COMPARISON AMONG THREE CRITERIA TO TEST THE EQUALITY OF CORRELATION COEFICIENTS[1]

JOSÉ RUY PORTO DE CARVALHO[2] and ROGER MEAD[3]

ABSTRACT - In this paper, the Likelihood Ratio Test is presented. It is shown to be an efficient procedure to test the hypothesis of equality of residual correlation coefficients. In comparing the Likelihood Ratio Test with the largely used U-Fisher statistic test, or the improved Rao version of it for testing the hypothesis of equality of residual correlation coefficients, very interesting result are found. The simulated and experimental results show that the Likelihood Ratio Test rejects the null hypothesis in more experimental situations than the U-Fisher or H-Rao test statistics, with grater observed power.

Index terms: maximum likelihood estimator, likelihood ratio test, split-plot analysis, correlation coefficients, observed power.

COMPARAÇÃO ENTRE TRÊS CRITÉRIOS PARA TESTAR IGUALDADE DE COEFICIENTES DE CORRELAÇÃO

RESUMO - Neste trabalho, o Teste da Razão de Máxima Verossimilhança é apresentado como um eficiente procedimento para testar a hipótese de igualdade de coeficientes de correlação residual. Comparando o Teste da Razão de Máxima Verossimilhança com o Teste U-Fisher e a versão de Rao do mesmo teste, interessantes conclusões são encontradas. Os resultados simulados e experimentais mostram que o Teste da Razão da Máxima Verossimilhança rejeita a hipótese de nulidade, com maior poder, em um número maior de vêzes quando comparado aos testes U-Fisher e H-Rao.

Termos para indexação: estimador de máxima verossimilhança, teste da razão de máxima verossimilhança, parcelas-subdivididas, coeficiente de correlação, poder observado.

INTRODUCTION

The objective of this paper is the development of a test to establish whether the main-plot and split-plot dispersion matrices have the same population correlation coefficients. The hypothesis stated below is not that two independent normally distributed bivariate populations are identical, but merely that their correlation coefficients are homogeneous, that is:

$$H_0 : \rho_1 = \rho_2 = \rho \quad \text{vs} \quad H_1 : \rho_1 \neq \rho_2 , \qquad (1)$$

with possibly differing variances and covariances.

The Likelihood Function (LF) under the null hypothesis $H_0 = \rho_1 = \rho_2 = \rho$ becomes:

$$LF(\sigma_{11}, \sigma_{12}, \sigma_{21}, \sigma_{22}, \rho) =$$

[1] Accepted for publication on October 23, 1991.
Extracted from a Dissertation submitted by the senior author in partial fulfillment of the requirements for the Ph.D. degree to the Department of Applied Statistics. The University of Reading, Berkshire, U.K.

[2] Statistician, Ph.D., EMBRAPA/NMA - P.O. Box 491 - 13001 - Campinas, SP, Brazil - Fone: (0192) 52.5977.

[3] Applied Statistics Professor, Department of Applied Statistics - The University of Reading, P.O. Box 217 - Reading, RG6 - 2AN - U.K.

02

$$\frac{(df_1+1)^{df_1}\, s_{11}^{df_1-1}\, s_{12}^{df_1-1}\, (1-r_1^2)^{(df_1-3)/2}}{\pi^2\, \Gamma(df_1-1)\, \sigma_{11}^{df_1}\, \sigma_{12}^{df_1}} \times$$

$$\frac{(df_2+1)^{df_2}\, s_{21}^{df_2-1}\, s_{22}^{df_2-1}\, (1-r_2^2)^{(df_2-3)/2}}{\Gamma(df_2-1)\, \sigma_{21}^{df_2}\, \sigma_{22}^{df_2}\, (1-\rho^2)^{(df_1+df_2)/2}} \times$$

$$\exp\left[-\frac{(df_1+1)}{2(1-\rho^2)}\left[\frac{s_{11}^2}{\sigma_{11}^2} - \frac{2\rho r_1 s_{11} s_{12}}{\sigma_{11}\sigma_{12}} + \frac{s_{12}^2}{\sigma_{12}^2}\right] - \frac{(df_2+1)}{2(1-\rho^2)}\left[\frac{s_{21}^2}{\sigma_{21}^2} - \frac{2\rho r_2 s_{21} s_{22}}{\sigma_{21}\sigma_{22}} + \frac{s_{22}^2}{\sigma_{22}^2}\right]\right] \quad (2)$$

where s_{11}^2, s_{12}^2 and $s_{11}s_{12}r_1$ are the residual sampling variances and covariance for main-plot with df_1 degrees of freedom and s_{21}^2, s_{22}^2 and $s_{21}s_{22}r_2$ are for the split-plot with df_2 degrees of freedom.

When applying logarithm to (2), and setting the derivatives to zero, the following equation in $\hat{\rho}$ is found:

$$(r_2 df_1 + r_1 df_2)\hat{\rho}^2 - [(df_1 + df_2)(r_1 r_2 + 1)]\hat{\rho} + (r_1 df_1 + r_2 df_2) = 0 \quad . \quad (3)$$

The only constraint which needs to be applied in the values of ρ is that if $(r_2 df_1 + r_1 df_2) = 0$ in (3), then

$$\hat{\rho} = \frac{(r_1 df_1 + r_2 df_2)}{(df_1 + df_2)(r_1 r_2 + 1)}$$

Otherwise, two real roots can be obtained, and the Maximum Likelihood (ML) is the one which corresponds to the largest value of the LF defined in (2), that is:

$$\hat{\rho} = \frac{(df_1+df_2)(r_1 r_2+1) - \left[(df_1+df_2)^2(r_1 r_2-1)^2 - 4 df_1 df_2 (r_1-r_2)^2\right]^{\frac{1}{2}}}{2(r_2 df_1 + r_1 df_2)} \quad (4)$$

The remaining four ML estimators are found by substituting (4) in the normal equations, and resulted:

$$\hat{\sigma}_{11}^2 = \frac{(df_1+1)(1-\hat{\rho} r_1)}{df_1(1-\hat{\rho}^2)}\, s_{11}^2,$$

$$\hat{\sigma}_{12}^2 = \frac{(df_1+1)(1-\hat{\rho} r_1)}{df_1(1-\hat{\rho}^2)}\, s_{12}^2,$$

$$\hat{\sigma}^2_{21} = \frac{(df_2+1)(1-\hat{\rho} r_2)}{df_2(1-\hat{\rho}^2)} s^2_{12} \quad \text{and}$$

$$\hat{\sigma}^2_{22} = \frac{(df_2+1)(1-\hat{\rho} r_2)}{df_2(1-\hat{\rho}^2)} s^2_{22} \qquad (5)$$

The 95% confidence interval for ρ is as follows:

$$Pr\left[\hat{\rho} - \frac{1.96(1-\hat{\rho}^2)}{(df_1+df_2)^{\frac{1}{2}}} < \rho < \hat{\rho} + \frac{1.96(1-\hat{\rho}^2)}{(df_1+df_2)^{\frac{1}{2}}}\right] = 0.95 \,,$$

where $V(\hat{\rho}) = \dfrac{[1-\hat{\rho}^2]^2}{df_1+df_2}$ is obtained by evaluating the 5x5 information matrix at the MLE point.

Substituting (5) in (2), the maximized restricted LF is as follows:

$$LF\{\hat{\sigma}_{11},\hat{\sigma}_{12},\hat{\sigma}_{21},\hat{\sigma}_{22},\hat{\rho}\} =$$

$$\frac{df_1^{df_1}(1-r_1^2)^{\frac{(df_1-3)}{2}}\, df_2^{df_2}(1-r_2^2)^{\frac{(df_2-3)}{2}}\,(1-\hat{\rho}^2)^{\frac{(df_1+df_1)}{2}}\exp(-(df_1+df_2))}{\pi^2\,\Gamma(df_1-1)\,\Gamma(df_2-1)\, s_{11}s_{12}s_{21}s_{22}\,(1-\hat{\rho} r_1)^{df_1}(1-\hat{\rho} r_2)^{df_2}} \qquad (6)$$

The Likelihood Ratio Test (LRT)

The λ LRT may serve as a reasonable statistic for testing H_0, with small values of λ leading to the rejection of H_0. Then, the LRT is defined as a criterion to test the null hypothesis of equality of main-plot and split-plot residual correlation coefficients, against the alternative hypothesis that they are different. Two different situations are considered:

i) The first situation examined is when both correlation coefficients come from two different quadratic-forms, with different df, $df_1 \neq df_2$, resulting in the following expressions for λ :

$$\lambda_{11} = \left[1-\left[\frac{\hat{\rho}-r_1}{1-\hat{\rho} r_1}\right]^2\right]^{\frac{df_1}{2}} \left[1-\left[\frac{\hat{\rho}-r_2}{1-\hat{\rho} r_2}\right]^2\right]^{\frac{df_2}{2}}$$

$$\text{for } r_2 df_1 + r_1 df_2 \neq 0$$

and

$$\lambda_{12} = (1-r_1^2)^{\frac{df_1}{2}}(1-r_2^2)^{\frac{df_2}{2}}$$

$$\text{for } r_2 df_1 + r_1 df_2 = 0, \qquad (7)$$

where $\hat{\rho}$ is defined as in (4).

ii) The second situation examined is when the ρ 's come from different quadratic-forms, with the same df, $df_1 = df_2 = df$, resulting in the following expressions for λ :

$$\lambda_{21} = \left[\left[1-\left[\frac{\hat{\rho}-r_1}{1-\hat{\rho} r_1}\right]^2\right]\left[1-\left[\frac{\hat{\rho}-r_2}{1-\hat{\rho} r_2}\right]^2\right]\right]^{\frac{df}{2}}$$

$$\text{for } r_1 + r_2 \neq 0$$

and

$$\lambda_{22} = [(1 - r_1^2)(1 - r_2^2)]^{\frac{df}{2}}$$

$$\text{for } r_1 + r_2 = 0, \qquad (8)$$

where

$$\hat{\rho} = \frac{(r_1 r_2 + 1) - [(r_1^2 - 1)(r_2^2 - 1)]^{\frac{1}{2}}}{(r_1 + r_2)}$$

This above equality of df is an unlikely situation to occur in practice, but it should be considered.

Comparison between the LRT and U-Fisher criteria to test the equality of correlation coefficients

The problem initially investigated in this section concerned a comparison between the statistic LRT [defined in (7) or (8)], which is a criterion to test the null hypothesis [(1), say] that the main-plot and split-plot variance-covariance matrices have the same population correlation coefficient against a known test statistic.

$$U = \frac{Z_1 - Z_2}{\sqrt{\frac{1}{N_1 - 3} + \frac{1}{N_2 - 3}}}$$

, introduced by Fisher (1921).

Fisher (1915) has noticed that the distribution of the sampling correlation coefficient r becomes extremely skew when the sample size N is small and the corresponding population parameter is near to either -1 or 1. He indicated a transformation of r,

$$Z = \frac{1}{2} \ln \left[\frac{1 + r}{1 - r}\right], \qquad (9)$$

which tends to normality much faster than r, even for moderate sample size N, with mean

$$E(Z) = \frac{1}{2} \ln \left[\frac{1 + \rho}{1 - \rho}\right] + \frac{\rho}{2(N-1)} \cong \frac{1}{2} \ln \left[\frac{1 + \rho}{1 - \rho}\right]$$

(when the sample size increases), and variance

$$V(Z) = \frac{1}{N-1} + \frac{4-\rho^2}{2(N-1)^2} \cong \frac{1}{N-3}$$

for small ρ.

The transformation of r may also be used to examine whether two observed correlation coefficients differ significantly. Let these be, say, N_1 and N_2 pairs of observations from which the correlation coefficients r_1 and r_2 have been calculated. The test hypothesis is $\rho_1 = \rho_2 = \rho$ or $\tau_1 = \tau_2 = \tau$,

$$\tau = \frac{1}{2} \ln \left[\frac{1 + \rho}{1 - \rho}\right].$$

If the hypothesis is true, $Z_1 = \frac{1}{2} \ln \left[\frac{1 + r_1}{1 - r_1}\right]$ is normally distributed about τ, with mean

$$\frac{1}{2} \ln \left[\frac{1 + \rho}{1 - \rho}\right] + \frac{\rho}{2(N_1 - 1)}$$

and variance $\frac{1}{N_1 - 3}$; $Z_2 = \frac{1}{2} \ln \left[\frac{1 + r_2}{1 - r_2}\right]$ is normally distributed about the same τ, with mean

$$\frac{1}{2} \ln \left[\frac{1 + \rho}{1 - \rho}\right] + \frac{\rho}{2(N_2 - 1)}$$

and variance $\frac{1}{N_2 - 3}$.

The statistic $Z_1 - Z_2$ is, following Rao (1968), distributed about the mean

$$\frac{\rho}{2(N_1 - 1)} - \frac{\rho}{2(N_2 - 1)}$$

and variance $\frac{1}{N_1-3} + \frac{1}{N_2-3}$

If the samples are not small, the test statistic.

$$U = \frac{Z_1 - Z_2}{\sqrt{\frac{1}{N_1-3} + \frac{1}{N_2-3}}} \qquad (10)$$

is approximately normally distributed with parameters 0 and 1. By adopting transformation (9), a test which is both easy to apply and completely independent of the unknown value of ρ, is obtained.

The best estimator of ρ, as suggested by Fisher (1915), is based on a weighting of Z_1 and Z_2 inversely proportional to the approximate values of their sampling variance, namely

$$\hat{\rho}_F = \tanh\left[\frac{(N_1-3)Z_1 + (N_2-3)Z_2}{(N_1-3) + (N_2-3)}\right] \qquad (11)$$

where tanh is the abbreviation for hyperbolic tangent.

Two different estimators of the population parameter ρ were found: the ML estimator ($\hat{\rho}$) defined in (4), and the Fisher estimator ($\hat{\rho}_F$) defined in (11). It is nevertheless of some interest to examine the logical basis for the choice between the two estimators.

Two situations should be considered:

i) the first situation is when the sampling residual correlations have been calculated from two bivariate samples, with the same degrees of freedom $df_1 = df_2 = df$.
From (4) we have

$$\hat{\rho} = \frac{(r_1r_2 + 1) - [(r_1r_2-1)^2 - (r_1-r_2)^2]^{\frac{1}{2}}}{(r_1 + r_2)}$$

and by applying (9) into (4) we found, for the ML estimator of ρ, the following expression

$$\hat{\rho} = \tanh\left[\frac{Z_1 + Z_2}{2}\right] \qquad (12)$$

which is identical to the Fisher estimator of the parameter ρ in expression (11). If the same substitution (9) is made in equation (8), it gives

$$\lambda_{21} = [\mathrm{sech}^2(Z-Z_1)$$

$$\mathrm{sech}^2(Z-Z_2)]^{df/2} =$$

$$\left[\mathrm{sech}\left[\frac{Z_1-Z_2}{2}\right]\right]^{2df}$$

and extracting the (2 df)th root gives

$$\lambda_{21}^{1/2df} = -\ \mathrm{sech}\left[\frac{Z_1-Z_2}{2}\right]$$

It follows that, if the LRT criterion used to test the hypothesis about equality of correlation coefficients is

$$Z_1 - Z_2 = -2\ \mathrm{sech}^{-1}\ \lambda_{21}^{1/2df} \qquad (13)$$

it is identical to that of Fisher ($Z_1 - Z_2$, say) for testing the same hypothesis, when the sampling residual correlations are calculated from two bivariate samples with the same df.

ii) the second situation is when the sampling residual correlations is calculated from two bivariate samples with different df, $df_1 \neq df_2$.

From (4) it was found, after applying (9) and some algebraic manipulations, that:

$$= \frac{\dfrac{\tanh Z_1 . \tanh Z_2 + 1}{\tanh Z_1 . \tanh Z_2 - 1} - \left[1 - \dfrac{4 df_1 df_2 (\tanh Z_1 - \tanh Z_2)^2}{(df_1 + df_2)^2 (\tanh Z_1 \tanh Z_2 - 1)} \right]^{\frac{1}{2}}}{\dfrac{2(df_1 \tanh Z_2 + df_2 \tanh Z_1)}{(df_1 + df_2)(\tanh Z_1 . \tanh Z_2 - 1)}}$$

which does not correspond to (11) suggested by Fisher.

The accuracy of $\hat{\rho}$ and $\hat{\rho}_F$ approximations of ρ, for widely values of r_1 and r_2 is verified through a simulated study, as can be in the next section.

Simulation study

A computer program written in SAS IML, using a pseudo-random number generator as specified in SAS (1979), was developed in order to evaluate the power of -2ln(λ) under an alternative hypothesis.

This program generates two independent sampling main-plot and split-plot variance-covariance matrices, needed for the simulation study from independent bivariate populations, with specified values of the parameters

$$\sigma^2_{11}, \ \sigma^2_{12}, \ \rho_1, \ \sigma^2_{21}, \ \sigma^2_{22}$$

and ρ_2 as follows:

If $X = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_p \end{bmatrix}$ is distributed such that the X_i,

$i = 1, 2, ..., p$

[p being the number of variables, and X_i being independent standardized Normal variates $X_i \sim N(0,1)$], then X is said to have a standardized Multivariate Normal distribution $X \sim MN(0,Ip)$.

If A is any p x p matrix of rank p, and μ is any p-vector, then $Y = \mu + AX$ has a Multivariate Normal distribution $Y \sim MN(\mu, \Sigma)$, with vector mean μ and variance-covariance matrix $\Sigma = A A'$. A is lower triangular and it is obtained by the Cholesky's decomposition of the population matrix Σ.

Suppose we consider, without loss of generality, that $\mu = 0$; then, $Y = AX \sim MN(O,\Sigma)$. To obtain two independent sample variance-covariance matrices, representing the main-plot and split-plot with some specified correlation structure, we need to generate two independent identically distributed N(0,1) for each sample and calculate the corresponding observation matrix $y = Ax$. The x and y sampling matrices are Bivariate Normally distributed, that is $x \sim BN(0,I)$ and $y \sim BN(0,\Sigma)$. Thus, $s = y'y/df$ is the corresponding bivariate estimate of the population variance-covariance matrix Σ.

To develop the likelihood Ratio Test λ in (7) or (8), two estimated variance-covariance matrices, corresponding to the main-plot and split-plot are needed, with their six statistics

$s^2_{11}, \ s^2_{12}, \ r_1, \ s^2_{21}, \ s^2_{22}, \ r_2$ and the respective degress of freedom. Let σ^2_{11}, σ^2_{12}, σ^2_{21}, and σ^2_{22} be the population variances chosen to have, as usual, the split-plot residual variances smaller than the correspondent main-plot variances and, fixed as below:

$$MP = \begin{bmatrix} \sigma^2_{11} = 15 & \rho_1 \sigma_{11} \sigma_{12} \\ \rho_1 \sigma_{11} \sigma_{12} & \sigma^2_{12} = 4 \end{bmatrix} \text{ and}$$

$$SP = \begin{bmatrix} \sigma^2_{21} = 5 & \rho_2 \sigma_{21} \sigma_{22} \\ \rho_2 \sigma_{21} \sigma_{22} & \sigma^2_{22} = 2 \end{bmatrix} \quad (14)$$

leading ρ_1 to vary from -0.9 to 0 by 0.1, and ρ_2 from -0.9 to 0.9 by 0.1. Three sets of df are considered, that is $df_1 = 3$ $df_2 = 3$, $df_1 = 3$ $df_2 = 16$ and $df_1 = 12$ $df_2 = 75$.

The first combination ($df_1 = 3$ and $df_2 = 3$) corresponds to the situation when no blocks are defined in the experiment, and there are three levels of the main plot factor and two levels of the split-plot factor. The second combination ($df_1 = 3$ and $df_2 = 16$) corresponds to a routine experiment with two blocks, four levels of main-plot factor and five levels of split-plot factor. Finally, the third combination ($df_1 = 12$ and $df_2 = 75$) corresponds to an extreme experiment with four blocks, five main-plot levels and six split-plot levels.

For each experimental situation corresponding to all possible combinations of the population coefficients, and for each set of df, one thousand replications have been generated, and the statistic -21n (λ) calculated. The sample size of one thousand observations was chosen, expecting that the approximate values of power be reasonably close to the true values. Requiring the standard error of our estimator to be smaller than 2%, the sample size N must be greater than 650.

Under the alternative hypothesis, the Power of the test is the probability of rejecting $H_0 : \rho_1 = \rho_2 = \rho$ when it is false. That is the number of cases when $-21n(\lambda) > 3.84$, where the value 3.84 is the critical value of the Chi-Square central distribution, with one df at 5% level of probability. Following the range of variation for correlation coefficients ρ_1 from -0.9 to 0, and for ρ_2 from -0.9 to 0.9, with the population variances fixed as in (14), the power values were calculated and presented in three dimensional plots, as shown in Fig. 1, 2 and 3.

Using the same procedure as above, the power of the U test, which is the probability of rejecting H_0: $\rho_1 = \rho_2 = \rho$ when it is false, that is, the number of cases where the absolute value of U is greater than 1.96 (the critical value of the standardized Normal distribution at 5% level of probability), was simulated.

Figures 4, 5 and 6 show the three dimensional plots for the observed power of the U test statistic, considering the range of variation of ρ_1 from -0.9 to 0.0, and of ρ_2 from -0.9 to 0.9, with the same fixed population variances as specified in (14). The same seed for each replication used for the simulation of the observed power for LRT [defined in (7) and (8)] was kept, with the intention of obtaining a base for comparison between the powers of these two different criteria.

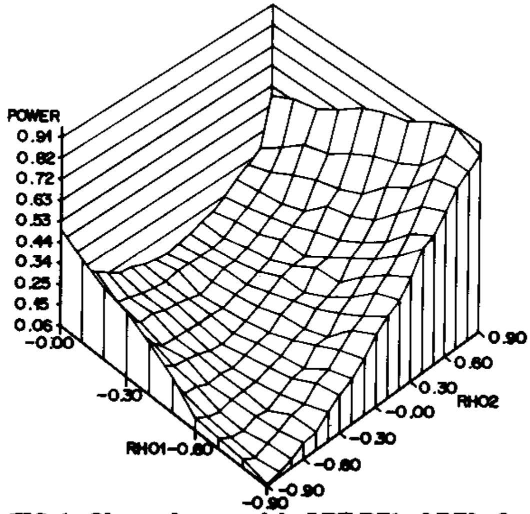

FIG. 1. Observed power of the LRT-DF1=3 DF2=3.

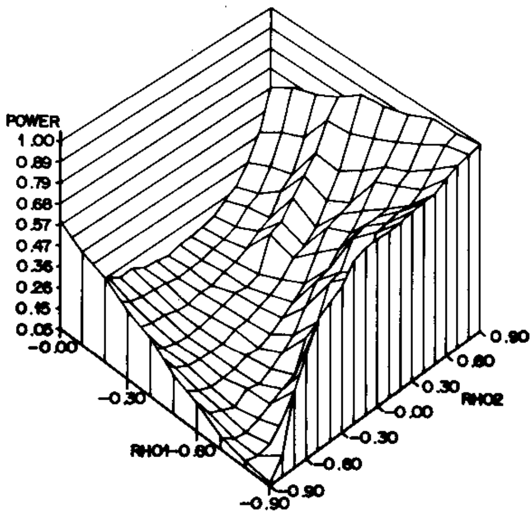

FIG. 2. Observed power of the LRT-DF1=3 DF2=16.

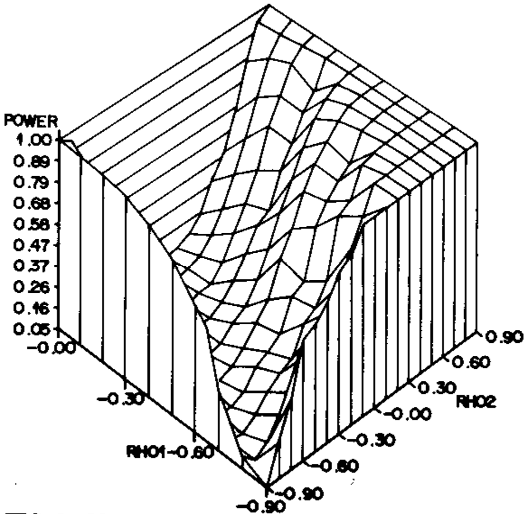

FIG. 3. Observed power of the LRT-DF1=12 DF2=75.

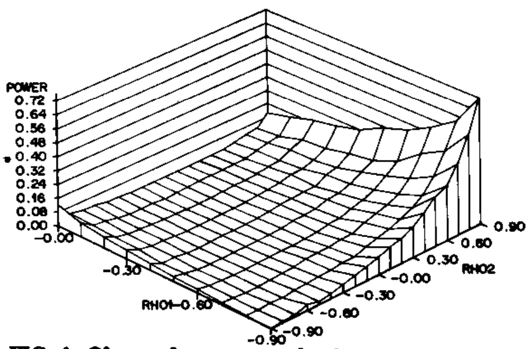

FIG. 4. Observed power of the U-test-DF1=3 DF2=3.

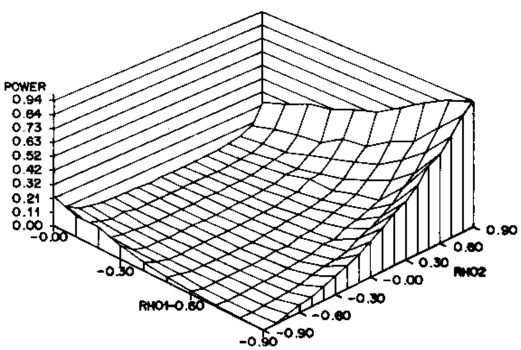

FIG. 5. Observed power of the U-test-DF1=3 DF2=16.

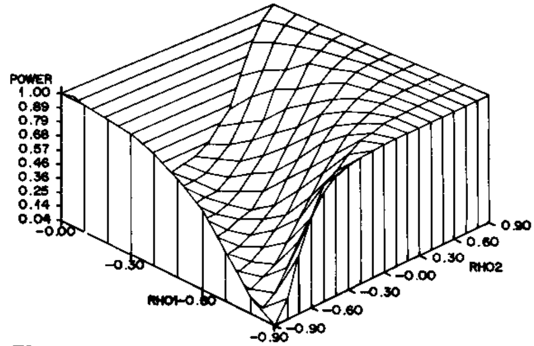

FIG. 6. Observed power of the U-test-DF1=12 DF2=75.

The visual comparison between the power of the U test in Fig. 4 and 5 with, either the observed power of the LRT statistic in Fig. 1 and 2, shows the loss of power for the U test to detect variation of the correlation coefficient. The wrong assumption that the U test statistic is normally distributed, for small and moderate sets of df, may be responsible for this loss of power. With increasing sample sizes (Fig. 6 with 3, say), the loss of power for the U test statistic is little, if any, since the data may be considered as normally distributed. For a reasonable sample size, such as the second combination of df (df_1 = 3 and df_2 = 16, say), the LRT statistic has greater probability of not rejecting H_0 when it is false than the U statistic. This is clearly visualized when comparing Fig. 2 with 5.

As Fisher (1921) pointed out, there is a small bias introduced in the estimate of ρ when the sample size is not large, due to the fact that the term $\frac{\rho}{2(N-1)}$ in the mean value of Z is neglected. He suggested to correct the bias by subtracting it from the value of Z.

Rao (1968) introduced a different procedure to test the homogeneity of correlation coefficients when the sample sizes are not large and not nearly equal, taking into account the bias in the best estimate of ρ. Since Z, defined in (9), is normally distributed with mean $\frac{1}{2}$ ln

$$\left[\frac{1+\rho}{1-\rho}\right] + \frac{\rho}{2(N-1)} \text{ and variance } \frac{1}{N-3},$$

he defined the statistics:

$$S = \sum_{i=1}^{2} (N_i - 3) \left[\frac{1}{1-\rho^2} + \frac{1}{2(N_i-1)}\right] \left[Z_i - \frac{1}{2} \ln \frac{1+\rho}{1-\rho} - \frac{\rho}{2(N_i-1)}\right]$$

with S being the score of ρ obtained from k samples, and the information

$$\varphi = E(S^2) = \sum_{i=1}^{2} (N_i - 3) \left[\frac{1}{1-\rho^2} + \frac{1}{2(N_i-1)}\right]^2$$

Using (11) as the starting point value of ρ, he also defined an additive correction $\delta\rho$ to ρ as $\delta\rho = \frac{S}{\varphi}$, and suggested the repetition of this process until the value for the correction becomes negligible, resulting in the best estimator $\hat{\rho}_R$ of ρ.

Then, the statistic H is defined to test the homogeneity of correlation coefficients from two samples, as follows:

$$H = \sum_{i=1}^{2} (N_i - 3) \left[Z_i - \frac{1}{2} \ln \frac{1+\hat{\rho}_R}{1-\hat{\rho}_R} - \frac{\hat{\rho}_R}{2(N_i-1)}\right]^2, \quad (15)$$

with H following a central Chi-Square distribution with one df.

The same simulation program, with the seeds for each replication kept constant, and varying the parameters ρ_1 from -0.9 to 0.0 by 0.3, and ρ_2 from -0.9 to 0.9 by 0.3, was used to calculate the observed power of the H test statistic in (15). The results are presented in the three dimensional plots shown in Fig. 7, 8 and 9.

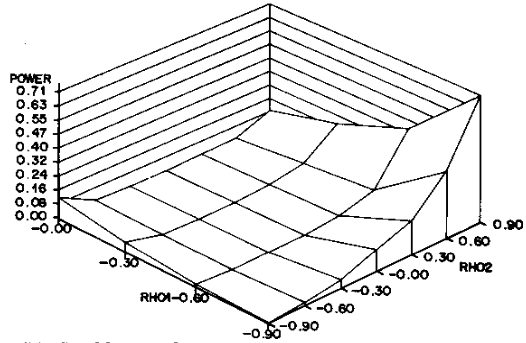

FIG. 7. Observed power of the H-test-DF1=3 DF2=3.

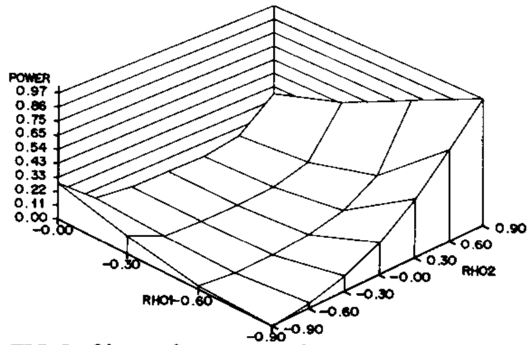

FIG. 8. Observed power of the H-test-DF1=3 DF2=16.

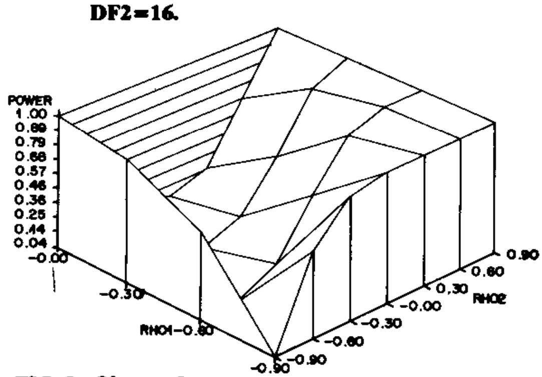

FIG. 9. Observed power of the H-test-DF1=12 DF2=75.

When comparing Fig. 4 with 7, and 6 with 9, the power of the U and H test statistics may be considered the same, as we expected. That is due to the fact that for these combinations of df, the inclusion of the bias in the mean of Z is unimportant. When Fig. 5 is compared with

Fig. 8, for moderate and different df, the power of the H test statistic has improved, but this improvement is not enough to have any distinguished difference between the powers. However, both powers are smaller when compared with the observed power of the LRT in Fig. 2. The little power of the H test statistic to detect variation of the correlation coefficient, was also verified by Dear & Mead (1984).

In summary, the results of this section indicate that the LRT might be used as a general approach for testing the null hypothesis of equality of correlation coefficients. It demonstrates greater power at least for the experimental situations considered in our study.

Assessment of the LRT and H test statistics using experimental data sets

Since the bivariate populations from where two bivariate samples are extracted, and the sampling residual correlations calculated are not as controlled in the experimental sets as in the simulation sets, some additional information on the performance of the criterion used to test the homogeneity of correlation is expected.

Table 1 gives the results for the LRT and H criteria used so far in testing if two bivariate population have the same correlation parameter for the eight experimental data sets described in Carvalho (1988), with the respective sampling residual correlations and df. These data sets are from split-plot experiments, and come from different Agricultural Research Institutes.

From Table 1, when data sets two and four are considered, no significant results are found for both criteria, when compared with 3.84, which is the critical value for the Chi-Square distribution with one df. This non-significance enables us to conclude that there is no reason to reject the null hypothesis about equality of correlations.

For all other data sets, the null hypothesis is always rejected at the same level of probability (5%) for the LRT statistic, demonstrating its value when compared with H test statistic (as verified from the simulated results). The H test statistic does not reject the null hypothesis in seven out of eight experimental data sets used in this Table.

CONCLUSIONS

1. As observed when comparing the MLE of the correlation coefficient ($\hat{\beta}$) with the Fisher estimator ($\hat{\rho}_F$) for the same parameter, the weights (df_1 and df_2) are more efficiently used in forming the LRT statistic than the H test statistic. As can be seen from Table 1, when the sample sizes are very different (as in data sets 5, 6, 8 and 9), the LRT statistic has values which

TABLE 1. Two different criteria, LRT and H, to test homogeneity of correlation coefficients using experimental data sets.

Data Set	df_1	df_2	r_1	r_2	LRT	H
2	12	15	-0.5602	-0.0415	2.2988	1.9626
3	6	16	0.9326	0.0152	10.1933	8.4098
4	9	36	0.1803	-0.1096	0.6108	0.5043
5	3	84	-0.8323	0.1669	4.3234	1.9062
6	3	84	-0.9156	-0.1242	4.7040	1.9757
7	15	18	0.3277	-0.3575	4.0861	3.6622
8	3	48	-0.8961	-0.0484	4.4693	1.8977
9	3	48	-0.6857	0.4292	3.9274	1.8198

are larger than the values of the H test statistic; when the sample sizes are nearly equal (as in data sets 2 and 7), both statistics have roughly the same values. Table 1 also indicates that the difference $r_1 - r_2$ plays an important role in the construction of both criteria. When the sample sizes are not sufficiently large and not nearly equal (as in data sets 3 and 4), the probability of rejecting H_0 increases proportionally to the increase of the difference $r_1 - r_2$.

2. In summary, it was concluded that, where experimental data sets were used for comparing both criteria LRT and H, agree with the simulated results. Consequently, there is no reason to expect, by using less controlled sets, to obtain results for the LRT criterion which could "mask" our choice of it as the best statistic for testing the null hypothesis of equality of correlation coefficients.

REFERENCES

CARVALHO, J.R.P. **Bivariate analysis intercropping with two levels of error variation.** Reading: Reading University, 1988. 245p. Ph.D. Thesis.

DEAR, K.B.G.; MEAD, R. **Testing assumptions and other topics in bivariate analysis.** Reading: [s.n.], 1984. 25p. (Statistics in Intercropping. Technical Report n.2. Department of Applied Statistics).

FISHER, R.A. Frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population. **Biometrika,** London, v.10, p.507-521, 1915. Published also in BENNETT, J.H. (Ed.). Collected papers of R.A. Fisher. Adelaide: The University of Adelaide, 1971. v.1.

FISHER, R.A. On the "Probable Error" of a coefficient of correlation deduced from a small sample. **Metron,** v.1, p.3-32, 1921. Published also in BENNETT, J.E. (Ed.). Collected papers of R.A. Fisher. Adelaide: The University of Adelaide, 1971. v.1.

RAO, C.R. **Linear statistical inference and its applications.** New York: John Wiley & Sons, Inc., 1968. 438p.

STATISTICAL ANALYSIS SYSTEM INSTITUTE (Cary, EUA). **SAS User's guide.** Cary, 1979. 630p.