PSYCHOMETRIKA—VOL. 6, NO. 5
OCTOBER. 1941

SYNTHESIS OF VARIANCE

FRANKLIN E. SATTERTHWAITE
UNIVERSITY OF IOWA

The distribution of a linear combination of two statistics distributed as is Chi-square is studied. The degree of approximation involved in assuming a Chi-square distribution is illustrated for several representative cases. It is concluded that the approximation is sufficiently accurate to use in many practical applications. Illustrations are given of its use in extending the Chi-square, the Student "*t*" and the Fisher "*z*" tests to a wider range of problems.

Introduction

In the analysis of normally distributed statistics, the variance occupies a prominent role. Therefore a knowledge of the distribution of the statistics used to estimate the variance is important. The Chi-square distribution furnishes an exact method for evaluating the significance of many of the estimates in common use. Thus, to determine the significance of an estimate of variance,

$$\bar{\sigma}^2 = \frac{\sum (x_i - \bar{x})^2}{n - 1},$$

we enter a table of the Chi-square distribution with

$$\chi^2 = \frac{(n-1)\bar{\sigma}^2}{\sigma^2} = \frac{\sum (x_i - \bar{x})^2}{\sigma^2}$$

and $n - 1$ degrees of freedom. We shall call estimates so evaluated *simple estimates of the variance.* However, in practical problems the best estimates available are often not simple estimates but are given by a linear combination of two or more simple estimates. We define such estimates as *complex estimates of variance.* In general, the Chi-square distribution does not give an exact test for a complex estimate of variance. The purpose of this paper is to give an approximation to the distributions of complex estimates of variance which is based on the Chi-square distribution and which is usually accurate enough for practical use in Chi-square, Student "*t*," and Fisher "*z*" tests.

Mathematical Development

Let x and y be two independent simple estimates of variance with expected values $\hat{x}$ and $\hat{y}$, and with degrees of freedom r_1 and r_2. By transforming the formula for the Chi-square distribution, we obtain for the distribution of x,

$$f_1(x) = \frac{1}{\Gamma(r_1/2)} \left(\frac{\hat{r}_1}{2x}\right)^{r_1/2} x^{(r_1-2)/2}\, e^{-r_1 x/2\hat{x}},$$

and a similar distribution, $f_2(y)$, for y. The distribution function of $z = x + y$ will then be

$$F(z) = \int_0^z f_1(x)\, f_2(z-x)\, dx$$

$$= K \int_0^z x^{a-1} (z-x)^{b-1}\, e^{-ax/\hat{x} - b(z-x)\hat{y}} dx,$$

where $a = r_1/2$, $b = r_2/2$. If we now specialize our distribution to the case where r_1 and r_2 are both even integers, we can expand $(z-x)^{b-1}$ by the binomial theorem and integrate on x by parts. The resulting function of z is a linear function of several Chi-square functions which when integrated gives for the probability that z will be greater than some fixed value, w,

$$P(z > w) = \sum_{i=0}^{a-1} \frac{(-1)^i (a+i-1)!}{i!(a-1)!} \Bigg[\left(\frac{a}{\hat{x}}\right)^a \left(\frac{b}{\hat{y}}\right)^i$$

$$\frac{P[x^2 > 2bw/\hat{y}, 2(b-i)\, d.f.]}{[a/\hat{x} - b/\hat{y}]^{a+i}} - \sum_{j=0}^{a+i-1} \frac{(a+b-j-2)!}{(b-i-1)!(a+i-j-1)!}$$

$$\left(\frac{a}{\hat{x}}\right)^{j+1-b} \left(\frac{b}{\hat{y}}\right)^b \frac{P[x^2 > 2aw/\hat{x}, 2(a+b-j-1)\, d.f.]}{[a/\hat{x} - b/\hat{y}]^{j+1}} \Bigg].$$

It is readily seen that, even for the simplified case here considered, the exact distribution of z is too complicated for practical use. We shall therefore approximate the distribution of z by use of a Chi-square distribution with r degrees of freedom. To determine r we impose the condition that both the theoretical and the approximate distributions of z shall have the same variance. On calculating the second moments about the means of $f_1(x)$ and $f_2(y)$, the variances of the distributions of x and y are found to be

$$\sigma_x^2 = \frac{\hat{x}^2}{r_1}, \qquad \sigma_y^2 = \frac{\hat{y}^2}{r_2}.$$

Similarly, the variance of the approximate distribution of z is

$$\sigma_z^2 = \frac{\hat{z}^2}{r} = \frac{(\hat{x}+\hat{y})^2}{r}.$$

Since the variance of a sum of independent variables is equal to the sum of the variances, we therefore have,

$$\sigma_z^2 = \sigma_x^2 + \sigma_y^2$$

or

$$\frac{(\hat{x}+\hat{y})^2}{r} = \frac{\hat{x}^2}{r_1} + \frac{\hat{y}^2}{r_2},$$

from which we may determine the *effective number of degrees of freedom*, r, of the approximate distribution of z. In practice, the expected values, $\hat{x}$ and $\hat{y}$, are usually unknown and are estimated by use of x and y.

In Figure 1 we have plotted both the exact and the approximate distributions of z for several values of the parameters. In each instance the agreement appears to be satisfactory enough for purposes of testing significance. On the basis of the few values so far investigated we should not be justified in making any general statements regarding the degree of approximation involved. However, from general reasoning we should expect the approximation to improve in the following situations:

1. As the r's increase, since both the exact and the approximate distributions approach the same normal distribution with the increase in the number of degrees of freedom.
2. As the ratio $r_1\hat{y}/r_2\hat{x}$ approaches unity, because the approximate and the exact distribution formulas are identical when this ratio is unity.

In general, the theoretical distribution is flatter topped than the Chi-square approximation. If we are going to use the approximation in a Chi-square test, we are interested in the upper end of the distribution and our rule slighly overestimates the best number of degrees of freedom to use. If for distribution (C_1) in Figure 1 we had assumed 5.9 effective degrees of freedom instead of 7.7 effective degrees of freedom, the approximate distribution would have been the dashed line, (C_2), and the theoretical would have been given by the associated dashed curve, a very close agreement at the upper end of the distribution. For the Student "t" test we are interested in the lower end of the Chi-square distribution (which affects the upper end of the "t" distribution) so that our rule slightly underestimates the effective number of degrees of freedom. In Figure 1, the dashed line,

(C_3), gives the approximation to the C distribution based on 9.0 degrees of freedom instead of the 7.7 given by the rule. The corresponding theoretical values given by the associated dashed curve show a very good fit at the lower end. For a Fisher "z" test our rule slightly

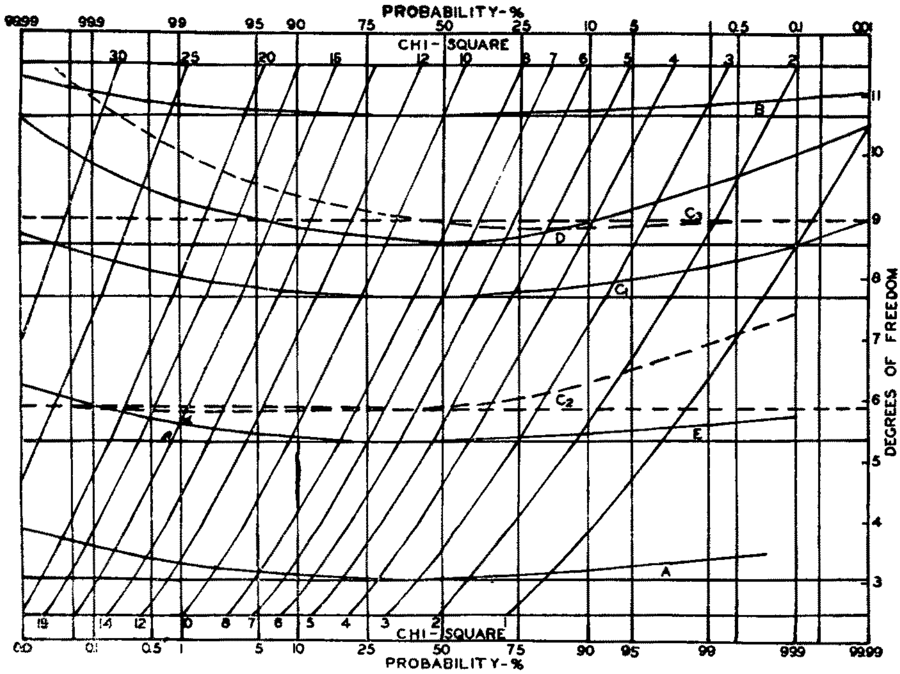

FIGURE 1

The exact distributions of the complex estimates of variance listed below are plotted on the chart as horizontal curved lines. The chart is so designed that any horizontal straight line corresponds to a Chi-square distribution with the number of degrees of freedom given by the vertical scale. The approximation suggested in this paper is given by the straight line tangent to the theoretical distribution. The degree of approximation may be measured numerically by following along lines of constant Chi-square. For example, a Chi-square of 16 for the (E) estimate of variance corresponds to a theoretical probability of 1% as is given by point, α, on the chart. However, if one uses the approximate distribution, he would read the point, β, obtaining a probability of 0.8% instead of the correct value of 1%. The distributions plotted are for complex estimates of variance with the following parameters:

	r_1	r_2	$\hat{x}/\hat{y}$	$r_1\hat{y}/r_2\hat{x}$	r
(A)	4	2	1/4	8	100/33
(B)	8	4	1	2	32/3
(C)	6	4	1/2	3	54/7
(D)	20	4	1/2	10	180/21
(E)	4	2	1	2	16/3

overestimates the effective degrees of freedom in the numerator and underestimates them in the denominator. No attempt has been made to estimate the foregoing adjustments without actually calculating the theoretical distribution. However, they are apparently of small enough magnitude to have little effect on the conclusions drawn in significance tests.

The analysis above is readily extended to the case when z is a linear function of several simple estimates of variance. If

$$z = a_1 x_1 + a_2 x_2 + \cdots$$

where x_i has the expected value, $\hat{x}_i$, and its distribution has r_i degrees of freedom, then $rz/\hat{z}$ is approximately distributed as is Chi-square, with r degrees of freedom, where r is determined by the equation,

$$\frac{(a_1\hat{x}_1 + a_2\bar{x}_2 + \cdots)^2}{r} = \frac{(a_1\,\hat{x}_1)^2}{r_1} + \frac{(a_2\hat{x}_2)^2}{r_2} + \cdots. \qquad (1)$$

If some of the a's are negative, special care should be exercised in using the x's to estimate the x's in (1). If the true value of $\hat{r}$ is small, major errors may result from such an approximation. Also the theoretical distribution in this case is not necessarily flatter topped than the approximation so that the foregoing remarks regarding the degree of approximation do not apply.

Applications

The first application we shall make is to the Student "t" test of the significance of the difference of the means of two samples. The usual estimate of variance used in this test is,

$$\bar{\sigma}_A{}^2 = \frac{\sum(x_{1j} - \bar{x}_1)^2 + \sum(x_{2j} - \bar{x}_2)^2}{N_1 + N_2 - 2} \cdot \sqrt{\frac{1}{N_1} + \frac{1}{N_2}},$$

with $N_1 + N_2 - 2$ degrees of freedom. The use of this estimate is based on the assumption that both samples were drawn from normal populations with the same variance. Frequently we do not have sufficient evidence to justify such an assumption of homogeneity of variance, and sometimes we have definite evidence of a lack of homogeneity. To avoid making such an assumption we can use as an estimate of the variance of the difference,

$$\begin{aligned}\bar{\sigma}_B{}^2 &= \frac{\bar{\sigma}_1{}^2}{N_1} + \frac{\bar{\sigma}_2{}^2}{N_2} \\ &= \frac{\sum(x_{1j} - \bar{x}_1)^2}{N_1(N_1 - 1)} + \frac{\sum(x_{2j} - \bar{x}_2)^2}{N_2(N_2 - 1)}.\end{aligned} \qquad (2)$$

We have shown that while the use of such complex estimates of variance in Student "t" tests is not strictly correct, the probabilities calculated are approximately correct when the number of degrees of freedom, r, is calculated from equation (1). Thus

$$\frac{\tilde{\sigma}_B^4}{r} = \frac{(\tilde{\sigma}_1^2/N_1)^2}{(N_1 - 1)} + \frac{(\tilde{\sigma}_2^2/N_2)^2}{(N_2 - 1)}.$$

The technique presented in this paper appears to have wide application when we are dealing with data subject to what we shall call "random classification." For example, if we are studying students and classify them according to schools, we have a random classification if the particular schools entering into our experiment are considered to be a random sample from a population of schools. On the other hand, if we classify the students by methods of teaching, we do not have random classification because the primary object of study is here the particular methods used, not some population of methods.

To illustrate the application of our theory to random classification, we shall examine the variance of certain normally distributed data, x_{ij}, $(i = 1, 2, \cdots, a; j = 1, 2, \cdots, b)$ which fall into classes identified by the subscript, i. We assume that there is significant variation between classes. Letting m_i be the expected value of the x's in the ith class, we assume that the m_i's are normally distributed about a mean $\hat{m}$. The following table lists the different variances entering into this problem with the corresponding estimates indicated by a wave, (~).

$$\sigma_1^2 = \text{variance of } (x_{ij} - m_i)\,;$$

$$\tilde{\sigma}_1^2 = \sum (x_{ij} - \bar{x}_i)^2/a(b-1).$$

$$\sigma_2^2 = \text{variance of } (\bar{x}_i - \hat{m}) \equiv (\bar{x}_i - m_i) + (m_i - \hat{m})\,;$$

$$\tilde{\sigma}_2^2 = \sum (\bar{x}_i - \bar{x})^2/(a-1).$$

$$\sigma_3^2 = \text{variance of } (\bar{x}_i - m_i)\,;$$

$$\tilde{\sigma}_3^2 = \tilde{\sigma}_1^2/b\,.$$

$$\sigma_4^2 = \text{variance of } (m_i - \hat{m})\,;$$

$$\tilde{\sigma}_4^2 = \tilde{\sigma}_2^2 - \tilde{\sigma}_1^2/b \quad \text{since } \sigma_2^2 = \sigma_3^2 + \sigma_4^2.$$

We see that σ_4^2 is the true variance between classes and that its estimate is a complex estimate. If we should desire to use $\tilde{\sigma}_4^2$ in a statistical test, we may assume that $r_4\tilde{\sigma}_4^2/\sigma_4^2$ is distributed approximately as is Chi-square with r_4, the number of degrees of freedom, determined from equation (1). Thus.

$$\frac{\hat{\sigma}_4{}^4}{r_4} = \frac{\hat{\sigma}_2{}^4}{a-1} + \frac{(\hat{\sigma}_1{}^2/b)^2}{a(b-1)}.$$

Given the data above, we frequently want to determine confidence limits within which we may expect an additional item in a new class (i.e., $x_{a+1,j}$) to fall. Two additional variances enter into this problem, namely,

$$\sigma_5{}^2 = \text{variance of } (\bar{x} - \hat{m}),$$

$$\hat{\sigma}_5{}^2 = \hat{\sigma}_2{}^2/a \text{ , since } \bar{x} \text{ is an average of the } \bar{x}_i\text{'s;}$$

$$\sigma_6{}^2 = \text{variance of } (x_{a+1,j} - \bar{x}) \equiv (x_{a+1,j} - m_{a+1})$$
$$+ (m_{a+1} - \hat{m}) - (\bar{x} - \hat{m});$$

$$\hat{\sigma}_6{}^2 = \hat{\sigma}_1{}^2 + (\hat{\sigma}_2{}^2 - \hat{\sigma}_1{}^2/b) + \hat{\sigma}_2{}^2/a$$
$$= (1 - 1/b)\hat{\sigma}_1{}^2 + (1 + 1/a)\hat{\sigma}_2{}^2.$$

The appropriate number; r_6; of degrees of freedom for the approximate distribution of $\hat{\sigma}_6{}^2$ is obtained from equation (1). Thus,

$$\frac{\hat{\sigma}_6{}^4}{r_6} = \frac{[(1 - 1/b)\hat{\sigma}_1{}^2]^2}{a(b-1)} + \frac{[(1 + 1/a)\hat{\sigma}_2{}^2]^2}{a-1}.$$

Then, determining the value of the Student "t" for the desired confidence level and r_6 degrees of freedom, we should expect $x_{a+1,j}$ usually to fall within the limits,

$$\bar{x} - t\,\hat{\sigma}_6 \leqslant x_{a+1,j} \leqslant x + t\,\hat{\sigma}_6.$$

For our last example we shall take two groups of students, x_{ij} and y_{ij}, each group being classified according to instructor. The instructors in the x group each used a different teaching method while those in the y group all used the same method. We should like to determine whether or not the differences observed between the methods in the x group are significantly greater than would have been expected because of variation in the ability of the instructors. We therefore compare by use of the Fisher "z" test the estimate ($\hat{\sigma}_4{}^2$ above) of the true variance between instructors for the x group with the corresponding estimate for the y group. Since $\hat{\sigma}_4{}^2$ is a complex estimate of variance, the appropriate number of degrees of freedom must be determined from equation (1).

Conclusion

This paper gives an extension of the Chi-square, Student "t," and Fisher "z" tests to situations where the Chi-square distribution does not give an exact evaluation of the estimates of variance used. This condition arises wherever the variance of a statistic can not be estimated directly but must be estimated by a linear function of two or more independent estimates. The analysis of variance provides us with a powerful method for splitting the variance of a statistic into its elementary factors. The synthesis of variance provides a method for constructing the variance of complex statistics out of such elementary factors.

Historical Note

The problem treated in this paper has been recognized by R. A. Fisher and others, particularly in connection with testing the significance of the difference between means. R. A. Fisher suggested a method of solution as an illustration of fiducial theory. B. L. Welch compared several suggested approximations and illustrated their bias as compared with the approximation used in this paper. He did not investigate the accuracy of the approximation here used. P. L. Hsu investigated the problem from the standpoint of errors of the second kind (Pearson-Neyman theory). He found the general distribution of $(\bar{x}_1 - \bar{x}_2)/\bar{\sigma}_B$ [equation (2)] which, as we would expect, is very complicated. Where comparable he confirmed the conclusions of Welch.

REFERENCES

Fisher, R. A. The fiducial argument in statistical inference. *Annals of Eugenics*, 1936, 6, 391.

Welch, B. L. The significance of the difference between two means when the population variances are unequal. *Biometrika*, 1938, 29, 350.

Hsu, P. L. Contribution to the theory of "Student's" *t*-test as applied to the problem of two samples. *Statistical Research Memoirs*, 1938, 2, 1.