INSTITUT DE STATISTIQUE

UNIVERSITÉ CATHOLIQUE DE LOUVAIN

DISCUSSION PAPER

0145

SOME DEPENDENCE CRITERIA FOR BIVARIATE ORDINAL VARIABLES WITH GIVEN MARGINAL DISTRIBUTIONS

F. VANDENHENDE and P. LAMBERT

http://www.stat.ucl.ac.be

Some dependence criteria for bivariate ordinal variables with given marginal distributions

François Vandenhende[1,2,*] and Philippe Lambert[2]

[1] Statistics and Information Sciences, Eli Lilly & Company, B-1348 Mont-Saint-Guibert, Belgium
[2] Institute of Statistics, Université Catholique de Louvain, B-1348 Louvain-la-Neuve, Belgium
* To whom correspondence should be addressed
email to Francois@lilly.com

November 27, 2001

Summary

This paper presents a family of dependence measures for bivariate ordinal data. The joint distribution is expressed as a function of the margins. That function characterizes their dependence in between complete positive and complete negative association. When the margins are given or known, all measures are presented as functions of the expectation of the coupling function. They are all rank-based and verify selected properties for dependence (Renyi, 1959) and concordance (Scarsini, 1984) measures. Our criteria extend Schweizer's σ (Schweizer & Wolff, 1981), Spearman's rho and Kendall's tau to the ordinal (non-continuous) case. Unlike these standard measures even after adjustment for ties, perfect dependence is always associated to the same values of ± 1 whatever the distribution of the margins. Estimates of these dependence measures are proposed together with their asymptotic distributions. A goodness of fit procedure is also given to evaluate the quality of parametric models for the dependence. An illustration is proposed using data from a migraine trial.

Some key words: Bivariate ordinal data; Dependence; Concordance; Goodness of fit test.

1 Introduction

Rank-based dependence criteria (e.g. Kendall's tau) are commonly used to measure the dependence in non-continuous ordinal data. Several corrections, as Kendall's tau-b (Goodman & Kruskal, 1954), have been introduced to adjust the statistic in the presence of ties. One limitation of rank-based statistics (even when corrected for ties) in the discrete case is that they do not necessarily have the same limiting values of ± 1 when perfect dependence (positive or negative) is found in the data.

In this paper, we construct multivariate distributions for discrete ordinal data and define rank-based dependence measures with ranges ($[0, 1]$ or $[-1, 1]$) that are independent of the considered marginal distributions. The joint distribution is characterized in Section 2 using a function that couples the marginal distributions and induces dependence between them (Sklar, 1959). Dependence criteria are defined in Section 3 by comparing the joint empirical distribution with three particular distributions that induce either independence, complete positive or complete negative dependence between the margins. Properties of these criteria are also examined. The usefulness of our criteria is illustrated and compared with Goodman's γ (Goodman & Kruskal, 1954), Spearman's rho and Kendall's tau-b in Section 4. Computational issues related to the estimation of these criteria and of their asymptotic distribution are discussed in Section 5. Section 6 proposes a goodness of fit test for the (parametric) model of the dependence. Finally, these estimation and test procedures are illustrated in Section 7 with the modeling of the dependence between baseline pain and nausea severity scores recorded in a trial on acute migraine.

2 Joint Distribution of Ordinal Variables

Let X and Y be two discrete ordinal random variables distributed on two (finite) sets of categories S_X and S_Y respectively. Without loss of generality, we assume that these are $S_X = \{1, ..., K\}$ and $S_Y = \{1, ..., L\}$, where K and L denote the number of possible values for X and Y respectively. We denote the marginal distributions of X and Y by $F_X(\cdot)$ and $F_X(\cdot)$ respectively where

$$F_X(x) = \Pr(X \leq x) \text{ and } F_Y(y) = \Pr(Y \leq y), \tag{2·1}$$

for $x, y \in \Re$. Note that $F_X(0) = F_Y(0) = 0$ given the chosen sets S_X and S_Y. The joint distribution $F_{XY}(\cdot)$ of X and Y can be expressed as a function of its marginal distributions

$$F_{XY}(k,l) = C\{F_X(k), F_Y(l)\}. \tag{2·2}$$

The coupling function C is uniquely defined on the range of the marginal distributions: Ran $F_X \times$ Ran F_Y (Sklar, 1959). When restricted to this range, C is called a *subcopula*. It has the following properties:

1. Domain: Dom C = Ran $F_X \times$ Ran F_Y.
2. Bounds: For every (u, v) in Dom C,
$$C(u, 0) = C(0, v) = 0,$$
$$C(u, 1) = u \text{ and } C(1, v) = v.$$
3. Supermodularity: For every u_1, u_2 in Ran F_X and every v_1, v_2 in Ran F_Y such that $u_1 \leq u_2$ and $v_1 \leq v_2$,
$$C(u_2, v_2) - C(u_2, v_1) - C(u_1, v_2) + C(u_1, v_1) \geq 0.$$

The distribution F_{XY} is bounded on $S_X \times S_Y$ by two other distributions introduced by Fréchet (1951). Indeed, for every (k, l) in $S_X \times S_Y$,

$$\max\{F_X(k) + F_Y(l) - 1, 0\} \leq F_{XY}(k,l) \leq \min\{F_X(k), F_Y(l)\}. \tag{2·3}$$

Note that the two Fréchet bounds give a valid bivariate distribution for X and Y. The lower bound for F_{XY} induces complete negative dependence between X and Y. It has a subcopula representation of the form

$$M(u, v) = \max(u + v - 1, 0).$$

The upper bound induces complete positive dependence and has a subcopula representation of the form

$$W(u, v) = \min(u, v).$$

A whole range of dependence is possible between these two bounds. Let F^1_{XY} and F^2_{XY} be two potential joint distributions for X and Y. F^1_{XY} is said to induce more positive dependence than F^2_{XY} (denoted by $F^2_{XY} \prec F^1_{XY}$) when

$$F^2_{XY}(k,l) \leq F^1_{XY}(k,l), \forall (k,l) \in S_X \times S_Y.$$

When the marginal distributions F_X and F_Y are given, the dependence structure is completely determined by the choice of the subcopula C in (2·2). A special choice for C is

$$\Pi(u, v) = uv.$$

It generates independence between X and Y, as $F_{XY}(k, l) = F_X(k)F_Y(l)$ is the product of the margins.

X and Y are positive (negative) quadrant dependent or PQD (NQD) when

$$F_{XY}(k, l) \geq (\leq) F_X(k)F_Y(l), \tag{2·4}$$

for every $(k, l) \in S_X \times S_Y$.

PQD induces the ordering $\Pi \prec C$ for the subcopula C in (2·2). With two joint distributions F^1_{XY} and F^2_{XY} for X and Y, we define the expectation of F^1_{XY} under F^2_{XY}, as

$$E_{F^2_{XY}}(F^1_{XY}) = \sum_{k \in S_X} \sum_{l \in S_Y} F^1_{XY}(k, l) dF^2_{XY}(k, l), \tag{2·5}$$

where

$$dF^2_{XY}(k, l) = F^2_{XY}(k, l) + F^2_{XY}(k-1, l-1) - F^2_{XY}(k-1, l) - F^2_{XY}(k, l-1)$$

is the joint probability mass function at (k, l).

When F^1_{XY} and F^2_{XY} have the same known marginal distributions F_X and F_Y and differ potentially by the form of their associated subcopula, denoted by C_1 and C_2 respectively, (2·5) can be re-expressed as :

$$E_{C_2}(C_1) = \sum_{k \in S_X} \sum_{l \in S_Y} C_1\{F_X(k), F_Y(l)\} dC_2\{F_X(k), F_Y(l)\}, \tag{2·6}$$

with

$$\begin{aligned} dC_2\{F_X(k), F_Y(l)\} &= C_2\{F_X(k), F_Y(l)\} + C_2\{F_X(k-1), F_Y(l-1)\} \\ &\quad - C_2\{F_X(k-1), F_Y(l)\} - C_2\{F_X(k), F_Y(l-1)\}. \end{aligned} \tag{2·7}$$

In fact, (2·6) is the expectation of the subcopula C_1 under C_2, both defined on Ran $F_X \times$ Ran F_Y.

One can show that for any subcopula C, (2·6) induces the ordering

$$E_M(M) \leq E_C(M) \leq E_C(C) \leq E_C(W) \leq E_W(W). \tag{2·8}$$

Inequalities $E_C(M) \leq E_C(C) \leq E_C(W)$ are a direct consequence of (2·3). The other two inequalities $E_M(M) \leq E_C(M)$ and $E_C(W) \leq E_W(W)$ were given by Tchen (1980) using the supermodularity property of subcopulas.

3 Dependence criteria

We use the ordering in (2·8) to build several dependence measures. In all criteria below, the numerator measures the difference in expectation between C and Π, the subcopula characterizing independence. The denominator is used as a scaling factor as it measures the maximum for the criterion numerator under complete dependence (i.e., when $C = M$ or $C = W$). The deviation towards W is considered in the denominator when the expectation of C is greater than or equal to that of Π. Otherwise, the deviation towards M is considered.

The first criterion is

$$\sigma_C^a(X,Y) = \begin{cases} \frac{E_\Pi|C-\Pi|^a}{E_\Pi(W-\Pi)^a} & \text{if } E_\Pi(C) \geq E_\Pi(\Pi) \\ \frac{E_\Pi|C-\Pi|^a}{E_\Pi(\Pi-M)^a} & \text{otherwise,} \end{cases} \tag{3·1}$$

with $a > 0$. It is non-negative and measures the *unsigned* dependence between X and Y. A signed version of (3·1) with $a = 1$ is

$$\rho_C(X,Y) = \begin{cases} \frac{E_\Pi(C-\Pi)}{E_\Pi(W-\Pi)} & \text{if } E_\Pi(C) \geq E_\Pi(\Pi) \\ \frac{E_\Pi(C-\Pi)}{E_\Pi(\Pi-M)} & \text{otherwise.} \end{cases} \tag{3·2}$$

In (3·1) and (3·2), all expectations are taken with respect to the subcopula of independence Π. A third measure is given in which any expectation of a subcopula is taken with respect to itself. It is

$$\tau_C(X,Y) = \begin{cases} \frac{E_C(C)-E_\Pi(\Pi)}{E_W(W)-E_\Pi(\Pi)} & \text{if } E_C(C) \geq E_\Pi(\Pi) \\ \frac{E_C(C)-E_\Pi(\Pi)}{E_\Pi(\Pi)-E_M(M)} & \text{otherwise.} \end{cases} \tag{3·3}$$

A positive (negative) value for τ_C and ρ_C indicates positive (negative) dependence between X and Y. τ_C and ρ_C measure the *concordance* between the random variables.

As detailed in the two theorems below, σ_C^a, τ_C and ρ_C are only sensitive to the ordering of the categories for X and Y. They are not affected by any (increasing) transformation of the category scores that preserve their order. They are rank-based measures that only depend on the sub-copula C when the margins F_X and F_Y are given or known.

3·1 *Dependence and Concordance Properties*

Criteria in (3·1) (with $a = 1$), (3·2) and (3·3) are well known in the continuous case. For continuous responses X and Y, the expectation in (2·6) is computed as a double integral over the whole $[0,1]$ range. Then $E_\Pi(W-\Pi) = E_\Pi(\Pi - M) = 1/12$ and $E_W(W) - E_\Pi(\Pi) = E_\Pi(\Pi) - E_M(M) = 1/4$. The measure in (3·1) with a=1 is then identical to Schweizer & Wolff (1981)'s σ, ρ in (3·2) is Spearman's rho and τ in (3·3) is Kendall's tau. The above criteria also share most properties of these well-known dependence measures.

Desired properties (axioms) for measures of dependence have been enumerated by Renyi (1959) and later revisited in Schweizer & Wolff (1981) and in Schweizer (1991) when focusing on continuous responses. Scarsini (1984) has presented a list of axioms for measures of concordance. Theorem 1 shows that σ_C^a has (selected) properties of a dependence measure between two discrete ordinal random variables X and Y. In Theorem 2, we present concordance properties for τ_C and ρ_C in the discrete case.

THEOREM 1. *Let X and Y be discrete ordinal random variables with marginal distributions F_X and F_Y given in (2·1) and joint distribution F_{XY} defined in (2·2) using subcopula C. The quantity $\sigma_C^a(X,Y)$ defined in (3·1) verifies the following properties :*

(A) $\sigma_C^a(X,Y)$ is defined when F_X and F_Y are not degenerated (i.e., Ran F_X and Ran $F_Y \neq \{0,1\}$).

(B) $\sigma_C^a(X,Y) = \sigma_C^a(Y,X)$.

(C) $0 \leq \sigma_C^a(X,Y) \leq 1$.

(D) $\sigma_C^a(X,Y) = 0$ if and only if X and Y are independent (ie, iff $C = \Pi$).

(E) $\sigma_C^a(X,Y) = 1$ if $C = W$ on Ran $F_X \times$ Ran F_Y or if $C = M$ on Ran $F_X \times$ Ran F_Y.

(F) If f and g are strictly increasing functions a.s. on V_X and V_Y respectively, then $\sigma_C^a[f(X), g(Y)] = \sigma_C^a(X,Y)$.

(G) if (X,Y) and (X_n,Y_n), $n = 1,2,...$ are pairs of r.v. with joint d.f. F_{XY} and $F_{X_nY_n}$ respectively, and if the sequence $\{F_{X_nY_n}\}$ converges weakly to F_{XY}, then

$$\lim_{n\to\infty} \sigma_C^a(X_n,Y_n) = \sigma_C^a(X,Y).$$

We propose a proof in the discrete case. A detailed proof of Theorem 1 can be found in Schweizer & Wolff (1981) in the continuous case for $a = 1$.

Proof of Theorem 1. (A) From (3·1), the denominator of σ_C^a is not zero when $M(u,v) < \Pi(u,v) < W(u,v)$ for at least one pair (u,v) in Ran $F_X \times$ Ran F_Y. As this is the case for every $(u,v) \in (0,1)^2$, σ_C^a is well defined in non-degenerated cases, i.e. when Ran F_X and Ran F_Y are not $\{0,1\}$. (B) is obvious. (C) follows from (2·8). (D) is due to the absolute value function in the numerator of (3·1). (E) is obtained by direct computation. Note that the condition expressed in (E) is not sufficient. Indeed, σ_C^a could be 1 for patterns of dependence that do not exhibit uniform monotonicity between the random variables. One example is when the bivariate distribution equals the lower Fréchet bound in some cells and the upper Fréchet bound in all the others. An additional condition for having $\sigma_C^a = 1$ in such cases is that $E_\Pi(W - \Pi) = E_\Pi(\Pi - M)$. (F) follows from the fact that strictly increasing functions of discrete variables preserve the definition of their distribution function: $F_{f(X)}[f(k)] = F_X(k)$, $F_{g(Y)}[g(l)] = F_Y(l)$ and $F_{f(X)g(Y)}[f(k), g(l)] = F_{XY}(k,l)$. As such, the sub-copula C' associated to $F_{f(X)g(Y)}[f(k), g(l)]$ is equal to the sub-copula C associated to $F_{XY}(k,l)$ and the result follows. But this is not the case with *decreasing* functions anymore. As a consequence, σ_C^a will not be preserved after a change in the ordering of the random variable categories. (G) comes from the fact that there is also weak convergence of the marginal distributions F_{X_n}, F_{Y_n} to F_X, F_Y. $\square$

THEOREM 2. *Let X and Y be discrete ordinal random variables with marginal distributions F_X and F_Y given in (2·1) and joint distribution F_{XY} defined in (2·2) using subcopula C. The quantities $\rho_C(X,Y)$ defined in (3·2) and $\tau_C(X,Y)$ defined in (3·3) verify the following properties :*

- *(A') $\rho_C(X,Y)$ and $\tau_C(X,Y)$ are defined when F_X and F_Y are not degenerated (i.e., Ran F_X and Ran $F_Y \neq \{0,1\}$).*
- *(B') $\rho_C(X,Y) = \rho_C(Y,X)$ and $\tau_C(X,Y) = \tau_C(Y,X)$.*
- *(C') $-1 \leq \rho_C(X,Y) \leq 1$ and $-1 \leq \tau_C(X,Y) \leq 1$.*
- *(D') $\rho_C(X,Y) = \tau_C(X,Y) = 0$ if X and Y are independent (ie, if $C = \Pi$).*
- *(E') $\rho_C(X,Y) = \tau_C(X,Y) = 1$ iff $C = W$ on Ran $F_X \times$ Ran F_Y.*
- *(F') $\rho_C(X,Y) = \tau_C(X,Y) = -1$ iff $C = M$ on Ran $F_X \times$ Ran F_Y.*
- *(G') If f and g are strictly increasing functions a.s. on V_X and V_Y respectively, then $\rho_C[f(X), g(Y)] = \rho_C(X,Y)$ and $\tau_C[f(X), g(Y)] = \tau_C(X,Y)$.*
- *(H') if (X,Y) and (X_n, Y_n), $n = 1, 2, ...$ are pairs of r.v. with joint d.f. F_{XY} and $F_{X_nY_n}$ respectively, and if the sequence $\{F_{X_nY_n}\}$ converges weakly to F_{XY}, then*

$$\lim_{n\to\infty} \rho_C(X_n, Y_n) = \rho_C(X,Y)$$

and

$$\lim_{n\to\infty} \tau_C(X_n, Y_n) = \tau_C(X, Y).$$

Proof of Theorem 2 is similar to that of Theorem 1. Notable differences between the two theorems include the fact that the sufficient condition in (D) for independence is not expressed in (D') anymore. Conversely, (E') and (F') exhibit sufficient conditions for complete dependence while this is not the case in (E).

Monotone quadrant dependence (2·4) is when $\sigma_C^1 = |\rho_C|$. In such cases, a positive (negative) sign for ρ_C indicates PQD (NQD).

4 Comparison with standard measures

σ_C^1, ρ_C and τ_C values are reported in Table 1 for various bivariate ordinal cross-classifications taken from Agresti (1990, p. 23). Results are compared to estimates of standard association measures: Kendall's tau-b, Spearman's rho and Goodman's γ. Table 1a has positive probabilities on the main diagonal only whereas Table 1b has some off-diagonal cells containing positive probabilities too. These two tables (1a and 1b) exhibit complete positive dependence between rows and columns. This is because the joint distribution of each cell (displayed in parentheses) is equal to the minimum of the marginal distributions of the related row and column ($C = W$). Table 1c shows a U-shaped pattern of cell probabilities. The bivariate distribution is non-monotonic-dependent, with complementary concordant and discordant proportions. Table 1d has only 4 (2 by 2) entries that exhibit complete negative dependence (i.e., $0 = \max\{0.03 + 0.30 - 1, 0\}$).

The performance of the three new measures is comparable to that of Goodman's γ under complete dependence (see Tables 1a and 1b for complete positive dependence and Table 1d for complete negative dependence). With complete positive (negative) dependence, τ_C and ρ_C reach their maximum (minimum) value of 1 (-1) whatever the margins. In contrast, Kendall's tau-b and Spearman's rho do not reach these limits in Tables 1b and 1d. Vandenhende & Lambert (2000) observed similarly that Kendall's tau-b does not necessarily reach the ± 1 limits under complete dependence ($C = M$ or $C = W$) and that the achievable limits are related to the marginal distributions. This is not the case anymore with our new criteria. Table 1c shows a non-monotonic-dependent distribution with complementary concordant and discordant proportions. Although standard measures do not detect the dependence, it is well characterized by the σ_C^1 estimate (0.40). Lack of

Table 1: New and standard (Goodman's γ, Kendall's tau-b and Spearman's rho) dependence measures for several cross-classification.

	Cell Probabilities (Cumulative Distributions)			New Criteria	Standard Criteria
a.	$\frac{1}{3}$ $(\frac{1}{3})$	0 $(\frac{1}{3})$	0 $(\frac{1}{3})$	$\sigma_C^1 = 1$	$\gamma = 1$
	0 $(\frac{1}{3})$	$\frac{1}{3}$ $(\frac{2}{3})$	0 $(\frac{2}{3})$	$\tau_C = 1$	tau-b=1
	0 $(\frac{1}{3})$	0 $(\frac{2}{3})$	$\frac{1}{3}$ (1)	$\rho_C = 1$	rho=1
b.	0.2 (0.2)	0 (0.2)	0 (0.2)	$\sigma_C^1 = 1$	$\gamma = 1$
	0.2 (0.4)	0.2 (0.6)	0 (0.6)	$\tau_C = 1$	tau-b=0.75
	0 (0.4)	0.2 (0.8)	0.2 (1)	$\rho_C = 1$	rho=0.81
c.	0.2 (0.2)	0 (0.2)	0.2 (0.4)	$\sigma_C^1 = 0.40$	$\gamma = 0$
	0.2 (0.4)	0 (0.4)	0.2 (0.8)	$\tau_C = 0.10$	tau-b=0
	0 (0.4)	0.2 (0.6)	0 (1)	$\rho_C = 0.13$	rho=0
d.	0 (0)	0.30 (0.30)		$\sigma_C^1 = 1$	$\gamma = -1$
	0.03 (0.03)	0.67 (1)		$\tau_C = -1$	tau-b=-0.12
				$\rho_C = -1$	rho=-0.12

monotone quadrant dependence is suggested in Table 1c as $\sigma_C^1 - |\rho_C| = 0.27$.

5 Computational Issues

This section presents a sample version and large sample properties of the measures defined in Section 3. A bootstrap approximation of their distribution is also provided.

5·1 *Large sample distribution*

Let us take a sample of n independent realizations of the random vector (X, Y). A representation of these data in a two-way contingency table produces a series of cell counts n_{kl} for every $k \in S_X$ and $l \in S_Y$, so that $\sum_{k,l} n_{kl} = n$. The probability π_{kl} to be into cell (k, l) can be estimated from the data by the sample proportion $p_{kl} = n_{kl}/n$. Using vector notations, the stacked $K \times L$ probabilities and proportions are denoted by π and $\mathbf{p}_n$ respectively, where the subscript n indicates the sample size. From the multivariate Central Limit Theorem, we have

$$\sqrt{n}(\mathbf{p_n} - \pi) \xrightarrow{d} N[\mathbf{0}, V_n], \tag{5·1}$$

where

$$V_n = \mathrm{Diag}(\pi) - \pi\pi'.$$

The empirical estimator of the joint distribution $F_{XY}(k,l)$ is

$$\widehat{F}_{XY}(k,l) = \sum_{k'=1}^{k}\sum_{l'=1}^{l} p_{k'l'}. \tag{5·2}$$

Marginal distributions $F_X(k)$ and $F_Y(l)$ are similarly estimated by $\widehat{F}_X(k) = \widehat{F}_{XY}(k,L)$ and $\widehat{F}_Y(l) = \widehat{F}_{XY}(K,l)$ respectively. The subcopula related to $\widehat{F}_{XY}(k,l)$ is estimated as

$$\widehat{C}\{\widehat{F}_X(k), \widehat{F}_Y(l)\} = \widehat{F}_{XY}(k,l).$$

The empirical estimators $\widehat{\sigma}^a_C$, $\widehat{\rho}_C$ and $\widehat{\tau}_C$ are then obtained from (3·1), (3·2) and (3·3) respectively using $\widehat{F}_X$, $\widehat{F}_Y$ and $\widehat{C}$ in (2·6).

From (5·2), $\widehat{\rho}_C$ and $\widehat{\tau}_C$ are functions (denoted below by g) of the cell proportions p_{kl} only. Both functions are differentiable with respect to $\mathbf{p_n}$. Let

$$\phi_{kl} = \frac{\partial g}{\partial p_{kl}}$$

denote the differential evaluated at $\mathbf{p}_n = \pi$ and let $\phi = (\phi_{11}, ..., \phi_{KL})'$. The asymptotic variance of g is approximately (Agresti, 1990, p. 424)

$$\phi' V_n \phi = \sum_{k,l} \pi_{kl}\phi_{kl}^2 - (\sum_{k,l} \pi_{kl}\phi_{kl})^2. \tag{5·3}$$

The delta method used in (5·3) cannot be readily applied to the computation of the asymptotic variance of $\widehat{\sigma}^a_C$ for any $a > 0$. This is because the absolute value functions in (3·1) are not differentiable at zero. However, the computation is possible when a is even (e.g. for $\widehat{\sigma}^2_C$) as $E_\Pi|C - \Pi|^a = E_\Pi(C - \Pi)^a$ is differentiable in such cases.

The three criteria defined in Section 3 have a bounded range ($[0,1]$ for σ^a_C and $[-1,1]$ for ρ_C and τ_C). The *exact* distributions of their estimators are not necessarily symmetric; the skewness of these distributions could even increase when the underlying population criteria are near the boundaries. This should be contrasted to the symmetry and the unboundedness of domain of the normal *asymptotic* distribution of the estimators (cfr Eq. (5·1)). The quality of the normal approximation increases as n, the number of observations, tends to infinity. With small sample sizes, the quality of the approximation can be poor over the whole statistic's domain and particularly, near the borders. Therefore, it is recommended to assess the quality of the normal approximation using simulations.

5·2 *Bootstrap method*

The distribution of $\widehat{\sigma}^a_C$, $\widehat{\rho}_C$ and $\widehat{\tau}_C$ can be approximated using basic bootstrap methods (Davison & Hinkley, 1997, pp. 16-20). Let $\{n^*_{11,r}, ..., n^*_{KL,r}\}$ be a simulated sample vector (for any $r = 1, ..., R$) of cell counts in a K by L cross-classification table counting a total of n records. Elements are simulated from a multinomial distribution with cell probabilities equal to p_{kl}, the observed cell proportions. A total of R such independent samples is obtained by repeating the sampling procedure R times. For each sample $r = 1, ..., R$, the simulated cell proportions are $p^*_{kl,r} = n^*_{kl,r}/n$ and the cumulative proportions are $F^*_{XY,r}(k,l) = \sum_{k'=1}^{k} \sum_{l'=1}^{l} p^*_{k'l',r}$. We denote by $\sigma^{a*}_{C,r}$, $\rho^*_{C,r}$ and $\tau^*_{C,r}$ the statistics defined in (3·1), (3·2) and (3·3) respectively that were calculated from the r^{th} simulated table using $F^*_{XY,r}$. The distribution of (say) $\widehat{\rho}_C$ can be approximated from the samples $\rho^*_{C,1}, ..., \rho^*_{C,R}$ using basic bootstrap methods when R is large. An estimator of $\widehat{\rho}_C$ is the sample mean $\bar{\rho}^*_C$, an estimator of its variance is the sample variance

$$\frac{1}{R-1} \sum_{r=1}^{R} (\rho^*_{C,r} - \bar{\rho}^*_C)^2,$$

and the q^{th} quantile of its distribution is estimated as $\rho^*_{C,([Rq])}$, where the function $[x]$ gives the integer part of x and $\rho^*_{C,(1)} \leq ... \leq \rho^*_{C,(R)}$ are the ordered estimates from the simulation. Estimators for the distribution of $\widehat{\sigma}^a_C$ and $\widehat{\tau}_C$ are similarly derived.

6 Goodness of Fit Test for Dependence Models

Criteria of Section 3 were defined based on the ordering given in (2·8) to measure the dependence between two ordinal random variables. The same ordering can be used to build criteria having a different purpose. In this section, one criterion is given to measure the adjustment of dependence models to bivariate ordinal data.

Let C be the subcopula associated to the joint distribution of X and Y and let C_θ be any (parametric) subcopula family defined on Ran $F_X\times$ Ran F_Y used in the modelling process. C_θ is (possibly) indexed by a vector of parameters θ controlling the strength of the dependence between X and Y. When θ is given or estimated (e.g. using maximum likelihood) by $\widehat{\theta}$, the adjustment of $C_{\widehat{\theta}}$ to C can be monitored using

$$\Delta = \frac{E_\Pi[C_{\widehat{\theta}} - C]^2}{E_\Pi[W - M]^2}. \tag{6·1}$$

From (2·8), one can see that $0 \leq \Delta \leq 1$ and that Δ is one when complete positive dependence ($C = W$) is modeled using the subcopula of complete negative dependence ($C_\theta = M$) or vice versa.

The null hypothesis $H_0 : C = C_{\widehat{\theta}}$ of a goodness of fit test assumes a perfect agreement between C and $C_{\widehat{\theta}}$ on the whole range of the marginal distributions. The considered alternative hypothesis $H_a : C \neq C_{\widehat{\theta}}$ states that there is at least one pair (u, v) in Ran $F_X\times$ Ran F_Y for which $C(u, v) \neq C_{\widehat{\theta}}(u, v)$.

In non degenerated cases (cfr. Theorem 1A), it is easy to see that the hypotheses can be equivalently expressed as

$$H_0 : \Delta = 0 \quad \text{and} \quad H_a : \Delta > 0. \tag{6·2}$$

An empirical estimator for Δ is

$$\widehat{\Delta} = \frac{\sum_{k,l} [C_{\widehat{\theta}}\{\widehat{F}_X(k), \widehat{F}_Y(l)\} - \widehat{F}_{XY}(k,l)]^2 d\Pi\{\widehat{F}_X(k), \widehat{F}_Y(l)\}}{\sum_{k,l} [W\{\widehat{F}_X(k), \widehat{F}_Y(l)\} - M\{\widehat{F}_X(k), \widehat{F}_Y(l)\}]^2 d\Pi\{\widehat{F}_X(k), \widehat{F}_Y(l)\}}. \tag{6·3}$$

Its asymptotic variance $\widehat{V}_\Delta$ can be estimated using the delta method in a similar way as what was described in Section 5 for the concordance measures (cfr Eq. 5·3). If the statistic $T = \sqrt{n}\widehat{\Delta}/\sqrt{\widehat{V}_\Delta}$ is considered to test (6·2), the associated large sample p-value is

$$p_{\text{as}} = 1 - \Phi(T_{\text{obs}}),$$

where $\Phi(\cdot)$ is the cumulative standard normal distribution.

The efficiency of the asymptotic test procedure is again questionable because of the poor normal approximation at the boundary ($\Delta = 0$). A bootstrap test can be implemented as an alternative procedure to overcome this problem. To do so, we simulate independently R contingency tables under

$\widehat{H}_0 : \widehat{F}_{XY}(k,l) = C_{\widehat{\theta}}\{\widehat{F}_X(k), \widehat{F}_Y(l)\}$. This means that vectors of cell counts $n^*_{\widehat{\theta},kl}$ are randomly generated from a multinomial distribution counting n observations and having cell probabilities equal to the predicted proportions $dC_{\widehat{\theta}}\{\widehat{F}_X(k), \widehat{F}_Y(l)\}$ from the estimated or known model $C_{\widehat{\theta}}$. Simulated cell proportions $p^*_{\widehat{\theta},kl,r}$ and empirical distributions $F^*_{\widehat{\theta},XY,r}(k,l)$ are then derived from the counts $n^*_{\widehat{\theta},kl}$, for each simulation and the estimates $\Delta^*_{\widehat{\theta},r}$ are calculated using (6·3). The vector $(\Delta^*_{\widehat{\theta},1}, ..., \Delta^*_{\widehat{\theta},R})$ is a random sample of $\widehat{\Delta}$ under $\widehat{H}_0$. The p-value of the bootstrap lack of fit test is then approximately (Davison & Hinkley, 1997, p. 161)

$$p_{\text{boot}} = \frac{1 + \sum_{r=1}^{R} I(\Delta^*_{\theta,r,\text{obs}} \geq \widehat{\Delta}_{\text{obs}})}{R+1},$$

where $I(\cdot)$ is the indicator function.

Bootstrap estimators for the distribution of $\widehat{\Delta}$ (such as $\bar{\Delta}^*$) can also be produced. To do so, contingency tables are simulated using the *observed* frequencies p_{kl} as cell probabilities. This contrasts to the above simulation procedure under $\widehat{H}_0$ that was based on the *predicted* cell proportions from the subcopula model $C_{\widehat{\theta}}$. A random sample Δ^*_r, $r = 1, ..., R$ is then calculated using (6·3) and the construction of bootstrap estimators is similar to what was given for $\widehat{\rho}_C$ in Section 5.1.

7 Illustration in a Migraine Trial

The estimation and test procedures are now illustrated in an example. Table 2 presents pain and nausea severity scores (no, mild, moderate or severe) recorded on 44 migraine patients during an attack. We report empirical estimates $\widehat{\sigma}^2_C$, $\widehat{\rho}_C$ and $\widehat{\tau}_C$ together with their asymptotic standard errors (a.s.e.) in Table 2. Bootstrap ($R = 10000$) means $\bar{\sigma}^{*2}_C$, $\bar{\rho}^*_C$ and $\bar{\tau}^*_C$ are also given with standard deviations (std.) in the same table.

These data exhibit strong positive dependence: $\widehat{\tau}_C = 0.65$ (0.14) so that a large pain score is more likely to occur with severe nausea. Empirical estimates (5·2) of the joint distribution are superior to the product of the marginal distribution estimates in all cells of Table 2 (result not shown). As a result, $\widehat{\sigma}^1_C = 0.74$ is equal to $\widehat{\rho}_C$, suggesting PQD.

The mean bootstrap estimates are comparable to their empirical counterparts; the a.s.e. and std. estimates are also similar. The asymptotic normal approximation is good as the sample size ($n = 44$) is reasonably large and as all three coefficients are estimated well inside of their respective domain.

The dependence between pain and nausea was modeled using Frank (1979) bivariate copula

$$C_\theta(u,v) = -\frac{1}{\theta}\ln[1 + \frac{(e^{-u\theta}-1)(e^{-v\theta}-1)}{(e^{-\theta}-1)}], \tag{7·1}$$

with a single dependence parameter $\theta \in \Re_0$.

Limiting cases for (7·1) include M $(\theta \to -\infty)$, Π $(\theta \to 0)$ and W $(\theta \to +\infty)$. C_θ induces PQD when $\theta > 0$ and NQD when $\theta < 0$. Note also that $C_{\theta_1}(\cdot,\cdot) \prec C_{\theta_2}(\cdot,\cdot)$ when $\theta_1 \le \theta_2$.

Maximum likelihood estimate (MLE) $\widehat{\theta}$ was obtained conditionally on the observed marginal distributions and the variance of $\widehat{\theta}$ was derived from the observed information matrix at the MLE: $\widehat{\theta}$ is 6.88 (2.22), highlighting the significant positive dependence found in the data. The statistic (6·3) for goodness of fit $\widehat{\Delta} = 0.0017$ (0.0038) suggests that Δ is not significantly different from 0 ($p_{\text{as}} = 0.32$) according to the asymptotic test. This indicates that the fit of model (7·1) with $\widehat{\theta} = 6.88$ is acceptable. The bootstrap mean (with $R = 10000$) is $\bar{\Delta}^* = 0.0083$ (std. $= 0.0089$). The bootstrap distribution of $\widehat{\Delta}$ is skewed to the right: quartiles are estimated by $\Delta^*_{(0.25)} = 0.0025$, $\Delta^*_{(0.50)} = 0.0055$ and $\Delta^*_{(0.75)} = 0.0110$. The p-value of the bootstrap goodness of fit test is $p_{\text{boot}} = 0.23$. Although the normal approximation is poor, the p-values of the asymptotic and bootstrap goodness of fit tests are comparable. In the two cases, the hypothesis that the underlying subcopula is Frank with $\theta = 6.89$ is not rejected at the 5% level.

Acknowledgement

We wish to thank Professor Michel Denuit for his review and many constructive contributions on an early version of the manuscript.

Research support from the contract "Projet d'Actions de Recherche Concertées" (PARC No. 98/03-217) of the Belgian Government is gratefully acknowledged.

References

Agresti, A. (1990). *Categorical Data Analysis.* John Wiley & Sons, Inc.

Davison, A. C. & Hinkley, D. V. (1997). *Bootstrap Methods and Their Application.* Cambridge University Press.

Frank, M. J. (1979). On the simultaneous associativity of F(x,y) and x+y-F(x,y). *Aequationes Mathematicae* **19**, 194–226.

Table 2: Empirical and bootstrap estimates of dependence measures for pain and nausea severity scores during a migraine attack. Frank's copula with dependence parameter θ is adjusted using maximum likelihoods. P-value of the asymptotic and bootstrap goodness of fit tests are reported.

Nausea	Pain Score			
Score	Mild	Moderate	Severe	
No	1	5	2	
Mild	0	7	3	
Moderate	0	3	18	
Severe	0	0	5	
Dependence measures	estimate	(a.s.e.)	bootstrap mean	(std.)
σ_C^2	0.48	(0.18)	0.52	(0.17)
ρ_C	0.74	(0.11)	0.75	(0.12)
τ_C	0.65	(0.14)	0.67	(0.14)
Dependence model	MLE	(a.s.e.)		
θ	6.89	(2.22)		
Goodness of fit test	estimate	(a.s.e.)	bootstrap mean	(std.)
statistic	0.0017	(0.0038)	0.0083	(0.0089)
p-values	p_{as}=0.32		p_{boot}=0.23	

Fréchet, M. (1951). Sur les tableaux de corrélation dont les marges sont données. *Ann. Univ. Lyon, Sci.* **14**, 53–77.

Goodman, L. A. & Kruskal, W. H. (1954). Measures of association for cross classifications. *Journal of the American Statistical Association* pages 732–764.

Renyi, A. (1959). On measures of dependence. *Acta. Math. Acad. Sci. Hungar.* **10**, 441–451.

Scarsini, M. (1984). On measures of concordance. *Stochastica* **8**, 201–218.

Schweizer, B. (1991). Thirty years of copulas. In Dall'Aglio, G., Kotz, S., & Satinetti, G., editors, *Advances in Probability Distributions with Given Marginals. Beyond the Copulas*, Mathematics and Its Applications, pages 13–50. Kluwer Academic Publishers, Dordrecht.

Schweizer, B. & Wolff, E. F. (1981). On nonparametric measures of dependence for random variables. *The Annals of Statistics* **9**, 879–885.

Sklar, M. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publication Institut Statistiques Université Paris* **8**, 229–231.

Tchen, A. (1980). Inequalities for distributions with given marginals. *The Annals of Probability* **8**, 814–827.

VANDENHENDE, F. & LAMBERT, P. (2000). Modeling repeated ordered categorical data using copulas. *Discussion Paper* **00–25**. Institut de Statistique, Université Catholique de Louvain, Louvain-la-Neuve, Belgium.