

Bounds on multivariate Kendall's tau and Spearman's
rho for zero-inflated continuous variables and their
application to insurance

Mhamed Mesfioui

Département de mathématiques et d'informatique
Université du Québec à Trois-Rivières
Trois-Rivières (Québec)
Canada G9A 5H7

Julien Trufin

Department of Mathematics
Université Libre de Bruxelles (ULB)
Bruxelles, Belgium

April 26, 2021

Abstract

In this note, we derive upper bounds on Kendall's tau and Spearman's rho for multivariate zero-inflated continuous variables often encountered in insurance. A lower bound for Spearman's rho is also established in the bivariate case. These bounds are easy to compute and can be estimated from a data set of zero-inflated random vectors as illustrated in this note with a motor insurance portfolio.

Keywords: Kendall's tau; Spearman's rho; multivariate zero-inflated data; upper bounds; insurance.

1 Introduction and motivation

Zero-inflated distributions are often encountered in insurance. They combine a continuous distribution on the positive real line and a point mass at zero. **In the literature, these distributions are also referred to as zero-augmented distributions or zero-adjusted distributions.** In insurance studies, the total claim amount related to a given contract is often equal to 0, when no claims have been filed against the insurer, but may also be strictly positive when one or several accidents occurred. The total claim cost is thus decomposed into the product of an indicator for the event “the policy produces at least one claim during the reference period” and a positive random variable representing the total claim amount produced by the policy when at least one claim has been filed.

There are many applications in insurance where we aim to measure the dependency between zero-inflated distributions. We can mention the need to assess the dependency between annual claim amounts related to different guarantees or products owned by the same policyholder or between annual claim amounts related to policyholders belonging to the same household as discussed for instance in Pechon et al. (2020).

Kendall’s tau and Spearman’s rho are classical measures of association between distributions. However, when at least one of the distributions is non-continuous, such as in the case of zero-inflated distributions, we know from Nešlehová (2007) that Kendall’s tau and Spearman’s rho are restricted to sub-intervals of $[-1, 1]$. Therefore, small positive values of Kendall’s tau or Spearman’s rho may actually support a strong positive dependence when getting close to its highest attainable value.

Denuit and Mesfioui (2017) derived lower and upper bounds on Kendall’s tau for bivariate zero-inflated distributions proposed by Pimentel et al. (2015). Mesfioui and Quessy (2010) generalized the notion of Kendall’s tau and Spearman’s rho to the multivariate non-continuous case and established upper bounds for their proposed versions of Kendall’s tau and Spearman’s rho.

In this note, we rely on the results of Mesfioui and Quessy (2010) to derive upper bounds for Kendall’s tau and Spearman’s rho in the case of multivariate zero-inflated distributions. Note that in the bivariate case, we recover the upper bound for Kendall’s tau established in Denuit and Mesfioui (2017) and we also derive a lower bound for Spearman’s rho. We illustrate the use of these upper bounds when assessing the dependency between the annual claim costs of three guarantees owned by the same policyholder in car insurance. In particular, we show that the bounds derived in this note may change the conclusions of the analyst when interpreting the values of Kendall’s tau and Spearman’s rho.

This note is organized as follows. In Section 2, we recall the notion of multivariate concordance and the population versions of Kendall’s tau and Spearman’s rho for non-continuous distributions proposed in Mesfioui and Quessy (2010). Sections 3 and 4 are respectively devoted to the derivation of upper bounds on Kendall’s tau and Spearman’s rho for multivariate zero-inflated distributions. In Section 4, we also establish a lower bound for Spearman’s rho in the bivariate case. The estimation from possibly multivariate zero-inflated observations is considered in Section 5. The final section deals with a numerical illustration performed on real data coming from a motor insurance portfolio of European policies with three guarantees.

2 Concordance for non-continuous random vectors

Let us consider the Fréchet space $\Gamma(F_1, \dots, F_d)$ of all d -variate distribution functions with marginal distributions F_1, \dots, F_d and let $\mathbf{X}_1 = (X_{11}, \dots, X_{1d})$ and $\mathbf{X}_2 = (X_{21}, \dots, X_{2d})$ be continuous random vectors with distribution functions $H_1, H_2 \in \Gamma(F_1, \dots, F_d)$. According to Nelsen (2002), the multivariate concordance among \mathbf{X}_1 and \mathbf{X}_2 is defined as the probability of concordance among all $d(d-1)/2$ possible pairs, namely,

$$Q(H_1, H_2) = P(\mathbf{X}_1 < \mathbf{X}_2) + P(\mathbf{X}_2 < \mathbf{X}_1)$$

where the above inequalities are taken componentwise.

It is clear that for non-continuous d -random vectors $\mathbf{X}_1 = (X_{11}, \dots, X_{1d})$ and $\mathbf{X}_2 = (X_{21}, \dots, X_{2d})$, the multivariate concordance may involve ties such as

$$P(X_{1i_1} = X_{2i_1}, X_{1i_2} = X_{2i_2}, \dots, X_{1i_s} = X_{2i_s}) > 0, \quad \{i_1, \dots, i_s\} \subset I_d = \{1, \dots, d\}, 1 \leq s \leq d.$$

To take into account the presence of ties in the expression of the concordance among \mathbf{X}_1 and \mathbf{X}_2 , we consider the operators T_1, \dots, T_d defined by

$$T_\ell P(\mathbf{X}_1 < \mathbf{X}_2) = P(X_{11} < X_{21}, \dots, X_{1\ell} \leq X_{2\ell}, \dots, X_{1d} < X_{2d}).$$

Let $S \subseteq I_d$, $\mathcal{T}_S = \prod_{\ell \in S} T_\ell$ and $\mathbf{T} = 2^{-d} \prod_{\ell=1}^d (T_\ell + 1)$. According to Mesfioui and Quessy (2010), the multivariate concordance for non-continuous distributions is defined as the mean of $\mathcal{T}_S Q(H_1, H_2)$ for all $S \subseteq I_d$, that is,

$$\tilde{Q}(H_1, H_2) = \frac{1}{2^d} \sum_{S \subseteq I_d} \mathcal{T}_S Q(H_1, H_2) = \mathbf{T} Q(H_1, H_2). \quad (2.1)$$

Hence, the population versions of Kendall's tau and Spearman's rho for a non-continuous d -random vector \mathbf{X} with distribution $H \in \Gamma(F_1, \dots, F_d)$ is

$$\tilde{\tau}_d(H) = \frac{2^{d-1} \tilde{Q}(H, H) - 1}{2^{d-1} - 1} \quad (2.2)$$

and

$$\tilde{\rho}_d(H) = \frac{d+1}{2^d - d - 1} \left\{ 2^{d-1} \tilde{Q}(\Pi_d, H) - 1 \right\} \quad (2.3)$$

where $\Pi_d(x_1, \dots, x_d) = \prod_{i=1}^d F_i(x_i)$. Notice that for $d = 2$ the previous expressions of Kendall's tau and Spearman's rho coincide with those examined by Denuit and Lambert (2005), Mesfioui and Tajar (2006), Niewiadomska-Bugaj and Kowalczyk (2005) and Nešlehová (2007).

In the continuous case, formulas (2.2) and (2.3) reduce to the expressions of Kendall's tau and Spearman's rho provided in Joe (1990). Also, in this case, Schmid and Schmidt (2007) established and studied other multivariate extensions of Spearman's rho and Taylor (2007) derived conditions that multivariate measures of concordance must satisfy.

In this note, we first provide an upper bound of Kendall's tau for zero-inflated continuous d -random vectors. When $d = 2$, we recover the upper bound given in Proposition 4.1 in

Denuit and Mesfioui (2017). An upper bound of Spearman's rho for zero-inflated continuous d -random vectors is also derived. Specifically, let $\mathbf{X} = (X_1, \dots, X_d)$ be a random vector such that X_ℓ is a random variable having a continuous distribution on \mathbb{R}^+ with distribution F_{h_ℓ} and a positive probability mass at 0, $P[X_\ell = 0] = p_\ell > 0$, $\ell = 1, \dots, d$. Recall that the distribution of X_ℓ is then given by

$$F_\ell(s) = p_\ell \mathbb{I}(s = 0) + \{p_\ell + (1 - p_\ell)F_{h_\ell}(s)\} \mathbb{I}(s > 0) \quad (2.4)$$

where F_{h_ℓ} denotes the distribution of the random variable $X_\ell | X_\ell > 0$. For such distribution, we denote $F_\ell \sim \text{ZID}(p_\ell, F_{h_\ell})$. Also, we denote by $\Gamma_{\text{ZID}}(F_1, \dots, F_d)$ the set of all d -variate distribution functions with marginal distributions $F_\ell \sim \text{ZID}(p_\ell, F_{h_\ell})$, $\ell = 1, \dots, d$.

3 Upper bound on Kendall's tau in $\Gamma_{\text{ZID}}(F_1, \dots, F_d)$

The following proposition provides an upper bound for Kendall's tau for zero-inflated continuous d -random vectors.

Proposition 3.1. *An upper bound of Kendall's tau in the space $\Gamma_{\text{ZID}}(F_1, \dots, F_d)$ is given by*

$$\tau_{d,\max} = 1 - \frac{2^{d-1}}{2^{d-1} - 1} \sum_{\ell=2}^d \left(\frac{1}{2}\right)^{d-\ell+1} p_{(\ell)}^2 \quad (3.1)$$

where $p_{(1)} \leq \dots \leq p_{(d)}$.

Proof. From Proposition 5 in Mesfioui and Quessy (2010), one has

$$\tau_{d,\max} = \frac{2^d}{2^{d-1} - 1} \sum_{\ell=1}^{d-1} \left(\frac{1}{2}\right)^\ell \mathbb{D}_{(\ell)} \quad (3.2)$$

where for all $\ell = 1, \dots, d$, $\mathbb{D}_{(\ell)} = \mathbb{E}(F_\ell^-(X_\ell))$ and $\mathbb{D}_{(1)} \leq \dots \leq \mathbb{D}_{(d)}$. Let $X_\ell \sim \text{ZID}(p_\ell, F_{h_\ell})$, it results from (2.4) that

$$F_\ell(s^-) = \{p_\ell + (1 - p_\ell)F_{h_\ell}(s)\} \mathbb{I}(s > 0). \quad (3.3)$$

Therefore,

$$\begin{aligned} \mathbb{E}(F_\ell^-(X_\ell)) &= \mathbb{E}[\{p_\ell + (1 - p_\ell)F_{h_\ell}(X_\ell)\} \mathbb{I}(X_\ell > 0)] \\ &= p_\ell(1 - p_\ell) + (1 - p_\ell) \mathbb{E}[F_{h_\ell}(X_\ell) \mathbb{I}(X_\ell > 0)] \\ &= p_\ell(1 - p_\ell) + (1 - p_\ell)^2 \mathbb{E}[F_{h_\ell}(X_\ell) | X_\ell > 0] \\ &= p_\ell(1 - p_\ell) + \frac{(1 - p_\ell)^2}{2} = \frac{1 - p_\ell^2}{2} \end{aligned}$$

which implies that $\mathbb{E}(F_\ell^-(X_\ell))$ is a decreasing function of p_ℓ . Consequently,

$$\mathbb{D}_{(\ell)} = \frac{1}{2} \{1 - p_{(d-\ell+1)}^2\}, \quad \ell = 1, \dots, d. \quad (3.4)$$

The result (3.1) is then deduced by inserting (3.4) in (3.2). \square

Note that for $d = 2$, we obtain $\tau_{d,\max} = 1 - p_{(2)}^2 = 1 - \max(p_1^2, p_2^2)$ which is exactly the upper bound derived in Denuit and Mesfioui (2017). For $d = 3$, the upper bound is

$$\tau_{d,\max} = 1 - \frac{1}{3}p_{(2)}^2 - \frac{2}{3}p_{(3)}^2 \quad \text{with} \quad p_{(1)} \leq p_{(2)} \leq p_{(3)}.$$

4 Upper bound on Spearman's rho in $\Gamma_{ZID}(F_1, \dots, F_d)$

The next proposition gives an upper bound for Spearman's rho for zero-inflated continuous d -random vectors.

Proposition 4.1. *An upper bound for Spearman's rho in the space $\Gamma_{ZID}(F_1, \dots, F_d)$ is given by*

$$\rho_{d,\max} = \frac{d+1}{2^d - d - 1} \sum_{A \in \mathcal{F}} \prod_{\ell \in A} \left\{ \frac{2(1+|A|)p_\ell(1-p_\ell)^{|A|} - (2p_\ell - 1)^{|A|+1} + 1}{2(1+|A|)} \right\}^{\frac{1}{|A|}} \quad (4.1)$$

where $\mathcal{F} = \{A \subset I_d : \text{such that } |A| \text{ is even and } |A| > 0\}$ and $|A|$ denotes the cardinal of the set A .

Proof. The general expression of the upper bound $\rho_{d,\max}$ is described in Proposition 6 in Mesfioui and Quesy (2010) as follows

$$\rho_{d,\max} = \frac{d+1}{2^d - d - 1} \sum_{A \in \mathcal{F}} \prod_{\ell \in A} \left\{ \mathbb{E} \left(\xi_\ell^{|A|} \right) \right\}^{\frac{1}{|A|}} \quad (4.2)$$

where $\xi_\ell = F_\ell(X_\ell^-) + F_\ell(X_\ell) - 1$. Let $X_\ell \sim \text{ZID}(p_\ell, F_{h_\ell})$, one observes from (2.4) and (3.3) that

$$\xi_\ell = p_\ell \mathbb{I}(X_\ell = 0) + 2\{p_\ell + (1-p_\ell)F_{h_\ell}(X_\ell)\} \mathbb{I}(X_\ell > 0) - 1.$$

It results that for $A \subset I_d$ such that $|A|$ is even,

$$\begin{aligned} \mathbb{E} \left(\xi_\ell^{|A|} \right) &= p_\ell(p_\ell - 1)^{|A|} + (1-p_\ell) \mathbb{E} \left[(2\{p_\ell + (1-p_\ell)F_{h_\ell}(X_\ell)\} - 1)^{|A|} \mid X_\ell > 0 \right] \\ &= p_\ell(p_\ell - 1)^{|A|} + (1-p_\ell) \mathbb{E} \left\{ [2(1-p_\ell)U + 2p_\ell - 1]^{|A|} \right\} \end{aligned}$$

where U is a random variable uniformly distributed over $[0, 1]$. Therefore,

$$\begin{aligned} \mathbb{E} \left\{ [2(1-p_\ell)U + 2p_\ell - 1]^{|A|} \right\} &= \frac{1}{2(1+|A|)(1-p_\ell)} \left[(2(1-p_\ell)u + 2p_\ell - 1)^{|A|+1} \right]_{u=0}^{u=1} \\ &= \frac{1}{2(1+|A|)(1-p_\ell)} \left\{ 1 - (2p_\ell - 1)^{|A|+1} \right\} \end{aligned}$$

which leads to

$$\mathbb{E} \left(\xi_\ell^{|A|} \right) = \frac{2(1+|A|)p_\ell(1-p_\ell)^{|A|} - (2p_\ell - 1)^{|A|+1} + 1}{2(1+|A|)}. \quad (4.3)$$

The announced result is then obtained by inserting (4.3) in (4.2). \square

Notice that the upper bounds (3.1) and (4.1) only involve the probability masses at zero, as expected.

For $d = 2$, simple calculations show from (4.1) that

$$\rho_{d,\max} = 3\sqrt{\mathbb{E}(\xi_1^2)\mathbb{E}(\xi_2^2)} = \sqrt{(1-p_1^3)(1-p_2^3)}$$

since we have

$$\mathbb{E}(\xi_\ell^2) = \frac{6p_\ell(1-p_\ell)^2 - (2p_\ell - 1)^3 + 1}{6} = \frac{1}{3}(1-p_\ell^3), \quad \ell = 1, 2.$$

Likewise, for $d = 3$, one gets

$$\rho_{d,\max} = \frac{1}{3} \left\{ \sqrt{(1-p_1^3)(1-p_2^3)} + \sqrt{(1-p_1^3)(1-p_3^3)} + \sqrt{(1-p_2^3)(1-p_3^3)} \right\}.$$

Denuit and Mesfioui (2017) have derived the best lower bound on Kendall's tau related to zero-inflated continuous random pairs. Let us then examine a lower bound on Spearman's rho for bivariate zero-inflated distributions.

Proposition 4.2. *A lower bound on Spearman's rho in the space $\Gamma_{ZID}(F_1, F_2)$ is given by*

$$\rho_{2,\min} = -\sqrt{(1-p_1^3)(1-p_2^3)}.$$

Proof. From (2.3), we have

$$\tilde{\rho}_2(H) = 6\tilde{Q}(\Pi_2, H) - 3$$

with

$$\tilde{Q}(\Pi_2, H) = \frac{1}{2} \sum_{A \in \mathcal{F}} \mathbb{E} \left(\prod_{\ell \in A} \xi_\ell \right) + \frac{1}{2} = \frac{1}{2} \mathbb{E}(\xi_1 \xi_2) + \frac{1}{2}.$$

Hence, we get

$$\tilde{\rho}_2(H) = 3\mathbb{E}(\xi_1 \xi_2).$$

Now it suffices to notice that $\mathbb{E}(\xi_1) = \mathbb{E}(\xi_2) = 0$, which yields

$$\tilde{\rho}_2(H) = 3\text{Cov}(\xi_1 \xi_2) \geq -3\sqrt{\text{var}(\xi_1)\text{var}(\xi_2)} = -3\sqrt{\mathbb{E}(\xi_1^2)\mathbb{E}(\xi_2^2)} = -\rho_{2,\max}.$$

□

As pointed out by Mesfioui and Quessy (2009) and Genest et al. (2013), the quantity $\tilde{Q}(\Pi_d, H)$ involved in the expression (2.3) of Spearman's rho can also be rewritten as

$$\tilde{Q}(\Pi_d, H) = \frac{1}{2^d} \mathbb{E} \left\{ \prod_{\ell=1}^d (1 + \xi_\ell) + \prod_{\ell=1}^d (1 - \xi_\ell) \right\} \quad (4.4)$$

where $\xi_\ell = F_\ell(X_\ell) + F_\ell(X_\ell^-) - 1$, $\ell = 1, \dots, d$. Furthermore, if $H \in \Gamma_{ZID}(F_1, \dots, F_d)$ one sees from (2.4) and (3.3) that

$$\xi_\ell = p_\ell \mathbb{I}(X_\ell = 0) + 2\{p_\ell + (1-p_\ell)F_{h_\ell}(X_\ell)\} \mathbb{I}(X_\ell > 0) - 1.$$

As shown in the next section, the previous formulas are convenient for deriving an appropriate estimator of Spearman's rho corresponding to zero inflated random vectors.

5 Empirical version

In this section, we describe the procedure allowing to estimate the previous concordance measures as well as their bounds.

To do this, let us consider a sample $\mathbf{X}_i = (X_{i1}, \dots, X_{id})$, $i = 1, \dots, n$, from a distribution $H \in \Gamma_{ZID}(F_1, \dots, F_d)$. In continuous setting, the concordance $Q(H, H)$ is naturally estimated by

$$Q_n(H, H) = \binom{n}{2}^{-1} \sum_{i < j} \{\mathbb{I}(\mathbf{X}_i < \mathbf{X}_j) + \mathbb{I}(\mathbf{X}_j < \mathbf{X}_i)\}.$$

In contrast, when $H \in \Gamma_{ZID}(F_1, \dots, F_d)$, one can use the relation (2.1) to estimate the concordance $\tilde{Q}(H, H)$ in (2.2) by

$$\tilde{Q}_n(H, H) = \mathbf{T}Q_n(H, H)$$

where

$$T_\ell \mathbb{I}(\mathbf{X}_i < \mathbf{X}_j) = \mathbb{I}(X_{i1} < X_{j1}, \dots, X_{i\ell} \leq X_{j\ell}, \dots, X_{id} < X_{jd}), \quad \ell = 1, \dots, d.$$

To establish the empirical version of Spearman's rho, we estimate the quantity $\tilde{Q}(\Pi_d, H)$ from (4.4) by

$$\tilde{Q}_n(\Pi_d, H) = \frac{1}{2^d n} \sum_{i=1}^n \left\{ \prod_{\ell=1}^d (1 + \hat{\xi}_{i\ell}) + \prod_{\ell=1}^d (1 - \hat{\xi}_{i\ell}) \right\}$$

where for all $i = 1, \dots, n$ and $\ell = 1, \dots, d$,

$$\hat{\xi}_{i\ell} = \hat{p}_\ell \mathbb{I}(X_{i\ell} = 0) + 2 \left\{ \hat{p}_\ell + (1 - \hat{p}_\ell) \hat{F}_{h_\ell}(X_{i\ell}) \right\} \mathbb{I}(X_{i\ell} > 0) - 1$$

with $\hat{p}_\ell = n^{-1} \sum_{k=1}^n \mathbb{I}(X_{k\ell} = 0)$ and

$$\hat{F}_{h_\ell}(X_{i\ell}) = \frac{\sum_{k=1}^n \mathbb{I}(0 < X_{k\ell} \leq X_{i\ell})}{\sum_{k=1}^n \mathbb{I}(X_{k\ell} > 0)}.$$

The upper bounds described in (3.1) and (4.1) can be immediately estimated using the empirical proportions \hat{p}_ℓ , $\ell = 1, \dots, d$.

6 Numerical illustration

We apply the proposed estimators for Kendall's tau, Spearman's rho and the upper bounds (3.1) and (4.1) to a real data set. The data relate to a motor insurance portfolio of an insurance company operating in the EU with three guarantees ($d = 3$). This portfolio has been observed from 2016 to 2019 and comprises 1158 policies in 2016, 1183 in 2017, 1208 in 2018 and 1312 in 2019. Note that a policy that is in the portfolio during a given year has been observed during the whole year. The guarantees under consideration are the Third-Party Liability (TPL) insurance ($\ell = 1$), the Material Damage (MD) insurance ($\ell = 2$) and the Glass Breakage (GB) insurance ($\ell = 3$). TPL insurance is compulsory and covers a

third-party’s loss caused by the insured car. MD is an optional guarantee that covers the cost of repairing or replacing the insured’s own vehicle. Typically, this guarantee is triggered when the policyholder is liable for the claim or could not identify the liable person. Finally, GB insurance covers the glass elements of the insured car in case of cracks, blows, breakage of the windshield, side windows or panorama roof for instance. For each policy and a given year, we have at our disposal the annual claim costs by guarantee.

Table 1 summarises the results for our data set. One sees that the estimates are quite stable over the four-year observation period, so that we only comment the 2019 results. **Note that a bootstrap procedure could be applied to derive standard deviations of the correlation coefficients estimates and their upper bounds. For each observation year, one could indeed use such an approach in order to get point estimates together with standard deviations on these estimates. In this example, the four years of observations reveal results that are stable over the years.**

The 2019 estimates of the probability masses at 0 for the three guarantees are $\hat{p}_1 = 0.965701$, $\hat{p}_2 = 0.873476$ and $\hat{p}_3 = 0.950457$, and Figure 1 depicts the 2019 estimates for the distribution functions F_{h_1} , F_{h_2} and F_{h_3} . There are thus many zero values for the three guarantees, which is typical in insurance as explained in the introduction. The estimated Kendall’s tau is equal to 0.014799 whereas the upper bound obtained in this paper is equal to 0.077158. Its normalized version obtained by dividing Kendall’s tau by its corresponding upper bound is then equal to 0.191801. Also, the estimated Spearman’s rho is equal to 0.022152 and its corresponding upper bound derived in this paper is equal to 0.172607. Its normalized version is thus equal to 0.128339. To summarize, the estimated Kendall’s tau (resp. Spearman’s rho) 0.014799 (resp. 0.022152) should be replaced by its normalized version estimate 0.191801 (resp. 0.128339). The conclusion that would have been that these guarantees are nearly independent becomes (with the normalized versions) that the three guarantees are actually (significantly) positively dependent. **Note that such a dependency could partially be explained by some explanatory variables or covariates (often available in practice) influencing in the same way the three guarantees. Hence, in practice, this analysis could also be conducted conditionally to the explanatory variables that are available in order to measure the residual dependency within the different risk classes.**

Year	Kendall’s tau ($\hat{\tau}$)	$\hat{\tau}_{3,\max}$	$\hat{\tau}/\hat{\tau}_{3,\max}$	Spearman’s rho ($\hat{\rho}$)	$\hat{\rho}_{3,\max}$	$\hat{\rho}/\hat{\rho}_{3,\max}$
2016	0.016696	0.085994	0.194147	0.025016	0.184998	0.135224
2017	0.019598	0.091866	0.213333	0.029259	0.196645	0.148790
2018	0.022080	0.097889	0.225561	0.033128	0.208008	0.159265
2019	0.014799	0.077158	0.191801	0.022152	0.172607	0.128339

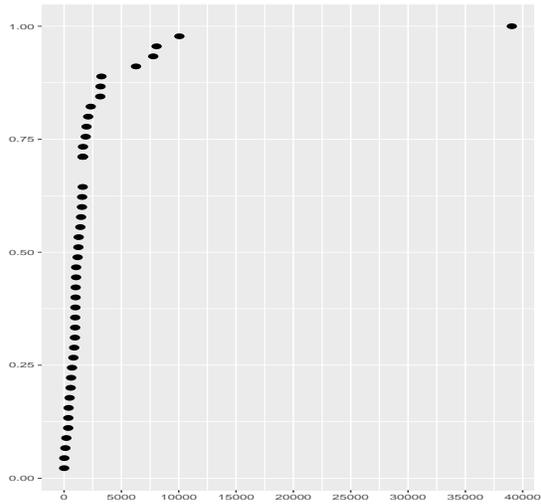
Table 1: Estimates of Kendall’s tau and Spearman’s rho together with their corresponding upper bounds.

Acknowledgements

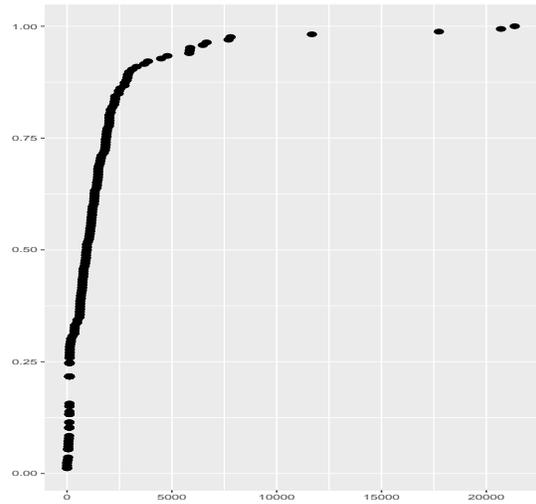
Mhamed Mesfioui acknowledges the financial support of the Natural Sciences and Engineering Research Council of Canada No 261968-2013.

References

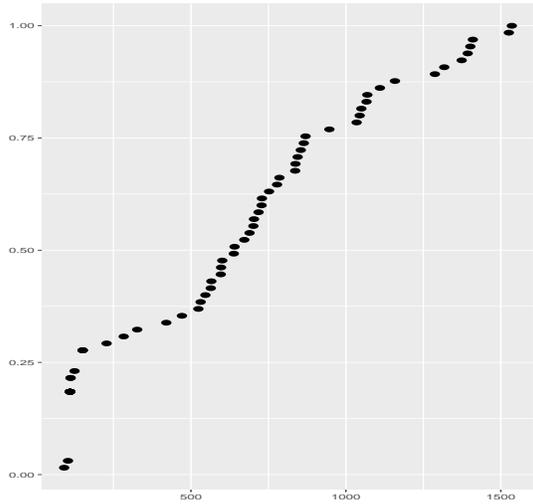
- Denuit, M., Lambert, P. (2005). Constraints on concordance measures in bivariate discrete data. *Journal of Multivariate Analysis* 93, 40-57.
- Denuit, M., Mesfioui, M. (2017). Bounds on Kendall's tau for zero-inflated continuous variables. *Statistics and Probability Letters*, 126, 173-178.
- Genest, C., Nešlehová, J. G., Rémillard, B. (2013). On the estimation of Spearman's rho and related tests of independence for possibly discontinuous multivariate data. *Journal of Multivariate Analysis*, 117, 214-228.
- Joe, H. (1990). Multivariate concordance. *Journal of multivariate analysis*, 35(1), 12-30.
- Mesfioui, M., Quessy, J.-F. (2010). Concordance measures for multivariate non-continuous random vectors. *Journal of Multivariate Analysis*, 101, 2398-2410.
- Mesfioui, M., Tajar, A. (2005). On the properties of some nonparametric concordance measures in the discrete case. *Nonparametric Statistics*, 17(5), 541-554.
- Nelsen, R.B. (2002). Concordance and copulas: a survey, in: *Distributions with Given Marginals and Statistical Modelling*, Kluwer Acad. Publ., Dordrecht, pp. 169-177.
- Nešlehová, J. (2007). On rank correlation measures for non-continuous random variables. *Journal of Multivariate Analysis*, 98(3), 544-567.
- Niewiadomska-Bugaj, M., Kowalczyk, T. (2005). On grade transformation and its implications for copulas. *Brazilian Journal of Probability and Statistics*, 125-137.
- Pechon, F., Denuit, M., Trufin, J. (2020). Home and Motor insurance joined at a household level using multivariate credibility. *Annals of Actuarial Science*, 1-33.
doi:10.1017/S1748499520000160
- Pimentel, R. S., Niewiadomska-Bugaj, M., Wang, J. C. (2015). Association of zero-inflated continuous variables. *Statistics and Probability Letters* 96, 61-67.
- Schmid, F., & Schmidt, R. (2007). Multivariate extensions of Spearman's rho and related statistics. *Statistics and Probability Letters*, 77(4), 407-416.
- Taylor, M.D. (2007). Multivariate measures of concordance, *Annals of the Institute of Statistical Mathematics* 59 (4) 789-806.



(a) TPL insurance: $\hat{F}_{h_1}(X_{i1})$.



(b) MD insurance: $\hat{F}_{h_2}(X_{i2})$.



(c) GB insurance: $\hat{F}_{h_3}(X_{i3})$.

Figure 1: 2019 estimates of the distribution functions F_{h_ℓ} for $\ell = 1, 2, 3$. X-axis: $X_{i\ell} | X_{i\ell} > 0$, Y-axis: $\hat{F}_{h_\ell}(X_{i\ell})$.