Kendall-type Tests in Multivariate Regression

K. Choi
Hong-Ik University, Department of General Study (Mathematics)
Cho-Chi-Won, Yun-Gi-Gun
Chung-Nam, 339-701, Korea
E-Mail : kmchoi@wow.hongik.ac.kr

1. The Model and Test Statistics

Suppose X_i's and Y_j's are row vectors of matrices X and Y respectively. Let the paired data be $(X_1, Y_1), \cdots, (X_n, Y_n)$, where $X_i \in R^p$ and $Y_i \in R^q$ for $p \geq 1$ and $q \geq 1$. Then we can consider a model

$$Y = XB/\sqrt{n} + U, \tag{1}$$

where X is an $n \times p$ known matrix, Y is an $n \times q$ observed matrix, and B is an unknown $p \times q$ parameter matrix. And U is a matrix of unobserved random disturbances whose rows for given X are uncorrelated with X, each with mean 0 and common covariance matrix Σ. We want to test $H_o : B/\sqrt{n} = 0$ against $H_A : B/\sqrt{n} \neq 0$. As a measure of association between X and Y, we extend Kendall's τ (Gibbon,1985 ; Choi and Marden, 1997) to a multivariate one as follows:

$$T = \frac{1}{C(n,2)} \sum \sum\nolimits_{i<j} b(X_i - X_j) b(Y_i - Y_j)'. \tag{2}$$

The statistic of our main interest based on T in (**??**) is defined as

$$\hat{\tau} \equiv \text{vec}(T). \tag{3}$$

For comparisons we express the linear model (**??**) in a vector form (Mardia et al.,1979) as follows :

$$\text{vec}(Y) = X^* \text{vec}(B)/\sqrt{n} + \text{vec}(U), \tag{4}$$

where $X^* \equiv X \otimes I_q$. Then projection matrix P^* on X^* is given by $X^*(X^{*'}X^*)^{-1}X^{*'} = P \otimes I_q$, where $P = X(X'X)^{-1}X'$. Under the null hypothesis the conventional test statistic is

$$LS = \text{vec}(Y)' P^* \text{vec}(Y) \sim \chi^2_{pq}. \tag{5}$$

2. Distributions and Asymptotic Relative Efficiencies

Theorem 1 *Suppose that Σ_X and Σ_Y are not zero, and the distribution of X_i's is not degenerate. Under the null hypothesis,*

$$\sqrt{n}\hat{\tau} \to N_{pq}(0, 4\Sigma_X \otimes \Sigma_Y), \tag{6}$$

in distribution. As n goes to infinity, with a consistent and equivariant estimator $\widehat{\text{Cov}(\hat{\tau})}$ *of* $\text{Cov}(\hat{\tau})$,

$$\hat{\tau}' \widehat{\text{Cov}(\hat{\tau})}^{-1} \hat{\tau} \to \chi^2_{pq}. \tag{7}$$

In order to obatain the asymptotic relative efficiency (ARE) of $\hat{\tau}$ with respect to LS, we use a probability density function (**??**) of X_i's or Y_i's, which has nice spherical feature and flexibility of the tail shape (Randles, 1989).

$$f(z) = k_p \exp\left(-[z'z/c_p]^\nu\right), \quad z \in R^p, \tag{8}$$

where

$$c_p = \frac{p\Gamma\left(\frac{p}{2\nu}\right)}{\Gamma\left(\frac{p+2}{2\nu}\right)}, \qquad k_p = \frac{\nu\Gamma\left(\frac{p}{2}\right)}{\Gamma\left(\frac{p}{2\nu}\right)[\pi c_p]^{p/2}}. \tag{9}$$

Here if $\nu > 1$, the pdf has heavy tails. It $\nu < 1$, it has light tails. If $\nu = 1$, it represents the normal distribution.

Theorem 2 *Both* $X_1, \cdots, X_n$ *and* $Y_1, \cdots, Y_n$ *are iid from spherically symmetric distributions defined in (**??**) with their tail shape parameters* ν_X *and* ν_Y *respectively. Then*

$$ARE(\hat{\tau}, LS) = \frac{1}{4\sigma_X\sigma_Y}\frac{(q-1)^2}{p^2q^2}E^2[||X_1 - X_2||]E^2\left[\frac{1}{||Y_1 - Y_2||}\right], \tag{10}$$

where $\sigma_X \equiv E\left[\frac{(X_1-X_2)'}{||X_1-X_2||}\frac{(X_1-X_3)'}{||X_1-X_3||}\right]/p$.

For any $p \geq 2$, $E[||X_1 - X_2||]$ is reduced to an integral of two-variable integrand. For any odd $q \geq 3$, $E[1/||Y_1 - Y_2||]$ is reduced to an integral of three-variable integrand, so that it can be calculated numerically.

REFERENCES

Choi,K. and Marden, J. (1997). An approach to multivariate rank transform tests in multivariate analysis of variance, *J. Ameri. Statist. Ass.*, **92**, 1581-1590.

Gibbons, J. D. (1985). *Nonparametric Statistical Inference, Dekker, Inc.*

Mardia, K.V., Kent, J.T. and Bibby, J.M. (1979). *Multivariate Analysis. Academic Press.*

Randles, R. H. (1989). A distribution-free multivariate sign test based on interdirections, *J. Ameri. Statist. Ass.*, **84**, 1045-1050.