ON SAMPLES FROM A NORMAL BIVARIATE POPULATION

By C. T. Hsu

1. Introduction. In a number of papers written during the last ten years, J. Neyman and E. S. Pearson[1] have discussed certain general principles underlying the choice of tests of statistical hypotheses. They have suggested that any formal treatment of the subject requires in the first place the specification of (*i*) the hypothesis to be tested, say H_0, (*ii*) the admissible alternative hypotheses. An appropriate test will then consist of a rule to be applied to observational data, for rejecting H_0 in such a way that (*iii*) the risk of rejecting H_0 when it is true is fixed at some desired value (e.g., 0.05 or 0.01), (*iv*) the risk of failing to reject H_0 when some one of the admissible alternatives is true is kept as small as possible. With these general principles in mind, they have investigated how best the condition (*iv*) may be satisfied in different classes of problems. In many cases, though not in all, it has been found that the conditions are satisfied by the test obtained from the use of what has been termed the likelihood ratio, [9], [10], [14]. Once the problem has been specified, the test criterion is usually very easily found, although its sampling distribution, if H_0 is true, often presents great difficulties. In the present paper, I propose to use this method to obtain appropriate tests for a number of hypotheses concerning two normally correlated variables. The investigation was suggested by a recent application of the method by W. A. Morgan [6] to a problem originally discussed by D. J. Finney [3].

2. The hypotheses and the appropriate criteria. A sample of two variables x_1 and x_2 is supposed to have been drawn at random from a normal bivariate population, with the distribution

$$(1)\quad p(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho_{12}^2}} \exp\left\{-\frac{1}{2(1-\rho_{12}^2)}\left[\left(\frac{x_1-\xi_1}{\sigma_1}\right)^2 - 2\rho_{12}\left(\frac{x_1-\xi_1}{\sigma_1}\right)\left(\frac{x_2-\xi_2}{\sigma_2}\right) + \left(\frac{x_2-\xi_2}{\sigma_2}\right)^2\right]\right\}$$

where ξ_1, ξ_2, σ_1, σ_2, and ρ_{12} are the population parameters.

Morgan tested the hypothesis that the variances of the two variables are equal, i.e.,

$$H_1: \qquad \sigma_1 = \sigma_2 .$$

---

[1] See bibliography at the end of the paper.

Other hypotheses that will be considered in the present paper are as follows:

H_2 : Assuming $\sigma_1 = \sigma_2$; to test $\rho_{12} = \rho_0$.
H_3 : Assuming $\sigma_1 = \sigma_2$; to test $\xi_1 = \xi_2$.
H_4 : To test simultaneously $\sigma_1 = \sigma_2$, $\rho_{12} = \rho_0$.
H_5 : To test simultaneously $\sigma_1 = \sigma_2$, $\xi_1 = \xi_2$.
H_6 : Assuming $\sigma_1 = \sigma_2$ and $\xi_1 = \xi_2$; to test $\rho_{12} = \rho_0$.
H_7 : Assuming $\sigma_1 = \sigma_2$, and $\rho_{12} = \rho_0$; to test $\xi_1 = \xi_2$.

Derivation of the criteria. Let x_{1i} , x_{2i} be the measurements of the two characters on the ith individual of the sample, then the joint elementary probability law of the two sets of n observations $E = (x_{11}, x_{12}, \cdots, x_{1n}; x_{21}, x_{22}, \cdots, x_{2n})$ is

$$p(E \mid \xi_1, \xi_2, \sigma_1, \sigma_2, \rho_{12}) = \left(\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1 - \rho_{12}^2}}\right)^n \cdot \exp\left\{-\frac{1}{2(1 - \rho_{12}^2)} \sum_{i=1}^{n} \left[\left(\frac{x_{1i} - \xi_1}{\sigma_2}\right)^2 - 2\rho_{12}\left(\frac{x_{1i} - \xi_1}{\sigma_1}\right)\left(\frac{x_{2i} - \xi_2}{\sigma_1}\right) + \left(\frac{x_{2i} - \xi_2}{\sigma_2}\right)^2\right]\right\}. \tag{2}$$

It will be convenient to denote by A, B, C, D, the following conditions of the population from which the sample is supposed to be drawn.

(A) that stated in equation (1).
(B) that stated in the equation for H_1, namely
$\sigma_1 = \sigma_2 = \sigma$($\sigma$ being unspecified).
(C) $\xi_1 = \xi_2 = \xi$(ξ being unspecified).
(D) $\rho_{12} = \rho_0$.

Neyman and Pearson's method affords a simple rule for obtaining appropriate test criteria once two sets of conditions have been defined. These are
(a) the conditions which can be assumed to be satisfied in any case, and
(b) the conditions which are satisfied if the hypothesis to be tested is true.
The conditions (a) define a class Ω of admissible populations, and the conditions (b) define a sub-class ω of Ω to which the population must belong if the hypothesis tested be true.

The maximum value of $p(E \mid \xi_1, \xi_2, \sigma_1, \sigma_2, \rho_{12})$ when the parameters vary in such a way that the population sampled always belongs to Ω, is called $p(\Omega \text{ max.})$. The maximum value when the population is restricted to ω is called $p(\omega \text{ max.})$. The likelihood ratio for testing the hypothesis specifying the subset ω has been defined to be

$$\lambda = \frac{p\ (\omega \text{ max.})}{p\ (\Omega \text{ max.})}. \tag{3}$$

It will be seen that $1 \leq \lambda \leq 0$. By referring λ, or a monotonic function of λ, to its sampling distribution when the hypothesis tested is true, we obtain a scale on which to assess our judgment of the truth of the hypothesis tested.

For each of the hypotheses H_1 to H_7, λ of (3) can be found. However, we shall use a more convenient criterion.

(4) $$L = \lambda^{2/n}$$

which is a monotonic function of λ.

Thus the respective test criteria are found to be:

For H_1:

(5) $$L_1 = \frac{4s_1^2 s_2^2(1 - r_{12}^2)}{(s_1^2 + s_2^2)^2(1 - R_1^2)}$$

where $R_1 = \frac{2r_{12}s_1s_2}{s_1^2 + s_2^2}$ is the estimate of ρ_{12} when σ_1 and σ_2 are assumed to be equal.

For H_2:

(6) $$L_2 = \frac{(1 - \rho_0^2)(1 - R_1^2)}{(1 - \rho_0 R_1)^2}.$$

For H_3:

(7) $$L_3 = 1 \bigg/ \left\{1 + \frac{(\bar{x}_1 - \bar{x}_2)^2}{s_1^2 + s_2^2 - 2r_{12}s_1s_2}\right\}.$$

For H_4:

(8) $$L_4 = \frac{4(1 - \rho_0^2)s_1^2 s_2^2(1 - r^2)}{(s_1^2 + s_2^2)^2(1 - \rho_0 R_1)^2} = L_1 \times L_2.$$

For H_5:

(9) $$L_5 = \frac{4s_1^2 s_2^2(1 - r_{12}^2)}{\{s_1^2 + s_2^2 + \frac{1}{2}(\bar{x}_1 - \bar{x}_2)^2\}(1 - R_2^2)} = L_1 \times L_2.$$

For H_6:

(10) $$L_6 = \frac{(1 - \rho_0^2)(1 - R_2^2)}{(1 - \rho^0 R_2)^2}$$

where $R_2 = \frac{2r_{12}s_1s_2 - \frac{1}{2}(\bar{x}_1 - \bar{x}_2)^2}{s_1^2 + s_2^2 + \frac{1}{2}(\bar{x}_1 - \bar{x}_2)^2}$ is the estimate of ρ_{12} when both the σ's and the ξ's are assumed to be equal.

For H_7:

(11) $$L_7 = 1 \bigg/ \left\{1 + \frac{(1 + \rho_0)(\bar{x}_1 - \bar{x}_2)^2}{2(s_1^2 - 2\rho_0 r_{12}s_1s_2 + s_2^2)}\right\}^2.$$

The different hypotheses are also given in Table V, at the end of this paper, together with the conditions defining sets of Ω and ω, and the appropriate likelihood criteria.

To complete the solution we must find the distributions of L or some monotonic function of L in each case when the hypothesis tested is true, in order to assess the significance of an observed value of L.

3. The distributions of the criteria. In order to simplify the problem of finding the distributions of the criteria, consider the following transformation:

$$x_{1i} = (X_i - Y_i)/\sqrt{2} \qquad (12)$$
$$x_{2i} = (X_i + Y_i)/\sqrt{2}.$$

It is clear that in view of (1) X and Y will be two normally correlated variables. We shall denote this property by A' corresponding to A. The conditions B', C', D' corresponding to B, C, D respectively are as follows:

B': $\rho_{XY} = 0,$

C': $\xi_X = 0,$

D': $\sigma_Y^2 = \gamma_0 \sigma_X^2$ (when $\rho_{XY} = 0$)

where

$$\gamma_0 = \frac{1 + \rho_0}{1 - \rho_0}. \qquad (13)$$

Thus we have the equivalent hypotheses H_1', $H_2' \cdots H_7'$ corresponding to H_1, H_2, $\cdots H_7$. The likelihood ratios L_1', $L_2' \cdots L_7'$ may be determined in the same way as before, and, in view of the transformation (12), it will be seen that they are equal to L_1, $L_2 \cdots L_7$ respectively.

The tests of the hypotheses H_1', H_2', H_3' are now seen to be well known.

The test of $H_1' : \rho_{XY} = 0$ is the test for significance of a correlation coefficient, and the criterion L_1 becomes

$$L_1 = \lambda_{H_1}^{2/n} = 1 - r_{XY}^2. \qquad (14)$$

This test has been dealt with by Morgan [6] and Pitman [15], and has been referred to above.

The test of $H_2' : \sigma_Y^2/\sigma_X^2 = \gamma_0$ *when* $\rho_{XY} = 0$ can be treated as an extension of Fisher's z-test [5], since γ_0 is specified. If we write

$$u = \frac{S_Y^2}{S_X^2} = \frac{1 + R_1}{1 - R_1} = \frac{s_1^2 + s_2^2 + 2r_{12}s_1s_2}{s_1^2 + s_2^2 - 2r_{12}s_1s_2} \qquad (15)$$

the test criterion L_2 of (6) may be written

$$L_2 = \frac{4u}{\gamma_0(1 + u/\gamma_0)^2}. \qquad (16)$$

It is well known that if H_2' is true, then

$$p(u) = \frac{1}{\gamma_0 B[\frac{1}{2}(n-1), \frac{1}{2}(n-1)]} \left(\frac{u}{\gamma_0}\right)^{\frac{1}{2}(n-3)} \left(1 + \frac{u}{\gamma_0}\right)^{-(n-1)} \tag{17}$$

and the test appropriate to H_2' and therefore of H_2 is the associated z-test ($z = \frac{1}{2} \log u/\gamma_0$) with degrees of freedom $f_1 = f_2 = n - 1$. It may be easily shown that the two values of u cutting off equal tail areas from the distribution $p(u)$ will correspond to a single value of L_2.

The test of H_3' : $\xi_X = 0$ *when* $\rho_{XY} = 0$ is in the form of "Student's" t test. If we write

$$\frac{t^2}{n-1} = \frac{\bar{X}^2}{s_X^2} = \frac{(\bar{x}_1 - \bar{x}_2)^2}{s_1^2 + s_2^2 - 2r_{12}s_1s_2} \tag{18}$$

it follows that the test criterion L_3 of (12) may be written

$$L_3 = 1 \bigg/ \left(1 + \frac{t^2}{n-1}\right). \tag{19}$$

But it is well known that if $\xi_X = 0$, then

$$p(t) = \frac{1}{\sqrt{n-1}B[\frac{1}{2}, \frac{1}{2}(n-1)]} \left(1 + \frac{t^2}{n-1}\right)^{-\frac{1}{2}n}. \tag{20}$$

The 5% or 1% points of significance of t may be obtained from Fisher's t-table [5] with degrees of freedom $f = n - 1$.

The tests of H_4 *and* H_5. We infer from (14), (16) and (19) that L_1 is a function of r_{XY}, L_2 a function of S_Y and S_Y, and L_3 a function of X and S_X. It is clear that if r_{XY} is distributed independently of S_X and S_Y, then L_1 and L_2 are independent, i.e.,

$$p(L_1, L_2) = p(L_1)p(L_2) \tag{21}$$

and that if r_{XY} is distributed independently of X and S_X, then L_1 and L_3 are independent, i.e.,

$$p(L_1, L_3) = p(L_1)p(L_3). \tag{22}$$

It is known that X, Y are independent of S_X, S_Y, r_{XY}; and in addition that r_{XY} is distributed independently of S_X, S_Y if $\rho_{XY} = 0$. Therefore, if H_1' is true, then the relations (21) and (22) hold. Hence, knowing $p(L_1)$ and $p(L_2)$, a very simple transformation and integration gives $p(L_4)$. Similarly, the distribution of L_5 may be readily derived from those of L_1 and L_3.

But from the distribution of r_{XY} when $\rho_{XY} = 0$, by transformation (14), the distribution of L_1 assuming H_1' true is found to be

$$p(L_1) = \frac{1}{B[\frac{1}{2}(n-2), \frac{1}{2}]} L_1^{\frac{1}{2}(n-4)}(1 - L_1)^{-\frac{1}{2}}. \tag{23}$$

If H_2' is true, from (17), by transformation (16) we have

(24) $$p(L_2) = \frac{1}{B[\frac{1}{2}(n-1), \frac{1}{2}]} L_2^{\frac{1}{2}(n-3)}(1 - L_2)^{-\frac{1}{2}}.$$

Again, if H_3' is true, from (20), by transformation (19), we have

(25) $$p(L_3) = \frac{1}{B[\frac{1}{2}(n-1), \frac{1}{2}]} L_3^{\frac{1}{2}(n-3)}(1 - L_3)^{-\frac{1}{2}}$$

which is the same as the distribution of L_2. Therefore by comparing (21) and (22) we see that the distribution of L_5 when H_5' is true will be exactly the same as that of L_4 when H_4' is true. We shall therefore confine ourselves to the problem of obtaining the distribution of L_4 from those of L_1 and L_2.

Now

(26) $$p(L_1, L_2) = \frac{1}{B[\frac{1}{2}(n-2), \frac{1}{2}]B[\frac{1}{2}(n-1), \frac{1}{2}]} L_1^{\frac{1}{2}(n-4)}(1 - L_1)^{-\frac{1}{2}} L_2^{\frac{1}{2}(n-3)}(1 - L_2)^{-\frac{1}{2}}.$$

Applying the transformation

(27) $$L_4 = L_1 L_2 \qquad Z = L_2$$

and integrating with respect to Z from 0 to 1, we obtain

(28) $$p(L_4) = \tfrac{1}{2}(n-2)L_4^{\frac{1}{2}(n-4)}, \qquad 0 \leq L_4 \leq 1.$$

Thus we can construct the values of L_4 at the 5% and 1% levels for different values of n as given in Table I.

TABLE I

5% and 1% values of L_4 (or L_5)

n	5%	1%
5	.1357	.0464
6	.2509	.1000
7	.3017	.1585
8	.3684	.2154
9	.4249	.2683
10	.4729	.3162
12	.5493	.3981
15	.6307	.4924
20	.7169	.5995
24	.7616	.6579
30	.8074	.7197
40	.8541	.7848
60	.9019	.8532
120	.9505	.9249
∞	1.0000	1.0000

The test of H_6. In the case of testing $H_6'(\sigma_Y^2 = \gamma_0\sigma_X^2)$, assuming ρ_{XY} and ρ_X each to be zero, the likelihood estimate of σ_X^2 becomes $\Sigma X^2/n$ or $S_X^2 + \bar{X}^2$. The distribution of this quantity is the same as that of S_X^2 but with degrees of freedom n instead of $n - 1$. Therefore, by analogy with the previous result (17) used in testing H_2, if we write

$$v = \frac{nS_Y^2}{\Sigma X^2} = \frac{S_Y^2}{S_X^2 + \bar{X}^2} = \frac{1 + R_2}{1 - R_2} \tag{29}$$

then the likelihood criterion of H_6 becomes

$$L_6 = \frac{4v}{\gamma_0\left(1 + \frac{v}{\gamma_0}\right)^2} \tag{30}$$

and

$$p\left(v \;\middle|\; \frac{\sigma_Y^2}{\sigma_X^2} = \gamma_0\right) = \frac{1}{\gamma_0 B[\frac{1}{2}(n-1), \frac{1}{2}n]}\left(\frac{v}{\gamma_0}\right)^{\frac{1}{2}(n-3)}\left(1 + \frac{v}{\gamma_0}\right)^{-(n-\frac{1}{2})}. \tag{31}$$

Hence the test appropriate to H_6 is the associated z-test $z = \frac{1}{2}\log\left\{\frac{v}{\gamma_0}\Big/\frac{n-1}{n}\right\}$ with $f_1 = n - 1$, $f_2 = n$. We can use the z-table as before.

The test of H_7. Here we test whether $\xi_X = 0$. It may be seen that L_7 is a function of $\bar{X}^2/(S_Y^2 + \gamma_0 S_X^2)$. Further, if we assume that $\rho_{XY} = 0$ and also that $\sigma_Y^2 = \gamma_0 S_X^2$, then it will follow that $\Sigma(X - \bar{X})^2$ and $\frac{1}{\gamma_0}\Sigma(Y - \bar{Y})^2$ are each distributed independently as $\chi^2\sigma_X^2$ with $n - 1$ degrees of freedom; and hence their sum is distributed as $\chi^2\sigma_X^2$ with $2n - 2$ degrees of freedom. Also if $\xi_X = 0$ (and H_7' is true) X will be distributed normally about zero with standard error $\sigma_{X/\sqrt{n}}$. Hence we may write

$$L_7 = 1\Big/\left\{1 + \frac{t^2}{2n-2}\right\}^2 \tag{32}$$

where

$$t^2 = \bar{X}\Big/\sqrt{\frac{\Sigma(X - \bar{X})^2 + \Sigma(Y - \bar{Y})^2/\gamma_0}{n(2n-2)}} \tag{33}$$

and is distributed in accordance with "Student's" distribution with $2n - 2$ degrees of freedom,

$$p(t_2) = \frac{1}{\sqrt{2n-2}\,B[\frac{1}{2}, \frac{1}{2}(2n-2)]}\left(1 + \frac{t^2}{2n-2}\right)^{-\frac{1}{2}(2n-1)}. \tag{34}$$

In terms of original variables

$$\frac{t_2^2}{2n-2} = \frac{\gamma_0\bar{X}^2}{\gamma_0 S_X^2 + S_Y^2} = \frac{(1 + \rho_0)(\bar{x}_1 - \bar{x}_2)^2}{2(s_1^2 - 2\rho_0 r_{12}\, s_1 s_2 + s_2^2)}. \tag{35}$$

4. Comparison of the R_1-test and R_2-test with the r_{12}-test in cases where H_2 and H_6 are true respectively. It will be noted that in the preceding discussion we have been concerned with three different tests of the hypothesis that ρ_{12} has some specified value ρ_0. When there is no information available regarding the means and standard deviations of x_1 and x_2, the test is based on the sampling distribution of the ordinary product-moment coefficient r_{12}. If it may be assumed that $\sigma_1 = \sigma_2$, then we have the estimate

$$R_1 = \frac{2r_{12}s_1s_2}{s_1^2 + s_2^2}.$$

If besides $\sigma_1 = \sigma_2$, it may also be assumed that $\xi_1 = \xi_2$, then we have the estimate

$$R_2 = \frac{2r_{12}s_1s_2 - \frac{1}{2}(\bar{x}_1 - \bar{x}_2)^2}{s_1^2 + s_2^2 + \frac{1}{2}(\bar{x}_1 - \bar{x}_2)^2}.$$

From the point of view of testing hypotheses, all these criteria r_{12}, R_1, R_2 follow from the application of the likelihood ratio method. It will be noted that if $\sigma_1 = \sigma_2$, either the r_{12} or the R_1 test may be used. But, insofar as the likelihood principle is accepted, the latter should be regarded as the "better" test. Again, if $\sigma_1 = \sigma_2$ and $\xi_1 = \xi_2$, all three tests may be used, but that based on R_2 will be the "best". A question of interest is to investigate just what is meant by the "better" or the "best" test. We may ask how far the improvements are sufficient to justify the use of the R_1 and R_2 tests in place of the more generally used r_{12} test. One method of comparison is to examine what Neyman and Pearson [12] have termed the "power function" of the tests.

For example, when testing the hypothesis that a parameter θ has the value θ_0 in the population sampled, the power of the test criterion T with regard to the alternative hypothesis that $\theta = \theta_1 > \theta_0$ is given by the expression $\beta(\theta_1) = P\{T > T_\alpha \mid \theta = \theta_1\}$ where $T($ is the value of T at the level of significance α. This quantity $\beta(\theta)$ measures the chance that the test as specified will detect the fact that $\theta = \theta_0$, i.e., the chance of rejecting the hypothesis when it is not true. A test whose power function is never less than that of any other test is termed the uniformly most powerful test.

If the permissible alternative hypotheses to $\theta = \theta_0$ are both $\theta < \theta_0$ and $\theta > \theta_0$, then the power of the test T is given by the expression

$$\beta(\theta_1) = 1 - p\{T'_\alpha < T < T''_\alpha \mid \theta_1\}$$

where T'_α and T''_α are the values of T at both ends of the distribution at the level of the significance α. When the test is such that the power function has a minimum value α at $\theta = \theta_0$, it is said to be unbiased.

A test is termed biased if, for certain alternative hypotheses $\theta \neq \theta_0$, the chance of rejecting the hypothesis $\theta = \theta_0$ is less than the chance of rejecting this hypothesis when it is true.

In what follows it is proposed to compare the power functions of the tests based on r_{12}, R_1, and R_2 in order to obtain more complete evidence of the extent to which one is "better" than the other.

The distribution of R_1.[2] We have obtained the distribution of n when H_2' and therefore H_2 is true. We are now able to find the distribution of R_1 by applying the transformation of (15). Thus the distribution of R_1 in terms of ρ_0 is

$$p(R_1 \mid \rho_0) = \frac{(1-\rho_0^2)}{2^{n-2}B[\frac{1}{2}(n-1), \frac{1}{2}(n-1)]} \frac{(1-R_1^2)^{\frac{1}{2}(n-3)}}{(1-\rho_0 R_1)^{n-1}}. \tag{36}$$

The significance of R_1 may be assessed by the z-test, where we take

$$\begin{aligned} Z &= \frac{1}{2}\log_e \frac{u}{\gamma_0} = \frac{1}{2}\log\frac{1+R_1}{1-R_1} - \frac{1}{2}\log\frac{1+\rho_0}{1-\rho_0} \\ &= z' - \zeta, \text{ say} \end{aligned} \tag{37}$$

with degrees of freedom $f_1 = f_2 = n - 1$. R. A. Fisher's z-table may be used in this connection.

When $\rho_{12} = 0$, the distribution simplifies to

$$\begin{aligned} p(R_1 \mid \rho_{12} = 0) &= \frac{1}{2^{n-1}B[\frac{1}{2}(n-1), \frac{1}{2}(n-1)]}(1-R_1^2)^{\frac{1}{2}(n-3)} \\ &= \frac{1}{B[\frac{1}{2}(n-1), \frac{1}{2}]}(1-R_1^2)^{\frac{1}{2}(n-3)} \end{aligned} \tag{38}$$

since 2^{2n-2} $B[\frac{1}{2}(n-1), \frac{1}{2}(n-1)]$ is equal to $B[\frac{1}{2}(n-1), \frac{1}{2}]$ by duplication formula [16, p. 240].

The distribution (38) is similar in form to that of $p(r_{12} \mid \rho_{12} = 0)$ with $n - 1$ degrees of freedom instead of $n - 2$. The significance levels of R_1 may then be obtained directly from the r-table [1] for the case $\rho_{12} = 0$, entering with degrees of freedom $n - 1$.

The distribution of R_2. The distribution of R_2 may be obtained from that of v when H_6' and therefore H_6 is true. It is

$$p(R_2 \mid \rho_{12} = \rho_0) = \frac{(1+\rho_0)^{\frac{1}{2}n}(1-\rho_0)^{\frac{1}{2}(n-1)}}{2^{n-\frac{1}{2}}B[\frac{1}{2}(n-1), \frac{1}{2}n]} \frac{(1+R_2)^{\frac{1}{2}(n-3)}(1-R_2)^{\frac{1}{2}(n-2)}}{(1-\rho_0 R_2)^{n-\frac{1}{2}}}. \tag{39}$$

This agrees with the result first obtained by R. A. Fisher [4] by a different method. The significance of R_2 may be assessed by the z-test, where we take

$$z = \frac{1}{2}\log\left(\frac{v}{\gamma_0} \Big/ \frac{n-1}{n}\right) \tag{40}$$

---

[2] Since finding the distribution of R_1 (36), (38) and the relation between R_1 and z' (37), my attention has been drawn to a recent paper by DeLury [2] in which the same results are obtained. Since my method of derivation is different from his, I have thought it worthwhile to retain it here.

with degrees of freedom $f_1 = n - 1, f_2 = n$. The tables for use with the z-test may be used in this connection.

When $\rho_{12} = 0$, the distribution is simplified to

$$(41) \qquad p(R_2 \mid \rho_{12} = 0) = \frac{1}{2^{n-\frac{1}{2}} \mathrm{B}[\frac{1}{2}(n-1), \frac{1}{2}n]} (1 + R_2)^{\frac{1}{2}(n-3)} (1 - R_2)^{\frac{1}{2}(n-2)}$$

which is simply a Pearson Type I curve.

Power functions of R_1 and R_2. In order to find the power functions of R_1 and R_2 with respect to alternative hypotheses H_t to H_2, specifying $\rho_{12} = \rho_t < \rho_0$, it will be convenient to consider the incomplete beta function distributions

$$(42) \qquad p(x_1) = \frac{1}{\mathrm{B}[\frac{1}{2}(n-1), \frac{1}{2}(n-1)]} x_1^{\frac{1}{2}(n-3)} (1 - x_1)^{\frac{1}{2}(n-3)}$$

$$(43) \qquad p(x_2) = \frac{1}{\mathrm{B}[\frac{1}{2}(n-1), \frac{1}{2}n]} x_2^{\frac{1}{2}(n-3)} (1 - x_2)^{\frac{1}{2}(n-2)}$$

where $x_1 = \dfrac{u}{\gamma_0(1 + u/\gamma_0)}$ and $x_2 = \dfrac{v}{\gamma_0(1 + v/\gamma_0)}$. From the *Tables of the Incomplete Beta Function* [13] we can find the values of x_1 and x_2 at the significance level α, i.e.

$$(44) \qquad I_{x_1'} [\tfrac{1}{2}(n-1), \tfrac{1}{2}(n-1)] = \alpha',$$

$$(45) \qquad I_{x_2'} [\tfrac{1}{2}(n-1), \tfrac{1}{2}n] = \alpha'.$$

The values of $R_1'(\alpha)$, and of $R_2'(\alpha)$, may then be calculated from the relations

$$(46) \qquad R_1 = \frac{u-1}{u+1} = \frac{-1 + x_1 + \gamma_0 x_1}{1 - x_1 + \gamma_0 x_1},$$

$$(47) \qquad R_2 = \frac{v-1}{v+1} = \frac{-1 + x_2 + \gamma_0 x_2}{1 - x_2 + \gamma_0 x_2}.$$

The power functions of R_1 and R_2 thus found may be given as follows:

$$(48) \qquad \beta'(\rho_t \mid R_1) = P\{R_1 < R_1'(\alpha) \mid \rho_t\},$$

$$(49) \qquad \beta'(\rho_t \mid R_2) = P\{R_2 < R_2'(\alpha) \mid \rho_t\}.$$

In the same way, for any alternative hypothesis H_t specifying $\rho_{12} = \rho_t > \rho_0$, we can find the values of x_1 and x_2 at the significance level α'', at the other end of the distribution, i.e.

$$(50) \qquad 1 - I_{x_1''} [\tfrac{1}{2}(n-1), \tfrac{1}{2}(n-1)] = \alpha'',$$

$$(51) \qquad 1 - I_{x_2''} [\tfrac{1}{2}(n-1), \tfrac{1}{2}n] = \alpha''.$$

Thence the corresponding values of $R_1''(\alpha)$ and $R_2''(\alpha)$ may be obtained, and their power functions are

$$(52) \qquad \beta''(\rho_t \mid R_1) = P\{R_1 > R_1''(\alpha) \mid \rho_t\},$$

(53) $$\beta''(\rho_t \mid R_2) = P\{R_2 > R_2''(\alpha) \mid \rho_t\}.$$

The power functions of R_1 and R_2 with respect to alternative hypotheses specifying $\rho_{12} = \rho_t < \rho_0$ and $> \rho_0$ may now be obtained by adding (48) and (52) or (49) and (53) or, more simply,

(54) $$\beta(\rho_t \mid R_1) = 1 - P\{R_1'(\alpha) < R_1 < R_1''(\alpha) \mid \rho_t\},$$

(55) $$\beta(\rho_t \mid R_2) = 1 - P\{R_2'(\alpha) < R_2 < R_2''(\alpha) \mid \rho_t\}$$

where $R_1'(\alpha)$, $R_1''(\alpha)$; $R_2'(\alpha)$, $R_2''(\alpha)$ are the values of R_1 and R_2 at the two ends of the distribution at the significance level $\alpha = \alpha' + \alpha''$.

In view of the fact that after transformation the tests based on R_1 and R_2 are equivalent to tests regarding the equality of variances, it follows from Neyman and Pearson's work [11] regarding the uniformly most powerful test of the hypothesis that $\sigma_Y^2/\sigma_X^2 = \gamma_0$, with alternatives $\sigma_Y^2/\sigma_X^2 = \gamma_t < \gamma_0$ (or $\gamma_t > \gamma_0$), that: (1) if $\sigma_1 = \sigma_2$ and alternative to $\rho_{12} = \sigma_0$ are that $\rho_{12} = \rho_t < \rho_0$ (or, in a second case, $\rho_t > \rho_0$) the test based on R_1 is the uniformly most powerful test, i.e., it is more powerful than that based on r_{12}; and (2) if $\sigma_1 = \sigma_2$ and $\xi_1 = \xi_2$, then the test based on R_2 is the uniformly most powerful test, i.e., it is more powerful than those based on either r_{12} or R_1.

For illustration, let us take a special case, say

(*a*) $$n = 10, \qquad \rho_0 = 0.6, \qquad \alpha' = \alpha'' = 0.025.$$

From the tables, we obtain the values

$$x_1' = .198902 \qquad x_2' = .184863$$
$$x_1'' = .801098 \qquad x_2'' = .772916$$

and by calculation the values

$$R_1'(\alpha) = -.0034 \qquad R_2'(\alpha) = -.0487$$
$$R_1''(\alpha) = .8831 \qquad R_2''(\alpha) = .8632.$$

The values of the power functions of R_1 and R_2 for specified values of ρ_t have been calculated and are given in Table II. For $\rho_t < \rho_0$, a comparison of columns 2 and 4 will show that the test based on R_2 is uniformly more powerful than that based on R_1 (or for $\rho_t > \rho_0$, a comparison of columns 3 and 5).

The unbiased test of H_2 and H_6. When however the alternatives are that $\rho_{12} = \rho_t < \rho_0$, and $\rho_t > \rho_0$, questions of bias may be introduced.

In the case of H_2, i.e. when R_1 is used, it was established by J. Neyman in his lecture courses [8], that if we test whether $\sigma_Y^2/\sigma_X^2 = \gamma_0$, where the alternatives are $\gamma_t < \gamma_0$ and $\gamma_t > \gamma_0$, and if the samples of X and Y are of equal size, then the test based on cutting off equal tail areas of the distribution of x_1 is unbiased and of the type B [7]. Therefore the same may be said of the R_1-test.

In the case of H_6, the equivalent transformed test is again whether $\sigma_Y^2/\sigma_X^2 = \gamma_0$. But the test now corresponds to that in which an estimate of σ_Y^2 is based on $f_1 = n - 1$ degrees of freedom and an estimate of σ_X^2 on $f_2 = n$ degrees of freedom. The degrees of freedom not being equal, it is known that if equal tail areas are cut off from the sampling distribution of x_2, this test will be biased. Neyman's result [8] shows that if the lower and upper significance levels are taken at x_2' and x_2'', then the equation

$$x_2''^{f_1}(1 - x_2'')^{f_2} = x_2'^{f_1}(1 - x_2')^{f_2} \tag{56}$$

should be satisfied if the test is unbiased. Since in the present case, with the test based on equal tail area critical region, the bias will be very small, the rejection levels $R_2'(\alpha)$ and $R_2''(\alpha)$ in the numerical investigation given in Table III have been selected taking equal tail areas for simplicity.

TABLE II

Values of the power functions of R_1 and R_2 with respect to alternative hypotheses $\rho_{12} = \rho_t < \rho_0$ or $\rho_t > \rho_0$

($n = 10$; $\rho_0 = 0.6$; $\alpha' = \alpha'' = 0.025$)

ρ_t	$\beta'(\rho_t \mid R_1)$	$\beta''(\rho_t \mid R_1)$	$\beta'(\rho_t \mid R_2)$	$\beta''(\rho_t \mid R_2)$
−0.8	.9984			
−0.6	.9739		.9807	
−0.4	.9867		.9005	
−0.2	.7189		.7360	
0.0	.4960	.0002	.5093	.0001
0.2	.2744	.0008	.2809	.0006
0.3	.1825	.0018	.1860	.0015
0.4	.1106	.0042	.1111	.0037
0.5	.0576	.0099	.0580	.0093
0.6	.025	.025	.025	.025
0.7	.0081	.0678	.0080	.0720
0.8	.0015	.1995	.0015	.2150
0.9	.0001	.5950	.0001	.6289
0.95		.8979		.9150
0.975		.9866		.9897

If we now take a special case, similar to (a) above, but taking equal tail areas, so that

$$n = 10 \qquad \rho = 0.6$$
$$\alpha = 0.5 \qquad (\alpha' = \alpha'' = \tfrac{1}{2}\alpha)$$

we can obtain the values of x's and of R's as before.

The values of the power functions of R_1 and R_2 for specified values of ρ_t are given in columns 3 and 4 of Table III. These values are equivalent to the sums of the corresponding values in Table II. The values of the power functions of R_1 and R_2 for the following additional cases are also given in Table III:

(*b*) $n = 10 \quad \rho_0 = 0.8 \quad \alpha = 0.05$

(*c*) $n = 20 \quad \rho_0 = 0.6 \quad \alpha = 0.05$

(*d*) $n = 20 \quad \rho_0 = 0.8 \quad \alpha = 0.05.$

Comparison of the power functions. We may now deal with the question raised at the beginning of this section, namely, as to what is meant by the "better" or "best" test. We shall proceed to compare for certain special cases the power functions of the three test, all of which are applicable where it may be assumed that $\sigma_1 = \sigma_2$, $\xi_1 = \xi_2$.

In the first place it will be noted that the power function of the test based on equal tail areas of the r_{12} distribution is

$$\beta(\rho_t \mid r_{12}) = 1 - p\{\gamma'_{12}(\alpha) < \gamma_{12} < \gamma''_{12}(\alpha) \mid \rho_t\} \tag{57}$$

where

$$\begin{aligned} P\{r_{12} < r'_{12}(\alpha) \mid \rho_0\} &= \int_{-1}^{r'_{12}(\alpha)} p(r_{12} \mid \rho_{12} = \rho_0)\, dr_{12} = \tfrac{1}{2}\alpha \\ P\{r_{12} > r'_{12}(\alpha) \mid \rho_0\} &= \int_{r''_{12}(\alpha)}^{1} p(r_{12} \mid \rho_{12} = \rho_0)\, dr_{12} = \tfrac{1}{2}\alpha \end{aligned} \tag{58}$$

and

$$p(r_{12} \mid \rho_{12} = \rho_0) = \frac{(1 - \rho_0^2)^{\frac{1}{2}(n-1)}}{\pi\Gamma[\frac{1}{2}(n-1)]}(1 - r_{12}^2)^{\frac{1}{2}(n-4)}\left(\frac{\partial}{\partial r_{12}}\right)^{n-2}\frac{\cos^{-1}(-\rho_0 r_{12})}{\sqrt{(1 - \rho_0^2 r_{12}^2)}}. \tag{59}$$

The probability that r_{12} is less than some specified value may be obtained from *Tables of the Correlation Coefficient* (F. N. David, [1]), or, where these are not sufficiently detailed, by using R. A. Fisher's z'-transformation for r_{12} [4].

The cases considered are (*a*), (*b*), (*c*), (*d*) as defined above. The power functions of the three different tests (all based upon the equal tail areas of their distributions) are given in Table III. The figures for r_{12} in the brackets are those obtained by the z'-transformation approximation.

An examination of Tables II and III brings out the following points:

(1) For reasons given above, the R_2 test based on equal tail area critical regions is very slightly biased; the amount of this bias for the case $n = 10$, $\rho_0 = 0.6$, $\alpha = 0.05$ is shown in Table IV. This shows that the power of the R_2 test is less than 0.05 in the fifth or sixth decimal places for $0.59 < \rho_t < 0.60$. As a result this test is very slightly less powerful than the other two tests for alternatives with ρ_t slightly less than ρ_0. The effect is, however, of little importance.

(2) Except in this short range of ρ_t, we find that

$$\beta(\rho_t \mid R_2) \geq \beta(\rho_t \mid R_1) \geq \beta(\rho_t \mid r_{12}).$$

TABLE III

Comparison of the power functions of r_{12}, R_1, and R_2 tests with respect to alternative hypotheses

ρ_t	$n = 10$ $\rho_0 = 0.6$			$n = 10$ $\rho_0 = 0.8$			$n = 20$ $\rho_0 = 0.6$			$n = 20$ $\rho_0 = 0.8$		
	$\beta(\rho_t \mid r_{12})$	$\beta(\rho_t \mid R_1)$	$\beta(\rho_t \mid R_2)$	$\beta(\rho_t \mid r_{12})$	$\beta(\rho_t \mid R_1)$	$\beta(\rho_t \mid R_2)$	$\beta(\rho_t \mid r_{12})$	$\beta(\rho_t \mid R_1)$	$\beta(\rho_t \mid R_2)$	$\beta(\rho_t \mid r_{12})$	$\beta(\rho_t \mid R_1)$	$\beta(\rho_t \mid R_2)$
−0.6	.9739	.9739	.9807									
−0.4	.8865	.8867	.9005	.9887	.9891	.9921	.9965	.9967	.9973			
−0.2	.7186	.7189	.7360	.9557	.9569	.9650	.9648	.9663	.9698			
0.0	.4960	.4962	.5094	.9742	.8766	.8909	.8328	.8369	.8449	.9952	.9959	.9966
0.2	.2753	.2752	.2815	.7158	.7189	.7360	.5412	.5456	.5534	.9624	.9663	.9698
0.4	.1142	.1148	.1148	.4727	.4750	.4877	.2026	.2036	.2061	.8062	.8170	.8254
0.5	.0679	.0675	.0673	.3330	.3345	.3427	.0915	.0917	.0922	.6309	.6432	.6520
0.6	.0500	.0500	.0500	.2005	.2010	.2047	.0500	.0500	.0500	.3920	.4011	.4085
0.7	.0735	.0759	.0800	.0969	.0965	.0971	.1096	.1119	.1147	.1589	.1617	.1635
0.8	.1890	.2010	.2165	.0500	.0500	.0500	.3886	.4010	.4134	.0500	.0500	.0500
0.9	.5656	.5951	.6290	.1466	.1771	.1904	.9034	.9106	.9181	.3272	.3493	.3604
	(.5569)			(.1454)			(.9004)			(.3270)		
0.95	(.8709)	.8979	.9150	(.4689)	.5426	.5763	(.9974)	.9974	.9978	(.8547)	.5871	.8844
0.975	(.9822)	.9866	.9897	(.8134)	.8692	.8896				(.9944)	.9947	.9960
0.99				(.9845)	.9908	.9938						
Levels	−.0039	−.0034	−.0487	.4004	.3817	.3423	.2289	.2253	.2041	.5671	.5613	.5459
	.8998	.8831	.8632	.9574	.9463	.9368	.8300	.8201	.8084	.9222	.9158	.9101

That is to say, the power function of the R_2 test never lies below those of the R_1 and r_{12} tests, and that of the R_1 test never lies below that of the r_{12} test.

(3) The gain in sensitivity as measured by the chance that the test will detect that $\rho_t \neq \rho_0$ is, however, very small. Further, R_1 may only be used if it is known that $\sigma_1 = \sigma_2$ and R_2 if it is known in addition that $\xi_1 = \xi_2$. It will only be in rather special problems that the statistician can feel confident that such assumptions are justified. We will therefore probably prefer the test based on the ordinary product moment correlation coefficient r_{12}, since the slight loss in power will be felt to be outweighed by the gain in simplicity. It is, however, only after an objective comparison of the consequences of applying the three tests that a definite opinion on these points can be reached.

TABLE IV

ρ_t	$\beta'(\rho_t \mid R_2)$	$\beta''(\rho_t \mid R_2)$	$\beta(\rho_t \mid R_2)$
0.5	.0580	.0093	.0673
0.590	.0274235	.0225806	.0500041
0.591	.0271778	.0228190	.0499968
0.592	.0269359	.0230578	.0499937
0.593	.0266934	.0232976	.0499910
0.594	.0264515	.0235337	.0499852
0.595	.0262096	.0237798	.0499894
0.596	.0259677	.0240222	.0499899
0.597	.0257257	.0242651	.0499908
0.598	.0254838	.0245107	.0499945
0.599	.0252419	.0247540	.0499959
0.6	.025	.025	.05

5. Summary. Various hypotheses relating to a population of two normal correlated variates have been considered and the appropriate test criteria for each hypothesis have been derived by the likelihood ratio method. The distributions of the likelihood ratio criteria or of monotonic functions of them have been obtained with the aid of transformation (14). References have been given to tables from which significance levels for use in conjunction with the tests may be obtained; a new table of significance levels for the tests of H_4 and H_5 was given.

The power functions of r_{12}, R_1 and R_2 have been compared; from these power functions it was concluded that R_1 and R_2 are suitable respectively for testing the hypothesis when $\sigma_1 = \sigma_2$ and when, in addition, $\xi_1 = \xi_2$.

In conclusion, I should like to express my indebtedness to Professor E. S. Pearson for continued advice and help in the preparation of this paper, to Dr. A. Wald and Professor S. S. Wilks for valuable suggestions.

TABLE V

Conditions defining Ω and ω together with the likelihood criteria appropriate for testing the hypotheses H_i

(1) Hypotheses H_i	(2) Initial Assumptions (Apart from Normality)	(3) To be tested	(4) Conditions defining (a) Ω	(5) Conditions defining (b) ω	(6) Criteria $L_i = \lambda_{H_i}^{2/n}$
H_1	None	$\sigma_1 = \sigma_2$	A	A, B	$\dfrac{4s_1^2 s_2^2(1 - r_{12}^2)}{\{(s_1^2 + s_2^2)^2 - 4r_{12}^2 s_1^2 s_2^2\}}$
H_2	$\sigma_1 = \sigma_2$	$\rho_{12} = \rho_0$	A, B	A, B, D	$\dfrac{(1 - \rho_0^2)(1 - R_1^2)}{(1 - \rho_0 R_1)^2}$ [1]
H_3	$\sigma_1 = \sigma_2$	$\xi_1 = \xi_2$	A, B	A, B, C	$1 \Big/ \left\{1 + \dfrac{(\bar{x}_1 - \bar{x}_2)^2}{s_1^2 + s_2^2 - 2r_{12} s_1 s_2}\right\}$
H_4	None	$\sigma_1 = \sigma_2$ $\rho_{12} = \rho_0$	A,	A, B, D	$\dfrac{4s_1^2 s_2^2(1 - \rho_0^2)(1 - r_{12}^2)}{(s_1^2 + s_2^2)(1 - \rho_0 R_1)^2}$
H_5	None	$\sigma_1 = \sigma_2$ $\xi_1 = \xi_2$	A,	A, B, C	$\dfrac{4s_1^2 s_2^2(1 - r_{12}^2)}{\{s_1^2 + s_2^2 + \frac{1}{2}(\bar{x}_1 - \bar{x}_2)^2\}(1 - R_2^2)}$ [2]
H_6	$\sigma_1 = \sigma_2$ $\xi_1 = \xi_2$	$\rho_{12} = \rho_0$	A, B, C	A, B, C, D	$\dfrac{(1 - \rho_0^2)(1 - R_2^2)}{(1 - \rho_0 R_2)^2}$
H_7	$\sigma_1 = \sigma_2$ $\rho_{12} = \rho_0$	$\xi_1 = \xi_2$	A, B, D	A, B, D, C	$1 \Big/ \left\{1 + \dfrac{(1 + \rho_0)(\bar{x}_1 - \bar{x}_2)^2}{2(s_1^2 - 2\rho_0 r_{12} s_1 s_2 + s_2^2)}\right\}^2$

[1] $R_1 = \dfrac{2r_{12} s_1 s_2}{s_1^2 + s_2^2}$ [2] $R_2 = \dfrac{2r_{12} s_1 s_2 - \frac{1}{2}(\bar{x}_1 - \bar{x}_2)^2}{s_1^2 + s_2^2 + \frac{1}{2}(\bar{x}_1 - \bar{x}_2)^2}$

REFERENCES

[1] F. N. DAVID, *Tables of the Correlation Coefficient*, London: Biometrika Office, 1938.
[2] D. B. DELURY, *Ann. Math. Stat.* Vol. 9 (1938) p. 149.
[3] D. J. FINNEY, *Biometrika*, Vol. 30 (1938), p. 190.
[4] R. A. FISHER, *Metron*, Vol. 1 (1921), p. 3.
[5] R. A. FISHER, *Statistical Methods for Research Workers*, 7th ed. Edinburgh: Oliver Boyd, 1938.
[6] W. A. MORGAN, *Biometrika*, Vol. 31 (1939), p. 13.
[7] J. NEYMAN, *Bull. Soc. Math. France*, Vol. 63 (1935), p. 246.
[8] J. NEYMAN, Lectures delivered in London, 1937-8, (Unpublished).
[9] J. NEYMAN AND E. S. PEARSON, *Biometrika*, Vol. 20A (1928), p. 175.
[10] J. NEYMAN AND E. S. PEARSON, *Bull. Acad. Polonaise Sci. Lettres*, A (1931), p. 460.
[11] J. NEYMAN AND E. S. PEARSON, *Phil. Trans.*, Ser. A, Vol. 236 (1933), p. 289.
[12] J. NEYMAN AND E. S. PEARSON, *Stat. Res. Mem.* Vol. 1 (1936), p. 1.
[13] K. PEARSON (Editor), *Tables of the Incomplete Beta-Function*, London: Biometrika Office.
[14] E. S. PEARSON AND J. NEYMAN, *Bull. Acad. Polonaise Sci. Lettres*, A (1930), p. 73.
[15] J. C. PITMAN, *Biometrika*, Vol. 31 (1939), p. 9.
[16] E. T. WHITTAKER AND G. N. WATSON, *Modern Analysis*, 4th Edition (1927).

UNIVERSITY COLLEGE,
LONDON, ENGLAND.