Communications to the Editor

A Simplified Test of Scedasticity.

In the field of quality control, the standard deviation of the measurements has been approximately estimated from its range[1], instead of the exact computation of it using the sums of squares, while this principle of range has also been applied for the study of biological assay.[2,3]

In the latter field of study, the routine statistical procedures such as parallel line assay or slope ratio assay include the assumption that the variances of the responses at the respective doses are homogeneous, or, in other words, in homoscedasticity. For the test of this assumption, the Bartlett test[4] has usually been used, but the procedure of the test seems to biologists rather complicated and troublesome, because the result of this analysis is not directly related to the relative potency in which the workers are most interested.

In this situation, it would be desirable to use a simplified test of scedasticity for experimental data, in order to obtain its information promptly, even though it is an approximate computation.*

1) Test of Scedasticity between Two Groups

The range of a set of values is defined as the difference between the maximum and the minimum, and therefore, R_1 and R_2, the ranges of A and B, can be written respectively as

$$R_1 \equiv x_{max} - x_{min} \qquad R_2 \equiv x'_{max} - x'_{min}$$

$$s_1 = \left\{ \sum_{i=1}^{m} (x_{1i} - \bar{x}_1)^2 \right\}^{1/2} \doteqdot \frac{R_1}{d_1} \quad \text{and} \quad s_2 = \left\{ \sum_{i=1}^{n} (x_{2i} - \bar{x}_2)^2 \right\}^{1/2} \doteqdot \frac{R_2}{d_2}$$

TABLE I.

Treatment	A	B
1	x_{11}	x_{21}
2	x_{12}	x_{22}
3	x_{13}	x_{23}
⋮	⋮	⋮
m	x_{1m}	x_{2m}
⋮	⋮	⋮
n	x_{1n}	x_{2n}
mean	x_1	x_2
min.	x_{min}	x'_{min}
max.	x_{max}	x'_{max}
Range	R_1	R_2

TABLE II.

Size of sample	Coefficient	Size of sample	Coefficient
n	d_i	n	d_i
2	1.13	12	3.26
3	1.69	13	3.34
4	2.06	14	3.40
5	2.33	15	3.48
6	2.53	16	3.52
7	2.70	17	3.58
8	2.85	18	3.64
9	2.97	19	3.68
10	3.08	20	3.74
11	3.18		

Then the standard deviations of these are approximately estimated by

1) Y. Ishida: "Applied Statistics," Kokuseido Co. Ltd. (Japan), 24 (1949).
2) The United States Pharmacopoeia XIV, Mark Publishing Co., U.S.A. (1950).
3) L. Kudsen Randolph: J. Am. Pharm. Assoc., Sci. Ed., 61, 438 (1952).
4) M. S. Bartlett: Suppl. J. Roy. Stat. Soc., 4., 137 (1937); G. W. Snedecor: "Statistical Methods", Iowa Univ. Press (1950); B. Tamaoki: J. Japan. Chem., Suppl. No. 11, Nankodo Co. Ltd. (Japan), 11 (1953).

* On the assumption that the distribution of the individual deviations is normal, the best estimate that can be formed from the range is less efficient than the root-mean square estimate, but the efficiency is high when the number of values involved is small.

NII-Electronic Library Service

where d_i is the coefficient shown in Table II, which depends upon the size of the sample. The test of the scedasticity between the two can be expressed as

$$F_0 = \frac{s_1^2}{s_2^2} \fallingdotseq \frac{\left(\frac{R_1}{d_1}\right)^2}{\left(\frac{R_2}{d_2}\right)^2} \sim F^{m-1}_{n-1}\ (\alpha)$$

when R_1/d_1 is greater than R_2/d_2.

When the sample sizes of the two groups are equal, or $m=n$, d_1 becomes equal to d_2, and accordingly this test can be simplified as

$$F_0 = (R_1/R_2)^2$$

where $R_1 > R_2$.

Therefore, when the ratio of these, R_1/R_2, is smaller than $\sqrt{F}$ obtained from F-Table[5] with $P=0.05$ or 0.01, it is concluded that there is no evidence of heterogeneity of variances, or heteroscedasticity.

2) Test of the Scedasticity of more than Two Variances

When the samples are equal in size, say m, the Bartlett's test can easily be simplified by the use of the range. The estimates of the standard deviations in the respective group are computed by the ranges and the corresponding coefficient, d_m.

In Table III, d_m is the coefficient obtained from Table II, and then the test of scedasticity can be written as

$$\chi_0^2 = 2.3026(m-1)\left(n \log \bar{s}^2 - \sum_{i=1}^{n} \log s_i^2\right)$$

where

$$\bar{s}^2 = \left(\sum_{i=1}^{n} \bar{s}_i^2\right)/n$$

When χ^2 lies close above one of the critical values from the χ^2-Table[5], it is necessary to make the correction by the 'Correction Factor,' C, as follows:

$$C = 1 + \frac{n+1}{3n(m-1)}$$

$$\text{Corrected } \chi_0^2 = \chi_0^2/C$$

This procedure is practicably applied for the data obtained by Kitagawa in the oxygen-consumption experiments[6], and the analysis of one of the results is shown in comparison with the one calculated by the original Bartlett's procedure.

TABLE III.

Group	Range	$s_i = R_i/d_m$	s_i^2	$\log s_i^2$
1	R_1	s_1	s_1^2	$\log s_1^2$
2	R_2	s_2	s_2^2	$\log s_2^2$
3	R_3	s_3	s_3^2	$\log s_3^2$
⋮	⋮	⋮	⋮	⋮
n	R_n	s_n	s_n^2	$\log s_n^2$

TABLE IV.

	Original	Simplified
χ_0^2	5.494	6.620
Corrected χ_0^2	5.023	6.046

cf. $\chi^2_{[7]}$ (0.05) $=14.067$

Though there is some but comparatively small difference between the two values of χ_0^2 obtained by the above procedures, the results of both analyses show that there is no evidence of heterogeneity of variances, or heteroscedasticity.

It is quite possible to apply this principle of range for the test of the scedasticity of the samples differing in size, but the computing procedure cannot be simplified as much as stated above.

Pharmaceutical Institute,
Medical Faculty,
University of Tokyo.
Hongo, Tokyo.

Bun-ichi Tamaoki

May 17, 1954

5) R. A. Fisher, F. Yates: "Statistical Tables," Oliver & Boyd Co. Ltd., England (1948).
6) H. Kitagawa: J. Pharm. Soc. Japan, 74, 271 (1954).

NII-Electronic Library Service