{n-1}$ are order statistics of a random sample of $n - 1$ from the uniform distribution over [0, 1].

For $j = 1, 2, \cdots, m$, set $Z_j = Z_j(\alpha) = s_j/s$. Then $\sum_{j=1}^{m} Z_j = 1$, and the joint density of $Z_1, \cdots, Z_{m-1}$ is

$$f_{Z_1,\cdots,Z_{m-1}}(z_1, \cdots, z_{m-1}) = (n-1)! \sum_{j=1}^{m} z_j^{\alpha_j - 1}/(\alpha_j - 1)! \tag{3.5}$$

where $z_m = 1 - \sum_{j=1}^{m-1} z_j$, $z_j > 0, j = 1, 2, \cdots, m-1$, $\sum_{j=1}^{m-1} z_j \leqq 1$. One finds then

$$E[\Lambda(\alpha)]^t = \{(n-1)! n^{tn}/[(t+1)n - 1]!\} \cdot \prod_{j=1}^{m} \{[(t+1)\alpha_j - 1]!/(\alpha_j - 1)! \alpha_j^{t\alpha_j}\}. \tag{3.6}$$

For fixed x, define coefficients $q_{t,r} = q_{t,r}(x)$ for nonnegative integers r and $t = 0, 1, \cdots, r+1$ by $\sum_{t=0}^{r+1} q_{t,r} y^t = 1 - y - (\frac{1}{2}) \sum_{t=1}^{r} [P_{t+1}(y) - P_{t-1}(y)] P_t(x)$. Then (3.2), (3.3), and (3.4) yield

$$F(x) = \lim_{r\to\infty} \sum_{m=1}^{n} \sum_{k\varepsilon\mathcal{K}_m} \sum_{t=0}^{r+1} q_{tr} E[\Lambda(\alpha)]^t / \prod_{i=1}^{m} k_i! i^{k_i}, \tag{3.7}$$

$E[\Lambda(\alpha)]^t$ being given by (3.6).

The exact values of $F(\lambda)$ in Table 1 were obtained by averaging for $r = 6, 7, 8$, and 9 in (3.7). (For $r = 10$ the round off error, even using a double-precision routine, became too large).

Table 2 was compiled by interpolation. We remark that a closed formula is available for $n = 2$:

$$F(\lambda) = (\tfrac{1}{2})(1 - (1 - \lambda)^{\frac{1}{2}}), \quad 0 \leqq \lambda < 1. \tag{3.8}$$

Indeed, this formula was used for $n = 2$ rather than the calculations described above for Tables 1 and 2. For $n = 2$ the calculations described were found to give values of $F(\lambda)$ agreeing with the formula (3.8) to within 0.001 except near $\lambda = 1$.

Table 1 indicates that the approximation by $\bar{\chi}^2$ is surprisingly good for $n = 2$, the maximum discrepancy being about 0.027. On the other hand, the fit appears not very much better for $n = 10$. In all cases it appears that $F(\lambda) < P[\bar{\chi}_r^2 > -2 \log \lambda]$, except near $\lambda = 1$, so that tests based on significance levels obtained from $\bar{\chi}^2$ tables rather than exact tables would be the opposite of conservative. For example, for $n = 10$ the significance level corresponding to an observed $\lambda = 0.03$ would be 0.05 rather than 0.04 as indicated by $\bar{\chi}^2$ tables. Still the approximation is clearly accurate enough to be useful.

TABLE 1

$F(\lambda) = P[\Lambda \leqq \lambda]$ *and* $G(\lambda) = P[\bar{\chi}_n^2 > -2 \ln \lambda]$

λ	$n = 2$ $F(\lambda)$	$G(\lambda)$	$n = 5$ $F(\lambda)$	$G(\lambda)$	$n = 10$ $F(\lambda)$	$G(\lambda)$
0.00	.000	.000	.000	.000	.000	.000
0.05	.013	.007	.046	.032	.083	.063
0.10	.026	.016	.082	.062	.143	.116
0.20	.053	.036	.153	.123	.243	.209
0.30	.082	.060	.218	.184	.328	.295
0.40	.113	.088	.282	.247	.408	.375
0.50	.146	.120	.345	.312	.482	.452
0.60	.184	.156	.410	.379	.553	.528
0.70	.226	.199	.478	.451	.624	.603
0.80	.276	.252	.550	.529	.694	.680
0.90	.342	.323	.635	.621	.771	.762

TABLE 2

Values of λ *defined by* $\alpha = F(\lambda)$

α	$n = 2$	3	4	5	6	7	8	9	10
.001	.004	.002	.001	.001	.000	.000	.000	.000	.000
.002	.008	.004	.002	.002	.001	.001	.000	.000	.000
.005	.020	.010	.004	.004	.002	.002	.001	.001	.001
.010	.040	.020	.011	.009	.006	.005	.004	.004	.003
.015	.058	.030	.018	.015	.011	.009	.006	.006	.006
.020	.076	.040	.025	.020	.016	.014	.010	.010	.009
.050	.175	.103	.070	.057	.046	.039	.035	.030	.026
.100	.390	.214	.153	.123	.101	.089	.077	.070	.064
.200	.760	.426	.330	.274	.235	.205	.185	.169	.155

TABLE 3

Values of λ *defined by* $\alpha = P[\bar{\chi}_n^2 > -2 \ln \lambda]$

α	$n = 2$	3	4	5	6	7	8	9	10
.001	.008	.004	.002	.001	.001	.001	.001	.001	.000
.002	.016	.007	.004	.003	.002	.002	.001	.001	.001
.005	.036	.017	.011	.008	.006	.005	.004	.003	.003
.010	.067	.033	.021	.015	.012	.010	.008	.007	.006
.015	.095	.048	.032	.023	.018	.015	.013	.011	.010
.020	.121	.064	.042	.031	.025	.020	.017	.015	.013
.050	.259	.148	.104	.080	.065	.055	.048	.042	.038
.100	.440	.275	.203	.162	.136	.117	.103	.092	.084
.200	.702	.494	.390	.325	.281	.249	.225	.205	.189

REFERENCES

BARTHOLOMEW, D. J. (1959a). A test of homogeneity for ordered alternatives. *Biometrika* **46** 36–48.

BARTHOLOMEW, D. J. (1959b). A test of homogeneity for ordered alternatives. II. *Biometrika* **46** 328–335.

BARTHOLOMEW, D. J. (1961). A test of homogeneity of means under restricted alternatives. *J. Roy. Statist. Soc. Ser. B* **23** 239–281.

BOSWELL, M. T. (1966). Estimating and testing trend in a stochastic process of Poisson type. *Ann. Math. Statist.* **37** 1564–1573.

BRUNK, H. D. (1955). Maximum likelihood estimation of monotone parameters. *Ann. Math. Statist.* **26** 607–616.

BRUNK, H. D. (1958). On the estimation of parameters restricted by inequalities. *Ann. Math. Statist.* **29** 437–454.

BRUNK, H. D. (1960). On a theorem of E. Sparre Andersen and its application to tests against trend. *Math. Scand.* **8** 305–326.

CHERNOFF, H. (1956). Large sample theory, parametric case. *Ann. Math. Statist.* **27** 1–22.

MADOW, W. G. (1940). Limiting distributions of quadratic and bilinear forms. *Ann. Math. Statist.* **11** 125–146.

MANN, H. B. and WALD, A. (1943). On stochastic limit and order relationships. *Ann. Math. Statist.* **14** 217–226.

MCSHANE, E. J. and BOTTS, T. A. (1959). *Real Analysis*. Van Nostrand, Princeton.

MILES, R. E. (1959). The complete amalgamation into blocks, by weighted means, of a finite set of real numbers. *Biometrika* **46** 317–327.

PROKHOROV, Y. V. (1956). Convergence of random processes and limit theorems in probability theory. (Russian). *Teorija Verojatnostei i ee Primenenija* **1** 177–238.

SVERDRUP, E. (1952). The limit distribution of a continuous function of random variables. *Skand. Akt.* 1–10.