Statistique et analyse des données

Jean François Ingenbleek

Tests simultanés de permutation des rangs pour bruit-blanc multivarié

Statistique et analyse des données, tome 6, nº 2 (1981), p. 60-65.
<http://www.numdam.org/item?id=SAD_1981__6_2_60_0>

Statistique et Analyse de données
1981 - Vol. 6 n° 2 - pp. 60-65

TESTS SIMULTANES DE PERMUTATION DES RANGS POUR BRUIT-BLANC MULTIVARIE

Jean François INGENBLEEK

Université Libre de Bruxelles
Institut de Statistique
1050 BRUXELLES.

RESUME : *Sous l'hypothèse nulle de bruit blanc gaussien, les coefficients d'autocorrélation observés d'ordres différents sont asymptotiquement indépendants. Ce résultat est bien souvent utilisé dans les tests de bruit-blanc basés sur une statistique tenant compte de plusieurs autocorrélations d'ordres différents. Nous démontrons ici qu'il peut s'étendre à une classe très générale de statistiques de rangs utilisées dans l'analyse des séries chronologiques. Cette classe comprend, notamment, les coefficients d'autocorrélation entre les rangs ainsi que les statistiques utilisées dans le test des séquences ou celui des différences premières positives.*

ABSTRACT : *Under the null hypothesis of white noise gaussian process, the observed autocorrelation coefficients of different orders are asymptotically independent. This property is very helpfull for testing for white noise taking into account autocorrelations of various orders. We show that this property is still valid for a large class of rank statistics used in time series analysis. This class includes among others the rank autocorrelations, the run test statistics and the difference-sign test statistics.*

1 . INTRODUCTION

Il est bien connu [1] que le coefficient d'autocorrélation d'ordre un est asymptotiquement normal de moyenne nulle et d'écart-type $1/\sqrt{n}$, sous l'hypothèse nulle de bruit-blanc gaussien. Il en est de même pour le coefficient d'autocorrélation d'ordre deux et, de plus, ces deux coefficients sont indépendants.

C'est un résultat analogue qu'on présente ici, mais portant sur des statistiques de rangs traitant des séries multivariées.

Notons $X_\alpha = (X_{1\alpha}, X_{2\alpha}, \ldots, X_{m\alpha})'$ $(\alpha = 1,2,\ldots,n)$ les éléments d'une série stationnaire m-variée de longueur n. Soient $R_{i1}^n, R_{i2}^n, \ldots, R_{in}^n$, les rangs des variables $X_{i1}, X_{i2}, \ldots, X_{in}$.

Posons :

$$R_n = (R_1^n, R_2^n, \ldots, R_n^n) = \begin{bmatrix} R_{11}^n & R_{12}^n \ldots & R_{1n}^n \\ R_{21}^n & R_{22}^n \ldots & R_{2n}^n \\ \vdots & \vdots & \vdots \\ R_{m1}^n & R_{m2}^n & R_{mn}^n \end{bmatrix}$$

$$v_\alpha = \begin{bmatrix} v_{1\alpha} \\ v_{2\alpha} \\ \vdots \\ v_{n\alpha} \end{bmatrix} = \begin{bmatrix} F_1(X_{1\alpha}) \\ F_2(X_{2\alpha}) \\ \vdots \\ F_m(X_{m\alpha}) \end{bmatrix} \quad (\alpha = 1,2,\ldots,n),$$

où F_i est la fonction de répartition continue de X_{i1} $(i = 1,2,\ldots,m)$.

2. STATISTIQUE ET PRINCIPE DE PERMUTATION UTILISES

Considérons le vecteur aléatoire $S_{1n}(R_n)$ à 2m composantes

$$S_{1n,j_1j_2} = \frac{1}{n} \sum_{\alpha=1}^{n} J\left(\frac{R^n_{j_1\alpha}}{n+1}, \frac{R^n_{j_2\alpha+1}}{n+1}\right) \qquad (j_1,j_2 = 1,2,\ldots,m).$$

où $J(u_1,u_2)$ est une fonction donnée définie sur $[0,1]^2$.
(et $R^n_{n+1} \equiv R_n$ etc ...).

Pour m = 1, et en particularisant J, il est possible de retrouver le test des séquences autour de la médiane, le test basé sur le nombre de différences premières positives ou le test basé sur le coefficient d'autocorrélation d'ordre un entre les rangs.

Le vecteur aléatoire $S_{1n}(R_n)$ n'est pas libre sous l'hypothèse nulle de bruit blanc : d'une manière générale, la distribution liée de R^n_1, dépend de celle de X_1. Par contre sous les permutations des colonnes de la matrice R_n, $S_n(R_n)$ prend n! valeurs équiprobables.

Pour chaque matrice R_n, notons R^*_n la matrice obtenue en permutant les colonnes de R_n de manière à faire apparaître sur la lère ligne les nombres 1,2,...,n dans l'ordre croissant. $S_n(R_n)$ est alors libre dans la distribution conditionnelle étant donné R^*_n ([3]). Soient $\mathscr{E}S_{1n}$ et $\mathscr{S}_{1n}$, l'espérance et la matrice variance covariance de S_{1n} sous les permutations des colonnes de R_n, ou dans la distribution conditionnelle étant donné R^*_n. Supposons $\mathscr{S}_{1n}$ inversible et considérons la forme quadratique en les composantes de S_{1n} :

$$L^2_{1n}(R_n) = (S_{1n} - \mathscr{E}S_{1n})' \, \mathscr{S}^{-1}_{1n}(S_{1n} - \mathscr{E}S_{1n}).$$

Dans la distribution conditionnelle étant donné R^*_n, $L^2_{1n}(R_n)$ est libre sous l'hypothèse nulle de bruit blanc et permet de construire un test au niveau α

Considérons à présent le vecteur S_{2n} à 2m composantes

$$S_{2n,j_1j_2} = \frac{1}{n} \sum_{\alpha=1}^{n} J\left(\frac{R^n_{J_1\alpha}}{n+1}, \frac{R^n_{J_2\alpha+1}}{n+1}\right) \quad (j_1,j_2 = 1,2,\ldots,m),$$

et la forme quadratique correspondante :

$$L^2_{2n}(R_n) = (S_{2n} - \mathscr{C} S_{2n})' \mathcal{S}^{-1}_{2n}(S_{2n} - \mathscr{C} S_{2n})$$

qui est également libre dans la distribution conditionnelle étant donné R^*_n (sous l'hypothèse nulle de bruit-blanc).

Le test basé sur L^2_{2n} peut être qualifié d'ordre deux en ce sens, qu'il considère par couple les rangs correspondants aux variables X_α, $X_{\alpha+2}$, par opposition à L^2_{1n} qui considère par couple les rangs correspondant aux variables X_α et $X_{\alpha+1}$. On peut montrer (pour plus de détail voir [2]) qu'en probabilité et sous les permutations des colonnes de R_n, L^2_{in} $(i = 1,2)$ converge en loi vers une distribution chi-carrée à 2.m degrés de liberté. On peut disposer en plus du théorème suivant :

Théorème

Pour une fonction $J(u_1,u_2)$ *p.s continue à l'intérieur de* $[0,1]^2$ *, majorée en valeur absolue par* $H(u_1)+H(u_2)$ *où* H *est de toutes puissances sommables, et pour une matrice* Σ *(voir plus loin) inversible,* L^2_{1n} *et* L^2_{2n} *sont asymptotiquement indépendants (sous l'hypothèse nulle de bruit-blanc).*

Démonstration

On a : $$L^2_{in} = \sqrt{n}(S_{in} - \mathscr{C} S_{in})' [n\, \mathcal{S}_{in}]^{-1}(S_{in} - \mathscr{C} S_{in}) \sqrt{n} \quad (i = 1,2)$$

Basons la démonstration sur le résultat suivant [2]:

Soit $J(\underline{u}_1, \underline{u}_2, \ldots, \underline{u}_q)$ une fonction définie pour $\underline{u}_\ell$ appartenant à $(0,1)^m$ $(\ell = 1,2,\ldots,q \geq 2)$. Supposons J p.s continue à l'intérieur de $[0,1]^{q.m}$ majorée en valeur absolue par $\Sigma^q_{r,\ell=1} H(u_{r,\ell})$ où H est de toutes puissances sommables. Définissons la variable

$$T_n = \frac{1}{n} \sum_{\alpha=1}^{n} J\left(\frac{R^n_\alpha}{n+1}, \frac{R^n_{\alpha+1}}{n+1}, \ldots, \frac{R^n_{\alpha+q-1}}{n+1}\right), \quad (R^n_{n+1} \equiv R_1 \text{ etc...}).$$

On peut alors affirmer qu'en probabilité et sous les permutations des colonnes de R_n, $T_n - \mathcal{C}\, T_n$ est asymptotiquement normale de moyenne nulle et d'écart-type $\sigma/\sqrt{n}$ où

$$\sigma^2 = \mathbb{D}^2 \overset{\circ}{J}(v_1,v_2,\ldots,v_q)+2\sum_{s=1}^{q-1} \mathrm{Cov}\{\overset{\circ}{J}(v_1,v_2,\ldots,v_q),\overset{\circ}{J}(v_{1+s},\ldots,v_{q+s})\}$$

avec

$$\overset{\circ}{J}(v_1,v_2,\ldots,v_q) = J(v_1,v_2,\ldots,v_q) - \sum_{k=1}^{q} \mathbb{E}[J(v_1,v_2,\ldots,v_q)\,|\,v_k]$$
$$= J(v_1,v_2,\ldots,v_q - \sum_{k=1}^{q} J_{(k)}(v_k).$$

Appliquons ce résultat au cas où

$$T_n = a\,\tilde{S}_{1n},\ b\,\tilde{S}_{2n} = \frac{1}{n}\sum_{\alpha=1}^{n}\{a\tilde{J}(\frac{R^n_\alpha}{n+1},\ \frac{R^n_{\alpha+1}}{n+1}) + b\,\tilde{J}(\frac{R^n_\alpha}{n+1},\ \frac{R^n_{\alpha+2}}{n+1})\},$$

$$\tilde{J}(\underline{u}_1,\ \underline{u}_2) = \Sigma_{1\leq \ell_1,\ell_2 \leq q}\ b_{\ell_1}\ b_{\ell_2}\ J(u_{\ell_1 1}, u_{\ell_2 2}),$$

où a,b et $b_{\ell_1\ell_2}$ sont des constantes réelles. L'expression de σ^2 est dans ce cas une forme quadratique en a et b.

Si le terme en a.b de cette forme est nul, $(S_{1n} - \mathcal{C}\, S_{1n})$ et $(S_{2n} - \mathcal{C}\, S_{2n})$ asymptotiquement normales seront indépendants. Un calcul direct montre que, d'une manière générale,

$$\sigma^2 = \mathbb{E}\ J^2+2\sum_{s=1}^{q-1} \mathbb{E}[J.J_{1+s}] - \sum_{k,\ell=1}^{q} \mathbb{E}[J.J_{(k)\ell}]+(q-1)m_o^2,$$

$$J_{1+s} = J(v_{1+s},v_{2+s},\ldots,v_{q+s}),$$

$$J_{(k)\ell} = \mathbb{E}[J(v_1,v_2,\ldots,v_{k-1},\ v_\ell,\ v_{k+1}\ ,\ldots,\ v_q)\,|\,v_\ell]\ ,$$

$$m_o = \mathbb{E}\ J(v_1,v_2,\ldots,v_q).$$

En supposant sans restreindre la généralité, $m_o = 0$, il est facile d'obtenir le terme en a.b :

$$2\,\mathbb{E}\,\tilde{J}(v_1,v_2)\,\tilde{J}(v_2,v_3) + 2\sum_{s=1}^{2}\mathbb{E}\,\tilde{J}(v_1,v_2)\,\tilde{J}(v_{1+s},v_{3+s})$$

$$+2\sum_{s=1}^{2}\mathbb{E}\,\tilde{J}(v_1,v_3)\,\tilde{J}(v_{1+s},v_{2+s})$$

$$-\sum_{\ell=1}^{3}\mathbb{E}\,\tilde{J}(v_1,v_2)\tilde{J}_{1(\ell)} - \sum_{\ell=1}^{3}\mathbb{E}\,\tilde{J}(v_1,v_2)\tilde{J}_{(2)\ell} - \sum_{\ell=1}^{3}\mathbb{E}\,\tilde{J}(v_1,v_3)\tilde{J}_{(1)\ell}$$

$$-\sum_{\ell=1}^{3}\mathbb{E}\,\tilde{J}(v_1,v_3)\tilde{J}_{(2)\ell}$$

$$= 2\,\mathbb{E}\,\tilde{J}^2_{(1)} + 2\,\mathbb{E}\,\tilde{J}_{(2)}\,J_{(1)} + 2\,\mathbb{E}\,\tilde{J}^2_1 + 2\,\mathbb{E}\,\tilde{J}_{(2)}\,\tilde{J}_{(1)} - 2\,\mathbb{E}\,\tilde{J}^2_{(1)} - 2\mathbb{E}\,\tilde{J}_{(1)}\tilde{J}_2$$

$$- 2\,\mathbb{E}\,\tilde{J}^2_1 - 2\,\mathbb{E}\,\tilde{J}_{(2)}\,\tilde{J}_{(1)} = 0.$$

Pour terminer la démonstration, il suffit d'appliquer le résultat auxiliaire ([2]) :

$$n\,\mathcal{S}_{1n} \xrightarrow{\mathbb{P}} \Sigma$$

$$n\,\mathcal{S}_{2n} \xrightarrow{\mathbb{P}} \Sigma$$

où Σ est une matrice $2m \times 2m$ d'éléments

$$\sigma_{i_1 i_2 j_1 j_2} = \operatorname{cov}\{J(v_{i_1 1}, v_{i_2 2}), J(v_{j_1 1}, v_{j_1 2})\} +$$

$$+ 2\operatorname{cov}\{J(v_{i_1 1}, v_{i_2 2}), J(v_{j_1 2}, v_{j_2 2})\} + 2\operatorname{cov}\{J(v_{j_1 1}, v_{j_2 2}), J(v_{i_1 2}, v_{i_2 3})\}$$

$$(i_1, i_2, j_1 j_2 = 1, 2, \ldots, m).$$

2. REFERENCES

[1] ANDERSON, T.W.(1971), The statistical analysis of time series, Wiley.

[2] INGENBLEEK, J.F., Normalité asymptotique d'une classe de statistiques de rang sérielles multivariées . Statistique et Analyse des Données 1981, Vol. 6 n°2 pp. 49-59

[3] PURI, M.L., SEN, P.K. (1971), Nonparametric methods in multivariate analysis, Wiley.