The Annals of Statistics
1996, Vol. 24, No. 3, 1386–1399

ON THE ASYMPTOTIC DISTRIBUTION OF A GENERAL MEASURE OF MONOTONE DEPENDENCE[1]

By Donato Michele Cifarelli,[2] Pier Luigi Conti and Eugenio Regazzini[2]

Università L. Bocconi, Università La Sapienza and Università L. Bocconi

In this paper the asymptotic normality of a class of statistics, including Gini's index of cograduation and Spearman's rank correlation coefficient, is proved. The asymptotic normality is stated under a large class of alternatives including the bivariate distributions corresponding to a condition of lack of association introduced in Section 3. The problems of testing the hypothesis of lack of association and of constructing confidence intervals for the population index of cograduation are also considered.

1. Introduction and summary. Gini (1914) introduced the concept of cograduation index to provide a summarizing measure of monotone dependence between two ordered statistical characteristics. Two characteristics, X and Y, say, are *monotone dependent* if the support S of the cumulative distribution function (c.d.f.) of (X, Y) fulfills one of the two following conditions.

Concordance condition. $(x_1, y_1), (x_2, y_2) \in S$ and $x_1 < x_2 \Rightarrow y_1 \le y_2$.

Discordance condition. $(x_1, y_1), (x_2, y_2) \in S$ and $x_1 < x_2 \Rightarrow y_1 \ge y_2$.

Gini's index ranks the data separately within each component, and it has been used, mainly by Italian statisticians, to test stochastic independence as an alternative to other common test statistics such as Kendall's τ and Spearman's rank correlation ρ.

Gini's cograduation index and Spearman's ρ are particular cases of a more general index—briefly described in Section 2 of the present paper—which was introduced by Cifarelli and Regazzini (1990) and more recently was generalized by Conti (1993). From now on, the term "cograduation index" will designate the general index $\gamma_g(H)$ described in Section 2.

If X and Y have continuous c.d.f.'s, then the cograduation index vanishes not only when X and Y are stochastically independent, but even when the

---

Received October 1992; revised June 1995.

[1]Research supported in part by MURST (40%, 1992, Progetto "Inferenza statistica: basi probabilistiche e sviluppi metodologici") and by Università L. Bocconi (fondi ricerca IMQ, 1992).

[2]Also affiliated with CNR-IAMI, Milano, Italy.

AMS 1991 *subject classifications*. 62H20, 62G10, 62E20.

Key words and phrases. Empirical processes, functional limit theorem, index of cograduation, lack of association (= indifference), monotone dependence, *U*-statistics.

c.d.f. of (X, Y) does not exhibit any "tendency" to concordance or to discordance. This last circumstance is made precise in Section 3, and it is called *indifference*. Then, in order to decide whether X and Y are monotone dependent, it is reasonable to define a procedure of testing the hypothesis of indifference in such a way that, at a given level of significance α, the probability of rejecting that hypothesis is not greater than α, under *each indifferent c.d.f. for* (X, Y). An asymptotic procedure of this kind is explained in Section 6, where an optimal property of Gini's index—within the class of all cograduation indices—is singled out. The implementation of that procedure motivates the main goal of the present paper, namely, the deduction of the asymptotic distribution of a sample estimate of the cograduation index under a wide class of generating c.d.f.'s. This is done in Section 4 via the theory of convergence of empirical processes.

Hoeffding (1948) via the theory of U-statistics, proved the asymptotic normality of Spearman's rank correlation coefficient. Subsequently, the theory of U-statistics was employed by Cifarelli and Regazzini (1974, 1977) to state the asymptotic distribution of Gini's simple cograduation index. Section 5 of the present paper includes a few remarks about the asymptotic equivalence between the general cograduation index and a suitable U-statistic of degree 2. Such an equivalence is used to obtain a consistent estimator of the variance of the sample cograduation index. This estimator is used in Section 6 for determining confidence intervals for the population index of cograduation.

2. Definition of cograduation indices in a bivariate population. Throughout this paper, F and G denote two *continuous* univariate c.d.f.'s and $\Gamma(F, G)$ designates the class of all bivariate c.d.f.'s with first and second marginal c.d.f.'s F and G, respectively (*Fréchet class*). Given H_1 and H_2 in $\Gamma(F, G)$, one says that H_1 is *more concordant*, or *more positively quadrant dependent* (PQD) than H_2, if $H_1(x, y) \geq H_2(x, y)$ for every (x, y) in $\mathbb{R}^2$ (in symbols, $H_1 \succcurlyeq H_2$). PQD ordering, originally introduced by Gini around 1915, was reconsidered, more recently, by Yanagimoto and Okamoto (1969), Cambanis, Simons and Stout (1976), Kimeldorf and Sampson (1978, 1987) and Tchen (1980). It is well known that H^+ and H^-, defined on $\mathbb{R}^2$ by

$$H^+(x, y) = \min\{F(x), G(y)\},$$

$$H^-(x, y) = \max\{0, F(x) + G(y) - 1\},$$

are elements of $\Gamma(F, G)$ and that

$$H^-(x, y) \leq H(x, y) \leq H^+(x, y)$$

holds for every H in $\Gamma(F, G)$ and for every (x, y) in $\mathbb{R}^2$.

Now, two random variables (r.v.'s) X and Y with c.d.f.'s F and G, respectively, turn out to be concordant (discordant, respectively) if, and only if, their joint c.d.f. coincides with H^+ (H^-, respectively). Hence, in the case of concordance, the random point $(F(X), G(Y))$ lies almost surely (a.s.) on the straight line $S_1 = \{(x, y)\colon y = x\}$, while in the case of discordance, the same point lies a.s. on the straight line $S_2 = \{(x, y)\colon y = 1 - x\}$.

We are now in a position to propose a summarizing measure of the monotone dependence of each element H in $\Gamma(F, G)$. It is a function C: $\Gamma(F, G) \to [-1, 1]$ such that:

1. $C(H) = 1$ (-1, respectively) if and only if $H = H^{+}$ (H^{-}, respectively);
2. $C(H_1) \geq C(H_2)$ whenever H_1 and H_2 belong to $\Gamma(F, G)$ and $H_1 \succcurlyeq H_2$.

Note that the horizontal (= vertical) distance from $(F(X), G(Y))$ to S_2 provides a picture of the concordance of the c.d.f. H of (X, Y). Hence, if g is any strictly increasing continuous function (with inverse g^{-1}) on $I = [0, 1]$, the quantity

$$\tilde{\nu}_g(H) = g^{-1}\left(\int_{\mathbb{R}^2} g(|F(x) + G(y) - 1|)\, dH(x, y)\right)$$

represents a (strictly) monotone mean of the above random distance. In Cifarelli and Regazzini (1990) it is shown that if g is *convex*, then $\tilde{\nu}_g$ agrees with the PQD ordering and therefore it can be considered as a measure of concordance on $\Gamma(F, G)$. Similarly, one can interpret

$$\tilde{\delta}_g(H) = g^{-1}\left(\int_{\mathbb{R}^2} g(|F(x) - G(y)|)\, dH(x, y)\right)$$

as a measure of discordance on $\Gamma(F, G)$. In order to obtain a measure of monotone dependence, we should combine $\tilde{\nu}_g$ and $\tilde{\delta}_g$. Without loss of generality, from now on we will assume that $g(0) = 0$ and, with no loss of information, we will consider $\nu_g = g(\tilde{\nu}_g)$ and $\delta_g = g(\tilde{\delta}_g)$ instead of $\tilde{\nu}_g$ and $\tilde{\delta}_g$, respectively. Then, since

$$\max_{H \in \Gamma} \nu_g(H) = \nu_g(H^{+}) = \delta_g(H^{-}) = \max_{H \in \Gamma} \delta_g(H) = \int_I g(x)\, dx,$$

one easily finds that

$$\gamma_g(H) = \frac{\nu_g(H) - \delta_g(H)}{\int_I g(x)\, dx} = \frac{\int_{I^2} [g(|x + y - 1|) - g(|x - y|)]\, dM_H(x, y)}{\int_I g(x)\, dx}$$

is a measure of monotone dependence, where M_H denotes the c.d.f. of $(F(X), G(Y))$, and provided that the following assumption holds.

ASSUMPTION 1. g: $I \to \mathbb{R}$ is a strictly increasing, continuous and convex function, such that $g(0) = 0$.

This assumption implies that there exists a nonnegative and nondecreasing function l: $I \to \mathbb{R}$ such that

$$g(x) = \int_0^x l(t)\, dt, \qquad x \in I, \tag{2.1}$$

and

$$g(|x - y|) = g(|z - y|) + \int_0^{x - z} \operatorname{sgn}(z + t - y)\, l(|t + z - y|)\, dt, \qquad (x, y, z) \in I^3. \tag{2.2}$$

In Section 4, the following additional technical assumptions will be used:

ASSUMPTION 2. $l\colon I \to \mathbb{R}$ is continuous.

ASSUMPTION 3. $\int_{S_1} dM_H(x, y) = \int_{S_2} dM_H(x, y) = 0.$

At any rate, under Assumption 1, γ_g is the *cograduation index* mentioned in Section 1. In fact, for $g(x) = x$ [$g(x) = x^2$, respectively], γ_g provides the continuous version of Gini's simple cograduation index (Spearman's rank correlation index, respectively). The reader is referred to Salvemini (1951) for a few remarks about Gini's and Spearman's indices. A nice property of Gini's index, connected with the problem of testing indifference of (X, Y), is shown in Section 6.

Precise and complete formulations of the statements reviewed in the present section can be found in Cifarelli and Regazzini (1990).

3. A definition of indifferent c.d.f. The aim of this section is to provide a formal definition of "lack of association" (indifference) of a bivariate distribution, and to single out the class of bivariate c.d.f.'s belonging to $\Gamma(F, G)$ which are "indifferent" with respect to (w.r.t.) the PQD ordering. In view of this ordering, a c.d.f. H in $\Gamma(F, G)$ is the more concordant (discordant, respectively) the higher the probability that M_H concentrates in rectangles of the type $\mathscr{R}_1$ and $\mathscr{R}_4$ ($\mathscr{R}_2$ and $\mathscr{R}_3$, respectively), where, given any (x, y) in I^2:

$\mathscr{R}_1$ has vertices $(0, 0)$, $(x, 0)$, (x, y) and $(0, y)$.

$\mathscr{R}_2$ has vertices $(0, 1 - y)$, $(x, 1 - y)$, $(x, 1)$ and $(0, 1)$.

$\mathscr{R}_3$ has vertices $(1 - x, 0)$, $(1, 0)$, $(1, y)$ and $(1 - x, y)$.

$\mathscr{R}_4$ has vertices $(1 - x, 1 - y)$, $(1, 1 - y)$, $(1, 1)$ and $(1 - x, 1)$.

Hence, in order that a c.d.f. H in $\Gamma(F, G)$ may be considered indifferent, one can require that M_H concentrates the same probability in each $\mathscr{R}_1$, $\mathscr{R}_2$ and $\mathscr{R}_3$. These remarks justify the following definition.

DEFINITION 3.1. The c.d.f. H in $\Gamma(F, G)$ is said to be *indifferent* if

$$M_H(x, y) = x - M_H(x, 1 - y) = y - M_H(1 - x, y) \tag{3.1}$$

for every (x, y) in I^2.

Clearly, (3.1) is the same as assuming that

$$H(x, y) = F(x) - H\big(x, G^{-1}(1 - G(y))\big) = G(y) - H\big(F^{-1}(1 - F(x)), y\big)$$

holds for every (x, y) in the interior of the support of H. Here, if V is a c.d.f. on $\mathbb{R}$, then $V^{-1}(x) = \inf\{t\colon V(t) \ge x\}$ for every x in $I^\circ = (0, 1)$. Moreover, (3.1) is equivalent to $(F(X), G(Y))$, $(1 - F(X), G(Y))$ and $(F(X), 1 - G(Y))$ being equal in distribution; in other words, under (3.1), the c.d.f. of $(F(X), G(Y)) =$ (grade X, grade Y) does not change when either of these grades is reversed.

Whenever (3.1) holds, one obtains $\gamma_g(H) = 0$, whatever g may be. It is worth noting that Kendall's τ also vanishes when (3.1) holds.

A few examples of indifferent c.d.f.'s in $\Gamma(F, G)$ are given below.

EXAMPLE 3.1. It is easy to show that the following c.d.f.'s are indifferent:

$$H(x, y) = F(x)G(y),$$

$$H(x, y) = \{H^+(x, y) + H^-(x, y)\}/2, \tag{3.2}$$

$$H(x, y) = \sum_{i=1}^{n} \lambda_i H_i(x, y) \quad \left(H_i \text{ indifferent}, \lambda_i > 0, i = 1, \dots, n, \sum_{i=1}^{n} \lambda_i = 1\right)$$

EXAMPLE 3.2. Given $0 < \varepsilon \le 1$, let us introduce the sets

$$A_\varepsilon = \{(x, y) \in I^2 : |x - y| \le \varepsilon\}, \qquad B_\varepsilon = \{(x, y) \in I^2 : |x + y - 1| \le \varepsilon\}$$

and the two absolutely continuous c.d.f.'s H_ε^+ and H_ε^-, whose density functions are given by

$$h_\varepsilon^+(x, y) = f(\varepsilon)\mathbf{1}_{A_\varepsilon}(x, y), \qquad h_\varepsilon^-(x, y) = f(\varepsilon)\mathbf{1}_{B_\varepsilon}(x, y),$$

respectively, with $f(\varepsilon) = \{1 - (1 - \varepsilon)^2\}^{-1}$ and $\mathbf{1}_A$ = indicator of A. After setting

$$F_\varepsilon(x) = \int_{(-\infty, x] \times \mathbb{R}} h_\varepsilon^+(u, y)\, du\, dy$$

for every x in $\mathbb{R}$, it is easy to check that H_ε^+ and H_ε^- belong to $\Gamma(F_\varepsilon, F_\varepsilon)$. Moreover, since the relation $1 - F_\varepsilon(x) = F_\varepsilon(1 - x)$ holds for every x in I, then

$$F_\varepsilon^{-1}(1 - F_\varepsilon(x)) = 1 - x \tag{3.3}$$

for every x in $I°$. In view of (3.3), the c.d.f.

$$H_\varepsilon^*(x, y) = [H_\varepsilon^+(x, y) + H_\varepsilon^-(x, y)]/2, \qquad (x, y) \in I^2,$$

is indifferent in $\Gamma(F_\varepsilon, F_\varepsilon)$.

REMARK 3.1. Since

$$\lim_{\varepsilon \to 0^+} M_{H_\varepsilon^*}(x, y) = \lim_{\varepsilon \to 0^+} H_\varepsilon^*(x, y) = [\max\{x + y - 1, 0\} + \min\{x, y\}]/2$$

for every (x, y), Example 3.2 shows that (3.2), with $F = G$ = uniform c.d.f. on I, can always be considered as the weak limit of a sequence of (absolutely) continuous bivariate c.d.f.'s whose marginals are uniform on I.

4. Asymptotic distribution of the sample cograduation index. Let $(X_1, Y_1), \dots, (X_n, Y_n)$ be a random sample from a bivariate c.d.f. H in $\Gamma(F, G)$. One can assume that each (X_i, Y_i) is a real random vector defined on a probability space, say $(\Omega, \mathcal{F}, P)$, such that

$$P\left(\bigcap_{i=1}^{n} \{\omega : X_i(\omega) \le x_i, Y_i(\omega) \le y_i\}\right) = \prod_{i=1}^{n} H(x_i, y_i)$$

for every (x_i, y_i) in $\mathbb{R}^2$, $i = 1, \dots, n$, and $n \geq 2$. Hence, letting

$$V = \bigcap_{n=2}^{\infty} \bigcap_{1 \leq i < j \leq n} \{\omega\colon X_i(\omega) \neq X_j(\omega) \text{ and } Y_i(\omega) \neq Y_j(\omega)\}$$

one has $P(V) = 1$, because F and G are continuous c.d.f.'s. Then, the cograduation index for the above sample is defined on V by

$$\gamma_{g,n} = \frac{1}{K_n} \sum_{i=1}^{n} \left\{ g\left(\frac{1}{n} |R_i(\omega) + S_i(\omega) - n - 1| \right) - g\left(\frac{1}{n} |R_i(\omega) - S_i(\omega)| \right) \right\},$$

where $R_i(\omega) = \operatorname{rank}(X_i(\omega))$ among the X_i's, $S_i(\omega) = \operatorname{rank}(Y_i(\omega))$ among the Y_i's and

$$K_n = \sum_{i=1}^{n} g\left(\frac{1}{n} |2i - n - 1| \right).$$

When $g(x) = x$, $\gamma_{g,n}$ coincides with the *simple cograduation index* originally defined by Gini for finite populations without ties. When $g(x) = x^2$, $\gamma_{g,n}$ becomes the classical *Spearman rank correlation coefficient.*

To obtain an extension of $\gamma_{g,n}$ to the whole space Ω, one can resort to Section 3 in Cifarelli and Regazzini (1990). Since $P(V^c) = 0$, the distribution of such an extension does not depend on its definition on V^c. From now on, $\gamma_{g,n}$ will designate the extension which vanishes on V^c.

We now show that $\sqrt{n}(\gamma_{g,n} - \gamma_g(H))$ converges in law to a normal r.v. as $n \to \infty$. In view of Assumption 1, it is possible to write

$$\lim_{n \to \infty} \frac{K_n}{n} = \int_0^1 g(x)\, dx = K$$

and

$$\gamma_{g,n} = \hat{\gamma}_{g,n} + O\left(\frac{1}{n} \right),$$

where

$$\hat{\gamma}_{g,n} = \frac{1}{K} \int_{\mathbb{R}^2} \{ g(|F_n(x) + G_n(y) - 1|) - g(|F_n(x) - G_n(y)|) \}\, dH_n(x, y)$$

and H_n, F_n and G_n denote the joint and marginal empirical c.d.f.'s of the random sample, respectively.

The asymptotic normality of $\sqrt{n}(\hat{\gamma}_{g,n} - \gamma_g(H))$ will be deduced from some basic results concerning two-dimensional empirical processes. A review is in Gaenssler and Stute [(1979), Section 2.1]. Let D_2 be the set of all real functions f on I^2, such that for each $\mathbf{t} = (t_1, t_2)$ the limit $\lim_{n \to \infty} f(\mathbf{t}_n)$ exists for all sequences $\{\mathbf{t}_n\}_{n \geq 1}$ converging to $\mathbf{t}$ in some quadrant with corner $\mathbf{t}$, and such that f is continuous from above. Then, there exists a separable and complete Skorohod's metric d_2 in D_2 [Neuhaus (1971)] whose corresponding σ-field $\mathscr{B}(d_2)$ of Borel sets equals the smallest σ-field of subsets of D_2 for which all coordinate mappings are measurable. Hence, $\beta_n(x, y) = \sqrt{n}(H_n(x, y) - H(x, y))$, with (x, y) in I^2 (the marginal c.d.f.'s F and G are assumed uniform on I), is a random element in D_2, and β_n converges weakly to β_0 on $(D_2, \mathscr{B}(d_2))$ as n goes to infinity, where β_0 is a Brownian sheet, that is, a centered Gaussian process with continuous sample paths (with probability 1) tied down to zero at (1, 1), and covariance function

$$\operatorname{Cov}(\beta_0(t_1,t_2),\beta_0(s_1,s_2)) = H(\mathbf{t}\wedge\mathbf{s}) - H(\mathbf{t})H(\mathbf{s}), \qquad \mathbf{t},\mathbf{s}\in I^2,$$

where $\mathbf{t}\wedge\mathbf{s} = (\min(t_1,s_1),\min(t_2,s_2))$. Finally, from Skorohod's representation theorem, there exist versions $\hat{\beta}_n$ of β_n (in the sense that $\hat{\beta}_n$ and β_n have the same distribution for every n) on an appropriate probability space $(\hat{\Omega},\hat{\mathscr{F}},\hat{P})$ such that $d_2(\hat{\beta}_n,\hat{\beta}_0)\to 0$. Then, since $\hat{\beta}_0$ has continuous sample paths a.s. $(\hat{P})$, (2.1) and (2.2) in Neuhaus (1971) imply that $\varrho(\hat{\beta}_n,\hat{\beta}_0) = \sup_{\mathbf{t}\in I^2}|\hat{\beta}_n(\mathbf{t}) - \hat{\beta}_0(\mathbf{t})| \to 0$ a.s. $(\hat{P})$ as $n\to\infty$.

THEOREM 4.1. *If $(X_1,Y_1),\dots,(X_n,Y_n),\dots$ is a sequence of i.i.d. random vectors with common c.d.f. H in $\Gamma(F,G)$ and if Assumptions* 1–3 *hold, then the asymptotic distribution* (*as $n\to\infty$*) *of*

$$\sqrt{n}\,(\gamma_{g,n} - \gamma_g(H))$$

is normal with mean 0 *and variance σ_g^2 defined by* (4.1).

PROOF. It suffices to prove the asymptotic normality of $\sqrt{n}\,(\hat{\gamma}_{g,n} - \gamma_g(H)) = (\Delta_{1,n} + \Delta_{2,n} + \Delta_{3,n})/K$, where F and G are supposed to be uniform on I and

$$\Delta_{1,n} = \sqrt{n}\int_{I^2}\Big\{g\big(\big|1-x-y-\psi_n(x,y)/\sqrt{n}\,\big|\big) - g(|1-x-y|)$$
$$-g\big(\big|x-y+\lambda_n(x,y)/\sqrt{n}\,\big|\big) + g(|x-y|)\Big\}\,dH(x,y),$$

$$\Delta_{2,n} = \int_{I^2}\{g(|1-x-y|) - g(|x-y|)\}\,d\beta_n(x,y),$$

$$\Delta_{3,n} = \int_{I^2}\Big\{g\big(\big|1-x-y-\psi_n(x,y)/\sqrt{n}\,\big|\big) - g(|1-x-y|)$$
$$-g\big(\big|x-y+\lambda_n(x,y)/\sqrt{n}\,\big|\big) + g(|x-y|)\Big\}\,d\beta_n(x,y),$$

$$\psi_n(x,y) = \beta_n(x,1) + \beta_n(1,y), \qquad \lambda_n(x,y) = \beta_n(x,1) - \beta_n(1,y).$$

After observing that $(\Delta_{1,n},\Delta_{2,n},\Delta_{3,n})$ and $(\hat{\Delta}_{1,n},\hat{\Delta}_{2,n},\hat{\Delta}_{3,n})$ have the same probability distribution (if the quantities $\hat{\Delta}_{j,n}$, $j=1,2,3$, are defined in the same way as the quantities $\Delta_{j,n}$, except for the terms β_n that are replaced by the corresponding terms $\hat{\beta}_n$) one splits the proof into three steps.

Claim 1. We have

$$\hat{\Delta}_{1,n} \to \Pi_1$$
$$= \int_{I^2}\big\{\hat{\beta}_0(x,1)[1-\varphi_{21}(x,y)] + \hat{\beta}_0(1,y)[1-\varphi_{12}(x,y)]\big\}$$
$$\times\, dg(|x+y-1|)$$

a.s. $(\hat{P})$ as $n\to\infty$, where

$$\varphi_{21}(x,y) = F_{21}(y|x) + F_{21}(1-y|x),$$
$$\varphi_{12}(x,y) = F_{12}(x|y) + F_{12}(1-x|y)$$

and $F_{21}(\cdot|x)$ [$F_{12}(\cdot|y)$, respectively] is the conditional c.d.f. of Y (X, respectively) given $X = x$ ($Y = y$, respectively).

To prove Claim 1, one observes that (2.1) implies

$$\Pi_1 = \int_{I^2}\Big\{\hat{\beta}_0(x,1)\big[\operatorname{sgn}(x+y-1)l(|x+y-1|) - \operatorname{sgn}(x-y)l(|x-y|)\big] - \hat{\beta}_0(1,y)\big[\operatorname{sgn}(1-x-y)l(|x+y-1|) + \operatorname{sgn}(y-x)l(|x-y|)\big]\Big\}\,dH(x,y)$$

and that (2.2) yields

$$g\left(\left|1-x-y-\frac{\hat{\psi}_n(x,y)}{\sqrt{n}}\right|\right) = g(|1-x-y|) + \frac{1}{\sqrt{n}}\left\{\int_0^{\hat{\beta}_n(x,1)}\operatorname{sgn}\left(x+y-1+\frac{t+\hat{\beta}_n(1,y)}{\sqrt{n}}\right) \times l\left(\left|x+y-1+\frac{t+\hat{\beta}_n(1,y)}{\sqrt{n}}\right|\right)dt + \int_0^{-\hat{\beta}_n(1,y)}\operatorname{sgn}\left(1-x-y+\frac{t}{\sqrt{n}}\right) \times l\left(\left|1-x-y+\frac{t}{\sqrt{n}}\right|\right)dt\right\}.$$

Thanks to Lebesgue's dominated convergence theorem, the term in braces converges a.s. ($\hat{P}$) to

$$\hat{\beta}_0(x,1)\operatorname{sgn}(x+y-1)l(|x+y-1|) - \hat{\beta}_0(1,y)\operatorname{sgn}(1-x-y)l(|x+y-1|)$$

for every (x, y) in $J = I^2 \setminus \{S_1 \cup S_2\}$. In the same way, one obtains

$$g\left(\left|x-y+\frac{\hat{\lambda}_n(x,y)}{\sqrt{n}}\right|\right) = g(|x-y|) + \frac{1}{\sqrt{n}}\left\{\int_0^{\hat{\beta}_n(x,1)}\operatorname{sgn}\left(x-y+\frac{t-\hat{\beta}_n(1,y)}{\sqrt{n}}\right) \times l\left(\left|x-y+\frac{t-\hat{\beta}_n(1,y)}{\sqrt{n}}\right|\right)dt + \int_0^{\hat{\beta}_n(1,y)}\operatorname{sgn}\left(y-x+\frac{t}{\sqrt{n}}\right)l\left(\left|y-x+\frac{t}{\sqrt{n}}\right|\right)dt\right\}$$

and the term under braces converges a.s. $(\hat{P})$ to

$$\hat{\beta}_0(x,1)\operatorname{sgn}(x-y)l(|x-y|) + \hat{\beta}_0(1,y)\operatorname{sgn}(y-x)l(|x-y|) \qquad (x,y) \in J.$$

At this stage, Claim 1 follows from Assumption 3 and the dominated convergence theorem.

Claim 2. $\hat{\Delta}_{3,n} \to 0$ a.s. $(\hat{P})$ as $n \to \infty$.

By resorting to a well-known multivariate version of the theorem of integration by parts [cf. Hildebrandt (1963), page 127], one has

$$\int_{I^2}\Big\{g\big(\big|x-y+\hat{\lambda}_n(x,y)/\sqrt{n}\,\big|\big) - g(|x-y|)\Big\}\, d\hat{\beta}_n(x,y) = W_{1,n} + W_{2,n} + W_{3,n},$$

where

$$W_{1,n} = -\int_I \hat{\beta}_n(x,1)\, d\Big\{g\big(\big|x-1+\hat{\beta}_n(x,1)/\sqrt{n}\,\big|\big) - g(1-x)\Big\},$$

$$W_{2,n} = -\int_I \hat{\beta}_n(1,y)\, d\Big\{g\big(\big|1-y+\hat{\beta}_n(1,y)/\sqrt{n}\,\big|\big) - g(1-y)\Big\},$$

$$W_{3,n} = \int_{I^2} \hat{\beta}_n(x,y)\, d\Big\{g\big(\big|x-y+\hat{\lambda}_n(x,y)/\sqrt{n}\,\big|\big) - g(|x-y|)\Big\}.$$

Taking into account that the function g has total variation $g(1)$ and, consequently, the function $g(|x-y+\lambda_n(x,y)/\sqrt{n}\,|) - g(|x-y|)$ has total variation less than or equal to $2g(1)$, one obtains

$$|W_{3,n}| \le \left|\int_{I^2} \hat{\beta}_0(x,y)\, d\Big\{g\big(\big|x-y+\hat{\lambda}_n(x,y)/\sqrt{n}\,\big|\big) - g(|x-y|)\Big\}\right| + 2g(1)\varrho\big(\hat{\beta}_n, \hat{\beta}_0\big),$$

and the conclusion $W_{3,n} \to 0$ a.s. $(\hat{P})$ easily follows.

In a similar way it can be shown that $W_{1,n} \to 0$ and $W_{2,n} \to 0$ a.s. $(\hat{P})$. This proves Claim 2.

Claim 3. $\hat{\Delta}_{2,n} \to \Pi_2 = \int_{I^2}\{\hat{\beta}_0(x,y) + \hat{\beta}_0(x,1-y) - \hat{\beta}_0(x,1) - \hat{\beta}_0(1,y)\}dg(|x+y-1|)$ a.s. $(\hat{P})$, as $n \to \infty$.

Integration by parts yields

$$\hat{\Delta}_{2,n} = -\int_I \hat{\beta}_n(x,1)\, d\{g(x) - g(1-x)\} - \int_I \hat{\beta}_n(1,y)\, d\{g(y) - g(1-y)\} + \int_{I^2} \hat{\beta}_n(x,y)\, d\{g(|x+y-1|) - g(|x-y|)\}$$

and Claim 3 follows from the Lebesgue dominated convergence theorem.

In conclusion, the asymptotic distribution of $\sqrt{n}(\hat{\gamma}_{g,n} - \gamma_g(H))$ coincides with the distribution of the r.v.

$$Y = \frac{1}{K}\int_{I^2}\{\hat{\beta}_0(x,y) + \hat{\beta}_0(x,1-y) - \hat{\beta}_0(x,1)\varphi_{21}(x,y) - \hat{\beta}_0(1,y)\varphi_{12}(x,y)\}\,dg(|x+y-1|),$$

which is normal with mean zero and variance

$$\sigma_g^2 = \frac{1}{K^2}\int_{I^2}\int_{I^2}\{\hat{E}(\xi(x,y)\xi(s,t))\}\,dg(|x+y-1|)\,dg(|s+t-1|), \tag{4.1}$$

where $\hat{E}$ denotes expectation w.r.t. $\hat{P}$,

$$\xi(x,y) = \hat{\beta}_0(x,y) + \hat{\beta}_0(x,1-y) - \hat{\beta}_0(x,1)\varphi_{21}(x,y) - \hat{\beta}_0(1,y)\varphi_{12}(x,y)$$

and

$$\hat{E}(\hat{\beta}_0(t_1,t_2)\hat{\beta}_0(s_1,s_2)) = M_H(t_1\wedge s_1, t_2\wedge s_2) - M_H(t_1,t_2)M_H(s_1,s_2). \quad \square$$

REMARK 4.1. If $H \neq H^+, H^-$ and $l(0) = 0$, then Assumption 3 can be dropped.

REMARK 4.2. Assumptions 2 and 3 can be weakened. In fact, Theorem 4.1 holds provided that the probability measure generated by H gives the set

$$\bigcup_{d\in D}\{(x,y)\in\mathbb{R}^2 : |F(x)-G(y)| = d \vee |F(x)+G(y)-1| = d\}$$

probability zero, D being the set of discontinuity points of l.

REMARK 4.3. If H is indifferent, then, from Theorem 4.1, $\hat{\Delta}_{1,n} \to 0$ a.s. $(\hat{P})$ as $n\to\infty$. Consequently, the asymptotic distribution of $\sqrt{n}(\gamma_{g,n} - \gamma_g(H))$ coincides with that of $\hat{\Delta}_{2,n}/K$, which is normal with mean zero and variance

$$\frac{2}{K^2}\int_{I^2}(g^2(|x-y|) - g(|x-y|)g(|1-x-y|))\,dM_H(x,y). \tag{4.2}$$

In particular, if $g(x) = x$ (Gini's cograduation index), (4.2) becomes

$$32\int_0^1 |1-2x| M_H(x,x)\,dx - 16/3$$

and if $g(x) = x^2$ (Spearman's rank correlation coefficient), (4.2) becomes

$$144\int_{I^2} x^2y^2\,dM_H(x,y) - 15.$$

5. Estimation of asymptotic variance. In this section we introduce and analyze the main properties of the U-statistic (of degree 2)

$$U_{g,n} = \binom{n}{2}^{-1}\sum_{i<j} k(x_i, y_i; x_j, y_j),$$

where

$$
\begin{aligned}
k(x_i, y_i; x_j, y_j) &= \left[v(x_i, y_i; x_j, y_j) + v(x_j, y_j; x_i, y_i)\right]/(2K),\\
v(x_i, y_i; x_j, y_j) &= g(|F(x_i) + G(y_i) - 1|) - g(|F(x_i) - G(y_i)|)\\
&\quad + \operatorname{sgn}(F(x_i) + G(y_i) - 1)l(|F(x_i) + G(y_i) - 1|)\\
&\qquad \times\left\{u(x_j, x_i) + u(y_j, y_i) - [F(x_i) + G(y_i)]\right\}\\
&\quad + \operatorname{sgn}(F(x_i) - G(y_i))l(|F(x_i) - G(y_i)|)\\
&\qquad \times\left\{u(y_j, y_i) - u(x_j, x_i) - [G(y_i) - F(x_i)]\right\}
\end{aligned}
$$

and $u(x, y) = 1$ if $y \geq x$, $u(x, y) = 0$ if $y < x$.

In fact, using this statistic, one can determine a consistent estimator of the asymptotic variance of $\gamma_{g,n}$. The basic result is stated in the following proposition. Its proof, almost identical to the one of Theorem 4.1, is omitted [cf. Conti (1993)].

THEOREM 5.1. *Under Assumptions* 1–3,

$$\sqrt{n}\left(U_{g,n} - \gamma_{g,n}\right) \to_P 0 \quad \textit{and} \quad \sqrt{n}\left(U_{g,n} - \hat{\gamma}_{g,n}\right) \to_P 0$$

as $n \to \infty$.

Thus, if the marginal c.d.f.'s F and G of H are assigned, then $U_{g,n}$ is an unbiased, consistent estimator of the population cograduation index $\gamma_g(H)$. From Theorem 5.1, $\gamma_{g,n}$ has the same asymptotic distribution as $U_{g,n}$, which is normal by virtue of well-known results by Hoeffding (1948). Consequently, $\gamma_{g,n}$ and $U_{g,n}$ have the same asymptotic variance and, from a theorem in Sen (1960), the estimator

$$S^2_{g,n} = \frac{1}{n}\sum_{i=1}^{n}\left(V_i - U_{g,n}\right)^2 = \frac{1}{n}\sum_{i=1}^{n} V_i^2 - U^2_{g,n}$$

$[V_i = \sum_{1 \leq j \neq i \leq n} k(x_i, y_i; x_j, y_j)/(n-1),\ i = 1, \dots, n]$ is a consistent estimator of $\sigma_g^2/4$. Unfortunately, in our case $S^2_{g,n}$ cannot be directly employed to estimate σ_g^2 because it involves the unknown marginal c.d.f.'s F and G. Instead of $S^2_{g,n}$ one can consider the statistic

$$\hat{S}^2_{g,n} = \frac{1}{n}\sum_{i=1}^{n}\left(\hat{V}_i - \hat{\gamma}_{g,n}\right)^2,$$

where the $\hat{V}_i$'s are defined in the same way as the V_i's, with F and G replaced by F_n and G_n, respectively. The following proposition holds.

THEOREM 5.2. *Under Assumptions* 1–3, $\hat{S}^2_{g,n} \to_P \sigma_g^2/4$ *as* $n \to \infty$.

A proof of Theorem 5.2 can be found in Conti (1993).

6. A few applications of the previous results. This section deals with two classical inferential problems in the presence of a large sample size:

PROBLEM A. Testing the hypothesis of indifference between two statistical characteristics.

PROBLEM B. Determining a confidence interval for the population index of cograduation.

As far as Problem A is concerned, define:

1. $\mathscr{G}$ to be the class of all functions satisfying Assumptions 1 and 2.
2. Γ' to be the class of all c.d.f.'s in $\Gamma(F, G)$ satisfying Assumption 3.
3. $\mathscr{I}_0$ to be the class of all indifferent c.d.f.'s in Γ'.

Thanks to Theorem 4.1, if H belongs to $\mathscr{I}_0$, if g is an element of $\mathscr{G}$ and if n is sufficiently large, then $\sqrt{n}\,\gamma_{g,n}$ can be (approximately) considered as normally distributed with mean zero and variance (4.2). Consequently, a conservative test of the composite hypothesis H_0: H belongs to $\mathscr{I}_0$ versus the alternative H_1: H belongs to $\Gamma' \setminus \mathscr{I}_0$ should reject H_0 whenever $|\gamma_{g,n}| > c$, where $\Phi(-c\sqrt{n}/\tilde{\sigma}_g) = \alpha/2$, α represents the significance level of the test, Φ is the standard normal c.d.f. and $\tilde{\sigma}_g = \sup\{\sigma_g(H)\colon H \in \mathscr{I}_0\}$. The conservative character of such a procedure could be mitigated by selecting the g's for which $\tilde{\sigma}_g$ attains its minimum value. Now, since the statements

$$\text{if } g(x) = x, \text{ then } \tilde{\sigma}_g = \sqrt{4/3} \tag{6.1}$$

and

$$\text{if } g \in \mathscr{G} \text{ and } g(x) \neq x \text{ for some } x, \text{ then } \tilde{\sigma}_g > \sqrt{4/3} \tag{6.2}$$

hold under the assumptions of Theorem 4.1, then the previous asymptotic procedure to test H_0 against H_1 leads us to prefer Gini's index as a test statistic, in the class of all cograduation indices. The resulting asymptotic rule consists in rejecting H_0 if $|\gamma_{g,n}| > c$, where, given the significance level α, the critical value c is determined by $\Phi(-c\sqrt{3n}/2) = \alpha/2$.

In particular, Gini's index should be preferred to the more popular Spearman rank correlation coefficient.

To complete the present point, it only remains to prove (6.1) and (6.2). One can start by observing that, for every H in $\mathscr{I}$ [= class of all indifferent c.d.f.'s in $\Gamma(F, G)$],

$$M_H(x, x) \le M_{H^*}(x, x) = \begin{cases} x/2, & \text{if } x \in [0, 1/2], \\ (3x - 1)/2, & \text{if } x \in [1/2, 1], \end{cases}$$

where H^* coincides with (3.2). Therefore, in view of Remark 3.1, one has

$$g(x) = x \ (x \in I) \quad \Rightarrow \quad \tilde{\sigma}_g = 32\int_0^1 |1 - 2x| M_{H^*}(x, x)\, dx - 16/3 = 4/3.$$

This proves (6.1). To prove (6.2), after indicating the value of (4.2) at M_{H^*} by $R(g)$, one observes that $R(g)$ is an accumulation point of the set $\{\sigma_g(H): H \in \mathcal{J}_0\}$ for every g satisfying Assumption 1; see Remark 3.1. Hence,

$$\tilde{\sigma}_g^2 \geq R(g) = \int_0^1 g(x)^2\, dx \Big/ \left(\int_0^1 g(x)\, dx \right)^2$$

and to prove (6.2) it suffices to show that

$$R(g) > 4/3 \tag{6.3}$$

holds for every g satisfying Assumption 1 and such that $g(x) \not\equiv x$ on I. Inequality (6.3) is true if and only if

$$\int_0^1 \phi(g(x)/2)\, dx > \int_0^1 \phi\left(x \int_0^1 g(u)\, du \right) dx \tag{6.4}$$

with $\phi(x) = x^2$ on I. In view of the strict convexity of ϕ and a refinement of a proposition by Hardy, Littlewood and Pólya (1929), (6.4) holds if

$$\tfrac{1}{2} \int_\xi^1 g(x)\, dx - \int_\xi^1 x \left(\int_0^1 g(u)\, du \right) dx \geq 0 \tag{6.5}$$

for all ξ in $[0, 1]$, and the set

$$\left\{ \xi \in [0,1]: \tfrac{1}{2} \int_\xi^1 g(x)\, dx - \int_\xi^1 x \left(\int_0^1 g(u)\, du \right) dx > 0 \right\} \tag{6.6}$$

has positive Lebesgue measure. In fact, thanks to (2.1), the left-hand side of (6.5) reduces to

$$\begin{aligned} &\frac{\xi^2}{2} \int_0^1 \left(\int_0^x l(u)\, du \right) dx - \frac{1}{2} \int_0^\xi \left(\int_0^x l(u)\, du \right) dx \\ &\quad = \frac{\xi^2}{2} \int_0^1 \int_0^x \{ l(u) - l(\xi u) \}\, du\, dx \geq 0 \end{aligned}$$

for all ξ in I. Furthermore, if $g(x) \not\equiv x$ on I, then there exists $\bar{\xi} \in (0, 1)$ such that

$$\int_0^1 \int_0^x \{ l(u) - l(\xi u) \}\, du\, dx > 0, \qquad \xi \in \left[0, \bar{\xi}\right],$$

and this yields (6.6).

Problem B can be solved thanks to the consistent estimator $\hat{S}^2_{g,n}$ of the asymptotic variance of $\gamma_{g,n}$, given in Section 5. In fact, in view of Theorems 5.1 and 5.2,

$$\gamma_{g,n} \pm 2 z_{\alpha/2} \hat{S}_{g,n} / \sqrt{n}$$

represent the extreme points of an asymptotic $(1 - \alpha)$ confidence interval for $\gamma_g(H)$, $z_{\alpha/2}$ being the $\alpha/2$th quantile of the standard normal c.d.f.

The same interval can be used to test hypotheses on the value of $\gamma_g(H)$ too. For instance, let us consider hypothesis H_0': $\gamma_g = \gamma^*$ versus H_1': $\gamma_g(H) \neq \gamma^*$,

γ^* being a fixed element in $(-1, 1)$. The test which rejects H_0' if and only if $|\gamma_{g,n} - \gamma^*| > 2|z_{\alpha/2}\hat{S}_{g,n}|/\sqrt{n}$, represents an asymptotic test of size α, for H_0' versus H_1'.

Acknowledgments. The authors are grateful to the referees for many valuable comments. A referee suggested study of the properties of $\sup_g \gamma_{g,n}$ as a test statistic. We were not able to do this, but, in any case, it would be desirable to pursue this line of thought.

REFERENCES

CAMBANIS, S., SIMONS, G. and STOUT, W. (1976). Inequalities for $\mathscr{E}k(X,Y)$ when the marginals are fixed. *Z. Wahrsch. Verw. Gebiete* **36** 285–294.

CIFARELLI, D. M. and REGAZZINI, E. (1974). Ancora sull'indice di cograduazione del Gini. Technical Report Serie III, No. 5, Istituto di Matematica Finanziaria dell'Università, Torino.

CIFARELLI, D. M. and REGAZZINI, E. (1977). On a distribution-free test of independence based on Gini's rank association coefficient. In *Recent Developments in Statistics* (J. R. Barra, F. Brodeau, G. Romier, B. Van Cutsem, eds.) 375–385. North-Holland, Amsterdam.

CIFARELLI, D. M. and REGAZZINI, E. (1990). Some contributions to the theory of monotone dependence. Technical Report 90.17, CNR-IAMI, Milano.

CONTI, P. L. (1993). Una classe di misure di dipendenza monotona tra due variabili continue: teoria descrittiva e problemi inferenziali non parametrici. Ph.D. dissertation, Dipartimento di Statistica, Probabilità e Statistiche Applicate, Università La Sapienza, Roma.

GAENSSLER, P. and STUTE, W. (1979). Empirical processes: a survey of results for independent and identically distributed random variables. *Ann. Probab.* **7** 193–243.

GINI, C. (1914). Di una misura delle relazioni tra le graduatorie di due caratteri. Tipografia, Cecchini, Roma.

HARDY, G. H., LITTLEWOOD, J. E. and PÓLYA, G. (1929). Some simple inequalities satisfied by convex functions. *Messenger of Math.* **58** 145–152.

HILDEBRANDT, T. H. (1963). *Introduction to the Theory of Integration*. Academic Press, New York.

HOEFFDING, W. (1948). A class of statistics with asymptotically normal distribution. *Ann. Math. Statist.* **19** 293–325.

KIMELDORF, G. and SAMPSON, A. R. (1978). Monotone dependence. *Ann. Statist.* **6** 895–903.

KIMELDORF, G. and SAMPSON, A. R. (1987). Positive dependence orderings. *Ann. Inst. Statist. Math.* **39** 113–128.

NEUHAUS, G. (1971). On weak convergence of stochastic processes with multidimensional parameter. *Ann. Math. Statist.* **42** 1285–1295.

SALVEMINI, T. (1951). Sui vari indici di cograduazione. *Statistica* **11** 133–154.

SEN, P. K. (1960). On some convergence properties of U-statistics. *Calcutta Statist. Assoc. Bull.* **10** 1–18.

TCHEN, A. H. (1980). Inequalities for distributions with given marginals. *Ann. Probab.* **8** 814–827.

YANAGIMOTO, T. and OKAMOTO, M. (1969). Partial orderings of permutations and monotonicity of a rank correlation statistic. *Ann. Inst. Statist. Math.* **21** 489–506.

D. M. CIFARELLI
E. REGAZZINI
IMQ—UNIVERSITÀ L. BOCCONI
VIA SARFATTI, 25
20135 MILANO
ITALY

P. L. CONTI
DIPARTIMENTO DI STATISTICA,
PROBABILITÀ E STATISTICHE APPLICATE
PIAZZALE ALDO MORO, 5
00185 ROMA
ITALY