Communications for Statistical Applications and Methods
2016, Vol. 23, No. 5, 385–397

http://dx.doi.org/10.5351/CSAM.2016.23.5.385
Print ISSN 2287-7843 / Online ISSN 2383-4757

Estimation of $P(X > Y)$ when X and Y are dependent random variables using different bivariate sampling schemes

Hani M. Samawi[1,a], Amal Helu[b], Haresh D. Rochani[a], Jingjing Yin[a], Daniel Linder[a]

[a]Department of Biostatistics, Jiann-Ping Hsu College of Public Health, Georgia Southern University, USA; [b]Carnegie Mellon University, Qatar

Abstract

The stress-strength models have been intensively investigated in the literature in regards of estimating the reliability $\theta = P(X > Y)$ using parametric and nonparametric approaches under different sampling schemes when X and Y are independent random variables. In this paper, we consider the problem of estimating θ when (X, Y) are dependent random variables with a bivariate underlying distribution. The empirical and kernel estimates of $\theta = P(X > Y)$, based on bivariate ranked set sampling (BVRSS) are considered, when (X, Y) are paired dependent continuous random variables. The estimators obtained are compared to their counterpart, bivariate simple random sampling (BVSRS), via the bias and mean square error (MSE). We demonstrate that the suggested estimators based on BVRSS are more efficient than those based on BVSRS. A simulation study is conducted to gain insight into the performance of the proposed estimators. A real data example is provided to illustrate the process.

Keywords: bivariate simple random sampling, bivariate ranked set sampling, empirical and kernel estimation, reliability, bias, mean square error

1. Introduction

In the literature, inference about $\theta = P(X > Y)$ has been extensively studied. In the area of reliability for a system with strength X and stress Y, inference about $\theta = P(X > Y)$ as a measure of component reliability is crucial (Kotz *et al.*, 2003). In medicine, the parameter θ can be interpreted as the effectiveness of treatment Y if X and Y are the outcomes of a control and an experimental treatment, respectively (Ventura and Racugno, 2011). This quantity is also related to the Receiver Operating Characteristic (ROC) curves, where θ is interpreted as an index of accuracy (Zhou, 2008). Therefore, the estimation of $\theta = P(X > Y)$ has a wide range of applications in the literature.

This problem has been investigated from different points of views. For parametric inference, assuming X and Y have independent exponential distributions, see Enis and Geisser (1971), Awad *et al.* (1981), Tong (1974), and Johnson (1975). Moreover, Li *et al.* (1999) studied the problem of estimating $\theta = P(X > c)$ based on simple random sampling (SRS) and ranked set sampling (RSS). They showed that the estimators of $\theta = P(X > c)$ based on RSS were more efficient than those based on SRS in terms of the variances.

In many situations the parameter θ may not be available in a closed form. This makes it difficult (if at all feasible) to find a reparameterization involving θ to use any classical approaches. In particular,

[1] Corresponding author: Department of Biostatistics, Karl E. Peace Center for Biostatistics, Jiann-Ping Hsu College of Public Health, Georgia Southern University, 8015, Statesboro, GA 30460, USA. E-mail: hsamawi@georgiasouthern.edu

Published 30 September 2016 / journal homepage: http://csam.or.kr
© 2016 The Korean Statistical Society, and Korean International Statistical Society. All rights reserved.

the use of the profile likelihood may be difficult if this reparameterization is not available (Díaz-Francés and Montoya, 2013). Alternative inferential approaches that overcome this difficulty are Bayesian inference, nonparametric estimation and the use of bootstrap methods, which can be used for obtaining confidence and credible intervals for the parameter of interest (AL-Hussaini *et al.*, 1997; Baklizi and Abu-Dayyeh, 2003; Baklizi and Eidous, 2006; Rubio and Steel, 2013; Zhou, 2008).

Currently, many authors have tried to estimate θ in the case where X and Y are dependent random variables. For example Barbiero (2012) assumed that (X, Y) are jointly normally distributed; Rubio and Steel (2013) assumed that X and Y are marginally distributed as a skewed scale mixture of normal and constructed the corresponding joint distribution using a Gaussian copula; Domma and Giordano (2013) constructed the joint distribution of (X, Y) using a Farlie-Gumbel-Morgenstern copula with marginal distributions belonging to the Burr system; Domma and Giordano (2012) considered Dagum distributed marginals and constructed their joint distribution using a Frank copula; among others (Gupta *et al.*, 2013; Nadarajah, 2005). In these papers, the importance of taking the assumption of dependence between X and Y into consideration is illustrated using simulated and real data sets.

In most cases SRS is considered for estimating θ, however, some variations of RSS and SRS with concomitant variable were considered for estimating θ, see for example Sengupta and Mukhuti (2008) and Muttlak *et al.* (2010). RSS has been applied in many fields, including but not limited to agricultural, environmental studies and recently in human populations. The motivation for using RSS is that in some cases quantification (the actual measurement) of a sampling unit can be more costly than the physical acquisition of the unit. For example, as stated by Samawi and Al-Sagheer (2001), the level of bilirubin in the blood of infants can be ranked visually by observing: i) color of the face, ii) color of the chest, iii) color of lower part of the body, iv) color of terminal parts of the whole body. As the level of bilirubin in the blood increases, the yellowish discoloration goes from i) to iv) (Samawi and Al-Sagheer, 2001). Also, in some circumstances, considerable cost savings can be achieved if the number of measured sampling units are only a small fraction of the number of available units, but all units contribute to the information content of the measured units, which is the case for RSS. RSS was first introduced by McIntyre (1952). RSS has been shown to be superior to the standard SRS for estimating some population parameters. More details about RSS and its variations are available in Kaur *et al.* (1995), Patil *et al.* (1999), Samawi *et al.* (1996) and Samawi and Muttlak (1996, 2001). However, all variations of RSS sampling methods, preformed their ranking on one and only one of the study variables.

For multiple characteristics estimation, few authors worked in this area such as Patil *et al.* (1993, 1994) and Norris *et al.* (1995). They used a bivariate ranked set sampling procedure by ranking only on one of the characteristics (X or Y). However, bivariate ranked set sampling (BVRSS) by ranking on both characteristics (X and Y) was introduced by Al-Saleh and Zheng (2002). They indicated that BVRSS procedure could easily be extended to a multivariate one. The focus of this paper is to show that the use of BVRSS will substantially improve the performance of the empirical and the kernel estimators for θ analytically and by simulation. Section 2 discusses the empirical estimator of θ and its properties, while Section 3 discusses the kernel estimator and its properties. Section 4 provides the simulation study. In Section 5 we illustrated the procedure using a real data set. Section 6 provides the final remarks.

2. Empirical estimation of $\theta = P(X > Y)$

2.1. Empirical estimation using BVSRS

First we consider the estimation of $[\theta = P(X > Y)]$ empirically, where (X, Y) is a bivariate random

variable with joint probability density function (p.d.f.) $f_{X,Y}(x, y)$. Thus

$$\theta = \int_{-\infty}^{\infty} \int_{-\infty}^{x} f_{X,Y}(x, y)\, dydx. \tag{2.1}$$

An alternative approach is to set $W = X - Y$ and then we have

$$\theta = P(W > 0) = \int_{0}^{\infty} f_W(w)\, dw = 1 - F_W(0) = S_W(0), \tag{2.2}$$

where $f_W(w)$, $F_W(0)$ and $S_W(0)$ are the density, cumulative distribution function (c.d.f.), and the survival function of W, respectively. Let $\{(X_1, Y_1), (X_2, Y_2), \ldots, (X_n, Y_n)\}$ be an independent BVSRS from $f_{X,Y}(x, y)$. The empirical estimator of θ based on these BVSRS for equation (2.1) is:

$$\hat{\theta}_{BVSRS} = \frac{\sum_{i=1}^{n} I(Y_i < X_i)}{n} \tag{2.3}$$

and for equation (2.2) is

$$\hat{\theta}_{BVSRS(W)} = \frac{\sum_{i=1}^{n} I(W_i > 0)}{n}. \tag{2.4}$$

Note that both estimators are unbiased and with variances $\theta(1 - \theta)/n$ and $S_W(0)(1 - S_W(0))/n$, respectively. Moreover, $\hat{\theta}_{BVSRS}$ and $\hat{\theta}_{BVSRS(W)}$ are strongly consistent estimators of θ, or $\hat{\theta}_{BVSRS} \xrightarrow{a.s.} \theta$ and $\hat{\theta}_{BVSRS(W)} \xrightarrow{a.s.} \theta$, and are also asymptotically normally distributed, implying that $\sqrt{n}(\hat{\theta}_{BVSRS} - \theta) \xrightarrow{d} N(0, \theta(1 - \theta))$ and $\sqrt{n}(\hat{\theta}_{BVSRS(W)} - \theta) \xrightarrow{d} N(0, S_W(0)(1 - S_W(0)))$, as n $\to \infty$, as shown by Montoya and Rubio (2014).

2.2. Empirical estimation using BVRSS

Based on Al-Saleh and Zheng (2002) description, a BVRSS can be obtained as follows: suppose (X, Y) is a bivariate random vector with the joint p.d.f. $f(x, y)$.

Step 1. A random sample of size r^4 is identified from the population and randomly allocated into r^4 pools of size r^4 each so that each pool is a square matrix with r rows and r columns.

Step 2. In the first pool, rank each set (row) by a suitable method of ranking with respect to (*w.r.t.*) the first characteristic (X). Then from each row identify the unit with the smallest rank *w.r.t.* X.

Step 3. Rank the r minima obtained in Step 2, in a similar manner but *w.r.t.* the second characteristic (Y). Then identify and measure the unit with the smallest rank *w.r.t.* Y. This pair of measurements (x, y), which is resembled by the label (1, 1), is the first element of the BVRSS sample.

Step 4. Repeat Steps 2 and 3 for the second pool, but in Step 3, the pair that corresponds to the second smallest rank *w.r.t.* the second characteristic (Y) is chosen for actual measurement (quantification). This pair is resembled by the label (1, 2).

Step 5. The process continues until the label (r, r) is resembled from the r^2-th (last) pool.

The above procedure produces a quantified BVRSS of size r^2. The procedure can be repeated m times to obtain a sample of size $n = mr^2$. In sampling notation, assume that a random sample of size mr^4 is identified (no measurements were taken) from a bivariate probability density function, say $f_{X,Y}(x,y) : (x,y) \in R^2$, with means μ_x and μ_y, variances σ_x^2 and σ_y^2 and correlation coefficient ρ. Following the Al-Saleh and Zheng (2002) definition of BVRSS, then $[(X_{[i](j)k}, Y_{(i)[j]k}),\ i = 1,2,\ldots,r;\ j = 1,2,\ldots,r;$ and $k = 1,2,\ldots,m]$ denotes the BVRSS. Now, let $f_{X_{[i](j)},Y_{(i)[j]}}(x,y)$ be the joint p.d.f. of $[(X_{[i](j)k}, Y_{(i)[j]k}),\ k = 1,2,\ldots,m]$. Al-Saleh and Zheng (2002), with $m = 1$, showed that

$$\frac{1}{r^2}\sum_{j=1}^{r}\sum_{i=1}^{r} f_{[i](j),(i)[j]}(x,y) = f_{X,Y}(x,y). \tag{2.5}$$

Then using these BVRSSs, for equation (2.1), we propose the following empirical estimators of θ:

$$\hat{\theta}_{BVRSS} = \frac{\sum_{k=1}^{m}\sum_{i=1}^{r}\sum_{j=1}^{r} I\left(X_{[i](j)k} > Y_{(i)[j]k}\right)}{n}; \quad n = mr^2 \tag{2.6}$$

and for equation (2.2)

$$\hat{\theta}_{BVRSS(W)} = \frac{\sum_{k=1}^{m}\sum_{i=1}^{r}\sum_{j=1}^{r} I(W_{i,j,k} > 0)}{n}, \tag{2.7}$$

where $W_{i,j,k} = X_{[i](j)k} - Y_{(i)[j]k}$.

Using equation (2.5) we have the following results.

Theorem 1.

(a) $\hat{\theta}_{BVRSS}$ *and* $\hat{\theta}_{BVRSS(W)}$ *are unbiased estimators of* θ.

(b)

$$Var\left(\hat{\theta}_{BVRSS}\right) = \frac{1}{n}\left\{\theta(1-\theta) - \frac{\sum_{i=1}^{r}\sum_{j=1}^{r}(\theta_{ij}-\theta)^2}{r^2}\right\}$$

and

$$Var\left(\hat{\theta}_{BVRSS}(W)\right) = \frac{1}{n}\left\{S_W(0)(1-S_W(0)) - \frac{\sum_{i=1}^{r}\sum_{j=1}^{r}\left(S_{W(i,j)}(0) - S_W(0)\right)^2}{r^2}\right\},$$

where $\theta_{ij} = P(X_{[i](j)} > Y_{(i)[j]})$, *and* $S_{W(i,j)(0)} = P(X_{[i](j)} - Y_{(i)[j]} > 0)$, $i = 1,2,\ldots,r,\ j = 1,2,\ldots,r$.

Proof: The proof of Theorem 1 is straightforward and it is omitted from the paper. □

Note that it is clear that the empirical estimate based on BVRSS has smaller variance than using BVSRS for estimating θ.

Theorem 2. *For fixed r and as* $m \to \infty$, *and hence* $n \to \infty$, *we have*

(a) $\hat{\theta}_{BVRSS}$ *and* $\hat{\theta}_{BVRSS}(W)$ *are strongly consist estimators of* θ, *or* $\hat{\theta}_{BVRSS} \xrightarrow{a.s.} \theta$ *and* $\hat{\theta}_{BVRSS}(W) \xrightarrow{a.s.} \theta$.

(b) $\sqrt{n}(\hat{\theta}_{BVRSS} - \theta) \xrightarrow{d} N(0,G)$ *and* $\sqrt{n}(\hat{\theta}_{BVRSS}(W) - \theta) \xrightarrow{d} N(0,G_S)$, *where* $G = \theta(1-\theta) - \{\sum_{i=1}^{r}\sum_{j=1}^{r}(\theta_{ij}-\theta)^2\}/r^2$ *and* $G = S_W(0)(1-S_W(0)) - \{\sum_{i=1}^{r}\sum_{j=1}^{r}(S_{W(i,j)}(0) - S_W(0))^2\}/r^2$.

Proof: The proof follows by the law of large number and the central limit theorem. □

3. Kernel estimation of $\theta = P(X > Y)$

3.1. Kernel estimation using BVSRS

Montoya and Rubio (2014) proposed nonparametric kernel estimators for θ from equations (2.1) and (2.2) as follows: Let H be a symmetric, positive definite, 2×2 bandwidth matrix and k_2 be a two-dimensional kernel function (Parzen, 1962). Define $k_H(\mathbf{t}) = (\det H)^{-1/2} k_2\left(H^{1/2}\mathbf{t}\right)$, $\mathbf{t} \in \mathbb{R}^2$, where, $\mathbf{t} = \{t_1 = x - X, t_2 = y - Y\}$. Using kernel density estimation, then we will have the kernel estimations for θ in equations (2.1) and (2.2) are defined as

$$\begin{aligned}\hat{\theta}_{K(BVSRS)} &= \int_{-\infty}^{\infty}\int_{-\infty}^{x} \frac{1}{n}\sum_{i=1}^{n} k_H(x - X_i, y - Y_i)\, dydx \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{t_1} \frac{1}{n}\hat{f}_{BVSRS}(t_1, t_2)\, dt_1 dt_2 \end{aligned} \tag{3.1}$$

and

$$\begin{aligned}\hat{\theta}_{K(BVSRS(W))} &= \frac{1}{nh}\sum_{i=1}^{n}\int_{0}^{\infty} k_1\left(\frac{w - W_i}{h}\right) dw \\ &= 1 - \frac{1}{n}\sum_{i=1}^{n} K_1\left(\frac{W_i}{h}\right) \\ &= \int_{0}^{\infty} \hat{f}_{BVSRS(W)}(w)\, dw \\ &= 1 - \hat{F}_{BVSRS}(0), \end{aligned} \tag{3.2}$$

where, $K_1(w/h) = (1/h)\int_0^{\infty} k_1(w/h)dw$, k_1 is one-dimensional kernel function with a bandwidth $h > 0$. One of the most common choices of kernel functions for one or two-dimensional kernels are the univariate and the bivariate normal densities, respectively. For the choice of the bandwidth matrix h and H we refer to Montoya and Rubio (2014) for full discussion. Moreover, they showed that, under some regularity conditions, $\hat{\theta}_{K(BVSRS)} \xrightarrow{P} \theta$ (weakly consistent estimator) and $\hat{\theta}_{K(BVSRS(W))} \xrightarrow{a.s.} \theta$ (strong consistent estimator).

3.2. Kernel estimation using BVRSS

Again, let $[(X_{[i](j)k}, Y_{(i)[j]k}),\ i = 1, 2, \ldots, r;\ j = 1, 2, \ldots, r;$ and $k = 1, 2, \ldots, m]$ be a BVRSS from (X, Y) with $f(x, y)$. Similarly, define $k_H(\mathbf{t}) = (\det H)^{-1/2} k_2(H^{1/2}\mathbf{t}), \mathbf{t} \in \mathbb{R}^2$. Then using the BVRSS samples, we propose the following kernel estimators of θ:

$$\begin{aligned}\hat{\theta}_{K(BVRSS)} &= \frac{\sum_{k=1}^{m}\sum_{i=1}^{r}\sum_{j=1}^{r}\int_{-\infty}^{\infty}\int_{-\infty}^{x} k_H\left(x - X_{[i](j)k}, y - Y_{(i)[j]k}\right) dydx}{n} \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{x} \hat{f}_{BVRSS}(t_1, t_2)\, dt_1 dt_2; \quad n = mr^2, \end{aligned} \tag{3.3}$$

for equation (2.1) and

$$\begin{aligned}\hat{\theta}_{K(BVRSS(W))} &= \frac{\sum_{k=1}^{m}\sum_{i=1}^{r}\sum_{j=1}^{r}\int_0^{\infty} K_1\left(\frac{w-W_{ij}}{h}\right)dw}{nh} \\ &= 1-\frac{1}{n}\sum_{k=1}^{m}\sum_{i=1}^{r}\sum_{j=1}^{r}K_1\left(\frac{w-W_{ij}}{h}\right) \\ &= \int_0^{\infty}\hat{f}_{BVRSS(W)}(w)dw \\ &= 1-\hat{F}_{BVRSS}\,(0) \end{aligned} \tag{3.4}$$

for equation (2.2), where $W_{ijk} = X_{[i](j)k} - Y_{(i)[j]k}$.

Theorem 3.

(a) $E\left(\hat{\theta}_{K(BVRSS)}\right) = E\left(\hat{\theta}_{K(BVSRS)}\right), E\left(\hat{\theta}_{K(BVRSS(W))}\right) = E\left(\hat{\theta}_{K(BVSRS(W))}\right),$

(b) $Var\left(\hat{\theta}_{K(BVRSS)}\right) = \left[Var\left(\hat{\theta}_{K(BVSRS)}\right) - \frac{1}{nr^2}\sum_{i=1}^{r}\sum_{j=1}^{r}\left(E_{ij}-E\right)^2\right] = V_1,$

(c) $Var\left(\hat{\theta}_{K(BVRSS(W))}\right) = \left[Var\left(\hat{\theta}_{K(BVSRS(W))}\right) - \frac{1}{nr^2}\sum_{i=1}^{r}\sum_{j=1}^{r}\left(D_{ij}-D\right)^2\right] = V_2,$

where,

$$E_{ij} = E_{X_{[i](j)},Y_{(i)[j]}}\left[\int_{-\infty}^{\infty}\int_{-\infty}^{u} K_H\left(u-X_{[i](j)}, v-Y_{(i)[j]}\right)dvdu\right], \quad E = E_{X,Y}\left[\int_{-\infty}^{\infty}\int_{-\infty}^{u} K_H\left(u-X, v-Y\right)dvdu\right],$$

$$D_{ij} = E_{W_{ij}}\left[\int_0^{\infty}\frac{1}{h}k_1\left(\frac{w-W_{ij}}{h}\right)dw\right], \quad \textit{and} \quad D = E_W\left[\int_0^{\infty}\frac{1}{h}k_1\left(\frac{w-W}{h}\right)dw\right].$$

Proof:

$$\begin{aligned}\text{(a)}\ E\left(\hat{\theta}_{K(BVRSS)}\right) &= E\left(\frac{\sum_{k=1}^{m}\sum_{i=1}^{r}\sum_{j=1}^{r}\int_{-\infty}^{\infty}\int_{-\infty}^{u} K_H\left(u-X_{[i](j)k}, v-Y_{(i)[j]k}\right)dvdu}{n}\right) \\ &= E\left(\frac{\sum_{i=1}^{r}\sum_{j=1}^{r}\int_{-\infty}^{\infty}\int_{-\infty}^{u} K_H\left(u-X_{[i](j)k}, v-Y_{(i)[j]k}\right)dvdu}{r^2}\right) \\ &= \frac{1}{r^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left[\sum_{i=1}^{r}\sum_{j=1}^{r}\int_{-\infty}^{\infty}\int_{-\infty}^{u} K_H\left(u-X_{[i](j)k}, v-Y_{(i)[j]k}\right)dvdu\right] f_{[i](j),(i)[j]}(x,y)dxdy \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty}\int_{-\infty}^{u} K_H\left(u-x, v-y\right)dudv\right]\frac{1}{r^2}\sum_{i=1}^{r}\sum_{j=1}^{r} f_{[i](j),(i)[j]}(x,y)dxdy, \end{aligned}$$

then by using (2.5), we have

$$\begin{aligned} E\left(\hat{\theta}_{K(BVRSS)}\right) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty}\int_{-\infty}^{u} K_H\left(u-x, v-y\right)dudv\right) f(x,y)dxdy \\ &= E\left(\hat{\theta}_{K(BVSRS)}\right) \end{aligned}$$

Similarly, we can show that $E(\hat{\theta}_{K(BVRSS(W))}) = E(\hat{\theta}_{K(BVSRS(W))})$.

(b) $\mathrm{Var}\left(\hat{\theta}_{K(BVRSS)}\right)$

$$= \mathrm{Var}\left[\frac{1}{mr^2}\sum_{k=1}^{m}\sum_{i=1}^{r}\sum_{j=1}^{r}\int_{-\infty}^{\infty}\int_{-\infty}^{u} K_H\left(u - X_{[i](j)k}, v - Y_{(i)[j]k}\right) dvdu\right]$$

$$= \frac{1}{mr^4}\mathrm{Var}\left[\sum_{i=1}^{r}\sum_{j=1}^{r}\int_{-\infty}^{\infty}\int_{-\infty}^{u} K_H\left(u - X_{[i](j)k}, v - Y_{(i)[j]k}\right) dvdu\right]$$

$$= \frac{1}{mr^4}\left(\sum_{i=1}^{r}\sum_{j=1}^{r}\mathrm{Var}\left[\int_{-\infty}^{\infty}\int_{-\infty}^{u} K_H\left(u - X_{[i](j)k}, v - Y_{(i)[j]k}\right) dvdu\right]\right)$$

$$= \frac{1}{mr^4}\sum_{i=1}^{r}\sum_{j=1}^{r}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty}\int_{-\infty}^{u} K_H\left(u - X_{[i](j)k}, v - Y_{(i)[j]k}\right) dudv - E_{ij}\right]^2 f_{[i](j),(i)[j]}(x,y)dxdy,$$

where,

$$E_{ij} = E_{X_{[i](j)},Y_{(i)[j]}}\left(\int_{-\infty}^{\infty}\int_{-\infty}^{u} K_H\left(u - X_{[i](j)}, v - Y_{(i)[j]}\right) dudv\right)$$

then,

$$\mathrm{Var}\left(\hat{\theta}_{K(BVRSS)}\right)$$

$$= \frac{1}{mr^4}\left(\sum_{i=1}^{r}\sum_{j=1}^{r}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty}\int_{-\infty}^{u} K_H\left(u - X_{[i](j)k}, v - Y_{(i)[j]k}\right) dudv - E_{ij} \pm E\right]^2 f_{[i](j),(i)[j]}(x,y)dxdy\right)$$

$$= \frac{1}{mr^4}\sum_{i=1}^{r}\sum_{j=1}^{r}\left\{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty}\int_{-\infty}^{u} K_H\left(u - X_{[i](j)k}, v - Y_{(i)[j]k}\right) dudv - E\right]^2 f_{[i](j),(i)[j]}(x,y)dxdy\right.$$

$$- 2\left(E_{ij} - E\right)\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty}\int_{-\infty}^{u} K_H\left(u - X_{[i](j)k}, v - Y_{(i)[j]k}\right) dudv - E\right] f_{[i](j),(i)[j]}(x,y)dxdy$$

$$\left. + \left(E_{ij} - E\right)^2\right\}$$

$$= \frac{1}{mr^4}\left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty}\int_{-\infty}^{u} K_H\left(u - x, v - y\right) dudv - E\right]^2 \frac{1}{r^2}\sum_{i=1}^{r}\sum_{j=1}^{r} f_{[i](j),(i)[j]}(x,y)\, dxdy\right)$$

$$- \frac{1}{mr^4}\sum_{i=1}^{r}\sum_{j=1}^{r}\left(E_{ij} - E\right)^2,$$

where, $E = E_{X,Y}(\int_{-\infty}^{\infty}\int_{-\infty}^{u} K_H\left(u - X, v - Y\right) dudv)$.

Using (2.5) again, we have

$$\mathrm{Var}\left(\hat{\theta}_{K(BVRSS)}\right) = \left[\mathrm{Var}\left(\hat{\theta}_{K(BVSRS)}\right) - \left(\frac{1}{mr^4}\right)\sum_{i=1}^{r}\sum_{j=1}^{r}\left(E_{ij} - E\right)^2\right].$$

(c) Can be proved similarly. □

As in Montoya and Rubio (2014), and under the same assumptions, we need to show that $\hat{\theta}_{K(BVSRS)}$ and $\hat{\theta}_{K(BVSRS(w))}$ consistent estimators for θ.

Theorem 4. *Assume that the two-dimensional kernel function k_2 is bounded on $\mathbb{R}^2$ and the one-dimensional kernel function k_1 is bounded on $\mathbb{R}$ with respect to L_2 and L_1 the distances*

$$L_2(u) = \sup_{\|\mathbf{t}\| \geq u} k_2(\mathbf{t}) \quad \text{and} \quad L_1(u) = \sup_{|\mathbf{t}| \geq u} k_1(\mathbf{t}),$$

for $u \geq 0$. Let the bandwidth satisfy $h > 0$, $h \longrightarrow 0$, $n = mr^2$, $nh^2 \longrightarrow \infty$ as $n \longrightarrow \infty$, and hence as $m \longrightarrow \infty$. For the bivariate case, define the bandwidth matrix $H = diag(h)$. Assuming that the same bandwidth is used in both $\hat{f}_{BVRSS}(x, y)$ and $\hat{f}_{BVSRS}(x, y)$ and in both $\hat{f}_{BVRSS(W)}(w)$ and $\hat{f}_{BVSRS(W)}(w)$, then if one of the following conditions stated by Montoya and Rubio (2014), Silverman (1986) and Wand and Jones (1995), holds

1. *$| t | k_1(t) \longrightarrow 0$ as $| t | \longrightarrow \infty$; $\| \mathbf{t} \|^2 k_2(\mathbf{t}) \longrightarrow 0$ as $\| \mathbf{t} \| \longrightarrow \infty$ and $f_{X,Y}$ and f_W are almost surely continuous.*
2. *f_W is bounded; $f_{X,Y}$ is bounded.*
3. *$\int_0^\infty L_1(u)\, du < \infty$; $\int_0^\infty u L_2(u)\, du < \infty$.*

We have $\hat{\theta}_{K(BVRSS)} \xrightarrow{p} \theta$ and $\hat{\theta}_{K(BVRSS(W))} \xrightarrow{p} \theta$.

Proof: First, we note that

$$\mathrm{MSE}\left(\hat{\theta}_{K(BVRSS)}\right) = E\left[\left|\hat{\theta}_{K(BVRSS)} - \theta\right|^2\right] = \mathrm{Bias}\left(\hat{\theta}_{K(BVRSS)}\right)^2 + \mathrm{Var}\left(\hat{\theta}_{K(BVRSS)}\right)$$

and

$$\mathrm{MSE}\left(\hat{\theta}_{K(BVRSS)(W)}\right) = E\left[\left|\hat{\theta}_{K(BVRSS)(W)} - \theta\right|^2\right] = \mathrm{Bias}\left(\hat{\theta}_{K(BVRSS)(W)}\right)^2 + \mathrm{Var}\left(\hat{\theta}_{K(BVRSS)(W)}\right).$$

However, by Theorem 3 we have

$$\mathrm{Bias}\left(\hat{\theta}_{K(BVRSS)}\right) = \mathrm{Bias}\left(\hat{\theta}_{K(BVSRS)}\right), \quad \mathrm{Bias}\left(\hat{\theta}_{K(BVRSS)(W)}\right) = \mathrm{Bias}\left(\hat{\theta}_{K(BVSRS)(W)}\right),$$

and

$$\mathrm{Var}\left(\hat{\theta}_{K(BVRSS)}\right) < \mathrm{Var}\left(\hat{\theta}_{K(BVSRS)}\right), \quad \mathrm{Var}\left(\hat{\theta}_{K(BVRSS)(W)}\right) < \mathrm{Var}\left(\hat{\theta}_{K(BVSRS)(W)}\right).$$

Montoya and Rubio (2014) showed that,

$$\mathrm{MSE}\left(\hat{\theta}_{K(BVRSS)}\right) \longrightarrow 0 \quad \text{and} \quad \mathrm{MSE}\left(\hat{\theta}_{K(BVRSS(W))}\right) \longrightarrow 0$$

and then

$$\hat{\theta}_{K(BVSRS)} \xrightarrow{p} \theta \quad \text{and} \quad \hat{\theta}_{K(BVSRS(W))} \xrightarrow{p} \theta.$$

However,

$$\mathrm{MSE}\left(\hat{\theta}_{K(BVRSS)}\right) < \mathrm{MSE}\left(\hat{\theta}_{K(BVSRS)}\right) \longrightarrow 0$$

and

$$\mathrm{MSE}\left(\hat{\theta}_{K(BVRSS(W))}\right) < \mathrm{MSE}\left(\hat{\theta}_{K(BVSRS(W))}\right) \longrightarrow 0.$$

Therefore,

$$\hat{\theta}_{K(BVRSS)} \xrightarrow{p} \theta \quad \text{and} \quad \hat{\theta}_{K(BVRSS(W))} \xrightarrow{p} \theta.$$

Moreover, by the central limit theorem, we have

$$\sqrt{n}\left(\hat{\theta}_{K(BVRSS)} - \theta\right) \xrightarrow{d} N\left(0, nV_1\right)$$

and

$$\sqrt{n}\left(\hat{\theta}_{K(BVRSS(W))} - \theta\right) \xrightarrow{d} N\left(0, nV_2\right).$$

□

4. Simulation studies

We conduct a computer simulations to gain insight into the efficiency of estimating θ. Placketts class of bivariate distribution with fixed marginal distribution functions $F(x)$ and $G(x)$ are used to investigate the performance of the proposed estimators. The Placketts joint c.d.f is given by

$$H(x,y) = \begin{cases} \dfrac{S(x,y) - \left[S^2(x,y) - 4\psi(\psi-1)F(x)G(y)\right]^{\frac{1}{2}}}{2(\psi-1)}, & \text{if } \psi \neq 1, \\ F(x)G(y), & \text{if } \psi = 1, \end{cases}$$

where $S(x,y) = 1 + (\psi - 1)\left[F(x) + G(y)\right]$ and the parameter ψ governs the dependence between X and Y. The reason for choosing this class of bivariate distributions is that it covers the full range of dependency. For example, in case of $U(0,1)$ marginal distributions, we have the following:

(a) $\psi \to 0 \Rightarrow X = 1 - Y$, (b) $\psi = 1 \Rightarrow X$ and Y are independent, (c) $\psi \to \infty \Rightarrow X = Y$.

For more detailed description of Placketts distribution and its random generation, see Johnson (1987).

Four types of dependencies from strongly negative to strongly positive corresponding to $\psi = 0.1, 1.0, 2.0, 10.0$ and two marginal distributions, exponential with mean $(\mu) = 0.5$ or 1.0, and gamma with scale parameter $\beta = 1$ or 2 and shape parameter $\lambda = 3$ are considered. The performance of the estimators of θ is investigated for $r = 2, 3, 4$ and $m = 20$, therefore the sample sizes used are $n = 80, 180$ and 320. However, we present the results of only set sizes $r = 3$ and 4 to reduce the number of tables in our simulation, since the other cases provide similar results. Using 5,000 replications, we estimate the bias and mean square errors (MSE) for the estimators of θ. The bound on the error of estimation (using 95% confidence level) is approximately ±0.013. The relative efficiency of using BVRSS relative to using BVSRS for estimating θ is defined by REF = MSE(BVSRS)/MSE (BVRSS).

Our simulation indicates that using BVRSS for estimating θ is at least more efficient for all presented cases in Tables 1 and 2. The relative efficiency is ranging from 1.02 to 2.28 depending on the strength and the direction of the dependency between X and Y and the underlying marginal distributions. However, increasing the set size r increases the efficiency. However, increasing the cycle size m has no effect on the efficiency of the proposed estimators but it decreases the absolute bias of the

Table 1: Performance of empirical and kernel estimators for two exponential marginal distribution and $m = 20$

r	Parameters				Empirical estimation			Kernal estimation				
	μ_x	μ_y	ψ	θ	BVSRS variance	BVRSS variance	REF	BVSRS |Bias|	BVSRS MSE	BVRSS |Bias|	BVRSS MSE	REF
3	1.0	0.5	0.1	0.627	0.0013	0.0009	1.44	0.0056	0.0008	0.0054	0.0005	1.60
3	1.0	0.5	1.0	0.667	0.0012	0.0008	1.50	0.0166	0.0010	0.0166	0.0008	1.25
3	1.0	0.5	2.0	0.695	0.0012	0.0009	1.33	0.0275	0.0015	0.0269	0.0014	1.07
3	1.0	0.5	10.0	0.797	0.0009	0.0006	1.50	0.0793	0.0074	0.0779	0.0071	1.04
3	1.0	1.0	0.1	0.500	0.0015	0.0009	1.67	0.0011	0.0009	0.0012	0.0005	1.80
3	1.0	1.0	1.0	0.500	0.0014	0.0009	1.56	0.0028	0.0009	0.0031	0.0005	1.80
3	1.0	1.0	2.0	0.500	0.0013	0.0009	1.44	0.0104	0.0010	0.0087	0.0006	1.67
3	1.0	1.0	10.0	0.500	0.0013	0.0008	1.63	0.0855	0.0091	0.0855	0.0090	1.01
4	1.0	0.5	0.1	0.627	0.0007	0.0004	1.75	0.0060	0.0005	0.0057	0.0003	1.67
4	1.0	0.5	1.0	0.667	0.0007	0.0004	1.75	0.0131	0.0007	0.0134	0.0005	1.40
4	1.0	0.5	2.0	0.695	0.0007	0.0004	1.75	0.0199	0.0010	0.0202	0.0009	1.11
4	1.0	0.5	10.0	0.797	0.0005	0.0003	1.67	0.0666	0.0052	0.0657	0.0051	1.02
4	1.0	1.0	0.1	0.500	0.0008	0.0004	2.00	0.0003	0.0005	0.0000	0.0002	2.50
4	1.0	1.0	1.0	0.500	0.0008	0.0004	2.00	0.0038	0.0005	0.0036	0.0003	1.67
4	1.0	1.0	2.0	0.500	0.0008	0.0005	1.60	0.0123	0.0007	0.0124	0.0005	1.40
4	1.0	1.0	10.0	0.500	0.0008	0.0005	1.60	0.0904	0.0092	0.0908	0.0091	1.01

BVRSS = bivariate ranked set sampling; BVSRS = bivariate simple random sampling; MSE = mean square error.

Table 2: Performance of empirical and kernel estimators for two gamma marginal distributions with ($\lambda_x = 3$ and $\lambda_y = 3$) and $m = 20$

r	Parameters				Empirical estimation			Kernal estimation				
	μ_x	μ_y	ψ	θ	BVSRS variance	BVRSS variance	REF	BVSRS |Bias|	BVSRS MSE	BVRSS |Bias|	BVRSS MSE	REF
3	2.0	1.0	0.1	0.730	0.0011	0.0007	1.57	0.0099	0.0009	0.0093	0.0005	1.80
3	2.0	1.0	1.0	0.790	0.0009	0.0006	1.50	0.0132	0.0008	0.0129	0.0006	1.33
3	2.0	1.0	2.0	0.826	0.0008	0.0004	2.00	0.0183	0.0009	0.0185	0.0007	1.29
3	2.0	1.0	10.0	0.918	0.0004	0.0004	1.00	0.0251	0.0010	0.0251	0.0009	1.11
3	1.0	1.0	0.1	0.500	0.0013	0.0010	1.30	0.0001	0.0010	0.0004	0.0006	1.67
3	1.0	1.0	1.0	0.500	0.0014	0.0010	1.40	0.0000	0.0010	0.0010	0.0006	1.67
3	1.0	1.0	2.0	0.500	0.0014	0.0009	1.56	0.0007	0.0009	0.0011	0.0005	1.80
3	1.0	1.0	10.0	0.500	0.0014	0.0009	1.56	0.0007	0.0009	0.0002	0.0005	1.80
4	2.0	1.0	0.1	0.730	0.0006	0.0004	1.50	0.0072	0.0004	0.0073	0.0003	1.33
4	2.0	1.0	1.0	0.790	0.0006	0.0004	1.50	0.0121	0.0005	0.0115	0.0003	1.67
4	2.0	1.0	2.0	0.826	0.0005	0.0003	1.67	0.0155	0.0006	0.0154	0.0005	1.20
4	2.0	1.0	10.0	0.918	0.0002	0.0002	1.00	0.0211	0.0006	0.0212	0.0006	1.00
4	1.0	1.0	0.1	0.500	0.0008	0.0005	1.60	0.0013	0.0006	0.0005	0.0003	2.00
4	1.0	1.0	1.0	0.500	0.0008	0.0005	1.60	0.0014	0.0005	0.0007	0.0003	1.67
4	1.0	1.0	2.0	0.500	0.0008	0.0005	1.60	0.0007	0.0005	0.0006	0.0003	1.67
4	1.0	1.0	10.0	0.500	0.0007	0.0004	1.75	0.0005	0.0005	0.0007	0.0002	2.50

BVRSS = bivariate ranked set sampling; BVSRS = bivariate simple random sampling; MSE = mean square error.

kernel estimators so we omitted the tables with $m = 10$.

5. Real data analysis

In order to illustrate estimation of $\theta = P(X > Y)$ under two different sampling schemes, i.e., BVSRS and BVRSS, the China Health and Nutrition Survey (CHNS) data set is used (Yan *et al.*, 2012). During the last several years, the clinicians have realized the importance of the lipid-transporting apolipoproteins, such as apoA and apoB which transport high-density lipoprotein (HDL, good) cholesterol and low-density lipoprotein (LDL, bad) cholesterol particles, respectively (Walldius *et al.*, 2004). It is

Table 3: Kernel smoothed and empirical estimates of $\theta = P(X > Y)$ under two sampling schemes (BVSRS vs. BVRSS)

	Population θ ($N = 10187$)	$\hat{\theta}_{BVSRS}$ ($n = 60$)	$\widehat{\mathrm{Var}}\left(\hat{\theta}_{BVSRS}\right)$	$\hat{\theta}_{BVSRS}$ ($n = 60$)	$\widehat{\mathrm{Var}}\left(\hat{\theta}_{BVSRS}\right)$
Kernal	0.7678	0.7572	0.0022	0.7553	0.0017
Empirical	0.7651	0.7661	0.0031	0.7650	0.0024

The estimates and variances are calculated based on 500 bootstrap samples.
BVRSS = bivariate ranked set sampling; BVSRS = bivariate simple random sampling.

expected that healthier individual should have larger apoA values than apoB, so they have less risk for cardiovascular disease. These apolipoproteins can be applied as alternative biomarkers to the traditional LDL and HDL biomarkers, which are sometimes more advantageous.

For example, compared to the traditional biomarker LDL-C and HDL-C, taking the measurement of apoB/apoA-I ratio does not require fasting, and the measurement of apoB and apoA-I are standardized and easy to compare across studies (Walldius and Jungner, 2006). Instead of the ratio between apoA and apoB, alternatively, we can consider the probability of apoA being greater than apoB, where both were from the same individual (i.e., $\theta = P(\text{apoA} > \text{apoB})$). If this probability is significantly larger than 0.5, then we can say apoA is stochastically larger than apoB for the study population, thus concluding the study population is relatively at low risk of cardiovascular disease.

The data set contains apoA and apoB biomarker values taken from 10,187 Chinese children and adults (aged ≥ 7) in year 2009. By assuming the existing full data set as the study population, we would want to collect a paired sub-sample of size 60 and each pair contains apoA and apoB values from the same individual. Henceforth, two sub-samples are drawn based on BVSRS and BVRSS. The kernel and empirical estimates and corresponding variances calculated by 500 resamples for $P(\text{apoA} > \text{apoB})$ under the two sampling schemes are listed in Table 3. Note that the estimates of the full data set are considered as the true value for the population parameter. For this data set, we observed that both the kernel and the empirical estimates under the two sampling schemes are very similar and are close to the true values, while BVRSS yields smaller estimated bootstrap variances. Table 3 shows that the estimated $P(\text{apoA} > \text{apoB})$ is about 0.76 which is larger than 0.5. Therefore, we can claim that the Chinese people (aged ≥ 7) are relatively at low risk of cardiovascular disease.

6. Final remarks and conclusions

The interest of drawing inferences about $\theta = P(X > Y)$ arises naturally in many areas of research, including but not limited to the reliability for a system with strength X and stress Y, to medicine when X and Y are the outcomes of a control and an experimental treatment where, the parameter θ can be interpreted as the effectiveness of the treatment Y, and this quantity is also related to the Receiver Operating Characteristic (ROC) curves, where θ is interpreted as an index of accuracy. Therefore, it is of interest to find a sampling strategy which provides more structured and representative samples to provide more efficient and reliable estimates of θ.

This paper shows that empirical and kernel estimates of θ based on BVRSS and BVSRS are equivalent in terms of bias, with both methods have small biases. As expected, bias improves as the sample size increases. However, using BVRSS is more efficient than using BVSRS in the case of empirical and kernel estimation. If the ranking cost of the two variables is negligible compared with that associated with taking their exact measurements, using BVRSS will result in reducing the number of subjects in the study and hence the overall cost of the study.

References

Al-Hussaini EK, Mousa MA, and Sultan KS (1997). Parametric and nonparametric simulation of $P(Y < X)$ for finite mixtures of lognormal components, *Communications in Statistics-Theory and Methods*, **26**, 1269–1289.

Al-Saleh MF and Zheng G (2002). Estimation of bivariate characteristics using ranked set sampling, *Australia and New Zealand Journal of Statistics*, **44**, 221–232.

Awad AM, Azzam MM, and Hamdan MA (1981). Some inference results on $\Pr(Y < X)$ in the bivariate exponential model, *Communications in Statistics-Theory and Methods*, **10**, 2515–2525.

Baklizi A and Abu-Dayyeh W (2003). Shrinkage estimation of $P(Y < X)$ in the exponential case, *Communications in Statistics-Simulation and Computation*, **32**, 31–42.

Baklizi A and Eidous O (2006). Nonparametric estimation of $P(X < Y)$ using kernel methods, *Metron*, **64**, 47–60.

Barbiero A (2012). Interval estimators for reliability: the bivariate normal case, *Journal of Applied Statistics*, **39**, 501–512.

Domma F and Giordano S (2012). A stress-strength model with dependent variables to measure household financial fragility, *Statistical Methods & Applications*, **21**, 375–389.

Domma F and Giordano S (2013). A copula based approach to account for dependence in stress-strength models, *Statistical Papers*, **54**, 807–826.

Díaz-Francés E and Montoya JA (2013). The simplicity of likelihood based inferences for $P(X < Y)$ and for the ratio of means in the exponential model, *Statistical Papers*, **54**, 499–522.

Enis P and Geisser S (1971). Estimation of the probability that $Y < X$, *Journal of the American Statistical Association*, **66**, 162–168.

Gupta RC, Ghitany ME, and Al-Mutairi DK (2013). Estimation of reliability from a bivariate log normal data, *Journal of Statistical Computation and Simulation*, **83**, 1068–1081.

Johnson ME (1987). *Multivariate Statistical Simulation*, John Wiley & Sons, New York.

Johnson NL (1975). Letter to the editor, *Technometrics*, **17**, 393.

Kaur A, Patil GP, Sinha AK, and Taillie C (1995). Ranked set sampling: an annotated bibliography, *Environmental and Ecological Statistics*, **2**, 25–54.

Kotz S, Lumelskii S, and Pensky M (2003). *The Stress-Strength Model and Its Generalizations: Theory and Applications*, World Scientific Publishing, Singapore.

Li D, Sinha BK, and Chuiv NN (1999). On estimation of $P(X > c)$ based on a ranked set sample. In UJ Dixit and MR Satam (Eds), *Statistical Inference and Design of Experiments* (pp. 47–54), Alpha Science International, Oxford, UK.

McIntyre GA (1952). A method for unbiased selective sampling, using ranked set, *Australian Journal of Agricultural Research*, **3**, 385–390.

Montoya JA and Rubio FJ (2014). Nonparametric inference for $P(X < Y)$ with paired variables, *Journal of Data Science*, **12**, 359–375.

Muttlak HA, Abu-Dayyeh WA, Saleh MF, and Al-Sawi E (2010). Estimating $P(Y < X)$ using ranked set sampling in case of the exponential distribution, *Communications in Statistics-Theory and Methods*, **39**, 1855–1868.

Nadarajah S (2005). Reliability for some bivariate beta distributions, *Mathematical Problems in Engineering*, **2005**, 101–111.

Norris RC, Patil GP, and Sinha AK (1995). Estimation of multiple characteristics by ranked set sampling methods, *Coenoses*, **10**, 95–111.

Parzen E (1962). On estimation of a probability density function and mode, *Annals of Mathematical Statistics*, **33**, 1065–1076.

Patil GP, Sinha AK, and Taillie C (1993). Relative precision of ranked set sampling: a comparison with the regression estimator, *Environmetrics*, **4**, 399–412.

Patil GP, Sinha AK, and Taillie C (1994). Ranked set sampling for multiple characteristics, *International Journal of Ecology and Environmental Sciences*, **20**, 94–109.

Patil GP, Sinha AK, and Taillie C (1999). Ranked set sampling: a bibliography, *Environmental and Ecological Statistics*, **6**, 91–98.

Rubio FJ and Steel MFJ (2013). Bayesian inference for $P(X < Y)$ using asymmetric dependent distributions, *Bayesian Analysis*, **8**, 43–62.

Samawi HM, Ahmed MS, and Abu-Dayyeh W (1996). Estimating the population mean using extreme ranked set sampling, *Biometrical Journal*, **38**, 577–586.

Samawi HM and Al-Sagheer OA (2001). On the estimation of the distribution function using extreme and median ranked set sampling, *Biometrical Journal*, **43**, 357–373.

Samawi HM and Muttlak HA (1996). Estimation of ratio using ranked set sampling, *Biometrical Journal*, **38**, 753–764.

Samawi HM and Muttlak HA (2001). On ratio estimation using median ranked set sampling, *Journal of Applied Statistical Science*, **10**, 89–98.

Sengupta S and Mukhuti S (2008). Unbiased estimation of $P(X > Y)$ for exponential populations using order statistics with application in ranked set sampling, *Communications in Statistics-Theory and Methods*, **37**, 898–916.

Silverman BW (1986). *Density Estimation for Statistics and Data Analysis*, **26**, CRC Press, New York.

Tong H (1974). A note on the estimation of $\Pr(Y < X)$ in the exponential case, *Technometrics*, **16**, 625.

Ventura L and Racugno W (2011). Recent advances on Bayesian inference for $P(X < Y)$, *Bayesian Analysis*, **6**, 411–428.

Walldius G and Jungner I (2006). The apoB/apoA-I ratio: a strong, new risk factor for cardiovascular disease and a target for lipid-lowering therapy: a review of the evidence, *Journal of Internal Medicine*, **259**, 493–519.

Walldius G, Jungner I, Aastveit AH, Holme I, Furberg CD, and Sniderman AD (2004). The apoB/apoA-I ratio is better than the cholesterol ratios to estimate the balance between plasma proatherogenic and antiatherogenic lipoproteins and to predict coronary risk, *Clinical Chemical Laboratory Medicine*, **42**, 1355–1363.

Wand MP and Jones MC (1995). *Kernel Smoothing*, Chapman and Hall, London.

Yan S, Li J, Li S, Zhang B, Du S, Gordon-Larsen P, Adair L, and Popkin B (2012). The expanding burden of cardiometabolic risk in China: the China Health and Nutrition Survey, *Obesity Reviews*, **13**, 810–821.

Zhou W (2008). Statistical inference for $P(X < Y)$, *Statistics in Medicine*, **27**, 257–279.

Received May 14, 2016; Revised July 2, 2016; Accepted August 26, 2016