Printed at the Mathematical Centre, 49, 2e Boerhaavestraat, Amsterdam.

The Mathematical Centre, founded the 11-th of February 1946, is a non-profit institution aiming at the promotion of pure mathematics and its applications. It is sponsored by the Netherlands Government through the Netherlands Organization for the Advancement of Pure Research (Z.W.O).

MATHEMATICAL CENTRE TRACTS 15

R. DOORNBOS

SLIPPAGE TESTS

SECOND EDITION

MATHEMATISCH CENTRUM AMSTERDAM 1976

AMS(MOS) subject classification scheme (1970): 62F05, 62G10

ISBN 90 6196 021 5

First printing 1966
Second edition 1976

CONTENTS

Chapter 1

HISTORICAL SURVEY

In this survey only the landmarks in the study of the slippage problem are presented. It does not aim at a complete description of all studies on this subject. Consequently the list of cited literature shows only a fraction of all the papers written in this field.

The necessity to decide whether one or more apparently aberrant observations come from a different population than that generating the other ones caused an interest in slippage tests as long as a century ago.

The problem is of a rather complicated nature. If $\underline{x}_1, \ldots, \underline{x}_n$ *) represent a series of observations of a random variable $\underline{x}$, the problem, in its most simple (one-sided) form, is to test whether the largest of these values is drawn from the same population as the other ones. The distribution may be fully specified or not. The case which has received most attention is that of the normal distribution with unknown mean and standard deviation.

The difficulty is that the index i of the outlying observation $\underline{x}_i$ is not known beforehand. Therefore it is much easier to find a good test statistic on intuitive grounds than to derive its exact distribution.

The first writer on the subject seems to have been BENJAMIN PEIRCE (1852) in the Astronomical Journal. He developed an outlier criterion and applied it to an example, viz. the case of "fifteen observations of the vertical semidiameters of Venus, made by Lieut. HERNDON, with the meridian circle at Washington, in the year 1846". This example has since become a classic as almost every writer refers to it and applies his test on it, for instance, about a century later, F.E. GRUBBS (1950).

---

*) Random variables will be denoted by underlined symbols

The criterion of PEIRCE presupposes the normal distribution whereas further the sample standard deviation is used where the test requires accurate knowledge of the population standard deviation σ. The following description is in the words of its author.

"It is proposed to determine in a series of m observations the limit of error, beyond which all observations involving so great an error may be rejected, provided there are as many as n such observations. The principle upon which it is proposed to solve this problem is, that the proposed observations should be rejected when the probability obtained by retaining them is less than that of the system of errors obtained by their rejection multiplied by the probability of making so many and no more, abnormal observations".

This formulation of the problem has clearly some resemblance to that of a likelihood-ratio test. The elaboration however is not obvious. According to W.A. CHAUVENET (1876) "the criterion involves some principles, derived from the theory of probabilities, which may seem obscure to those not familiar with that branch of science". Even for the initiated however the criterion lacks clarity, especially for more than one observation. Moreover it requires rather heavy computations, because the limit of rejection has to be recalculated for each new set of observations.

Another astronomer, E.J. STONE (1867), postulates that for a given class of observations, and for a given observer there exists a number m which expresses the average number of observations which that person makes with one mistake. This number m he calls the *modulus of carelessness*. Then the rejection limit k is defined by:

$$P\left[\left|\frac{\underline{x} - \mu}{\sigma}\right| > k \right] = \psi(k) = \frac{1}{m} , \tag{1.1}$$

where $\underline{x}$ is normally distributed with mean μ and standard deviation σ. Now all observations with a deviation from their mean greater in absolute value than $k\sigma$ are discarded.

We note that this criterion is independent of the number of observations. A fraction $\frac{1}{m}$ of all observations will on the average be rejected under the hypothesis that all observations are sampled from the same distribution.

The criterion of CHAUVENET (1876) is numerically the same as that of STONE with $m = 2n$, where n is the number of observations. The argument of CHAUVENET is that if the number of observations to be expected outside the limit is smaller than $\frac{1}{2}$ "an error of this magnitude will have a greater probability against than for it, and may therefore be rejected". This is true, because if the probability that an arbitrary observation exceeds the limit is $\frac{1}{2n}$, the probability P_n that at least one observation exceeds this limit is (for $n > 1$) smaller than $n \times \frac{1}{2n} = \frac{1}{2}$, whereas the expected number of exceedances is exactly equal to $\frac{1}{2}$. The first mentioned probability P_n is equal to $1 - (1 - \frac{1}{2n})^n$. When n tends to infinity we have $\lim_{n\to\infty} P_n = 1 - \frac{1}{\sqrt{e}} \sim 0.4$. It is not clear whether CHAUVENET has observed that P_n is not equal to 0.5 but apart from this point it is interesting to note that he aims at an error of the first kind with a probability at least approximately equal to 0.5.

F.Y. EDGEWORTH (1887) would, referring to STONE, not use equation (1.1) but

$$\left[1 - \psi(k)\right]^n = 1 - \frac{1}{m}, \tag{1.2}$$

n being the number of observations *). Apparently EDGEWORTH postulates a fixed a priori probability of making at least one mistake, not per observation, but per sample of n observations. This test can of course also be regarded as giving a rejection limit for discarding an observation such that the probability that at least one observation exceeds the limit is equal to $\frac{1}{m}$, without any reference to a postulated probability of making errors.

This probability of making an error of the first kind is exact and moreover EDGEWORTH is the first author who points out that the knowledge of μ and σ from a small sample is inaccurate.

---

*) Actually EDGEWORTH writes (in our notation) $\left[\psi(k)\right]^n = \frac{1}{m}$, but from the context it is clear that he has (1.2) in mind.

He correctly reasons that using estimates for μ and σ from the sample has the effect of widening the rejection limits. On the other hand he asks himself whether the mean and standard deviation should not be computed after excluding the suspected observation. This would result in contracting the limits. For the modern reader it is obvious of course that the latter suggestion is not right because the probability of exceeding the rejection limits should be calculated under the null hypothesis that all observations have the same distribution.

Now there follows a period during which no new contributions appear. Published literature is mainly restricted to criticism on the criteria mentioned before.

A different type of statistic, but also requiring accurate knowledge of the standard deviation, is introduced by J.O. IRWIN (1925). He uses

$$(1.3)\qquad \underline{\lambda}_1 = \frac{\underline{x}_{(n)} - \underline{x}_{(n-1)}}{\sigma} \quad \text{and} \qquad \underline{\lambda}_2 = \frac{\underline{x}_{(n-1)} - \underline{x}_{(n-2)}}{\sigma},$$

where $\underline{x}_{(n)}$, $\underline{x}_{(n-1)}$ and $\underline{x}_{(n-2)}$ are the three largest observations in descending order. He presents tables of the cumulative distributions of these criteria for sample sizes up to 1000.

Like EDGEWORTH, IRWIN points out that in dealing with small samples the standard deviation of the sample is an unreliable measure of the population standard deviation.

In a paper which is mainly dedicated to the distribution of the range of samples from a normal distribution, L.H.C. TIPPET (1925) proposes as a criterion the one-sided analogue to (1.2), viz.

$$(1.4)\qquad \left[1 - \phi(k)\right]^n = 1 - \alpha,$$

where

$$(1.5)\qquad \phi(k) = P\left[\frac{\underline{x} - \mu}{\sigma} > k\right].$$

STUDENT (1927) introduces the range of small samples for testing the significance of outlying observations. He proposes to use this criterion in the case of routine analyses where the standard deviation is known.

In the case where an appreciable fraction of the observations may be affected by abnormal errors, H. JEFFREYS (1932) proposes an alternative method. He supposes two normal distributions, one for each of the two fractions of the sample. He assumes prior distributions for the standard deviations of the two fractions and then applies Bayes' rule. The computations involved for the exact solution are excessive. An approximate solution for the estimate of the true mean appears to be a weighted average of the observations, the weight of an observation being a continuous fraction of its deviation from the sample mean.

An important step further than EDGEWORTH and TIPPET comes A.T. MAC KAY (1935). He still supposes σ to be known but he takes into account the stochastic nature of the sample mean. His statistic is

$$(1.6) \qquad \frac{\underline{x}_{(n)} - \bar{\underline{x}}}{\sigma} .$$

Tables of its distribution were given later by NAIR (1948).

In most practical cases however σ will also be unknown and therefore the necessity arises either to estimate the population standard deviation from the sample involved or from a second independent sample, the socalled "studentisation" *).

---

*) There does not seem to be a general agreement on the meaning of the term "studentisation". Some writers suppose that the estimate s has to be derived from an independent sample (f.i. K.R. NAIR (1948), M. HALPERIN et.al. (1955)). Others speak of studentisation in all cases where σ is replaced by an estimate either from the single sample involved or from an independent source (E.S. PEARSON and C. CHANDRA SEKAR (1936)). H.O. HARTLEY (1943) says that a statistic is "studentized" when the estimate of σ in the denominator is independent from the numerator, like in STUDENT's t-test.

The "studentized" form of MC KAY's criterion, with s calculated from a second sample, was developed by K.R. NAIR (1948). The two-sided version of this test was tabulated by M. HALPERIN et.al. (1955), the one-sided form by H.A. DAVID (1956). The "studentized" range, in the same sense, was tabulated by E.S. PEARSON and H.O. HARTLEY (1943).

In most practical situations however we need the other line of attack, using the information about σ from the only sample available. W.R. THOMPSON (1935) was the first to devise an exact test based on these considerations. He showed that if

$$(1.7) \qquad \underline{T}_i = \frac{\underline{x}_i - \bar{\underline{x}}}{\underline{s}} ,$$

where $\bar{\underline{x}} = \frac{1}{n} \sum_{j=1}^{n} \underline{x}_j$, $\underline{s}^2 = \frac{1}{n} \sum (\underline{x}_j - \bar{\underline{x}})^2$ and $\underline{x}_i$ is an observation selected at random from a sample $\underline{x}_1, \ldots, \underline{x}_n$ from a normal distribution, then

$$(1.8) \qquad \underline{t}_i = \frac{\underline{T}_i \sqrt{n-2}}{\sqrt{n - 1 - \underline{T}_i^2}}$$

has a Student's t-distribution with (n-2) degrees of freedom. This is easily verified by applying the two-sample t-test to $\underline{x}_i$ against the other (n-1) observations. It should be noted however that (1.7) does not yet give us an exact test for one outlying observation, because we need therefore the distribution not of an arbitrary $\underline{T}_i$ but of the $\underline{T}_i$ largest in absolute value.

In an important paper of E.S. PEARSON and C. CHANDRA SEKAR (1936) the criterion of THOMPSON is studied comprehensively. The authors point out that $\underline{T}_{(n)}$, the largest value of the $\underline{T}_i$ has a distribution which is known in the upper tail, where

$$(1.9) \qquad P\left[\underline{T}_{(n)} > k\right] = n\, P\left[\underline{T}_i > k\right],$$

when k is so large that

$$P\left[\underline{T}_i > k \text{ and } \underline{T}_j > k\right] = 0.$$

Therefore PEARSON and CHANDRA SEKAR were able to give a table with a number of percentage points of the distribution of $\underline{T}_{(n)}$.

They also discussed the power of this test and concluded that for the alternative of more than one outlier the criterion will be quite ineffective whereas for one outlier it would be very useful. In 1940 N. ARLEY gave a generalisation of the results of THOMPSON for the following case. The $\underline{x}_1, \ldots, \underline{x}_n$ are independent normal variates with mean values $\underline{x}_i = \sum_{j=1}^{n} a_{ij}p_j$, where the a's are known coefficients and the p_j unknown parameters. He further supposes that the variance of $\underline{x}_i$ is $\sigma_i^2 = \sigma^2/P_i$, where σ^2 is unknown but the weights P_i are known.

The development of test criteria based on the statistic $\underline{T}_{(n)}$ ends for the time being with F.E. GRUBBS (1950) who calculates numerically the exact distribution and presents extensive tables of percentage points of this distribution. At the same time W.J. DIXON (1950) used sampling methods to compare the efficiency of a number of criteria in the case of a normal distribution with one or two outliers. His conclusions in the most important case, viz. σ unknown, are that the criteria

$$(1.10) \qquad \underline{r}_{10} = \frac{\underline{x}_{(n)} - \underline{x}_{(n-1)}}{\underline{x}_{(n)} - \underline{x}_{(1)}},$$

for small samples,

$$(1.11) \qquad \underline{r}_{21} = \frac{\underline{x}_{(n)} - \underline{x}_{(n-2)}}{\underline{x}_{(n)} - \underline{x}_{(2)}},$$

for sample sizes from 8-13 and

$$(1.12) \qquad \underline{r}_{22} = \frac{\underline{x}_{(n)} - \underline{x}_{(n-2)}}{\underline{x}_{(n)} - \underline{x}_{(3)}},$$

for larger samples, have approximately the same performance as $\underline{T}_{(n)}$, when one outlier is present, and are to be preferred because of the simpler computation. Moreover (1.11) and (1.12) are more efficient when two outliers may be present.

E. PAULSON (1952) generalizes the normal case to k samples of n observations each. He proposes, in order to find outlying samples, a test statistic equivalent to the one that will be discussed in Chapter 2 of this tract and the same approximation to its critical values, without giving exact bounds for the obtained level of significance. The test statistic is

$$(1.13) \qquad \frac{(\underline{\bar{x}}_{(k)} - \underline{\bar{\bar{x}}})}{\sqrt{n \sum_{i=1}^{k} (\underline{\bar{x}}_i - \underline{\bar{\bar{x}}})^2 + \sum_{i,j} (\underline{x}_{ij} - \underline{\bar{x}}_i)^2}},$$

where $\underline{\bar{x}}_i$ is the mean value of the i-th sample and $\underline{\bar{x}}_{(k)}$ is the largest mean, whereas $\underline{\bar{\bar{x}}}$ is the mean of all k n observations. Moreover PAULSON proves that the obtained solution has certain optimum properties when the slippage problem is formulated as a multiple decision problem in maximising the probability of making the correct decision when one of the k categories has slipped to one side.

In an important paper A. KUDO (1956) gives a generalisation in another direction. He considers the case where we have three groups of observations

$$(1.14) \quad \begin{cases} \underline{x}_i^{(1)} & (i = 1,\dots,N_1), \text{ distributed as } N(m_i,\sigma^2) \\ \underline{x}_i^{(2)} & (i = 1,\dots,N_2), \text{ distributed as } N(m^{(2)},\sigma^2) \\ \text{and } \underline{x}_i^{(3)} & (i = 1,\dots,N_3), \text{ distributed as } N(m^{(3)},\sigma^2), \end{cases}$$

where m_i $(i = 1,\dots,N_1)$, $m^{(2)}$, $m^{(3)}$ and σ are unknown. The null hypothesis is:

$$H_0\colon m_1 = m_2 = \dots = m_{N_1} = m^{(2)},$$

against the alternatives

$$H_i\colon m_1 = \dots = m_{i-1} = m_{i+1} = \dots = m_{N_1} = m^{(2)} = m_i - \Delta,\ \Delta > 0.$$

He shows that the following test statistic has certain optimum properties:

$$\frac{\underline{x}_M - \bar{\underline{x}}}{\underline{s}} , \qquad (1.15)$$

where

$$\underline{x}_M = \max \underline{x}_j^{(1)} \quad (j = 1,\ldots,N_1),$$

$$\bar{\underline{x}} = \left(\sum_1^{N_1} \underline{x}_i^{(1)} + \sum_1^{N_2} \underline{x}_i^{(2)}\right)/(N_1 + N_2),$$

$$\bar{\underline{x}}^{(3)} = \sum_1^{N_3} \underline{x}_i^{(3)}/N_3,$$

$$\underline{s}^2 = \left[\sum_1^{N_1} (\underline{x}_i^{(1)} - \bar{\underline{x}})^2 + \sum_1^{N_2} (\underline{x}_i^{(2)} - \bar{\underline{x}})^2 + \sum_1^{N_3} (\underline{x}_i^{(3)} - \bar{\underline{x}}^{(3)})^2\right]/(N_1+N_2+N_3).$$

Clearly the test statistic of PEARSON and CHANDRA SEKAR forms the special case for $N_2 = N_3 = 0$. Also the PAULSON test is equivalent to the special case $N_2 = 0$, because it is not essential whether the extra information about the variance of the population comes from an independent sample like here, or from the within sample variance as in the PAULSON case.

Recently C.P. QUESENBERRY and H.A. DAVID (1961) developed a method for computing percentage points of the statistic (1.15) for $N_2 = 0$.

Remarkably little has been published on non-normal variates. Closely related to the normal case are the slippage tests for variances. In the paper of W.C. COCHRAN (1941) the largest sample variance of a set divided by the sum of all variances is used as a test statistic:

$$\underline{a} = \frac{\underline{s}_{max}^2}{\sum_1^k \underline{s}_i^2} . \qquad (1.16)$$

Independently from Cochran the same test was described by C. CHANDRA SEKAR and M.G. FRANCIS (1941). They give only critical values of $\underline{a}$ larger that $\frac{1}{2}$, because in this domain the distribution of $\underline{a}$ can easily be evaluated.

COCHRAN uses the same approximation to the critical values that will be used throughout this tract.

R. DOORNBOS and H.J. PRINS (1956) generalized COCHRAN's test to different sample sizes and gave approximate power functions with respect to the alternative hypothesis that one of the variances has slipped to the right or to the left.

C.I. BLISS, W.G. COCHRAN and J.W. TUKEY (1956) consider a closely related case, with equal sample sizes. Their test criterion is the largest range divided by the sum of the ranges. They have an alternative hypothesis in mind, however, which is different from the one considered by COCHRAN. Their alternative is that one of the samples contains an outlying observation.

D.A. DARLING (1952) gives a general method for handling the problem of outlying observations. He obtains an integral form for the cumulative distribution of

$$\underline{z}_{(n)} = \frac{\sum_{1}^{n} \underline{x}_i}{\underline{x}_{(n)}}, \tag{1.17}$$

where the n observations $\underline{x}_i$ are independent and positive-valued variates with a fully specified distribution, which is the same for all i.

In R. DOORNBOS and H.J. PRINS (1958) a number of tests are given for various distributions that will further be dealt with in Chapter 2.

Chapter 2

SLIPPAGE TESTS FOR ONE OUTLIER

2.1. INTRODUCTION

This chapter is largely based on a series of papers by the author and H.J. PRINS (1958). The same method was applied for the first time by W.G. COCHRAN (1941) in the case of estimated normal variances. The same principle was also applied by E. PAULSON (1952) and by H.A. DAVID (1956). These authors indicated the same lower limit for the level of significance, but the first proof of its correctness was given by R. DOORNBOS and H.J. PRINS (1958).

The tests are of the following general type.
Suppose

$$\vec{\underline{y}}_1, \ldots, \vec{\underline{y}}_k$$

are k random vectors. Thus

$$\vec{\underline{y}}_i = (\underline{y}_{i1}, \ldots \underline{y}_{in_i}), \ (i = 1,\ldots,k).$$

The variates $\underline{y}_{ih}$ are distributed independently and have all the same type of distribution function. These distribution functions contain an unknown parameter θ_i and possibly other unknown parameters as well. The test serves to decide whether one of the θ_i has slipped. The simultaneous distribution function of the $\underline{y}_{ih}$ is

$$F(\vec{y}_1, \ldots, \vec{y}_k; \vec{\theta}, \vec{\theta}'),$$

where

$$\vec{\theta} = (\theta_1, \ldots, \theta_k)$$

and $\vec{\theta}'$ is the vector of the other unknown parameters.

We want to test

$$H_0: \theta_1 = \ldots = \theta_k$$

against the alternatives

H_{1i}: θ_i slipped to the right (i unknown)

or against the alternatives

H_{2i}: θ_i slipped to the left (i unknown)

or, in the two-sided case:

H_{3i}: θ_i slipped to the right or to the left (i unknown).

In order to get rid of the unknown parameters θ', in all but the distributionfree cases sufficient estimates are used. This implies using new, one-dimensional variates, which are functions of the original variates and which have a simultaneous distribution (in the discrete case a conditional distribution) which does not contain the parameters θ' and, when H_0 is true, not the parameters θ_i either.

We state the test criterion in terms of the new variates

(2.1.1) $$\underline{x}_1, \ldots, \underline{x}_k,$$

which are, under H_0, the hypothesis tested, distributed simultaneously with a known distribution function $F(x_1, \ldots, x_k)$, which may be continuous or not.

Suppose the observed values of $\underline{x}_1, \ldots, \underline{x}_k$ are $x_1, \ldots, x_k$ respectively. When testing against slippage to the right we determine the right hand tail probabilities

(2.1.2) $$d_i \overset{def}{=} P\left[\underline{x}_i \geq x_i\right], \ (i = 1,\ldots,k).$$ *)

We reject H_0 and decide that the m-th population has slipped to the right if

(2.1.3) $$D = \min_i d_i \leq \alpha/k.$$

---

*) The symbol $\overset{def}{=}$ denotes an equality, defining the left hand member.

Testing against slippage to the left requires computing

$$e_i \stackrel{\text{def}}{=} P\left[\underline{x}_i \le x_i\right], \quad (i = 1,\ldots,k). \tag{2.1.4}$$

Now H_0 is rejected and it is concluded that the m-th population has slipped to the left if

$$E = \min_i e_i \le \alpha/k. \tag{2.1.5}$$

A two-sided test is obtained when H_0 is rejected if

$$\min (D,E) \le \alpha/2k.$$

The probability that an error of the first kind occurs when this procedure is applied, is derived along the following general lines. Consider a set of k real numbers $g_1, \ldots, g_k$ and the probabilities defined by

$$\left\{\begin{aligned} p_i &\stackrel{\text{def}}{=} P\left[\underline{x}_i \le g_i\right], \\ p_{i,j} &\stackrel{\text{def}}{=} P\left[\underline{x}_i \le g_i \text{ and } \underline{x}_j \le g_j\right], \quad (i \ne j) \\ q_i &\stackrel{\text{def}}{=} P\left[\underline{x}_i > g_i\right], \\ q_{i,j} &\stackrel{\text{def}}{=} P\left[\underline{x}_i > g_i \text{ and } \underline{x}_j > g_j\right], \quad (i \ne j) \end{aligned}\right. \tag{2.1.6}$$

all computed under H_0.

Denoting by P the probability that at least one of the $\underline{x}_i$ does not exceed the corresponding value g_i, it follows from BONFERRONI's inequality (cf. W. FELLER (1950), chapter 4) that

$$\sum_i p_i - \sum_{i<j} p_{i,j} \le P \le \sum_i p_i. \tag{2.1.7}$$

For Q, the probability that at least one $\underline{x}_i$ exceeds g_i, we have

(2.1.8) $$\sum_i q_i - \sum_{i<j} q_{i,j} \leq Q \leq \sum_i q_i.$$

Then, in each case separately, we proceed to prove the inequality

(2.1.9) $$p_{i,j} \leq p_i\, p_j$$

or

(2.1.10) $$q_{i,j} \leq q_i\, q_j.$$

It is easily seen that (2.1.9) and (2.1.10) are equivalent. We have

$$p_i = 1 - q_i \quad \text{and} \quad p_j = 1 - q_j$$

and consequently

(2.1.11) $$p_i(1 - p_j) = q_j(1 - q_i).$$

Further

(2.1.12) $$p_i - p_{i,j} = q_j - q_{i,j} \;(= P\left[\underline{x}_i \leq g_i \text{ and } \underline{x}_j > g_j\right]).$$

From (2.1.11) and (2.1.12) we obtain

(2.1.13) $$p_i\, p_j - p_{i,j} = q_i\, q_j - q_{i,j},$$

which proves the equivalence of (2.1.9) and (2.1.10). Assuming that (2.1.9) and (2.1.10) are true we get immediately from (2.1.7) and (2.1.8) the inequalities

(2.1.14) $$\sum_i p_i - \sum_{i<j} p_i\, p_j \leq P \leq \sum_i p_i$$

and

(2.1.15) $$\sum_i q_i - \sum_{i<j} q_i\, q_j \leq Q \leq \sum_i q_i$$

respectively. Denoting $\sum_i p_i$ by p (p is not necessarily ≤ 1), we have

$$p^2 = (\sum_i p_i)^2 = 2 \sum_{i<j} p_i\, p_j + \sum_i p_i^2 \geq 2 \sum_{i<j} p_i\, p_j,$$

or

$$\sum_{i<j} p_i \, p_j \leq \tfrac{1}{2} \, p^2 .$$

Thus from (2.1.14) follows

$$p - \tfrac{1}{2} p^2 \leq P \leq p \tag{2.1.16}$$

and similarly from (2.1.15)

$$q - \tfrac{1}{2} q^2 \leq Q \leq q, \tag{2.1.17}$$

where $\sum_i q_i = q$.

Now, when testing H_0 against slippage to the left of one of the k variates the critical region is of the form

$$\{x_1 \leq g_{1\alpha} \text{ or } \ldots \text{ or } x_k \leq g_{k\alpha}\},$$

where the values $g_{i\alpha}$ are determined such that all p_i are equal to α/k where α is the prescribed level of significance. In the discontinuous case this will in géneral not be possible; there $g_{i\alpha}$ is the largest value which can be attained by $\underline{x}_i$ with a positive probability satisfying

$$\alpha_i' \overset{\text{def}}{=} P\left[\underline{x}_i \leq g_{i\alpha}\right] \leq \alpha/k. \tag{2.1.18}$$

So from (2.1.16) it follows that the probability P of rejecting H_0 when H_0 is true, satisfies

$$\alpha - \alpha^2/2 \leq P_\alpha \leq \alpha \tag{2.1.19}$$

or

$$\alpha' - (\alpha')^2/2 \leq P_\alpha \leq \alpha' \quad (\alpha' \overset{\text{def}}{=} \sum_i \alpha_i') \tag{2.1.20}$$

respectively, according to whether the continuous or the discrete case is considered.

<u>REMARK 2.1.1</u>

We will always have $k \geq 3$, because for $k = 2$ our problem reduces to the two sample case. Further α will be relatively small, say

$\alpha \leq 0.30$. Therefore in a practical case we can assume that $\alpha/k \leq 0.10$. Therefore it will not be necessary to prove (2.1.9) for all values g_i and g_j. In fact in the case of the slippage test for normal distributions, which will be described in section 2.3 (2.1.9) will only hold when either both p_i and p_j are ≤ 0.50 or both ≥ 0.50, which means that q_i and q_j are both ≤ 0.50.

REMARK 2.1.2

When (2.1.9) and (2.1.10) do not hold we can still apply the same method to obtain a slippage test, but in this case we will only know that $P_\alpha \leq \alpha$ and we cannot give a lower bound for P_α. In this case only the second half of the BONFERRONI inequality (2.1.7) is used.

When testing against slippage to the right we get similar bounds for the probability Q_α of rejecting H_0 when H_0 is true.

In the two-sided case we have only an upper bound for the probability R_α of an error of the first kind, because (cf. remark 2.1.2)

(2.1.21) $$R_\alpha \leq P_{\frac{1}{2}\alpha} + Q_{\frac{1}{2}\alpha} \leq \alpha.$$

There is of course a trivial lower bound

(2.1.22) $$R_\alpha \geq \frac{1}{2}\alpha - \frac{1}{8}\alpha^2,$$

because both $P_{\frac{1}{2}\alpha}$ and $Q_{\frac{1}{2}\alpha}$ are $\geq \frac{1}{2}\alpha - \frac{1}{8}\alpha^2$ in virtue of (2.1.9).

2.2. A GENERAL CONDITION FOR THE VALIDITY OF THE INEQUALITY (2.1.9) FOR CONTINUOUS DISTRIBUTIONS

In this section we prove the following theorem:

THEOREM 2.2.1

Suppose the random variables $\underline{x}$ and $\underline{y}$ have a joint distribution which is given by their density function $f(x,y)$. Now the inequality

(2.2.1) $$P\left[\underline{x} \le a \quad \text{and} \quad \underline{y} \le b\right] \le P\left[\underline{x} \le a\right] \cdot P\left[\underline{y} \le b\right]$$

holds if

(2.2.2) $$f(x_1,y_1)\cdot f(x_2,y_2) \le f(x_1,y_2)\cdot f(x_2,y_1),$$

for $$x_1 \le a \le x_2 \quad \text{and} \quad y_1 \le b \le y_2.$$

PROOF

From (2.2.2) it follows that

$$\int_{x_1=-\infty}^{a} dx_1 \int_{y_1=-\infty}^{b} dy_1 \int_{x_2=a}^{\infty} dx_2 \int_{y_2=b}^{\infty} dy_2 \left[f(x_1,y_1)f(x_2,y_2)-f(x_1,y_2)f(x_2,y_1)\right] \le 0.$$

Or:

(2.2.3) $$P\left[\underline{x} \le a \quad \text{and} \quad \underline{y} \le b\right] \cdot P\left[\underline{x} \ge a \quad \text{and} \quad \underline{y} \ge b\right] \le$$
$$P\left[\underline{x} \le a \quad \text{and} \quad \underline{y} \ge b\right] \cdot P\left[\underline{x} \ge a \quad \text{and} \quad \underline{y} \le b\right].$$

Or:

(2.2.4) $P_3P_1 \le P_2P_4$, say (cf. fig.(2.2.1)).

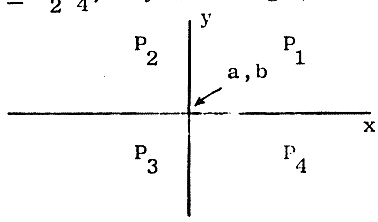

Figure (2.2.1)

Or:

(2.2.5) $$P_3(1 - P_2 - P_3 - P_4) \le P_2P_4,$$

which gives

(2.2.6) $$P_3 \le P_2P_3 + P_3^2 + P_3P_4 + P_2P_4 = (P_2 + P_3)(P_3 + P_4),$$

which is the same as (2.2.1).

REMARK 2.2.1

The condition (2.2.2) is certainly satisfied in the special case where $\frac{\partial^2 \log f(x,y)}{\partial x \partial y}$ exists and is non-positive everywhere inside the rectangle $x_1 \leq x \leq x_2$; $y_1 \leq y \leq y_2$.

For (2.2.2) says

$$\frac{f(x_1,y_1)}{f(x_2,y_1)} \leq \frac{f(x_1,y_2)}{f(x_2,y_2)}, \tag{2.2.7}$$

which holds if

$$\frac{\partial}{\partial y} \frac{f(x_1,y)}{f(x_2,y)} \geq 0 \quad \text{if} \quad y_1 \leq y \leq y_2 \tag{2.2.8}$$

or

$$\{\frac{\partial}{\partial y} f(x_1,y)\} f(x_2,y) - f(x_1,y) \frac{\partial}{\partial y} f(x_2,y) \geq 0 \text{ if } y_1 \leq y \leq y_2. \tag{2.2.9}$$

Inequality (2.2.9) may be written as

$$\frac{\partial \log f(x_1,y)}{\partial y} \geq \frac{\partial \log f(x_2,y)}{\partial y} \quad \text{if} \quad y_1 \leq y \leq y_2, \tag{2.2.10}$$

which is certainly satisfied if

$$\frac{\partial^2 \log f(x,y)}{\partial x \partial y} \leq 0 \quad \text{for} \quad x_1 \leq x \leq x_2,\ y_1 \leq y \leq y_2. \tag{2.2.11}$$

As a simple illustration we may consider the bivariate normal distribution where the density function has the form

$$f(x,y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_1}{\sigma_1}\right)^2 - 2\rho\frac{x-\mu_1}{\sigma_1}\cdot\frac{y-\mu_2}{\sigma_2} + \left(\frac{y-\mu_2}{\sigma_2}\right)^2\right]}. \tag{2.2.12}$$

Here we have

$$\frac{\partial^2}{\partial x \partial y} \log f(x,y) = \frac{\rho}{\sigma_1\sigma_2(1-\rho^2)}. \tag{2.2.13}$$

Thus inequality (2.2.1) is true if the correlation coefficient is negative. This case of the inequality (2.2.1) was used in H.A. DAVID (1956 [a]) without proof.

2.3. THE SLIPPAGE TEST FOR THE NORMAL DISTRIBUTION

We consider k normal distributions with unknown means $\mu_1, \mu_2, \ldots, \mu_k$ and common unknown variance σ^2. From these distributions we have samples of $n_1, n_2, \ldots, n_k$ observations respectively.

We want to test the hypothesis

$$H_0: \mu_1 = \mu_2 = \ldots = \mu_k = \mu \text{ say,} \tag{2.3.1}$$

against the alternatives

$$H_{1i}: \begin{cases} \mu_1 = \ldots = \mu_{i-1} = \mu_{i+1} = \ldots = \mu_k = \mu \\ \mu_i = \mu + \Delta \ (\Delta > 0), \end{cases} \tag{2.3.2}$$

for one value of i, which, however is not known, or

$$H_{2i}: \begin{cases} \mu_1 = \ldots = \mu_{i-1} = \mu_{i+1} = \ldots = \mu_k = \mu \\ \mu_i = \mu - \Delta \ (\Delta > 0), \end{cases} \tag{2.3.3}$$

for one unknown value of i. From the observations

$$\begin{cases} \underline{y}_{11}, \ldots, \underline{y}_{1n_1} \\ \underline{y}_{21}, \ldots, \underline{y}_{2n_2} \\ \ldots \\ \underline{y}_{k1}, \ldots, \underline{y}_{kn_k} \end{cases} \tag{2.3.4}$$

the variables

$$\underline{b}_i = \sqrt{\frac{n_i N}{N-n_i}} \frac{\underline{y}_i - \underline{y}}{\sqrt{\sum_{i,j} (\underline{y}_{ij} - \underline{y})^2}} \quad (i = 1, \ldots, k) \tag{2.3.5}$$

are formed, where

$$(2.3.6)\qquad \begin{cases} \underline{N} \overset{\text{def}}{=} \sum_i n_i \\ \underline{y}_i \overset{\text{def}}{=} \frac{1}{n_i} \sum_j \underline{y}_{ij} \\ \underline{y} \overset{\text{def}}{=} \frac{1}{N} \sum_{i,j} \underline{y}_{ij} . \end{cases}$$

As was remarked in Chapter 1 this leads in the case $n_1 = \ldots = n_k = n$ to the same test as the one proposed by E. PAULSON (1952) (cf. 1.13).

The $\underline{b}_i$ take the place of the variables $\underline{x}_i$ in (2.1.1). Once inequality (2.1.9) has been proved for the $\underline{b}_i$ we may apply the procedure described in section 2.1. For the purpose of obtaining critical values it is easier however to use the variables

$$(2.3.7)\qquad \underline{t}_i \overset{\text{def}}{=} \underline{b}_i \sqrt{\frac{N-2}{1-\underline{b}_i^2}} .$$

Because $\underline{t}_i$ is a monotone increasing function of $\underline{b}_i$ it is equivalent as a test statistic. Straightforward computation shows that $\underline{t}_i$ can be written as

$$(2.3.8)\qquad \underline{t}_i = \sqrt{\frac{(N-2)n_i(N-n_i)}{N}} \; \frac{\underline{y}_i - \underline{y}^i}{\sqrt{\sum_j (\underline{y}_{ij}-\underline{y}_i)^2 + \sum_{\substack{s,l \\ s\neq i}} (\underline{y}_{sl}-\underline{y}^i)^2}} ,$$

where

$$(2.3.9)\qquad \underline{y}^i \overset{\text{def}}{=} \frac{1}{N-n_i} \sum_{\substack{s,l \\ s\neq i}} \underline{y}_{sl} .$$

It is clear that, when H_0 is true, $\underline{t}_i$ has a Students t-distribution with N-2 degrees of freedom because it is the test statistic of Students two sample test, where one sample consists of $\underline{y}_{i1},\ldots,\underline{y}_{in_i}$ and the other sample of all the $N-n_i$ other observations. The minimum values of d_i and e_i as defined by (2.1.2) and (2.1.4) can then be tested at their significance by means of a table of extreme percentage points of Students t-distribution as the one presented by E.T. FEDERIGHI (1959). The minimum value of d_i corresponds to the largest $\underline{t}_i$, the minimum value of e_i to the smallest $\underline{t}_i$.

2.4. PROOF OF THE INEQUALITY (2.1.9) IN THE NORMAL CASE

In this section we give a proof of the inequality

$$P\left[\underline{b}_i \le g_i \text{ and } \underline{b}_j \le g_j\right] \le P\left[\underline{b}_i \le g_i\right] \cdot P\left[\underline{b}_j \le g_j\right], \tag{2.4.1}$$

provided that g_i and g_j have the same sign. First the simultaneous distribution of $\underline{b}_i$ and $\underline{b}_j$ has to be derived. The first derivation of this distribution was given in R. DOORNBOS, H. KESTEN and H.J. PRINS (1956). C.P. QUESENBERRY and H.A. DAVID (1961) used essentially the same derivation but in a strongly simplified form. Professor HEMELRIJK pointed out the following, more heuristic, approach to the present author. For definiteness we take i = 1 and j = 2 and further we assume that k = 3. This is no restriction on the generality as pooling of the samples 3, ..., k does not affect $\underline{b}_1$ or $\underline{b}_2$.
We put

$$\left\{\begin{array}{l} \underline{S}^2 = \sum_{i=1}^{3} \sum_{j=1}^{n_i} (\underline{y}_{ij} - \underline{y})^2, \\ \\ \underline{S}_1^2 = \sum_{i=1}^{3} \sum_{j=1}^{n_i} (\underline{y}_{ij} - \underline{y}_i)^2. \end{array}\right. \tag{2.4.2}$$

From the analysis of variance we know the following decomposition:

$$\underline{S}^2 = \underline{S}_1^2 + \sum_{i=1}^{3} n_i (\underline{y}_i - \underline{y})^2, \tag{2.4.3}$$

where $\underline{S}_1^2$ and $\sum_i n_i (\underline{y}_i - \underline{y})^2$ are distributed as $\chi^2 \cdot \sigma^2$ with N-3 and 2

degrees of freedom respectively. Moreover these two terms are independent according to COCHRAN's theorem (W.G. COCHRAN (1934)).
From the second term we split off a term proportional to $(\underline{y}_1 - \underline{y})^2$, i.e. the squared numerator of $\underline{b}_1$. The variance of $(\underline{y}_1 - \underline{y})$ is easily calculated as $\sigma^2 \cdot \frac{n_2 + n_3}{n_1 N}$, therefore

$$(2.4.4) \qquad \underline{r}^2 \overset{\text{def}}{=} \frac{n_1 N}{n_2 + n_3} (\underline{y}_1 - \underline{y})^2$$

is distributed as $\chi^2 . \sigma^2$ with 1 degree of freedom.
We find now that

$$(2.4.5) \quad \sum_{i=1}^{3} n_i (\underline{y}_i - \underline{y})^2 = \underline{r}^2 + \frac{n_2(n_2+n_3)}{n_3} (\underline{y}_2 - \underline{y}^1)^2 = \underline{r}^2 + \underline{s}^2, \text{ say,}$$

where $\underline{y}^1$ is defined by (2.3.9) and therefore equal to $(n_2\underline{y}_2 + n_3\underline{y}_3)/(n_2 + n_3)$.

According to COCHRAN's theorem we have now that $\underline{S}_1^2$, $\underline{r}^2$ and $\underline{s}^2$ are stochastically independent and have therefore a simultaneous distribution which is, apart from a constant factor, equal to the product of three χ^2 distributions with N-3, 1 and 1 degrees of freedom respectively:

$$(2.4.6) \quad f_1(S_1^2, r^2, s^2) :: (S_1^2)^{\frac{N-5}{2}} (r^2)^{-\frac{1}{2}} (s^2)^{-\frac{1}{2}} e^{-S^2/2\sigma^2} \; {}^{*)} ,$$

where $S^2 = S_1^2 + r^2 + s^2$.
We now transform the variates $\underline{S}_1^2$, $\underline{r}^2$ and $\underline{s}^2$ into the new ones $\underline{S}_1^2$, $\underline{r}$ and $\underline{s}$ Now also $\underline{r}$ and $\underline{s}$ are independent, as may easily be verified from the defining equations (2.4.4) and (2.4.5), by calculating $\mathcal{E}\underline{r}\,\underline{s}$ which is equal to 0. This proves the independence because $\underline{r}$ and $\underline{s}$ have a simultaneous normal distribution with means 0. And because $d(r^2) = 2r\,dr$ and $d(s^2) = 2s\,ds$ we have immediately:

---

*) The symbol :: means: "but for a constant factor equal to".

$$(2.4.7) \qquad f_2(S_1^2,\ r,\ s)\ ::\ (S_1^2)^{\frac{N-5}{2}} e^{-S^2/2\sigma^2} .$$

Next we introduce the following new variates:

$$\begin{cases} \underline{S}^2 = \underline{S}_1^2 + \underline{r}^2 + \underline{s}^2 \\ \underline{s}_1 = \underline{s}\ (\underline{S}^2)^{-\frac{1}{2}} \\ \underline{b}_1 = \underline{r}\ (\underline{S}^2)^{-\frac{1}{2}} . \end{cases}$$

The Jacobian of this transformation is equal to S^2 and therefore

$$(2.4.8) \qquad f_3(S^2,\ s_1,\ b_1)\ ::\ (S^2)^{\frac{N-3}{2}} (1-s_1^2-b_1^2)^{\frac{N-5}{2}} e^{-S^2/2\sigma^2} ,$$

from which we conclude that $\underline{S}^2$ is stochastically independent of $\underline{s}_1$ and $\underline{b}_1$ simultaneously, and the distribution function of $\underline{b}_1$ and $\underline{s}_1$ reads

$$(2.4.9) \quad f_4(b_1,s_1)\ ::\ (1-s_1^2-b_1^2)^{\frac{N-5}{2}},\ \text{if } 1-s_1^2-b_1^2 \geq 0, \text{ elsewhere } f_4 = 0.$$

Now

$$(2.4.10) \quad \underline{s}_1 = \underline{s}\cdot(\underline{S}^2)^{-\frac{1}{2}} = \sqrt{\frac{n_2(n_2+n_3)}{n_3}}\ (\underline{y}_2 - \underline{y}^1)\cdot(\underline{S}^2)^{-\frac{1}{2}} =$$

$$= \sqrt{\frac{n_2(n_2+n_3)}{n_3}}\ \left\{\frac{n_1}{n_2+n_3}\ (\underline{y}_1 - \underline{y}) + (\underline{y}_2 - \underline{y})\right\}\cdot(\underline{S}^2)^{-\frac{1}{2}} =$$

$$= \sqrt{\frac{n_1 n_2}{n_3\ N}}\ \underline{b}_1 + \sqrt{\frac{(n_1+n_3)(n_2+n_3)}{n_3\ N}}\ \underline{b}_2 .$$

Finally $\underline{b}_1$ and $\underline{s}_1$ are transformed into $\underline{b}_1$ and $\underline{b}_2$. The Jacobian is a constant, thus $f_4(b_1,s_1)$ changes into

$$(2.4.11)\quad f(b_1,b_2) :: \{1 - \frac{(n_1 + n_3)(n_2 + n_3)}{n_3 N} b_1^2 -$$

$$2\frac{\sqrt{n_1 n_2 (n_1 + n_3)(n_2 + n_3)}}{n_3 N} b_1 b_2 - \frac{(n_1 + n_3)(n_2 + n_3)}{n_3 N} b_2^2\}^{\frac{N-5}{2}} ,$$

in the region where the expression between brackets is positive and zero everywhere else.

This region is bounded by an ellipse (cf. fig.(2.4.1)) with principal axes on the lines

$$(2.4.12)\qquad \begin{cases} b_1 + b_2 = 0, \\ b_1 - b_2 = 0. \end{cases}$$

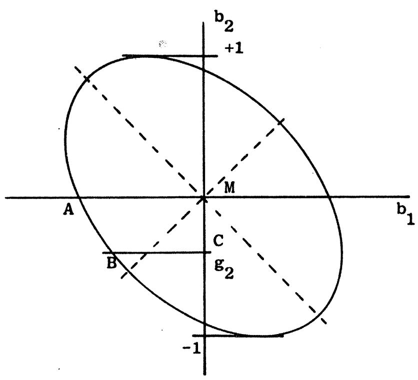

Figure (2.4.1)
The region where $f(b_1,b_2) > 0$.

We now proceed to prove the inequality (2.4.1).

At first sight it seems as if this proof follows easily from theorem (2.2.1). For, as will be shown in the following lemma, condition (2.2.11) is satisfied when b_1 and b_2 are either both ≥ 0 or both ≤ 0 and $f(b_1,b_2) > 0$.

LEMMA 2.4.1

If $N \geq 5$, $b_1 \leq 0$ and $b_2 \leq 0$, or $b_1 \geq 0$ and $b_2 \geq 0$, the inequality

$$\frac{\partial^2 \log f(b_1,b_2)}{\partial b_1 \; \partial b_2} \leq 0,$$

where $f(b_1,b_2)$ is given by 2.4.11, holds.

PROOF

We have

$$f(b_1,b_2) = C\{1-c_1b_1^2-2c_2b_1b_2-c_1b_2^2\}^{\frac{N-5}{2}},$$

where
$$c_1 = \frac{(n_1 + n_3)(n_2 + n_3)}{n_3 \; N}$$

and

$$c_2 = \frac{\sqrt{n_1n_2(n_1+n_3)(n_2+n_3)}}{n_3 \; N} \quad .$$

Therefore

$$\log f(b_1,b_2) = \log C + \frac{N-5}{2} \log \{1-c_1b_1^2-2c_2b_1b_2-c_1b_2^2\}.$$

$$\frac{\partial \log f(b_1,b_2)}{\partial b_1} = \frac{N-5}{2} \; \frac{-2c_1b_1-2c_2b_2}{\{1-c_1b_1^2-2c_2b_1b_2-c_1b_2^2\}} \quad .$$

$$\frac{\partial^2 \log f(b_1,b_2)}{\partial b_1 \partial b_2} = \frac{N-5}{2} \left[\frac{-2c_2}{\{1-c_1b_1^2-2c_2b_1b_2-c_1b_2^2\}}\right.$$

$$\left.\frac{-4(c_1b_1+c_2b_2)(c_2b_1+c_1b_2)}{\{1-c_1b_1^2-2c_2b_1b_2-c_1b_2^2\}^2}\right] = \frac{N-5}{2} \left[\frac{-2c_2}{\{1-c_1b_1^2-2c_2b_1b_2-c_1b_2^2\}}\right.$$

$$\left.\frac{-4\{c_1c_2(b_1+b_2)^2+(c_1-c_2)^2b_1b_2\}}{\{1-c_1b_1^2-2c_2b_1b_2-c_1b_2^2\}^2}\right],$$

which is clearly negative if (but not only if) b_1b_2 remains positive.

This lemma means that condition (2.2.11) is fulfilled in every rectangle which lies completely inside one of the regions ($b_1 \leq 0$, $b_2 \leq 0$) or ($b_1 \geq 0$, $b_2 \geq 0$). But the proof of inequality (2.4.1) is only complete for the point (g_1, g_2), when (2.2.11) holds in every rectangle which contains (g_1, g_2) and this does not follow from the lemma.

Therefore another type of proof is needed in this case. We suppose that both g_1 and g_2 are negative. This is no restriction, for when (2.4.1) holds for a pair of values g_1 and g_2 the inequality

$$P\left[\underline{b}_1 > -g_1 \text{ and } \underline{b}_2 > -g_2\right] \leq P\left[\underline{b}_1 > -g_1\right] \cdot P\left[\underline{b}_2 > -g_2\right]$$

holds as well for reasons of symmetry. Consequently (2.4.1) is also true for $-g_1$ and $-g_2$ because of the equivalence of (2.1.9) and (2.1.10). Further we may assume that the point (g_1, g_2) lies inside the ellipse of fig. (2.4.1), because otherwise $P\left[\underline{b}_1 \leq g_1 \text{ and } \underline{b}_2 \leq g_2\right] = 0$ and (2.4.1) is trivial. We shall prove that in the (g_1, g_2) region considered (2.4.1) holds with the < sign.

The distribution $f_4(b_1, s_1)$ found in (2.4.9), where $\underline{s}_1$ is given by (2.4.10) suggests the following transformation:

$$\text{(2.4.13)} \qquad \underline{b}_2' = \frac{\underline{s}_1}{\sqrt{1-\underline{b}_1^2}} = \frac{\sqrt{\frac{n_1 n_2}{n_3 N}}\,\underline{b}_1 + \sqrt{\frac{(n_1+n_3)(n_2+n_3)}{n_3 N}}\,\underline{b}_2}{\sqrt{1-\underline{b}_1^2}} .$$

This transformation changes $f_4(b_1, s_1)$ into

$$\text{(2.4.14)} \qquad f_5(b_1, b_2') :: (1-b_1^2)^{\frac{N-4}{2}} (1-(b_2')^2)^{\frac{N-5}{2}} .$$

In other words $\underline{b}_1$ and $\underline{b}_2'$ are independent of each other and the distribution of $\underline{b}_1$ is

$$\text{(2.4.15)} \qquad f(b_1) :: (1 - b_1^2)^{\frac{N-4}{2}} .$$

This distribution can of course also directly be obtained from the t-distribution by applying the transformation (2.3.7).
Further we introduce the negative valued function h(t), which is defined for $t \leq 0$ by the property that the points $\{h(b_2), b_2\}$ and $\{b_1, h(b_1)\}$ belong to the ellipse of fig. (2.4.1).
Now we have

$$(2.4.16)\quad P\left[\underline{b}_1 \leq g_1 \text{ and } \underline{b}_2 \leq g_2\right] = \int_{h(g_2)}^{g_1} db_1 \int_{h(b_1)}^{g_2} f(b_1, b_2)\, db_2 =$$

(applying (2.4.13)) =

$$k_1 \int_{h(g_2)}^{g_1} (1-b_1^2)^{\frac{N-4}{2}} db_1 \cdot k_2 \int_{-1}^{b_2'(b_1,g_2)} \{1-(b_2^1)^2\}^{\frac{N-5}{2}} db_2' ,$$

where $b_2'(b_1, g_2)$ stands for

$$(2.4.17)\qquad b_2'(b_1,g_2) = \frac{\sqrt{\dfrac{n_1 n_2}{n_3\, N}}\, b_1 + \sqrt{\dfrac{(n_1+n_3)(n_2+n_3)}{n_3\, N}}\, g_2}{\sqrt{1 - b_1^2}}$$

and k_1 and k_2 are two constants appearing in the distribution of $\underline{b}_1$ and $\underline{b}_2'$ respectively.
We have to prove

$$(2.4.18)\quad \phi(g_1,g_2) \overset{\text{def}}{=} P\left[\underline{b}_1 \leq g_1\right] \cdot P\left[\underline{b}_2 \leq g_2\right] - P\left[\underline{b}_1 \leq g_1 \text{ and } \underline{b}_2 \leq g_2\right] > 0.$$

From (2.4.16) and (2.4.15) we have

$$(2.4.19)\quad \phi(g_1,g_2) = k_1 \int_{-1}^{g_1} (1-b_1^2)^{\frac{N-4}{2}} db_1\, k_1 \int_{-1}^{g_2} (1-b_2^2)^{\frac{N-4}{2}} db_2$$

$$- k_1 \int_{h(g_2)}^{g_1} (1-b_1^2)^{\frac{N-4}{2}} db_1\, k_2 \int_{-1}^{b_2'(b_1,g_2)} \{1-(b_2^1)^2\}^{\frac{N-5}{2}} db_2' .$$

Differentiation with respect to g_1 gives

$$(2.4.20)\qquad \frac{\partial\phi(g_1,g_2)}{\partial g_1} = k_1^2(1-g_1^2)^{\frac{N-4}{2}} \int_{-1}^{g_2} (1-b_2^2)^{\frac{N-4}{2}} db_2$$

$$-k_1(1-g_1^2)^{\frac{N-4}{2}} k_2 \int_{-1}^{b_2'(g_1,g_2)} \{1-(b_2')^2\}^{\frac{N-5}{2}} db_2' =$$

$$= k_1(1-g_1^2)^{\frac{N-4}{2}} \cdot \phi_1(g_1,g_2), \text{ say.}$$

First we remark that

$$(2.4.21)\qquad \phi_1(0,0) = \frac{1}{2} - \frac{1}{2} = 0.$$

Further we see that because

$$(2.4.22)\qquad \frac{\partial b_2'(g_1,g_2)}{\partial g_1} = \frac{\sqrt{\frac{n_1 n_2}{n_3\, N}} + g_1 g_2\sqrt{\frac{(n_1+n_3)(n_2+n_3)}{n_3\, N}}}{(1-g_1^2)^{3/2}} \quad ,$$

which is positive when $g_1 g_2 \geq 0$, that $\phi_1(g_1,g_2)$ is a decreasing function of g_1.

Now we return to fig. (2.4.1). In point A $\phi(g_1,g_2)$ is positive, because $P\left[\underline{b}_1 \leq g_1 \text{ and } \underline{b}_2 \leq g_2\right] = 0$. $\phi_1(g_1,0)$ is a decreasing function of g_1 and $\phi_1(0,0) = 0$, therefore $\phi_1(g_1,0) \geq 0$ on the b_1-axis between A and M.

Because $(1-g_1^2)^{\frac{N-4}{2}}$ is also positive on the segment AM, it follows that $\phi(g_1,g_2)$ is increasing from A to M and therefore > 0 on the negative part of the b_1-axis. For symmetry reasons $\phi(g_1,g_2)$ is also > 0 on the negative part of the b_2-axis.

Therefore $\phi(g_1,g_2) > 0$ in B (where $P\left[\underline{b}_1 \leq g_1 \text{ and } \underline{b}_2 \leq g_2\right] = 0$) and in C (the point $0,g_2$).

Now $\phi_1(g_1,g_2)$ is decreasing in g_1 and $(1-g_1^2)^{\frac{N-4}{2}}$ is positive, therefore between B and C $\frac{\partial\phi(g_1,g_2)}{\partial g_1}$ is everywhere negative, everywhere positive, or positive up to a certain point g_0 (depending on g_2), say, and negative thereafter. Thus $\phi(g_1,g_2)$ is necessarily positive in every point (g_1,g_2) between B and C.

2.5. THE SLIPPAGE TEST FOR THE GAMMA DISTRIBUTION

Suppose we have a set of random variables

$$\underline{u}_1, \ldots, \underline{u}_k \tag{2.5.1}$$

distributed independently of one another according to gamma distributions with parameters ε_1, β_1; ...; ε_k, β_k respectively; that is to say the density function of $\underline{u}_i$ is

$$f_i(u) = \frac{1}{\Gamma(\varepsilon_i)\beta_i^{\varepsilon_i}} u^{\varepsilon_i - 1} e^{-u/\beta_i}, \quad 0 \leq u \leq \infty, \tag{2.5.2}$$

where ε_i and β_i are real positive numbers.

A special case is formed by a set of estimated normal variances, which are, when multiplied by their respective degrees of freedom, distributed as $\underline{u} = \underline{\chi}^2 \cdot \sigma^2$. Now, when $\underline{\chi}^2$ has a chi-squared distribution with ν degrees of freedom, $\underline{u}$ has a gamma distribution with parameters $\varepsilon = \nu/2$ and $\beta = 2\sigma^2$.

We want to test the hypothesis

$$H_0: \beta_1 = \ldots = \beta_k = \beta, \text{ say,} \tag{2.5.3}$$

against the alternatives

$$\begin{cases} H_{1i}: \beta_1 = \ldots = \beta_{i-1} = \beta_{i+1} = \ldots = \beta_k = \beta, \\ \beta_i = C\beta, \; C > 1, \end{cases} \tag{2.5.4}$$

for one unknown value of i and

$$\begin{cases} H_{2i}: \beta_1 = \ldots = \beta_{i-1} = \beta_{i+1} = \ldots = \beta_k = \beta, \\ \beta_i = c\beta, \; c < 1, \end{cases} \tag{2.5.5}$$

for one unknown value of i.

For both tests we compute the ratios

$$\underline{x}_j = \frac{\underline{u}_j}{\sum \underline{u}_i}, \quad (j = 1, \ldots, k). \tag{2.5.6}$$

In order to be able to apply the procedure described in section 2.1 we have to prove the inequality (2.1.9).

The joint distribution of $\underline{x}_i$ and $\underline{x}_j$ under H_0 can be derived in the following way.
It is known that $\sum_{r\neq i,j} \underline{u}_r$ has also a gamma distribution with parameters $\sum_{r\neq i,j} \varepsilon_r$ and β. We define

$$\underline{U} \overset{\text{def}}{=} \underline{u}_1 + \dots + \underline{u}_k.$$

We have

$$(2.5.7) \qquad \begin{cases} u_i = x_i U, \\ u_j = x_j U, \\ \sum_{r\neq i,j} u_r = (1 - x_i - x_j)U. \end{cases}$$

The Jacobian of this transformation is U^2 and the joint distribution of $\underline{x}_i$, $\underline{x}_j$ and $\underline{U}$ is therefore

$$(2.5.8) \quad g(x_i,x_j,U) = C_{ij} x_i^{\varepsilon_i-1} x_j^{\varepsilon_j-1} (1-x_i-x_j)^{A-\varepsilon_i-\varepsilon_j-1} \frac{U^{A-1}e^{-U/\beta}}{\Gamma(A)\beta^A},$$

where

$$C_{ij} = \Gamma(A)\{\Gamma(\varepsilon_i)\Gamma(\varepsilon_j)\Gamma(A-\varepsilon_i-\varepsilon_j)\}^{-1}$$

and

$$A \overset{\text{def}}{=} \varepsilon_1 + \dots + \varepsilon_k.$$

Thus we see, as is well known, that $\underline{U}$ has also a gamma distribution with parameters $\varepsilon_1 + \dots + \varepsilon_k = A$ and β and moreover that the joint distribution of $\underline{x}_i$ and $\underline{x}_j$ is given by

$$(2.5.9) \qquad f(x_i,x_j) = C_{ij} x_i^{\varepsilon_i-1} x_j^{\varepsilon_j-1} (1-x_i-x_j)^{A-\varepsilon_i-\varepsilon_j-1},$$

$$0 \leq x_i \leq 1, \quad 0 \leq x_j \leq 1, \quad x_i + x_j \leq 1.$$

A similar derivation gives

$$(2.5.10) \qquad f(x_i) = C_i x_i^{\varepsilon_i-1} (1-x_i)^{A-\varepsilon_i-1},$$

where

$$C_i = \Gamma(A)\{\Gamma(\varepsilon_i)\Gamma(A-\varepsilon_i)\}^{-1}.$$

In this case theorem (2.2.1) provides us with a partial proof of inequality (2.1.9).

We consider 4 values

$$\begin{cases} 0 \le x_{i1} \le x_{i2} \le 1, \\ 0 \le x_{j1} \le x_{j2} \le 1. \end{cases}$$

If $x_{i2} + x_{j2} > 1$, $f(x_{i2},x_{j2}) = 0$ and (2.2.2) is fulfilled. Therefore we assume that $x_{i2} + x_{j2} \le 1$, which implies that also $x_{i1} + x_{j2} \le 1$, $x_{i2} + x_{j1} \le 1$ and $x_{i1} + x_{j1} \le 1$. Hence we may apply condition (2.2.11), which says in our case that

$$(2.5.11) \qquad \frac{\partial^2}{\partial x_i\, \partial x_j} \log \left[x_i^{\varepsilon_i-1}\, x_j^{\varepsilon_j-1}(1-x_i-x_j)^{A-\varepsilon_i-\varepsilon_j-1} \right] = - \frac{A-\varepsilon_i-\varepsilon_j-1}{(1-x_i-x_j)^2} \le 0,$$

which is only true if $A-\varepsilon_i-\varepsilon_j \ge 0$. In most practical cases, particularly in the case of estimated normal variances with degrees of freedom ≥ 2 this will be fulfilled, but a more general proof is needed. [1)]

It is easily seen that (2.1.9) is equivalent with

$$(2.5.12) \qquad \frac{p_{i,j}}{p_j} \le \frac{p_i - p_{i,j}}{q_j}.$$

From (2.5.9) and (2.5.10) it follows that the left hand member $L(g_i,g_j)$ of (2.5.12) equals

$$C \frac{\int_0^{g_j} \int_0^{g_i} x_j^{\varepsilon_j-1}\, x_i^{\varepsilon_i-1}\, (1-x_i-x_j)^{A-\varepsilon_i-\varepsilon_j-1}\, dx_i\, dx_j}{\int_0^{g_j} x_j^{\varepsilon_j-1}\, (1-x_j)^{A-\varepsilon_j-1}\, dx_j},$$

---

1) The following proof was given by professor H. Kesten, formerly at the Mathematical Centre, Amsterdam.

where $C = C_{ij} C_j^{-1}$.

For the same reason as mentioned above we need only consider values g_i and g_j such that $g_i + g_j \leq 1$.

Putting $x_i = v(1-x_j)$ we get

$$(2.5.13)\quad L(g_i,g_j) = C \frac{\int_0^{g_j}\int_0^{g_i/(1-x_j)} v^{\varepsilon_i-1}(1-v)^{A-\varepsilon_i-\varepsilon_j-1} x_j^{\varepsilon_j-1}(1-x_j)^{A-\varepsilon_j-1}\,dv\,dx_j}{\int_0^{g_j} x_j^{\varepsilon_i-1}(1-x_j)^{A-\varepsilon_j-1}\,dx_j}$$

$$\leq C \int_0^{g_i/(1-g_j)} v^{\varepsilon_i-1}(1-v)^{A-\varepsilon_i-\varepsilon_j-1}\,dv.$$

Similarly the right hand member $R(g_i,g_j)$ of (2.5.12) is found to be

$$(2.5.14)\qquad R(g_i,g_j) \geq C \int_0^{g_i/(1-g_j)} v^{\varepsilon_i-1}(1-v)^{A-\varepsilon_i-\varepsilon_j-1}\,dv.$$

So it follows from (2.5.13) and (2.5.14) that (2.5.12) holds.

The values d_j and e_j as defined by (2.1.2) and (2.1.4) can be obtained from tables or nomograms of the incomplete B-function as

$$(2.5.15)\qquad \begin{cases} d_j = 1 - I_{x_j}(\varepsilon_j, A-\varepsilon_j), \\ e_j = 1 - d_j = I_{x_j}(\varepsilon_j, A-\varepsilon_j). \end{cases}$$

When all ε_i are equal, the smallest d_j corresponds to the smallest x_j and the smallest e_j to the largest x_j. When the common parameter value ε is a multiple of 1/2 (normal variances estimated from samples of the same size) the tables may be found in C. EISENHART, M.W. HASTAY and W.A. WALLIS (1947) for the first test and in R. DOORNBOS (1956) for the second one.

In all other cases approximated d- and e-values can be obtained from the nomograms of H.O. HARTLEY and E.R. FITCH (1951) (Table 17 in BIOMETRIKA TABLES (1956)).

Finally we shall derive upper and lower bounds for the probability of making a correct decision under the hypotheses H_{1i} and H_{2i} when following our test procedure.

In the first case, we assume that H_{1i} is true, i.e. $\beta_i = C\beta$, $C > 1$. Then we prove that Q_i, the probability of making the correct decision lies between the limits

(2.5.16) $$\{1 - I_{B_i}(\varepsilon_i, A-\varepsilon_i)\}(1-\alpha) \leq Q_i \leq \{1 - I_{B_i}(\varepsilon_i, A-\varepsilon_i)\},$$

where

(2.5.17) $$B_i = \frac{G_{i,\alpha}}{C-(C-1)G_{i,\alpha}} \quad ,$$

where $G_{i,\alpha}$ is determined such as to make

(2.5.18) $$I_{G_{i,\alpha}}(\varepsilon_i, A-\varepsilon_i) = 1 - \alpha/k.$$

When C becomes large Q_i converges towards the upper bound given by the right hand member of (2.5.16).

When H_{2i} is true, i.e. $\beta_j = c\beta$, $0 \leq c < 1$, the following limits can be derived for P_j, the probability of making the correct decision in this case.

(2.5.19) $$\{I_{b_j}(\varepsilon_j, A-\varepsilon_j)\}(1-\alpha) \leq P_j \leq I_{b_j}(\varepsilon_j, A-\varepsilon_j),$$

where

$$b_j = \frac{g_{j,\alpha}}{c+(1-c)g_{j,\alpha}}$$

and $g_{j,\alpha}$ is determined from

$$I_{g_{j,\alpha}}(\varepsilon_j, A-\varepsilon_j) = \alpha/k.$$

Again, now for small values of c

$$P_j \approx I_{b_j}(\varepsilon_j, A-\varepsilon_j).$$

In order to prove (2.5.16) we may assume that i = 1 and then we put $\underline{u}_1/C = \underline{v}_1$, thus $\underline{v}_1$ has a gamma distribution with parameters ε_1 and β. The probability Q_1 of making the correct decision is

$$Q_1 = P\left[\underline{d}_1 = \min \underline{d}_j \quad \text{and} \quad \underline{d}_1 < \alpha/k\right]$$

$$\geq P\left[\underline{d}_1 < \alpha/k \quad \text{and} \quad \underline{d}_2 > \alpha/k \quad \text{and} \ \ldots \ \text{and} \quad \underline{d}_k > \alpha/k\right]$$

$$= P\left[\underline{d}_1 < \alpha/k\right] - P\left[(\underline{d}_1 < \alpha/k \quad \text{and} \quad \underline{d}_2 < \alpha/k); \ \ldots \ \text{or} \ (\underline{d}_1 < \alpha/k \quad \text{and} \quad \underline{d}_k < \alpha/k)\right].$$

Thus the following inequality holds

$$(2.5.20) \qquad P\left[\underline{d}_1 < \alpha/k\right] - \sum_{j=2}^{k} P\left[\underline{d}_1 < \alpha/k \quad \text{and} \quad \underline{d}_j < \alpha/k\right] \leq Q_1 \leq P\left[\underline{d}_1 < \alpha/k\right].$$

We have

$$P\left[\underline{d}_1 < \alpha/k\right] = (\text{cf. } (2.5.18)) = P\left[\underline{x}_1 > G_{1,\alpha}\right] =$$

$$= P\left[\frac{C\,\underline{v}_1}{\underline{v}_1 + \underline{u}_2 + \ldots + \underline{u}_k + (C-1)\underline{v}_1} > G_{1,\alpha}\right]$$

$$= P\left[\frac{\underline{v}_1}{\underline{v}_1 + \underline{u}_2 + \ldots + \underline{u}_k} > \frac{G_{1,\alpha}}{C-(C-1)G_{1,\alpha}}\right]$$

$$= P\left[\frac{\underline{v}_1}{\underline{v}_1 + \underline{u}_2 + \ldots + \underline{u}_k} > B_1\right] \quad (\text{cf. } (2.5.17)).$$

The distribution of $\dfrac{\underline{v}_1}{\underline{v}_1 + \underline{u}_2 + \ldots + \underline{u}_k}$ is the distribution of $\underline{x}_1$ under H_0 and thus we have

$$(2.5.21) \qquad P\left[\underline{d}_1 < \alpha/k\right] = 1 - I_{B_1}(\varepsilon_1, A-\varepsilon_1).$$

Further

(2.5.22) $P\left[\underline{d}_1 < \alpha/k \text{ and } \underline{d}_j < \alpha/k\right] =$

$$= P\left[\frac{\underline{v}_1}{\underline{v}_1 + \underline{u}_2 + \ldots + \underline{u}_k} > B_1 \text{ and } \frac{\underline{u}_j}{\underline{v}_1 + \ldots + \underline{u}_k + (C-1)\underline{v}_1} > G_{j,\alpha}\right]$$

$$\leq P\left[\frac{\underline{v}_1}{\underline{v}_1 + \underline{u}_2 + \ldots + \underline{u}_k} > B_1 \text{ and } \frac{\underline{u}_j}{\underline{v}_1 + \underline{u}_2 + \ldots + \underline{u}_k} > G_{j,\alpha}\right]$$

(according to (2.1.10))

$$\leq P\left[\frac{\underline{v}_1}{\underline{v}_1 + \underline{u}_2 + \ldots + \underline{u}_k} > B_1\right] \cdot P\left[\frac{\underline{u}_j}{\underline{v}_1 + \underline{u}_2 + \ldots + \underline{u}_k} > G_{j,\alpha}\right] =$$

$$= \left\{1 - I_{B_1}(\varepsilon_1, A-\varepsilon_1)\right\} \cdot \alpha/k.$$

Substituting (2.5.21) and (2.5.22) into (2.5.20) we get

$$\left[1 - I_{B_1}(\varepsilon_1, A-\varepsilon_1)\right](1-\alpha) < \left[1 - I_{B_1}(\varepsilon_1, A-\varepsilon_1)\right]\left\{1 - \frac{k-1}{k}\alpha\right\}$$

$$\leq Q_1 \leq \left[1 - I_{B_1}(\varepsilon_1, A-\varepsilon_1)\right],$$

which proves (2.5.16). When C is large $P\left[\underline{d}_j < \alpha/k\right]$ will for $j \neq 1$ be much smaller than α/k and therefore in that case Q_1 converges towards its upper bound.

The inequalities (2.5.19) can be derived in the same way.

2.6. SLIPPAGE TESTS FOR SOME DISCRETE VARIABLES

First we prove the following theorem.

THEOREM 2.6.1

Suppose the discrete random variables

(2.6.1) $$\underline{u}_1, \ldots, \underline{u}_k$$

are distributed independently and can take integer values only (the latter assumption is not essential but simplifies the notation).
If

$$(2.6.2) \qquad \frac{P[\sum \underline{u}_l - \underline{u}_i - \underline{u}_j = a]}{P[\sum \underline{u}_l - \underline{u}_i - \underline{u}_j = a+1]} ,$$

where a is an integer, is a non decreasing function of a, then

$$(2.6.3) \quad P\left[\underline{u}_i \geq u_i \text{ and } \underline{u}_j \geq u_j \;\middle|\; \sum \underline{u}_l = N\right] \leq$$

$$\leq P\left[\underline{u}_i \geq u_i \;\middle|\; \sum \underline{u}_l = N\right] \cdot P\left[\underline{u}_j \geq u_j \;\middle|\; \sum \underline{u}_l = N\right],$$

for every pair of integers u_i and u_j and for every non negative integer N.

PROOF

According to (2.6.2) we have that

$$(2.6.4) \quad \frac{P[\underline{u}_i = y] \cdot P[\underline{u}_j = x] \cdot P[\sum \underline{u}_l - \underline{u}_i - \underline{u}_j = N-x-y]}{P[\underline{u}_i = y] \cdot P[\underline{u}_j = x+1] \cdot P[\sum \underline{u}_l - \underline{u}_i - \underline{u}_j = N-x-y-1]}$$

is non increasing in y. Dividing (2.6.4) by the factor

$$(2.6.5) \qquad \frac{P[\sum \underline{u}_l = N \text{ and } \underline{u}_j = x]}{P[\sum \underline{u}_l = N \text{ and } \underline{u}_j = x+1]} ,$$

which does not depend on y, (2.6.4) changes into

$$(2.6.6) \qquad \frac{P[\underline{u}_i = y \mid \sum \underline{u}_l = N \text{ and } \underline{u}_j = x]}{P[\underline{u}_i = y \mid \sum \underline{u}_l = N \text{ and } \underline{u}_j = x+1]} .$$

Thus also (2.6.6) is non increasing in y for all x. This means that there exists a value y_0, which may depend on x, with the property that

$$(2.6.7) \quad P\left[\underline{u}_i = y \;\middle|\; \sum \underline{u}_l = N \text{ and } \underline{u}_j = x\right] \geq$$

$$\geq P\left[\underline{u}_i = y \;\middle|\; \sum \underline{u}_l = N \text{ and } \underline{u}_j = x+1\right], \text{ if } y \geq y_0$$

and

$$P\left[\underline{u}_i = y \;\middle|\; \sum \underline{u}_l = N \text{ and } \underline{u}_j = x\right] <$$
$$< P\left[\underline{u}_i = y \;\middle|\; \sum \underline{u}_l = N \text{ and } \underline{u}_j = x+1\right], \text{ if } y < y_0.$$

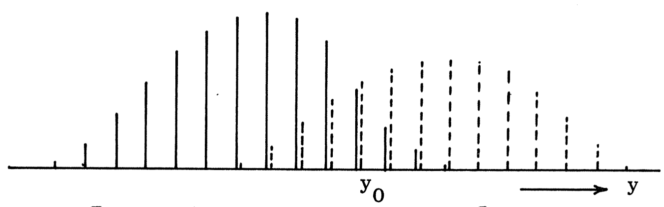

Fig. (2.6.1) $P\left[\underline{u}_i = y \;\middle|\; \sum \underline{u}_l = N \text{ and } \underline{u}_j = x\right]$ (dotted lines) and $P\left[\underline{u}_i = y \;\middle|\; \sum \underline{u}_l = N \text{ and } \underline{u}_j = x+1\right]$ (full lines).

This situation is sketched in fig.(2.6.1). It follows for each value of u_i

(2.6.8) $$P(x) \overset{\text{def}}{=} \sum_{y=u_i}^{\infty} P\left[\underline{u}_i = y \;\middle|\; \sum \underline{u}_l = N \text{ and } \underline{u}_j = x\right]$$

is a non increasing function of x. Now

(2.6.9) $$\frac{P\left[\underline{u}_i \geq u_i \text{ and } \underline{u}_j \geq u_j \mid \sum \underline{u}_l = N\right]}{P\left[\underline{u}_j \geq u_j \mid \sum \underline{u}_l = N\right]} =$$

$$= \frac{\sum_{x=u_j}^{\infty} P\left[\underline{u}_j = x \mid \sum \underline{u}_l = N\right] \sum_{y=u_i}^{\infty} P\left[\underline{u}_i = y \mid \sum \underline{u}_l = N \text{ and } \underline{u}_j = x\right]}{\sum_{x=u_j}^{\infty} P\left[\underline{u}_j = x \mid \sum \underline{u}_l = N\right]} \leq$$

$$\leq \sum_{y=u_i}^{\infty} P\left[\underline{u}_i = y \;\middle|\; \sum \underline{u}_l = N \text{ and } \underline{u}_j = u_j\right].$$

In the same way we have

$$(2.6.10) \qquad \frac{P\left[\underline{u}_i \geq u_i \text{ and } \underline{u}_j < u_j \mid \sum \underline{u}_1 = N\right]}{P\left[\underline{u}_j < u_j \mid \sum \underline{u}_1 = N\right]} \geq$$

$$\geq \sum_{y=u_i}^{\infty} P\left[\underline{u}_i = y \;\middle|\; \sum \underline{u}_1 = N \text{ and } \underline{u}_j = u_j\right].$$

From (2.6.9) and (2.6.10) it follows that, in the notation of (2.1.6), where $u_i = g_i + 1$ and $u_j = g_j + 1$, whilst $\underline{u}_i$ under the condition $\sum \underline{u}_1 = N$ stands for $\underline{x}_i$ and $\underline{u}_j$ under the condition $\sum \underline{u}_1 = N$ for $\underline{x}_j$,

$$(2.6.11) \qquad \frac{q_{i,j}}{q_j} \leq \frac{q_i - q_{i,j}}{1 - q_j} ,$$

or

$$(2.6.12) \qquad q_{i,j} \leq q_i \, q_j,$$

which proves the theorem, because (2.6.12) is the same as (2.6.3).

As the first application of theorem (2.6.1) we shall consider the Poisson case. Suppose we have a set of independent random variables

$$(2.6.13) \qquad \underline{z}_1, \ldots, \underline{z}_k,$$

distributed according to Poisson distributions, i.e.:

$$P\left[\underline{z}_i = z_i\right] = \frac{e^{-\mu_i} \mu_i^{z_i}}{z_i!}, \quad (i = 1,\ldots,k), \; \mu_i > 0.$$

Now we want to test the hypothesis H_0 that the means μ_i have known ratios

$$(2.6.14) \qquad H_0: \frac{\mu_i}{\sum\limits_j \mu_j} = p_i \quad (i = 1,\ldots,k).$$

This situation occurs for instance if from k Poisson-populations with, under H_0, equal means unequal numbers of observations are present and $\underline{z}_1, \ldots, \underline{z}_k$ represent the sums of these observations.

In this case the p_i are proportional to the number of observations. Also k Poisson processes with the same parameter may be observed during different lengths of time. Then the p_i are proportional to these lengths of time.

The alternatives against which H_0 is tested are

$$(2.6.15)\quad H_{1i}: \frac{\mu_i}{\sum_j \mu_j} = Cp_i, \frac{\mu_l}{\sum_j \mu_j} = \frac{1 - Cp_i}{1 - p_i} p_l \ (l \neq i),$$

$$1 < C < \frac{1}{p_i}, \text{ C unknown, for one unknown value of i or}$$

$$(2.6.16)\quad H_{2i}: \frac{\mu_i}{\sum_j \mu_j} = cp_i, \frac{\mu_l}{\sum_j \mu_j} = \frac{1 - cp_i}{1 - p_i} p_l \ (l \neq i),$$

$$0 < c < 1, \text{ c unknown, for one unknown value of i.}$$

A well known property of Poisson-variates is:
If $\underline{z}_1, \ldots, \underline{z}_k$ are independent Poisson-variates with means $\mu_1, \ldots, \mu_k$, then the simultaneous conditional distribution of $\underline{z}_1, \ldots, \underline{z}_k$ gives their sum (i.e. $\sum \underline{z}_i = N$, N a constant), is a multinomial distribution with probabilities $p_i = \mu_i / \sum \mu_j$ and number of trials $\sum \underline{z}_i = N$. As the hypotheses (2.6.14), (2.6.15) and (2.6.16) only contain the ratios p_i it seems natural to use a conditional test for H_0, using only the multinomial distribution

$$(2.6.17)\qquad P\left[\underline{z}_1 = z_1, \ldots, \underline{z}_k = z_k \;\middle|\; \sum \underline{z}_i = N\right] = \frac{N!}{\Pi z_i!} \Pi p_i^{z_i}.$$

From this it is clear that a test against slippage for Poisson variates is closely related to a similar test for multinomial variates and the test stated here may easily be translated into tests for the multinomial case.

Now we apply theorem 2.6.1. In the case under consideration the sum of k-2 of the variables $\underline{z}_i$ of (2.6.13) has a Poisson distribution with mean μ, say. So condition 2.6.2 states that

(2.6.18) $$\frac{e^{-\mu}\mu^{a}}{a!}\,\frac{(a+1)!}{e^{-\mu}\mu^{a+1}} = \frac{a+1}{\mu},$$

is non decreasing in a, which is clearly true. Thus the inequality (2.6.3) holds for every pair $\underline{z}_i$, $\underline{z}_j$ and the test procedure of section (2.1) may be applied to the variates $\underline{z}_1, \ldots, \underline{z}_k$ under the condition $\sum \underline{z}_i = N$.

The marginal distribution of $\underline{z}_i$ under the condition $\sum \underline{z}_i = N$ is a binomial one, so when testing H_0 against H_{1i} $(i = 1,\ldots,k)$ we compute, if $z_1, \ldots, z_k$ are the observed values and $\sum z_i = N$,

(2.6.20) $$r_i \overset{def}{=} P\left[\underline{z}_i \geq z_i \;\middle|\; \sum \underline{z}_i = N\right] = \sum_{x=z_i}^{N} \binom{N}{x} p_i^{\,x}(1-p_i)^{N-x} =$$

$$= I_{p_i}(z_i, N - z_i + 1),$$

where $I_{p_i}(z_i, N-z_i+1)$ stands for the incomplete B-function

$$\frac{N!}{(z_i-1)!(N-z_i)!} \int_0^{p_i} u^{z_i-1}(1-u)^{N-z_i}\,du.$$

Now H_0 is rejected if

(2.6.21) $$\min_i r_i \leq \alpha/k$$

and then we decide that $\mu_j/\sum \mu_i > p_j$ if $r_j = \min r_i$ *).

If under H_0 $\mu_1 = \ldots = \mu_k$, all p_i are equal and the smallest r_i corresponds to the largest value z_i. Otherwise the minimum r-value can be found along the lines described for the e-values in section 2.5.

The test for slippage to the left is completely analogous. Table I gives critical values for max z_i in the case $p_1 = \ldots = p_k$.

---

*) If the minimum is reached for more than one value j, one can be selected at random. A. WALD (1950) takes the smallest index for which the minimum is attained. As the alternative hypotheses under consideration allow for one outlier only, the rejection of more than one value is not considered here.

Along the same lines as followed in section 2.5 in the case of gamma-variates it can be shown that the probability Q_j of making the correct decision when the j-th population has slipped to the right (i.e. H_{1i} is true with i = j) satisfies the inequality

$$I_{Cp_j}(G_{j,\alpha}, N-G_{j,\alpha}+1)(1-\alpha) \le$$

(2.6.22) $$\le I_{Cp_j}(G_{j,\alpha}, N-G_{j,\alpha}+1)\left[1 - \sum_{i\neq j} I_{\frac{1-Cp_j}{1-p_j}p_i}(G_{i,\alpha}, N-G_{i,\alpha}+1)\right]$$

$$\le Q_j \le I_{Cp_j}(G_{j,\alpha}, N-G_{j,\alpha}+1).$$

Here $G_{l,\alpha}$ $(l = 1,\dots,k)$ is the smallest number which satisfies

(2.6.23) $$P\left[\underline{z}_l \ge G_{l,\alpha} \;\middle|\; \sum \underline{z}_i = N,\ H_0\right] \le \alpha/k$$

or

(2.6.24) $$I_{p_l}(G_{l,\alpha}, N-G_{l,\alpha}+1) \le \alpha/k.$$

Clearly Q_j converges towards its upper bound when $C \to 1/p_j$ and for each $C \ge 1$ the factor between square brackets is larger than $1 - (k-1)\alpha/k$ according to (2.6.24).

In the case of slippage to the left we have analogously

(2.6.25) $$\begin{cases} \left[1 - I_{cp_j}(g_{j,\alpha}+1, N-g_{j,\alpha})\right](1-\alpha) \le \\ \le \left[1-I_{cp_j}(g_{j,\alpha}+1, N-g_{j,\alpha})\right]\left[1-\sum_{i\neq j}\left\{1-I_{\frac{1-cp_j}{1-p_j}p_i}(g_{i,\alpha}+1, N-g_{i,\alpha})\right\}\right] \\ \le P_j \le 1 - I_{cp_j}(g_{j,\alpha}+1, N-g_{j,\alpha}), \end{cases}$$

where $g_{l,\alpha}$ is the largest number satisfying

$$1 - I_{p_l}(g_{l,\alpha}+1, N-g_{l,\alpha}) \le \alpha/k.$$

We can apply theorem 2.6.1 also to the case of independent variables

(2.6.26) $$\underline{v}_1, \ldots, \underline{v}_k$$

which are distributed according to binomial laws with numbers of trials $n_1, \ldots, n_k$ and probabilities of success $p_1, \ldots, p_k$. Now the hypothesis H_0 is

(2.6.27) $$H_0: p_1 = \ldots = p_k = p, \text{ say}$$

and the alternatives are

(2.6.28) $$H_{1i}: \begin{cases} p_1 = \ldots = p_{i-1} = p_{i+1} = \ldots = p_k = p, \\ p_i = Cp \; (1 < C \leq 1/p), \end{cases}$$

for one unknown value of i and

(2.6.29) $$H_{2i}: \begin{cases} p_1 = \ldots = p_{i-1} = p_{i+1} = \ldots = p_k = p, \\ p_i = cp \; (0 < c \leq 1), \end{cases}$$

for one unknown value of i.

Because, under H_0, the sum of (k-2) of the variates (2.6.26) has again a binomial distribution with number of trials n, say, and probability of a success in each trial p, the condition (2.6.2) of theorem (2.6.1) reads:

(2.6.30) $$\frac{\binom{n}{a} p^a (1-p)^{n-a}}{\binom{n}{a+1} p^{a+1} (1-p)^{n-a-1}} = \frac{a+1}{n-a} \cdot \frac{1-p}{p}$$

is a non decreasing function of a, which is true. So in this case also the approximation procedure described in section (2.1) can be applied to obtain a conditional test for slippage under the condition that the sum of the variates $\sum \underline{v}_i$ has a constant value N. The conditional distribution of $\underline{v}_i$ is a hypergeometrical one

(2.6.31) $$P\left[\underline{v}_i = v_i \;\middle|\; \sum \underline{v}_j = N\right] = \binom{n_i}{v_i} \binom{\sum n_j - n_i}{N - v_i} \binom{\sum n_j}{N}^{-1} \quad (v_i \geq 0).$$

From these distributions the exceedance probabilities

$$s_i \stackrel{\text{def}}{=} P\left[\underline{v}_i \geq v_i \;\middle|\; \sum \underline{v}_j = N\right]$$

have to be determined and it is decided that v_j has slipped to the right when

$$s_j = \min s_i \leq \alpha/k$$

and similarly for slippage to the left.

Recently the hypergeometrical distribution was extensively tabulated by G.J. LIEBERMAN and D.B. OWEN (1961). In this book tables are given for n_i up to 50 and $\sum n_j$ up to 100. So in most practical cases critical values can be found there.

In the special case $n_1 = \ldots = n_k = n$, the test procedure for slippage to the right reduces to comparing the largest variate $\underline{v}_m$ with a constant v_0, where v_0 is the largest integer satisfying

$$P\left[\underline{v}_i \geq v_0 \;\middle|\; \sum \underline{v}_i = N\right] \leq \alpha/k,$$

α being the level of significance.

The test for slippage to the left is found in a similar way.

Now we can consider the variates

(2.6.32) $$\underline{w}_1, \ldots, \underline{w}_k,$$

which are independently distributed according to negative binomial laws, with parameters $r_1, \ldots, r_k$ and probabilities $p_1, \ldots, p_k$, i.e.

(2.6.33) $$P\left[\underline{w}_i = w_i\right] = \binom{w_i+r_i-1}{r_i-1} p_i^{r_i} q_i^{w_i},$$

where r_i is an integer ≥ 1 and $0 \leq p_i \leq 1$, whilst $p_i+q_i = 1$.

The hypothesis H_0 is

(2.6.34) $$H_0: q_1 = \ldots = q_k = q, \text{ say}$$

and the alternatives are

$$(2.6.35)\qquad \begin{cases} H_{1i}: q_1 = \dots = q_{i-1} = q_{i+1} = \dots = q_k = q, \\ q_i = Cq \ (1 \le C \le 1/q), \end{cases}$$

for one unknown value of i or

$$(2.6.36)\qquad \begin{cases} H_{2i}: q_1 = \dots = q_{i-1} = q_{i+1} = \dots = q_k = q, \\ q_i = cq \ (0 \le c \le 1), \end{cases}$$

for one unknown value of i.

The hypotheses are stated in terms of the q_i and not in terms of the p_i in order to obtain that slippage to the right of the i-th population corresponds to a large value of $\underline{w}_i$.

Under H_0 the term of a set of negative binomial variates has again a negative binomial distribution with the same probability p (or q) and a parameter r which is the sum of the r_i of the individual variates. So condition (2.6.2) gives here

$$(2.6.37)\qquad \frac{\binom{a+r-1}{r-1} p^r q^a}{\binom{a+r}{r-1} p^r q^{a+1}} = \frac{a+1}{a+r}\frac{1}{q}$$

is a non decreasing function of a, which is true if $r \ge 1$. Thus again the method of section (2.1) can be applied. The conditional distribution of $\underline{w}_i$ under the condition $\sum \underline{w}_i = N$, has the form

$$(2.6.38)\qquad P\left[\underline{w}_i = w \,\middle|\, \sum \underline{w}_i = N\right] = \frac{\binom{w_i+r_i-1}{r_i-1}\binom{N+\sum r_j-w_i-r_i-1}{\sum r_j-r_i-1}}{\binom{N+\sum r_j-1}{\sum r_j-1}}, \qquad (w_i = 0, \dots, N).$$

The critical region for the test against H_{1i} $(i = 1,\dots,k)$ (2.6.35) consists of large values of $\underline{w}_i$. In the case where $r_1 = \dots = r_k$ the test statistic is the largest variate $\underline{w}_m$, when testing against slippage to the right and the smallest when testing against slippage to the left.

2.7. A SLIPPAGE TEST FOR THE METHOD OF m RANKINGS

In the well known method of m rankings due to M. FRIEDMAN (1937) (cf. M.G. KENDALL (1955), chapters 6 and 7) m "observers" are considered. Each observer ranks k "objects". The method of m rankings enables us to investigate whether the observers agree in their opinion about the objects. For that reason one tests the hypothesis H_0, which states that the rankings are chosen at random from the collection of all permutations of the numbers 1, ..., k and that they are independent.

Here we present tests which are powerful especially against the alternative that one of the objects has larger probability than the other ones of being ranked high (or low), whilst the other (k-1) objects are ranked in a random order. We denote the sums of the m ranks of each object by

(2.7.1) $$\underline{s}_1, \ldots, \underline{s}_k \quad (m \leq \underline{s}_i \leq km).$$

Obviously we have

(2.7.2) $$\sum_{i=1}^{k} \underline{s}_i = \tfrac{1}{2}mk(k+1).$$

First we prove the following theorem

<u>THEOREM 2.7.1</u>

For each pair $\underline{s}_i$, $\underline{s}_j$ of the variates (7.1) and for every pair of integers s_i, s_j the following inequality holds under H_0

(2.7.3) $$P\left[\underline{s}_i \leq s_i \text{ and } \underline{s}_j \leq s_j\right] \leq P\left[\underline{s}_i \leq s_i\right] \cdot P\left[\underline{s}_j \leq s_j\right].$$

<u>PROOF</u>

We suppose that $m \leq s_i, s_j \leq km$, because otherwise (2.7.3) obviously holds with the equality sign. For m = 1 we have

$$(2.7.4)\quad\begin{cases} P\left[\underline{s}_i \le s_i \text{ and } \underline{s}_j \le s_j \;\middle|\; m = 1\right] = \dfrac{s_i s_j - \min(s_i, s_j)}{k(k-1)}, \\ P\left[\underline{s}_i \le s_i \;\middle|\; m = 1\right] = \dfrac{s_i}{k}, \\ P\left[\underline{s}_j \le s_j \;\middle|\; m = 1\right] = \dfrac{s_j}{k}, \end{cases}$$

so in that case (2.7.3) is true. Now let us suppose that (2.7.3) is true for m observers, then we have

$$(2.7.5)\quad\begin{cases} P\left[\underline{s}_i \le s_i \text{ and } \underline{s}_j \le s_j \;\middle|\; m+1\right] = \\ = \sum_{a\neq b} P\left[\underline{s}_i \le s_i - a \text{ and } \underline{s}_j \le s_j - b \;\middle|\; m\right] \cdot P\left[\text{the i-th object has rank a and the j-th object rank b in the (m+1)-st ranking}\right] = \\ = \sum_{a\neq b} P\left[\underline{s}_i \le s_i - a \text{ and } \underline{s}_j \le s_j - b \;\middle|\; m\right] \cdot \dfrac{1}{k(k-1)} \le \\ \le \sum_{a\neq b} P\left[\underline{s}_i \le s_i - a \middle| m\right] \cdot P\left[\underline{s}_j \le s_j - b \middle| m\right] \cdot \dfrac{1}{k(k-1)} = \\ = \sum_{a=1}^{k} P\left[\underline{s}_i \le s_i - a \middle| m\right] \cdot \dfrac{1}{k} \cdot \sum_{b=1}^{k} P\left[\underline{s}_j \le s_j - b \middle| m\right] \cdot \dfrac{1}{k} + \\ + \dfrac{1}{k^2(k-1)} \sum_{a=1}^{k} P\left[\underline{s}_i \le s_i - a \middle| m\right] \cdot \sum_{b=1}^{k} P\left[\underline{s}_j \le s_j - b \middle| m\right] + \\ - \dfrac{1}{k(k-1)} \sum_{a=1}^{k} P\left[\underline{s}_i \le s_i - a \middle| m\right] \cdot P\left[\underline{s}_j \le s_j - a \middle| m\right] = \\ = P\left[\underline{s}_i \le s_i \middle| m+1\right] \cdot P\left[\underline{s}_j \le s_j \middle| m+1\right] + \\ - \dfrac{1}{k(k-1)} \sum_{a=1}^{k} \left\{ P\left[\underline{s}_i \le s_i - a \middle| m\right] - \dfrac{\sum_{b=1}^{k} P\left[\underline{s}_i \le s_i - b \middle| m\right]}{k} \right\}. \end{cases}$$

$$(2.7.5)\quad \begin{cases} \cdot \left\{ P\left[\underline{s}_j \le s_j - a \middle| m\right] - \dfrac{\sum_{b=1}^{k} P\left[\underline{s}_j \le s_j - b | m\right]}{k} \right\} \le \\ \le P\left[\underline{s}_i \le s_i \middle| m+1\right] \cdot P\left[\underline{s}_j \le s_j \middle| m+1\right]. \end{cases}$$

So theorem 2.7.1 is proved by induction.

So we can apply our approximation method of section 2.1 for obtaining slippage tests for $\underline{s}_1, \ldots, \underline{s}_k$. Because the marginal distributions of the $\underline{s}_i$ are all equal under H_0, the test statistic for the test against slippage to the right is max $\underline{s}_i$ and for testing against slippage to the left min $\underline{s}_i$. The critical values are determined by the smallest integer S_α satisfying

$$(2.7.6)\qquad P\left[\underline{s}_i \ge S_\alpha\right] \le \alpha/k$$

and the largest integer s_α satisfying

$$(2.7.7)\qquad P\left[\underline{s}_i \le s_\alpha\right] \le \alpha/k,$$

respectively.

The distribution of $\underline{s}_i$ is easily seen to be symmetric with respect to the mean value $\frac{1}{2}m(k+1)$, so we have

$$(2.7.8)\qquad s_\alpha = m(k+1) - S_\alpha .$$

Now we will show that the distribution of $\underline{s}_i$, under H_0, reads

$$(2.7.9)\qquad P\left[\underline{s}_i = n\right] = \sum_{x=0}^{\infty} \binom{m}{x} \binom{n-kx-1}{m-1} (-1)^x k^{-m}, \quad (i = 1,\ldots,k;\ m \le n \le km)$$

where the binomial coefficient $\binom{a}{b}$ as usual is defined to be 0 for $b > a$.

The cumulative probability $P\left[\underline{s}_i \le n\right]$ is consequently

$$(2.7.10)\qquad P\left[\underline{s}_i \le n\right] = \sum_{x=0}^{\infty} \binom{m}{x} \binom{n-kx}{m} (-1)^x k^{-m}.$$

Table II of critical values s_α and S_α is based on this formula.

Formula (2.7.9) can be proved in the following way:

$k^m P\left[\underline{s}_i = n \middle| m\right]$ = the number of partitions of n into m positive integers, no one being larger than k (different permutations of the same integers are counted as different partitions).

Thus

$k^m P\left[\underline{s}_i = n \middle| m\right]$ = coefficient of z^n in $(z + z^2 + \ldots + z^k)^m$ = coefficient of z^{n-m} in $\left(\frac{1-z^k}{1-z}\right)^m$ = coefficient of z^{n-m} in

$$\sum_{x=0}^{\infty} \binom{m}{x} (-1)^x z^{kx} \sum_{r=0}^{\infty} \binom{m+r-1}{r} z^r =$$

$$= \sum_{x=0}^{\infty} \binom{m}{x} \binom{n-kx-1}{m-1} (-1)^x,$$

which proves (2.7.9).

REMARK

W.J. YOUDEN (1963) proposed a test which was developed by W.A. THOMPSON and T.A. WILKE (1963) which is closely related to the one described in this section. Their test is two-sided but the result is virtually the same as when our test is used two-sided by halving the levels in the right hand sides of (2.7.6) and (2.7.7) to $\alpha/2k$ and using two critical values simultaneously.

2.8. A DISTRIBUTION FREE K SAMPLE SLIPPAGE TEST

We consider the independent variates

(2.8.1) $$\underline{u}_1, \ldots, \underline{u}_k,$$

which have, under H_0, the same continuous distribution function. From the i-th population we have n_i independent observations $\underline{u}_{ij}$ $(j = 1,\ldots,n_i)$.

We want to test H_0 against the alternatives

(2.8.2) $$H_{1i} \begin{cases} P\left[\underline{u}_i > \underline{u}_j\right] > \frac{1}{2} \quad (j \neq i) \\ \underline{u}_j \ (j = 1,\ldots,i-1,i+1,\ldots,k) \text{ follow the same distribution,} \end{cases}$$

for one unknown value of i and

$$(2.8.3)\qquad H_{2i}\begin{cases} P\left[\underline{u}_i > \underline{u}_j\right] < \frac{1}{2} \quad (j = i), \\ \underline{u}_j \quad (j = 1,\dots,i-1,i+1,\dots,k) \text{ follow the same distribution,} \end{cases}$$

for one unknown value of i.

Now the following test procedure is proposed. If all the observations $\underline{u}_{ij}$ $(i = 1,\dots,k;\ j = 1,\dots,n_i)$ are ranked, we denote by $\underline{T}_i$ the sum of the ranks of the observations $\underline{u}_{ij}$ $(j = 1,\dots,n_i)$. As $\underline{T}_i$ is a linear function of WILCOXON's test statistic applied to the i-th sample and the other k-1 samples together, its distribution function under H_0 is known. So for each set of observed values T_1, ..., T_k we can, under H_0, compute

$$(2.8.4)\qquad q_i \overset{\text{def}}{=} P\left[\underline{T}_i \geq T_i\right].$$

Now, when testing H_0 against H_1, H_0 is rejected when $q_i \leq \alpha/k$ and it is decided that H_{1j} is true when $q_j = \min q_i$.
A similar procedure is followed for slippage to the left. This procedure, based upon the method of section 2.1 is valid if the following theorem holds:

THEOREM 2.8.1

For every pair of integers i, j $(i \leq k;\ j \leq k;\ i \neq j)$ and for every pair of integers T_i and T_j the following inequality holds under H_0

$$(2.8.5)\qquad P\left[\underline{T}_i \geq T_i \text{ and } \underline{T}_j \geq T_j\right] \leq P\left[\underline{T}_i \geq T_i\right] \cdot P\left[\underline{T}_j \geq T_j\right].$$

Before proving this theorem we will first prove the following lemma.

LEMMA 2.8.1

If between the random variables $\underline{x}$ and $\underline{y}$ exists a relationship

$$\underline{y} = \phi(\underline{x}),$$

ϕ being a monotone non-increasing function, then

$$\mathcal{E}\,\underline{x}\,\underline{y} \leq \mathcal{E}\,\underline{x}\;\mathcal{E}\,\underline{y}\,.$$

PROOF

For every pair of values $\{x,y = \phi(x)\}$, we have

$$(x - \mathcal{E}\,\underline{x})\{\phi(x) - \phi(\mathcal{E}\,\underline{x})\} \leq 0.$$

Therefore

$$\begin{aligned} 0 &\geq \mathcal{E}\,(\underline{x} - \mathcal{E}\,\underline{x})\{\phi(\underline{x}) - \phi(\mathcal{E}\,\underline{x})\} = \\ &= \mathcal{E}\,(\underline{x} - \mathcal{E}\,\underline{x})(\underline{y} - \mathcal{E}\,\underline{y}) + \{\mathcal{E}\,\underline{y} - \phi(\mathcal{E}\,\underline{x})\}\,\mathcal{E}(\underline{x} - \mathcal{E}\,\underline{x}) = \\ &= \mathcal{E}\,\underline{x}\,\underline{y} - \mathcal{E}\,\underline{x}\;\mathcal{E}\,\underline{y} + 0, \end{aligned}$$

which proves the lemma.

Next we give the proof of theorem 2.8.1. [1)]
For simplicity we take i = 1, j = 2. We also take k = 3. This is no restriction on the generality as pooling of the samples 3, ..., k does not affect the probabilities under consideration.

Let

(2.8.6)
$$\begin{cases} P\left[T_i\right] \overset{\text{def}}{=} P\left[\underline{T}_i \geq T_i\right] \\ P\left[T_i, T_j\right] \overset{\text{def}}{=} P\left[\underline{T}_i \geq T_i \text{ and } \underline{T}_j \geq T_j\right] \\ P\left[T_i \middle| 1\right] \overset{\text{def}}{=} \text{the conditional probability of } \underline{T}_i \geq T_i \text{ given that the largest observation belongs to the 1-th sample.} \\ P\left[T_i, T_j \middle| 1\right] \overset{\text{def}}{=} \text{the conditional probability of } \underline{T}_i \geq T_i \text{ and } \underline{T}_j \geq T_j \text{ given that the largest observation belongs to the 1-th sample.} \end{cases}$$

---

1) For this proof thanks are due to Prof. H. KESTEN.

We shall prove (2.8.5) by induction with respect to $n_1 + n_2 + n_3$. Therefore from now on we provide all probabilities with subscripts indicating the sample sizes. So we have to prove

$$(2.8.7) \qquad P_{n_1,n_2,n_3}\left[T_1,T_2\right] \leq P_{n_1,n_2,n_3}\left[T_1\right] \cdot P_{n_1,n_2,n_3}\left[T_2\right].$$

It is easily verified that (2.8.7) holds for $n_1 + n_2 + n_3 = 3$ ($n_1 = n_2 = n_3 = 1$). Now suppose (2.8.7) holds if $n_1 + n_2 + n_3 \leq N-1$, then we prove that the inequality holds for $n_1 + n_2 + n_3 = N$. We have

$$(2.8.8) \qquad P_{n_1,n_2,n_3}\left[T_1,T_2\right] = \sum_{i=1}^{3} \frac{n_i}{N} P_{n_1,n_2,n_3}\left[T_1,T_2 \middle| i\right].$$

For the first term of the sum in the right hand member we get

$$(2.8.9) \qquad P_{n_1,n_2,n_3}\left[T_1,T_2 \middle| 1\right] = P_{n_1-1,n_2,n_3}\left[T_1 - N,\ T_2\right] \leq$$

$$\text{(according to our assumption)} \leq P_{n_1-1,n_2,n_3}\left[T_1-N\right]$$

$$P_{n_1-1,n_2,n_3}\left[T_2\right] = P_{n_1,n_2,n_3}\left[T_1 \middle| 1\right] \cdot P_{n_1,n_2,n_3}\left[T_2 \middle| 1\right].$$

In exactly the same way we find that

$$(2.8.10) \qquad P_{n_1,n_2,n_3}\left[T_1,T_2 \middle| 2\right] \leq P_{n_1,n_2,n_3}\left[T_1 \middle| 2\right] \cdot P_{n_1,n_2,n_3}\left[T_2 \middle| 2\right].$$

Further

$$(2.8.11) \qquad P_{n_1,n_2,n_3}\left[T_1,T_2 \middle| 3\right] = P_{n_1,n_2,n_3-1}\left[T_1,T_2\right] \leq$$

$$\leq P_{n_1,n_2,n_3-1}\left[T_1\right] \cdot P_{n_1,n_2,n_3-1}\left[T_2\right] =$$

$$= P_{n_1,n_2,n_3}\left[T_1 \middle| 3\right] \cdot P_{n_1,n_2,n_3}\left[T_2 \middle| 3\right].$$

So, combining (2.8.8), (2.8.9), (2.8.10) and (2.8.11) we find, dropping the subscripts

$$(2.8.12) \qquad P\left[T_1,T_2\right] \leq \sum_{i=1}^{3} \frac{n_i}{N} P\left[T_1 \middle| i\right] \cdot P\left[T_2 \middle| i\right].$$

We now define

$$x_i \overset{\text{def}}{=} P\left[T_1 \middle| i\right] \qquad (i = 1,\dots,3) \tag{2.8.13}$$

and

$$y_i \overset{\text{def}}{=} P\left[T_2 \middle| i\right] \qquad (i = 1,\dots,3) \tag{2.8.14}$$

and consider $\underline{x}$ as a random variable which takes the values x_i and $\underline{y}$ as a random variable which takes the values y_i, both with probabilities

$$p_i \overset{\text{def}}{=} \frac{n_i}{N} \qquad (i = 1,\dots,3). \tag{2.8.15}$$

Now we have proved (2.8.12)

$$P\left[T_1, T_2\right] \leq \sum_1^3 p_i x_i y_i = \mathcal{E}\, \underline{x}\, \underline{y}. \tag{2.8.16}$$

Further we have

$$\left\{ \begin{aligned} P\left[T_1\right] &= \sum_1^3 p_i x_i = \mathcal{E}\, \underline{x}, \\ P\left[T_2\right] &= \sum_1^3 p_i y_i = \mathcal{E}\, \underline{y}. \end{aligned} \right. \tag{2.8.17}$$

Therefore our proof is completed when we have proved

$$\mathcal{E}\, \underline{x}\, \underline{y} \leq \mathcal{E}\, \underline{x}\; \mathcal{E}\, \underline{y}. \tag{2.8.18}$$

We have

$$x_2 = x_3 \tag{2.8.19}$$

and

$$y_1 = y_3. \tag{2.8.20}$$

Further we can prove that

$$x_1 \geq x_2 \tag{2.8.21}$$

and its equivalent

(2.8.22) $$y_1 \leq y_2.$$

That (2.8.21) holds can be seen in the following way. (2.8.21) is equivalent to

(2.8.23) $$n_1 \; P\left[T_1, 2\right] \leq n_2 \; P\left[T_1, 1\right],$$

where

$P\left[T_i, l\right] \overset{def}{=}$ the probability that $\underline{T}_i \geq T_i$ and that the largest observation belongs to the l-th sample.

Consider now a ranking which gives T_1 and 2 (i.e. the largest element belongs to the 2nd sample and $\underline{T}_1 \geq T_1$). From this ranking we get one with T_1 and 1 by interchanging the largest element with an element of the first sample. In this way we get $n_1 N(T_1, 2)$ rankings with T_1 and 1, when we denote by $N(T_1, 2)$ the number of rankings that give T_1 and 2. Among this set each ranking with T_1 and 1 can occur at most n_2 different times, because in a ranking with T_1 and 1 the largest element can be interchanged at most with n_2 different elements of the second sample under the condition that it becomes one with T_1 and 2.
Therefore

(2.8.24) $$n_1 N(T_1, 2) \leq n_2 N(T_1, 1),$$

which is equivalent with (2.8.23).
We have therefore

$$x_1 \geq x_3 = x_2$$

and

$$y_1 = y_3 \leq y_2.$$

Thus lemma 2.8.1 can be applied, which proves (2.8.18) and with that theorem 2.8.1.

Chapter 3

TESTS FOR MORE THAN ONE OUTLIER

3.1. INTRODUCTION

One line of attack in case of more than one suspected outlier is to apply one of the tests described in Chapter 2 more times, that is, after rejecting for instance the largest observation, testing the second one and so on.

Clearly the probability of an error of the first kind of this composite test is the same as for the simple test for one outlier, if the null hypothesis is that no outlier is present and if an error of the first kind means deciding that one or more variates have slipped when H_0 is true.

The whole procedure can be applied also in a two-sided version, i.e. for outliers to the right and to the left simultaneously.

A serious drawback of this procedure is that it may never get started because the test for one outlier is not very powerful when more outliers are present, especially when these outliers all have the same parameter value $\theta + \Delta$.

In the case of a normal distribution this was already pointed out by E.S. PEARSON and C. CHANDRA SEKAR (1936). It is therefore more efficient to devise special tests for two and also for three and more outliers. We have to bear in mind however that when a test is selected from a set of tests after the observations have been made, the critical region of the test procedure consists of the sum of the critical regions of all the tests of the set. We propose therefore to adopt the following rule:

Admit among the parameters

$$\theta_1, \ldots, \theta_k$$

a maximum number of m_k outliers either to the right, or to the left or to both sides *). Apply separate tests for 1, 2, ... and m_k outliers. Calculate the probabilities for the m_k test statistics to exceed their observed values. If the smallest of these P-values goes with the test for m_0 outliers and if this probability is smaller than or equal to α/m_k, the null hypothesis is rejected in favour of the alternative that m_0 outliers are present.

The probability P of rejecting H_0 when H_0 is true, is, according to BONFERRONI's inequality smaller than or equal to the sum of the probabilities that the individual P-values are smaller than or equal to α/m_k, i.e.

$$P \leq m_k \cdot \alpha/m_k = \alpha.$$

Now we need a test for any given number, say m, outliers. For this purpose we propose the following method, which can be expected to attain its maximum power when m of the variates $\vec{\underline{y}}_i$ of section 2.1 have the same parameter $\theta + \Delta$.

The variates $\vec{\underline{y}}_i$ are divided in all $\binom{k}{m}$ possible ways into two groups of (k-m) and m variates. An appropriate two-sample test is applied to all $\binom{k}{m}$ pairs of samples. The smallest of the $\binom{k}{m}$ P-values is multiplied by $\binom{k}{m}$, which gives an upper limit to the P-value of the slippage test.

Now we only know an upper bound to the probability of rejecting H_0 when H_0 is true, we cannot give a lower bound as in the case of one outlier.

3.2. THE CASE OF THE NORMAL DISTRIBUTION

We consider the same k normal distributions as described in section 2.3.

---

*) For $m_k = 1$ we obtain the tests described in Chapter 2 as a special case.

The same null hypothesis

$$H_0: \mu_1 = \mu_2 = \dots = \mu_k = \mu \tag{3.2.1}$$

is tested, now however against the alternatives

$$H_1: \mu_j = \mu + \Delta;\ j \in I_0,\ \Delta > 0, \qquad \mu_j = \mu \quad ;\ j \notin I_0, \tag{3.2.2}$$

where

$$I_0 \overset{def}{=} \{i_1, \dots, i_m\},$$

for m unknown values $i_1, \dots, i_m \in \{1,\dots,k\}$, or

$$H_2: \mu_j = \mu - \Delta;\ j \in I_0,\ \Delta > 0, \qquad \mu_j = \mu \quad ;\ j \notin I_0. \tag{3.2.3}$$

Now we consider instead of the variables $\underline{t}_i$ of (2.3.7) the variables

$$\underline{t}_I. \tag{3.2.4}$$

The statistic $\underline{t}_I$ is the test-statistic of STUDENT's two sample test when testing the observations of the samples $i_1, i_2, \dots, i_m$, considered as one sample, against all the other observations considered as the second sample. It is clear that $\underline{t}_I$ has a STUDENT's t-distribution with N-2 degrees of freedom.

When testing against the alternative (3.2.2) the values

$$d_I = P\left[\underline{t}_I \geq t_I\right] \tag{3.2.5}$$

are calculated and H_0 is rejected when the smallest of these d-values is smaller than $\alpha / \binom{k}{m}$.

When testing against the alternative (3.2.3), the values

$$e_I = P\left[\underline{t}_I \leq t_I\right] \tag{3.2.6}$$

are considered.

As an example let us consider the following case presented by F.E. GRUBBS (1950):

"The following ranges (horizontal distances from gun muzzle to point of impact) were obtained in firing projectiles from a weapon at a constant angle of elevation and at the same weight of charge of propellant powder.

Distances in yards (arranged in increasing order of magnitude). 4420, 4549, 4730, 4765, 4782, 4803, 4833, 4838".

The first two observations are suspected to be outliers. Applying a t-test to these two values against the six others gives a t-value of 7.09. The probability of exceeding this value for a t with 6 degrees of freedom lies between 0.0001 and 0.00025 (cf. the tables of FEDERIGHI (1959)). Multiplying these values with $\binom{8}{2} = 28$ gives an upper bound for the P-value between 0.0028 and 0.0070.

Application of the criterion of GRUBBS gives a result which is significant at the 0.01 level.

In this example the maximum number of admitted outliers m_k was not stated. It can be argued that in this case, with 8 observations, the maximum number of outliers is 3. This means that the P-values for both tests should be multiplied by 3 or even by 6 in the two-sided case.

When only the smallest value is tested with the test for one outlier we get a t-value of 3.17 which is near the percentage point 0.01. This value has to be multiplied by 8, so the lowest value would not be rejected at the 0.05 level, which illustrates the point made in section 3.1.

In the case where, as in this example, all n_i are equal to 1 we can compare our test criterion with the one given by F.E. GRUBBS (1950).

Let us consider the case of two outliers to the right and denote the k variables

(3.2.7) $$\underline{y}_1, \ldots, \underline{y}_k$$

arranged in ascending order by

(3.2.8) $$\underline{z}_1, \ldots, \underline{z}_k.$$

We define

$$(3.2.9)\qquad \begin{cases} \bar{\underline{z}}_1 \overset{\text{def}}{=} \frac{1}{k-2} \sum_{1}^{k-2} \underline{z}_i \\ \bar{\underline{z}}_2 \overset{\text{def}}{=} \frac{1}{2} (\underline{z}_{k-1} + \underline{z}_k) \\ \bar{\underline{z}} \overset{\text{def}}{=} \frac{1}{k} \sum_{1}^{k} \underline{z}_i \end{cases}$$

and

$$(3.2.10)\qquad \begin{cases} \underline{S}_1^2 = \sum_{1}^{k-2} (\underline{z}_i - \bar{\underline{z}}_1)^2 \\ \underline{S}_2^2 = \sum_{k-1}^{k} (\underline{z}_i - \bar{\underline{z}}_2)^2 \\ \underline{S}_3^2 = \frac{2(k-2)}{k} (\bar{\underline{z}}_2 - \bar{\underline{z}}_1)^2 . \end{cases}$$

Now our test criterion reads

$$(3.2.11)\qquad \underline{T} = (k-2)\sqrt{\frac{2}{k}}\,\frac{\bar{\underline{z}}_2 - \bar{\underline{z}}_1}{\sqrt{\underline{S}_1^2 + \underline{S}_2^2}} ,$$

whereas GRUBBS' test statistic is

$$(3.2.12)\qquad \underline{G} = \frac{\underline{S}_2^2}{\underline{S}^2} ,$$

small values being rejected. After some calculations we find

$$(3.2.13)\qquad 1 + \frac{\underline{T}^2}{k-2} = \frac{\underline{S}_1^2 + \underline{S}_2^2 + \underline{S}_3^2}{\underline{S}_1^2 + \underline{S}_2^2}$$

and

$$\frac{1}{\underline{G}} = \frac{\underline{S}_1^2 + \underline{S}_2^2 + \underline{S}_3^2}{\underline{S}_1^2} \qquad (3.2.14)$$

This gives some insight in the relative powers of the two methods in some situations. When the two outliers come closer together, their average remaining in the same position, $\underline{S}_2^2$ decreases and both $\underline{S}_1^2$ and $\underline{S}_3^2$ remain constant. This makes that $\underline{T}$ increases, making the power larger, but $\frac{1}{\underline{G}}$ decreases making this test less sensitive.

3.3. THE TESTS FOR THE GAMMA DISTRIBUTION

Here we consider again the variates

$$\underline{u}_1, \ldots, \underline{u}_k \qquad (3.3.1)$$

of (2.5.1) which have gamma distributions with parameters $\varepsilon_1, \beta_1; \ldots; \varepsilon_k, \beta_k$.

The hypothesis H_0

$$H_0: \beta_1 = \ldots = \beta_k = \beta \qquad (3.3.2)$$

is now tested against the alternatives

$$H_1: \beta_j = \beta \text{ for } j \notin I_0, \text{where } I_0 = \{i_1, \ldots, i_m\} \qquad (3.3.3)$$

$$\beta_j = C\beta \text{ for } j \in I_0, C > 1$$

for m unknown values $i_1, \ldots, i_m$ and

$$H_2: \beta_j = \beta \text{ for } j \notin I_0 \qquad (3.3.4)$$

$$\beta_j = c\beta \text{ for } j \in I_0, c < 1$$

for m unknown values $i_1, \ldots, i_m$.

For both tests the ratios

$$\underline{x}_I = \frac{\sum_{j \in I} \underline{u}_j}{\sum_1^k \underline{u}_i} \qquad (3.3.5)$$

are computed.

The distribution of $\underline{x}_I$ is

$$(3.3.6)\quad f(x) = \frac{\Gamma(A)}{\Gamma(\sum_{j\in I} \varepsilon_j)\Gamma(A-\sum_{j\in I} \varepsilon_j)} x^{\sum_{j\in I} \varepsilon_j} (1-x)^{A-\sum_{j\in I} \varepsilon_j - 1},$$

where $A \overset{\text{def}}{=} \sum_1^k \varepsilon_i$.

When testing against the alternative H_1 the values

$$d_I \overset{\text{def}}{=} P\left[\underline{x}_I > x_I\right]$$

are calculated for all sets I. H_0 is rejected when the smallest d-value is smaller than or equal to $\alpha/\binom{k}{m}$. A similar procedure is followed for testing against H_2.

3.4. THE CASE OF THE POISSON DISTRIBUTION

The null hypothesis H_0 states that the variates

$$(3.4.1)\qquad \underline{z}_1, \ldots, \underline{z}_k$$

(cf. 2.6.13) are distributed according to Poisson distributions with means μ_i which have the following ratios

$$(3.4.2)\quad H_0: \frac{\mu_i}{\sum \mu_j} = p_i \quad (i = 1, \ldots, k).$$

The alternatives are

$$(3.4.3)\quad H_1: \frac{\mu_i}{\sum \mu_j} = Cp_i \text{ for } i \in I_0, \text{ where } I_0 = \{i_1,\ldots,i_m\}$$

$$\frac{\mu_i}{\sum \mu_j} = \frac{1 - Cp_{I_0}}{1 - p_{I_0}}, \text{ for } i \notin I_0, \text{ with } p_{I_0} \overset{\text{def}}{=} \sum_{i\in I_0} p_i$$

and $1 < C < 1/p_{I_0}$, for m unknown values $i_1,\ldots,i_m$

and:

$$(3.4.4)\quad H_2: \frac{\mu_i}{\sum \mu_j} = cp_i \text{ for } i \in I_0$$

$$\frac{\mu_i}{\sum \mu_j} = \frac{1 - cp_{I_0}}{1 - p_{I_0}}, \text{ for } i \notin I_0,\ 0 < c < 1,$$

for m unknown values $i_1, \dots, i_m$.

If the observations are $z_1, \dots, z_k$, with $\sum z_i = N$, we calculate, when testing against H_1

$$(3.4.5)\quad r_I = P\left[\underline{z}_I \geq z_I \;\middle|\; \sum \underline{z}_i = N\right] = \sum_{x=z_I}^{N} \binom{N}{x} p_I^x (1-p_I)^{N-x} =$$

$$= I_{p_I}(z_I, N-z_I+1) \quad \text{(cf. 2.6.20)},$$

where $z_I \overset{\text{def}}{=} \sum_{i \in I} z_i$.

If min $r_I \leq \alpha/\binom{k}{m}$, H_0 is rejected. If under H_0 $\mu_1 = \dots = \mu_k$, all p_i are equal and the smallest r_I corresponds to the set I consisting of the indices of the largest m observations.

The test against the alternative H_2 is completely analogous.

For the tests against two outliers with H_0: $\mu_1 = \dots = \mu_k$ exact critical values have been calculated on an I.B.M. 1401 computer at the Computer Department of Unilever at Rotterdam. These critical values are presented in the tables III and IV.

It is interesting to compare these exact critical values with those obtained by our approximative method. In the case of two outliers to the right, a small number of critical values would have been 1 higher when using the approximation. These values are indicated by an asterisk in table III.

In table IV the values which would have been 1 lower when using the approximation are indicated in the same way.

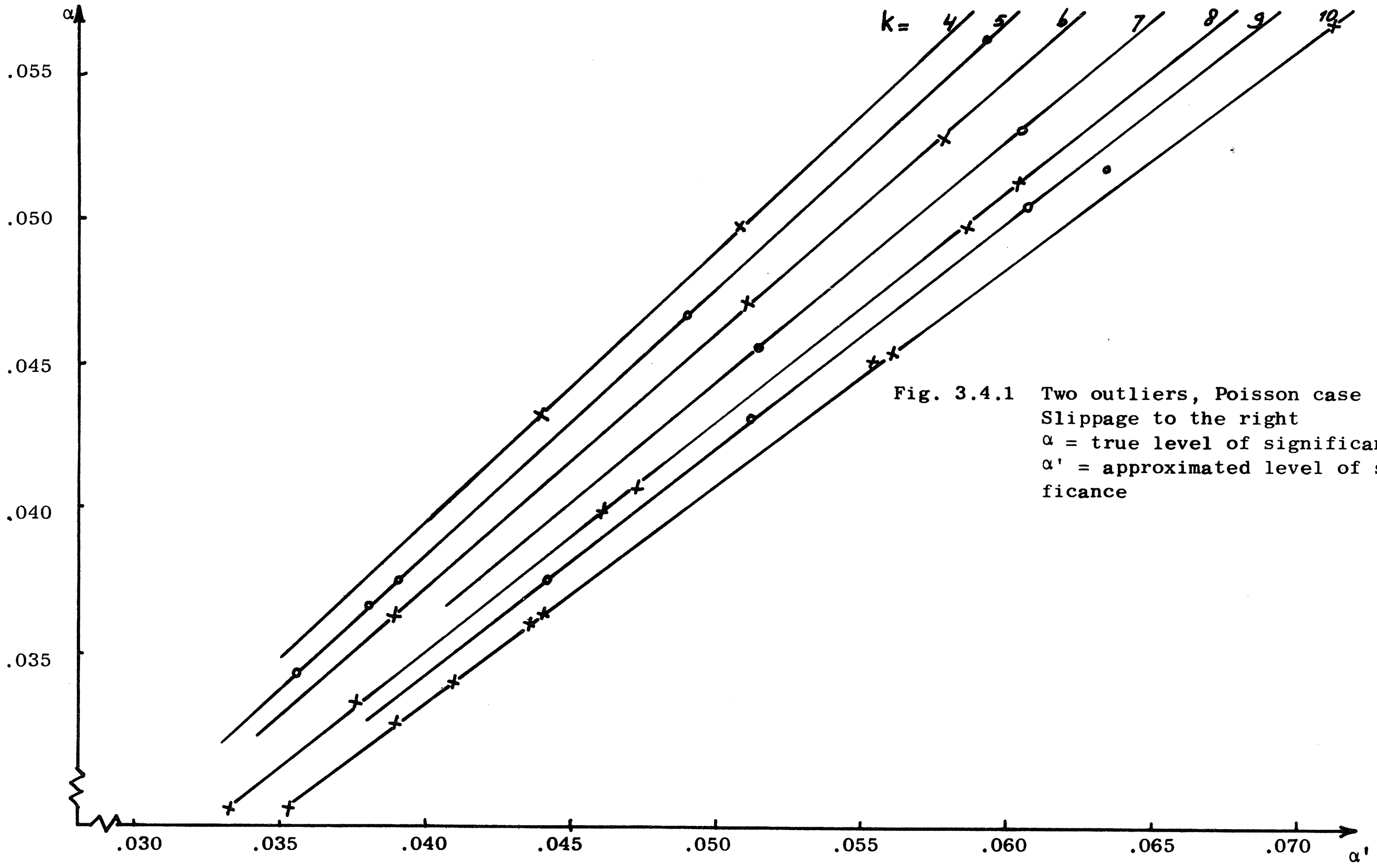

Fig. 3.4.1 Two outliers, Poisson case
Slippage to the right
α = true level of significance
α' = approximated level of significance

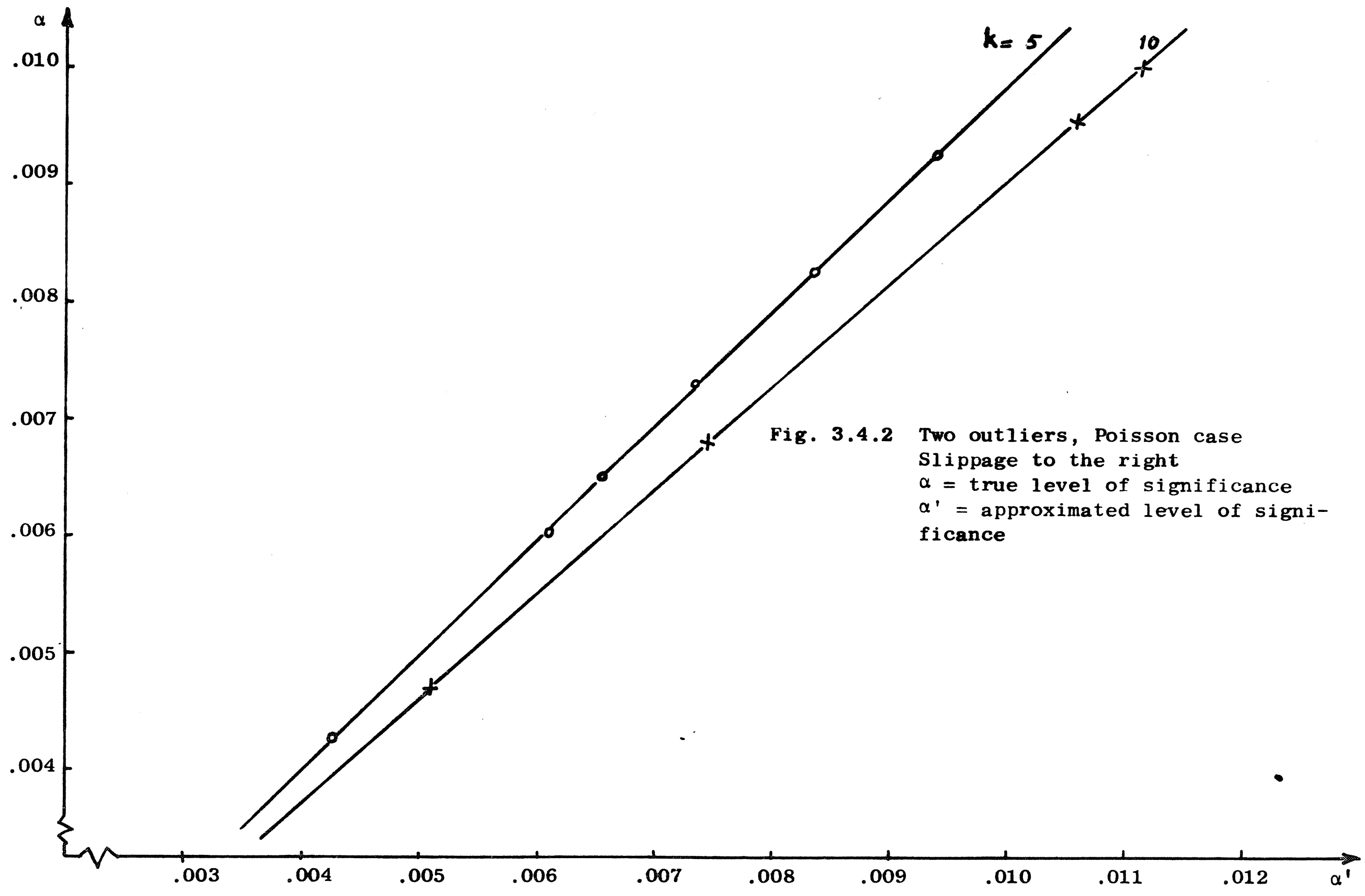

Fig. 3.4.2 Two outliers, Poisson case
Slippage to the right
α = true level of significance
α' = approximated level of significance

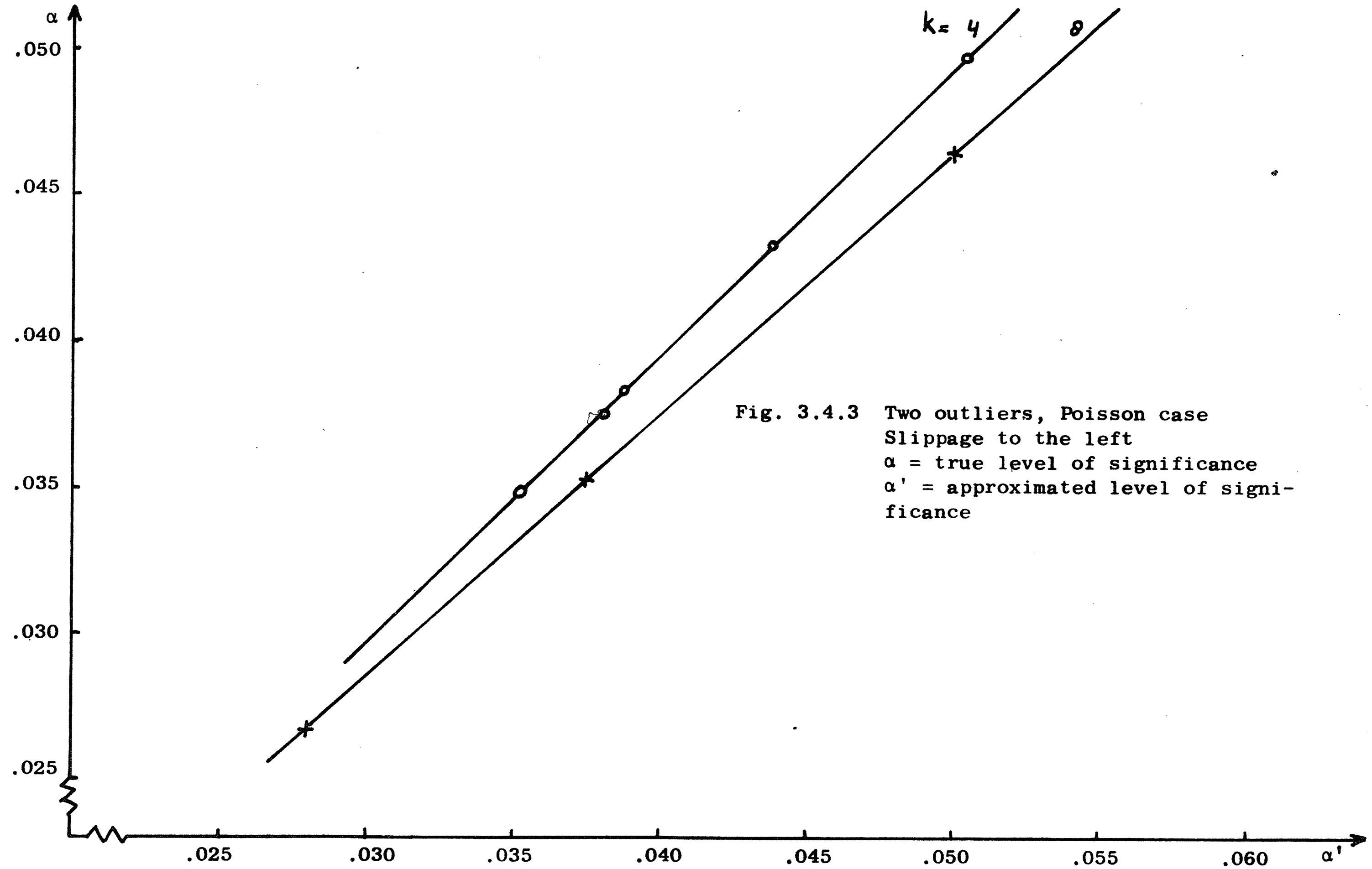

Fig. 3.4.3 Two outliers, Poisson case
Slippage to the left
α = true level of significance
α' = approximated level of significance

A closer analysis learns that there exists a relation between the exact values $P\left[\underline{z}_1 + \underline{z}_2 \geq z \mid \sum \underline{z}_i = N\right]$, where $\underline{z}_1$ and $\underline{z}_2$ respectively are the largest and the second largest observation respectively and between the approximation, which depends only on k and not on N. A similar conclusion can be drawn for the P-values to the left-hand side.

In the figures 3.4.1, 3.4.2 and 3.4.3 both cases have been plotted. It appears that the approximation grows worse when k increases. Further we see that the approximation is better in the case of slippage to the left than in the case of slippage to the right.

In the latter case for k = 10 the real size α of the test is 0.041 when the approximation α' is 0.050.

It is possible now to get exact critical values for larger values of N and for k up to 10 by using as approximate significance levels the values corresponding to an exact level of 0.05 or 0.01 read off from the graphs.

3.5. OTHER DISCRETE VARIATES

The generalisations of the tests for binomial (cf. 2.6.26) and negative binomial (cf. 2.6.32) variates are obvious.

Like in the case of the Poisson variates the test statistics are the sums of the suspected observations. In the binomial case the conditional distribution of the sum of a number of variates is again hypergeometric and in the negative binomial case the conditional distribution of a sum

$$\underline{w}_I \overset{\text{def}}{=} \sum_{i \in I} \underline{w}_i$$

under the condition

$$\sum_{1}^{k} \underline{w}_i = N,$$

is, analogous to (2.6.38)

$$(3.5.1)\quad P\left[\underline{w}_I = w \;\middle|\; \sum_1^k \underline{w}_i = N\right] = \frac{\binom{w_I+r_I-1}{r_I-1}\binom{N+\sum r_j-w_I-r_I-1}{\sum r_j-r_I-1}}{\binom{N+\sum r_j-1}{\sum r_j-1}} .$$

3.6. THE METHOD OF M RANKINGS

As in section 2.7 we consider k "objects" which have been ranked by m "observers". The sums of the m ranks are denoted by (cf. 2.7.1)

$$(3.6.1)\qquad \underline{s}_1, \ldots, \underline{s}_k \qquad (m \le \underline{s}_i \le km),$$

with

$$(3.6.2)\qquad \sum_1^k \underline{s}_i = \frac{1}{2}\, mk(k+1).$$

The distribution of $\underline{s}_i$ under the hypothesis H_0 that all m rankings are independent of each other and chosen at random from all permutations of 1, ..., k has been derived in section 2.7.

For a test against more outliers we need the distribution of a sum of a number, say r, of the variates $\underline{s}_i$.

We propose to use the normal approximation of this distribution. It is easy to see that

$$(3.6.3)\qquad \begin{cases} \mathcal{E}\,\underline{s}_i = \frac{1}{2}\, m(k+1), \\ \mathcal{E}\,\underline{s}_i^2 - (\mathcal{E}\,\underline{s}_i)^2 = \frac{1}{12}\, m(k^2-1), \\ \rho(\underline{s}_i, \underline{s}_j) = -1/(k-1). \end{cases}$$

Therefore the sum of r variates $\underline{s}_{i_1}, \ldots, \underline{s}_{i_r}$ has the mean value

$$(3.6.4)\qquad \frac{1}{2}\, mr(k+1)$$

and the variance

$$(3.6.5)\qquad \left[r - \frac{r(r-1)}{k-1}\right] \cdot \frac{1}{12}\, m(k^2-1) = \frac{1}{12}\, mr(k-r)(k+1).$$

We now have a test for any number r of outliers by taking as a test statistic the sum of the largest, or smallest, r observations and applying the procedure described in section 3.1.

3.7. THE DISTRIBUTIONFREE k-SAMPLE TEST

We consider the independent variates from (2.8.1):

$$\underline{u}_1, \ldots, \underline{u}_k, \tag{3.7.1}$$

with the same continuous distribution, from which $t_1, \ldots, t_k$ observations have been taken.

We want to test H_0 against the alternatives

$$H_1: \begin{cases} P\left[\underline{u}_i > \underline{u}_j\right] > \frac{1}{2},\ i \in I_0,\ I_0 \stackrel{\text{def}}{=} \{i_1, \ldots, i_m\} \\ \underline{u}_j,\ j \notin I_0 \text{ follow the same distribution,} \end{cases}$$

for m unknown values $i_1, \ldots, i_m$, or

$$H_2: \begin{cases} P\left[\underline{u}_i < \underline{u}_j\right],\ i \in I_0,\ j \notin I_0 \\ \underline{u}_j,\ j \notin I_0 \text{ follow the same distribution,} \end{cases}$$

for m unknown values $i_1, \ldots, i_m$.

We consider now the statistics $\underline{T}_I$, the sums of the ranks of the observations of the samples $i_1, \ldots, i_m$, for all sets I.

When the smallest of the values

$$d_I \stackrel{\text{def}}{=} P\left[\underline{T}_I \geq T_I\right]$$

is smaller than $\alpha / \binom{k}{m}$ H_0 is rejected in favour of H_1.

When testing against H_2 a similar procedure is followed.

Chapter 4

OPTIMUM PROPERTIES AND EFFICIENCY

4.1. INTRODUCTION

In this chapter the slippage problems for one outlier for normal, gamma, Poisson, binomial and negative binomial distributions are formulated as multiple decision problems and it is shown that the tests described in chapter 2 have certain optimum properties.

The method used was developed by E. PAULSON (1952) and also used by D.R. TRUAX (1953).

For the non-parametric cases of m rankings and of the k sample test the consistency is proved and the asymptotic relative efficiency is calculated as compared with the methods based on normal distribution.

This chapter is partly based on a report of the Mathematical Centre by R. DOORNBOS, H. KESTEN and H.J. PRINS (1956).

4.2. THE SLIPPAGE PROBLEM FOR THE NORMAL DISTRIBUTION

For the test for the normal distribution as described in section 2.3 PAULSON (1952) proved the following optimum property in the special case $n_1 = \ldots = n_k = n$.

Let D_0 denote the decision that the k means are all equal, and let D_j $(j = 1, 2, \ldots, k)$ denote the decision that D_0 is incorrect and that $\mu_j = \max(\mu_1, \ldots, \mu_k)$.

Now the procedure

$$\text{(4.2.1)} \qquad \text{If } \underline{b}_m > \lambda_\alpha \text{ select } D_m, \qquad \text{if } \underline{b}_m \leq \lambda_\alpha \text{ select } D_0,$$

where m is the index of the maximum $\underline{b}$-value ($\underline{b}$ defined by (2.3.5)), maximizes the probability of making a correct decision, subject to the following restrictions:

(a) when all means are equal, D_0 should be selected with probability $1-\alpha$,

(b) the decision procedure must be invariant if a constant is added to all the observations,

(c) the decision must be invariant when all the observations are multiplied by a positive constant, and

(d) the decision procedure must be symmetric in the sense that the probability of making a correct decision when the i-th mean has slipped to the right by an amount Δ must be the same for $i = 1, 2, \ldots, k$.

The constant λ_α in (4.2.1) is determined by requirement (a). Our critical value for $\underline{b}_m$ is an approximation of λ_α. The case of slippage to the left is completely analogous and the same optimum property holds there.

The main part of the proof consists in showing that for any Δ and σ there exists a set of nonzero a priori probabilities $g_0, g_1, \ldots, g_k$ which are functions of Δ and σ so that the procedure (4.2.1) will maximize the probability of making the correct decision among the set $(D_0, D_1, \ldots, D_k)$ when g_i is the a priori probability that D_i is the correct decision.

Assuming that this has been demonstrated, it follows easily that (4.2.1) must be the optimum solution. For suppose there existed another allowable procedure, which for some Δ and σ has a greater probability than (4.2.1) of making the correct decision when some mean had slipped to the right by an amount Δ. Then this procedure would have a greater probability than (4.2.1) of making the correct decision when D_i is the right one for $i = 1, \ldots, k$ (according to (d)) and the same probability of choosing D_0 rightly (because of (a)).

Thus this other procedure would have a greater probability of making the correct decision than (4.2.1) with respect to any set of a priori probabilities, which would be a contradiction.

4.3. THE SLIPPAGE PROBLEM FOR THE GAMMA DISTRIBUTION

Using the method of PAULSON, D.R. TRUAX (1953) proved a similar optimum property for the slippage test for estimated normal variances of W.G. COCHRAN (1941). Consequently our tests for the gamma-distribution as described in section 2.5 are optimal in the following sense if $\varepsilon_1 = \ldots = \varepsilon_k$.

Let D_0 be the decision that H_0 is true and let D_j be the decision that H_0 is false and that $\beta_j = \max(\beta_1, \ldots, \beta_k)$.
Then the procedure

$$\text{(4.3.1)} \qquad \begin{aligned} &\text{If } \underline{x}_m > \lambda_\alpha \text{ select } D_m, \\ &\text{if } \underline{x}_m \leq \lambda_\alpha \text{ select } D_0, \end{aligned}$$

where m is the index of the maximum $\underline{x}$-value ($\underline{x}$ defined by (2.5.6)), maximizes the probability of making the correct decision, subject to the following conditions:

(a) when H_0 is true, D_0 should be selected with probability $1 - \alpha$ (this determines λ_α),
(b) the decision procedure must be invariant when all the observations are multiplied by a positive constant,
(c) the decision procedure must be symmetric in the sense that the probability of making a correct decision when the i-th parameter β_i is multiplied by C (C > 1) must be the same for i = 1, ..., k. Here again the case of slippage to the left is analogous.

4.4. AN OPTIMUM PROPERTY OF THE TESTS FOR DISCRETE VARIATES

We consider the independent discrete random variables

$$\underline{x}_1, \ldots, \underline{x}_k,$$

which have distributions $P(\underline{x}_i = x) = h_i(x) \quad (i = 1, \ldots, k)$.
We want to test the hypothesis H_0:

$$H_0\colon h_1(x) = \ldots = h_k(x) = h(x)$$

against the alternatives

$$H_i\colon h_1(x) = \ldots = h_{i-1}(x) = h_{i+1}(x) = \ldots = h_k(x) = h(x),$$

$$h_i(x) = h^*(x) \neq h(x), \text{ for some unknown value } i.$$

We denote by D_0 the decision that H_0 is true and by D_i the decision that H_i is true.

Now the following theorem holds:

THEOREM 4.1

Under the condition that $\sum \underline{x}_i = N$ and if

$$\frac{h^*(x)}{h(x)} \tag{4.4.1}$$

is an increasing function of x, the procedure

$$\begin{array}{l} \text{if } \underline{x}_m \geq \lambda_{\alpha,N} \text{ select } D_i, \\ \text{if } \underline{x}_m < \lambda_{\alpha,N} \text{ select } D_0, \end{array} \tag{4.4.2}$$

where m is the index of the maximum $\underline{x}$-value, maximizes the probability of making a correct decision subject to the following restrictions:

(a) when H_0 is true, D_0 should be selected with probability $\geq 1-\alpha$,

(b) the probability of selecting D_i when H_i is true must be the same for $i = 1, \ldots, k$.

PROOF

As indicated in section 4.2 in the case of the normal distribution the proof consists of finding a set of a priori probabilities under which the procedure 4.4.2 maximizes the probability of making the correct decision. According to A. WALD (1950), p. 128, the optimum solution relative to a set of a priori probabilities $g_0, g_1, \ldots, g_k$ is given by the rule: "For each j (j = 0, 1, ..., k) decide D_j for all points in the sample space where j is the smallest *) integer for which $g_j f_j = \max(g_0 f_0, g_1 f_1, \ldots, g_k f_k)$, where f_j is the joint elementary probability law of $\underline{x}_1, \ldots, \underline{x}_k$ under the hypothesis H_j".

*) Compare the note on page 40

We consider the special a priori distribution

$$g_0 = 1 - kg; \; g_1 = \ldots = g_k = g.$$

According to WALD the region where e.g. D_1 is selected is given by the points in the sample space where $f_1 > f_i$ $(i = 2, \ldots, k)$ and $gf_1 > (1-kg)f_0$, each f_i $(i = 0, \ldots, k)$ being computed under the relevant hypothesis H_i. The region where $f_1 > f_i$ is given by

$$(4.4.3) \qquad \frac{P[\underline{x}_1 = x_1 | H_1] \ldots P[\underline{x}_k = x_k | H_1]}{P[\sum \underline{x}_i = N | H_1]} > \frac{P[\underline{x}_1 = x_1 | H_i] \ldots P[\underline{x}_k = x_k | H_i]}{P[\sum \underline{x}_i = N | H_i]}.$$

Because $\underline{x}_1, \ldots, \underline{x}_k$ have the same distribution under H_0 and on account of the form of the hypotheses H_i for $i \neq 0$ we have

$$(4.4.4) \qquad \begin{cases} P\left[\sum \underline{x}_i = N \;\middle|\; H_j\right] \text{ is the same for } j = 1, \ldots, k; \\ P\left[\underline{x}_i = x \;\middle|\; H_j\right] = P\left[\underline{x}_i = x \;\middle|\; H_0\right] \text{ for } j \neq i; \\ P\left[\underline{x}_t = x \;\middle|\; H_i\right] = P\left[\underline{x}_t = x \;\middle|\; H_j\right] \text{ for } t = 1, \ldots, k;\; t \neq i, j. \end{cases}$$

With the help of these relations (4.4.3) reduces to

$$(4.4.5) \qquad P\left[\underline{x}_1 = x_1 \middle| H_1\right] P\left[\underline{x}_i = x_i \middle| H_1\right] > P\left[\underline{x}_1 = x_1 \middle| H_i\right] P\left[\underline{x}_i = x_i \middle| H_i\right]$$

or

$$h^*(x_1) h(x_i) > h(x_1) h^*(x_i),$$

or

$$(4.4.6) \qquad \frac{h^*(x_1)}{h(x_1)} > \frac{h^*(x_i)}{h(x_i)},$$

which is, according to condition (4.4.1), equivalent to

$$x_1 > x_i.$$

The region where $gf_1 > (1-kg)f_0$ is given by

$$(4.4.7) \qquad g \frac{P[\underline{x}_1 = x_1 | H_1] \ldots P[\underline{x}_k = x_k | H_1]}{P[\sum \underline{x}_i = N | H_1]} > (1-kg) \frac{P[\underline{x}_1 = x_1 | H_0] \ldots P[\underline{x}_k = x_k | H_0]}{P[\sum \underline{x}_i = N | H_0]}$$

which is, according to (4.4.4), the same as

$$g \cdot \frac{P[\underline{x}_1 = x_1 \mid H_1]}{P[\sum \underline{x}_i = N \mid H_1]} > (1-kg) \frac{P[\underline{x}_1 = x_1 \mid H_0]}{P[\sum \underline{x}_i = N \mid H_0]} ,$$

or as

$$\text{(4.4.8)} \qquad \frac{h^*(x_1)}{h(x_1)} > \frac{1-kg}{g} \frac{P[\sum \underline{x}_i = N \mid H_1]}{P[\sum \underline{x}_i = N \mid H_0]} .$$

In virtue of (4.4.1) this is equivalent to $x_1 > L$, where L is a number depending on N. Thus the Bayes solution is: if m is the smallest integer for which x_m is the maximum of $x_1, \ldots, x_k$ select D_m if $x_m > L$, otherwise select D_0. Define the function F(g) by the equation

$$\text{(4.4.9)} \qquad F(g) = \frac{P[\underline{x}_1 = \lambda_{\alpha,N} \mid H_1]}{P[\underline{x}_1 = \lambda_{\alpha,N} \mid H_0]} - \frac{1-kg}{g} \cdot \frac{P[\sum \underline{x}_i = N \mid H_1]}{P[\sum \underline{x}_i = N \mid H_0]} ,$$

where $\lambda_{\alpha,N}$ is the constant used in (4.4.2). Obviously F(g) is a continuous function of g, with $F\left(\frac{1}{k}\right) > 0$.

Further F(g) = 0, for $g = g^*$ satisfying

$$0 < g^* = \frac{P[\sum \underline{x}_i = N | H_1]}{P[\sum \underline{x}_i = N | H_0]} \left\{ \frac{P[\underline{x}_1 = \lambda_{\alpha,N} | H_1]}{P[\underline{x}_1 = \lambda_{\alpha,N} | H_0]} + k \cdot \frac{P[\sum \underline{x}_i = N | H_1]}{P[\sum \underline{x}_i = N | H_0]} \right\}^{-1} < \frac{1}{k} .$$

To get the Bayes solution relative to $(1-kg^*, g^*, \ldots, g^*)$ it is only necessary to replace L by $\lambda_{\alpha,N}$ in the solution given above.

Thus the procedure (4.4.2) is the Bayes solution relative to $(1-kg^*, g^*, \ldots, g^*)$, which proves that it is optimal.

REMARK

A completely analogous theorem can be stated, where in (4.4.2) the maximum x-value is replaced by the minimum value, the signs are reversed and (4.4.1) is decreasing instead of increasing.

4.5. THE SLIPPAGE TESTS FOR THE POISSON, THE BINOMIAL AND THE NEGATIVE BINOMIAL DISTRIBUTION

Let us first consider the Poisson case (cf. 2.6.13), with under H_0 (2.6.14) $p_1 = \ldots = p_k = \frac{1}{k}$, in other words with equal means. We cannot apply Theorem 4.1 directly in this case because of the way in which the hypotheses (2.6.14) and (2.6.15) are formulated. Rather than reformulating the hypotheses we apply WALD's rule directly here.

The joint distribution of $\underline{z}_1, \ldots, \underline{z}_k$ under H_0 and H_{1i} respectively are

$$(4.5.1) \quad f^0(z_1, \ldots, z_k) = \frac{N!}{\Pi z_i!} \left(\frac{1}{k}\right)^N$$

and

$$(4.5.2) \quad f_i(z_1, \ldots, z_k) = \frac{N!}{\Pi z_i!} \left(\frac{1}{k}\right)^N C^{z_i} \left(\frac{k-C}{k-1}\right)^{N-z_i} \quad (1 < C < \frac{1}{k}).$$

Because

$$(4.5.3) \quad C^{z_i} \left(\frac{k-C}{k-1}\right)^{N-z_i}$$

is increasing in z_i for $C > 1$, the region where $f_1 > f_i$ is given by $z_1 > z_i$ and the region where $gf_1 > (1-kg)f_0$ by $z_1 > L$, L depending on N and C. Therefore WALD's rule may be applied in the same way as was done in the proof of theorem 4.1 and therefore our Poisson-test is optimal in the sense described in this theorem.

In the case of the binomial distribution (2.6.26) with $n_1 = \ldots = n_k$, theorem 4.1 can be applied directly. Condition (4.4.1) gives here

$$(4.5.4) \quad \frac{\binom{n}{x} (Cp)^x (1-Cp)^{n-x}}{\binom{n}{x} p^x (1-p)^{n-x}} = \left(\frac{C-Cp}{1-Cp}\right)^x \left(\frac{1-Cp}{1-p}\right)^n, \quad (C > 1)$$

which is increasing in x as required.

The same applies in the negative binomial case (2.6.32) with $r_1 = \ldots = r_k$, where the condition reads

$$(4.5.5)\qquad \frac{\binom{x+r-1}{r-1}\ (1-Cq)^r (Cq)^x}{\binom{x+r-1}{r-1}\ (1-q)^r q^x} = \left(\frac{1-Cq}{1-q}\right)^r C^x, \quad (C > 1),$$

which is also increasing in x.

4.6. THE METHOD OF M RANKINGS

We will first prove the consistency of our test procedure as described in section 2.7. We denote the ranks of the objects 1, ..., k in the i-th ranking (i = 1, ..., m) by

$$(4.6.1)\qquad \underline{r}_{i1}, \ldots, \underline{r}_{ik} .$$

The null hypothesis that the m rankings are independent and chosen with equal probabilities from all permutations of the numbers 1, ..., k is tested against the hypotheses

$$(4.6.2)\qquad H_{1j} \begin{cases} P\left[\underline{r}_{ij} > \underline{r}_{il}\right] \overset{def}{=} p > \frac{1}{2} \quad (l \neq j); \\ \text{All rankings of the objects } 1, \ldots, j-1, j+1, \ldots \\ \ldots, k \text{ are equally probable} \end{cases}$$

and

$$(4.6.3)\qquad H_{2j} \begin{cases} P\left[\underline{r}_{ij} > \underline{r}_{il}\right] \overset{def}{=} p < \frac{1}{2} \quad (l \neq j); \\ \text{All rankings of the objects } 1, \ldots, j-1, j+1, \ldots \\ \ldots, k \text{ are equally probable.} \end{cases}$$

We have now the theorem:

THEOREM 4.6.1

If H_{1j} is true, the probability of making a correct decision with the procedure described in section 2.7 tends to 1 if $m \to \infty$.

PROOF

For simplicity we assume that j = 1.

Further we will consider for the present only one ranking, i.e. $m = 1$ and we denote the ranks by

$$\underline{r}_1, \ldots, \underline{r}_k.$$

Then we have

$$(4.6.4) \quad \begin{cases} P\left[\underline{r}_1 > \underline{r}_j\right] = p > \frac{1}{2} & (j = 2, \ldots, k); \\ P\left[\underline{r}_i > \underline{r}_j\right] = \frac{1}{2} & (i \neq j;\ i = 2,\ldots,k;\ j = 2,\ldots,k). \end{cases}$$

Next we introduce the variates $\underline{t}_{ij}$ which are defined by:

$$(4.6.5) \quad \begin{cases} \underline{t}_{ij} = 1 & \text{if } \underline{r}_i > \underline{r}_j \\ \underline{t}_{ij} = 0 & \text{if } \underline{r}_i < \underline{r}_j. \end{cases}$$

From (4.6.4) and (4.6.5) it follows that

$$(4.6.6) \qquad \mathcal{E}\,\underline{t}_{ij} = \mathcal{E}\,\underline{t}_{1j}^2 = p \quad (j = 2, \ldots, k).$$

Further

$$(4.6.7) \qquad \mathcal{E}\,\underline{t}_{1i}\underline{t}_{1j} = P\left[\underline{r}_1 > \underline{r}_i \text{ and } \underline{r}_1 > \underline{r}_j\right] \stackrel{\text{def}}{=} p_2 < p \quad (i \neq j).$$

Now we can calculate the mean and the variance of $\underline{r}_1$.
For

$$(4.6.8) \qquad \mathcal{E}\,\underline{r}_1 = \mathcal{E}(1 + \sum_{i=2}^{k} \underline{t}_{1i}) = 1 + p(k-1) = \frac{1}{2}(k+1) + (p - \frac{1}{2})(k-1)$$

and

$$(4.6.9) \qquad \mathcal{E}(\sum_{2}^{k} \underline{t}_{1i})^2 = \sum_{2}^{k} \mathcal{E}\,\underline{t}_{1i}^2 + \sum_{i \neq j} \mathcal{E}\,\underline{t}_{1i}\underline{t}_{1j} =$$

(according to (4.6.6) and (4.6.7))

$$= (k-1)p + (k-1)(k-2)p_2.$$

Therefore

(4.6.10) $$\operatorname{var} \underline{r}_1 = \operatorname{var}(\sum_{i=2}^{k} \underline{t}_{1i}) = \mathcal{E}(\sum t_{1i})^2 - \{\mathcal{E}(\sum \underline{t}_{1i})\}^2 =$$

$$= (k-1)p + (k-1)(k-2)p_2 - (k-1)^2 p^2 =$$

$$= (k-1)\{p - (k-1)p^2 + (k-2)p_2\} \leq \quad (\text{because } p_2 \leq p)$$

$$\leq (k-1)^2 \{p - p^2\} \leq \frac{1}{4}(k-1)^2.$$

In a similar way we calculate mean and variance of $\underline{r}_2, \ldots, \underline{r}_k$, which have all identical distributions.

We have

(4.6.11) $$\begin{cases} \mathcal{E}\underline{t}_{21} = \mathcal{E}\underline{t}_{21}^2 = 1-p, \\ \mathcal{E}\underline{t}_{2j} = \mathcal{E}\underline{t}_{2j}^2 = \frac{1}{2} & (j = 3, \ldots, k), \\ \mathcal{E}\underline{t}_{2i}\underline{t}_{2j} = \frac{1}{3} & (i \neq j;\ i = 3,\ldots,k;\ j = 3,\ldots,k), \end{cases}$$

and

(4.6.12) $$\mathcal{E}\underline{t}_{21}\underline{t}_{2j} = P\left[\underline{r}_2 > \underline{r}_1 \text{ and } \underline{r}_2 > \underline{r}_j\right] \overset{\text{def}}{=} p_3 < 1-p.$$

Therefore

(4.6.13) $$\mathcal{E}\underline{r}_2 = \mathcal{E}(1 + \sum_{i\neq 2} \underline{t}_{2i}) = \mathcal{E}(1 + \underline{t}_{21} + \sum_{i=3}^{k} \underline{t}_{2i}) =$$

$$= 1 + (1-p) + \frac{1}{2}(k-2) = \frac{1}{2}(k+1) - (p - \frac{1}{2}).$$

Further

(4.6.14) $$\mathcal{E}(\sum_{i\neq 2} \underline{t}_{2i})^2 = \mathcal{E}\{\underline{t}_{21}^2 + 2\underline{t}_{21}\sum_{i=3}^{k} \underline{t}_{2i}^2 + \sum_{i\neq j=3}^{k} \underline{t}_{2i}\underline{t}_{2j}\} =$$

$$= (1-p) + 2(k-2)p_3 + \frac{1}{2}(k-2) + \frac{1}{3}(k-2)(k-3).$$

Therefore

$$\text{(4.6.15)} \qquad \text{var}\ \underline{r}_2 = \text{var}(\sum_{i\neq 2} \underline{t}_{2i}) = \mathcal{E}(\sum \underline{t}_{2i})^2 - \{\mathcal{E}(\sum \underline{t}_{2i})\}^2 =$$

$$= (1-p) + 2(k-2)p_3 + \frac{1}{2}(k-2) + \frac{1}{3}(k-2)(k-3)$$

$$- \{(1-p) + \frac{1}{2}(k-2)\}^2 =$$

$$= \frac{1}{12}k^2 - \frac{7}{6}k + 2 + p(k-1-p) + 2(k-2)p_3 \leq \text{(because}$$

$$p_3 \leq 1-p)$$

$$\leq \frac{1}{12}k^2 + \frac{5}{6}k - 2 - p^2 - p(k-3).$$

Because $\frac{1}{2} < p \leq 1$, we get an upper bound when we take $p = \frac{1}{2}$. So we have finally

$$\text{(4.6.16)} \qquad \text{var}\ \underline{r}_2 \leq \frac{1}{12}k^2 + \frac{1}{3}k - \frac{3}{4}.$$

Returning now to the case of m rankings we can conclude from (4.6.8), (4.6.10), (4.6.13) and (4.6.16)

$$\text{(4.6.17)} \qquad \begin{cases} \mathcal{E}\ \underline{s}_1 = \frac{1}{2}m(k+1) + m(p - \frac{1}{2})(k-1), \\ \text{var}\ \underline{s}_1 \leq \frac{1}{4}m(k-1)^2, \\ \mathcal{E}\ \underline{s}_i = \frac{1}{2}m(k+1) - m(p - \frac{1}{2}), \qquad (i = 2, \ldots, k) \\ \text{var}\ \underline{s}_i \leq m(\frac{1}{12}k^2 + \frac{1}{3}k - \frac{3}{4}). \end{cases}$$

Now for $m \to \infty$ and fixed k the distributions of $\underline{s}_1, \underline{s}_2, \ldots, \underline{s}_k$ tend to normal distributions, both under H_0 and under $H_{1,1}$ because of the central limit theorem.

The difference between the means of $\underline{s}_1$ and $\underline{s}_i$ $(i = 2, \ldots, k)$ is equal to

$$mk(p - \frac{1}{2})$$

and therefore proportional to m.

The standard deviations of $\underline{s}_1$ and $\underline{s}_i$ are proportional to $\sqrt{m}$. Therefore

(4.6.18) $$\lim_{m\to\infty} P\left[\underline{s}_1 > \underline{s}_i\right] = 1 \quad (i = 2,\ldots,k).$$

Further

$$\mathcal{E}(\underline{s}_1 \mid H_{1,1}) - \mathcal{E}(\underline{s}_1 \mid H_0) = m(p - \tfrac{1}{2})(k-1),$$

therefore the probability that H_0 will be rejected at a level α $(0 \leq \alpha \leq 1)$ tends also to 1 if $m \to \infty$, which proves our theorem.

Next we determine the asymptotic relative efficiency (ARE) in the sense of E.J.G. PITMAN (1948) of our test as compared with a method based on the normal distribution. The latter method has not been discussed in chapter 2 and therefore we will indicate here in short how we would go to work in that case.

Suppose the independent variates

(4.6.19) $$\underline{x}_{ij} \quad (i = 1, \ldots, m;\ j = 1, \ldots, k)$$

are normally distributed and have all the same variance σ^2 and means $\mu + \nu_i + \tau_j$ $(\sum \nu_i = 0,\ \sum \tau_j = 0)$.

We wish to test the hypothesis

(4.6.20) $$H_0:\ \tau_j = 0 \qquad (j = 1, \ldots, k)$$

against the alternatives

(4.6.21) $$H_{1j}:\ \tau_j - \tau_1 = \theta, \text{ for all } 1 \neq j,\ \theta > 0,$$

for one unknown value of j

and

(4.6.22) $$H_{2j}:\ \tau_j - \tau_1 = -\theta, \text{ for all } 1 \neq j,\ \theta > 0,$$

for one unknown value of j.

The appropriate test is an F-test for the largest sum $\underline{x}_j = \sum_{i=1}^{m} \underline{x}_{ij}$ against the other when testing against H_1 and for the smallest sum $\underline{x}_j$

when testing against H_2. This is an F-test with 1 and m(k-1) degrees of freedom, or equivalently a t-test with m(k-1)-1 degrees of freedom. The two means that are compared are means of m and m(k-1) values respectively, therefore as has been shown by PITMAN the efficacy of the t-test is equal to

$$\frac{m.m(k-1)}{\{m + m(k-1)\}\sigma^2} = \frac{m(k-1)}{k\sigma^2} . \tag{4.6.23}$$

Next we have to find the relationship between the parameters p of (4.6.2) and θ of (4.6.21).

Because $\underline{x}_{ij}$ and $\underline{x}_{il}$ are independent normal variates with $\mathcal{E}\,\underline{x}_{ij} - \mathcal{E}\,\underline{x}_{il} = \theta$ and common variance σ^2, $\underline{x}_{ij} - \underline{x}_{il}$ has mean 0 and variance $2\sigma^2$.

We have therefore

$$p = P\left[\underline{r}_{ij} > \underline{r}_{il}\right] = P\left[\underline{x}_{ij} > \underline{x}_{il}\right] = P\left[\underline{x}_{ij} - \underline{x}_{il} > 0\right] = \tag{4.6.24}$$

$$= \frac{1}{\sqrt{2\pi}} \int_{\frac{-\theta}{\sigma\sqrt{2}}}^{\infty} e^{-\frac{1}{2}x^2}\, dx$$

and

$$\frac{dp}{d\theta} = \frac{1}{2\sigma\sqrt{\pi}}\, e^{-\frac{\theta^2}{4\sigma^2}} . \tag{4.6.25}$$

If $H_{1,1}$ is true the test statistic is, for large m with probability 1, $\underline{s}_1$. From (4.6.17) we find

$$\frac{\partial}{\partial\theta}(\mathcal{E}\underline{s}_1) = m(k-1)\frac{dp}{d\theta} = \frac{m(k-1)}{2\sigma\sqrt{\pi}}\, e^{-\frac{\theta^2}{4\sigma^2}}$$

and therefore

$$\frac{\partial}{\partial\theta}(\mathcal{E}\underline{s}_1 \;\Big|\; \theta = 0) = \frac{m(k-1)}{2\sigma\sqrt{\pi}} . \tag{4.6.26}$$

We further have to verify the regularity conditions that if $\theta \to 0$ as $m \to \infty$

$$\lim_{m\to\infty} \frac{\mathcal{E}(\underline{s}_1 \mid \theta)}{\mathcal{E}(\underline{s}_1 \mid \theta=0)} = 1 \quad \text{and} \quad \lim_{m\to\infty} \frac{\mathrm{var}(\underline{s}_1 \mid \theta)}{\mathrm{var}(\underline{s}_1 \mid \theta=0)} = 1. \tag{4.6.27}$$

The first condition states that (cf. (4.6.17)) the limiting value of

$$1 + \frac{2(k-1)}{k+1}\,(p - \tfrac{1}{2})$$

equals 1.

It does not contain m and for $\theta \to 0$, $p \to \frac{1}{2}$, therefore it is true.

The second condition says (cf. (4.6.15)):

$$\lim_{\theta\to 0} \frac{\frac{1}{12}k^2 - \frac{7}{6}k + 2 + p(k-1-p) + 2(k-2)p_3}{\frac{1}{12}(k^2 - 1)} = 1.$$

For $\theta \to 0$, $p \to \frac{1}{2}$ and $p_3 \to \frac{1}{3}$ and when these values are filled in we find the desired value.

Further it is required that the test statistic is asymptotically normal. As was remarked in the proof of theorem 4.6.1 this follows from the central limit theorem.

Finally we should have

$$\lim_{m\to\infty} \frac{\frac{\partial}{\partial\theta}\,(\mathcal{E}(\underline{s}_1 \mid \theta=0))}{\sigma(s_1 \mid \theta=0)} = c\sqrt{m}, \text{ where c is a constant.}$$

We find

$$\lim_{m\to\infty} \frac{\frac{\partial}{\partial\theta}\,(\mathcal{E}(\underline{s}_1 \mid \theta=0))}{\sigma(\underline{s}_1 \mid \theta=0)} = \frac{\frac{m(k-1)}{2\sigma\sqrt{\pi}}}{\sqrt{\frac{1}{12}\,m(k^2-1)}} = \sqrt{\frac{3(k-1)}{\pi\sigma^2(k+1)}}\;\sqrt{m}$$

Therefore also this condition is satisfied.

The efficacy of the test is

$$\frac{\left(\frac{\partial}{\partial\theta}\,\mathcal{E}(\underline{s}_1 \mid \theta=0)\right)^2}{\mathrm{var}(\underline{s}_1 \mid \theta=0)} = \frac{3m(k-1)}{\pi\sigma^2(k+1)} \quad . \tag{4.6.28}$$

The asymptotic relative efficiency is found from (4.6.23) and (4.6.28):

$$\frac{3m(k-1)}{\pi\sigma^2(k+1)} \times \frac{k\sigma^2}{m(k-1)} = \frac{3k}{\pi(k+1)} .$$

It is interesting to note that the same value was found for the general m rankings test by Ph. VAN ELTEREN and G.E. NOETHER (1959).

4.7. THE DISTRIBUTION FREE K-SAMPLE TEST

In this case we can prove the following theorem.

THEOREM 4.7.1

If H_{1j} is true (cf. 2.8.2) the probability of a correct decision, with the procedure described in section 2.8 tends to 1 if $n_1 \to \infty, \ldots, n_k \to \infty$, such that

$$\liminf \frac{n_i}{\sum n_j} > 0 \quad (i = 1, \ldots, k).$$

PROOF.

Let $H_{1,j}$ be true and define

$$N \overset{\text{def}}{=} \sum_{i=1}^{k} n_i .$$

If $n_i \to \infty$ $(i = 1,\ldots,k)$ such that

$$\liminf \frac{n_i}{N} > 0 \qquad (i = 1,\ldots,k),$$

we know that WILCOXON's test comparing sample j with the other samples pooled is consistent. This means

(4.7.1) $$\lim_{n_i \to \infty} P[q_j \leq \eta \mid H_{1,j}] = 1 \qquad \text{for every } \eta \ (0 < \eta \leq 1),$$

q_j defined by (2.8.4), or the exceedance probability for the j-th sample converges to 0 in probability (cf. D. VAN DANTZIG (1951) or E.J.G. PITMAN (1948)).

In a similar way as in D. VAN DANTZIG (1951) we find, if

$$p \overset{\text{def}}{=} P[\underline{u}_j > \underline{u}_i \mid H_{1,j}] > \frac{1}{2},$$

that

$$(4.7.2) \qquad \mu_{i,0} \overset{\text{def}}{=} \mathcal{E}(\underline{T}_i \mid H_0) = \tfrac{1}{2}n_i(N-n_i) + \tfrac{1}{2}n_i(n_i+1),$$

$$(4.7.3) \qquad \mu_{i,j} \overset{\text{def}}{=} \mathcal{E}(T_i \mid H_{1,j}) = \tfrac{1}{2}n_i(N-n_i-n_j) + (1-p)n_in_j + \tfrac{1}{2}n_i(n_i+1) < \mu_{i,0} \quad (i \neq j)$$

and

$$(4.7.4) \qquad \sigma^2_{i,j} \overset{\text{def}}{=} \mathrm{var}(\underline{T}_i \mid H_{1,j}) \leq 3\ \mathrm{var}(\underline{T}_i \mid H_0).$$

Now for any η $(0 < \eta \leq 1)$ we have that $\underline{q}_i \leq \eta$ when $\underline{T}_i \geq \mu_{i,0} + c\sigma_{i,0}$, where $\sigma^2_{i,0} \overset{\text{def}}{=} \mathrm{var}(\underline{T}_i \mid H_0)$ and where c depends on n_i, N and η. Because of the asymptotic normality of $\underline{T}_i$,

$$\lim_{n_i \to \infty} c = \xi_\eta,$$

where

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

Therefore c is bounded for $n_i \to \infty$ (under the restrictions of the theorem). Now, because of (4.7.2) and (4.7.3), we have

$$(4.7.5) \qquad P[\underline{q}_i \leq \eta \mid H_{1,j}] = P[\underline{T}_i \geq \mu_{i,0} + c\sigma_{i,0} \mid H_{1,j}] = P[\underline{T}_i \geq \mu_{i,j} + (p-\tfrac{1}{2})n_jn_i + c\sigma_{i,0} \mid H_{1,j}].$$

When n_j and n_i are sufficiently large $(p-\frac{1}{2})n_jn_i + c\sigma_{i,0}$ will be positive and because of Bienaymé-Cebysev's inequality and (4.7.4)

$$(4.7.6) \qquad P[\underline{q}_i \leq \eta \mid H_{1,j}] \leq \frac{\sigma^2_{i,j}}{\{(p-\frac{1}{2})n_jn_i + c\sigma_{i,0}\}^2} \leq 3\left\{\frac{(p-\frac{1}{2})n_jn_i}{\sigma_{i,0}} + c\right\}^{-2}.$$

Because $\sigma_{i,0}^2 = \frac{1}{12} n_i(N - n_i)(N + 1)$, the last member tends to 0, so for $0 < \eta < 1$ we have

$$\lim_{n_i \to \infty} P[\underline{q}_j \leq \eta \mid H_{1,j}] = 1$$

$$\lim_{n_i \to \infty} P[\underline{q}_i \leq \eta \mid H_{1,j}] = 0 \qquad (i \neq j) .$$

This means that for $n_i \to \infty$ H_0 will be rejected for the right reason with a probability tending to 1 and theorem 4.7.1 is proved.

From the consistency as proved above it follows that the asymptotic relative efficiency with respect to the method based on the normal distribution as described in section 2.2 is the same as for WILCOXON's two sample test with respect to the STUDENT test. For the latter case PITMAN (1948) found the value $\frac{3}{\pi}$.

TABLE I

Critical values for the slippage test to the right in the Poisson-case with H_0: $\mu_1 = \mu_2 = \ldots = \mu_k$. Test statistic: max z_i. Approximate significance level 0.05 (upper values) and 0.01 (lower values). The approximated true level of significance, under the condition $\sum z_i = N$, is written behind the critical value. Number of observations k, sum of the observations N.

N \ k	2	3	4	5	6	7	8	9	10
2									
3				3 .040	3 .028	3 .020	3 .016	3 .012	3 .010 3 .010
4		4 .037	4 .016	4 .008 4 .008	4 .005 4 .005	4 .003 4 .003	4 .002 4 .002	3 .045 4 .001	3 .037 4 .001
5		5 .012	5 .004 5 .004	4 .034 5 .002	4 .020 5 .001	4 .013 5 .000	4 .009 4 .009	4 .006 4 .006	4 .005 4 .005
6	6 .031	6 .004 6 .004	5 .019 6 .001	5 .008 5 .008	5 .004 5 .004	4 .035 5 .002	4 .024 5 .001	4 .017 5 .001	4 .013 5 .001
7	7 .016	6 .021 7 .001	6 .005 6 .005	5 .023 6 .002	5 .012 6 .001	5 .007 5 .007	5 .004 5 .004	4 .037 5 .003	4 .027 5 .002
8	8 .008 8 .008	7 .008 7 .008	6 .017 7 .002	6 .006 6 .006	5 .028 6 .003	5 .016 6 .001	5 .010 5 .010	5 .006 5 .006	5 .004 5 .004
9	8 .039 9 .004	7 .025 8 .003	6 .040 7 .005	6 .015 7 .002	6 .007 6 .007	5 .032 6 .003	5 .020 6 .002	5 .013 6 .001	5 .009 5 .009
10	9 .021 10 .002	8 .010 9 .001	7 .014 8 .002	6 .032 7 .004	6 .015 7 .002	6 .008 6 .008	5 .036 6 .004	5 .024 6 .002	5 .016 6 .001
11	10 .012 11 .001	8 .027 9 .004	7 .030 8 .005	7 .010 7 .010	6 .028 7 .004	6 .015 7 .002	6 .008 6 .008	5 .040 6 .005	5 .028 6 .003
12	10 .039 11 .006	9 .012 10 .002	8 .011 9 .002	7 .020 8 .003	6 .048 7 .008	6 .026 7 .004	6 .015 7 .002	6 .009 6 .009	5 .043 6 .005
13	11 .022 12 .003	9 .027 10 .005	8 .023 9 .004	7 .035 8 .006	7 .015 8 .002	6 .042 7 .007	6 .024 7 .003	6 .015 7 .002	6 .009 6 .009
14	12 .013 13 .002	10 .012 11 .002	8 .041 9 .009	8 .012 9 .002	7 .025 8 .004	7 .012 8 .002	6 .038 7 .006	6 .023 7 .003	6 .015 7 .002
15	12 .035 13 .007	10 .026 11 .005	9 .017 10 .003	8 .021 9 .004	7 .040 8 .008	7 .019 8 .003	7 .010 8 .001	6 .035 7 .005	6 .022 7 .003
16	13 .021 14 .004	10 .048 12 .002	9 .030 10 .007	8 .035 9 .007	8 .013 9 .002	7 .030 8 .005	7 .016 8 .002	7 .009 7 .009	6 .033 7 .005

N \ k	2	3	4	5	6	7	8	9	10
17	13 .049 15 .002	11 .024 12 .006	9 .050 11 .002	9 .013 10 .002	8 .021 9 .004	7 .045 8 .009	7 .024 8 .004	7 .013 8 .002	6 .047 7 .008
18	14 .031 15 .008	11 .044 13 .003	10 .022 11 .005	9 .021 10 .005	8 .032 9 .007	8 .014 9 .003	7 .035 8 .007	7 .020 8 .003	7 .012 8 .002
19	15 .019 16 .004	12 .022 13 .006	10 .036 11 .009	9 .033 10 .008	8 .048 10 .002	8 .021 9 .004	7 .050 9 .002	7 .028 8 .005	7 .017 8 .003
20	15 .041 17 .003	12 .039 14 .003	11 .016 12 .004	9 .050 11 .003	9 .017 10 .004	8 .031 9 .007	8 .015 9 .003	7 .040 8 .008	7 .024 8 .004
21	16 .027 17 .007	13 .021 14 .006	11 .026 12 .007	10 .020 11 .005	9 .026 10 .006	8 .044 10 .002	8 .022 9 .004	8 .011 9 .002	7 .033 8 .006
22	17 .017 18 .004	13 .035 15 .003	11 .040 13 .003	10 .031 11 .008	9 .037 10 .009	9 .015 10 .003	8 .031 9 .007	8 .016 9 .003	7 .044 8 .009
23	17 .035 19 .003	14 .019 15 .005	12 .019 13 .005	10 .045 12 .003	10 .014 11 .003	9 .022 10 .005	8 .042 9 .010	8 .022 9 .004	8 .012 9 .002
24	18 .023 19 .007	14 .031 15 .010	12 .029 13 .008	11 .019 12 .005	10 .020 11 .005	9 .030 10 .007	9 .014 10 .003	8 .030 9 .006	8 .017 9 .003
25	18 .043 20 .004	14 .049 16 .005	12 .043 14 .004	11 .028 12 .008	10 .029 11 .008	9 .041 11 .002	9 .019 10 .004	8 .040 9 .009	8 .023 9 .005

TABLE II

Critical values s_α and S_α of the test statistics min s_i and max s_i for the slippage tests for the method of m rankings. Levels of significance α = 0.05 and 0.01, number of rankings m, number of ranked objects k.

m	α \ k	3	4	5	6	7	8	9	10	11	12
3	0.05			3-15	3-18	3-21	3-24	4-26	4-29	4-32	4-35
	0.01								3-30	3-33	3-36
4	0.05	4-12	4-16	5-19	5-23	6-26	6-30	7-33	7-37	7-41	8-44
	0.01			4-20	4-24	4-28	5-31	5-35	5-39	5-43	6-46
5	0.05	5-15	6-19	7-23	8-27	8-32	9-36	10-40	11-44	11-49	12-53
	0.01		5-20	6-24	6-29	7-33	7-38	8-42	8-47	9-51	9-56
6	0.05	7-17	8-22	9-27	10-32	11-37	12-42	13-47	14-52	16-56	17-61
	0.01	6-18	7-23	8-28	8-34	9-39	10-44	11-49	12-54	12-60	13-65
7	0.05	9-19	10-25	11-31	13-36	14-42	16-47	17-53	18-59	20-64	21-70
	0.01	7-21	9-26	10-32	11-38	12-44	13-50	14-56	15-62	16-68	17-74
8	0.05	10-22	12-28	14-34	16-40	17-47	19-53	21-59	23-65	24-72	26-78
	0.01	9-23	10-30	12-36	13-43	15-49	16-56	18-62	19-69	20-76	22-82
9	0.05	12-42	14-31	16-38	18-45	21-51	23-58	25-65	27-72	29-79	31-86
	0.01	11-25	12-33	14-40	16-47	18-54	20-61	21-69	23-76	25-83	26-91
10	0.05	14-26	16-34	19-41	21-49	24-56	26-64	29-71	31-79	34-86	36-94
	0.01	12-28	14-36	17-43	19-51	21-59	23-67	25-75	27-83	29-91	31-99
11	0.05	15-29	18-37	21-45	24-53	27-61	30-69	33-77	35-86	38-94	41-102
	0.01	14-30	16-39	19-47	21-56	24-64	26-73	29-81	31-90	33-99	36-107
12	0.05	17-31	20-40	24-48	27-57	30-66	33-75	37-83	40-92	43-101	46-110
	0.01	16-32	18-42	21-51	24-60	27-69	30-78	32-88	35-97	38-106	41-115
13	0.05	19-33	23-42	26-52	30-61	34-70	37-80	41-89	44-99	48-108	51-118
	0.01	17-35	20-45	24-54	27-64	30-74	33-84	36-94	39-104	42-114	45-124
14	0.05	21-35	25-45	29-55	33-65	37-75	41-85	45-95	49-105	53-115	56-126
	0.01	19-37	23-47	26-58	30-68	33-79	37-89	40-100	44-110	47-121	50-132
15	0.05	22-38	27-48	31-59	36-69	40-80	45-90	49-101	53-112	57-123	62-133
	0.01	21-39	25-50	29-61	33-72	36-84	40-95	44-106	48-117	52-128	55-140
16	0.05	24-40	29-51	34-62	39-73	44-84	48-96	53-107	58-118	62-130	67-141
	0.01	22-42	27-53	31-65	35-77	40-88	44-100	48-112	52-124	56-136	61-147
17	0.05	26-42	31-54	37-65	42-77	47-89	52-101	57-113	62-125	67-137	72-149
	0.01	24-44	29-56	34-68	38-81	43-93	48-105	52-118	57-130	61-143	66-155
18	0.05	28-44	33-57	39-69	45-81	51-93	56-106	61-119	67-131	72-144	78-156
	0.01	26-46	31-59	36-72	41-85	46-98	51-111	56-124	61-137	66-150	71-163
19	0.05	29-47	36-59	42-72	48-85	54-98	60-111	66-124	72-137	77-151	83-164
	0.01	28-48	33-62	39-75	44-89	49-103	55-116	60-130	66-143	71-157	76-171
20	0.05	31-49	38-62	44-76	51-89	57-103	64-116	70-130	76-144	82-158	89-171
	0.01	29-51	35-65	41-79	47-93	53-107	59-121	64-136	70-150	76-164	81-179

TABLE III

Critical values for the slippage test to the right against two outliers in the Poisson-case with $H_0: \mu_1 = \ldots = \mu_k$.
Significance level 0.05 (right hand values) and 0.01 (left hand values).
Number of observations k, sum of the observations N.
Test statistic: sum of largest two observations.
The values with an asterisk would have been 1 higher when using the approximation of section 3.4.

N \ k	4		5		6		7		8		9		10	
5	--	--	--	--	--	--	--	5	--	5	--	5	--	5
6	--	--	--	6	--	6	--	6	6	6	6	6	6	6
7	--	7	--	7	7	7	7	7	7	6	7	6	7	6
8	--	8	8	8	8	7	8	7	7*	7	7	7	7	6*
9	--	9	9	8	9	8	8	8	8	7	8	7	8	7
10	10	10	10	9	9	8*	9	8	9	8	8	8	8	7
11	11	10	10	10	10	9	9	9	9	8	9	8	8*	8
12	12	11	11	10	10	10	10	9	10	9	9	8	9	8
13	13	12	12	11	11	10	10	10	10	9	10	9	9	8*
14	13	12	12	11	12	11	11	10	10	10	10	9	10	9
15	14	13	13	12	12	11	11	10*	11	10	10*	10	10	9
16	15	14	13	12	13	12	12	11	11	10	11	10	10*	10
17	15	14	14	13	13	12	12	11	12	11	11	10	11	10
18	16	15	15	14	14	13	13	12	12	11	12	11	11	10
19	17	16	15	14	14	13	13	12	13	12	12	11	12	11
20	17	16	16	15	15	14	14	13	13	12	12*	11*	12	11
21	18	17	16	15	15	14	14	13	13*	12	13	12	12	11
22	19	17*	17	16	16	14*	15	13*	14	13	13	12	13	12
23	19	18	18	16	16	15	15	14	14	13	14	12*	13	12
24	20	19	18	17	17	15	16	14	15	13*	14	13	13	12
25	21	19	19	17	17	16	16	15	15	14	14	13	14	13

TABLE IV

Critical values for the slippage test to the left against two outliers in the Poisson-case with $H_0: \mu_1 = \ldots = \mu_k$.
Significance level 0.05 (right hand values) and 0.01 (left hand values).
Number of observations k, sum of the observations N.
Test statistic: sum of smallest two observations.
The values with an asterisk would have been 1 lower when using the approximation of section 3.4.

N \ k	4	5	6	7	8	9	10
4	- -	- -					
5	- -	- -					
6	- -	- -					
7	- 0	- -					
8	- 0	- -					
9	- 0	- -					
10	0 0	- -					
11	0 1	- 0					
12	0 1	- 0					
13	0 1	- 0					
14	1 2	0 0	- 0*				
15	1 2	0 1*	- 0				
16	1 2	0 1	- 0				
17	2 3	0 1	- 0	- -	- -		
18	2 3	0 1	0* 0	- 0	- -		
19	2 3	1 1	0 0	- 0	- -		
20	3 4	1 2	0 1	- 0	- -		
21	3 4	1 2	0 1	- 0	- -		
22	3 5*	1 2	0 1	- 0	- 0		
23	4 5	1 3*	0 1	0 0	- 0		
24	4 5	2 3	0 1	0 0	- 0		
25	4 6	2 3	1 2*	0 1	- 0	- -	

REFERENCES

ARLEY, N., On the distribution of relative errors from a normal population of errors. A discussion of some problems in the theory of errors, Danske Vid. Selsk. Matt-Fys. Medd., 18, no. 3, (1940).

BIOMETRIKA TABLES FOR STATISTICIANS, Volume I, Cambridge (1956).

BLISS, C.I., W.G. COCHRAN and J.W. TUKEY, A rejection criterion based upon the range, Biometrika 43, 418-422 (1956).

CHANDRA SEKAR, C. and M.G. FRANCIS, A method to get the significance limit of a type of test criteria, Sankhya, 5, 165-168 (1941).

CHAUVENET, W.A., A manual of spherical and practical astronomy, 5th ed., London 1876, 558-566.

COCHRAN, W.G., The distribution of quadratic forms in a normal system, with applications to the analysis of covariance, Proceedings of the Cambridge Philosophical Society, 30, 178-191 (1934).

COCHRAN, W.G., The distribution of the largest of a set of estimated variances as a fraction of their total, Annals of Eugenics, 11, 47-52 (1941).

VAN DANTZIG, D., On the consistency and the power of Wilcoxon's two sample test, Indagationes Mathematicae, 13, 1-8 (1951).

DARLING, D.A., On a test for homogenity and extreme values, Annals of Mathematical Statistics, 23, 450-456 (1952).

DAVID, H.A., On the application to statistics of an elementary theorem in probability, Biometrika, 43, 85-91 (1956a).

DAVID, H.A., Revised upper percentage points of the extreme studentized deviate from the sample mean, Biometrika, 43, 449-451, (1956b).

DIXON, W.J., Analysis of extreme values, Annals of Mathematical Statistics, 21, 488-506, (1950).

DOORNBOS, R., Significance of the smallest of a set of estimated normal variances, Statistica Neerlandica, 10, 17-26 (1956).

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

DOORNBOS, R. and H.J. PRINS, Slippage tests for a set of gamma variates, Indagationes Mathematicae, 18, 329-337 (1956).

DOORNBOS, R. and H.J. PRINS, On slippage tests, I, II and III, Indagationes Mathematicae, 20, 38-55 and 438-447 (1958).

EDGEWORTH, F.Y., On discordant observations, Philosophical Magazine Series 5, 23, 364-375 (1887).

EISENHART, C., M.W. HASTAY and W.A. WALLIS, Selected Techniques of Statistical Analysis, New York and London (1947).

VAN ELTEREN, Ph. and G.E. NOETHER, The asymptotic efficiency of the χ_r^2-test for a balanced incomplete block design, Biometrika, 46, 475-477 (1959).

FEDERIGHI, E.T., Extended tables of the percentage points of Student's t-distribution, Journal of the American Statistical Association, 54, 683-688 (1959).

FELLER, W., An Introduction to Probability Theory and its Applications, Volume I, New York and London (1950).

GRUBBS, F.E., Sample criteria for testing outlying observations, Annals of Mathematical Statistics, 21, 27-58 (1950).

HALPERIN, M., S. GREENHOUSE, J. CORNFIELD and JULIA ZALOKAR, Tables of percentage points for the studentized maximum absolute deviate in normal samples, Journal of the American Statistical Association, 50, 185-195 (1955).

HARTLEY, H.O., Studentization, Biometrika, 33, 173-180 (1943).

HARTLEY, H.O. and E.R. FITCH, A chart for the incomplete Beta-function and the cumulative binomial distribution, Biometrika, 38, 423-426 (1951).

IRWIN, J.O., On a criterion for the rejection of outlying observations, Biometrika, 17, 238-250 (1925).

JEFFREYS, H., An alternative to the rejection of observations, Proceedings of the Royal Society, 137, 78-87 (1932).

KUDO, A., On the testing of outlying observations, Sankhya, 17, 67-76 (1956).

LIEBERMAN, G.J. and D.B. OWEN, Tables of the hypergeometric probability distribution, Stanford (1961).

MACKAY, A.T., The distribution of the difference between the extreme observation and the sample mean in samples of n from a normal universe, Biometrika, 27, 466-471 (1935).

NAIR, K.R., The distribution of the extreme deviate from the sample mean and its studentized form, Biometrika, 35, 118-144 (1948).

PAULSON, E., An optimum solution to the k-sample slippage problem for the normal distribution, Annals of Mathematical Statistics, 23, 610-616 (1952).

PEARSON, E.S. and C. CHANDRA SEKAR, The efficiency of statistical tests and a criterion for the rejection of outlying observations, Biometrika, 28, 308-320 (1936).

PEARSON, E.S. and H.O. HARTLEY, Tables of the probability integral of the studentized range, Biometrika, 33, 89-99 (1943).

PEIRCE, B., Criterion for the rejection of doubtful observations, Astronomical Journal, 2, 161-163 (1852).

PITMAN, E.J.G., Lecture notes on nonparametric statistical inference (lectures given for the University of North Carolina, Institute of Statistics) (1948).

QUESENBERRY, C.P. and H.A. DAVID, Some tests for outliers, Biometrika, 48, 379-390 (1961).

STONE, E.J., On the rejection of discordant observations, Monthly Notices of the Royal Astronomical Society, 28, 165-168 (1867).

STUDENT, Errors of routine analysis, Biometrika, 19, 151-165 (1927).

THOMPSON, W.A. and T.A. WILKE, On an extreme rank sum test for outliers, Biometrika, 50, 375-383 (1963).

THOMPSON, W.R., On a criterion for the rejection of observations and the distribution of the ratio of deviation to sample standard deviation, Annals of Mathematical Statistics, 6, 214-219 (1935).

TIPPET, L.H.C., On the extreme individuals and the range of samples taken from a normal population, Biometrika, 17, 151-164 (1925).

TRUAX, D.R., An optimum slippage test for the variances of normal distributions, Annals of Mathematical Statistics, 24, 669-673 (1953).

WALD, A., Statistical Decision Functions, New York and London (1950).

YOUDEN, W.J., Ranking laboratories by round robin tests, Materials Research and Standards, 3, 9-13 (1963).

INDEX

OTHER TITLES IN THE SERIES MATHEMATICAL CENTRE TRACTS

A leaflet containing an order-form and abstracts of all publications mentioned below is available at the Mathematisch Centrum, Tweede Boerhaavestraat 49, Amsterdam-1005, The Netherlands. Orders should be sent to the same address.

MCT 1 T. VAN DER WALT, *Fixed and almost fixed points*, 1963. ISBN 90 6196 002 9.

MCT 2 A.R. BLOEMENA, *Sampling from a graph*, 1964. ISBN 90 6196 003 7.

MCT 3 G. DE LEVE, *Generalized Markovian decision processes, part I: Model and method*, 1964. ISBN 90 6196 004 5.

MCT 4 G. DE LEVE, *Generalized Markovian decision processes, part II: Probabilistic background*, 1964. ISBN 90 6196 005 3.

MCT 5 G. DE LEVE, H.C. TIJMS & P.J. WEEDA, *Generalized Markovian decision processes, Applications*, 1970. ISBN 90 6196 051 7.

MCT 6 M.A. MAURICE, *Compact ordered spaces*, 1964. ISBN 90 6196 006 1.

MCT 7 W.R. VAN ZWET, *Convex transformations of random variables*, 1964. ISBN 90 6196 007 X.

MCT 8 J.A. ZONNEVELD, *Automatic numerical integration*, 1964. ISBN 90 6196 008 8.

MCT 9 P.C. BAAYEN, *Universal morphisms*, 1964. ISBN 90 6196 009 6.

MCT 10 E.M. DE JAGER, *Applications of distributions in mathematical physics*, 1964. ISBN 90 6196 010 X.

MCT 11 A.B. PAALMAN-DE MIRANDA, *Topological semigroups*, 1964. ISBN 90 6196 011 8.

MCT 12 J.A.TH.M. VAN BERCKEL, H. BRANDT CORSTIUS, R.J. MOKKEN & A. VAN WIJNGAARDEN, *Formal properties of newspaper Dutch*, 1965. ISBN 90 6196 013 4.

MCT 13 H.A. LAUWERIER, *Asymptotic expansions*, 1966, out of print; replaced by MCT 54 and 67.

MCT 14 H.A. LAUWERIER, *Calculus of variations in mathematical physics*, 1966. ISBN 90 6196 020 7.

MCT 15 R. DOORNBOS, *Slippage tests*, 1966. ISBN 90 6196 021 5.

MCT 16 J.W. DE BAKKER, *Formal definition of programming languages with an application to the definition of ALGOL 60*, 1967. ISBN 90 6196 022 3.

MCT 17 R.P. VAN DE RIET, *Formula manipulation in ALGOL 60, part 1*, 1968. ISBN 90 6196 025 8.

MCT 18 R.P. VAN DE RIET, *Formula manipulation in ALGOL 60, part 2*, 1968. ISBN 90 6196 038 X.

MCT 19 J. VAN DER SLOT, *Some properties related to compactness*, 1968. ISBN 90 6196 026 6.

MCT 20 P.J. VAN DER HOUWEN, *Finite difference methods for solving partial differential equations*, 1968. ISBN 90 6196 027 4.

MCT 21 E. WATTEL, *The compactness operator in set theory and topology*, 1968. ISBN 90 6196 028 2.

MCT 22 T.J. DEKKER, *ALGOL 60 procedures in numerical algebra, part 1*, 1968. ISBN 90 6196 029 0.

MCT 23 T.J. DEKKER & W. HOFFMANN, *ALGOL 60 procedures in numerical algebra, part 2*, 1968. ISBN 90 6196 030 4.

MCT 24 J.W. DE BAKKER, *Recursive procedures*, 1971. ISBN 90 6196 060 6.

MCT 25 E.R. PAERL, *Representations of the Lorentz group and projective geometry*, 1969. ISBN 90 6196 039 8.

MCT 26 EUROPEAN MEETING 1968, *Selected statistical papers, part I*, 1968. ISBN 90 6196 031 2.

MCT 27 EUROPEAN MEETING 1968, *Selected statistical papers, part II*, 1969. ISBN 90 6196 040 1.

MCT 28 J. OOSTERHOFF, *Combination of one-sided statistical tests*, 1969. ISBN 90 6196 041 X.

MCT 29 J. VERHOEFF, *Error detecting decimal codes*, 1969. ISBN 90 6196 042 8.

MCT 30 H. BRANDT CORSTIUS, *Excercises in computational linguistics*, 1970. ISBN 90 6196 052 5.

MCT 31 W. MOLENAAR, *Approximations to the Poisson, binomial and hypergeometric distribution functions*, 1970. ISBN 90 6196 053 3.

MCT 32 L. DE HAAN, *On regular variation and its application to the weak convergence of sample extremes*, 1970. ISBN 90 6196 054 1.

MCT 33 F.W. STEUTEL, *Preservation of infinite divisibility under mixing and related topics*, 1970. ISBN 90 6196 061 4.

MCT 34 I. JUHÁSZ, A. VERBEEK & N.S. KROONENBERG, *Cardinal functions in topology*, 1971. ISBN 90 6196 062 2.

MCT 35 M.H. VAN EMDEN, *An analysis of complexity*, 1971. ISBN 90 6196 063 0.

MCT 36 J. GRASMAN, *On the birth of boundary layers*, 1971. ISBN 90 6196 064 9.

MCT 37 J.W. DE BAKKER, G.A. BLAAUW, A.J.W. DUIJVESTIJN, E.W. DIJKSTRA, P.J. VAN DER HOUWEN, G.A.M. KAMSTEEG-KEMPER, F.E.J. KRUSEMAN ARETZ, W.L. VAN DER POEL, J.P. SCHAAP-KRUSEMAN, M.V. WILKES & G. ZOUTENDIJK, *MC-25 Informatica Symposium*, 1971. ISBN 90 6196 065 7.

MCT 38 W.A. VERLOREN VAN THEMAAT, *Automatic analysis of Dutch compound words*, 1971. ISBN 90 6196 073 8.

MCT 39 H. BAVINCK, *Jacobi series and approximation*, 1972. ISBN 90 6196 074 6.

MCT 40 H.C. TIJMS, *Analysis of (s,S) inventory models*, 1972. ISBN 90 6196 075 4.

MCT 41 A. VERBEEK, *Superextensions of topological spaces*, 1972. ISBN 90 6196 076 2.

MCT 42 W. VERVAAT, *Success epochs in Bernoulli trials (with applications in number theory)*, 1972. ISBN 90 6196 077 0.

MCT 43 F.H. RUYMGAART, *Asymptotic theory of rank tests for independence*, 1973. ISBN 90 6196 081 9.

MCT 44 H. BART, *Meromorphic operator valued functions*, 1973. ISBN 90 6196 082 7.

MCT 45 A.A. BALKEMA, *Monotone transformations and limit laws*, 1973. ISBN 90 6196 083 5.

MCT 46 R.P. VAN DE RIET, *ABC ALGOL, A portable language for formula manipulation systems, part 1: The language*, 1973. ISBN 90 6196 084 3.

MCT 47 R.P. VAN DE RIET, *ABC ALGOL, A portable language for formula manipulation systems, part 2: The compiler*, 1973. ISBN 90 6196 085 1.

MCT 48 F.E.J. KRUSEMAN ARETZ, P.J.W. TEN HAGEN & H.L. OUDSHOORN, *An ALGOL 60 compiler in ALGOL 60, Text of the MC-compiler for the EL-X8*, 1973. ISBN 90 6196 086 X.

MCT 49 H. KOK, *Connected orderable spaces*, 1974. ISBN 90 6196 088 6.

MCT 50 A. VAN WIJNGAARDEN, B.J. MAILLOUX, J.E.L. PECK, C.H.A. KOSTER, M. SINTZOFF, C.H. LINDSEY, L.G.L.T. MEERTENS & R.G. FISKER (Eds), *Revised report on the algorithmic language ALGOL 68*, 1976. ISBN 90 6196 089 4.

MCT 51 A. HORDIJK, *Dynamic programming and Markov potential theory*, 1974. ISBN 90 6196 095 9.

MCT 52 P.C. BAAYEN (ed.), *Topological structures*, 1974. ISBN 90 6196 096 7.

MCT 53 M.J. FABER, *Metrizability in generalized ordered spaces*, 1974. ISBN 90 6196 097 5.

MCT 54 H.A. LAUWERIER, *Asymptotic analysis, part 1*, 1974. ISBN 90 6196 098 3.

MCT 55 M. HALL JR. & J.H. VAN LINT (Eds), *Combinatorics, part 1: Theory of designs, finite geometry and coding theory*, 1974. ISBN 90 6196 099 1.

MCT 56 M. HALL JR. & J.H. VAN LINT (Eds), *Combinatorics, part 2: graph theory, foundations, partitions and combinatorial geometry*, 1974. ISBN 90 6196 100 9.

MCT 57 M. HALL JR. & J.H. VAN LINT (Eds), *Combinatorics, part 3: Combinatorial group theory*, 1974. ISBN 90 6196 101 7.

MCT 58 W. ALBERS, *Asymptotic expansions and the deficiency concept in statistics*, 1975. ISBN 90 6196 102 5.

MCT 59 J.L. MIJNHEER, *Sample path properties of stable processes*, 1975. ISBN 90 6196 107 6.

MCT 60 F. GÖBEL, *Queueing models involving buffers*, 1975. ISBN 90 6196 108 4.

* MCT 61 P. VAN EMDE BOAS, *Abstract resource-bound classes, part 1*. ISBN 90 6196 109 2.

* MCT 62 P. VAN EMDE BOAS, *Abstract resource-bound classes, part 2*. ISBN 90 6196 110 6.

MCT 63 J.W. DE BAKKER (ed.), *Foundations of computer science*, 1975. ISBN 90 6196 111 4.

MCT 64 W.J. DE SCHIPPER, *Symmetrics closed categories*, 1975. ISBN 90 6196 112 2.

MCT 65 J. DE VRIES, *Topological transformation groups 1 A categorical approach*, 1975. ISBN 90 6196 113 0.

MCT 66 H.G.J. PIJLS, *Locally convex algebras in spectral theory and eigenfunction expansions*. ISBN 90 6196 114 9.

* MCT 67 H.A. LAUWERIER, *Asymptotic analysis, part 2*. ISBN 90 6196 119 X.

MCT 68 P.P.N. DE GROEN, *Singularly perturbed differential operators of second order*. ISBN 90 6196 120 3.

MCT 69 J.K. LENSTRA, *Sequencing by enumerative methods*. ISBN 90 6196 125 4.

MCT 70 W.P. DE ROEVER JR., *Recursive program schemes: semantics and proof theory*, 1976. ISBN 90 6196 127 0.

MCT 71 J.A.E.E. VAN NUNEN, *Contracting Markov decision processes*, 1976. ISBN 90 6196 129 7.

* MCT 72 J.K.M. JANSEN, *Simple periodic and nonperiodic Lamé functions and their applications in the theory of conical waveguides*. ISBN 90 6196 130 0.

* MCT 73 D.M.R. LEIVANT, *Absoluteness of intuitionistic logic*. ISBN 90 6196 122 x.

MCT 74 H.J.J. TE RIELE, *A theoretical and computational study of generalized aliquot sequences*. ISBN 90 6196 131 9.

* MCT 75 A.E. BROUWER, *Treelike spaces and related connected topological spaces*. ISBN 90 6196 132 7.

MCT 76 M. REM, *Associons and the closure statement*. ISBN 90 6196 135 1.

* MCT 77

MCT 78 E. DE JONGE, A.C.M. VAN ROOIJ, *Introduction to Riesz spaces*, 1977. ISBN 90 6196 133 5.

MCT 79 M.C.A. VAN ZUIJLEN, *Empirical distributions and rankstatistics*, 1977. ISBN 90 6196 145 9.

* MCT 80 P.W. HEMKER, *A numerical study of stiff two-point boundary problems*, 1977. ISBN 90 6196 146 7.

MCT 81 K.R. APT & J.W. DE BAKKER (eds), *Foundations of computer science II, part I*, 1976. ISBN 90 6196 140 8.

MCT 82 K.R. APT & J.W. DE BAKKER (eds), *Foundations of computer science II, part II*, 1976. ISBN 90 6196 141 6.

* MCT 83 L.S. VAN BENTEM JUTTING, *Checking Landau's "Grundlagen" in the automath system*, 1977. ISBN 90 6196 147 5.

MCT 84 H.L.L. BUSARD, *The translation of the elements of Euclid from the Arabic into Latin by Hermann of Carinthia books vii-xii*, 1977. ISBN 90 6196 148 3.

* MCT 85 J. VAN MILL, *Supercompactness and Wallman spaces*, 1977. ISBN 90 6196 151 3.

An Asterisk before the number means "to appear".