Symmetric Gini covariance and correlation

Yongli Sang[1], Xin Dang[1] and Hailin Sang[1]*

[1]*Department of Mathematics, University of Mississippi, University, MS 38677, USA.*

Abstract: Standard Gini covariance and Gini correlation play important roles in measuring the dependence of random variables with heavy tails. However, the asymmetry brings a substantial difficulty in interpretation. In this paper, we propose a symmetric Gini-type covariance and a symmetric Gini correlation (ρ_g) based on the joint rank function. The proposed correlation ρ_g is more robust than the Pearson correlation but less robust than the Kendall's τ correlation. We establish the relationship between ρ_g and the linear correlation ρ for a class of random vectors in the family of elliptical distributions, which allows us to estimate ρ based on estimation of ρ_g. The asymptotic normality of the resulting estimators of ρ are studied through two approaches: one from influence function and the other from U-statistics and the delta method. We compare asymptotic efficiencies of linear correlation estimators based on the symmetric Gini, regular Gini, Pearson and Kendall's τ under various distributions. In addition to reasonably balancing between robustness and efficiency, the proposed measure ρ_g demonstrates superior finite sample performance, which makes it attractive in applications. *The Canadian Journal of Statistics* xx: 1–19; 2016 © 2016 Statistical Society of Canada

Résumé: Insérer votre résumé ici. We will supply a French abstract for those authors who can't prepare it themselves. *La revue canadienne de statistique* xx: 1–19; 2016 © 2016 Société statistique du Canada

1. INTRODUCTION

Let X and Y be two non-degenerate random variables with marginal distribution functions F and G, respectively, and a joint distribution function H. To describe dependence correlation between X and Y, the Pearson correlation (denoted as ρ_p) is probably the most frequently used measure. This measure is based on the covariance between two variables, which is optimal for the linear association between bivariate normal variables. However, the Pearson correlation performs poorly for variables with heavily-tailed or asymmetric distributions, and may be seriously impacted even by a single outlier (e.g., Shevlyakov and Smirnov, 2011). Under the assumption that F and G are continuous, the Spearman correlation, a robust alternative, is a multiple (twelve) of the covariance between the cumulative functions (or ranks) of two variables; the Gini correlation is based on the covariance between one variable and the cumulative distribution of the other (Blitz and Brittain, 1964). Two Gini correlations can be defined as

$$\gamma(X,Y) := \frac{\mathrm{cov}(X, G(Y))}{\mathrm{cov}(X, F(X))} \quad \text{and} \quad \gamma(Y,X) := \frac{\mathrm{cov}(Y, F(X))}{\mathrm{cov}(Y, G(Y))}$$

to reflect different roles of X and Y. The representation of Gini correlation $\gamma(X,Y)$ indicates that it has mixed properties of those of the Pearson and Spearman correlations. It is similar to Pearson in X (the variable taken in its variate values) and similar to Spearman in Y (the

---

* *Author to whom correspondence may be addressed.*
 E-mail: sang@olemiss.edu

variable taken in its ranks). Hence Gini correlations complement the Pearson and Spearman correlations (Schechtman and Yitzhaki, 1987; 1999; 2003). Two Gini correlations are equal if X and Y are exchangeable up to a linear transformation. However, Gini covariances are not symmetric in X and Y in general. On one hand, this asymmetrical nature is useful and can be used for testing bivariate exchangeability (Schechtman, Yitzhaki and Artsev, 2007). On the other hand, such asymmetry violates the axioms of correlation measurement (Mari and Kotz, 2001). Although some authors (e.g., Xu *et al.*, 2010) dealt with asymmetry by a simple average $(\gamma(X,Y) + \gamma(Y,X))/2$, it is difficult to interpret this measure, especially when $\gamma(X,Y)$ and $\gamma(Y,X)$ have different signs.

The asymmetry of $\gamma(X,Y)$ and $\gamma(Y,X)$ stems from the usage of marginal rank function $F(x)$ or $G(y)$. A remedy is to utilize a joint rank function. To do so, let us look at a representation of the Gini mean difference (GMD) under continuity assumption: $\Delta(F) = 4\text{cov}(X, F(X)) = 2\text{cov}(X, 2F(X) - 1)$ (Stuart, 1954; Lerman and Yitzhaki, 1984). The second equality rewrites GMD as twice of the covariance of X and the *centered rank* function $r(X) := 2F(X) - 1$. If F is continuous, $\mathbb{E}r(X) = 0$. Hence

$$\Delta(F) = 2\text{cov}(X, r(X)) = 2\mathbb{E}(Xr(X)). \tag{1}$$

The rank function $r(X)$ provides a center-orientated ordering with respect to the distribution F. Such a rank concept is of vital importance for high dimensions where the natural linear ordering on the real line no longer exists. A generalization of the centered rank in high dimension is called the spatial rank. Based on this joint rank function, we are able to propose a symmetric Gini covariance (denoted as cov_g) and a corresponding symmetric correlation (denoted as ρ_g). That is, $\text{cov}_g(X,Y) = \text{cov}_g(Y,X)$ and $\rho_g(X,Y) = \rho_g(Y,X)$.

We study properties of the proposed Gini correlation ρ_g. In terms of the influence function, ρ_g is more robust than the Pearson correlation ρ_p. However, ρ_g is not as robust as the Spearman correlation and Kendall's τ correlation. Kendall's τ is another commonly used nonparametric measure of association. The Kendall correlation measure is more robust and more efficient than the Spearman correlation (Croux and Dehon, 2010). For this reason, in this paper we do not consider Spearman correlation for comparison.

As Kendall's τ has a relationship with the linear correlation ρ under elliptical distributions (Kendall and Gibbons, 1990; Lindskog, Mcneil and Schmock, 2003), we also set up a function between ρ_g and ρ under elliptical distributions. This provides us an alternative method to estimate ρ based on estimation of ρ_g. The asymptotic normality of the estimator based on the symmetric Gini correlation is established. Its asymptotic efficiency and finite sample performance are compared with those of Pearson, Kendall's τ and the regular Gini correlation coefficients under various elliptical distributions.

As any quantity based on spatial ranks, ρ_g is only invariant under translation and homogeneous change. In order to gain the invariance property under heterogeneous changes, we provide an affine invariant version.

The paper is organized as follows. In Section 2, we introduce a symmetric Gini covariance and the corresponding correlation. Section 3 presents the influence function. Section 4 gives an estimator of the symmetric Gini correlation and establishes its asymptotic properties. In Section 5, we present the affine invariant version of the symmetric Gini correlation and explore finite sample efficiency of the corresponding estimator. We present a real data application of the proposed correlation in Section 6. Section 7 concludes the paper with a brief summary. All proofs are reserved to the Appendix.

2. SYMMETRIC GINI COVARIANCE AND CORRELATION

The main focus of this section is to present the proposed symmetric Gini covariance and correlation, and to study the corresponding properties.

2.1. Spatial rank

Given a random vector $\mathbf{Z}$ in $\mathbb{R}^d$ with distribution H, the *spatial rank* of $\mathbf{z}$ with respect to the distribution H is defined as

$$\mathbf{r}(\mathbf{z}, H) := \mathbb{E}\mathbf{s}(\mathbf{z} - \mathbf{Z}) = \mathbb{E}\frac{\mathbf{z} - \mathbf{Z}}{\|\mathbf{z} - \mathbf{Z}\|},$$

where $\mathbf{s}(\cdot)$ is the *spatial sign* function defined as $\mathbf{s}(\mathbf{z}) = \mathbf{z}/\|\mathbf{z}\|$ with $\mathbf{s}(\mathbf{0}) = \mathbf{0}$. The solution of $\mathbf{r}(\mathbf{z}, H) = \mathbf{0}$ is called the *spatial median* of H, which minimizes $\mathbb{E}_H\|\mathbf{z} - \mathbf{Z}\|$. Obviously, $\mathbb{E}\mathbf{r}(\mathbf{Z}, H) = \mathbf{0}$ if H is continuous. For more comprehensive account on the spatial rank, see Oja (2010).

In particular, for $d = 2$ with $\mathbf{Z} = (X, Y)^T$, the bivariate spatial rank function of $\mathbf{z} = (x, y)^T$ is

$$\mathbf{r}(\mathbf{z}, H) = \mathbb{E}\frac{(x - X, y - Y)^T}{\|\mathbf{z} - \mathbf{Z}\|} := (R_1(\mathbf{z}), R_2(\mathbf{z}))^T,$$

where $R_1(\mathbf{z}) = \mathbb{E}(x - X)/\|\mathbf{z} - \mathbf{Z}\|$ and $R_2(\mathbf{z}) = \mathbb{E}(y - Y)/\|\mathbf{z} - \mathbf{Z}\|$ are two components of the joint rank function $\mathbf{r}(\mathbf{z}, H)$.

2.2. Symmetric Gini covariance

Our new symmetric covariance and correlation are defined based on the bivariate spatial rank function. Replacing the univariate centered rank in (1) with $R_2(\mathbf{z})$, we define the *symmetric Gini covariance* as

$$\mathrm{cov}_g(X, Y) := 2\mathbb{E}X R_2(\mathbf{Z}). \tag{2}$$

Note that $\mathrm{cov}_g(X, Y) = 2\mathrm{cov}(X, R_2(\mathbf{Z}))$ if H is continuous. Dually, $\mathrm{cov}_g(Y, X) = 2\mathbb{E}Y R_1(\mathbf{Z})$ can also be taken as the definition of the symmetric Gini covariance between X and Y. Indeed,

$$\mathrm{cov}_g(X, Y) = 2\mathbb{E}X R_2(\mathbf{Z}) = 2\mathbb{E}\left(X_1 \mathbb{E}\left[\frac{Y_1 - Y_2}{\|\mathbf{Z}_1 - \mathbf{Z}_2\|}\Big|\mathbf{Z}_1\right]\right) = 2\mathbb{E}X_1\frac{Y_1 - Y_2}{\|\mathbf{Z}_1 - \mathbf{Z}_2\|}$$

$$= -2\mathbb{E}X_2\frac{Y_1 - Y_2}{\|\mathbf{Z}_1 - \mathbf{Z}_2\|} = \mathbb{E}\left[\frac{(X_1 - X_2)(Y_1 - Y_2)}{\|\mathbf{Z}_1 - \mathbf{Z}_2\|}\right] = \mathrm{cov}_g(Y, X), \tag{3}$$

where $\mathbf{Z}_1 = (X_1, Y_1)^T$ and $\mathbf{Z}_2 = (X_2, Y_2)^T$ are independent copies of $\mathbf{Z} = (X, Y)^T$ from H. In addition, we define

$$\mathrm{cov}_g(X, X) := 2\mathbb{E}X R_1(\mathbf{Z}) = \mathbb{E}\frac{(X_1 - X_2)^2}{\|\mathbf{Z}_1 - \mathbf{Z}_2\|}; \tag{4}$$

$$\mathrm{cov}_g(Y, Y) := 2\mathbb{E}Y R_2(\mathbf{Z}) = \mathbb{E}\frac{(Y_1 - Y_2)^2}{\|\mathbf{Z}_1 - \mathbf{Z}_2\|}. \tag{5}$$

We see that not only the Gini covariance between X and Y but also Gini variances of X and of Y are defined jointly through the spatial rank. Dang, Sang and Weatherall (2015) considered

the Gini covariance matrix $\mathbf{\Sigma}_g = 2\mathbb{E}\mathbf{Z}\mathbf{r}^T(\mathbf{Z})$. The covariances defined above in (2), (4) and (5) are elements of $\mathbf{\Sigma}_g$ for two dimensional random vectors. Rather than the assumption on a finite second moment in the usual covariance and variance, the Gini counterparts assume only the first moment, hence being more suitable for heavy-tailed distributions. A related covariance matrix is spatial sign covariance matrix (SSCM), which requires a location parameter to be known but no assumption on moments (Visuri, Koivunen and Oja, 2000).

Particularly, if Z is a one dimensional random variable, we have $\mathrm{cov}_g(Z, Z) = \mathbb{E}|Z_1 - Z_2|$, which reduces to GMD. In this sense, we may view the symmetric Gini covariance as a direct generalization of GMD to two variables.

2.3. Symmetric Gini correlation

Using the symmetric Gini covariance defined by (2), we propose a symmetric *Gini correlation coefficient* as follows.

Definition 2. $\mathbf{Z} = (X, Y)^T$ *is a bivariate random vector from the distribution H with finite first moment and non-degenerate marginal distributions, then the symmetric Gini correlation between X and Y is*

$$\rho_g(X, Y) := \frac{cov_g(X, Y)}{\sqrt{cov_g(X, X)}\sqrt{cov_g(Y, Y)}} = \frac{\mathbb{E}XR_2(\mathbf{Z})}{\sqrt{\mathbb{E}XR_1(\mathbf{Z})}\sqrt{\mathbb{E}YR_2(\mathbf{Z})}}. \tag{6}$$

Theorem 1.   *For a bivariate random vector $(X, Y)^T$ from H with finite first moment, ρ_g has the following properties:*

1. $\rho_g(X, Y) = \rho_g(Y, X)$.
2. $-1 \leq \rho_g(X, Y) \leq 1$.
3. *If X, Y are independent, then $\rho_g(X, Y) = 0$.*
4. *If $Y = aX + b$ and $a \neq 0$, then $\rho_g = \mathrm{sgn}(a)$.*
5. $\rho_g(aX + b, aY + d) = \rho_g(X, Y)$ *for any constants b, d and $a \neq 0$. Measure ρ_g is sensitive to a heterogeneous change, i.e., $\rho_g(aX, cY) \neq \rho_g(X, Y)$ for $a \neq c$. In particular, $\rho_g(X, Y) = -\rho_g(aX, -aY) = -\rho_g(-aX, aY)$.*

The proof is placed in the Appendix. Theorem 1 shows that the symmetric Gini correlation has all properties of Pearson correlation coefficient except for Property 5. It loses the invariance property under heterogeneous changes because of the Euclidean norm in the spatial rank function. To overcome this drawback, we give the affine invariant version of the ρ_g in Section 5. Comparing with Pearson correlation, as we will see in Section 3, the Gini correlation is more robust in terms of the influence function.

2.4. Symmetric Gini correlation for elliptical distributions

The relationship between Kendall's τ and the linear correlation coefficient ρ, $\tau = 2/\pi \arcsin(\rho)$, holds for all elliptical distributions. So $\rho = \sin(\pi\tau/2)$ provides a robust estimation method for ρ by estimating τ (Lindskog *et al.*, 2003). This motivates us to explore the relationship between the symmetric Gini correlation ρ_g and the linear correlation coefficient ρ under elliptical distributions.

A d-dimensional continuous random vector $\mathbf{Z}$ has an elliptical distribution if its density function is of the form

$$f(\mathbf{z}|\mu, \mathbf{\Sigma}) = |\mathbf{\Sigma}|^{-1/2}g\{(\mathbf{z} - \mu)^T\mathbf{\Sigma}^{-1}(\mathbf{z} - \mu)\}, \tag{7}$$

where $\boldsymbol{\Sigma}$ is the scatter matrix, μ is the location parameter and the nonnegative function g is the density generating function. An important property for the elliptical distribution is that the non-negative random variable $R = ||\boldsymbol{\Sigma}^{-1/2}(\mathbf{Z} - \mu)||$ is independent of $\mathbf{U} = \{\boldsymbol{\Sigma}^{-1/2}(\mathbf{Z} - \mu)\}/R$, which is uniformly distributed on the unit sphere. When $d = 1$, the class of elliptical distributions coincides with the location-scale class. For $d = 2$, let $\mathbf{Z} = (X, Y)^T$ and Σ_{ij} be the (i, j) element of $\boldsymbol{\Sigma}$, then the *linear correlation coefficient* of X and Y is $\rho = \rho(X, Y) := \frac{\Sigma_{12}}{\sqrt{\Sigma_{11}\Sigma_{22}}}$. If the second moment of $\mathbf{Z}$ exists, then the scatter parameter $\boldsymbol{\Sigma}$ is proportional to the covariance matrix. Thus the Pearson correlation ρ_p is well defined and is equal to the parameter ρ in the elliptical distributions. If $\Sigma_{11} = \Sigma_{22} = \sigma^2$, we say X and Y are homogeneous, and $\boldsymbol{\Sigma}$ can then be written as $\boldsymbol{\Sigma} = \sigma^2 \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$. In this case, if $\rho = \pm 1$, $\boldsymbol{\Sigma}$ is singular and the distribution reduces to an one-dimensional distribution.

The following theorem states the relationship between ρ_g and ρ under elliptical distributions.

Theorem 2. *If $\mathbf{Z} = (X, Y)^T$ has an elliptical distribution H with finite first moment and the scatter matrix $\boldsymbol{\Sigma} = \sigma^2 \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$, then we have*

$$\rho_g = k(\rho) = \begin{cases} \rho & \rho = 0, \pm 1, \\ \dfrac{1}{\rho} + \dfrac{\rho - 1}{\rho} \dfrac{EK(\frac{2\rho}{\rho+1})}{EE(\frac{2\rho}{\rho+1})}, & otherwise, \end{cases} \tag{8}$$

where

$$EK(x) = \int_0^{\pi/2} \frac{1}{\sqrt{1 - x^2 \sin^2 \theta}} \, d\theta \quad and \quad EE(x) = \int_0^{\pi/2} \sqrt{1 - x^2 \sin^2 \theta} \, d\theta$$

are the complete elliptic integral of the first kind and the second kind, respectively.

The relationship (8) holds only for $\boldsymbol{\Sigma}$ with $\Sigma_{11} = \Sigma_{22}$ because of the loss of invariance property of ρ_g under the heterogeneous changes (Theorem 1). Note that for any elliptical distribution, the regular Gini correlations are equal to ρ. Schechtman and Yitzhaki (1987) proved that $\gamma(X, Y) = \gamma(Y, X) = \rho$ for bivariate normal distributions, but their proof can be modified for all elliptical distributions. Based on the spatial sign covariance matrix, Dürre, Vogel and Fried (2015) considered a spatial sign correlation coefficient, which equals to ρ for elliptical distributions.

Figure 1 plots the proposed symmetric Gini correlation ρ_g as function of ρ under homogeneous elliptical distributions with finite second moment. In comparison, we also plot Pearson ρ_p and Kendall's τ against ρ. All correlations are increasing in ρ. It is clear that $|\tau| < |\rho_g| < |\rho_p| = |\rho|$.

With (8), the estimate $\hat{\rho}_g$ of ρ_g can be corrected to ensure Fisher consistency by using the inversion transformation $k^{-1}(\hat{\rho}_g)$, denoted as $\hat{\rho}^g$. In the next section, we study the influence function of ρ_g, which can be used to evaluate robustness and efficiency of the estimators $\hat{\rho}_g$ in any distribution and that of $\hat{\rho}^g$ under elliptical distributions.

3. INFLUENCE FUNCTION

The influence function (IF) introduced by Hampel (1974) is now a standard tool in robust statistics for measuring effects on estimators due to infinitesimal perturbations of sample distribution
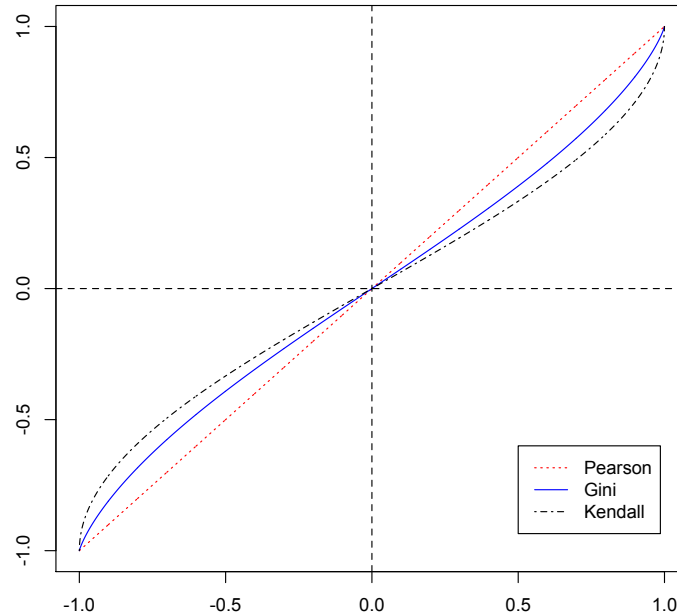
FIGURE 1: Pearson ρ_p, Kendall's τ and symmetric Gini ρ_g correlation coefficients versus ρ, the correlation parameter of homogeneous elliptical distributions with finite second moment.

functions (Hampel *et al.*, 1986). For a cdf H on $\mathbb{R}^d$ and a functional $T : H \mapsto T(H) \in \mathbb{R}^m$ with $m \geq 1$, the IF of T at H is defined as $\mathrm{IF}(\mathbf{z}; T, H) = \lim\limits_{\varepsilon \downarrow 0} \dfrac{T((1-\varepsilon)H + \varepsilon\delta_{\mathbf{z}}) - T(H)}{\varepsilon}, \quad \mathbf{z} \in \mathbb{R}^d$, where $\delta_{\mathbf{z}}$ denotes the point mass distribution at $\mathbf{z}$. Under regularity conditions on T (Hampel *et al.*, 1986; Serfling, 1980), we have $\mathbb{E}_H\{\mathrm{IF}(\mathbf{Z}; T, H)\} = \mathbf{0}$ and the von Mises expansion

$$T(H_n) - T(H) = \frac{1}{n}\sum_{i=1}^{n} \mathrm{IF}(\mathbf{z}_i; T, H) + o_p(n^{-1/2}), \tag{9}$$

where H_n denotes the empirical distribution based on sample $\mathbf{z}_1,...,\mathbf{z}_n$. This representation shows the connection of the IF with robustness of T, observation by observation. Furthermore, (9) yields asymptotic m-variate normality of $T(H_n)$,

$$\sqrt{n}(T(H_n) - T(H)) \xrightarrow{d} N(\mathbf{0}, \mathbb{E}_H(\mathrm{IF}(\mathbf{Z}; T, H)\mathrm{IF}(\mathbf{Z}; T, H)^T) \ \text{ as } n \to \infty. \tag{10}$$

To find the influence function of the symmetric Gini correlation defined in (6), let $T_1(H) = 2\mathbb{E}XR_1(\mathbf{Z})$, $T_2(H) = 2\mathbb{E}XR_2(\mathbf{Z})$, $T_3(H) = 2\mathbb{E}YR_2(\mathbf{Z})$ and $h(t_1, t_2, t_3) = t_2/\sqrt{t_1 t_3}$. Then $\rho_g = T(H) = h(T_1, T_2, T_3)$. Denote the influence function of T_i as $L_i(x, y) = \mathrm{IF}((x, y)^T; T_i, H)$ for $i = 1, 2, 3$.

Theorem 3. *For any distribution H with finite first moment, the influence function of $\rho_g = T(H)$ is given by*

$$IF((x,y)^T; \rho_g, H) = -\frac{\rho_g}{2}\left(\frac{L_1(x,y)}{T_1} - \frac{2L_2(x,y)}{T_2} + \frac{L_3(x,y)}{T_3}\right)$$

$$= -\frac{\rho_g}{2}\left(\frac{1}{T_1}\int\frac{2(x-x_1)^2}{\sqrt{(x-x_1)^2+(y-y_1)^2}}dH(x_1,y_1)\right.$$

$$-\frac{1}{T_2}\int\frac{4(x-x_1)(y-y_1)}{\sqrt{(x-x_1)^2+(y-y_1)^2}}dH(x_1,y_1)$$

$$\left.+\frac{1}{T_3}\int\frac{2(y-y_1)^2}{\sqrt{(x-x_1)^2+(y-y_1)^2}}dH(x_1,y_1)\right).$$

Note that each of $L_i(x,y)$ is approximately linear in x or y. Comparing with the quadratic effects in the Pearson's correlation coefficient (Devlin, Gnanadesikan and Kettering, 1975),

$$\text{IF}((x,y)^T; \rho_p, H) = \frac{(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y} - \frac{1}{2}\rho\left[\frac{(x-\mu_X)^2}{\sigma_X^2} + \frac{(y-\mu_Y)^2}{\sigma_Y^2}\right],$$

ρ_g is more robust than the Pearson correlation. However, ρ_g is not as robust as Kendall's τ correlation since the influence function of ρ_g is unbounded. Kendall's τ correlation has a bounded influence function (Croux and Dehon, 2010), which is $\text{IF}((x,y)^T; \tau, H) = 2\{2P_H[(x-X)(y-Y) > 0] - 1 - \tau\}$. In this sense, ρ_g is more robust than ρ_p but less robust than τ.

IF of ρ_p	IF of ρ_g	IF of τ
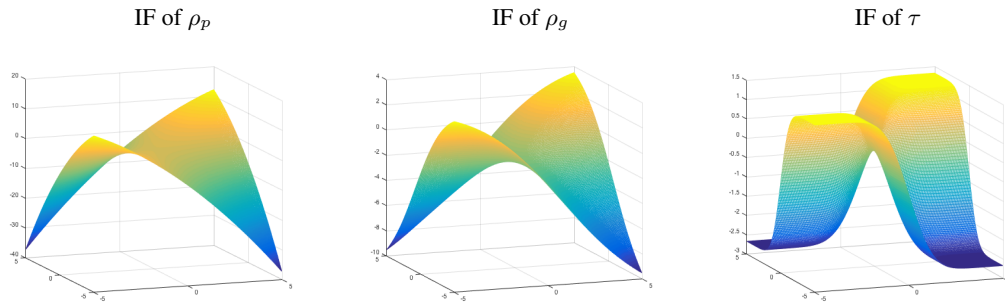
FIGURE 2: Influence functions of correlation correlations ρ_p, ρ_g and τ for the bivariate normal distribution with $\mu_x = \mu_y = 0$, $\sigma_x = \sigma_y = 1$ and $\rho = 0.5$.

Figure 2 displays the influence function of each correlation coefficient for the bivariate normal distribution with $\mu_X = \mu_Y = 0$, $\sigma_X = \sigma_Y = 1$ and $\rho = 0.5$. Note that scales of the value of the influence functions in the three plots are quite different.

4. ESTIMATION

Let $\mathbf{z}_i = (x_i, y_i)^T$, and $\mathcal{Z} = (\mathbf{z}_1, \mathbf{z}_2, ..., \mathbf{z}_n)$ be a random sample from a continuous distribution H with an empirical distribution H_n. Replacing H in (6) with H_n, we have the sample counter-

part of the symmetric Gini correlation coefficient $\rho_g(H_n) = \hat{\rho}_g$:

$$\hat{\rho}_g = \frac{\sum_{1 \leq i < j \leq n} \frac{(x_i - x_j)(y_i - y_j)}{\sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}}}{\sqrt{\sum_{1 \leq i < j \leq n} \frac{(x_i - x_j)^2}{\sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}}} \sqrt{\sum_{1 \leq i < j \leq n} \frac{(y_i - y_j)^2}{\sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}}}}.$$

Using the same notations in Section 3, we have the following central limit theorem of the sample symmetric Gini correlation $\hat{\rho}_g$.

Theorem 4. *Let $\mathbf{z}_1, \mathbf{z}_2, ..., \mathbf{z}_n$ be a random sample from 2-dimensional distribution H with finite second moment. Then $\hat{\rho}_g$ is an unbiased, $\sqrt{n}$-consistent estimator of ρ_g. Furthermore, $\sqrt{n}(\hat{\rho}_g - \rho_g) \xrightarrow{d} N(0, v_g)$ as $n \to \infty$, where*

$$v_g = \mathbb{E}[IF((X,Y)^T, \rho_g, H)]^2 = \frac{\rho_g^2}{4} \left(\frac{1}{T_1^2} \mathbb{E}[L_1^2(X,Y)] + \frac{4}{T_2^2} \mathbb{E}[L_2^2(X,Y)] \right.$$

$$+ \frac{1}{T_3^2} \mathbb{E}[L_3^2(X,Y)] - \frac{4}{T_1 T_2} \mathbb{E}L_1(X,Y)L_2(X,Y) + \frac{2}{T_1 T_3} \mathbb{E}L_1(X,Y)L_3(X,Y)$$

$$\left. - \frac{4}{T_2 T_3} \mathbb{E}L_2(X,Y)L_3(X,Y) \right).$$

Although (10) implies Theorem 4, it is hard to check regularity conditions for the von Mises expansion (9). Instead, we prove it in the Appendix using the multivariate delta method and the asymptotic normality of the sample Gini covariance matrix, which is based on the U-statistics theory (Dang *et al.*, 2015).

For an elliptical distribution H, Theorem 2 shows that $\hat{\rho}_g$ is not a Fisher consistent estimator of ρ. We need to consider the inverse transformation $\hat{\rho}^g = k^{-1}(\hat{\rho}_g)$, where the function k is given in (8). Applying the delta method, we obtain the $\sqrt{n}$-consistency of estimator $\hat{\rho}^g$ for ρ.

Theorem 5. *Let $\mathbf{z}_1, \mathbf{z}_2, ..., \mathbf{z}_n$ be a sample from elliptical distribution H with finite second moment and $\boldsymbol{\Sigma} = \sigma^2 \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$. Then $\hat{\rho}^g = k^{-1}(\hat{\rho}_g)$ is unbiased and a $\sqrt{n}$-consistent estimator of ρ. Moreover, $\sqrt{n}(\hat{\rho}^g - \rho) \xrightarrow{d} N(0, [1/k'(\rho)]^2 v_g)$ as $n \to \infty$, where the function k is given in (8), v_g is given in Theorem 4, and $k'(\rho)$ is*

$$k'(\rho) = \frac{-3(\rho+1)EE^2(\frac{2\rho}{\rho+1}) + 4EE(\frac{2\rho}{\rho+1})EK(\frac{2\rho}{\rho+1}) + (\rho-1)EK^2(\frac{2\rho}{\rho+1}))}{2(\rho+1)\rho^2 EE^2(\frac{2\rho}{\rho+1})}.$$

Theorem 5 provides an estimator based on $\hat{\rho}_g$ for the correlation parameter for elliptical distributions. The asymptotic variance $[k'(\rho)]^{-2} v_g$ can be used to evaluate asymptotic efficiency of $\hat{\rho}^g$.

4.1. Asymptotic efficiency

To compare relative efficiency, we present the asymptotic variances (ASV) of four estimators of ρ including Pearson's estimator $\hat{\rho}_p$, $\hat{\rho}^g$, the regular Gini correlation estimator, and the estimator through Kendall's τ estimator.

Witting and Müller-Funk (1995) established asymptotic normality for the regular sample Pearson correlation coefficient $\hat{\rho}_p$:

$$\sqrt{n}(\hat{\rho}_p - \rho) \xrightarrow{d} N(0, v_p) \ as \ n \to \infty,$$

where

$$v_p = (1 + \frac{\rho^2}{2})\frac{\sigma_{22}}{\sigma_{20}\sigma_{02}} + \frac{\rho^2}{4}(\frac{\sigma_{40}}{\sigma_{20}^2} + \frac{\sigma_{04}}{\sigma_{02}^2} - \frac{4\sigma_{31}}{\sigma_{11}\sigma_{20}} - \frac{4\sigma_{13}}{\sigma_{11}\sigma_{02}}),$$

and $\sigma_{kl} = \mathbb{E}[(X - \mathbb{E}X)^k(Y - \mathbb{E}Y)^l]$. The Pearson correlation estimator requires a finite fourth moment on the distribution to evaluate its asymptotic variance. For bivariate normal distributions, the asymptotic variance v_p simplifies to $(1 - \rho^2)^2$.

An estimator $\hat{\rho}_\gamma$ of the regular Gini correlation $\gamma(X, Y)$ is

$$\hat{\rho}_\gamma = \frac{\binom{n}{2}^{-1} \sum_{1 \le i < j \le n} h_1(\mathbf{z}_i, \mathbf{z}_j)}{\binom{n}{2}^{-1} \sum_{1 \le i < j \le n} h_2(\mathbf{z}_i, \mathbf{z}_j)},$$

where

$$h_1(\mathbf{z}_1, \mathbf{z}_2) = [(x_1 - x_2)I(y_1 > y_2) + (x_2 - x_1)I(y_2 > y_1)]/4$$

and $h_2(\mathbf{z}_1, \mathbf{z}_2) = |x_1 - x_2|/4$. Using U-statistic theory, Schechtman and Yitzhak (1987) provided the asymptotic normality:

$$\sqrt{n}(\hat{\rho}_\gamma - \rho_\gamma) \xrightarrow{d} N(0, v_\gamma) \ as \ n \to \infty,$$

with

$$v_\gamma = (4/\theta_2^2)\zeta_1(\theta_1) + (4\theta_1^2/\theta_2^4)\zeta_2(\theta_2) - (8\theta_1/\theta_2^3)\zeta_3(\theta_1, \theta_2),$$

where

$$\theta_1 = \text{cov}(X, G(Y)), \ \ \theta_2 = \text{cov}(X, F(X)),$$

$$\zeta_1(\theta_1) = \mathbb{E}_{\mathbf{z}_1} \{\mathbb{E}_{\mathbf{z}_2}[h_1(\mathbf{Z}_1, \mathbf{Z}_2)]\}^2 - \theta_1^2,$$

$$\zeta_2(\theta_2) = \mathbb{E}_{\mathbf{z}_1} \{\mathbb{E}_{\mathbf{z}_2}[h_2(\mathbf{Z}_1, \mathbf{Z}_2)]\}^2 - \theta_2^2$$

and

$$\zeta_3(\theta_1, \theta_2) = \mathbb{E}_{\mathbf{z}_1} \{\mathbb{E}_{\mathbf{z}_2}[h_1(\mathbf{Z}_1, \mathbf{Z}_2)]\mathbb{E}_{\mathbf{z}_2}[h_2(\mathbf{Z}_1, \mathbf{Z}_2)]\} - \theta_1\theta_2.$$

Under elliptical distributions, $\gamma(X, Y) = \gamma(Y, X) = \rho$, hence the asymptotic variance of $\hat{\rho}_\gamma$ is v_γ. For a normal distribution, Xu *et al.* (2010) provided an explicit formula of v_γ, given by $v_\gamma = \pi/3 + (\pi/3 + 4\sqrt{3})\rho^2 - 4\rho \arcsin(\rho/2) - 4\rho^2 \sqrt{4 - \rho^2}$.

Borovskikh (1996) presented the asymptotic normality of the estimator $\hat{\tau}$:

$$\sqrt{n}(\hat{\tau} - \tau) \xrightarrow{d} N(0, v_\tau) \ as \ n \to \infty,$$

with

$$v_\tau = 4\mathbb{E}\{\mathbb{E}^2_{\mathbf{z}_1}(\mathrm{sgn}[(X_2 - X_1)(Y_2 - Y_1)])\} - 4\mathbb{E}^2\{\mathrm{sgn}[(X_2 - X_1)(Y_2 - Y_1)]\}.$$

Applying the delta method to $\hat{\rho}_\tau = \sin(\pi\hat{\tau}/2)$, we obtain the asymptotic variance of $\hat{\rho}_\tau$ to be $\frac{\pi^2}{4}(1 - \rho^2)v_\tau$. Under a normal distribution, the asymptotic variance of $\hat{\rho}_\tau$ is $\pi^2(1 - \rho^2)(\frac{1}{9} - \frac{4}{\pi^2}\arcsin^2(\frac{\rho}{2}))$ (Croux and Dehon, 2010).

We compare asymptotic efficiency of the four estimators $\hat{\rho}^g$, $\hat{\rho}_\gamma$, $\hat{\rho}_\tau$ and $\hat{\rho}_p$ under three bivariate elliptical distributions (7) with different fatness on the tail regions: the normal distributions with $g(t) = \frac{1}{2\pi}e^{-t/2}$; the t-distributions with $g(t) = \frac{1}{2\pi}(1 + t/\nu)^{-\nu/2-1}$, where ν is the degree of freedom; and the Kotz type distribution with $g(t) = \frac{1}{2\pi}e^{-\sqrt{t}}$. The normal distribution is the limiting distribution of the t-distributions as $\nu \to \infty$. The Kotz type distribution is a bivariate generalization of the Laplace distribution with the tail region fatness between that of the normal and t distributions (Fang, Kotz and Hg, 1987). We consider only elliptical distributions because all four estimators $\hat{\rho}^g$, $\hat{\rho}_\gamma$, $\hat{\rho}_\tau$ and $\hat{\rho}_p$ are Fisher consistent for parameter ρ. The estimators for non-elliptical distributions may estimate different quantities, resulting in their asymptotical variances incomparable.

TABLE 1: Asymptotic relative efficiencies (ARE) of estimators $\hat{\rho}^g$, $\hat{\rho}_\gamma$ and $\hat{\rho}_\tau$ relative to $\hat{\rho}_p$ for different distributions, with asymptotic variance (ASV($\hat{\rho}_p$)) of Pearson estimator $\hat{\rho}_p$.

Distribution		ARE($\hat{\rho}^g, \hat{\rho}_p$)	ARE($\hat{\rho}_\gamma, \hat{\rho}_p$)	ARE($\hat{\rho}_\tau, \hat{\rho}_p$)	ASV($\hat{\rho}_p$)
	$\rho = 0.1$	0.9321	0.9558	0.9125	0.9816
Normal	$\rho = 0.5$	0.9769	0.9398	0.8925	0.5631
	$\rho = 0.9$	0.9601	0.9004	0.8439	0.0361
	$\rho = 0.1$	1.0182	1.0304	1.0146	1.1558
$t(15)$	$\rho = 0.5$	1.0560	0.9852	0.9896	0.6643
	$\rho = 0.9$	1.0289	0.9468	0.8804	0.0427
	$\rho = 0.1$	2.0095	1.9502	2.2586	2.8800
$t(5)$	$\rho = 0.5$	1.9795	1.7666	2.1060	1.5961
	$\rho = 0.9$	1.8629	1.5346	1.7940	0.1019
	$\rho = 0.1$	1.2081	1.1385	1.2171	1.6382
Kotz	$\rho = 0.5$	1.1850	1.0854	1.1510	0.9378
	$\rho = 0.9$	1.1599	0.9789	1.0256	0.0602

Without loss of generality, we consider only cases with $\rho > 0$. Listed in Table 1 are asymptotic variances (ASV) of Pearson estimator $\hat{\rho}_p$, and asymptotic relative efficiencies (ARE) of estimators $\hat{\rho}^g$, $\hat{\rho}_\gamma$ and $\hat{\rho}_\tau$ relative to $\hat{\rho}_p$ for different elliptical distributions under the homogeneous assumption, where the asymptotic relative efficiency of an estimator with respect to another is defined as $\mathrm{ARE}(\hat{\rho}_1, \hat{\rho}_2) = \mathrm{ASV}(\hat{\rho}_2)/\mathrm{ASV}(\hat{\rho}_1)$. The asymptotic variance of each estimator is obtained based on a combination of numeric integration and the Monte Carlo simulation.

Table 1 shows that the asymptotic variances of $\hat{\rho}_p$, $\hat{\rho}^g$, $\hat{\rho}_\gamma$ and $\hat{\rho}_\tau$ all decrease as ρ increases. When $\rho = 1$, every estimator is equal to 1 without any estimation error. Asymptotic variances increase for t distributions as the degrees of freedom ν decreases. Under normal distributions, the Pearson correlation estimator is the maximum likelihood estimator of ρ, thus is most efficient asymptotically. The symmetric Gini estimator $\hat{\rho}^g$ is high in efficiency with ARE's greater than 93%; it is more efficient than Kendall's estimator $\hat{\rho}_\tau$. For heavy-tailed distributions, the sym-

metric Gini estimator is more efficient than Pearson's estimator $\hat{\rho}_p$. The AREs of the symmetric Gini estimator are close to those of Kendall's estimator $\hat{\rho}_\tau$ for Kotz samples. Comparing with the regular Gini correlation estimator, the proposed measure has higher efficiency for all cases except for $\rho = 0.1$ under normal and $t(15)$ distributions, in which the efficiency is about 2.4% and 1.2% lower. These results may be explained by that the joint spatial rank used in $\hat{\rho}^g$ takes more dependence information than the marginal rank used in $\hat{\rho}_\gamma$.

In summary, the proposed symmetric Gini estimator has nice asymptotic behavior that well balances between efficiency and robustness. It is more efficient than the regular Gini, which is also symmetric under elliptical distributions.

4.2. Finite sample efficiency

We conduct a small simulation to study the finite sample efficiencies of the correlation estimators for the symmetric Gini, regular Gini, Kendall's τ and Pearson correlations. $M = 3000$ samples of two different sample sizes, $n = 30, 300$, are drawn from t-distributions with $1, 3, 5, 15$ and ∞ degrees of freedoms and from the Kotz distribution. We use R Package "mnormt" to generate samples from multivariate t and normal distributions (referred as $t(\infty)$ in Table 2). For the Kotz sample, we first generate uniformly distributed random vectors on the unit circle by $\mathbf{u} = (\cos\theta, \sin\theta)^T$ with θ in $[0, 2\pi]$, then generate r from a Gamma distribution with $\alpha = 2$ (the shape parameter) and $\beta = 1$ (the scale parameter) and hence $\mathbf{\Sigma}^{1/2}r\mathbf{u} + \mu$ is a sample from bivariate Kotz$(\mu, \mathbf{\Sigma})$. For more details, see Dang *et al.* (2015).

For each sample m, each estimator $\hat{\rho}^{(m)}$ is calculated and the root of mean squared error (RMSE) of the estimator $\hat{\rho}$ is computed as

$$\text{RMSE}(\hat{\rho}) = \sqrt{\frac{1}{M}\sum_{m=1}^{M}(\hat{\rho}^{(m)} - \rho)^2}.$$

The procedure is repeated 100 times. In Table 2, we report the mean and standard deviation (in parentheses) of $\sqrt{n}$RMSEs of correlation estimators $\hat{\rho}^g$, $\hat{\rho}_\gamma$, $\hat{\rho}_\tau$ and $\hat{\rho}_p$ when the scatter matrix is homogeneous with $\mathbf{\Sigma} = \sigma^2 \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$. The case of $n = \infty$ corresponds to the asymptotic standard deviation of each estimator that can be obtained from Table 1. Since $\hat{\rho}^g$ cannot be given explicitly due to the inverse transformation involved in $\hat{\rho}^g = k^{-1}(\hat{\rho}_g)$, we use a numerical way to obtain $\hat{\rho}^g$ by creating a correspondence between s and t, where $s = k(t)$ and t is a very fine grid on $[0, 1]$. $\hat{\rho}_g$ is computed by using R package "ICSNP" for spatial.rank function.

In Table 2, the $\sqrt{n}$RMSEs demonstrate an increasing trend as ρ decreases or as the degree of freedom ν decreases for t distributions. For $n = 300$, the behavior of each estimator is similar to its asymptotic efficiency behavior. For example, for $n = 300$ and $\rho = 0.5$ under the normal distribution, the $\sqrt{n}$RMSE of $\hat{\rho}_p$ is 0.7534 close to the asymptotic standard deviation 0.7504. We include heavy-tailed $t(1)$ and $t(3)$ distributions in the simulation to demonstrate finite sample behavior of Pearson and Gini estimators when their asymptotic variances may not exist. $\sqrt{n}$RMSE of $\hat{\rho}_p$ is about twice as that of $\hat{\rho}^g$ for $n = 300$ in both $t(1)$ and $t(3)$ distributions. For $t(1)$ distribution, $\hat{\rho}_\tau$ is much better than others in terms of $\sqrt{n}$RMSE. When the sample size is small ($n = 30$), $\hat{\rho}^g$ performs the best. The $\sqrt{n}$RMSEs of $\hat{\rho}^g$ are smaller than that of $\hat{\rho}_\tau$ even under heavy-tailed $t(3)$ distribution. $\hat{\rho}^g$ has a smaller $\sqrt{n}$RMSE than the Pearson correlation estimator for the normal distribution with $\rho = 0.1$ and all other distributions. The symmetric Gini estimator $\hat{\rho}^g$ has smaller $\sqrt{n}$RMSE than the regular Gini estimator $\hat{\rho}_\gamma$ for all cases we consider. The simulation demonstrates superior finite sample behavior of the proposed estimator.

TABLE 2: The mean and standard deviation (in parentheses) of $\sqrt{n}$RMSE of $\hat{\rho}^g$, $\hat{\rho}_\gamma$, $\hat{\rho}_\tau$ and $\hat{\rho}_p$ under different distributions when the scatter matrix is homogeneous.

Dist	ρ	n	$\sqrt{n}$RMSE($\hat{\rho}^g$)	$\sqrt{n}$RMSE($\hat{\rho}_\gamma$)	$\sqrt{n}$RMSE($\hat{\rho}_\tau$)	$\sqrt{n}$RMSE($\hat{\rho}_p$)
	$\rho = 0.1$	$n = 30$	0.7767 (.0115)	1.0418 (.0115)	1.0785 (.0120)	1.0095 (.0120)
		$n = 300$	0.9648 (.0104)	1.0184 (.0121)	1.0427 (.0139)	0.9925 (.0121)
$t(\infty)$	$\rho = 0.5$	$n = 30$	0.7887 (.0110)	0.8150 (.0115)	0.8517 (.0126)	0.7827 (.0115)
		$n = 300$	0.7638 (.0087)	0.7777 (.0104)	0.8002 (.0104)	0.7534 (.0104)
	$\rho = 0.9$	$n = 30$	0.2147 (.0044)	0.2306 (.0044)	0.2541 (.0049)	0.2103 (.0044)
		$n = 300$	0.1957 (.0017)	0.2026 (.0035)	0.2113 (.0035)	0.1923 (.0017)
	$\rho = 0.1$	$n = 30$	0.8013 (.0120)	1.0828 (.0120)	1.1026 (.0115)	1.0735 (.0115)
		$n = 300$	1.0011 (.0104)	1.0669 (.0121)	1.0721 (.0139)	1.0756 (.0121)
$t(15)$	$\rho = 0.5$	$n = 30$	0.8177 (.0115)	0.8506 (.0126)	0.8731 (.0131)	0.8347 (.0126)
		$n = 300$	0.7985 (.0104)	0.8227 (.0104)	0.8279 (.0104)	0.8193 (.0104)
	$\rho = 0.9$	$n = 30$	0.2251 (.0044)	0.2432 (.0044)	0.2635 (.0164)	0.2262 (.0044)
		$n = 300$	0.2044 (.0035)	0.2165 (.0035)	0.2200 (.0035)	0.2078 (.0035)
	$\rho = 0.1$	$n = 30$	0.8698 (.0137)	1.2083 (.0126)	1.1562 (.0131)	1.2987 (.0137)
		$n = 300$	1.1085 (.0121)	1.2246 (.0156)	1.1310 (.0139)	1.5155 (.0242)
$t(5)$	$\rho = 0.5$	$n = 30$	0.9032 (.0110)	0.9580 (.0126)	0.9202 (.0126)	1.0221 (.0159)
		$n = 300$	0.9007 (.0121)	0.9492 (.0121)	0.8764 (.0121)	1.1535 (.0208)
	$\rho = 0.9$	$n = 30$	0.2569 (.0164)	0.2859 (.0066)	0.2832 (.0164)	0.2908 (.0088)
		$n = 300$	0.2338 (.0069)	0.2615 (.0035)	0.2408 (.0069)	0.2996 (.0087)
	$\rho = 0.1$	$n = 30$	0.9706 (.0137)	1.3923 (.0170)	1.2050 (.0142)	1.6459 (.0214)
		$n = 300$	1.2921 (.0156)	1.5329 (.0191)	1.1865 (.0156)	2.7782 (.0554)
$t(3)$	$\rho = 0.5$	$n = 30$	1.0231 (.0131)	1.1201 (.0170)	0.9651 (.0148)	1.3343 (.0246)
		$n = 300$	1.1068 (.0173)	1.2142 (.0208)	0.9284 (.0121)	2.1876 (.0675)
	$\rho = 0.9$	$n = 30$	0.3127 (.0104)	0.3642 (.0131)	0.3051 (.0066)	0.4289 (.0236)
		$n = 300$	0.2944 (.0104)	0.3672 (.0173)	0.2615 (.0035)	0.6564 (.0658)
	$\rho = 0.1$	$n = 30$	1.7418 (.0301)	2.7222 (.0285)	1.3704 (.0170)	3.3104 (.0279)
		$n = 300$	4.3423 (.0814)	6.7879 (.0918)	1.3735 (.0173)	10.256 (.0918)
$t(1)$	$\rho = 0.5$	$n = 30$	1.6706 (.0153)	2.3892 (.0361)	1.1184 (.0164)	2.9687 (.0466)
		$n = 300$	4.2574 (.0485)	5.9357 (.1057)	1.0999 (.0156)	9.1781 (.1472)
	$\rho = 0.9$	$n = 30$	0.9065 (.0361)	1.2083 (.0586)	0.4004 (.0088)	1.5917 (.0728)
		$n = 300$	2.1616 (.1074)	2.9947 (.1784)	0.3464 (.0052)	4.9589 (.2182)
	$\rho = 0.1$	$n = 30$	0.8692 (.0126)	1.2083 (.0148)	1.1842 (.0148)	1.2389 (.0148)
		$n = 300$	1.0947 (.0139)	1.2055 (.0173)	1.1639 (.0156)	1.2713 (.0173)
Kotz	$\rho = 0.5$	$n = 30$	0.9037 (.0137)	0.9569 (.0148)	0.9465 (.0142)	0.9711 (.0170)
		$n = 300$	0.8903 (.0121)	0.9318 (.0121)	0.9059 (.0121)	0.9665 (.0121)
	$\rho = 0.9$	$n = 30$	0.2563 (.0164)	0.2832 (.0164)	0.2952 (.0060)	0.2706 (.0060)
		$n = 300$	0.2304 (.0035)	0.2529 (.0035)	0.2494 (.0035)	0.2477 (.0035)

5. THE AFFINE INVARIANT VERSION OF SYMMETRIC GINI CORRELATION

The proposed ρ_g in Section 2.3 is only invariant under translation and homogeneous change. We now provide an affine invariant version of ρ_g, denoted as ρ_G, in order to gain the invariance property under heterogeneous changes. This is based on the affine equivariant (AE) Gini covariance matrix $\mathbf{\Sigma}_G$ proposed by Dang *et al.* (2015).

The basic idea of $\mathbf{\Sigma}_G$ is that the Gini covariance matrix on standardized data should be proportional to the identity matrix $\mathbf{I}$. That is, $\mathbb{E}(\mathbf{\Sigma}_G^{-1/2}\mathbf{Z})\mathbf{r}^T(\mathbf{\Sigma}_G^{-1/2}\mathbf{Z}) = c\mathbf{I}$, where c is a positive constant. In other words, the AE version of the Gini covariance matrix is the solution of

$$\mathbb{E}\frac{\mathbf{\Sigma}_G^{-1/2}(\mathbf{Z}_1 - \mathbf{Z}_2)(\mathbf{Z}_1 - \mathbf{Z}_2)^T\mathbf{\Sigma}_G^{-1/2}}{\sqrt{(\mathbf{Z}_1 - \mathbf{Z}_2)^T\mathbf{\Sigma}_G^{-1}(\mathbf{Z}_1 - \mathbf{Z}_2)}} = c(H)\mathbf{I}, \tag{11}$$

where $c(H)$ is a constant depending on H. In this way, the matrix valued functional $\mathbf{\Sigma}_G(\cdot)$ is a scatter matrix in the sense that for any nonsingular matrix A and vector $\mathbf{b}$, $\mathbf{\Sigma}_G(A\mathbf{Z} + \mathbf{b}) = A\mathbf{\Sigma}_G(\mathbf{Z})A^T$.

Let $\mathbf{Z} = (X, Y)^T$ be a bivariate random vector with distribution function H and $\mathbf{\Sigma}_G := \begin{pmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{pmatrix}$ be the solution of (11). Then the affine invariant version of ρ_g is defined as $\rho_G(X, Y) = \frac{G_{21}}{\sqrt{G_{11}}\sqrt{G_{22}}}$. Since the value of $c(H)$ in (11) does not change the value of $\rho_G(X, Y)$, without loss of generality, assume $c(H) = 1$.

Theorem 6. *For any bivariate random vector $\mathbf{Z} = (X, Y)^T$ having an elliptical distribution H with finite first moment, $\rho_G(aX, bY) = sgn(ab)\rho_G(X, Y)$ for any $ab \neq 0$.*

Remark 1. *Under elliptical distributions, $\rho_G = \rho$. This is true since $\mathbf{\Sigma}_G = \mathbf{\Sigma}$ for elliptical distributions.*

When a random sample $\mathbf{z}_1, \mathbf{z}_2, ..., \mathbf{z}_n$ is available, replacing H with its empirical distribution H_n in (11) yields the sample counterpart $\hat{\mathbf{\Sigma}}_G$, and hence the sample $\hat{\rho}_G$ is obtained accordingly. We obtain $\hat{\mathbf{\Sigma}}_G$ by a common re-weighted iterative algorithm:

$$\hat{\mathbf{\Sigma}}_G^{(t+1)} \longleftarrow \frac{2}{n(n-1)} \sum_{1 \leq i < j \leq n} \frac{(\mathbf{z}_i - \mathbf{z}_j)(\mathbf{z}_i - \mathbf{z}_j)^T}{\sqrt{(\mathbf{z}_i - \mathbf{z}_j)^T(\hat{\mathbf{\Sigma}}_G^{(t)})^{-1}(\mathbf{z}_i - \mathbf{z}_j)}}.$$

The initial value can take $\hat{\mathbf{\Sigma}}_G^{(0)} = \mathbf{I}_d$. The iteration stops when $\|\hat{\mathbf{\Sigma}}_G^{(t+1)} - \hat{\mathbf{\Sigma}}_G^{(t)}\| < \varepsilon$ for a pre-specified number $\varepsilon > 0$, where $\|\cdot\|$ can take any matrix norm.

Next, we study finite sample efficiency of $\hat{\rho}_G$ under the same simulation setting as in Section 4.2 except that the scatter matrix is heterogeneous. The scatter matrix of each elliptical distribution is $\mathbf{\Sigma} = \begin{pmatrix} 1 & 2\rho \\ 2\rho & 4 \end{pmatrix}$. Table 3 reports $\sqrt{n}$RMSE of correlation estimators $\hat{\rho}_G$, $\hat{\rho}_\gamma$, $\hat{\rho}_\tau$ and $\hat{\rho}_p$. The numbers in the last three columns are very close to those in Table 2 because $\hat{\rho}_\gamma$, $\hat{\rho}_\tau$ and $\hat{\rho}_p$ are affine invariant. $\sqrt{n}$RMSEs of $\hat{\rho}_G$ are also close to $\sqrt{n}$RMSE of $\hat{\rho}^g$ for $n = 300$, but are larger than those for $n = 30$ and $\rho = 0.1$. The loss of finite sample efficiency of $\hat{\rho}_G$ for a small size under low dependence ρ is probably caused by the iterative algorithm in the computation of $\hat{\rho}_G$. The problem is even worse in $t(1)$ distribution where the first moment does not exist. As the value of ρ increases, $\sqrt{n}$RMSE of each estimator decreases for all distributions. Under Kotz and $t(15)$ distributions, the affine invariant Gini estimator $\hat{\rho}_G$ is the most efficient; under $t(5)$ distribution, the $\sqrt{n}$RMSE of $\hat{\rho}_G$ is smaller than that of Kendall's $\hat{\rho}_\tau$ when $\rho = 0.9$. For

TABLE 3: The mean and standard deviation (in parentheses) of $\sqrt{n}$RMSE of $\hat{\rho}_G$, $\hat{\rho}_\gamma$, $\hat{\rho}_\tau$ and $\hat{\rho}_p$ under different distributions with a heterogeneous scatter matrix.

Dist	ρ	n	$\sqrt{n}$RMSE($\hat{\rho}_G$)	$\sqrt{n}$RMSE($\hat{\rho}_\gamma$)	$\sqrt{n}$RMSE($\hat{\rho}_\tau$)	$\sqrt{n}$RMSE($\hat{\rho}_p$)
	$\rho = 0.1$	$n = 30$	1.0171 (.0126)	1.0401 (.0126)	1.0768 (.0131)	1.0073 (.0120)
		$n = 300$	1.0011 (.0139)	1.0133 (.0139)	1.0392 (.0156)	0.9890 (.0139)
$t(\infty)$	$\rho = 0.5$	$n = 30$	0.7887 (.0120)	0.8123 (.0126)	0.8501 (.0137)	0.7800 (.0120)
		$n = 300$	0.7621 (.0104)	0.7794 (.0104)	0.8002 (.0104)	0.7534 (.0104)
	$\rho = 0.9$	$n = 30$	0.2125 (.0022)	0.2306 (.0044)	0.2541 (.0049)	0.2098 (.0044)
		$n = 300$	0.1940 (.0035)	0.2026 (.0035)	0.2113 (.0035)	0.1923 (.0017)
	$\rho = 0.1$	$n = 30$	1.0582 (.0126)	1.0839 (.0131)	1.1042 (.0126)	1.0741 (.0126)
		$n = 300$	1.0496 (.0121)	1.0687 (.0121)	1.0739 (.0121)	1.0756 (.0121)
$t(15)$	$\rho = 0.5$	$n = 30$	0.8221 (.0099)	0.8506 (.0099)	0.8731 (.0110)	0.8353 (.0099)
		$n = 300$	0.7967 (.0104)	0.8210 (.0121)	0.8279 (.0104)	0.8175 (.0121)
	$\rho = 0.9$	$n = 30$	0.2224 (.0049)	0.2437 (.0049)	0.2635 (.0060)	0.2262 (.0049)
		$n = 300$	0.2026 (.0035)	0.2165 (.0035)	0.2200 (.0035)	0.2078 (.0035)
	$\rho = 0.1$	$n = 30$	1.1727 (.0164)	1.2072 (.0148)	1.1557 (.0153)	1.2981 (.0192)
		$n = 300$	1.1847 (.0156)	1.2246 (.0156)	1.1310 (.0139)	1.5155 (.0242)
$t(5)$	$\rho = 0.5$	$n = 30$	0.9169 (.0120)	0.9585 (.0115)	0.9213 (.0120)	1.0226 (.0137)
		$n = 300$	0.8989 (.0139)	0.9492 (.0139)	0.8764 (.0121)	1.1553 (.0242)
	$\rho = 0.9$	$n = 30$	0.2520 (.0060)	0.2865 (.0071)	0.2832 (.0060)	0.2919 (.0110)
		$n = 300$	0.2304 (.0035)	0.2615 (.0035)	0.2408 (.0035)	0.2979 (.0087)
	$\rho = 0.1$	$n = 30$	1.3540 (.0519)	1.3918 (.0159)	1.2039 (.0142)	1.6475 (.0203)
		$n = 300$	1.4497 (.0225)	1.5346 (.0225)	1.1847 (.0156)	2.7782 (.0606)
$t(3)$	$\rho = 0.5$	$n = 30$	1.0670 (.0159)	1.1190 (.0170)	0.9629 (.0148)	1.3321 (.0219)
		$n = 300$	1.1033 (.0139)	1.2090 (.0173)	0.9249 (.0121)	2.1910 (.0606)
	$\rho = 0.9$	$n = 30$	0.3095 (.0099)	0.3681 (.0137)	0.3062 (.0066)	0.4376 (.0230)
		$n = 300$	0.2841(.0069)	0.3655 (.0156)	0.2615 (.0035)	0.6461 (.0675)
	$\rho = 0.1$	$n = 30$	2.7622 (.0274)	2.7244 (.0268)	1.3693 (.0192)	3.3148 (.0268)
		$n = 300$	6.8381 (.0970)	6.7879 (.0797)	1.3770 (.0173)	10.259 (.0901)
$t(1)$	$\rho = 0.5$	$n = 30$	2.4133 (.0433)	2.3831 (.0372)	1.1206 (.0164)	2.9643 (.0466)
		$n = 300$	5.8768 (.1386)	5.9132 (.1178)	1.0947 (.0139)	9.1522 (.1455)
	$\rho = 0.9$	$n = 30$	1.1875 (.0608)	1.2148 (.0537)	0.4009 (.0088)	1.6015 (.0635)
		$n = 300$	2.7747 (.2148)	2.9930 (.1853)	0.3481 (.0052)	4.9727 (.2113)
	$\rho = 0.1$	$n = 30$	1.1672 (.0131)	1.2066 (.0131)	1.1831 (.0142)	1.2368 (.0142)
		$n = 300$	1.1674 (.0139)	1.2038 (.0139)	1.1605 (.0139)	1.2731 (.0156)
Kotz	$\rho = 0.5$	$n = 30$	0.9136 (.0148)	0.9574 (.0148)	0.9454 (.0153)	0.9706 (.0148)
		$n = 300$	0.8885 (.0121)	0.9336 (.0121)	0.9059 (.0121)	0.9665 (.0121)
	$\rho = 0.9$	$n = 30$	0.2503 (.0049)	0.2815 (.0060)	0.2941 (.0060)	0.2684 (.0055)
		$n = 300$	0.2269 (.0035)	0.2546 (.0035)	0.2511 (.0035)	0.2477 (.0035)

the normal distributions, $\hat{\rho}_G$ is almost as efficient as $\hat{\rho}_p$ when $\rho = 0.9$. The affine invariant Gini correlation estimator shows a good finite sample efficiency. Again, the proposed Gini has smaller $\sqrt{n}$RMSEs than the regular Gini in all cases.

6. APPLICATION

For the purpose of illustration, we apply the symmetric Gini correlations to the famous Fisher's Iris data which is available in R. The data set consists of 50 samples from each of three species of Iris (Setosa, Versicolor and Virginica). Four features are measured in centimeters from each sample: sepal length (Sepal L.), sepal width (Sepal W.), petal length (Petal L.), and petal width (Petal W.). The mean and standard deviation of each of the variables for all data and each species data are listed in Table 4. All the three species have similar sizes in sepals. But Setosa has a much smaller petal size than the other two species. Hence we shall study the correlation of the variables for each Iris species.

TABLE 4: Summary Statistics of Variables in Iris Data

	Mean				Standard Deviation			
	All	Setosa	Vesicolor	Virginica	All	Setosa	Vesicolor	Virginica
Sepal L.	5.843	5.006	5.936	6.588	0.828	0.352	0.516	0.636
Sepal W.	3.057	3.428	2.770	2.974	0.436	0.379	0.314	0.322
Petal L.	3.758	1.462	4.260	5.552	1.765	0.174	0.470	0.552
Petal W.	1.199	0.246	1.326	2.026	0.762	0.105	0.198	0.275

For each Iris species, we compute different correlation measures for all pairs of variables. Since variations of variables are quite different, the affine equivariant version of symmetric gini correlation estimator $\hat{\rho}_G$ is used. For each pair of variables X and Y, we also calculate Pearson correlation, Kendall's τ and two regular gini correlation estimators, denoted as $\hat{\gamma}_{1,2}$ ($\hat{\gamma}(X,Y)$) and $\hat{\gamma}_{2,1}$ ($\hat{\gamma}(Y,X)$). All correlation estimators are listed in Table 5.

From Table 5, we observe that comparing with other two species, Iris Setosa has high correlation between sepal length and sepal width, but has low correlation between sepal length and petal length. Versicolor has much larger correlation between petal length and petal width than the other two species do. Virginica has the highest correlation between sepal length and petal length among the three species.

Kendall's τ correlation value is the smallest among all correlation estimators across all pairs and across all species. Two regular Gini correlation estimators are quite different especially between sepal width and petal length in Iris Virginica species. The difference is as high as 0.159. One might perform a hypothesis test on exchangeability of two variables by testing $\gamma_{1,2} = \gamma_{2,1}$ (Schechtman, Yitzhaki and Artsev, 2007). The p-value of the test is 0.0113, which serves as a strong evidence to reject exchangeability of two variables sepal width and petal length in Iris Virginica. We also observe that $\hat{\rho}_G$ and $\hat{\rho}_p$ tend to have a same pattern across variable pairs and across species. For example, for all six pairs of variables in Iris Setosa, $\hat{\rho}_G$ is large or small whenever $\hat{\rho}_p$ is large or small. In other words, the correlation ranking across variable pairs provided by the Pearson correlation is the same as the ranking by the proposed symmetric Gini correlation. However, such a pattern is not shared by any two correlations from $\hat{\rho}_G$, $\hat{\tau}$, $\hat{\gamma}_{1,2}$ and $\hat{\gamma}_{2,1}$. Also, values of $\hat{\rho}_G$ are larger than values of $\hat{\rho}_p$ in most cases.

TABLE 5: Pearson correlation, Kendal's τ, Affine equivariant symmetric Gini correlation and Regular Gini correlations of variables for Iris data set.

Species	Correlations	Sepal L. & Sepal W.	Sepal L. & Petal L.	Sepal L. & Petal W.	Sepal W. & Petal L.	Sepal W. & Petal W.	Petal L. & Petal W.
Setosa	$\hat{\rho}_P$	0.743	0.267	0.278	0.178	0.233	0.332
	$\hat{\tau}$	0.597	0.217	0.231	0.143	0.234	0.222
	$\hat{\rho}_G$	0.742	0.274	0.285	0.182	0.256	0.312
	$\hat{\gamma}_{1,2}$	0.759	0.283	0.261	0.211	0.214	0.280
	$\hat{\gamma}_{2,1}$	0.781	0.295	0.358	0.174	0.350	0.384
Versicolor	$\hat{\rho}_P$	0.526	0.754	0.546	0.561	0.664	0.787
	$\hat{\tau}$	0.398	0.567	0.403	0.430	0.551	0.646
	$\hat{\rho}_G$	0.546	0.756	0.551	0.584	0.687	0.790
	$\hat{\gamma}_{1,2}$	0.533	0.744	0.542	0.580	0.658	0.787
	$\hat{\gamma}_{2,1}$	0.523	0.766	0.559	0.572	0.682	0.809
Virginica	$\hat{\rho}_P$	0.457	0.864	0.281	0.401	0.538	0.322
	$\hat{\tau}$	0.307	0.670	0.219	0.291	0.419	0.271
	$\hat{\rho}_G$	0.687	0.820	0.455	0.621	0.623	0.519
	$\hat{\gamma}_{1,2}$	0.406	0.867	0.278	0.467	0.567	0.304
	$\hat{\gamma}_{2,1}$	0.476	0.832	0.315	0.308	0.548	0.355

7. CONCLUSION

In this paper we propose symmetrized Gini correlation ρ_g and study its properties. The relationship between ρ_g and ρ is established when the scatter matrix, Σ, is homogeneous. The affine invariant version ρ_G is also proposed to deal with the case when Σ is heterogeneous. Asymptotic normality of the proposed estimators are established. The influence function reveals that ρ_g is more robust than the Pearson correlation while it is less robust than the Kendall's τ correlation. Comparing with the Pearson correlation estimator, the regular Gini correlation estimator and the Kendall's τ estimator of ρ, the proposed estimators balance well between efficiency and robustness and provide an attractive option for measuring correlation. Numerical studies demonstrate that the proposed estimators have satisfactory performance under a variety of situations. In particular, the symmetric Gini estimators are more efficient than the regular Gini estimators. This can be explained by the fact that the multivariate spatial rank used in the symmetrized Gini correlations takes more dependence information than the marginal ranks in the traditional ones.

We comment that the symmetric Gini correlation ρ_g is not limited to elliptical distributions. Theorems 1, 3 and 4 hold for any bivariate distribution with a finite first moment. Under elliptical distributions, the linear correlation parameter ρ is well defined and all the four estimators are Fisher consistent. Hence their asymptotical variances are comparable and can be used for evaluating relative asymptotic efficiency among the estimators.

The proposed symmetric Gini correlation has some disadvantages. Although its formulation is natural, the symmetric Gini loses an intuitive interpretation. It is more difficult to compute than the Pearson correlation, especially when X and Y are heterogeneous. In that case, an iterative scheme is required to obtain the affine invariant version of symmetric Gini correlation. When applying the proposed measure, one may consider the trade-off among efficiency, robustness, computation and interpretability.

BIBLIOGRAPHY

Blitz, R.C. and Brittain, J.A. (1964). An extension of the Lorenz diagram to the correlation of two variables. *Metron* **XXIII** (1-4) 137-143.

Borovskikh, Y.V. (1996). *U-statistics in Banach spaces*, VSP, Utrecht.

Croux, C. and Dehon, C. (2010). The influence function of the Spearman and Kendall correlation measures. *Stat. Methods Appl.* **19** (4) 497-515.

Devlin, S.J., Gnanadesikan, R. and Kettering, J.R. (1975). Robust estimation and outlier detection with correlation coefficients. *Biometrika* **62** 531-545.

Dürre, A., Vogel, D. and Fried, R. (2015). Spatial sign correlation. *J. Multivariate Anal.* **135** 89-105.

Fang, K.T., Kotz, S. and Hg, K.W. (1987). *Symmetric Multivariate and Related Distributions*, Chapman & Hall, London.

Hampel, F.R. (1974). The influence curve and its role in robust estimation. *J. Amer. Statist. Assoc.* **69** 383-393.

Hampel, F.R., Ronchetti, E. M., Rousseeuw, P. J. and Stahel, W. J. (1986). *Robust Statistics: The Approach Based on Influence Functions*. Wiley, New York.

Kendall, M.G. and Gibbons, J.D. (1990). *Rank Correlation Methods*, 5th edition, Griffin, London.

Lerman, R.I. and Yitzhaki, K. (1984). A note on the calculation and interpretation of the Gini index. *Econom. Lett.* **15** 363-368.

Lindskog, F., Mcneil, A. and Schmock, U. (2003). Kendall's tau for elliptical distributions. In Bol, et al. (Eds), *Credit Risk: Measurement, Evaluation and Management*, 149-156, Springer-Verlag, Heidelberg.

Mari, D.D. and Kotz, S. (2001). *Correlation and Dependence*, Imperial College Press, London.

Oja, H. (2010). *Multivariate Nonparametric Methods with R: An Approach Based on Spatial Signs and Ranks*. Springer

Schechtman, E. and Yitzhaki, S. (1987). A measure of association based on Gini's mean difference. *Comm. Statist. Theory Methods* **16** (1) 207-231.

Schechtman, E. and Yitzhaki, S. (1999). On the proper bounds of the Gini correlation. *Econom. Lett.* **63** 133-138.

Schechtman, E. and Yitzhaki, S. (2003). A Family of Correlation Coefficients Based on the Extended Gini Index. *J. Econ. Inequal.* **1** (2) 129-146.

Schechtman, E., Yizhaki, S. and Artsev, Y. (2007). The similarity between mean-variance and mean-Gini: Testing for equality of Gini correlations. *Advances in Investment Analysis and Portfolio Management (AIAPM)*, **3**, 103-128.

Serfling, R. (1980). *Approximation Theorems of Mathematical Statistics*. Wiley, New York.

Shevlyakov G.L. and Smirnov P.O. (2011). Robust Estimation of the Correlation Coefficient: An Attempt of Survey. *Aust. J. Stat.* **40** 147-156.

Stuart, A. (1954). The correlation between variate-values and ranks in samples from a continuous distribution. *Brit. J. Statist. Psych.* **7** 37-44.

Visuri, S., Koivunen, V. and Oja, H. (2000). Sign and rank covariance matrices. *J. Stat. Plan. Infer.* **91** 557-575.

Witting, H. and Müller-Funk, U. (1995). *Mathematische Statistik II*, B.G. Teubner, Stuttgart.

Xu, W., Huang, Y.S., Niranjan, M. and Shen, M. (2010). Asymptotic mean and variance of Gini correlation for bivariate normal samples. *IEEE Trans. Signal Process.* **58** (2) 522-534.

Dang, X., Sang, H. and Weatherall, L. (2015). Gini Covariance Matrix and its Affine Equivariant Version. Submitted.

APPENDIX

Proof of Theorem 1. The first property is obvious. Hölder's inequality implies $\left| \mathbb{E} \frac{(X_1-X_2)(Y_1-Y_2)}{\|\mathbf{Z}_1-\mathbf{Z}_2\|} \right| \leq \sqrt{ \mathbb{E} \frac{(X_1-X_2)^2}{\|\mathbf{Z}_1-\mathbf{Z}_2\|} \mathbb{E} \frac{(Y_1-Y_2)^2}{\|\mathbf{Z}_1-\mathbf{Z}_2\|} }$ and hence $|\rho_g(X,Y)| \leq 1$.

Let (X_i, Y_i), $i = 1, 2$, be independent copies of (X, Y), then $\text{cov}_g(X, Y) = 2\mathbb{E}\frac{X_1(Y_1 - Y_2)}{\|\mathbf{Z}_1 - \mathbf{Z}_2\|} = 2\mathbb{E}\frac{X_1 Y_1}{\|\mathbf{Z}_1 - \mathbf{Z}_2\|} - 2\mathbb{E}\frac{X_1 Y_2}{\|\mathbf{Z}_1 - \mathbf{Z}_2\|} = 0$ by symmetry. Hence $\rho_g(X, Y) = 0$. If $Y = aX + b$, then

$$\rho_g(X, Y) = \frac{\mathbb{E}\frac{(X_1 - X_2)(Y_1 - Y_2)}{\|\mathbf{Z}_1 - \mathbf{Z}_2\|}}{\sqrt{\mathbb{E}\frac{(X_1 - X_2)^2}{\|\mathbf{Z}_1 - \mathbf{Z}_2\|}\mathbb{E}\frac{(Y_1 - Y_2)^2}{\|\mathbf{Z}_1 - \mathbf{Z}_2\|}}} = \frac{\frac{a}{\sqrt{a^2 + 1}}\mathbb{E}|X_1 - X_2|}{\frac{|a|}{\sqrt{a^2 + 1}}\mathbb{E}|X_1 - X_2|} = \text{sgn}(a).$$

$\rho_g(aX + b, aY + d) = \rho_g(X, Y)$ can be obtained from

$$\text{cov}_g(aX + b, aY + d) = \mathbb{E}\frac{a^2(X_1 - X_2)(Y_1 - Y_2)}{|a|\|\mathbf{Z}_1 - \mathbf{Z}_2\|} = |a|\text{cov}_g(X, Y),$$

$$\text{cov}_g(aX + b, aX + b) = |a|\text{cov}_g(X, X)$$

and

$$\text{cov}_g(aY + d, aY + d) = |a|\text{cov}_g(Y, Y).$$

By (3), (4) and (5), it is easy to see the rest part of property 5. $\qquad\square$

Proof of Theorem 2. To prove the theorem, we need a result from Dang *et al.* (2015). They consider the Gini covariance matrix $\mathbf{\Sigma}_g = 2\mathbb{E}\mathbf{Z}\mathbf{r}^T(\mathbf{Z})$. Their Theorem 2.1 states that if the scatter matrix $\mathbf{\Sigma}$ has the spectral decomposition $V\Lambda V^T$ with $\Lambda = \text{diag}(\lambda_1, \lambda_2)$, then $\mathbf{\Sigma}_g = V\Lambda_g V^T$ with $\Lambda_g = \text{diag}(\lambda_{g,1}, \lambda_{g,2})$ and

$$\lambda_{g,i} = c(H)\mathbb{E}\left[\frac{\lambda_i u_i^2}{\sqrt{\lambda_1 u_1^2 + \lambda_2 u_2^2}}\right], \quad i = 1, 2 \qquad (A1)$$

where $\mathbf{u} = (u_1, u_2)^T$ is uniformly distributed on the unit circle, λ_i's are eigenvalues of $\mathbf{\Sigma}$ and $c(H)$ is a constant depending on the distribution H. Here the eigenvalues of $\mathbf{\Sigma}$ are $\lambda_1 = \sigma^2(1 - \rho)$ and $\lambda_2 = \sigma^2(1 + \rho)$, and the corresponding eigenvectors are $(1, -1)^T$ and $(1, 1)^T$. Consequently, $\rho_g = \frac{\lambda_{g,2} - \lambda_{g,1}}{\lambda_{g,2} + \lambda_{g,1}}$. Obviously, if $\rho = \pm 1$, either λ_1 or λ_2 is zero. With (A1), we have $\rho_g = \pm 1 = \rho$. If $\rho = 0$, then $\lambda_1 = \lambda_2$, and hence we have $\lambda_{g,1} = \lambda_{g,2}$ and $\rho_g = 0 = \rho$. When $|\rho| < 1$ and $\rho \neq 0$, let $u_1 = \cos\theta$ and $u_2 = \sin\theta$, then θ is uniformly distributed in $[0, 2\pi]$. With (A1), we have

$$\rho_g = \frac{\int_0^{2\pi}\frac{1}{2\pi}\frac{(1-\rho)\cos^2\theta - (1+\rho)\sin^2\theta}{\sqrt{(1-\rho)\cos^2\theta + (1+\rho)\sin^2\theta}}d\theta}{\int_0^{2\pi}\frac{1}{2\pi}\sqrt{(1-\rho)\cos^2\theta + (1+\rho)\sin^2\theta}d\theta} = \frac{\int_0^{\pi/2}\frac{\rho - \cos 2\theta}{\sqrt{1 - \rho\cos 2\theta}}d\theta}{\int_0^{\pi/2}\sqrt{1 - \rho\cos 2\theta}d\theta}$$

$$= \frac{1}{\rho} + \frac{\rho - 1}{\rho}\frac{EK(2\rho/(\rho + 1))}{EE(2\rho/(\rho + 1))}.$$

$\qquad\square$

Proof of Theorem 3. Let $\tilde{H} = (1 - \varepsilon)H + \varepsilon\delta_{(x,y)}$, then

$$T_1(\tilde{H}) = 2\iint\frac{x_1(x_1 - x_2)}{\sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}}d\tilde{H}(x_2, y_2)d\tilde{H}(x_1, y_1)$$

$$= 2(1 - \varepsilon)^2 T_1(H) + 2\varepsilon(1 - \varepsilon)\int\frac{(x - x_2)^2}{\sqrt{(x - x_2)^2 + (y - y_2)^2}}dH(x_2, y_2).$$

We have

$$L_1(x,y) = 2\int \frac{(x-x_2)^2}{\sqrt{(x-x_2)^2+(y-y_2)^2}} dH(x_2,y_2) - 4T_1(H),$$

$$L_2(x,y) = 2\int \frac{(x-x_2)(y-y_2)}{\sqrt{(x-x_2)^2+(y-y_2)^2}} dH(x_2,y_2) - 4T_2(H),$$

$$L_3(x,y) = 2\int \frac{(y-y_2)^2}{\sqrt{(x-x_2)^2+(y-y_2)^2}} dH(x_2,y_2) - 4T_3(H).$$

Hence,

$$\mathrm{IF}((x,y)^T; \rho_g, H) = \sum_{i=1}^{3} \frac{\partial h}{\partial t_i}\big|_T L_i(x,y)$$

$$= -\frac{T_2}{2\sqrt{T_1^3 T_3}} L_1(x,y) - \frac{T_2}{2\sqrt{T_1 T_3^3}} L_3(x,y) + \frac{1}{\sqrt{T_1 T_3}} L_2(x,y).$$

Replacing $T_2/\sqrt{T_1 T_3}$ with ρ_g completes the proof. $\square$

Proof of Theorem 4. Let Σ_g be the Gini covariance matrix of $Z = (X,Y)^T$ and $\hat{\Sigma}_g = \begin{pmatrix} G_x^2 & G_{xy} \\ G_{xy} & G_y^2 \end{pmatrix}$ be the sample Gini covariance matrix from sample $\{Z_i\}_{i=1}^n$. Let $vec(M)$ be the operator that stacks columns of M to form a vector. According to the Theorem 4.1 in Dang *et al.* (2015), we get

$$\sqrt{n}(vec(\hat{\Sigma}_g) - vec(\Sigma_g)) \xrightarrow{d} N_4(\mathbf{0}, \mathbf{V}),$$

where $\mathbf{V} = 4\mathbb{E}[\psi(\mathbf{Z})\psi^T(\mathbf{Z})]$, $\psi(\mathbf{z}) = vec\left(\mathbb{E}[\frac{(\mathbf{z}-\mathbf{Z})(\mathbf{z}-\mathbf{Z})^T}{\|\mathbf{z}-\mathbf{Z}\|}] - \Sigma_g\right)$. Then

$$\sqrt{n}\left\{(G_x^2, G_{xy}, G_y^2)^T - (\mathrm{cov}_g(X,X), \mathrm{cov}_g(X,Y), \mathrm{cov}_g(Y,Y))^T\right\} \xrightarrow{d} N_3(\mathbf{0}, \mathbf{V}^*)$$

with $\mathbf{V}^*$ being the matrix of $\mathbf{V}$ deleting the third row and third column. Now, since $\hat{\rho}_g = h(G_x^2, G_{xy}, G_y^2) = G_{xy}/\sqrt{G_x^2 G_y^2}$, and the derivative of h is

$$\dot{h}(a,b,c) = -b/(2\sqrt{ac})(1/a, -2/b, 1/c),$$

we have

$$\dot{h}\left(\mathrm{cov}_g(X,X), \mathrm{cov}_g(X,Y), \mathrm{cov}_g(Y,Y)\right) = \frac{-\rho_g}{2}\left(\frac{1}{\mathrm{cov}_g(X,X)}, \frac{-2}{\mathrm{cov}_g(X,Y)}, \frac{1}{\mathrm{cov}_g(Y,Y)}\right)$$

denoted as B. Applying the delta method yields the asymptotic normality of $\hat{\rho}_g$ with the asymptotic variance $v_g = B\mathbf{V}^*B^T$. Working out the explicit form of v_g completes the proof. $\square$

Proof of Theorem 6. The proof is straightforward. Let A be the diagonal matrix with the diagonal elements being a and b. Since Σ_G is affine equivariant, $\Sigma_G(A\mathbf{Z}) = A\Sigma_G(\mathbf{Z})A^T$. As a result, $\rho_G(aX, bY) = \frac{abG_{21}}{\sqrt{a^2 G_{11}}\sqrt{b^2 G_{22}}} = \mathrm{sgn}(ab)\rho_G(X,Y)$. $\square$