MATH. SCAND. 8 (1960), 305–326

ON A THEOREM OF E. SPARRE ANDERSEN AND ITS APPLICATION TO TESTS AGAINST TREND

H. D. BRUNK

0. Introduction and summary.

Let n be a positive integer, to be fixed throughout, and let $X_1,\ldots,X_n$ be random variables. Let

$$S_0 = 0, \qquad S_r = \sum_{i=1}^{r} X_i, \qquad r = 1,2,\ldots,n\,.$$

The lower part of the boundary of the convex hull of the set of points $\{(r,S_r)\}_{r=0}^{n}$ in the cartesian plane is the *greatest convex minorant* of this set of points, and will be referred to as the gcm. E. Sparre Andersen [2] found the distribution of the random variable M defined as the number of sides of the gcm, when the random variables $X_1,\ldots,X_n$ are symmetrically dependent, and observed that it is the same distribution as that of the number of cycles in a randomly chosen permutation of the integers $1,2,\ldots,n$.

The gcm can be described in terms of averages of successive terms of the finite sequence $X_1,X_2,\ldots,X_n$. Determine a_1 so that

$$\sum_{i=1}^{a_1} X_i/a_1 = \min_r \sum_{i=1}^{r} X_i/r\,;$$

determine a_2 so that

$$\sum_{i=a_1+1}^{a_2} X_i/(a_2-a_1) = \min_r \sum_{i=a_1+1}^{a_1+r} X_i/r\,; \qquad \text{etc.}$$

Then the index m for which $a_m=n$ is the number of sides, M, of the gcm. It is found in Section 1 that the random variable M will have the same distribution if the arithmetic averages $\sum_{i=a+1}^{a+r} X_i/r$ are replaced by $E_r(X_{a+1},\ldots,X_{a+r})$ for less specialized functions E_r. The resulting theo-

Received August 17, 1959.

Part of the research for this paper was done while the author was a recipient of a United States Government Award under Public Law 584 of the 79th Congress, The Fulbright Act, and part was supported by the United States Air Force through the Air Force Office of Scientific Research of the Air Research and Development Command under Contract No. AF 18(600)1108.

rem contains Andersen's theorem on the distribution of the number of sides in the gcm, and the theorem giving the distribution of the number of cycles in a randomly chosen permutation of $1, 2, \ldots, n$, as special cases.

In the developing this result, a combinatorial theorem of Spitzer (Theorem 2.1 of [12]) is generalized so as to apply when the functions E_r replace arithmetic averages. This combinatorial theorem may be applied in the same way as was Spitzer's [12] to yield a corresponding generalization of another theorem of Andersen [1].

The principal theorem is extended somewhat so as to cover applications made in Sections 2 and 3. The applications in Section 3 involve only the special case

$$E_r(x_1, \ldots, x_r) \equiv \sum_{i=1}^{r} x_i/r .$$

Section 2 is devoted to the distributions of certain random variables determined by the gcm. Section 3 contains a brief discussion of the application of the result of Section 2 to tests against trend. It contains also a remark as to the possible utility of the number of sides of the gcm as a statistic in a distribution-free test against trend.

The writer gratefully acknowledges the benefit of conversations with Professor E. Sparre Andersen.

1. The principal theorem.

The theorems of this section are largely combinatorial in nature; their statements require the prior introduction of a rather large number of special notations and terms.

DEFINITION 1.1. *Let $x = (x_1, \ldots, x_n)$ denote the generic point of euclidean n-space $\mathscr{E}^n$. Let $E_r(t_1, t_2, \ldots, t_r)$ be a real-valued function of the real variables $t_1, t_2, \ldots, t_r$, $r = 1, 2, \ldots, n$. Let $E_r(x)$ denote $E_r(x_1, x_2, \ldots, x_r)$ when $x \in \mathscr{E}^n$.*

In further definitions, the symbols $\equiv_D$ and $\Leftrightarrow_D$ will be used in defining that which appears on the left of the symbol by that which appears on the right.

DEFINITIONS 1.2. *Sets $\mathscr{K}_m$, $\mathscr{A}^k$, $\mathscr{A}_m$, $\mathscr{A}$, $\mathscr{V}_m$ and tuples a, u, y are defined by* (1.1–7) *below.*

For $m = 1, \ldots, n$, the set $\mathscr{K}_m$ of n-vectors k with non-negative integer-valued coordinates is defined by

$$(1.1) \qquad k \in \mathscr{K}_m \Leftrightarrow_D k = \{k_i\}_{i=1}^{n}, \qquad \sum_{i=1}^{n} ik_i = n, \qquad \sum_{i=1}^{n} k_i = m .$$

(In particular: one may think of k_i as the number of sides of horizontal extent i in the gcm of a set of points $(0,0),(1,s_1),\ldots,(n,s_n)$; then m would be the number of sides of the gcm.)

For $m=1,\ldots,n$, the set $\mathscr{A}^k$ of m-vectors α with positive integer-valued coordinates is defined by

$$(1.2)\qquad \alpha \in \mathscr{A}^k \Leftrightarrow_D \begin{cases} k \in \mathscr{K}_m;\ \alpha = \{\alpha_i\}_{i=1}^m;\ \textit{for any } i = 1,\ldots,n\,, \\ k_i \textit{ of the coordinates of } \alpha \textit{ are equal to } i. \end{cases}$$

It follows from the definition that

$$\sum_{i=1}^m \alpha_i = n \quad \text{and} \quad \prod_{i=1}^m \alpha_i = \prod_{i=1}^n i^{k_i}\,.$$

(In particular: $\alpha_1,\ldots,\alpha_m$ may be thought of as a specification of the horizontal extents of the sides of the gcm referred to above.)

The set $\mathscr{A}_m$, $m=1,\ldots,n$, and the set $\mathscr{A}$ are defined by

$$(1.3)\qquad \mathscr{A}_m \equiv_D \bigcup_{k\in\mathscr{K}_m} \mathscr{A}^k, \qquad \mathscr{A} \equiv_D \bigcup_{m=1}^n \mathscr{A}_m\,.$$

(In particular: $\mathscr{A}_m$ is the set of all α which specify horizontal extents of sides of the gcm when it has m sides.)

For $m=1,\ldots,n$, the set $\mathscr{V}_m$ is defined by

$$(1.4)\quad v \in \mathscr{V}_m \Leftrightarrow \begin{cases} v \equiv_D (v_1,\ldots,v_m),\ v_j \equiv_D [\alpha_j, w_j],\ j = 1,\ldots,m; \\ \alpha = \ (\alpha_1,\ldots,\alpha_m) \in \mathscr{A}_m; \\ w = \ (w_1,\ldots,w_m) \textit{ is a point of euclidean } m\textit{-space } \mathscr{E}^m. \end{cases}$$

(In particular: v may be thought of as a specification of horizontal extents, α_j, and slopes, w_j, of sides of the gcm.)

For $m=1,\ldots,n$, and $\alpha\in\mathscr{A}_m$, the $(m+1)$-tuple $a=a(\alpha)=(a_0,a_1,\ldots,a_m)$ is defined by

$$(1.5)\qquad \begin{cases} a_0 = 0; \\ a_j = a_j(\alpha) \equiv_D \sum_{\nu=1}^j \alpha_\nu, \quad j = 1,\ldots,m\,. \end{cases}$$

(In particular: The a_j may be thought of as abscissas of endpoints of the sides of the gcm.)

For $\alpha\in\mathscr{A}_m$ and $x\in\mathscr{E}^n$, the m-tuple $u=u(\alpha,x)$ is defined by (cf. Definition 1.1)

$$(1.6)\qquad \begin{cases} u = u(\alpha,x) \equiv_D (u_1,\ldots,u_m)\,, \\ u_j = u_j(\alpha,x) \equiv_D E_{\alpha_j}(x_{a_{j-1}+1},\ldots,x_{a_j}), \quad j = 1,\ldots,m\,. \end{cases}$$

(In particular: if E is a simple arithmetic average, then n_j represents the slope of a segment of the graph of $\{(i,s_i)\}_{i=0}^n$ between points with abscissa a_{j-1} and a_j.)

For $\alpha \in \mathscr{A}_m$ *and* $x \in \mathscr{E}^n$, *the m-tuple* $y = y(\alpha, x)$ *is defined by*

$$(1.7)\qquad \begin{cases} y = y(\alpha,x) \equiv_D v(\alpha,u) = (y_1,\ldots,y_m)\,, \\ y_j = y_j(\alpha,x) \equiv_D (\alpha_j,u_j), \quad j = 1,\ldots,m\,. \end{cases}$$

It follows from the definition that $y \in \mathscr{V}_m$. (In particular: if E is an arithmetic average, again y is a specification of horizontal extents and slopes of segments of the gcm.)

DEFINITION 1.3. *Let* $\mathscr{E}_*$ *be a subset of* $\mathscr{E}^n$, *invariant under permutation operators, such that at each point* $x = (x_1, \ldots, x_n)$ *of* $\mathscr{E}_*$, *for* $1 \leqq r < t \leqq n$,

$$(1.8)\qquad \min[E_r(x_1,\ldots,x_r),\ E_{t-r}(x_{r+1},\ldots,x_t)] \leqq E_t(x_1,\ldots,x_t) \leqq \max[E_r(x_1,\ldots,x_r),\ E_{t-r}(x_{r+1},\ldots,x_t)]\,,$$

and

$$(1.9)\qquad E_r(x_1,\ldots,x_r) \neq E_t(x_1,\ldots,x_t)\,.$$

The writer's attention was called to property (1.8) by Mr. H. Brøns of the Statistics Institute, University of Copenhagen, in conversations in which Mr. Brøns suggested that such functions might be of rather wide statistical interest.

It is property (1.8) of the functions E_r on $\mathscr{E}_*$ which is critical in the present investigations. It is noted that both

$$E_r(x_1,\ldots,x_r) \equiv \sum_{i=1}^{r} x_i/r \qquad \text{and} \qquad E_r(x_1,\ldots,x_r) \equiv x_r$$

satisfy (1.8). Property (1.9) may most appropriately be considered a property of the set $\mathscr{E}_*$ for given functions E_r satisfying (1.8) on that set.

DEFINITION 1.4. *The sets* A_m, $A(\alpha)$, $B(\alpha)$, $D(\alpha)$, D^k, D_m *are defined by the equations* (1.10–14) *below.*

For $m \leqq n$,

$$(1.10)\qquad A_m \equiv_D \{w \in \mathscr{E}^m\colon\ w_1 < w_2 < \ldots < w_m\}\,.$$

(Note that the slopes of the sides of the gcm of a set $\{(i,s_i)\}_{i=0}^n$ of points in the cartesian plane increase to the right.)

For $\alpha \in \mathscr{A}_m$,

$$(1.11)\qquad \begin{cases} A(\alpha) \equiv_D \{x \in \mathscr{E}_*\colon\ u(\alpha,x) \in A_m\} = \{x \in \mathscr{E}_*\colon\ u_1 < \ldots < u_m\}\,, \\ u_j = u_j(\alpha,x), \quad j = 1,2,\ldots,m; \quad \textit{if } m = 1,\ A(\alpha) \equiv_D \mathscr{E}_*\,. \end{cases}$$

$$
(1.12)\quad \begin{cases} B_j(\alpha) \equiv_D \{x \in \mathscr{E}_*:\ u_j < \min\limits_{1 \leq r < \alpha_j} E_r(x_{a_{j-1}+1}, \ldots, x_{a_{j-1}+r})\}\,, \\ \qquad j = 1, \ldots, m;\ a_j = a_j(\alpha),\ u_j = u_j(\alpha, x)\,, \\ B(\alpha) \equiv_D \bigcap\limits_{j=1}^{m} B_j(\alpha)\,. \end{cases}
$$

$$
(1.13)\qquad D(\alpha) \equiv_D A(\alpha) B(\alpha)\,.
$$

For $k \in \mathscr{K}_m$,

$$
(1.14)\qquad D^k \equiv_D \bigcup_{\alpha \in \mathscr{A}^k} D(\alpha), \qquad D_m \equiv_D \bigcup_{k \in \mathscr{K}_m} D^k = \bigcup_{\alpha \in \mathscr{A}_m} D(\alpha)\,.
$$

We pause to interpret some of the variables and sets defined above in each of two cases: Case 1, that of Andersen's theorem [2], and Case 2, that of randomly chosen permutations of $1, \ldots, n$.

Definition 1.5. *Case* 1 *is specified by setting*

$$
E_r(t_1, \ldots, t_r) \equiv_D \sum_{i=1}^{r} t_i / r\,,
$$

and $\mathscr{E}_*$ *equal to the set of all points* x *of* $\mathscr{E}^n$ *such that the arithmetic averages of distinct sets of coordinates of* x *are unequal.*

Set $s_0 = 0$, $s_r = \sum_{\nu=1}^{r} x_\nu$, $r = 1, \ldots, n$, and plot the points $\{(r, s_r)\}_{r=0}^{n}$ in the cartesian plane. Then x_i is the slope of the line segment joining $(i-1, s_{i-1})$ to (i, s_i); u_j is the slope of the segment joining $(a_{j-1}, s_{a_{j-1}})$ to (a_j, s_{a_j}). For $x \in \mathscr{E}_*$, the segments joining two distinct pairs of points will have different slopes. The set $B(\alpha)$ is the set of all $x \in \mathscr{E}_*$ such that for $j = 1, \ldots, m$ the segment of slope u_j lies below all intermediate points. $A(\alpha)$ is the set of those x in $\mathscr{E}_*$ for which the slopes u_j strictly increase with j. The set $D(\alpha)$ consist of all points in $\mathscr{E}_*$ such that the segments of slopes $u_1, \ldots, u_m$ form the greatest convex minorant (gcm) of $\{(r, s_r)\}_{r=0}^{n}$. Therefore, $D(\alpha)$ may be interpreted as the set of points x in $\mathscr{E}_*$ such that the gcm has precisely m sides, the first of horizontal extent α_1, the second of horizontal extent α_2, etc. Correspondingly, D^k may be interpreted as the set of points x in $\mathscr{E}_*$ such that the gcm has k_1 sides of horizontal extent 1, k_2 of horizontal extent 2, etc. In particular, D_m is the set of points x in $\mathscr{E}_*$ for which the gcm has m sides.

Definition 1.6. *Case* 2 *is specified by setting*

$$
E_r(t_1, \ldots, t_r) \equiv_D t_r\,,
$$

and taking for $\mathscr{E}_*$ *the set of permutations of* $1, 2, \ldots, n$; *more precisely,* $\mathscr{E}_*$ *is the set of points* $x \in \mathscr{E}^n$ *whose coordinates are distinct positive integers not greater than* n.

Here $B_j(\alpha)$ is the set of points $x \in \mathscr{E}_*$ such that x_{a_j} is the smallest of the integers $x_{a_{j-1}+1}, x_{a_{j-1}+2}, \ldots, x_{a_j}$; and $A(\alpha)$ is the set of $x \in \mathscr{E}_*$ such that $x_{a_1} < x_{a_2} < \ldots < x_{a_n}$. Thus if, for $x \in \mathscr{E}_*$, one picks a_1 as the index of the smallest coordinate (this coordinate is 1), a_2 as the index of the smallest of those with greater index, etc., and then sets $\alpha_j = a_j - a_{j-1}$, $j = 1, 2, \ldots, m$, one has $x \in D(\alpha)$.

For each $m \leqq n$ we have to deal below with permutations of $1, 2, \ldots, m$; for each such permutation, and for fixed $\alpha \in \mathscr{A}_m$, we consider a certain permutation of the integers $1, 2, \ldots, n$.

Definitions 1.7. *For $m = 1, \ldots, n$, we denote by Π_m the set of all permutations*

$$\pi\colon\ (1, \ldots, m) \to (i_1, \ldots, i_m)\,.$$

We denote also by π the permutation operator carrying an ordered m-tuple

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

and by Π_m the class of such permutation operators. For $\pi \in \Pi_m$, $\alpha \in \mathscr{A}_m$,

$$\begin{aligned} A(\alpha,\pi) &\equiv_D \{x \in \mathscr{E}_*\colon\ \pi u(\alpha,x) \in A_m\} \\ &= \{x \in \mathscr{E}_*\colon\ u_{i_1} < u_{i_2} < \ldots < u_{i_m}\}\,. \end{aligned} \tag{1.15}$$

Remark 1.1. If $\alpha \in \mathscr{A}^k$, then as π ranges over Π_m, $\pi\alpha$ ranges over $\mathscr{A}^k$, each element of $\mathscr{A}^k$ appearing exactly $\prod_{i=1}^n k_i!$ times.

For greater clarity, we illustrate the notation. Suppose that $n = 7$, $m = 4$, $k_1 = 2$, $k_2 = 1$, $k_3 = 1$, $\alpha = (2, 1, 3, 1)$, and that π carries $(1, 2, 3, 4)$ into $(2, 4, 1, 3)$. Then

$$\begin{aligned} \pi\alpha &= (\alpha_2, \alpha_4, \alpha_1, \alpha_3) = (1, 1, 2, 3); \\ v(\alpha, w) &= \big([2, w_1], [1, w_2], [3, w_2], [1, w_4]\big); \\ (a_0, a_1, a_2, a_3, a_4) &= (0, 2, 3, 6, 7); \\ u = (u_1, u_2, u_3, u_4) &= \big(E_2(x_1, x_2), E_1(x_3), E_3(x_4, x_5, x_6), E_1(x_7)\big)\,; \\ \pi u &= \big(E_1(x_3), E_1(x_7), E_2(x_1, x_2), E_3(x_4, x_5, x_6)\big)\,. \end{aligned}$$

The permutation operator $p = p(\alpha, \pi)$ defined in Definition 1.8 below carries

$$x = (x_1, x_2, \ldots, x_7) \quad \text{into} \quad px = (x_3, x_7, x_1, x_2, x_4, x_5, x_6)\,.$$

Since $\pi\alpha = (1, 1, 2, 3)$ we have

$$u(\pi\alpha, px) = [E_1(x_3), E_1(x_7), E_2(x_1, x_2), E_3(x_4, x_5, x_6)] = \pi u(\alpha, x)$$

verifying (1.16) below. Also

$$B(\alpha) = \{x \in \mathscr{E}_*:\ E_2(x_1,x_2) < E_1(x_1),\ E_3(x_4,x_5,x_6) < \min[E_1(x_4),\ E_2(x_4,x_5)]\}\ ,$$

$$B(\pi\alpha) = \{x \in \mathscr{E}_*:\ E_2(x_3,x_4) < E_1(x_3),\ E_3(x_5,x_6,x_7) < \min[E_1(x_5),\ E_2(x_5,x_6)]\}\ .$$

Thus

$$x \in B(\alpha) \Leftrightarrow px \in B(\pi\alpha);$$

that is, $B(\pi\alpha) = pB(\alpha)$, verifying (1.17) below.

Definition 1.8. *For $m \leqq n$, $\alpha \in \mathscr{A}_m$, $\pi \in \Pi_m$, let $u = u(\alpha,x)$ and think of the coordinates of x appearing in the definition of $u_j(\alpha,x)$ as being written in the order of increasing index j, $j = 1,2,\ldots,m$. Let $j_1, j_2, \ldots, j_n$ be the indices of the coordinates of x in the order in which they appear when the u_j are rearranged to form $\pi u = (u_{i_1}, u_{i_2}, \ldots, u_{i_m})$ without rearranging the coordinates of x within u_{i_j}, $(j = 1,2,\ldots,m)$. Let $p = p(\alpha,\pi)$ carry $x = (x_1,\ldots,x_n)$ into $px = (x_{j_1}, x_{j_2}, \ldots, x_{j_n})$.*

Lemma 1.1. *For $m \leqq n$, $\alpha \in \mathscr{A}_m$, $\pi \in \Pi_m$, $p = p(\alpha,\pi)$, we have*

$$\pi u(\alpha,x) = u(\pi\alpha,px) \tag{1.16}$$

and

$$pB(\alpha) = B(\pi\alpha)\ . \tag{1.17}$$

This lemma is immediate from the definitions of the symbols involved. Continuing the illustration of the notation, we have

$$\begin{aligned} y = y(\alpha,x) &= v(\alpha,u) \\ &= \big([2,\ E_2(x_1,x_2)],\ [1,\ E_1(x_3)],\ [3,\ E_3(x_4,x_5,x_6)],\ [1,\ E_1(x_7)]\big)\ . \end{aligned}$$

Also

$$\begin{aligned} y(\pi\alpha,px) &= v\big(\pi\alpha, u(\pi\alpha,px)\big) \\ &= \big([1,\ E_1(x_3)],\ [1,\ E_1(x_7)],\ [2,\ E_2(x_1,x_2)],\ [3,\ E_3(x_4,x_5,x_6)]\big) \\ &= \pi y(\alpha,x)\ , \end{aligned}$$

verifying (1.18) below.

Corollary 1.1.

$$\pi y(\alpha,x) = y(\pi\alpha,px)\ . \tag{1.18}$$

Proof:

$$\begin{aligned} \pi y(\alpha,x) &= \pi v[\alpha,u(\alpha,x)] = v[\pi\alpha,\pi u(\alpha,x)] \\ &= v[\pi\alpha,u(\pi\alpha,px)] = y(\pi\alpha,px)\ . \end{aligned}$$

Lemma 1.2. *If $m \leqq n$, $\alpha \in \mathscr{A}_m$, the sets $A(\alpha,\pi)$, $\pi \in \Pi_m$, are disjoint and their union is $\mathscr{E}_*$. If $p = p(\alpha,\pi)$ for $\pi \in \Pi_m$, then*

$$pA(\alpha,\pi) = A(\pi\alpha)\ . \tag{1.19}$$

PROOF. The first statement is immediate from Definition 1.3 and (1.15). Further

$$A(\alpha,\pi) = \{x\in\mathscr{E}_*:\ \pi u(\alpha,x)\in A\} = \{x\in\mathscr{E}_*:\ u(\pi\alpha,px)\in A\}\,.$$

That is, $x\in A(\alpha,\pi)$ if and only if $px\in A(\pi\alpha)$; this is (1.19).

Motivated by the applications in Sections 2 and 3, we introduce some definitions and a property of classes $\{H(\alpha)\}$ of subsets of $\mathscr{E}_*$, one set corresponding to each $\alpha\in\mathscr{A}$.

DEFINITION 1.9. *The class* $\{H(\alpha)\}$, $\alpha\in\mathscr{A}$, *of subsets of* $\mathscr{E}_*$ *will be called* π*-invariant if* $\alpha\in\mathscr{A}_m$, $\pi\in\Pi_m$, $p=p(\alpha,\pi)$ (cf. (1.3), Def. 1.7–8) *imply*

$$pH(\alpha) = H(\pi\alpha)\,. \tag{1.20}$$

It is part of the conclusion of Lemma 1.1 that the sets $\{B(\alpha)\}$ are π-invariant.

DEFINITIONS 1.10. *For* $m\leqq n$, $k\in\mathscr{K}_m$,

$$\begin{cases} H^k \equiv_D \bigcup_{\alpha\in\mathscr{A}^k} H(\alpha)D(\alpha)\,, \\ H_m \equiv_D \bigcup_{k\in\mathscr{K}_m} H^k, \qquad H\equiv_D \bigcup_{m=1}^{n} H_m\,. \end{cases} \tag{1.21}$$

DEFINITION 1.11. *The random variables* $X_1,\ldots,X_n$ *are symmetrically dependent* (E. Sparre Andersen) *if for every Borel set* $J\subset\mathscr{E}^n$ *and every permutation operator* p *we have*

$$\Pr\{X\in J\} = \Pr\{pX\in J\}; \tag{1.22}$$

or equivalently, if for every event $J\subset\mathscr{E}^n$ *and every permutation operator* p,

$$P(J) = P(pJ)\,. \tag{1.23}$$

In general, we shall use the same symbol for a subset of $\mathscr{E}^n$ and for the event which occurs if X falls in that subset.

LEMMA 1.3. *If* $\{H(\alpha)\}$, $\alpha\in\mathscr{A}$, *is* π*-invariant, if* $m\leqq n$, $k\in\mathscr{K}_m$, $\alpha\in\mathscr{A}^k$, *and if the random variables* $X_1,\ldots,X_n$ *are symmetrically dependent, then*

$$P[H(\alpha)] = \left(\prod_{i=1}^{n} k_i!\right)\sum_{\beta\in\mathscr{A}^k} P[A(\beta)H(\beta)]\,. \tag{1.24}$$

PROOF. $\{A(\alpha,\pi)\}$, $\pi\in\Pi_m$, is a partition of $\mathscr{E}_*$ into disjoint sets. Hence

$$P[H(\alpha)] = \sum_{\pi\in\Pi_m} P[A(\alpha,\pi)H(\alpha)]\,.$$

Since the random variables $X_1,\ldots,X_n$ are symmetrically dependent, we have

$$P[A(\alpha,\pi)H(\alpha)] = P[pA(\alpha,\pi) \cap pH(\alpha)] = P[A(\pi\alpha)H(\pi\alpha)]$$

by (1.19) and (1.20). The conclusion now follows from Remark 1.1.

LEMMA 1.4. *For distinct* $\alpha \in A$, *the sets* $D(\alpha)$ *are disjoint, and* $\bigcup_{\alpha\in\mathscr{A}} D(\alpha) = \mathscr{E}_*$.

PROOF. Given $x\in\mathscr{E}_*$, choose $\alpha_1 = a_1$ so that

$$u_1 = E_{\alpha_1}(x_1,\ldots,x_{\alpha_1}) = \min_{1\le r\le n} E_r(x_1,\ldots,x_r)\,.$$

Then choose α_2 so that

$$u_2 = E_{\alpha_2}(x_{a_1+1},\ldots,x_{a_2}) = \min_{1\leqq r\leqq n-a_1} E_r(x_{a_1+1},\ldots,x_{a_1+r}),\ \text{etc.}$$

For $\alpha=(\alpha_1,\alpha_2,\ldots,\alpha_m)$ thus determined we clearly have $x\in B(\alpha)$ (cf. (1.12)). Further,

$$u_j < E_{\alpha_j+\alpha_{j+1}}(x_{a_{j-1}+1},\ldots,x_{a_{j+1}})\,,$$

hence from (1.8) it follows that $u_j < u_{j+1}$, $j=1,2,\ldots,m-1$. Thus also $x\in A(\alpha)$ (cf. (1.11)), hence $x\in D(\alpha)$. Further, if $x\in D(\beta)$ for $\beta\in\mathscr{A}$, one verifies that

$$E_{\beta_1}(x_1,\ldots,x_{b_1}) = \min_{1\leqq r\leqq n} E_r(x_1,\ldots,x_r),\ \text{etc.},$$

so that $\beta=\alpha$.

DEFINITIONS 1.12. *By Lemma* 1.4, *to each* $x\in\mathscr{E}_*$ *corresponds a unique* $m=m(x)$ *and an* $\alpha=\alpha(x)=[\alpha_1(x),\ldots,\alpha_m(x)]\in\mathscr{A}_m$ *such that* $x\in D[\alpha(x)]$. *Combining this fact with* (1.5), *we obtain the functions* $a_j[\alpha(x)]$, $j=1,2,\ldots,$ $m(x)$, *and, in conjunction with* (1.6), *the function*

$$u(x) \equiv_D u[\alpha(x),x]\,. \tag{1.25}$$

We also define

$$M = m(X)\,. \tag{1.26}$$

REMARK 1.2. $\{x\in\mathscr{E}_* : m(x)=m\} = D_m$ (cf. 1.14).

In Case 1, $m=m(x)$ is the number of sides in the gcm of $\{(r,s_r)\}_{r=0}^{n}$; $\alpha_1(x)$ is the horizontal extent of the first side, etc. The random variable M is the function of X obtained on replacing the point x by the random variable X.

In order to interpret M in Case 2 as the number of cycles in a randomly chosen permutation of the integers $1,2,\ldots,n$, we first introduce the representation $x=(x_1,x_2,\ldots,x_n)$ of a permutation $(1,2,\ldots,n)\to(j_1,j_2,\ldots,j_n)$.

The first coordinate x_1 is the integer j_1 which replaces 1; x_2 is the integer which replaces x_1, etc., this phase of the definition terminating when a coordinate equal to 1 is reached. The following coordinate of x is the integer replacing the smallest integer which doesn't already appear among the coordinates, etc. It will be recognized that this representation is a familiar device for exhibiting the cycles of the permutation. For example, if the permutation carries $(1,2,3,4,5,6,7)$ into $(5,1,4,7,6,2,3)$, then $1 \to 5$, $5 \to 6$, $6 \to 2$, $2 \to 1$; $3 \to 4$, $4 \to 7$, $7 \to 3$; hence the permutation has 2 cycles, and we have $x_1 = 5$, $x_2 = 6$, $x_3 = 2$, $x_4 = 1$; $x_5 = 4$, $x_6 = 7$, $x_7 = 3$. Here $m(x) = 2$, $\alpha_1(x) = 4$, $\alpha_2(x) = 3$.

Since to each permutation corresponds one and only one such representation in cycles, it is clear that each point in $\mathscr{E}_*$ occurs with probability $1/n!$ when the permutation is chosen at random. It follows that the random variables $X_1, \ldots, X_n$ are symmetrically dependent (Definition 1.11).

THEOREM 1.1. *If* $\{H(\alpha)\}$, $\alpha \in \mathscr{A}$, *is π-invariant (Definition* 1.9), *if* $m \leqq n$, $k \in \mathscr{K}_m$, $\alpha \in \mathscr{A}^k$, *and if the random variables* $X_1, \ldots, X_n$ *are symmetrically dependent, then*

$$P(H^k) = P[H(\alpha)B(\alpha)] \Big/ \prod_{i=1}^{n} k_i! \, . \tag{1.27}$$

PROOF. It suffices to replace $H(\alpha)$ in Lemma 1.3 by $H(\alpha)B(\alpha)$ and to observe that since $D(\beta) = A(\beta)B(\beta)$ one has (cf. (1.21))

$$P(H^k) = \sum_{\beta \in \mathscr{A}^k} P[A(\beta)B(\beta)H(\beta)] \, .$$

It is shown below that

$$P[H(\alpha)B(\alpha)] = P[H(\alpha)] \prod_{i=1}^{n} (1/i)^{k_i} \, .$$

For Case 1 and for $H(\alpha) = \mathscr{E}_*$ this result is essentially a consequence of a theorem of Andersen [1]. This latter theorem Spitzer [12] subsequently obtained as an immediate application of a combinatorial theorem on the averages $(\sum_{i=1}^{r} x_i)/r$. This theorem is generalized here so as to apply to the functions E_r.

Let t be a positive integer not greater than n. Then $\mathscr{E}_*{}^t$ will denote the projection of $\mathscr{E}_*$ on $\mathscr{E}^t$:

$$\mathscr{E}_*{}^t \equiv_D \{w = (w_1, \ldots, w_t) \in \mathscr{E}^t \colon \ \exists\, x_{t+1}, x_{t+2}, \ldots, x_n \text{ such that } (w_1, \ldots, w_t, x_{t+1}, \ldots, x_n) \in \mathscr{E}_*\} \, .$$

Let C denote the cyclic shift transformation in $\mathscr{E}^t$:

$$w = (w_1, w_2, \ldots, w_t) \to Cw = (w_2, w_3, \ldots, w_t, w_1) \, .$$

THEOREM 1.2. *For fixed t, $1 \leqq t \leqq n$, for fixed r, $0 \leqq r \leqq t-1$, and $w \in \mathscr{E}_*^t$, there is exactly one integer $\nu = \nu(r)$, $0 \leqq \nu \leqq t-1$, such that exactly r of the inequalities*

(1.28) $E_t(C^\nu w) < E_1(C^\nu w),\ E_t(C^\nu w) < E_2(C^\nu w), \ldots, E_t(C^\nu w) < E_{t-1}(C^\nu w)$

hold.

The method of proof is suggested by the graphical interpretation given by Spitzer for Case 1. In Case 1, define

$$x_{n+i} \equiv_D x_i, \quad i = 1,2,\ldots,n, \qquad s_j \equiv_D \sum_{i=1}^{j} x_i/j, \quad j = 1,2,\ldots,2n\,,$$

and plot the points (j,s_j), $j=0,1,\ldots,2n$. There will be exactly one integer ν such that exactly r points lie above the segment joining the points (ν,s_ν) and $(\nu+n,s_{\nu+n})$. For this ν, exactly r of the inequalities specified in the conclusion of the theorem obtain.

PROOF OF THEOREM 1.2. It is useful to reinterpret (1.8) and (1.9) in terms of the present notation. For $1 \leqq t \leqq n$, $1 \leqq j \leqq t$,

$$w = (w_1,w_2,\ldots,w_t), \quad E_j(w) \equiv_D E_j(w_1,w_2,\ldots,w_j)\,.$$

Then (1.8) may be rewritten:

$$\text{(1.8}'\text{)} \qquad \min[E_j(w),E_{t-j}(C^j w)] \leqq E_t(w) \leqq \max[E_j(w),E_{t-j}(C^j w)];$$

and (1.9) becomes

$$\text{(1.9}'\text{)} \qquad E_j(w) \neq E_t(w)\,.$$

The proof of Theorem 1.2 depends on showing first:

1° *If an inequality $E_t(C^\nu w) < E_j(C^\nu w)$ obtains, then fewer of the inequalities* (1.28) *are satisfied if ν is replaced by $\nu+j$* (mod t) *than are satisfied for ν.* Thus in this case it is not so that the same number of inequalities (1.28) are satisfied for ν as for $\nu+j$. Suppose 1° established, and suppose further that $E_j(C^\nu w) < E_t(C^\nu w)$. From (1.8′),

$$E_t(C^\nu w) \leqq E_{t-j}(C^{\nu+j} w)\,.$$

If $E_{t-j}(C^{\nu+j}w) < E_t(C^{\nu+j}w)$, then by (1.8′) again,

$$E_t(C^{\nu+j} w) \leqq E_j(C^{\nu+t} w) = E_j(C^\nu w)\,.$$

Then

$$E_{t-j}(C^{\nu+j} w) \leqq E_j(C^\nu w) < E_t(C^\nu w)\,,$$

a contradiction. Thus

$$E_j(C^\nu w) < E_t(C^\nu w) \quad \text{implies} \quad E_t(C^{\nu+j} w) < E_{t-j}(C^{\nu+j} w)\,.$$

It then follows from 1° that fewer of the inequalities (1.28) are satisfied for $\nu+j+t-j\equiv\nu \pmod{t}$ than for $\nu+j$. Thus again it is not so that the same number of inequalities (1.28) are satisfied for ν as for $\nu+j$. Thus for distinct ν there are distinct numbers r of the inequalities (1.28) satisfied, and the conclusion of Theorem 1.2 follows.

It remains to establish 1°. It suffices to establish it when $\nu=0$, as 1° may then be obtained by replacing w by $C^{\nu}w$. We show that if

$$E_t(w) < E_j(w) \tag{1.29}$$

and if

$$E_t(C^j w) < E_k(C^j w)\,, \tag{1.30}$$

then

$$E_t(w) < E_h(w), \qquad 1 \leqq h < t, \quad h \equiv j+k \pmod{t}\,. \tag{1.31}$$

Thus if (1.29) obtains, corresponding to each of the inequalities (1.28) satisfied for $\nu=j$ there is one other than (1.29) satisfied for $\nu=0$; this establishes 1°. Suppose (1.31) does not follow from (1.29) and (1.30), and that instead

$$E_h(w) < E_t(w)\,. \tag{1.32}$$

From (1.30) and (1.8′), it follows that

$$E_{t-k}(C^{j+k}w) \leqq E_t(C^j w) < E_k(C^j w)\,. \tag{1.33}$$

Case (i). Suppose now that $j+k<t$. From (1.29) and (1.32), $E_{j+k}(w)<E_j(w)$. From (1.8′),

$$E_k(C^j w) \leqq E_{j+k}(w) < E_j(w)\,. \tag{1.34}$$

From (1.33) it now follows that

$$E_{t-k}(C^{j+k}w) < E_j(w) = E_j(C^t w) \tag{1.35}$$

and from (1.8′) that

$$E_{t-k-j}(C^{j+k}w) < E_{t-k}(C^{j+k}w) \leqq E_j(C^t w)\,. \tag{1.36}$$

From (1.36), (1.33), (1.34) and (1.32) we have $E_{t-k-j}(C^{j+k}w)<E_t(w)$. Application of (1.8′) yields

$$E_{t-k-j}(C^{j+k}w) < E_t(w) < E_{j+k}(w)\,,$$

contradicting (1.32).

Case (ii). Suppose, finally, that $j+k>t$ (note that (1.32) forbids $j+k=t$). The inequalities (1.33) may then be written

$$E_{t-k}(C^h w) \leqq E_t(C^j w) < E_k(C^j w)\,. \tag{1.33′}$$

Also (1.8′) yields

$$(1.37)\qquad E_h(w) = E_h(C^t w) \leqq E_k(C^j w) < E_{t-j}(C^j w) .$$

Now

$$E_t(C^j w) < E_{t-j}(C^j w)$$

together with (1.8′) yields

$$(1.38)\qquad E_j(w) = E_j(C^t w) \leqq E_t(C^j w) < E_{t-j}(C^j w) .$$

From (1.29) and (1.38) we have $E_t(w) < E_j(w) < E_{t-j}(C^j w)$, contradicting (1.8′). This completes the proof of Theorem 1.2.

Andersen's theorem [1] is now seen still to hold when the arithmetic averages are replaced by the present functions E_r:

THEOREM 1.3. *Let C denote the cyclic shift transformation in $\mathscr{E}^n$. Let L denote an event (subset of $\mathscr{E}^n$) invariant under C. Let $N = N(x)$ denote the number of indices i $(i = 1, 2, \ldots, n-1)$ such that $E_n(x) < E_i(x)$. If $P(\mathscr{E}_*) = 1$, then*

$$\Pr\{N(X) = r \mid L\} = 1/n, \qquad r = 0, 1, \ldots, n-1 .$$

PROOF. Theorem 1.3 is immediate from Theorem 1.2, which implies that the sets

$$\{x\colon N(C^\nu x) = r\} \cap L \cap \mathscr{E}_*, \qquad \nu = 0, 1, \ldots, n-1$$

are disjoint sets which have the same probability and whose union is $L \cap \mathscr{E}_*$.

DEFINITIONS 1.13. *Let $n_1, n_2, \ldots, n_m$ be positive integers with $n = \sum_{j=1}^m n_j$. For $j = 1, 2, \ldots, m$, let $x^j = (x_{j,1}, x_{j,2}, \ldots, x_{j,nj})$ denote the generic point of $\mathscr{E}^{n_j}$, and let $x = (x^1, \ldots, x^m)$ denote the generic point of the cartesian product $\mathscr{E}^n = X_{j=1}^m \mathscr{E}^{n_j}$. For each j, define the cyclic shift transformation*

$$(1.39)\qquad C_j\colon\ x^j = (x_{j,1}, \ldots, x_{j,n_j}) \to C_j x^j = (x_{j,2}, \ldots, x_{j,n_j}, x_{j,1});$$

and define

$$(1.40)\qquad C_j x = C_j(x^1, \ldots, x^j, \ldots, x^m) \equiv_D (x^1, \ldots, C_j x^j, \ldots, x^m) .$$

Let X_{ji} be a random variable, $i = 1, 2, \ldots, n_j$, $j = 1, 2, \ldots, m$. Let L be a Borel subset of $\mathscr{E}^n$ invariant under each C_j, $j = 1, 2, \ldots, m$. Then the event

$$\{(X_{1,1}, \ldots, X_{1,n_1}; \ldots; X_{m,1}, \ldots, X_{m,n_m}) \in L\}$$

(denoted also by L) will be termed *cyclically symmetric* in $X_{j,1}, \ldots, X_{j,n_j}$ for $j = 1, 2, \ldots, m$.

DEFINITIONS 1.14.

$$E_r^j(x) \equiv_D E_r(x^j) = E_r(x_{j,1},\ldots,x_{j,r}), \qquad r = 1,2,\ldots,n_j,\ j = 1,2,\ldots,m; \tag{1.41}$$

$$F_j \equiv_D \{x \in \mathscr{E}_*\colon\ E_{n_j}^j(x) < E_r^j(x) \text{ for } r = 1,2,\ldots,n_j-1\}, \qquad j = 1,2,\ldots,m. \tag{1.42}$$

THEOREM 1.4. *Let the random variables X_{ji}, $i=1,2,\ldots,n_j$, $j=1,2,\ldots,m$, be symmetrically dependent, and let L be an event which in cyclically symmetric in $X_{j,1},\ldots,X_{j,m_j}$ for $j=1,2,\ldots,m$. Then*

$$P\left(L \bigcap_{j=1}^{m} F_j\right) = P(\mathscr{E}_* L) \Big/ \prod_{j=1}^{m} n_j. \tag{1.43}$$

Note that if the random variables are independent as well as identically distributed then Theorem 1.4 follows from the special case $m=1$.

PROOF. By Theorem 1.2, to each $x \in \mathscr{E}_*$ corresponds exactly one m-tuple $(\nu_1,\nu_2,\ldots,\nu_m)$, $0 \leqq \nu_j \leqq n_j-1$, $j=1,2,\ldots,m$, such that $C_j^{\nu_j}x \in F_j$, $j=1,2,\ldots,m$. Thus for distinct m-tuples $(\mu_1,\mu_2,\ldots,\mu_m)$, $0 \leqq \mu_j \leqq n_j-1$, $j=1,2,\ldots,m$, the sets $\bigcap_{j=1}^m C_j^{\mu_j}F_j$ are disjoint, and their union is $\mathscr{E}_*$. Hence

$$L\mathscr{E}_* = \bigcup \bigcap_{j=1}^{m} [C_j^{\mu_j}F_j]L,$$

the union being carried out over all distinct m-tuples $(\mu_1,\mu_2,\ldots,\mu_m)$, $0 \leqq \mu_j \leqq n_j-1$, $j=1,2,\ldots,m$. Since L is cyclically symmetric in $X_{j,1},\ldots,X_{j,n_j}$ for $j=1,2,\ldots,m$, we have also

$$L\mathscr{E}_* = \bigcup \bigcap_{j=1}^{m} C_j^{\mu_j}[F_jL].$$

Since the random variables X_{ji} are symmetrically dependent, each term in the union has the same probability, so that

$$P(L\mathscr{E}_*) = \left[P\left(\bigcap_{j=1}^{m} F_jL\right)\right] \prod_{j=1}^{m} n_j,$$

which is (1.43).

COROLLARY 1.2. *Let $m \leqq n$, $k \in \mathscr{K}_m$, $\alpha \in \mathscr{A}^k$. If $X_1,\ldots,X_n$ are symmetrically dependent, if $H(\alpha)$ is cyclically symmetric in each of*

$$X_1,X_2,\ldots,X_{a_1};\quad X_{a_1+1},\ldots,X_{a_2};\quad \ldots;\quad X_{a_{m-1}+1},\ldots,X_{a_m},$$

and if $P(\mathscr{E}_*)=1$, then

$$P[H(\alpha)B(\alpha)] = P[H(\alpha)] \prod_{i=1}^{n} (1/i)^{k_i}. \tag{1.44}$$

PROOF. In Theorem 1.4 set

$$n_1 = \alpha_1,\ n_2 = \alpha_2, \ldots,\ n_m = \alpha_m,$$
$$X_{j,i} = X_{a_{j-1}+i}, \qquad i = 1,2,\ldots,\alpha_j; \quad j = 1,2,\ldots,m,$$

and $L = H(\alpha)$. We have $B(\alpha) = \bigcap_{j=1}^{m} F_j$, and find that

$$P[H(\alpha)B(\alpha)] = P[H(\alpha)]\Big/\prod_{j=1}^{m}\alpha_j = P[H(\alpha)]\prod_{i=1}^{n}(1/i)^{k_i}.$$

COROLLARY 1.3. *Under the hypotheses of Corollary* 1.2 (*omitting that on* $H(\alpha)$),

$$P[B(\alpha)] = \prod_{i=1}^{n}(1/i)^{k_i}.$$

PROOF. It suffices to set $H(\alpha) = \mathscr{E}_*$.

THEOREM 1.5. *If the random variables* $X_1,\ldots,X_n$ *are symmetrically dependent, if* $P(\mathscr{E}_*) = 1$, *if* $\{H(\alpha)\}$, $\alpha \in \mathscr{A}$, *is* π*-invariant* (Def. 1.9), *if* $m \leqq n$, $k \in \mathscr{K}_m$, $\alpha \in \mathscr{A}^k$, *and if* $H(\alpha)$ *is cyclically symmetric in each of*

$$X_1,\ldots,X_{a_1};\quad X_{a_1+1},\ldots,X_{a_2};\quad \ldots;\quad X_{a_{m-1}+1},\ldots,X_{a_m},$$

then

$$P(H^k) = P[H(\alpha)]\prod_{i=1}^{n} 1/(k_i!\, i^{k_i}). \tag{1.45}$$

This, the "principal theorem" referred to in the section heading, is now immediate from Theorem 1.1 and Corollary 1.2.

THEOREM 1.6. *If the random variables* $X_1,\ldots,X_n$ *are symmetrically dependent and if* $P(\mathscr{E}_*) = 1$ *then*

$$\Pr\{M = m\} = \sum_{k\in\mathscr{K}_m}\prod_{i=1}^{n} 1/(k_i!\, i^{k_i}), \quad m = 1,2,\ldots,n. \tag{1.46}$$

Thus each such random variable M (cf. (1.26)) has the same distribution as in the special Case 2 (Def. 1.6), where M is the number of cycles in a randomly chosen permutation of the integers $1,2,\ldots,n$. For Case 1, Theorem 1.6 is Andersen's theorem [2]. (Andersen's theorem has been generalized in another way by Miles [11]. Let $W_1, W_2, \ldots, W_n$ be fixed positive numbers. Then the role of the functions E_r is played by weighted averages $\sum_{i=1}^{r} W_{\varrho_i'} X_{\varrho_i} / \sum_{i=1}^{r} W_{\varrho_i'}$, where $\varrho_1', \varrho_2', \ldots, \varrho_n'$ and $\varrho_1, \varrho_2, \ldots, \varrho_n$ are randomly chosen permutations of $1,2,\ldots,n$. This clearly does not fit the pattern of a function E_r fixed for each r.)

PROOF OF THEOREM 1.6. By Remark 1.2,

$$\{M = m\} = D_m = \bigcup_{k\in\mathscr{K}_m} D^k$$

(cf. (1.14)). On setting $H(\alpha)=\mathscr{E}_*$, we have $H^k=D^k$ (cf. (1.21)), and (1.46) now follows from Theorem 1.5.

In the following paragraph we denote by M_n the random variable designated M above, and by M_t the corresponding random variable when n is replaced by a positive integer t less than n. Under the hypotheses of Theorem 1.6 we have

$$(1.47)\quad \Pr\{M_n=m\} = (1/n)\sum_{t=m-1}^{n-1}\Pr\{M_t=m-1\}, \qquad m = 2,3,\ldots,n\,,$$

as can be verified from formula (1.48). This formula has an interesting interpretation. Let A_t denote the event that t of the functions E_j, $j=1,2,\ldots,n-1$, are smaller than E_n. We have

$$\{M_n=m\} = \bigcup_{t=m-1}^{n-1} A_t\cap\{M_t=m\};$$

and since $P(A_t)=1/n$, formula (1.47) may be interpreted as asserting that $\Pr\{M_n=m\}$ may be computed as if $\Pr\{M_n=m \mid A_t\}=\Pr\{M_t=m-1\}$. The author does not know whether or not this last equation is in fact valid.

Probably the formulas

$$(1.48)\quad \Pr\{M_n=m\} = \Pr\{M_{n-1}=m-1\}/n + \Pr\{M_{n-1}=m\}(n-1)/n$$

and

$$(1.49)\quad \Pr\{M_n\geqq m\} = \Pr\{M_{n-1}\geqq m-1\}/n + \Pr\{M_{n-1}\geqq m\}(n-1)/n$$

are more suitable than (1.47) for computation (cf. [1]).

2. Distribution of a function determined by the $\mathfrak{g}$cm.

Theorem 2.1 in this section will be applied in Section 3 to the problem of determining the distribution of the likelihood ratio in a test against trend. In Section 3 our attention is restricted to Case 1. However, while the title of the present section refers to Case 1, Theorem 2.1 is not restricted to that situation; the functions E_r envisaged here may be any, subject to (1.8) and (1.9).

DEFINITIONS 2.1. *Let $f_m(v,x)$ be a (real) function defined for $v\in\mathscr{V}_m$ (cf. (1.4)), $x\in\mathscr{E}^n$ (Def. 1.1). The function $z(x)$, the random variable Z, and the sets G, $G(\alpha)$, G^k, G_m are explained by equation (2.1–5) below.*

First let

$$(2.1)\qquad z = z(x) \equiv_D y[\alpha(x),x]$$

(cf. Definitions 1.2, 1.12), and

$$Z \equiv_D z(X)\,. \tag{2.2}$$

For real q, let

$$G \equiv_D \{x \in \mathscr{E}_*:\ f_{m(x)}[z(x), x] < q\}; \tag{2.3}$$

thus if $P(\mathscr{E}_*)=1$, $P(G)$ is the value at q of the distribution function of the random variable $f_M(Z, X)$. For $m \leqq n$, $\alpha \in \mathscr{A}_m$, let

$$G(\alpha) \equiv_D \{x \in \mathscr{E}_*:\ f_m[y(\alpha, x), x] < q\}\,. \tag{2.4}$$

One may think of $G(\alpha)$ as "the form the event G takes if $D(\alpha)$ occurs" (i.e., in Case 1, if the gcm has m sides, the first of horizontal extent α_1, the second of horizontal extent α_2, etc.). For $m \leqq n$, $k \in \mathscr{K}_m$, let

$$\begin{cases} G^k \equiv_D \bigcup_{\alpha \in \mathscr{A}^k} G(\alpha) D(\alpha)\,, \\ G_m \equiv_D \bigcup_{k \in \mathscr{K}_m} G^k \quad \text{(cf. (1.21))}\,. \end{cases} \tag{2.5}$$

Since $x \in D(\alpha)$ if and only if $\alpha = \alpha(x)$, $m = m(x)$, we have

$$G = \bigcup_{m=1}^{n} G_m\,. \tag{2.6}$$

LEMMA 2.1. *If $f_m(v, x)$ is symmetric in the components of $v \in \mathscr{V}_m$ and in the components of $x \in \mathscr{E}^n$, $m = 1, 2, \ldots, n$, then the family $\{G(\alpha)\}$ defined in (2.4) is π-invariant* (Def. 1.9). *Also, for fixed $\alpha \in \mathscr{A}_m$, $G(\alpha)$ is then cyclically symmetric in*

$$X_1, \ldots, X_{a_1}; \quad X_{a_1+1}, \ldots, X_{a_2}; \quad \ldots; \quad X_{a_{m-1}+1}, \ldots, X_{a_m}; \qquad m = 1, 2, \ldots, n\,.$$

PROOF. For $m \leqq n$, let $\pi \in \Pi_m$, $\alpha \in \mathscr{A}_m$, $p = p(\alpha, \pi)$ (cf. (1.2), Definitions 1.7, 1.8). Let p' denote the permutation operator inverse to p. Then

$$\begin{aligned} pG(\alpha) &= \{x \in \mathscr{E}_*:\ f_m[y(\alpha, p'x),\, p'x] < q\} \\ &= \{x \in \mathscr{E}_*:\ f_m[\pi y(\alpha, p'x),\, x] < q\} \end{aligned}$$

by hypothesis on f_m. By (1.18),

$$pG(\alpha) = \{x \in \mathscr{E}_*:\ f_m[y(\pi\alpha, x),\, x] < q\} = G(\pi\alpha)\,.$$

Since $f_m(v, x)$ is symmetric in the components of x, it is clear from (2.4) and Definitions 1.2 with $E_r(x) = \sum_{i=1}^{r} x_i / r$ that, $G(\alpha)$ is cyclically symmetric in the variables mentioned in Lemma 2.1.

THEOREM 2.1. *Let $X_1, \ldots, X_n$ be independent and identically distributed, and let $f_m(v, x)$ be symmetric in the components of $v \in \mathscr{V}_m$ and in the components of $x \in \mathscr{E}^n$, $m = 1, 2, \ldots, n$. If*

(2.7) *the common distribution function of the* $\{X_i\}_{i=1}^n$ *is continuous,*

or if

(2.8) $f_{m(x)}[z(x),x]$, *and, for* $\alpha \in \mathscr{A}$, $f_m[y(\alpha,x),x]$ *are continuous in* x *on* $\mathscr{E}^n$,

then for $m \leqq n$, $k \in \mathscr{K}_m$, *and* α *chosen arbitrarily from* $\mathscr{A}^k$, *the conditional distribution of* $f_M(Z,X)$ *given* D^k *is the distribution of* $f_m[y(\alpha,X),X]$.

We recall (cf. (1.14)) that $D^k \equiv_D \bigcup_{\alpha \in \mathscr{A}^k} D(\alpha)$; in Case 1, this is the event which occurs if the gcm has k_1 sides of horizontal extent 1, k_2 of horizontal extent 2, etc.

PROOF OF THEOREM 2.1. Since $H^k = HD^k$ (cf. (1.21) and (1.14)), the conclusion of Theorem 1.5 may be rewritten as follows: for $m \leqq n$, $k \in \mathscr{K}_m$, and $\alpha \in \mathscr{A}_m$,

$$P(D^k) = \prod_{i=1}^{n} 1/(k_i!\, i^{k_i}), \quad P(H \mid D^k) = P[H(\alpha)]\,, \tag{2.9}$$

the first equation being the special case $H(\alpha) \equiv \mathscr{E}_*$ of (1.45). If the common distribution of $X_1, \ldots, X_n$ is continuous, then not only are these random variables symmetrically dependent, but also $P(\mathscr{E}_*) = 1$. Using Lemma 2.1 to apply Theorem 1.5 with H replaced by G and $H(\alpha)$ by $G(\alpha)$, we obtain the desired conclusion. In case the common distribution function F of $X_1, \ldots, X_n$ is not continuous, let $\{X_i^\nu\}$, $\nu = 1, 2, \ldots$, be independent and identically distributed for fixed ν $(i = 1, 2, \ldots, n)$ with a common continuous distribution function F_ν. Let the sequence $\{F_\nu\}$ converge completely to F as $\nu \to \infty$ (i.e., at each point of continuity of F; cf. [9, p. 178]). Set

$$X^\nu = (X_1^\nu, \ldots, X_n^\nu), \qquad W = f_M(Z, X), \qquad W^\nu = f_{m(X^\nu)}[z(X^\nu), X^\nu]\,.$$

By the Helly–Bray Theorem ([9, p. 182]), the characteristic function of W^ν converges to that of W, hence the distribution function of W^ν converges completely to that of W. A similar argument applies to

$$W^\nu(\alpha) \equiv_D f_m[y(\alpha, X^\nu), X^\nu] \qquad \text{and} \qquad W(\alpha) \equiv_D f_m[y(\alpha, X), X]$$

for fixed $\alpha \in \mathscr{A}$. The conclusion of Theorem 2.1 now follows for the general case by letting $\nu \to \infty$.

3. Applications to tests against trend.

We shall first apply Theorem 2.1 to the problem of determining the distribution of the likelihood ratio when the population belongs to an exponential family, and the ultimate aim is testing against trend.

Let $F_0(x)$ be a non-degenerate distribution function, and let

$$\Theta(\tau) \equiv_D \log \int e^{x\tau}\, dF_0(x)\,,$$

the integral being assumed convergent in a neighborhood of $\tau = 0$. The distribution function

$$F(x;\,\tau) \equiv_D \int_{(-\infty,\,x)} \exp[z\tau - \Theta(\tau)]\, dF_0(z)$$

is the member corresponding to the parameter value τ of the *exponential family* generated by F_0. Let

$$\theta(\tau) \equiv_D \int x \exp[x\tau - \Theta(\tau)]\, dF_0(x)\,;$$

then $\theta(\tau) = \Theta'(\tau)$ is the mean of the distribution. The function $\theta(\tau)$ is strictly increasing.

Let n be a positive integer. For $i = 1, 2, \ldots, n$, let h_i be a positive integer. Let

$$\{\Xi_{i\nu}\}, \quad \nu = 1, 2, \ldots, h_i, \quad i = 1, 2, \ldots, n\,,$$

be independent random variables. For $\nu = 1, 2, \ldots, h_i$, let the distribution of $\Xi_{i\nu}$ be the member of a common exponential family (the same family for $i = 1, 2, \ldots, n$) corresponding to a parameter value τ_i, $i = 1, 2, \ldots, n$. Set $\theta_i \equiv_D \theta(\tau_i)$, $i = 1, 2, \ldots, n$. Let Ω denote the set of all n-tuples $\theta \equiv_D (\theta_1, \ldots, \theta_n)$, with each coordinate θ_i in the range of the function $\theta(\tau)$. Set

$$\Omega_0 \equiv_D \{\theta\colon\ \theta_1 = \theta_2 = \ldots = \theta_n\} \quad \text{and} \quad \Omega_1 \equiv_D \{\theta\colon\ \theta_1 \leqq \ldots \leqq \theta_n\}\,.$$

Consider the hypotheses

$$H_0\colon\ \theta_1 = \ldots = \theta_n \qquad (\text{that is, } \theta \in \Omega_0)$$

and

$$H_1\colon\ \theta_1 \leqq \ldots \leqq \theta_n \qquad (\text{that is, } \theta \in \Omega_1)\,.$$

The maximum for $\theta \in \Omega_1$ of the joint density function of the random variables

$$\{\Xi_{i\nu}\}, \quad \nu = 1, 2, \ldots, h_i, \quad i = 1, 2, \ldots, n$$

is found in [5], [4], and [6]. These results yield simple expressions, in terms of the gcm of the points

$$\left\{\left(\sum_{i=1}^{r} h_i,\ \sum_{i=1}^{r} \sum_{\nu=1}^{h_i} \xi_{i\nu}\right)\right\}_{r=0}^{n},$$

for the likelihood ratios in likelihood ratio tests of H_0 within Ω_1 and of

H_1 within Ω. Under the null hypothesis H_0, and *when the sample sizes are equal*, $h_1 = \ldots = h_n$, these likelihood ratios are found to be of the form $f_M(Z, X)$. Theorem 2.1 can then be used to effect an important reduction in the labor of computing the distribution of such a likelihood ratio under H_0.

Particularly simple and interesting results are obtained when the random variables $\{\Xi_{i\nu}\}$ are normally distributed with means θ_i and common standard deviation σ, known or unknown. This case is treated by Bartholomew in [3], *without the restriction to equal sample sizes.* (Dr. P. Armitage kindly called the attention of the writer to this interesting paper by Bartholomew.) It turns out that the general theory in Section 1, and Theorem 2.1, are not required in this situation; special properties of the normal distribution suffice. Let Λ_0 and Λ_1 denote the likelihood ratios for testing H_0 within Ω_1 and H_1 within Ω, respectively. Then if σ is known, the conditional distribution under H_0 of $-2\log\Lambda_0$ given $M = m$ is χ^2 with $m-1$ degrees of freedom, and that of $-2\log\Lambda_1$ is χ^2 with $n-m$ degrees of freedom. Let $N = \sum_{i=1}^n h_i$. If σ is unknown, the conditional distribution under H_0 of

$$[(N-m)/(m-1)]\,[(1/\Lambda_0)^{2/N} - 1]$$

given $M = m$ is F with $m-1$ and $N-m$ degrees of freedom, and that of

$$[(N-n)/(n-m)]\,[(1/\Lambda_1)^{2/N} - 1]$$

is F with $n-m$ and $N-n$ degrees of freedom.

Andersen's theorem is of interest also in connection with a possible distribution-free test against trend.

It is intuitively clear that if there is in fact an upward trend in the distributions of $X_1, \ldots, X_n$, then one should expect a larger number, M, of sides in the gcm than if they are identically distributed. One is thus led to a distribution-free test based on the statistic M, rejecting the hypothesis that the random variables are identically distributed if too large a value of M is observed. The question of interest is: against which alternatives, if any, is this, the *GCM* test, a "good" test?

It is easy to construct alternatives at which Mann's T test [10] is more powerful than the *GCM* test. For $1 \leqq i < j \leqq n$, let

$$Y_{ij} = 1 \quad \text{if} \quad X_i < X_j, \qquad Y_{ij} = 0 \quad \text{if} \quad X_i > X_j\,.$$

Then

$$T \equiv_D \sum_{j=1}^{n} \sum_{i=1}^{j-1} Y_{ij}\,.$$

Let the null hypothesis, H_0, assert that the random variables $X_1, \ldots, X_n$

are independent and have the same continuous distribution function. Under H_0, T is asymptotically normally distributed ([10], [8]) with

$$E(T) \sim n^2/4, \qquad V(T) \sim n^3/36 ,$$

while M is also asymptotically normally distributed with

$$E(M) \sim \log n, \qquad V(M) \sim \log n$$

(cf. [2], and [7, p. 242]).

Let r be a positive integer, and $n = 2r$. Let H_1 denote an alternative to H_0 according to which $X_1, \ldots, X_r$ possess a common distribution, while $X_{r+1}, \ldots, X_{2r}$ have in common another distribution such that $\Pr\{X_j > X_i\} = 1$ for $1 \leqq i \leqq r < j \leqq 2r$. Under H_1, as $r \to \infty$, the statistic M is asymptotically normal with

$$E(M) \sim 2 \log r, \qquad V(M) \sim 2 \log r .$$

The power of the GCM test at H_1 is therefore given asymptotically by

$$\begin{aligned} (3.1) \qquad & 1 - \Phi\{[\log 2r + c(\log 2r)^{\frac{1}{2}} - 2 \log r]/(2 \log r)^{\frac{1}{2}}\} \\ & \sim 1 - \Phi\{[c - (\log r)^{\frac{1}{2}}]/2^{\frac{1}{2}} , \end{aligned}$$

where

$$\Phi(u) \equiv_D (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{u} e^{-\frac{1}{2}z^2}\, dz ,$$

and where c (independent of r) determines the size of the test. Under H_1, as $r \to \infty$, the statistic T is asymptotically normal with

$$E(T) \sim 3r^2/2, \qquad V(T) \sim r^3/18 .$$

The power of the T test at H_1 is therefore given asymptotically by

$$\begin{aligned} (3.2) \qquad & 1 - \Phi\{[c(2r^3/q)^{\frac{1}{2}} + r^2 - 3r^2/2]/[(2r^3)^{\frac{1}{2}}/6]\} \\ & = 1 - \Phi[2c - 3(r/2)^{\frac{1}{2}}] . \end{aligned}$$

A comparison of (3.1) with (3.2) shows that as $r \to \infty$ the power at H_1 of the T test approaches 1 more rapidly than the power of the GCM test.

Mann remarks that the K test [10] is more powerful than the T test at alternatives specifying rapidly increasing trend. It would be interesting to compare the three tests at such an alternative.

REFERENCES

1. E. Sparre Andersen, *On sums of symmetrically dependent random variables*, Skand. Aktuarietidskr. 36 (1953), 123–138.

2. E. Sparre Andersen, *On the fluctuations of sums of random variables* II, Math. Scand. 2 (1954), 195–223.
3. D. J. Bartholomew, *A test of homogeneity for ordered alternatives*, Biometrika 46 (1959), 36–48.
4. H. R. Brunk, *Maximum likelihood estimates of monotone parameters*, Ann. Math. Statist. 26 (1955), 607–616.
5. H. D. Brunk, G. M. Ewing, W. R. Utz, *Minimizing integrals in certain classes of monotone functions*, Pacific J. Math. 7 (1957), 833–847.
6. C. van Eeden, *Maximum likelihood estimation of partially or completely ordered parameters* I–II, Nederl. Akad. Wetensch. Proc. Ser. A 60 (= Indag. Math. 19 (1957)), 128–136 and 201–211.
7. W. Feller, *An introduction to probability theory and its applications* I, Second edition, New York, 1957.
8. M. G. Kendall, *A new measure of rank correlation*, Biometrika 30 (1938), 81–93.
9. M. Loève, *Probability theory*, New York, 1955.
10. H. B. Mann, *Non-parametric tests against trend*, Econometrica 13 (1945), 245–259.
11. R. E. Miles, *The complete amalgamation into blocks, by weighted means, of a finite set of real numbers*, Biometrika 46 (1959), 317–327.
12. F. Spitzer, *A combinatorial lemma and its application to probability theory*, Trans. Amer. Math. Soc. 82 (1956), 323–339.

UNIVERSITY OF MISSOURI, MISSOURI, U. S. A.,
AND
UNIVERSITY OF COPENHAGEN, DENMARK