

STICHTING
MATHEMATISCH CENTRUM
2e BOERHAAVESTRAAT 49
AMSTERDAM

ZW 1958 - 019

Voordracht in de serie "Actualiteiten"

29 november 1958

A class of tests for the hypothesis that k parameters
 $\theta_1, \dots, \theta_k$ satisfy the inequalities $\theta_1 \leq \dots \leq \theta_k$

Dr. Constance van Eeden



1958

Voordracht in de serie
"Actualiteiten"

29 november 1958

A class of tests for the hypothesis that k parameters
 $\theta_1, \dots, \theta_k$ satisfy the inequalities $\theta_1 \leq \dots \leq \theta_k$

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 $\theta_1, \dots, \theta_k$ satisfy the inequalities

$$(1;1) \quad \theta_1 \leq \dots \leq \theta_k$$

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 $\theta_1, \dots, \theta_k$, known to satisfy

$$(1.2) \quad \begin{cases} 1. & \text{inequalities of the type: } \varphi_i(\theta_i) \leq \varphi_j(\theta_j), \\ 2. & \text{inequalities of the type: } c_i \leq \varphi_i(\theta_i) \leq d_i, \end{cases}$$

where, for each $i=1, \dots, k$, $\varphi_i(\theta_i)$ 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 $\theta_1, \dots, \theta_k$, known to satisfy the equalities $\theta_1 \leq \dots \leq \theta_k$.

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.

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 $\underline{x}_1, \dots, \underline{x}_k$ ¹⁾ be k independent random variables and let, for each $i=1, \dots, k$, $x_{i,j}$ ($j=1, \dots, n_i$) be n_i independent observations of \underline{x}_i . Let further, for each $i=1, \dots, k$, θ_i denote an unknown parameter of the distribution of \underline{x}_i .

The hypothesis

$$(2;1) \quad H_0 : \theta_1 \leq \dots \leq \theta_k$$

will be tested against the alternative hypothesis

$$(2;2) \quad H : \text{at least one value of } i \text{ exists with } \theta_i > \theta_{i+1}.$$

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

$$(2;3) \quad H_{0,i} : \theta_i \leq \theta_{i+1}$$

against the alternative hypothesis

$$(2;4) \quad H_i : \theta_i > \theta_{i+1}.$$

Let, for each $i=1, \dots, k-1$, t_i denote the test statistic and Z_i the critical region of this test. Then t_i is a function of $x_{i,1}, \dots, x_{i,n_i}, x_{i+1,1}, \dots, x_{i+1,n_{i+1}}$, and $H_{0,i}$ is rejected if and only if $t_i \in Z_i$.

The test for the hypothesis H_0 then consists of rejecting H_0 if and only if a value of i exists with $t_i \in Z_i$.

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

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

$$(2;5) \quad \begin{cases} \alpha_i \stackrel{\text{def}}{=} P\{\underline{t}_{-i} \in Z_i | \theta_i = \theta_{i+1}\}, & 2) \\ N_i \stackrel{\text{def}}{=} n_i + n_{i+1} \end{cases}$$

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

$$(2;6) \quad \begin{cases} \lim_{N_i \rightarrow \infty} n_i = \infty, \\ \lim_{N_i \rightarrow \infty} n_{i+1} = \infty, \end{cases}$$

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

$$(2;7) \quad \begin{cases} 1. P\{\underline{t}_{-i} \in Z_i | \theta_i < \theta_{i+1}\} \leq \alpha_i, \\ 2. \lim_{N_i \rightarrow \infty} P\{\underline{t}_{-i} \in Z_i | \theta_i < \theta_{i+1}\} = 0, \\ 3. \lim_{N_i \rightarrow \infty} P\{\underline{t}_{-i} \in Z_i | \theta_i > \theta_{i+1}\} = 1. \end{cases}$$

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 $n \rightarrow \infty$ be taken under the conditions

$$(2;9) \quad \lim_{n \rightarrow \infty} n_i = \infty \quad \text{for each } i=1, \dots, k,$$

then we have

$$(2;10) \quad \begin{cases} 1. \alpha_0 \leq \sum_{i=1}^{k-1} \alpha_i, \\ 2. \text{ the probability of rejecting } H_0, \text{ under the hypothesis } \theta_1 < \dots < \theta_k, \text{ tends to zero for } n \rightarrow \infty, \\ 3. \text{ the probability of rejecting } H_0, \text{ under the hypothesis } H \text{ tends to 1 for } n \rightarrow \infty. \end{cases}$$

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

$$(2;11) \quad P\{\underline{t}_{-i} \in Z_i \text{ and } \underline{t}_{-j} \in Z_j | \theta_i = \theta_{i+1}, \theta_j = \theta_{j+1}\} \leq \\ \leq P\{\underline{t}_{-i} \in Z_i | \theta_i = \theta_{i+1}\} \cdot P\{\underline{t}_{-j} \in Z_j | \theta_j = \theta_{j+1}\},$$

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

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

$$(2;12) \quad \left\{ \begin{array}{l} \text{the probability of rejecting } H_0, \text{ under the hypothesis} \\ \theta_1 = \dots = \theta_k, \text{ is } \geq \sum_{i=1}^{k-1} \alpha_i - \frac{1}{2} \left\{ \sum_{i=1}^{k-1} \alpha_i \right\}^2 . \end{array} \right.$$

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 $\leq 0,05$,
2. the probability of rejecting H_0 , under the hypothesis $\theta_1 = \dots = \theta_k$, is $\geq 0,05 - \frac{1}{2}(0,05)^2 = 0,04875$.

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 $i=1, \dots, k$, an exponential distribution with parameter θ_i , i.e.

$$(3.1;1) \quad P\{\underline{x}_i \leq x\} = 1 - e^{-\theta_i x} \quad (x \geq 0) .$$

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

$$(3.1;2) \quad \bar{x}_i \stackrel{\text{def}}{=} \frac{1}{n_i} \sum_{j=1}^{n_i} x_{i,j}$$

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

$$(3.1;3) \quad t_i = \frac{\bar{x}_{i+1}}{\bar{x}_i}$$

and for Z_i we take a critical region of the form $t_i \geq t_{i, \alpha_i}$ where (cf. (2;5)) t_{i, α_i} satisfies

$$(3.1;4) \quad P\{t_i \geq t_{i, \alpha_i} \mid \theta_i = \theta_{i+1}\} = \alpha_i .$$

Now (3.1;1) entails that, for each $i=1, \dots, k$, $2\theta_i n_i \bar{x}_i$ possesses a χ^2 -distribution with $2n_i$ degrees of freedom, thus t_i possesses, for each $i=1, \dots, k-1$, 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, \dots, T_{k-1} satisfy the conditions (2;7) and (2;11).

3.2 A normal distribution with variance θ_i

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

$$(3.2;1) \quad \begin{cases} \bar{x}_i \stackrel{\text{def}}{=} \sum_{j=1}^{n_i} x_{i,j} , \\ s_i^2 \stackrel{\text{def}}{=} \frac{1}{n_i-1} \sum_{j=1}^{n_i} (x_{i,j} - \bar{x}_i)^2 , \end{cases} \quad (i=1, \dots, k)$$

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

$$(3.2;2) \quad t_i = \frac{s_i^2}{s_{i+1}^2} \quad (i=1, \dots, k-1) .$$

Now $\frac{(n_i-1)s_i^2}{\theta_i}$ possesses, for each $i=1, \dots, k$, a χ^2 -distribution with n_i-1 degrees of freedom; thus, for each $i=1, \dots, k-1$, 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 $t_i \geq t_{i,\alpha_i}$, 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 s_i^2 is replaced by $s_i'^2 \stackrel{\text{def}}{=} \frac{1}{n_i} \sum_{j=1}^{n_i} (x_{i,j} - \mu_i)^2$, where $\frac{n_i s_i'^2}{\theta_i}$ 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 $i=1, \dots, k$, \underline{x}_i possesses a normal distribution with mean θ_i and known variance σ_i^2 . Let, for each $i=1, \dots, k$,

$$(3.3;1) \quad \bar{x}_i \stackrel{\text{def}}{=} \frac{1}{n_i} \sum_{j=1}^{n_i} x_{i,j}$$

then we take

$$(3.3;2) \quad t_i = \bar{x}_{i+1} - \bar{x}_i \quad (i=1, \dots, k-1) .$$

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

$$(3.3;3) \quad \sigma^2(\underline{t}_i | \theta_i = \theta_{i+1}) = \frac{\sigma_i^2}{n_i} + \frac{\sigma_{i+1}^2}{n_{i+1}} \quad (i=1, \dots, k-1)$$

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

$$(3.3;4) \quad t_{i, \alpha_i} = \xi_{\alpha_i} \sqrt{\frac{\sigma_i^2}{n_i} + \frac{\sigma_{i+1}^2}{n_{i+1}}},$$

where ξ_{α} is defined by

$$(3.3;5) \quad \frac{1}{\sqrt{2\pi}} \int_{\xi_{\alpha}}^{\infty} e^{-\frac{1}{2}x^2} dx = \alpha.$$

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 correlation coefficient 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 $i=1, \dots, k$, x_i possess a rectangular distribution between 0 and $\theta_i > 0$. Let, for each $i=1, \dots, k$,

$$(3.4;1) \quad z_i \stackrel{\text{def}}{=} \max_{1 \leq j \leq n_i} x_{i,j},$$

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

$$(3.4;2) \quad t_i = \frac{z_i}{z_{i+1}}$$

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

Now we have (cf. [4])

$$(3.4;3) \quad t_{i, \alpha_i} = \begin{cases} \left(\frac{n_i}{N_i \alpha_i} \right)^{\frac{1}{n_{i+1}}} & \text{if } \alpha_i \leq \frac{n_i}{N_i}, \\ \left\{ \frac{N_i}{n_{i+1}} (1 - \alpha_i) \right\}^{\frac{1}{n_i}} & \text{if } \alpha_i \geq \frac{n_i}{N_i}. \end{cases}$$

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 $\underline{x}_1, \dots, \underline{x}_k$ be independent random variables, possessing continuous probability distributions. Let further, for each $i=1, \dots, k$, $x_{i,1}, \dots, x_{i,n_i}$ be independent observations of \underline{x}_i and let (cf. [1])

$$(3.5;1) \quad W_i \stackrel{\text{def}}{=} \sum_{j=1}^{n_i} \sum_{\lambda=1}^{n_{i+1}} \text{sgn}(x_{i,j} - x_{i+1,\lambda}) \quad 3)$$

where

$$(3.5;2) \quad \text{sgn } z \stackrel{\text{def}}{=} \begin{cases} 1 & \text{if } z > 0, \\ 0 & \text{if } z = 0, \\ -1 & \text{if } z < 0. \end{cases}$$

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

$$(3.5;3) \quad P\{\underline{W}_i \in Z'_i | H'_{0,i}\} = P\{W_i \geq W_{i,\alpha_i} | H'_{0,i}\} = \alpha_i.$$

Then the hypothesis H'_0 is rejected if and only if a value of i exists with $W_i \in Z'_i$.

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 distribution¹ 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

$$(3.5;4) \quad \sigma^2(\underline{W}_i | H'_{0,i}) = \frac{1}{3} n_i n_{i+1} (N_i + 1).$$

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

$$(3.5;5) \quad \alpha_0 \stackrel{\text{def}}{=} P\{\underline{W}_i \in Z'_i \text{ for at least one value of } i | H'_0\}$$

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

$$(3.5;6) \quad \sum_{i=1}^{k-1} \alpha_i - \frac{1}{2} \left\{ \sum_{i=1}^{k-1} \alpha_i \right\}^2 \leq \alpha_0 \leq \sum_{i=1}^{k-1} \alpha_i .$$

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

$$(3.5;7) \quad \theta_i' \stackrel{\text{def}}{=} P\{\underline{x}_i > \underline{x}_{i+1}\} \quad (i=1, \dots, k-1) ,$$

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

$$(3.5;8) \quad \lim_{n \rightarrow \infty} n_i = \infty \quad \text{for each } i=1, \dots, k$$

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

$$(3.5;9) \quad \left\{ \begin{array}{l} 1. H'_1 : \text{for each value of } i : \theta_i' < \frac{1}{2} , \\ 2. H'_2 : \text{at least one value of } i \text{ exists with } \theta_i' > \frac{1}{2} , \\ 3. H'_3 : \left\{ \begin{array}{l} \text{for each value of } i : \theta_i' \leq \frac{1}{2} , \\ \text{at least one value of } i \text{ exists with } \theta_i' = \frac{1}{2} . \end{array} \right. \end{array} \right.$$

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

$$(3.5;10) \quad \left\{ \begin{array}{l} 1. \text{ the probability of rejecting } H'_0 \text{ under the hypothesis } H'_1, \text{ tends to zero,} \\ 2. \text{ the probability of rejecting } H'_0 \text{ under the hypothesis } H'_2, \text{ tends to 1,} \\ 3. \text{ if } \alpha_i \text{ is sufficiently small for each value of } i \text{ with } \theta_i' = \frac{1}{2}, \text{ the probability of rejecting } H'_0 \text{ under the hypothesis } H'_3 \text{ tends to a limit } < 1. \end{array} \right.$$

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 CONSTANCE VAN EEDEN, Handleiding voor de toets van Wilcoxon, Report S 176 (M 65) of the Statistical Department of the Mathematical Centre, Amsterdam.