# 285

# THE WEIGHTED MEAN OF TWO NORMAL SAMPLES WITH UNKNOWN VARIANCE RATIO

SIR RONALD A. FISHER, F.R.S.

*Division of Mathematical Statistics, CSIRO, Adelaide*

*SUMMARY.* The exact small-sample solution of the problem of the weighted mean, flows from an analysis parallel with that required for Behrens' problem. The algebraic forms are, however, more complex, involving one more parameter, for which it is convenient to choose Sukhatme's $d$. The logical basis of choice between weighted and unweighted means is considered.

## 1. ANALYTICAL PROCEDURE

Since the analytic demonstration of Behrens' test of significance for a difference between the means of two Normal samples has been set out explicitly (Fisher, 1935), it has been apparent that the problem of the distribution of the weighted mean of two such samples could be resolved, in the exact terms appropriate to small samples, by a similar analysis. In its mathematical generality, however, this problem is a more difficult one than the simple test of significance of the difference, for in addition to involving the two degrees of freedom, $n_1$ and $n_2$, for the two samples, the modular angle

$$\tan\,\theta = s_1/s_2$$

of the ratio of the two standard deviations, and the level of significance required, it involved also as a fifth parameter, the measure of discrepancy of Behrens' test as defined by Sukhatme, namely

$$d = (\bar{x}_1 - \bar{x}_2)/\sqrt{s_1^2 + s_2^2}$$

and this to no unimportant an extent.

In a recent paper in *Sankhyā* (Fisher, 1961) I have shown how this test, like that of Behrens, can be exhibited as a verifiable assertion about frequencies, as must indeed always be the case when the Reference Set adverted to in any statements of probability has been explicitly specified.



494 *Sankhyā*, 23: 103-114, (1961).

Whereas Behrens' distribution is symmetrical, and so can be conveniently tabulated by giving the deviate, positive or negative, determining limits outside which $d$ shall fall with a specified frequency, in random samples from the reference set, as is the case with Student's $t$, the distribution of the weighted mean is unsymmetrical in the general case. The analytical method of asymptotic expansion, which will be used in this paper, is however, just as applicable as it is in the case of Behrens' test (Fisher, 1941).

## 2. Specification of the analysis

The ordinate of a Student's random variable $t$, for $n$ degrees of freedom, may be expanded in Student polynomials, $P_r$ in the form

$$y = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2} \sum_0^\infty P_r\, n^{-r}, \qquad \ldots \quad (1)$$

in which the first six polynomials are

$$\left.\begin{array}{lr}
P_0 = 1 & \\
P_1 = t^4-2t^2-1 & \div \quad 4 \\
P_2 = 3t^8-28t^6+30t^4+12t^2+3 & \div \quad 96 \\
P_3 = t^{12}-22t^{10}+113t^8-92t^6-33t^4-6t^2+15 & \div \quad 384 \\
P_4 = 15t^{16}-600t^{14}+7100t^{12}-26616t^{10}+18330t^8+6360t^6+1980t^4 & \\
\qquad\qquad -1800t^2-945 & \div\ 92160 \\
P_5 = 3t^{20}-190t^{18}+4025t^{16}-3397t^{14}+103702t^{12}-63444t^{10}-21270t^8 & \\
\qquad\qquad -7800t^6+4455t^4+1890t^2-17955. & \div 368640
\end{array}\right\} \ldots \quad (2)$$

The simultaneous distribution of two Student variates is exhibited by the product of two expressions of the form (1).

The distribution of the weighted mean is determined from the Statistics observable in the two samples:

$$\left.\begin{array}{l}
\bar{x}_1,\ s_1^2 = \dfrac{1}{n_1(n_1+1)}\, S(x_1-\bar{x}_1)^2, \\[2ex]
\bar{x}_2,\ s_2^2 = \dfrac{1}{n_2(n_2+1)}\, S(x_2-\bar{x}_2)^2,
\end{array}\right\} \qquad \ldots \quad (3)$$

by putting

$$\left.\begin{array}{l}
\bar{x} = (s_2^2\,\bar{x}_1+s_1^2\,\bar{x}_2)/(s_1^2+s_2^2\,) \\[2ex]
\dfrac{1}{S^2} = \dfrac{1}{s_1^2}+\dfrac{1}{s_2^2},
\end{array}\right\} \qquad \ldots \quad (4)$$

and

and obtaining finally

$$\mu_p = \bar{x}+S \sum_0^\infty \sum_0^\infty u_{rs}\, n_1^{-r}\, n_2^{-s}. \qquad \ldots \quad (5)$$

To obtain explicit expressions for $u_{rs}$, for any given probability we set

$$\mu - \bar{x}_1 = t_1 s_1$$

$$\mu - \bar{x}_2 = t_2 s_2, \quad s_1/s_2 = \tan\theta$$

whence

$$\mu - \bar{x} = \frac{s_1 s_2}{s_1^2 + s_2^2}(s_2 t_1 + s_1 t_2) = S(t_1 \cos\theta + t_2 \sin\theta). \qquad \ldots (6)$$

Let

$$\begin{aligned} t_1 \cos\theta + t_2 \sin\theta &= \xi \\ t_1 \sin\theta - t_2 \cos\theta &= d \end{aligned} \qquad \ldots (7)$$

where $d$ is Sukhatme's criterion for the significance of the observed difference between $\bar{x}_1$ and $\bar{x}_2$. Since $d$ is known, the distribution of $\mu$ is required conditional upon a given value of $d$.

## 3. Integration

We should then substitute

$$\left.\begin{aligned} t_1 &= u \cos\theta + d \sin\theta \\ t_2 &= u \sin\theta - d \cos\theta \end{aligned}\right\} \qquad \ldots (8)$$

in Student's polynomials. Then

$$\int_{\xi}^{\infty} du \div \int_{-\infty}^{\infty} du$$

will give the probability of the deviation of the weighted mean (standardized by dividing by $S$) exceeding any chosen value $\xi$, the integrand being the product of two expressions in the form (1) for $n_1$ and $n_2$ degrees of freedom respectively.

From the way the variates have been transformed from $t_1$, $t_2$ to $u$, $d$ the external factor

$$e^{-\frac{1}{2}(t_1^2 + t_2^2)} \qquad \ldots (9)$$

becomes

$$e^{-\frac{1}{2}u^2} \cdot e^{-\frac{1}{2}d^2} \qquad \ldots (10)$$

of which the latter factor is common both to numerator and denominator, and may, therefore, be omitted with the corresponding constant divisor $\sqrt{2\pi}$. The integral of

$$\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} \; P_r(u \cos\theta + d \sin\theta) P_s(u \sin\theta - d \cos\theta) \qquad \ldots (11)$$

being that of the exponential multiplied by a polynomial in $u$, when taken from $-\infty$ to $\infty$ is a function of $d$ and $\theta$ which may be designated by $D_{rs}$, so that the complete divisor is

$$D = \sum_{r=0}^{\infty} \sum_{s=0}^{\infty} D_{rs}\, n_1^{-r}\, n_2^{-s}. \qquad \ldots (12)$$

We may at once note, writing $c$ and $s$ for the cosine and sine of $\theta$,

$$D_{00} = 1$$

$$\left.\begin{aligned} D_{10} &= \{4c^2s^2(d^2-1)+s^4(d^4-2d^2-1)\}\div 4 \\ D_{01} &= \{c^4(d^4-2d^2-1)+4c^2s^2(d^2-1)\}\div 4 \end{aligned}\right\} \qquad \dots \quad (13.1)$$

$D_{20}$, in tabular form

| | $1$ | $d^2$ | $d^4$ | $d^6$ | $d^8$ |
|---|---|---|---|---|---|
| $c^6s^2$ | $-192$ | $192$ | . | . | . |
| $c^4s^4$ | $48(+3$ | $-18$ | $+5$ | . | $.)$ |
| $c^2s^6$ | $8(+3$ | $+27$ | $-45$ | $+7$ | $.)$ |
| $s^8$ | $+3$ | $+12$ | $+30$ | $-28$ | $+3$ |

$\div 96$ … (13.2)

while $D_{11}$ expressed similarly is

| | $1$ | $d^2$ | $d^4$ | $d^6$ | $d^8$ |
|---|---|---|---|---|---|
| $c^6s^2$ | $4(-3$ | $21$ | $-11$ | $1$ | $.)$ |
| $c^4s^4$ | $81$ | $-252$ | $122$ | $-20$ | $1$ |
| $c^2s^6$ | $4(-3$ | $21$ | $-11$ | $1$ | $.)$ |

$\div 16$ … (13.3)

Expressions for $D_{30}$, $D_{21}$, $D_{12}$, $D_{03}$ are easily found, but will only be needed if adjustment beyond the third are to be calculated.

The incomplete integral in the numerator involves terms of the form,

$$\int_{\xi}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} u^m \, du, \qquad \dots \quad (14)$$

which when $m$ is odd is simply

$$z\{\xi^{m-1} + (m-1)\,\xi^{m-3} + \dots + (m-1)(m-3)\dots 2\}, \qquad \dots \quad (15)$$

but when $m$ is even has in addition

$$q\{(m-1)(m-3)\ \dots 3\}, \qquad \dots \quad (16)$$

in which expressions $z$ stands for

$$\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\xi^2}$$

and $q$ for … (17)

$$\int_{\xi}^{\infty} z \, du$$

## 4. Division

We may, therefore, express the incomplete integral of the numerator as

$$z \sum_r \sum_s I_{rs} n_1^{-r} n_2^{-s} + q \sum_r \sum_s D_{rs} n_1^{-r} n_2^{-s}, \quad \dots \quad (18)$$

and if dividing by $$D = 1 + \frac{1}{n_1} D_{10} + \frac{1}{n_2} D_{01} + \dots \quad \dots \quad (19)$$

there remains the quotient $$q + z \sum_r \sum_s q_{rs} n_1^{-r} n_2^{-s}, \quad \dots \quad (20)$$

then the first operation (division) takes the form

$$\left.\begin{aligned} q_{10} &= I_{10}, \quad q_{01} = I_{01} \\ q_{20} &= I_{20} - q_{10} D_{10}, \\ q_{11} &= I_{11} - q_{10} D_{01} - q_{01} D_{10}, \\ q_{30} &= I_{30} - q_{20} D_{10} - q_{10} D_{20}, \\ q_{21} &= I_{21} - q_{11} D_{10} - q_{20} D_{01} - q_{10} D_{11} - q_{01} D_{20}. \end{aligned}\right\} \quad \dots \quad (21)$$

and so on, adjusting the direct integral $I$ at each stage by subtracting all products of the same weight $(r, s)$ of $D$ and $q$.

## 5. Adjustment of deviation

The expression

$$zQ = z \sum_r \sum_s q_{rs} n_1^{-r} n_2^{-s} \quad \dots \quad (22)$$

gives the excess probability beyond the deviate $\xi$ defined in terms of $q$.

If the probability evaluated is exactly

$$q(x) = q(\xi - F)$$

where $F$ is some function of $\xi$, then

$$\xi = x + F$$

and $$zQ = q(\xi - F) - q(\xi)$$

or $$z\left\{F + \frac{1}{2}\xi F^2 + \frac{1}{6}(\xi^2 - 1)F^3 + \frac{1}{24}(\xi^3 - 3\xi)F^4 + \dots\right\} \quad \dots \quad (23)$$

using the expansion in Hermite polynomials. Hence $F$ may be expressed in terms of $\xi$ as

$$F = Q - \frac{1}{2}\xi Q^2 + \frac{1}{6}(2\xi^2+1)Q^3 - \frac{1}{24}(6\xi^3+7\xi)Q^4 + \qquad \dots \quad (24)$$

from which may be derived

$$f_{10} = q_{10}, \quad f_{01} = q_{01},$$

$$f_{20} = q_{20} - \tfrac{1}{2}\xi q_{10}^2$$

$$f_{11} = q_{11} - \xi q_{10} q_{01} \qquad \dots \quad (25)$$

$$f_{30} = q_{30} - \xi q_{10} q_{20} + \tfrac{1}{6}(2\xi^2+1) q_{10}^3$$

$$f_{21} = q_{21} - \xi(q_{10}\, q_{11} + q_{01}\, q_{20}) + \tfrac{1}{2}(2\xi^2+1) q_{10}^2\, q_{01}$$

and

$$x + \sum_r \sum_s f_{rs}\, n_1^{-r}\, n_2^{-s} \qquad \dots \quad (26)$$

is the deviate cutting off exactly the assigned probability

$$\int_x^\infty \frac{1}{\sqrt{2\pi}}\, e^{-\frac{1}{2}u^2}\, du,$$

but in which the polynomials in $F$ are expressed in terms of $\xi$ .

Finally to express the successive adjustments in terms wholly of the normal deviate, known in advance, denoted by $x$, we put $x$ for $\xi$ and calculate

$$U = F + \frac{1}{2}\,\frac{d}{dx}F^2 + \frac{1}{6}\,\frac{d^2}{dx^2}F^3 + \dots \qquad \dots \quad (27)$$

with, in detail

$$u_{10} = f_{10}, \quad u_{01} = f_{01}$$

$$u_{20} = f_{20} + \frac{1}{2}\,\frac{d}{dx} f_{10}^2 \qquad \dots \quad (28)$$

$$u_{11} = f_{11} + \frac{d}{dx}(f_{10} f_{01})$$

and so on.

The three first adjustments calculated in this way are set out in the accompanying Table.

## TABLE OF THE FIRST THREE ADJUSTMENTS

### I

The adjustments are successively of the 4-th, 8-th and 12-th degrees in $\cos\theta$ and $\sin\theta$, and of the 3-rd, 5-th and 7-th degrees in $x$ and $d$

$$u_{10}/n_1 = \{(x^3+x)c^4+4d(x^2+1)c^3s+2(3d^2-1)xc^2s^2+4d(d^2-1)cs^3\} \div 4n_1$$

$$u_{01}/n_2 = \{-4d(d^2-1)c^3s+2(3d^2-1)xc^2s^2-4d(x^2+1)cs^3+(x^3+x)s^4\} \div 4n_2$$

where, as always, $u_{sr}$ may be obtained from $u_{rs}$, by interchanging $c$ and $s$ and reversing the sign of $d$, or by writing $c$ for $s$, $(-s)$ for $c$ as if $(\pi/2+\theta)$ were written for $\theta$.

### II

For higher adjustments it is convenient to separate the odd and even powers of $d$. The terms to be divided by $96n_1^2$ are

| even | | | | | odd | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | $x^5$ | $x^3$ | $x$ | | | $x^4$ | $x^2$ | 1 |
| $c^8$ | $+1$ | 5 | 16 | 3 | | | | | |
| | | | | | $c^7s$ | $+48d$ | 1 | 3 | 2 |
| $c^6s^2$ | $-4$ | 4 | 23 | 51 | | | | | |
| | $+12d^2$ | . | 15 | 31 | $c^5s^3$ | $-96d$ | 1 | 5 | 8 |
| | | | | | | $+32d^3$ | . | 11 | 13 |
| $c^4s^4$ | $+12$ | . | 2 | 9 | | | | | |
| | $-24d^2$ | . | 10 | 39 | $c^3s^5$ | $+96d$ | . | 1 | 3 |
| | $+372d^4$ | . | . | 1 | | $-64d^3$ | . | 5 | 13 |
| | | | | | | $+192d^5$ | . | . | 1 |
| $c^2s^6$ | $+144d^2$ | . | . | 1 | | | | | |
| | $-240d^4$ | . | . | 1 | $cs^7$ | $+96d^3$ | . | . | 1 |
| | | | | | | $-96d^5$ | . | . | 1 |

with conjugate terms divided by $96n_2^2$, while with divisor $16n_1n_2$ they are

| even | | | | | odd | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | $x^5$ | $x^3$ | $x$ | | | $x^4$ | $x^2$ | 1 |
| $c^4s^4$ | $+1$ | 7 | 24 | 69 | | | | | |
| | $-40d^2$ | . | 2 | 7 | $(c^5s^3-c^3s^5)$ | $-8d$ | 3 | 13 | 22 |
| | $+108d^4$ | . | . | 1 | | $+16d^3$ | . | 6 | 13 |
| | | | | | | $-48d^5$ | . | . | 1 |
| $(c^6s^2+c^2s^6)$ | $-2$ | . | 5 | 9 | | | | | |
| | $+6d^2$ | . | 5 | 17 | $(c^7s-cs^7)$ | $+16d$ | . | 1 | 1 |
| | $-48d^4$ | . | . | 1 | | $-16d^3$ | . | 1 | 1 |

# III

The third adjustment has four parts in two complementary pairs for $u_{30}$ and $u_{21}$ respectively divided by $384n_1^3$

| even | | $x^7$ | $x^5$ | $x^3$ | $x$ |
|---|---|---|---|---|---|
| $c^{12}$ | $+1$ | 3 | 19 | 17 | $-15$ |
| $c^{10}s^2$ | $-2$ | 24 | 181 | 632 | 915 |
| | $+6d^2$ | . | 77 | 368 | 475 |
| $c^8s^4$ | $+4$ | 12 | 134 | 649 | 1695 |
| | $-12d^2$ | . | 196 | 1146 | 2486 |
| | $+4d^4$ | . | . | 965 | 2341 |
| $c^6s^6$ | $-8$ | . | 8 | 70 | 225 |
| | $+24d^2$ | . | 56 | 480 | 1441 |
| | $-40d^4$ | . | . | 250 | 839 |
| | $+4488d^6$ | . | . | . | 1 |
| $c^4s^8$ | $-96d^2$ | . | . | 10 | 51 |
| | $+96d^4$ | . | . | 35 | 181 |
| | $-6624d^6$ | . | . | . | 1 |
| $c^2s^{10}$ | $-960d^4$ | . | . | . | 1 |
| | $+1344d^6$ | . | . | . | 1 |

| odd | | $x^6$ | $x^4$ | $x^2$ | 1 |
|---|---|---|---|---|---|
| $c^{11}s$ | $+26d$ | 4 | 23 | 39 | 24 |
| $c^9s^3$ | $-64d$ | 8 | 55 | 165 | 198 |
| | $+192d^3$ | . | 9 | 33 | 28 |
| $c^7s^5$ | $+384d$ | 1 | 10 | 40 | 71 |
| | $-2048d^3$ | . | 3 | 14 | 19 |
| | $+768d^5$ | . | . | 7 | 9 |
| $c^5s^7$ | $-384d$ | . | 1 | 7 | 15 |
| | $+128d^3$ | . | 21 | 146 | 286 |
| | $-10368d^5$ | . | . | 1 | 2 |
| | $+1920d^7$ | . | . | . | 1 |
| $c^3s^9$ | $-256d^3$ | . | . | 5 | 16 |
| | $+384d^5$ | . | . | 7 | 23 |
| | $-2304d^7$ | . | . | . | 1 |
| $cs^{11}$ | $-384d^5$ | . | . | . | 1 |
| | $+384d^7$ | . | . | . | 1 |

Normal deviations for chosen probabilities in a single tail are given below for ready reference.

| P | $x$ | P | $x$ |
|---|---|---|---|
| .5 | 0 | .01 | 2.32635 |
| .25 | .67449 | .005 | 2.57583 |
| .10 | 1.28155 | .0025 | 2.80703 |
| .05 | 1.64485 | .001 | 3.09023 |
| ·025 | 1.95996 | .0005 | 3.29053 |

and by $384n_1^2n_2$

even

| | | $x^7$ | $x^5$ | $x^3$ | $x$ |
|---|---|---|---|---|---|
| $c^{10}s^2$ | $-2$ | . | 77 | 328 | 435 |
| | $+6d^2$ | . | 77 | 568 | 931 |
| | $-96d^4$ | . | . | 15 | 31 |
| $c^8s^4$ | $+1$ | 153 | 1177 | 5099 | 11355 |
| | $-144d^2$ | . | 28 | 183 | 427 |
| | $+24d^4$ | . | . | 515 | 1679 |
| | $-5952d^6$ | . | . | . | 1 |
| $c^6s^6$ | $-4$ | 36 | 361 | 1580 | 4065 |
| | $+180d^2$ | . | 35 | 224 | 607 |
| | $-8d^4$ | . | . | 3040 | 11099 |
| | $+17880d^6$ | . | . | . | 1 |
| $c^4s^8$ | $+12$ | . | 14 | 103 | 225 |
| | $-120d^2$ | . | 14 | 121 | 319 |
| | $+4d^4$ | . | . | 3085 | 11813 |
| | $-11808d^6$ | . | . | . | 1 |
| $c^2s^{10}$ | $+144d^2$ | . | . | 5 | 13 |
| | $-48d^4$ | . | . | 25 | 93 |
| | $+1728d^6$ | . | . | . | 1 |

odd

| | | $x^6$ | $x^4$ | $x^2$ | 1 |
|---|---|---|---|---|---|
| $c^{11}s$ | $+192d$ | . | 1 | 3 | 2 |
| | $-192d^3$ | . | 1 | 3 | 2 |
| $c^9s^3$ | $-16d$ | 28 | 251 | 819 | 1080 |
| | $+384d^3$ | . | 10 | 49 | 57 |
| | $-384d^5$ | . | . | 11 | 13 |
| $c^7s^5$ | $+32d$ | 52 | 377 | 1341 | 2148 |
| | $-512d^3$ | . | 27 | 136 | 209 |
| | $+384d^5$ | . | . | 53 | 101 |
| | $-3840d^7$ | . | . | . | 1 |
| $c^5s^7$ | $-192d$ | 4 | 39 | 145 | 246 |
| | $+128d^3$ | . | 93 | 487 | 818 |
| | $-384d^5$ | . | . | 61 | 137 |
| | $+7680d^7$ | . | . | . | 1 |
| $c^3s^9$ | $+192d$ | . | 3 | 17 | 24 |
| | $-128d^3$ | . | 15 | 103 | 161 |
| | $+6912d^5$ | . | . | 1 | 2 |
| | $-2304d^7$ | . | . | . | 1 |
| $cs^{11}$ | $+384d^3$ | . | . | 1 | 1 |
| | $-384d^5$ | . | . | 1 | 1 |

## 6. Discussion

It appears unmistakably that the quantity $d$, defined by Sukhatme for implementing Behrens' test of the significance of the difference between the two observed means, is a major factor in determining the precision with which the weighted mean, with appropriate adjustments, can be used for the estimation of the common mean, assumed by hypothesis, of the two populations sampled. The analysis implies that the result of applying Behrens' test is not such that the hypothesis of a common mean is abandoned; it does not imply that the difference is not formally significant at some of the standard levels.

The importance of $d$ in the formulae here developed is that the experimenter is led to realize that, if the populations have a common mean, but at the same time there has by chance occurred a somewhat wide discrepancy between the means observed, then the precision with which this common mean can be estimated from the two samples available is much lower than it would have been had the two means come in close agreement. This is really to be expected for two reasons:

(a) If the two means are close together, the weights attached to each in making a common estimate make little difference to the estimate obtained, which will be in this case principally liable only to nearly equal errors in the same direction of the means of two independent samples. Whereas with a large discrepancy the value of the common estimate will be much affected by the relative importance attached to the two pieces of evidence, and this, with small samples, is not well determined.

(b) A large discrepancy between the two means, in relation to the variation within samples, is in itself evidence that the mean squares observed within samples are both too low, and that the precision will be overestimated if the discrepancy is ignored. This information is absent on the hypothesis that the true means have an unknown difference.

A numerical example will perhaps make the position clearer. If we wish to locate the point which the true common mean has a probability of one in forty of exceeding, we should take, on large sample theory the value

$$\bar{x}+(1.95996)S.$$

The coefficient of $S$ is modified by the first three corrective terms as follows in the case $n_1 = 15$, $n_2 = 20$, $\cos^2 \theta = 2/3$, and a range of values of $d$, as shown below.

The successive values for the adjusted coefficient seem to show a satisfactory convergence for the values of $(n_1, n_2)$ chosen. The coefficients after three adjustments seem usually to be correct to two places of decimals and in the central region to three. As in other cases, it is to be expected that convergence will be slower at higher levels of significance, but more rapid for larger samples, with which, however, the influence of $d$ may still be very considerable.

| $d$ | first adjustment | first adjusted coefficient | second adjustment | second adjusted coefficient | third adjustment | third corrected coefficient |
|---|---|---|---|---|---|---|
| 3.0 | .80853 | 2.76849 | .15687 | 2.92536 | .00226 | 2.92762 |
| 2.5 | .62419 | 2.58415 | .08344 | 2.66759 | −.00968 | 2.65791 |
| 2.0 | .45831 | 2.41827 | .03568 | 2.45395 | −.01012 | 2.44383 |
| 1.5 | .31483 | 2.27479 | .00861 | 2.28340 | −.00623 | 2.27717 |
| 1.0 | .19768 | 2.15764 | −.00280 | 2.15484 | −.00234 | 2.15250 |
| 0.5 | .11078 | 2.07074 | −.00435 | 2.06639 | −.00028 | 2.06611 |
| 0.0 | .05806 | 2.01802 | −.00166 | 2.01636 | .00006 | 2.01642 |
| -0.5 | .04345 | 2.00341 | +.00013 | 2.00354 | −.00018 | 2.00336 |
| -1.0 | .07089 | 2.03085 | −.00323 | 2.02762 | .00013 | 2.02775 |
| -1.5 | .14478 | 2.10424 | −.01460 | 2.08964 | .00095 | 2.09059 |
| -2.0 | .26758 | 2.22754 | −.03506 | 2.19248 | −.00007 | 2.19241 |
| -2.5 | .44470 | 2.40466 | −.06345 | 2.34121 | −.00880 | 2.33241 |
| -3.0 | .67957 | 2.63953 | −.09618 | 2.54335 | −.03556 | 2.50779 |

An experimenter who can obtain values of a physical constant by two different methods, the sources of error in which are unrelated, may wish to review his data from two distinct points of view :

(a) That he has full assurance that his theoretical formulation is correct, and that his experimental procedures are both free from systematic error. In this case any apparent discrepancy between the two mean values is ascribed confidently to random sampling errors only, and fiducial limits for the common mean at any chosen level of probability may be obtained by the formulae of the present paper.

(b) That he does not exclude the possibility that his two methods would lead, if indefinitely repeated, to two different average values; that he will admit this possibility without knowing to what such a discrepancy might be due, or being able to compensate or allow for it. In this case, a large value for Sukatme's $d$ is not evidence of lower precision. Indeed, as the discrepancy is not known to be wholly due to errors of random sampling the *weighted* mean has no special merit.

A statistical formulation appropriate to this case is that if $\mu_1$ and $\mu_2$ are the two population means of the two methods, there is one relevant quantity of which estimation is possible, namely

$$(\mu_1+\mu_2)/2,$$

which presents a simpler problem than the weighted mean, for its error curve is symmetrical, and is indeed supplied by Behrens' test of significance. The error of the estimate,

$$(\bar{x}_1+\bar{x}_2)/2$$

being $$(t_1 s_1 + t_2 s_2)/2,$$

which, since the $t$ distribution is symmetrical, is just half the variate tabulated for Behrens' test, when multiplied by $\sqrt{s_1^2 + s_2^2}$.


### Acknowledgement

I am indebted to L. C. Payne for detailed checking, so far as the second adjustment from the unpublished portion of his Ph.D. thesis.


### References


Cornish, E. A. and Fisher, R. A. (1937): Moments and cumulants in the specification of distribution. *Revue de l' Institut International de Statistique*, **5**, 307-352.

——— (1960): The percentile points of distributions having known cumulants. *Technometrics*, **2**, 209-226.

Fisher, R. A. (1926): Applications of 'Student's' distribution. *Metron*, **5**, 90-112.

——— (1935): The fiducial argument in statistical inference. *Ann. Eug.*, **6**, 391-398.

——— (1941): The asymptotic approach to Behren's integral, with further tables for the $d$ test of significance. *Ann. Eug.*, **II**, 141-172.

——— (1961): Sampling the reference set. *Sankhyā.*, Series A, **23**, 1-6.

Fisher, R. A. and Yates, F. (1957): *Statistical Tables for Biological Agricultural and Medical Research.* (Fifth edition)

Payne, L. C. (1957): Unpublished portion of Ph.D. thesis. Cambridge.



*Paper received : December, 1960.*