ON COMPARING DIFFERENT TESTS OF THE SAME HYPOTHESIS

by

Carmen Adela Pérez

Thesis submitted to the Graduate Faculty of the

Virginia Polytechnic Institute

in candidacy for the degree of

MASTER OF SCIENCE

in

Statistics

October 1959

Blacksburg, Virginia

TABLE OF CONTENTS

I. INTRODUCTION AND SUMMARY

The usual procedure for comparing tests of the same hypothesis is based on their power functions. This is done by plotting the power curves of each test for a fixed Type I error. The power curve of the best test (if such a test exists) will lie above the curve of any other test with the same level of significance.

Suppose we have the standard test of a null hypothesis H_o based on the statistic u_1. Sometimes it is convenient to use a quicker, although less powerful, test based on a statistic u_2. We want to determine the degree of agreement between these two tests. The method of power curves, however, only gives an indication of the long-run behavior of each test and not of the extent to which their results actually agree when applied to the same problem.

In this thesis two methods for comparing standard and quick tests are proposed. The first method consists of determining the probability of establishing significance with u_2 at a level β, given that u_1 is just significant at level α. The second approach determines how significant the expected value of u_2 is, given that u_1 is just significant at level α.

We apply these two alternative procedures to compare:

(a) tests of location in a sample from a normal population with known variance;

(b) tests of dispersion, namely those based on the standard deviation, mean deviation, and range in a normal sample with unknown mean;

(c) the sign test and the paired t-test, again in a sample from a normal population.

The subject of this thesis and some further issues have also been treated in a slightly more advanced fashion in a Report by David and Perez (1959).

II. TESTS OF LOCATION

Let us consider a random sample $x_1, x_2, \ldots, x_n$ taken from a normal population with variance equal to 1. It is desired to test the null hypothesis H_0: $\mu = 0$ against the alternative H_1: $\mu > 0$. The standard procedure in this case is to use the statistic $u_1 = \bar{x}\sqrt{n}$, which is a unit normal variable, and for an α level of significance reject H_0 when $u_1 > u_\alpha$. This is a uniformly most powerful test of H_0.

Restricting ourselves to the first k observations only, a quicker test would consist of rejecting H_0 when $u_2 > u_\alpha$, where $u_2 = \bar{x}_k\sqrt{k}$. We shall now apply two methods to compare the statistics u_1 and u_2 of the efficient and quick tests respectively.

(a) To determine the probability that u_2 exceeds its β significance point given that u_1 is just significant at level α. That is, to find

$$P = Pr(u_2 > u_\beta \mid u_1 = u_\alpha).$$

From the normal density function of two variables x and y, with means μ_x, μ_y, variances σ_x^2, σ_y^2, and correlation coefficient ρ, the conditional density $f(x|y)$ is found to be a univariate normal with mean $\mu_x + \rho\frac{\sigma_x}{\sigma_y}(y - \mu_y)$ and variance $\sigma_x^2(1 - \rho^2)$. Thus, under H_0, $f(u_2|u_1{=}u_\alpha)$ is $N(\rho u_\alpha, 1-\rho^2)$ and therefore we have

$$P = Pr[u > \frac{u_\beta - \rho\, u_\alpha}{\sqrt{(1-\rho^2)}}] \qquad (1)$$

The correlation coefficient ρ is, by definition,

$$\rho(u_1,u_2) = \frac{\operatorname{cov}(u_1,u_2)}{\sqrt{\operatorname{var} u_1 \operatorname{var} u_2}} = \operatorname{cov}(u_1,u_2),$$

and

$$\operatorname{cov}(u_1,u_2) = \sqrt{nk}\operatorname{cov}(\bar{x}, \bar{x}_k)$$

$$= \frac{1}{\sqrt{nk}} \operatorname{cov}\left[\sum_{i=1}^{n} x_i, \sum_{i=1}^{k} x_i\right]$$

$$= \frac{1}{\sqrt{nk}} \sum_{i,j=1}^{k} \operatorname{cov}(x_i,x_j) ;$$

but $\operatorname{cov}(x_i, x_j) = \operatorname{var}(x_i) = 1$, $(i=j)$, and since the observations are independent, $\operatorname{cov}(x_i,x_j) = 0$, $(i \neq j)$. Therefore,

$$\rho = \operatorname{cov}(u_1, u_2) = \frac{k}{\sqrt{nk}} = \sqrt{\frac{k}{n}} .$$

From equation (1), it can be seen that $P \leq \frac{1}{2}$ when $\alpha = \beta$. Hence it seems appropriate to choose $\beta > \alpha$ $(u_\beta < u_\alpha)$.

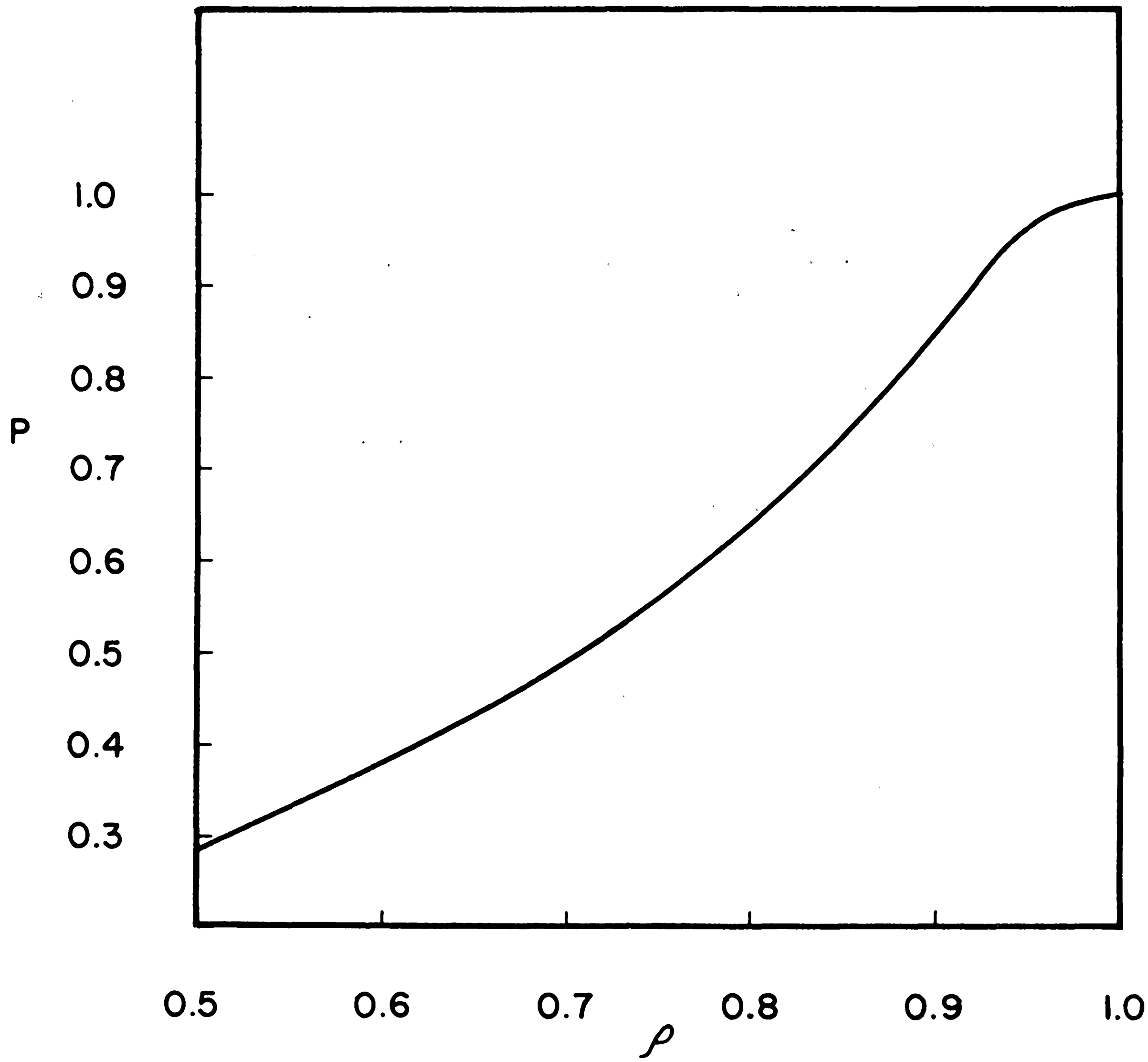
In Table 1 we have the values of P obtained for $\alpha = 0.01$, $\beta = 0.05$ and for values of $\rho \geq 0.5$.

Table 1

ρ	0.5	0.6	0.7	0.8	0.9	0.95	0.98	1
P	.289	.378	.491	.641	.848	.965	.999	1

The approach of P to unity is better illustrated in Fig. 1, where P is plotted as a function of ρ.

Fig. 1. $P = Pr(u_2 > 1.645 | u_1 = 2.326)$ as a function of ρ, the correlation coefficient of the unit normal variates u_1, u_2.



(b) The second approach, which is generally simpler, is to determine the significance level corresponding to the expected value of u_2 given that u_1 is just significant at a level α. This conditional expectation is given by

$$\mathcal{E}\ (u_2|u_1 = u_\alpha) = \rho u_\alpha .$$

The corresponding level of significance or "equivalent Type I error" of the quick test is then

$$\gamma = \Pr\ (u_2 > \rho u_\alpha).$$

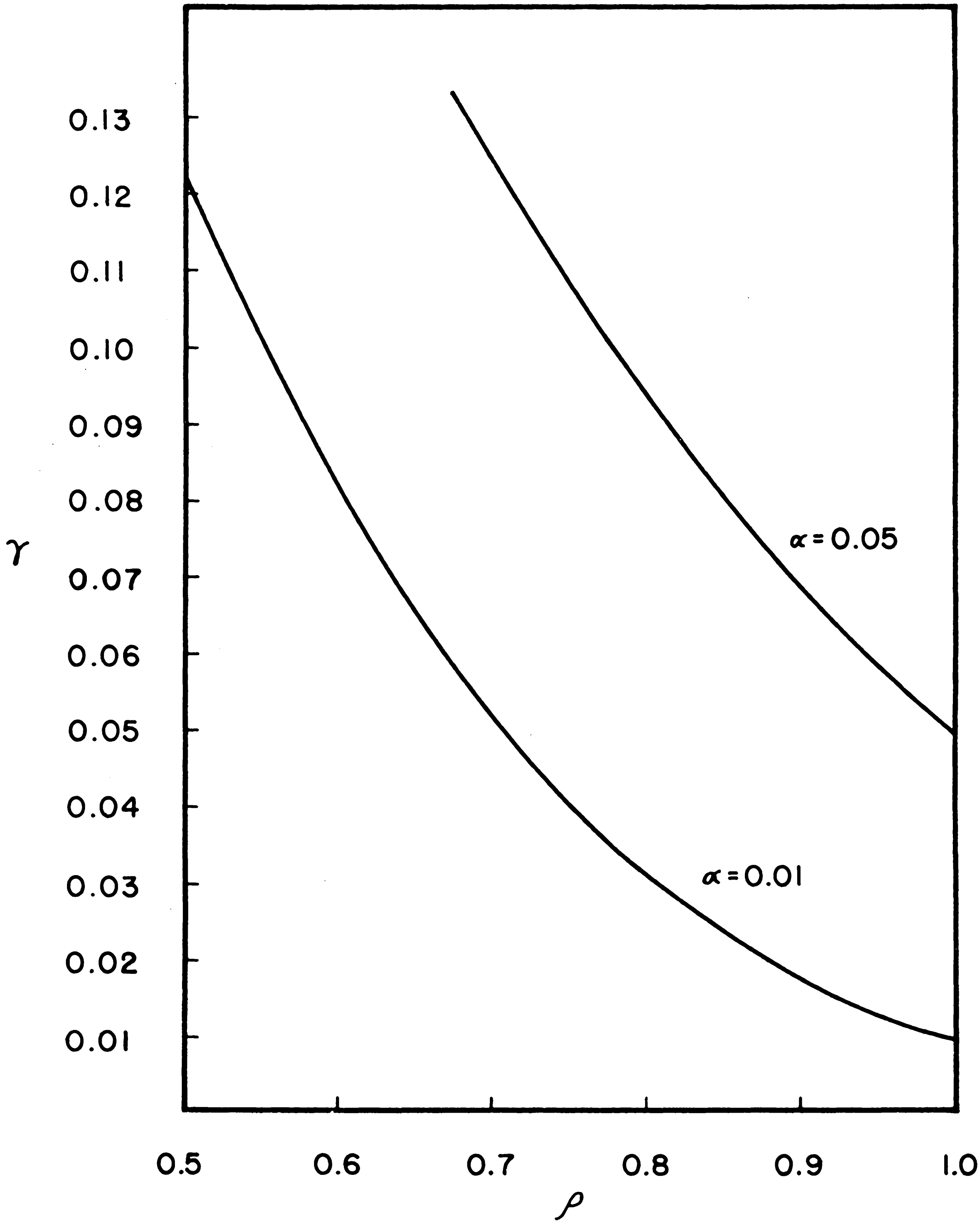
Table 2 gives the values of γ for $\rho \geq 0.5$ and $\alpha = 0.01,\ 0.05$.

Table 2

ρ	α=0.01 γ	α=0.05 γ
0.5	.1224	.2054
0.6	.0814	.1618
0.7	.0517	.1248
0.8	.0314	.0941
0.9	.0181	.0694
0.95	.0136	.0591
0.98	.0113	.0535
1.00	.0100	.0500

In Fig. 2 γ is plotted as a function of ρ for both values of α.

Fig. 2. $\gamma = \Pr(u_2 > \rho u_\alpha)$



2.1 Generalizations

The above procedure can be applied to the more general case of comparing $\bar{x}$ with any linear unbiased estimator of μ (e.g. median, mid-range) which we will term m.

We will first determine P which can be expressed as

$$P = Pr(m > m_\beta | \bar{x} = \bar{x}_\alpha) \qquad (\bar{x}_\alpha = u_\alpha/\sqrt{n})$$
$$= Pr(m - \bar{x} > m_\beta - \bar{x} | \bar{x} = \bar{x}_\alpha)$$

Now $m - \bar{x}$ can be shown to be independent of $\bar{x}$. We have

$$f(x_1, x_2, \ldots, x_n) = \frac{1}{(2\pi)^{n/2}} e^{-\frac{1}{2}\Sigma(x_i-\mu)^2}$$

$$= \frac{1}{(2\pi)^{n/2}} e^{-\frac{1}{2}\Sigma(x_i-\bar{x})^2} e^{-\frac{1}{2}n(\bar{x}-\mu)^2} \tag{1}$$

Applying Helmert's orthogonal transformation to the x's we get

$$x_1 = \frac{y_1}{\sqrt{n}} + \frac{y_2}{\sqrt{1.2}} + \frac{y_3}{\sqrt{2.3}} + \ldots + \frac{y_n}{\sqrt{(n-1)n}}$$
$$x_2 = \frac{y_1}{\sqrt{n}} - \frac{y_2}{\sqrt{1.2}} + \frac{y_3}{\sqrt{2.3}} + \ldots + \frac{y_n}{\sqrt{(n-1)n}}$$
$$x_3 = \frac{y_1}{\sqrt{n}} \qquad - \frac{2y_3}{\sqrt{2.3}} + \ldots + \frac{y_n}{\sqrt{(n-1)n}}$$
$$\vdots$$
$$x_n = \frac{y_1}{\sqrt{n}} \qquad\qquad - \frac{(n-1)y_n}{\sqrt{(n-1)n}}$$

where $y_1 = \sqrt{n}\,\bar{x}$ and $\Sigma(x_i - \bar{x})^2$ is formed by

$$x_1 - \frac{y_1}{\sqrt{n}} = \frac{y_2}{\sqrt{1.2}} + \frac{y_3}{\sqrt{2.3}} + \cdots + \frac{y_n}{\sqrt{(n-1)n}}$$

$$x_2 - \frac{y_1}{\sqrt{n}} = -\frac{y_2}{\sqrt{1.2}} + \frac{y_3}{\sqrt{2.3}} + \cdots + \frac{y_n}{\sqrt{(n-1)n}}$$

$$x_3 - \frac{y_1}{\sqrt{n}} = \qquad -\frac{2y_3}{\sqrt{2.3}} + \cdots + \frac{y_n}{\sqrt{(n-1)n}}$$

$$\vdots$$

$$x_n - \frac{y_1}{\sqrt{n}} = \qquad\qquad -\frac{(n-1)y_n}{\sqrt{(n-1)n}}$$

Therefore $\Sigma(x_i - \bar{x})^2 = y_2^2 + y_3^2 + \ldots\ y_n^2$ is a function of the $y_j (j = 2, \ldots, n)$. Now (1) can be written as

$$f(y_1, y_2, \ldots, y_n) = \frac{1}{(2\pi)^{n/2}}\, e^{-\frac{1}{2}\sum_{2}^{n} y_j^2}\, e^{-\frac{1}{2}n(y_1 - \mu\sqrt{n})^2}$$

$$= f(y_2, y_3, \ldots, y_n) f(y_1).$$

Therefore y_1 is independent of the joint distribution of $y_2, y_3, \ldots, y_n$ and hence of any function of $y_2, \ldots, y_n$. Since $m - \bar{x}$ can be expressed as

$$m - \bar{x} = \frac{1}{n}(m - x_1 + m - x_2 + \ldots + m - x_n)$$

which is a linear function of the differences of the observations and hence of the y_j, it follows that $m - \bar{x}$ is also independent of $y_1 = \bar{x}\sqrt{n}$. Hence

$$P = Pr(m - \bar{x} > m_\beta - \bar{x}_\alpha).$$

The second approach is usually easier to apply. Thus, the conditional expectation of m given that $\bar{x}$ is just significant at level α is given by

$$\begin{aligned}\mathcal{E}(m|\bar{x} = \bar{x}_\alpha) &= \mathcal{E}(m-\bar{x}+\bar{x}|\bar{x} = \bar{x}_\alpha) \\ &= \mathcal{E}(m-\bar{x}) + \bar{x}_\alpha \\ &= \bar{x}_\alpha .\end{aligned}$$

Therefore the "equivalent Type I error" of m is

$$\gamma = \Pr(m > \bar{x}_\alpha).$$

It should be noted that since we are considering conditional probabilities and expectations given the statistic $\bar{x}$ which is sufficient for μ, P and γ do not depend on μ, that is, hold on H_1 as well as on H_0.

III. TESTS OF DISPERSION

Again, let us consider a random sample $x_1, x_2, \ldots, x_n$ from a normal population with mean unknown. We wish to test the composite null hypothesis H_o: $\sigma = 1$ against the alternative H_1: $\sigma > 1$.

We will consider three estimators of σ, namely,

i) The standard deviation $s = \sqrt{\dfrac{\Sigma(x_i - \bar{x})^2}{n-1}}$

ii) The range $w = x_{max} - x_{min}$

iii) The mean deviation $d = \frac{1}{n}\Sigma|x_i - \bar{x}|$,

and apply the two procedures to compare the relative merits of s and w and of s and d. According to H_1, the right tails of the distributions of these estimators will be used as critical regions, i.e. we will reject H_o when $s > s_\alpha$, $w > w_\alpha$ or $d > d_\alpha$ for an α level of significance.

Since $\dfrac{(n-1)s^2}{\sigma^2}$ is distributed as a χ^2 with n-1 degrees of freedom, the percentage points of s^2, and hence of s, can be obtained from those of the χ^2 distribution. The percentage points of w and d are given in Biometrika Tables (tables 22,21).

The probability of obtaining significance with w at level β, given that s is just significant at level α, is

$$\begin{aligned} P_w &= Pr(w > w_\beta \mid s = s_\alpha) \\ &= Pr(\tfrac{w}{s} > \tfrac{w_\beta}{s} \mid s = s_\alpha) \\ &= Pr(\tfrac{w}{s} > \tfrac{w_\beta}{s_\alpha}) \qquad (2) \end{aligned}$$

This last step is justified by the independence of $\frac{w}{s}$ and s which can be proved as follows:

By applying Helmert's orthogonal transformation to the x's of our sample we showed that

$$y_1 = \bar{x}\sqrt{n}$$

and

$$\Sigma(x_i - \bar{x})^2 = \sum_{j=2}^{n} y_j^2$$

y_1 being independent of $y_2, y_3, \ldots, y_n$.

Since $\quad \Sigma(x_i - \bar{x})^2 = (n-1)s^2$

hence $\quad \Sigma y_j^2 = (n-1)s^2$

and $x_i - \bar{x}$ is a function of the y_j.

The range w can be written as

$$w = (x_{max} - \bar{x}) - (x_{min} - \bar{x})$$

$$= \max(x_i - \bar{x}) - \min(x_i - \bar{x})$$

which is also a function of the y_j and hence independent of $y_1 = \bar{x}\sqrt{n}$.

We now make the polar transformation

$$\begin{aligned}
y_2 &= R \cos\theta_1 \cos\theta_2 \cdots\cdots\cdots \cos\theta_{n-2} \\
y_3 &= R \cos\theta_1 \cos\theta_2 \cdots \cos\theta_{n-3} \sin\theta_{n-2} \\
y_4 &= R \cos\theta_1 \cos\theta_2 \cdots \sin\theta_{n-3} \\
&\vdots \\
y_n &= R \sin\theta_1.
\end{aligned}$$

Therefore

$$\sum_{i=2}^{n} y_i^2 = (n-1)s^2 = R^2$$

and

$$s = \frac{R}{\sqrt{(n-1)}}$$

Since

$$\frac{w}{s} = \frac{f(y_2,\ldots,y_n)}{s} = \frac{Rf_1(\theta_1,\theta_2,\ldots,\theta_{n-2})}{\frac{R}{\sqrt{(n-1)}}}$$

$$= \sqrt{(n-1)}\ f(\theta_1,\theta_2,\ldots,\theta_{n-2})$$

then $\frac{w}{s}$ is independent of R and of s.

In order to evaluate (2) it is necessary to have some knowledge of the distribution of $\frac{w}{s}$. This distribution can be approximated by a Pearson type frequency curve if its first four moments are set equal to the first four moments of a particular Pearson-type curve.

To determine what type of Pearson distribution is appropriate we use the chart given as Table 43 (Biometrika Tables). In it, the "moment coefficients", $\beta_1 = \frac{\mu_3^2}{\mu_2^3}$, $\beta_2 = \frac{\mu_4}{\mu_2^2}$, are plotted on rectangular axes and the seven main types of Pearson curves can be associated with points, lines, or areas in that plane. Hence, our problem is to find $\beta_1(\frac{w}{s})$ and $\beta_2(\frac{w}{s})$.

The moments of $\frac{w}{s}$ about the origin can be obtained from the relation

$$\mu'_r(\frac{w}{s}) = \mathcal{E}(\frac{w}{s})^r = \mathcal{E}(\frac{w^r}{s^r}) = \frac{\mathcal{E}(w^r)}{\mathcal{E}(s^r)} = \frac{\mu'_r(w)}{\mu'_r(s)} \; . \qquad (3)$$

This can be justified by noting that

$$\mathcal{E}(w^r) = \mathcal{E}(\frac{w^r}{s^r} \cdot s^r) = \mathcal{E}(\frac{w^r}{s^r}) \cdot \mathcal{E}(s^r)$$

which follows from the independence of $\frac{w}{s}$ and s.

The moments $\mu'_r(s)$ are calculated directly from the frequency distribution of s, which can be derived from the χ^2 function. The moments $\mu'_r(w)$ are determined from the moments about the mean $\mu_r(w)$ by applying familiar relationships. After (3) has been evaluated the reverse relationships are used to obtain $\mu_r(\frac{w}{s})$ from $\mu'_r(\frac{w}{s})$.

For n = 10, $\beta_1(\frac{w}{s})$ and $\beta_2(\frac{w}{s})$ were found to be .013 and 2.62 respectively, determining a point which falls in the Pearson Type I area of the chart mentioned above.(*) Since the probability integral of this curve can be obtained from the probability integral of the incomplete B - function by making a simple change of variables, it is possible to evaluate P_w.

When n = 10, $\alpha = 0.01$, $\beta = 0.05$, we have

$$P_w = Pr(\frac{w}{s} > 2.881) = .809.$$

(*) The calculation of the constants of the Pearson type I distribution was made from formulas given by W.P. Elderton (1938).

The distribution of $\frac{w}{s}$ has been shown to be of a bounded nature (Thomson) and for n=3, $\frac{w_\beta}{s_\alpha}$ is less than the lower bound of $\frac{w}{s}$ which is equal to $\sqrt{3}$. Therefore in this case P_w is unity.

P_d, the probability of establishing significance with the mean deviation at a level β when s is just significant at level α, can be determined by the same approach as P_w. Thus, for $n = 10$, $\alpha = 0.01$, $\beta = 0.05$

$$P_d = \Pr\left(\frac{d}{s} > \frac{d_\beta}{s_\alpha}\right)$$

$$= \Pr\left(\frac{d}{s} > 0.6997\right)$$

$$= 0.905.$$

As in the case of $\frac{w}{s}$, the distribution of $\frac{d}{s}$ is approximated by a Pearson Type I curve, and P_d is evaluated from the probability integral of the Incomplete B-function.

For n = 3, the bounds of the distribution $\frac{d}{s}$ can be found by maximizing - or minimizing - the ratio

$$\frac{d}{s} = \frac{\frac{1}{3}\sum_{i=1}^{3} | x_i - \bar{x} |}{\sqrt{\frac{\sum_{i=1}^{3} (x_i - \bar{x})^2}{2}}} \quad .$$

Since the maximum (or minimum) does not depend on the origin and scale of the x_i, we may assume $x_1 = 0$ and $x_3 = 1$. We may also assume, without loss of generality, that $x_2 \geq \frac{1}{2}$ which enables us to write d as follows:

$$d = \frac{1}{3}\left\{\left|0 - \frac{x_2+1}{3}\right| + \left|x_2 - \frac{x_2+1}{3}\right| + \left|1 - \frac{x_2+1}{3}\right|\right\}$$

$$= \frac{2}{9}(x_2 + 1).$$

Hence

$$\frac{d}{s} = \frac{2\sqrt{3}\,(x_2 + 1)}{9\sqrt{x_2^2 - x_2 + 1}} .$$

Taking the derivative of $\frac{d}{s}$ with respect to x_2 and setting it equal to zero, a maximum is found for $x_2 = 1$. The upper bound of $\frac{d}{s}$ is then $\frac{4}{3\sqrt{3}}$. Since the function is increasing, a minimum is obtained when $x_2 = \frac{1}{2}$, the lower bound being then equal to $\frac{2}{3}$. In this case $\frac{d_\beta}{s_\alpha} = 0.595$ and therefore P_d is also 1.

The "equivalent Type I error" of the range, γ_w, can be determined by noting that

$$\begin{aligned}\mathcal{E}(w \mid s = s_\alpha) &= \mathcal{E}\left(\frac{w}{s} \cdot s \mid s = s_\alpha\right)\\ &= s_\alpha \mathcal{E}\left(\frac{w}{s} \mid s = s_\alpha\right)\\ &= s_\alpha \mathcal{E}\left(\frac{w}{s}\right)\\ &= s_\alpha \frac{\mathcal{E}(w)}{\mathcal{E}(s)} ,\end{aligned}$$

the third and fourth steps being justified by the independence of $\frac{w}{s}$ and s.

Hence

$$\gamma_w = \Pr\left[w > s_\alpha \frac{\mathcal{E}(w)}{\mathcal{E}(s)}\right].$$

Similarly, we obtain

$$\gamma_d = \Pr[d > s_\alpha \frac{\mathcal{E}(d)}{\mathcal{E}(s)}].$$

The probability integral of the range is given in Table 23 (Biometrika Tables) and tables of the probability integral of the mean deviation are given by Godwin and Hartley (Biometrika, 1945). The values of $\mathcal{E}(w)$, $\mathcal{E}(d)$ are taken from Table 20 and $\mathcal{E}(s)$ from Table 35 of Biometrika Tables.

Table 3 below gives γ_w and γ_d for $\alpha = 0.05, 0.01$ and $n = 3(1)10$. The values of γ_d for $n = \infty$ were obtained from the formula $\gamma = \Pr(u > \rho u_\alpha)$ remembering that the mean deviation tends to normality as n becomes larger. The correlation between s and d is

$$\rho = \sqrt{E_d} , \tag{4}$$

where E_d is the relative efficiency of the two estimators.

Formula (4) can be proved as follows: (*)

$$\rho(s,d) = \frac{\mathcal{E}(sd) - \mathcal{E}(s)\,\mathcal{E}(d)}{\sigma(s)\,\sigma(d)} .$$

Now, since $\frac{d}{s}$ is independent of s,

$$\mathcal{E}(sd) = \mathcal{E}(s^2 \cdot \frac{d}{s}) = \mathcal{E}(s^2)\cdot \mathcal{E}(\frac{d}{s})$$

$$= \mathcal{E}(s^2)\cdot \frac{\mathcal{E}(d)}{\mathcal{E}(s)} .$$

(*) Compare H. O. Hartley (1955)

Therefore

$$\rho\ (s,d) = \frac{\mathcal{E}(s^2)\,\frac{\mathcal{E}(d)}{\mathcal{E}(s)} - \mathcal{E}(s)\,\mathcal{E}(d)}{\sigma(s)\ \sigma(d)}$$

$$= \frac{\frac{\mathcal{E}(d)}{\mathcal{E}(s)}\left\{\mathcal{E}(s^2) - [\mathcal{E}(s)]^2\right\}}{\sigma(s)\ \sigma(d)}$$

$$= \frac{\mathcal{E}(d)}{\mathcal{E}(s)} \cdot \frac{\sigma^2(s)}{\sigma(s)\sigma(d)}$$

$$= \frac{\mathcal{E}(d)}{\mathcal{E}(s)} \cdot \frac{\sigma(s)}{\sigma(d)}$$

$$= \sqrt{E_d}\,.$$

Table 3. Values of the "equivalent Type I errors" γ_w, γ_d, of tests based on the range w and the mean deviation d corresponding to Type I error α in standard (χ^2) test

n \ α	0.05 γ_w	0.01 γ_w	0.05 γ_d	0.01 γ_d
3	0.0508	0.0105	0.0508	0.0105
4	.0526	.0114	.0536	.0122
5	.0548	.0125	.0557	.0130
6	.0571	.0136	.0569	.0136
7	.0594	.0149	.0578	.0139
8	.0617	.0161	.0585	.0141
9	.0641	.0174	.0589	.0143
10	.0664	.0186	.0593	.0145
∞	-	-	.0618	.0147

IV. THE SIGN TEST AND THE t-TEST

If we have a random sample of n observations from a normal population with σ^2 unknown and want to test H_0: $\mu=0$ against H_1: $\mu > 0$, the use of the statistic

$$t = \frac{\bar{x}\sqrt{n}}{s}$$

provides a uniformly most powerful test for that null hypothesis. The critical region in this case is

$$t > t_\alpha$$

where t_α is the upper α significance point of a t-distribution with n-1 degrees of freedom.

The sign test is useful when the assumptions required for the t-test cannot be made (e.g. assumption of normality), and because of its simplicity is sometimes preferred even when the t-test is applicable. To apply the sign test to the above situation, we count the number, r', of positive observations. If H_0 is true, r' is a binomial variate with $p = \frac{1}{2}$. We now apply our first method to a comparison of the sign test and the t-test.

The probability that $x_i > 0$, given that the t-test is just significant at level α, is

$$\pi_c = Pr(x_i > 0 \mid t = t_\alpha) \qquad (5)$$

$$= Pr(\frac{x_i}{s} > 0 \mid \frac{\bar{x}\sqrt{n}}{s} = t_\alpha)$$

$$= \Pr\left(\frac{x_i - \bar{x}}{s} + \frac{\bar{x}}{s} > 0 \;\middle|\; \frac{\bar{x}\sqrt{n}}{s} = t_\alpha\right).$$

Since $\frac{x_i - \bar{x}}{s}$ is independent of $\bar{x}$ and s and therefore independent of $\underline{t}$ (a function of $\bar{x}$, s) it follows that

$$\pi_c = \Pr\left(\frac{x_i - \bar{x}}{s} + \frac{t_\alpha}{\sqrt{n}} > 0\right)$$

$$= \Pr\left(\frac{x_i - \bar{x}}{s} > - \frac{t_\alpha}{\sqrt{n}}\right).$$

However, the probability of $\underline{r}'$ "successes" out of n trials is difficult to evaluate because the n trials are not independent. But the conditional probability of n straight successes is

$$P' = \Pr\left(\frac{x_{min} - \bar{x}}{s} > - \frac{t_\alpha}{\sqrt{n}}\right) = \Pr\left(\frac{x_{max} - \bar{x}}{s} < \frac{t_\alpha}{\sqrt{n}}\right).$$

P' can be evaluated by integrating the probability density function of $\frac{x_{max} - \bar{x}}{s}$ which has been determined by Borenius (1958).

The function $\frac{x_{max} - \bar{x}}{s} = r$ is bounded and for sample sizes 4, 5, 6 takes the following forms:

$n = 4$

$$f(r) = \frac{2}{3} \qquad 1 \leq r \leq \sqrt{3}$$

$$= \frac{2}{\sqrt{3}} \left[1 - \frac{3}{\pi} \tan^{-1} \frac{\sqrt{3(1-r^2)}}{r\sqrt{2}}\right] \qquad \frac{1}{\sqrt{3}} \leq r \leq 1$$

$n = 5$

$$f(r) = \frac{5}{2\pi}\left\{2r\sqrt{\frac{5}{3}} - \sqrt{4-r^2}\right\} \qquad \sqrt{\frac{2}{3}} \leq r \leq \sqrt{\frac{3}{2}}$$

$$= \frac{5}{2\pi}\sqrt{4-r^2} \qquad \sqrt{\frac{3}{2}} \leq r \leq 2$$

$n = 6$

$$f(r) = \frac{9}{10\sqrt{5}}(5-r^2) \qquad \sqrt{2} \leq r \leq \sqrt{5}$$

$$= \frac{9}{10\sqrt{5}}(5-r^2) + \frac{9\sqrt{3}}{4\pi} r\sqrt{2-r^2}$$

$$- \frac{9}{2\pi\sqrt{5}}(5-r^2)\tan^{-1}\frac{\sqrt{5(2-r^2)}}{r\sqrt{3}} \qquad 1 \leq r \leq \sqrt{2}$$

$$= \frac{9}{10\sqrt{5}}(5-r^2) + \frac{9\sqrt{3}}{2\pi}\left\{r^2 - \frac{r}{2}\sqrt{2-r^2}\right\}$$

$$+ \frac{9}{2\pi\sqrt{5}}(5-r^2)\left\{\tan^{-1}\frac{\sqrt{5(2-r^2)}}{r\sqrt{3}} - 2\tan^{-1}\sqrt{\frac{5}{3}}\right\}$$

$$\frac{1}{\sqrt{2}} \leq r \leq 1$$

Most of these integrals are readily evaluated by standard procedures.

The integral of $\tan^{-1}\frac{\sqrt{5(2-r^2)}}{r\sqrt{3}}$ can be evaluated by letting $\frac{\sqrt{5(2-r^2)}}{r\sqrt{3}} = t$ and integrating by parts which reduces it to the form

$$k\int\frac{dt}{(1+t^2)\sqrt{5+3t^2}} \quad .$$

This can then be integrated by performing a second transformation, viz. $1 + t^2 = \frac{1}{w}$ to give

$$- \frac{1}{2\sqrt{2}} \sin^{-1} \left(\frac{12w + 1}{5}\right).$$

The final result is then

$$\sqrt{10} \left[\frac{\tan^{-1} t}{\sqrt{5+3t^2}} + \frac{1}{2\sqrt{2}} \sin^{-1} \left(\frac{t^2 + 5}{5(t^2+1)}\right)\right].$$

By the same procedure

$$\int r^2 \tan^{-1} \frac{\sqrt{5(2-r^2)}}{r\sqrt{3}} \, dr$$

can be evaluated, but in this case a third transformation is needed. Thus, after integration by parts we are left with an integral of the form

$$k \int \frac{dt}{(1+t^2)(5+3t^2)^{3/2}}$$

which, by letting $1 + t^2 = \frac{1}{w}$, is transformed into

$$k' \int \frac{w \, dw}{\sqrt{1-w} \, (2w+3)^{3/2}}. \tag{6}$$

Our last step is to let $2w + 3 = y^2$. Equation (6) is then equal to

$$k'' \left[\sin^{-1} \frac{y}{\sqrt{5}} + \frac{3\sqrt{5-y^2}}{5y}\right]$$

and the final result is

$$10\sqrt{10} \left\{ \frac{\tan^{-1}}{3(5+3t^2)^{3/2}} + \frac{1}{6\sqrt{2}} \sin^{-1} \sqrt{\frac{5+3t^2}{5(1+t^2)}} + \frac{1}{10\sqrt{2}} \sqrt{\frac{2t^2}{5 + 3t^2}} \right\}$$

where $t = \frac{\sqrt{5(2-r^2)}}{r\sqrt{3}}$ as before.

Table 4 gives P' for α = .05, .01 and n = 4, 5, 6. In the case n = 4, α = .01, the value of $\frac{t_\alpha}{\sqrt{n}}$ falls beyond the upper bound, $\sqrt{3}$, of the function and therefore P' equals unity.

TABLE 4

n	P' α=0.05	α=0.01	Significance level of sign test
4	0.359	1	1/16
5	0.074	0.809	1/32
6	0.055	0.417	1/64

Using the second approach, we will now determine the "equivalent Type I error" of the sign test. From (5), the expected number of positive x's is

$$\mathcal{E}(r) = n\pi_c$$

$$= n\Pr\left(\frac{x_1 - \bar{x}}{s} > -\frac{t_\alpha}{\sqrt{n}}\right),$$

where t_α is based on n-1 degrees of freedom. Let $y = \frac{x_1 - \bar{x}}{s}$; then

$$y^2 = \frac{(x_1 - \bar{x})^2}{\frac{1}{n-1}\Sigma(x_i-\bar{x})^2} .$$

But

$$\sum_{i=1}^{n}(x_i-\bar{x})^2 = \sum_{j=2}^{n}(x_j-\bar{x}_{n-1})^2 + \frac{n}{n-1}(x_1-\bar{x})^2$$

where $$\bar{x}_{n-1} = \frac{\sum_{j=2}^{n} x_j}{n-1} .$$

The first term on the right hand side is distributed as $\chi^2_{n-2}\sigma^2$ and the second as $\chi^2_1\sigma^2$, i.e., is the square of a unit normal variable. Since the left hand term is distributed as $\chi^2_{n-1}\sigma^2$, it follows that the two terms on the right are statistically independent. We have, therefore,

$$y^2 = \frac{(n-1)(x_1-\bar{x})^2}{\chi^2_{n-2}\sigma^2 + \frac{n}{n-1}(x_1-\bar{x})^2}$$

$$= \frac{\frac{(n-1)^2}{n} \cdot \frac{\frac{n}{n-1}(x_1 - \bar{x})^2}{\frac{\chi^2_{n-2}\sigma^2}{n-2}}}{(n-2) + \frac{\frac{n}{n-1}(x_1-\bar{x})^2}{\frac{\chi^2_{n-2}\sigma^2}{n-2}}}$$

$$= \frac{\frac{(n-1)^2}{n} \quad t^2_{n-2}}{(n-2) + t^2_{n-2}} = \frac{(n-1)^2 \; t^2_{n-2}}{n(n-2+t^2_{n-2})} . \qquad (7)$$

Let $c = \frac{t_\alpha}{\sqrt{n}}$. Then

$$\pi_c = \Pr(y > -c)$$
$$= \Pr(y > 0) + \Pr(-c < y < 0)$$
$$= \tfrac{1}{2} + \tfrac{1}{2}\Pr(y^2 < c^2).$$

From (7) we have

$$\frac{n(n-2)y^2}{(n-1)^2-ny^2} = t^2_{n-2};$$

and since this is a monotone increasing function of y^2, we can write

$$\Pr(y^2 < c^2) = \Pr\left(\frac{n(n-2)y^2}{(n-1)^2-ny^2} < \frac{n(n-2)c^2}{(n-2)^2-nc^2}\right)$$

$$= \Pr(t^2_{n-2} < k^2)$$

$$= 2\Pr(0 < t_{n-2} < k),$$

provided $(n-1)^2 - ny^2 > 0$

$(n-1)^2 - nc^2 > 0.$

Hence

$$\pi_c = \tfrac{1}{2} + Pr(0 < t_{n-2} < k)$$

$$= Pr(t_{n-2} < k)$$

$$= Pr(t_{n-2} < \sqrt{\frac{(n-2)t_\alpha^2}{(n-1)^2 - t_\alpha^2}})$$

$$= Pr\ (t_{n-2} < t_\alpha \sqrt{\frac{n-2}{(n-1)^2 - t_\alpha^2}}).$$

The minimum of $\frac{x_1 - \bar{x}}{s}$ is $-(\frac{n-1}{\sqrt{n}})$ so when $n-1 < t_\alpha$ we get $\pi_c = 1$.

We can now evaluate $\gamma = Pr(r > n\pi_c)$ by making the histogram representation of the binomial distribution of r (parameters n, $\frac{1}{2}$). This representation is necessary since $n\pi_c$ is, in general, not an integer. If we denote by I the integral part of $n\pi_c + \frac{1}{2}$,

$$\gamma = Pr(r \geq I + 1) + (n\pi_c + \tfrac{1}{2} - I)\ Pr(r = I).$$

For large values of n the normal approximation to the binomial is quite satisfactory. Hence

$$\gamma \doteq Pr[u > \sqrt{n}\ (2\pi_c - 1)] \text{ for large n.}$$

In Table 5 γ is given for several values of $\gamma = n-2$ and α. Some irregularities in the values obtained can be attributed to the discreteness of the variable $\underline{\gamma}$.

Table 5. Equivalent Type I error γ of the sign test corresponding to the paired t-test in samples of n with Type I error α

$\nu = n - 2$ \ α	0.05	0.025	0.01	0.005
2	0.0582	0.0312	0.0312	0.0312
3	.0918	.0308	.0171	.0156
4	.0892	.0522	.0145	.0103
5	.0912	.0501	.0240	.0077
6	.0971	.0518	.0245	.0142
7	.0872	.0570	.0211	.0138
8	.0972	.0511	.0278	.0119
9	.0945	.0574	.0262	.0161
10	.0939	.0568	.0264	.0152
12	.0896	.0587	.0263	.0171
15	.0956	.0577	.0271	.0174
20	.0964	.0576	.0281	.0182
30	.0950	.0565	.0301	.0187
60	.0950	.0588	.0308	.0193
∞	.0947	.0589	.0317	.0199

ACKNOWLEDGMENTS

The author wishes to express her sincere appreciation to   for his guidance and advice in preparing this thesis. His invaluable assistance made this paper possible.

The author wishes also to thank   for his helpful criticism and suggestions. Thanks are also extended to   who prepared the final typewritten copies.

REFERENCES

Borenius, G. (1958). On the distribution of the extreme values in a sample from a normal distribution. Skand. Aktuarietedskr. 131-66.

David, H. A., Perez, C. A. (1959). On comparing tests of the same hypothesis. Technical Report No. 44. Dept. of Statistics and Statistical Laboratory. Virginia Polytechnic Institute.

David, H. A., Hartley, H. O., and Pearson, E. S. (1954). The distribution of the ratio, in a single normal sample, of range to standard deviation. Biometrika, 41, 482-93.

Elderton, W. P. (1938). Frequency Curves and Correlation. Cambridge University Press, p. 57-58.

Godwin, H. J. and Hartley, H. O. (1945). Probability integral and percentage points of the mean deviation in normal samples. Biometrika, 33, 254-65.

Hartley, H. O. (1955). Some recent developments in analysis of variance. Comm. Pure and Appl. Math. 8, 47-72.

Thomson, G. W. (1955). Bounds for the ratio of range to standard deviation. Biometrika, 42, 268-269.

The vita has been removed from the scanned document

ABSTRACT

This thesis presents two alternative procedures for comparing standard and quick tests of a null hypothesis H_o. This comparison is usually made by plotting the power curves of each test for a fixed Type I error. However, the power curves give only an indication of the individual performances of each test and not of the extent to which they agree when applied to the same problem. The procedures discussed in this paper deal with determining this degree of agreement.

The first method determines the probability, P, that the quick test leads to a significant result at a level β, given that the standard test is just significant at level α. If the standard and quick tests are based on the statistics u_1 and u_2, respectively, the second approach determines the level of significance corresponding to the expected value of u_2 given that u_1 is just significant at level α. This level of significance is termed the "equivalent Type I error" of the quick test and denoted by γ.

Both methods are applied to compare tests of location, dispersion, and the paired t-test with the sign test, all in samples taken from a normal population. In the first two cases, values of P and γ are given for different sample sizes, and in the third case only the "equivalent Type I error" of the sign test is given, P being rather difficult to evaluate.