Testing Independence with Additional Information

by

François Perron
Université de Montréal and Purdue University

Technical Report # 90–24C

Department of Statistics
Purdue University

May, 1990
TESTING INDEPENDENCE WITH ADDITIONAL INFORMATION

by

François Perron
Université de Montréal and Purdue University

ABSTRACT

Summary. Let $X = (X^j : j = 1, \ldots, n)$ be $n$ row vectors of dimension $p$ independently and identically distributed multinormal. For each $j$, $X^j$ is partitioned as $X^j = (X^j_1, X^j_2, X^j_3)$ where $p_i$ is the dimension of $X^j_i$ with $p_1 = 1, p_1 + p_2 + p_3 = p$. In addition, consider $Y^j_i, i = 1, 2, j = 1, \ldots, n$, where the vectors $Y^j_i$ are independent and distributed as $X^j_i$. We treat here the problem of testing independence between $X^1_1$ and $X^1_2$ knowing that $X^1_1$ and $X^1_2$ are uncorrelated. A locally best invariant test is proposed for this problem.

Key words: Invariance, group, orbits, maximal invariant.

Résumé. À partir de vecteurs aléatoires lignes indépendants entre eux $(X^j : j = 1, \ldots, n)$ distribués selon une loi multinormale de dimension $p$ on cherche à tester s’il y a indépendance ou non entre la première composante et les $p_3$ dernières composantes des vecteurs $X^j$ sachant que la première composante est indépendante des $p_2$ composantes suivantes. Pour cela, on partitionne tous les vecteurs $X^j$ en sous-vecteurs $(X^j = X^j_1, X^j_2, X^j_3)$ de dimensions $p_1 = 1, p_2, p_3$ respectivement. Dans ce problème on suppose qu’il existe également des vecteurs $(Y^j_i : i = 1, 2, j = 1, \ldots, n_i)$ dont les distributions en loi sont les mêmes que celles des marginales $X^j_i$ de $X^1$. Ces vecteurs forment l’information additionnelle. Dans cet article, on trouve un test localement le plus puissant pour ce problème.

AMS 1985 subject classifications: Primary 62H15; secondary 62A05, 62F03.

Supported by the National Sciences and Engineering Research Council of Canada.
1. Introduction

Consider a random vector in $\mathbb{R}^p$ having a multivariate normal distribution with mean 0 and covariance $\Sigma$ ($\Sigma$ positive definite). Decompose this vector into three subvectors of dimensions $p_1 = 1, p_2$ and $p_3$. Assuming that the first subvector is independent of the second subvector we wish to test whether the first subvector is independent of the third subvector or not. Partitioning $\Sigma$ in an appropriate way, the problem corresponds to testing $H_0 : \Sigma_{13} = 0$ versus $H_1 : \Sigma_{13} \neq 0$ with

$$
\Sigma = \begin{pmatrix}
\Sigma_{11} & 0 & \Sigma_{13} \\
0 & \Sigma_{22} & \Sigma_{23} \\
\Sigma_{31} & \Sigma_{32} & \Sigma_{33}
\end{pmatrix}
$$

where $\Sigma_{ij}$ is a $p_i \times p_j$ matrix $i, j = 1, 2, 3$.

In order to test this hypothesis a sample consisting of three random matrices $(X, Y_1, Y_2)$ of dimensions $n \times p, n_1 \times p_1$ and $n_2 \times p_2$ respectively is observed. It is assumed that $X, Y_1$ and $Y_2$ are mutually independent and distributed as $N(0, I_n \otimes \Sigma), N(0, I_{n_1} \otimes \Sigma_{11})$ and $N(0, I_{n_2} \otimes \Sigma_{22})$ respectively. In other words, the row vectors of $X$ are independent and identically distributed $N_p(0, \Sigma)$ and, similarly, the row vectors of $Y_i$ are independent and identically distributed $N(0, \Sigma_{ii}), i = 1, 2$. The matrix $X$ is partitioned as $X = (X_1X_2X_3)$ where $X_i$ is $n \times p_i, i = 1, 2, 3$. Let $S_{ij} = X_i^tX_j$ and $W_k = Y_k^tY_k, k = 1, 2, i, j = 1, 2, 3$. The matrices $S, W_1, W_2$ are independent Wishart and they form a sufficient statistic. The couple $(W_1, W_2)$ is what we call the additional information.

In practice, additional information appears in many circumstances. It occurs in sample surveys when two types of questionnaires (one partial and the other one complete) are distributed. It may also occur when we combine data obtained from different experiments.
The likelihood ratio test for testing $H_0$ versus $H_1$ is given by rejecting $H_0$ if and only if $R$ is large where
\[
R = \frac{S_{13.2}S_{33.2}^{-1}S_{31.2}/p_3}{S_{11.2}/(n - p_2)}
\]
and $S_{ij,k} = S_{ij} - S_{ik}S_{kk}^{-1}S_{kj}$. This test does not take into account the additional information. When no additional information is available the locally best invariant test (Giri 1979) is given by rejecting $H_0$ if and only if $\varphi_1$ is large where
\[
\varphi_1 = (R - 1) \left( \frac{S_{11.2}/(n - p_2)}{S_{11}/n} \right).
\]
The purpose of this article is to find the locally best invariant test when additional information is available. It is expected that this test will make use of $(W_1, W_2)$. For instance, consider the following related problem where $p_1 \geq 1, p_2 = 0$ and $W_3$ is $W_{p_3}(\Sigma_{33}, n_3)$. In this case the locally best invariant test (Eaton and Kariya 1983) consists in rejecting $H_0$ if and only if $\varphi_2$ is large where
\[
\varphi_2 = \frac{(n + n_1)(n + n_3)}{p_1 p_3} \text{tr}\left((S_{11} + W_1)^{-1}S_{13}(S_{33} + W_3)^{-1}S_{31}\right)
\]
\[ - \sum_{i=1,3} \frac{(n + n_i)}{p_i} \text{tr}\left((S_{ii} + W_i)^{-1}S_{ii}\right).\]
Notice how the statistic $\varphi_2$ makes use of $(W_1, W_3)$.

The invariance context is introduced in Section 2 and the locally bets invariant test is given. Computations are reported in Section 3. Throughout this text the multiplicative group of $p \times p$ non-singular matrices is denoted by $G_{\ell}(p)$. The subgroups of orthogonal matrices and lower triangular matrices with positive elements on the diagonal are respectively denoted by $O(p)$ and $T(p)$. 

3
2. Invariance context

Let $S, W_1, W_2$ be independently distributed Wishart with parameters $(p, \Sigma, n), (p_1, \Sigma_{11}, n_1)$ and $(p_2, \Sigma_{22}, n_2)$ respectively. Let $P_{\Sigma}$ be the probability measure associated with $(S, W_1, W_2)$ when $\Sigma$ is the parameter and let $G$ be the group of transformations given by
\[
G = \left\{ g = \begin{pmatrix} g_{11} & 0 & 0 \\ 0 & g_{22} & 0 \\ 0 & g_{32} & g_{33} \end{pmatrix} : g \in G_\ell(p), g_{ii} \in G_\ell(p_i) \text{ for } i = 1, 2, 3 \right\}.
\]
Corresponding to $g \in G$ the transformations on the sufficient statistic $(S, W_1, W_2)$ and the parameter $\Sigma$ are given by
\[
(S, W_1, W_2) \rightarrow \tilde{g}(S, W_1, W_2) = (gSg^t, g_{11}^2W_1, g_{22}W_2g_{22}^t),
\]
and a maximal invariant in the parameter space is given by $\rho$ where $\rho(\Sigma) = \Sigma_{13}\Sigma_{33,2}^{-1}\Sigma_{31}\Sigma_{11}^{-1}$.

The power function of any invariant test is constant on each orbit of the parameter space so there is no loss of generality in working on a class of representatives instead of working on the original parameter space. Let
\[
A(\delta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{pmatrix}^{-1} \quad (2.1)
\]
with $D^t = (0, \ldots, 0, 1)$. The set $\{A(\delta) : \delta \in [0, 1)\}$ consists in a class of representatives for the orbits of the parameter space and $\rho(A(\delta)) = \delta^2$. The last inequality indicates a bijection between the class of representatives and the range of $\rho$.

Define $P^I_\delta$ as the restriction of $P_{A(\delta)}$ over the class of all measurable sets which are invariant. According to Theorem 1 Section 3
\[
\frac{dP^I_\delta}{dP_\delta}(S, W_1, W_2) = (1-\delta^2)^{n/2}[1 + \frac{n + n_1}{2} + \frac{(n - p_2)}{2}(R - 1) \frac{S_{11,2}}{(S_{11} + W_1)/(n + n_1)} + 0(\delta^3)].
\]
Therefore, the locally best invariant test is given by rejecting $H_0$ if and only if $\varphi_3$ is large where

$$\varphi_3 = (R - 1) \frac{S_{11.2}/(n - p_2)}{(S_{11} + W_1)/(n + n_1)}.$$

The statistic $\varphi_3$ is the product of two factors. The first factor is equivalent to the likelihood ratio statistic. It essentially measures the multiple correlation between $X_1$ and $X_3$ after removing the effect of $X_2$. The second factor is a ratio of two estimators of $\Sigma_{11}$. The additional information is used to get an improved estimator in the denominator. (Giri’s test has $n_1 = 0$ and $W_1 = 0$.) The second factor provides a measure of orthogonality between $X_1$ and the columns of $X_2$. The fact that this test is locally most powerful suggests that, as $X_1$ becomes more nearly orthogonal to the columns of $X_2$, the first factor becomes more effective in detecting near-zero correlation.

A comment on the fact that the test does not involve $Y_2$ : Giri’s test uses $X_2$ only through the projection matrix $X_2(X_2'X_2)^{-1}X_2'$, which contains no information about $\Sigma_{22}$. Thus it is not surprising that additional information about $\Sigma_{22}$ is ignored.

3. Computations

In this section, we derive expression (2.2). Using Wijsman’s representation theorem (ref. Wijsman 1967)

$$\frac{dP_\delta^f(S, W_1, W_2)}{dP_0^f} = \frac{r(\delta, S, W_1, W_2)}{r(0, S, W_1, W_2)}$$

with

$$r(\delta, S, W_1, W_2) = \int_G f(\hat{\gamma}(S, W_1, W_2)|A(\delta))\lambda(d\gamma)$$

where $\lambda$ is a left invariant measure on $G$ and $f(\cdot|\Sigma)$ is the density function of $(S, W_1, W_2)$.
with respect to an invariant measure $\mu$ when the parameter is $\Sigma$. The measures $\lambda$ and $\mu$ are unique up to a multiplicative factor. Let

$$\lambda(dg) = |g_{33}g_{33}^t|^{-p_2/2} \prod_{i=1}^{3} \lambda_{p_i}(dg_{ii})$$

and

$$\mu(d(S,W_1,W_2)) = \mu_p(dS)\mu_{p_1}(dW_1)\mu_{p_2}(dW_2)$$

where $\lambda_q(dh) = |hh^t|^{-q/2}dh$ is a left invariant measure on the space of all $q \times q$ matrices and $\mu_q(dW) = |W|^{-(q+1)/2} \prod_{1 \leq i \leq j \leq q} dw_{ij}$ is an invariant measure on the space of all $q \times q$ positive definite matrices. The joint density of $(S,W_1,W_2)$ with respect to the measure $\mu$ is given by

$$f(S,W_1,W_2|\Sigma) = K|\Sigma^{-1}S|^{n/2} \prod_{i=1}^{2} |\Sigma_{ii}W_i|^{n_i/2}$$

$$\exp\left\{-\frac{1}{2}\left\{tr(\Sigma^{-1}S) + \sum_{i=1}^{2} tr(\Sigma_{ii}^{-1}W_i)\right\}\right\}$$

where $K$ is a normalization constant independent of $\Sigma$.

**Theorem 1.**

$$\frac{dP^I}{dP^0}(S,W_1,W_2) = (1 - \delta^2)^{n/2} \left[ 1 + \frac{(n + n_1)}{2} + \frac{(n - p_2)}{2} \{R - 1\} \right.$$ \[ \frac{S_{11,2}/(n - p_2)}{(S_{11} + W_1)/(n + n_1)} + O(\delta^3) \left. \right] . $$

**Proof.** First, begin with

$$f(\tilde{g}(S,W_1,W_2|A(\delta)) = K(1 - \delta^2)^{n/2}|S|^{n/2}W_1^{n_1/2}|W_2|^{n_2/2}\left(\begin{array}{c}
g_{11}^2 \left(n + n_1\right)/2 \quad g_{22}g_{22}^t \left(n + n_2\right)/2 \quad g_{33}g_{33}^t \left(n/2\right) \\
\end{array}\right)$$

$$\exp\left\{-\frac{1}{2}\{(S_{11} + (1 - \delta^2)W_1)g_{11}^2 + tr(g_{22}(S_{22} + W_2)g_{22}^t) + tr(g_{33}S_{33}g_{33}^t) + tr(g_{32}S_{22}g_{32}^t) + 2tr(g_{32}S_{23}g_{33}^t) + 2\delta g_{11}Dt g_{32}S_{21} + 2\delta g_{11}Dt g_{33}S_{31}\right\} .$$
Introducing the variable transformation
\[ h_{32} = [g_{32} + g_{33}S_{22}^{-1} + \delta D g_{11} S_{12} S_{22}^{-1}] S_{22}^{1/2} \]
\[ h_{ii} = g_{ii} V_i, \quad V_1 = (S_{11} + W_1)^{1/2}, \quad V_2 = (S_{22} + W_2)^{1/2}, \quad V_3 = S_{33,2}^{1/2} \]
\[ \lambda(dg) = |S_{22}|^{-p_{a}/2} \lambda(dh). \]
We get
\[ r(\delta, S, W_1, W_2) = K(1 - \delta^2)^{n/2} |S|^{n/2} |S_{22}|^{-p_{a}/2} V_1^{-(n+n_1)} V_2^{-(n+n_2)} \]
\[ |V_3|^{-n} W_1^{n_{11}/2} W_2^{n_{22}/2} \int_G (h_{i1}^2)^{(n+n_1)/2} (h_{22} h_{i2}^t)^{(n+n_2)/2} (h_{33} h_{3i}^t)^{(n-p_2)/2} \]
\[ \exp \left( -\frac{1}{2} \{ h_{i1}^2 + \text{tr}(h_{22} h_{i2}^t) + \text{tr}(h_{33} h_{3i}^t) + \text{tr}(h_{32} h_{3i}^t) \right) \]
\[ + 2\delta D^t h_{33} V_3^{-1} S_{31,2} V_1^{-1} h_{11} - \delta^2 h_{i1}^2 S_{11} V_1^{-2} (R_1 + Z) \} \lambda(dh) \]
with \( Z = S_{11}^{-1} W_1 \) and \( R_1 = S_{12} S_{22}^{-1} S_{21}/S_{11} \).

Integrating over \( h_{22}, h_{32} \) and developing the exponential for \( \delta \) close to zero we find
\[ r(\delta, S, W_1, W_2) = (1 - \delta^2)^{n/2} b(S, W_1, W_2) \]
\[ \{ \int_{\Omega_{\ell}(p_1)} \int_{\Omega_{\ell}(p_2)} [1 - D^t h_{33} V_3^{-1} S_{31,2} V_1^{-1} h_{11} \delta + \frac{1}{2} ((D^t h_{33} V_3^{-1} S_{31,2})^2 + \]
\[ S_{11}(R_1 + Z)) \delta^2 V_1^{-2} h_{i1}^2 (h_{i1}^2)^{(n+n_1)/2} (h_{33} h_{3i}^t)^{(n-p_2)/2} \exp \left( -\frac{1}{2} \{ h_{i1}^2 + \text{tr}(h_{33} h_{3i}^t) \right) \]
\[ \lambda_{p_1}(dh_{11}) \lambda_{p_2}(dh_{33}) + Q(\delta, S, W_1, W_2) \}
\]
where \( Q(\delta, S, W_1, W_2) = O(\delta^3) \) uniformly in \((S, W_1, W_2)\) (can be proven).

Next, decompose \( G_{\ell}(p) \) as a product space \( G_{\ell}(p) = T(p) \times O(p) \) where \( g = \phi, g \in G_{\ell}(p), t \in T(p) \) and \( o \in O(p) \). According to this decomposition the left invariant measure \( \lambda_p \) is decomposed as \( \lambda_p(dg) = \tau_p(dt) \times \nu_p(do) \) where \( \tau_p \) is a left invariant measure on \( T(p) \) and \( \nu_p \) is a left invariant probability measure on \( O(p) \). Moreover, corresponding to \( \nu_p \) the following identities (James 1954).
\[ \int_{O(p)} \text{tr}(AO) \nu_p(dO) = 0, \int_{O(p)} \text{tr}^2(AO) \nu_p(dO) = \text{tr}(AA^t)/p, A(p \times p) \]
are used to obtain
\begin{align*}
  r(\delta, S, W_1, W_2) &= (1 - \delta^2)^{n/2} b(S, W_1, W_2) \left\{ \int_{T(p_1)} \int_{T(p_2)} [1 + \frac{1}{2} (D^t T_{33} T_{33}^t D R_2 / p_3 + Z + R_1)(1 + Z)^{-1} T_{11}^2 \delta^2]^n |T_{33}|^{n-p_2} \right. \\
  & \left. \exp \left\{ -\frac{1}{2} \{ T_{11}^2 + tr(T_{33} T_{33}^t) \} \tau_{p_1} (d T_{11}) \tau_{p_2} (d T_{33}) + O(\delta^3) \} \right. \right\}.
\end{align*}

Finally, recognizing the Bartlett decomposition of a Wishart distribution (ref. Giri 1977) we obtain
\begin{align*}
  r(\delta, S, W_1, W_2) / r(0, S, W_1, W_2) &= (1 - \delta^2)^{n/2} \\
  \{1 + \frac{1}{2} E\{D^t U_1 D R_2 / p_3 + Z + R_1\} U_2 (1 + Z)^{-1} \delta^2 + O(\delta^3)\} \\
  &= (1 - \delta^2)^{n/2} \{1 + \frac{(n + n_1)}{2(1 + Z)} \{ \frac{(n - p_2)}{p_3} R_2 + Z + R_1 \} \delta^2 + O(\delta^3)\} \\
  &= (1 - \delta^2)^{n/2} \{1 + \frac{(n + n_1)}{2} + \frac{(n - p_2)}{2} (R - 1) \frac{S_{1.2}}{(S_{11} + W_1)/(n + n_1)} + O(\delta^3)\}
\end{align*}

where $U_1$ is $W_{p_3}(n - p_2, I)$ and $U_2$ is $\chi^2_{n+n_1}$ with $R_2 = S_{13.2} S_{33.2}^{-1} S_{31.2} / S_{11}$ and $R = (n - p_2) R_2 / p_3 (1 - R_1)$. Q.E.D.

Acknowledgement: I am grateful to the referees for the valuable suggestions concerning the intuitive interpretation of the statistic $\varphi_3$.

REFERENCES


