Journal of Statistical Planning and Inference 141 (2011) 924–936

Contents lists available at ScienceDirect

Journal of Statistical Planning and Inference

journal homepage: www.elsevier.com/locate/jspi

Robustness and monotonicity properties of generalized correlation coefficients

Vivian Yi-Ju Chen [a], Vernon M. Chinchilli [b,*], Donald St. P. Richards [c]

[a] Department of Statistics, Tamkang University, Tamsui, Taipei 251, Taiwan, ROC
[b] Department of Public Health Sciences, College of Medicine, The Pennsylvania State University, Hershey, PA 17033, United States
[c] Department of Statistics, The Pennsylvania State University, University Park, PA 16802, United States

ARTICLE INFO

Article history:
Received 21 December 2008
Received in revised form
20 January 2010
Accepted 25 August 2010
Available online 8 September 2010

Keywords:
Clinical trials
Correlation coefficient
Cross-over design
Influence function
Robustness
Strictly sign regular kernel

ABSTRACT

A new class of generalized correlation coefficients that contains the Pearson and Kendall statistics as special cases was defined by Chinchilli et al. (2005) and applied to the estimation of correlations coefficients within the context of 2×2 cross-over designs for clinical trials. In this paper, we determine the infinitesimal robustness and local stability properties of these generalized correlation coefficients by deriving their corresponding influence functions. For cases in which the population distribution is a bivariate normal or a mixture of bivariate normal distributions we obtain explicit formulas, and establish monotonicity and sign-reverse rule properties of the generalized correlation coefficients.

1. Introduction

In the area of clinical trials, cross-over designs, *viz.*, repeated measures designs in which the study is divided into different periods and each subject receives a different treatment during each period, are often used to compare the efficacy of different treatments. We refer to Jones and Kenward (1998) or Senn (2002) for comprehensive treatments of cross-over designs in clinical trials research. Although many classical analyses of clinical trials data have focused on comparisons of the treatments while controlling for nuisance effects such as period effects, sequence effects, and carryover effects, the correlation between responses from different treatment periods is also of interest to researchers. In this regard, because the nuisance effects can be aliased with direct treatment effects in cross-over designs, the estimation of correlations without accounting for these nuisance effects may result in an inappropriate statistical analysis. To address this difficulty, Chinchilli et al. (2005) introduced a new class of generalized correlation coefficients that contains the Pearson and Kendall statistics as special cases, and developed a method for estimating those coefficients within the context of cross-over designs.

Let (X,Y) be a pair of continuous random variables with joint distribution function $F_{X,Y}$ and marginal distribution functions F_X and F_Y, respectively. For $\gamma \in [0,1]$, let g_γ denote the function $g_\gamma(t) = |t|^\gamma \operatorname{sgn}(t)$, $t \in \mathbb{R}$. Chinchilli et al. (2005) defined the *generalized correlation coefficient* (GCC) between X and Y as

$$\rho_\gamma = \frac{E[g_\gamma(X_1-X_2)g_\gamma(Y_1-Y_2)]}{\sqrt{E[g_\gamma^2(X_1-X_2)]E[g_\gamma^2(Y_1-Y_2)]}}, \tag{1.1}$$

* Corresponding author. Tel./fax: +1 717 531 4262.
E-mail address: vchinchi@hes.hmc.psu.edu (V.M. Chinchilli).

0378-3758/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.jspi.2010.08.016

where the pairs (X_1,Y_1) and (X_2,Y_2) are independent and identically distributed with joint cumulative distribution function $F_{X,Y}$. The GCC ρ_γ provides simultaneously a class of correlation coefficients as γ ranges over [0,1].

Chinchilli et al. (2005) showed that ρ_γ reduces to Kendall's correlation coefficient when $\gamma=0$ and to Pearson's coefficient when $\gamma=1$. Chinchilli et al. also applied the theory of U-statistics (Lee, 1990) to construct $\hat{\rho}_\gamma$, an estimator of ρ_γ, obtained the asymptotic distribution and other properties of $\hat{\rho}_\gamma$, and showed that $\hat{\rho}_\gamma$ can be modified for application to 2×2 cross-over designs in such a way that the correlation between treatment responses can be estimated without being affected by nuisance effects. In an investigation of the values of ρ_γ, Chinchilli et al. also observed among certain distribution functions $F_{X,Y}$ a tendency for the monotonicity property $|\hat{\rho}_\gamma| < |\hat{\rho}_\delta|$ to hold whenever $0 \le \gamma < \delta \le 1$.

In this paper, we derive explicit formulas for ρ_γ for cases in which the distribution of (X,Y) is bivariate normal or is a mixture of bivariate normal distributions. These results provide insight about conditions under which the above monotonicity property is valid and also about the robustness characteristics of ρ_γ. In particular, we investigate the local stability properties of ρ_γ by deriving the influence function for the estimator $\hat{\rho}_\gamma$, and we address the issue of assessing robustness through the computed influence function, particularly in cases in which the underlying population distribution is a bivariate normal or a mixture of bivariate normal distributions.

This paper is organized as follows. Section 2 summarizes our results on explicit formulas and monotonicity properties of the GCC under bivariate normality and mixture of bivariate normals. In Section 3, we provide results for the influence function for $\hat{\rho}_\gamma$. Following on those derivations, we consider in Section 4 numerical features of the GCCs and their robustness properties, Section 5 provides a discussion of the preceding results, and we also include two appendices with rigorous proofs of results in Sections 2 and 3.

2. Properties of the GCC under bivariate normality

Throughout this section, we assume that the pair (X,Y) has a bivariate normal distribution. We have noted in Section 1 that Chinchilli et al. (2005) observed a tendency for $|\rho_\gamma|$ to increase as γ increases. To determine the complete nature of that tendency, we now obtain an explicit formula for ρ_γ for the case in which (X,Y) is normally distributed.

2.1. Explicit formulas for $\boldsymbol{\rho}_\gamma$

Let (X_1,Y_1) and (X_2,Y_2) be independent and identically distributed with bivariate normal distributions $\mathcal{N}_2(\boldsymbol{\mu},\boldsymbol{\Sigma})$ with mean vector and covariance matrix,

$$\boldsymbol{\mu} = \begin{pmatrix} \mu_X \\ \mu_Y \end{pmatrix}, \quad \boldsymbol{\Sigma} = \begin{pmatrix} \sigma_X^2 & \rho\sigma_X\sigma_Y \\ \rho\sigma_X\sigma_Y & \sigma_Y^2 \end{pmatrix}, \tag{2.1}$$

respectively, where μ_X, μ_Y, σ_X^2, and σ_Y^2 are, respectively, the means and marginal variances of X and Y, and ρ is the Pearson coefficient of correlation between X and Y. Also, we use the standard notation ${}_2F_1(a,b;c;x)$ for the classical Gaussian hypergeometric series (Andrews et al., 2000, p. 64). We shall establish in Appendix A the following explicit formulas for ρ_γ.

Property 2.1. *Under bivariate normality*,

$$\rho_\gamma = 2\pi^{-1/2}\frac{\left[\Gamma\left(\frac{1}{2}\gamma+1\right)\right]^2}{\Gamma\left(\gamma+\frac{1}{2}\right)}\rho(1-\rho^2)^{\gamma+1/2}{}_2F_1\left(\frac{1}{2}\gamma+1,\frac{1}{2}\gamma+1;\frac{3}{2};\rho^2\right) \tag{2.2}$$

$$= 2\pi^{-1/2}\frac{\left[\Gamma\left(\frac{1}{2}\gamma+1\right)\right]^2}{\Gamma\left(\gamma+\frac{1}{2}\right)}\rho\,{}_2F_1\left(\frac{1}{2}(1-\gamma),\frac{1}{2}(1-\gamma);3/2;\rho^2\right). \tag{2.3}$$

It is interesting that the Gaussian hypergeometric series arise in the context of the generalized correlation coefficients. It has been known, since Fisher (1915) and Hotelling (1953), that there is a connection between the Gaussian hypergeometric series and the distributions of sample correlation coefficients; however, we are unaware of any results in which any correlation coefficients themselves are expressible in terms of the Gaussian hypergeometric series.

As partial verification of (2.2), we set $\gamma=0$ and apply the well-known result,

$${}_2F_1(1,1;3/2;\rho^2) = \frac{\sin^{-1}\rho}{\rho\sqrt{1-\rho^2}},$$

(Andrews et al., 2000, p. 64), to obtain

$$\rho_0 = 2\pi^{-1}\sin^{-1}\rho. \tag{2.4}$$

It is well known that ρ is related to the Kendall correlation coefficient, τ, by the equation $\rho = \sin(\pi\tau/2)$; cf. Joe (1997), p. 54, Exercise 2.14. Therefore, by (2.4), $\rho_0 \equiv \tau$. For the case in which $\gamma = 1$ we apply to (2.3) the elementary formula, ${}_2F_1(0,0;c;x) \equiv 1$, obtaining $\rho_1 \equiv \rho$.

2.2. Monotonicity and reverse rule properties

We have noted in Section 1 that Chinchilli et al. (2005) observed a tendency of $|\rho_\gamma|$ to increase monotonically as γ increases on [0,1]. That tendency was observed by examining the values of sample estimates, $\hat{\rho}_\gamma$, from illustrative examples including a case in which $\hat{\rho}_0 = 0.32 < \hat{\rho}_{0.5} = 0.51 < \hat{\rho}_1 = 0.61$; cf. Chinchilli et al. (2005), Table 1. To determine the exact nature of that tendency, we begin by obtaining an explicit formula for ρ_γ for the case in which (X,Y) is normally distributed or distributed as a mixture of bivariate normal distributions.

As a consequence of (2.3), we obtain the following result.

Property 2.2. *The GCC $|\rho_\gamma|$ is a strictly increasing function of $|\rho|$.*

To establish this result we observe that, by (2.3), it suffices to show that the hypergeometric series in (2.3) is strictly increasing in $|\rho|$. Since $0 \le \gamma \le 1$ then, by the definition (A.5) of the Gaussian hypergeometric series, all terms in the hypergeometric series in (2.3) are nonnegative. Moreover, since that series is a function of ρ^2 then the resulting expression is a strictly increasing function of $|\rho|$.

A more intricate property of the GCC under bivariate normality is the following.

Property 2.3. *For $\rho \ge 0$, the GCC ρ_γ is a strictly increasing function of γ.*

In Appendix A, we obtain this result by means of the theory of total positivity (Karlin, 1968), a theory which has figured prominently in studies of the positivity properties of correlation coefficients.

Let us change notation temporarily and denote the right-hand side of (2.2) or (2.3) by $K(\rho,\gamma)$. Then we shall prove the following result.

Property 2.4. *For $0 \le \rho_2 < \rho_1 \le 1$, the ratio $K(\rho_2,\gamma)/K(\rho_1,\gamma)$ is a strictly increasing function of $\gamma \in [0,1]$.*

In order to obtain Property 2.3, we now substitute $\rho_1 = 1$ in Property 2.4, and use an earlier observation that $K(1,\gamma) = \rho$, which does not depend on γ.

Table 1
Parameter values for cases 1–9.

Cases	$\gamma=0$	$\gamma=0.25$	$\gamma=0.50$	$\gamma=0.75$	$\gamma=1$
Values of $E[g_\gamma^2(X_1-X_2)]$					
1–3,5–9	1.0	0.97774	1.12838	1.44641	2.00
4	1.0	1.12900	0.150451	2.22689	3.56
Values of $E[g_\gamma^2(Y_1-Y_2)]$					
1–3,5–9	1.0	0.97774	1.12838	1.44641	2.00
4	1.0	1.19748	1.69257	2.65722	4.50
Values of $E[g_\gamma(X_1-X_2)g_\gamma(Y_1-Y_2)]$					
1	0.79784	0.87524	1.05342	1.36988	1.900
2	0.33333	0.40962	0.52896	0.71339	1.000
3	0.12819	0.16030	0.20956	0.28468	0.400
4	0.12819	0.19063	0.29637	0.47877	0.800
5	0.75513	0.83304	1.00642	1.31141	1.820
6	0.71805	0.78771	0.94808	1.23289	1.710
7	0.63827	0.70019	0.84274	1.09590	1.520
8	0.51859	0.56890	0.68472	0.89042	1.235
9	0.23935	0.26257	0.31603	0.41096	0.570
Values of ρ_γ					
1	0.79784	0.89516	0.93357	0.94709	0.9500
2	0.33333	0.41894	0.46878	0.49321	0.5000
3	0.12819	0.16395	0.18572	0.19682	0.2000
4	0.12819	0.16395	0.18572	0.19682	0.2000
5	0.75513	0.85201	0.89192	0.90667	0.9100
6	0.71805	0.80565	0.84021	0.85238	0.8550
7	0.63827	0.71613	0.74686	0.75767	0.7600
8	0.51859	0.58186	0.60682	0.61561	0.6175
9	0.23935	0.26855	0.28007	0.28413	0.2850

It is straightforward to see that the strictly increasing nature of the ratio in Property 2.4 is equivalent to the fact that the determinant

$$\begin{vmatrix} K(\rho_1,\gamma_1) & K(\rho_1,\gamma_2) \\ K(\rho_2,\gamma_1) & K(\rho_2,\gamma_2) \end{vmatrix} < 0$$

for $1>\rho_1>\rho_2>0$ and $1\geq\gamma_1>\gamma_2\geq 0$. In the terminology of the theory of total positivity (Karlin, 1968, p. 12), the negative sign of these determinants signifies that for $\rho\in(0,1)$ and $\gamma\in[0,1]$, the kernel $K(\rho,\gamma)$ is *strictly reverse rule of order* 2 (SRR_2).

Generalizing Property 2.4, we also shall prove in Appendix A the following result.

Property 2.5. *The kernel $K(\rho,\gamma)$ is strictly reverse rule of order infinity (SRR_∞), i.e., for any positive integer r, if $1>\rho_1>\cdots>\rho_r>0$ and $1\geq\gamma_1>\cdots>\gamma_r\geq 0$ then the sign of the determinant*

$$\begin{vmatrix} K(\rho_1,\gamma_1) & K(\rho_1,\gamma_2) & \cdots & K(\rho_1,\gamma_r) \\ K(\rho_2,\gamma_1) & K(\rho_2,\gamma_2) & \cdots & K(\rho_2,\gamma_r) \\ \vdots & \vdots & & \vdots \\ K(\rho_r,\gamma_1) & K(\rho_r,\gamma_2) & \cdots & K(\rho_r,\gamma_r) \end{vmatrix}$$

is $(-1)^{r(r-1)/2}$.

We remark that Wijsman (1959) long ago established a connection between the theory of total positivity and the sampling distribution of the Pearson correlation coefficient based on samples from a bivariate normal distribution; cf. Karlin (1968, p. 120).

2.3. Mixtures of bivariate normal distributions

Departures from multivariate normality are well known to arise naturally in applied work (Kowalski, 1972; Wilcox, 1993). In such instances, a classic approach is to apply mixtures of normal distributions as the underlying model when it is of interest to determine the effects of outlying observations (Tukey, 1960). Let $\varepsilon_1,\ldots,\varepsilon_m$ be constants representing the mixing proportions of component sub-populations, where $0\leq\varepsilon_i\leq 1$ and $\sum_{i=1}^m\varepsilon_i=1$. A mixture of m bivariate normal distributions is a random vector (X,Y) having distribution function

$$F_{X,Y}=\varepsilon_1F_1+\cdots+\varepsilon_mF_m,$$

where F_i is the joint cumulative distribution function of a bivariate normal distribution with mean vector $(\mu_{i,X},\mu_{i,Y})$, variances $(\sigma_{i,X}^2,\sigma_{i,Y}^2)$, correlation coefficient ρ_i, $i=1,\ldots,m$, and it is assumed that the corresponding sub-populations are mutually independent. For independent, identically distributed, bivariate random variables (X_1,Y_1) and (X_2,Y_2), having the same distribution as (X,Y), we apply (A.6) to obtain

$$E[g_\gamma(X_1-X_2)g_\gamma(Y_1-Y_2)]=\frac{2^{2\gamma+1}\left[\Gamma\left(\frac{1}{2}\gamma+1\right)\right]^2}{\pi}\sum_{i=1}^m\varepsilon_i\sigma_{i,X}^\gamma\sigma_{i,Y}^\gamma\rho_i(1-\rho_i^2)^{\gamma+1/2}{}_2F_1\left(\frac{1}{2}\gamma+1,\frac{1}{2}\gamma+1;\frac{3}{2};\rho_i^2\right). \tag{2.5}$$

By applying (A.7) we also find that

$$E[g_\gamma^2(X_1-X_2)]=\frac{\Gamma(2\gamma+1)}{\Gamma(\gamma+1)}\sum_{i=1}^m\varepsilon_i\sigma_{i,X}^{2\gamma}, \tag{2.6}$$

and

$$E[g_\gamma^2(Y_1-Y_2)]=\frac{\Gamma(2\gamma+1)}{\Gamma(\gamma+1)}\sum_{i=1}^m\varepsilon_i\sigma_{i,Y}^{2\gamma}. \tag{2.7}$$

Therefore, the GCC of (X,Y) is obtained in the form

$$\rho_\gamma=\frac{E[g_\gamma(X_1-X_2)g_\gamma(Y_1-Y_2)]}{\sqrt{E[g_\gamma^2(X_1-X_2)]E[g_\gamma^2(Y_1-Y_2)]}}=2\pi^{-1/2}\frac{\left[\Gamma\left(\frac{1}{2}\gamma+1\right)\right]^2}{\Gamma\left(\gamma+\frac{1}{2}\right)}\frac{\sum_{i=1}^m\varepsilon_i\sigma_{i,X}^\gamma\sigma_{i,Y}^\gamma\rho_i(1-\rho_i^2)^{\gamma+1/2}{}_2F_1\left(\frac{1}{2}\gamma+1,\frac{1}{2}\gamma+1;\frac{3}{2};\rho_i^2\right)}{\sqrt{\left(\sum_{i=1}^m\varepsilon_i\sigma_{i,X}^{2\gamma}\right)\left(\sum_{i=1}^m\varepsilon_i\sigma_{i,Y}^{2\gamma}\right)}}. \tag{2.8}$$

For the case in which each F_i is a standard bivariate normal distribution with $\sigma_{i,X}=\sigma_{i,Y}=1$, $i=1,\ldots,m$, (2.8) reduces to

$$\rho_\gamma=2\pi^{-1/2}\frac{\left[\Gamma\left(\frac{1}{2}\gamma+1\right)\right]^2}{\Gamma\left(\gamma+\frac{1}{2}\right)}\sum_{i=1}^m\varepsilon_i\rho_i(1-\rho_i^2)^{\gamma+1/2}{}_2F_1\left(\frac{1}{2}\gamma+1,\frac{1}{2}\gamma+1;\frac{3}{2};\rho_i^2\right). \tag{2.9}$$

Setting $\gamma=1$ in this latter formula, we obtain $\rho_{X,Y}\equiv\rho_\gamma|_{\gamma=1}=\sum_{i=1}^m \varepsilon_i\rho_i$, a result due originally to Kowalski (1972). If the underlying population follows a two-component mixture of bivariate normal distribution then, by (2.8),

$$\rho_\gamma|_{\gamma=1}=\frac{\varepsilon\sigma_{1X}^\gamma\sigma_{1Y}^\gamma\rho_1+(1-\varepsilon)\sigma_{2X}^\gamma\sigma_{2Y}^\gamma\rho_2}{\sqrt{(\varepsilon\sigma_{1,X}^{2\gamma}+(1-\varepsilon)\sigma_{2,X}^{2\gamma})(\varepsilon\sigma_{1,Y}^{2\gamma}+(1-\varepsilon)\sigma_{2,Y}^{2\gamma})}}. \tag{2.10}$$

For general values of γ, it is clear from the expressions in (2.5)–(2.10) that the population parameters ρ_γ, $E[g_\gamma(X_1-X_2)g_\gamma(Y_1-Y_2)]$, $E[g_\gamma^2(X_1-X_2)]$, and $E[g_\gamma^2(Y_1-Y_2)]$) are dependent on the parameters ε_i, ρ_i, $\sigma_{i,X}$, and $\sigma_{i,Y}$ corresponding to the component bivariate normal distributions in the mixture.

3. The influence function of the GCC

In research on robust statistical methods, the influence function originated by Hampel (1974) (cf. Hampel et al., 1986) is a basic tool for describing the infinitesimal stability of an estimator and quantifying the approximate effect of a single observation on the estimator. When the influence function is bounded, then the corresponding estimator is said to have *infinitesimal robustness*; in such instances, the estimator displays less sensitivity to variations in the characteristics of the observations, e.g., non-normality, small amounts of outliers or influential points, or badly placed observations.

Let $\boldsymbol{x}=((x_1,y_1),(x_2,y_2))$ denote a pair of bivariate points at the values (x_1,y_1) and (x_2,y_2). Also, let $F_{X,Y}$ be the distribution function of (X,Y), and (X_1,Y_1) and (X_2,Y_2) be a random sample from (X,Y). We shall show in Appendix B that the asymptotic influence function for $\hat\rho_\gamma$ is

$$IF(\boldsymbol{x};\hat\rho_\gamma,F_{X,Y})=\frac{1}{2}\rho_\gamma\left(\frac{2g_\gamma(x_1-x_2)g_\gamma(y_1-y_2)}{E[g_\gamma(X_1-X_2)g_\gamma(Y_1-Y_2)]}-\frac{[g_\gamma(x_1-x_2)]^2}{E[g_\gamma(X_1-X_2)]^2}-\frac{[g_\gamma(y_1-y_2)]^2}{E[g_\gamma(Y_1-Y_2)]^2}\right). \tag{3.1}$$

Let u=x_1-x_2, v=y_1-y_2, U=X_1-X_2 and V=Y_1-Y_2. Substituting

$$\rho_\gamma=\frac{E[g_\gamma(U)g_\gamma(V)]}{\sqrt{E[g_\gamma^2(U)]E[g_\gamma^2(V)]}},$$

we find that the influence function (3.1) can be reexpressed as

$$IF((u,v);\hat\rho_\gamma,F_{U,V})=\frac{g_\gamma(u)g_\gamma(v)}{\sqrt{E[g_\gamma(U)]^2E[g_\gamma(V)]^2}}-\frac{1}{2}\rho_\gamma\left(\frac{[g_\gamma(u)]^2}{E[g_\gamma(U)]^2}+\frac{[g_\gamma(v)]^2}{E[g_\gamma(V)]^2}\right). \tag{3.2}$$

For a given γ and distribution function $F_{X,Y}$, we can see from (3.2) that observations (x_1,y_1) and (x_2,y_2) that are far from each other tend to increase the influence function $IF(\boldsymbol{x};\hat\rho_\gamma,F_{X,Y})$. Also, the influence function is unbounded unless $\gamma=0$, the case of the Kendall correlation coefficient. For various values of γ, the Kendall correlation coefficient ($\gamma=0$) is the most robust, whereas the Pearson correlation coefficient ($\gamma=1$) is the least robust.

4. Robustness properties of the GCC

In this section, numerical investigations are conducted to demonstrate what we have observed in Section 2 and to explore the robustness behavior of the GCC using the influence function derived in Section 3. Denoting the bivariate normal distribution in (2.1) by $\mathcal{N}_2(\mu_X,\mu_Y;\sigma_X^2,\sigma_Y^2;\rho)$, we shall consider several bivariate population distributions $F_{X,Y}$ as follows:

Case 1: A normal distribution, $\mathcal{N}_2(0,0;1,1;0.95)$, representing examples in which the random variables are highly correlated.

Case 2: A normal distribution, $\mathcal{N}_2(-\sqrt{0.1}/2,\sqrt{0.1}/2;1,1;0.5)$, representing examples in which the random variables are moderately correlated.

Case 3: A normal distribution, $\mathcal{N}_2(-0.5,0.5;1,1;0.2)$, representing examples in which the random variables are weakly correlated.

Case 4: A normal distribution, $\mathcal{N}_2(\sqrt{0.25}/2,\sqrt{0.25}/2;(4/3)^2,(3/2)^2;0.2)$, representing examples in which two weakly correlated variables both have large variances.

Case 5: A mixture, $0.9\times\mathcal{N}_2(0,0;1,1;0.95)+0.1\times\mathcal{N}_2(0,0;1,1;0.55)$, in which 10% of the population is generated by a standard bivariate normal distribution with moderate correlation.

Case 6: A mixture, $0.9\times\mathcal{N}_2(0,0;1,1;0.95)+0.1\times\mathcal{N}_2(0,0;1,1;0)$, where the 10% of the population is through a standard bivariate normal distribution with uncorrelated variables.

Case 7: A mixture, $0.9\times\mathcal{N}_2(0,0;1,1;0.95)+0.1\times\mathcal{N}_2(0,0;1,1;-0.95)$, in which 10% of the population is generated by a standard bivariate normal distribution with highly negative correlated variables.

Case 8: A mixture, $0.65\times\mathcal{N}_2(0,0;1,1;0.95)+0.35\times\mathcal{N}_2(0,0;1,1;0)$, in which 35% of the population is drawn from a standard bivariate normal distribution with uncorrelated variables.

Case 9: A mixture, $0.3\times\mathcal{N}_2(0,0;1,1;0.95)+0.7\times\mathcal{N}_2(0,0;1,1;0)$, in which 70% of the population is through an uncorrelated bivariate standard normal distribution.

In each case, we tabulate the corresponding values of the four parameters $E[g_\gamma(Y_1-Y_2)]^2$, $E[g_\gamma(X_1-X_2)]^2$, $E[g_\gamma(X_1-X_2)g_\gamma(Y_1-Y_2)]$, and ρ_γ by means of the explicit formulas provided in Section 2.

4.1. Bivariate normal distributions

The parameter values for the distributions of Cases 1–4 are listed in Table 1. In Cases 1–3, we obtain identical values for $E[g_\gamma(X_1-X_2)]^2$ and $E[g_\gamma(Y_1-Y_2)]^2$ for given choices of γ, and we also find that the values of $E[g_\gamma(X_1-X_2)g_\gamma(Y_1-Y_2)]$ and ρ_γ decrease as ρ, the associated Pearson correlation coefficient, decreases. These results also reflect the fact that if two normal distributions have the same Pearson correlation coefficient, as in Cases 3–4, then the resulting values of ρ_γ are unchanged irrespective of the values of σ_X and σ_Y. By comparing Cases 3–4 we see that, an increase in σ_X^2, σ_Y^2, or γ tends to increase the values of the parameters $E[g_\gamma(X_1-X_2)]^2$, $E[g_\gamma(Y_1-Y_2)]^2$, and $E[g_\gamma(X_1-X_2)g_\gamma(Y_1-Y_2)]$.

The strict sign-regularity property derived in Property 2.5 is also exhibited by the values of the GCC for Cases 1–3. In Table 1, it is straightforward to verify that the 4×4 array of values of ρ_γ satisfies the SRR properties, e.g., the determinant of any 2×2 or 3×3 sub-matrix constructed from that array, with rows and columns indexed by ρ and γ in increasing order, has sign -1, as predicted by Property 2.5.

To illustrate further the behavior of the associated influence function of the GCCs for these various normal distributions, the graphs of $IF(\boldsymbol{x};T,F_{X,Y})$ formed by (3.2) are pictured in Fig. 1 for Cases 1–3 and for various values of γ. In these graphs, the horizontal and vertical axes list values of $u=x_1-x_2$ and $v=y_1-y_2$, respectively, ranging from -5 to 5. Cells having similar or identical color as their neighbors represent regions of bounded influence. We also consider pairs $\boldsymbol{x}=((x_1,y_1),(x_2,y_2))$ of points for which the difference vector $(u,v)\equiv(x_1-y_1,x_2-y_2)$, takes the values (3,3), $(-3,-3)$, $(3,-3)$, or $(-3,3)$. Such points will be viewed as pairs of "bad points" or "gross errors" for the underlying population, and we shall interpret the effect of an additional observation on the GCC ρ_γ. The values of the influence function at such pairs of observations for Cases 1–4 are provided in Table 2.

By Fig. 1 and as illustrated in Table 2, several findings are observed. First, Kendall's correlation coefficient (the case in which $\gamma=0$) is, as expected, the most robust one with good stability against small perturbations of the underlying distribution. Second, under the same distribution, the tendency toward bounded influence declines as γ increases (see Fig. 1). In other words, the larger the value of γ the greater the influence (see Table 2). Finally, the influence function is sign-symmetric, and for u and v of the same sign, the calculated influence function values are less pronounced than those for u and v of opposite sign. For brevity, we have omitted the graphs of $IF(\boldsymbol{x};T,F_{X,Y})$ for Case 4 because those graphs have shape similar to those of Case 3, the only difference being that they have a smaller range of influence function values due to the larger variances chosen for the underlying distribution.

On the other hand, comparisons in Cases 1–3 indicate that the values of the influence function of the GCC are inversely related to the Pearson correlation coefficient in the associated bivariate normal distribution. For highly correlated normal random variables with unit variances, the outcome is a smaller influence at an additional pair of bivariate points than is the case for a bivariate normal random vector with weakly correlated components. For example, the influence function values at a pair of bivariate observations with difference vector (u,v)=(3,3) for $\gamma=0.5$ are 0.17661, 1.41235, and 2.16491 for the bivariate normal distributions having Pearson correlation coefficients 0.95 (in Case 1), 0.5 (in Case 2), and 0.2 (in Case 3), respectively.

4.2. Contaminated bivariate normal distributions

Comparisons for the two-component mixture of bivariate normal distributions are performed in a similar fashion. True parameter values for Cases 5–9 are also given in Table 1.

By (2.6) and (2.7), and since all the population variances are equal to 1, it follows that $E[g_\gamma^2(X_1-X_2)]$ and $E[g_\gamma^2(Y_1-Y_2)]$ are equal for all such cases. It also can be seen that these parameters increase as γ increases. Cases 5–7 are instances in which $1-\varepsilon$, the mixing proportion, is fixed at 0.1; both ρ_γ and $E[g_\gamma(X_1-X_2)g_\gamma(Y_1-Y_2)]$ are dependent on ρ and ρ', the Pearson correlation coefficients corresponding to the standard bivariate normal distributions in the mixture. It can be seen also that the larger the value of $|\rho-\rho'|$, the smaller the values of the parameters. The combination $(\rho,\rho')=(0.95,-0.95)$ considered in Case 7 results in even smaller values of the parameters $E[g_\gamma^2(X_1-X_2)]$ and $E[g_\gamma^2(Y_1-Y_2)]$. In addition, by comparing Cases 8 and 9 with Case 6, we see that as the mixing proportion $1-\varepsilon$ increases for a given pair (ρ,ρ') in the mixture, the values of ρ_γ and $E[g_\gamma(X_1-X_2)g_\gamma(Y_1-Y_2)]$ decrease.

Table 2 also provides values of the influence function at a pair of bivariate observations for Cases 5–9. In comparing Cases 5–7, we see from Table 2 that small values of $|\rho-\rho'|$ at a fixed mixing proportion result in small influence function values. For Cases 6, 8 and 9, the values of the influence function increase with mixing proportion $1-\varepsilon$ for cases in which ρ and ρ' are fixed. These implies that influence function tends to be unbounded (note that for brevity, the graphs of $IF(\boldsymbol{x};T,F_{X,Y})$ for these case comparisons are not provided here) with increased mixing proportion $1-\varepsilon$ for certain mixtures of two standard bivariate normal distributions (i.e., ρ, ρ' fixed), or with $|\rho-\rho'|$ large and where the mixing proportion, 10%, is held fixed.

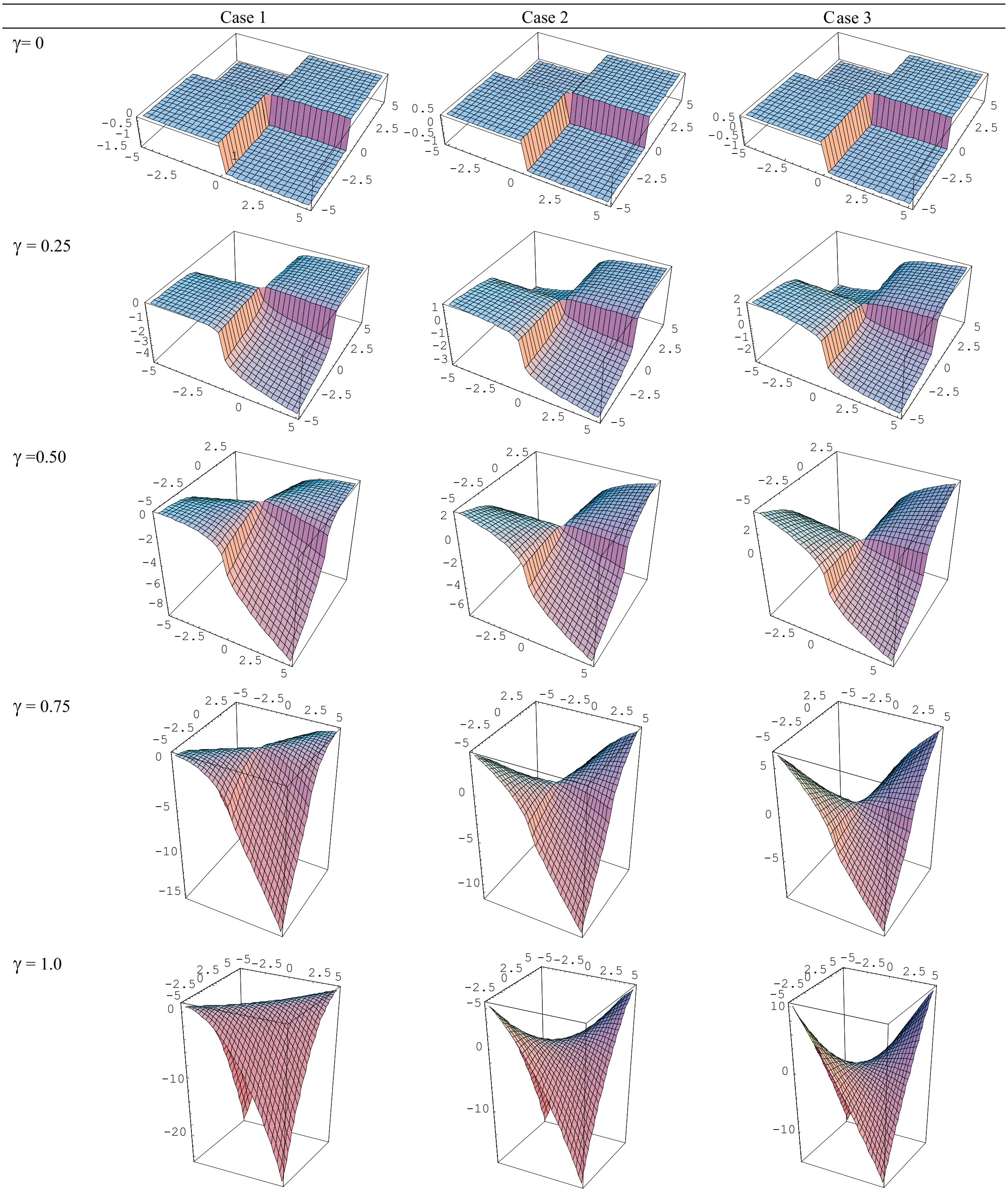

Fig. 1. Three-dimensional graphs of the influence functions for Cases 1–3.

5. Concluding remarks

In this article, we have explored in a twofold way the robustness properties of the GCCs introduced by Chinchilli et al. (2005). We derived explicit formulas for the GCC for bivariate normal and contaminated normal distributions, and established monotonicity and sign-regularity properties of the GCC.

Table 2
Values of $IF(\boldsymbol{x};T;F_{X,Y})$ for Cases 1–9 at given pairs of bivariate points.

Cases	$\gamma=0$	$\gamma=0.25$	$\gamma=0.50$	$\gamma=0.75$	$\gamma=1$
$(x_1-x_2,y_1-y_2)=\pm(3,3)$					
1	0.20217	0.18572	0.17661	0.19009	0.22500
2	0.66667	1.02933	1.41235	1.82061	2.25000
3	0.87181	1.48104	2.16491	2.88539	3.60000
4	0.87181	1.24530	1.53022	1.71402	1.79688
5	0.24487	0.26216	0.28736	0.33529	0.40500
6	0.28195	0.34430	0.42482	0.53032	0.65250
7	0.36173	0.50287	0.67303	0.87056	1.08000
8	0.48141	0.74074	1.04534	1.38091	1.72125
9	0.76065	1.29575	1.91406	2.57174	3.21750
$(x_1-x_2,y_1-y_2)=\pm(-3,3)$					
1	−1.79783	−3.35724	−5.14075	−6.99481	−8.77500
2	−1.33333	−2.51363	−3.90501	−5.36429	−6.75000
3	−1.12819	−2.06912	−3.15245	−4.29951	−5.40000
4	−1.12819	−1.73397	−2.29730	−2.55815	−2.70313
5	−1.75513	−3.28080	−5.03000	−6.84960	−8.59500
6	−1.71805	−3.19867	−4.89254	−6.65458	−8.34750
7	−1.63827	−3.04009	−4.64434	−6.31434	−7.92000
8	−1.51859	−2.80223	−4.27202	−5.80399	−7.27875
9	−1.23935	−2.24721	−3.40330	−4.61316	−5.78250

We have also obtained the influence function for assessing robustness of the GCC, thereby evaluating the robustness when the underlying distribution is a normal or mixture of normal distributions. Except for the case in which $\gamma=0$, the influence function for the GCC is unbounded; this justifies a phenomenon observed earlier in empirical studies, as well as in Chinchilli et al. (2005), that the Kendall correlation coefficient is robust and is a good alternative to the Pearson correlation coefficient when bivariate normality is suspect.

We have noted in Section 3 that if the distance between observations (x_1,y_1) and (x_2,y_2) is large then the influence function for the estimator $\hat{\rho}_\gamma$ is commensurately large, and we have illustrated this phenomenon with nine examples of normal or contaminated normal distributions.

Kocherlakota and Kocherlakota (1981) showed that the non-normality of a mixture of two standard bivariate normal distributions, similar to those considered in Cases 5–7 of Section 4, can be demonstrated by an examination of the coefficient of kurtosis. Kocherlakota and Kocherlakota noted that marked departures from bivariate normality arise for large values of $|\rho-\rho'|$, and especially so for cases in which ρ and ρ' are of opposite signs. Thus, when the parent population is a mixture of two standard bivariate normal distributions, the results of Section 4 suggest an intuitive relationship between the robustness of the GCC, as assessed by the influence function, and departures from normality.

The inverse hyperbolic tangent transformation (or Fisher's Z-transformation) ordinarily applied to sample correlation statistics has been also described in Chinchilli et al. (2005) for $\hat{\rho}_\gamma$. In this paper, we do not, however, investigate the robustness properties based on the transformed scale. As the $\hat{\rho}_\gamma$ is bounded between -1 and $+1$, an unbounded influence function will result in a large relative, instead of absolute, change in the correlation coefficient when a data point is perturbed. Therefore, it would be unnecessary to develop the influence function for the Fisher-transformed correlation coefficient, when the relative change is of higher concern than the absolute change.

In data analysis, the choice of a specific GCC, as defined by a particular choice of γ, must be determined by an investigator based on their experience and familiarity with the underlying data. In practice, as the retention of a potential outlier becomes more costly or precarious, bearing in mind that the GCC ρ_γ behaves more robustly as γ decreases, we would recommend that an investigator apply correspondingly smaller values of γ. Conversely, if the data analysis is able to tolerate a higher possibility of a potential outlier then an investigator may find it appropriate to use higher values of γ in their choice of a GCC. The suggestion above, however, is more reasonable for the data that do not exhibit extreme departures from a bivariate normal distribution.

It is important to realize that the GCC is not invariant under the class of monotonic transformations (e.g., exponential transformation). The ρ_γ calculated on the original data is not equal to that calculated on the transformed data for $0<\gamma\leq 1$. Unlike Kendall's correlation coefficient and Spearman's correlation coefficient, the GCC is not a measure of monotone association. To construct the invariant GCC, one can replace X by F_X and Y by F_Y in Eq. (1.1). With F_X and F_Y being uniformly distributed between zero and one, the sample estimate then would be based on the ranks of the X's and the Y's. This new class of GCC yields a family of monotone association measures which interpolates between Kendall's and Spearman's correlation coefficients. While we currently do not include the Spearman rank correlation coefficient as a case, the invariant GCC may provide an alternative particularly for nonparametric correlation analysis.

The relationship between Kendall's and Spearman's correlation coefficients has attracted attention in recent years (Capéraá and Genest, 1993; Chen, 2006; Li and Li, 2007; Fredricks and Nelsen, 2007). It has been suggested that Spearman's correlation coefficient is, under certain regularity conditions, larger than Kendall's correlation coefficient. This relationship, therefore, may still hold for the rank or invariant GCC due to its strictly increasing property (Property 2.4), and further investigation will be required to settle this issue.

Chinchilli et al. (2005) have also shown how the GCC may be adapted to estimate correlation coefficients within the context of a 2×2 cross-over design. It is believed that the GCC can be applied also to more general types of two-treatment cross-over designs that include at least two sequences or treatment periods. Such generalizations are expected to involve corresponding asymptotic theory and associated robustness properties more complicated than in the present paper, and those possibilities will be described in forthcoming work.

Appendix A. Proofs for Section 2

By changing location and scale, it follows from the definition (1.1) of ρ_γ that, without loss of generality, we may assume $\mu_X = \mu_Y = 0$ and $\sigma_X^2 = \sigma_Y^2 = 1$. Now define the random variables $U = 2^{-1/2}(X_1 - X_2)$ and $V = 2^{-1/2}(Y_1 - Y_2)$; then U and V have the joint density function

$$f(u,v) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2(1-\rho^2)}(u^2 - 2\rho uv + v^2)\right), \tag{A.1}$$

$u,v \in \mathbb{R}$, and this density satisfies the symmetry property, $f(u,-v) = f(-u,v)$ for all u, v. By the definition of U and V,

$$Eg_\gamma(X_1 - X_2)g_\gamma(Y_1 - Y_2) = 2^\gamma E|UV|^\gamma \operatorname{sgn}(UV) = 2^\gamma \int\int_{\mathbb{R}^2} |uv|^\gamma \operatorname{sgn}(uv) f(u,v)\, du\, dv.$$

By decomposing $\mathbb{R}^2$ into the four quadrants $\{(\pm u, \pm v) : u \ge 0, v \ge 0\}$, and applying the above symmetry property of $f(u,v)$, we obtain

$$E|UV|^\gamma \operatorname{sgn}(UV) = 2\int_0^\infty \int_0^\infty (uv)^\gamma [f(u,v) - f(u,-v)]\, du\, dv. \tag{A.2}$$

Applying (A.1), we find that (A.2) equals

$$\begin{aligned} &\frac{2^\gamma}{\pi\sqrt{1-\rho^2}} \int_0^\infty \int_0^\infty (uv)^\gamma \exp\left(-\frac{u^2+v^2}{2(1-\rho^2)}\right) \sinh\left(\frac{\rho}{1-\rho^2} uv\right) du\, dv \\ &= \frac{2^\gamma}{\pi}(1-\rho^2)^{(2\gamma+1)/2} \int_0^\infty \int_0^\infty (uv)^\gamma \exp[-(u^2+v^2)/2] \sinh(\rho uv)\, du\, dv, \end{aligned} \tag{A.3}$$

where the latter expression is obtained by replacing (u,v) by $(1-\rho^2)^{1/2}(u,v)$. On applying the Taylor–Maclaurin series expansion,

$$\sinh(\rho uv) = \sum_{k=0}^\infty \frac{(\rho uv)^{2k+1}}{(2k+1)!}$$

to (A.3), verifying *via* Fubini's theorem that the interchange of integrals and summation is justified, and evaluating each integral using the formula:

$$\int_0^\infty t^{2\alpha-1} \exp(-t^2/2)\, dt = 2^{\alpha-1}\Gamma(\alpha),$$

$\alpha > 0$, we obtain

$$\begin{aligned} E|UV|^\gamma \operatorname{sgn}(UV) &= \frac{2(1-\rho^2)^{\gamma+1/2}}{\pi} \sum_{k=0}^\infty \frac{\rho^{2k+1}}{(2k+1)!} \int_0^\infty \int_0^\infty (uv)^{2k+\gamma+1} e^{-(u^2+v^2)/2}\, du\, dv \\ &= \frac{2^{\gamma+1}}{\pi}\rho(1-\rho^2)^{\gamma+1/2} \sum_{k=0}^\infty \frac{\rho^{2k}}{(2k+1)!}\left[\Gamma\left(\frac{1}{2}\gamma + k + 1\right)\right]^2 2^{2k}. \end{aligned}$$

Recall the notation for the *rising factorial*,

$$(a)_k = \frac{\Gamma(a+k)}{\Gamma(a)} = a(a+1)(a+2)\cdots(a+k-1), \tag{A.4}$$

$k = 0,1,2,\ldots$; then it is straightforward to verify that

$$(2k+1)! = k!\,2^{2k}(3/2)_k,$$

and that

$$\Gamma(\tfrac{1}{2}\gamma + k + 1) = (\tfrac{1}{2}\gamma + 1)_k \Gamma(\tfrac{1}{2}\gamma + 1).$$

Therefore,

$$E|UV|^{\gamma}\mathrm{sgn}(UV)=\frac{2^{\gamma+1}\left[\Gamma\left(\frac{1}{2}\gamma+1\right)\right]^2}{\pi}\rho(1-\rho^2)^{\gamma+1/2}\sum_{k=0}^{\infty}\frac{\left(\frac{1}{2}\gamma+1\right)_k\left(\frac{1}{2}\gamma+1\right)_k}{(3/2)_k k!}\rho^{2k}$$
$$=\frac{2^{\gamma+1}\left[\Gamma\left(\frac{1}{2}\gamma+1\right)\right]^2}{\pi}\rho(1-\rho^2)^{\gamma+\frac{1}{2}}\cdot {}_2\mathrm{F}_1\left(\frac{1}{2}\gamma+1,\frac{1}{2}\gamma+1;3/2;\rho^2\right),$$

where ${}_2\mathrm{F}_1$ is the Gauss hypergeometric function,

$$ {}_2\mathrm{F}_1(a,b;c;x)=\sum_{k=0}^{\infty}\frac{(a)_k(b)_k}{(c)_k}\frac{x^k}{k!}, \tag{A.5}$$

$|x|<1$. Consequently,

$$E[g_{\gamma}(X_1-X_2)g_{\gamma}(Y_1-Y_2)]=2^{\gamma}E[|UV|^{\gamma}\mathrm{sgn}(UV)]=\frac{2^{2\gamma+1}\left[\Gamma\left(\frac{\gamma}{2}+1\right)\right]^2}{\pi}\rho(1-\rho^2)^{\gamma+1/2}\cdot {}_2\mathrm{F}_1\left(\frac{\gamma}{2}+1,\frac{\gamma}{2}+1;\frac{3}{2};\rho^2\right). \tag{A.6}$$

By a similar argument, we obtain

$$E[g_{\gamma}^2(X_1-X_2)]=E[g_{\gamma}^2(Y_1-Y_2)]=\frac{\Gamma(2\gamma+1)}{\Gamma(\gamma+1)}. \tag{A.7}$$

Substituting (A.6) and (A.7) into (1.1), and applying Legendre's duplication formula for the gamma function (Andrews et al., 2000, p. 22) to simplify the constant term, we obtain the explicit formula (2.2).

Finally, (2.3) is obtained from (2.2) by applying the Euler identity,

$$ {}_2\mathrm{F}_1(a,b;c,x)=(1-x)^{c-a-b}{}_2\mathrm{F}_1(c-a,c-b;c,x),$$

to the hypergeometric series in (2.2).

Proof of Property 2.5. Let $\delta_i=1/2(1-\gamma_i)$, $i=1,\dots,r$. Then, we need to determine the sign of the $r\times r$ determinant

$$\det(K(\rho_i,\gamma_j))=\left[\prod_{j=1}^{r}2\pi^{-1/2}\frac{\left[\Gamma\left(\frac{1}{2}\gamma_j+1\right)\right]^2}{\Gamma\left(\gamma_j+\frac{1}{2}\right)}\rho_j\right]\det({}_2\mathrm{F}_1(\delta_i,\delta_i;3/2;\rho_j^2)) \tag{A.8}$$

for $1>\rho_1>\cdots>\rho_r>0$ and $1\geq\gamma_1>\cdots>\gamma_r\geq 0$. Define the functions $h_i(k)=[(\delta_i)_k]^2/k!(3/2)_k$ and $g_i(k)=\rho_i^{2k}$, k=0,1,2,..., i=1,...,r. By the Binet–Cauchy formula (Karlin, 1968, p. 1),

$$\det({}_2\mathrm{F}_1(\delta_i,\delta_i;3/2;\rho_j^2))\equiv\det\left(\sum_{k=0}^{\infty}h_i(k)g_j(k)\right)=\sum_{k_1>\cdots>k_r\geq 0}\det(h_i(k_j))\det(g_i(k_j)). \tag{A.9}$$

To determine the sign of the $r\times r$ determinant red $\det(h_i(k_j))$, it suffices to determine the sign of $\det([(\delta_i)_{k_j}]^2)$. Define a measure $\overline{v}$ on $(0,\infty)$ by $d\overline{v}(t)=e^{-t}t^{-1}\,dt$; then, for $\delta>0$ and $k\geq 0$, it is elementary that

$$[(\delta)_k]^2=\frac{1}{[\Gamma(\delta)]^2}\int_0^{\infty}\int_0^{\infty}(t_1t_2)^{\delta+k}\,d\overline{v}(t_1)\,d\overline{v}(t_2). \tag{A.10}$$

We now define a measure v on $(0,\infty)$ by convolving the measure $\overline{v}$ multiplicatively, *viz.*,

$$dv(t)=\int_{t_1t_2=t}d\overline{v}(t_1)\,d\overline{v}(t_2).$$

The measure v is sigma-finite and, since the double integral (A.10) converges absolutely, we may apply Fubini's theorem to deduce that

$$[(\delta)_k]^2=\frac{1}{[\Gamma(\delta)]^2}\int_0^{\infty}t^{\delta+k}\,dv(t).$$

Therefore

$$\det([(\delta_i)_{k_j}]^2)=\det\left(\frac{1}{[\Gamma(\delta_i)]^2}\int_0^{\infty}t^{\delta_i+k_j}\,dv(t)\right)=\left[\prod_{j=1}^{r}\frac{1}{[\Gamma(\delta_i)]^2}\right]\det\left(\int_0^{\infty}t^{\delta_i}t^{k_j}\,dv(t)\right). \tag{A.11}$$

Applying the continuous form of the Binet–Cauchy formula (Karlin, 1968, pp. 16–17) to the latter determinant, we have

$$\det\left(\int_0^\infty t^{\delta_i} t^{k_j}\, d\nu(t)\right) = \underset{\infty > t_1 > \cdots > t_r > 0}{\int \cdots \int} \det(t_i^{\delta_j})\det(t_i^{k_j}) \prod_{j=1}^{r} d\nu(t_j).$$

Each determinant in the above integrand is a well-known generalized Vandermonde determinant. It is well known that $\det(t_i^{k_j}) > 0$ whenever $t_1 > \cdots > t_r$ and $k_1 > \cdots > k_r$. As for $\det(t_i^{\delta_j})$, since $\delta_1 < \cdots < \delta_r$, then this determinant would be positive if we were to reverse the order of all its rows, a process that would require $r(r-1)/2$ row interchanges. Hence, the sign of $\det(t_i^{\delta_j})$ is $(-1)^{r(r-1)/2}$, and the same holds for $\det(\phi_i(k_j))$.

Finally, returning to (A.9), the determinant $\det(g_i(k_j)) \equiv \det(\rho_i^{k_j})$ also is a generalized Vandermonde determinant, hence is positive for $\rho_1 > \cdots > \rho_r$ and $k_1 > \cdots > k_r \geq 0$; see Karlin (1968, p. 15). Therefore, the sign of the determinant (A.8) is $(-1)^{r(r-1)/2}$. □

We remark that for small values of k, the determinant (A.11) can be calculated explicitly and then its sign is seen directly to be positive. For instance, for the case in which k=2, we have

$$\begin{vmatrix} [(\delta_1)_{k_1}]^2 & [(\delta_1)_{k_2}]^2 \\ [(\delta_2)_{k_1}]^2 & [(\delta_2)_{k_2}]^2 \end{vmatrix} = [(\delta_1)_{k_2}]^2[(\delta_2)_{k_2}]^2\left(\frac{[(\delta_1)_{k_1}]^2}{[(\delta_1)_{k_2}]^2} - \frac{[(\delta_2)_{k_1}]^2}{[(\delta_2)_{k_2}]^2}\right). \tag{A.12}$$

For $k_1 > k_2$,

$$\frac{(\delta)_{k_1}}{(\delta)_{k_2}} = (\delta+k_2)\cdots(\delta+k_1-1) \equiv (\delta+k_2)_{k_1-k_2},$$

hence the last term in (A.12) equals $[(\delta_1+k_2)_{k_1-k_2}]^2 - [(\delta_2+k_2)_{k_1-k_2}]^2$, which is seen easily to be positive for $\delta_1 > \delta_2$ and $k_1 > k_2$; therefore, the determinant (A.12) is positive under the same conditions.

Appendix B. Proofs for Section 3

Let (X_1,Y_1) and (X_2,Y_2) be independent and identically distributed according to the distribution function $F_{X,Y}$. Define

$$\phi_1((X_1,Y_1),(X_2,Y_2)) = [g_\gamma(X_1-X_2)]^2,$$

$$\phi_2((X_1,Y_1),(X_2,Y_2)) = [g_\gamma(Y_1-Y_2)]^2,$$

$$\phi_3((X_1,Y_1),(X_2,Y_2)) = g_\gamma(X_1-X_2)g_\gamma(Y_1-Y_2),$$

and let

$$\Phi((X_1,Y_1),(X_2,Y_2)) = \begin{pmatrix} \phi_1((X_1,Y_1),(X_2,Y_2)) \\ \phi_2((X_1,Y_1),(X_2,Y_2)) \\ \phi_3((X_1,Y_1),(X_2,Y_2)) \end{pmatrix}. \tag{B.1}$$

Also, define $\psi_{\gamma,j} = E[\phi_j((X_1,Y_1),(X_2,Y_2))]$, j=1,2,3, and set

$$\psi_\gamma = \begin{pmatrix} \psi_{\gamma,1} \\ \psi_{\gamma,2} \\ \psi_{\gamma,3} \end{pmatrix} \equiv E\Phi((X_1,Y_1),(X_2,Y_2)), \quad A_\gamma = \frac{1}{2}\rho_\gamma \begin{pmatrix} -1/\psi_{\gamma,1} \\ -1/\psi_{\gamma,2} \\ 2/\psi_{\gamma,3} \end{pmatrix}. \tag{B.2}$$

For random samples (X_{j1}, Y_{j1}) and (X_{j2}, Y_{j2}), $j=1,\ldots,n$, define

$$U_\gamma \equiv \begin{pmatrix} U_{\gamma,XX} \\ U_{\gamma,YY} \\ U_{\gamma,XY} \end{pmatrix} = \frac{2}{n(n-1)} \sum_{j_1 < j_2} \Phi((X_{j_1},Y_{j_1}),(X_{j_2},Y_{j_2})), \tag{B.3}$$

a vector of U-statistics. Let $\Omega_{n,\gamma}$ be the matrix defined by Chinchilli et al. (2005, Eq. (A.25)); and set $\sigma^2_{n,\gamma} = A'_\gamma \Omega_{n,\gamma} A_\gamma$; then Chinchilli et al. (2005) have shown that the asymptotic distribution of $\hat{\rho}_\gamma = U_{\gamma,XY}/\sqrt{U_{\gamma,XX}U_{\gamma,YY}}$ is given by

$$(\hat{\rho}_\gamma - \rho_\gamma)/\sigma_{n,\gamma} \xrightarrow{L} \mathcal{N}(0,1), \tag{B.4}$$

as $n \to \infty$. Further, by Chinchilli et al. (2005), the joint distribution of the vector U_γ is asymptotically trivariate normal, and $A_\gamma'(U_\gamma - \psi_\gamma)$ is asymptotically normal with asymptotic variance $A_\gamma' \Omega_{n,\gamma} A_\gamma \equiv \sigma^2_{n,\gamma}$, i.e., as $n \to \infty$,

$$A_\gamma'(U_\gamma - \psi_\gamma)/\sigma_{n,\gamma} \xrightarrow{L} \mathcal{N}(0,1). \tag{B.5}$$

By (B.2), we have $A_\gamma{}' \psi_\gamma = 0$. Therefore, it follows from (B.4) and (B.5) that $\hat{\rho}_\gamma - \rho_\gamma$ and $A_\gamma{}' U_\gamma$ are asymptotically equivalent in probability, and therefore we can derive the asymptotic influence function for $\hat{\rho}_\gamma$ from that of $A_\gamma' U_\gamma$. On applying (B.3), we obtain

$$A_\gamma' U_\gamma = \frac{2}{n(n-1)} \sum_{j_1 < j_2} A_\gamma' \Phi_\gamma((X_{j_1}, Y_{j_1}), (X_{j_2}, Y_{j_2})),$$

which shows that $A_\gamma{}' U_\gamma$ is an unbiased estimator of the functional

$$T(F) = E_F[A_\gamma' \Phi_\gamma((X_1, Y_1), (X_2, Y_2))] = \int A_\gamma' \Phi_\gamma(\boldsymbol{w})\, dF(\boldsymbol{w}), \tag{B.6}$$

where, similar to the foregoing, if $\boldsymbol{w} = ((s_1, t_1), (s_2, t_2))$ is a pair of bivariate variables then we denote $dF_{X,Y}(s_1, t_1)\, dF_{X,Y}(s_2, t_2)$ by $dF(\boldsymbol{w})$.

Let $\delta_{\boldsymbol{x}}$ denote the point mass 1 at a pair of bivariate observations with observed value $\boldsymbol{x} = ((x_1, y_1), (x_2, y_2))$. For $\varepsilon \in (0,1)$ consider the distribution

$$F_{\varepsilon, \boldsymbol{x}}(\boldsymbol{w}) = (1-\varepsilon) F_{X,Y}(\boldsymbol{w}) + \varepsilon \delta_{\boldsymbol{x}}(\boldsymbol{w}) \tag{B.7}$$

representing the mixture distribution under which a pair of bivariate observations $\boldsymbol{w}$ is randomly sampled from the distribution $F_{X,Y}$ with probability $1-\varepsilon$, and otherwise the observed value is $\boldsymbol{x}$ with probability ε.

By definition, the influence function of the functional $T(F_{X,Y})$ in (B.6), is

$$IF(\boldsymbol{x}; T, F) = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} [T(F_{\varepsilon, \boldsymbol{x}}) - T(F)], \tag{B.8}$$

and it is this limit that we need to calculate. By (B.6) and (B.7) we have

$$\begin{aligned} T(F_{\varepsilon,\boldsymbol{x}}) - T(F) &= \int A_\gamma' \Phi_\gamma(\boldsymbol{w})\, d[(1-\varepsilon) F_{X,Y}(\boldsymbol{w}) + \varepsilon \delta_{\boldsymbol{x}}(\boldsymbol{w})] - \int A_\gamma' \Phi_\gamma(\boldsymbol{w})\, dF(\boldsymbol{w}) \\ &= (1-\varepsilon) \int A_\gamma' \Phi_\gamma(\boldsymbol{w})\, dF_{X,Y}(\boldsymbol{w}) + \varepsilon A_\gamma' \Phi_\gamma(\boldsymbol{x}) - \int A_\gamma' \Phi_\gamma(\boldsymbol{w})\, dF(\boldsymbol{w}) = \varepsilon \left[A_\gamma' \Phi_\gamma(\boldsymbol{x}) - \int A_\gamma' \Phi_\gamma(\boldsymbol{w})\, dF(\boldsymbol{w}) \right]. \end{aligned}$$

Substituting this result into (B.8), we obtain

$$IF(\boldsymbol{x}; T, F_{X,Y}) = A_\gamma' \Phi_\gamma(\boldsymbol{x}) - \int A_\gamma' \Phi_\gamma(\boldsymbol{w})\, dF(\boldsymbol{w}).$$

By (B.5), $\int A_\gamma' \Phi_\gamma(\boldsymbol{w})\, dF(\boldsymbol{w})$ is asymptotically close to 0; therefore (B.8) reduces to

$$IF(\boldsymbol{x}; T, F_{X,Y}) \simeq A_\gamma{}' \Phi_\gamma(\boldsymbol{x}). \tag{B.9}$$

By the asymptotic equivalence of $A_\gamma' U_\gamma$ and $\hat{\rho}_\gamma$, the asymptotic influence function for $\hat{\rho}_\gamma$ also is given by (B.9). On applying (B.1) and (B.2) to (B.9), we obtain

$$IF(\boldsymbol{x}; T, F_{X,Y}) \simeq A_\gamma{}' \Phi_\gamma(\boldsymbol{x}) = \frac{1}{2} \rho_\gamma \left(\frac{2 g_\gamma(x_1 - x_2) g_\gamma(y_1 - y_2)}{E[g_\gamma(X_1 - X_2) g_\gamma(Y_1 - Y_2)]} - \frac{g_\gamma^2(x_1 - x_2)}{E[g_\gamma^2(X_1 - X_2)]} - \frac{g_\gamma^2(y_1 - y_2)}{E[g_\gamma^2(Y_1 - Y_2)]} \right),$$

which establishes (3.1).

References

Andrews, G.E., Askey, R., Roy, R., 2000. Special Functions. Cambridge University Press, New York.

Capéraá, P., Genest, C., 1993. Spearman's ρ is larger than Kendall's τ for positively dependent random variables. Nonparametric Statistics 2, 183–194.

Chen, Y.-P., 2006. A note on the relationship between Spearman's ρ and Kendall's τ. Journal of Statistical Planning and Inference 137, 2165–2171.

Chinchilli, V.M., Phillips, B.R., Mauger, D.T., Szefler, S.J., 2005. A general class of correlation coefficients for the 2×2 cross-over design. Biometrical Journal 47, 644–653.

Fisher, R.A., 1915. Frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population. Biometrika 10, 507–521.

Fredricks, G.A., Nelsen, R.B., 2007. On the relationship between Spearman's ρ and Kendall's τ for pairs of continuous random variables. Journal of Statistical Planning and Inference 137, 2143–2150.

Hampel, F.R., 1974. The influence curve and its role in robust estimation. Journal of the American Statistical Association 69, 383–393.

Hampel, F.R., Ronchetti, E.M., Rousseeuw, P.J., Stahel, W.A., 1986. Robust Statistics. Wiley, New York.

Hotelling, H., 1953. New light on the correlation coefficient and its transforms with discussion. Journal of the Royal Statistical Society, Series B 15, 193–232.

Joe, H., 1997. Multivariate Models and Dependence Concepts. Chapman & Hall, London.

Jones, B., Kenward, M.G., 1998. Design and Analysis of Cross-Over Trials. Chapman & Hall, London.

Karlin, S., 1968. Total Positivity. Stanford University Press, Stanford, CA.

Kocherlakota, K., Kocherlakota, S., 1981. On the distribution of r in samples from the mixtures of bivariate normal populations. Communications in Statistics—Theory and Methods 10, 1943–1966.
Kowalski, C.J., 1972. On the effects of non-normality on the distribution of the sample product moment correlation coefficient. Applied Statistics 21, 1–12.
Lee, A.J., 1990. U-Statistics. Marcel Dekker, New York.
Li, X., Li, Z., 2007. Proof of a conjecture on Spearman's ρ and Kendall's τ for sample minimum and maximum. Journal of Statistical Planning and Inference 137, 359–361.
Senn, S., 2002. Cross-Over Trials in Clinical Research. Wiley, Chichester, UK.
Tukey, J.W., 1960. A survey of sampling from contaminated normal distributions. In: Olkin, I. (Ed.), Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling. Stanford University Press, Stanford, CA, pp. 448–485.
Wijsman, R.A., 1959. Applications of a certain representation of the Wishart matrix. Annals of Mathematical Statistics 30, 597–601.
Wilcox, R.R., 1993. Some results on a Winsorized correlation coefficient. British Journal of Mathematical and Statistical Psychology 46, 339–349.