A BIVARIATE SIGN TEST FOR LOCATION

By Shoutir Kishore Chatterjee

University of Calcutta

0. Summary. A strictly distribution-free test has been proposed for testing that several independent pairs of random variables have locations as specified. Under the null hypothesis, the test statistic has asymptotically a chi-square distribution with degrees of freedom two. The test has been shown to be unbiased and consistent against reasonable classes of alternatives. Asymptotic power and efficiency have been examined.

1. Introduction. In the univariate case there are two distinct hypotheses testing problems which may be tackled by the well-known sign test. The first is the problem of testing whether several random variables have specified medians; and the second is that of testing whether these are distributed symmetrically about specified medians (see [13], pp. 147, 242). In the bivariate case also, we would have analogous problems of location and symmetry. In this paper we develop a distribution-free test which may be considered as a bivariate extension of the sign test for location.

Let

$$(X_1, Y_1), (X_2, Y_2), \cdots, (X_n, Y_n) \tag{1.1}$$

be n independent pairs of random variables and let $F_k(x, y)$ be the cumulative distribution function (cdf) of (X_k, Y_k), $k = 1, \cdots, n$. We shall assume $(F_1, F_2, \cdots, F_n) \varepsilon \Omega_n$ where Ω_n is the set of all n-tuplets of bivariate cdf's. Our problem is to test whether $F_1, F_2, \cdots, F_n$ have n specified pairs of (marginal) medians. By suitably choosing the origins we can assume that the pair of hypothetical median values is $(0, 0)$ for each (X_k, Y_k). Then if we write ω_n for the subset of Ω_n within which

$$F_k(0, \infty) = F_k(\infty, 0) = \tfrac{1}{2}, \qquad k = 1, \cdots, n, \tag{1.2}$$

we can formally state the null hypothesis as

$$H_0 : (F_1, F_2, \cdots, F_n) \varepsilon \omega_n . \tag{1.3}$$

In the literature, several tests go by the name of bivariate sign test, notably those of Hodges [7] and Blumen [4] (see also Klotz [11], Joffe and Klotz [10], Klotz [12], Bennett [2] and Bickel [3]). But it is to be emphasized that the tests of both Hodges and Blumen are concerned with the problem of symmetry as follows: Let us call the distribution of a pair of random variables (X, Y) as diagonally symmetric about (a, b) if the distribution of $(X - a, Y - b)$ is same as that of $(a - X, b - Y)$. The tests of Hodges and Blumen require that, under the null hypothesis, all the distributions $F_k(x, y)$ be diagonally symmetric about

Received 14 June 1965; revised 3 March 1966.

Institute of Mathematical Statistics is collaborating with JSTOR to digitize, preserve, and extend access to *The Annals of Mathematical Statistics.* www.jstor.org

the origin. These can be used for testing (1.3) provided each $F_k(x, y)$ is known to be diagonally symmetric about its median point (assumed unique).

To develop a distribution-free test in the general case, we shall require some alternative characterizations of the subset ω_n. First we note that (1.2) is equivalent to

$$(1.4)\quad \begin{aligned} P\{X_k \leqq 0,\ Y_k \leqq 0\} &= P\{X_k > 0,\ Y_k > 0\};\\ P\{X_k \leqq 0,\ Y_k > 0\} &= P\{X_k > 0,\ Y_k \leqq 0\}, \qquad k = 1, 2, \cdots, n. \end{aligned}$$

For the pair (X_k, Y_k) let us call the events $X_k \leqq 0$, $Y_k \leqq 0$ and $X_k > 0$, $Y_k > 0$ as concordances of the first and second kind, and the events $X_k \leqq 0$, $Y_k > 0$ and $X_k > 0$, $Y_k \leqq 0$ as discordances of the first and second kind respectively. Writing γ_k for the probability of concordance of (X_k, Y_k), let us assume

$$(1.5)\qquad 0 < \gamma_k < 1, \qquad k = 1, 2, \cdots, n.$$

Then for (X_k, Y_k) we may define the conditional probabilities of a concordance of the first kind given concordance and a discordance of the first kind given discordance, and let these be θ_k and τ_k respectively. In terms of these (1.4) may equivalently be stated as

$$(1.6)\qquad \theta_k = \tfrac{1}{2}, \qquad \tau_k = \tfrac{1}{2}, \qquad k = 1, 2, \cdots, n.$$

ω_n can be characterized by any one of the equivalent conditions (1.2), (1.4) and (1.6).

A non-parametric solution to the above problem would be developed in the following sections. We mention here that as regards the principle of solution this paper has some similarity with an earlier paper (Chatterjee and Sen [6]) where the analogous two-sample problem was solved.

2. Formulation of the test. Let us denote by C_1, C_2 and D_1, D_2 the numbers of concordances of the first and second kind and numbers of discordances of the first and second kind respectively among the n pairs (1.1). Also, let $C = C_1 + C_2$ and $D = D_1 + D_2$ denote the total numbers of concordances and discordances. We shall denote the running variables corresponding to the random variables C_1, C_2, C, D_1, D_2, D by using lower-case letters. From the conditions (1.4) it seems that a reasonable test for H_0 would be obtained by comparing C_1 with C_2 and D_1 with D_2. However, the joint distribution of C_1, C_2, D_1, D_2 is seen to be not distribution-free even under H_0. To circumvent this difficulty, we consider the conditional distribution of these, given which pairs among $(X_1, Y_1) \cdots (X_n, Y_n)$ are concordant and which discordant.

Let c be any integer ($0 \leqq c \leqq n$). Consider any partition

$$(2.1)\quad (i_1, i_2, \cdots, i_c), (i_{c+1}, \cdots, i_n), \quad i_1 < \cdots < i_c, \quad i_{c+1} < \cdots < i_n,$$

of the set of numbers $1, 2, \cdots, n$ into two disjoint subsets containing c and $n - c$ numbers respectively. (If $c = 0$, or n, the first or second subset is of course empty.) Let $\mathcal{E}_{i_1 \cdots i_c}$ be the event that among (1.1), the i_1th, $\cdots$, i_cth pairs are concordant, and the rest are discordant. Then

$$P\{\mathcal{E}_{i_1\cdots i_c}\} = \gamma_{i_1}\cdots\gamma_{i_c}(1-\gamma_{i_{c+1}})\cdots(1-\gamma_{i_n}). \tag{2.2}$$

The probability that $\mathcal{E}_{i_1\cdots i_c}$ will occur and there will be just c_1 concordance of the first kind and just d_1 discordances of the first kind ($0 \leqq c_1 \leqq c$, $0 \leqq d_1 \leqq n - c$) is

$$\begin{aligned} &\gamma_{i_1}\cdots\gamma_{i_c}(1-\gamma_{i_{c+1}})\cdots(1-\gamma_{i_n}) \\ &\qquad\cdot\{(1-\theta_{i_1})\cdots(1-\theta_{i_c})\textstyle\sum_1\prod_1[\theta_{i_j}/(1-\theta_{i_j})]\} \\ &\qquad\cdot\{(1-\tau_{i_{c+1}})\cdots(1-\tau_{i_n})\textstyle\sum_2\prod_2[\tau_{i_j}/(1-\tau_{i_j})]\}\end{aligned} \tag{2.3}$$

where $\prod_1$ denotes a product over a subset of c_1 of the values $1, 2, \cdots, c$ of j, and $\sum_1$ denotes the sum over all the $\binom{c}{c_1}$ such subsets; similarly, $\prod_2$ denotes a product over a subset of d_1 of the values $c + 1, \cdots, n$ of j, and $\sum_2$ denotes the sum over all the $\binom{n-c}{d_1}$ such subsets. From (2.2) and (2.3), the required conditional distribution of C_1 and D_1 is given by

$$\begin{aligned} &P\{C_1 = c_1, D_1 = d_1 \mid \mathcal{E}_{i_1\cdots i_c}\} \\ &\qquad = \{(1-\theta_{i_1})\cdots(1-\theta_{i_c})\textstyle\sum_1\prod_1[\theta_{i_j}/(1-\theta_{i_j})]\} \\ &\qquad\cdot\{(1-\tau_{i_{c+1}})\cdots(1-\tau_{i_n})\textstyle\sum_2\prod_2[\tau_{i_j}/(1-\tau_{i_j})]\}, \\ &\qquad\qquad 0 \leqq c_1 \leqq c,\ 0 \leqq d_1 \leqq n - c.\end{aligned} \tag{2.4}$$

Under H_0 given by (1.6), whatever $(F_1, \cdots, F_n) \,\varepsilon\, \omega_n$, (2.4) gives

$$\begin{aligned} &P\{C_1 = c_1, D_1 = d_1 \mid \mathcal{E}_{i_1\cdots i_c}, H_0\} \\ &\qquad = P\{C_1 = c_1, D_1 = d_1 \mid C = c, H_0\} = \binom{c}{c_1}\binom{n-c}{d_1}2^{-n}, \\ &\qquad\qquad 0 \leqq c_1 \leqq c,\ 0 \leqq d_1 \leqq n - c.\end{aligned} \tag{2.5}$$

Thus, under H_0, given $C = c$, C_1 and D_1 are independently distributed as binomial random variables with parameters $(c, \frac{1}{2})$ and $(n - c, \frac{1}{2})$ respectively. Hence, for testing H_0, it seems reasonable to use the statistic

$$T = (4/C)(C_1 - C/2)^2 + [4/(n - C)][D_1 - (n - C)/2]^2. \tag{2.6}$$

(For $C = 0$ or n, one of the terms in T is absent.) Given $C = c$, the conditional distribution of T under H_0 would be clearly distribution-free. Let $F_n(t \mid c)$ denote the cdf of this distribution as obtained by summing (2.5) over those combinations (c_1, d_1) for which the value of T does not exceed t. For any α $(0 < \alpha < 1)$ let $t_{n,\alpha}(c)$ be the value of t for which $1 - F_n(t \mid c) \leqq \alpha < 1 - F_n(t - 0 \mid c)$. We then define a critical function $\varphi(t, c)$ which assumes the value 0 for $t < t_{n,\alpha}(c)$ and 1 for $t > t_{n,\alpha}(c)$. For $t = t_{n,\alpha}(c)$ we take $\varphi(t, c) = a_{n,\alpha}(c)$ where $0 \leqq a_{n,\alpha}(c) < 1$ is chosen such that

$$E\{\varphi(T, C) \mid C = c, H_0\} = \alpha. \tag{2.7}$$

We then consider the randomized test:

$$\text{(2.8)} \qquad \text{reject } H_0 \text{ with probability } \varphi(T, C).$$

By (2.7) the level of significance $E\varphi(T, C)$ of this test would be α whatever $(F_1, \cdots, F_n) \varepsilon \omega_n$ and therefore the test would be strictly distribution-free. In practice one would usually prefer the non-randomized test:

$$\text{(2.9)} \qquad \text{reject } H_0 \text{ if } T > t_{n,\alpha}(C); \text{ otherwise, accept } H_0.$$

Whatever $(F_1, \cdots, F_n) \varepsilon \omega_n$, the level of significance of this would not exceed α.

Although the above development of the test is based on heuristic reasoning, it is of interest to note that, from considerations of invariance and sufficiency, we are led to a class of tests similar to and including (2.8). The problem of testing the hypothesis (1.3) remains invariant under all transformations $X_k' = g_k^{(1)}(X_k)$, $Y_k' = g_k^{(2)}(Y_k)$, $k = 1, \cdots, n$, with strictly increasing continuous functions $g_k^{(i)}(x)$ for which $g_k^{(i)}(0) = 0$, $i = 1, 2$, and all permutations of the n pairs (1.1). Therefore, if we restrict ourselves to invariant tests, we need consider only tests based on the set of maximal invariants C_1, D_1, C. The joint distribution of these is obtained by summing (2.3) over all the $\binom{n}{c}$ partitions (2.1). From this, the power of any critical function $\psi(c_1, d_1, c)$ is seen to be continuous in the parameters, so that a test $\psi(c_1, d_1, c)$ for (1.6) is unbiased only if it is similar (see [13], p. 125). Now it may be noted that C is a sufficient statistic for the family of distributions of C_1, D_1, and C under H_0. Further, as the induced family of distributions of C includes all binomial distributions with parameter n, the statistic C may be shown to be complete (see Lehmann and Scheffé [14], pp. 312, 315). Hence by a well known result ([14], p. 318) $\psi(c_1, d_1, c)$ would give a similar level-α test for H_0, if and only if

$$\text{(2.10)} \qquad E\{\psi(C_1, D_1, C) \mid C = c, H_0\} = \alpha, \qquad \text{for all } c, \qquad 0 \leqq c \leqq n.$$

Thus we are led to consider the class of critical functions which satisfy (2.10). By (2.7), the critical function $\varphi(t, c)$ belongs to this class. The choice of the particular critical function $\varphi(t, c)$ is made from considerations of practical convenience and is to some extent justified by its properties of unbiasedness and consistency to be established later.

To apply the tests (2.8) or (2.9) in practice, we require knowledge of $t_{n,\alpha}(c)$ (and also of $a_{n,\alpha}(c)$ for (2.8)). For small n, $t_{n,\alpha}(c)$ may be found easily with the help of binomial tables. We now show that, for large n, we may approximate $F_n(t \mid c)$ by the cdf $F_{\chi_2^2}(t)$ of the χ_2^2-distribution and use the upper α-point $\chi^2_{2,\alpha}$ of the latter for $t_{n,\alpha}(c)$.

From (2.5) and (2.6), by DeMoivre-Laplace limit theorem, we get that $F_n(t \mid c) \to F_{\chi_2^2}(t)$ if c and $n - c$ both tend to infinity. Now, writing $C = \sum_1^n Z_k$, $n - C = \sum_1^n (1 - Z_k)$ where Z_k assumes the value 1 or 0 according as (X_k, Y_k) is concordant or discordant, by Kolmogorov's three series criterion ([15], p. 237) we see that, as $n \to \infty$, if

$$\text{(2.11)} \qquad \textstyle\sum_{k=1}^{\infty} \gamma_k = \infty, \qquad \sum_{k=1}^{\infty} (1 - \gamma_k) = \infty,$$

both $C \to \infty$ and $n - C \to \infty$ almost surely. (If the first or second series in (2.11) is convergent C or $n - C$ converges almost surely.) Hence, assuming (2.11), we have $F_n(t \mid C) \to_{\text{a.s.}} F_{\chi_2^2}(t)$ so that $t_{n,\alpha}(C) \to_{\text{a.s.}} \chi^2_{2,\alpha}$, $a_{n,\alpha}(C) \to_{\text{a.s.}} 0$. Therefore, for large n, the tests (2.8) and (2.9) both approach the test:

$$\text{reject } H_0, \text{ if } T > \chi^2_{2,\alpha}; \text{ otherwise, accept } H_0. \tag{2.12}$$

The level of significance $1 - EF_n(\chi^2_{2,\alpha} \mid C)$ of the test (2.12) depends on the particular $(F_1, \cdots, F_n) \, \varepsilon \, \omega_n$ through the distribution of C. However, as $n \to \infty$, under (2.11), it tends to α whatever F_k, $k = 1, 2, \cdots$.

3. Unbiasedness of the test. In this section we shall show that the test (2.8) is unbiased against a specific class of alternatives. For that we first prove two lemmas.

In what follows we shall use S_r generically for a random variable representing the number of successes in r independent trials. For any set E of possible values, $P\{S_r \, \varepsilon \, E \mid p_1, p_2, \cdots, p_r\}$ will stand for the probability of the event $S_r \, \varepsilon \, E$ when the probabilities of success in the r trials taken in order are $p_1, p_2, \cdots, p_r$.

LEMMA 3.1. *If $p_1, p_2, \cdots, p_n$ satisfy either*

$$0 \leqq p_1, p_2, \cdots, p_n \leqq \tfrac{1}{2}, \quad \textit{or} \quad \tfrac{1}{2} \leqq p_1, p_2, \cdots, p_n \leqq 1, \tag{3.1}$$

and s be an integer $0 \leqq s \leqq n/2$,

$$\begin{aligned} P\{s \leqq S_n \leqq n - s \mid p_1, p_2, \cdots, p_n\} \\ \leqq P\{s \leqq S_n \leqq n - s \mid \tfrac{1}{2}, \tfrac{1}{2}, \cdots, \tfrac{1}{2}\}. \end{aligned} \tag{3.2}$$

PROOF. We prove the lemma for the case $\frac{1}{2} \leqq p_1, p_2, \cdots, p_n \leqq 1$. The case $0 \leqq p_1, p_2, \cdots, p_n \leqq \frac{1}{2}$ then follows by interchanging the roles of success and failure.

First suppose $1 \leqq s \leqq n/2$. The event $s \leqq S_n \leqq n - s$ can materialize in three mutually exclusive and exhaustive ways: (i) the first trial gives a success and the last $(n - 1)$ trials give just $(s - 1)$ successes, (ii) the first gives a failure and the last $(n - 1)$ trials give $n - s$ successes, (iii) the number of successes in the last $(n - 1)$ trials is not less than s and not more than $n - s - 1$. Hence,

$$\begin{aligned} P\{s \leqq S_n \leqq n - s \mid p_1, p_2, \cdots, p_n\} &= p_1 P\{S_{n-1} = s - 1 \mid p_2, \cdots, p_n\} \\ &+ (1 - p_1) P\{S_{n-1} = n - s \mid p_2, \cdots, p_n\} \\ &+ P\{s \leqq S_{n-1} \leqq n - s - 1 \mid p_2, \cdots, p_n\}. \end{aligned} \tag{3.3}$$

Let p_1' be any number, $0 \leqq p_1' \leqq p_1$. Replacing p_1 in (3.3) by p_1' and from the resulting expression subtracting (3.3), we get

$$\begin{aligned} P\{s \leqq S_n \leqq n - s \mid p_1', p_2, \cdots, p_n\} \\ - P\{s \leqq S_n \leqq n - s \mid p_1, p_2, \cdots, p_n\} \\ = (p_1 - p_1')[P\{S_{n-1} = n - s \mid p_2, \cdots, p_n\} \\ P\{S_{n-1} = s - 1 \mid p_2, \cdots, p_n\}]. \end{aligned} \tag{3.4}$$

Now it is known (see [13], p. 220) that for $p_2, \cdots, p_n \geqq \frac{1}{2}$, $\binom{n-1}{r}^{-1}P\{S_{n-1} = r \mid p_2, \cdots, p_n\}$ is a non-decreasing function of r. As $s - 1 \leqq n - s$ we conclude that the difference (3.4) is non-negative. This statement is also trivially true for $s = 0$. Further we can always renumber the trials so that the ith trial becomes the first. From these considerations it follows that for any integer i $(1 \leqq i \leqq n)$, if $0 \leqq p_i' \leqq p_i$,

$$P\{s \leqq S_n \leqq n - s \mid p_1, \cdots, p_{i-1}, p_i, p_{i+1}, \cdots, p_n\}$$
$$\leqq P\{s \leqq S_n \leqq n - s \mid p_1, \cdots, p_{i-1}, p_i', p_{i+1}, \cdots, p_n\},$$

irrespective of the values of $p_1, \cdots, p_{i-1}, p_{i+1}, \cdots, p_n$ provided these are all $\geqq \frac{1}{2}$. Therefore as $p_1, p_2, \cdots, p_n \geqq \frac{1}{2}$, successively taking $i = 1, 2, \cdots, n$ and putting $p_i' = \frac{1}{2}$ in every case, we get (3.2).

In the following, for any two random variables with cdf's $F(x)$, $G(x)$ respectively, we shall say that the first random variable (or its distribution) is stochastically larger than the second (or its distribution) when $F(x) \leqq G(x)$ for all x (c.f. [13], p. 73). Then we get the following corollary to the above lemma.

Corollary. *Whatever $a > 0$, the distribution of $a(S_n - \frac{1}{2}n)^2$ when $p_1, p_2, \cdots, p_n$ satisfy* (3.1), *is stochastically larger than the distribution of the same when $p_1 = \cdots = p_n = \frac{1}{2}$.*

This follows immediately from the above lemma, since under (3.1) whatever $x \geqq 0$,

$$P\{a(S_n - \tfrac{1}{2}n)^2 \leqq x \mid p_1, \cdots, p_n\} \leqq P\{a(S_n - \tfrac{1}{2}n)^2 \leqq x \mid \tfrac{1}{2}, \cdots, \tfrac{1}{2}\}.$$

Lemma 3.2. *If U_1, U_2 and V_1, V_2 are pairs of independent random variables, such that U_i is stochastically larger than V_i, $i = 1, 2$, then $U_1 + U_2$ is stochastically larger than $V_1 + V_2$.*

Proof. Let the cdf's of U_2 and V_2 be $F(x)$ and $G(x)$ respectively. Then, for any x, $G(x) \geqq F(x)$, so that

$$P\{V_1 + V_2 \leqq x\} = EG(x - V_1) \geqq EF(x - V_1). \tag{3.5}$$

Again, as U_1 is stochastically larger than V_1, and for a fixed x, $F(x - v)$ is non-increasing in v, by a well known result (see [13], p. 73), we have

$$EF(x - V_1) \geqq EF(x - U_1) = P\{U_1 + U_2 \leqq x\}. \tag{3.6}$$

Combining (3.5) and (3.6), the lemma follows.

Now we consider the unbiasedness of the test (2.8). For this, we define the particular subclass $\Omega_n^{(1)}$ of alternatives $(F_1, F_2, \cdots, F_n) \varepsilon\, \Omega_n - \omega_n$ for which the following two conditions both hold:

$$\text{either} \quad \theta_1, \cdots, \theta_n \geqq \tfrac{1}{2}, \quad \text{or} \quad \theta_1, \cdots, \theta_n \leqq \tfrac{1}{2}; \tag{3.7}$$

$$\text{either} \quad \tau_1, \cdots, \tau_n \geqq \tfrac{1}{2}, \quad \text{or} \quad \tau_1, \cdots, \tau_n \leqq \tfrac{1}{2}. \tag{3.8}$$

We now prove the following:

Theorem 3.1. *The test* (2.8) *is unbiased against the class of alternatives $\Omega_n^{(1)}$ given by* (3.7) *and* (3.8).

PROOF. Take any $(F_1, \cdots, F_n) \varepsilon \Omega_n^{(1)}$ and let the corresponding alternative be called H_1. For any partition (2.1) of $\{1, 2, \cdots, n\}$ let the event $\mathcal{E}_{i_1 \cdots i_c}$ be defined as in Section 2. By (2.4), given $\mathcal{E}_{i_1 \cdots i_c}$ C_1 and D_1 are generally independently distributed as numbers of successes in sets of c and $n - c$ independent trials with probabilities of success $\theta_{i_1}, \cdots, \theta_{i_c}$ and $\tau_{i_{c+1}}, \cdots, \tau_{i_n}$ respectively. So by (1.6), (3.7) and (3.8), and the Corollary to Lemma 3.1, given $\mathcal{E}_{i_1 \cdots i_c}$ the conditional distributions of $4c^{-1}\{C_1 - \frac{1}{2}c\}^2$ and $4(n-c)^{-1}\{D_1 - \frac{1}{2}(n-c)\}^2$ under H_1 are respectively stochastically larger than those under H_0. (In the particular cases $c = 0, n$, the statement holds for only one of the expressions). Hence applying Lemma 3.2, from (2.6) it follows that, given $\mathcal{E}_{i_1 \cdots i_c}$, the conditional distribution of T under H_1 is stochastically larger than that under H_0. As given $\mathcal{E}_{i_1 \cdots i_c}$, c is fixed and the critical function $\varphi(t, c)$ defined in Section 2 is non-decreasing in t for a fixed c, by a well known result (see [13], p. 73), it follows that

$$E\{\varphi(T, C) \mid \mathcal{E}_{i_1 \cdots i_c}, H_1\} \geqq E\{\varphi(T, C) \mid \mathcal{E}_{i_1 \cdots i_c}, H_0\}. \tag{3.9}$$

As (3.9) holds whatever c $(0 \leqq c \leqq n)$ and whatever the partition (2.1), unconditionally also,

$$E\{\varphi(T, C) \mid H_1\} \geqq E\{\varphi(T, C) \mid H_0\}.$$

This proves the theorem.

To see the significance of the conditions (3.7)–(3.8), we note that from (1.7),

$$\begin{aligned} \theta_k &\gtreqless \tfrac{1}{2} \Leftrightarrow F_k(0, \infty) + F_k(\infty, 0) - 1 \gtreqless 0; \\ \tau_k &\gtreqless \tfrac{1}{2} \Leftrightarrow F_k(0, \infty) - F_k(\infty, 0) \gtreqless 0. \end{aligned} \tag{3.10}$$

Hence a little consideration shows that (3.7)–(3.8) hold if and only if either (i) $F_k(0, \infty) - \frac{1}{2}$, $k = 1, \cdots, n$, are of the same sign, and $|F_k(0, \infty) - \frac{1}{2}| \geqq |F_k(\infty, 0) - \frac{1}{2}|$, $k = 1, \cdots, n$, or (ii) $F_k(\infty, 0) - \frac{1}{2}$, $k = 1, \cdots, n$, are of the same sign, and $|F_k(0, \infty) - \frac{1}{2}| \leqq |F_k(\infty, 0) - \frac{1}{2}|$, $k = 1, \cdots, n$. When F_1, $F_2, \cdots, F_n$ are identical, this holds for every alternative, and thus the test (2.13) is unbiased against all alternatives.

Finally, it may be noted that as the tests (2.9) and (2.12) are not similar these can not be unbiased against the alternatives $\Omega_n^{(1)}$ (see [13], p. 125).

4. Consistency of the tests. We shall first prove the consistency of the large-sample test (2.12). As in studying consistency we make $n \to \infty$, we consider now the set Ω_∞ of all infinite sequences $\{F_k(x, y)\}$ of bivariate cdf's such that $0 < \gamma_k < 1$ for all k. The null hypothetical set ω_∞ will be the subset of sequences for which (1.2), or equivalently, (1.6) holds for all k. Consider the subset Ω_∞^* of $\Omega_\infty - \omega_\infty$ such that, for any $\{F_k(x, y)\} \varepsilon \Omega_\infty^*$, at least one of the following two conditions holds as $n \to \infty$:

$$\liminf |n^{-1} \textstyle\sum_1^n \gamma_k(\theta_k - \frac{1}{2})| > 0; \tag{4.1}$$

$$\liminf |n^{-1} \textstyle\sum_1^n (1 - \gamma_k)(\tau_k - \frac{1}{2})| > 0. \tag{4.2}$$

From the definitions of γ_k, θ_k, and τ_k in Section 1, we see that conditions (4.1) and (4.2) are respectively equivalent to

$$\liminf |n^{-1} \textstyle\sum_1^n \{F_k(0, \infty) + F_k(\infty, 0) - \frac{1}{2}\}| > 0; \tag{4.3}$$

$$\liminf |n^{-1} \textstyle\sum_1^n \{F_k(0, \infty) - F_k(\infty, 0)\}| > 0. \tag{4.4}$$

If all the distributions F_k are identical then at least one of (4.3) and (4.4) is satisfied by every alternative to H_0.

Now consider the test (2.12). Let $\{F_k^*(x, y)\}$ be a sequence belonging to the class Ω_∞^* and suppose (X_k, Y_k) follows the distribution $F_k^*(x, y)$, $k = 1, 2, \cdots$.

From (2.6), we have

$$(4n)^{-1}T \geqq n^{-2}\{C_1 - \tfrac{1}{2}C\}^2 + n^{-2}\{D_1 - \tfrac{1}{2}(n - C)\}^2. \tag{4.5}$$

By Tshebysheff's inequality, as $n \to \infty$, with probability approaching unity $n^{-1}\{C_1 - \frac{1}{2}C\} \sim n^{-1}\sum_1^n \gamma_k(\theta_k - \frac{1}{2})$ and $n^{-1}\{D_1 - \frac{1}{2}(n - C)\} \sim n^{-1}\sum_1^n (1 - \gamma_k)\cdot(\tau_k - \frac{1}{2})$. Now, as at least one of (4.1) and (4.2) holds, by (4.5), as $n \to \infty$, with probability approaching unity, T/n is bounded away from zero. Hence

$$\lim_{n\to\infty} P\{T > \chi^2_{2,\alpha} \mid \{F_k^*\}\} = 1.$$

As, with n increasing, $t_{n,\alpha}(C) \to_p \chi^2_{2,\alpha}$ and $a_{n,\alpha}(C) \to_p 0$, we may conclude the following (see [8], p. 171–172).

THEOREM 4.1. *The tests* (2.8), (2.9) *and* (2.12) *are all consistent against the class of alternatives* Ω_∞^*.

5. Asymptotic power and relative efficiency. As, with increasing n, the tests (2.8) and (2.9) approach the test (2.12), for studying the asymptotic power, we confine our attention to (2.12). Throughout this section we shall consider the case when the observations represent independent samples from the same population with cdf, say, $F(x, y)$. $F(x, y)$ is known to lie in the set Ω of all bivariate cdf's $F(x, y)$ for which (1.5) holds. The null hypothetical set $\omega \subset \Omega$ consists of the cdf's $F(x, y)$ satisfying (1.2).

As the test (2.12) is consistent, to study its asymptotic power we take an $F_0(x, y) \in \omega$ and consider the sequence of alternatives represented by

$$F_{(n)}(x, y) = F_0(x + n^{-\frac{1}{2}}\xi_1, y + n^{-\frac{1}{2}}\xi_2), \qquad n = 1, 2, \cdots, \tag{5.1}$$

where $(\xi_1, \xi_2) \neq (0, 0)$. For each n, let the test (2.12) be based on a sample of size n taken from $F_{(n)}(x, y)$. Then, applying a form of bivariate central limit theorem due to Hoeffding and Robbins [9] to the joint distribution of $n^{-\frac{1}{2}}(C_2 - C_1)$ and $n^{-\frac{1}{2}}(D_2 - D_1)$, and using a well-known result [16] regarding the limiting distribution of continuous functions, it follows that, as $n \to \infty$, T is asymptotically distributed as a non-central χ_2^2 with non-centrality parameter

$$\Delta = \gamma_0^{-1}\{\xi_1 f_{0\cdot1}(0) + \xi_2 f_{0\cdot2}(0)\}^2 + (1 - \gamma_0)^{-1}\{\xi_1 f_{0\cdot1}(0) - \xi_2 f_{0\cdot2}(0)\}^2, \tag{5.2}$$

where $\gamma_0 = 2F_0(0, 0) - F_0(0, \infty) - F_0(\infty, 0) + 1$, and $F_0(x, \infty)$ and $F_0(\infty, y)$ are assumed absolutely continuous with densities $f_{0\cdot1}(x)$ and $f_{0\cdot2}(y)$. The asymptotic power is given by the tail-probability of this non-central χ^2-distribution.

As pointed out by a referee, in this case of identically distributed sample observations, Bickel [3] considers a class of asymptotically equivalent test statistics (the class $\hat{\mathfrak{M}}_n^2$) which includes T, and the above result follows from a general theorem of Bickel (see [3], pp. 163, 167).

Efficiency relative to Hotelling's T^2. When the parent cdf $F(x, y)$ is known *a priori* to be bivariate normal, taking the cdf of the distribution $N(0, 0, \sigma_1, \sigma_2, \rho)$ for $F_0(x, y)$ and writing $\xi_i' = \xi_i/\sigma_i$, $i = 1, 2$, (5.2) reduces to

$$\Delta = \tfrac{1}{2}[(\xi_1' + \xi_2')^2(\pi - \cos^{-1}\rho)^{-1} + (\xi_1' - \xi_2')^2(\cos^{-1}\rho)^{-1}].$$

In this case, the standard T^2-statistic of Hotelling is asymptotically distributed as a non-central χ_2^2 with non-centrality parameter $\Delta^* = (1 - \rho^2)^{-1} \cdot (\xi_1'^2 - 2\rho\xi_1'\xi_2' + \xi_2'^2)$. Hence for the given ρ, ξ_1', and ξ_2', the efficiency of the test (2.12) relative to the comparable Hotelling test is given by $e = \Delta/\Delta^*$. Bickel [3], [3a] gives the details of these results and the expression for e in connexion with his class of statistics $\hat{\mathfrak{M}}_n^2$ which includes T. He also obtains the maximum e^M and minimum e^m of e for a fixed ρ and studies their behaviour. It turns out that both these are symmetric about $\rho = 0$, and at this point both assume the value $2/\pi \cong .64$, which incidentally is the Pitman efficiency against normal alternatives of the univariate sign test. As $|\rho|$ increases e^M first rises slightly to a value of about .72 and then gradually decreases to $2/\pi$ as $|\rho| \to 1$. e^m monotonically falls off to 0 as $|\rho|$ increases to 1. For very high values of $|\rho|$, e will be high or low according as the deviation of the standardized median point is in a direction along or orthogonal to the axis of scatter of the distribution. Thus taking $\xi_1' = \xi_2'$ and making $\rho \to \pm 1$, we see $e \to 2/\pi$ and 0 respectively. On the other hand for moderate values of $|\rho|$, say $|\rho| < .50$, the efficiency will be satisfactory whatever the direction of deviation.

Efficiency relative to the Hodges test. As noted in Section 1, when it is known that the observations are independently taken from a distribution $F(x, y)$ diagonally symmetric about the median point, for testing H_0, we may also use the alternative bivariate sign test of Hodges. In this case it is of interest to compare the efficiency of the test (2.12) relative to the Hodges test. But the test-statistic for the latter is not asymptotically distributed as a central χ^2 under H_0 (see Joffe and Klotz [10]), and its asymptotic non-null distribution is not known. Therefore comparison of efficiency in terms of power, as in the preceding discussion is not possible. However, as in [10], we may derive the relative efficiency in the sense of Bahadur [1] in the particular case when $F(x, y)$ is bivariate normal.

Let all observations be taken from the distribution $F(x, y)$ which is bivariate normal with parameters $(\mu_1, \mu_2, \sigma_1, \sigma_2, \rho)$ $(|\rho| < 1)$, and let $\mu_1/\sigma_1 = \mu_1'$, $\mu_2/\sigma_2 = \mu_2'$. Also let T be defined as in (2.6) on the basis of a sample size n. Let us write

$$\begin{aligned} &F(0, 0) = P_1, \qquad 1 - F(0, \infty) - F(\infty, 0) + F(0, 0) = P_2,\\ (5.3)\quad &F(0, \infty) - F(0, 0) = P_3, \qquad F(\infty, 0) - F(0, 0) = P_4,\\ &P_1 + P_2 = \gamma, \qquad P_3 + P_4 = 1 - \gamma. \end{aligned}$$

Then from the results of Section 2, it follows that T satisfies the conditions of Bahadur [1], and its slope in the sense of Bahadur is

$$\begin{aligned}\operatorname{plim} n^{-1}T &= \gamma^{-1}(P_1 - P_2)^2 + (1-\gamma)^{-1}(P_3 - P_4)^2 \\ &= \gamma^{-1}\{\Phi(\mu_1') + \Phi(\mu_2') - 1\}^2 + (1-\gamma)^{-1}\{\Phi(\mu_1') - \Phi(\mu_2')\}^2\end{aligned}$$

where $\Phi(x)$ stands for the standard univariate normal cdf. The slope of the Hodges test in the present case has been shown to be (see [10]) $\{2\Phi(\Lambda) - 1\}^2$, where $\Lambda^2 = (1-\rho^2)^{-1}\{\mu_1'^2 - 2\rho\mu_1'\mu_2' + \mu_2'^2\}$. Hence, the Bahadur efficiency of the test (2.12) relative to the Hodges test would be

$$(5.4)\quad E = \{2\Phi(\Lambda) - 1\}^{-2}[\gamma^{-1}\{\Phi(\mu_1') + \Phi(\mu_2') - 1\}^2 + (1-\gamma)^{-1}\{\Phi(\mu_1') - \Phi(\mu_2')\}^2].$$

To calculate the limiting value of E as $\mu_1', \mu_2' \to 0$, let us put $\mu_1' = \delta h_1, \mu_2' = \delta h_2$, where $(h_1, h_2) \neq (0, 0)$ is fixed, and make $\delta \to 0$. Writing $\lambda^2 = (1-\rho^2)^{-1} \cdot \{h_1^2 - 2\rho h_1 h_2 + h_2^2\}$, as, whatever h_1, h_2, $\lim \gamma = \frac{1}{2} + \pi^{-1}\sin^{-1}\rho$, from (5.4), we derive

$$\lim_{\delta\to 0} E = \tfrac{1}{2}\lambda^{-2}\{(1 + (2/\pi)\sin^{-1}\rho)^{-1}(h_1 + h_2)^2 + (1 - (2/\pi)\sin^{-1}\rho)^{-1}(h_1 - h_2)^2\}.$$

This limiting efficiency clearly depends on h_1, h_2 which represent the direction in which (μ_1', μ_2') approaches $(0, 0)$. In particular we have, for $h_1 = h_2$, $\lim E = \{1 + (2/\pi)\sin^{-1}\rho\}^{-1}(1+\rho)$ and this is seen to be greater than or less than 1 depending on whether ρ is positive or negative. Similarly it is seen that for $h_1 = -h_2$, $\lim E$ is greater than or less than 1 depending on whether ρ is negative or positive. When either h_1 or h_2 is zero, E is seen to be less than 1 although it is close to 1 unless $|\rho|$ is very large.

The above asymptotic studies of the performance of the test (2.12) relative to the tests of Hotelling and Hodges of course do not tell us anything regarding the performance in small samples. For any parent cdf $F(x, y)$, if we use the notations (5.3), we can write the exact power of the test (2.12) as the multinomial sum

$$\sum [n!\, P_1^{c_1} P_2^{c-c_1} P_3^{d_1} P_4^{n-c-d_1} / c_1!\,(c - c_1)!\, d_1!\,(n - c - d_1)!],$$

the summation being taken over all values $0 \leqq c_1 \leqq c$, $0 \leqq d_1 \leqq n - c$, $0 \leqq c \leqq n$ for which the value of T given by (2.6) exceeds $\chi^2_{2,\alpha}$. For different parent cdf's $F(x, y)$, numerical studies such as in Klotz [12] would reveal the performance of the proposed test in small samples.

6. Concluding remarks. In the preceding sections we confined ourselves solely to the bivariate problem of testing whether several pairs of variables have specified location. However, some problems involving sets of three or more variables which can be reduced to bivariate problems by suitable transformations, would be tractable by identical techniques. Thus, if we have independent triplets $(Z_k^{(1)}, Z_k^{(2)}, Z_k^{(3)})$, $k = 1, 2, \cdots, n$, to test whether the variables within each triplet are identically located we may set $X_k = Z_k^{(1)} - Z_k^{(2)}$, $Y_k = Z_k^{(1)} - Z_k^{(3)}$ and apply the methods of Section 2. This gives an exact non-parametric test for identity of location without any assumption of symmetry of the original trivariate distribution. Another familiar problem is related to quadrivariate distributions in which the four variables group into natural pairs. Thus if we have n independent sets of observations $(X_k', Y_k', X_k'', Y_k'')$, $k = 1, \cdots, n$, and our problem is to test whether X' and Y' have respectively same locations as X'' and Y'' we can apply the proposed test on the differences $X_k = X_k' - X_k''$, $Y_k = Y_k' - Y_k''$. Here also in using this method we do not require any assumption of symmetry or interchangeability of the variables to be able to test the hypothesis.

Lastly one may note that, whereas the bivariate sign tests of Hodges [7] and Blumen [4] are invariant under non-singular transformations of the pairs (X_k, Y_k), $k = 1, \cdots, n$, the test proposed here is not. On the other hand it may be readily seen from the expression of the proposed test-statistic given by (2.6) that the latter is invariant under transformations in each variable which are zero-preserving and monotonic increasing (when the underlying cdf's are continuous the transformations need only be zero-preserving and monotonic). As such transformations will not necessarily preserve diagonal symmetry in the joint distribution, the tests of Hodges and Blumen will not be generally invariant under these. Invariance under a non-singular transformation will be meaningful when the choice of co-ordinate axes in the bivariate plane is more or less arbitrary. On the other hand, when one or the other of the variables is being measured on an indirect scale (such as on a calibrated instrument or on a psychophysical scale) invariance under a monotonic transformation preserving zero may be important.

Acknowledgment. The author is indebted to Mr. H. K. Nandi of Calcutta University for many useful discussions which helped, in particular, to simplify the proofs of Section 3. Some of the ideas in the paper germinated during the author's collaboration with Dr. P. K. Sen in the preparation of the paper [6]. Thanks are also due to the referee for his comments, and particularly, for pointing out a short-cut in the proof of Lemma 3.1.

REFERENCES

[1] Bahadur, R. R. (1960). Stochastic comparison of tests. *Ann. Math. Statist.* **31** 276–295.

[2] Bennett, B. M. (1962). On multivariate sign tests. *J. Roy. Statist. Soc. Ser. B* **24** 159–161.

[3] Bickel, Peter J. (1965). On some asymptotically nonparametric competitors of Hotelling's T^2. *Ann. Math. Statist.* **36** 160–173.

[3a] Bickel, Peter J. (1965). Corrections to 'On some asymptotically nonparametric competitors of Hotelling's T^2'. *Ann. Math. Statist.* **36** 1583.

[4] Blumen, I. (1958). A new bivariate sign test. *J. Amer. Statist. Assoc.* **53** 448–456.

[5] Chatterjee, S. K. (1964). An alternative bivariate sign test (abstract). *Proc. Indian Sci. Congress* (Pt. III), *Combined 51st and 52nd Session* **23**.

[6] Chatterjee, S. K. and Sen, P. K. (1964). Nonparametric tests for the bivariate two-sample location problem. *Calcutta Statist. Assoc. Bull.* **13** 18–58.

[7] Hodges, H. L., Jr. (1955). A bivariate sign test. *Ann. Math. Statist.* **26** 523–527.

[8] HOEFFDING, WASSILY (1952). The large-sample power of tests based on permutations of observations. *Ann. Math. Statist.* **23** 169–192.

[9] HOEFFDING, WASSILY and ROBBINS, H. (1948). A central limit theorem for dependent random variables. *Duke Math. J.* **15** 773–780.

[10] JOFFE, A. and KLOTZ, J. (1962). Null distribution and Bahadur efficiency of the Hodges bivariate sign test. *Ann. Math. Statist.* **33** 803–807.

[11] KLOTZ, JEROME (1959). Null distribution of the Hodges bivariate sign test. *Ann. Math. Statist.* **30** 1029–1033.

[12] KLOTZ, JEROME (1964). Small sample power of the bivariate sign tests of Blumen and Hodges. *Ann. Math. Statist.* **35** 1576–1582.

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

[14] LEHMANN, E. L. and SCHEFFÉ, H. (1950). Completeness, similar regions, and unbiased estimation—part I. *Sankhyā* **10** 305–340.

[15] LOÈVE, MICHEL (1955). *Probability Theory.* Van Nostrand, New York.

[16] SVERDRUP, EARLING (1952). The limit distribution of a continuous function of random variables. *Skand. Aktuarietidscr.* **35** 1–10.