TEST IN MULTIVARIATE ANALYSIS

José A. Díaz-García and Francisco José Caro-Lopera

Comunicación Técnica No I-05-09/22-08-2005
(PE/CIMAT)

Test in Multivariate Analysis*

By José A. Díaz-García
Universidad Autónoma Agraria Antonio Narro
and
Francisco José Caro-Lopera
Universidad de Antioquia

This work proposes a method for finding the critical values of several linear test hypothesis criteria in the context of the general multivariate linear model by using the existing tables in the literature. The exact distribution of a certain criterion announced as new by Olson (1974), but really defined by Wilks (1932), is studied. Some errors in two of the criteria obtained by Wilks (1932) are detected and corrected. The moments and the exact distribution of Dempster's test criterion are found. At the end, an example of the literature determines all of the criteria and their test.

1. Introduction. Consider the general multivariate linear model

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon} \tag{1}$$

where: $\mathbf{Y} \in \Re^{n\times p}$ is the matrix of the observed values; $\boldsymbol{\beta} \in \Re^{q\times p}$ is the parameter matrix; $\mathbf{X} \in \Re^{n\times q}$ is the design matrix or the regression matrix of rank $r \leq q$; $\boldsymbol{\epsilon} \in \Re^{n\times p}$ is the error matrix which has a matrix variate normal distribution, specifically $\boldsymbol{\epsilon} \sim \mathcal{N}_{n\times p}(\mathbf{0}, \mathbf{I}_n \otimes \boldsymbol{\Sigma})$; $\otimes$ denotes the Kronecker product; and $\boldsymbol{\Sigma} \in \Re^{p\times p}$, $\boldsymbol{\Sigma} > \mathbf{0}$. For this model, we want to test the hypothesis

$$H_0 : \mathbf{C}\boldsymbol{\beta}\mathbf{M} = \mathbf{0} \text{ versus } H_a : \mathbf{C}\boldsymbol{\beta}\mathbf{M} \neq \mathbf{0} \tag{2}$$

where $\mathbf{C} \in \Re^{\nu_H\times q}$ of rank $\nu_H \leq r$ and $\mathbf{M} \in \Re^{p\times g}$ of rank $g \leq p$. As in the univariate case, the matrix $\mathbf{C}$ concerns to the hypothesis among the elements of the parameter matrix columns, while the matrix $\mathbf{M}$ allows hypothesis among the different response parameters. The matrix $\mathbf{M}$ plays a role in profile analysis, for example; in ordinary hypothesis test it is taken to be the identity matrix, $\mathbf{M} = \mathbf{I}_p$.

Let $\mathbf{S}_H$ be the matrix of sums of squares and sums of products due to the hypothesis and let $\mathbf{S}_E$ be the matrix of sums of squares and sums of products due to the error, and both defined like this

$$\begin{aligned}\mathbf{S}_H &= (\mathbf{C}\widetilde{\boldsymbol{\beta}}\mathbf{M})'(\mathbf{C}(\mathbf{X}'\mathbf{X})^{-}\mathbf{C}')^{-1}(\mathbf{C}\widetilde{\boldsymbol{\beta}}\mathbf{M})\\ \mathbf{S}_E &= \mathbf{M}'\mathbf{Y}'(\mathbf{I}_n - \mathbf{X}\mathbf{X}^{-})\mathbf{Y}\mathbf{M}\end{aligned}$$

respectively; where $\widetilde{\boldsymbol{\beta}} = \mathbf{X}^{-}\mathbf{Y}$ and $\mathbf{X}^{-}$ is any generalised inverse of $\mathbf{X}$ such that $\mathbf{X} = \mathbf{X}\mathbf{X}^{-}\mathbf{X}$. Besides, under a null hypothesis, $\mathbf{S}_H$ has a g-dimensional Wishart distribution with ν_H degrees of freedom and parameter matrix $\mathbf{M}'\boldsymbol{\Sigma}\mathbf{M}$, i.e. $\mathbf{S}_H \sim \mathcal{W}_g(\nu_H, \mathbf{M}'\boldsymbol{\Sigma}\mathbf{M})$; similarly $\mathbf{S}_E \sim \mathcal{W}_g(\nu_E, \mathbf{M}'\boldsymbol{\Sigma}\mathbf{M})$[1].

---

* *AMS 2000 Subject Classification*: Primary 62E15: Secondary 62H10
Key word and phrases: Multivariale linear model, likelihood ratio test, Wilks's criteria, Olson's criterion, Gnanadesikan's criterion, Pillai's citerion, Roy's criteria, Anderson's criterion, Pillai's trace, Lawlwy-Hotelling trace, Dempster's criterion, new Pillai's criteria.

[1] For this time we will take $\mathbf{M} = \mathbf{I}_p$. But, if this is not the case under consideration, the following results of the paper just appear by substituting p by g.

Now, let $\lambda_1, \cdots, \lambda_s$ be the $s = \min(\nu_H, g)$ non null eigenvalues of the matrix $\mathbf{S}_H \mathbf{S}_E^{-1}$ such that $0 < \lambda_s < \cdots < \lambda_1 < \infty$ and let $\theta_1, \cdots, \theta_s$ be the s non null eigenvalues of the matrix $\mathbf{S}_H(\mathbf{S}_H + \mathbf{S}_E)^{-1}$ with $0 < \theta_s < \cdots < \theta_1 < 1$; here we note $\lambda_i = \theta_i/(1-\theta_i)$ and $\theta_i = \lambda_i/(1+\lambda_i)$, $i = 1, \cdots, s$. Various authors have proposed a number of different criteria for test the hypothesis (2). But as it is known, see for example Kres (1983), all the test can be expressed in function of the eigenvalues $\lambda's$ or $\theta's$. In our experience, a reason for which many of these test statistics are not used is due to lack and/or inaccessibility of tables for the respective critical values. In this work the three test statistics proposed by Wilks (1932) are studied after correcting some errors in the published density functions. We emphasize that two of those statistics were proposed as new by Roy *et al.* (1971) (U-statistics) and Olson (1974) (S-statistics). Besides we show how to obtain the critical values for the U-statistics starting form the tables of Wilks's Λ statistics. The density of the S-statistics is derived by three different methods. The exact distribution of another test criterion proposed by Pillai (1955) is found. The moments and the exact distribution for the Dempster statistics are gotten. At the end, this work solves a problem of the literature computing all the published test statistics studied including the S-statistics, also we propose the way for finding the critical values for the remaining test criteria. We will begin from the published tables in literature; the critical values of the S-statistics are found by using the tables here derived.

2. Wilks's criteria. Unfortunately, there is not homogeneity in the symbol of the test statistics, moreover, some of them were renamed creating more confusion. For example, the most known statistics of Wilks is the W, but in the literature is defined as Wilks's Λ, however Anderson (1982, p. 299) denoted it by U, but Wilks (1932) named with that symbol another of the statistics. In order to avoid any confusion in notation we return to the original notation of Wilks (1932) and we define the three criteria in this way:

$$\begin{aligned}
\Lambda = W &= \frac{|\mathbf{S}_E|}{|\mathbf{S}_E + \mathbf{S}_H|} = \prod_{i=1}^{s} \frac{1}{1+\lambda_i} = \prod_{i=1}^{s} (1-\theta_i) && \text{Wilks (1932, p. 485)}\\
U &= \frac{|\mathbf{S}_H|}{|\mathbf{S}_E + \mathbf{S}_H|} = \prod_{i=1}^{s} \frac{\lambda_i}{1+\lambda_i} = \prod_{i=1}^{s} \theta_i && \text{Wilks (1932, p. 482)}\\
S &= \frac{|\mathbf{S}_H|}{|\mathbf{S}_E|} = \prod_{i=1}^{s} \lambda_i = \prod_{i=1}^{s} \frac{\theta_i}{(1-\theta_i)} && \text{Wilks (1932, p. 486)}
\end{aligned}$$

Curiously the U-statistics was proposed as a new statistics in the literature by Roy *et al.* (1971, last paragraph p. 72) with the same original notation of Wilks (1932). Similarly, the third statistics was proposed as new by Olson (1974); here we used that notation.

Moreover, Wilks (1932) proposed integral expressions for the densities of the three statistics, but even when the general expression for the W and the U statistics are correct (Wilks (1932, eq. (5), p. 475)), the density for W (Wilks (1932, eq. (35), p. 486)) is wrong. Maybe this fact explains the inconsistences of some particular expressions for the densities of W published by Wilks (1935), such as it is corroborated by Consul (1966) when the results are compared with the results obtained by Anderson (1982, p. 308). The error in the density of W in Wilks (1932) is the exponent of the term $(v_1 v_2 \cdots v_{n-1})$ which appears as $(p-2)/2$ and it should be$(p-3)/2$. By using our notation

Wilks's notation	Our notation
N	$\nu_H + \nu_E + 1$
p	$\nu_H + 1$
n	p

where $\nu_E = N - p$. The right density of W is

$$\begin{aligned} f_W(w) &= \frac{\Gamma_p[(\nu_H+\nu_E)/2]}{\Gamma_p[\nu_H/2]\Gamma_p[\nu_E/2]} w^{(\nu_E-p+1)/2-1}(1-w)^{p\nu_H/2-1}\int_0^1\cdots\int_0^1\prod_{i=1}^{p-1} v_i^{(\nu_H-2)/2} \\ &\times \prod_{i=1}^{p-1}(1-v_i)^{(p-i)\nu_H)/2-1} \\ &\times [1-v_1(1-w)]^{-(\nu_H-1)/2}[1-\{v_1+v_2(1-v_1)\}(1-w)]^{-(\nu_H-1)/2} \\ &\times [1-\{v_1+v_2(1-v_1)+v_3(1-v_1)(1-v_2)\}(1-w)]^{-(\nu_H-1)/2}\cdots \\ &\times [1-\{v_1+v_2(1-v_1)+\cdots+v_{p-1}(1-v_1)(1-v_2)\dots(1-v_{p-2})\}(1-w)]^{-(\nu_H-1)/2} \\ &\times dv_1dv_2\dots dv_{p-1}, \end{aligned} \tag{3}$$

where $\Gamma_p[a] = \pi^{p(p-1)/4}\prod_{i=1}^p \Gamma(a-(i-1)/2)$ is the multivariate gamma function, see Muirhead (1982, pp. 61-62). Note the distribution of U can be found as a function of the distribution of W (and vice versa), just change the rules of ν_H and ν_E. This is, by making the transformation

$$(\nu_H,\nu_E) \to (\nu_E,\nu_H)$$

Observe in Wilks's notation the density of U can be obtained form the density of W by making the transformation

$$(N-p, p-1) \to (p-1, N-p)$$

here it is detected again the error above-mentioned in the density of W. This equivalency can be easily seen by replacing particular values of ν_H and ν_E in the densities of U and W, derived by Hsu (1940) form the joint density of the eigenvalues of the θ's. For proving this equivalency, let us denote the density of $\mathbf{\Theta} = (\theta_1,\dots,\theta_s)'$ by $p(\mathbf{\Theta}; s, m, h)$, where m and h are functions of ν_H, ν_E, p, see Muirhead (1982, pp. 451 and 454-455) or Díaz-García and Gutiérrez-Jáimez (1997). See also Srivastava & Khatri (1979, Theorem 3.6.2, p. 93)[2].

Now, it is known that $\Lambda \sim$ Wilks's Λ. If $\mathbf{\Theta}^* = ((1-\theta_1),\dots,(1-\theta_s))' = (\theta_1^*,\dots,\theta_s^*)'$, the distribution of $\mathbf{\Theta}^*$is the same as that of $\mathbf{\Theta}$, by interchanging m and h, see Nanda (1948, Section 5). Then,

$$\Lambda^* = \prod_{i=1}^s \theta_i^* \sim \text{Wilks's } \Lambda, \quad \text{with } m \text{ and } h \text{ interchanged}$$

but note $\Lambda^* = U$. Therefore,

$$U \sim \text{Wilks's } \Lambda, \quad \text{with } m \text{ and } h \text{ interchanged}$$

Given that $m = (|\nu_H - p| - 1)/2$ and $h = (\nu_E - p - 1)/2$, see Nanda (1948), Pillai (1955) or Rencher (1995, p. 165). But $\nu_E > p$ and $\nu_H \geq p$ or $\nu_H < p$, then the interchange of m and h is equivalent to the interchange of ν_H and ν_E.

In summary,

Theorem 2.1. *The distribution of U-statistics, can be obtained from the distribution of the Λ-statistics, by interchanging ν_H and ν_E; this is*

$$U_{\nu_H,\nu_E} \stackrel{d}{=} \Lambda_{\nu_E,\nu_H}$$

where $\stackrel{d}{=}$ denotes equally distributed.

[2]But first note some minor errors appear there: the exponent of π must be $p^2/2$ instead of $p/2$ and the exponent of the l_i should be $(n_2 - p - 1)/2$ in place of $(n_2 - p - 1)$.

Note, in general, all the distributions of the test statistics and the tabulation of the correspondents critical values are assuming that $p \leq \nu_H$; if $p > \nu_H$ the associated densities and their respective critical values can be obtained by making the following transformations in the parameters, see Muirhead (1982, eq. (7), p. 455), Srivastava & Khatri (1979, p. 96) or Rencher (1995, p. 167),

$$(p, \nu_H, \nu_E) \rightarrow (\nu_H, p, \nu_E + \nu_H - p)$$

3. Third Wilks's statistics. ***S*-statistics.** Wilks's S-statistics have been used rare, maybe because its exact and asymptotic distribution have not been derived and of course no tables of its critical values have been constructed, except the Table H in Olson (1973) where the critical values were obtained via Monte Carlo. Recently that statistics have been used in the context of the sensibility analysis in regression, see Díaz-García *et al.* (2005). In fact, Wilks never found its particular distribution. But the k-moments were derived and a suggestion for determine its distribution was given starting from general expressions eqs. (13) and (16), see Wilks (1932). However the equation (16) contains two errors; the first one we think is a typographic error and the second one appeared when the integration respect to the θ's (in Wilks's notation) where carried out. In notation of Wilks (1932), the errors in the expression (16) are respectively

1. typographic error: the exponent of $(1 + \psi)$ must be$-((M + N)/2 - n)$ instead of $-((M + n)/2 - n)$, and
2. the argument in the gamma function of the numerator should be $(M + N)/2 - i$ in place of $(M + N - 1 - i)/2$

Then, the right density function of the S-statistics is (in our notation)

$$\begin{aligned} f_S(s) &= \frac{\pi^{p(p-1)/2} \prod_{i=1}^{p} (\Gamma[(\nu_H + \nu_E)/2 - i + 1])}{\Gamma_p[\nu_H/2]\Gamma_p[\nu_E/2]} \ s^{(\nu_H - p - 1)/2} \ (1 + s)^{-((\nu_H + \nu_E)/2 - p + 1)} \\ &\times \int_0^1 \cdots \int_0^1 \left\{ 1 - \frac{\prod_{i=1}^{p-1}(1 - r_i)}{1 + s} - \left[1 - \prod_{i=1}^{p-1}(1 - r_i) - \frac{\prod_{i=1}^{p-1} r_i}{1 + s} \right] \right\}^{-((\nu_H + \nu_E)/2 - p + 1)} \\ &\times \prod_{i=1}^{p-1} [r_i(1 - r_i)]^{(\nu_H + \nu_E)/2 - (p+i)/2} dr_1 dr_2 \ldots dr_{(p-1)}, \ s > 0 \end{aligned}$$

thus we have that

$$\frac{1 - \prod_{i=1}^{p-1}(1 - r_i) - \frac{\prod_{i=1}^{p-1} r_i}{1+s}}{1 - \frac{\prod_{i=1}^{p-1}(1 - r_i)}{1+s}} < 1$$

and

$$\frac{\prod_{i=1}^{p-1}(1 - r_i)}{1 + s} < 1$$

for $r_i \in [0, 1]$ and $s > 0$. This allows to expand in a double series of powers the term between brace and later integrating term by term.

- For $p = 1$ we get

$$f_S(s) = \frac{\Gamma[(\nu_H + \nu_E)/2]}{\Gamma[\nu_H/2]\Gamma[\nu_E/2]} \ s^{(\nu_H - 2)/2} \ (1 + s)^{-(\nu_H + \nu_E)/2}, \ s > 0$$

- For $p = 2$ we have

$$f_S(s) = k_2\ s^{(\nu_H-3)/2} \int_0^1 [(1-r_1)s + r_1]^{-(\nu_H+\nu_E-2)/2} [r_1(1-r_1)]^{(\nu_H+\nu_E-3)/2} dr_1,\ s > 0$$

$$k_2 = \frac{\Gamma[(\nu_H+\nu_E)/2]\Gamma[(\nu_H+\nu_E-2)/2]}{\Gamma[\nu_H/2]\Gamma[(\nu_H-1)/2]\Gamma[\nu_E/2]\Gamma[(\nu_E-1)/2]}$$

- For $p = 3$ we obtain

$$f_S(s) = k_3\ s^{(\nu_H-4)/2} \int_0^1 \int_0^1 [(1-r_1)(1-r_2)s + r_1 r_2]^{-(\nu_H+\nu_E-4)/2} [r_1(1-r_1)]^{(\nu_H+\nu_E-4)/2}$$
$$\times [r_2(1-r_2)]^{(\nu_H+\nu_E-5)/2} dr_1 dr_2,\ s > 0$$

$$k_3 = \frac{\Gamma[(\nu_H+\nu_E)/2]\Gamma[(\nu_H+\nu_E-2)/2]\Gamma[(\nu_H+\nu_E-4)/2]}{\Gamma[\nu_H/2]\Gamma[(\nu_H-1)/2]\Gamma[(\nu_H-2)/2]\Gamma[\nu_E/2]\Gamma[(\nu_E-1)/2]\Gamma[(\nu_E-2)/2]}$$

- And for $p = 4$ we get

$$f_S(s) = k_4\ s^{(\nu_H-5)/2} \int_0^1 \int_0^1 \int_0^1 [(1-r_1)(1-r_2)(1-r_3)s + r_1 r_2 r_3]^{-(\nu_H+\nu_E-6)/2}$$
$$\times [r_1(1-r_1)]^{(\nu_H+\nu_E-5)/2} [r_2(1-r_2)]^{(\nu_H+\nu_E-6)/2} [r_3(1-r_3)]^{(\nu_H+\nu_E-7)/2} dr_1 dr_2,\ s > 0$$

$$k_4 = \frac{\Gamma[(\nu_H+\nu_E)/2]\Gamma[(\nu_H+\nu_E-2)/2]\Gamma[(\nu_H+\nu_E-4)/2]\Gamma[(\nu_H+\nu_E-6)/2]}{\Gamma[\nu_H/2]\Gamma[(\nu_H-1)/2]\Gamma[(\nu_H-2)/2]\Gamma[(\nu_H-3)/2]\Gamma[\nu_E/2]\Gamma[(\nu_E-1)/2]\Gamma[(\nu_E-2)/2]\Gamma[(\nu_E-3)/2]}$$

Alternatively, the exact distribution of the S-statistics can be determined via the approach of Hsu (1940), i.e. by the joint distribution of the λ's. In particular a simplified expression for $p = 2$ can be obtained as follows: from Muirhead (1982, pp. 451 and 454-455) we have that

$$f_{\lambda_1,\lambda_2}(\lambda_1,\lambda_2) = k\ (\lambda_1\lambda_2)^{(\nu_H-3)/2}[(1+\lambda_1)(1+\lambda_2)]^{-(\nu_H+\nu_E)/2}(\lambda_1-\lambda_2)$$

where

$$k = \frac{\pi\Gamma_2[(\nu_H+\nu_E)/2]}{\Gamma_2[\nu_H/2]\Gamma_2[\nu_E/2]};$$

if we define $S = \lambda_1\lambda_2$ and $R = (1+\lambda_1)(1+\lambda_2)$, then $dsdr = (\lambda_1-\lambda_2)d\lambda_1 d\lambda_2$. Thus

$$f_{S,R}(s,r) = k\ s^{(\nu_H-3)/2} r^{-(\nu_H+\nu_E)/2}$$

by integrating with respect to R ranging from $(1+\sqrt{s})$ to ∞ we get the density function

$$f_S(s) = k_2^*\ s^{(\nu_H-3)/2}\left(1+\sqrt{s}\right)^{-(\nu_H+\nu_E-2)},\ s > 0$$

with

$$k_2^* = \frac{2k}{(\nu_H+\nu_E-2)} = \frac{2\sqrt{\pi}\ \ \Gamma[(\nu_H+\nu_E)/2]\Gamma[(\nu_H+\nu_E-1)/2]}{(\nu_H+\nu_E-2)\Gamma[\nu_H/2]\Gamma[(\nu_H-1)/2]\Gamma[\nu_E/2]\Gamma[(\nu_E-1)/2]}$$

A third approach for deriving the distribution of the U-statistics is based on the Mellin transformation proposed by Consul (1966). From Wilks (1932, p. 486) (in our notation),

$$E(S^h) = \frac{\Gamma_p[\nu_H/2+h]\Gamma_p[\nu_E/2+h]}{\Gamma_p[\nu_H/2]\Gamma_p[\nu_H/2]}.$$

Then the exact density function of S is

$$f_S(s) = \frac{1}{\Gamma_p[\nu_H/2]\Gamma_p[\nu_H/2]}\frac{1}{2\pi\ i}\int_{c'-i\infty}^{c'+i\infty} s^{-h-1}\Gamma_p[\nu_H/2+h]\Gamma_p[\nu_E/2+h]dh$$

which, on putting $h+(\nu_E+1-p)/2=t$, we obtain

$$f_S(s) = \frac{s^{(\nu_E-1-p)/2}}{\Gamma_p[\nu_H/2]\Gamma_p[\nu_H/2]}\frac{1}{2\pi\ i}\int_{c-i\infty}^{c+i\infty} s^{-t}\Gamma_p[t-(\nu_E-\nu_H+1-p)/2]\Gamma_p[\nu_E-t+(1-p)/2]dt$$

4. Exact distribution of new Pillai criterion, $W^{(s)}$. Pillai (1955) proposed another test criterion and its approximated distribution, apart from the other three criteria exposed in that work. The new criterion is defined like this

$$W^{(s)} = 1 - V^{(s)}/s = \frac{s-\sum_{i=1}^{s}\theta_i}{s} = \frac{\sum_{i=1}^{s}(1-\theta_i)}{s}$$

where $V^{(s)}$ is the Pillai's statistics, see Rencher (1995, p. 166) or Seber (1984, p. 414), among many others. Then

$$sW^{(s)} = \sum_{i=1}^{s}(1-\theta_i) = \operatorname{tr}(\mathbf{S}_E(\mathbf{S}_E+\mathbf{S}_H)^{-1}).$$

Now, as in Section 2, a similar result for the exact distribution of the $W^{(s)}$-statistics can be derived

Exactly as before, if $\mathbf{\Theta}^* = ((1-\theta_1),\dots,(1-\theta_s))' = (\theta_1^*,\dots,\theta_s^*)'$, the distribution of $\mathbf{\Theta}^*$ is the same as that of $\mathbf{\Theta}$, by interchanging m and h

$$s{V^{(s)}}^* = \sum_{i=1}^{s}\theta_i^* \sim \text{Pillai's } V^{(s)}, \quad \text{with } m \text{ and } h \text{ interchanged}$$

and by noting that $s{V^{(s)}}^* = sW^{(s)}$, we have

$$sW^{(s)} \sim \text{Pillai's } V^{(s)}, \quad \text{with } m \text{ and } h \text{ interchanged}$$

In summary,

THEOREM 4.2. *The distribution of the $W^{(s)}$-statistics can be obtained from the distribution of $V^{(s)}$-statistics, by interchanging ν_H and ν_E. This is*

$$W^{(s)}_{\nu_H,\nu_E} \stackrel{d}{=} \frac{1}{s}\, V^{(s)}_{\nu_E,\nu_H}$$

Note, this behavior of the parameters can be seen in the approximated distributions of both statistics given in Pillai (1955, eqs. (5) and (6), respectively).

5. Exact distribution of the Dempster criterion. For the case of one or two samples Dempster (1958) and Dempster (1960) propose a non exact proof for testing the hypothesis (2). For the general case ($p>2$), Fujikoshi *et al.* (2004) propose the following statistics

$$T_D = (\operatorname{tr}\mathbf{S}_H)/(\operatorname{tr}\mathbf{S}_E)$$

which is termed Dempster trace criterion.

Dempster's criterion is rarely used, it is difficult to find a book at least refereing it. Perhaps, this is given because its exact and asymptotic distribution are in terms of the matrix of parameters $\mathbf{\Sigma}$.

Fujikoshi *et al.* (2004) derive asymptotic null and nonnull distributions of Dempster trace criterion when $n \to \infty$ and $p \to \infty$. They prove that

$$\frac{\widetilde{T}_D}{\sigma_D} \xrightarrow{d} \mathcal{N}(0,1), \tag{4}$$

where $\xrightarrow{d}$ denotes convergence in distribution, and

$$\widetilde{T}_D = \sqrt{p}\left\{n\frac{\mathrm{tr}\mathbf{S}_H}{\mathrm{tr}\mathbf{S}_E} - \nu_H\right\},$$

and

$$\sigma_D = \frac{\sqrt{2\nu_H\left(\mathrm{tr}\mathbf{\Sigma}^2\right)/p}}{(\mathrm{tr}\mathbf{\Sigma})/p}.$$

For a practical situation, an (n,p)-consistent estimator is given by

$$\widehat{\sigma}_D = \frac{\sqrt{2\nu_H\{(\mathrm{tr}\mathbf{S}_E^2)/n^2 - (\mathrm{tr}\mathbf{S}_E)^2/n^3\}/p}}{(\mathrm{tr}\mathbf{S}_E)/(np)}$$

Next we derive the exact null distribution and the moments of Dempster trace criterion.

THEOREM 5.3. *When $\nu_H > p-1$ and $\nu_E > p-1$, the exact null distribution of T_D is*

$$f_{T_D}(t) = |\delta^{-1}\mathbf{\Sigma}|^{-(\nu_H+\nu_E)/2}\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}\frac{\beta^{II}(t;(p\nu_H+2k)/2,(p\nu_E+2l)/2)}{k!\ l!}$$

$$\times\sum_{\kappa}\sum_{\mu}\left(\frac{1}{2}\nu_H\right)_{\kappa}\left(\frac{1}{2}\nu_E\right)_{\mu}C_{\kappa}\left(\mathbf{I}-\delta\mathbf{\Sigma}^{-1}\right)C_{\mu}\left(\mathbf{I}-\delta\mathbf{\Sigma}^{-1}\right),\quad t>0$$

where $\beta^{II}(t;b,c)$ denote the density function of a univariate type II beta distribution of parameters b and c; $\sum_{\kappa}$ denotes summation over all partition $\kappa = (k_1,\dots,k_p)$, $k_1 \geq \cdots \geq k_p \geq 0$, of k, $C_{\kappa}(X)$ is the zonal polynomial of X corresponding to κ and the generalised hypergeometric coefficient $(a)_{\kappa}$ is given by

$$(a)_{\kappa} = \prod_{i=1}^{p}(a-(i-1)/2)_{k_i}$$

$(r)_k = r(r+1)\cdots(r+k-1)$, $(a)_0 = 1$ (see Muirhead (1982, 258))and $\delta \in (0,\infty)$ is an arbitrary parameter. Muirhead (1982, p. 341) propose to $\delta = 2\delta_1\delta_p/(\delta_1+\delta_p)$ as a close value to the optimal, where δ_1, δ_p are the largest and smallest eigenvalue of $\mathbf{\Sigma}$ respectively.

Proof. Remember that $S_H \sim \mathcal{W}_p(\nu_H,\mathbf{\Sigma})$ and $S_E \sim \mathcal{W}_p(\nu_E,\mathbf{\Sigma})$ are independent. Let $X = \mathrm{tr}S_H$ and $Y = \mathrm{tr}S_E$, then X and Y are independent too. Using Theorem 8.3.4 in Muirhead (1982, p. 339), the joint density function of X and Y is

$$f_{X,Y}(x,y) = |\delta^{-1}\mathbf{\Sigma}|^{-(\nu_H+\nu_E)/2}\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}\frac{1}{k!\ l!}g(x;p\nu_H/2+k,2\delta)g(y;p\nu_E/2+k,2\delta)$$

$$\times \sum_{\kappa} \sum_{\mu} \left(\frac{1}{2}\nu_H\right)_{\kappa} \left(\frac{1}{2}\nu_E\right)_{\mu} C_{\kappa}\left(\mathbf{I}-\delta\mathbf{\Sigma}^{-1}\right) C_{\mu}\left(\mathbf{I}-\delta\mathbf{\Sigma}^{-1}\right),$$

where

$$g(x;r,2\delta) = \frac{\exp(-x/(2\delta))x^{r-1}}{(2\delta^r)\Gamma[r]}.$$

Making the change of variables

$$T_D = X/Y, \quad Z = X \quad (T_D > 0,\ Z > 0)$$

with $dxdy = zdzdt$, the joint density function of T_D and Z is

$$f_{T_D,Z}(t,z) = |\delta^{-1}\mathbf{\Sigma}|^{-(\nu_H+\nu_E)/2} \sum_{k=0}^{\infty}\sum_{l=0}^{\infty} \frac{z}{k!\ l!} g(t/z; p\nu_H/2+k, 2\delta) g(z; p\nu_E/2+k, 2\delta)$$

$$\times \sum_{\kappa} \sum_{\mu} \left(\frac{1}{2}\nu_H\right)_{\kappa} \left(\frac{1}{2}\nu_E\right)_{\mu} C_{\kappa}\left(\mathbf{I}-\delta\mathbf{\Sigma}^{-1}\right) C_{\mu}\left(\mathbf{I}-\delta\mathbf{\Sigma}^{-1}\right).$$

Now integrating with respect to z over $z \in (0,\infty)$ gives the desired marginal density function of T_D.
■

COROLLARY 5.1. *Observe if in Theorem 5.3,* $\mathbf{\Sigma} = \delta\mathbf{I}$*, then*

$$f_{T_D}(t) = \beta^{II}(t; p\nu_H/2, p\nu_E)/2)$$

or alternatively

$$\frac{\nu_E}{\nu_H} T_D \sim \mathcal{F}(p\nu_H, p\nu_E)$$

where $\mathcal{F}(b,c)$ *is a central* F*-distribution with* b *and* c *degrees of freedom.*

COROLLARY 5.2. *Under the condition of Theorem 5.3 the moments of* T_D *are given by*

$$E(T_D^h) = |\delta^{-1}\mathbf{\Sigma}|^{-(\nu_H+\nu_E)/2} \sum_{k=0}^{\infty}\sum_{l=0}^{\infty} \frac{\Gamma[(p\nu_H+2k)/2+h]\Gamma[(p\nu_E+2l)/2-h]}{k!\ l!\Gamma[(p\nu_H+2k)/2]\Gamma[(p\nu_E+2l)/2]}$$

$$\times \sum_{\kappa} \sum_{\mu} \left(\frac{1}{2}\nu_H\right)_{\kappa} \left(\frac{1}{2}\nu_E\right)_{\mu} C_{\kappa}\left(\mathbf{I}-\delta\mathbf{\Sigma}^{-1}\right) C_{\mu}\left(\mathbf{I}-\delta\mathbf{\Sigma}^{-1}\right).$$

similarly, if $\mathbf{\Sigma} = \delta\mathbf{I}$

$$E(T_D^h) = \frac{\Gamma[p\nu_H/2+h]\Gamma[p\nu_E/2-h]}{\Gamma[p\nu_H/2]\Gamma[p\nu_E/2]}.$$

Proof. The proof follows easily from the moments of univariate type II beta distribution. ■

REMARK 5.1. Alternative expressions of the density function and the moments of T_D given in the Theorem 5.3 and the Corollary 5.2 can be derived in function of the *Invariant Polynomials*, Davis (1980); specifically, by the eq. (5.1) and (5.10) in Davis (1980), the following result are obtained (or see also Chikuse (1980)):

$$f_{T_D}(t) = |\delta^{-1}\mathbf{\Sigma}|^{-(\nu_H+\nu_E)/2} \sum_{k=0}^{\infty}\sum_{l=0}^{\infty} \frac{\beta^{II}(t;(p\nu_H+2k)/2,(p\nu_E+2l)/2)}{k!\ l!}$$

$$\times \sum_{\kappa} \sum_{\mu} \sum_{\phi \in \kappa \cdot \mu} \left(\frac{1}{2}\nu_H\right)_{\kappa} \left(\frac{1}{2}\nu_E\right)_{\mu} \left(\theta_{\phi}^{\kappa,\mu}\right)^2 C_{\phi}\left(\mathbf{I} - \delta \mathbf{\Sigma}^{-1}\right), \quad t > 0,$$

and

$$E(T_D^h) = |\delta^{-1}\mathbf{\Sigma}|^{-(\nu_H+\nu_E)/2} \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \frac{\Gamma[(p\nu_H + 2k)/2 + h]\Gamma[(p\nu_E + 2l)/2 - h]}{k!\ l!\Gamma[(p\nu_H + 2k)/2]\Gamma[(p\nu_E + 2l)/2]}$$

$$\times \sum_{\kappa} \sum_{\mu} \sum_{\phi \in \kappa \cdot \mu} \left(\frac{1}{2}\nu_H\right)_{\kappa} \left(\frac{1}{2}\nu_E\right)_{\mu} \left(\theta_{\phi}^{\kappa,\mu}\right)^2 C_{\phi}\left(\mathbf{I} - \delta \mathbf{\Sigma}^{-1}\right),$$

where

$$\theta_{\phi}^{\kappa,\mu} = \frac{C_{\phi}^{\kappa,\mu}(\mathbf{I},\mathbf{I})}{C_{\phi}(\mathbf{I})}.$$

6. Example. The following application is a modification of the example 9.4.3 of Srivastava (2002, p. 294), where the first two dependent variables Y_1 and Y_2 are considered. The matrices of sums of squares and sums of products due to the error and due to hypothesis are, respectively,

$$\mathbf{S}_E = \begin{pmatrix} 255.8029 & 112.6230 \\ 112.6230 & 415.2567 \end{pmatrix} \text{ and } \mathbf{S}_H = \begin{pmatrix} 10.05464 & 27.55231 \\ 27.55231 & 81.30097 \end{pmatrix}.$$

The test statistics for all de known criteria are tabulated in the Table 1.

Table 1: Criteria to test the null hypothesis

Criteria	Statistics	$\alpha(= 0.05)$ Critical value
Wilks's Λ	0.8324688	0.626917
Wilks's U	0.0005190485	0.0251245
Wilks's S	0.000623505	0.038469
Lawley-Hotelling's $U^{(s)}$	0.2006227	0.5487446
Pillai's $V^{(s)}$	0.1680502	0.4156
Pillai's $W^{(s)}$	0.9159749	0.7922
Pillai's $H^{(s)}$	0.9088337	0.9690216
Pillai's $R^{(s)}$	0.006177302	0.9690216
Pillai's $T^{(s)}$	0.1678238	3.168246
Roy's $\lambda_{\max}$	0.1974652	0.489425
Roy's $\theta_{\max}$	0.1649026	0.3286
Anderson's $\lambda_{\min}$	0.04104808	0.1173425
Roy's $\theta_{\min}$	0.003147606	0.105019
Dempster's T_D	0.1361364	0.1825111

Remark 6.2. Some comments about the results in the Table 1:

1. Wilks's Λ statistics, see Wilks (1932), Rencher (1995, p. 161) and Kres (1983, p. 5) among many other;

$$\Lambda = \frac{|\mathbf{S}_E|}{|\mathbf{S}_E + \mathbf{S}_H|} = \prod_{i=1}^{s} \frac{1}{1+\lambda_i} = \prod_{i=1}^{s} (1 - \theta_i)$$

The critical value was taken from table 1 in Kres (1983, pp. 14-51), besides, it was computed with the expression (3) by using the software called Mathematica.

2. Wilks's U statistics, see Wilks (1932), Roy *et al.* (1971, p. 72), Seber (1984, p. 413) and Kres (1983, p. 6) among many other;

$$U = \frac{|\mathbf{S}_H|}{|\mathbf{S}_E + \mathbf{S}_H|} = \prod_{i=1}^{s} \frac{\lambda_i}{1 + \lambda_i} = \prod_{i=1}^{s} \theta_i$$

This criterion is also known as Gnanadesikan's U statistics. The critical value was computed with the expression (31) in Wilks (1932) by the use of Mathematica. Observe, also, this statistics is wrong defined as a function of the eigenvalues λ's and θ's in Kres (1983, p.6).

3. Wilks's S statistics, see Wilks (1932), Olson (1974) and Kres (1983, p. 8)

$$S = \frac{|\mathbf{S}_H|}{|\mathbf{S}_E|} = \prod_{i=1}^{s} \lambda_i = \prod_{i=1}^{s} \frac{\theta_i}{(1 - \theta_i)}$$

Remember that this statistics is also known as Olson's S statistics. The critical values was computed by the use of Mathematica.

4. Lawley-Hotelling's $U^{(s)}$ statistics, see Muirhead (1982, p. 466), Rencher (1995, 167) and Kres (1983, p. 6) among many others;

$$U^{(s)} = \operatorname{tr}(\mathbf{S}_E^{-1}\mathbf{S}_H) = \sum_{i=1}^{s} \lambda_i = \sum_{i=1}^{s} \frac{\theta_i}{(1 - \theta_i)}$$

Unfortunately, the tables for the critical values do not include the minimum required possible combinations between de parameters s, m and h; see Table 6 in Kres (1983, pp.118-135). For the example we use an F approximation, see equation (6.30) in Rencher (1995, p. 167), see also Pillai (1955, eq. (7)). Observe this approximation is satisfactory for practical use for $h + s \geq 30$ when $s = 2$; when s increases by 1, $h + s$ must increase by 10 to give a satisfactory results.

5. Pillai's $V^{(s)}$ statistics, see Muirhead (1982, p. 466), Rencher (1995, 168) and Kres (1983, p. 6) among many others;

$$V^{(s)} = \operatorname{tr}((\mathbf{S}_E + \mathbf{S}_H)^{-1}\mathbf{S}_H) = \sum_{i=1}^{s} \frac{\lambda_i}{(1 + \lambda_i)} = \sum_{i=1}^{s} \theta_i$$

The critical values was taken from Table 7 in Kres (1983, pp. 136-153).

6. Pillai's $W^{(s)}$ statistics, see Pillai (1955);

$$W^{(s)} = \operatorname{tr}((\mathbf{S}_E + \mathbf{S}_H)^{-1}\mathbf{S}_E) = \sum_{i=1}^{s} \frac{1}{(1 + \lambda_i)} = \sum_{i=1}^{s} (1 - \theta_i) = (1 - V^{(s)}/s)$$

For the critical values we can use a type I beta approximation, see equation (6) in Pillai (1955). For practical use, this approach is satisfactory for $m+h \geq 30$ when $s = 2$; but, if s increases by 1 then, $m+h$ must be increased by 10, for getting satisfactory results. However, note the exact critical value can be obtained from $V^{(s)}$-statistics and the expression $W^{(s)} = (1 - V^{(s)}/s)$. In fact, Table 1 contains the exact value.

7. Pillai's $H^{(s)}$ statistics, see Pillai (1955) and Kres (1983, p. 8);

$$H^{(s)} = \frac{s}{\sum_{i=1}^{s}(1+\lambda_i)} = s\left\{\sum_{i=1}^{s}(1-\theta_i)^{-1}\right\}^{-1} = (1+U^{(s)}/s)^{-1}$$

The critical values was obtained using a type I beta approximation, see equation (9) in Pillai (1955). Here we have to apply the above mentioned practical conditions for the correct use of the approximation of the $U^{(s)}$ statistics.

8. Pillai's $R^{(s)}$ statistics, see Pillai (1955) and Kres (1983, p. 8);

$$R^{(s)} = \frac{s}{\sum_{i=1}^{s}\frac{1+\lambda_i}{\lambda_i}} = s\left\{\sum_{i=1}^{s}\theta_i^{-1}\right\}^{-1} = (1+U^{'(s)}/s)^{-1}$$

where $U^{'(s)}$ is the same $U^{(s)}$ but with m and h interchanged. For this criterion, the critical values are computed using a type I beta approximation, see equation (11) in Pillai (1955). Again, the same conditions explained before for $W^{(s)}$ statistics have to be applied for a satisfactory result in the approximations. Besides, we need the conditions $m \geq 0$, or $|\nu_H - p| \geq 1$ for getting satisfactory approximation. The last condition was not considered by Pillai (1955), but it is required, because for a beta distribution $\beta(a,b)$ it is known that $a > 0$, which is guaranteed when $m \geq 0$ in the approximation.

9. Pillai's $T^{(s)}$ statistics, see Pillai (1955) and Kres (1983, p. 8);

$$T^{(s)} = s\left\{\sum_{i=1}^{s}\lambda_i^{-1}\right\}^{-1} = \frac{s}{\sum_{i=1}^{s}\frac{1-\theta_i}{\theta_i}} = \frac{R^{(s)}}{1-R^{(s)}}$$

The critical values was obtained using a type II beta approximation, see equation (13) in Pillai (1955). Again, we use the same for rules of $R^{(s)}$ statistics for a satisfactory approximation including the restriction over m.

10. Roy's $\lambda_{\max}$, see Roy (1957) and Kres (1983, p. 7);

$$\lambda_{\max} = \frac{\theta_{\max}}{1-\theta_{\max}}$$

The critical values were obtained from table 3 in Kres (1983, pp. 62-86). Besides, we got the critical value by integrating the joint distribution of the λ's via Mathematica.

11. Roy's $\theta_{\max}$, see Roy (1957), Muirhead (1982, p. 481), Rencher (1995, p. 164) and Kres (1983, p. 7) among many others;

$$\theta_{\max} = \frac{\lambda_{\max}}{1+\lambda_{\max}}$$

For this criterion the critical values can be obtained from table 2, 4 or 5 in Kres (1983, pp. 52-61, 87-104 and 105-117, respectively). Again, Mathematica was used for finding the critical value of that criterion by integrating the joint distribution of the eigenvalues θ's.

12. Anderson's $\lambda_{\min}$, see Roy (1957), Anderson (1982) and Kres (1983, p. 7) among many others;

$$\lambda_{\min} = \frac{\theta_{\min}}{1-\theta_{\min}}$$

As above, the critical values were computed via Mathematica by integrating the joint distribution of the eigenvalues λ's. However, note the critical value can be determined as a function of the critical value for $\theta_{\min}$.

13. Roy's $\theta_{\min}$, see Pillai (1955) and Roy (1957);

$$\theta_{\min} = \frac{\lambda_{\min}}{1+\lambda_{\min}}$$

Similarly to Anderson's criterion, the integration of the joint distribution of the eigenvalues λ's via Mathematica gave the critical values. However those values can be obtained from the distribution of $\theta_{\max}$ by $\theta_{\min}(\alpha, s, \nu_H, \nu_E) = 1 - \theta_{\max}(\alpha, s, \nu_E, \nu_H)$, see Nanda (1948), but, again, the published tables do not allow to read the values because, they do not incorporate such combination of the parameters; in fact there are a lot of similar particular cases for which the critical value can not be found from those tables.

14. Dempster's T_D, see Dempster (1958), Dempster (1960) and Fujikoshi *et al.* (2004);

$$T_D = (\mathrm{tr}\mathbf{S}_H)/(\mathrm{tr}\mathbf{S}_E)$$

In this criterion, the critical value was obtained using the normal approximation (4).

15. General remark:

    (a) In general, the decision rule for all criteria is:
    reject H_0 if the statistics $\geq$ critical value
    However, for Wilks's Λ and Pillai's $W^{(s)}$ criteria, the decision rule is (this class of test are known in statistical literature as **inverse test**, see Rencher (1995, p. 162)):
    reject H_0 if the statistics $\leq$ critical value.

    (b) In general, the tables for critical values of all the criterion are tabulated in terms of the parameters (p, ν_H, ν_E) or in terms of the parameters (s, m, h), where

    $$s = \min(p, \nu_H), \quad m = (|\nu_H - p| - 1)/2 \quad \text{and} \quad h = (\nu_E - p - 1)/2.$$

    Besides, the tables (in general) have been computed by assuming that $p \leq \nu_H$ and $p \leq \nu_E$. If $p > \nu_H$ then use the combination of parameters $(\nu_H, p, \nu_E + \nu_H - p)$ in place of (p, ν_H, ν_E), see Muirhead (1982, eq. (7), p. 455), Srivastava & Khatri (1979, p. 96) or Rencher (1995, p. 167).

    (c) Finally, observe the null hypothesis is not rejected under any criterion.

Acknowledgment. This work was supported partially by the research project 39017E of CONACYT-México.

REFERENCES

Anderson, T. W. (1982). *An introduction to multivariate statistical Analysis*, 2nd edition, John Wiley & Sons, New York.

Consul, P. C. (1966). On the exact distributions of the likelihood ratio criteria for testing linear hypotheses about regresion coefficients, *Ann. Math. Statist.* **37**(5), 1319-1330.

Chikuse, Y. (1980). Invariant polynomials with matrix arguments and their applications. In: Gupta, R. P. (ed.) *Multivariate Statistical Analysis*. North-Holland Publishing Company, pp. 53-68.

Davis, A. W. (1980). Invariant polynomials with two matrix arguments, extending the zonal polynomials. In: P. R. Krishnaiah (ed.) *Multivariate Statistical Analysis*. North-Holland Publishing Company, pp. 287-299.

Dempster, A. P. (1958). A high dimensional two sample significance test, *Ann. Math. Statist.* **29**, 995-1010.

Dempster, A. P. (1960). A significance test for the separation of two highly multivariate small samples, *Biometrics* **16**, 41-50.

Díaz-García J. A. and Gutiérrez-Jáimez, R. (1997). Proof of conjectures of H. Uhlig on the singular multivariate Beta and the Jacobian of a certain matrix transformation, *Ann. of Statist.* **25**: 2018-2023.

Díaz-García J. A., González-Farías, G. and Alvarado-Castro, V. M. (2005). Sensitivity analysis in linear regression. *Comunicación Técnica*, No. **I-05-05** (PE/CIMAT), Guanajuato, México.

Fujikoshi, Y., Himeneo, T., and Wakaki, H. (2004). Asymptotic results of a high dimensional MANOVA test and power comparation when the dimension is large compared to the sample size, *J. Japan Statist. Soc.* **43**(1), 19-26.

Hsu, P. L. (1940). On generalized analysis of variance. (I), *Biometrika* **31**, 221-237.

Kres, H. (1983). *Statistical tables for multivariate analysis*, Springer- Verlag, New York.

Morrison, D. F. (1978). *Multivariate statistical methods*, McGraw Hill, London.

Muirhead, R. J. (1982). *Aspects of multivariate statistical theory*, John Wiley & Sons, New York.

Nanda, D. N. (1948). Distribution of root of determinantal equation, *Ann. Math. Statist.* **19**, 47-57.

Olson, Ch. L. (1973). A Monte Carlo investigation of the robustness of multivariate analysis of variance, *Unpubilshed Ph.D. dissertation*, University of Toronto, Canada.

Olson, Ch. L. (1974). Comparative robustness of six test in multivariate analysis of variance, *JASA* **69**, 894-908.

Pillai, K. C. S. (1955). Some new test criteria in multivariate analysis, *Ann. Math. Statist.* **26**, 117-121.

Rencher, A. C. (1995). *Methods of multivariate analysis.* John Wiley & Sons, New York.

Roy, S. N. (1957). *Some Aspect of multivariate analysis.* John Wiley & Sons, New York.

Roy, S. N., Gnanadesikan, R., and Sirivastava, J. N. (1971). *Analysis and design of certain quantitative multiresponse experiments.* Oxford: Pergamon Press Ltd.

Seber, G. A. F. (1984). *Multivariate observations.* John Wiley & Sons, New York.

Schatzoff, M. (1966). Exact distributions of Wilks's likelihood ratio criterion, *Biometrika* **54**(3-4), 347-358.

Srivastava, M. S. (2002). *Methods of multivariate statistics.* John Wiley & Sons, New York.

Srivastava, S. M., & Khatri, C. G. (1979). *An introduction to multivariate statistics.* North Holland, New York.

Wilks, S. S. (1932). Certain generalizations in the analysis of variance, *Biometrika* **24**, 471-494.

Wilks, S. S. (1935). On the independence of k sets of normally distributed statistical variables, *Econometrica* **3**(3), 309-326.

José A. Díaz-García
Department of Statistics and Computation
Universidad Autónoma Agraria Antonio Narro
25350 Buenavista, Saltillo, Coahuila
México
E-mail: jadiaz@uaaan.mx

Francisco José Caro-Lopera
Department of Mathematics
Universidad de Antioquia
Medellín
Colombia
E-mail: fjcaro@matematicas.udea.edu.co