MATHEMATICS

ON SLIPPAGE TESTS 1)

I. A GENERAL TYPE OF SLIPPAGE TEST AND A SLIPPAGE TEST FOR NORMAL VARIATES

BY

R. DOORNBOS AND H. J. PRINS

(Communicated by Prof. D. VAN DANTZIG at the meeting of September 28, 1957)

1. Summary

In this paper slippage tests for variates following various specified distributions, viz the normal, the Poisson, the binomial and the negative binomial, as well as a slippage test for the method of m rankings and a distributionfree k-sample slippage test, are discussed. These tests are all of the general type discussed in section 2. The choice of a test criterion for this type is a plausible one, but in some cases the tests can be proved to be optimal in a sense as described by a theorem of WALD.

For discrete variates the tests are derived as special cases of a slippage test for a general class of distribution functions. The class of distribution functions consists of all distribution functions, for which a close approximation to the true significance levels using a specified method is possible.

In the case of a test for Poisson variates it is possible to give the powerfunctions of the test in very good approximation, using the same method.

The same techniques were used previously for obtaining slippage tests for gamma variates by W. G. COCHRAN (1941), R. DOORNBOS (1956), and R. DOORNBOS and H. J. PRINS (1956) and for normal variates by E. PAULSON (1952). The slippage test for normal variates given here is a generalization of the one given by PAULSON. H. A. DAVID (1956) applied the same principle, without proof however, in two other cases.

2. Introduction

The general type of slippage test considered in this paper serves to decide whether one variate (or a group of variates if the variates occur in groups) slipped or no slippage occurred. These tests arise from the demands of a practical problem which is of a more general type, than the tests describe. For instance in industrial quality control in investigating a process one does not want to decide whether one variate slipped but one wants to decide if variates slipped and if so, how many and which ones.

Thus the tests described here have a restricted practical usefulness, as under the hypotheses considered at most one variate slipped. Still

1) Report SP 62 of the Statistical Department of the Mathematical Centre.
until a more general solution is found to the practical problem, these tests may serve their purpose.

Mosteller (1948) and Pearson and Chandra Sekar (1936) already pointed out this difficulty and Tukey (without date) and Rose and Roy (1953) tried to find a solution for the general problem for normal variates.

The tests dealt with in this paper are of the following general type. Suppose

\[ \vec{y}_1, \ldots, \vec{y}_k \]

are \( k \) random vectors. Thus

\[ \vec{y}_i = (y_{i1}, \ldots, y_{in}) \quad (i = 1, \ldots, k) \]

The variates \( y_{ij} \) are distributed independently and have all the same type of distribution function. These distribution functions contain an unknown parameter \( \theta_i \) as well as other unknown parameters. The test serves to decide whether one of the \( \theta_i \) has slipped.

The simultaneous distribution of the \( y_{ij} \) 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 for the other unknown parameters.

We want to test

\[ H_0 : \theta_1 = \ldots = \theta_k \]

with the \( k \) alternatives

\[ H_i : \theta_i \text{ slipped to the right} \quad (i = 1, \ldots, k) \]

or we want to test \( H_0 \) with the \( k \) alternatives

\[ H_i : \theta_i \text{ slipped to the left} \quad (i = 1, \ldots, k). \]

In order to get rid of the unknown parameters in all but the distributionfree cases sufficient estimates are used.

This sometimes implies using new, one-dimensional, variates, which are functions of the original variates and which have a simultaneous distribution function (in the discrete case a conditional distribution) which does not contain the unknown parameters.

We state the test criterion in terms of the new variates. These variates are

\[ x_1, \ldots, x_k. \]

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

\[ 2) \quad \text{Symbols printed in bold type denote random variables.} \]
Suppose the observed values of \( x_1, \ldots, x_k \) are \( x_1, \ldots, x_k \) respectively. When testing against slippage to the right we determine the right hand tail probabilities

\[
d_j \stackrel{\text{def}}{=} P[x_j \geq x_j], \quad (j = 1, \ldots, k) \quad (2.10)
\]

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

\[
d_m = \min_j d_j \leq \alpha/k. \quad (2.3)
\]

Testing against slippage to the right requires computing

\[
e_j = P[x_j \leq x_j], \quad (j = 1, \ldots, k). \quad (2.4)
\]

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

\[
e_m = \min_j e_j \leq \alpha/k. \quad (2.5)
\]

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

\[
\left\{
\begin{array}{ll}
p_i \stackrel{\text{def}}{=} P[x_i \leq g_i], \\
p_{i,j} \stackrel{\text{def}}{=} P[x_i \leq g_i \text{ and } x_j \leq g_j], & (i \neq j) \\
q_i \stackrel{\text{def}}{=} P[x_i > g_i], \\
q_{i,j} \stackrel{\text{def}}{=} P[x_i > g_i \text{ and } x_j > g_j], & (i \neq j)
\end{array}
\right. \quad (2.6)
\]

all computed under \( H_0 \).

Denoting by \( P \) the probability that at least one of the \( x_i \) does not exceed the corresponding value \( g_i \), it follows from Bonferroni’s inequality (cf. W. Feller (1950), chapter 4) that

\[
\sum_i p_i - \sum_{i<j} p_{i,j} \leq P \leq \sum_i p_i. \quad (2.7)
\]

For \( Q \), i.e. the probability that at least one \( x_i \) exceeds \( g_i \), we have

\[
\sum_i q_i - \sum_{i<j} q_{i,j} \leq Q \leq \sum_i q_i. \quad (2.8)
\]

Then in each case separately we proceed to prove the inequality

\[
p_{i,j} \leq p_i p_j, \quad (2.9)
\]

or

\[
q_{i,j} \leq q_i q_j, \quad (2.10)
\]

which is equivalent with (2.9) (cf. R. Doornbos and H. J. Prins (1956)). Of course, (2.9) and (2.10) to be only hold for a class of distribution functions \( F(x_1, \ldots, x_k) \). The problem of finding general conditions to be

\[\]
imposed on $F(x_1, \ldots, x_k)$, sufficient for the validity of (2.9) has only partly been solved in this paper. Besides in some cases (2.9) only holds for some sets $g_1, \ldots, g_k$, for instance for $\nu^{11} g_i \geq 0$.

Assuming that (2.9) and (2.10) are true we get immediately from (2.7) and (2.8) respectively

$$(2.11) \quad \sum_i p_i - \sum_{i<j} p_i p_j \leq P \leq \sum_i p_i$$

and

$$(2.12) \quad \sum_i q_i - \sum_{i<j} q_i q_j \leq Q \leq \sum_i q_i$$

respectively. Denoting $\sum_i p_i$ ($p$ needs not be $\leq 1$) we have

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

where the equality sign only holds if all $p_i$ vanish, or

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

Thus

$$(2.13) \quad p - \frac{1}{2} p^2 \leq P \leq p$$

and similarly

$$(2.14) \quad q - \frac{1}{2} q^2 \leq Q \leq q,$$

where $\sum_i q_i = q$.

Now, when testing $H_0$ against slippage to the left of one of the $k$ variables the critical region is of the form \{${x_1 \leq g_{1i}}$ or $\ldots$ or $x_k \leq g_{ki}$\}.

The values $g_{1i}$ are determined so as to make all $p_i$ equal to $\alpha/k$ where $\alpha$ is the prescribed level of significance. In the discontinuous case this will in general not be possible; there $g_{1i}$ is the largest value which can be attained by $x_i$ with a positive probability, satisfying

$$(2.15) \quad \alpha_i \overset{\text{def}}{=} P[x_i \leq g_{1i}] \leq \frac{\alpha}{k}.$$

So from (2.13) it follows that the probability $P_\alpha$ of rejecting $H_0$, if $H_0$ is true, satisfies

$$(2.16) \quad \alpha - \frac{\alpha^2}{2} \leq P_\alpha \leq \alpha$$

or

$$(2.17) \quad \alpha' - (\alpha')^2/2 \leq P_\alpha \leq \alpha' \quad (\alpha' = \sum_i \alpha'_i)$$

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

Testing slippage to the right we get similar bounds for the probability of rejecting $H_0$ when $H_0$ is true.
3. The slippage test for normal distributions

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

We want to test the hypothesis

\[
H_0: \mu_1 = \ldots = \mu_k = \mu \quad \text{say},
\]

against the alternatives

\[
H_{11}: \begin{cases} 
\mu_1 = \ldots = \mu_{i-1} = \mu_{i+1} = \ldots = \mu_k = \mu \\
\mu_i = \mu + \Delta \quad (\Delta > 0),
\end{cases}
\]

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

\[
H_{21}: \begin{cases} 
\mu_1 = \ldots = \mu_{i-1} = \mu_{i+1} = \ldots = \mu_k = \mu \\
\mu_i = \mu - \Delta \quad (\Delta > 0),
\end{cases}
\]

for one unknown value of \( i \). From the observations

\[
\begin{align*}
&\{ y_{11}, \ldots, y_{1n_1}, \\
&y_{21}, \ldots, y_{2n_2}, \\
&\vdots \\
&y_{ki}, \ldots, y_{kn_k},
\end{align*}
\]

the variables

\[
b_i = \frac{\sqrt{n_i} (y_i - \bar{y})}{\sqrt{\sum_{j, l} (y_{jl} - \bar{y})^2}}, \quad (i = 1, \ldots, k)
\]

are formed, where

\[
\begin{align*}
&y_i \overset{\text{def}}{=} \frac{1}{n_i} \sum_{l} y_{il}, \\
&\bar{y} \overset{\text{def}}{=} \frac{1}{N} \sum_{j, l} y_{jl},
\end{align*}
\]

and where \( N \) is defined by

\[
N \overset{\text{def}}{=} \sum_{j} n_j.
\]

The \( b_i \) take the place of the variables \( x_i \) in (2.1).

In the following section we shall prove the inequality corresponding to (2.9) if \( g_i \) and \( g_j \) have the same sign and it will be proved that

\[
u_i = \frac{1}{2} \left( 1 + \sqrt{\frac{N}{N - n_i}} b_i \right)
\]

has a \( B \)-distribution with parameters \((N-2)/2\) and \((N-2)/2\) or, that

\[
t_i = \sqrt{N - 2} \frac{\sqrt{\frac{N}{N - n_i}} b_i}{\sqrt{1 - \frac{N}{N - n_i} b_i^2}},
\]

has a Students' \( t \)-distribution with \( N - 2 \) degrees of freedom, for \( i = 1, \ldots, k \).
Thus the procedure described in section 2 can be applied and the $d_i$ and $e_i$ values as defined by (2.2) and (2.4) may be obtained for instance by means of (3.8) and the methods described in section 6 of R. Doornbos and H. J. Prins (1956).

In the present case the determination of the minimal $d$ and $e$ values is much simpler, however, because these minimum values correspond to the largest and the smallest of the $u_i$ respectively and thus of

$$\sqrt{\frac{n_i}{N-n_i}} (y_i - y)$$

and consequently only one incomplete B-integral has to be computed. The critical values $g_{ia}$ for the $b_i$ are determined from

$$g_{ia} = \sqrt{\frac{N-n_i}{N}} (2u_{a/k} - 1),$$

(3.10)

where $u_{a/k}$ is defined by

$$P[u_i \leq u_{a/k}] = \alpha/k.$$  

(3.11)

Because of the symmetry of the distribution of $u_i$ with respect to the point $\frac{1}{2}$, the critical values $G_{ia}$ for the test against slippage to the right are

$$G_{ia} = \sqrt{\frac{N-n_i}{N}} (2u_{1-a/k} - 1) = -g_{ia}.$$  

(3.12)

In the simplest case, i.e. $n_1 = \ldots = n_k = 1$ our test-statistic reduces to the one suggested already by E.S. Pearson and C. Chandra Sekar (1936), but for a constant factor. Using previous work of W. R. Thompson (1935), who derived in this special case the distribution of $t_i$ as defined by (3.9), Pearson and Chandra Sekar were able to derive certain percentage points of $\max b_i$ and $\min b_i$ without deriving the exact distribution. They used the same approximation as is done here, but only up to

$$g_{1a} = \ldots = g_{ka} = g_a \leq -\sqrt{\frac{k-2}{2k}} \left( \text{or } G_a \leq \sqrt{\frac{k-2}{2k}} \right),$$

because, if all $n_i$ are equal, in that region the probability that two of the variables, e.g. $b_i$ and $b_j$, both do not exceed $g_a$ or exceed $G_a$ is equal to zero. Thus the level of significance is then exactly equal to $\alpha$.

The exact distribution for $n_1 = \ldots = n_k = 1$ has been computed numerically by F. E. Grubbs (1950), who gave tables of exact percentage points up to $k=25$ for $\epsilon=0.10$, 0.05, 0.025 and 0.01.

E. Paulson (1952) proposed the same test statistic (but for a constant factor) for slippage to the right and the same approximation as suggested here in the special case $n_1 = \ldots = n_k = n$ but he gives no bounds for the corresponding level of significance. Paulson proved that in this case the use of $\max b_i$ as test statistic has the following optimum property. Let
$D_0$ denote the decision that the $k$ means are equal and let $D_i$ ($j = 1, \ldots, k$) denote the decision that $D_0$ is incorrect and the $\mu_j = \max(\mu_1, \ldots, \mu_k)$. Now the procedure:

$$
(3.13) \begin{cases}
\text{If } b_m > \lambda_\alpha \text{ select } D_m, \\
\text{if } b_m \leq \lambda_\alpha \text{ select } D_0,
\end{cases}
$$

where $m$ is the index of the maximum $b$-value 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 the observations,

(c) the decision procedure 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 $\Delta$ must be the same for $i = 1, 2, \ldots, k$.

The constant $\lambda_\alpha$ in (3.13) is determined by requirement (a). Our critical value $G_\alpha$ is an approximation of $\lambda_\alpha$.

The case of slippage to the left, although not mentioned explicitly by Paulson is completely analogous and the same optimum property holds there.

4. Outline of a proof of the results stated in 3

In this section we merely sketch the proof of the inequality

$$
(4.1) \quad P[b_i \leq g_i \text{ and } b_j \leq g_j] \leq P[b_i \leq g_i] \cdot P[b_j \leq g_j],
$$

provided $g_i g_j \geq 0$,

where $b_i$ and $b_j$ are defined by (3.5) for all pairs $i, j (i \neq j; i, j = 1, \ldots, k)$. Giving all details would require too much space.

First the marginal distributions of $b_i$ and $b_j$ and their simultaneous distribution have to be derived. These are

$$
(4.2) \quad f(b_i) = \sqrt{\frac{N}{N - n_i}} f(\frac{N - 1}{2}) \int \frac{1}{\sqrt{\pi}} \left( 1 - \frac{N}{N - n_i} b_i^2 \right)^{\frac{N - 4}{2}}, \quad (i = 1, \ldots, k)
$$

and

$$
(4.3) \quad g(b_i, b_j) = \sqrt{\frac{N}{N - n_i - n_j}} \int \frac{N - 3}{2\pi} \left[ 1 - \frac{N - n_j}{N - n_i - n_j} b_i^2 - \frac{2\sqrt{n_i n_j}}{N - n_i - n_j} b_i b_j - \frac{N - n_i}{N - n_i - n_j} b_j^2 \right]^{\frac{N - 5}{2}}.
$$

Both formulae are valid in the regions where the expressions between braces are positive, outside these regions the respective density functions are zero. The region where $g(b_i, b_j)$ differs from zero is bounded by an
ellipse in the \((b_i, b_j)\) plane, with principle axes of length \(1\) and \(\sqrt{\frac{N-n_i-n_j}{N}}\) and equations

\[
\begin{align*}
 b_i + \sqrt{\frac{n_j}{n_i}} b_j &= 0, \\
 b_i - \sqrt{\frac{n_j}{n_i}} b_j &= 0.
\end{align*}
\]  
\(4.4\)

When proving (4.1) we may obviously assume that the point \((g_i, g_j)\) lies within this ellipse, because otherwise \(P[b_i \leq g_i \text{ and } b_j \leq g_j] = 0\). Further we suppose that both \(g_i\) and \(g_j\) are \(\leq 0\). This is no restriction, for, when (4.1) holds for a pair of values \(g_i\) and \(g_j\), the inequality \(P[b_i > -g_i \text{ and } b_j > -g_j] \leq P[b_i > -g_i] \cdot P[b_j > -g_j]\) also holds for reasons of symmetry. Consequently (4.1) is also true for \(-g_i\) and \(-g_j\) because of the equivalence of (2.9) and (2.10). We shall see that in the \((g_i, g_j)\) region considered (4.1) holds with the < sign. We have to prove

\[
\phi(g_i, g_j) \overset{\text{def}}{=} P[b_i \leq g_i] \cdot P[b_j \leq g_j] - P[b_i \leq g_i \text{ and } b_j \leq g_j] > 0.
\]  
\(4.5\)

The proof consists of showing consecutively

\[
\phi \left(-\sqrt{\frac{N-n_i-n_j}{N-n_i}}, g_j\right) > 0, 
\]  
\(4.6\)

\[
\phi \left(0, -\sqrt{\frac{N-n_i-n_j}{N-n_i}}\right) > 0
\]  
\(4.7\)

and

\[
\frac{d \phi(0, g_j)}{d g_j} \geq 0 \quad \left(-\sqrt{\frac{N-n_i-n_j}{N-n_i}} \leq g_j \leq 0\right).
\]  
\(4.8\)

From (4.7) and (4.8) it follows that

\[
\phi(0, g_j) > 0 \quad \left(-\sqrt{\frac{N-n_i-n_j}{N-n_i}} \leq g_j \leq 0\right).
\]  
\(4.9\)

Further we can derive

\[
\frac{\partial \phi(g_i, g_j)}{\partial g_i} = \sqrt{\frac{N}{N-n_i}} \frac{1}{\Gamma\left(\frac{N-2}{2}\right)} \left(1 - \frac{N}{N-n_i} g_i^2\right)^{\frac{N-4}{2}} \phi_1(g_i, g_j),
\]  
\(4.10\)

where \(\phi_1(g_i, g_j)\) is a decreasing function of \(g_i\) if \(g_i, g_j \geq 0\), thus \(\frac{\partial \phi(g_i, g_j)}{\partial g_i}\) is everywhere positive, everywhere negative, or positive up to a certain point \(g_{0i}\) (depending upon \(g_i\)), say, and negative thereafter. So in virtue of (4.6) and (4.9) we may conclude

\[
\phi(g_i, g_j) > 0, \text{ if } g_i \leq 0,
\]  
\(4.11\)

\[
g_i \leq 0 \text{ and } 1 - \frac{N-n_j}{N-n_i-n_j} g_i^2 - \frac{2}{N-n_i-n_j} g_i g_j - \frac{N-n_i}{N-n_i-n_j} g_j^2 \geq 0.
\]  

A detailed proof can be found in R. Doornbos, H. Kesten and H. J. Prins (1956).
REFERENCES


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


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


PEARSON, E. S. and C. CHANDRA SEKAR, The efficiency of statistical tools and a criterion for the rejection of outlying observations, Biometrika, 28, 308–320 (1936).


(To be continued).