[3] J. LEV, "The point-biserial coefficient of correlation," *Ann. Math. Stat.*, Vol. 20 (1949), pp. 125–126.
[4] J. NEYMAN AND B. TOKARSKA, "Errors of the second kind in testing 'Student's' hypothesis," *J. Amer. Stat. Assn.*, Vol. 31 (1936), pp. 318–326.
[5] K. PEARSON, *Tables for Statisticians and Biometricians*, Part I, Cambridge University Press, 1945, p. 20.

A COMPUTING FORMULA FOR THE POWER OF THE ANALYSIS OF VARIANCE TEST

BY W. L. NICHOLSON

University of Oregon[1] *and University of Illinois*[2]

1. Summary. A formula for the power of the analysis of variance test is derived for the case when the denominator of the F ratio has an even number of degrees of freedom. The form employed is particularly adapted to computation of the power as a function of the alternative hypothesis with arbitrary fixed level of significance and fixed degrees of freedom. For m degrees of freedom in the numerator and 2, 4, 6, 8 and 10 in the denominator, the power functions are deduced from the general formula, with an indication of their use.

2. The power function. In the classical analysis of variance test we are interested in a ratio of the form

$$F = n \sum_{i=1}^{m} x_i^2 \Big/ m \sum_{j=1}^{n} y_j^2, \tag{1}$$

where x_i $(i = 1, 2, \cdots, m)$ and y_j $(j = 1, 2, \cdots, n)$ are distributed $N(\theta_i, \sigma^2)$ and $N(0, \sigma^2)$, respectively. If the null hypothesis, $\theta_i = 0$ $(i = 1, 2, \cdots, m)$, is false, it is well known that the distribution of F is completely specified by m, n, and the single additional parameter

$$\lambda = \frac{1}{2\sigma^2} \sum_{i=1}^{m} \theta_i^2. \tag{2}$$

Therefore, for a predetermined level of significance α, the power of the test is a function of m, n, and λ. It is [1]

$$P(\lambda \mid a, b; \alpha) = 1 - \sum_{k=0}^{\infty} e^{-\lambda} \frac{\lambda^k}{k!} I_x(a + k, b), \tag{3}$$

where $m = 2a$ and $n = 2b$, and

$$I_x(a + k, b) = \frac{1}{\beta(a + k, b)} \int_0^x t^{a+k-1}(1 - t)^{b-1}\, dt \tag{4}$$

Received 2/24/52, revised 2/12/53.

[1] Research under contract with the Office of Naval Research, U. S. Navy.
[2] Research under contract with the Office of Ordnance Research, U. S. Army.

is the incomplete beta function [2] with $x = x(a, b, 1 - \alpha)$ defined as

$$I_x(a, b) = 1 - \alpha. \tag{5}$$

Limiting the argument to integral values of b, it follows that

$$I_x(a + k, b) = \Gamma(a + b + k) \sum_{j=0}^{b-1} \frac{x^{a+k+j}(1 - x)^{b-j-1}}{\Gamma(a + k + j + 1)\,\Gamma(b - j)}. \tag{6}$$

As in [1] we obtain

$$\begin{aligned} &P(\lambda \mid a, b; \alpha) \\ &= 1 - e^{-\lambda(1-x)} \sum_{k=0}^{b-1} \frac{\Gamma(a + b)x^{a+k}(1 - x)^{b-k-1}}{\Gamma(a + k + 1)\Gamma(b - k)} F(k + 1 - b, a + k + 1, -\lambda x) \end{aligned} \tag{7}$$

where

$$F(\alpha, \beta, z) = \sum_{j=0}^{\infty} \frac{\alpha^{(j)}}{\beta^{(j)}} \frac{z^j}{j!}. \tag{8}$$

Here, for any number d, the symbol $d^{(j)}$ is defined by

$$d^{(j)} = \begin{cases} \prod_{i=0}^{j-1} (d + i) & j \geqq 1 \\ 1 & j = 0 \end{cases}$$

Interchange the order of summation, utilize the identity

$$\sum_{j=0}^{c} (-1)^j \binom{c}{j} \frac{z^j}{d + j} = \sum_{j=0}^{c} \frac{c_{(j)}}{d^{(j+1)}} z^j(1 - z)^{c-j}, \quad c_{(j)} = \begin{cases} \prod_{i=0}^{j-1} (c - i), & j \geqq 1, \\ 1 & j = 0, \end{cases} \tag{9}$$

and note that $P(0 \mid a, b; \alpha) = \alpha$, then (7) reduces to

$$\begin{aligned} &P(\lambda \mid a, b; \alpha) = 1 - e^{-\lambda(1-x)} \{(1 - \alpha) \\ &+ x^a \sum_{j=1}^{b-1} \frac{x^j(1 - x)^j}{(b - j - 1)!\,j!} \left[\sum_{k=0}^{b-j-1} (-1)^k \binom{b - j - 1}{k} \frac{(a + b - 1)_{(b-j)}}{a + j + k} x^k \right] \lambda^j \Big\}. \end{aligned} \tag{10}$$

For fixed m, n and α the power may be rapidly calculated as a function of λ.

3. An error term. The identity (9) may be replaced by

$$\sum_{j=0}^{c} \frac{c_{(j)}}{d^{(j+1)}} z^j(1 - z)^{c-j} = \frac{\beta(d, c + 1)}{z^d} I_z(d, c). \tag{11}$$

Then a similar argument as before gives from (7)

$$P(\lambda \mid a, b; \alpha) = 1 - e^{-\lambda(1-x)} \sum_{k=0}^{b-1} \frac{[\lambda(1 - x)]^k}{k!} I_x(a + k, b - k). \tag{12}$$

Since $I_x(a+k, b-k)$ is a nonincreasing function of k ($k = 0, 1, 2, \cdots$), neglecting the last $b - r$ terms in the general formula (10) will result in an error of

$$R(\lambda \mid a, b, r; \alpha) < I_x(a+r, b-r)e^{-\lambda(1-x)} \sum_{k=r}^{b-1} \frac{[\lambda(1-x)]^k}{k!}. \tag{13}$$

The finite sum of Poisson terms may be evaluated with the aid of the incomplete Γ-function,

$$\Gamma(u, p) = \frac{1}{\Gamma(p+1)} \int_0^{u\sqrt{p+1}} t^p e^{-t}\, dt, \tag{14}$$

which is tabulated [4]. Maximizing (13) with respect to λ gives as a uniform bound for the error

$$R(\lambda \mid a, b, r; \alpha) < I_x(a+r, b-r)\,[\Gamma(L/\sqrt{r}, r-1) - \Gamma(L/\sqrt{b}, b-1)], \tag{15}$$

where $L = [\prod_{j=1}^{r} (b-j)]^{1/(b-r-1)}$.

4. Special cases. For $b = 1$, $x^a = (1 - \alpha)$, so (10) simplifies to

$$P(\lambda \mid a, 1; \alpha) = 1 - (1-\alpha)\exp\{-\lambda[1 - (1-\alpha)^{1/a}]\}. \tag{16}$$

For $b = 2, 3, 4$ and 5 the power functions are

$$P(\lambda \mid a, 2; \alpha) = 1 - e^{-\lambda(1-x)}\,[(1-\alpha) + x^{a+1}(1-x)\lambda], \tag{17}$$

$$\begin{aligned} P(\lambda \mid a, 3; \alpha) = 1 &- e^{-\lambda(1-x)} \\ &\cdot\{(1-\alpha) + x^{a+1}(1-x)[(a+2) - (a+1)x]\lambda + \tfrac{1}{2}x^{a+2}(1-x)^2\lambda^2\}, \end{aligned} \tag{18}$$

$$\begin{aligned} P(\lambda \mid a, 4; \alpha) = 1 - e^{-\lambda(1-x)}\{&(1-\alpha) \\ &+ \tfrac{1}{2}x^{a+1}(1-x)[(a+3)(a+2) - 2(a+3)(a+1)x \\ &+ (a+2)(a+1)x^2]\lambda + \tfrac{1}{2}x^{a+2}(1-x)^2[(a+3) - (a+2)x]\lambda^2 \\ &+ \tfrac{1}{6}x^{a+3}(1-x)^3\lambda^3\}, \end{aligned} \tag{19}$$

$$\begin{aligned} P(\lambda \mid a, 5; \alpha) = 1 &- e^{-\lambda(1-x)}\{(1-\alpha) \\ &+ \tfrac{1}{6}x^{a+1}(1-x)[(a+4)(a+3)(a+2) - 3(a+4)(a+3)(a+1)x \\ &+ 3(a+4)(a+2)(a+1)x^2 - (a+3)(a+2)(a+1)x^3]\lambda \\ &+ \tfrac{1}{4}x^{a+2}(1-x)^2[(a+4)(a+3) - 2(a+4)(a+2)x \\ &+ (a+3)(a+2)x^2]\lambda^2 + \tfrac{1}{6}x^{a+3}(1-x)^3[(a+4) - (a+3)x]\lambda^3 \\ &+ x^{a+4}(1-x)^4\lambda^4/24\}. \end{aligned} \tag{20}$$

Values of the parameter $x = x(a, b, 1-\alpha)$ corresponding to $\alpha = 0.50, 0.25, 0.10, 0.05, 0.025, 0.01$ and 0.005 are tabled [3]. Other values may be interpolated from [2].

EXAMPLE 1. The power function for $\alpha = 0.05$, $m = 2$ and $n = 8$ would be obtained from (19) with $a = 1$. From [3], $x(1, 4, 0.95) = 0.52713$; substituting gives the power function as

$$P(\lambda \mid 1, 4; 0.05) = 1 - e^{-0.47287\lambda}(0.95000 + 0.34381\,\lambda + 0.03961\,\lambda^2 + 0.00136\,\lambda^3). \tag{21}$$

EXAMPLE 2. Suppose that two-figure accuracy is desired in calculating the power function for $\alpha = 0.05$, $m = 8$ and $n = 30$. The unabridged form of (10) with $a = 4$ and $b = 15$ would entail evaluating 15 terms. From (15),

$$R(\lambda \mid 4, 15, 8; 0.05) < 0.003.$$

Thus using the first eight terms of (10) would certainly secure the necessary accuracy.

REFERENCES

[1] P. C. TANG, "The power function of the analysis of variance test with tables and illustrations of their use," *Statistical Research Memoirs*, Vol. 2 (1938), p. 143.

[2] KARL PEARSON, *Tables of the Incomplete Beta Function*, Cambridge University Press, 1934.

[3] CATHERINE M. THOMPSON, "Tables of percentage points of the incomplete beta-function," *Biometrika*, Vol. 32 (1941), pp. 151–153.

[4] KARL PEARSON, *Tables of the Incomplete Γ-Function*, His Majesty's Stationery Office, London, 1922.

POWER UNDER NORMALITY OF SEVERAL NONPARAMETRIC TESTS

BY W. J. DIXON

University of Oregon[1]

1. Summary. Presented are tabulations of the power and power efficiency of four nonparametric tests (rank-sum, maximum deviation, median, and total number of runs) for the difference in means of two samples drawn from normal populations with equal variance. The cases considered are for equal sample sizes of three, four and five observations and alternatives $\delta = |\mu_1 - \mu_2|/\sigma$.

2. Introduction. One method of comparison of various nonparametric tests is a study of their performance under the assumption of normality. An advantage of this method is the wide use of the normal assumption. Disadvantages are the limitation to a particular type of distribution and the extensive computation necessary.

The computation of power under normality is simplest for small samples and small levels of significance. This fact has guided the present study, but it is

Received 6/8/53, revised 2/6/54.

[1] Research sponsored in part by the Office of Naval Research.