65

31 SESSION DE L'INSTITUT INTERNATIONAL DE STATISTIQUE
SESSION OF THE INTERNATIONAL STATISTICAL INSTITUTE
SESSIONE DELL'ISTITUTO INTERNAZIONALE DI STATISTICA

C. VAN EEDEN

A class of tests for the hypothesis that K parameters $\theta_1, \ldots, \theta_k$ satisfy the inequalities $\theta_1 \leqslant \ldots \leqslant \theta_k$

BRUXELLES
1960

TIRE A PART DU
BULLETIN DE L'INSTITUT INTERNATIONAL DE STATISTIQUE
TOME XXXVII - 3ème LIVRAISON

S.A. IMPRIMERIE L. VANMELLE - GAND

A CLASS OF TESTS FOR THE HYPOTHESIS THAT K PARAMETERS $\theta_1, \ldots, \theta_k$ SATISFY THE INEQUALITIES $\theta_1 \leqslant \ldots \leqslant \theta_k$ (*)

by

Dr. Constance van EEDEN

1. — INTRODUCTION

In this paper a description will be given of a class of tests treated in chapter 4 of my thesis [4]. By means of these tests the hypothesis H_0 that k parameters $\theta_1, \ldots, \theta_k$ satisfy the inequalities

$$\theta_1 \leqslant \ldots \leqslant \theta_k \tag{1.1}$$

may be tested against the alternative hypothesis that at least one value of i exists with $\theta_i > \theta_{i+1}$.
In the chapters 1-3 of my thesis a related problem is treated namely the problem of estimating k unknown parameters $\theta_1, \ldots, \theta_k$, known to satisfy

(1.2)
1. inequalities of the type : $\varphi_i(\theta_i) \leqslant \varphi_j(\theta_j)$,
2. inequalities of the type : $c_i \leqslant \varphi_i(\theta_i) \leqslant d_i$,

where, for each $i = 1, \ldots, k$, $\varphi_i(\theta_i)$ is a given function of θ_i, whereas c_i and d_i are given numbers. A special case of this problem is e.g. the estimation of k parameters $\theta_1, \ldots, \theta_k$, known to satisfy the equalities $\theta_1 \leqslant \ldots \leqslant \theta_k$.

A description of this estimationproblem and its solution has been given by J. HEMELRIJK [5]. The proofs may be found in [4].

A description of the class of tests for the hypothesis (1.1) will be given in this paper in section 2. Section 3 contains the special cases where θ_i is

1. the parameter of an exponential distribution,
2. the variance of a normal distribution,
3. the mean of a normal distribution with known variance,
4. the length of the interval of a rectangular distribution.

(*) Report SP 65 of the Statistical Department of the Mathematical Centre, Amsterdam.

Further an analogous distributionfree test, based on WILCOXON's two sample test, will be described.

In this paper no proofs will be given; these may be found in [4].

2. — DESCRIPTION OF THE TESTS

The situation to be considered may be described as follows. Let $\underline{x}_1, \ldots, \underline{x}_k$ [1]) be k independent random variables and let, for each $i = 1, \ldots, k$, $x_{i,\gamma}$ $(\gamma = 1, \ldots, n_i)$ be n_i independent observations of $\underline{x}_i$. Let further, for each $i = 1, \ldots, k$, θ_i denote an unknown parameter of the distribution of $\underline{x}_i$.
The hypothesis

$$H_0 : \theta_1 \leqslant \ldots \leqslant \theta_k \tag{2.1}$$

will be tested against the alternative hypothesis

$$H : \text{at least one value of i exists with } \theta_i > \theta_{i+1}. \tag{2.2}$$

This test is performed as follows. Let, for each $i = 1, \ldots, k-1$, T_i denote a test for the hypothesis

$$H_{0,i} : \theta_i \leqslant \theta_{i+1} \tag{2.3}$$

against the alternative hypothesis

$$H_i : \theta_i > \theta_{i+1}. \tag{2.4}$$

Let, for each $i = 1, \ldots, k-1$, t_i denote the test statistic and Z_i the critical region of this test. Then t_i is a function of $x_{i,1}, \ldots, x_{i,n_i}, x_{i+1,1}, \ldots, x_{i+1,n_{i+1}}$ and $H_{0,i}$ is rejected if and only if $t_i \varepsilon Z_i$.
The test for the hypothesis H_0 then consists of rejecting H_0 if and only if a value of i exists with $t_i \varepsilon Z_i$.

Now suppose that the tests $T_1, \ldots, T_{k-1}$ possess the following properties. Let

$$\left\{ \begin{array}{l} \alpha_i \overset{\text{def}}{=} \mathrm{P}\{\underline{t}_i \,\varepsilon\, Z_i \mid \theta_i = \theta_{i+1}\}, \;^{2)} \\ N_i \overset{\text{def}}{=} n_i + n_{i+1} \end{array} \right. \tag{2.5}$$

and let, for each $i = 1, \ldots, k-1$, the limit $N_i \to \infty$ be taken under the conditions

$$\left\{ \begin{array}{l} \lim\limits_{N_i \to \infty} n_i = \infty, \\ \lim\limits_{N_i \to \infty} n_{i+1} = \infty, \end{array} \right. \tag{2.6}$$

then we suppose that, for each $i = 1, \ldots, k-1$,

[1]) Random variables are distinguished from numbers (e.g. from the values they take in an experiment) by underlining their symbols.

[2]) $\mathrm{P}\{A\}$ denotes the probability of event A.

$$(2.7)\quad \begin{cases} 1. \ \mathrm{P}\{\underline{t}_i \,\varepsilon\, Z_i \mid \theta_i < \theta_{i+1}\} \leqslant \alpha_i, \\ 2. \ \lim_{N_i \to \infty} \mathrm{P}\{\underline{t}_i \,\varepsilon\, Z_i \mid \theta_i < \theta_{i+1}\} = 0, \\ 3. \ \lim_{N_i \to \infty} \mathrm{P}\{\underline{t}_i \,\varepsilon\, Z_i \mid \theta_i > \theta_{i+1}\} = 1. \end{cases}$$

Now it may easily be proved (cf. [4]) that the test for the hypothesis H_0 possesses the following properties. Let α_0 denote the size of the critical region of the test for H_0 (i.e. let α_0 denote the probability, if H_0 is true, of rejecting H_0), let

$$(2.8)\quad n \stackrel{\text{def}}{=} \sum_{i=1}^{k} n_i$$

and let the limit $n \to \infty$ be taken under the conditions

$$(2.9)\quad \lim_{n\to\infty} n_i = \infty \text{ for each } i = 1, \ldots, k,$$

then we have

(2.10)
1. $\alpha_0 \leqslant \sum_{i=1}^{k-1} \alpha_i$,
2. the probability of rejecting H_0, under the hypothesis $\theta_1 < \ldots < \theta_k$, tends to zero for $n \to \infty$,
3. the probability of rejecting H_0, under the hypothesis H, tends to 1 for $n \to \infty$.

If, moreover, we suppose that, for each pair of values (i,j) with $i < j$

$$(2.11)\quad \mathrm{P}\{\underline{t}_i \,\varepsilon\, Z_i \text{ and } \underline{t}_j \,\varepsilon\, Z_j \mid \theta_i = \theta_{i+1},\, \theta_j = \theta_{j+1}\} \leqslant \leqslant \mathrm{P}\{\underline{t}_i \,\varepsilon\, Z_i \mid \theta_i = \theta_{i+1}\} \,.\, \mathrm{P}\{\underline{t}_j \,\varepsilon\, Z_j \mid \theta_j = \theta_{j+1}\},$$

then we have also (cf. [3] and [4])

(2.12) the probability of rejecting H_0, under the hypothesis $\theta_1 = \ldots = \theta_k$, is $\geqslant \sum_{i=1}^{k-1} \alpha_i - \frac{1}{2} \{\sum_{i=1}^{k-1} \alpha_i\}^2$.

Thus if we take e.g. $\sum_{i=1}^{k-1} \alpha_i = 0{,}05$ then we have

1. the probability of rejecting H_0, if H_0 is true, is $\leqslant 0{,}05$,
2. the probability of rejecting H_0, under the hypothesis

   $\theta_1 = \ldots = \theta_k$, is $\geqslant 0{,}05 - \frac{1}{2}(0{,}05)^2 = 0{,}04875$.

Tests T_i satisfying the conditions (2.7) and (2.11) will be described in section 3.

3. — EXAMPLES

3.1. — *An exponential distribution with parameter* θ_i

We first consider the case that $\underline{x}_i$ possesses, for each $i = 1, \ldots, k$, an exponential distribution with parameter θ_i, i. e.

(3.1.1) $$P\{\underline{x}_i \leqslant x\} = 1 - e^{-\theta_i x} \quad (x \geqslant 0).$$

Now let, for each $i = 1, \ldots, k$,

(3.1.2) $$\bar{x}_i \stackrel{\text{def}}{=} \frac{1}{n_i} \sum_{\gamma=1}^{n_i} x_{i,\gamma}$$

then we take, for each $i = 1, \ldots, k-1$, as a test statistic for the hypothesis $H_{0,i}$

(3.1.3) $$t_i = \frac{\bar{x}_{i+1}}{\bar{x}_i}$$

and for Z_i we take a critical region of the form $t_i \geqslant t_{i,\alpha_i}$ where [cf. (2.5)] t_{i,α_i} satisfies

(3.1.4) $$P\{\underline{t}_i \geqslant \underline{t}_{i,\alpha_i} \mid \theta_i = \theta_{i+1}\} = \alpha_i.$$

Now (3.1.1) entails that for each $i = 1, \ldots, k$, $2\theta_i n_i \underline{\vec{x}}_i$ possesses a χ^2-distribution with $2n_i$ degrees of freedom, thus $\underline{t}_i$ possesses, for each $i = 1, \ldots, k-1$, under the hypothesis $\theta_i = \theta_{i+1}$, an F-distribution with $2n_{i+1}$, and $2n_i$ degrees of freedom. Thus the critical values t_{i,α_i} may be found from a table of the F-distribution.

It may easily be proved (cf. [4]) that these tests $T_1, \ldots, T_{k-1}$ satisfy the conditions (2.7) and 2.11).

3.2. — *A normal distribution with variance* θ_i

Now let, for each $i = 1, \ldots, k$, $\underline{x}_i$ possess a normal distribution with unknown mean μ_i and variance θ_i. Then, if

(3.2.1) $$\begin{cases} \bar{x}_i \stackrel{\text{def}}{=} \sum_{\gamma=1}^{n_i} x_{i,\gamma} & (i = 1, \ldots, k) \\ s_i^2 \stackrel{\text{def}}{=} \frac{1}{n_i - 1} \sum_{\gamma=1}^{n_i} (x_{i,\gamma} - \bar{x}_i)^2, & \end{cases}$$

we take, as a test statistic for the hypothesis $H_{0,i}$

(3.2.2) $$t_i = \frac{s_i^2}{s_{i+1}^2} (i = 1, \ldots, k-1).$$

Now $\dfrac{(n_i - 1)\underline{s}_i^2}{\theta_i}$ possesses, for each $i = 1, \ldots, k$, a χ^2-distribution with $n_i - 1$ degrees of freedom ; thus, for each $i = 1, \ldots, k-1$, $\underline{t}_i$ possesses, under the hypothesis $\theta_i = \theta_{i+1}$, an F-distribution with $n_i - 1$ and $n_{i+1} - 1$ degrees of freedom. We again take critical regions of the form $t_i \geqslant t_{i,\alpha_i}$, where t_{i,α_i} may be found from a table of the F-distribution.

The proofs of (2.7) and (2.11) are identical with those of the foregoing example.

Remark (3.2.1)

If μ_i is known then $\underline{s}_i^2$ is replaced by $\underline{s}_i'^2 \overset{\text{def}}{=} \frac{1}{n_i} \sum_{\gamma=1}^{n_i} (\underline{x}_{i,\gamma} - \mu_i)^2$, where $\frac{n_i \underline{s}_i'^2}{\theta_i}$ possesses a χ^2-distribution with n_i degrees of freedom.

3.3. — *A normal distribution with mean θ_i and known variance*

We now consider the case that, for each $i = 1, \ldots, k$, $\underline{x}_i$ possesses a normal distribution with mean θ_i and known variance σ_i^2. Let, for each $i = 1, \ldots, k$,

$$\bar{x}_i \overset{\text{def}}{=} \frac{1}{n_i} \sum_{\gamma=1}^{n_i} x_{i,\gamma} \tag{3.3.1.}$$

then we take

$$t_i = \bar{x}_{i+1} - \bar{x}_i \quad (i = 1, \ldots, k-1). \tag{3.3.2}$$

The statistic $\underline{t}_i$ possesses, under the hypothesis $\theta_i = \theta_{i+1}$, a normal distribution with zero mean and variance

$$\sigma^2(\underline{t}_i \mid \theta_i = \theta_{i+1}) = \frac{\sigma_i^2}{n_i} + \frac{\sigma_{i+1}^2}{n_{i+1}} \quad (i = 1, \ldots, k-1). \tag{3.3.3}$$

We take a critical region of the form $t_i \geqslant t_{i,\alpha_i}$; then

$$t_{i,\alpha_i} = \xi_{\alpha_i} \sqrt{\frac{\sigma_i^2}{n_i} + \frac{\sigma_{i+1}^2}{n_{i+1}}}, \tag{3.3.4}$$

where ξ_α is defined by

$$\frac{1}{\sqrt{2\pi}} \int_{\xi_\alpha}^{\infty} e^{-\frac{1}{2}x^2}\, dx = \alpha. \tag{3.3.5}$$

Thus t_{i,α_i} may be found by means of a table of the normal distribution. It may easily be seen that this test satisfies (2.7). Further $\underline{t}_i$ and $\underline{t}_j$ are, for $j > i + 1$, independently distributed, i.e. (2.11) holds for each pair of values (i, j) with $j > i + 1$. For $j = i + 1$, $\underline{t}_i$ and $\underline{t}_j$ possess a two-dimensional normal distribution with negative correlationcoefficient and it may easily be proved (cf. [2]) that (2.11) holds in this case.

3.4. — *A rectangular distribution between 0 and θ_i*

Finally, let, for each $i = 1, \ldots, k$, $\underline{x}_i$ possess a rectangular distribution between 0 and $\theta_i > 0$. Let, for each $i = 1, \ldots, k$, (3.4.1) $$z_i \overset{\text{def}}{=} \max_{1 \leqslant \gamma \leqslant n_i} x_{i,\gamma},$$

then (cf. [4], chapter 2) z_i is the maximum likelihood estimate of θ_i. In this case we take, for $i = 1, \ldots, k-1$,

(3.4.2) $$t_i = \frac{z_i}{z_{i+1}}$$

with critical regions of the form $t_i \geqslant t_{i,\alpha_i}$.

Now we have (cf. [4])

(3.4.3) $$t_{i,\alpha_i} = \begin{cases} \left(\dfrac{n_i}{N_i \alpha_i} \right)^{\frac{1}{n_{i+1}}} & \text{if } \alpha_i \leqslant \dfrac{n_i}{N_i}, \\ \left\{ \dfrac{N_i}{n_{i+1}} (1-\alpha_i) \right\}^{\frac{1}{n_i}} & \text{if } \alpha_i \geqslant \dfrac{n_i}{N_i}. \end{cases}$$

The proof of (2.7) and (2.11) may be found in [4].

3.5. — *An analogous distributionfree test*

In this section an analogous distributionfree test based on WILCOXON's two sample test will be described. Let $\underline{x}_1, \ldots, \underline{x}_k$ be independent random variables, possessing continuous probability distributions. Let further, for each $i = 1, \ldots, k$, $x_{i,1}, \ldots, x_{i,n_i}$ be independent observations of $\underline{x}_i$ and let (cf. [1])

(3.5.1) $$W_i \overset{\text{def}}{=} \sum_{\gamma=1}^{n_i} \sum_{\lambda=1}^{n_{i+1}} sgn(x_{i,\gamma} - x_{i+1,\lambda}),$$ [3])

where

(3.5.2) $$sgn\ z \overset{\text{def}}{=} \begin{cases} 1 & \text{if } z > 0, \\ 0 & \text{if } z = 0, \\ -1 & \text{if } z < 0. \end{cases}$$

In the sequel of this section a test will be described for the hypothesis H'_0 that $\underline{x}_1, \ldots, \underline{x}_k$ possess the same probability distribution. This test is based on $W_1, \ldots, W_{k-1}$ and is performed as follows. Let, for $i = 1, \ldots, k-1$, $H'_{0,i}$ denote the hypothesis that $\underline{x}_i$ and $\underline{x}_{i+1}$ possess the same probability distribution and let Z'_i denote a critical region of the form $W_i \geqslant W_{i,\alpha_i}$ where

(3.5.3) $$\mathrm{P}\{\underline{W}_i \,\varepsilon\, Z'_i \,|\, H'_{0,i}\} = \mathrm{P}\{\underline{W}_i \geqslant W_{i,\alpha_i} \,|\, H'_{0,i}\} = \alpha_i.$$

[3]) If U_i is the test statistic of WILCOXON's two sample test, according to H.B. MANN and D.R. WHITNEY [6] then $W_i = 2U_i - n_i n_{i+1}$.

Then the hypothesis H'_0 is rejected if and only if a value of i exists with $\underline{W}_i \varepsilon Z'_i$.

For small values of n_i and n_{i+1} the critical values W_{i,α_i} may be found from a table of the exact probability distribution of $\underline{W}_i$ under the hypothesis $H'_{0,i}$ (cf. e.g. [6] and [7]). For large values of n_i and n_{i+1} $\underline{W}_i$ is under the hypothesis $H'_{0,i}$ approximately normally distributed with zero mean and variance

$$(3.5.4) \qquad \sigma^2(\underline{W}_i \mid H'_{0,i}) = \frac{1}{3} n_i n_{i+1}(N_i + 1).$$

Thus in this case an approximation to W_{i,α_i} may be found from a table of the normal distribution.

Now let α_0 denote the size of the critical region of the test for H'_0, i.e. let

$$(3.5.5) \qquad \alpha_0 \overset{\text{def}}{=} \mathrm{P}\{\underline{W}_i \varepsilon Z'_i \text{ for at least on value of } i \mid H'_0\}$$

then it may be proved (cf. [4]) that

$$(3.5.6) \qquad \sum_{i=1}^{k-1} \alpha_i - \tfrac{1}{2}\{\sum_{i=1}^{k-1} \alpha_i\}^2 \leqslant \alpha_0 \leqslant \sum_{i=1}^{k-1} \alpha_i.$$

Further the test for the hypothesis H'_0 possesses the following properties. Let

$$(3.5.7) \qquad \theta'_i \overset{\text{def}}{=} \mathrm{P}\{\underline{x}_i > \underline{x}_{i+1}\} \qquad (i = 1, \ldots, k-1),$$

let the limit $n \to \infty$ be taken under the conditions

$$(3.5.8) \qquad \lim_{n\to\infty} n_i = \infty \qquad \text{for each } i = 1, \ldots, k$$

and let H'_1, H'_2 *and* H'_3 denote the hypotheses

(3.5.9)

1. H'_1 : for each value of i : $\theta'_i < \frac{1}{2}$,
2. H'_2 : at least one value of i exists with $\theta'_i > \frac{1}{2}$,
3. H'_3 : for each values of i : $\theta'_i \leqslant \frac{1}{2}$, at least one values of i exists with $\theta'_i = \frac{1}{2}$.

Then we have, (cf. [4]), for $n \to \infty$

(3.5.10)

1. the probability of rejecting H'_0 under the hypothesis H'_1 tends to zero,
2. the probability of rejecting H'_0 under the hypothesis H'_2 tends to 1,
3. if α_i is sufficiently small for each value of i with $\theta'_i = \frac{1}{2}$, the probability of rejecting H'_0 under the hypothesis H'_3 tends to a limit < 1.

REFERENCES

[1] VAN DANTZIG, D. and J. HEMELRIJK: Statistical methods based on few assumptions, Bull. of the Intern. Stat. Inst. *34* (1954) 239-267.

[2] DOORNBOS, R., H. KESTEN and H.J. PRINS: A class of slippage tests, Report S 206 (VP 8) of the Statistical Department of the Mathematical Centre, Amsterdam (1956).

[3] DOORNBOS, R. and H.J. PRINS: On slippage tests, Proc. Kon. Ned. Akad. v. Wet. A *61* (1958), Indagationes Mathematicae *20* (1958) 38-46 and 47-55.

[4] VAN EEDEN, Constance: Testing and estimating ordered parameters of probability distributions, thesis, Amsterdam (1958).

[5] HEMELRIJK, J. : Distributionfree tests against trend and maximum likelihood estimates of ordered parameters, I.S.I. congress, Stockholm, 1957.

[6] 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* (1947) 50-60.

[7] WABEKE, Doraline en VAN EEDEN, Constance: Handleiding voor de toets van Wilcoxon, Report S 176 (M 65) of the Statistical Department of the Mathematical Centre, Amsterdam.

RESUME

Des tests pour l'hypothèse $\theta_1 \leqslant \ldots \leqslant \theta_k$ concernant k paramètres $\theta_1, \ldots, \theta_k$ inconnus.

Soient $\theta_1, \ldots, \theta_k$ des paramètres inconnus de k lois de distributions. Le problème, dont une solution est donnée ici, est de tester l'hypothèse

$$\theta_1 \leqslant \ldots \leqslant \theta_k$$

contre les hypothèses alternatives qu'il y a au moins un pair (θ_i, θ_j) avec $i < j$ *et*

$$\theta_i > \theta_j.$$

Le test se compose d'une série de tests de deux échantillons pour l'hypothèse $\theta_i \leqslant \theta_{i+1}$ $(i = 1, \ldots, k-1)$. Le type de ces tests pour deux échantillons dépend de l'information disponible sur la forme des lois de distribution dont les échantillons ont été prélevés.