Probability in the Engineering and Informational Sciences, **28**, 2014, 401–417.
doi:10.1017/S0269964814000072

BIVARIATE EXTENSIONS OF SKELLAM'S DISTRIBUTION

Christian Genest
Department of Mathematics and Statistics, McGill University, 805, rue Sherbrooke ouest, Montréal (Québec) Canada H3A 0B9
E-mail: cgenest@math.mcgill.ca

Mhamed Mesfioui
Département de mathématiques et d'informatique, Université du Québec à Trois-Rivières, 3351, boulevard des Forges, Trois-Rivières (Québec) Canada G9A 5H7
E-mail: Mhamed.Mesfioui@uqtr.ca

Skellam's name is traditionally attached to the distribution of the difference of two independent Poisson random variables. Many bivariate extensions of this distribution are possible, e.g., through copulas. In this paper, the authors focus on a probabilistic construction in which two Skellam random variables are affected by a common shock. Two different bivariate extensions of the Skellam distribution stem from this construction, depending on whether the shock follows a Poisson or a Skellam distribution. The models are nested, easy to interpret, and yield positive quadrant-dependent distributions, which share the convolution closure property of the univariate Skellam distribution. The models can also be adapted readily to account for negative dependence. Closed form expressions for Pearson's correlation between the components make it simple to estimate the parameters via the method of moments. More complex formulas for Kendall's tau and Spearman's rho are also provided.

1. INTRODUCTION

Let Y_1, Y_2 be two independent Poisson random variables with $\mathrm{E}(Y_1) = \lambda_1 > 0$ and $\mathrm{E}(Y_2) = \lambda_2 > 0$. The difference $X = Y_1 - Y_2$ is then said to have a Skellam distribution, denoted $\mathcal{S}(\lambda_1, \lambda_2)$. This distribution is due to the ecologist and statistician John Gordon Skellam (1914–1979), who studied its basic properties [14]. If $X \sim \mathcal{S}(\lambda_1, \lambda_2)$, then, for all $x \in \mathbb{Z}$,

$$\Pr(X = x) = \lambda_2^{-x} e^{-(\lambda_1+\lambda_2)} \sum_{k=\max(0,x)}^{\infty} \frac{(\lambda_1\lambda_2)^k}{k!(k-x)!}, \tag{1}$$

which can also be expressed in terms of Bessel functions; see, e.g., Prékopa [12]. The symmetric case $\lambda_1 = \lambda_2$ was considered earlier by Irwin [6].

The Skellam distribution arises naturally in Kendall's famous taxicab problem [4,9]. It has found applications in various fields such as physics, medicine, and sports statistics; see, e.g., [5,7,15]. Its properties were reviewed by Karlis and Ntzoufras [8], who note that (1) remains valid when the random vector (Y_1, Y_2) follows a bivariate Poisson distribution and its components share a common additive random contribution that is canceled by the differencing.

This paper considers various bivariate extensions of the Skellam distribution. One such extension was recently proposed by Bulla et al. [2]. In their model, a pair (X_1, X_2) is said to have a bivariate Skellam distribution, denoted $\mathcal{BS}_0(\lambda_0, \lambda_1, \lambda_2)$, if there exist mutually independent Poisson random variables Y_0, Y_1, Y_2 with $\mathrm{E}(Y_j) = \lambda_j > 0$ for $j \in \{0, 1, 2\}$ such that

$$X_1 = Y_1 - Y_0, \quad X_2 = Y_2 - Y_0.$$

The case of independence is encompassed if, by convention, $Y_0 \equiv 0$ when $\lambda_0 = 0$. While this construction induces dependence $\lambda_0 = \mathrm{cov}(X_1, X_2) \geq 0$ between the components, their margins also involve λ_0, because

$$X_1 \sim \mathcal{S}(\lambda_1, \lambda_0), \quad X_2 \sim \mathcal{S}(\lambda_2, \lambda_0).$$

In this model, therefore, λ_0 is not a margin-free dependence parameter.

In Section 2, alternative bivariate extensions of the Skellam distribution are proposed. After briefly considering copula-based constructions, we propose and study two different shock models. In the first model, positive dependence is governed by a real-valued parameter that does not affect the marginal Skellam distributions of X_1 and X_2. The second model is a two-parameter extension of the first which provides a greater range of positive dependence between the components. Section 3 describes how these two models can be adapted to account for negative dependence.

The models proposed here are easy to interpret and simulate. Their parameters are linked in an explicit way to the marginal distributions and Pearson's correlation between the variables. As shown in Section 4, however, their connection with Kendall's tau and Spearman's rho is only expressible in series form. Because the formula for Pearson's correlation is simple, explicit moment-based estimators exist for the dependence parameters, as shown in Section 5. Multivariate extensions of the model are sketched in Section 6.

2. BIVARIATE SKELLAM DISTRIBUTIONS

Given positive parameters λ_{11}, λ_{12}, λ_{21} and λ_{22}, suppose that it is desired to construct a model for a pair (X_1, X_2) such that

$$X_1 \sim \mathcal{S}(\lambda_{11}, \lambda_{12}), \quad X_2 \sim \mathcal{S}(\lambda_{21}, \lambda_{22}). \tag{2}$$

Possibly the simplest analytic way of constructing a joint distribution with such marginal distributions, denoted F_1 and F_2, is to set, for all $x_1, x_2 \in \mathbb{Z}$,

$$\Pr(X_1 \leq x_1, X_2 \leq x_2) = C\{F_1(x_1), F_2(x_2)\}, \tag{3}$$

where C is a copula, i.e., a bivariate cumulative distribution function with uniform margins on $(0, 1)$; see Nelsen [10] for an introduction to the subject.

For example, the Farlie–Gumbel–Morgenstern (FGM) copula with parameter $\theta \in [-1, 1]$ is defined, for all $u_1, u_2 \in [0, 1]$, by

$$C_\theta(u_1, u_2) = u_1u_2 + \theta u_1u_2(1 - u_1)(1 - u_2).$$

For any fixed θ, a bivariate Skellam distribution H_θ with margins F_1 and F_2 is thus obtained by setting, for all $x_1, x_2 \in \mathbb{R}$,

$$H_\theta(x_1, x_2) = F_1(x_1)F_2(x_2) + \theta F_1(x_1)F_2(x_2)\{1 - F_1(x_1)\}\{1 - F_2(x_2)\}. \tag{4}$$

In addition, the FGM family is ordered by positive quadrant dependence (PQD), i.e., $\theta_1 \le \theta_2 \Rightarrow C_{\theta_1}(u_1, u_2) \le C_{\theta_2}(u_1, u_2)$ for all $u_1, u_2 \in (0, 1)$. Consequently, $\theta_1 \le \theta_2 \Rightarrow H_{\theta_1}(x_1, x_2) \le H_{\theta_2}(x_1, x_2)$ for all $x_1, x_2 \in \mathbb{R}$ and thus the map

$$\theta \mapsto \text{cov}_\theta(X_1, X_2) = \iint x_1x_2 \mathrm{d}H_\theta(x_1, x_2) - \mathrm{E}(X_1)\mathrm{E}(X_2)$$

is non-decreasing, allowing θ to be interpreted as a genuine, margin-free parameter governing dependence between X_1 and X_2.

The FGM is only one of dozens of parametric families of bivariate copulas that are ordered by PQD; again, see Nelsen [10]. Each such family yields a different extension of the Skellam distribution. It is relatively easy to simulate samples from any such distribution using the R package `copula` and the R implementation of the univariate Skellam distribution by Jerry W. Lewis.

For example, to generate 1000 pairs $X = (X_1, X_2)$ from the FGM model (4) with $\theta = .5$, $\lambda_{11} = 1$, $\lambda_{12} = 2$, $\lambda_{21} = 3$ and $\lambda_{22} = 4$, one would begin by simulating 1000 pairs $U = (U_1, U_2)$ from the FGM(.5), and one would then set $X_1 = F_1^{-1}(U_1)$ and $X_2 = F_2^{-1}(U_2)$, viz.

```
U <- rCopula(1000,fgmCopula(0.5))
X1 <- qskellam(U[,1],1,2)
X2 <- qskellam(U[,2],3,4)
X <- cbind(X1,X2).
```

However, copula-based extensions of the Skellam distribution cannot be interpreted in simple probabilistic terms and do not share with the univariate Skellam distribution its closure under convolution. In addition, they typically do not lead to practical expressions for the joint cumulative distribution or probability mass function, let alone moments. Further note that when X_1 and X_2 are modeled through relation (3), the copula C is uniquely defined only on $\mathcal{R} = \text{Ran}(F_1) \times \text{Ran}(F_2)$; in other words, two copulas that agree on $\mathcal{R}$ lead to the same model. This lack of uniqueness raises identifiability issues and other complications; see Genest and Nešlehová [3] for a review.

In what follows, two probabilistic approaches are proposed for the construction of bivariate Skellam distributions. The first model is a special case of the second. In both constructions, the dependence between the margins is induced by a common shock, much as in the standard bivariate extension of the Poisson distribution. The parameter governing the amplitude of this shock also regulates the dependence between the variables; it can be tuned without affecting the marginal distributions. In addition, the models have a simple interpretation, they are closed under convolution and easy to simulate.

2.1. First Model

Let $\lambda_1 = \min(\lambda_{11}, \lambda_{21}) > 0$ and for fixed $\theta \in [0, \lambda_1]$, let Y_0, Y_1, Y_2 be mutually independent random variables such that

$$Y_1 \sim \mathcal{S}(\lambda_{11} - \theta, \lambda_{12}), \quad Y_2 \sim \mathcal{S}(\lambda_{21} - \theta, \lambda_{22}),$$

and $Y_0 \sim \mathcal{P}(\theta)$ is a Poisson random variable with mean $\theta \geq 0$, with the understanding that $Y_0 \equiv 0$ if $\theta = 0$. Let $G_{1\theta}$ and $G_{2\theta}$ denote the cumulative distribution functions of Y_1 and Y_2, respectively.

Definition 2.1: *A pair (X_1, X_2) is said to have a bivariate Skellam distribution of the first kind, denoted $\mathcal{BS}_1(\theta; \lambda_{11}, \lambda_{12}; \lambda_{21}, \lambda_{22})$, if and only if*

$$X_1 = Y_1 + Y_0, \quad X_2 = Y_2 + Y_0.$$

It is clear that this construction meets condition (2). In other words, the parameter θ does not affect the marginal distributions of X_1 and X_2. To see that θ is a dependence parameter, it suffices to show that, for all $x_1, x_2 \in \mathbb{R}$,

$$H_\theta(x_1, x_2) = \Pr(X_1 \leq x_1, X_2 \leq x_2) = \sum_{k=0}^{\infty} G_{1\theta}(x_1 - k) G_{2\theta}(x_2 - k) \, \frac{e^{-\theta}\theta^k}{k!}$$

is non-decreasing in θ. This is the object of the following result.

Proposition 2.2: *For all $x_1, x_2 \in \mathbb{R}$ and $\theta_1, \theta_2 \in [0, \lambda_1]$, one has*

$$\theta_1 \leq \theta_2 \Rightarrow H_{\theta_1}(x_1, x_2) \leq H_{\theta_2}(x_1, x_2).$$

Proof: Fix $\lambda > 0$ and for given $\theta \in (0, \lambda)$, let F_θ denote the cumulative distribution function of a Poisson random variable with mean $\lambda - \theta$. Then for arbitrary $k \in \mathbb{N}$, $\partial F_\theta(k)/\partial\theta = F_\theta(k) - F_\theta(k-1) = e^{-(\lambda-\theta)}(\lambda - \theta)^k/k!$. Using this fact, one can see that, for all $x \in \mathbb{Z}$ and $j \in \{1, 2\}$,

$$g_{j\theta}(x) = \frac{\partial}{\partial\theta} G_{j\theta}(x) = G_{j\theta}(x) - G_{j\theta}(x-1) \geq 0.$$

It follows that, for all $x_1, x_2 \in \mathbb{Z}$,

$$\begin{aligned}
\frac{\partial}{\partial\theta} H_\theta(x_1, x_2) = & \sum_{k=0}^{\infty} g_{1\theta}(x_1 - k) G_{2\theta}(x_2 - k) \, \frac{e^{-\theta}\theta^k}{k!} \\
& + \sum_{k=0}^{\infty} G_{1\theta}(x_1 - k) g_{2\theta}(x_2 - k) \, \frac{e^{-\theta}\theta^k}{k!} \\
& + \sum_{k=0}^{\infty} G_{1\theta}(x_1 - k - 1) G_{2\theta}(x_2 - k - 1) \, \frac{e^{-\theta}\theta^k}{k!} \\
& - \sum_{k=0}^{\infty} G_{1\theta}(x_1 - k) G_{2\theta}(x_2 - k) \, \frac{e^{-\theta}\theta^k}{k!}.
\end{aligned}$$

Upon simplification, one finds

$$\frac{\partial}{\partial\theta} H_\theta(x_1, x_2) = \sum_{k=0}^{\infty} g_{1\theta}(x_1 - k) g_{2\theta}(x_2 - k) \, \frac{\theta^k e^{-\theta}}{k!} \geq 0.$$

Hence for arbitrary $x_1, x_2 \in \mathbb{R}$, the map $\theta \mapsto H_\theta(x_1, x_2)$ is non-decreasing, whence the result. ■

Given that, for all $x_1, x_2 \in \mathbb{R}$,

$$H_0(x_1, x_2) = F_1(x_1)F_2(x_2),$$

it follows from Proposition 2.2 that if $(X_1, X_2) \sim \mathcal{BS}_1(\theta; \lambda_{11}, \lambda_{12}; \lambda_{21}, \lambda_{22})$, its components are PQD. An extension allowing for negative dependence between X_1 and X_2 is considered in Section 3.

It is also clear from Definition 2.1 that the Skellam distribution of the first kind is closed under convolution, i.e., the component-wise sum of mutually independent Skellam random pairs has a Skellam distribution, as stated below.

Proposition 2.3: *Let $\mathbf{X}_1, \ldots, \mathbf{X}_n$ be mutually independent random pairs such that, for $i \in \{1, \ldots, n\}$, $\mathbf{X}_i \sim \mathcal{BS}_1(\theta_i; \lambda_{i11}, \lambda_{i12}; \lambda_{i21}, \lambda_{i22})$. Then*

$$\sum_{i=1}^n \mathbf{X}_i \sim \mathcal{BS}_1\left(\sum_{i=1}^n \theta_i; \sum_{i=1}^n \lambda_{i11}, \sum_{i=1}^n \lambda_{i12}; \sum_{i=1}^n \lambda_{i21}, \sum_{i=1}^n \lambda_{i22}\right).$$

In addition, all moments of the Skellam distribution of the first kind can be deduced easily from its probability generating function given below.

Proposition 2.4: *Suppose that $(X_1, X_2) \sim \mathcal{BS}_1(\theta; \lambda_{11}, \lambda_{12}; \lambda_{21}, \lambda_{22})$. Its probability generating function is then given, for all $s_1, s_2 \in (0,1)$, by*

$$\begin{aligned} \mathrm{E}(s_1^{X_1} s_2^{X_2}) = \exp\{(\lambda_{11} - \theta)(s_1 - 1) + \lambda_{12}(1/s_1 - 1)\} \times \exp\{(\lambda_{21} - \theta)(s_2 - 1) \\ + \lambda_{22}(1/s_2 - 1)\} \times \exp\{\theta(s_1 s_2 - 1)\}. \end{aligned}$$

Proof: By definition, one has $X_1 = Y_1 + Y_0$ and $X_2 = Y_2 + Y_0$, where Y_0, Y_1, Y_2 are mutually independent with $Y_0 \sim \mathcal{P}(\theta)$, $Y_1 \sim \mathcal{S}(\lambda_{11} - \theta, \lambda_{12})$, and $Y_2 \sim \mathcal{S}(\lambda_{21} - \theta, \lambda_{22})$. Consequently,

$$\mathrm{E}(s_1^{X_1} s_2^{X_2}) = \mathrm{E}(s_1^{Y_1})\mathrm{E}(s_2^{Y_2})\mathrm{E}\{(s_1 s_2)^{Y_0}\}. \quad \textbf{(5)}$$

Using the fact for $j \in \{1, 2\}$, Y_j is the difference of two independent Poisson random variables with parameters $\lambda_{j1} - \theta$ and λ_{j2}, one finds

$$\mathrm{E}(s_j^{Y_j}) = \exp\{(\lambda_{j1} - \theta)(s_j - 1) + \lambda_{j2}(1/s_j - 1)\}.$$

As $\mathrm{E}\{(s_1 s_2)^{Y_0}\} = \exp\{\theta(s_1 s_2 - 1)\}$, the argument is complete. ■

In particular, $\mathrm{E}(X_j) = \lambda_{j1} - \lambda_{j2}$ and $\mathrm{var}(X_j) = \lambda_{j1} + \lambda_{j2}$ for $j \in \{1, 2\}$. Furthermore,

$$\mathrm{corr}(X_1, X_2) = \frac{\theta}{\sqrt{(\lambda_{11} + \lambda_{12})(\lambda_{21} + \lambda_{22})}} \le \frac{\lambda_1}{\sqrt{(\lambda_{11} + \lambda_{12})(\lambda_{21} + \lambda_{22})}}.$$

The largest possible value of $\mathrm{corr}(X_1, X_2)$ occurs in the special case where X_1 and X_2 are identically distributed, i.e., $\lambda_{11} = \lambda_{21} = \lambda_1$ and $\lambda_{12} = \lambda_{22} = \lambda_2$. In this case, one finds $\mathrm{corr}(X_1, X_2) = \lambda_1/(\lambda_1 + \lambda_2)$ which is strictly less than 1. In particular, $\mathrm{corr}(X_1, X_2) = 1/2$ when $\lambda_1 = \lambda_2$.

To generate 1000 pairs $X = (X_1, X_2)$ from the bivariate Skellam distribution of the first kind with parameters $\theta = .5$, $\lambda_{11} = 1$, $\lambda_{12} = 2$, $\lambda_{21} = 3$ and $\lambda_{22} = 4$, one can proceed as follows:

```
Y0 <- rpois(1000,0.5)
Y1 <- rskellam(1000,1-0.5,2)
Y2 <- rskellam(1000,3-0.5,4)
X <- cbind(Y0+Y1,Y0+Y2).
```

While this model is easy to interpret and simulate from, it provides a limited range of dependence between X_1 and X_2. This problem can be alleviated in part through the following generalization.

2.2. Second Model

Introduce $\lambda_2 = \min(\lambda_{12}, \lambda_{22}) > 0$. For fixed $\Theta = (\theta_1, \theta_2) \in [0, \lambda_1] \times [0, \lambda_2]$, let Y_0, Y_1, Y_2 be mutually independent random variables with $Y_0 \sim \mathcal{S}(\theta_1, \theta_2)$,

$$Y_1 \sim \mathcal{S}(\lambda_{11} - \theta_1, \lambda_{12} - \theta_2), \quad Y_2 \sim \mathcal{S}(\lambda_{21} - \theta_1, \lambda_{22} - \theta_2).$$

Denote the cumulative distribution functions of Y_1 and Y_2 by $G_{1\Theta}$ and $G_{2\Theta}$, respectively.

Definition 2.5: *A pair (X_1, X_2) is said to have a bivariate Skellam distribution of the second kind, denoted $\mathcal{BS}_2(\theta_1, \theta_2; \lambda_{11}, \lambda_{12}; \lambda_{21}, \lambda_{22})$, if and only if*

$$X_1 = Y_1 + Y_0, \quad X_2 = Y_2 + Y_0.$$

In this construction, the vector parameter $\Theta = (\theta_1, \theta_2)$ does not affect the marginal distributions of X_1 and X_2. Furthermore, taking $\theta_2 = 0$ reduces Y_0 to a Poisson random variable with mean θ_1 while one then has $Y_1 \sim \mathcal{S}(\lambda_{11} - \theta, \lambda_{12})$ and $Y_2 \sim \mathcal{S}(\lambda_{21} - \theta, \lambda_{22})$. Accordingly,

$$\mathcal{BS}_1(\theta_1; \lambda_{11}, \lambda_{12}; \lambda_{21}, \lambda_{22}) \equiv \mathcal{BS}_2(\theta_1, 0; \lambda_{11}, \lambda_{12}; \lambda_{21}, \lambda_{22}).$$

The cumulative distribution function of the bivariate Skellam distribution of the second kind is given, for all $x_1, x_2 \in \mathbb{R}$, by

$$\Pr(X_1 \le x_1, X_2 \le x_2) = H_\Theta(x_1, x_2) = \sum_{k=-\infty}^{\infty} G_{1\Theta}(x_1 - k) G_{2\Theta}(x_2 - k) g_\Theta(k),$$

where g_Θ denotes the probability mass function of Y_0. In what follows, vector algebra is applied component-wise.

Proposition 2.6: *For all $x_1, x_2 \in \mathbb{R}$ and $\Theta_1, \Theta_2 \in [0, \lambda_1] \times [0, \lambda_2]$, one has*

$$\Theta_1 \le \Theta_2 \Rightarrow H_{\Theta_1}(x_1, x_2) \le H_{\Theta_2}(x_1, x_2).$$

Proof: Proceeding as in the proof of Proposition 2.2, one finds, for all $x \in \mathbb{Z}$,

$$\frac{\partial}{\partial \theta_1} g_\Theta(x) = g_\Theta(x-1) - g_\Theta(x)$$

and, for $j \in \{1,2\}$,

$$g_{j\Theta}(x) = \frac{\partial}{\partial \theta_1} G_{j\Theta}(x) = G_{j\Theta}(x) - G_{j\Theta}(x-1) \geq 0.$$

It follows that for all $x_1, x_2 \in \mathbb{Z}$,

$$\begin{aligned}\frac{\partial}{\partial \theta_1} H_\Theta(x_1,x_2) = &\sum_{k=-\infty}^{\infty} g_{1\Theta}(x_1-k)G_{2\Theta}(x_2-k)g_\Theta(k)\\ &+ \sum_{k=-\infty}^{\infty} G_{1\Theta}(x_1-k)g_{2\Theta}(x_2-k)g_\Theta(k)\\ &+ \sum_{k=-\infty}^{\infty} G_{1\Theta}(x_1-k)G_{2\Theta}(x_2-k)g_\Theta(k-1)\\ &- \sum_{k=-\infty}^{\infty} G_{1\Theta}(x_1-k)G_{2\Theta}(x_2-k)g_\Theta(k).\end{aligned}$$

Simple manipulations then lead to

$$\frac{\partial}{\partial \theta_1} H_\Theta(x_1,x_2) = \sum_{k=-\infty}^{\infty} g_{1\Theta}(x_1-k)g_{2\Theta}(x_2-k)g_\Theta(k) \geq 0.$$

Similarly, $\partial H_\Theta(x_1,x_2)/\partial\theta_2 \geq 0$. It follows that the map $\Theta \mapsto H_\Theta(x_1,x_2)$ is non-decreasing in both of its arguments, which completes the proof. ■

Given that $Y \sim \mathcal{S}(0,0)$ corresponds to the case where $Y \equiv 0$, one has, for all $x_1, x_2 \in \mathbb{R}$,

$$H_{(0,0)}(x_1,x_2) = F_1(x_1)F_2(x_2).$$

Thus Proposition 2.6 implies that H_Θ is PQD for all $\Theta \in [0,\lambda_1] \times [0,\lambda_2]$. For an extension allowing for negative dependence, see Section 3. The following results generalize Propositions 2.3 and 2.4, respectively.

Proposition 2.7: *Let $\mathbf{X}_1 \ldots, \mathbf{X}_n$ be mutually independent random pairs such that, for $i \in \{1,\ldots,n\}$, $\mathbf{X}_i \sim \mathcal{BS}_2(\theta_{i1},\theta_{i2};\lambda_{i11},\lambda_{i12};\lambda_{i21},\lambda_{i22})$. Then*

$$\sum_{i=1}^n \mathbf{X}_i \sim \mathcal{BS}_2\left(\sum_{i=1}^n \theta_{i1}, \sum_{i=1}^n \theta_{i2}; \sum_{i=1}^n \lambda_{i11}, \sum_{i=1}^n \lambda_{i12}; \sum_{i=1}^n \lambda_{i21}, \sum_{i=1}^n \lambda_{i22}\right).$$

Proposition 2.8: *Suppose that $(X_1,X_2) \sim \mathcal{BS}_2(\theta_1,\theta_2;\lambda_{11},\lambda_{12};\lambda_{21},\lambda_{22})$. Then its probability generating function is given, for all $s_1, s_2 \in (0,1)$, by*

$$\begin{aligned}\mathrm{E}(s_1^{X_1}s_2^{X_2}) = &\exp\{(\lambda_{11}-\theta_1)(s_1-1) + (\lambda_{12}-\theta_2)(1/s_1-1)\} \times \exp\{(\lambda_{21}-\theta_1)(s_2-1)\\ &+ (\lambda_{22}-\theta_2)(1/s_2-1)\} \times \exp[\theta_1(s_1s_2-1) + \theta_2\{1/(s_1s_2)-1\}].\end{aligned}$$

https://doi.org/10.1017/S0269964814000072 Published online by Cambridge University Press

Proof: By definition, one has $X_1 = Y_1 + Y_0$ and $X_2 = Y_2 + Y_0$, where Y_0, Y_1, Y_2 are mutually independent with $Y_1 \sim \mathcal{S}(\lambda_{11} - \theta_1, \lambda_{12} - \theta_2)$, $Y_2 \sim \mathcal{S}(\lambda_{21} - \theta_1, \lambda_{22} - \theta_2)$, and $Y_0 \sim \mathcal{S}(\theta_1, \theta_2)$. Now for $j \in \{1, 2\}$, Y_j is the difference of two independent Poisson random variables with parameters $\lambda_{j1} - \theta_1$ and $\lambda_{j2} - \theta_2$. Therefore,

$$\mathrm{E}(s_j^{Y_j}) = \exp\{(\lambda_{j1} - \theta_1)(s_j - 1) + (\lambda_{j2} - \theta_2)(1/s_j - 1)\}.$$

Furthermore, $\mathrm{E}\{(s_1 s_2)^{Y_0}\} = \exp[\theta_1(s_1 s_2 - 1) + \theta_2\{1/(s_1 s_2) - 1\}]$. The conclusion now follows, upon substitution into Eq. (5). ■

In particular, $\mathrm{E}(X_j) = \lambda_{j1} - \lambda_{j2}$ and $\mathrm{var}(X_j) = \lambda_{j1} + \lambda_{j2}$ for $j \in \{1, 2\}$. Furthermore,

$$\mathrm{corr}(X_1, X_2) = \frac{\theta_1 + \theta_2}{\sqrt{(\lambda_{11} + \lambda_{12})(\lambda_{21} + \lambda_{22})}} \le \frac{\lambda_1 + \lambda_2}{\sqrt{(\lambda_{11} + \lambda_{12})(\lambda_{21} + \lambda_{22})}}\,.$$

The range of values for $\mathrm{corr}(X_1, X_2)$ is thus larger than under the first model. In fact, one can get $\mathrm{corr}(X_1, X_2) = 1$ when X_1 and X_2 are identically distributed, i.e., $\lambda_{11} = \lambda_{21} = \lambda_1$ and $\lambda_{12} = \lambda_{22} = \lambda_2$. Indeed if $\theta_1 \to \lambda_1$ and $\theta_2 \to \lambda_2$, one has $Y_1 = Y_2 \equiv 0$ and $X_1 = X_2 = Y_0$ almost surely. The variables X_1 and X_2 are then said to be comonotonic.

To generate 1000 pairs $X = (X_1, X_2)$ from the bivariate Skellam distribution of the second kind with parameters $\theta_1 = .5$, $\theta_2 = 1.5$, $\lambda_{11} = 1$, $\lambda_{12} = 2$, $\lambda_{21} = 3$ and $\lambda_{22} = 4$, one can proceed as follows:

```
Y0 <- rskellam(1000,0.5,1.5)
Y1 <- rskellam(1000,1-0.5,2-1.5)
Y2 <- rskellam(1000,3-0.5,4-1.5)
X <- cbind(Y0+Y1,Y0+Y2).
```

3. MODELS WITH NEGATIVE DEPENDENCE

The two shock models described in Section 2 can be adapted easily to account for negative dependence between X_1 and X_2. This can be done by setting

$$X_1 = Y_1 + Y_0, \quad X_2 = Y_2 - Y_0, \tag{6}$$

where Y_0, Y_1, Y_2 are mutually independent random variables such that

$$Y_1 \sim \mathcal{S}(\lambda_{11} - \theta_1, \lambda_{12} - \theta_2), \quad Y_2 \sim \mathcal{S}(\lambda_{21} - \theta_2, \lambda_{22} - \theta_1),$$

and $Y_0 \sim \mathcal{S}(\theta_1, \theta_2)$ with $\theta_1 \le \kappa_1 = \min(\lambda_{11}, \lambda_{22})$ and $\theta_2 \le \kappa_2 = \min(\lambda_{12}, \lambda_{21})$.

Let $G_{1\Theta}$ and $G_{2\Theta}$ be the cumulative distribution functions of Y_1 and Y_2, respectively. Let also g_Θ stand for the probability mass function of Y_0. The joint distribution of the pair (X_1, X_2) is then given, for all $x_1, x_2 \in \mathbb{R}$, by

$$H_\Theta(x_1, x_2) = \sum_{k=-\infty}^{\infty} G_{1\Theta}(x_1 - k) G_{2\Theta}(x_2 + k) g_\Theta(k).$$

This model clearly induces negative dependence between the variables X_1 and X_2 because, by construction,

$$(X_1, -X_2) \sim \mathcal{BS}_2(\theta_1, \theta_2; \lambda_{11}, \lambda_{12}; \lambda_{22}, \lambda_{21}). \tag{7}$$

Note that in the above, λ_{21} and λ_{22} do not appear in the same order as in Definition 2.5 because $-Y_2 \sim \mathcal{S}(\lambda_{22} - \theta_1, \lambda_{21} - \theta_2)$. In view of Proposition 2.6, $(X_1, -X_2)$ is PQD and hence (X_1, X_2) is negative quadrant dependence for all possible values of Θ. In addition, the class of distributions is ordered as follows.

Proposition 3.1: *For all* $x_1, x_2 \in \mathbb{R}$ *and* $\Theta_1, \Theta_2 \in [0,\kappa_1] \times [0,\kappa_2]$, *one has*

$$\Theta_1 \leq \Theta_2 \Rightarrow H_{\Theta_1}(x_1, x_2) \geq H_{\Theta_2}(x_1, x_2).$$

This model is sufficiently broad to allow for the random variables X_1 and $-X_2$ to be decreasing functions of one another when they are identically distributed, i.e., $\lambda_{11} = \lambda_{22}$ and $\lambda_{12} = \lambda_{21}$. In this situation, the smallest possible dependence is obtained by taking $\theta_1 = \lambda_{11} = \lambda_{22}$ and $\theta_2 = \lambda_{12} = \lambda_{21}$. One then has $X_1 = -X_2 = Y_0$ almost surely, i.e., X_1 and X_2 are countermonotonic.

Using relation (7), one can show that the probability generating function of any pair (X_1, X_2) of the form (6) is given, for all $s_1, s_2 \in (0, 1)$, by

$$\begin{aligned} \mathrm{E}(s_1^{X_1} s_2^{X_2}) &= \mathrm{E}\{s_1^{Y_1} s_2^{Y_2} (s_1/s_2)^{Y_0}\} \\ &= \exp\{(\lambda_{11} - \theta_1)(s_1 - 1) + (\lambda_{12} - \theta_2)(1/s_1 - 1)\} \\ &\quad \times \exp\{(\lambda_{21} - \theta_1)(s_2 - 1) + (\lambda_{22} - \theta_2)(1/s_2 - 1)\} \\ &\quad \times \exp\{\theta_1(s_1/s_2 - 1) + \theta_2(s_2/s_1 - 1)\}. \end{aligned}$$

Here again, $\mathrm{E}(X_j) = \lambda_{j1} - \lambda_{j2}$, $\mathrm{var}(X_j) = \lambda_{j1} + \lambda_{j2}$ for $j \in \{1, 2\}$, and

$$\mathrm{corr}(X_1, X_2) = -\frac{\theta_1 + \theta_2}{\sqrt{(\lambda_{11} + \lambda_{12})(\lambda_{21} + \lambda_{22})}}.$$

4. NON-PARAMETRIC MEASURES OF DEPENDENCE

As seen in Section 2, a simple formula is available for $\mathrm{corr}(X_1, X_2)$ whenever $(X_1, X_2) \sim \mathcal{BS}_2(\theta_1, \theta_2; \lambda_{11}, \lambda_{12}; \lambda_{21}, \lambda_{22})$. A fortiori, the same holds true in the special case where $(X_1, X_2) \sim \mathcal{BS}_1(\theta; \lambda_{11}, \lambda_{12}; \lambda_{21}, \lambda_{22})$. Alas, closed form expressions for Kendall's tau and Spearman's rho are not available for these models. As shown below, however, these two non-parametric measures of dependence can be expressed in terms of covariances between non-decreasing functions of underlying variables distributed as $\mathcal{S}(\theta, \theta)$, where $\theta = \theta_1 + \theta_2$.

4.1. Kendall's Tau

In order to compute the (raw) value of Kendall's tau, consider independent pairs (X_1, X_2) and (X_1', X_2') from $\mathcal{BS}_2(\theta_1, \theta_2; \lambda_{11}, \lambda_{12}; \lambda_{21}, \lambda_{22})$. By definition, one can then write, for $j \in \{1, 2\}$,

$$X_j = Y_j + Y_0, \quad X_j' = Y_j' + Y_0'$$

using mutually independent random variables Y_0, Y_1, Y_2, Y_0', Y_1', Y_2' such that

$$Y_0 \sim \mathcal{S}(\theta_1, \theta_2), \quad Y_1 \sim \mathcal{S}(\lambda_{11} - \theta_1, \lambda_{12} - \theta_2), \quad Y_2 \sim \mathcal{S}(\lambda_{21} - \theta_1, \lambda_{22} - \theta_2)$$

and

$$Y_0' \sim \mathcal{S}(\theta_1, \theta_2), \quad Y_1' \sim \mathcal{S}(\lambda_{11} - \theta_1, \lambda_{12} - \theta_2), \quad Y_2' \sim \mathcal{S}(\lambda_{21} - \theta_1, \lambda_{22} - \theta_2).$$

For $j \in \{0, 1, 2\}$, introduce $Z_j = Y_j - Y_j'$ and let

$$T_1 = X_1 - X_1' = Z_1 + Z_0, \quad T_2 = X_2 - X_2' = Z_2 + Z_0.$$

Further set $\theta = \theta_1 + \theta_2$ and $\Lambda_j = \lambda_{j1} + \lambda_{j2}$ for $j \in \{1,2\}$. Then

$$Z_0 \sim \mathcal{S}(\theta, \theta), \quad Z_1 \sim \mathcal{S}(\Lambda_1 - \theta, \Lambda_1 - \theta), \quad Z_2 \sim \mathcal{S}(\Lambda_2 - \theta, \Lambda_2 - \theta)$$

and hence $(T_1, T_2) \sim \mathcal{BS}_2(\theta, \theta; \Lambda_1, \Lambda_1; \Lambda_2, \Lambda_2)$.

Denote by K_θ the joint cumulative distribution function of (T_1, T_2), and let $K_{1\theta}$ and $K_{2\theta}$ be the cumulative distribution functions of Z_1 and Z_2, respectively. Let also g_θ denote the probability mass function of Z_0. The following proposition leads, if desired, to a series expansion for the (raw) population value of Kendall's tau in the pair (X_1, X_2).

Proposition 4.1: *If $(X_1, X_2) \sim \mathcal{BS}_2(\theta_1, \theta_2; \lambda_{11}, \lambda_{12}; \lambda_{21}, \lambda_{22})$, the (unscaled) value of Kendall's tau for the pair (X_1, X_2) is then given by*

$$\tau(X_1, X_2) = \operatorname{cov}\{K_{1\theta}(Z_0) + K_{1\theta}(Z_0 - 1), K_{2\theta}(Z_0) + K_{2\theta}(Z_0 - 1)\}.$$

Proof: By definition, $\tau(X_1, X_2) = \Pr(T_1T_2 > 0) - \Pr(T_1T_2 < 0) = \Delta$, say. Given that $(T_1, T_2) = (Z_1 + Z_0, Z_2 + Z_0)$, the difference

$$\Delta = K_\theta(0,0) + K_\theta(-1,0) + K_\theta(0,-1) + K_\theta(-1,-1) - 1$$

can be developed as a series upon conditioning by Z_0. One finds

$$\begin{aligned}\Delta &= \sum_{k=-\infty}^{\infty} \{K_{1\theta}(-k)K_{2\theta}(-k) + K_{1\theta}(-k-1)K_{2\theta}(-k)\}g_\theta(k) \\ &\quad + \sum_{k=-\infty}^{\infty} \{K_{1\theta}(-k)K_{2\theta}(-k-1) + K_{1\theta}(-k-1)K_{2\theta}(-k-1)\}g_\theta(k) - 1 \\ &= \sum_{k=-\infty}^{\infty} \{K_{1\theta}(-k) + K_{1\theta}(-k-1)\}\{K_{2\theta}(-k) + K_{2\theta}(-k-1)\}g_\theta(k) - 1.\end{aligned}$$

Now use the fact that for $j \in \{0,1,2\}$, $-Z_j$ has the same distribution as Z_j to deduce that, for $j \in \{1,2\}$ and all $k \in \mathbb{Z}$,

$$K_{j\theta}(-k) = 1 - K_{j\theta}(k-1), \quad K_{j\theta}(-k-1) = 1 - K_{j\theta}(k).$$

Upon substitution, one finds

$$\begin{aligned}\Delta &= \sum_{k=-\infty}^{\infty} \{K_{1\theta}(k) + K_{1\theta}(k-1) - 2\}\{K_{2\theta}(k) + K_{2\theta}(k-1) - 2\}g_\theta(k) - 1 \\ &= \mathrm{E}[\{K_{1\theta}(Z_0) + K_{1\theta}(Z_0 - 1) - 2\}\{K_{2\theta}(Z_0) + K_{2\theta}(Z_0 - 1) - 2\}] - 1,\end{aligned}$$

which yields the desired conclusion because $\mathrm{E}\{K_{1\theta}(Z_0) + K_{j\theta}(Z_0 - 1)\} = 1$ for $j \in \{1,2\}$. ∎

Given that the mapping $t \mapsto K_{j\theta}(t) + K_{j\theta}(t-1)$ is non-decreasing for $j \in \{1,2\}$, it follows from Proposition 4.1 that $\tau(X_1, X_2) \geq 0$. The lower bound (i.e., 0) is reached when $\min(\theta_1, \theta_2) \to 0$, because in the limit $Z_0 \equiv 0$. As for the upper bound, it is reached when $\theta_j \to \min(\lambda_{1j}, \lambda_{2j})$ for $j \in \{1,2\}$.

An explicit expression for the upper bound on $\tau(X_1, X_2)$ can be obtained when the random variables X_1 and X_2 are identically distributed, i.e., when $\lambda_{1j} = \lambda_{2j} = \lambda_j$ for $j \in$

$\{1,2\}$. In that case, one has $Z_1 = Z_2 \equiv 0$, and hence $K_{j\theta}(k) = \mathbf{1}(k \geq 0)$ for $j \in \{1,2\}$ and all $k \in \mathbb{Z}$. In view of Proposition 4.1, the upper bound is then given by

$$\begin{aligned}\operatorname{var}\{\mathbf{1}(Z_0 \geq 0) + \mathbf{1}(Z_0 - 1 \geq 0)\} &= \operatorname{var}\{\mathbf{1}(Z_0 > 0) - \mathbf{1}(Z_0 < 0)\} \\ &= \mathrm{E}[\{\mathbf{1}(Z_0 > 0) - \mathbf{1}(Z_0 < 0)\}^2],\end{aligned}$$

because the distribution of Z_0 is symmetric with respect to the origin. Upon expanding the square, the latter expectation is found to be

$$2\,\mathrm{E}\{\mathbf{1}(Z_0 > 0)\} = 2\Pr(Y_0 > Y_0') = 2\,\mathrm{E}\{F_1(X_1 - 1)\}.$$

This is in concordance with Proposition 8 of Nešlehová [11].

As another immediate consequence of Proposition 4.1, note that if a random pair (X_1, X_2) is defined as in (6), it follows from (7) that

$$\tau(X_1, X_2) = -\operatorname{cov}\{K_{1\theta}(Z_0) + K_{1\theta}(Z_0 - 1), K_{2\theta}(Z_0) + K_{2\theta}(Z_0 - 1)\}.$$

4.2. Spearman's Rho

Turning to the computation of the (raw) theoretical value of Spearman's rho, let (X_1, X_2), (X_1', X_2') and (X_1'', X_2'') be mutually independent and identically distributed pairs from $\mathcal{BS}_2(\theta_1, \theta_2; \lambda_{11}, \lambda_{12}; \lambda_{21}, \lambda_{22})$. By definition, one can then write, for $j \in \{1,2\}$,

$$X_j = Y_j + Y_0, \quad X_j' = Y_j' + Y_0', \quad X_j'' = Y_j'' + Y_0'',$$

where for $j \in \{0,1,2\}$, Y_j, Y_j' and Y_j'' be mutually independent and identically distributed random variables such that $Y_0 \sim \mathcal{S}(\theta_1, \theta_2)$ and

$$Y_1 \sim \mathcal{S}(\lambda_{11} - \theta_1, \lambda_{12} - \theta_2), \quad Y_2 \sim \mathcal{S}(\lambda_{21} - \theta_1, \lambda_{22} - \theta_2).$$

Consider the random variables

$$Z_1 = Y_1 - Y_1', \quad Z_2 = Y_2 - Y_2'', \quad Z_0 = Y_0 - Y_0', \quad Z_0' = Y_0 - Y_0''$$

and set

$$T_1 = X_1 - X_1' = Z_1 + Z_0, \quad T_2 = X_2 - X_2'' = Z_2 + Z_0'.$$

Clearly, Z_0 and Z_0' are two dependent observations from distribution $\mathcal{S}(\theta, \theta)$, where $\theta = \theta_1 + \theta_2$. For $j \in \{1,2\}$, one also has $Z_j \sim \mathcal{S}(\Lambda_j - \theta, \Lambda_j - \theta)$, where $\Lambda_j = \lambda_{j1} + \lambda_{j2}$, as before.

Denote by K_θ the joint cumulative distribution function of (T_1, T_2), and let $K_{1\theta}$ and $K_{2\theta}$ be the cumulative distribution functions of Z_1 and Z_2, respectively. Let also g_θ denote the probability mass function of $(Z_0, Z_0') \sim \mathcal{BS}_2(\theta_1, \theta_2; \theta, \theta; \theta, \theta)$. The following result leads, if desired, to a series expansion for the (raw) population value of Spearman's rho for the pair (X_1, X_2).

PROPOSITION 4.2: *If $(X_1, X_2) \sim \mathcal{BS}_2(\theta_1, \theta_2; \lambda_{11}, \lambda_{12}; \lambda_{21}, \lambda_{22})$, the (unscaled) value of Spearman's rho for the pair (X_1, X_2) is then given by*

$$\rho(X_1, X_2) = 3\operatorname{cov}\left\{K_{1\theta}(Z_0) + K_{1\theta}(Z_0 - 1), K_{2\theta}(Z_0') + K_{2\theta}(Z_0' - 1)\right\}.$$

Proof: By definition, $\rho(X_1, X_2)/3 = \Pr(T_1T_2 > 0) - \Pr(T_1T_2 < 0) = \Delta$, say. Given that $(T_1, T_2) = (Z_1 + Z_0, Z_2 + Z_0')$, the difference

$$\Delta = K_\theta(0,0) + K_\theta(-1,0) + K_\theta(0,-1) + K_\theta(-1,-1) - 1$$

can be developed as a series upon conditioning by Z_0 and Z_0'. One finds

$$\begin{aligned}\Delta = &\sum_{k=-\infty}^{\infty}\sum_{k'=-\infty}^{\infty}\{K_{1\theta}(-k)K_{2\theta}(-k') + K_{1\theta}(-k-1)K_{2\theta}(-k')\}g_\theta(k,k')\\ &+ \sum_{k=-\infty}^{\infty}\sum_{k'=-\infty}^{\infty} K_{1\theta}(-k)K_{2\theta}(-k'-1)g_\theta(k,k')\\ &+ \sum_{k=-\infty}^{\infty}\sum_{k'=-\infty}^{\infty} K_{1\theta}(-k-1)K_{2\theta}(-k'-1)g_\theta(k,k') - 1.\end{aligned}$$

Upon simplification, one gets

$$\begin{aligned}\Delta = &\sum_{k=-\infty}^{\infty}\sum_{k'=-\infty}^{\infty}\{K_{1\theta}(-k) + K_{1\theta}(-k-1)\}\\ &\times\{K_{2\theta}(-k') + K_{2\theta}(-k'-1)\}g_\theta(k,k') - 1.\end{aligned}$$

Now use the fact that for $j \in \{1,2\}$, $-Z_j$ has the same distribution as Z_j, and $-Z_j'$ has the same distribution as Z_j'. Thus, for $j \in \{1,2\}$ and all $k \in \mathbb{Z}$, one has $K_{j\theta}(-k-1) = 1 - K_{j\theta}(k)$. Consequently,

$$\begin{aligned}\Delta = &\sum_{k=-\infty}^{\infty}\sum_{k'=-\infty}^{\infty}\{K_{1\theta}(k) + K_{1\theta}(k-1) - 2\}\\ &\times\{K_{2\theta}(k') + K_{2\theta}(k'-1) - 2\}g_\theta(k,k') - 1.\end{aligned}$$

Furthermore,

$$\mathrm{E}\{K_{1\theta}(Z_0) + K_{1\theta}(Z_0 - 1)\} = 1, \quad \mathrm{E}\{K_{2\theta}(Z_0') + K_{2\theta}(Z_0' - 1)\} = 1.$$

Consequently,

$$\Delta = \operatorname{cov}\{K_{1\theta}(Z_0) + K_{1\theta}(Z_0-1), K_{2\theta}(Z_0') + K_{2\theta}(Z_0'-1)\},$$

whence the conclusion. ∎

From Proposition 2.6, it is known that the pair (Z_0, Z_0') is PQD. Given that the mapping $t \mapsto K_{j\theta}(t) + K_{j\theta}(t-1)$ is non-decreasing for $j \in \{1,2\}$, one can then conclude from Proposition 4.2 that $\rho(X_1, X_2) \geq 0$. The lower bound (i.e., 0) is reached when $\min(\theta_1, \theta_2) \to 0$, because in the limit $Z_0 = Z_0' \equiv 0$. The upper bound occurs when $\theta_j \to \lambda_j = \min(\lambda_{1j}, \lambda_{2j})$ for $j \in \{1,2\}$.

As for Kendall's tau, an explicit upper bound on Spearman's rho can be found when the random variables X_1 and X_2 are identically distributed, i.e., when $\lambda_{1j} = \lambda_{2j} = \lambda_j$ for $j \in \{1,2\}$. The largest possible value, $\rho_{\max}$, of ρ then occurs when $\theta_j \to \lambda_j$ for $j \in \{1,2\}$. In

https://doi.org/10.1017/S0269964814000072 Published online by Cambridge University Press

that case, one has $Z_1 = Z_2 \equiv 0$ and hence $K_{j\theta}(k) = \mathbf{1}(k \geq 0)$ for $j \in \{1,2\}$ and all $k \in \mathbb{Z}$. Hence

$$\begin{aligned}\rho_{\max}/3 &= \operatorname{cov}\{\mathbf{1}(Z_0 > 0) - \mathbf{1}(Z_0 < 0), \mathbf{1}(Z_0' > 0) - \mathbf{1}(Z_0' < 0)\} \\ &= \Pr(Z_0 > 0, Z_0' > 0) + \Pr(Z_0 < 0, Z_0' < 0) \\ &\quad - \Pr(Z_0 < 0, Z_0' > 0) - \Pr(Z_0 > 0, Z_0' < 0).\end{aligned}$$

Note that if $\theta_j \to \lambda_j$ for $j \in \{1,2\}$, then $X_1 = X_2 = Y_0$ almost surely because $Y_1 = Y_2 \equiv 0$. Consequently,

$$\Pr(Z_0 > 0, Z_0' > 0) = \Pr(Y_0' < Y_0, Y_0'' < Y_0) = \mathrm{E}[\{F_1(X_1 - 1)\}^2].$$

Similarly, one has

$$\Pr(Z_0 > 0, Z_0' < 0) = \Pr(Z_0 < 0, Z_0' > 0) = \mathrm{E}[F_1(X_1 - 1)\{1 - F_1(X_1)\}]$$

and $\Pr(Z_0 < 0, Z_0' < 0) = \mathrm{E}[\{1 - F_1(X_1)\}^2]$. Finally, by using the identity

$$\mathrm{E}\{F_1(X_1 - 1)\} + \mathrm{E}\{F_1(X_1)\} = 1,$$

one gets

$$\rho_{\max} = 3\operatorname{var}\{F_1(X_1) + F_1(X_1 - 1)\}.$$

This is in accordance with the discussion in Section 4.3 of Nešlehová [11].

As another immediate consequence of Proposition 4.2, note that if a random pair (X_1, X_2) is defined as in (6), it follows from (7) that

$$\rho(X_1, X_2) = -3\operatorname{cov}\left\{K_{1\theta}(Z_0) + K_{1\theta}(Z_0 - 1), K_{2\theta}(Z_0') + K_{2\theta}(Z_0' - 1)\right\}.$$

5. PARAMETER ESTIMATION

In addition to being easy to interpret and simulate, the bivariate Skellam distributions defined here are simple to fit by the method of moments. Maximum-likelihood estimation is also possible but is not considered, as it is even more involved than for the univariate Skellam distribution; see, e.g., [1,7].

First observe that whether the random sample $(X_{11}, X_{12}), \ldots, (X_{n1}, X_{n2})$ arises from the bivariate Skellam distribution of the first kind or the second kind, the parameters of the margins $\mathcal{S}(\lambda_{11}, \lambda_{12})$ and $\mathcal{S}(\lambda_{21}, \lambda_{22})$ can be estimated in the standard way [13]. Setting

$$\bar{X}_1 = \frac{1}{n}\sum_{i=1}^n X_{i1}, \quad \bar{X}_2 = \frac{1}{n}\sum_{i=1}^n X_{i2},$$

and

$$S_1 = \frac{1}{n-1}\sum_{i=1}^n (X_{i1} - \bar{X}_1)^2, \quad S_2 = \frac{1}{n-1}\sum_{i=1}^n (X_{i2} - \bar{X}_2)^2,$$

the moment estimators are easily found to be

$$\hat{\lambda}_{11} = \frac{S_1^2 + \bar{X}_1}{2}, \quad \hat{\lambda}_{12} = \frac{S_1^2 - \bar{X}_1}{2}, \quad \hat{\lambda}_{21} = \frac{S_2^2 + \bar{X}_2}{2}, \quad \hat{\lambda}_{22} = \frac{S_2^2 - \bar{X}_2}{2}. \tag{8}$$

These estimators are explicit and consistent by the Law of Large Numbers. However, note that in order for all of them to be non-negative, one must have $S_j^2 \geq |\bar{X}_j|$ for $j \in \{1,2\}$. If this condition fails for some $j \in \{1,2\}$, one may proceed as Alzaid and Omair [1] by setting the negative estimate to zero and the other one to $|\bar{X}_j|$.

5.1. First Model

When the data arise from the $\mathcal{BS}_1(\theta; \lambda_{11}, \lambda_{12}; \lambda_{21}, \lambda_{22})$, a moment-based estimator of the dependence parameter θ is given by

$$\hat{\theta} = S_{12} = \frac{1}{n-1} \sum_{i=1}^{n} (X_{i1} - \bar{X}_1)(X_{i2} - \bar{X}_2).$$

This estimator is consistent by the Law of Large Numbers. When the sample size is small, however, there is a non-zero probability that $S_{12} \leq 0$, in which case one might set $\hat{\theta} = 0$.

In practice, a negative value for S_{12} may also suggest that a model of the form (6) is more appropriate for the pair (X_1, X_2). As seen in Section 3, this amounts to assuming that $(X_1, -X_2) \sim \mathcal{BS}_1(\theta; \lambda_{11}, \lambda_{12}; \lambda_{22}, \lambda_{21})$. Note that this does not affect the estimates of the marginal parameters given in (8) but as $\text{cov}(X_1, -X_2) = \theta$, a consistent estimator of θ is now given by $-S_{12} > 0$.

5.2. Second Model

When the data arise from the broader model $\mathcal{BS}_2(\theta_1, \theta_2; \lambda_{11}, \lambda_{12}; \lambda_{21}, \lambda_{22})$, two equations are needed in order to estimate θ_1 and θ_2 by the method of moments. The identity $\text{cov}(X_1, X_2) = \theta_1 + \theta_2$ leads to the estimating equation

$$S_{12} = \hat{\theta}_1 + \hat{\theta}_2,$$

from which a consistent estimator of $\theta = \theta_1 + \theta_2$ can be deduced. As $n \to \infty$, the probability that $S_{12} < 0$ becomes negligible. Thus if $S_{12} < 0$, it may be that a model of the form (6) is more appropriate.

5.2.1 Case $S_{12} > 0$.
In order to estimate θ_1 and θ_2 subject to $\hat{\theta}_1 + \hat{\theta}_2 = S_{12}$, a second equation must be called upon. The following proposition will be used to this end.

Proposition 5.1: *Suppose that* $(X_1, X_2) \sim \mathcal{BS}_2(\theta_1, \theta_2; \lambda_{11}, \lambda_{12}; \lambda_{21}, \lambda_{22})$. *Then*

$$\text{cov}(X_1, X_2^2) = (\theta_1 - \theta_2) + 2(\lambda_{21} - \lambda_{22})(\theta_1 + \theta_2),$$
$$\text{cov}(X_1^2, X_2) = (\theta_1 - \theta_2) + 2(\lambda_{11} - \lambda_{12})(\theta_1 + \theta_2).$$

Proof: By definition, one has $X_1 = Y_1 + Y_0$ and $X_2 = Y_2 + Y_0$, where Y_0, Y_1, Y_2 are mutually independent with $Y_0 \sim \mathcal{S}(\theta_1, \theta_2)$,

$$Y_1 \sim \mathcal{S}(\lambda_{11} - \theta_1, \lambda_{12} - \theta_2), \quad Y_2 \sim \mathcal{S}(\lambda_{21} - \theta_1, \lambda_{22} - \theta_2).$$

Consequently,

$$\begin{aligned}\text{cov}(X_1, X_2^2) &= \text{cov}(Y_0, Y_0^2) + 2\text{E}(Y_2)\text{var}(Y_0) \\ &= \text{E}(Y_0^3) + \{2\text{E}(Y_2) - \text{E}(Y_0)\}\text{var}(Y_0) - \{\text{E}(Y_0)\}^3.\end{aligned}$$

The desired expression for $\text{cov}(X_1, X_2^2)$ results from the fact that

$$\text{E}(Y_0) = \theta_1 - \theta_2, \quad \text{E}(Y_2) = \lambda_{21} - \lambda_{22} - (\theta_1 - \theta_2), \quad \text{var}(Y_0) = \theta_1 + \theta_2,$$

and

$$\text{E}(Y_0^3) = (\theta_1 - \theta_2) + 3(\theta_1^2 - \theta_2^2) + (\theta_1 - \theta_2)^3.$$

The formula for $\text{cov}(X_1^2, X_2)$ can be obtained in a similar way. ■

Plug-in estimates of $\mathrm{cov}(X_1^2, X_2)$ and $\mathrm{cov}(X_1, X_2^2)$ are given by

$$T_{12} = \frac{1}{n-1}\sum_{i=1}^{n}(X_{i2} - \bar{X}_2)(X_{i1}^2 - S_1^2 - \bar{X}_1^2) = \frac{1}{n-1}\sum_{i=1}^{n}(X_{i2} - \bar{X}_2)X_{i1}^2,$$

$$T_{21} = \frac{1}{n-1}\sum_{i=1}^{n}(X_{i1} - \bar{X}_1)(X_{i2}^2 - S_2^2 - \bar{X}_2^2) = \frac{1}{n-1}\sum_{i=1}^{n}(X_{i1} - \bar{X}_1)X_{i2}^2,$$

respectively. Given that $\hat{\lambda}_{11} - \hat{\lambda}_{12} = \bar{X}_1$ and $\hat{\lambda}_{21} - \hat{\lambda}_{22} = \bar{X}_2$ from Eq. (8), Proposition 5.1 suggests that moment estimators of θ_1 and θ_2 can be obtained by solving the equation

$$\frac{1}{2}(T_{12} + T_{21}) = \hat{\theta}_1 - \hat{\theta}_2 + (\bar{X}_1 + \bar{X}_2)(\hat{\theta}_1 + \hat{\theta}_2),$$

subject to $S_{12} = \hat{\theta}_1 + \hat{\theta}_2$. The solution happens to be explicit, viz.

$$\hat{\theta}_1 = \frac{1}{2}\left\{S_{12}(1 - \bar{X}_1 - \bar{X}_2) + \frac{1}{2}(T_{12} + T_{21})\right\},$$

$$\hat{\theta}_2 = \frac{1}{2}\left\{S_{12}(1 + \bar{X}_1 + \bar{X}_2) - \frac{1}{2}(T_{12} + T_{21})\right\}.$$

Again, these estimators are consistent by the Law of Large Numbers. By assumption, at most one of them can be negative; when this happens, it can be set equal to zero.

5.2.1. Case $S_{12} < 0$.
As mentioned earlier, it may be preferable to assume a model of the form (6) for (X_1, X_2) when S_{12} is negative. This amounts to assuming that $(X_1, -X_2) \sim \mathcal{B}_2(\theta_1, \theta_2; \lambda_{11}, \lambda_{12}; \lambda_{22}, \lambda_{21})$, in which λ_{12} and λ_{22} have been interchanged. In view of this fact, it is easy to see from Proposition 5.1 that

$$\mathrm{cov}(X_1, X_2^2) = (\theta_1 - \theta_2) - 2(\lambda_{21} - \lambda_{22})(\theta_1 + \theta_2),$$

$$\mathrm{cov}(X_1^2, X_2) = (\theta_2 - \theta_1) - 2(\lambda_{11} - \lambda_{12})(\theta_1 + \theta_2).$$

Estimators for these parameters are then obtained by solving the equations

$$S_{12} = -(\hat{\theta}_1 + \hat{\theta}_2), \quad \frac{1}{2}(T_{12} - T_{21}) = \hat{\theta}_1 - \hat{\theta}_2 + (\bar{X}_1 - \bar{X}_2)(\hat{\theta}_1 + \hat{\theta}_2).$$

Consequently,

$$\hat{\theta}_1 = \frac{1}{2}\left\{S_{12}(\bar{X}_1 - \bar{X}_2 - 1) + \frac{1}{2}(T_{12} - T_{21})\right\},$$

$$\hat{\theta}_2 = \frac{1}{2}\left\{S_{12}(\bar{X}_2 - \bar{X}_1 - 1) - \frac{1}{2}(T_{12} - T_{21})\right\}.$$

These estimators are consistent by the Law of Large Numbers. By assumption, at most one of them is negative; when this happens, it can be set equal to zero.

A study of the sampling properties of the estimators proposed herein will be the object of subsequent work.

6. DISCUSSION

The main purpose of this paper was to introduce bivariate Skellam distributions that could be easily interpreted and simulated. Two such models were proposed, based on a probabilistic construction involving a common shock. The models, which are nested, share the convolution closure property of the univariate Skellam distribution. Their basic features were studied, and moment-based estimators of their parameters were derived.

While the discussion was limited to the bivariate case, it is obvious that multivariate extensions of these models are possible. A general d-variate Skellam distribution could be defined, e.g., as the distribution of a random vector $\mathbf{X} = (X_1, \ldots, X_d)$ with components

$$X_1 = Y_1 + Y_0, \ldots, X_d = Y_d + Y_0,$$

where the random variables $Y_0, \ldots, Y_d$ are mutually independent with $Y_0 \sim \mathcal{S}(\theta_1, \theta_2)$ and, for all $j \in \{1, \ldots, d\}$, $Y_j \sim \mathcal{S}(\lambda_{1j} - \theta_1, \lambda_{2j} - \theta_2)$. In this model, $X_j \sim \mathcal{S}(\lambda_{1j}, \lambda_{2j})$ for each $j \in \{1, \ldots, d\}$ and the dependence between them is governed by non-negative parameters

$$\theta_1 < \min(\lambda_{11}, \ldots, \lambda_{1d}), \quad \theta_2 < \min(\lambda_{21}, \ldots, \lambda_{2d}).$$

Another fruitful way of inducing dependence in multi-class models is through shocks that can affect either the entire portfolio or subclasses thereof. For example, suppose that $\mathbf{X} = (\mathbf{X}_1, \ldots, \mathbf{X}_\ell)$, where $\mathbf{X}_j = (X_{k1}, \ldots, X_{kd})$ for each $k \in \{1, \ldots, \ell\}$. One could assume that for each k and all $j \in \{1, \ldots, d\}$,

$$X_{kj} = Y_{kj} + Y_j + Y_0,$$

where $Y_0 \sim \mathcal{S}(\theta_1, \theta_2)$, $Y_k \sim \mathcal{S}(\theta_{1k} - \theta_1, \theta_{2k} - \theta_2)$, and

$$Y_{kj} \sim \mathcal{S}(\lambda_{1kj} - \theta_{1k} - \theta_1, \lambda_{2kj} - \theta_{2k} - \theta_2).$$

Clearly, $X_{kj} \sim \mathcal{S}(\lambda_{1kj}, \lambda_{2kj})$ for all $k \in \{1, \ldots, \ell\}$ and $j \in \{1, \ldots, d\}$. In this model, the parameters θ_1 and θ_2 govern the global dependence, while θ_{1k} and θ_{2k} account for the dependence in class $k \in \{1, \ldots, \ell\}$. These models may be the subject of future study.

Acknowledgements

This research was supported by the Canada Research Chairs Program and grants from the Natural Sciences and Engineering Research Council of Canada and from the Fonds de recherche du Québec — Nature et technologies.

References

1. Alzaid, A.A. & Omair, M.A. (2010). On the Poisson difference distribution: inference and applications. *Bulletin of the Malaysian Mathematical Sciences Society* 33: 17–45.
2. Bulla, J., Chesneau, C. & Kachour, M. (2012). On the bivariate Skellam distribution. http://hal.archives-ouvertes.fr/hal-00744355
3. Genest, C. & Nešlehová, J. (2007). A primer on copulas for count data. *ASTIN Bulletin* 37: 475–515.
4. Giveen, S.M. (1963). A taxicab problem with time-dependent arrival rates. *SIAM Review* 5: 119–127.
5. Hwang, Y., Kim, J.-S. & Kweon, I.-S. (2007). Sensor noise modeling using the Skellam distribution: Application to the color edge detection. *IEEE Conference on Computer Vision and Pattern Recognition* pp. 1–8.
6. Irwin, W. (1937). The frequency distribution of the difference between two independent variates following the same Poisson distribution. *Journal of the Royal Statistical Society* 100: 415–416.
7. Karlis, D. & Ntzoufras, I. (2003). Analysis of sports data by using bivariate Poisson models. *The Statistician* 52: 381–393.

8. Karlis, D. & Ntzoufras, I. (2006). Bayesian analysis of the differences of count data. *Statistics in Medicine* 25: 1885–1905.
9. Kendall, D.G. (1951). Some problems in the theory of queues (with discussion). *Journal of the Royal Statistical Society Series B* 13: 151–185.
10. Nelsen, R.B. (2006). *An Introduction to Copulas*, 2nd ed. Berlin: Springer.
11. Nešlehová, J. (2007). On rank correlation measures for non-continuous random variables. *Journal of Multivariate Analysis* 98: 544–567.
12. Prékopa, A. (1952). On composed Poisson distributions. IV. Remarks on the theory of differential processes. *Acta Mathematica Academiae Scientiarum Hungaricae* 3: 317–325.
13. Romani, J. (1956). Distribución de la suma algebraica de variables de Poisson. *Trabajos de Estadística* 7: 175–181.
14. Skellam, J.G. (1946). The frequency distribution of the difference between two Poisson variates belonging to different populations. *Journal of the Royal Statistical Society* 109: 296.
15. Strackee, J. & Denier van der Gon, J.J.D. (1962). The frequency distribution of the difference between two Poisson variates. *Statistica Neerlandica* 16: 17–23.