RANK TESTS OF PARTIAL CORRELATION

Shirahata, Shingo
Department of Statistics, Faculty of General Education, Osaka University

https://doi.org/10.5109/13144

出版情報：統計数理研究. 19 (3/4), pp.9-18, 1981-03. Research Association of Statistical Sciences

RANK TESTS OF PARTIAL CORRELATION

By

Shingo SHIRAHATA*

(Received March 24, 1980)

Abstract

Rank statistics to test the null hypothesis that X and Y are conditionally, given Z, independent are given and their asymptotic properties are investigated under the model $(X, Y, Z)=(U+a_nW, V+b_nW, W)$ where (U, V) and W are independent. It is shown that linear rank tests given by (X, Y) based on the random sample of size n are asymptotically distribution-free when $(a_n, b_n)=n^{-1/2}(a, b)$. It is also shown that Spearman's coefficient of rank correlation and Kendall's coefficient of rank correlation given by $(X-\hat{a}Z, Y-\hat{b}Z)$ are asymptotically distribution-free when $(a_n, b_n)=(a, b)$ where $(\hat{a}, \hat{b})$ is some consistent estimator of (a, b).

1. Introduction

Let (X_i, Y_i, Z_i), $i=1, 2, \cdots, n$ be a random sample of size n. Suppose we want to investigate the association between X and Y. When the hypothesis that X and Y are independent is rejected, can we say that X and Y are really correlated? This is not necessarily true since X and Y may be correlated only through the third variable Z. For example, the number of vocabulary of a child and his height are heavily correlated, but they clearly depend on his age. Thus, we need to consider the problem of testing the null hypothesis "H; X and Y are conditionally, given Z, independent of each other". When H is rejected, it can be said that X and Y are associated truely.

If (X, Y, Z) is normally distributed, it is enough to consider the partial correlation coefficient.

(1.1) $$r_{xy\cdot z}=(r_{xy}-r_{xz}r_{yz})/\{(1-r_{xz}^2)(1-r_{yz}^2)\}^{1/2},$$

see Anderson [1], where r_{xy} is the sample correlation coefficient of X and Y and r_{xz} and r_{yz} are defined similarly. However, if the distribution is unknown, we need distribution-free or at least asymptotically distribution-free procedures to test H.

AMS 1970 Subject Classifications. Primary 62G10: Secondary 62H15.

Key words and phrases. Test of partial correlation, linear rank statistics, Spearman's rank correlation, Kendall's rank correlation, Pitman efficiency.

* Department of Statistics, Faculty of General Education, Osaka University, Toyonaka, Osaka 560, Japan.

Kendall [8] proposed a rank analogue $\tau_{xy\cdot z}$ of (1.1) with using Kendall's rank correlation coefficient. Moran [10], Hoflund [6] and Maghsoodloo [9] investigated its null properties in the case that all possible sets of rankings are equally probable and Johnson [7] considered non-null properties. However, many of the works are concerned only the case that X, Y and Z are independent and is a proper subhypothesis of H. From Hoeffding [5], the asymptotic null variance of $\tau_{xy\cdot z}$ depends on the distribution even under H and hence $\tau_{xy\cdot z}$ is not asymptotically distribution-free. It is very difficult to standardize $\tau_{xy\cdot z}$ with an estimator of the variance.

Exactly or asymptotically distribution-free procedures will not be constructed without assumptions on the model of (X, Y, Z).

Let us consider the model that (X_i, Y_i, Z_i)'s satisfy

$$(X_i, Y_i, Z_i)=(U_i+a_nW_i, V_i+b_nW_i, W_i), \qquad i=1, 2, \cdots, n \tag{1.2}$$

where (U, V) and W are independent. The model (1.2) is a slight generalization of Shirahata [14] in which it is assumed that

$$(U, V)=(U^*, V^*+cU^*) \tag{1.3}$$

where U^* and V^* are indpendent. In the model (1.2) the influences of Z to X and Y are supposed to be linear and we may consider the null hypothesis "H'; U and V are independent of each other". In this paper, asymptotically distribution-free tests of H' based on ranks are considered when the nuisance parameters satisfy $(a_n, b_n)=n^{-1/2}(a, b)$ or $(a_n, b_n)=(a, b)$ for arbitrary but fixed constants a and b.

Let R_{iX}, Q_{iY}, R_{iU} and Q_{iV} be the ranks of X_i, Y_i, U_i and V_i among X's, Y's, U's and V's, respectively. Let $a_n(i)$ and $b_n(i)$ for $i=1, 2, \cdots, n$ be given constants and consider the linear rank statistics

$$S_{nXY}=\sum_{i=1}^{n} a_n(R_{iX})b_n(Q_{iY}) \tag{1.4}$$

and the random variable

$$S_{nUV}=\sum_{i=1}^{n} a_n(R_{iU})b_n(Q_{iV}). \tag{1.5}$$

In Section 2 the asymptotic equivalence of S_{nXY} and S_{nUV} is proved when $(a_n, b_n)=n^{-1/2}(a, b)$. The random variable S_{nUV} can not be observed but its asymptotic distributions are known and normal under some regularity conditions and yet it is exactly distribution-free under H'. Hence S_{nXY} which is usually adopted to test the independence of X and Y can be used as an asymptotically distribution-free test of H under (1.2) with $(a_n, b_n)=n^{-1/2}(a, b)$.

In Section 3 the case $(a_n, b_n)=(a, b)$ is considered. In this case S_{nXY} and S_{nUV} are not equivalent even in the asymptotic sense. Let $(\hat{a}, \hat{b})$ be a consistent estimator of (a, b) with order $n^{-1/2}$ and consider $(X_i^*, Y_i^*)=(X_i-\hat{a}Z_i, Y_i-\hat{b}Z_i)$. Denote by R_{iX}^* and Q_{iY}^* the rank of X_i^* and Y_i^* among X^*'s and Y^*'s, respectively. Put

$$S_{nXY}^*=\sum_{i=1}^{n} a_n(R_{iX}^*)b_n(Q_{iY}^*). \tag{1.6}$$

From the result in Section 2, it can be conjectured that S^*_{nXY} and S_{nUV} are asymptotically equivalent. Unfortunately we can not prove this for general $a_n(i)$ and $b_n(i)$ but the conjecture is found to be true in the Spearman's rank correlation case $a_n(i)=b_n(i)=i$. The equivalence also holds in the Kendall's rank correlation case although it is not a linear rank statistic.

In Section 4, the asymptotic relative efficiency of the Spearman's rank correlation case in (1.6) with respect to the usual partial correlation coefficient given by (1.1) is calculated under (1.3). It is found that the efficiency is the same with that of Spearman's rank correlation with the usual correlation coefficient in the test of independence in the bivariate model (1.3).

2. Asymptotic equivalence of S_{nXY} and S_{nUV} when $(a_n, b_n)=n^{-1/2}(a, b)$

Let us denote the model (1.2) by $M_n(a, b)$ when $(a_n, b_n)=n^{-1/2}(a, b)$. Denote, furthermore, by $H(u, v)$, $F(u)$ and $G(v)$ the distribution functions of (U, V), U and V, respectively. Put $J_n(s)=a_n(i)$ and $K_n(s)=b_n(i)$ for $(i-1)/n<s\leqq i/n$. The step functions J_n and K_n are assumed to satisfy

ASSUMPTION 2.1. There exist functions $J(s)$ and $K(s)$ such that

$$\lim_{n\to\infty} J_n(s)=J(s) \quad \text{and} \quad \lim_{n\to\infty} K_n(s)=K(s)$$

for almost everywhere s in the unit interval (0, 1).

Bhuchongkul [2] and Ruymgaart, Shorack and van Zwet [12] showed the asymptotic normality

$$n^{-1/2}(S_{nUV}-n\mu) \longrightarrow {}_d N(0, \eta^2) \tag{2.1}$$

under suitable conditions where

$$\mu=\iint J(F)K(G)dH$$

and

$$\eta^2=\mathrm{Var}\left\{J(F(U))K(G(V))+\iint(\phi_U-F)J'(F)K(G)dH+\iint(\phi_V-G)J(F)K'(G)dH\right\}$$

for $\phi_X(y)\equiv u(y-x)\equiv 1$ or 0 according as $y-x\geqq 0$ or $y-x<0$. Therefore, the asymptotic distribution of S_{nXY} is independent of (a, b) and hence the test of H' based on S_{nXY} is asymptotically distribution-free provided S_{nXY} and S_{nUV} are asymptotically equivalent. To show the equivalence, we need the following assumptions.

ASSUMPTION 2.2. $H(u, v)$ has a density function $h(u, v)$ such that there exist

$$h_1(u, v)=(\partial/\partial u)h(u, v) \quad \text{and} \quad h_2(u, v)=(\partial/\partial v)h(u, v)$$

and furthermore h_1 and h_2 satisfy

$$\iint h_1^2(u, v)/h(u, v)dudv<\infty \quad \text{and} \quad \iint h_2^2(u, v)/h(u, v)dudv<\infty .$$

ASSUMPTION 2.3. The variance of W is finite.

The result of this section is the following

THEOREM 2.1. *If Assumptions* 2.1—2.3 *hold and* $(F, G, H, J_n, K_n, J, K)$ *satisfies the assumptions of Theorem* 2.1 *of Ruymgaart, Shorack and van Zwet* [12], *then* S_{nXY} *and* S_{nUV} *are asymptotically equivalent in probability and the convergence*

$$n^{-1/2}(S_{nXY}-n\mu) \longrightarrow {}_dN(0, \eta^2) \tag{2.2}$$

holds under $M_n(a, b)$ *for arbitrary but fixed* (a, b).

PROOF. It sufficies to show that $n^{-1/2}(S_{nXY}-S_{nUV})$ converges to zero in probability since S_{nUV} enjoys the convergence (2.1). Clearly

$$P(|S_{nXY}-S_{nUV}|\neq 0)=0$$

under $M_n(0, 0)$ and hence it sufficies to show that $M_n(a, b)$ is contiguous to $M_n(0, 0)$ for any (a, b). The proof of the contiguity follows the usual method of Hájek and Šidák [4].

Define

$$\log L_n = \log \text{ likelihood ratio of } M_n(a, b) \text{ to } M_n(0, 0)$$

$$= \sum_{i=1}^{n} \log \{h(X_i-a_nZ_i,\ Y_i-b_nZ_i)/h(X_i,\ Y_i)\}\ ,$$

$$W_n=2\sum_{i=1}^{n} \{(h(X_i-a_nZ_i,\ Y_i-b_nZ_i)/h(X_i,\ Y_i))^{1/2}-1\}$$

and

$$T_n=-\sum_{i=1}^{n} (a_nZ_ih_1(X_i,\ Y_i)+b_nZ_ih_2(X_i,\ Y_i))/h(X_i,\ Y_i)\,.$$

Put

$$\sigma^2=E(W^2)E[\{(ah_1(U,\ V)+bh_2(U,\ V))/h(U,\ V)\}^2]\,.$$

Using Le Cam's second lemma in [4] and the asymptotic normality of T_n, the result will follow from the following two lemmas.

LEMMA 2.1. *If Assumptions* 2.2 *and* 2.3 *hold, then*

$$\lim_{n\to\infty} E_0(W_n)=\sigma^2/4\,. \tag{2.3}$$

Here the sign E_0 denote the expectation under $M_n(0, 0)$.

PROOF. Denote by $M(z)$ the distribution function of $Z=W$. Then

$$E_0(W_n)=-n\iiint(h^{1/2}(x,\ y)-h^{1/2}(x-a_nz,\ y-b_nz))^2dxdydM(z)\,. \tag{2.4}$$

Put $s(x,\ y)=h^{1/2}(x,\ y)$. Then (2.4) is

$$-\iiint\left[\frac{s(x,\ y)-s(x-a_nz,\ y-b_nz)}{n^{-1/2}}\right]^2 dxdydM(z)\,. \tag{2.5}$$

The function in the integrand (2.5) converges to

$$(azh_1(x,\ y)+bzh_2(x,\ y))^2/4h(x,\ y).$$

Furthermore, (2.5) is, after changing the sign,

$$\iiint n^{1/2}\left\{\int_0^{n^{-1/2}}(\partial/\partial t)s(x-azt,\ y-bzt)dt\right\}^2 dxdydM(z)$$

(2.6)
$$\leqq n^{1/2}\iiiint_0^{n^{-1/2}}((\partial/\partial t)s(x-azt,\ y-bzt))^2 dtdxdydM(z)\,.$$

Noting that

$$(\partial/\partial t)s(x-azt,\ y-bzt)=-\{azh_1(x-azt,\ y-bzt)+bzh_2(x-azt,\ y-bzt)\}$$
$$/2h^{1/2}(x-azt,\ y-bzt)\,,$$

(2.6) is

$$E(Z^2)\iint(ah_1(x,\ y)+bh_2(x,\ y))^2/4h(x,\ y)dxdy=\sigma^2/4\,.$$

Hence from the convergence theorem II 4.2 of [4], (2.3) follows.

LEMMA 2.2. *Under the same assumptions in Lemma* 2.1, *it holds that*

(2.7)
$$\lim_{n\to\infty} Var_0(T_n-W_n)=0\,.$$

PROOF. We have

$$Var_0(T_n-W_n)\leqq nE_0\{(2s(X-a_nZ,\ Y-b_nZ)/s(X,\ Y)-2+(a_nZh_1(X,\ Y)$$
$$+b_nZh_2(X,\ Y))/h(X,\ Y))^2\}$$
$$=4\iiint\left[\frac{s(x-a_nz,\ y-b_nz)-s(x,\ y)}{n^{-1/2}}\right.$$
$$\left.+\frac{azh_1(x,\ y)+bzh_2(x,\ y)}{2s(x,\ y)}\right]^2 dxdydM(z)\,.$$

From the process of the proof of Lemma 2.1 and the convergence theorem VI.3 of [4], (2.7) follows.

The theorem is also true for discontinuous J and K with modifying the variance when the conditions of [12] are replaced with that of Ruymgaart [11] or Shirahata [13]. This is because they are used to ensure only the asymptotic normality of S_{nUV} and the contiguity is guaranteed by Assumptions 2.2 and 2.3.

3. Asymptotic properties of Spearman rank correlation and Kendall rank correlation given by $(X^*,\ Y^*)$ when $(a_n,\ b_n)=(a,\ b)$

From the result of the previous section we can use S_{nXY} as an asymptotically distribution-free test of H' in the model (1.2) if $(a_n,\ b_n)=n^{-1/2}(a,\ b)$. However, when a_n and b_n are fixed, the asymptotic distribution of S_{nXY} depends on $(a_n,\ b_n)=(a,\ b)$ even under H' and the test is not asymptotically distribution-free. To correct this defect, let us consider S_{nXY}^* given by (1.6). Many estimators used to estimate $(a,\ b)$ are consistent with order $n^{-1/2}$. Therefore it can be conjectured from theorem 2.1 that S_{nXY}^* is asymptotically equivalent to S_{nUV} under some regularity conditions. The test based on S^*_{nXY} in the Spearman's rank correlation case

$$S^*_{nS}=\sum_{i=1}^{n} R^*_{iX}Q^*_{iY} \tag{3.1}$$

together with the test based on Kendall's rank correlation

$$S^*_{nK}=\sum\sum_{i\neq j} sgn(R^*_{iX}-R^*_{jX})sgn(Q^*_{iY}-Q^*_{jY}) \tag{3.2}$$

was proposed in Shirahata [14]. The estimator

$$(\hat{a},\ \hat{b})=\left(\sum_{i=1}^{n} X_i(Z_i-\bar{Z})/\sum_{i=1}^{n}(Z_i-\bar{Z})^2,\quad \sum_{i=1}^{n} Y_i(Z_i-\bar{Z})/\sum_{i=1}^{n}(Z_i-\bar{Z})^2\right) \tag{3.3}$$

was adopted and some simulation results were given in [14]. The simulation results were fairly satisfactory but the proofs were not given. Here, let us show that S^*_{nS} and S^*_{nK} are asymptotically equivalent to

$$S_{nS}=\sum_{i=1}^{n} R_{iU}Q_{iV} \tag{3.4}$$

and

$$S_{nK}=\sum\sum_{i\neq j} sgn(R_{iU}-R_{jU})sgn(Q_{iV}-Q_{jV}), \tag{3.5}$$

respectively.

We need the following assumptions.

ASSUMPTION 3.1. There exists a consistent estimator $(\hat{a},\ \hat{b})$ of $(a,\ b)$ with order $n^{-1/2}$ such that $(\hat{a},\ \hat{b})$ is symmetric with respect to $(X_i,\ Y_i,\ Z_i)$'s.

ASSUMPTION 3.2. The estimator $(\hat{a},\ \hat{b})$ satisfies that $E\{(n^{1/2}(\hat{a}-a))^8\}$ and $E\{(n^{1/2}(\hat{b}-b))^8\}$ are bounded.

ASSUMPTION 3.3. The estimator $(\hat{a},\ \hat{b})$ satisfies that $E\{(n^{1/2}(\hat{a}-a))^4\}$ and $E\{(n^{1/2}(\hat{b}-b))^4\}$ are bounded.

ASSUMPTION 3.4. U and V have respective density functions $f(u)$ and $g(v)$ such that they are boundedly first differentiable.

ASSUMPTION 3.5. The distribution function of $(U,\ V)$ is boundedly twice partially differentiable.

ASSUMPTION 3.6. $E(W^4)<\infty$.

ASSUMPTION 3.7. $E(W^2)<\infty$.

ASSUMPTION 3.8. Let $(\hat{a}_k,\ \hat{b}_k)$ be the estimator given by $(X_i,\ Y_i,\ Z_i)$, $i=k+1,\ \cdots,\ n$. Then for any fixed k

(i) $\hat{a}_k-\hat{a}=O_p(n^{-1})$ and $\hat{b}_k-\hat{b}=O_p(n^{-1})$,

(ii) For any event N which is concerning to $\hat{a}$, $\hat{a}_k$ and $(U_i,\ V_i,\ W_i)$, $i=1,\ \cdots,\ k-1$ and random variable I symmetric around the origin given by $(U_i,\ V_i,\ W_i)$, $i=1,\ \cdots,\ k-1$, it holds that

$$P(N,\ \hat{b}_k-b<I<\hat{b}-b)-P(N,\ \hat{b}-b<I<\hat{b}_k-b)=o(n^{-1}). \tag{3.6}$$

(iii) The same fact as in (ii) holds with replacing $(\hat{a},\ \hat{a}_k)$ and $(\hat{b},\ \hat{b}_k)$ with $(\hat{b},\ \hat{b}_k)$ and $(\hat{a},\ \hat{a}_k)$, respectively.

Assumption 3.8(i) is a natural one for an estimator. The estimator (3.3) satisfies it if $E(U^2)$, $E(V^2)$ and $E(W^2)$ are finite. Assumption 3.8 (ii) and (iii) are rather technical and are very hard to check. However, considering the fact that each term in (3.6) is $O(n^{-1})$, it is not so curious.

The results in this section are the following theorems.

THEOREM 3.1. *When the model* (1.2) *is satisfied with* $(a_n, b_n)=(a, b)$ *and Assumptions* 3.1, 3.2, 3.4, 3.6 *and* 3.8 *hold, then* S_{nS}^* *and* S_{nS} *given by* (3.1) *and* (3.4), *respectively are asymptotically equivalent in probability. More precisely, it holds that*

$$\lim_{n\to\infty} n^{-5}E\{(S_{nS}^*-S_{nS})^2\}=0 . \tag{3.7}$$

THEOREM 3.2. *When the model* (1.2) *is satisfied with* $(a_n, b_n)=(a, b)$ *and Assumptions* 3.1, 3.3, 3.5, 3.7 *and* 3.8 *hold, then* S_{nK}^* *and* S_{nK} *given by* (3.2) *and* (3.5), *respectively are asymptotically equivalent in probability. More precisely, it holds that*

$$\lim_{n\to\infty} n^{-3}E\{(S_{nK}^*-S_{nK})^2\}=0 . \tag{3.8}$$

PROOF of the theorems. The proof of (3.8) is very similar to that of (3.7) and hence we shall prove only (3.7). It is easy to show that

$$\begin{aligned} n^{-5}E\{(S_{nS}-S_{nS}^*)^2\} \leqq\ & 2n^{-2}E\{(R_{1U}-R_{1X}^*)^2+(Q_{1V}-Q_{1Y}^*)^2\} \\ & +n^{-4}(n-1)E\{(R_{1U}Q_{1V}-R_{1X}^*Q_{1Y}^*)(R_{2U}Q_{2V}-R_{2X}^*Q_{2Y}^*)\} . \end{aligned} \tag{3.9}$$

Now

$$\begin{aligned} R_{1U}-R_{1X}^* &= \sum_{i=1}^{n} \{u(U_1-U_i)-u(U_1-U_i+(a-\hat{a})(W_1-W_i))\} \\ &\equiv \sum_{i=1}^{n} A_i , \qquad \text{say,} \end{aligned}$$

where $u(x)=1$ or 0 according as $x\geqq 0$ or $x<0$. Thus,

$$\begin{aligned} E\{(R_{1U}-R_{1X}^*)^2\} &\leqq n^2E(A_2^2) \\ &= n^2\{P((a-\hat{a})(W_1-W_2)>U_1-U_2>0) \\ &\quad + P((a-\hat{a})(W_1-W_2)<U_1-U_2<0)\} . \end{aligned} \tag{3.10}$$

From Assumption 3.1, $a-\hat{a}=O_p(n^{-1/2})$ and hence (3.10) is $O(n^{3/2})$. Similarly we have $E\{(Q_{1V}-Q_{1Y}^*)^2\}=O(n^{3/2})$. Therefore the first term of (3.9) is $O(n^{-1/2})$.

On the other hand, the second term of (3.9) except the multiplicative constant is

$$\begin{aligned} & \sum_{i,j,k,m}\sum\sum\sum E[\{u(U_1-U_i)u(V_1-V_j)-u(X_1^*-X_i^*)u(Y_1^*-Y_j^*)\} \\ & \quad \cdot\{u(U_2-U_k)u(V_2-V_m)-u(X_2^*-X_k^*)u(Y_2^*-Y_m^*)\}] \\ & \equiv \sum_{i,j,k,m}\sum\sum\sum A_{ijkm} , \qquad \text{say} . \end{aligned} \tag{3.11}$$

It can be shown that (3.11) is

$$(n-2)(n-3)(n-4)(n-5)A_{3456}+O(n^{5/2}) .$$

Therefore, we may show that $A_{3456}=o(n^{-1})$. Now

$$A_{3456}=E\{u(U_1-U_3)u(V_1-V_4)u(U_2-U_5)u(V_2-V_6)\}$$
$$-2E\{u(U_1-U_3)u(V_1-V_4)u(U_2-U_5+(a-\hat{a})(W_2-W_5))$$
$$\times u(V_2-V_6+(b-\hat{b})(W_2-W_6))\}$$
$$+E\{u(U_1-U_3+(a-\hat{a})(W_1-W_3))u(V_1-V_4+(b-\hat{b})(W_1-W_4))$$
$$\times u(U_2-U_5)+(a-\hat{a})(W_2-W_5))u(V_2-V_6+(b-\hat{b})(W_2-W_6))\}\ . \tag{3.12}$$

The first term of (3.12) is

$$\left(\iint F(u)G(v)dH(u,\ v)\right)^2 . \tag{3.13}$$

Consider $(\hat{a}_6,\ \hat{b}_6)$ defined in Assumption 3.8 and denote the distribution function of $n^{1/2}(a-\hat{a}_6,\ b-\hat{b}_6)$ by Q_n. Then

$$E\{u(U_1-U_3)u(V_1-V_4)u(U_2-U_5+(a-\hat{a}_6)(W_2-W_5))u(V_2-V_6+(b-\hat{b}_6)(W_2-W_6))\}$$
$$=\iint FG\,dH\iiiiii F(u+n^{-1/2}tw_1)G(v+n^{-1/2}sw_2)dH(u,\ v)dM(w_1,\ w_2)dQ_n(t,\ s) \tag{3.14}$$

where $M(w_1,\ w_2)$ is the distribution function of $(W_2-W_5,\ W_2-W_6)$.

Denote the events $\{U_1-U_3>0,\ V_1-V_4>0\}$, $\{U_2-U_5+(a-\hat{a})(W_2-W_5)>0\}$, $\{V_2-V_6+(b-\hat{b})(W_2-W_6)>0\}$, $\{U_2-U_5+(a-\hat{a}_6)(W_2-W_5)>0\}$ and $\{V_2-V_6+(b-\hat{b}_6)(W_2-W_6)>0\}$ by B, C, D, C_6 and D_6, respectively. Then the second term of (3.12) except the constant -2 minus (3.14) is

$$\{P(B,\ C,\ D,\ C_6^c,\ D_6^c)-P(B,\ C^c,\ D^c,\ C_6,\ D_6)\}+\{P(B,\ C,\ D,\ C_6,\ D_6^c)$$
$$-P(B,\ C,\ D^c,\ C_6,\ D_6)\}+\{P(B,\ C,\ D,\ C_6^c,\ D_6)-P(B,\ C^c,\ D,\ C_6,\ D_6)\}\ . \tag{3.15}$$

The first term of (3.15) is $O(n^{-2})$ from Assumption 3.8(i). The second and the third terms of (3.15) are, from Assumption 3.8(ii) and (iii), $o(n^{-1})$. Hence we may consider (3.14) in the place of the second term of (3.12).

As in the second term, the third term of (3.12) can be replaced by

$$E\{u(U_1-U_3+(a-\hat{a}_6)(W_1-W_3))u(V_1-V_4+(b-\hat{b}_6)(W_1-W_4))$$
$$\times u(U_2-U_5+(a-\hat{a}_6)(W_2-W_5))u(V_2-V_6+(b-\hat{b}_6)(W_2-W_6))\}$$
$$=\iint\left\{\iiiint F(u+n^{-1/2}tw_1)G(v+n^{-1/2}sw_2)dH(u,\ v)dM(w_1,\ w_2)\right\}^2 dQ_n(t,\ s)\ . \tag{3.16}$$

Combining (3.13), (3.14) and (3.16), we have

$$A_{3456}=\iint\left\{\iiiint(F(u+n^{-1/2}tw_1)G(v+n^{-1/2}sw_2)-F(u)G(v))\right.$$
$$\left.\cdot dH(u,\ v)dM(w_1,\ w_2)\right\}^2 dQ_n(t,\ s)+o(n^{-1})\ . \tag{3.17}$$

Taylor expansions give

$$F(u+n^{-1/2}tw_1)=F(u)+n^{-1/2}tw_1f(u)+O(n^{-1})$$

and

$$G(v+n^{-1/2}sw_2)=G(v)+n^{-1/2}sw_2g(v)+O(n^{-1})\,.$$

Using the facts $\iint w_1dM(w_1,\ w_2)=\iint w_2dM(w_1,\ w_2)=0$ and the above expansions, it is shown that the formula in the square in (3.17) is $O(n^{-1})$ and hence $A_{3456}=o(n^{-1})$. Thus, (3.7) is established.

Unfortunately we can not prove Theorem 3.1 for general linear rank statistics $(S^*_{nXY},\ S_{nUV})$. The calculations in the proof of Theorem 3.1 depend on the fact that the statistic is of Spearman. However, the theorem will hold for fairly general rank statistics. It is also conjectured that the same conclusion holds for a more general model $(X,\ Y,\ Z)=(U+a_n(W),\ V+b_n(W),\ W)$ if the functions $a_n(\cdot)$ and $b_n(\cdot)$ allow consistent estimators of order $n^{-1/2}$.

4. Asymptotic relative efficiency

In the usual normal distribution theory, the partial correlation coefficient $r_{xy\cdot z}$ given by (1.1) is used to test H. In this section we shall compare S^*_{nS} given by (3.1) with $r_{xy\cdot z}$ in the model (1.2) with additional assumption (1.3). The random variable U and V are positively or negatively correlated according as $c>0$ or $c<0$ and they are independent when $c=0$. Let us consider the one-sided alternative "$c>0$". Then the critical regions of the tests are of the forms $\{S^*_{nS}\geqq\text{constant}\}$ and $\{r_{xy\cdot z}\geqq\text{constant}\}$. For simplicity the distribution of $(a-\hat{a},\ b-\hat{b})$ is assumed to be independent of $(a,\ b)$. The estimator (3.3) has this property. Then, since $r_{xy\cdot z}$ and S^*_{nS} are location-free, we may assume that $EU^*=EV^*=EW=0$ without loss of generality.

Denote by F, G and G^* the distribution functions of $U=U^*$, V and V^* and by f and g^* the density functions of U^* and V^*, respectively. Furthermore, let us assume that every regularity conditions such as the exchangeability of the differentiation and the integration and the continuity of the asymptotic variances of test statistics at $c=0$ are satisfied. Then $n^{-5/2}(S^*_{nS}-n(n+1)^2/4)$ is asymptotically normal with mean

$$m_c(S^*_{nS})=n^{1/2}\iint\Big(F(u)-\frac{1}{2}\Big)\Big(G^*(v)-\frac{1}{2}\Big)f(u)g^*(v-cu)dudv \tag{4.1}$$

and null variance 1/144. Differentiate (4.1) with c, we have

$$(d/dc)m_c(S^*_{nS})/_{c=0}=n^{1/2}\int(g^*(v))^2dv\int uF(u)dF(u)\,. \tag{4.2}$$

Here we used the assumption $EU^*=0$. Hence the efficacy of S^*_{nS} is

$$e(S^*_{nS})=144\Big\{\int(g^*(v))^2dv\int uF(u)dF(u)\Big\}^2\,. \tag{4.3}$$

On the other hand, $n^{1/2}r_{xy\cdot z}$ is asymptotically equivalent to

$$n^{1/2}\Big\{n^{-1}\sum_{i=1}^n X_iY_i-n^{-2}\sum_{i=1}^n X_iW_i\sum_{i=1}^n Y_iW_i/Var(W)\Big\}/(Var(U)Vav(V))^{1/2}$$

which is asymptotically

$$n^{-3/2}(Var(U)Var(V)Var(W))^{-1/2} \cdot \sum\sum_{i<j}(U_iV_iW_j^2+U_jV_jW_i^2-U_iV_jW_iW_j-U_jV_iW_iW_j)\,. \tag{4.4}$$

From U-statistic theory in Hoeffding [5], (4.4) is asymptotically normal with mean $n^{1/2}Cor(U, V)$ and unit variance where by Cor we denote the correlation coefficient. Simple calculation shows that the efficacy of $r_{xy\cdot z}$ is

$$e(r_{xy\cdot z})=(Var(U^*)Var(V^*))^{1/2}\,. \tag{4.5}$$

Therefore, from (4.3) and (4.5), Pitman efficiency of S_{nS}^* with respect to $r_{xy\cdot z}$ is

$$e(S_{nS}^*, r_{xy\cdot z})=144\,Var(V^*)\left(\int(g^*(v))^2dv\int uF(u)dF(u)\right)^2/Var(U^*)\,. \tag{4.6}$$

It is worth noted that (4.6) is the same with Pitman efficiency of Spearman rank correlation with respect to the usual correlation coefficient for the bivariate population model (1.3). This can be established by easy calculations. Also note that it is enough to consider the case $Var(U^*)=Var(V^*)=1$ to calculate (4.6). If the underlying distribution of (X, Y, Z) is normal, (4.6) is found to be $9/\pi^2$.

References

[1] ANDERSON, T. W.: An Introduction to Multivariate Analysis. (1958). John Wiley and Sons, New York.

[2] BHUCHONGKUL, S.: *A class of nonparametric tests for independence in bivariate populations.* Ann. Math. Statist., **35** (1964), 138-149.

[3] GOODMAN, L. A.: *Partial tests for partial taus.* Biometrika, **46** (1959), 425-432.

[4] HÁJEK, J. and ŠIDÁK, Z.: Theory of Rank Tests. (1967). Academic Press, New York.

[5] HOEFFDING, W.: *A class of statistics with asymptotically normal distribution.* Ann. Math. Statist., **19** (1948), 293-325.

[6] HOFLUND, O.: *Simulated distributions for small n of Kendall's partial rank correlation coefficient.* Biometrika, **50** (1963), 520-522.

[7] JOHNSON, N. S.: *Nonnull properties of Kendall's partial rank correlation coefficient.* Biometrika, **66** (1979), 333-337.

[8] KENDALL, M. G.: *Partial rank correlation.* Biometrika, **32** (1942), 277-283.

[9] MAGHSOODLOO, S.: *Estimates of the quantiles of Kendall's partial rank correlation coefficient.* J. Statist. Comput. Simul., **4** (1975), 155-164.

[10] MORAN, P. A. P.: *Partial and multiple rank correlation.* Biometrika, **38** (1951), 26-32.

[11] RUYMGAART, F. H.: *Asymptotic normality of nonparametric tests for independence.* Ann. Statist., **2** (1974), 892-910.

[12] RUYMGAART, F. H., SHORACK, G. R. and VAN ZWET, W. R.: *Asymptotic normality of nonparametric tests for independence.* Ann. Math. Statist., **43** (1972), 1122-1135.

[13] SHIRAHATA, S.: *Locally most powerful rank tests for independence with censored data.* Ann. Statist., **3** (1975), 241-245.

[14] SHIRAHATA, S.: *Tests of partial correlation in a linear model.* Biometrika, **64** (1977), 162-164.