The Generalization of Student's Ratio*

Harold Hotelling
Columbia University

The accuracy of an estimate of a normally distributed quantity is judged by reference to its variance, or rather, to an estimate of the variance based on the available sample. In 1908 "Student" examined the ratio of the mean to the standard deviation of a sample.[1] The distribution at which he arrived was obtained in a more rigorous manner in 1925 by R.A. Fisher,[2] who at the same time showed how to extend the application of the distribution beyond the problem of the significance of means, which had been its original object, and applied it to examine regression coefficients and other quantities obtained by least squares, testing not only the deviation of a statistic from a hypothetical value but also the difference between two statistics.

Let ξ be any linear function of normally and independently distributed observations of equal variance, and let s be the estimate of the standard error of ξ derived by the method of maximum likelihood. If we let t be the ratio to s of the deviation of ξ from its mathematical expectation, Fisher's result is that the probability that t lies between t_1 and t_2 is

$$\frac{1}{\sqrt{\pi n}} \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)} \int_{t_1}^{t_2} \frac{dt}{\left(1+\frac{t^2}{n}\right)^{(n+1)/2}} \tag{1}$$

where n is the number of degrees of freedom involved in the estimate s.

* Presented at the meeting of the American Mathematical Society at Berkeley, April 11, 1931.

[1] Biometrika, vol. 6 (1908), p. 1.

[2] Applications of Student's Distribution, Metron, vol. 5 (1925), p. 90.

S. Kotz et al. (eds.), *Breakthroughs in Statistics*

It is easy to see how this result may be extended to cases in which the variances of the observations are not equal but have known ratios and in which, instead of independence among the observations, we have a known system of intercorrelations. Indeed, we have only to replace the observations by a set of linear functions of them which are independently distributed with equal variance. By way of further extension beyond the cases discussed by Fisher, it may be remarked that the estimate of variance s^2 may be based on a body of data not involved in the calculation of ξ. Thus the accuracy of a physical measurement may be estimated by means of the dispersion among similar measurements on a different quantity.

A generalization of quite a different order is needed to test the simultaneous deviations of several quantities. This problem was raised by Karl Pearson in connection with the determination whether two groups of individuals do or do not belong to the same race, measurements of a number of organs or characters having been obtained for all the individuals. Several "coefficients of racial likeness" have been suggested by Pearson and by V. Romanovsky with a view to such biological uses. Romanovsky has made a careful study[1] of the sampling distributions, assuming in each case that the variates are independently and normally distributed. One of Romanovsky's most important results is the exact sampling distribution of L, a constant multiple of the sum of the squares of the values of t for the different variates. This distribution function is given by a somewhat complex infinite series. For large samples and numerous variates it slowly approximates to the normal form; for 500 individuals, Romanovsky considers that an adequate approach to normality requires that no fewer than 62 characters be measured in each individual. When it is remembered that all these characters must be entirely independent, and that it is usually hard to find as many as three independent characters, the difficulties in application will be apparent. To avoid these troubles, Romanovsky proposes a new coefficient of racial likeness, H, the average of the ratios of variances in the two samples for the several characters. He obtains the exact distribution of H, again as an infinite series, though it approaches normality more rapidly than the distribution of L. But H does not satisfy the need for a comparison between magnitudes of characters, since it concerns only their variabilities.

Joint comparisons of correlated variates, and variates of unknown correlations and standard deviations, are required not only for biologic purposes, but in a great variety of subjects. The eclipse and comparison star plates used in testing the Einstein deflection of light show deviations in right ascension and in declination; an exact calculation of probability combining the two least-square solutions is desirable. The comparison of the prices of a list of

[1] V. Romanovsky, On the criteria that two given samples belong to the same normal population (on the different coefficients of racial likeness), Metron, vol. 7 (1928), no. 3, pp. 3–46; K. Pearson, On the coefficient of racial likeness, Biometrika, vol. 18 (1926), pp. 105–118.