ESTIMATING AND TESTING TREND IN A STOCHASTIC PROCESS OF POISSON TYPE[1]

By M. T. Boswell[2]

University of Missouri

1. Introduction and summary. Let $\{T_i: i = 1, 2, \cdots\}$ be a stochastic process of Poisson type, with $\lambda(t)$, the rate of occurrence of the events, depending on time. We may interpret T_i as the time of occurrence of the ith event. In Section 2, starting with the joint density function of $T_1, \cdots, T_n$, the maximum likelihood estimate of $\lambda(t)$ subject to $0 \leqq \lambda(t) \leqq M$ being a non-decreasing function of time (M some positive number) is found.

In Section 3, starting with the conditional joint density of $T_1, \cdots, T_n$ given there are n events in $(0, t^*]$, the conditional maximum likelihood estimate of $\lambda(t)$ subject to $0 \leqq \lambda(t)$ being a non-decreasing function of time is found. In Section 4, the conditional likelihood ratio test of the hypothesis that $\lambda(t)$ is constant against the alternate hypothesis that $\lambda(t)$ is not constant but is non-decreasing is found, and a limiting distribution is found which may be used to approximate the probability of a type I error for large sample size.

Theorem 2.1 (Brunk-van Eeden), I believe is important in its own right. It is contained in the works of Brunk and van Eeden, although it is not explicitly stated. This theorem can be used as a basis for tests of hypotheses for constant parameters against increasing parameters or for increasing parameters against all other alternatives.

2. The maximum likelihood estimate of $\lambda(t)$. Let $T_1, T_2, \cdots$ be the times of occurrence of a stochastic process of Poisson type. It is known for such a process that the joint density function of the first n times is

$$f_{T_1,\cdots,T_n}(t_1, \cdots, t_n) = [\exp\{-\Lambda(t_n)\}]\prod_{k=1}^{n}\lambda(t_k), \tag{2.1}$$

where $\Lambda(t) = \int_0^t \lambda(u)\,du$ and where $\lambda(t)$ is the rate of occurrence of the Poisson events. The problem is to find a function $\lambda(t)$ which maximizes (2.1) for fixed $t_1, \cdots, t_n$ subject to

$$0 \leqq \lambda(t) \quad \text{is non-decreasing.} \tag{2.2}$$

However this problem as stated has no solution since (2.1) can be made arbitrarily large by setting $\lambda(t) = 0$ for $t < t_n$ and setting $\lambda(t_n)$ arbitrarily large. We assume $\lambda(t) \leqq M$ for some fixed positive number M. The product $\prod_{k=1}^{n}$

Received 9 June 1965; revised 11 April 1966.

[1] This paper is based on the author's dissertation submitted in partial fulfillment of the requirements for the Ph.D. degree in Mathematics at the University of California, Riverside, January, 1965. The research was supported in part by the U. S. Army Research Office under contract DA-31-124-ARO(D)-158.

[2] Now at The Pennsylvania State University.

$\lambda(t_k)$ is unaffected by values $\lambda(t)$ for $t \neq t_k$, $k = 1, \cdots, n$. Thus if we know $\lambda(t_k)$, $k = 1, 2, \cdots, n$, we need find the rest of the values $\lambda(t)$ which minimize the area, $\Lambda(t_n)$, under $\lambda(t)$ between 0 and t_n, subject of course, to the restrictions (2.2). This occurs if

$$\begin{aligned} \lambda(t) &= 0 && \text{if } 0 \leqq t < t_1 \\ &= \lambda(t_k) && \text{if } t_k \leqq t < t_{k+1}, \quad k = 1, 2, \cdots, n-1 \\ &= \lambda(t_n) && \text{if } t_n \leqq t. \end{aligned}$$

Furthermore $\lambda(t_n) = M$. Therefore the problem reduces to finding $(x_1, \cdots, x_{n-1})$ which maximizes

$$[\exp(-\textstyle\sum_{k=1}^{n-1} a_k x_k)][\prod_{k=1}^{n-1} x_k] = \prod_{k=1}^{n-1} x_k \exp(-a_k x_k) \tag{2.3}$$

subject to

$$0 \leqq x_1 \leqq \cdots \leqq x_{n-1} \leqq M, \tag{2.4}$$

where $a_k = t_{k+1} - t_k$ and $x_k = \lambda(t_k)$, $k = 1, 2, \cdots, n-1$. We will need the following theorem of Brunk-van Eeden.

THEOREM 2.1. *Let $f_n(\theta)$ be a function unimodal in θ with a unique maximum at θ_n^*, $n = 1, 2, \cdots$. If the product $\prod f_j(\theta)$ of any finite number of these functions is unimodal with a unique maximum, then $(\hat{\theta}_1, \cdots, \hat{\theta}_n)$ maximizes $\prod_{k=1}^{n} f_k(\theta_k)$ subject to $0 \leqq \theta_1 \leqq \cdots \leqq \theta_n$ if*

$$\hat{\theta}_k = \max_{1 \leqq \alpha \leqq k} \min_{k \leqq \beta \leqq n} M(\alpha, \beta) \tag{2.5}$$

where $M(\alpha, \beta)$ is the maximizing value of $\prod_{k=\alpha}^{\beta} f_k(\theta)$.

PROOF. This theorem is not stated explicitly in the works of Brunk and van Eeden, but may be extracted in the following manner. In [6] van Eeden makes the assumptions of the above theorem but allows a partial ordering of the parameters $\theta_1, \cdots, \theta_n$. She finds a procedure for maximizing $\prod_{k=1}^{n} f_k(\theta_k)$. In [3] Brunk assumes an "exponential family" of functions $f_k(\theta)$. His result is the form (2.5). Then in [7] van Eeden proves that her result is under the conditions of the above theorem equivalent to Brunk's result (i.e. the same as (2.5)).

COROLLARY 2.1. *Under the hypoiheses and definitions of Theorem 2.1, $(\theta_1, \cdots, \theta_n)$ maximizes $\prod_{k=1}^{n} f_k(\theta)$ subject to $0 \leqq \theta_1 \leqq \cdots \leqq \theta_n$ if $\theta_k = \min\{\hat{\theta}_k, M\}$.*

PROOF. This theorem is easy to see from the fact that the product of any finite number of the $\{f_k\}$ is unimodal with a unique maximum.

THEOREM 2.2. *The maximum-liklihood estimate of $\lambda(t)$ over $(0, t_n]$ where $\lambda(t)$ satisfies (2.2) is*

$$\begin{aligned} \hat{\lambda}(t) &= 0 && \text{if } 0 \leqq t < t_1 \\ &= \min\{M, \hat{\lambda}(t_k)\}, && \text{if } t_k \leqq t < t_{k+1}, \ k = 1, \cdots, n-1, \\ &= M && \text{if } t = t_k \end{aligned} \tag{2.5}$$

where $\hat{\lambda}(t_k) = \max_{1 \leqq \alpha \leqq k} \min_{k \leqq \beta \leqq n} (\beta - \alpha + 1)/(a_\alpha + \cdots + a_\beta)$.

PROOF. From Corollary 2.1, setting $f_k(x) = \exp(\alpha_k x)$, one can show that $\hat{\lambda}(t_k) = \max_{1 \leqq \alpha \leqq k} \min_{k \leqq \beta \leqq n} (\beta - \alpha + 1)/(a_\alpha + \cdots + a_\beta)$ for $k = 1, 2, \cdots, n - 1$. The functions $\{f_k\}$ satisfy the hypotheses of Corollary 2.1; for example $\prod_{k=\alpha}^{\beta} f_k(x) = x^{\beta-\alpha+1} \exp[-(a_\alpha + \cdots + a_\beta)x]$ is unimodal with a unique maximum at $x = (\beta - \alpha + 1)/(a_\alpha + \cdots + a_\beta)$. The result then follows from the fact that the product of any number of the f_k's is unimodal.

3. The conditional maximum likelihood estimate of $\lambda(t)$. Let $T_1, T_2, \cdots$ be as before. It is known that the conditional joint density of the first n time points given there are n occurrences in time $(0, t^*]$ is

$$f_{T_1,\cdots,T_n}(t_1, \cdots, t_n \mid n) = n![\prod_{k=1}^{n} \lambda(t_k)]/\Lambda^n(t^*), \tag{3.1}$$

if $0 \leqq t_1 \leqq t_2 \leqq \cdots \leqq t_{n+1} \equiv t^*$. Once again the problem is to find a function $\lambda(t)$, $0 \leqq t \leqq t^*$ which maximizes (3.1) subject to $\lambda(t)$ being non-negative and non-decreasing for fixed $t_1, \cdots, t_n$. Reasoning similar to that used in Section 2 tells us that we need only find values of λ at t_k, $k = 1, 2, \cdots, n$. Then

$$\begin{aligned} \hat{\lambda}(t) &= 0 && \text{if } 0 \leqq t < t_1 \\ &= \hat{\lambda}(t_k) && \text{if } t_k \leqq t < t_{k+1}, \qquad k = 1, 2, \cdots, n. \end{aligned}$$

The value of λ at t^* is immaterial. Thus the problem reduces to finding values of $x_1, \cdots x_n$ which maximize

$$f(x) = [\prod_{k=1}^{n} x_k]/[\sum_{k=1}^{n} a_k x_k]^n \tag{3.2}$$

subject to $0 \leqq x_1 \leqq x_2 \leqq \cdots \leqq x_n$, where $x_k = \hat{\lambda}(t_k)$ and $a_k = t_{k+1} - t_k$, $k = 1, 2, \cdots, n$. We observe that if $(x_1, \cdots, x_n)$ maximizes (3.2), then so does $(cx_1, \cdots, cx_n)$ for any $c > 0$.

LEMMA 3.1. *If* $x_k = \max_{1 \leqq \alpha \leqq k} \min_{k \leqq \beta \leqq n} [(\beta + 1 - \alpha)/(a_\alpha + \cdots + a_\beta)] \cdot (c/n)$, $k = 1, 2, \cdots, n$. *Then* $(x_1, \cdots, x_n)$ *maximizes* $\prod_{k=1}^{n} x_k$ *subject to* $0 \leqq x_1 \leqq x_2 \leqq \cdots \leqq x_n$ *and to* $\sum_{k=1}^{n} a_k x_k = c$.

PROOF. Grenander in [5] pp. 138–140, by the use of Lagrange multipliers, solves this problem with the sense of the inequalities reversed and with $c = 1$. The lemma follows by a change of scale and a change in the labeling of the variables.

THEOREM 3.1. *A maximizing point for* (3.2) *is* $(\bar{x}_1, \cdots, \bar{x}_n)$ *where*

$$\bar{x}_k = \max_{1 \leqq \alpha \leqq k} \min_{k \leqq \beta \leqq n} [(\beta - \alpha + 1)/(t_{\beta+1} - t_\alpha)] \cdot (c/n),$$

for any $c > 0$. *Furthermore* $\sum_{k=1}^{n} \alpha_k \bar{x}_k = c$.

PROOF. The proof of this theorem is a direct application of Lemma 3.1, by maximizing (3.2) restricted to the hyperplane $\sum_{k=1}^{n} \alpha_k x_k = c$. The value of (3.2) turns out to be independent of the value c.

4. The conditional likelihood-ratio test against trend. Let H_0 be the composite hypothesis that $\lambda(t) = \lambda > 0$, in which case the density function (3.1) becomes

$$f_{T_1}, \cdots, {}_{T_n}(t_1, \cdots, t_n \mid n) = n!/(t^*)^n$$

if $0 \leqq t_1 < t_2 < \cdots < t_n \leqq t^*$. Let H_1 be the hypothesis that $\lambda(t)$ is not constant, is non-negative, and is non-decreasing. Then the likelihood-ratio test of H_0 against H_1 tells us to reject H_0 if

$$[\sup_{H_0} n!/(t^*)^n]/[\sup_{H_1} n!\lambda(t_1)\lambda(t_2) \cdots \lambda(t_n)/\Lambda^n(t^*)] \leqq k_1 ,$$

where k_1 is made as small as possible consistent with the level of significance. Here is a test of Neyman structure. For the question of monotonicity of λ, a constant of proportionality is a nuisance parameter which is eliminated by the use of the sufficient statistic n (i.e. we will take c equal to n). Theorem 3.1 gives us

$$\hat{\lambda}(t_k) = \bar{x}_k = \max_{1 \leqq \alpha \leqq k} \min_{k \leqq \beta \leqq n} (\beta - \alpha + 1)/(t_{\beta+1} - t_\alpha), \tag{4.1}$$

and $\Lambda(t^*) = n$. Therefore we reject H_0 if

$$1/\prod_{k=1}^{n} \bar{x}_k \leqq k_0 . \tag{4.2}$$

We wish to find, at least approximately, $P[\prod_{k=1}^{n} (\bar{x}_k^{-1}) \leqq k_0]$ for various k_0. In order to do this we will introduce notation most of which comes from [3].

Definitions and Notation. It is known that the $\bar{x}_k$, $k = 1, \cdots, n$, may be calculated by the following iterative procedure:

$$\bar{x}_1 = \min_{1 \leqq \beta \leqq n} \beta/(t_{\beta+1} - t_1) = \alpha_1/(t_{\alpha_1+1} - t_1)$$

for some integer α_1, and then $\bar{x}_k = \bar{x}_1$ for $k = 2, 3, \cdots, \alpha_1$;

$$\bar{x}_{\alpha_1+1} = \min_{\alpha_1+1 \leqq \beta \leqq n} (\beta - \alpha_1)/(t_{\beta+1} - t_{\alpha_1+1})$$
$$= \alpha_2/(t_{\alpha_1+\alpha_2+1} - t_{\alpha_1+1})$$

for some integer α_2, and then $\bar{x}_k = \bar{x}_{\alpha_1+1}$ for $k = \alpha_1 + 2, \cdots, \alpha_1 + \alpha_2$, etc. Let $\bar{X}_k$ be the random variable formed by replacing $(t_1, \cdots, t_n)$ by $(T_1, \cdots, T_n)$ in (4.1), the formula for $\bar{x}_k$. Let $\bar{x} = (\bar{x}_1, \cdots, \bar{x}_n)$ be an observation of the random vector $X = (\bar{X}_1, \cdots, \bar{X}_n)$. For this $\bar{x}$ let $\alpha_1, \alpha_2, \cdots, \alpha_m, m$ be the numbers found in the iterative procedure given above. Here m is an observation on the random variable M, the number of distinct values which the components of $\bar{X}$ assume. Let A_k be the number of components of $\bar{X}$ which are equal to the kth distinct value, and let $\mathcal{A}_m$ be the set of all possible outcomes of $A = (A_1, \cdots, A_m)$ when $M = m$. Let $K = (K_1, \cdots, K_n)$ be the random vector where K_i is the number of components of A which are equal to i, $i = 1, 2, \cdots, n$. Let $\mathcal{K}_m$ be the set of all possible outcomes of K for which $\sum_{j=1}^{n} k_j = m$; that is $\mathcal{K}_m$ is the set of all ordered n-triples $(k_1, \cdots, k_n)$ of non-negative integers such that $\sum_{j=1}^{n} k_j = m$ and $\sum_{j=1}^{n} jk_j = n$. Finally let $(\alpha_1, \cdots, \alpha_m)$ be an outcome of A, and let $a_k = \alpha_1 + \alpha_2 + \cdots + \alpha_k$, $k = 1, 2, \cdots, m$.

Intuitively the outcome of $\bar{X}$ consists of groups of equal values; the outcome of K is a specification of the number of elements which are in the various groups without regard to the order of the groups. The outcome of A is a specification of the number of elements in each of the groups as well as the order of the groups.

Further $\mathcal{K}_m$ is the collection of all possible outcomes of K with m the number of groups $(M = m)$.

From the iterative procedure for computing $\bar{X}$, we see that $\bar{X}_{a_k} = \alpha_k/(T_{a_k+1} - T_{a_{k-1}+1})$. We will denote the reciprocal of this by U_k. Let $Y_k = T_{k+1} - T_k$. Then U_k is the average of the $\{Y_j\}$ for $j = a_{k-1}+1, a_{k-1}+2, \cdots, a_k$. Given an outcome α for A and an outcome y of $Y = (Y_1, \cdots, Y_n)$ we define $u(\alpha, y) = (u_1, \cdots, u_m)$ by

$$u_k = (y_{a_{k-1}+1} + y_{a_{k-1}+2} + \cdots + y_{a_k})/(a_k - a_{k-1}). \tag{4.3}$$

If the components of $u(\alpha, y)$ are increasing, we may think of $u(\alpha, y)$ as an outcome of $U = (U_1, \cdots, U_n)$. Let $\mathfrak{A}^k$ be the collection of all possible outcomes of A when $K = k$. Then $\mathfrak{A}_m = \bigcup_{k\varepsilon\mathcal{K}_m}\mathfrak{A}^k$. Let $C(\alpha)$ be the set of all y for which $u(\alpha, y)$ has increasing components. Let $B_k(\alpha)$ be the set of all y for which u_k is less than the average of $\{y_j\}$ for $j = a_{k-1}+1, a_{k-1}+2, \cdots, a_{k-1}+r$ for any r such that $1 \leqq r \leqq a_k - 1$. Let $B(\alpha) = \bigcap_{k=1}^{m} B_k(\alpha)$, let $D(\alpha) = B(\alpha) \cap C(\alpha)$, let $D^k = \bigcup_{\alpha\varepsilon\mathfrak{A}^k} D(\alpha)$, and let

$$H(\alpha) = \{y: \textstyle\prod_{k=1}^{m}(u_k)^{\alpha_k} \leqq k_0\}. \tag{4.4}$$

We observe that an outcome of Y completely determines an outcome of $\bar{X}$. Furthermore, using the iterative procedure for computing $\bar{x}$, we see that $D(\alpha)$ is the collection of all possible outcomes of Y which determine outcomes of $\bar{X}$ with a corresponding α.

Lemma 4.1. *Let* $H^k = \bigcup_{\alpha\varepsilon\mathfrak{A}^k} H(\alpha) \cap D(\alpha)$. *Then*

$$P[\text{Rej. } H_0] = \textstyle\sum_{m=1}^{n} \sum_{k\varepsilon\mathcal{K}_m} P[Y \,\varepsilon\, H^k]. \tag{4.5}$$

Proof. Now $P[\text{Rej. } H_0] = P[\prod_{k=1}^{n}(\bar{X}_k)^{-1} \leqq k_0] = P[\bigcup_{m=1}^{n} \bigcup_{\alpha\varepsilon\mathfrak{A}_m} \{\prod_{k=1}^{n}(\bar{X}_k)^{-1} \leqq k_0, A = \alpha\}] = \sum_{m=1}^{n} \sum_{\alpha\varepsilon\mathfrak{A}_m} P[\prod_{k=1}^{n}(\bar{X}_k)^{-1} \leqq k_0, A = \alpha]$. Recall the $\bar{X}_k$'s are groups of constant values (equal to $(U_k)^{-1}$'s). Therefore if $A = \alpha$, $\prod_{k=1}^{n} (\bar{X}_k)^{-1} = \prod_{k=1}^{m}(U_k)^{\alpha_k}$. Then in the notation defined above $P[\prod_{k=1}^{n}(U_k)^{\alpha_k} \leqq k_0, A = \alpha] = P\{[\prod_{k=1}^{n}(U_k)^{\alpha_k} \leqq k_0] \cap D(\alpha)\} = P[H(\alpha) \cap D(\alpha)]$. The conclusion follows from the definition of H^k.

It is known that $Y_1, \cdots, Y_n$ are exchangeable; that is for every Borel set $J \subseteq R^n$ and every permutation operator p we have $P[Y \,\varepsilon\, J] = P[pY \,\varepsilon\, J]$. We need the following definitions, which come directly from [3].

Definition 4.1. For $m = 1, \cdots, n$, we denote by Π_m the set of all permutations $\pi: (1, \cdots, m) \rightarrow (i_1, \cdots, i_m)$. We denote also by π the permutation operator carrying an ordered m-tuple

$$w = (w_1, \cdots, w_m) \quad \text{into} \quad \pi w = (w_{i_1}, w_{i_2}, \cdots, w_{i_m})$$

and by Π_m the class of such permutation operators.

Definition 4.2. For $m \leqq n$, $\alpha \,\varepsilon\, \mathfrak{A}_m$, $\pi \,\varepsilon\, \Pi_m$, let $u = u(\alpha, y)$ and think of the coordinates of y appearing in the definition of $u_j(\alpha, y)$ as being written in the order of increasing index j, $j = 1, 2, \cdots, m$. Let $j_1, j_2, \cdots j_n$ be the indices of the coordinates of y in the order in which they appear when the u_j are rearranged to form $\pi u = (u_{i_1}, u_{i_2}, \cdots, u_{i_m})$ without rearranging the coordinates of y within u_{ij}, $(j = 1, 2, \cdots, m)$. Let $p = p(\alpha, \pi)$ carry $y = (y_1, \cdots, y_n)$ into $py = (y_{j_1}, y_{j_2}, \cdots, y_{j_n})$.

DEFINITION 4.3. The class $\{H(\alpha)\}$, $\alpha \varepsilon \mathfrak{A}$ will be called π-invariant if $\alpha \varepsilon \mathfrak{A}_m$, $\pi \varepsilon \Pi_m$, $p = p(\alpha, \pi)$ imply $pH(\alpha) = H(\pi\alpha)$.

LEMMA 4.2. *If $Y_1, \cdots, Y_n$ are exchangeable, if $\{H(\alpha)\}$, $\alpha \varepsilon \mathfrak{A}$ is π-invariant, if $m \leqq n$, $k \varepsilon \mathfrak{K}_m$, $\alpha \varepsilon \mathfrak{A}^k$, and if $H(\alpha)$ is cyclically symmetric in each of*

$$Y_1, \cdots, Y_{a_1}; Y_{a_1+1}, \cdots, Y_{a_2}; \cdots; Y_{a_{m-1}+1}, \cdots, Y_{a_m},$$

then $P(H^k) = P[H(\alpha)] \prod_{k=1}^n (k_i ! \, i^{k_i})^{-1}$

PROOF. This lemma is proved in [3], p. 319, by Brunk.

LEMMA 4.3. *For α an arbitrary but fixed element of $\mathfrak{A}^k$,*

$$P[Y \varepsilon H^k] = P[Y \varepsilon H(\alpha)] \prod_{k=1}^n (k_i ! \, i^{k_i})^{-1}.$$

PROOF. It is obvious that $H(\alpha)$ is cyclically symmetric. Since $Y_1, \cdots, Y_n$ are exchangeable all that we need show is that $\{H(\alpha)\}$, $\alpha \varepsilon \mathfrak{A}$ is π-invariant. This too is fairly obvious and can be seen by looking closely at the definition of $p(\alpha, \pi)$ in Definition 4.2 and at (4.4), the definition of $H(\alpha)$.

Combining Lemma 4.1 and Lemma 4.3, we have:

THEOREM 4.1. *For an arbitrary but fixed element α of $\mathfrak{A}^k$*

$$P[\text{Rej. } H_0] = \sum_{m=1}^n \sum_{k \varepsilon \mathfrak{K}_m} P[Y \varepsilon H(\alpha)] \prod_{k=1}^n (k_i ! \, i^{k_i})^{-1}. \tag{4.6}$$

Recall $H(\alpha) = \{y: \prod_{k=1}^m (U_k)^\alpha \leqq k_0\}$. The above theorem results in a large saving in the work necessary to calculate the probability of a type I error, but the work left often is large.

DEFINITION 4.4. Let $U(\alpha) = \prod_{k=1}^m (U_k)^{\alpha_k}$, and let $U(A)$ be the random variable whose observation is the observed value of $U(\alpha)$ when $A = \alpha$ ($U(\alpha)$ is a random variable).

We will reject H_0 if $U(A) \leqq k_0$; we would like to approximate $P[U(A) \leqq k_0]$. In order to do this we will find the limiting moment sequence of $U(A)$. Let $Z_k = \alpha_k U_k$. Starting from the conditional joint density function of $T_1, \cdots, T_n$ (or $Y_1, \cdots, Y_n$) given n, one can show by the usual change of variable techniques that the conditional joint density function of $Z_1, \cdots, Z_n$ given n is

$$f_{Z_1, \cdots, Z_n}(z_1, \cdots, z_n \mid n) = n ! \, (t^*)^{-n} \prod_{k=1}^m [(z_k)^{\alpha_k - 1}/(\alpha_k - 1) !],$$

if $0 \leqq z_k$, $k = 1, 2, \cdots, m$ and if $\sum_{k=1}^m Z_k \leqq t^*$. Using the above joint density function one finds

$$E[U^j(\alpha)] = \{n ! \, (t^*)^{jn}/[(j+1)n] !\} \prod_{k=1}^m \cdot \{[(j+1)\alpha_k - 1] ! / (\alpha_k - 1) ! \, (\alpha_k)^{j\alpha_k}\}. \tag{4.8}$$

In the notation of [3], for $\alpha \varepsilon \mathfrak{A}_m$, $\omega \varepsilon E^m$, let $\nu = ((\alpha_1, \omega_1), (\alpha_2, \omega_2), \cdots, (\alpha_m, \omega_m))$. We define $f_m(\nu, y) \equiv f_m(\nu) = \prod_{k=1}^m (\omega_k)^{j\alpha_k}$ where j is a fixed positive integer. For an observation y of Y we define $\nu(\alpha, y) = ((\alpha_1, u_1), (\alpha_2, u_2), \cdots, (\alpha_m, u_m))$. We also define the vector $\alpha(Y)$ to be the value of A if $Y = y$ and define $m(y)$ to be the value of M if $Y = y$. Let $G = \{y: f_{m(y)}(\nu[\alpha(y), y]) < q\}$, and for $\alpha \varepsilon \mathfrak{A}_m$ let $G(\alpha) = \{y: f_m(\nu[\alpha, y]) < q\}$. We observe that $U^j(\alpha) = f_m(\nu(\alpha, y))$ and that $f_m(\nu, y) \equiv f_m(\nu)$ is symmetric in the components of ν and, vacuously, in the components of y.

LEMMA 4.3. *Let $Y_1, \cdots, Y_n$ be exchangeable, and let $f_m(\nu, y)$ be symmetric in the components of ν and the components of y for $m = 1, 2, \cdots, n$. If the common distribution function of $\{Y_i: i = 1, 2, \cdots, n\}$ is continuous, then for $m < n$, $k \varepsilon \mathcal{K}_m$, and α chosen arbitrarily from $\mathfrak{A}^k$, the conditional distribution of $f_M(\nu(A, Y), Y)$ given D^k, is the distribution of $f_m(\nu(\alpha, Y), Y)$.*

PROOF. Brunk in [3] pp. 321, 322 (Theorem 2.1) proves this lemma. However, instead of exchangeable, Brunk assumes $Y_1, \cdots, Y_n$ are independent and identically distributed. However, he uses this assumption only to prove that $Y_1, \cdots, Y_n$ are exchangeable.

THEOREM 4.2. *For an arbitrary α in $\mathfrak{A}^k$,*

$$E[U^j(A) \mid D^k] = E[U^j(\alpha)]. \tag{4.9}$$

PROOF. The proof is a simple application of Lemma 4.3.

This result will be used in obtaining the limiting distribution of the likelihood statistic, $U(A)$. To this end we require some observations on the asymptotic properties of randomly chosen permutations. Also, we will need the following lemma.

LEMMA 4.4. *For $K \varepsilon \mathcal{K}_m$, $P(D^k) = \prod_{i=1}^n (k_i ! \, i^{k_i})^{-1}$.*

PROOF. If in Lemma 4.2 we let $H(\alpha)$ be the whole space for each α, then it can easily be shown that $\{H(\alpha)\}$, $\alpha \varepsilon \mathfrak{A}$ is π-invariant. The conclusion follows.

E. Sparre Andersen in [1] found that the distribution of M coincides with that of the number of cycles in a random permutation of the first n positive integers. For any y in the space of all n-tuples of distinct positive integers less than or equal to n, let a_1 be the index of the smallest coordinate, 1; let a_2 be the index of the smallest coordinate with index greater than a_1; etc. The process ends when for some $m \leqq n$, $a_m = n$. Let $a_0 = 0$, and let $\alpha_k = a_k - a_{k-1}$, $k = 1, 2, \cdots, m$. Since y is a random permutation of the first n positive integers, we may think of $\alpha = (\alpha_1, \cdots, \alpha_m)$ as the specification of the lengths of the cycles, in the order that they appear, of the permutation which carries $(1, 2, \cdots, n)$ into y, where the first cycle contains 1; the second cycle contains the smallest integer not contained in the first; etc.

This example was also considered by Brunk in [3]; it comes about by defining $u(\alpha, y)$ by $u_j = y_{a_j}$ instead of (4.3). All the theory developed so far still holds with this new definition of $u(\alpha, y)$.

Let W_j be the indicator function of the event that a cycle ends on the jth term, and let $X_j = W_{n-j+1}$, $j = 1, 2, \cdots, n$. Then as Feller points out (cf. [4] pp. 205, 206) $P[X_j = 1] = j^{-1}$, $P[X_j = 0] = 1 - j^{-1}$, $j = 1, 2, \cdots$, and $\{X_j\}$ are independent. Let $1 < k_1 < k_2 < \cdots < k_r \leqq n$ be r positive integers and let

$S_n = \sum_{k=1}^{n} X_k$. It is easy to show since the X_i's are independent that

$$P[X_{k_1} = \cdots = X_{k_r} = 1,\ S_n = r + 1] = n^{-1} \prod_{j=1}^{r} (k_j - 1)^{-1}. \tag{4.10}$$

We observe that

$$P[S_n = r] = n^{-1} \sum_{\Omega} \prod_{j=1}^{r-1} (k_j - 1)^{-1}, \tag{4.11}$$

where Ω is the set of all $(k_1, \cdots, k_{r-1})$ such that $1 < k_1 < \cdots < k_{r-1} \leqq n$.

LEMMA 4.5. *Let*

$$Q_r = \int_r^n \int_{r-1}^{x_r} \cdots \int_1^{x_2} (x_1 x_2 \cdots x_r)^{-1}\, dx_1\, dx_2 \cdots dx_r,$$

where $r < n$ *are fixed positive integers. Then*

(i) $Q_r \leqq (\log n)^r$, *and*

(ii) $Q_r \geqq (\log n)^r / r! - \sum_{j=2}^{r} c_j (\log n)^{r-j}$,

where $c_j = (\log j)^j / j!$.

PROOF. The proof of (i) depends only on replacing all of the lower limits on the integral defining Q_r by 1's. The proof of (ii) depends on the following two facts: first

$$\int_k^{x_{k+1}} [(\log x_k)^{k-1} / (k-1)!\, x_k]\, dx_k = [(\log x_{k+1})^k / k!] - [(\log k)^k / k!]$$

and second

$$\int_r^n \int_{r-1}^{x_r} \cdots \int_{k+1}^{x_{k+2}} (x_{k+1} \cdots x_r)^{-1} dx_{k+1} \cdots dx_r \leqq Q_{r-k}.$$

THEOREM 4.3. *For fixed* $m < n$, $P[S_n = m - 1] / P[S_n = m] \to 0$ *as* $n \to \infty$.

PROOF. Replacing the upper limits by n and the lower limits by 1 in (4.11) one finds that

$$nP[S_n = r] \leqq \left(1 + \sum_{k=2}^{n} k^{-1}\right)^{r-1} \leqq \left(1 + \int_1^n x^{-1} dx\right)^{r-1} = (1 + \log n)^{r-1}.$$

Furthermore $nP[S_n = r] \geqq Q_r$. Using these inequalities and Lemma 4.5 we get the result if $m \geqq 2$. However $S_n \geqq 1$, which gives us the result when $m = 1$.

For $K \in \mathcal{K}_m$ let $d(k)$ be the index of the first nonzero k_i, and let $d(K)$ be the corresponding random variable (i.e. $d(K)$ is the length of the shortest cycle). Let $\mathcal{K}_{m,r} = \{k \in \mathcal{K}_m : d(K) \leqq r\}$.

THEOREM 4.4. *The limit as n goes to infinity of*

$$\sum_{k \in \mathcal{K}_{m,r}} \sum_{\alpha \in \mathcal{a}^k} P[A = \alpha \mid M = m] \equiv \sum_{k \in \mathcal{K}_{m,r}} P[D^k \mid M = m]$$

is zero, where r is an arbitrary but fixed positive integer.

PROOF. The above is equivalent to showing that $P[d(K) \leqq r \mid M = m]$ converges to zero as n goes to infinity, which is itself equivalent to showing that the probability that there exist r adjacent X's whose sum is greater than or equal to 2 given $S_n = m$ goes to zero as n goes to infinity. Let R_c be the event that there exist r adjacent X's whose sum is equal to c. It is sufficient to show that $P[R_c \mid S_n = m]$ converges to zero as n goes to infinity. Recall $P[X_1 = 1] = 1$. Therefore we may write $P[R_c, S_n = m] \leqq L_1 + L_2$ where

$$L_1 = P\left[\sum_{k=2}^{r} X_k \geqq 1,\ S_n = m\right]$$

and $L_2 = \sum_{j=r+1}^{n-r+1} P[\sum_{k=j}^{j+r-1} X_k = c, \sum_{k=2}^{r} X_k = 0, S_n = m]$. One can show that $L_1 \leqq \binom{r}{j} P[S_n = m - j]$ and that $L_2 \leqq \xi \binom{r}{c} P[S_n = m - c]$, where $\xi = \sum_{j=1}^{\infty} (1/j)^2$. The conclusion follows from Theorem 4.3 and the definition of conditional probability (for the rest of this paper we will make the change of scale, $t^* = 1$).

LEMMA 4.6. *Let* $V(A) = n^n U(A)$. *Then for* $j = 1, 2, \cdots$

$$\lim_{d(\alpha)\to\infty} E[V^j(\alpha)] = (j + 1)^{-(m+1)/2}.$$

PROOF. We observe that the components of α go to ∞ as $d(\alpha) \to \infty$. The conclusion then follows from formula (4.8) by using Stirling's approximation for factorials and the fact that $(1 - x^{-1})^{x+a}$ converges to e^{-1} as x goes to infinity.

THEOREM 4.5. *For* $j = 1, 2, \cdots, \lim_{n\to\infty} E[V^j(A) \mid M = m] = (j + 1)^{-(m+1)/2}$.

PROOF. Now

$$\begin{aligned} E[V^j(A) \mid M = m] &= \sum_{k\varepsilon\mathcal{K}_m} \sum_{\alpha\varepsilon\mathfrak{A}^k} E[V^j(\alpha) \mid A = \alpha]P[A = \alpha \mid M = m] \\ &= \sum_{k\varepsilon\mathcal{K}_m} n^{jn} \sum_{\alpha\varepsilon\mathfrak{A}^k} E[U^j(\alpha) \mid A = \alpha]P[A = \alpha]/P[M = m] \\ &= \sum_{k\varepsilon\mathcal{K}_m} n^{jn} E[U^j(A) \mid D^k]P[D^k]/P[M = m]. \end{aligned}$$

Using (4.9) one obtains $E[V^j(A) \mid M = m] = \sum_{k\varepsilon\mathcal{K}_m} E[V^j(\alpha)]P[D^k \mid M = m]$ where α is an arbitrary element of $\mathfrak{A}^k$. This sum can be broken up into a sum over $\mathcal{K}_{m,r}$ and a sum over $\mathcal{K}_m - \mathcal{K}_{m,r}$. One can show that $0 \leqq (n/t^*)^{jn} E[U^j(\alpha)] = (t^*)^{-jn} E[V^j(\alpha)] \leqq 1$. Using this bound and Theorem 4.4 one can show that the sum over $\mathcal{K}_{m,r}$ converges to 0 as n goes to infinity. The conclusion follows from Lemma 4.6 by using an ϵ-type argument with r chosen sufficiently large.

It is known for a sequence of random variables $\{X_n\}$ that if the limit moment sequence uniquely determines a distribution function, then it is the limit as $n \to \infty$ of the distribution functions of $\{X_j\}$ (cf. [8] p. 128). Also if for the limit moment sequence $\{m_j\}$, $\sum m_j c^j/j!$ is absolutely convergent for some $c > 0$, then there is at most one distribution function with these moments (cf. [8] p. 125). The moment sequence $\{(j + 1)^{-(m+1)/2}\}$ satisfies the latter condition with $c = 1$. Therefore we have the following.

THEOREM 4.6. *The limiting distribution given* $M = m$ *of* $V_n \equiv V(A)$ *as* $n \to \infty$ *is given by*

$$\begin{aligned} F(x) &= 0, && \text{if } x \leqq 0 \\ &= 1 - F_{m+1}(-2\cdot\ln x), && \text{if } 0 < x < 1 \\ &= 1, && \textit{otherwise}, \end{aligned}$$

where F_{m+1} *is the* χ^2-*distribution function with* $m + 1$ *degrees of freedom.*

Recall that $V_n(A) = n^n U(A)$ and that we reject H_0 in favor of H_1 if $U(A) \leqq k_0$ or equivalently if $V_n(A) \leqq c(c = n^{-n}k_0)$. We may approximate the probability of a type one error for large n by $\sum_{m=1}^{n} [1 - F_{m+1}(-2\cdot\ln x)]P[M = m]$, where it is known that $P[M = m] = |S_n^m|/n!$, $|S_n^m|$ is a Stirling's number of the first kind (cf. [2] p. 129).

Professor H. D. Brunk pointed out that the limiting distribution found in

Theorem 4.6 is exact if $m = 1$. This can be seen as follows. From Theorem 4.5 we have $E[V^j(A) \mid M = 1] = E[V^j(\alpha)]P[D^k \mid M = 1]$ for $\alpha = (n)$, where the only k in $\mathcal{K}_1$ is $k = (0, \cdots, 0, 1)$ and for this k the only α in $\mathcal{A}^k$ is $\alpha = (n)$. Also for this k, $P[D^k \mid M = 1] = 1$. From (4.8), recall $t^* = 1$, we see that for $\alpha = (n)$, $E[V^j(\alpha)] = n^{jn}E[U^j(\alpha)] = (j + 1)^{-(j+1)/2}$. That is the limit obtained in Theorem 4.5 is exact when $m = 1$, and therefore Theorem 4.6 gives the exact conditional distribution given $M = 1$.

5. Acknowledgment. I wish to express my gratitude to Professor H. D. Brunk for suggesting this problem and for continued help and encouragement. Without his help this paper would not have been possible. Also I wish to thank the referee for his many useful comments and suggestions.

REFERENCES

[1] ANDERSEN, E. SPARRE (1954). On the fluctuations of sums of random variables II. *Math. Scand.* **2** 195-223.

[2] BARTON, D. E. and DAVID, F. D. (1962). *Combinatorial Chance.* Charles Griffin and Co., London.

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

[4] FELLER, WILLIAM (1950). *An Introduction to Probability Theory and Its Applications.* Wiley, New York.

[5] GRENANDER, ULF (1956). On the theory of mortality measurement. *Skand. Aktuarietidskr.* 71–79 and 126–153.

[6] VAN EEDEN, CONSTANCE (1957). Maximum likelihood estimation of partially or completely ordered parameters I and II. *Nederl. Akad. Wetensch Proc. Ser. A* **60**, *Indag. Math.* **19** 128-136, 201-211.

[7] VAN EEDEN, CONSTANCE (1957). Note on two methods for estimating ordered parameters of probability distributions. *Nederl. Akad. Wetensch. Proc. Ser. A* **60**, *Indag. Math.* **19** 506–512

[8] WILKS, SAMUEL S. (1962). *Mathematical Statistics*, Wiley, New York.