New Tables of Behrens' Test of Significance

By R. A. Fisher
[*Department of Genetics, Cambridge University*]
and
M. J. R. Healy
Rothamsted Experimental Station

[Received July, 1956]

SUMMARY

The Tables for using Behren's test of significance of the difference between the means of two Normal samples are already available to reasonable accuracy for all cases save those involving very small samples (*Statistical Tables*, Tables VI, V2). In some experimental work, however, such small samples occur in conditions in which only Behren's test is appropriate. The calculation of values for such cases is greatly facilitated by the explicit expressions developed in this note for the probability integral, when the numbers of degrees of freedom are both odd.

1. *The Characteristic Function of Student's t*

The characteristic function of Student's distribution,

$$M(u) = \frac{\frac{n-1}{2}\,!}{\frac{n-2}{2}\,!\,\sqrt{\pi}} \int_{-\infty}^{\infty} \frac{\cos(ux\sqrt{n})\,du}{(1+x^2)^{(n+1)/2}},$$

is, when n is odd, expressible in the form

$$e^{-|u\sqrt{n}|}\, S_n(|u\sqrt{n}|),$$

in which S is a polynomial of degree,

$$\frac{1}{2}(n-1),$$

with the recurrence relation, demonstrable for any value of m

$$S_{m+3}(u) = S_{m+1}(u) + \frac{u^2}{m^2-1} S_{m-1}(u)$$

beginning with

$$S_1 = 1,$$

$$S_3 = 1 + u;$$

whence, applying the recurrence equation

$$S_5 = 1 + u + \frac{1}{3}u^2,$$

$$S_7 = 1 + u + \frac{2}{5}u^2 + \frac{1}{15}u^3,$$

$$S_9 = 1 + u + \frac{3}{7}u^2 + \frac{2}{21}u^3 + \frac{1}{105}u^4,$$

$$S_{11} = 1 + u + \frac{4}{9}u^2 + \frac{1}{9}u^3 + \frac{1}{63}u^4 + \frac{1}{945}u^5$$

$$S_{13} = 1 + u + \frac{5}{11}u^2 + \frac{4}{33}u^3 + \frac{2}{99}u^4 + \frac{1}{495}u^5 + \frac{1}{10395}u^6,$$

or, more generally

$$S_m(u) = 1 + u + \frac{m-3}{2(m-2)}u^2 + \frac{m-5}{6(m-2)}u^3 + \ldots + \frac{(m-r-1)!\,\frac{m-1}{2}!\,2^r}{(m-1)!\left(\frac{m-1}{2}-r\right)!\,r!}u^r + \ldots$$

with values of r from zero to $\frac{1}{2}(m-1)$

2. *Behrens' Distribution as a "Mixture" of Student's Distributions*

Since the characteristic function of Student's t based on an odd number of degrees of freedom is of the simple form

$$e^{-|\sqrt{n}u|}\, S_n(\,|\sqrt{n}u|\,),$$

that of any multiple $s\,t$ must be

$$e^{-|s\sqrt{n}u|}\, S_n(\,|s\sqrt{n}u|\,),$$

and that of any compound

$$s_1t_1 \pm s_2t_2$$

is simply

$$e^{-(s_1\sqrt{n_1}+s_2\sqrt{n_2})|u|}\, S_{n_1}(s_1\sqrt{n_1}\,|u|\,)\, S_{n_2}(s_2\sqrt{n_2}\,|u|\,),$$

where s_1 and s_2 are positive.

Since the product,

$$S_{n_1} S_{n_2},$$

is simply a polynomial in u of degree $\frac{1}{2}(n_1 + n_2) - 1$, it can be expanded in a finite series of terms

$$S_m\{(s_1\sqrt{n_1} + s_2\sqrt{n_1})\,|u|\,\},$$

in which the highest value of m is

$$n_1 + n_2 - 1$$

For example

$$S_3(\sqrt{3}\sin\theta\,.\,u)\,S_5(\sqrt{5}\cos\theta\,.\,u)$$

$$= (1 + u\sqrt{3}\sin\theta)(1 + u\sqrt{5}\cos\theta + \frac{1}{3}u^2\,5\cos^2\theta)$$

$$= 1 + (\sqrt{3}\sin\theta + \sqrt{5}\cos\theta)u + (\sqrt{15}\sin\theta\cos\theta + \frac{5}{3}\cos^2\theta)u^2 + \frac{5\sqrt{3}}{3}\sin\theta\cos^2\theta\,.\,u^3$$

$$= \frac{25\sqrt{3}\tan\theta}{(\sqrt{5}+\sqrt{3}\tan\theta)^3} S_7\{(\sqrt{3}\sin\theta + \sqrt{5}\cos\theta)\,u\}$$

$$+ \frac{5\sqrt{5} - 10\sqrt{3}\tan\theta + 9\sqrt{5}\tan^2\theta}{(\sqrt{5}+\sqrt{3}\tan\theta)^3} S_5\{(\sqrt{3}\sin\theta + \sqrt{5}\cos\theta)\,u\}$$

$$+ \left\{\left(\frac{\sqrt{3}\tan\theta}{\sqrt{5}+\sqrt{3}\tan\theta}\right)^3\right\} S_3\{(\sqrt{3}\sin\theta + \sqrt{5}\cos\theta)\,u\}$$

the number of terms required being $\frac{1}{2}(n_2 + 1)$ where n_2 is the higher of the two suffices in the product.

3. *Expansion of the Probability Integral of Behrens' d*

For odd degrees of freedom the probability integral of Student's distribution is expressible in terms of the parametric angle, α, such that

$$t = \sqrt{n} \, . \tan \alpha$$

in a terminating series, such that the probability in a single tail is

$$\frac{1}{2} - \frac{1}{\pi}\left\{\alpha + \sin\alpha\cos\alpha + \frac{2}{3}\sin\alpha\cos^3\alpha + \ \ldots \ + \frac{2\,.\,4\ \ldots\ (n-3)}{3\,.\,5\ \ldots\ (n-2)}\sin\alpha\cos^{n-2}\alpha\right\};$$

so if

$$d = (\sqrt{n_1}\sin\theta + \sqrt{n_2}\cos\theta)\tan\alpha,$$

the probability integral of d must be, for the case

$$n_1 = 3_1\, n_2 = 5$$

$$P = -\frac{25\sqrt{3}\tan\theta}{(\sqrt{5}+\sqrt{3}\tan\theta)^3}\cdot\frac{8}{15\pi}\sin\alpha\cos^5\alpha$$

$$-\left\{1 - \left(\frac{\sqrt{3}\tan\theta}{\sqrt{5}+\sqrt{3}\tan\theta}\right)^3\right\}\cdot\frac{2}{3\pi}\sin\alpha\cos^3\alpha$$

$$+\frac{1}{2} - \frac{1}{\pi}(\alpha + \sin\alpha\cos\alpha)$$

so giving an explicit expression, similar to that of "Student", for the probability of exceeding any chosen value $+\,d$, on the null hypothesis that the two distributions sampled have in fact the same mean.

4. *Tabulation of Algebraic Forms for Simple Cases*

The explicit expressions for the probability integral, for the numerous cases provided by different combinations of the numbers n_1 and n_2, are sufficiently heavy for it to be advantageous to use a certain amount of abbreviation in their tabulation.

The expansion appearing in the integral of Student's distribution, for a single tail,

$$\frac{1}{2} - \frac{1}{\pi}\left\{\alpha + \sin\alpha\cos\alpha + \frac{2}{3}\sin\alpha\cos^3\alpha + \ \ldots \right.$$

$$\left. + \frac{2\,.\,4\ \ldots\ (n-3)}{3\,.\,5\ \ldots\ (n-2)}\sin\alpha\cos^{n-2}\alpha\right\}$$

may be shortly written

$$\frac{1}{2} - \frac{1}{\pi}(a_0 + a_1 + a_2 + \ \ldots \ + a_{(n-1)/2})$$

and it is only the coefficients of the terms in which $a_1, a_2, \ \ldots$ occur that depend on the values n_1, n_2, and θ, with which the table is entered. These coefficients are sufficiently simple functions of the ratio

$$\sqrt{(n_1/n_2)}\tan\theta,$$

so that if we write u for

$$\sqrt{n_2}/(\sqrt{n_1}\tan\theta + \sqrt{n_2})$$

and v for

$$\sqrt{n_1}\tan\theta/(\sqrt{n_1}\tan\theta + \sqrt{n_2})$$

then the expansions going no further than a_6 may be shown as in Table 1.

5. *Arithmetical Tabulation*

The significance levels were computed on the NRDC-Elliott 401 computer of which a brief description has been published (Lipton 1955).

The expression for the probability integral can be written

$$P = \frac{1}{2} - \frac{1}{\pi}[\alpha + \sin\alpha\cos\alpha(\lambda_0 + \lambda_2\cos^2\alpha + \lambda_4\cos^4\alpha + \ \ldots \ )]$$

TABLE 1

Coefficients of a_1 to a_6 in the Expansion of Behrens' Integral

$n_1 . n_2$	a_1	a_2	a_3	a_4	a_5	a_6
1 . 1	—	—	—	—	—	—
1 . 3	u	—	—	—	—	—
1 . 5	u	u^2	—	—	—	—
1 . 7	u	$u^2(1+v/5)$	u^3	—	—	—
1 . 9	u	$u^2(1+2v/7)$	$u^3(1+3v/7)$	u^4	—	—
1 . 11	u	$u^2(1+v/3)$	$u^3(21+13v+v^2)/21$	$u^4(1+2v/3)$	u^5	—
1 . 13	u	$u^2(1+4v/11)$	$u^3(11+8v+v^2)/11$	$u^4(33+32v+5v^2)/33$	$u^5(1+10v/11)$	u^6
3 . 3	1	$3u$	—	—	—	—
3 . 5	1	$1-v^3$	$5u^2v$	—	—	—
3 . 7	1	$1-v^3$	$u^2(1+2v+3v^2)$	$7u^3v$	—	—
3 . 9	1	$1-v^3$	$u^2(7+14v+21v^2+3v^3)/7$	$u^3(1+3v+6v^2)$	$9u^4v$	—
3 . 11	1	$1-v^3$	$u^2(3+6v+9v^2+2v^3)/3$	$u^3(3+9v+18v^2+15v^3)/3$	$u^4(1+4v+10v^2)$	$11u^5v$
5 . 5	1	1	$5uv(1-v+v^2)$	$35u^2v^2/3$	—	—
5 . 7	1	1	$1-v^5$	$7u^2v(1-v+2v^2)$	$21u^3v^2$	—
5 . 9	1	1	$1-v^5$	$u^2(1+2v+3v^2+4v^3+5v^4)$	$3u^3v(3-3v+10v^2)$	$33u^4v^2$
7 . 7	1	1	1	$7uv(1-v+v^2)^2$	$42u^2v^2(3-5v+5v^2)/5$	$231u^3v^3/5$

where the coefficients λ are functions of n_1, n_2 and θ. We require the value of α (and hence of $d = (\sqrt{n_1} \sin \theta + \sqrt{n_2} \cos \theta) \tan \alpha$) which makes $P = P^*$, where P^* is pre-assigned. The method used was straightforward. The function $F = P^* - P$ was first evaluated at $\alpha = \pi/4$, and its value at the point compared with that at $\alpha = \pi/2$. If the signs of these two values differed, F was next evaluated at $3\pi/8$, and so on. At each stage, the new value of F was compared with the old value, and a step of half the previous length taken in the appropriate direction. When the step length was zero to the accuracy used (31 binary places, or about 9½ decimals), d was computed from the current value of α and printed out, together with the value of F as a check on machine operation.

The values of the λ's were computed by hand and fed into the machine before the computations. The programme was organized so that all the results for given n_1 and n_2, comprising seven values of θ for each of four significance levels, were printed out consecutively, the whole process taking about 3½ minutes.

The numerical values for 10 combinations of (n_1, n_2) are to appear in the fifth edition of Fisher and Yates' *Statistical Tables* for the application of Behrens' test of the means of two normal samples.

Reference

LIPTON, S. (1955), "A note on the electronic computer at Rothamsted", *Math. Tab., Wash.*, **9**, 69–70.