

# On Blest's measure of rank correlation

Christian GENEST and Jean-François PLANTE

**Key words and phrases:** Concordance measure; copula; correlation; Kendall's tau; Pitman's asymptotic relative efficiency; Spearman's rho; tests of independence.

**MSC 2000:** Primary 62H15, 62H20; Secondary 62G20, 62G30.

**Abstract:** Blest (2000) proposed a new nonparametric measure of correlation between two random variables. His coefficient, which is dissymmetric in its arguments, emphasizes discrepancies observed among the first ranks in the orderings induced by the variables. The authors derive the limiting distribution of Blest's index and suggest symmetric variants whose merits as statistics for testing independence are explored using asymptotic relative efficiency calculations and Monte Carlo simulations.

## À propos de la mesure de corrélation des rangs de Blest

**Résumé :** Blest (2000) a proposé une nouvelle mesure non paramétrique de corrélation entre deux variables aléatoires. Son coefficient, qui est asymétrique en ses arguments, met l'accent sur les écarts observés dans les premiers rangs des classements induits par les aléas. Les auteurs déterminent la loi limite de l'indice de Blest et en suggèrent des variantes symétriques dont ils explorent les mérites en tant que statistiques de tests d'indépendance au moyen de calculs d'efficacité relative asymptotique et d'une étude de Monte-Carlo.

## 1. INTRODUCTION

Although Pearson's correlation coefficient is one of the most ubiquitous concepts in the scientific literature, it is now widely recognized, at least among statisticians, that the degree of stochastic dependence between continuous random variables  $X$  and  $Y$  with joint cumulative distribution function  $H$  and marginals  $F$  and  $G$  is much more appropriately characterized and measured in terms of the joint distribution  $C$  of the pair  $(F(X), G(Y))$ , namely

$$C(u, v) = P\{F(X) \leq u, G(Y) \leq v\}, \quad 0 \leq u, v \leq 1$$

whose marginals are uniform on the interval  $[0, 1]$ . Indeed, most modern concepts and measures of dependence, not to mention stochastic orderings (see, for instance, Joe 1997, Nelsen 1999 or Drouet-Mari & Kotz 2001), are functions of the so-called “copula”  $C$ , which is uniquely determined on  $\text{Range}(F) \times \text{Range}(G)$  in general, and hence everywhere in the special case where  $F$  and  $G$  are continuous, as will be assumed henceforth.

In particular, classical nonparametric measures of dependence such as Spearman's rho

$$\begin{aligned} \rho(X, Y) &= 12 \int_{\mathbf{R}^2} F(x)G(y) dH(x, y) - 3 \\ &= 12 \int_{[0,1]^2} uv dC(u, v) - 3 = 12 \int_{[0,1]^2} C(u, v) dv du - 3 \end{aligned}$$

and Kendall's tau

$$\tau(X, Y) = 4 \int_{\mathbf{R}^2} H(x, y) dH(x, y) - 1 = 4 \int_{[0,1]^2} C(u, v) dC(u, v) - 1$$

are superior to Pearson's coefficient in that while they vanish when the variables are independent, they always exist and take their extreme values  $\pm 1$  when  $X$  and  $Y$  are in perfect monotone functional dependence, that is, when either  $Y = G^{-1}\{F(X)\}$  or  $Y = G^{-1}\{1 - F(X)\}$  with probability one. Except in special circumstances (as when  $F$  and  $G$  are members of the same

location-scale family, for example), these cases of complete dependence are not instances of *linear* dependence. Thus when considering random data  $(X_1, Y_1), \dots, (X_n, Y_n)$  from an unknown distribution  $H$  whose support is  $[0, \infty)^2$ , for instance, the values of  $\rho$  and  $\tau$  are unconstrained, whereas Pearson's correlation can only span an interval  $[r, s]$  whose bounds  $-1 < r \leq s \leq 1$  depend on the choice of marginals  $F$  and  $G$ . When the latter are unknown, it is thus difficult to know what to make of an observed Pearson correlation of  $-0.2$ , say. For additional discussion on the advantages of rank-based measures of dependence over Pearson's traditional correlation coefficient, the reader may refer to the nice survey paper written by Embrechts, McNeil & Straumann (2002).

Because copulas are margin free and ranks are maximally invariant statistics of the observations under monotone transformations of the marginal distributions,  $\rho$ ,  $\tau$  and indeed all other copula-based measures of dependence (for example, the index of Schweizer & Wolff 1981) should be estimated by functions of the ranks  $R_i$  and  $S_j$  of the  $X_i$  and the  $Y_j$ . Letting  $F_n$  and  $G_n$  stand for the empirical distribution functions of  $X$  and  $Y$  respectively, the classical estimate for  $\rho$  is

$$\rho_n = \frac{12}{n^3 - n} \sum_{i=1}^n R_i S_i - 3 \frac{n+1}{n-1},$$

namely the Pearson correlation between the components of the pairs  $(F_n(X_i), G_n(Y_i)) = (R_i/n, S_i/n)$ ,  $1 \leq i \leq n$ . Likewise,  $\tau$  is traditionally estimated by the scaled difference in the numbers of concordant and discordant pairs, or equivalently by

$$\tau_n = \frac{2}{n^2 - n} \sum_{1 \leq i < j \leq n} \text{sign}(R_i - R_j) \text{sign}(S_i - S_j).$$

In comparing  $\rho_n$  and  $\tau_n$  in terms of their implicit weighing of differences  $R_i - S_i$ , Blest (2000) was led to propose an alternative measure of rank correlation that "... attaches more significance to the early ranking of an initially given order." Assume, for example, that  $X_i$  and  $Y_i$  represent the running times of sprinter  $i = 1, \dots, n$  in two successive track-and-field meetings. The correlation in the pairs  $(R_i, S_i)$  then gives an idea of the consistency between the two rankings. However, differences in the top ranks would seem to be more critical, in that they matter in awarding medals. As a result, Blest suggests that these discrepancies should be emphasized, whereas all rank reversals are given the same importance in Spearman's or Kendall's coefficient.

To be specific, Blest's index is defined by

$$\nu_n = \frac{2n+1}{n-1} - \frac{12}{n^2 - n} \sum_{i=1}^n \left(1 - \frac{R_i}{n+1}\right)^2 S_i.$$

The constants are rigged so that the coefficient varies between 1 and  $-1$  and is most extreme when the rankings coincide ( $S_i = R_i$ ) or are antithetic ( $S_i = n+1 - R_i$ ). Numerical results and calculations reported by Blest (2000) indicate that his measure can discriminate more easily between individual permutations than either  $\rho_n$  or  $\tau_n$  while being highly correlated with both of them. Furthermore, partial evidence is provided which indicates that the large-sample distribution of  $\nu_n$  is normal.

The first objective of this paper is to show that  $\nu_n$  is an asymptotically unbiased estimator of the population parameter

$$\nu(X, Y) = 2 - 12 \int_{\mathbf{R}^2} \{1 - F(x)\}^2 G(y) dH(x, y) = 2 - 12 \int_{[0,1]^2} (1-u)^2 v dC(u, v) \quad (1)$$

and that  $\sqrt{n}(\nu_n - \nu)$  converges in distribution to a centered normal random variable whose variance is specified in Section 2. While  $\nu_n$  may be appropriate as an index of discrepancy

between two rankings, it is pointed out in Section 3 that since in general  $\nu(X, Y) \neq \nu(Y, X)$ , the parameter (1) estimated by  $\nu_n$  is not a measure of concordance, in the sense of Scarsini (1984). The properties of a symmetrized version of  $\nu_n$  are considered in Section 4, and the relative merits of this new statistic (and a natural complement thereof) for testing independence are then examined in Section 5 through asymptotic relative efficiency calculations and a small Monte Carlo simulation study. Brief concluding comments are given in Section 6 and the Appendix contains some technical details.

## 2. LIMITING DISTRIBUTION OF BLEST'S COEFFICIENT

Define  $J(t) = (1-t)^2$  and  $K(t) = t$  for all  $0 \leq t \leq 1$ , and for arbitrary integers  $n$  and  $i \in \{1, \dots, n\}$ , let  $J_n(t) = J\{i/(n+1)\}$  and  $K_n(t) = K\{i/(n+1)\}$  if  $(i-1)/n < t \leq i/n$ . The large-sample behaviour of Blest's sample measure of association is clearly the same as that of

$$2 - \frac{12}{n} \sum_{i=1}^n \left(1 - \frac{R_i}{n+1}\right)^2 \frac{S_i}{n+1},$$

which may be written as

$$2 - 12 \int J_n(F_n) K_n(G_n) dH_n$$

in terms of the empirical cumulative distribution function  $H_n$  of  $H$  and its marginals  $F_n$  and  $G_n$ . Since the functions  $J$  and  $K$  and their derivatives are bounded on  $[0, 1]$ , and in view of Remark 2.1 of Ruymgaart, Shorack & van Zwet (1972), a direct application of Theorem 2.1 of these authors implies that  $\sqrt{n}(\nu_n - \nu)$  is asymptotically normally distributed with mean and variance as specified below.

**PROPOSITION 1.** *Under the assumption of random sampling from a continuous bivariate distribution  $H$  with underlying copula  $C$ ,  $\sqrt{n}(\nu_n - \nu)$  converges weakly, as  $n \rightarrow \infty$ , to a normal random variable with zero mean and the same variance as*

$$12 \left[ (1-U)^2 V - 2 \int_U^1 (1-u) E(V | U=u) du + \int_V^1 E\{(1-U)^2 | V=v\} dv \right], \quad (2)$$

where the pair  $(U, V)$  is distributed as  $C$ . In particular, the variance of the latter expression equals  $16/15$  when  $U$  and  $V$  are independent.

As is the case for Spearman's rho, compact algebraic formulas for  $\nu$  can be found for relatively few models. In two special cases of interest, Examples 1 and 2 illustrate the explicit calculations that can sometimes be made using a symbolic calculator such as MAPLE; see Plante (2002) for details. Example 3, which concerns the pervasive normal model, is somewhat more subtle.

**Example 1.** Suppose that  $(X, Y)$  follows a Farlie–Gumbel–Morgenstern distribution with marginals  $F$  and  $G$ , namely

$$H_\theta(x, y) = F(x)G(y) + \theta F(x)G(y)\{1 - F(x)\}\{1 - G(y)\}, \quad x, y \in \mathbb{R},$$

with parameter  $\theta \in [-1, 1]$ . Then

$$\nu_\theta(X, Y) = \rho_\theta(X, Y) = \frac{3}{2} \tau_\theta(X, Y) = \frac{\theta}{3}$$

and the variance of (2) equals  $16/15 - 16\theta^2/63$ .

**Example 2.** For arbitrary  $a, b \in [0, 1]$  and  $\theta \in [-1, 1]$ , let

$$\begin{aligned} H_{\theta,a,b}(x, y) &= F(x)G(y)[1 + \theta\{1 - F(x)^a\}\{1 - G(y)^b\}] \\ &= \Pi\{F(x)^{1-a}, G(y)^{1-b}\}C_\theta\{F(x)^a, G(y)^b\}, \quad x, y \in \mathbb{R}, \end{aligned}$$

where  $C_\theta(u, v) = H_\theta\{F^{-1}(u), G^{-1}(v)\}$  is the Farlie–Gumbel–Morgenstern copula and  $\Pi(u, v) = uv$  denotes the independence copula. By Khoudraji's device (see Genest, Ghoudi & Rivest 1998),  $H_{\theta,a,b}$  is an asymmetric extension of the Farlie–Gumbel–Morgenstern distribution, recently considered in a somewhat more general form by Bařamov, Kotz & Bekçi (2000). For any pair  $(X, Y)$  distributed as  $H_{\theta,a,b}$ , one finds

$$\nu_{\theta,a,b}(X, Y) = \frac{2ab(a+5)\theta}{(a+2)(a+3)(b+2)},$$

while

$$\rho_{\theta,a,b}(X, Y) = \frac{3ab\theta}{(a+2)(b+2)} = \frac{3}{2}\tau_{\theta,a,b}(X, Y).$$

Note that

$$\nu_{\theta,a,b}(X, Y) \neq \nu_{\theta,a,b}(Y, X) = \nu_{\theta,b,a}(X, Y),$$

unless of course  $a = b$ , in which case the copula is symmetric in its arguments. An explicit but long formula (not shown here) for the asymptotic variance of  $\sqrt{n}(\nu_n - \nu)$  is also available in this case.

**Example 3.** Suppose that  $(X, Y)$  has a bivariate normal distribution and that  $\text{corr}(X, Y) = r \in [-1, 1]$ . Then

$$\nu_r(X, Y) = \rho_r(X, Y) = \frac{6}{\pi} \arcsin\left(\frac{r}{2}\right),$$

while

$$\tau_r(X, Y) = \frac{2}{\pi} \arcsin(r).$$

The formulas for  $\rho_r(X, Y)$  and  $\tau_r(X, Y)$  are standard normal theory; see, for instance, Exercise 2.14 of Joe (1997, p. 54). Unfortunately, the variance of (2) can only be computed numerically in this case.

The fact that  $\nu(X, Y) = \rho(X, Y)$ , as in the third example, arises whenever a copula  $C$  is radially symmetric, that is, when its associated survival function  $\bar{C}(u, v) = 1 - u - v + C(u, v)$  satisfies the condition

$$\bar{C}(u, v) = C(1 - u, 1 - v), \quad 0 \leq u, v \leq 1. \tag{3}$$

Indeed, a simple calculation shows that in general, one has

$$\nu(X, Y) = \rho(X, Y) - 6E\{U^2V - (1 - U)^2(1 - V)\}$$

with  $(U, V)$  distributed as  $C$ , the underlying copula of the pair  $(X, Y)$ . Clearly, the second summand vanishes when  $C$  is radially symmetric, since condition (3) implies that the pairs  $(U, V)$  and  $(1 - U, 1 - V)$  have the same distribution.

Additional properties of  $\nu(X, Y)$  are described next.

### 3. PROPERTIES OF BLEST'S INDEX

According to Scarsini (1984), the following are fundamental properties that any measure  $\kappa$  of concordance should satisfy:

- (a)  $\kappa$  is defined for every pair  $(X, Y)$  of continuous random variables.
- (b)  $-1 \leq \kappa(X, Y) \leq 1$ ,  $\kappa(X, X) = 1$ , and  $\kappa(X, -X) = -1$ .
- (c)  $\kappa(X, Y) = \kappa(Y, X)$ .
- (d) If  $X$  and  $Y$  are independent, then  $\kappa(X, Y) = 0$ .
- (e)  $\kappa(-X, Y) = \kappa(X, -Y) = -\kappa(X, Y)$ .
- (f) If  $(X, Y) \prec (X^*, Y^*)$  in the positive quadrant dependence ordering, then  $\kappa(X, Y) \leq \kappa(X^*, Y^*)$ .
- (g) If  $(X_1, Y_1), (X_2, Y_2), \dots$  is a sequence of continuous random vectors that converges weakly to a pair  $(X, Y)$ , then  $\kappa(X_n, Y_n) \rightarrow \kappa(X, Y)$  as  $n \rightarrow \infty$ .

It is well known that both  $\rho$  and  $\tau$  meet all these conditions, and it is easy to check that the index  $\nu$  defined in (1) has properties (a), (b), (d), (f), and (g). To show the latter two, it is actually more convenient to use the alternative representation

$$\nu(X, Y) = -2 + 24 \int_{[0,1]^2} (1-u)C(u, v) \, du \, dv,$$

which follows immediately from the general identity

$$\int_{[0,1]^2} K(u, v) \, dC(u, v) - \int_{[0,1]^2} K(u, v) \, du \, dv = \int_{[0,1]^2} \{C(u, v) - uv\} \, dK(u, v) \quad (4)$$

shown by Quesada-Molina (1992) to be valid for all right-continuous and quasi-monotone functions  $K: [0, 1]^2 \rightarrow \mathbb{R}$ . The appropriate choice here is  $K(u, v) = (1-u)^2v$ , while taking  $K(u, v) = uv$  yields the standard form of Hoeffding's identity.

As was already pointed out in Example 2, however, Blest's measure does not satisfy condition (c). The same example shows that property (e) is not verified either. In fact, if  $(X, Y)$  has copula  $C$ , then  $C^*(u, v) = v - C(1-u, v)$  is the copula associated with the pair  $(-X, Y)$ , and hence

$$\nu(-X, Y) = \nu(X, Y) - 2\rho(X, Y), \quad (5)$$

so that  $\nu(-X, Y) \neq -\nu(X, Y)$  except in the special case where  $\nu(X, Y) = \rho(X, Y)$ , as when  $C$  is radially symmetric, for instance. Note, however, that  $\nu(X, -Y) = -\nu(X, Y)$  by a similar argument involving the copula  $C^{**}(u, v) = u - C(u, 1-v)$  of the pair  $(X, -Y)$ .

A second illustration of the failure of condition (e) for Blest's index is provided below.

**Example 4.** Suppose that  $X$  is uniform on the interval  $[0, 1]$  and that either

$$Y = \begin{cases} \theta - X & \text{if } 0 \leq X < \theta, \\ X & \text{if } \theta \leq X \leq 1. \end{cases} \quad (6)$$

or

$$Y = \begin{cases} X & \text{if } 0 \leq X \leq 1 - \theta, \\ 2 - \theta - X & \text{if } 1 - \theta < X \leq 1. \end{cases} \quad (7)$$

for some constant  $0 \leq \theta \leq 1$ . The joint distribution of the pair  $(X, Y)$  is then a “shuffle of min” copula in the sense of Mikusiński, Sherwood & Taylor (1992), and the probability mass of this singular distribution is spread uniformly on two line segments depending on  $\theta$ , as illustrated in Figure 1.

When  $(X, Y)$  is distributed as (6), one finds  $\nu_\theta(X, Y) = 1 - 4\theta^3 + 2\theta^4$ , while if  $(X, Y)$  is distributed as (7), one gets  $\nu_\theta(X, Y) = 1 - 2\theta^4$ .

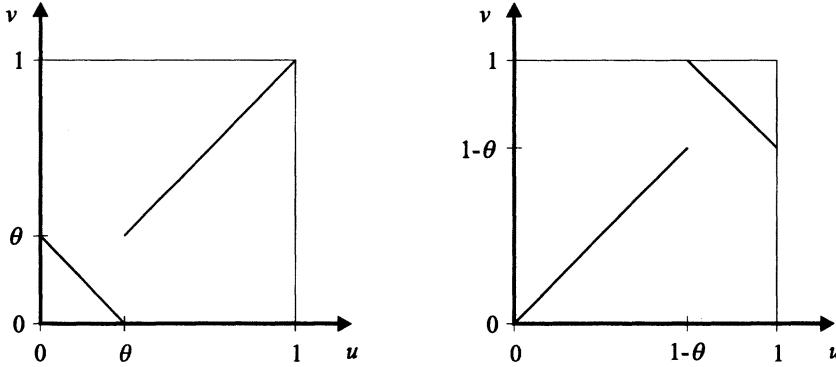


FIGURE 1: Support of the “shuffle of min” copulas (6) and (7) discussed in Example 4.

Had Blest’s coefficient met condition (e) above, the two expressions would have been equal, since the two supports displayed in Figure 1 are equivalent up to a rotation of 180 degrees about the point  $(1/2, 1/2)$ , that is, a reflection through the line  $u = 1/2$  followed by another reflection through the line  $v = 1/2$ .

#### 4. A SYMMETRIZED VERSION OF BLEST’S MEASURE OF ASSOCIATION

In the light of Example 4, Scarsini’s axiom (e) is clearly incompatible with the notion that a nonparametric measure of dependence  $\kappa$  could emphasize differences in early ranks. Indeed, if  $(U, V)$  and  $(U^*, V^*)$  follow distributions (6) and (7) with the same parameter, say  $\theta < 1/2$ , one would then expect to have  $\kappa(U, V) < \kappa(U^*, V^*)$ , since it is plain from Figure 1 that reversals occur in small ranks under (6), while they occur in large ranks under (7). Should property (e) hold true, however, it would then follow that

$$\kappa(U, V) < \kappa(U^*, V^*) = \kappa(-U^*, -V^*),$$

but this is impossible since  $\kappa(-U^*, -V^*) = \kappa(1 - U^*, 1 - V^*) = \kappa(U, V)$  because these three pairs of variables have the same underlying copula.

By contrast, it is relatively easy to adapt Blest’s suggestion in order to meet Scarsini’s symmetry requirement (c) in addition to conditions (a), (b), (d), (f) and (g). Specifically, let

$$\tilde{\nu}_n = \frac{2n+1}{n-1} - \frac{12}{n^2-n} \sum_{i=1}^n \left(1 - \frac{S_i}{n+1}\right)^2 R_i$$

and consider the empirical rank coefficient

$$\xi_n = \frac{\nu_n + \tilde{\nu}_n}{2} = -\frac{4n+5}{n-1} + \frac{6}{n^3-n} \sum_{i=1}^n R_i S_i \left(4 - \frac{R_i + S_i}{n+1}\right) \quad (8)$$

and its theoretical counterpart

$$\xi(X, Y) = \frac{\nu(X, Y) + \nu(Y, X)}{2} = -4 + 6 \int_{[0,1]^2} uv(4 - u - v) dC(u, v),$$

whose alternative representation

$$\xi(X, Y) = -2 + 12 \int_{[0,1]^2} (2 - u - v) C(u, v) \, du \, dv$$

derives from an application of identity (4). Because of an obvious connection between English history and the concatenation of their names, the authors got in the habit of referring jokingly to these quantities as the (empirical and theoretical) “Plantagenet” coefficients. Nevertheless, they believe  $\xi_n$  and  $\xi$  would be more appropriately called symmetrized versions of Blest’s index.

As explained in Section 4 of Blest (2000),  $E(\nu_n) = E(\tilde{\nu}_n) = 0$  under the null hypothesis of independence, and hence  $E(\xi_n) = 0$  as well. Under  $H_0$ , it is also a simple matter to prove (see Appendix for details) that

$$\text{var}(\xi_n) = \frac{31n^2 + 60n + 26}{30(n+1)^2(n-1)}$$

and that  $\text{cov}(\rho_n, \xi_n) = \text{var}(\rho_n) = 1/(n-1)$ , so that

$$\text{corr}(\rho_n, \xi_n) = \sqrt{\frac{30(n+1)^2}{31n^2 + 60n + 26}} \rightarrow \sqrt{\frac{30}{31}} \approx 0.9837,$$

while  $\text{corr}(\rho_n, \nu_n) \rightarrow \sqrt{15/16} \approx 0.9682$ .

Although it is not generally true that  $\xi_n$  is an unbiased estimator of  $\xi$  under other distributional hypotheses between  $X$  and  $Y$ , it is asymptotically unbiased, as implied by the following result.

**PROPOSITION 2.** *Under the assumption of random sampling from a continuous bivariate distribution  $H$  with underlying copula  $C$ ,  $\sqrt{n}(\xi_n - \xi)$  converges weakly, as  $n \rightarrow \infty$ , to a normal random variable with zero mean and the same variance as*

$$6 \left[ UV(4 - U - V) + \int_U^1 \{2(2-u)E(V|U=u) - E(V^2|U=u)\} \, du \right. \\ \left. + \int_V^1 \{2(2-v)E(U|V=v) - E(U^2|V=v)\} \, dv \right],$$

where the pair  $(U, V)$  is distributed as  $C$ . In particular, the variance of the latter expression equals  $31/30$  when  $U$  and  $V$  are independent.

*Proof.* The asymptotic behaviour of  $\xi_n$  is obviously the same as that of

$$-4 + \frac{6}{n} \sum_{i=1}^n \frac{R_i}{n+1} \frac{S_i}{n+1} \left( 4 - \frac{R_i}{n+1} - \frac{S_i}{n+1} \right),$$

which may be written alternatively as

$$-4 + 6 \int_{[0,1]^2} uv(4 - u - v) \, dC_n(u, v)$$

in terms of the rescaled empirical copula function, as defined by Genest, Ghoudi & Rivest (1995). The conclusion is then an immediate consequence of their Proposition A.1 when  $J(u, v) = uv(4 - u - v)$ ,  $\delta = 1/4$  and  $M = p = q = 2$ , say, are chosen to satisfy conditions (i) and (ii) of their result.

Proceeding as above but with  $J(u, v) = auv + buv(4 - u - v)$  with arbitrary reals  $a$  and  $b$  actually shows that any linear combination of  $\sqrt{n}(\rho_n - \rho)$  and  $\sqrt{n}(\xi_n - \xi)$  is normally distributed with zero mean and the same variance as

$$6 \left[ UV\{2a + b(4 - U - V)\} \right. \\ \left. + \int_U^1 [2\{a + b(2 - u)\}\text{E}(V | U = u) - b\text{E}(V^2 | U = u)] du \right. \\ \left. + \int_V^1 [2\{a + b(2 - v)\}\text{E}(U | V = v) - b\text{E}(U^2 | V = v)] dv \right]. \quad (9)$$

Consequently, the joint distribution of  $\sqrt{n}(\rho_n - \rho)$  and  $\sqrt{n}(\xi_n - \xi)$  must be asymptotically normal with zero mean and a covariance matrix whose diagonal entries correspond to the choices  $(a, b) = (1, 0)$  and  $(0, 1)$ . The limiting covariance between these two quantities can also be derived from these large-sample variances and that of the linear combination corresponding to  $a = 2$  and  $b = -1$ , for instance. These observations are formally gathered in the following proposition.

**PROPOSITION 3.** *Under the assumption of random sampling from a continuous bivariate distribution  $H$  with underlying copula  $C$ ,  $\sqrt{n}(\xi_n - \xi, \rho_n - \rho)'$  converges weakly, as  $n \rightarrow \infty$ , to a normal random vector with zero mean and covariance matrix*

$$\begin{pmatrix} \sigma_\xi^2 & \kappa \\ \kappa & \sigma_\rho^2 \end{pmatrix} \quad \text{with} \quad \kappa = \frac{4\sigma_\rho^2 + \sigma_\xi^2 - \sigma_{2\rho-\xi}^2}{4},$$

where  $\sigma_\xi^2$ ,  $\sigma_\rho^2$  and  $\sigma_{2\rho-\xi}^2$  are the variances of (9) corresponding respectively to the choices  $(a, b) = (1, 0)$ ,  $(0, 1)$ , and  $(2, -1)$ , and where the pair  $(U, V)$  is distributed as  $C$ . In particular,  $\sigma_\xi^2 = 31/30$  and  $\sigma_\rho^2 = \kappa = 1$  under independence.

**Remark 1.** When condition (3) holds, making the change of variables  $(x, y) = (1 - u, 1 - v)$  in Equation (9) shows that one must then have  $\sigma_\xi^2 = \sigma_{2\rho-\xi}^2$  and hence  $\kappa = \sigma_\rho^2$  in Proposition 3. Models for which  $C = \bar{C}$  include the Farlie–Gumbel–Morgenstern, the Gaussian, the Plackett (1965), and Frank's copula (Nelsen 1986; Genest 1987); note that the latter is the only Archimedean copula that is radially symmetric (see, for instance, Nelsen 1999, p. 97).

**Remark 2.** Since the joint distribution of  $\sqrt{n}(\rho_n - \rho)$  and  $\sqrt{n}(\tau_n - \tau)$  is also known to be normal with limiting correlation equal to 1, Proposition 3 implies that

$$\lim_{n \rightarrow \infty} \text{corr}(\xi_n, \tau_n) = \lim_{n \rightarrow \infty} \text{corr}(\xi_n, \rho_n) = \sqrt{\frac{30}{31}}$$

under the null hypothesis of independence.

The following examples provide numerical illustrations of these various facts.

**Example 1 (continued).** If  $(X, Y)$  follows a Farlie–Gumbel–Morgenstern distribution with marginals  $F$  and  $G$  and parameter  $\theta \in [-1, 1]$ , then

$$\xi_\theta(X, Y) = \rho_\theta(X, Y) = \frac{3}{2}\tau_\theta(X, Y) = \frac{\theta}{3}$$

and

$$\sqrt{n}(\xi_n - \xi_\theta) \longrightarrow N\left(0, \frac{1}{450}\theta^3 - \frac{157}{630}\theta^2 + \frac{2}{225}\theta + \frac{31}{30}\right)$$

as  $n \rightarrow \infty$ . Furthermore,  $\text{cov}(\sqrt{n}\xi_n, \sqrt{n}\rho_n) \rightarrow 1 - 11\theta^2/45$  and

$$\text{corr}(\xi_n, \rho_n) \longrightarrow \sqrt{\frac{3150 - 770\theta^2}{3255 + 28\theta - 785\theta^2 + 7\theta^3}}.$$

The latter is a decreasing function of  $\theta$  taking values in the interval [0.9747, 0.9887].

**Example 2** (continued). For any pair  $(X, Y)$  distributed as  $H_{\theta,a,b}$ , one finds

$$\xi_{\theta,a,b}(X, Y) = \frac{2ab(ab + 4a + 4b + 15)\theta}{(a+2)(a+3)(b+2)(b+3)},$$

which is (happily!) symmetric in  $a$  and  $b$ . An algebraic expression for the asymptotic variance of  $\sqrt{n}(\xi_n - \xi)$  exists but is rather unwieldy.

**Example 3** (continued). When the population is bivariate normal with Pearson correlation  $r$ , it was seen earlier that  $\nu_r(X, Y) = \rho_r(X, Y)$ , and hence  $\xi_r(X, Y) = \rho_r(X, Y)$ . Note, however, that  $\xi_n$  is not necessarily equal to  $\rho_n$ . In this case, numerical integration must be used to compute the asymptotic variance of  $\sqrt{n}(\xi_n - \xi)$  or the correlation between that statistic and  $\rho_n$ .

**Example 4** (continued). Whether the pair  $(X, Y)$  is distributed as a shuffle of min (6) or (7), one has  $\xi_\theta(X, Y) = \nu_\theta(X, Y)$ , since both of these copulas are symmetric in their arguments, and the symmetrized version  $\xi$  of Blest's index satisfies all the conditions listed by Scarsini (1984), except for (e). If  $(X, Y)$  is distributed as (6), then

$$\sqrt{n}(\xi_n - \xi_\theta) \rightarrow N[0, 16\theta^5(1-\theta)(3-2\theta)^2]$$

while if  $(X, Y)$  is distributed as (7), then

$$\sqrt{n}(\xi_n - \xi_\theta) \rightarrow N[0, 64\theta^7(1-\theta)].$$

In both cases,  $\rho_\theta(X, Y) = 1 - 2\theta^3$  [because Spearman's rho satisfies condition (e)] and  $\text{corr}(\xi_n, \rho_n) \rightarrow 1$  as  $n \rightarrow \infty$ , for all values of  $0 \leq \theta \leq 1$ .

## 5. PERFORMANCE AS A TEST OF INDEPENDENCE

While Blest's index may be valuable as a measure of discrepancy between two sets of ranks, because of its asymmetric character it would seem inappropriate as a test statistic for the null hypothesis of stochastic independence between random variables  $X$  and  $Y$ . In some circumstances, acceptance or rejection of  $\mathcal{H}_0$  might conceivably depend on the rather arbitrary choice of  $\nu_n$  or  $\bar{\nu}_n$  as a test statistic.

The symmetrized statistic  $\xi_n$  escapes this criticism and thus yields a more easily defendable test procedure, whose potential is assessed below by comparing its performance with that of Spearman's rho and Kendall's tau, which are the two most common rank statistics used to this end. Asymptotic and finite-sample comparisons are presented in turn.

### 5.1. Pitman efficiency.

Using Proposition 3, it is a simple matter to compute Pitman's asymptotic relative efficiency (ARE) of tests  $T_\xi$  and  $T_\rho$  based on  $\xi_n$  and  $\rho_n$ , respectively. Given a family  $(C_\theta)$  of copulas with  $\theta = \theta_0$  corresponding to independence, standard theory (see, for instance, Lehmann 1998, p. 371 ff.) implies that

$$\text{ARE}(T_\xi, T_\rho) = \frac{30}{31} \left( \frac{\xi'_0}{\rho'_0} \right)^2,$$

where  $\xi'_0 \equiv d\xi_\theta/d\theta$  evaluated at  $\theta = \theta_0$  and  $\rho'_0$  is defined *mutatis mutandis*. The factor 30/31 comes about because as mentioned in Proposition 3, the asymptotic variances of  $\sqrt{n}\xi_n$  and  $\sqrt{n}\rho_n$  are 30/31 and 1, respectively.

It was already pointed out in Section 2 that  $\xi(X, Y) = \rho(X, Y)$  when the copula associated with the pair  $(X, Y)$  is radially symmetric. Consequently,

$$\text{ARE}(T_\xi, T_\rho) = \frac{30}{31} \approx 96.77\%$$

for such models. There is thus no reason to base a test of independence on  $\xi_n$  (whose variance is larger than that of  $\rho_n$ ) if the alternative satisfies condition (3), as is the case for the normal distribution and the Farlie–Gumbel–Morgenstern, Plackett and Frank copulas, for instance. Other examples in which  $\text{ARE}(T_\xi, T_\rho)$  can be computed explicitly are given below.

*Example 5.* Suppose that the copula of a pair  $(X, Y)$  is of the form

$$C_\theta(u, v) = (u^{-\theta} + v^{-\theta} - 1)^{-1/\theta}, \quad \theta > 0$$

with  $C_0(u, v) = \lim_{\theta \rightarrow 0} C_\theta(u, v) = uv$  for all  $0 \leq u, v \leq 1$ . This Archimedean copula model, generally attributed to Clayton (1978), is quite popular in survival analysis, where it provides a natural bivariate extension of Cox's proportional hazards model; see, for instance, Oakes (2001, Section 7.3). One can easily check that

$$\dot{C}_0(u, v) \equiv \lim_{\theta \rightarrow 0} \frac{\partial C_\theta(u, v)}{\partial \theta} = uv \log(u) \log(v),$$

so that

$$\rho'_0 = 12 \int_{[0,1]^2} \dot{C}_0(u, v) du dv = \frac{3}{4},$$

while

$$\xi'_0 = 12 \int_{[0,1]^2} (2 - u - v) \dot{C}_0(u, v) du dv = \frac{5}{6},$$

yielding  $\text{ARE}(T_\xi, T_\rho) = 1000/837 \approx 119.47\%$ . Consequently, an improvement of some 20% can be achieved in large samples by using  $\xi_n$  instead of  $\rho_n$  when testing for independence. The same would be true of  $\tau_n$ , since  $\text{ARE}(T_\tau, T_\rho) = 1$  in this case (and all subsequent ones considered here).

*Example 6.* Suppose that the pair  $(X, Y)$  is distributed as Gumbel's bivariate exponential distribution (Gumbel 1960a), whose copula (Nelsen 1999, p. 94) is

$$C_\theta(u, v) = uv \exp\{-\theta \log(u) \log(v)\}, \quad 0 \leq \theta < 1$$

with  $C_0(u, v) = uv$  corresponding to independence. In this case,  $\dot{C}_0(u, v) = -uv \log(u) \log(v)$ , so that  $\rho'_0 = -3/4$  and  $\xi'_0 = -5/6$ , and  $\text{ARE}(T_\xi, T_\rho) = 1000/837$ , as above.

*Example 7.* Suppose that the pair  $(X, Y)$  follows a bivariate logistic distribution as defined by Ali, Mikhail & Haq (1978). The corresponding copula is then (Nelsen 1999, p. 25)

$$C_\theta(u, v) = \frac{uv}{1 + \theta(1-u)(1-v)}, \quad -1 < \theta < 1$$

with  $\theta = 0$  corresponding once again to independence. In this case,  $\rho'_0 = \xi'_0 = 1/3$ , whence  $\text{ARE}(T_\xi, T_\rho) = 30/31$ , as is the case for radially symmetric copulas (which this one is not).

**Example 8.** Suppose that the distribution of  $(X, Y)$  is a member of the bivariate exponential family of Marshall & Olkin (1967), considered as a prime example of “common shock” model in reliability theory. Its associated copula, known as the generalized Cuadras–Augé copula (Nelsen 1999, §3.1.1), is

$$C_{a,b}(u, v) = \min(u^{1-a}v, uv^{1-b})$$

with  $0 \leq a, b \leq 1$ . For this family, independence occurs whenever  $\min(a, b) = 0$ . Furthermore,

$$\rho_{a,b} = \frac{3ab}{2a + 2b - ab}$$

with the convention that  $\rho_{a,b} = 0$  when  $a = b = 0$ , so that

$$\left. \frac{\partial \rho_{a,b}}{\partial b} \right|_{b=0} = \left. \frac{\partial \rho_{a,b}}{\partial a} \right|_{a=0} = \frac{3}{2}$$

when the other parameter is fixed. A closed-form expression (not reproduced here) is also available for  $\xi_{a,b}$ , and symbolic calculation yields

$$\left. \frac{\partial \xi_{a,b}}{\partial b} \right|_{b=0} = \left. \frac{\partial \xi_{a,b}}{\partial a} \right|_{a=0} = \frac{4}{3},$$

whence  $\text{ARE}(T_\xi, T_\rho) = 640/837 \approx 76.46\%$ .

As an additional example, consider the one-parameter family obtained by setting  $b = ka$  for some fixed  $0 < k < 1/a$ . Then  $\rho'_0 = 3k/(2k + 2)$  and

$$\xi'_0 = \frac{k(8 + 19k + 8k^2)}{(2 + 3k)(3 + 2k)(1 + k)},$$

so that Pitman’s ARE is a function of the constant  $k$  that is plotted in the left panel of Figure 2. The ARE is seen to reach its minimum value of  $640/837$  when  $k \rightarrow 0$  or  $\infty$ . Its maximum, namely  $392/465 \approx 84.30\%$ , occurs when  $k = 1$ , which corresponds to the standard copula of Cuadras & Augé (1981).

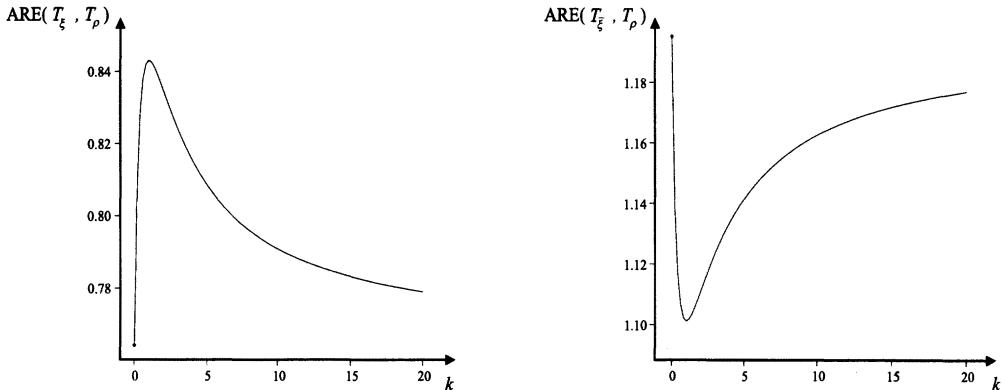


FIGURE 2: Pitman’s asymptotic relative efficiency for the generalized Cuadras–Augé copula of Example 8 when  $b = ka$ , plotted as a function of  $k$ ;  $\text{ARE}(T_\xi, T_\rho)$  is depicted in the left panel, while  $\text{ARE}(\hat{T}_\xi, T_\rho)$  is displayed in the right panel.

Examples 5 to 8 show that a test of independence based on the symmetrized version of Blest’s coefficient is sometimes, but not always, preferable to Spearman’s test in large samples. Interestingly, however, in general (since  $\rho_n$  and  $\tau_n$  are usually equivalent asymptotically) the test

based on  $\rho_n$  or  $\tau_n$  can be outperformed asymptotically by either the test involving  $\xi_n$  or a similar procedure founded on

$$\begin{aligned}\bar{\xi}_n &= -\frac{4n+5}{n-1} + \frac{6}{n^3-n} \sum_{i=1}^n \bar{R}_i \bar{S}_i \left( 4 - \frac{\bar{R}_i + \bar{S}_i}{n+1} \right) \\ &= -\frac{2n+1}{n-1} + \frac{6}{n^3-n} \sum_{i=1}^n R_i S_i \left( \frac{R_i + S_i}{n+1} \right),\end{aligned}$$

defined as in (8), but with  $\bar{R}_i = n+1-R_i$  and  $\bar{S}_i = n+1-S_i$  instead of  $R_i$  and  $S_i$ , respectively. Since using the reverse ranks amounts to working with the transformed data  $(-X_1, -Y_1), \dots, (-X_n, -Y_n)$ , this new statistic is complementary to  $\xi_n$  in that it emphasizes discrepancies observed in the greatest ranks induced by the original variables.

In view of the discussion surrounding equation (5), it is plain that  $\bar{\xi}_n$  is an asymptotically unbiased estimator of

$$\bar{\xi}(X, Y) = \xi(-X, -Y) = 2\rho(X, Y) - \xi(X, Y). \quad (10)$$

Calling on Proposition 2, one can also check readily that the limiting distribution of  $\bar{\xi}_n$  is actually the same as that which  $\xi_n$  would have if the underlying dependence function were what Nelsen (1999, p. 28) calls the survival copula, that is,  $\hat{C}(u, v) = u+v-1+C(1-u, 1-v)$ . Furthermore,  $\text{var}(\bar{\xi}_n) = \text{var}(\xi_n)$  for any sample size  $n \geq 1$  under the null hypothesis of independence.

The right panel of Figure 2 shows  $\text{ARE}(T_{\bar{\xi}}, T_{\rho})$  as a function of  $k$  for the generalized Cuadras–Augé copula of Example 8 with parameter  $b = ka$ . The curve reaches its minimum value of  $512/465 \approx 110.11\%$  at  $k = 1$ . The ARE tends to  $1000/837 \approx 119.47\%$  as  $k \rightarrow 0$  or  $\infty$ .

More generally, it follows from relation (10) that  $\bar{\xi}'_{\theta} = 2\rho'_{\theta} - \xi'_{\theta}$ , whence

$$\frac{31}{30} \text{ARE}(T_{\bar{\xi}}, T_{\rho}) = \left( \frac{\bar{\xi}'_0}{\rho'_0} \right)^2 = \left( 2 - \frac{\xi'_0}{\rho'_0} \right)^2 = \left\{ 2 - \sqrt{\frac{31}{30} \text{ARE}(T_{\xi}, T_{\rho})} \right\}^2,$$

so that the two AREs are monotone decreasing functions of each other. Accordingly,

$$\frac{31}{30} \max\{\text{ARE}(T_{\bar{\xi}}, T_{\rho}), \text{ARE}(T_{\xi}, T_{\rho})\} = \max\left\{x, (2 - \sqrt{x})^2\right\},$$

where  $x = (31/30) \times \text{ARE}(T_{\xi}, T_{\rho}) \geq 0$ . Since the right-hand side is minimized when  $x = 1$ ,

$$\max\{\text{ARE}(T_{\bar{\xi}}, T_{\rho}), \text{ARE}(T_{\xi}, T_{\rho})\} \geq \frac{30}{31} \approx 96.77\%.$$

Moreover, at least one of  $T_{\xi}$  or  $T_{\bar{\xi}}$  provides an improvement over Spearman's test unless

$$\sqrt{x} = \frac{\xi'_0}{\rho'_0} \in \left( 2 - \sqrt{\frac{31}{30}}, \sqrt{\frac{31}{30}} \right) \approx (0.9834, 1.0165).$$

From the above examples and the authors' experience with other copula models, it would appear that for "smooth" families of distributions, the largest possible asymptotic relative efficiency attainable with either  $T_{\xi}$  or  $T_{\bar{\xi}}$  is  $1000/837$ , that is, when  $\sqrt{x} = 9/10$  or  $10/9$ , as in Examples 5, 6 and 8. The exact conditions under which this occurs remain to be determined, however.

### 5.2. Power comparisons in finite samples.

To compare the performance of tests of independence based on  $\xi_n$ ,  $\bar{\xi}_n$ ,  $\rho_n$  and  $\tau_n$ , Monte Carlo simulations were carried out by Plante (2002) for various sample sizes and families of distributions spanning all possible degrees of association between stochastic independence ( $\rho = 0$ )

and complete positive dependence ( $\rho = 1$ ), whose underlying copula is the Fréchet upper bound  $M(u, v) = \min(u, v)$ .

Results reported herein are for pseudo-random samples of size  $n = 25$  and 100 for the normal (which is archetypical of radially symmetric copulas), the Clayton (Example 5), and the bivariate extreme value distributions of Gumbel (1960b) and Galambos (1975). The latter two have copulas of the form

$$C_\theta(u, v) = \exp \left[ \log(uv) A \left\{ \frac{\log(u)}{\log(uv)} \right\} \right]$$

with

$$A(t) = \{t^\theta + (1-t)^\theta\}^{1/\theta}, \quad \theta \geq 1$$

for Gumbel's model and

$$A(t) = 1 - \{t^{-\theta} + (1-t)^{-\theta}\}^{-1/\theta}, \quad \theta \geq 0$$

for Galambos's model; see, for instance, Ghoudi, Khoudraji & Rivest (1998).

Figures 3 to 6 compare the power of the (two-sided) tests based on  $\xi_n$ ,  $\bar{\xi}_n$ ,  $\rho_n$  and  $\tau_n$  under the four selected models when  $n = 25$  (left panel) and  $n = 100$  (right panel). These curves are based on 5000 replicates. In each case, the test statistic was standardized using its exact variance under the null hypothesis of independence and compared to the 97.5th centile of the asymptotic standard normal distribution. [Out of curiosity, the authors also carried out simulations for tests of independence based on  $\nu_n$  and  $\bar{\nu}_n$ ; they found that in all cases, the power curve nearly matched that of  $\xi_n$ .]

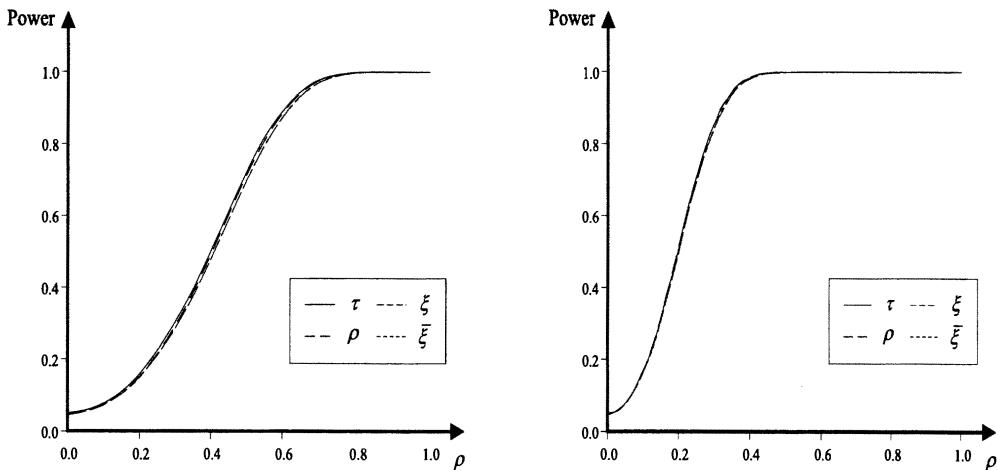


FIGURE 3: Power curve for rank tests of independence of level  $\alpha = 5\%$  based on  $\rho_n$ ,  $\tau_n$ ,  $\xi_n$  and  $\bar{\xi}_n$ , drawn as a function of Spearman's rho for random samples of size  $n = 25$  (left panel) and  $n = 100$  (right panel) from the bivariate normal distribution.

According to Figure 3, there is very little evidence for choosing between the four procedures when the underlying dependence structure is normal, even when  $n$  is small. Although strictly speaking, the test based on Kendall's tau is best, its slight advantage tends to be attenuated as the sample size increases; and as expected, power generally increases with the sampling effort.

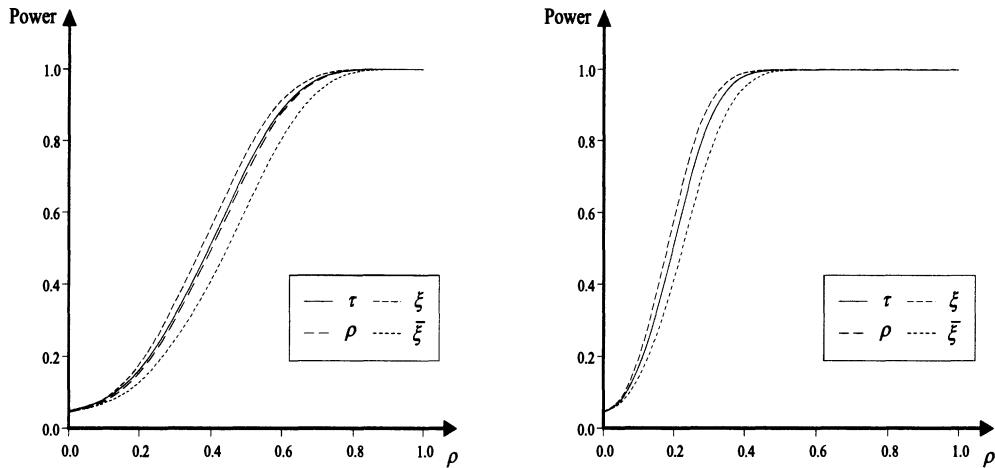


FIGURE 4: Power curve for rank tests of independence of level  $\alpha = 5\%$  based on  $\rho_n$ ,  $\tau_n$ ,  $\xi_n$  and  $\bar{\xi}_n$ , drawn as a function of Spearman's rho for random samples of size  $n = 25$  (left panel) and  $n = 100$  (right panel) from Clayton's bivariate copula.

As suggested by Example 5, the test based on  $\xi_n$  should be preferable to those based on Spearman's rho or Kendall's tau in Clayton's model. This is confirmed in Figure 4, where the good performance of  $T_\xi$  is compensated by the comparative lack of power of  $T_{\bar{\xi}}$ , as already discussed in Section 5.1.

Finally, Figures 5 and 6 provide examples of extreme value distributions in which greater power accrues from the use of the statistic  $\xi_n$  than from either Spearman's rho, Kendall's tau or the symmetrized version  $\bar{\xi}_n$  of Blest's coefficient.

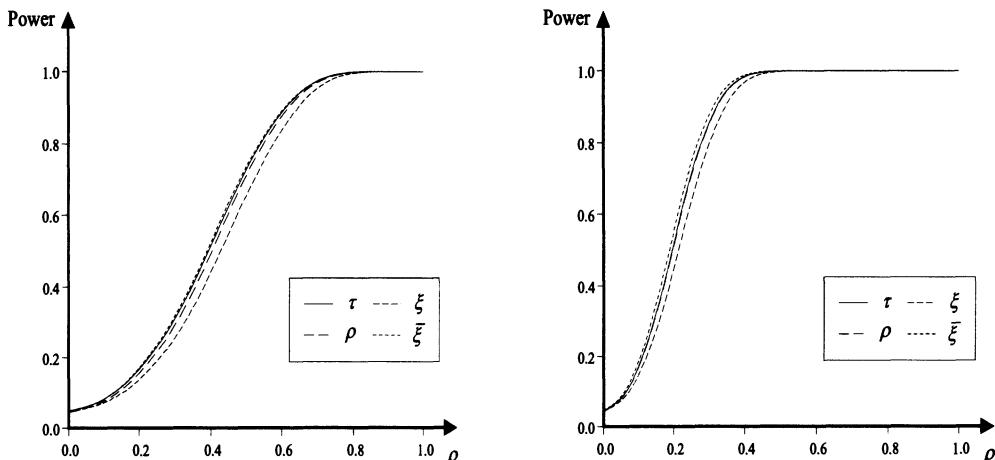


FIGURE 5: Power curve for rank tests of independence of level  $\alpha = 5\%$  based on  $\rho_n$ ,  $\tau_n$ ,  $\xi_n$  and  $\bar{\xi}_n$ , drawn as a function of Spearman's rho for random samples of size  $n = 25$  (left panel) and  $n = 100$  (right panel) from Gumbel's bivariate extreme value copula.

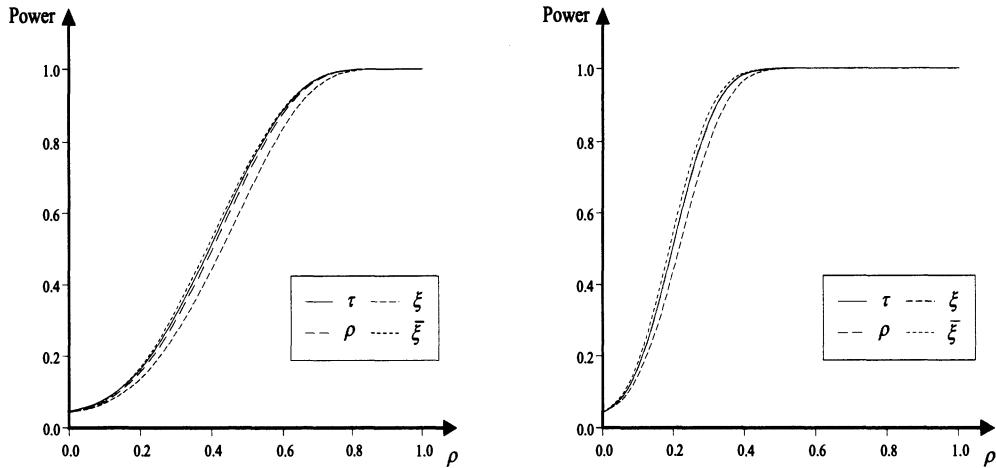


FIGURE 6: Power curve for rank tests of independence of level  $\alpha = 5\%$  based on  $\rho_n$ ,  $\tau_n$ ,  $\xi_n$  and  $\bar{\xi}_n$ , drawn as a function of Spearman's rho for random samples of size  $n = 25$  (left panel) and  $n = 100$  (right panel) from the bivariate extreme value copula of Galambos.

## 6. CONCLUSION

This paper continues the work of Blest (2000) by showing that his coefficient is asymptotically normal with parameters for which an explicit form is given in several instances. A symmetric version of his measure is also proposed which is highly correlated with Spearman's rho while retaining Blest's idea that greater emphasis should be given to discrepancies in the small ranks induced by two variables observed on the same set of individuals. The new measure, whose limiting distribution is also normal, is compared to Spearman's rho and Kendall's tau as a test statistic for independence, both through simulations in small samples and in terms of asymptotic relative efficiency. It is shown that nonnegligible improvements in power are possible, either when the test is based on the symmetrized version  $\xi_n$  of Blest's coefficient, or on a complementary statistic  $\bar{\xi}_n$  involving reverse ranks.

It would be of interest, in future work, to characterize the type of alternatives to independence for which  $\xi_n$  is preferable to  $\bar{\xi}_n$ . A more ambitious project would be to identify polynomials  $Q(u, v)$  and dependence structures for which an empirical coefficient of the form  $\int Q(u, v) dC_n(u, v)$  would be a powerful test statistic. Finally, in the spirit of Hallin & Puri (1992), Ferguson, Genest & Hallin (2000) or Genest, Quessy & Rémillard (2002), the merits of different variants of Blest's index could also be investigated as measures of serial dependence or as tests of randomness in a time series context.

## APPENDIX

Explicit formulas for

$$\text{var}(\xi_n) = \frac{36}{(n^3 - n)^2} \text{var} \left\{ \sum_{i=1}^n R_i S_i \left( 4 - \frac{R_i + S_i}{n+1} \right) \right\}$$

and

$$\text{corr}(\xi_n, \rho_n) = \text{corr} \left\{ \sum_{i=1}^n R_i S_i \left( 4 - \frac{R_i + S_i}{n+1} \right), \sum_{i=1}^n R_i S_i \right\}$$

can be found under the assumption of independence through repeated use of the following elementary result.

**LEMMA.** Let  $(X_1, Y_1), \dots, (X_n, Y_n)$  be a random sample from some continuous distribution  $H(x, y) = F(x)G(y)$ , and let  $(R_1, S_1), \dots, (R_n, S_n)$  be its associated set of ranks. If  $J, K, L$  and  $M$  are real-valued functions defined on the integers  $\{1, \dots, n\}$ , then

$$\mathbb{E} \left\{ \sum_{i=1}^n J(R_i)K(S_i) \right\} = \frac{1}{n} \left\{ \sum_{i=1}^n J(i) \right\} \left\{ \sum_{j=1}^n K(j) \right\}.$$

and

$$\mathbb{E} \left\{ \sum_{i \neq k} J(R_i)K(S_i)L(R_k)M(S_k) \right\} = \frac{1}{n(n-1)} \left\{ \sum_{i \neq k} J(i)L(k) \right\} \left\{ \sum_{j \neq \ell} K(j)M(\ell) \right\}.$$

*Proof.* The first identity can be found, for example, in the book of Hájek (1969, Th. 24B, p. 117). The second one is undoubtedly known as well, but harder to locate. A proof is included here for completeness.

Without loss of generality, one may write

$$\sum_{i \neq k} J(R_i)K(S_i)L(R_k)M(S_k) = \sum_{i \neq k} J(i)K(q_{t,i})L(k)M(q_{t,k})$$

for a specific element  $Q_t = (q_{t,1}, \dots, q_{t,n})$  in the collection  $Q = \{Q_1, \dots, Q_{n!}\}$  of permutations of the vector  $(1, \dots, n)$ . Under the hypothesis of independence, all points in  $Q$  are equally likely. Thus, if  $Q_T$  denotes a random permutation in this set, one has

$$\begin{aligned} \mathbb{E} \left\{ \sum_{i \neq k} J(R_i)K(S_i)L(R_k)M(S_k) \right\} &= \mathbb{E} \left\{ \sum_{i \neq k} J(i)K(q_{T,i})L(k)M(q_{T,k}) \right\} \\ &= \frac{1}{n!} \sum_{t=1}^{n!} \sum_{i \neq k} J(i)K(q_{t,i})L(k)M(q_{t,k}) \\ &= \frac{1}{n!} \sum_{i \neq k} J(i)L(k) \sum_{t=1}^{n!} K(q_{t,i})M(q_{t,k}). \end{aligned}$$

Now for arbitrary integers  $i, j, k, \ell \in \{1, \dots, n\}$  with  $i \neq k$  and  $j \neq \ell$ , the event  $(q_{t,i}, q_{t,k}) = (j, \ell)$  occurs exactly  $(n-2)!$  times as  $t$  ranges over  $1, \dots, n!$ . Therefore,

$$\sum_{t=1}^{n!} K(q_{t,i})M(q_{t,k}) = (n-2)! \sum_{j \neq \ell} K(j)M(\ell),$$

which yields the second identity.

As an example of application,

$$\begin{aligned} \mathbb{E} \left\{ \sum_{i=1}^n R_i S_i \left( 4 - \frac{R_i + S_i}{n+1} \right) \right\} &= 4 \mathbb{E} \left( \sum_{i=1}^n R_i S_i \right) - \frac{2}{n+1} \mathbb{E} \left( \sum_{i=1}^n R_i^2 S_i \right) \\ &= \frac{4}{n} \left( \sum_{i=1}^n i \right)^2 - \frac{2}{n^3 - n} \left( \sum_{i=1}^n i \right) \left( \sum_{i=1}^n i^2 \right) \\ &= \frac{n(n+1)(4n+5)}{6}, \end{aligned}$$

from which it follows that  $\mathbb{E}(\xi_n) = 0$  under the assumption of independence. All other computations are similar and are easily performed with symbolic-computation software such as MAPLE.

## ACKNOWLEDGEMENTS

This paper constitutes part of the second author's Master's thesis, prepared under the supervision of the first author at Université Laval. For graduate scholarships and research grants in partial support of this work, thanks are due to the Natural Sciences and Engineering Research Council of Canada and to the Fonds québécois de la recherche sur la nature et les technologies.

## REFERENCES

- M. M. Ali, N. N. Mikhail & M. S. Haq (1978). A class of bivariate distributions including the bivariate logistic. *Journal of Multivariate Analysis*, 8, 405–412.
- I. G. Bairamov, S. Kotz & M. Bekçi (2000). *New Generalized Farlie–Gumbel–Morgenstern Distributions and Concomitants of Order Statistics*. Technical Report, George Washington University, Washington, DC.
- D. C. Blest (2000). Rank correlation: an alternative measure. *Australian and New Zealand Journal of Statistics*, 42, 101–111.
- D. G. Clayton (1978). A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. *Biometrika*, 65, 141–151.
- C. M. Cuadras & J. Augé (1981). A continuous general multivariate distribution and its properties. *Communications in Statistics: Theory and Methods*, 10, 339–353.
- D. Drouet-Mari & S. Kotz (2001). *Correlation and Dependence*. Imperial College Press, London.
- P. Embrechts, A. J. McNeil & D. Straumann (2002). Correlation and dependence in risk management: properties and pitfalls. In *Risk Management: Value at Risk and Beyond* (M. A. H. Dempster, ed.), Cambridge University Press, pp. 176–223.
- T. S. Ferguson, C. Genest & M. Hallin (2000). Kendall's tau for serial dependence. *The Canadian Journal of Statistics*, 28, 587–604.
- J. Galambos (1975). Order statistics of samples from multivariate distributions. *Journal of the American Statistical Association*, 70, 674–680.
- C. Genest (1987). Frank's family of bivariate distributions. *Biometrika*, 74, 549–555.
- C. Genest, K. Ghoudi & L.-P. Rivest (1995). A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. *Biometrika*, 82, 543–552.
- C. Genest, K. Ghoudi & L.-P. Rivest (1998). Comment on the paper by E. W. Frees and E. A. Valdez entitled “Understanding relationships using copulas.” *North American Actuarial Journal*, 2, 143–149.
- C. Genest, J.-F. Quesy & B. Rémillard (2002). Tests of serial independence based on Kendall's process. *The Canadian Journal of Statistics*, 30, 441–461.
- K. Ghoudi, A. Khoudraji & L.-P. Rivest (1998). Propriétés statistiques des copules de valeurs extrêmes bidimensionnelles. *The Canadian Journal of Statistics*, 26, 187–197.
- E. J. Gumbel (1960a). Bivariate exponential distributions. *Journal of the American Statistical Association*, 55, 698–707.
- E. J. Gumbel (1960b). Distributions des valeurs extrêmes en plusieurs dimensions. *Bulletin de l'Institut de statistique de l'Université de Paris*, 9, 171–173.
- J. Hájek (1969). *A Course in Nonparametric Statistics*. Holden-Day, San Francisco.
- M. Hallin & M. L. Puri (1992). Rank tests for time series analysis: a survey. In *New Directions in Time Series Analysis* (D. R. Brillinger, E. Parzen & M. Rosenblatt, eds.), Springer, New York, pp. 111–153.
- H. Joe (1997). *Multivariate Models and Dependence Concepts*. Chapman & Hall, London.
- E. L. Lehmann (1998). *Nonparametrics: Statistical Methods Based on Ranks*. Revised First Edition. Prentice Hall, Upper Saddle River, New Jersey.
- A. W. Marshall & I. Olkin (1967). A generalized bivariate exponential distribution. *Journal of Applied Probability*, 4, 291–302.
- P. Mikusínski, H. Sherwood & M. D. Taylor (1992). Shuffles of min. *Stochastica*, 13, 61–74.
- R. B. Nelsen (1986). Properties of a one-parameter family of bivariate distributions with specified marginals. *Communications in Statistics: Theory and Methods*, 15, 3277–3285.

- R. B. Nelsen (1999). *An Introduction to Copulas*. Lecture Notes in Statistics No. 139. Springer-Verlag, New York.
- D. Oakes (2001). *Biometrika* centenary: Survival analysis. *Biometrika*, 88, 99–142.
- R. L. Plackett (1965). A class of bivariate distributions. *Journal of the American Statistical Association*, 60, 516–522.
- J.-F. Plante (2002). *À propos de la mesure de corrélation des rangs de Blest*. Mémoire de maîtrise no 20160, Université Laval, Québec.
- J. J. Quesada-Molina (1992). A generalization of an identity of Hoeffding and some applications. *Journal of the Italian Statistical Society*, 1, 405–411.
- F. H. Ruymgaart, G. R. Shorack & W. R. van Zwet (1972). Asymptotic normality of nonparametric tests for independence. *The Annals of Mathematical Statistics*, 43, 1122–1135.
- M. Scarsini (1984). On measures on concordance. *Stochastica*, 8, 201–218.
- B. Schweizer & E. F. Wolff (1981). On nonparametric measures of dependence for random variables. *The Annals of Statistics*, 9, 879–885.

---

Received 13 December 2001

Accepted 2 November 2002

Christian GENEST: genest@mat.ulaval.ca

Jean-François PLANTE: jplanete@mat.ulaval.ca

Département de mathématiques et de statistique

Université Laval, Québec (Québec), Canada G1K 7P4