The Annals of Mathematical Statistics
1969, Vol. 40, No. 1, 259–269

BIVARIATE SYMMETRY TESTS: PARAMETRIC AND NONPARAMETRIC

BY C. B. BELL[1] AND H. SMITH HALLER

Case Western Reserve University

1. Introduction and summary. In testing whether a treatment has an effect or not, the experimenter is often obliged to use the same subjects for control and treated groups. In such a case it is generally unrealistic to assume independence and one is led to tests of bivariate symmetry. The object of this paper is to show that bivariate symmetry is not equivalent to univariate symmetry; that there exists a feasible procedure different from the likelihood ratio test in the normal case; that there is no unique "natural" concept of rank; that all distribution-free (DF) procedures are based on permutations; and that optimal DF procedures for a simple alternative are based on permutations of the likelihood function.

2. Terminology and notation. The tests are based on random samples $z = [x_1, y_1; x_2, y_2; \cdots; x_n, y_n]$ from bivariate continuous distribution F. In such a case the joint distribution is $F^{(n)}(z) = \prod_1^n F(x_i, y_i)$. The hypotheses classes are then

$$\Omega(H_0 \cup H_1) = \{F^{(n)}: F, \quad \text{continuous}\};$$

and $\Omega(H_0) = \{F^{(n)}: F, \text{continuous; and } F(x, y) = F(y, x) \text{ for all } (x, y)\}$.

Of prime interest are functions, sets and statistics which exhibit some invariance wrt $\Omega(H_0)$, and, hence, the invariance properties of $\Omega(H_0)$ itself.

DEFINITION 2.1. Wrt $\Omega(H_0)$

(i) A set B is *similar* if there exists α such that $P(B \mid J) = \alpha$ for all J in $\Omega(H_0)$;

(ii) A test function ψ is *similar* if there exists α with $\int \psi \, dJ = \alpha$ for all J in $\Omega(H_0)$.

(iii) A statistic T is DF (distribution-free) if there exists a univariate distribution Q such that $P\{T \leqq t \mid J\} = Q(t)$ for all t and for all J in $\Omega(H_0)$.

In order to construct such sets, tests and statistics one extends the ideas of Pitman (1937), Lehmann and Stein (1949), Hajek and Sidak (1967) and Bell and Doksum (1967) in constructing the maximal permutation group under which each element J of $\Omega(H_0)$ is invariant. It is easily seen that this group is

$$S' = S_2 \int S_n' \quad \text{the wreath product (Hall, (1959))}$$

of the symmetric group S_n' of the $n!$ permutations of the intact pairs with S_2.

The set $S'(z)$ of images of z under S' is called the *orbit* of z, and one sees that each J in $\Omega(H_0)$ is invariant over each orbit. By Scheffé (1943), one needs to

Received 8 August 1967.

[1] Current address: University of Michigan.

select points from the orbits; and this is facilitated by the use of the functions defined below. The first type of function distinguishes among all points on a.e. orbit, and the second type ignores the chronological order of the data.

DEFINITION 2.2.

(i) t is a B-Pitman function if $P\{t(z) = t(\delta(z)) | J\} = 0$ for all J in $\Omega(H_0)$ and all non-identity elements of δ of S'.

(ii) t^* is a BNS-Pitman function if (a) $t^*(z) = t^*(\gamma(z))$ for all γ in S_n', and (b) $P\{t^*(z) = t^*(\tau(z)) | J\} = 0$ for all non-identity elements τ of $\times_i^n S_{2,i}$.

It will be seen that all DF statistics can be expressed in terms of rankings of such functions. The two pertinent ranking statistics are

$$\tilde{R}(t(z)) = \sum_{\delta \varepsilon S'} \epsilon\{t(z) - t(\delta(z))\}; \quad \text{and}$$

$$R^*(t^*(z)) = \sum_{\gamma \varepsilon S_n'} \epsilon\{t(z) - t(\gamma(z))\},$$

where $\epsilon(U) = 0$ or 1 according as $U < 0$ or $U \geqq 0$.

One can now proceed to study the various formulations of symmetry.

3. Formulations of symmetry. In order to give directions to the tests and to demonstrate that several definitions of symmetry are not equivalent, the following are presented. Table 3.1 gives several formulations of symmetry under the assumptions of continuity, absolute continuity, and normality. As is customary (X, Y) is the "basic" bivariate random variable and (W, V) is the transformed random variable with $W = X + Y$ and $V = Y - X$.

For the table below F and G are the distributions of (X, Y) and (W, V), respectively. F_x, F_y and G_w are marginals, while $G(\cdot | w)$ is a conditional distribution. f, g, f_x, f_y, g_v and $g(\cdot | w)$ are the corresponding densities.

TABLE 3.1

Formulations of symmetry

	Continuous	Absolutely Continuous	Normal
(A)	$F(x, y) = F(y, x)$ for all x and y.	$f(x, y) = f(y, x)$ for all x and y.	$E(X) = E(Y)$ and Var $(X) =$ Var (Y)
(B)	$2G(w, 0) = G(w, v) + G(w, -v)$ for all w and v.	$g(w, v) = g(w, -v)$ for all w and v.	$E(V) = 0$ and Cov $(W, V) = 0$
(C)	$G(v \mid w) + G(-v \mid w) = 1$ for all w and v.	$g(v \mid w) = g(-v \mid w)$ for all w and v.	$E(V \mid w) = 0$ for all w.
(D)	$F_x = F_y$, i.e. the x- and y-marginals are equal.	$f_x = f_y$, i.e. the x- and y-marginal densities are equal.	$E(X) = E(Y)$ and Var $(X) =$ Var (Y).
(E)	$G_v(t) + G_v(-t) = 1$ for all t. (Univariate Symmetry).	$g_v(t) = g_v(-t)$ for all t.	$E(V) = 0$
(F)	$P\{X > Y\} = \frac{1}{2} = G_v(0)$	$\int_{-\infty}^{0} g_v(t)\, dt = \frac{1}{2}$	$E(V) = 0$
(G)	$G(0 \mid w) = \frac{1}{2}$ for all w.	$\int_{-\infty}^{0} g(v \mid w)\, dv = \frac{1}{2}$ for all w.	$E(V \mid w) = 0$ for all w.

One can readily establish

THEOREM 3.1. *For bivariate normals*

(i) *conditions* (A), (B), (C), (D) *and* (G) *are equivalent;*

(ii) *conditions* (E) *and* (F) *are equivalent; and*

(iii) *the conditions in* (i) *imply those in* (ii) *but not vice versa.*

The theorem in the continuous case is

THEOREM 3.2. *For continuous (and absolutely continuous) bivariate distributions*

(i) *conditions* (A), (B) *and* (C) *are equivalent;*

(ii) *no other equivalences hold;*

(iii) (A), (B) *and* (C) *imply each of the other conditions; and*

(iv) (F) *is implied by* (G), *and also by* (C).

In summary one should say that

(a) in the normal case symmetry is equivalent to equal means and variances of the (X, Y) distributions; and independence and $E(V) = 0$ for the (W, V) distributions; and

(b) in the continuous case symmetry is equivalent to the symmetry of the conditional distribution of V for every value of W.

4. Tests in the normal case. The usual first procedure in such a case is to derive the likelihood ratio statistic.

THEOREM 4.1. *The likelihood ratio test is of the form* $\psi(z) = 1$ *if* $T_1 \leqq \lambda_0(n, \alpha)$; $= 0$ *otherwise, where* $T_1 = [1 - \{r(W, V)^2\}][1 + U_1]^{-1}$, $r(W, V)$ *is the correlation coefficient of* W *and* V; *and* $U_1^2 = n(\bar{V})^2[\sum V_i^2 - n\bar{V}^2]^{-1}$.

To obtain the null distribution of T_1, one makes use of the following lemma from Kendall and Stuart (1958), p. 385.

LEMMA 4.1. *If* Cov $(W, V) = 0$, *then* $\bar{W}$, $\bar{V}$, s_W^2, s_V^2 *and* r^2 *are mutually independent.*

Now from elementary considerations it follows that

THEOREM 4.2.

(i) $(n - 1)U_1^2$ *is distributed as an* F*-statistic with* 1 *and* $(n - 1)$ *degrees of freedom.*

(ii) *Under* H_0, r *and* U_1^2 *are independent.*

(iii) *Under* H_0, T_1 *has density which vanishes outside of* $(0, 1)$ *and for* $0 < t < 1$ *satisfies*

$$\begin{aligned} h(t) &= \Gamma(\tfrac{1}{2}n)(\tfrac{1}{2})t^{\frac{1}{2}(n-4)}\Pi^{-1}(\Gamma(\tfrac{1}{2}(n - 2)))^{-1} \int_1^{t^{-1}} x^{\frac{1}{2}(n-2)} \\ &\quad \cdot (x - (n - 1)^{-1})^{-\frac{1}{2}}(1 - xt)^{-\frac{1}{2}}[x + ((n - 2)/(n - 1))]^{-\frac{1}{2}n}\, dx \\ &= \Gamma(m)[2\Pi\Gamma(m - 1)b^{\frac{1}{2}}]^{-1}a^{m-1}t^{m-\frac{1}{2}5} \int_{B_1}^{B_2} (c - \sin\theta)^{m-1}\, d\theta \end{aligned}$$

where $a = (2m - 2)/(2m - 1)$, $b = t^{-1} + a$, $B_1 = -\frac{1}{2}\Pi$,

$$B_2 = \text{arc sin}\, \{[(1 + a)^{-1} - \tfrac{1}{2}(1 + b^{-1})]2b(b - 1)^{-1}\}.$$

In view of Theorem 4.2 and well-known results it is seen that

$$T_2 = r(W, V)(n - 2)^{\frac{1}{2}}[1 - r^2(W, V)]^{-\frac{1}{2}} \quad \text{and} \quad U_1 = \bar{V}n^{\frac{1}{2}}s_V^{-1}$$

are mutually independent, and have t-distributions with $(n - 2)$ and $(n - 1)$ degrees of freedom, respectively. Further it is clear, T_2 is a "good" statistic to test "cov $(W, V) = 0$" (or $\sigma_1^2 = \sigma_2^2$), while U_1 is directed toward "$E(V) = 0$" (or $\mu_1 = \mu_2$), i.e. each statistic is directed toward one of the facets of symmetry in the normal case. Consequently, another test of bivariate symmetry in the normal case can be given by

THEOREM 4.3. *Let* $\psi = 1$ *if* $|T_2| > t(\beta_1, n - 2)$ or $|U_1| > t(\beta_2, n - 1)$, *and* $\psi = 0$ *otherwise, where* $t(p, m)$ *is the* 100_pth *percentile of a t-distribution with m degrees of freedom and* T_2 *and* U_1 *are as in the preceding paragraph. Then* ψ *is a test function of size* $\alpha = (1 - 2\beta_1)(1 - 2\beta_2)$.

Of course, infinitely many tests are possible in the normal case, but attention will now be turned to the continuous case. Initial considerations will be of rank tests.

5. Rank Sets. The usual (e.g. 1, 2, 3, 5) procedure for characterizing rank sets (and, hence, rank tests) is as follows:

(1) Find a group G of 1-1 monotone transformations (of the sample space onto itself) which generates $\Omega(H_0)$.

For such a group three conditions are satisfied

(1.1) $\Omega(H_0)$ is an equivalence class of $\Omega(H_0 \cup H_1)$ under G.

If B is invariant under G, then

(1.2) B is similar wrt $\Omega(H_0)$; and

(1.3) B is SDF (strongly distribution-free), I.E., $P\{B \mid J_g\} = P\{B \mid J\}$ for J in $\Omega(H_0 \cup H_1)$ and for all g in G.

(2) If $\Omega(H_0 \cup H_1)$ is chosen to be appropriately complete, then one finds

(2.1) if A is SDF, A is equivalent to a set invariant under G.

The sets invariant under G are then referred to as "the rank sets."

For the bivariate symmetry problem both an (X, Y) and a (W, V) sample space have been considered and so there are in some sense two classes of candidates for "the rank sets."

Let $B(s, \gamma) = \{\gamma(s(x_1, y_1)) < \cdots < \gamma(s(x_n, y_n))\}$ for each continuous real-valued function s on R_g. Then the "most natural" sets to consider are (for each γ in S_n):

$$B(x; \gamma) = \{\gamma(x_1) < \cdots < \gamma(x_n)\}; \qquad B(y; \gamma) = \{\gamma(y_1) < \cdots < \gamma(y_n)\};$$

$$B(w; \gamma) = \{\gamma(w_1) < \cdots < \gamma(w_n)\}; \qquad B(v; \gamma) = \{\gamma(v_1) < \cdots < \gamma(v_n)\}.$$

However, under H_0 the V-distribution is symmetric about zero, and one will wish to consider

$$B(|v|; \gamma) = \{\gamma(|v_1|) < \cdots < \gamma(|v_n|)\}.$$

Further, let $\underline{m}_i = \min (x_i, y_i)$, $\bar{m}_i = \max (x_i, y_i)$ and $\epsilon_i = \epsilon(y_i - x_i) = \epsilon(v_i)$, where $\epsilon(v) = 1$ if $v \geqq 0$, and 0 otherwise. These lead to sets of the form

$$B(\underline{m}; \gamma) = \{\gamma(\underline{m}_1) < \cdots < \gamma(\underline{m}_1)\};$$

$$B(\bar{m}; \gamma) = \{\gamma(\bar{m}_1) < \cdots < \gamma(\bar{m}_n)\}$$

and

$$B(\epsilon_1', \cdots, \epsilon_n') = \{\epsilon_1 = \epsilon_1', \cdots, \epsilon_n = \epsilon_n'\}.$$

The pertinent groups are

$$G_2^*(2n) = \{\bar{g}: \bar{g}(r_1, \cdots, r_{2n}) = [\hat{g}(r_1), \cdots, \hat{g}(r_{2n})]\},$$

where $\hat{g}$ is a 1-1 monotone increasing continuous transformation of the real line R_1 onto R_1, to be considered on the (X, Y) sample space as well as the (W, V) sample space; and [see Lehmann (1959), p. 234].

$$G_2^*(n) \times H_2^*(n) = \{(\bar{g}, \bar{h}): (\bar{g}, \bar{h})(w_1, \cdots, w_n, v_1, \cdots, v_n)$$
$$= [\hat{g}(w_1), \cdots, \hat{g}(w_n); \hat{h}(v_1), \cdots, \hat{h}(v_n)]\},$$

where $\hat{g}$ is as above and $\hat{h}$ is a monotone increasing continuous *odd* transformation of R_1 onto R_1, to be considered on the (W, V) sample space alone.

From elementary considerations the similarity results below follow.

LEMMA 5.1.

(i) *Each* $B(x;\gamma), B(y;\gamma), B(w;\gamma), B(v;\gamma), B(|v|;\gamma), B(\underline{m};\gamma)$ *and* $B(\bar{m};\gamma)$ *is similar of size* $(n!)^{-1}$ wrt $\Omega(H_0)$.

(ii) *Each* $B(\epsilon_1', \cdots, \epsilon_n')$ *is similar of size* 2^{-n}.

(iii) *Each* $B(\epsilon_1', \cdots, \epsilon_n')B(s;\gamma)$ *for* $s = \underline{m}, \bar{m}, |v|$ *and* w, *is similar of size* $(n!)^{-1}2^{-n} = K_n^{-1}$.

(iv) *Further, if* S *is an arbitrary continuous real-valued function on the plane* R_2, *each of the sets* $B(s;\gamma)$ *is similar of size* $(n!)^{-1}$.

For the groups the following results are valid.

LEMMA 5.2.

(i) *Neither* $G_2^*(2n)$ *nor* $G_2^*(n) \times H_2^*(n)$ *generates* $\Omega(H_0)$ *or its subfamily with strictly monotone marginals*;

(ii) *the sets in Lemma* 5.1 (i) *except those involving* $B(|v|;\gamma)$ *are invariant under* $G_2^*(2n)$;

(iii) *the* $B(\epsilon_1', \cdots, \epsilon_n')$ *are invariant under* $G_2^*(2n)$ *on the* (X, Y) *space*;

(iv) *the* $B(|v|, \gamma)$ *and* $B(\epsilon_1', \cdots, \epsilon_n')$ *are invariant under* $G_2^*(n) \times H_2^*(n)$ *on the* (W, V) *space*; *and*

(v) *intersections are invariant if their factors are.*

That the "2-sample" ranks sets in the (X, Y)-space are not similar is seen from

EXAMPLE 5.1. Let $H(x_1, y_1, x_2, y_2) = \frac{1}{4}[x_1y_1 + x_1^2y_1^2][x_2y_2 + x_2^2y_2^2]$ on the unit hypercube in R_4. Then $\{x_1 < y_1 < x_2 < y_2\}$ and its 7 images under S' have probability $\frac{11}{240}$. While for $\{x_1 < x_2 < y_1 < y_2\}$ and $\{x_1 < y_2 < x_2 < y_1\}$ the probabilities are $\frac{29}{720}$ and $\frac{7}{180}$, respectively.

The authors were unable (see OPEN PROBLEMS) to find a group G which generates $\Omega(H_0)$ (or some dense subfamily). Hence no characterization of rank sets is possible here. However, one can consider several rank statistics.

6. Rank tests.

(1) *Sign test.* The simplest rank statistic for the bivariate symmetry hypothe-

sis is:

$$S_n = \sum_1^n \epsilon(y_i - x_i) = \sum_1^n \epsilon(v_i).$$

Under H_0, S_n has a binomial distribution with parameters n and $\frac{1}{2}$. However, it is clear that for any non-symmetric distribution with $P\{Y \geqq X\} = \frac{1}{2}$, or Median $(V) = 0$, the power of any test based solely on S_n is equal to its size.

(2) *Two-sample procedure.* Lehmann (1959), p. 234, and several other authors suggest tests which are based only on the $v_i = y_i - x_i$ in the following manner. If $S_n = k$, let $\{Y_1', \cdots, Y_k'\} = \{|V_i|: V_i \geqq 0\}$ and $\{X_1', \cdots, X_{n-k}'\} = \{|V_i|: V_i < 0\}$. Then apply any 2-sample DF test to the resulting two samples. For example the Mann-Whitney-Wilcoxon statistic becomes $\sum_{i=1}^n R(|v_i|)\epsilon(v_i)$, where $R(|v_i|)$ is the rank of $|v_i|$ among $|v_1|, |v_2|, \cdots, |v_n|$.

One sees immediately that such tests have power equal to size for alternatives with the V-marginal symmetric about zero. Further, there is some difficulty in applying the test if S_n is near 0 or n. For this latter reason, one introduces the modification below.

(3) *Modified two-sample procedure.* Let $D(r, m)$ be a 2-sample DF statistic for sample sizes r, and m, respectively; $C(\alpha, r, m)$ be a critical region of size related to α for each α, r and m; and $k_1 = k_1(\alpha, n)$, $k_2(\alpha, n)$ and $\{C(\alpha, r, m)\}$ be chosen such that

$$\begin{aligned} \alpha &= P\{S_n < k_1\} + S_n > k_2\} \\ &\quad + \sum_{r=k_1}^{k_2} P\{S_n = r\} P\{D(r, n-r) \,\varepsilon\, C(\alpha, r, n-r)\} \\ &= \sum_{r=1}^{k_1-1} \binom{n}{r} 2^{-n} + \sum_{r=k_2+1}^{n} \binom{n}{r} 2^{-n} \\ &\quad + \sum_{r=k_1}^{k_2} \binom{n}{r} 2^{-n} P\{D(r, n-r) \,\varepsilon\, C(\alpha, r, n-r)\}. \end{aligned}$$

Then let

$$\begin{aligned} \psi(z) &= 1 \quad \text{if} \quad S_n < k_1 \quad \text{or} \quad > k_2 \\ &= 1 \quad \text{if} \quad k_1 \leqq S_n \leqq k_2 \quad \text{and} \quad D(S_n, n - S_n) \,\varepsilon\, C(\alpha, S_n, n - S_n) \\ &= 0 \quad \text{otherwise.} \end{aligned}$$

This procedure eliminates the second difficulty above, but the power is still α when the V-marginal is symmetric about zero, i.e., when one has univariate symmetry.

From Section 3 it is clear that univariate and bivariate symmetry are not equivalent. It is therefore desirable to bring into the test procedure some dependence on W.

(4) *A chi-square procedure.* Let $-\infty = w_0' < w_1' < w_2' < \cdots < w_k' = +\infty$ be arbitrary points on the w-axis; let $T_j'(t) = \{\epsilon(w_j' - t) - \epsilon(w_{j-1}' - t)\}$; and $p_j' = P\{w_{j-1}' < w < w_j'\}$. Then $n_j' = \sum_{i=1}^n T_j'(W_i)\epsilon(V_i)$ is the number of sample points (W_i, V_i) with $w_{j-1}' < W_i \leqq w_j'$ and $V_i \geqq 0$ and $n_j'' = \sum_i T_j'(W_i)[1 - \epsilon(V_i)]$ is analogously interpreted. Under H_0, n_j' has a binomial

distribution with parameters n and $\frac{1}{2}p_j$. Therefore, the statistic

$$\sum_{j=1}^{k} (n_j' - \tfrac{1}{2}np_j')^2(np_j')^{-1} + \sum_{j=1}^{k} (n_j'' - np_j')^2(np_j')^{-1}$$

is DF, and has as asymptotic null distribution a chi-square distribution with $2k - 1$ degrees of freedom.

Unfortunately, the $\{p_j'\}$ will, in general, be unknown. Further, an arbitrary choice of $\{w_i'\}$ and estimation of the probabilities may lead to a poor approximation of the asymptotic distribution, due to small $E(n_j')$. An alternative procedure is to choose the $\{w_i'\}$ as appropriate order statistics of the W's.

(5) *A chi-square procedure based on order statistics.* Let $p_1, p_2, \cdots, p_k$ be probabilities which sum to one: $P_i = \sum_1^i p_r$; $-\infty = W(0), W(1), \cdots, W(n), W(n+1) = \infty$ be the order statistics of the W-sample; $T_j(t) = \{\epsilon(W(P_jn) - t) - (W(P_{j-1}n) - t)\}$; and $n_j = \sum_{i=1}^{n} T_j(W_i)\epsilon(V_i)$. Then one has

LEMMA 6.1. *Under H_0 and conditionally given $W(P_1n), W(P_2n), \cdots, W(P_{k-1}n)$, the $n_1, \cdots, n_k$ are independent and binomially distributed with parameters $np_1, np_2, \cdots, np_k$, respectively, and $\frac{1}{2}$.*

Therefore, for the statistic $S_1' = 4\sum_{j=1}^{k} (n_j - \frac{1}{2}np_j)^2(np_j)^{-1}$ it is immediate that

THEOREM 6.1.

(i) *S_1' is a* DF ("rank") *statistic;*

(ii) *any test based solely on S_1' has power equal to size for all alternatives with the property $P\{V \geqq 0 \mid w\} = \frac{1}{2}$ for all w; and*

(iii) *under H_0, S_1' has asymptotically a chi-square distribution with k degrees of freedom.*

In order to rectify the situation in Theorem 5.1 (ii), one first thinks of dividing the V-axis as was initially done with the W-axis. A personal discussion with H. Rubin resulted in the consideration of V order statistics in the manner to be described below.

(6) *A modified chi-square procedure.* In addition to the $p_1, \cdots, p_k$, one selects $\{q_{11}, \cdots, q_{m_1,1}; \cdots q_{1k}, \cdots, q_{m_k,k}\}$ with the property that $\sum_{r=1}^{m_j} q_{rj} = 1$ for $j = 1, 2, \cdots, k$. Now let $Q_{kj} = \sum_{r=1}^{k} q_{rj}$, let the collection

$$\{V'_{rj} : r = 1, 2, \cdots, np_j\}$$

represent the collection,

$$\{|V_i| : W(P_{j-1}n) < W_i \leqq W(P_jn)\},$$

the jth subsample; $V_j'(Q_{ij}P_jn)$ be the $(Q_{ij}Pn)$th order statistic of this latter jth subsample for all j. Further, let

$$S_{ij}(s) = \{\epsilon[V_j'(Q_{ij}P_jn) - s] - \epsilon[V_j'(Q_{i-1,j}P_jn) - s]\}.$$

Then

$$n_{rj} = \sum_{i=1}^{n} T_j(W_i)S_{ij}(|V_i|)\epsilon(V_i)$$

is the number of sample points whose W-values lie in $[V_j'(Q_{r-1,j}P_jn), V_j'(Q_{r,j}P_jn)]$.

Analogous to Lemma 6.1 one has

LEMMA 6.2. *Under* H_0 *and conditionally given* $W(P_1n), \cdots, W(P_{j-1}n)$ *and* $\{V_j'(Q_{ij}P_jn): i = 1, 2, \cdots, m_j - 1, j = 1, 2, \cdots, k\}$,

(i) *each* n_{rj} *is binomially distributed with parameters* $q_{ij}p_jn$ *and* $\frac{1}{2}$; *and*

(ii) *the* $\{n_{rj} : r = 1, 2, \cdots, m_j, j = 1, 2, \cdots, k\}$ *are independent.*

Now let $S_2' = 4 \sum_{j=1}^{k} \sum_{r=1}^{m_j} (n_{rj} - q_{ij}p_jn2^{-1})^2[np_jq_{ij}]^{-1}$, the "natural chi-square" statistics for this partition of the plane. It now follows that

THEOREM 6.2.

(i) S_2' *is a* DF (*"rank"*) *statistic, and*

(ii) *under* H_0, S_2' *has asymptotically a chi-square distribution with* $\sum_{j=1}^{k} m_j$ *degrees of freedom.*

(iii) *Further for an arbitrary specified alternative with the* $P\{V \geqq 0 \mid w\} = \frac{1}{2}$ *for all* w, *there exist* $\{p_j\}$ *and* $\{q_{ij}\}$ *such that* S_2' *has asymptotic power greater than its size.*

Of course, the number of possible "rank" procedures is unlimited. However, attention will now be given to constructing the class of all DF statistics.

7. The family of all DF statistics. In order to characterize the class of all similar test functions, one proceeds as in Bell (1967), and finds

THEOREM 7.1.

$$\int \psi \, dF^{(n)} = \alpha \quad \textit{for all} \quad F^{(n)} \quad \text{in} \quad \Omega(H_0) \text{ iff } \textstyle\sum_{\delta \varepsilon S'} \psi(\delta(z)) = \alpha K_n$$

for a.a.z, *where* $K_n = (n!)(2^n)$.

This theorem characterizes the class of all DF test functions and from this theorem one derives immediately the fundamental result concerning similar sets (analogous to that of Bell and Doksum (1967)).

COROLLARY 7.1.

(i) *If* A *is similar* wrt $\Omega(H_0)$, *then* $P(A)$ *is one of the values* $K_n^{-1}k$ *for* $k = 0, 1, 2, \cdots, K_n = (n!)(2^n)$.

(ii) *The following conditions are equivalent*:

(a) A *is similar of size* α;

(b) *there exists a B-Pitman function* t *such that* $A \equiv \{\tilde{R}(t) \leqq \alpha K_n\}$; *and*

(c) A *contains a proportion of* α *of* a.e. *orbit.*

(iii) *If one replaces "similar" by* "NS *similar"*; *"B-Pitman" by* "BNS-*Pitman"* *and "*$(n!)(2^n)$*" by "*(2^n)*"; "*t*" by "*t^**"; and "*$\tilde{R}$*" by "*R^**", then* (i) *is valid, and* (ii)(a) *and* (ii)(b) *are equivalent.*

One notes that the sets $\{A_1, \cdots, A_{K_n}\}$ constitute an (essentially) maximal similar partition of R_{2n}, if $A_i = \{\tilde{R}(t) = i\}$ and $K_n = (n!)(2^n)$. This is so in the sense that none of these contains a proper similar subset of smaller probability and the complement of their union is a null set.

The characterization of the family of all non-randomized DF statistics follows immediately.

THEOREM 7.2.

(i) *If* T *is* DF [NS *and* DF] wrt $\Omega(H_0)$, *then* T *has a discrete null hypothesis* *distribution with probabilities which are integral multiples of* $(n!)^{-1}2^{-n} = K_n^{-1}[2^{-n}]$.

(ii) T *is* DF [NS *and* DF] wrt $\Omega(H_0)$ *if there exists a B-Pitman function* t[BNS-*Pitman function* t^*] *and a measurable function* $U_2[U_2^*]$ *such that* $T \equiv U_2\{R(t)\}[\equiv U_2^*\{R^*(t^*)\}]$.

(iii) *For each discrete distribution* F_0 *with probabilities which are integral multiples of* $(n!)^{-1}2^{-n}[2^{-n}]$, *there exists a* DF [NS *and* DF] *statistic with null distribution* F_0.

Now one turns to the "best" DF test for a simple alternative.

8. Optimal tests and alternatives. For specific simple alternatives one can obtain the (most powerful) MPDF test by simply applying the Neyman-Pearson Lemma to each *orbit*. This is what Lehmann and Stein (1949) and Lehmann (1959), p. 185 did in obtaining the MP permutation tests for several hypotheses. However, now it is known (Theorem 6.2) that all DF statistics are functions of permutation statistics. Hence one has

THEOREM 8.1. *The* MPDF *level* α *test of* H_0 *against simple* H_1 *is of the form*

$$\begin{aligned}\psi(z) &= 1 \quad \text{if} \quad \tilde{R}[t(z)] > k\\ &= \lambda \quad \text{if} \quad \tilde{R}[t(z)] = k\\ &= 0 \quad \text{if} \quad \tilde{R}[t(z)] < k\end{aligned}$$

where t *is a B-Pitman function whose ordering of the points on the orbits is consistent with that of* L_1, *the joint density of* Z *under* H_1.

[One should note here that if the sample is non-random in the sense of non-identical marginals or dependence, but L_1 is invariant under S_n', then the power of the MPDF test is α. See Bell and Donoghue (1968)].

EXAMPLE 8.1. For the general bivariate normal distribution it is easy to see that a MPDF test is based on the NS Pitman function

$$\begin{aligned}t(z) = (\sigma_1^2 - \sigma_2^2)[\{\textstyle\sum (x_i - y_i)(x_i + y_i)\}/(\sigma_1^2 + \sigma_2^2 + 2\sigma_1\sigma_2\rho)^{\frac{1}{2}}]\\ + \{(\mu_1 - \mu_2)(\sigma_1^2 + \sigma_2^2 - 2\sigma_1\sigma_2\rho)^{-\frac{1}{2}}\\ - (\sigma_1^2 - \sigma_2^2)(\mu_1 + \mu_2)(\sigma_1^2 + \sigma_2^2 + 2\sigma_1\sigma_2\rho)^{-\frac{1}{2}}\}n(\bar{x} - \bar{y}).\end{aligned}$$

When $\sigma_1 = \sigma_2$ it is clear that $t(z)$ reduces to $\bar{x} - \bar{y}$ or equivalently $\bar{v}$. However for arbitrary σ_1 and σ_2, note that the NS Pitman function depends upon V and W.

EXAMPLE 8.2. For the case where X and Y are independent with unequal marginals, consider the pdf $h(x, y) = \lambda \exp(-x - \lambda y)$. From Theorem 8.1 it is immediate that $t_3 = [\text{sgn}(1 - \lambda)](\sum x_i)$. One notes here as in the preceding example that t_3 is a B-NS Pitman function.

One notices that in both examples above the MPDF test depends on the parameters as well as the form of the distribution of the simple alternatives. In order to obtain "optimal" tests which do not depend on θ, one introduces the concept of L (locally) MPDF tests. (This parallels the development in Bell and Doksum (1967)).

To this end one needs the following notation:

(a) A, an interval containing 0;

(b) $\{Q(\theta, \cdot, \cdot) : \theta \in A\}$, a class of absolutely continuous bivariate distributions with

(i) $Q(0; x, y) = Q(0; y, x)$ for all x and y; and

(ii) regularity conditions given below in terms of the power function $\beta(\phi, \cdot)$

(c) $L(\theta; z) = \prod_i q(x_i, y_i)$, where q is the density of Q;

(d) $\beta(\psi; \theta)$, the power of the test ϕ against alternative F_θ ; and

(e) $\beta^{(r)}(\psi; \cdot)$, the rth derivative wrt θ of $\beta(\psi; \theta)$.

DEFINITION 8.1. A level α test ϕ_0 is LMP for testing $\theta = 0$ vs. $\theta > 0$ if, given any other level α test ψ_1 there exists $\Lambda(\psi_1)$, such that $\beta(\phi_0; \theta) \geqq \beta(\psi_1; \theta)$ for all θ with $0 < \theta < \Lambda(\psi_1)$.

It follows, e.g. from the proof of Theorem 4.1 of Bell and Doksum (1967), that

THEOREM 8.2. *Let there exist $a > 0$ such that for all θ in $(-a, a)$, all Z in R_{2n} and all level α* DF *tests*

(i) $L^{(r)}(\theta; z)$ *exists and is continuous, and*

(ii) $\beta^{(r)}(\psi; \theta)$ *exists, is continuous and can be obtained by differentiating inside the integral sign*

$$\beta(\psi; \theta) = \int \{[\textstyle\sum_\gamma \psi[\gamma(s)]]L(\theta; \gamma(s))][\sum_\gamma L(\theta; \gamma(s))]^{-1}\} \, dP^S(s)$$

where S is the vector (min (X_1, Y_1); max (X_1, Y_1); min (X_2, Y_2); max (X_2, Y_2); $\cdots$, max (X_n, Y_n); *and*

(iii) $r \geqq 1$ *be the smallest integer for which $L^{(r)}(0; z)$ is not invariant under S_n'.*

Then, the LMPDF *level α test against the family of alternatives $\{Q(\theta; \cdot, \cdot) : \theta > 0\}$ is of the form of Theorem* 8.1 *with $t(z) = L^{(r)}(0; z)$.*

Of interest at this point are the families of alternatives wrt which a given DF test is MPDF and LMPDF.

THEOREM 8.3. *Let $\theta \geqq \theta_0$ index a family of distributions with densities $q(\theta; x, y) \exp\{a(\theta)\bar{t}(x, y) + b(\theta) + \bar{S}(x, y, \theta)\}$, with $a(\theta) > 0$. Then against this family of alternatives the tests based on $\tilde{R}(t(z))$; where $t(z) = \sum_1^n \bar{t}(x_i, y_i)$ are*

(i) MPDF *if $\sum_{i=1}^n \bar{S}(x_i, y_i, \theta)$ is invariant under S'; and are*

(ii) LMPDF *if $a(\theta_0) = 0$ and $\sum_{i=1}^n \bar{S}(x_i, y_i, \theta) = S_1(z, \theta) + S_2(z, \theta)$ where $S_1(z, \theta)$ is invariant under S' and $S_2(z, \theta) = o(a(\theta))$.*

Unfortunately, the preceding theorem does not give the largest families wrt which $\tilde{R}(t)$ is MPDF and LMPDF. This and other open problems are mentioned in the next section.

9. Open problems. There are, of course, many open problems related to the symmetry hypothesis. The authors have chosen to state several problems which are more or less closely related to the development above.

(A) *Rank sets.* Which are the "natural" rank sets for the symmetry problem? Several families of sets are discussed in Section 5 but none seems more desirable than the others. Does there exist a MP rank test in some reasonable sense?

(B) *Group generating* $\Omega(H_0)$. Closely related ot question (a) is: Which group G of transformations on R_{2n} generates $\Omega(H_0)$ or an appropriately dense subfamily thereof? H. Rubin conjectured that in general there is difficulty in finding a transformation which is invertible and which can transform the distribution.

$$H_1(x, y) = \tfrac{1}{2}\{F(x)F(y) + G(x)G(y)\} \quad \text{into}$$

$$H_2(x, y) = \tfrac{1}{2}\{F(x)G(y) + F(y)G(y)\}.$$

(C) *All randomized tests.* Theorem 6.1 gives all randomized test functions. Can all randomized test functions be represented in terms of randomized statistics?

(D) *Optimality for large families of alternatives.* Can power bounds, as in Chapman (1958) or Bell, Moser and Thompson (1966) be found? Are there minimax tests in the sense of Bell and Doksum (1967) and Doksum (1966) available? What is known about ARE's? Very few power results are available for finite sample size.

(E) *Koopman-Pitman-type theory.* What is the largest family wrt which a given DF statistic is MPDF? LMPDF? In Section 8 the classical Koopman-Pitman development is paralleled, but use has not been made, e.g., of all monotone functions of the statistic. Also, one might consider classes of alternatives for non-random samples, e.g. as in Bell and Donoghue (1968).

REFERENCES

[1] Bell, C. B. and Donoghue, John (1968). Distribution-Free Tests of Randomness. To be published, *Sankhyā*.

[2] Bell, C. B. (1967). Completeness Theorems for Characterizing Distribution-Free Statistics. Submitted for publication.

[3] Bell, C. B. (1964). A Characterization of Multisample Distribution-Free Statistics. *Ann. Math. Statist.* **35** 735–738.

[4] Bell, C. B., Moser, J. M., and Thompson, R. (1966). Goodness Criteria for Two Sample Distribution-Free Tests. *Ann. Math. Statist.* **37** 133–142.

[5] Bell, C. B. and Doksum, K. A. (1967). Distribution-Free Tests of Independence. *Ann. Math. Statist.* **38** 429–446.

[6] Chapman, D. G. (1958). A Comparative Study of Several One Sided Goodness-Of-Fit Tests. *Ann. Math. Statist.* **29** 655–674.

[7] Doksum, K. A. (1966). Asymptotically Minimax Distribution-Free Procedures. *Ann. Math. Statist.* **37** 619–628.

[8] Hájek, Jaroslav and Šidák, Zbyněk. (1967). *Theory of Rank Tests*. Academia Publishing House of the Czechoslovak Academy of Sciences, Prague.

[9] Hall, Marshall, Jr. (1959). *The Theory of Groups*. Macmillan, New York.

[10] Kendall, Maurice G. and Stuart, Alan (1958). The Advanced Theory of Statistics **1**. Griffin, London.

[11] Lehmann, E. L. and Stein, C. (1949). On the Theory of Some Nonparametric hypotheses. *Ann. Math. Statist.* **20** 28–45.

[12] Lehmann, E. L. (1959). *Testing Statistical Hypotheses*. Wiley, New York.

[13] Pitman, E. J. G. (1937). Significance Tests Which May Be Applied to Samples from Any Population. *Suppl. J. Roy. Statist. Soc. L.* 119–130.

[14] Scheffé, H. (1943). On a Measure Problem Arising in the Theory of Nonparametric Tests. *Ann. Math. Statist.* **14** 227–233.

[15] Wormleighton, R. (1959). Some tests of permutation symmetry. *Ann. Math. Statist.* **30** 1005–1017.