SOME TWO SAMPLE TESTS

By Douglas G. Chapman[1]

University of Washington

1. Introduction and summary. Stein [4] has exhibited a double sampling procedure to test hypotheses concerning the mean of normal variables with power independent of the unknown variances. This procedure is here adapted to test hypotheses concerning the ratio of means of two normal populations, also with power independent of the unknown variances. The use of a two sample procedure in a regression problem is also considered.

Let $\{X_{ij}\}$ $(i = 1, 2)$ $(j = 1, 2, 3, \cdots)$ be independent random variables distributed according to $N(m_i, \sigma_i)$: all parameters are assumed to be unknown.

Defining k by the equation

$$m_1 = km_2 \tag{1}$$

we wish to test the hypothesis H that k has a specified value k_0.

If $k_0 = 1$ the hypothesis H reduces to a classical problem, often referred to in the literature as the Behrens-Fisher-problem (cf. Scheffé [3] for a bibliography). At the present time it is still an open question whether it is possible (or desirable) to find a non-trivial single sample test for H with the *size* of the critical region independent of σ_1 and σ_2. In any case it is a simple extension of the result of Dantzig [1] (cf. also Stein [4]) to show that no non-trivial single sample test exists whose *power* is independent of σ_1 and σ_2.

On the other hand the case $k_0 \neq 1$ may be expected to occur frequently in fields of application where a choice must be made between different products, methods of experimentation etc. which involve different costs. The statistician must make a choice on the basis of results relative to the ratio of costs involved. Nevertheless this problem appears to have received little attention in the literature.

In general tests based on a two-sample procedure may not be as "efficient" in the sense of Wald [5] as a strict sequential procedure. On the other hand the two sample procedure reduces the number of decisions to be made by the experimenter and it will, in certain fields, simplify the experimental procedure.

2. The two sample procedure. Stein's double sampling procedure (which may be denoted procedure S) to test a hypothesis concerning the mean of a normal population consists briefly in the following steps:

(a) Choose "a priori" a positive number z and a preliminary sample size n.

(b) Take n independent observations $x_1, \cdots, x_n$ of the random variable X

---

[1] This research was carried out while the author was at the University of California, Berkeley, and was supported in part by the Office of Naval Research.

which is assumed to be distributed according to $N(m, \sigma^2)$ with unknown mean m and unknown variance σ^2, and calculate

$$u^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x}).}{n - 1} \tag{2}$$

(c) Let $N = \max\left(\left[\frac{u^2}{z}\right] + 1, n + 1\right)$ where $[r]$ = largest integer $\leq r$

(d) Take $N - n$ more independent observations of X and choose a set of constants $a_1, \cdots a_N$ such that

$$\text{(i) } \sum_{i=1}^{N} a_i = 1, \qquad \text{(ii) } a_1 = a_2 = \cdots = a_n, \qquad \text{(iii) } \sum_{i=1}^{N} a_i^2 = \frac{z}{u^2}. \tag{3}$$

(e) Then $\dfrac{\sum_{i=1}^{N} a_i x_i - m}{\sqrt{z}}$ has Student's t-distribution with $n - 1$ degrees of freedom.

Stein further showed that the procedure may be modified to some advantage in problems dealing with a single population. This modification is not applicable in the problems under consideration here.

There remains to be discussed briefly the choice of n, z and the a's. The preliminary sample size n may be determined by other considerations or it may be chosen as part of the design of the experiment. Hodges [2] has shown that the expected value of the total sample size N and the power of the test both depend on the choice of n and he has discussed the optimum choice of n with respect to the modified procedure of Stein. In general this optimum choice of n depends upon prior knowledge concerning the variance.

The power of the test will depend upon z: some considerations concerning the choice of z will be dealt with after discussing the tables upon which the two sample tests are based.

The arbitrariness involved in choosing the a's may be eliminated by placing the additional requirement that

$$a_{n+1} = a_{n+2} = \cdots = a_N = b \quad \text{(say)}. \tag{4}$$

Letting $a_1 = a_2 = \cdots = a_n = a$ it is elementary to solve for a and b explicitly viz.,

$$\begin{aligned} na + (N - n)b &= 1, \\ na^2 + (N - n)b^2 &= \frac{z}{u^2}. \end{aligned} \tag{5}$$

The solutions are

$$b = \frac{1}{N}\left(1 + \sqrt{\frac{n(Nz - u^2)}{(N - n)u^2}}\right), \tag{6}$$

$$a = \frac{1 - (N - n)b}{n}. \tag{7}$$

3. Test for H. The steps involved in testing the hypothesis H are

(a) Choose the preliminary sample size n, and positive numbers z_1, z_2 subject to the restriction

$$\frac{z_1}{z_2} = k_0^2. \tag{8}$$

(b) Carry out procedure S with the same n for each population, determining two statistics T_1, T_2, i.e.

$$T_i = \frac{\sum_{j=1}^{N_i} a_{ij} x_{ij}}{\sqrt{z_i}} \qquad (i = 1, 2). \tag{9}$$

Then $T_1 - T_2$ has, under the hypothesis tested, the distribution of the difference of two independent Student variables.

If s denotes the difference of two independent random variables t_1 and t_2 each distributed according to Student's t-distribution with $n - 1$ degrees of freedom and if s_0 is defined by the equation

$$P(| s | > s_0) = \alpha,$$

then a test of size α is given by the rule: H is rejected if $| T_1 - T_2 | > s_0$.

4. The distribution of differences of Student variables. The distribution of s is easily found by the method of characteristic functions, in case n is even.

Let $m = n - 1$ and to simplify slightly put

$$y_i = \frac{t_i}{\sqrt{m}} \qquad (i = 1, 2). \tag{10}$$

Then the density function of y_i is

$$f(y) = \frac{\Gamma\left(\frac{m+1}{2}\right)}{\sqrt{\pi}\, \Gamma\left(\frac{m}{2}\right)} \frac{1}{(1 + y^2)^{(m+1)/2}} \tag{11}$$

and its characteristic function

$$\varphi_y(t) = \int_{-\infty}^{+\infty} e^{iyt} f(y)\, dy \tag{12}$$

$$= \frac{\sqrt{\pi}}{\Gamma\left(\frac{m}{2}\right)} \frac{e^{-|t|}}{2^{m-1}} \left(\sum_{r=0}^{(m-1)/2} \frac{\left(\frac{m-1}{2} + r\right)!}{m! \left(\frac{m-1}{2} - r\right)!} [2(| t |)]^{(m-1)/2-r} \right). \tag{13}$$

Formula (13) may be obtained by contour integration; it is, however, a standard formula in connection with Bessel functions of the second kind of purely imaginary argument (cf. Watson [6], pp. 80, 185–188).

While it is not possible to obtain a simple general expression for

$$f(w) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-int} [\varphi_y(t)]^2 \, dt, \tag{14}$$

the density function of $w = \frac{s}{\sqrt{m}}$, this integral may be evaluated for $m = 1, 3, 5$ etc. and furthermore the density function of s may be integrated in a closed form for such values of m, and consequently tabulated fairly easily.

In case n is odd it is possible to express $\varphi_y(t)$ in terms of Bessel functions but the Bessel functions obtained are not expressible in a closed form. While the problem may be attacked directly by numerical integration, it will generally be sufficient to interpolate in Table I where necessary, for such values of n.

Table I gives the distribution of s for $n = 2, 4, 6, 8, 10, 12$. For larger values of n it may be sufficiently accurate to use the normal approximation to the distribution of s. In virtue of the asymptotic normality of the t-distribution s will be distributed approximately normally with mean zero and variance $\frac{2(n-1)}{n-3}$ for n sufficiently large.

5. Power of the test. Writing

$$\Delta = \frac{m_1}{\sqrt{z_1}} - \frac{m_2}{\sqrt{z_2}} \quad \text{and} \quad T = T_1 - T_2 \tag{15}$$

it is seen that $T = s + \Delta$ and hence

$$P(H \text{ is rejected}) = P(|T| > s_0) = P(s < -s_0 - \Delta) + P(s > s_0 - \Delta). \tag{16}$$

Since

$$\Delta = \frac{m_2}{\sqrt{z_2}} \left(\frac{k}{k_0} - 1 \right)$$

equation (16) may be used as a guide in choosing z_2 so that a certain minimum power is attained; the presence of the nuisance parameter m_2 makes impossible the determination of z_2 so as to give exactly some preassigned power.

Since s is distributed independently of σ_1, σ_2, it follows that the power of the test is independent of these parameters. Using the addition formula to express the frequency function of s in terms of the frequency function of Students' t-distribution, it may be shown that $f(s)$ in unimodal and symmetrical about $s = 0$. Hence the test is unbiased. It also follows from (16) that if z_2 is made to approach zero the probability of rejecting H when it is false tends to 1: i.e. the test is consistent.

It may be observed that tests for the one-sided hypotheses

$$\frac{m_1}{m_2} \geq k \quad \text{or} \quad \frac{m_1}{m_2} \leq k$$

may easily be formulated. Table II provides a table useful for such tests also, at half the indicated significance levels.

TABLE I

Distribution of s: difference of two independent student-variables with n − 1 degrees of freedom

The value tabled is $P(0 \leqq s \leqq s_o)$

s_o \ n	2	4	6	8	10	12	Normal Approximation for $n = 12$
0.50	0.0780	0.1014	0.1222	0.1265	0.1290	0.1306	0.1254
1.00	.1476	.1922	.2311	.2392	.2438	.2467	.2388
1.50	.2048	.2660	.3185	.3290	.3349	.3386	.3313
2.00	.2500	.3243	.3825	.3939	.4002	.4041	.3996
2.50	.2852	.3620	.4260	.4364	.4415	.4465	.4451
3.00	.3128	.3903	.4542	.4637	.4687	.4724	.4725
3.50	.3348	.4104	.4726	.4796	.4834	.4856	.4874
4.00	.3524	.4247	.4825	.4884	.4914	.4929	.4947
4.50	.3669	.4352	.4890	.4936	.4956	.4966	.4980
5.00	.3789	.4431	.4930	.4964	.4977		
5.50	.3890	.4491	.4955	.4980	.4988		
6.00	.3976	.4539	.4970	.4988			
6.50	.4050	.4578	.4980				
7.00	.4114	.4611	.4986				
7.50	.4170	.4638					
8.00	.4220	.4661					
10.00	.4372	.4730					
12.00	.4474	.4774					
21.00	.4698	.4870					
30.00	.4788	.4908					
50.00	.4873						
100.00	.4936						

TABLE II

The 5% and 1% significance points of the distribution of s

The value tabled is s_o

Significance Level \ n	2	4	6	8	10	12	Normal Approximation for $n = 12$
$P(\|s\| \geqq s_o) = .05$	25.41	10.82	3.62	3.34	3.18	3.10	3.06
$P(\|s\| \geqq s_o) = .01$	127.3	36.8	5.38	4.72	4.42	4.26	4.03

6. A regression problem. We consider the problem where x_i are values of a sure variable, Y_i are independent random variables with

$$E(Y_i) = a + bx_i \tag{17}$$

and σ_{Y_i} is unknown. It is desired to estimate a and b and to test the hypothesis $b = b_0$.

The usual procedure is to assume $\sigma_{Y_i}^2$ constant, and use the Markov theorem (i.e. the standard least squares formulae). In this way unbiased estimates of a and b are obtained, whether or not this assumption is fulfilled. However the usual significance test for b is not valid if this assumption (plus normality of the Y's) is not fulfilled.

The two sample procedure leads to a valid test of the hypothesis $b = b_0$, with power independent of the unknown variance. Since linearity of the expected value of Y on x is assumed, the optimum procedure is to observe Y for only two values of x, at opposite ends of the range. Let these points be x_1, x_2. For these values of x, procedure S may be used (choosing $z_1 = z_2$) to determine T_1, T_2 where $T_i - (a + bx_i)/\sqrt{z}$ has Student's t-distribution with $n - 1$ degrees of freedom.

Then the following estimates of a, b are unbiased, for $n \geq 3$,

$$\hat{b} = \left(\frac{T_2 - T_1}{x_2 - x_1}\right) \sqrt{z}, \tag{18}$$

$$\hat{a} = \left(\frac{x_2 T_1 - x_1 T_2}{x_2 - x_1}\right) \sqrt{z}. \tag{19}$$

To test the hypothesis $H_1 : b = b_0$ it is necessary only to calculate the statistic $\zeta = [(T_1 - T_2)\sqrt{z} - b_0(x_1 - x_2)]/\sqrt{z}$ and reject H_1, at the α level of significance if $|\zeta| > s_0$, where s_0 was defined above (Section 3).

It is seen that if b' is the true value of b, then the power of the test is a function of $(b' - b_0)(x_1 - x_2)/\sqrt{z}$ and z maybe determined to obtain any prescribed power desired. It is also immediate that the power of the test is independent of σ_{Y_i}.

The author wishes to express thanks to the members of the computing staff of the Statistical Laboratory, University of California, Mrs. E. Putz, Miss J. Linton, and Mr. J. Blum, for assistance in preparing Tables I and II.[2]

REFERENCES

[1] GEORGE B. DANTZIG, "On the non-existence of tests of 'Student's' hypothesis having power functions independent of σ," *Annals of Math. Stat.*, Vol. 11 (1940), p. 186.

[2] JOSEPH L. HODGES, JR., "The selection of initial sample size in the Stein two sample procedure", unpublished dissertation, University of California, Berkeley, 1948.

[3] HENRY SCHEFFÉ, "On solutions of the Behrens-Fisher Problem based on the t-distribution", *Annals of Math. Stat.*, Vol. 14 (1943), p. 35.

[4] CHARLES STEIN, "A two sample test for a linear hypothesis whose power is independent of the variance", *Annals of Math. Stat.*, Vol. 16 (1945), p. 243.

[5] ABRAHAM WALD, *Sequential Analysis*, John Wiley and Sons, Inc., 1947.

[6] G. N. WATSON, *A Treatise on the Theory of Bessel Functions*, Cambridge University Press, 1944.

[2] It has been pointed out to the writer that percent points of linear combinations of two independent Student t's are given in Table VI (by P. V. Sukatme) in R. A. FISHER AND F. YATES, *Statistical Tables for Biological, Medical and Agricultural Research*, Oliver and Boyd, Edinburgh, 1943 (added in page proof).