MATHEMATISCH CENTRUM

2e BOERHAAVESTRAAT 49

AMSTERDAM

STATISTISCHE AFDELING

р

Leiding: Prof. Dr D. van Dantzig Chef van de Statistische Consultatie: Prof. Dr J. Hemelrijk

Report S 187 (VP 4)

A slippage test for a set of Gamma-variates

bу

R. Doornbos and H.J. Prins

January

1956

1. Summary

In this report a generalization is given of the significance tests for the largest and the smallest respectively of a set of estimated normal variances as suggested by W.G. COCHRAN (1941) and one of the present authors (cf. R. Doornbos (1956)) respectively. These tests only deal with the case where the samples from which the variances are estimated all have the same size.

The present report gives a treatment which is also valid for different sample sizes. Further we consider the power function of the tests with respect to the alternative hypothesis that one of the variances has slipped to the right or, in the case of the test for the smallest variances to the left. Slippage tests for a set of Poisson-variates which appear to lead to the same distribution-functions will be discussed in a separate report by the second author.

Finally the construction of a nomogram is suggested to faciliate the application of the tests.

2. Introduction and description of the tests

Suppose we have a set of random variables

$$(2.1) \qquad \qquad \underline{u}_1, \dots, \underline{u}_k \qquad \qquad 1$$

distributed independently of one another according to gamma distributions with parameters $\alpha_1,\beta_1,\dots,\alpha_k,\beta_k$ respectively; that is to say the density function of $\mathcal{U}_{\mathcal{L}}$ is

(2.2)
$$f(u_i) = \frac{1}{\Gamma(\alpha_i)\beta_i^{\alpha_i}} u_i^{\alpha_i-1} e^{-u_i/\beta_i}, \quad 0 \le u \le \infty,$$

where \ll_{i} and β_{i} are real positive numbers. As is well known the distribution of $t=\chi^{2}\sigma^{2}$, where χ^{2} is distributed as a chi-square with ν degrees of freedom, is a special case of a gamma distribution, with parameters $\ll=\nu/2$ and $\beta=2\sigma^{2}$.

Now our problem is to find tests for the hypothesis

(2.3)
$$H_0: \beta_1 = \dots = \beta_k = \beta_k, say,$$

against the alternatives

(2.4)
$$H_{i}: \beta_{i} = \cdots = \beta_{i-i} = \beta_{i+1} = \cdots = \beta_{k} = \beta_{i},$$

$$\beta_{i} = c_{i} \beta_{i}, c_{i} > 1,$$

¹⁾ Random variables are denoted by underlined symbols.

for one unknown value of i and

(2.5)
$$H_{\lambda}: \beta_{1} = \cdots = \beta_{i-1} = \beta_{i+1} = \cdots = \beta_{k} = \beta_{k},$$

$$\beta_{i} = c_{2i}\beta_{1}, \quad 0 < c_{2i} < 1,$$
for one unknown value of i .

For both tests we compute the ratios

Then, if we are testing H_o against H_i , the following incomplete ${\mathcal B}$ integrals are determined:

(2.7)
$$\underline{\alpha}_{j} = \frac{1}{B(\alpha_{j}, A - \alpha_{j})} \int_{x_{j}}^{1} x^{\alpha_{j}-1} (J - x)^{A - \alpha_{j}-1} dx = \frac{1}{B(\alpha_{j}, A - \alpha_{j})} \int_{x_{j}}^{1} x^{\alpha_{j}-1} (J - x)^{A - \alpha_{j}-1} dx = \frac{1}{A(\alpha_{j}, A - \alpha_{j})}, \qquad (J = 1, \dots, K),$$
 where $A = \sum_{i=1}^{k} \alpha_{i}$ Next we define the test statistic $\underline{\alpha}$ by

If we reject H_0 when Δ takes a value $\Delta \leq \frac{\mathcal{E}}{\mathcal{E}}$, the level of significance lies between \mathcal{E} and $\mathcal{E} - \frac{1}{2} \mathcal{E}^2$ as will be shown in section 4.

Testing H_0 against H_2 requires computation of the integrals

Testing
$$H_0$$
 against H_2 requires computation of the integral (2.9)

$$= \frac{1}{B(\alpha_j, A - \alpha_j)} \int_0^{\infty} 2^{\alpha_j - 1} (1 - 2\epsilon)^{A - \alpha_j - 1} dz = 1 - 2\epsilon = 1$$
We reject H_0 if

$$(2.10) \qquad \underline{e} = \min_{\underline{e}_{j}} \leq \frac{\underline{\varepsilon}}{k}.$$

The level of significance is again a number between $oldsymbol{\mathcal{E}}$ and $oldsymbol{\mathcal{E}}\sim rac{1}{\lambda}\, \widehat{\boldsymbol{\xi}}^{\,\lambda}$

3. An optimum property of the tests if $\alpha_1 = \cdots = \alpha_k$

D.R. TRUAX (1953) proved an optimum property of COCHRAN's test. In exactly the same way one can prove that our tests are optimal in the following sence if $lpha_{,=}$. = $lpha_{\not\leftarrow}$. Let \mathcal{D}_{o} be the decision that \mathcal{H}_0 is true and let $\mathcal{D}_{i,j}$ be the decision that \mathcal{H}_0 is false and that $\mathcal{B}_i = \max(\mathcal{B}_1, \ldots, \mathcal{B}_k)$. Then, if $\mathcal{A} = \mathcal{A}_m$, i.e. if \mathcal{A}_m is the smallest of $\mathcal{A}_i, \ldots, \mathcal{A}_k$ the procedure

(3.1) if
$$d \le L_{\varepsilon}$$
 select D_{im} , if $d > L_{\varepsilon}$ select D_{o} ,

where \angle_{ε} is a constant determined by the condition that $\mathcal{D}_{\varepsilon}$ should be selected with probability $/-\varepsilon$ if $\mathcal{H}_{\varepsilon}$ is true, maximizes the probability of making the correct decision if the hypothesis $/\varepsilon$ is true.

When the hypothesis \mathcal{H}_2 is true the analogous optimum property holds for our second test. In both cases ξ is an approximation of the critical values of $\underline{\mathscr{A}}$ and \underline{e} .

4. Proofs of the results stated in 2.

To obtain the joint distribution of $2, \ldots, 2c_{k}$, as given by (2.6) and of $\mathcal{U} = \mathcal{U}_1 + \cdots + \mathcal{U}_k$ we put

$$(4.1) \begin{cases} \alpha_{i} = \kappa_{i} \mathcal{U} \\ \alpha_{k-i} = \kappa_{k}, \mathcal{U} \\ \alpha_{k} = \mathcal{U}(1-\kappa_{i}-\cdots-\kappa_{k}). \end{cases}$$

The Jacobian of this transformation becomes

Under \mathcal{H}_o the simultaneous distribution of $\mathcal{U}_1, \ldots, \mathcal{U}_k$ is (2.2)

$$(4.3) \quad f(u_1, ..., u_k) = \frac{u_1^{\alpha_{r-1}} ... u_k^{\alpha_{k-1}} e^{-(u_1 + ... + u_k)/\beta}}{\Gamma(\alpha_1) ... \Gamma(\alpha_k)/\beta^A},$$

where $A = \alpha_1 + \cdots + \alpha_k$.

Thus the joint distribution of $\underline{x}_1,\ldots,\underline{x}_{k-1}$ and $\underline{\mathcal{U}}$ is given by the density function

$$g(x_{1},...,x_{k-1},\mathcal{U}) = \frac{x_{1}^{\alpha_{1}-1}...x_{k-1}^{\alpha_{k}-1}^{-1}(1-x_{1}...-x_{k-1})^{\alpha_{k}}\mathcal{U}^{A-1}.\mathcal{U}^{B}}{\Gamma(\alpha_{1})...\Gamma(\alpha_{k})\beta^{A}} = \frac{\Gamma(A)}{\Gamma(\alpha_{1})...\Gamma(\alpha_{k})} \frac{x_{1}^{\alpha_{1}-1}...x_{k-1}^{\alpha_{k}-1}(1-x_{1}...-x_{k-1})^{\alpha_{k}}\mathcal{U}^{A-1}.\mathcal{U}^{B}}{\Gamma(A)\beta^{A}}$$

Thus we see, as is well known that \mathcal{U} has also a gamma distribution, with parameters $\alpha_1+\cdots+\alpha_k=A$ and β and moreover that the joint distribution of $\alpha_1,\ldots,\alpha_{k-1}$ is given by

(4.5)
$$h(x_1,...,x_{k-1}) = \frac{\Gamma(A)}{\Gamma(\alpha_1)...\Gamma(\alpha_k)} x_1^{\alpha_{j-1}}...x_{k-1}^{\alpha_{k-1}} (1-x_1...x_{k-1})^{\alpha_{k-1}}$$

But the same derivation gives us the general result

$$(4.6) \quad p(x_1, \dots, x_i) = \frac{\Gamma(A)}{\Gamma(d_i) \cdot \Gamma(d_i) \cdot \Gamma(d_{i+1} + \dots + d_k)} x_i^{d_i - 1} \cdot x_i^{d_i - 1} (-x_1 \dots - x_i)$$

$$(i = 1, \dots, k-1),$$

if we consider instead of $\mathcal{U}_1, \ldots, \mathcal{U}_k$ the \mathcal{C}_{+} variables $\mathcal{U}_1, \ldots, \mathcal{U}_{\ell}$ and $\mathcal{U}_{\ell+1} + \cdots + \mathcal{U}_k$ which are also independent from one another and which have gamma distributions with parameters $\mathcal{A}_1, \ldots, \mathcal{A}_{\ell}, \mathcal{A}_{\ell+1} + \cdots + \mathcal{A}_k$ and \mathcal{B}_n .

We consider now a set of k real numbers g_1, \ldots, g_k ($0 \le g_i \le j$) and the probabilities defined by

$$(4.7) \begin{cases} p_i = P[x_i \leq g_i] \\ p_{i,j} = P[x_i \leq g_i \text{ and } x_j \leq g_j], (i \neq j) \\ q_i = P[x_i > g_i] \\ q_{i,j} = P[x_i > g_i \text{ and } x_j > g_j], (i \neq j) \end{cases}$$

all computed under \mathcal{H}_o . If we denote by \mathcal{P} the probability that at least one of the ratios \mathcal{X}_i does not exceed the corresponding value \mathcal{G}_i , we have

where the z^{th} summation is extended over all β 's with z subscripts; hence the z^{th} sum has $\binom{k}{z}$ terms (G.W. FELLER (1950), Chapter 4).

For Q , the probability that at least one of the $\underline{\mathcal{X}}_{\mathcal{C}}$ exceeds $Q_{\mathcal{C}}$, we have

It follows from BONFERRONI's inequality (cf. W. FELLER (1950)), chapter 4) that

$$(4.10) \qquad \qquad Z p_i - Z p_{i,j} \leq P \leq Z p_i$$

and

$$(4.11) \qquad \qquad Zq_i - Zq_{i,j} \leq Q \leq Zq_i.$$

In the following we shall prove the inequalities

and

9in = 9i9; (4.13)

If we now determine the numbers $g_{i,\xi}$ that all p_i are equal to ξ , then we get from (4.10) and (4.12)

$$\sum p_i - \sum p_i p_j \leq \sum p_i - \sum p_{i,j} \leq P_{\epsilon} \leq \sum p_i$$
,

or

$$\varepsilon - \frac{1}{2} \frac{k-1}{k} \varepsilon^2 \leq \mathcal{P}_{\varepsilon} \leq \varepsilon$$

or for $k \ge 2$

$$(4.14) \qquad \mathcal{E} - \frac{1}{2} \mathcal{E}^2 \leq \mathcal{P}_{\mathcal{E}} \leq \mathcal{E}$$

In the same way we get

$$(4.15) \qquad \qquad \varepsilon - \frac{1}{2} \varepsilon^2 \leq Q_{\varepsilon} \leq \varepsilon$$

if the numbers $g_{i,\epsilon}^{(j)}$ are determined so as to make all g_i equal to $\frac{\epsilon}{I}$. As the procedure described in section 2 to test H_o against the sets of alternatives H_2 and H_3 respectively gives us the probabilities R_2 and R_3 respectively of rejecting H_3 when H_4 is true, these probabilities lie between the bounds stated there.

We now proceed to prove the inequalities (4.12) and (4.13). First it is shown that (4.12) and (4.13) are equivalent. We have

and consequently

$$(4.16) p_{i}(1-p_{j}) = q_{j}(1-q_{i})$$

Further

From (4.16) and (4.17) we obtain

which proves the equivalence of (4.12) and (4.13). Thus it is sufficient to prove $(4.12)^{-1}$ and we need only consider values g_i and g_j such that $g_{ij}+g_{j}\leq 1$, for when $g_{ij}+g_{j}>1$, $g_{i,j}=0$ and so (4.13) and (4.12) are obviously true.

It is easily seen that (4.12) is equivalent with

$$(4.19) \qquad \frac{\beta_{i,j}}{\beta_{j}} \leq \frac{\beta_{i} - \beta_{i,j}}{\gamma_{j}},$$

or

$$(4.20) \frac{P[x_i \leq g_i \text{ and } x_j \leq g_j]}{P[x_j \leq g_j]} \leq \frac{P[x_i \leq g_i \text{ and } x_j > g_j]}{P[x_j > g_j]}.$$

Fron (4.6) it follows that the left hand member $\angle(g_i,g_j)$ of

equals
$$C = \frac{\int_{0}^{\pi} \int_{0}^{g_{i}} x_{j} \, dx_{j}^{2} - 2c_{i} dx_{i}^{2} \, dx_{i}^{2} - (1-x_{i}-x_{j})^{A-\alpha_{i}-\alpha_{j}^{2}} \, dx_{i} \, dx_{j}^{2}}{\int_{0}^{g_{i}} x_{j}^{2} \, dx_{j}^{2} - (1-x_{j})^{A-\alpha_{j}^{2}} \, dx_{j}^{2}}$$

where $C = \frac{\Gamma(A - \alpha_i)}{\Gamma(\alpha_i) \Gamma(A - \alpha_i - \alpha_i)}$ Putting $\mathcal{X}_i = \mathcal{V}(I - 2c_i)$ we get where

$$(4.21) L(g;g_{j}) = C \frac{\int_{0}^{g_{j}} \int_{1-2g_{j}}^{1-2g_{j}} v^{d_{i}-1}(1-v)^{A-\alpha_{i}-\alpha_{j}-1} x_{j}^{\alpha_{j}-1}(1-x_{j})^{A-\alpha_{j}-1} dv dx_{j}}{\int_{0}^{g_{j}} x_{j}^{\alpha_{j}-1}(1-x_{j})^{A-\alpha_{j}-1} dx_{j}} \\ \leq C \int_{1-g_{j}}^{1-g_{j}} v^{\alpha_{i}-1}(1-v)^{A-\alpha_{i}-\alpha_{j}-1} dv.$$

On the other hand the right hand member $\mathcal{R}(g_i,g_i)$ of (4.20)

is equal to
$$\lim_{x \to \infty} (g_i, 1-x_j)$$

$$C = \underbrace{g_i \circ x_i \cdot x_j \cdot x_{i-1} \cdot x_{i-2} \cdot x_{i-1}}_{x_i \cdot x_i \cdot x_{i-1} \cdot x_{i-1} \cdot x_{i-1} \cdot x_{i-1}} \underbrace{dx_i \cdot dx_i}_{dx_i \cdot x_i \cdot x_i \cdot x_{i-1} \cdot x_i \cdot x_{i-1} \cdot x_{i-1}}_{dx_i \cdot x_i \cdot x_{i-1} \cdot x_i \cdot x_{i-1} \cdot x_$$

¹⁾ The following proof, which is substantially simpler than another one which was developed by the authors, has been found by H. KESTEN, assistant of the Statistical Department, oas a speclia case of the proof of the more general inequality

So it follows from (4.21) and (4.22) that (4.20) holds.

5. The power of the tests

In this section we shall derive upper and lawer bounds for the probabilities of making a correct decision, following the procedure described in section 3, under the hypotheses $\frac{1}{2}$ and $\frac{1}{2}$

In the first case, i.e. when \mathcal{A}_{i} is true, we assume that \mathcal{A}_{i} is the parameter which has slipped to the right. i.e. $\mathcal{A}_{i} = \mathcal{A}_{i}/\mathcal{B}_{i}$, $\mathcal{C}_{ii} > 1$. Then we prove that \mathcal{Q}_{i} , the probability of making the correct decision lies between the limits

$$(5.1) \left[1 - I_{c_i^{(0)}}(\alpha_i, A - \alpha_i) \right] (1 - \epsilon) \leq Q_i \leq \left[1 - I_{c_i^{(0)}}(\alpha_i, A - \alpha_i) \right],$$

where

(5.2)
$$C_{i}^{(i)} = \frac{g_{i,\epsilon}}{g_{i,\epsilon}}$$
,

where $g_{i,\epsilon}^{(\prime)}$ is determined so as to make

$$(5.3) \qquad I_{q_{i,\epsilon}^{(i)}} \left(\alpha_{i}, A - \alpha_{i}\right) = 1 - \frac{\varepsilon}{k}$$

When C_{ii} becomes large C_{ii} converges to the upper bound given by the right hand member of (5.1).

When \mathcal{H}_2 is true and \mathcal{J}_2 has slipped to the left, i.e. $\mathcal{J}_2 = \mathcal{C}_{2/2}$, $0 \le \mathcal{C}_{2/2}$, the following limits can be derived for \mathcal{F}_2 , the probability of making the correct decision in this case.

$$(5.4) \left[I_{c_{\alpha}}(\alpha_{j}, A - \alpha_{j}) \right] (1 - \varepsilon) \leq P_{j} \leq I_{c_{\alpha}}(\alpha_{j}, A - \alpha_{j}),$$

where

(5.5)
$$C_{j}^{(2)} = \frac{g_{j, \varepsilon}^{(2)}}{C_{2j} + (1 - C_{2j}) g_{j, \varepsilon}^{(2)}}$$

and $Q_{j,\xi}^{(2)}$ is determined from

$$(5.6) \qquad I_{gis}(\alpha_j, A - \alpha_j) = \frac{\varepsilon}{k}.$$

Again for small values of C_2 , C_2 , C_2 , C_3 , C_4 , C_4 , C_5 , C_6 , C_7 , C_8 , $C_$

In order to prove (5.1) we may assume without loss of generality that $\ell=\ell$ and then we put $\underline{\mathcal{U}}_{\ell}/c_{\ell\ell}=\underline{\mathcal{U}}_{\ell}$, thus $\underline{\mathcal{U}}_{\ell}$ has a gamma distribution with parameters α_{ℓ} and β . The probability Q_{ℓ} , of making the correct decision is

$$Q_{i} = P\left[d_{i} = \min_{j=1,\dots,k} d_{j} \text{ and } d_{i} \leq \frac{\mathbb{E}}{L}\right] \geq 2$$

$$\geq P\left[d_{i} \leq \frac{\mathbb{E}}{L} \text{ and } d_{2} > \frac{\mathbb{E}}{L} \dots \text{ and } d_{k} > \frac{\mathbb{E}}{L}\right] = 2$$

$$= P\left[d_{i} \leq \frac{\mathbb{E}}{L}\right] - P\left[d_{i} \leq \frac{\mathbb{E}}{L} \text{ and } d_{2} \leq \frac{\mathbb{E}}{L}\right]; \text{ or } \dots; \text{ or } \left(d_{i} \leq \frac{\mathbb{E}}{L} \text{ and } d_{k} \leq \frac{\mathbb{E}}{L}\right).$$

Thus the following inequality holds

(5.7)
$$P[d_i \leq \frac{\varepsilon}{k}] - \frac{k}{2k} P[d_i \leq \frac{\varepsilon}{k} \text{ and } d_i \leq \frac{\varepsilon}{k}] \leq Q_i \leq P[d_i \leq \frac{\varepsilon}{k}].$$

We have

$$\begin{aligned}
\mathcal{P}[d_{1} \leq \frac{\varepsilon}{k}] &= \mathcal{P}[2_{1} \geq g_{1,\varepsilon}^{(0)}] = \\
&= \mathcal{P}\left[\frac{c_{1} \, 2_{1}}{u_{1} + u_{2} + \dots + u_{k} + (c_{N-1}) \, u_{1}} \geq g_{1,\varepsilon}^{(0)}] = \\
&= \mathcal{P}\left[\frac{v_{1}}{v_{1} + u_{2} + \dots + u_{k}} \geq \frac{g_{1,\varepsilon}^{(0)}}{c_{N-1} - (c_{N-1}) g_{1,\varepsilon}^{(0)}}\right] =
\end{aligned}$$

$$(cf. (5.2)) = P\left[\frac{2i}{2i+4i} + \frac{2i}{2k} \ge c^{(i)}\right].$$

The distribution of $\frac{v_{i}}{v_{i}+u_{2}+\cdots+u_{k}}$ is the distribution of x_{i} under x_{i} and therefore known. In fact

$$(5.8) \quad P\left[\alpha_{\ell} \leq \frac{\varepsilon}{k}\right] = 1 - I_{c0}\left(\alpha_{k}, A - \alpha_{\ell}\right).$$

Further we have

(5.9)
$$P\left[\alpha_{i} \leq \frac{\varepsilon}{k} \text{ and } \alpha_{j} \leq \frac{\varepsilon}{k}\right] =$$

$$= P\left[\frac{v_{i}}{v_{i}+u_{2}+\cdots+u_{k}} \geq c_{i}^{(i)} \text{ and } \frac{u_{j}}{v_{j}+\cdots+u_{k}} + (c_{i-1})v_{j} \geq g_{j,\varepsilon}^{(i)} \leq e^{-\frac{v_{i}}{k}} + e^{-\frac{v_{$$

Substituting (5.8) and (5.9) into (5.7) we get

$$(5.10) \left[I - I_{c0} \left(\alpha_{1}, A - \alpha_{1} \right) \right] \left(I - \epsilon \right) \leq \left[I - I_{c0} \left(\alpha_{1}, A - \alpha_{1} \right) \right] \left(I - \frac{\epsilon}{2} \right) \leq \left[I - I_{c0} \left(\alpha_{1}, A - \alpha_{1} \right) \right].$$

 $\leq Q_{1} \leq \left[1 - \frac{1}{2}CQ_{1}(\alpha_{1}, A - \alpha_{2})\right].$ which proves (5.1). When C_{1} is large $P[A] \leq \frac{1}{2}$ will for the be much smaller than $\frac{1}{2}$ and therefore in that case Q_{1} converges to its upper bound.

The inequalities (5.4) can be derived in the same way.

6. Tables and nomograms

To obtain the values d_i and e_i as defined by (2.7) and (2.9) and to evaluate the power functions (5.4) and (5.5) we need suitable tables or nomograms of the incomplete ${\cal B}$ function.

When all \mathcal{L}_i are equal, the smallest \mathcal{L}_i corresponds to the largest ratio \mathcal{L}_i and the smallest \mathcal{L}_i corresponds to the smallest ratio \mathcal{L}_i . Further the critical values \mathcal{L}_i of \mathcal{L}_i when testing the largest ratio and \mathcal{L}_i for testing the smallest ratio are then all equal:

$$g_{ij\epsilon}^{(2)} = g_{\epsilon}^{(2)} \quad (i=1,...,k).$$

Therefore in this case it suffices to have tables with these critical values with entries k and the common parameter value k. These tables may be found in C. EISENHART, M.W. HASTAY and W.A. WALLIS (1947) ($\epsilon = aos and aoi$) for the first test and in R. DOORN-BOS (1956) ($\epsilon = aos$) for the second one.

When there are unequal values among the α_c the minimum α value may be found in most cases by means of PEARSON's tables of the incomplete B-function (K. PEARSON (1934)).

The smallest e value, however, will, when it lies in the neighbourhood of $\frac{\xi}{k}$ and k is not very small, correspond to such a small ratio χ that neither these tables nor the nomograms constructed by H.O. HARTLEY and E.R. FITCH (1951) are suitable for our purpose.

An extension of the tables or nomograms to cover this case seems useful.

7. References

- W.G. COCHRAN (1941), The distribution of the largest of a set of estimated variances as a fraction of their total, Annals of Enginics, <u>11</u> (1941) pp 47-52.
- R. DOORNBOS (1956), Significance of the smallest of a set of estimated normal variances, Statistica

 Neerlandica, 10 (1956), pp
- C. EISENHART, M.W. HASTAY and W.A. WALLIS (1947), Selected

 Techniques of Statistical Analysis, Chapter

 15, New York and London, 1947.
- W. FELLER (1950), An Introduction to Probability Theory and its Applications, Volume I, New York and London, 1950.
- H.O. HARTLEY and E.R. FITCH (1951), A chart for the incomplete Bêta-function and the cumulative binomial distribution, Biometrika, 38 (1951), pp 423-426.
- K. PEARSON (1934), Tables of the incomplete Bêta-function, Cambridge, 193^{4} .