MATHEMATICS

# A FAMILY OF PARAMETERFREE TESTS FOR SYMMETRY WITH RESPECT TO A GIVEN POINT. II

BY

J. HEMELRIJK

(Communicated by Prof. D. VAN DANTZIG at the meeting of June 24, 1950)

## 6. *Introduction.*

6. 1. In a previous paper on this subject [1]) an exact test has been given for the hypothesis $H_0$, that $n$ random variables $\mathbf{z}_i (i = 1, \ldots, n)$ are distributed independently, each with a probability distribution, which is symmetrical with respect to zero. We shall now give a generalisation of this test by describing a family of tests for $H_0$, which contains this one as a special case. The computations involved in the application of the test are described in section 11 and an example is given at the end of this paper in section 12.

6. 2. These tests will be based on the simultaneous application of the sign test, which depends on the number of positive and negative values among $z_1, \ldots, z_n$, and on the application of a parameterfree two sample test to the two groups of values $x_1, \ldots, x_{n_1}$ and $y_1, \ldots, y_{n_2}$ defined in section 3.

A two sample test is a test for the hypothesis $H'$, that two random samples $x_1, \ldots, x_{n_1}$ and $y_1, \ldots, y_{n_2}$ have been drawn independently from the same population. We shall mainly be concerned with a "family" of two-sample tests, consisting of those two-sample tests, which are based on the fact, that, assuming $H'$ to be true, all partitions of the $n_1 + n_2$ values $x_j$ and $y_k$ of the two samples, taken together, into two samples of $n_1$ and $n_2$ values respectively, have the same probability. This fact may also be expressed by saying, that, if $H'$ is true and the samples are drawn in a fixed order, all permutations of the obtained values are equally probable.

6. 3. Several two-sample tests have been developed on this basis, e.g. by E. J. PITMAN (1937), N. SMIRNOFF (1939) (using a theorem developed by A. KOLMOGOROFF (1933)), A. WALD and J. WOLFOWITZ (1940) and F. WILCOXON (1945). Wilcoxon's test was studied in detail by H. B. MANN and D. R. WHITNEY (1947).

## 7. *The main theorem.*

7. 1. Let $T$ be a two sample test of the type described above and let

---

[1]) These Proceedings 53, 941—955 (1950).

$\mathbf{u}_1, \ldots, \mathbf{u}_\nu$ be the statistics, on which $T$ is based. These statistics are known functions of the random variables $\mathbf{x}_1, \ldots, \mathbf{x}_{n_1}$ and $\mathbf{y}_1, \ldots, \mathbf{y}_{n_2}$ and $n_1$ and $n_2$ are given numbers. Usually $\nu = 1$, but this is by no means necessary. We shall therefore give the main theorem in the more general form with $\nu \geqq 1$.

Since we are using two-sample tests, which have been developed previously, we may assume the simultaneous distribution of $\mathbf{u}_1, \ldots, \mathbf{u}_\nu$, under the assumption that $H'$ is true, to be known. We shall denote by

$$G^* (u_1, \ldots, u_\nu) \tag{11}$$

the conditional simultaneous distribution function [2]) of $\mathbf{u}_1, \ldots, \mathbf{u}_\nu$, under the condition (denoted by the asterisk), that the two samples, taken together, assume the set of values $x_1, \ldots, x_{n_1}, y_1, \ldots, y_{n_2}$ and under the assumption, that $H'$ is true.

7. 2. If, instead of the two samples, we take the values $x_1, \ldots, x_{n_1}$ and $y_1, \ldots, y_{n_2}$ defined in section 3, it follows from lemma 3, that, if $H_0$ is true, if $\mathbf{n}_1 = n_1$ and if condition $Z$ is satisfied, the conditions indicated by the asterisk in (11) are satisfied too, and that (11) is the conditional distribution function of $\mathbf{u}_1, \ldots, \mathbf{u}_n$. We express this fact by changing the notation of this distribution function into

$$G (u_1, \ldots, u_\nu | Z ; \mathbf{n}_1 = n_1 ; H_0) \tag{12}$$

where $u_1, \ldots, u_\nu$ are derived from $x_1, \ldots, x_{n_1}$ and $y_1, \ldots, y_{n_2}$, the group of positive values and the group of negative values (taken positively) of the original observations $z_1, \ldots, z_n$, which are available to test the symmetry of the variables $\mathbf{z}_1, \ldots, \mathbf{z}_n$.

For $n_1 = 0$ and $n_1 = n - m$ (i.e. $n_2 = 0$) the statistics $\mathbf{u}_1, \ldots, \mathbf{u}_\nu$ have not yet been defined, since one of the groups $x_1, \ldots, x_{n_1}$ or $y_1, \ldots, y_{n_2}$ is empty in that case. Defining for this case $\mathbf{u}_1 = \ldots = \mathbf{u}_\nu = 0$, we find from lemma 2 and 3:

Theorem III: *If $H_0$ is true, the conditional simultaneous probability distribution of $\mathbf{n}_1$ and $\mathbf{u}_1, \ldots, \mathbf{u}_\nu$, under the condition $Z$, is given by*

$$\left\{ \begin{aligned} & \mathrm{P}\,[\mathbf{n}_1 = n_1 ; \mathbf{u}_1 \leqq u_1 ; \ldots ; \mathbf{u}_\nu \leqq u_\nu | Z ; H_0] = \\ & = 2^{-n+m} \binom{n-m}{n_1} G (u_1, \ldots, u_\nu | Z ; \mathbf{n}_1 = n_1 ; H_0), \end{aligned} \right. \tag{13}$$

with $0 \leqq n_1 \leqq n - m$.

Remarks: 1. If we want to test the hypothesis $H_0'$, that all $\mathbf{z}_i$ are distributed independently according to the *same* symmetrical probability distribution, $T$ need not be restricted to the family of tests described in 6. 2. For it is easy to prove, that under the hypothesis $H_0'$ and under the conditions $\mathbf{n}_1 = n_1$ and $\mathbf{m} = m$ the values $x_1, \ldots, x_{n_1}$ and $y_1, \ldots, y_{n_2}$ may

[2]) We use the term "distribution function" in the sense sometimes denoted by the term "cumulative distribution function".

be regarded as independent random samples from a common population. This may be of importance, if additional information about the common probability distribution of the $\mathbf{z}_i$ is available, or is contained in the hypothesis to be tested, since we may then use any two sample test based on this information.

2. Theorems I and III enable us to give a test for $H_0$, based on the statistics $\mathbf{n}_1$ and $\mathbf{u}_1, \ldots, \mathbf{u}_\nu$. Since a family of tests $T$ may be used (cf. section 6. 2), we have a family of tests for $H_0$. The exact test, described in part I of this paper, is a member of this family as may be seen from remark 3 of section 4. 2. $T$ is then a two sample test based on the statistic $\mathbf{u}$.

8. *The critical region.*

8. 1. In section 7. 1 we have supposed the conditional probability distribution of $\mathbf{u}_1, \ldots, \mathbf{u}_\nu$, under the conditions $Z$, $\mathbf{n}_1 = n_1$ and hypothesis $H_0$, to be known, since $T$ is a known two sample test. For the same reason we now assume a critical region for $\mathbf{u}_1, \ldots, \mathbf{u}_\nu$ to have been chosen already. We shall, however, want to make a distinction between *bilateral* and *unilateral* critical regions. To make this clear, the critical regions of some of the two-samples tests mentioned in 6. 3 will be described.

8. 2. WILCOXON's test depends on the number of pairs $(x_j, y_k)$ $(j = 1, \ldots, n_1;\ k = 1, \ldots, n_2)$ with $x_j > y_k$. This statistic, usually denoted by $\mathbf{U}$, can take all values $0, 1, \ldots, n_1 + n_2$. A unilateral critical region has either the form

$$\mathbf{U} - \frac{n_1 n_2}{2} \geqq U_a$$

or

$$\mathbf{U} - \frac{n_1 n_2}{2} \leqq - U_a$$

where $U_a$ depends on $n_1$, $n_2$ and the chosen significance level $a$. The first of these critical regions is suitable for testing the hypothesis $H'$, that $x_1, \ldots, x_{n_1}$ and $y_1, \ldots, y_{n_2}$ are random samples taken independently from the same population, against the alternative (composite) hypothesis, that $x_1, \ldots x_{n_1}$ are independent observations of a random variable $\mathbf{x}$ and $y_1, \ldots, y_{n_2}$ of a random variable $\mathbf{y}$, with

$$\mathrm{P}\,[\mathbf{x} < \mathbf{y}] < \tfrac{1}{2}$$

and the second critical region is suitable for testing $H'$ against the alternative hypothesis, that

$$\mathrm{P}\,[\mathbf{x} < \mathbf{y}] > \tfrac{1}{2} \quad ^{3)}$$

---

[3]) This has been proved recently by Prof. D. VAN DANTZIG as a generalisation of MANN and WHITNEY's theorem, according to which WILCOXON's test is consistent against alternatives with $P\,[\mathbf{x} \leqq a] < P\,[\mathbf{y} \leqq a]$ for all $a$, if the first of the above mentioned unilateral critical regions is used, and consistent against alternatives with $P\,[\mathbf{x} \leqq a] > P\,[\mathbf{y} \leqq a]$ if the second critical region is used. Cf. D. VAN DANTZIG (1947—1950), Chapter 6, § 3.

A symmetrical bilateral critical region for $\boldsymbol{U}$ has the form

$$\left|\boldsymbol{U}-\frac{n_1 n_2}{2}\right| \geqq U_{\frac{1}{2}\alpha}$$

and is suitable for testing $H'$ against the alternative hypothesis, that

$$\mathrm{P}\,[\mathbf{x} < \mathbf{y}] \neq \tfrac{1}{2}.\ ^{3)}$$

The probability distribution of $\boldsymbol{U}$ can be computed exactly with the aid of a recursion formula given by MANN and WHITNEY, under the assumption, that $H'$ is true and that $\mathbf{x}$ and $\mathbf{y}$ have a continuous probability distribution. It has been tabulated by them for $n_1 \leqq 8$ and $n_2 \leqq 8$ and for larger values the normal distribution with mean $\frac{n_1 n_2}{2}$ and variance $\frac{1}{12} n_1 n_2 (n_1 + n_2 + 1)$ (which is the asymptotic distribution of $\boldsymbol{U}$ for $n_1 \to \infty$ and $n_2 \to \infty$, $n_1/n_2$ and $n_2/n_1$ being bounded) is a good approximation.

8. 3. The statistic of PITMAN's test, which we shall also denote by $\boldsymbol{U}$, is the difference of the means of $\mathbf{x}_1, \ldots, \mathbf{x}_{n_1}$ and $\mathbf{y}_1, \ldots, \mathbf{y}_{n_2}$:

$$\boldsymbol{U} = \frac{1}{n_1} \sum_{j=1}^{n_1} \mathbf{x}_j - \frac{1}{n_2} \sum_{k=1}^{n_2} \mathbf{y}_k.$$

The unilateral critical regions

$$\boldsymbol{U} \leqq -U'_\alpha$$

and

$$\boldsymbol{U} \geqq U'_\alpha$$

where $U'_\alpha$ depends on the observed values $x_1, \ldots, x_{n_1}, y_1, \ldots, y_{n_2}$ and the chosen significance level $\alpha$, are suitable for testing $H'$ against the alternative hypotheses

$$\mathcal{E}\mathbf{x} < \mathcal{E}\mathbf{y}$$

and

$$\mathcal{E}\mathbf{x} > \mathcal{E}\mathbf{y}$$

respectively.

A bilateral critical region

$$|\boldsymbol{U}| \geqq U'_{\frac{1}{2}\alpha}$$

is suitable for testing $H'$ against the alternative hypothesis

$$\mathcal{E}\mathbf{x} \neq \mathcal{E}\mathbf{y}$$

The probability distribution of $\boldsymbol{U}$ can be derived exactly from the values $x_1, \ldots, x_{n_1}, y_1, \ldots, y_{n_2}$. This, however, is only practicable, if $n_1$ and $n_2$ are very small. For larger values of $n_1$ and $n_2$ PITMAN has given an approximation for the distribution of $\boldsymbol{U}$. The assumption of continuity is not necessary.

8. 4. The test of WALD and WOLFOWITZ is based on the number of runs in the sequence of values $x_j$ and $y_k$ $(j = 1, \ldots, n_1;\ k = 1, \ldots, n_2)$

when arranged according to decreasing magnitude. Small values of this statistic are critical. Its probability distribution is known exactly as well as asymptotically, under the assumption that $H'$ is true and that the probability distribution of $\boldsymbol{x}$ and $\boldsymbol{y}$ is continuous. F. S. SWED and C. EISENHART (1943) have given tables of this distribution for $n_1 \leqq n_2 \leqq 20$.

For this test we shall not try to make a distinction between unilateral and bilateral critical regions, since the class of alternative hypotheses, for which the test is consistent contains nearly all possible alternative hypotheses. It is difficult to see how this class could be divided into two mutually exclusive classes of a kind similar to the two classes of alternatives for Wilcoxon's test or Pitman's test, which have been described in 8. 2 and 8. 3. As far as our application of the test of Wald and Wolfowitz is concerned, its critical region can therefore be taken to be a *bilateral* one.

8. 5. The probability distribution of the statistic of the test of KOLMOGOROFF–SMIRNOFF is known asymptotically only. An exact method for determining the confidence limits for an unknown distribution function (the problem, which had been solved asymptotically by A. KOLMOGOROFF (1933)) has been given by A. WALD and J. WOLFOWITZ (1939). Possibly the method applied by Smirnoff to derive a two sample test from Kolmogoroff's theorem might be applied to this theorem of Wald and Wolfowitz and give an exact two sample test of this type.

So far, however, we have no knowledge either of the exact probability distribution of the statistic of this test, nor of the amount of the divergence between this exact distribution and the asymptotical one, derived by Smirnoff. Apart from this the remarks, made above about the critical region of the test of Wald and Wolfowitz, apply to this test also. No attempt will be made to make a distinction between unilateral and bilateral critical regions. The only difference is, that in this case *large* values of this statistic are critical, and that no continuity of the probability distribution of $\boldsymbol{x}$ and $\boldsymbol{y}$ is needed.

8. 5. We shall now consider the choice of a *critical region* for testing $H_0$, *if no alternative hypothesis is specified*. In order to simplify the notation, we confine ourselves to $\nu = 1$, i.e. to the case, that the two sample test $T$ is based on one statistic $\boldsymbol{U}$. Denoting the bilateral critical region for $T$ with size $\varepsilon$ by $R_{n-m,\, n_1}(\varepsilon)$, we propose the following construction of a critical region $R_1^*$ with size $\leqq a$ for testing $H_0$, applicable if $a \geqq 2^{-n+m+1}$:

*A*. Let $k$ be the largest positive integer $\leqq \frac{n-m}{2}$, for which the relation

$$2^{-n+m}\binom{n-m}{k} \leqq \frac{a}{n-m+1} \tag{14}$$

holds (where $m$ is the value of $\boldsymbol{m}$ following from the observations $z_1, \ldots, z_n$) or, if no positive integer satisfies (14), $k = 0$.

*B.* Put

(15) $$\beta = \beta(n-m, \alpha) = 2^{-n+m+1} \sum_{i=0}^{k} \binom{n-m}{i}$$

and

(16) $$\varepsilon = \varepsilon(n-m, n_1, \alpha) = \frac{\alpha-\beta}{n-m-2k-1} \cdot \frac{2^{n-m}}{\binom{n-m}{n_1}}$$

*C.* Then the critical region $R_1^*$ consists of those points $(n_1, U)$, for which at least one of the following two conditions holds:

$$C_1: \quad n_1 \leqq k \text{ or } n_1 \geqq n-m-k$$
$$C_2: \quad U \in R_{n-m,n_1}(\delta)$$

where $\delta \leqq \varepsilon$, and $\varepsilon - \delta$ is as small as possible. [It is clear, that the size of $R_1^*$ is then $\leqq \alpha$.

8.6. For $n-m = 12$ and $\alpha = 0{,}10$ ($\alpha$ has been chosen rather large to get better diagrams) $R_1^*$ has been outlined in fig. 3 for the case that

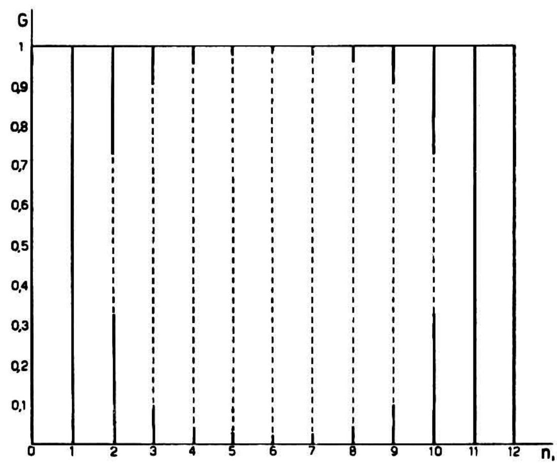


Fig. 3. Critical region $R_1^*$, when no alternative hypothesis is specified and when $T$ is WILCOXON's test or PITMAN's test; $\alpha = 0{,}10$.

Wilcoxon's test or Pitman's test is used for $T$ and in fig. 4 if the test of Wald and Wolfowitz is used. In these figures $G(U|Z;\ \mathbf{n}_1 = n_1;\ H_0)$, the conditional distribution function of $\boldsymbol{U}$, has been plotted on vertical lines above the points $n_1 = 1, \ldots, n_1 = 11$. $R_1^*$ consists of the points $(n_1, U)$ on those parts of these lines, which have been *drawn*. The points with $n_1 = 0$ and $n_1 = 12$ belong to the critical region according to 8.5.*A*. This has been indicated by drawing the vertical lines above these points. The *broken* parts of the vertical lines indicate the region, where $H_0$ is not rejected. The reader should bear in mind, that in reality $G$ is discontinuous.

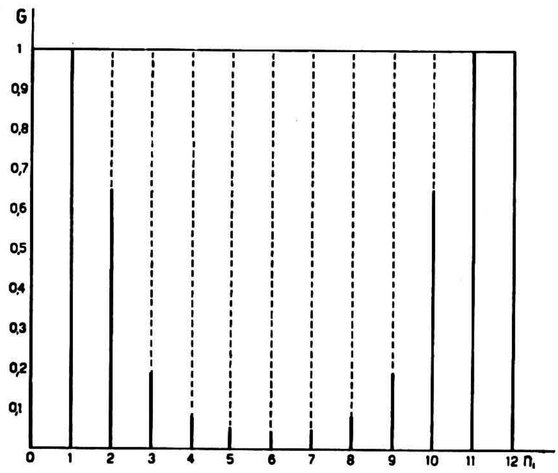


Fig. 4. Critical region $R_1^*$, when no alternative hypothesis is specified and $T$ is the test of WALD and WOLFOWITZ; $\alpha = 0{,}10$.

**Remark:** The critical region $R_1^*$ for the case, that the test of Kolmogoroff–Smirnoff is used for $T$, can be constructed in an analogous way, large values of $G$ being critical instead of small values. Strictly speaking, however, we do not know much about $\alpha$ in that case.

8. 7. As a *special alternative hypothesis*, against which $H_0$ can be tested, we consider the hypothesis $H$ (cf. section 5. 3) of a *displacement* of one or more of the variables $\mathbf{z}_i$ in one direction along the $z$-axis. In this case we restrict the "family" of tests $T$ to those tests, where a *unilateral* critical region can be indicated (cf. 8. 2, 8. 3 and 8. 4). We shall denote unilateral critical regions of the types $\mathbf{U} \leqq U_1$ and $\mathbf{U} \geqq U_1'$ with size $\varepsilon$ by $R'_{n-m,n_1}(\varepsilon)$ and $R''_{n-m,n_1}(\varepsilon)$ respectively. These critical regions may also be defined by the relations

$$G(\mathbf{U}|Z;\mathbf{n}_1 = n_1; H_0) \leqq G_1 = G(U_1|Z;\mathbf{n}_1 = n_1; H_0)$$

and

$$G(\mathbf{U}|Z;\mathbf{n}_1 = n_1; H_0) \geqq G_1' = G(U_1'|Z;\mathbf{n}_1 = n_1; H_0).$$

For reasons given in section 5. 3 we exclude, if $n - m$ is even, the points $(n_1, U)$ with $n_1 = \frac{n-m}{2}$ from the critical region $R_2^*$ for testing $H_0$ against $H$. We further remark, that for Wilcoxon's test and Pitman's test the probability of small (large) values of $\mathbf{U}$ increases if some of the variables $\mathbf{z}_i$ are shifted towards the left (right) and decreases, if the displacement is towards the right (left) along the $z$-axis (cf. section 5. 3). We therefore propose for these cases the following construction for $R_2^*$ (with size $\leqq \alpha$), applicable if $\alpha \geqq 2^{-n+m+1}$:

*A*. Let $k$ be the largest positive integer $< \frac{n-m}{2}$, for which the relation (14) holds, or, if no positive integer satisfies (14), $k = 0$.

*B'*. Define $\beta$ by (15) and (cf. (16)):

$$\varepsilon' = \varepsilon' (n-m, n_1, a) = \varepsilon (n-m, n_1, a) \tag{16'}$$

if $n - m$ is odd, and

$$\varepsilon' = \varepsilon' (n-m, n_1, a) = \frac{a-\beta}{n-m-2k-2} \frac{2^{n-m}}{\binom{n-m}{n_1}} \tag{16''}$$

if $n - m$ is even.

*C'*. Then $R_2^*$ consists of those points $(n_1, U)$, for which at least one of the following three conditions holds:

$$C_1: \quad n_1 \leqq k \text{ or } n_1 \geqq n-m-k$$

$$C_2': \quad n_1 < \frac{n-m}{2} \text{ and } U \in R'_{n-m,n_1}(\delta)$$

$$C_3': \quad n_1 > \frac{n-m}{2} \text{ and } U \in R''_{n-m,n_1}(\delta)$$

where $\delta \leqq \varepsilon$, and $\varepsilon - \delta$ is as small as possible.

For $n - m = 12$ and $a = 0{,}10$ $R_2^*$ has been given in figure 5 for the case that Wilcoxon's test or Pitman's test has been used for $T$.

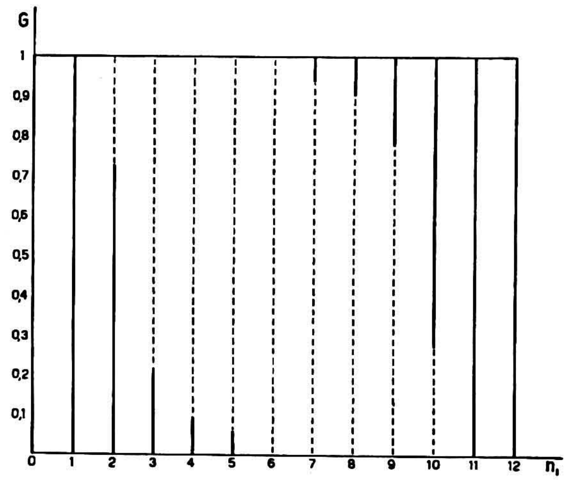


Fig. 5. Critical region $R_2^*$, when the alternative is a displacement of at least one of the distributions in one direction along the $z$-axis; $a = 0{,}10$.

If the direction of the displacement is specified in the alternative hypothesis, the critical region may be confined either to the left or to the right half of the diagram only, using $2a$ instead of $a$ in (14) and (15).

8. 8. The computations are now comparatively simple. A table of $k$, $2^{n-m}$ and of the quantities

$$\gamma = \frac{a-\beta}{n-m-2k-1} \cdot 2^{n-m} \tag{17}$$

and, for even values of $n - m$, of

$$\gamma' = \frac{a-\beta}{n-m-2k-2} \cdot 2^{n-m} \tag{18}$$

has been computed by the Computing Department of the "Mathematisch Centrum" for $a = 0{,}025$; $0{,}05$ and $0{,}10$ and for $n - m = 10(1)50$ (cf. section 10). From this table the value of $\varepsilon(n - m, n_1, a)$ or $\varepsilon'(n - m, n_1, a)$ is easily computed with the aid of a table of the binomial coefficients (cf. 5. 2). If then condition $C$ (or $C'$) is satisfied, the result is significant with significance level $\leqq a$.

Moreover, if $n_1 \neq 0$ and $\neq n - m$, an upper bound for the size of the smallest critical region of type $R_1^*$ or $R_2^*$, which contains the point $(n_1, U)$ following from the observations, may be found as follows:

Let $\eta$ be the size of the smallest critical region for $\boldsymbol{U}$, given $n - m$ and $n_1$ (either bilateral or unilateral), which contains the observed value $U$, then

$$a^* = 2^{-n+m}(n-m+1)\binom{n-m}{n_1} \cdot \eta \geqq a. \tag{19}$$

8. 9. Sections 8. 5 and 8. 6 may be applied to the special case, described in sections 4 and 5. According to remark 3 of section 4. 2, $\boldsymbol{u}$ has, if $H_0$ is true, for given $n_1$ and under the condition $Z$, a hypergeometric distribution. For this distribution we have

$$\mathcal{E}\,(\boldsymbol{u}\,|\,Z;\,\boldsymbol{n}_1 = n_1;\,H_0) = \frac{rn_1}{n_1+n_2} \tag{20}$$

and

$$\sigma^2_{\boldsymbol{u}\,|\,Z;\,\boldsymbol{n}_1=n_1;\,H_0} = \frac{r\,n_1\,n_2\,(n_1+n_2-r)}{(n_1+n_2)^2\,(n_1+n_2-1)} \tag{21}$$

with $n_1 + n_2 = n - m$. A normal probability distribution with (20) and (21) as mean and variance respectively is a good approximation of this probability distribution of $\boldsymbol{u}$, especially if a continuity correction is applied.

If $n - m$ is so large, that the exact method of section 5 becomes too laborious, this approximate method may be used instead. It should, however, be born in mind, that the construction of the critical regions $R_1^*$ and $R_2^*$ is different from the construction of $R_1$ and $R_2$, and that $R_1^*$ and $R_2^*$ should therefore not be regarded as approximations of $R_1$ and $R_2$, but as an approximate method using a slightly different form of critical regions.

On the other hand, if the number of observations is small and $T$ is a test, such that the exact distribution of $\boldsymbol{U}$ is known, the critical region may, for the general case, be chosen according to a system analogous to the method described for the special case in section 5. We shall not go into the details of this method for other special cases, since the principle remains unchanged for every $T$.

9. *Remarks*.

Of the two-sample tests, mentioned in 6. 2, the tests of Wilcoxon and of Wald and Wolfowitz can only be applied to our problem if $x_1, \ldots, x_{n_1}$ and $y_1, \ldots, y_{n_2}$ are all different. This is not required for the tests of Pitman and Kolmogoroff–Smirnoff. On the other hand, the latter is not an exact test, as has been pointed out already in section 8. 5. Since small values of $n_1$ or $n_2$ will often occur in the application of the test, this is a serious drawback. The same applies, to a certain extent, to Pitman's test, since the computation of the exact distribution of its statistic is impracticable for values of $n_1$ and $n_2$, which are at all large. Furthermore little is known about the accuracy of the approximation to the distribution of $\mathbf{U}$, given by Pitman, especially in the case of discontinuous random variables.

So far the only exact test for $H_0$, developed untill now, which is valid if there are equal values among $x_1, \ldots, x_{n_1}, y_1, \ldots, y_{n_2}$, and which can be used for reasonably large values of $n_1 + n_2$, seems to be the one described in sections 4 and 5. Moreover for large values of $n_1 + n_2$ the accuracy of the approximate method, described in 8. 9 is independent of the number of equal values among $x_1, \ldots, x_{n_1}$ and $y_1, \ldots, y_{n_2}$.

If no equal values occur among the $x_j$ and $y_k$, Wilcoxon's test seems a very suitable one for $T$, especially when the alternative hypothesis is the hypothesis $H$ of a displacement along the $z$-axis.

The number of values $z_i$, which are equal to zero, is of no consequence whatever as far as the choice of $T$ is concerned.

11. *Explanation of the table and of the practical application of the test.*

The use of the table in applying the test may be facilitated by the following indications:

$n$ denotes the number of observations $z_1, \ldots, z_n$ and $m$ the number of values $z_i$ which are equal to zero;

$n_1$ denotes the number of positive values $x_1, \ldots, x_{n_1}$ among $z_1, \ldots, z_n$.

If $n_1 \leqq k$ or $n_1 \geqq n - m - k$, $H_0$ is rejected with significance level $\leqq \alpha$. If $k < n_1 < n - m - k$, two cases are to be distinguished:

I. If no alternative hypothesis to $H_0$ is specified, the chosen two sample test $T$ is applied to the two sets of values $x_1, \ldots, x_{n_1}$ and $y_1, \ldots, y_{n_2}$ (the $y_k$ are the negative values among $z_1, \ldots, z_n$ taken positively). This results in a value $U$ of the statistic $\mathbf{U}$ of $T$. Let $\eta$ be the size of the smallest bilateral critical region for $\mathbf{U}$, belonging to $T$, which contains $U$. If then

$$\eta \leqq \gamma \binom{n-m}{n_1}$$

$H_0$ is rejected with significance level $\leqq \alpha$; cf. (17) for $\gamma$.

In case Wilcoxon's test is used for $T$, we have

$$\eta = 2G(U \mid Z; \mathbf{n}_1 = n_1; H_0)$$

if $U < \frac{n_1 n_2}{2}$ and

$$\eta = 2\{1 - G(U \mid Z; \mathbf{n}_1 = n_1; H_0)\}$$

if $U > \frac{n_1 n_2}{2}$.

10. Table of $k$, $2^{n-m}$, $\gamma$ and $\gamma'$.

$\alpha = 0,025;\ 0,05;\ 0,10$

| $n-m$ | $2^{n-m}$ | $\alpha = 0,025$ | | | $\alpha = 0,05$ | | | $\alpha = 0,10$ | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | $k$ | $\gamma$ | $\gamma'$ | $k$ | $\gamma$ | $\gamma'$ | $k$ | $\gamma$ | $\gamma'$ |
| 10 | $1024.10^{0}$ | 0 | $2,622.10^{0}$ | $2,950.10^{0}$ | 0 | $5,467.10^{0}$ | $6,150.10^{0}$ | 0 | $1,116.10^{1}$ | $1,255.10^{1}$ |
| 11 | $2048.10^{0}$ | 0 | $4,919.10^{0}$ | | 0 | $1,004.10^{1}$ | | 1 | $2,260.10^{1}$ | |
| 12 | $4096.10^{0}$ | 0 | $9,126.10^{0}$ | $1,004.10^{1}$ | 1 | $1,987.10^{1}$ | $2,235.10^{1}$ | 1 | $4,262.10^{1}$ | $4,795.10^{1}$ |
| 13 | $8192.10^{0}$ | 1 | $1,768.10^{1}$ | | 1 | $3,816.10^{1}$ | | 1 | $7,912.10^{1}$ | |
| 14 | $1638.10^{1}$ | 1 | $3,450.10^{1}$ | $3,796.10^{1}$ | 1 | $7,175.10^{1}$ | $7,892.10^{1}$ | 2 | $1,585.10^{2}$ | $1,783.10^{2}$ |
| 15 | $3277.10^{1}$ | 1 | $6,560.10^{1}$ | | 1 | $1,339.10^{2}$ | | 2 | $3,035.10^{2}$ | |
| 16 | $6554.10^{1}$ | 1 | $1,234.10^{2}$ | $1,337.10^{2}$ | 2 | $2,730.10^{2}$ | $3,003.10^{2}$ | 2 | $5,709.10^{2}$ | $6,280.10^{2}$ |
| 17 | $1311.10^{2}$ | 2 | $2,475.10^{2}$ | | 2 | $5,205.10^{2}$ | | 3 | $1,144.10^{3}$ | |
| 18 | $2621.10^{2}$ | 2 | $4,776.10^{2}$ | $5,175.10^{2}$ | 2 | $9,817.10^{2}$ | $1,064.10^{3}$ | 3 | $2,203.10^{3}$ | $2,424.10^{3}$ |
| 19 | $5243.10^{2}$ | 2 | $9,091.10^{2}$ | | 3 | $1,991.10^{3}$ | | 3 | $4,175.10^{3}$ | |
| 20 | $1049.10^{3}$ | 3 | $1,809.10^{3}$ | $1,959.10^{3}$ | 3 | $3,825.10^{3}$ | $4,144.10^{3}$ | 4 | $8,405.10^{3}$ | $9,246.10^{3}$ |
| 21 | $2097.10^{3}$ | 3 | $3,521.10^{3}$ | | 3 | $7,267.10^{3}$ | | 4 | $1,622.10^{4}$ | |
| 22 | $4194.10^{3}$ | 3 | $6,749.10^{3}$ | $7,231.10^{3}$ | 4 | $1,473.10^{4}$ | $1,596.10^{4}$ | 4 | $3,086.10^{4}$ | $3,344.10^{4}$ |
| 23 | $8389.10^{3}$ | 3 | $1,285.10^{4}$ | | 4 | $2,840.10^{4}$ | | 5 | $6,248.10^{4}$ | |
| 24 | $1678.10^{4}$ | 4 | $2,624.10^{4}$ | $2,812.10^{4}$ | 4 | $5,421.10^{4}$ | $5,807.10^{4}$ | 5 | $1,205.10^{5}$ | $1,306.10^{5}$ |
| 25 | $3355.10^{4}$ | 4 | $5,053.10^{4}$ | | 5 | $1,101.10^{5}$ | | 5 | $2,299.10^{5}$ | |
| 26 | $6711.10^{4}$ | 4 | $9,657.10^{4}$ | $1,026.10^{5}$ | 5 | $2,125.10^{5}$ | $2,278.10^{5}$ | 6 | $4,679.10^{5}$ | $5,069.10^{5}$ |
| 27 | $1342.10^{5}$ | 5 | $1,970.10^{5}$ | | 5 | $4,068.10^{5}$ | | 6 | $9,019.10^{5}$ | |
| 28 | $2684.10^{5}$ | 5 | $3,804.10^{5}$ | $4,043.10^{5}$ | 6 | $8,281.10^{5}$ | $8,874.10^{5}$ | 6 | $1,723.10^{6}$ | $1,845.10^{6}$ |
| 29 | $5369.10^{5}$ | 5 | $7,291.10^{5}$ | | 6 | $1.600.10^{6}$ | | 7 | $3,523.10^{6}$ | |
| 30 | $1074.10^{6}$ | 6 | $1,488.10^{6}$ | $1,582.10^{6}$ | 6 | $3,068.10^{6}$ | $3,260.10^{6}$ | 7 | $6,785.10^{6}$ | $7,269.10^{6}$ |
| 31 | $2148.10^{6}$ | 6 | $2,878.10^{6}$ | | 7 | $6,264.10^{6}$ | | 7 | $1,298.10^{7}$ | |
| 32 | $4295.10^{6}$ | 6 | $5,528.10^{6}$ | $5,837.10^{6}$ | 7 | $1,210.10^{7}$ | $1,286.10^{7}$ | 8 | $2,663.10^{7}$ | $2,853.10^{7}$ |
| 33 | $8590.10^{6}$ | 7 | $1,130.10^{7}$ | | 7 | $2,323.10^{7}$ | | 8 | $5,125.10^{7}$ | |
| 34 | $1718.10^{7}$ | 7 | $2,187.10^{7}$ | $2,307.10^{7}$ | 8 | $4,755.10^{7}$ | $5,053.10^{7}$ | 8 | $9,808.10^{7}$ | $1,042.10^{8}$ |
| 35 | $3536.10^{7}$ | 8 | $4,412.10^{7}$ | | 8 | $9,184.10^{7}$ | | 9 | $2,019.10^{8}$ | |
| 36 | $6872.10^{7}$ | 8 | $8,611.10^{7}$ | $9,092.10^{7}$ | 8 | $1,765.10^{8}$ | $1,864.10^{8}$ | 9 | $3,883.10^{8}$ | $4,126.10^{8}$ |
| 37 | $1374.10^{8}$ | 8 | $1,667.10^{8}$ | | 9 | $3,623.10^{8}$ | | 10 | $7,934.10^{8}$ | |
| 38 | $2749.10^{8}$ | 9 | $3,376.10^{8}$ | $3,565.10^{8}$ | 9 | $6,993.10^{8}$ | $7,383.10^{8}$ | 10 | $1,534.10^{9}$ | $1,630.10^{9}$ |
| 39 | $5498.10^{8}$ | 9 | $6,581.10^{8}$ | | 10 | $1,424.10^{9}$ | | 10 | $2,951,10^{9}$ | |
| 40 | $1100.10^{9}$ | 9 | $1,273.10^{9}$ | $1,337.10^{9}$ | 10 | $2,765.10^{9}$ | $2,918.10^{9}$ | 11 | $6,052.10^{9}$ | $6,430.10^{9}$ |
| 41 | $2199.10^{9}$ | 10 | $2,590.10^{9}$ | | 10 | $5.339.10^{9}$ | | 11 | $1,169.10^{10}$ | |
| 42 | $4398.10^{9}$ | 10 | $5,040.10^{9}$ | $5,291.10^{9}$ | 11 | $1,090.10^{10}$ | $1,151.10^{10}$ | 11 | $2,248.10^{10}$ | $2,373.10^{10}$ |
| 43 | $8796.10^{9}$ | 10 | $9,755.10^{9}$ | | 11 | $2,115.10^{10}$ | | 12 | $4,623.10^{10}$ | |
| 44 | $1759.10^{10}$ | 11 | $1.988.10^{10}$ | $2,088.10^{10}$ | 11 | $4,083.10^{10}$ | $4,287.10^{10}$ | 12 | $8,920.10^{10}$ | $9,415.10^{10}$ |
| 45 | $3518.10^{10}$ | 11 | $3,867.10^{10}$ | | 12 | $8,363.10^{10}$ | | 13 | $1,825.10^{11}$ | |
| 46 | $7037.10^{10}$ | 11 | $7,487.10^{10}$ | $7,825.10^{10}$ | 12 | $1,621.10^{11}$ | $1,702.10^{11}$ | 13 | $3,536.10^{11}$ | $3,732.10^{11}$ |
| 47 | $1407.10^{11}$ | 12 | $1,530.10^{11}$ | | 13 | $3,302.10^{11}$ | | 13 | $6,820.10^{11}$ | |
| 48 | $2815.10^{11}$ | 12 | $2,972.10^{11}$ | $3,107.10^{11}$ | 13 | $6.420.10^{11}$ | $6,744.10^{11}$ | 14 | $1,400.10^{12}$ | $1,477.10^{12}$ |
| 49 | $5629.10^{11}$ | 13 | $6,040.10^{11}$ | | 13 | $1,244.10^{12}$ | | 14 | $2,708.10^{12}$ | |
| 50 | $1126.10^{12}$ | 13 | $1,177.10^{12}$ | $1,232.10^{12}$ | 14 | $2,541.10^{12}$ | $2,668.10^{12}$ | 14 | $5,222.10^{12}$ | $5,483.10^{12}$ |

According to continental usage the comma designates the decimal sign (e.g. $0,5 = \frac{1}{2}$).

II. If $H_0$ is tested against the alternative hypothesis $H$ of a displacement along the $z$-axis in one direction (which one not being specified) of some of the variables $\mathbf{z}_i$, a two sample test $T$ is applied to $x_1, \ldots, x_{n_1}$ and $y_1, \ldots y_{n_2}$, which allows a distinction between two unilateral critical regions. Taking e.g. Wilcoxon's test, the size $\eta'$ of the smallest unilateral critical region containing the value $U$, found from the observations, is computed, using the unilateral critical region of the form

$$\mathbf{U} - \frac{n_1 n_2}{2} \geqq U_a$$

if $n_1 > \frac{n-m}{2}$, and the unilateral critical region of the type

$$\mathbf{U} - \frac{n_1 n_2}{2} \leqq - U_a$$

if $n_1 < \frac{n-m}{2}$ (cf. 8. 2).

In the first case, we have

$$\eta' = 1 - G(U | Z; \mathbf{n}_1 = n_1; H_0)$$

and in the second case

$$\eta' = G(U | Z; \mathbf{n}_1 = n_1; H_0).$$

If $n - m$ is odd, and

$$\eta' \leqq \gamma \binom{n-m}{n_1}$$

or if $n - m$ is even, and

$$\eta' \leqq \gamma' \binom{n-m}{n_1}$$

$H_0$ is rejected; cf. (17) and (18) for $\gamma$ and $\gamma'$.

The values of $2^{n-m}$ have been included in the table to facilitate the computation of the size $a^*$ of smallest critical region (either of unilateral or bilateral type), which contains the point $(n_1, U)$, following from the observations. This computation has been described in 8. 8.

12. *Example.*

Let us consider a set of observed values $z_1, \ldots, z_{22}$:

$-8{,}0$; $-5{,}0$; $-4{,}5$; $-3{,}0$; $-2{,}7$; $-2{,}3$; $-2{,}1$; $-1{,}3$; $-1{,}2$; $-1{,}0$; $-0{,}9$; $-0{,}5$; $-0{,}2$; 0; 0; 1,8; 2,5; 3,5; 6,2; 7,3; 7,4; 9,5.

We then have $n = 22$, $m = 2$, $n_1 = 7$. From the table of section 11 we find (for $a = 0{,}05$)

$$\varepsilon = \gamma \Big/ \binom{20}{7} = \frac{3825}{77520} = 0{,}049$$

and $k = 3$. Therefore $k < n_1 < n - m - k$ and a two sample test must be applied. Let us take Wilcoxon's test for this. The number of pairs $(x_j, y_k)$ with $x_j > y_k$ is 73. According to section 8. 2 we have

$$\mathcal{E}\,\mathbf{U} = \frac{7.13}{2} = 45{,}5$$

and

$$\sigma_u = \sqrt{\tfrac{1}{12}\, 7.13(7 + 13 + 1)} = 12{,}62.$$

Applying a correction for continuity, we find

$$\frac{U-\frac{1}{2}-\mathcal{E}\boldsymbol{U}}{\sigma_u}=\frac{72{,}5-45{,}5}{12{,}62}=2{,}14.$$

From a table of the normal distribution we find therefore, that

$$\eta=\mathrm{P}\,[\,|\boldsymbol{U}-\mathcal{E}\boldsymbol{U}|\geqq|73-45{,}5|\,|Z\,\boldsymbol{n}_1=n_1;H_0]=0{,}032.$$

Since $\eta<\varepsilon$, $H_0$ is rejected with significance level 0,05.

If, however, $H_0$ is tested against the alternative hypothesis $H$ of a displacement of some of the variables $\boldsymbol{z}_i$ in one direction along the $z$-axis, $H_0$ is *not* rejected, since

$$n_1<\frac{n-m}{2}=\mathcal{E}\,(\boldsymbol{n}_1|H_0)$$

and

$$U>\frac{n_1 n_2}{2}=\mathcal{E}\,(\boldsymbol{U}|\boldsymbol{n}_1=n_1;H_0)$$

thus $G(U|Z;\boldsymbol{n}_1=n_1;H_0)$ having the value $1-0{,}016=0{,}984$. The point $(n_1, U)$ corresponding with this result is not contained in the critical region $R_2^*$ (cf. figure 5). This means, that the observations do not indicate a displacement of some of the $\boldsymbol{z}_i$ in one direction along the $z$-axis. They do, however, suggest displacements in both directions, or asymmetry of some of the distributions or a combination of displacements and asymmetry. This follows from the fact, that $H_0$ is rejected if no special alternative hypothesis is specified.

My thanks are due to Prof. D. van Dantzig for his remarks and criticism, which have been very helpful, and for reading the manuscript.

*Publication of the Statistical Department of the "Mathematisch Centrum", Amsterdam.*

REFERENCES

Dantzig, D. van, Kadercursus mathematische statistiek, Mathematisch Centrum, Amsterdam (1947–1950).

Kolmogoroff, A., Sulla determinatione empirica di una legge di distribuzione, Giornale dell'Inst. Ital. degli Attuari **4**, 83–91 (1933).

———, Confidence limits for an unknown distribution function, Ann. Math. Stat. **12**, 461–463 (1941).

Mann, H. B. and D. R. Whitney, On a test of whether one of two random variables is stochastically larger than the other, Ann. Math. Stat. **18**, 50–60 (1947).

Pitman, E. J., Significance tests which may be applied to samples from any populations, I., Jrn. Roy. Stat. Soc. Suppl. **4**, 119–129 (1937).

Smirnoff, N., Sur les écarts de la courbe de distribution empirique, Receuil Math. de Moscou **6**, 3–26 (1939).

———, On the estimation of the discrepancy between empirical curves of distribution for two independent samples, Bull. de l'Univ. de Moscou, Série intern. (Mathématiques) **2**, fasc. 2 (1939).

Swed, F. S. and C. Eisenhart, Tables for testing randomness of grouping in a sequence of alternatives, Ann. Math. Stat. **14**, 66–87 (1943).

Wald, A. and J. Wolfowitz, Confidence intervals for continuous distribution functions, Ann.-Math. Stat. **10**, 105–118 (1939).

Wilcoxon, F., Individual comparisons by ranking methods, Biometrics **1**, 80–82 (1945).