STICHTING
MATHEMATISCH CENTRUM
2e BOERHAAVESTRAAT 49
AMSTERDAM

ZW-001

(S 157 (VP 3))

A test for the equality of probabilities against a class of specified alternative hypotheses, including trend

Constance van Eeden

1955

MATHEMATICAL CENTRE,
2e Boerhaavestraat 49,
Amsterdam - O.

Statistical Department

Report S 157(VP 3)

Errata

page	line			
4	11 f.b.[1]	N-dimensional	shoud read:	(N-1)dimensional
5	14 f.t.	$\frac{t_{1\nu}}{N} = O(1)$, $\frac{t_{2\nu}}{N} = O(1)$	" :	$\frac{N_\nu}{t_{1\nu}} = O(1)$, $\frac{N_\nu}{t_{2\nu}} = O(1)$
9	5 f.t.	necessary	" :	necessarily
9	3 f.b.	$\frac{t_1}{N} = O(1)$, $\frac{t_2}{N} = O(1)$	" :	$\frac{N}{t_1} = O(1)$, $\frac{N}{t_2} = O(1)$
11	6 f.b.			
14	2 f.t.	hypothese	" :	hypothesis

1) f.b.= from below.
f.t.= from the top.

MATHEMATICAL CENTRE
2e Boerhaavestraat 49,
A M S T E R D A M.

STATISTICAL DEPARTMENT

Head of Department: Prof.Dr.D.van Dantzig
Chief of Statistical Consultation: Prof.Dr.J.Hemelrijk

Report S 157(VP3)

A test for the equality of probabilities against a class of specified alternative hypotheses, including trend.

by

Constance van Eeden

Lecture in the series „Actualiteiten."

on January 29. 1955.

1954.

1. Introduction.

We consider k ($k \geq 2$) independent series of independent trials, each trial resulting in a success or a failure. The i-th series consists of n_i trials with $\underline{a}_i$ [1] successes and $\underline{b}_i$ failures [2]; $\underline{t}_1 = \sum_i \underline{a}_i$, $\underline{t}_2 = \sum_i \underline{b}_i$, $N = \sum_i n_i$ and p_i is the probability of a success for each trial of the i-th series.

The observations may be summarized in the following table.

Series	Number of		Total
	successes	failures	
1	$\underline{a}_1$	$\underline{b}_1$	n_1
2	$\underline{a}_2$	$\underline{b}_2$	n_2
.	.	.	.
.	.	.	.
.	.	.	.
k	$\underline{a}_k$	$\underline{b}_k$	n_k
Total	$\underline{t}_1$	$\underline{t}_2$	N

We want to test the hypothesis:

$$(1.1) \qquad H_0 \; : \quad p_1 = p_2 = \ldots = p_k$$

against an upward or downward trend. This may be done e.g. in the following way:

We consider the n_i trials of the i-th series as n_i observations of a random variable $\underline{x}_i$, where $\underline{x}_i$ takes the values 0 and 1 with

$$(1.2) \qquad P[\underline{x}_i = 1] = p_i \quad , \quad P[\underline{x}_i = 0] = 1 - p_i \quad , \; i = 1, 2, \ldots, k.$$

Then H_0 is identical with the hypothesis that $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_k$ possess the same probability distribution and this hypothesis may be tested against the above mentioned alternatives by applying TERPSTRA's [5] test against trend to the observations of $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_k$. This test is executed as follows:

We apply WILCOXON's two-sample-test to the samples of $\underline{x}_i$ and $\underline{x}_j$. Then, if we denote WILCOXON's test-statistic for these two samples by $\underline{u}_{i,j}$:

---

1) Random variables will be denoted by underlined characters; values taken by a random variable are denoted by the same character, not underlined.

2) Unless explicity stated otherwise i and j take the values $1, 2, \ldots, k$.

$$(1.3) \qquad \underline{W}_{i,j} \overset{\text{def}}{=} 2\left[\underline{U}_{i,j} - \mathcal{E}(\underline{U}_{i,j} \mid H_0)\right] = \underline{a}_i n_j - \underline{a}_j n_i$$

and for TERPSTRA's test statistic $\underline{T}$ we have

$$(1.4) \qquad \underline{W} \overset{\text{def}}{=} 2\left[\underline{T} - \mathcal{E}(\underline{T} \mid H_0)\right] = \sum_{i<j}\sum \underline{W}_{i,j}.$$

Consequently

$$(1.5) \qquad \underline{W} = \sum_{i<j}\sum (\underline{a}_i n_j - \underline{a}_j n_i)$$

with (cf. [5]):

$$(1.6) \qquad \sigma^2\left[\underline{W} \mid t_1, H_0\right] = \frac{t_1 t_2 (N^3 - \Sigma n_i^3)}{3N(N-1)}.$$

In section 6 we shall prove that this test is consistent for the class of alternative hypotheses:

$$(1.7) \qquad H : \lim_{N\to\infty} \frac{\sum_{i<j}\sum n_i n_j (p_i - p_j)}{\sum_i n_i \left| \sum_{i<j} n_j - \sum_{i>j} n_j \right|} \neq 0 \quad {}^{3)}$$

and, for sufficiently small α , for no other alternatives.

Consequently if we apply TERPSTRA's test the class of alternative hypotheses for which the test is consistent depends on the sample-sizes n_i . This means, that as soon as at least one of the p_i differs from the others, the n_i may be chosen such, that the test is consistent, even if the p_i do not show a trend at all. According to a remark of J.HEMELRIJK this disagreable property ought to be avoided by choosing the test-statistic in such a way that the alternative hypotheses, for which the test is consistent, do not depend on the ratios of the numbers of observations taken from the different random variables, except possibly for boundary conditions of a general nature.

Taking this into account, the general form of our problem may be stated as follows. Consider N independent trials, each trial resulting in a success or a failure. The total number of successes is $\underline{t}_1$, $\underline{t}_2 = N - \underline{t}_1$, $\underline{x}_\lambda$ [4] is the number of successes and p_λ the probability of a success for the λ-th trial. We now want a test for the hypothesis

$$(1.8) \qquad H_0 : \quad p_1 = p_2 = \ldots = p_N ,$$

---

3) If $\lim_{N\to\infty} \sum_i \frac{n_i}{N} \left| \sum_{i<j} \frac{n_j}{N} - \sum_{i>j} \frac{n_j}{N} \right| \neq 0$ then (1.7) is identical with

$$\lim_{N\to\infty} \sum_{i<j}\sum \frac{n_i n_j}{N^2} (p_i - p_j) \neq 0 .$$

4) Unless explicitely stated otherwise λ and μ take the values $1, 2, \ldots, N$.

which is consistent for the class of alternative hypotheses

(1.9) $$H \ : \ \lim_{N \to \infty} \sum_{\lambda} q_{\lambda} p_{\lambda} \neq 0$$

and if possible for no other alternatives, where q_{λ} $(\lambda = 1, 2, \ldots, N)$ are given numbers.

These numbers must satisfy the condition

(1.10) $$\sum_{\lambda} q_{\lambda} = 0$$

because, if H_0 is true $\sum_{\lambda} q_{\lambda} p_{\lambda}$ must be equal to zero, in accordance with our wishes as to the consistency (cf.(1.9)). Imposing without any loss of generality, the condition

(1.11) $$\sum_{\lambda} | q_{\lambda} | = 1$$

we have

(1.12) $$| \sum_{\lambda} q_{\lambda} p_{\lambda} | \leq 1 .$$

In the special case (1.7) the class of admissible hypotheses consists of those values of $p_1, p_2, \ldots, p_N$ which satisfy

(1.13) $$\begin{cases} p_1 = p_2 = \ldots = p_{n_1}, \\ p_{n_1+1} = \ldots = p_{n_1+n_2}, \\ \vdots \\ p_{n_1+\ldots+n_{k-1}+1} = \ldots = p_{n_1+\ldots+n_k} \end{cases} \qquad \begin{array}{l} n_i > 0 \ , \ i = 1, 2, \ldots k \\ \\ k \geq 2 \ , \ \sum_i n_i = N \end{array}$$

and thus if we take

(1.14) $$q_{\lambda} = \frac{q_i'}{n_i} \qquad \begin{cases} n_1 + \ldots + n_{i-1} < \lambda \leq n_1 + \ldots + n_i \\ i = 1, 2, \ldots, k \ ; \ \lambda = 1, 2, \ldots, N \end{cases}$$

where q_i' are given numbers and if we put

(1.15) $$p_{\lambda} = p_i' \qquad \begin{cases} n_1 + \ldots + n_{i-1} < \lambda \leq n_1 + \ldots + n_i \\ i = 1, 2, \ldots, k \ ; \ \lambda = 1, 2, \ldots, N \end{cases}$$

then

(1.16) $$\sum_{\lambda} q_{\lambda} p_{\lambda} = \sum_i q_i' p_i' .$$

Condition (1.10) and (1.11) reduce to

(1.17) $$\sum_i q_i' = 0$$

and

(1.18) $$\sum_i | q_i' | = 1$$

respectively.

Consequently in the case (1.17) q_i' is proportional to $n_i \left(\sum_{i<j} n_j - \sum_{i>j} n_j \right)$, which introduces the n_i into (1.9). If we take q_i' proportional to $(k+1-2i)$ the above mentioned drawback of TERPSTRA's test is avoided and the alternatives, for which the test to be developed is consistent, are those, for which $\sum_i (k+1-2i)\, p_i' = \sum_{i<j}\sum (p_i' - p_j') \neq 0$.

In this paper we shall consider the general case (1.9). We test the hypothesis H_0 conditionally under the condition $\underline{t}_1 = t_1$ and we choose, on intuitive grounds as a test-statistic a linear combination of the random variables $\underline{x}_\lambda$:

(1.19) $$\underline{W} = \sum_\lambda h_\lambda \underline{x}_\lambda .$$

The $h_\lambda (\lambda = 1, 2, \ldots, N)$ will later on be expressed in terms of $g_1, g_2, \ldots, g_N$ such that the test is consistent for the class of alternative hypotheses (1.9) and for no other alternatives. In the special case of TERPSTRA's test against trend $h_\lambda (\lambda = 1, 2, \ldots, N)$ is proportional to $\sum_{i<j} n_j - \sum_{i>j} n_j$ $(n_1 + \ldots + n_{i-1} < \lambda \leq n_1 + \ldots n_i ;\ i = 1, 2, \ldots, k)$.

Without any loss of generality we can suppose

(1.20) $$\sum_\lambda h_\lambda = 0$$

which means that $\underline{W}$ is chosen in such a way that $\mathcal{E}[\underline{W} | t_1, H_0] = 0$ (cf. (2.6)).

2. The mean and variance of $\underline{W}$ under the hypothesis H_0.

Under H_0 and under the condition $\underline{t}_1 = t_1$ the simultanious distribution of the $\underline{x}_\lambda$ is an N-dimensional hypergeometric distribution, i.e.

(2.1) $$P[\underline{x}_1 = x_1 \wedge \underline{x}_2 = x_2 \wedge \ldots \wedge \underline{x}_N = x_N | t_1, H_0] = \frac{\prod_\lambda \binom{1}{x_\lambda}}{\binom{N}{t_1}} = \binom{N}{t_1}^{-1},$$

and

(2.2) $$\mathcal{E}[\underline{x}_\lambda | t_1, H_0] = \frac{t_1}{N},$$

(2.3) $$\sigma^2[\underline{x}_\lambda | t_1, H_0] = \frac{t_1 t_2}{N^2},$$

(2.4) $$\mathrm{cov}[\underline{x}_\lambda, \underline{x}_\mu | t_1, H_0] = -\frac{t_1 t_2}{N^2(N-1)} \qquad \lambda \neq \mu .$$

Consequently

(2.5) $$\sigma^2[\underline{W} | t_1, H_0] = \sum_\lambda h_\lambda^2 \sigma^2[\underline{x}_\lambda | t_1, H_0] + \sum_{\lambda \neq \mu}\sum h_\lambda h_\mu \mathrm{cov}[\underline{x}_\lambda, \underline{x}_\mu | t_1, H_0] =$$
$$= \frac{t_1 t_2}{N(N-1)} \sum_\lambda h_\lambda^2 \qquad \text{(cf. (1.20))},$$

(2.6) $$\mathcal{E}[\underline{W} | t_1, H_0] = \sum_\lambda h_\lambda \cdot \frac{t_1}{N} = 0 \qquad \text{(cf. (1.20))}.$$

3. The asymptotic distribution of $\underline{W}$ under the hypothesis H_0.

We consider a sequence of groups of trials, the ν-th group of which consists of N_ν trials of the kind described in section 1 and where

(3.1) $$\lim_{\nu\to\infty} N_\nu = \infty .$$

Then we have for each ν: $\underline{t}_{1\nu}$ successes, $\underline{t}_{2\nu}$ failures and a test-statistic

(3.2) $$\underline{W}_\nu = \sum_\lambda h_{\lambda\nu} \underline{x}_\lambda$$ 5)

with

(3.3) $$\mathcal{E}[\underline{W}_\nu | t_{1\nu}, H_0] = 0$$

(3.4) $$\sigma^2[\underline{W}_\nu | t_{1\nu}, H_0] = \frac{t_{1\nu}t_{2\nu}}{N(N-1)} \sum_\lambda h_{\lambda\nu}^2 .$$

We shall now prove the following theorem:
If the conditions

(3.5) $$\begin{cases} 1. & \dfrac{t_{1\nu}}{N} = O(1) , \quad \dfrac{t_{2\nu}}{N} = O(1) \\ 2. & \max\limits_{1 \leq \lambda \leq N_\nu} h_{\lambda\nu}^2 \Big/ \sum\limits_\lambda h_{\lambda\nu}^2 = o(1) \end{cases}$$

or the conditions

(3.6) $$\begin{cases} 1. & N_\nu^{r/2-1} \dfrac{\sum_\lambda h_{\lambda\nu}^r}{\left[\sum_\lambda h_{\lambda\nu}^2\right]^{r/2}} = O(1) \qquad \text{for each integer } r>2 \\ 2. & \lim t_1 = \infty , \quad \lim t_2 = \infty \end{cases}$$

are fulfilled the random variable

$$\frac{\underline{W}_\nu}{\sigma[\underline{W}_\nu | t_{1\nu}, H_0]}$$

is under the sequence of conditions $\underline{t}_{1\nu} = t_{1\nu}$ and under the hypothesis H_0, for ν tending to infinity, asymptotically nomally distributed with mean 0 and variance 1.

Proof 6).

For the proof we use theorems by WALD and WOLFOWITZ [6], NOETHER [4] and HOEFFDING [3]. To apply these theorems to our problem we consider the N trials as one observation

---

5) In this and the following section λ and μ take the values $1, 2, \ldots, N_\nu$ and all limits are for $\nu \to \infty$.

6) To simplify the notation we shall omit the index ν.

of each of the random variables $y_1, y_2, \ldots, y_N$ where the values taken by these variables form a permutation of the numbers $c_1, c_2, \ldots, c_N$. If we take for these numbers a row consisting of the numbers $h_1, h_2, \ldots, h_N$ and if a second row $d_1, d_2, \ldots, d_N$ consists of t_1 times the number 1 and t_2 times the number 0, then

$$(3.7) \qquad \underline{L}_N \overset{\text{def}}{=} \sum_\lambda d_\lambda \underline{y}_\lambda = \underline{W}.$$

The above mentioned theorems state that if

(3.8)

1. all permutations of $c_1, c_2, \ldots, c_N$ have the same probability,
2. the row $\{d_\lambda\}$ satisfies the condition

$$\frac{\mu_\tau\{d_\lambda\}}{[\mu_2\{d_\lambda\}]^{\tau/2}} = O(1) \qquad \text{for each integer } \tau > 2$$

where

$$\mu_\tau\{d_\lambda\} \overset{\text{def}}{=} \frac{1}{N} \sum_\lambda \Big\{d_\lambda - \frac{1}{N}\sum_\mu d_\mu\Big\}^\tau,$$

3. the row $\{c_\lambda\}$ satisfies the condition

$$\frac{\max\limits_{1\le\lambda\le N} \Big\{c_\lambda - \frac{1}{N}\sum_\mu c_\mu\Big\}^2}{\sum_\lambda \Big\{c_\lambda - \frac{1}{N}\sum_\mu c_\mu\Big\}^2} = o(1)$$

then the random variable

$$\frac{\underline{L}_N - \mathcal{E}(\underline{L}_N)}{\sigma(\underline{L}_N)}$$

is for ν tending to infinity asymptotically normally distributed with mean 0 and variance 1.

The condition (3.8.1.) is, given the independence of the trials, fulfilled if and only if H_0 is true and it is easy to see that the conditions (3.8.2.) and (3.8.3) reduce to (3.5.1) and (3.5.2) respectively.

The above mentioned theorems may also be applied in the following way:
If a row $\{c'_\lambda\}$ consists of t_1 times the number 1 and t_2 times the number 0 and a row $\{d'_\lambda\}$ consists of the numbers $h_1, h_2, \ldots, h_N$ then

$$(3.9) \qquad \underline{W} = \sum_\lambda d'_\lambda \underline{x}_\lambda$$

where the values taken by $\underline{x}_\lambda$ $(\lambda = 1, 2, \ldots, N)$ form a permutation of the numbers $c'_1, c'_2, \ldots, c'_N$. (cf. section 1).

Consequently $\frac{\underline{W}}{\sigma[\underline{W}|t_1,H_0]}$ is under the hypothesis H_0 and under the condition $\underline{t}_1 = t_1$, for ν tending to infinity, asymptotically normally distributed with mean 0 and variance 1 if the row $\{d'_\lambda\}$ satisfies condition (38.2) and the row $\{c'_\lambda\}$ the condition (3.8.3).

It is easy to see that in this case (3.8.2) reduces to (3.6.1) and (3.8.3) to (3.6.2).

4. The consistency of the test.

In this section we shall investigate the consistency of the test for the hypothesis H_0 if we take a one-sided critical region consisting of positive value of W. We again consider a sequence, the ν-th term of which consists of N_ν trials with

$$\lim N_\nu = \infty \qquad \text{(cf. section 3).}$$

We suppose that the conditions (3,5) or the conditions (3.6) are fullfilled; then for large ν the conditional critical region under the condition $\underline{t}_1 = t_1$ 7) consists of those values of W which satisfy

$$(4.1) \qquad \frac{W}{\sigma[\underline{W}|t_1,H_0]} \geq \xi_\alpha ,$$

where α is the level of significance and ξ_α follows from

$$\frac{1}{\sqrt{2\pi}} \int_{\xi_\alpha}^{\infty} e^{-\frac{1}{2}x^2} dx = \alpha.$$

If an alternative hypothesis H is true, $\sigma^2[\underline{W}|\underline{t}_1,H_0]$ converges in probability, for ν tending to infinity to

$$(4.2) \qquad \lim \frac{\sum_\lambda p_\lambda \sum_\lambda q_\lambda}{N(N-1)} \sum_\lambda h_\lambda^2 \qquad (= \lim \sigma_0^2 \text{ for short})$$
(cf. (2.5).

We define

$$(4.3) \qquad \mu \overset{\text{def}}{=} \mathcal{E}[\underline{W}|H] = \sum_\lambda h_\lambda p_\lambda ,$$

$$(4.4) \qquad \sigma^2 \overset{\text{def}}{=} \sigma^2[\underline{W}|H] = \sum_\lambda h_\lambda^2 p_\lambda q_\lambda .$$

We can suppose without any loss of generality

$$(4.5) \qquad \sum_\lambda |h_\lambda| = 1 ,$$

---

7) We again omit the index ν .

then

(4.6) $$\lim|\mu| = \lim\left|\sum_\lambda h_\lambda p_\lambda\right| \leqq 1,$$

and from (3.5.2) or (3.6.1) follows

(4.7) $$\lim \sum_\lambda h_\lambda^2 = 0.$$

Consequently

(4.8) $$\lim \sigma_0^2 = 0 \qquad \text{(cf.(4.2))}$$
(4.9) $$\lim \sigma^2 = 0 \qquad \text{(cf.(4.4))}$$

If now

(4.10) $$\lim \sum_\lambda h_\lambda p_\lambda > 0$$

then the probability of not-rejecting H_0 converges in probability for $\nu \to \infty$ to:

(4.11) $$\lim P\left[\frac{W}{\sigma_0} < \xi_\alpha\right] = \lim P[\underline{W} - \mu < \xi_\alpha\sigma_0 - \mu].$$

From (4.8) and (4.10) follows that $\xi_\alpha\sigma_0 - \mu$ is negative for sufficiently large ν ; consequently

(4.12) $$\lim P\left[\frac{W}{\sigma_0} < \xi_\alpha\right] \leqq \lim P[|\underline{W} - \mu| > \mu - \xi_\alpha\sigma_0] \leqq$$
$$\leqq \lim \frac{\sigma^2}{(\mu - \xi_\alpha\sigma_0)^2} = 0 \qquad \text{(cf. (4.9))}.$$

If

(4.13) $$\lim \sum_\lambda h_\lambda p_\lambda < 0$$

we see in the same way that the probability of rejecting H_0 converges in probability for $\nu \to \infty$ to 0. If finally

(4.14) $$\lim \sum_\lambda h_\lambda p_\lambda = 0$$

the probability of rejecting H_0 converges in probability for $\nu \to \infty$ to

(4.15) $$\lim P[\underline{W} \geqq \xi_\alpha \sigma_0]$$

Consequently if

(4.16) $$\lim \frac{\sigma_0^2}{\sigma^2} > 0$$

and

(4.17) $$\xi_\alpha^2 > \lim \frac{\sigma^2}{\sigma_0^2}$$

then

$$(4.18) \qquad \lim P[\underline{W} \geq \xi_\alpha \sigma_0] \leqq \lim \frac{\sigma^2}{\xi_\alpha^2 \sigma_0^2} < 1 .$$

The condition (4.16) is satisfied, if

$$(4.19) \qquad \lim \frac{\sum_\lambda p_\lambda}{N} \cdot \lim \frac{\sum_\lambda q_\lambda}{N} > 0$$

but this is not necessary the case if (4.16) is fullfilled. The class of alternative hypotheses with $\lim \frac{\sigma_0^2}{\sigma^2} = 0$ and $\lim \sum_\lambda h_\lambda p_\lambda = 0$ is of a rather unusual character, but it may be worthly of further investigation, because probably for at least a part of this class the test is also consistent.

5. Summary.

Substituting g_λ for h_λ in the above formulae, we get the following results. If we use the test-statistic

$$(5.1) \qquad \underline{W} = \sum_\lambda g_\lambda \underline{x}_\lambda ,$$

where g_λ $(\lambda = 1, 2, \dots, N)$ are given numbers, satisfying

$$(5.2) \qquad \sum_\lambda g_\lambda = 0$$

then

$$(5.3) \qquad \mathcal{E}[\underline{W} \mid t_1, H_0] = 0$$

$$(5.4) \qquad \sigma^2[\underline{W} \mid t_1, H_0] = \frac{t_1 t_2}{N(N-1)} \sum_\lambda g_\lambda^2 .$$

If

$$(5.5) \qquad \sum_\lambda |g_\lambda| = 1$$

and if the conditions

$$(5.6) \begin{cases} 1. \ \frac{t_1}{N} = O(1) \text{ and } \frac{t_2}{N} = O(1) \text{ for } N \to \infty \\ 2. \ \frac{\max_{1 \leq \lambda \leq N} g_\lambda^2}{\sum_\lambda g_\lambda^2} = o(1) \quad \text{for} \quad N \to \infty \end{cases} \qquad \text{(cf. (3.5))}$$

or the conditions:

$$(5.7)\begin{cases} 1.\ N^{r/2-1}\dfrac{\sum_\lambda g_\lambda^r}{[\sum_\lambda g_\lambda^2]^{r/2}} = O(1) & \text{for } \nu\to\infty \text{ and each integer } r>2 \quad \text{(cf. (3.6))} \\ 2.\ \lim t_1 = \infty\ ,\ \lim t_2 = \infty & \end{cases}$$

are satisfied and

$$(5.8)\qquad \lim \frac{\sigma_0^2}{\sigma^2} = \lim \frac{\sum_\lambda p_\lambda \sum_\lambda q_\lambda \sum_\lambda g_\lambda^2}{N(N-1)\sum_\lambda g_\lambda^2 p_\lambda q_\lambda} > 0 \qquad \text{(cf. (4.16))}$$

and

$$(5.9)\qquad \xi_\alpha^2 > \lim \frac{\sigma_0^2}{\sigma^2} \qquad \text{(cf. (4.17)).}$$

then the test is consistent for the class of alternative hypotheses

$$(5.10)\qquad \lim \sum_\lambda g_\lambda p_\lambda \neq 0$$

and for no other alternatives, and the conditional distribution of $\dfrac{\underline{W}}{\sigma[\underline{W}\,|t_1, H_0]}$ is under the hypothesis H_0 for $\nu\to\infty$ asymptotically normal.

6. Examples.

1. Suppose we want to test the hypothesis H_0 against the alternative hypotheses of trend, where a trend is defined by

$$\sum_{\lambda<\mu}\sum (p_\lambda - p_\mu) \neq 0.$$

From $\sum\sum_{\lambda<\mu}(p_\lambda - p_\mu) = \sum_\lambda (N+1-2\lambda)p_\lambda$ follows that g_λ is proportional to $N+1-2\lambda$ $(\lambda = 1,2,\dots,N)$ and therefore condition (5.2) is satisfied.

If we take

$$(6.1)\qquad g_\lambda = 2\,\frac{N+1-2\lambda}{N^2} \qquad \lambda = 1,2,\dots,N$$

then the conditions (5.5), (5.6.2) and (5.7.1) are all fullfilled.

The test-statistic is

$$(6.2)\qquad \underline{W} = \sum_\lambda (N+1-2\lambda)\underline{x}_\lambda = \sum_{\lambda<\mu}\sum (\underline{x}_\lambda - \underline{x}_\mu)$$

with

$$(6.3)\qquad \mathcal{E}[\underline{W}\,|t_1, H_0] = 0,$$

$$(6.4)\qquad \sigma^2[\underline{W}\,|t_1, H_0] = \frac{t_1 t_2}{N(N-1)}\sum_\lambda (N+1-2\lambda)^2 = \frac{1}{3}t_1 t_2 (N+1),$$

and the test is consitent for the class of alternative hypotheses

(6.5) $$\lim_{N\to\infty} \frac{1}{N^2} \sum_{\lambda<\mu}\sum (p_\lambda - p_\mu) \neq 0$$

and for no other alternative hypotheses if (5.8) and (5.9) are satisfied.

2. Suppose the class of admissible hypotheses consists of those values of $p_1, p_2, \ldots, p_N$ which satisfy (1.13). From section 5 it follows then that if we take as a test-statistic

(6.6) $$\underline{W} = \sum_i g_i' \frac{\underline{a}_i}{n_i},$$

where g_i' $(i = 1, 2, \ldots, k)$ are given numbers, satisfying

(6.7) $$\sum_i g_i' = 0$$

then

(6.8) $$\mathcal{E}[\underline{W} \mid t_1, H_0] = 0,$$

(6.9) $$\sigma^2[\underline{W} \mid t_1, H_0] = \frac{t_1 t_2}{N(N-1)} \sum_i \frac{g_i'^2}{n_i}.$$

The test is then consistent for the class of alternative hypotheses

(6.10) $$\lim_{N\to\infty} \sum_i g_i' p_i' \neq 0 \qquad \text{(cf. (1.16))}$$

and for no other alternatives, if

(6.11) $$\sum_i |g_i'| = 1 \qquad \text{(cf. (5.5)),}$$

the conditions

(6.12) $$\begin{cases} 1.\ \dfrac{t_1}{N} = O(1) \quad \text{and} \quad \dfrac{t_2}{N} = O(1) \quad \text{for} \quad N \to \infty \\ 2.\ \dfrac{\max\limits_{1\le i\le k} \dfrac{g_i'^2}{n_i^2}}{\sum\limits_i \dfrac{g_i'^2}{n_i}} = o(1) \quad \text{for } N \to \infty \end{cases} \qquad \text{(cf. (5.6))}$$

or the conditions

(6.13) $$\begin{cases} 1.\ N^{r/2-1} \dfrac{\sum\limits_i \dfrac{g_i'^r}{n_i^{r-1}}}{\left[\sum\limits_i \dfrac{g_i'^2}{n_i}\right]^{r/2}} = O(1) \quad \text{for} \quad N \to \infty \\ 2.\ \lim t_1 = \infty \quad \text{and} \quad \lim t_2 = \infty \end{cases} \qquad \text{(cf. (5.8))}$$

are satisfied and if furthermore the conditions

(6.14) $$\lim \frac{\sigma_a^2}{\sigma^2} = \lim \frac{\sum_i n_i p_i \sum_i n_i q_i \sum_i \frac{g_i'^2}{n_i}}{N(N-1) \sum_i \frac{g_i'^2}{n_i} p_i q_i} > 0$$ (cf.(5.8)

(6.15) $$\xi_\alpha^2 > \lim \frac{\sigma^2}{\sigma_a^2}$$ (cf.(5.9)

are fullfilled. Consequently if g_i' is independent of $n_1, n_2, \ldots, n_k$ the test is consistent for a class of alternative hypotheses which does not depend on the sample sizes E.g. if we take

(6.16) $$g_i' = 2\,\frac{k+1-2i}{k^2} \qquad i = 1, 2, \ldots, k$$

then the conditions (6.7) and (6.11) are satisfied. If k is finite condition (6.12.2) is fullfilled if

(6.17) $$\begin{cases} \lim\limits_{N\to\infty} n_i = \infty & \text{for } i \neq \frac{k+1}{2} \\ \lim\limits_{N\to\infty} n_i \leqq \infty & \text{for } i = \frac{k+1}{2} \end{cases}$$

and (6.13.1) is fullfilled if

(6.18) $$\begin{cases} \lim\limits_{N\to\infty} \frac{n_i}{N} > 0 & \text{for } i \neq \frac{k+1}{2} \\ \lim\limits_{N\to\infty} \frac{n_i}{N} \geqq 0 & \text{for } i = \frac{k+1}{2}\,. \end{cases}$$

The test-statistic is

(6.19) $$\underline{W} = \sum_{i<j}\sum \left(\frac{\underline{a}_i}{n_i} - \frac{\underline{a}_j}{n_j} \right) = \sum_{i<j}\sum \frac{\underline{W}_{i,j}}{n_i n_j}$$ (cf.(1.3)),

with

(6.20) $$\mathcal{E}[\underline{W} \mid t_1, H_0] = 0$$

(6.21) $$\sigma^2[\underline{W} \mid t_1, H_0] = \frac{t_1 t_2}{N(N-1)} \sum_i \frac{(k+1-2i)^2}{n_i},$$

and the test is consistent for the class of alternative hypothe-

(6.22) $$\lim_{N\to\infty} \frac{1}{k^2} \sum_{i<j}\sum (p_i - p_j) \neq 0$$

and for no other alternatives if (6.14) and (6.15) are fullfilled. For TERPSTRA's test-statistic we have

(6.23) $$g_i' = \frac{n_i\left(\sum_{i<j} n_j - \sum_{i>j} n_j\right)}{\sum_i n_i \left| \sum_{i<j} n_j - \sum_{i>j} n_j \right|}$$ (cf. (1.3) and (1.4))

Condition (6.12.2) reduces to

(6.24) $$\frac{\max\limits_{1\le i\le k} \left(\sum_{i<j} n_j - \sum_{i>j} n_j\right)^2}{N^3 - \sum_i n_i^3} = o(1) \qquad \text{for } N \to \infty$$

and (6.13.1) reduces to

$$(6.25)\quad N^{r/2-1}\,\frac{\sum_i n_i\left(\sum_{i<j} n_j-\sum_{i>j} n_j\right)^r}{\left[N^3-\sum_i n_i^3\right]^{r/2}}=O(1)$$ for $N\to\infty$ and each integer $r>2$.

From (6.9) it follows, that

$$(6.26)\quad \sigma^2[\underline{W}\,|\,t_1,H_0]=\frac{t_1t_2}{N(N-1)}\sum_i n_i\left(\sum_{i<j} n_j-\sum_{i>j} n_j\right)^2=\frac{t_1t_2(N^3-\sum_i n_i^3)}{3N(N-1)}$$

with

$$(6.27)\qquad \underline{W}=\sum_{i<j}\sum(\underline{a}_i n_j-\underline{a}_j n_i)\qquad \text{(cf. (1.6)).}$$

The test is consistent for the class of alternative hypotheses

$$(6.28)\qquad \lim_{N\to\infty}\frac{\sum_{i<j}\sum n_in_j(p_i-p_j)}{\sum_i n_i\left|\sum_{i<j} n_j-\sum_{i>j} n_j\right|}\neq 0$$

and for no other alternatives if (6.14) and (6.15) are fullfilled
If

$$(6.29)\qquad \lim_{N\to\infty}\sum_i\frac{n_i}{N}\left|\sum_{i<j}\frac{n_j}{N}-\sum_{i>j}\frac{n_j}{N}\right|\neq 0$$

then (6.28) is identical with

$$(6.30)\qquad \lim_{N\to\infty}\sum_{i<j}\sum\frac{n_in_j}{N^2}(p_i-p_j)\neq 0$$

and as

$$(6.31)\qquad p_i-p_j=P[\underline{x}_i>\underline{x}_j]-P[\underline{x}_i<\underline{x}_j]$$

(6.30) is identical with

$$(6.32)\quad \lim_{N\to\infty}\sum_{i<j}\sum\frac{n_in_j}{N^2}\left\{P[\underline{x}_i>\underline{x}_j]-P[\underline{x}_i<\underline{x}_j]\right\}\neq 0.$$

7. Remarks.

1. The test of example 1 and TERPSTRA's test may also be derived in the following way:

Consider two random variables $\underline{x}$ and $\underline{y}$ where, in two samples of size t_1 and t_2 respectively, $\underline{x}$ and $\underline{y}$ have taken the value i $\underline{a}_i$ and $\underline{b}_i$ times respectively. If $\underline{u}$ is the test-statistic of WILCOXON's test applied to these two samples then

(7.1) $$2\underline{U} = \underline{W} + t_1 t_2 .$$

Hemelrijk [1] and [2] has proved that the hypothese H_0 under the condition $t_1 = t_1$ is identical with the hypothesis that $\underline{x}$ and $\underline{y}$ possess the same probability distribution. Consequently

(7.2) $$\mathcal{E}[\underline{W} \mid t_1, H_0] = 2\mathcal{E}[\underline{U} \mid t_1, H_0] - t_1 t_2 = 0$$

(7.3) $$\sigma^2[\underline{W} \mid t_1, H_0] = 4\sigma^2[\underline{U} \mid t_1, H_0] = 4 \cdot \frac{t_1 t_2 (N^3 - \sum_i n_i^3)}{12 N (N-1)}$$

or, if all n_i are equal to 1 (example 1):

(7.4) $$\sigma^2[\underline{W} \mid t_1, H_0] = 4 \cdot \frac{t_1 t_2 (N+1)}{12} .$$

For the case that all n_i are equal to 1 the exact distribution of $\underline{U}$ under the hypothesis H_0 is known for small values of t_1 and t_2; therefore in this case an exact test is possible.

2. If $n_i = n$ for each i TERPSTRA's test is identical with the test given by (6.19). Consequently

a. if $n_i = n$ for each i Terpstra's test is consistent for a class of alternative hypotheses which does not depend on the sample-sizes,

b. if $n_i = n$ for each i the test given by (6.19) is identical with Wilcoxon's two sample test applied to the samples of $\underline{x}$ and $\underline{y}$ (cf. remark 1).

3. In the preceding sections we proved that if we take as a test-statistic

(7.5) $$\underline{W} = \sum_{i<j}\sum \frac{\underline{W}_{ij}}{n_i n_j}$$

instead of TERPSTRA's test-statistic $\sum_{i<j}\sum \underline{W}_{ij}$ the test is consistent for the class of alternative hypotheses

$$\lim_{N\to\infty} \frac{1}{k^2} \sum_{i<j}\sum (p_i' - p_j') \neq 0$$

which is independent of the sample-sizes.

If now $\underline{x}_i$ possesses a continuous or a discrete distribution function and if n_i observations of $\underline{x}_i$ are given $(i = 1, 2, \dots, k)$ it may be proved that if we take (7.5) as a test-statistic to test the hypothesis H_0 that $\underline{x}_1, \underline{x}_2, \dots, \underline{x}_k$ possess the same distributionfunction instead of TERPSTRA's test-statistic the test is consistent for the class of alternative hypotheses

$$\lim_{N\to\infty} \frac{1}{k^2} \sum_{i<j}\sum \left\{ P[\underline{x}_i > \underline{x}_j] - P[\underline{x}_i < \underline{x}_j] \right\} \neq 0$$

which again does not depend on the n_i.

4. The test described in the preceding sections may be generalised e.g. in the following way:

Consider N random variables $x_1, x_2, \ldots, x_N$ where x_λ takes the values $1, 2, \ldots, \ell$ with probabilities $p_{\lambda 1}, p_{\lambda 2}, \ldots, p_{\lambda \ell}$ respectively ($\lambda = 1, 2, \ldots, N$; $\sum_{i=1}^{\ell} p_{\lambda i} = 1$ for each λ). Given one observation of each of these random variables we want to test the hypotheses H_0 that $x_1, x_2, \ldots, x_N$ possess the same probability distribution, which is identical with

$$p_{1i} = p_{2i} = \ldots = p_{Ni} \qquad \text{for each } i \ (i = 1, 2, \ldots, \ell),$$

against the alternative hypotheses

$$\sum_\lambda \sum_i q_{\lambda i} \, p_{\lambda i} \neq 0,$$

where $q_{\lambda i}$ ($\lambda = 1, 2, \ldots, N$; $i = 1, 2, \ldots, \ell$) are given numbers. If H_0 is true

$$\sum_\lambda \sum_i q_{\lambda i} \, p_{\lambda i} = \sum_i p_i \sum_\lambda q_{\lambda i}$$

and this must be equal to zero, Consequently the numbers $q_{\lambda i}$ must satisfy the conditions

$$\sum_\lambda q_{\lambda i} = 0 \qquad \text{for each } i \quad (i = 1, 2, \ldots, \ell).$$

If $\ell = 2$

$$\sum_\lambda \sum_i q_{\lambda i} \, p_{\lambda i} = \sum_\lambda (q_{\lambda 1} p_{\lambda 1} + q_{\lambda 2} p_{\lambda 2}) = \sum_\lambda (q_{\lambda 1} - q_{\lambda 2}) \, p_{\lambda 1} + \sum_\lambda q_{\lambda 2} =$$
$$= \sum_\lambda (q_{\lambda 1} - q_{\lambda 2}) \, p_{\lambda 1} = \sum_\lambda q_\lambda \, p_\lambda$$

where

$$q_\lambda = q_{\lambda 1} - q_{\lambda 2} \quad , \ \lambda = 1, 2, \ldots, N \ ; \ \sum_\lambda q_\lambda = 0$$

and

$$p_\lambda = p_{\lambda 1} \quad , \ \lambda = 1, 2, \ldots, N.$$

References.

[1] Hemelrijk, J., Note on Wilcoxon's two sample test when ties are present, Ann. Math.Stat 23 (1952) p. 133-135.

[2] Hemelrijk, J., Exemple d'application des méthodes non-paramétriques et un nouveau test pour l'égalité de plusieurs probabilités, Report SP 38 of the Mathematical Centre, Amsterdam (1954).

[3] Hoeffding, W., A combinatorial central limit theorem, Ann.Math.Stat. 22 (1951), p. 558-566.

[4] Noether, G.E., On a theorem by Wald and Wolfowitz, Ann.Math.Stat. 20 (1949) p. 455-461.

[5] Terpstra, T.J., The assymptotic normality and consistency of Kendall's test against trend when ties are present in one ranking, Proc. Kon.Ned.Akad.v.Wet. A 55(1952),p.327-333.

[6] Wald, A. and Wolfowitz, J., Statistical tests based on permutations of the observations, Ann. Math. Stat. 15 (1944), p. 358-372.