

Table des matières

Résumé	ii
Table des matières	iv
1 La loi de Skellam univariée	3
1.1 Définition et propriétés de base	3
1.2 Autres représentations de la loi de Skellam	5
1.3 Estimation des paramètres	7
1.3.1 La méthode des moments	7
1.3.2 La méthode du maximum de vraisemblance	7
1.4 Loi asymptotique des estimateurs à vraisemblance maximale	9
1.5 Tests du rapport des vraisemblances maximales	11
2 La loi de Skellam bivariée	14
2.1 Définition et propriétés de base	14
2.2 Méthodes d'estimation des paramètres	16
2.2.1 La méthode des moments	16
2.2.2 La méthode du maximum de vraisemblance	16
2.3 Modèles de Skellam avec véritables paramètres de dépendance	19
2.3.1 Modèle de Skellam bivarié de première espèce	19
2.3.2 Modèle de Skellam bivarié de deuxième espèce	21
2.3.3 Modèles à dépendance négative	24
2.4 Estimation des paramètres des lois $\mathcal{BS}_1$ et $\mathcal{BS}_2$	25
2.4.1 Estimation du paramètre θ pour le modèle $\mathcal{BS}_1$	26
2.4.2 Estimation des paramètres θ_1 et θ_2 pour le modèle $\mathcal{BS}_2$	27
3 Contribution principale	30
3.1 Loi asymptotique des estimateurs des moments des paramètres du modèle $\mathcal{BS}_1$	31
3.2 Loi asymptotique des estimateurs des moments des paramètres du modèle $\mathcal{BS}_2$	34
3.3 Estimateurs à vraisemblance maximale des paramètres du modèle $\mathcal{BS}_1$	40

3.4	Estimateurs à vraisemblance maximale des paramètre du modèle $\mathcal{BS}_2$.	42
3.5	Loi asymptotique des estimateurs du maximum de vraisemblance des paramètres du modèle $\mathcal{BS}_1$	46
3.6	Loi asymptotique des estimateurs du maximum de vraisemblance des paramètres du modèle $\mathcal{BS}_2$	47
3.7	Comparaison entre les deux approches	50
Bibliographie		57

Introduction

On dit qu'une variable aléatoire obéit à une loi de Skellam si elle s'exprime comme la différence de deux variables aléatoires de Poisson indépendantes. Cette loi de probabilité est ainsi nommée en l'honneur du statisticien et biologiste John Gordon Skellam (1914–1979), qui a été parmi les premiers à l'étudier (Skellam, 1946).

La loi de Skellam trouve des applications dans divers domaines. Elle a notamment été utilisée pour modéliser la différence entre les scores de deux équipes sportives (Karlis & Ntzoufras, 2003) ou la différence, avant et après un traitement préventif, dans l'indice épidémiologique du cumul de nombre de dents cariées, absentes ou obturées. Hwang et coll. (2007) ont employé la loi de Skellam pour mesurer la différence d'intensité des pixels entre deux images numériques. Strackee & van der Gon (1962) évoquent une application plus ancienne de cette loi en physique.

L'objectif principal de ce mémoire est d'étudier le comportement asymptotique des estimateurs des paramètres de dépendance des lois de Skellam bivariées de première et de deuxième espèce introduites récemment par Genest & Mesfioui (2014). Dans un premier temps, la loi limite des estimateurs des moments des paramètres de dépendance des lois $\mathcal{BS}_1$ et $\mathcal{BS}_2$ sera déterminée. Dans un second temps, nous aborderons l'estimation de ces paramètres au moyen de la méthode du maximum de vraisemblance. Des résultats asymptotiques relatifs à ces estimateurs seront présentés et des simulations seront employées afin de comparer l'efficacité de ces deux approches d'estimation dans le cadre des lois de Skellam bivariées de première et de deuxième espèce.

Ce mémoire est structuré comme suit. Au chapitre 1, nous rappellerons la définition et les principales propriétés de la loi de Skellam univariée, qui comporte deux paramètres. Nous verrons ensuite comment il est possible d'estimer ces paramètres par la méthode des moments et par la méthode du maximum de vraisemblance. Les résultats asymptotiques de Alzaid & Omair (2010) concernant le comportement limite gaussien de ces deux types d'estimateurs seront rappelés. Nous montrerons aussi comment il est possible d'exploiter ces résultats pour tester l'égalité des paramètres de la loi de Skellam

au moyen d'un test du rapport des vraisemblances maximales.

Le chapitre 2 donne une description détaillée de trois modèles de Skellam bivariés récemment proposés. Dans un premier temps, nous décrirons et critiquerons le modèle introduit par Bulla et coll. (2012). Nous entamerons ensuite une étude exhaustive de deux modèles de Skellam bivariés alternatifs, notés $\mathcal{BS}_1$ et $\mathcal{BS}_2$, qui ont été suggérés par Genest & Mesfioui (2014). Nous décrirons plusieurs propriétés de ces modèles et verrons comment il est possible d'en estimer les différents paramètres par la méthode des moments.

Les principales contributions de ce mémoire sont présentées au chapitre 3. Nous y verrons comment il est possible d'estimer les paramètres des lois $\mathcal{BS}_1$ et $\mathcal{BS}_2$ par la méthode du maximum de vraisemblance. Nous préciserons en outre les lois asymptotiques des estimateurs des moments et à vraisemblance maximale des paramètres de dépendance de ces modèles. Nous donnerons en outre des expressions explicites des variances asymptotiques de ces estimateurs, dont nous explorerons l'efficacité asymptotique relative en termes numériques et graphiques.

Le mémoire se termine sur une courte conclusion et une annexe technique. Tous les travaux cités sont énumérés dans la bibliographie.

Chapitre 1

La loi de Skellam univariée

Ce chapitre rassemble un certain nombre de faits connus concernant la loi de Skellam univariée et les méthodes d'inférence qui lui sont propres. À la section 1.1, nous rappelons d'abord la définition et les principales propriétés de cette loi discrète à deux paramètres. Nous présentons ensuite à la section 1.3 d'autres représentations de cette loi. À la section 1.4, nous expliquons comment il est possible d'estimer les paramètres de la loi de Skellam univariée par la méthode des moments et la méthode du maximum de vraisemblance. Nous nous intéressons plus particulièrement au comportement asymptotique des estimateurs à vraisemblance maximale. Nous verrons en outre comment il est possible de tester l'égalité des deux paramètres de la loi de Skellam univariée au moyen de tests du rapport des vraisemblances maximales. La plupart des résultats présentés ici sont repris de l'article de Alzaid & Omair (2010).

1.1 Définition et propriétés de base

Rappelons d'abord la définition de la loi de Skellam univariée, popularisée par Skellam (1946) mais dont l'origine remonte en fait aux travaux de Irwin (1937).

Définition 1.1.1. Soient Y_1, Y_2 deux variables aléatoires de Poisson indépendantes telles que $E(Y_1) = \lambda_1$ et $E(Y_2) = \lambda_2$. On dit alors que la variable aléatoire $X = Y_1 - Y_2$ obéit à la loi de Skellam de paramètres (λ_1, λ_2) , notée $S(\lambda_1, \lambda_2)$. La fonction de masse de X est donnée par l'expression suivante, valable pour tout $x \in \mathbb{Z}$:

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

Par définition, la loi de Skellam est symétrique, au sens où

$$X \sim \mathcal{S}(\lambda_1, \lambda_2) \quad \Leftrightarrow \quad -X \sim \mathcal{S}(\lambda_2, \lambda_1).$$

De plus, la fonction génératrice des moments de la loi $\mathcal{S}(\lambda_1, \lambda_2)$ est donnée de façon explicite en tout $t \in \mathbb{R}$ par

$$M_X(t) = \mathbb{E}(e^{tX}) = \exp\{-(\lambda_1 + \lambda_2) + \lambda_1 e^t + \lambda_2 e^{-t}\}. \quad (1.1)$$

Cette expression permet de calculer facilement les premiers moments de la loi de Skellam univariée. En particulier, on trouve

$$\mathbb{E}(X) = \lambda_1 - \lambda_2, \quad \text{var}(X) = \lambda_1 + \lambda_2. \quad (1.2)$$

De plus, les coefficient d'asymétrie et d'aplatissement normalisé de la loi $\mathcal{S}(\lambda_1, \lambda_2)$ sont respectivement donnés par

$$\beta_1 = \frac{\lambda_1 - \lambda_2}{(\lambda_1 + \lambda_2)^{3/2}}, \quad \beta_2 = 3 + \frac{1}{\lambda_1 + \lambda_2}.$$

La loi de Skellam possède en outre une intéressante propriété de stabilité par convolution, tel qu'énoncé ci-dessous.

Proposition 1.1.1. *Soient $X_1, \dots, X_n$ des variables aléatoire telles que $X_i \sim \mathcal{S}(\lambda_{i1}, \lambda_{i2})$ pour tout $i \in \{1, \dots, n\}$. Alors*

$$\sum_{i=1}^n X_i \sim \mathcal{S}\left(\sum_{i=1}^n \lambda_{i1}, \sum_{i=1}^n \lambda_{i2}\right).$$

Démonstration. Le fait que $X_i \sim \mathcal{S}(\lambda_{i1}, \lambda_{i2})$ pour tout $i \in \{1, \dots, n\}$ entraîne l'existence de variables aléatoires de Poisson indépendantes, disons $Y_{i1} \sim \mathcal{P}(\lambda_{i1})$ et $Y_{i2} \sim \mathcal{P}(\lambda_{i2})$, telles que $X_i = Y_{i1} - Y_{i2}$. Il en résulte que

$$\sum_{i=1}^n X_i = \sum_{i=1}^n Y_{i1} - \sum_{i=1}^n Y_{i2}.$$

Or, les variables $Y_{1k}, \dots, Y_{nk}$ étant mutuellement indépendantes pour $k \in \{1, 2\}$, on a

$$\sum_{i=1}^n Y_{i1} \sim \mathcal{P}\left(\sum_{i=1}^n \lambda_{i1}\right).$$

La conclusion s'ensuit. □

1.2 Autres représentations de la loi de Skellam

La fonction de masse de la loi de Skellam peut aussi s'exprimer en termes de la fonction de Bessel modifiée de première espèce définie en tous $x \in \mathbb{Z}$ et $y > 0$ par

$$I_x(y) = \left(\frac{y}{2}\right)^x \sum_{k=0}^{\infty} \frac{(y^2/4)^k}{k!(x+k)!}.$$

Prékopa (1952) montre en effet que si $X \sim \mathcal{S}(\lambda_1, \lambda_2)$, alors pour tout $x \in \mathbb{Z}$, on a

$$\Pr(X = x) = e^{-\lambda_1 - \lambda_2} \left(\frac{\lambda_1}{\lambda_2}\right)^{x/2} I_x\left(2\sqrt{\lambda_1 \lambda_2}\right). \quad (1.3)$$

Cette expression est rarement utilisée en pratique, compte tenu des difficultés de calcul qu'elle présente. En conjuguant les formules (1.1) et (1.3), on peut toutefois déduire certaines identités propres à la fonction de Bessel modifiée de première espèce.

Commençons par rappeler quelques propriétés connues de cette fonction spéciale. Pour tous $\lambda > 0$ et $x \in \mathbb{Z}$, on a

$$I_x(\lambda) = I_{-x}(\lambda), \quad \sum_{y \in \mathbb{Z}} I_y(\lambda) = e^\lambda, \quad \sum_{y \in \mathbb{Z}} y I_y(\lambda) = 0,$$

ainsi que

$$I_x(\lambda) = \left(\frac{\lambda}{2}\right)^x {}_0\tilde{F}_1\left(; x+1; \frac{\lambda^2}{4}\right), \quad (1.4)$$

où

$${}_0\tilde{F}_1(; b; z) = \sum_{k=0}^{\infty} \frac{z^k}{k! \Gamma(b+k)}$$

est une fonction hypergéométrique régularisée et $\Gamma(\cdot)$ dénote la célèbre fonction Gamma d'Euler, qui généralise la notion de factorielle.

Le résultat suivant énonce d'autres relations satisfaites par la fonction de Bessel qui peuvent être facilement déduites des formules (1.1) et (1.3).

Proposition 1.2.1. *Pour tous $\lambda > 0$, $\lambda_1 > 0$ et $\lambda_2 > 0$, on a*

$$\sum_{x \in \mathbb{Z}} \left(\frac{\lambda_1}{\lambda_2}\right)^{x/2} I_x\left(2\sqrt{\lambda_1 \lambda_2}\right) = e^{\lambda_1 + \lambda_2},$$

$$\sum_{x \in \mathbb{Z}} x \left(\frac{\lambda_1}{\lambda_2}\right)^{x/2} I_x\left(2\sqrt{\lambda_1 \lambda_2}\right) = (\lambda_1 - \lambda_2) e^{\lambda_1 + \lambda_2},$$

$$\sum_{x \in \mathbb{Z}} x^2 \left(\frac{\lambda_1}{\lambda_2} \right)^{x/2} I_x \left(2\sqrt{\lambda_1 \lambda_2} \right) = \{\lambda_1 + \lambda_2 + (\lambda_1 - \lambda_2)^2\} e^{\lambda_1 + \lambda_2},$$

$$\sum_{x \in \mathbb{Z}} x^2 I_x(\lambda) = \lambda e^\lambda$$

et

$$\sum_{x \in \mathbb{Z}} x^4 I_x(\lambda) = \lambda e^\lambda (3\lambda + 1).$$

Démonstration. La première identité découle du fait que l'expression (1.3) est celle d'une fonction de probabilité. Les deux identités suivantes sont des réécritures des formules données en (1.2) pour la moyenne et la variance de la loi de Skellam univariée. La quatrième identité est un cas spécial de la troisième obtenu en posant $\lambda_1 = \lambda_2 = \lambda/2$. La dernière expression découle du fait que le quatrième cumulant de la loi est $3\mu_2^2$. $\square$

Puisque la fonction de Bessel modifiée de première espèce est liée à la fonction hypergéométrique ${}_0\tilde{F}_1$ par l'identité (1.4), on peut en déduire d'autres expressions pour la loi de Skellam univariée. En invoquant le fait que pour toute valeur entière de x ,

$${}_0\tilde{F}_1(; -x + 1; \lambda) = \lambda^x {}_0\tilde{F}_1(; x + 1; \lambda),$$

on trouve facilement que

$$\begin{aligned} \Pr(X = x) &= e^{-\lambda_1 - \lambda_2} \lambda_1^x {}_0\tilde{F}_1(; x + 1; \lambda_1 \lambda_2) \\ &= e^{-\lambda_1 - \lambda_2} \lambda_2^{-x} {}_0\tilde{F}_1(; -x + 1; \lambda_1 \lambda_2) \\ &= e^{-\lambda_1 - \lambda_2} (\lambda_1 \lambda_2)^{\max(0, -x)} \lambda_1^x {}_0\tilde{F}_1(; |x| + 1; \lambda_1 \lambda_2). \end{aligned} \quad (1.5)$$

De fait, la formule (1.5) est plus naturelle que celle fondée sur la fonction de Bessel modifiée de première espèce puisqu'en procédant par simple conditionnement, on voit directement que si la variable $X = Y_1 - Y_2$ est tel que précisé à la définition 1.1.1, alors

$$\begin{aligned} \Pr(X = x) &= \sum_{y \in \mathbb{N}} \Pr(Y_1 - Y_2 = x | Y_2 = y) \Pr(Y_2 = y) \\ &= \sum_{y \in \mathbb{N}} \Pr(Y_1 = x + y) \Pr(Y_2 = y) \\ &= e^{-\lambda_1 - \lambda_2} \lambda_1^x \sum_{k=\max(-x, 0)}^{\infty} \frac{(\lambda_1 \lambda_2)^k}{k! (x + k)!} \\ &= e^{-\lambda_1 - \lambda_2} \lambda_1^x {}_0\tilde{F}_1(; x + 1; \lambda_1 \lambda_2) \end{aligned}$$

en tout $x \in \mathbb{Z}$. De plus, on a immédiatement

$$\Pr(X = x) = e^{-\lambda_1} \lambda_1^x {}_0\tilde{F}_1(; x + 1; 0) = \frac{e^{-\lambda_1} \lambda_1^x}{x!}$$

en tout $x \in \mathbb{Z}$ dans le cas particulier où $\lambda_2 = 0$, c'est-à-dire quand $\mathcal{S}(\lambda_1, 0) \equiv \mathcal{P}(\lambda_1)$. Cette correspondance est moins évidente lorsque l'on se réfère à la représentation fondée sur la fonction de Bessel modifiée de première espèce.

1.3 Estimation des paramètres

Nous allons maintenant voir comment il est possible d'estimer les paramètres de la loi de Skellam par la méthode des moments et la méthode du maximum de vraisemblance.

1.3.1 La méthode des moments

Soit $X_1, \dots, X_n$ un échantillon aléatoire de la loi de $\mathcal{S}(\lambda_1, \lambda_2)$. Soient en outre

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i, \quad S_n^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2,$$

la moyenne et la variance de l'échantillon. La méthode des moments consiste à estimer les paramètres λ_1 et λ_2 par les formules

$$\hat{\lambda}_{1MM} = \frac{1}{2} (S_n^2 + \bar{X}_n), \quad \hat{\lambda}_{2MM} = \frac{1}{2} (S_n^2 - \bar{X}_n).$$

Ces estimateurs des moments sont sans biais. Toutefois, ils n'ont aucun sens lorsque $S_n^2 - |\bar{X}_n| < 0$, puisqu'ils conduisent alors à une estimation négative de λ_1 ou de λ_2 . Cette situation peut se produire, notamment quand l'un des paramètres est très petit par rapport à l'autre, par exemple lorsque $\max(\lambda_1/\lambda_2, \lambda_2/\lambda_1) \geq 10$ comme mentionné dans l'article de Alzaid & Omair (2010). Afin de résoudre ce problème, on peut convenir de poser l'estimation négative égale à zéro, dans lequel cas l'autre estimation est alors égale à la valeur absolue de la moyenne.

1.3.2 La méthode du maximum de vraisemblance

Considérons à nouveau un échantillon aléatoire $X_1, \dots, X_n$ de la loi de Skellam $\mathcal{S}(\lambda_1, \lambda_2)$. La fonction de vraisemblance est donnée par

$$L(x_1, \dots, x_n, \lambda_1, \lambda_2) = \prod_{i=1}^n \Pr(X_i = x_i) = \prod_{i=1}^n \left\{ e^{-\lambda_1 - \lambda_2} \left(\frac{\lambda_1}{\lambda_2} \right)^{x_i/2} I_{x_i} \left(2\sqrt{\lambda_1 \lambda_2} \right) \right\}.$$

Les équations à résoudre sont obtenues en dérivant son logarithme naturel par rapport à λ_1 et à λ_2 . On utilise pour ce fait les relations

$$\begin{aligned}\frac{\partial}{\partial \lambda} I_x(\lambda) &= \frac{x}{\lambda} I_x(\lambda) + I_{x+1}(\lambda), \\ I_x(\lambda) &= \frac{2(x+1)}{\lambda} I_{x+1}(\lambda) + I_{x+2}(\lambda),\end{aligned}$$

valables en tout $x \in \mathbb{Z}$. On obtient ainsi

$$\begin{aligned}\frac{\partial}{\partial \lambda_1} \ln\{L(x_1, \dots, x_n, \lambda_1, \lambda_2)\} \\ = -n + \frac{n\bar{x}}{2\lambda_1} + \frac{\lambda_2}{\sqrt{\lambda_1\lambda_2}} \sum_{i=1}^n \frac{\frac{x_i}{2\sqrt{\lambda_1\lambda_2}} I_{x_i}(2\sqrt{\lambda_1\lambda_2}) + I_{x_i+1}(2\sqrt{\lambda_1\lambda_2})}{I_{x_i}(2\sqrt{\lambda_1\lambda_2})} \\ = -n + \frac{n\bar{x}}{\lambda_1} + \frac{\lambda_2}{\sqrt{\lambda_1\lambda_2}} \sum_{i=1}^n \frac{I_{x_i+1}(2\sqrt{\lambda_1\lambda_2})}{I_{x_i}(2\sqrt{\lambda_1\lambda_2})}\end{aligned}$$

et

$$\begin{aligned}\frac{\partial}{\partial \lambda_2} \ln\{L(x_1, \dots, x_n, \lambda_1, \lambda_2)\} \\ = -n - \frac{n\bar{x}}{2\lambda_2} + \frac{\lambda_1}{\sqrt{\lambda_1\lambda_2}} \sum_{i=1}^n \frac{\frac{x_i}{2\sqrt{\lambda_1\lambda_2}} I_{x_i}(2\sqrt{\lambda_1\lambda_2}) + I_{x_i+1}(2\sqrt{\lambda_1\lambda_2})}{I_{x_i}(2\sqrt{\lambda_1\lambda_2})} \\ = -n + \frac{\lambda_1}{\sqrt{\lambda_1\lambda_2}} \sum_{i=1}^n \frac{I_{x_i+1}(2\sqrt{\lambda_1\lambda_2})}{I_{x_i}(2\sqrt{\lambda_1\lambda_2})},\end{aligned}$$

expressions dans lesquelles $\bar{x} = (x_1 + \dots + x_n)/n$. Les estimateurs à vraisemblance maximale $\hat{\lambda}_{1EMV}$ et $\hat{\lambda}_{2EMV}$ peuvent donc être obtenus en résolvant les deux équations non linéaires suivantes :

$$\frac{n\bar{X}_n}{\hat{\lambda}_{1EMV}} + \frac{\hat{\lambda}_{2EMV}}{\sqrt{\hat{\lambda}_{1EMV}\hat{\lambda}_{2EMV}}} \sum_{i=1}^n \frac{I_{X_i+1}(2\sqrt{\hat{\lambda}_{1EMV}\hat{\lambda}_{2EMV}})}{I_{X_i}(2\sqrt{\hat{\lambda}_{1EMV}\hat{\lambda}_{2EMV}})} = n \quad (1.6)$$

et

$$\frac{\hat{\lambda}_{1EMV}}{\sqrt{\hat{\lambda}_{1EMV}\hat{\lambda}_{2EMV}}} \sum_{i=1}^n \frac{I_{X_i+1}(2\sqrt{\hat{\lambda}_{1EMV}\hat{\lambda}_{2EMV}})}{I_{X_i}(2\sqrt{\hat{\lambda}_{1EMV}\hat{\lambda}_{2EMV}})} = n. \quad (1.7)$$

En multipliant l'équation (1.6) par $\hat{\lambda}_{1EMV}$ et l'équation (1.7) par $\hat{\lambda}_{2EMV}$, on obtient, après soustraction,

$$n\bar{X}_n - n\hat{\lambda}_{1EMV} + n\hat{\lambda}_{2EMV} = 0$$

ou encore

$$\hat{\lambda}_{1EMV} = \hat{\lambda}_{2EMV} + \bar{X}_n. \quad (1.8)$$

Puis, en substituant l'équation (1.8) dans l'équation (1.7), on trouve

$$\frac{\hat{\lambda}_{2EMV} + \bar{X}_n}{\sqrt{(\hat{\lambda}_{2EMV} + \bar{X}_n)\hat{\lambda}_{2EMV}}} \sum_{i=1}^n \frac{I_{X_i+1} \left(2\sqrt{(\hat{\lambda}_{2EMV} + \bar{X}_n)\hat{\lambda}_{2EMV}} \right)}{I_{X_i}(2\sqrt{(\hat{\lambda}_{2EMV} + \bar{X}_n)\hat{\lambda}_{2EMV}})} = n. \quad (1.9)$$

Par conséquent, on peut déterminer $\hat{\lambda}_{2EMV}$ en résolvant l'équation non linéaire (1.9) et, par suite, calculer $\hat{\lambda}_{1EMV}$ en utilisant l'équation (1.8). En faisant appel à l'identité

$$\frac{\partial}{\partial \lambda} {}_0\tilde{F}_1(; -x+1; \lambda) = {}_0\tilde{F}_1(; -x+2; \lambda),$$

valide en tout $x \in \mathbb{Z}$, on peut aussi déduire l'estimation à vraisemblance maximale λ_{2EMV} de la formule (1.5) en résolvant l'équation non linéaire ci-dessous :

$$(\hat{\lambda}_{2EMV} + \bar{X}_n) \sum_{i=1}^n \frac{{}_0\tilde{F}_1\{; X_i+2; (\hat{\lambda}_{2EMV} + \bar{X}_n)\hat{\lambda}_{2EMV}\}}{{}_0\tilde{F}_1\{; X_i+1; (\hat{\lambda}_{2EMV} + \bar{X}_n)\hat{\lambda}_{2EMV}\}} = n.$$

1.4 Loi asymptotique des estimateurs à vraisemblance maximale

Alzaid & Omair (2010) ont montré que quand $n \rightarrow \infty$, la loi de l'estimateur $\hat{\lambda}_n = (\hat{\lambda}_{1EMV}, \hat{\lambda}_{2EMV})$ du couple de paramètres $\lambda = (\lambda_1, \lambda_2)$ est approximativement gaussienne, d'espérance λ et de matrice de covariance I_λ^{-1}/n , où I_λ est la matrice d'information de Fisher dont l'élément (j, k) est donné en tous $j, k \in \{1, 2\}$ par

$$I_{j,k}(\lambda) = E \left[\frac{\partial^2}{\partial \lambda_j \partial \lambda_k} \ln\{L(\lambda)\} \right].$$

De façon plus formelle, ces auteurs ont montré que lorsque $n \rightarrow \infty$, le vecteur $\sqrt{n}(\hat{\lambda}_n - \lambda)$ converge faiblement vers la loi $\mathcal{N}_2(0, I^{-1}(\lambda))$.

Sous des conditions de régularité faibles, la matrice d'information observée $I(\hat{\lambda})$ est un estimateur convergent de $I(\lambda)$. Cet estimateur de type "plug-in", viz.

$$I(\hat{\lambda}) = \begin{pmatrix} I_{11}(\hat{\lambda}) & I_{12}(\hat{\lambda}) \\ I_{21}(\hat{\lambda}) & I_{22}(\hat{\lambda}) \end{pmatrix},$$

peut être déduit comme suit au moyen de la formule (1.5) :

$$\begin{aligned}
 I_{11}(\hat{\lambda}) &= -\frac{\partial^2}{\partial \lambda_1^2} \ln L \big|_{\lambda_1=\hat{\lambda}_1} \\
 &= \frac{n\bar{X}_n}{\hat{\lambda}_{1EMV}^2} + \hat{\lambda}_{1EMV}^2 \sum_{i=1}^n \left\{ \frac{{}_0\tilde{F}_1(; X_i + 1; \hat{\lambda}_{1EMV} \hat{\lambda}_{2EMV}) {}_0\tilde{F}_1(; X_i + 3; \hat{\lambda}_{1EMV} \hat{\lambda}_{2EMV})}{{}_0\tilde{F}_1(; X_i + 1; \hat{\lambda}_{1EMV} \hat{\lambda}_{2EMV})} \right. \\
 &\quad \left. - \frac{{}_0\tilde{F}_1(; X_i + 2; \hat{\lambda}_{1EMV} \hat{\lambda}_{2EMV})^2}{{}_0\tilde{F}_1(; X_i + 1; \hat{\lambda}_{1EMV} \hat{\lambda}_{2EMV})^2} \right\},
 \end{aligned}$$

$$\begin{aligned}
 I_{22}(\hat{\lambda}) &= -\frac{\partial^2}{\partial \lambda_2^2} \ln L \big|_{\lambda_2=\hat{\lambda}_2} \\
 &= -\hat{\lambda}_{1EMV}^2 \sum_{i=1}^n \left\{ \frac{{}_0\tilde{F}_1(; X_i + 1; \hat{\lambda}_{1EMV} \hat{\lambda}_{2EMV}) {}_0\tilde{F}_1(; X_i + 3; \hat{\lambda}_{1EMV} \hat{\lambda}_{2EMV})}{{}_0\tilde{F}_1(; X_i + 1; \hat{\lambda}_{1EMV} \hat{\lambda}_{2EMV})} \right. \\
 &\quad \left. - \frac{{}_0\tilde{F}_1(; X_i + 2; \hat{\lambda}_{1EMV} \hat{\lambda}_{2EMV})^2}{{}_0\tilde{F}_1(; X_i + 1; \hat{\lambda}_{1EMV} \hat{\lambda}_{2EMV})^2} \right\},
 \end{aligned}$$

et

$$\begin{aligned}
 I_{12}(\hat{\lambda}) &= I_{21}(\hat{\lambda}) \\
 &= -\frac{\partial^2}{\partial \lambda_1 \partial \lambda_2} \ln L \big|_{\lambda_1=\hat{\lambda}_1, \lambda_2=\hat{\lambda}_2} \\
 &= -\sum_{i=1}^n \frac{{}_0\tilde{F}_1(; X_i + 2; \hat{\lambda}_{1EMV} \hat{\lambda}_{2EMV})}{{}_0\tilde{F}_1(; X_i + 1; \hat{\lambda}_{1EMV} \hat{\lambda}_{2EMV})} \\
 &= -\hat{\lambda}_{1EMV} \hat{\lambda}_{2EMV} \sum_{i=1}^n \left\{ \frac{{}_0\tilde{F}_1(; X_i + 1; \hat{\lambda}_{1EMV} \hat{\lambda}_{2EMV}) {}_0\tilde{F}_1(; X_i + 3; \hat{\lambda}_{1EMV} \hat{\lambda}_{2EMV})}{{}_0\tilde{F}_1(; X_i + 1; \hat{\lambda}_{1EMV} \hat{\lambda}_{2EMV})} \right. \\
 &\quad \left. - \frac{{}_0\tilde{F}_1(; X_i + 2; \hat{\lambda}_{1EMV} \hat{\lambda}_{2EMV})^2}{{}_0\tilde{F}_1(; X_i + 1; \hat{\lambda}_{1EMV} \hat{\lambda}_{2EMV})^2} \right\}.
 \end{aligned}$$

Au vu de ces résultats asymptotiques, on peut donc obtenir des intervalles de confiance approximatifs pour λ_1 et λ_2 . Si $z_{\alpha/2} = \Phi^{-1}(\alpha/2)$ dénote le quantile d'ordre $\alpha/2$ de la loi normale centrée réduite, les intervalles ayant pour bornes

$$\hat{\lambda}_1 \pm z_{\alpha/2} \sqrt{\frac{I_{22}}{I_{11}I_{22} - I_{12}^2}}, \quad \hat{\lambda}_2 \pm z_{\alpha/2} \sqrt{\frac{I_{11}}{I_{11}I_{22} - I_{12}^2}}$$

ont un degré de couverture approximatif de $100 \times (1 - \alpha)\%$.

1.5 Tests du rapport des vraisemblances maximales

Dans leur article, Alzaid & Omais (2010) expliquent aussi comment il est possible de confronter les hypothèses

$\mathcal{H}_0$: Les données sont issues de la loi $\mathcal{S}(\lambda, \lambda)$,

$\mathcal{H}_1$: Les données sont issues de la loi $\mathcal{S}(\lambda_1, \lambda_2)$

au moyen d'un test statistique.

Étant donné un échantillon aléatoire $X_1, \dots, X_n$ de même loi que X , la statistique du test consiste à évaluer le rapport des vraisemblances maximales, à savoir

$$\Lambda_n = \frac{L(X_1, \dots, X_n, \hat{\lambda})}{L(X_1, \dots, X_n, \hat{\lambda}_1, \hat{\lambda}_2)},$$

où $L(x_1, \dots, x_n, \hat{\lambda}_n)$ est la probabilité de l'échantillon calculée sous $\mathcal{H}_0$ en supposant que λ est égal à l'estimateur du maximum de vraisemblance, tandis que $L(x_1, \dots, x_n, \hat{\lambda}_1, \hat{\lambda}_2)$ représente la probabilité de l'échantillon calculée sous $\mathcal{H}_1$ au moyen des estimations à vraisemblance maximale de λ_1 et λ_2 . Notons que sous $\mathcal{H}_0$, la vraisemblance est

$$L(x_1, \dots, x_n, \lambda) = e^{-2n\lambda} \lambda^{n\bar{x}} \prod_{i=1}^n {}_0\tilde{F}_1(; x_i + 1; \lambda^2),$$

où $\bar{x} = (x_1 + \dots + x_n)/n$. La log-vraisemblance est donc de la forme

$$\ln\{L(x_1, \dots, x_n, \lambda)\} = -2n\lambda + n\bar{x} \ln(\lambda) + \sum_{i=1}^n \ln\{{}_0\tilde{F}_1(; x_i + 1; \lambda^2)\}.$$

L'estimation à vraisemblance maximale $\hat{\lambda}$ de λ est obtenue en résolvant l'équation non linéaire

$$\frac{\partial}{\partial \lambda} \ln\{L(x_1, \dots, x_n, \lambda)\} = -2n + \frac{n\bar{x}}{\lambda} + 2\lambda \sum_{i=1}^n \frac{{}_0\tilde{F}_1(; x_i + 2; \lambda^2)}{{}_0\tilde{F}_1(; x_i + 1; \lambda^2)} = 0.$$

Par conséquent,

$$L(x_1, \dots, x_n, \hat{\lambda}) = e^{-2n\hat{\lambda}} \hat{\lambda}^{n\bar{x}} \prod_{k=1}^n {}_0\tilde{F}_1(; x_k + 1; \hat{\lambda}^2).$$

En revanche sous $\mathcal{H}_1$, la vraisemblance maximale est donnée par

$$L(x_1, \dots, x_n, \hat{\lambda}_1, \hat{\lambda}_2) = e^{-n\hat{\lambda}_1 - n\hat{\lambda}_2} \hat{\lambda}_1^{n\bar{x}} \prod_{i=1}^n {}_0\tilde{F}_1(; x_i + 1; \hat{\lambda}_1 \hat{\lambda}_2).$$

Il s'ensuit que la statistique du test est

$$\begin{aligned} -2 \ln(\Lambda_n) = -2 \Big[& -2n\hat{\lambda} + n\bar{X}_n \ln(\hat{\lambda}) + \sum_{i=1}^n \ln\{{}_0\tilde{F}_1(; X_i + 1; \hat{\lambda}^2)\} \\ & + n\hat{\lambda}_1 + n\hat{\lambda}_2 - n\bar{X}_n \ln(\hat{\lambda}_1) - \sum_{i=1}^n \ln\{{}_0\tilde{F}_1(; X_i + 1; \hat{\lambda}_1 \hat{\lambda}_2)\} \Big]. \end{aligned}$$

Wilks (1938) a étudié le comportement asymptotique de la statistique $-2 \ln(\Lambda_n)$. Il a montré que cette dernière converge vers une loi du khi-deux à un degré de liberté lorsque n tend vers l'infini. On rejette donc $\mathcal{H}_0$ au seuil α si $-2 \ln(\Lambda_n) > \chi_{1-\alpha,1}^2$.

Lorsque l'on veut vérifier si le modèle de Poisson est plus approprié que le modèle de Skellam, on peut construire d'une manière semblable un test du rapport des vraisemblances maximales permettant de confronter les hypothèses

$\mathcal{H}_0$: Les données sont issues de la loi $\mathcal{S}(\lambda, 0)$, à savoir une $\mathcal{P}(\lambda)$,

$\mathcal{H}_1$: Les données sont tirées de $\mathcal{S}(\lambda_1, \lambda_2)$.

Supposons pour ce faire que les valeurs prises par $X_1, \dots, X_n$ soient toutes entières et non négatives. La statistique du test est alors donnée par

$$\Lambda_n = \frac{L(X_1, \dots, X_n, \hat{\lambda})}{L(X_1, \dots, X_n, \hat{\lambda}_1, \hat{\lambda}_2)},$$

où $L(x_1, \dots, x_n, \hat{\lambda})$ représente la probabilité de l'échantillon sous $\mathcal{H}_0$ calculée en utilisant l'estimation du maximum de vraisemblance de λ et $L(x_1, \dots, x_n, \hat{\lambda}_1, \hat{\lambda}_2)$ représente la probabilité de l'échantillon sous $\mathcal{H}_1$ calculée aux valeurs estimées de λ_1 et λ_2 obtenues par la méthode du maximum de vraisemblance. Sous $\mathcal{H}_0$, $\hat{\lambda} = \bar{x} = (x_1 + \dots + x_n)/n$ puisque les données proviennent d'une loi de Poisson. Ainsi,

$$L(x_1, \dots, x_n, \hat{\lambda}) = e^{-2n\bar{x}} \bar{x}^{n\bar{x}} / \prod_{i=1}^n x_i!$$

Sous $\mathcal{H}_1$, on a toutefois

$$L(x_1, \dots, x_n, \hat{\lambda}_1, \hat{\lambda}_2) = e^{-n\hat{\lambda}_1 - n\hat{\lambda}_2} \hat{\lambda}_1^{n\bar{x}} \prod_{i=1}^n {}_0\tilde{F}_1(; x_i + 1; \hat{\lambda}_1 \hat{\lambda}_2).$$

Par conséquent, la statistique du test est

$$\begin{aligned} -2 \ln(\Lambda_n) = -2 \Big[& -2n\bar{X}_n + n\bar{X}_n \ln(\bar{X}_n) - \sum_{i=1}^n \ln(X_i!) \\ & + n\hat{\lambda}_1 + n\hat{\lambda}_2 - n\bar{X}_n \ln(\hat{\lambda}_1) - \sum_{i=1}^n \ln\{{}_0\tilde{F}_1(; X_i + 1; \hat{\lambda}_1 \hat{\lambda}_2)\} \Big]. \end{aligned}$$

Pour n assez grand, la loi de la statistique $-2\ln(\Lambda_n)$ s'approche d'une khi-deux à un degré de liberté. On rejette donc $\mathcal{H}_0$ au seuil α si $-2\ln(\Lambda_n) > \chi^2_{1-\alpha,1}$.

Chapitre 2

La loi de Skellam bivariée

Quelques généralisations possibles de la loi de Skellam univariée ont récemment été proposées. Ce chapitre donne un compte rendu de ces modèles. Nous commençons par étudier le modèle de Skellam bivarié introduit dans l'article de Bulla et coll. (2012). Après une brève critique de ce modèle, nous examinons deux modèles de Skellam alternatifs récemment proposés par Genest & Mesfioui (2014). Nous présentons ensuite plusieurs propriétés de ces lois de Skellam bivariées dites de première et de deuxième espèce. En particulier, nous montrons que ces modèles sont stables par convolution et monotones par rapport aux paramètres de dépendance. Finalement, nous exposons une façon d'estimer les paramètres de ces modèles par la méthode des moments.

2.1 Définition et propriétés de base

Nous commençons par définir la version de la loi de Skellam bivarié définie dans l'article de Bulla et coll. (2012).

Définition 2.1.1. Soient $\lambda_0 \geq 0$, $\lambda_1 \geq 0$ et $\lambda_2 \geq 0$. On dit que le couple aléatoire (X_1, X_2) suit une loi de Skellam bivariée, notée $\mathcal{BS}(\lambda_0, \lambda_1, \lambda_2)$, si et seulement si

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

où Y_0 , Y_1 et Y_2 sont trois variables aléatoires de Poisson mutuellement indépendantes telles que $E(Y_j) = \lambda_j$ pour tout $j \in \{0, 1, 2\}$.

La fonction de masse de (X_1, X_2) est donnée en tout $(x_1, x_2) \in \mathbb{Z}^2$ par

$$\Pr(X_1 = x_1, X_2 = x_2) = e^{-(\lambda_1 + \lambda_2 + \lambda_0)} \lambda_1^{x_1} \lambda_2^{x_2} \sum_{k=\max(0, -x_1, -x_2)}^{\infty} \frac{(\lambda_0 \lambda_1 \lambda_2)^k}{(x_1 + k)! (x_2 + k)! k!}.$$

Il est clair que les variables aléatoires marginales du couple (X_1, X_2) suivent des lois de Skellam univariées, à savoir $X_1 \sim \mathcal{S}(\lambda_1, \lambda_0)$ et $X_2 \sim \mathcal{S}(\lambda_2, \lambda_0)$. Puisque le paramètre λ_0 affecte les lois marginales X_1 et X_2 , il ne peut pas être considéré comme un véritable paramètre de dépendance. Nous allons voir plus loin que les paramètres de dépendance des lois de Skellam bivariées définies dans le travail de Genest & Mesfioui (2014) n'affectent pas les marges.

Voici quelques résultats liés au modèle de Skellam bivarié de Bulla et coll. (2012).

Proposition 2.1.1. *Soit (X_1, X_2) un couple de loi $\mathcal{BS}(\lambda_0, \lambda_1, \lambda_2)$ pour certains choix de paramètres $\lambda_0 \geq 0$, $\lambda_1 \geq 0$, $\lambda_2 \geq 0$.*

a) *On a $E(X_1, X_2) = (\lambda_1 - \lambda_0, \lambda_2 - \lambda_0)$ et*

$$\text{cov}(X_1, X_2) = \Sigma = \begin{pmatrix} \lambda_1 + \lambda_0 & \lambda_0 \\ \lambda_0 & \lambda_2 + \lambda_0 \end{pmatrix}.$$

b) *Le coefficient de corrélation du couple (X_1, X_2) est donné par*

$$\text{corr}(X_1, X_2) = \frac{\lambda_0}{\sqrt{(\lambda_1 + \lambda_0)(\lambda_2 + \lambda_0)}}.$$

c) *La fonction caractéristique de (X_1, X_2) est donnée en tout $(t_1, t_2) \in \mathbb{Z}^2$ par*

$$\phi_{(X_1, X_2)}(t_1, t_2) = e^{\lambda_1(e^{it_1} - 1)} e^{\lambda_2(e^{it_2} - 1)} e^{\lambda_0(e^{-it_1} - 1)(e^{-it_2} - 1)}.$$

En utilisant la fonction caractéristique de $(X_1, X_2) \sim \mathcal{BS}(\lambda_0, \lambda_1, \lambda_2)$, la fonction de masse de (X_1, X_2) peut être exprimée en tout $(x_1, x_2) \in \mathbb{Z}^2$ par

$$\Pr(X_1 = x_1, X_2 = x_2) = \frac{1}{(2\pi)^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi_{(X_1, X_2)}(t_1, t_2) e^{-kx_1 t_1} e^{-kx_2 t_2} dt_1 dt_2.$$

En particulier, on a l'identité suivante :

$$\begin{aligned} & \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} e^{\lambda_1(e^{kt_1} - 1)} e^{\lambda_2(e^{kt_2} - 1)} e^{\lambda_0(e^{kt_1} - 1)(e^{kt_2} - 1)} e^{-kx_1 t_1} e^{-kx_2 t_2} dt_1 dt_2 \\ &= (2\pi)^2 e^{-(\lambda_1 + \lambda_2 + \lambda_0)} \lambda_1^{x_1} \lambda_2^{x_2} \sum_{k=0}^{\infty} \frac{(\lambda_0 \lambda_1 \lambda_2)^k}{(x_1 + k)! (x_2 + k)! k!}. \end{aligned}$$

La proposition suivante, due à Bulla et coll. (2012), donne une approximation de la loi $\mathcal{BS}(\lambda_0, \lambda_1, \lambda_2)$ par la loi normale bivariée.

Proposition 2.1.2. *Soient $\lambda_0 \geq 0$, $\lambda_1 \geq 0$ et $\lambda_2 \geq 0$. Pour λ_1 et λ_2 assez grand, on a l'approximation suivante :*

$$\mathcal{BS}(\lambda_0, \lambda_1, \lambda_2) \approx \mathcal{N}_2 \left(\begin{pmatrix} \lambda_1 - \lambda_0 \\ \lambda_2 - \lambda_0 \end{pmatrix}, \begin{pmatrix} \lambda_1 - \lambda_0 & \lambda_0 \\ \lambda_0 & \lambda_2 - \lambda_0 \end{pmatrix} \right).$$

2.2 Méthodes d'estimation des paramètres

Cette section est consacrée à l'estimation des paramètres de la loi $\mathcal{BS}(\lambda_0, \lambda_1, \lambda_2)$ par la méthode des moments et celle du maximum de vraisemblance. Ces procédures ont été développées dans l'article de Bulla et coll. (2012).

2.2.1 La méthode des moments

Soit $(X_{11}, X_{12}), \dots, (X_{n1}, X_{n2})$ un échantillon aléatoire de loi $\mathcal{BS}(\lambda_0, \lambda_1, \lambda_2)$. Posons

$$\bar{X}_{n1} = \frac{1}{n} \sum_{i=1}^n X_{i1}, \quad \bar{X}_{n2} = \frac{1}{n} \sum_{i=1}^n X_{i2}$$

et

$$S_{n12} = \frac{1}{n-1} \sum_{i=1}^n (X_{i1} - \bar{X}_{n1})(X_{i2} - \bar{X}_{n2}).$$

La méthode des moments consiste à trouver les valeurs des paramètres pour lesquelles les moments théoriques sont égaux aux moments échantillonnaires. Au vu de la proposition 2.1.1, on trouve

$$\hat{\lambda}_0 = S_{n12}, \quad \hat{\lambda}_1 = \bar{X}_{n1} + \hat{\lambda}_0, \quad \hat{\lambda}_2 = \bar{X}_{n2} + \hat{\lambda}_0.$$

Notons que l'estimation par la méthode des moments de λ_0 n'existe pas si $S_{n12} < 0$. En effet, ce modèle induit toujours une dépendance positive entre les variables.

2.2.2 La méthode du maximum de vraisemblance

La fonction de masse du modèle de Bulla et coll. (2012) est donnée en tous $\lambda = (\lambda_0, \lambda_1, \lambda_2) \in (0, \infty)^3$ et $(x_1, x_2) \in \mathbb{Z}^2$ par

$$G(\lambda, x_1, x_2) = \lambda_1^{x_1} \lambda_2^{x_2} \sum_{k=0}^{\infty} \frac{(\lambda_0 \lambda_1 \lambda_2)^k}{(x_1 + k)! (x_2 + k)! k!}.$$

Afin de maximiser la vraisemblance, on doit déterminer les points critiques de cette fonction. Le résultat suivant est utile à cette fin.

Proposition 2.2.1. *Pour tout $(x_1, x_2) \in \mathbb{Z}^2$, on a*

$$\begin{aligned}\frac{\partial}{\partial \lambda_1} G(\lambda, x_1, x_2) &= \frac{x_1}{\lambda_1} G(\lambda, x_1, x_2) + \frac{\lambda_0}{\lambda_1} G(\lambda, x_1 + 1, x_2 + 1), \\ \frac{\partial}{\partial \lambda_2} G(\lambda, x_1, x_2) &= \frac{x_2}{\lambda_2} G(\lambda, x_1, x_2) + \frac{\lambda_0}{\lambda_2} G(\lambda, x_1 + 1, x_2 + 1)\end{aligned}$$

et

$$\frac{\partial}{\partial \lambda_0} G(\lambda, x_1, x_2) = G(\lambda, x_1 + 1, x_2 + 1).$$

Étant donné des observations $(X_{11}, X_{12}) = (x_{11}, x_{12}), \dots, (X_{n1}, X_{n2}) = (x_{n1}, x_{n2})$ du modèle de Bulla et coll. (2012), la fonction de vraisemblance du paramètre $\lambda = (\lambda_0, \lambda_1, \lambda_2)$ est donnée par

$$L(\lambda) = \prod_{i=1}^n \Pr(X_1 = x_{i1}, X_2 = x_{i2}) = e^{-n(\lambda_0 + \lambda_1 + \lambda_2)} \prod_{i=1}^n G(\lambda, x_{i1}, x_{i2}).$$

Lorsque l'échantillon est de grande taille, cette fonction de vraisemblance peut prendre des valeurs extrêmement grandes ou petites : elle prend souvent des valeurs qui sont bien au-delà des possibilités des nombres à virgule flottante que les ordinateurs manipulent. C'est pour cette raison, entre autres, qu'il est d'usage de maximiser le logarithme de la fonction de vraisemblance plutôt que la fonction de vraisemblance elle-même. On obtient évidemment la même solution en procédant ainsi, car le logarithme de la fonction de vraisemblance $\ell(\lambda) = \ln\{L(\lambda)\}$ est une fonction monotone croissante de $L(\lambda)$. Par conséquent, si la fonction $\ell(\lambda)$ est maximisée en $\hat{\lambda}$, il en va de même pour $L(\lambda)$.

Pour $j \in \{0, 1, 2\}$, la dérivée partielle de la log-vraisemblance par rapport à λ_j est donnée par

$$\frac{\partial}{\partial \lambda_j} \ln\{L(\lambda)\} = -n + \sum_{i=1}^n \frac{1}{G(\lambda, x_{i1}, x_{i2})} \frac{\partial}{\partial \lambda_j} G(\lambda, x_{i1}, x_{i2}).$$

En invoquant la proposition 2.2.1, on trouve

$$\begin{aligned}\frac{\partial}{\partial \lambda_1} \ln\{L(\lambda)\} &= -n + \frac{1}{\lambda_1} \sum_{i=1}^n x_{i1} + \frac{\lambda_0}{\lambda_1} \sum_{i=1}^n \frac{G(\lambda, x_{i1} + 1, x_{i2} + 1)}{G(\lambda, x_{i1}, x_{i2})}, \\ \frac{\partial}{\partial \lambda_2} \ln\{L(\lambda)\} &= -n + \frac{1}{\lambda_2} \sum_{i=1}^n x_{i2} + \frac{\lambda_0}{\lambda_2} \sum_{i=1}^n \frac{G(\lambda, x_{i1} + 1, x_{i2} + 1)}{G(\lambda, x_{i1}, x_{i2})}\end{aligned}$$

et

$$\frac{\partial}{\partial \lambda_0} \ln\{L(\lambda)\} = -n + \sum_{i=1}^n \frac{G(\lambda, x_{i1} + 1, x_{i2} + 1)}{G(\lambda, x_{i1}, x_{i2})}.$$

L'estimateur à vraisemblance maximale $\hat{\lambda}_n = (\hat{\lambda}_{n0}, \hat{\lambda}_{n1}, \hat{\lambda}_{n2})$ de $(\lambda_0, \lambda_1, \lambda_2)$ est donc la solution du système d'équations suivant :

$$\begin{cases} \frac{1}{\hat{\lambda}_{n1}} \sum_{i=1}^n X_{i1} + \frac{\hat{\lambda}_{n0}}{\hat{\lambda}_{n1}} \sum_{i=1}^n \frac{G(\hat{\lambda}_n, X_{i1} + 1, X_{i2} + 1)}{G(\hat{\lambda}_n, X_{i1}, X_{i2})} = n, \\ \frac{1}{\hat{\lambda}_{n2}} \sum_{i=1}^n X_{i2} + \frac{\hat{\lambda}_{n0}}{\hat{\lambda}_{n2}} \sum_{i=1}^n \frac{G(\hat{\lambda}_n, X_{i1} + 1, X_{i2} + 1)}{G(\hat{\lambda}_n, X_{i1}, X_{i2})} = n, \\ \sum_{i=1}^n \frac{G(\hat{\lambda}_n, X_{i1} + 1, X_{i2} + 1)}{G(\hat{\lambda}_n, X_{i1}, X_{i2})} = n. \end{cases} \quad (2.1)$$

Au vu de la troisième équation, on conclut que

$$\hat{\lambda}_{n1} = \bar{X}_{n1} + \hat{\lambda}_{n0}, \quad \hat{\lambda}_{n2} = \bar{X}_{n2} + \hat{\lambda}_{n0}.$$

Si on tient compte de ces identités dans le système (2.1), on voit que le problème est entièrement résolu dans la mesure où l'on peut trouver numériquement la valeur de $\hat{\lambda}_{n0}$ telle que

$$\sum_{i=1}^n \frac{G\{(\hat{\lambda}_{n0}, \hat{\lambda}_{n0} + \bar{X}_{n1}, \hat{\lambda}_{n0} + \bar{X}_{n2}), X_{i1} + 1, X_{i2} + 1\}}{G\{(\hat{\lambda}_{n0}, \hat{\lambda}_{n0} + \bar{X}_{n1}, \hat{\lambda}_{n0} + \bar{X}_{n2}), X_{i1}, X_{i2}\}} = n.$$

Le résultat suivant est démontré par Bulla et coll. (2012).

Proposition 2.2.2. *Quand $n \rightarrow \infty$, $\sqrt{n}(\hat{\lambda}_n - \lambda)$ converge en loi vers une $\mathcal{N}_3(0, I^{-1}(\lambda))$, où $I(\lambda)$ est la matrice d'information de Fisher 3×3 dont l'élément (j, k) est donné en tous $j, k \in \{1, 2, 3\}$ par*

$$I_{j,k}(\lambda) = -E \left[\frac{\partial^2}{\partial \lambda_{j-1} \partial \lambda_{k-1}} \ln[L\{\lambda, (X_1, X_2)\}] \right].$$

Cette matrice peut être estimée de façon convergente par $I(\hat{\lambda})$, ce qui permet de construire des intervalles de confiance et des tests d'hypothèses.

2.3 Modèles de Skellam avec véritables paramètres de dépendance

Comme on l'a vu plus haut, le modèle de Skellam bivarié proposé par Bulla et coll. (2012) présente un problème conceptuel. En effet, le paramètre de dépendance λ_0 affecte les lois marginales. Idéalement, il est préférable qu'il n'en soit pas ainsi. Pour éviter ce problème, Genest & Mesfioui (2014) proposent deux modèles de Skellam bivariés qui exploitent la notion de choc commun. Nous commençons par présenter le modèle de Skellam à un seul paramètre de dépendance.

2.3.1 Modèle de Skellam bivarié de première espèce

Soient λ_{11} , λ_{12} , λ_{21} et λ_{22} des paramètres positifs. Posons $\lambda_1 = \min(\lambda_{11}, \lambda_{21}) > 0$. Pour tout $\theta \in [0, \lambda_1]$, considérons des variables aléatoires Y_0 , Y_1 et Y_2 mutuellement indépendantes telles que

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

Par convention, $Y_0 \equiv 0$ lorsque $\theta = 0$. Soient $G_{1\theta}$ et $G_{2\theta}$ les fonctions de répartition respectives de Y_1 et Y_2 .

Définition 2.3.1. *On dit que le couple aléatoire (X_1, X_2) obéit à la loi de Skellam bivariée de première espèce, que l'on note $\mathcal{BS}_1(\theta; \lambda_{11}, \lambda_{12}; \lambda_{21}, \lambda_{22})$, si et seulement si*

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

Notons que dans ce modèle, le paramètre θ n'affecte pas les lois marginales de X_1 et X_2 . Ce paramètre peut donc être considéré comme un indice de dépendance, d'autant plus que, pour tous $x_1, x_2 \in \mathbb{R}$,

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

est une fonction croissante de θ , tel qu'on le montre dans la proposition suivante.

Proposition 2.3.1. *Pour tous $x_1, x_2 \in \mathbb{R}$ et $\theta_1, \theta_2 \in [0, \lambda_1]$, on a*

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

Démonstration. Désignons par F_θ la fonction de répartition de la loi de Poisson de paramètre $\lambda - \theta$, où $\lambda > 0$ et $\theta \in (0, \lambda)$. Il est facile de voir que pour tout $k \in \mathbb{N}$,

$$\frac{\partial}{\partial \theta} F_\theta(k) = F_\theta(k) - F_\theta(k-1) = \frac{e^{-(\lambda-\theta)}(\lambda-\theta)^k}{k!}.$$

Il en résulte que pour tous $x \in \mathbb{Z}$ et $j \in \{1, 2\}$,

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

Par conséquent, pour tout $(x_1, x_2) \in \mathbb{Z}^2$,

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

Après simplification, on trouve

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

Il s'ensuit que la fonction $\theta \mapsto H_\theta(x_1, x_2)$ est croissante. $\square$

À l'instar de la loi de Skellam univariée, ce modèle est stable par convolution. C'est là l'objet du résultat suivant, dont la démonstration est omise, puisqu'elle est en tout point semblable à celle de la proposition 1.1.1.

Proposition 2.3.2. *Soient $X_1, \dots, X_n$ des variables aléatoires mutuellement indépendantes telles que, pour tout $i \in \{1, \dots, n\}$, $X_i \sim \mathcal{BS}_1(\theta_i; \lambda_{i11}, \lambda_{i12}; \lambda_{i21}, \lambda_{i22})$. Alors*

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

Le résultat suivant donne une expression explicite pour la fonction génératrice de probabilité du modèle $\mathcal{BS}_1(\theta; \lambda_{11}, \lambda_{12}; \lambda_{21}, \lambda_{22})$. Cette formule peut être utilisée pour calculer les moments mixtes de ce modèle.

Proposition 2.3.3. *La fonction génératrice de probabilité de la loi de Skellam bivariée $\mathcal{BS}_1(\theta; \lambda_{11}, \lambda_{12}, \lambda_{21}, \lambda_{22})$ est donnée, en tout $(t_1, t_2) \in (0, \infty)^2$, par*

$$\begin{aligned} E(t_1^{X_1} t_2^{X_2}) &= \exp\{(\lambda_{11} - \theta)(t_1 - 1) + \lambda_{12}(1/t_1 - 1)\} \times \exp\{(\lambda_{21} - \theta)(t_2 - 1) \\ &\quad + \lambda_{22}(1/t_2 - 1)\} \times \exp\{\theta(t_1 t_2 - 1)\}. \end{aligned}$$

Démonstration. Par définition, $X_1 = Y_1 + Y_0$ et $X_2 = Y_2 + Y_0$, où Y_0, Y_1, Y_2 sont des variables aléatoires mutuellement indépendantes telles que $Y_0 \sim \mathcal{P}(\theta)$, $Y_1 \sim \mathcal{S}(\lambda_{11} - \theta, \lambda_{12})$ et $Y_2 \sim \mathcal{S}(\lambda_{21} - \theta, \lambda_{22})$. Par conséquent, on a

$$E(t_1^{X_1} t_2^{X_2}) = E(t_1^{Y_1}) E(t_2^{Y_2}) E\{(t_1 t_2)^{Y_0}\}$$

pour tous $t_1, t_2 \in (0, \infty)$. En utilisant le fait que pour $j \in \{1, 2\}$, Y_j s'exprime comme la différence de deux variables aléatoires de Poisson indépendantes de paramètres $\lambda_{j1} - \theta$ et λ_{j2} , on déduit en outre que

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

Le résultat découle du fait que $E\{(t_1 t_2)^{Y_0}\} = \exp\{\theta(t_1 t_2 - 1)\}$. □

On a entre autres $E(X_j) = \lambda_{j1} - \lambda_{j2}$ et $\text{var}(X_j) = \lambda_{j1} + \lambda_{j2}$ pour $j \in \{1, 2\}$. De plus,

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

Notons que la plus grande valeur possible de $\text{corr}(X_1, X_2)$ se produit dans le cas particulier où X_1 et X_2 sont de même loi, c'est-à-dire quand $\lambda_{11} = \lambda_{21} = \lambda_1$ et $\lambda_{12} = \lambda_{22} = \lambda_2$. Dans ce cas particulier, on trouve

$$\text{corr}(X_1, X_2) = \frac{\lambda_1}{\lambda_1 + \lambda_2},$$

qui est strictement inférieur à 1. En particulier, $\text{corr}(X_1, X_2) = 1/2$ si $\lambda_1 = \lambda_2$.

2.3.2 Modèle de Skellam bivarié de deuxième espèce

Soient $\lambda_{11}, \lambda_{12}, \lambda_{21}$ et λ_{22} des paramètres positifs. Posons $\lambda_1 = \min(\lambda_{11}, \lambda_{21}) > 0$. Pour tout $\theta \in [0, \lambda_1]$, considérons des variables aléatoires Y_0, Y_1 et Y_2 mutuellement indépendantes telles que

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

Définition 2.3.2. On dit que le couple aléatoire (X_1, X_2) obéit à une loi de Skellam bivariée de deuxième espèce, notée $\mathcal{BS}_2(\theta_1, \theta_2; \lambda_{11}, \lambda_{12}; \lambda_{21}, \lambda_{22})$, si et seulement si

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

Dans cette construction, le vecteur de paramètre $\Theta = (\theta_1, \theta_2)$ n'affecte pas les lois marginales de X_1 et X_2 , respectivement notées F_1 et F_2 . Lorsque $\theta_2 = 0$, on voit que Y_0 suit une loi de Poisson d'espérance θ_1 , tandis que $Y_1 \sim \mathcal{S}(\lambda_{11} - \theta, \lambda_{12})$ et $Y_2 \sim \mathcal{S}(\lambda_{21} - \theta, \lambda_{22})$. Par conséquent,

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

La fonction de répartition de la loi de Skellam bivariée de deuxième espèce est donnée, en tout $(x_1, x_2) \in \mathbb{Z}^2$, par

$$\Pr(X_1 \leq x_1, X_2 \leq x_2) = H_\Theta(x_1, x_2) = \sum_{k \in \mathbb{Z}} G_{1\Theta}(x_1 - k) G_{2\Theta}(x_2 - k) g_\Theta(k),$$

où g_Θ représente la fonction de masse de Y_0 . La proposition suivante montre la croissance de la fonction $\Theta \mapsto H_\Theta$ en ses deux arguments, où la comparaison entre vecteurs se fait composante par composante.

Proposition 2.3.4. Pour tous $(x_1, x_2) \in \mathbb{R}^2$ et $\Theta_1, \Theta_2 \in [0, \lambda_1] \times [0, \lambda_2]$, on a

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

Démonstration. En procédant de la même manière que pour la démonstration de la proposition 2.3.1, on voit que pour tout $x \in \mathbb{Z}$,

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

et que, pour $j \in \{1, 2\}$,

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

Il s'ensuit que pour tout $(x_1, x_2) \in \mathbb{Z}^2$, on a

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

Au terme de quelques calculs algébriques simples, on trouve

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

De même, $\partial H_{\Theta}(x_1, x_2) / \partial \theta_2 \geq 0$. On en conclut que la fonction $\Theta \mapsto H_{\Theta}(x_1, x_2)$ est bel et bien croissante en ses deux arguments, tel qu'annoncé. $\square$

Étant donné que $Y \sim \mathcal{S}(0, 0)$ correspond au cas où $Y \equiv 0$, on a, pour tous $x_1, x_2 \in \mathbb{R}$,

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

Ainsi la proposition 2.3.4 entraîne que quel que soit $\Theta \in [0, \lambda_1] \times [0, \lambda_2]$, H_{Θ} est en dépendance positive par quadrant (DPQ). Toutefois, il est possible de construire une généralisation du modèle qui permette de rendre compte d'une éventuelle dépendance négative. Cette question fera l'objet de la section 2.3.3.

Les résultats suivants sont des généralisations des propositions 2.3.2 et 2.3.3.

Proposition 2.3.5. *Soient $X_1, \dots, X_n$ des variables aléatoires mutuellement indépendantes telles que $X_i \sim \mathcal{BS}_2(\theta_{i1}, \theta_{i2}; \lambda_{i11}, \lambda_{i12}, \lambda_{i21}, \lambda_{i22})$ pour tout $i \in \{1, \dots, n\}$. Alors*

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

Proposition 2.3.6. *La fonction génératrice de probabilité de la loi de Skellam bivariée $\mathcal{BS}_2(\theta_1, \theta_2; \lambda_{11}, \lambda_{12}, \lambda_{21}, \lambda_{22})$ est donnée, en tout $(t_1, t_2) \in (0, \infty)^2$, par*

$$\begin{aligned} E(t_1^{X_1} t_2^{X_2}) &= \exp\{(\lambda_{11} - \theta_1)(t_1 - 1) + (\lambda_{12} - \theta_2)(1/t_1 - 1)\} \times \exp\{(\lambda_{21} - \theta_1) \\ &\quad + (\lambda_{22} - \theta_2)(1/t_2 - 1)\} \times \exp[\theta_1(t_1 t_2 - 1) + \theta_2\{1/(t_1 t_2) - 1\}]. \end{aligned}$$

Démonstration. Par définition, on sait que l'on peut écrire $X_1 = Y_1 + Y_0$ et $X_2 = Y_2 + Y_0$ en termes de variables aléatoires mutuellement indépendantes Y_0, Y_1, Y_2 telles que

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

Or pour $j \in \{1, 2\}$, Y_j est la différence de deux variables aléatoires indépendantes ayant des lois de Poisson de paramètres respectifs $\lambda_{j1} - \theta_1$ et $\lambda_{j2} - \theta_2$, ce qui permet de déduire que pour tout $(t_1, t_2) \in (0, \infty)^2$,

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

Il s'ensuit que pour tout $(t_1, t_2) \in (0, \infty)^2$,

$$E\{(t_1 t_2)^{Y_0}\} = \exp[\theta_1(t_1 t_2 - 1) + \theta_2\{1/(t_1 t_2) - 1\}],$$

ce qui permet de conclure. $\square$

En particulier, $E(X_j) = \lambda_{j1} - \lambda_{j2}$ et $\text{var}(X_j) = \lambda_{j1} + \lambda_{j2}$ pour $j \in \{1, 2\}$. On peut voir en outre que

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

La plage de valeurs pour $\text{corr}(X_1, X_2)$ est donc plus grande que celle du premier modèle. En fait, $\text{corr}(X_1, X_2) = 1$ est atteint lorsque X_1 et X_2 sont identiquement distribuées, c'est-à-dire lorsque $\lambda_{11} = \lambda_{21} = \lambda_1$ et $\lambda_{12} = \lambda_{22} = \lambda_2$. En effet, si $\theta_1 \rightarrow \lambda_1$ et $\theta_2 \rightarrow \lambda_2$, on a $Y_1 = Y_2 \equiv 0$ et $X_1 = X_2 = Y_0$ presque sûrement. Dans ce cas, les variables X_1 et X_2 sont comonotones.

2.3.3 Modèles à dépendance négative

Les deux modèles de Skellam bivariés décrits dans la section précédente peuvent être facilement adaptés pour tenir compte d'une éventuelle dépendance négative entre X_1 et X_2 . Ceci peut être réalisé en exprimant X_1 et X_2 comme suit

$$X_1 = Y_1 + Y_0, \quad X_2 = Y_2 - Y_0, \quad (2.2)$$

en termes de variables aléatoires mutuellement indépendantes Y_0, Y_1, Y_2 telles que

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

et $Y_0 \sim \mathcal{S}(\theta_1, \theta_2)$ avec $\theta_1 \leq \kappa_1 = \min(\lambda_{11}, \lambda_{22})$ et $\theta_2 \leq \kappa_2 = \min(\lambda_{12}, \lambda_{21})$.

Notons par $G_{1\Theta}$ et $G_{2\Theta}$ les fonctions de répartition respectives de Y_1 et Y_2 . Soit g_Θ la fonction de masse de la variable aléatoire Y_0 . La loi jointe de la paire (X_1, X_2) est donnée en tous $x_1, x_2 \in \mathbb{R}$ par

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

Il est clair que ce modèle engendre une dépendance négative entre les variables X_1 et X_2 . En effet, par construction

$$(X_1, -X_2) \sim \mathcal{BS}_2(\theta_1, \theta_2; \lambda_{11}, \lambda_{12}; \lambda_{22}, \lambda_{21}). \quad (2.3)$$

Remarquons que pour ce modèle, les paramètres λ_{21} et λ_{22} n'apparaissent pas dans le même ordre que dans la définition du modèle initial décrit par la construction 2.3.2. Ceci vient du fait que $-Y_2 \sim \mathcal{S}(\lambda_{22} - \theta_1, \lambda_{21} - \theta_2)$. En utilisant la proposition 2.3.4, on voit que $(X_1, -X_2)$ est DPQ et donc que la famille des couples aléatoires (X_1, X_2) obéit à la propriété de dépendance négative par quadrant. De plus, cette famille est décroissante par rapport au paramètre de dépendance θ , tel qu'énoncé ci-dessous.

Proposition 2.3.7. *Pour tous $x_1, x_2 \in \mathbb{R}$ et $\Theta_1, \Theta_2 \in [0, \kappa_1] \times [0, \kappa_2]$, on a*

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

Ce modèle permet aux variables aléatoires X_1 et $-X_2$ de décroître l'une en fonction de l'autre quand elles sont identiquement distribuées, c'est-à-dire lorsque $\lambda_{11} = \lambda_{22}$ et $\lambda_{12} = \lambda_{21}$. Dans ce cas, la plus faible dépendance possible est obtenue en prenant $\theta_1 = \lambda_{11} = \lambda_{22}$ et $\theta_2 = \lambda_{12} = \lambda_{21}$. Il en résulte que $X_1 = -X_2 = Y_0$ presque sûrement ; en d'autres mots, X_1 et X_2 sont alors comonotones.

Grâce à la relation (2.3), on peut montrer que la fonction génératrice de probabilité du couple aléatoire (X_1, X_2) définie par (2.2) est donnée en tout $(t_1, t_2) \in (0, \infty)^2$, par

$$\begin{aligned} E(t_1^{X_1} t_2^{X_2}) &= E\{t_1^{Y_1} t_2^{Y_2} (t_1/t_2)^{Y_0}\} \\ &= \exp\{(\lambda_{11} - \theta_1)(t_1 - 1) + (\lambda_{12} - \theta_2)(1/t_1 - 1)\} \\ &\quad \times \exp\{(\lambda_{21} - \theta_1)(t_2 - 1) + (\lambda_{22} - \theta_2)(1/t_2 - 1)\} \\ &\quad \times \exp\{\theta_1(t_1/t_2 - 1) + \theta_2(t_2/t_1 - 1)\}. \end{aligned}$$

En particulier, $E(X_j) = \lambda_{j1} - \lambda_{j2}$, $\text{var}(X_j) = \lambda_{j1} + \lambda_{j2}$ pour $j \in \{1, 2\}$ et

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

2.4 Estimation des paramètres des lois $\mathcal{BS}_1$ et $\mathcal{BS}_2$

Soit $(X_{11}, X_{12}), \dots, (X_{n1}, X_{n2})$ un échantillon aléatoire de la loi de Skellam bivariée de première ou de deuxième espèce. Comment peut-on en estimer les paramètres ?

Comme Romani (1956) l'a montré, les paramètres des lois marginales $\mathcal{S}(\lambda_{11}, \lambda_{12})$ et $\mathcal{S}(\lambda_{21}, \lambda_{22})$ peuvent être estimés par la méthode des moments en résolvant les équations suivantes :

$$\hat{\lambda}_{n11} - \hat{\lambda}_{n12} = \bar{X}_{n1}, \quad \hat{\lambda}_{n21} - \hat{\lambda}_{n22} = \bar{X}_{n2}$$

et

$$\hat{\lambda}_{n11} + \hat{\lambda}_{n12} = S_{n1}^2, \quad \hat{\lambda}_{n21} + \hat{\lambda}_{n22} = S_{n2}^2.$$

Dans ces équations,

$$\bar{X}_{n1} = \frac{1}{n} \sum_{i=1}^n X_{i1}, \quad \bar{X}_{n2} = \frac{1}{n} \sum_{i=1}^n X_{i2}$$

et

$$S_{n1}^2 = \frac{1}{n-1} \sum_{i=1}^n (X_{i1} - \bar{X}_{n1})^2, \quad S_{n2}^2 = \frac{1}{n-1} \sum_{i=1}^n (X_{i2} - \bar{X}_{n2})^2.$$

Il en résulte que

$$\begin{aligned} \hat{\lambda}_{n11} &= \frac{S_{n1}^2 + \bar{X}_{n1}}{2}, & \hat{\lambda}_{n12} &= \frac{S_{n1}^2 - \bar{X}_{n1}}{2}, \\ \hat{\lambda}_{n21} &= \frac{S_{n2}^2 + \bar{X}_{n2}}{2}, & \hat{\lambda}_{n22} &= \frac{S_{n2}^2 - \bar{X}_{n2}}{2}. \end{aligned} \quad (2.4)$$

La loi des grands nombres assure la convergence de ces estimateurs. Notons que pour que ces estimateurs soient positifs, il faut que $S_{nj}^2 \geq |\bar{X}_{nj}|$ pour tout $j \in \{1, 2\}$. Si cette condition n'est pas vérifiée, on peut procéder comme Alzaid & Omair (2010) en réduisant l'estimation négative à zéro, et en posant l'autre égale à $|\bar{X}_{nj}|$.

Il reste donc à trouver une façon d'estimer les paramètres de dépendance des deux modèles. C'est la question qui sera abordée dans les deux prochaines sections, au moyen de la méthode des moments. Nous commençons par traiter le cas du modèle de Skellam bivarié de première espèce.

2.4.1 Estimation du paramètre θ pour le modèle $\mathcal{BS}_1$

Étant donné un échantillon aléatoire $(X_{11}, X_{12}), \dots, (X_{n1}, X_{n2})$ provenant du modèle $\mathcal{BS}_1(\theta; \lambda_{11}, \lambda_{12}; \lambda_{21}, \lambda_{22})$, l'estimateur des moments du paramètre de dépendance θ est donné par

$$\hat{\theta}_n = S_{n12} = \frac{1}{n-1} \sum_{i=1}^n (X_{i1} - \bar{X}_{n1})(X_{i2} - \bar{X}_{n2}). \quad (2.5)$$

Ceci découle du fait que la covariance du modèle $\mathcal{BS}_1(\theta; \lambda_{11}, \lambda_{12}; \lambda_{21}, \lambda_{22})$ est exactement θ . Ici encore, la loi des grands nombres nous assure que cet estimateur est convergent. Lorsque la taille de l'échantillon est faible, il existe toutefois une probabilité non nulle que $S_{n12} \leq 0$; dans ce cas, on peut poser $\hat{\theta}_n = 0$.

Dans la pratique, une valeur négative de S_{n12} suggère de considérer le modèle de dépendance négative décrit par (2.2). Comme nous avons indiqué à la section 2.3.3, cela revient à supposer que $(X_1, -X_2) \sim \mathcal{BS}_1(\theta; \lambda_{11}, \lambda_{12}; \lambda_{22}, \lambda_{21})$. Noter que cela n'affecte pas les estimations des paramètres marginaux donnés par (2.4). Comme $\text{cov}(X_1, -X_2) = \theta$, $-S_{n12} > 0$ sera alors un estimateur convergent de θ .

2.4.2 Estimation des paramètres θ_1 et θ_2 pour le modèle $\mathcal{BS}_2$

Lorsque les données proviennent du modèle plus général $\mathcal{BS}_2(\theta_1, \theta_2; \lambda_{11}, \lambda_{12}; \lambda_{21}, \lambda_{22})$, deux équations sont nécessaires pour estimer θ_1 et θ_2 par la méthode des moments. L'identité $\text{cov}(X_1, X_2) = \theta_1 + \theta_2$ nous conduit à l'équation d'estimation

$$S_{n12} = \hat{\theta}_{n1} + \hat{\theta}_{n2},$$

à partir de laquelle un estimateur convergent de $\theta = \theta_1 + \theta_2$ peut être déduit. Quand $n \rightarrow \infty$, la probabilité que $S_{n12} < 0$ devient négligeable. S'il s'avère néanmoins que $S_{n12} < 0$, c'est vraisemblablement que le modèle (2.2) est plus approprié.

2.4.2.1 Cas $S_{n12} > 0$

Afin d'estimer θ_1 et θ_2 sous la contrainte $\hat{\theta}_{n1} + \hat{\theta}_{n2} = S_{n12}$, on doit disposer d'une seconde équation. Celle-ci peut être obtenue au moyen du résultat suivant.

Proposition 2.4.1. *Supposons que $(X_1, X_2) \sim \mathcal{BS}_2(\theta_1, \theta_2; \lambda_{11}, \lambda_{12}; \lambda_{21}, \lambda_{22})$. Alors*

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

Démonstration. Par définition, on a $X_1 = Y_1 + Y_0$ et $X_2 = Y_2 + Y_0$, où Y_0, Y_1, Y_2 sont des variables aléatoires mutuellement indépendantes telles que $Y_0 \sim \mathcal{S}(\theta_1, \theta_2)$,

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

Par conséquent,

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

Ainsi l'expression souhaitée pour $\text{cov}(X_1, X_2^2)$ est donnée par

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

et

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

La formule pour $\text{cov}(X_1^2, X_2)$ peut être obtenue d'une manière semblable. □

Les estimations des quantités $\text{cov}(X_1^2, X_2)$ et $\text{cov}(X_1, X_2^2)$ sont respectivement données par

$$T_{n12} = \frac{1}{n-1} \sum_{i=1}^n (X_{i2} - \bar{X}_{n2})(X_{i1}^2 - S_{n1}^2 - \bar{X}_{n1}^2)$$

et

$$T_{n21} = \frac{1}{n-1} \sum_{i=1}^n (X_{i1} - \bar{X}_{n1})(X_{i2}^2 - S_{n2}^2 - \bar{X}_{n2}^2).$$

Étant donné que $\hat{\lambda}_{n11} - \hat{\lambda}_{n12} = \bar{X}_{n1}$ et $\hat{\lambda}_{n21} - \hat{\lambda}_{n22} = \bar{X}_{n2}$, les estimateurs de θ_1 et θ_2 s'obtiennent en résolvant l'équation

$$\frac{1}{2} (T_{n12} + T_{n21}) = \hat{\theta}_{n1} - \hat{\theta}_{n2} + (\bar{X}_{n1} + \bar{X}_{n2})(\hat{\theta}_{n1} + \hat{\theta}_{n2})$$

et en tenant compte de la contrainte $S_{n12} = \hat{\theta}_{n1} + \hat{\theta}_{n2}$. Les estimateurs possèdent donc une forme explicite, à savoir

$$\begin{aligned} \hat{\theta}_{n1} &= \frac{1}{2} \left\{ S_{n12}(1 - \bar{X}_{n1} - \bar{X}_{n2}) + \frac{1}{2} (T_{n12} + T_{n21}) \right\}, \\ \hat{\theta}_{n2} &= \frac{1}{2} \left\{ S_{n12}(1 + \bar{X}_{n1} + \bar{X}_{n2}) - \frac{1}{2} (T_{n12} + T_{n21}) \right\}. \end{aligned}$$

Encore une fois, il découle de la loi des grands nombres que ces estimateurs sont convergents. Notons que si l'un d'eux est négatif, on peut le poser égal à zéro.

2.4.2.2 Cas $S_{n12} < 0$

Tel que mentionné plus haut, s'il avère que S_{n12} est négatif, il peut alors être préférable de postuler un modèle de dépendance négative pour (X_1, X_2) , défini par (2.2). Cela revient à supposer que $(X_1, -X_2) \sim \mathcal{B}_2(\theta_1, \theta_2; \lambda_{11}, \lambda_{12}; \lambda_{22}, \lambda_{21})$. Noter que dans l'expression de la loi, les paramètres λ_{12} et λ_{22} ont été intervertis. Compte tenu de cette situation, on peut facilement déduire de la proposition 2.4.1 que

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

Les estimateurs de ces paramètres sont obtenus en résolvant les équations suivantes :

$$S_{n12} = -(\hat{\theta}_{n1} + \hat{\theta}_{n2}), \quad \frac{1}{2} (T_{n12} - T_{n21}) = \hat{\theta}_{n1} - \hat{\theta}_{n2} + (\bar{X}_{n1} - \bar{X}_{n2})(\hat{\theta}_{n1} + \hat{\theta}_{n2}).$$

Il vient

$$\hat{\theta}_{n1} = \frac{1}{2} \left\{ S_{n12}(\bar{X}_{n1} - \bar{X}_{n2} - 1) + \frac{1}{2} (T_{n12} - T_{n21}) \right\}, \quad (2.6)$$

$$\hat{\theta}_{n2} = \frac{1}{2} \left\{ S_{n12}(\bar{X}_{n2} - \bar{X}_{n1} - 1) - \frac{1}{2} (T_{n12} - T_{n21}) \right\}. \quad (2.7)$$

Ici aussi, la convergence des estimateurs découle de la loi des grands nombres. Dans le cas où l'un de ces deux estimateurs est négatif, on peut le poser égal à zéro.

Chapitre 3

Contribution principale

L'objectif principal de ce chapitre est d'étudier le comportement asymptotique des estimateurs des paramètres de dépendance des modèles de Skellam bivariés $\mathcal{BS}_1$ et $\mathcal{BS}_2$. Le premier volet de ce chapitre est consacré à l'étude des lois asymptotiques des estimateurs des moments des paramètres de dépendance des lois $\mathcal{BS}_1$ et $\mathcal{BS}_2$ identifiés par Genest & Mesfioui (2014). Des expressions explicites des variances asymptotiques de ces estimateurs sont données. La vitesse à laquelle ces approximations deviennent valables est ensuite mesurée par voie de simulation.

Le deuxième volet du chapitre concerne l'étude des estimateurs à vraisemblance maximale des paramètres des lois $\mathcal{BS}_1$ et $\mathcal{BS}_2$. Dans un premier temps, nous présentons les équations normales qui conduisent à ces estimateurs. Nous montrons ensuite la convergence et la normalité asymptotique des estimateurs. Nous présentons en outre des résultats de simulation permettant de valider les formules de variances obtenues.

Ce chapitre est structuré comme suit. Les sections 3.1 et 3.2 décrivent le comportement asymptotique des estimateurs des moments des paramètres de dépendance des modèles $\mathcal{BS}_1$ et $\mathcal{BS}_2$. Les sections 3.3 et 3.4 montrent ensuite comment calculer les estimateurs à vraisemblance maximale des paramètres de ces deux modèles. Puis, leur comportement asymptotique respectif est étudié dans les sections 3.5 et 3.6. Enfin, des comparaisons d'efficacité entre les estimations des moments et du maximum de vraisemblance sont présentées à la section 3.7. Certains détails concernant le calcul de la variance limite des estimateurs à vraisemblance maximale des paramètres du modèle $\mathcal{BS}_2$ sont consignés dans une annexe.

3.1 Loi asymptotique des estimateurs des moments des paramètres du modèle $\mathcal{BS}_1$

Le but de cette section est de déterminer la loi asymptotique de l'estimateur des moments du paramètre de dépendance θ du modèle $\mathcal{BS}_1(\theta; \lambda_{11}, \lambda_{21}; \lambda_{12}, \lambda_{22})$. Pour ce faire, nous aurons recours au résultat suivant, qui énonce une version multivariée du théorème central limite et la transformation de sa limite par la méthode Delta.

Proposition 3.1.1. *Soit $\mathbf{X}_1, \mathbf{X}_2, \dots$ une suite de vecteurs aléatoires mutuellement indépendants issus d'une loi en dimension d dont tous les moments d'ordre 2 sont finis. Pour tout $n \in \mathbb{N}$, soit $\bar{X}_n = (\mathbf{X}_1 + \dots + \mathbf{X}_n)/n$. Soit en outre $\mu = E(\bar{X}_n)$. Alors il existe une matrice Σ de taille $d \times d$ telle que, quand $n \rightarrow \infty$,*

$$\sqrt{n}(\bar{X}_n - \mu) \rightarrow \mathcal{N}_d(0, \Sigma).$$

Soit en outre $h : \mathbb{R}^d \rightarrow \mathbb{R}^p$ une fonction dont la dérivée h' existe et est non nulle au point μ . Alors, quand $n \rightarrow \infty$,

$$\sqrt{n}\{h(\bar{X}_n) - h(\mu)\} \rightarrow \mathcal{N}_p(0, h'(\mu)^\top \Sigma h'(\mu)).$$

En utilisant ce résultat, nous pouvons établir la loi asymptotique de l'estimateur des moments défini par l'équation (2.5).

Proposition 3.1.2. *Soit $(X_1, X_2), (X_{11}, X_{12}), (X_{21}, X_{22}), \dots$ une suite d'observations mutuellement indépendantes de loi $\mathcal{BS}_1$. Soit $\hat{\theta}_n$ l'estimateur des moments du paramètre de dépendance θ défini par l'équation (2.5). Quand $n \rightarrow \infty$,*

$$\sqrt{n}(\hat{\theta}_n - \theta) \rightarrow \mathcal{N}(0, \sigma_\theta^2),$$

où

$$\sigma_\theta^2 = \theta + \theta^2 + (\lambda_{11} + \lambda_{12})(\lambda_{21} + \lambda_{22}).$$

Démonstration. On peut supposer sans perte de généralité que $E(X_{i1}) = E(X_{i2}) = 0$, et que $\text{var}(X_{i1}) = \text{var}(X_{i2}) = 1$. Posons $\text{cov}(X_{i1}, X_{i2}) = \gamma$ et, pour tout $i \in \mathbb{N}$, $W_i = (X_{i1}, X_{i2}, X_{i1}X_{i2})^\top$. Pour tout $n \in \mathbb{N}$, soit en outre $\bar{W}_n = (W_1 + \dots + W_n)/n$.

Il est clair que $W_1, W_2, \dots$ forment une suite de vecteurs mutuellement indépendants de même loi dont tous les moments d'ordre 2 sont finis. Il découle donc de la proposition 3.1.1 que, quand $n \rightarrow \infty$,

$$\sqrt{n}(\bar{W}_n - \Gamma) \rightarrow \mathcal{N}_3(0, \Sigma),$$

où $\Gamma = (0, 0, \gamma)^\top$ et

$$\Sigma = \begin{pmatrix} 1 & \gamma & E(X_1^2 X_2) \\ \gamma & 1 & E(X_1 X_2^2) \\ E(X_1^2 X_2) & E(X_1 X_2^2) & E(X_1^2 X_2^2) - \gamma^2 \end{pmatrix}.$$

Considérons ensuite une application $h : \mathbb{R}^3 \rightarrow \mathbb{R}$ définie, pour tout $(x_1, x_2, x_3) \in \mathbb{R}^3$, par $h(x_1, x_2, x_3) = x_3 - x_1 x_2$. Alors $h(\theta) = \gamma$ et

$$\hat{\theta} = h(\bar{W}_n) = \frac{1}{n} \sum_{i=1}^n X_{i1} X_{i2} - \bar{X}_{n1} \bar{X}_{n2},$$

où $\bar{X}_{n1} = (X_{11} + \dots + X_{n1})/n$ et $\bar{X}_{n2} = (X_{21} + \dots + X_{n2})/n$. En faisant appel à la proposition 3.1.1, on se convainc facilement que, quand $n \rightarrow \infty$,

$$\sqrt{n} \{h(\bar{W}_n) - h(\Gamma)\} \rightarrow \mathcal{N}(0, c^2),$$

où $c^2 = \epsilon^\top \Sigma \epsilon$ et

$$\epsilon = \left. \frac{\partial}{\partial w} h(w) \right|_{w=\theta} = (0, 0, 1)^\top.$$

Par conséquent, $c^2 = E(X_{11}^2 X_{12}^2) - \gamma^2 = \text{var}(X_{11} X_{12})$.

Pour déduire l'expression de la variance asymptotique $\sigma_{\hat{\theta}}^2$, il suffit de remplacer X_{11} et X_{12} par

$$\frac{X_1 - E(X_1)}{\sqrt{\text{var}(X_1)}} \quad \text{et} \quad \frac{X_2 - E(X_2)}{\sqrt{\text{var}(X_2)}}$$

dans l'expression de c^2 . Il vient

$$\sigma_{\hat{\theta}}^2 = \text{var}(X_1) \text{var}(X_2) \text{var} \left[\frac{\{X_1 - E(X_1)\}}{\sqrt{\text{var}(X_1)}} \frac{\{X_2 - E(X_2)\}}{\sqrt{\text{var}(X_2)}} \right].$$

Il s'ensuit que

$$\begin{aligned} \sigma_{\hat{\theta}}^2 &= \text{var}(X_1 X_2) + \{E(X_1)\}^2 \text{var}(X_2) + \{E(X_2)\}^2 \text{var}(X_1) \\ &\quad - 2E(X_1) \text{cov}(X_2, X_1 X_2) - 2E(X_2) \text{cov}(X_1, X_1 X_2) \\ &\quad + 2E(X_1)E(X_2) \text{cov}(X_1, X_2). \end{aligned}$$

Afin de simplifier cette expression, on peut exploiter le fait que puisque (X_1, X_2) obéit à une loi de Skellam de première espèce, il existe des variables mutuellement indépendantes Y_0, Y_1, Y_2 telles que

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

On trouve alors

$$\begin{aligned}
 \text{var}(X_1 X_2) = & 4\text{var}(Y_0)E(Y_1)E(Y_2) + 2E(Y_0^3)E(Y_1) \\
 & - 2\text{var}(Y_0)E(Y_0)E(Y_1) - \{E(Y_0)\}^3E(Y_1) \\
 & + 2E(Y_0^3)E(Y_2) - 2\text{var}(Y_0)E(Y_0)E(Y_2) \\
 & - 2\{E(Y_0)\}^3E(Y_2) - [\text{var}(Y_0) + \{E(Y_0)\}^2]^2 \\
 & + E(Y_0^4) - E(Y_1)E(Y_2) + E(Y_0)E(Y_1) + E(Y_0)E(Y_2) \\
 & + \{E(Y_0^2)\}^2 + [\text{var}(Y_0) + \text{var}(Y_1) + \{E(Y_0) + E(Y_1)\}^2] \\
 & \times [\text{var}(Y_0) + \text{var}(Y_2) + \{E(Y_0) + E(Y_2)\}^2],
 \end{aligned}$$

De plus,

$$\begin{aligned}
 \text{cov}(X_1, X_1 X_2) = & E(Y_0)\text{var}(Y_1) + E(Y_1)\text{var}(Y_0) + E(Y_2)\text{var}(Y_0) \\
 & + E(Y_2)\text{var}(Y_1) + E(Y_0^3) - E(Y_0)\text{var}(Y_0) + \{E(Y_0)\}^3
 \end{aligned}$$

et, par symétrie,

$$\begin{aligned}
 \text{cov}(X_2, X_1 X_2) = & E(Y_0)\text{var}(Y_2) + E(Y_2)\text{var}(Y_0) + E(Y_1)\text{var}(Y_0) \\
 & + E(Y_1)\text{var}(Y_2) + E(Y_0^3) - E(Y_0)\text{var}(Y_0) + \{E(Y_0)\}^3.
 \end{aligned}$$

Or, sachant que Y_0 est une variable de Poisson de paramètre θ , on a

$$E(Y_0) = \text{var}(Y_0) = \theta$$

et aussi

$$E(Y_0^3) = \theta^3 + 3\theta^2 + \theta, \quad E(Y_0^4) = \theta^4 + 6\theta^3 + 7\theta^2 + \theta.$$

Faisant en outre appel au fait que

$$\begin{aligned}
 E(X_{11}) &= \lambda_{11} - \lambda_{12}, & \text{var}(X_{11}) &= \lambda_{11} + \lambda_{12}, \\
 E(X_{12}) &= \lambda_{21} - \lambda_{22}, & \text{var}(X_{12}) &= \lambda_{21} + \lambda_{22}
 \end{aligned}$$

et

$$\begin{aligned}
 E(Y_1) &= \lambda_{11} - \theta - \lambda_{12}, & \text{var}(Y_1) &= \lambda_{11} - \theta + \lambda_{12}, \\
 E(Y_2) &= \lambda_{21} - \theta - \lambda_{22}, & \text{var}(Y_2) &= \lambda_{21} - \theta + \lambda_{22},
 \end{aligned}$$

on trouve, après simplification,

$$\sigma_{\hat{\theta}}^2 = \theta^2 + \theta + (\lambda_{11} + \lambda_{12})(\lambda_{21} + \lambda_{22}).$$

C'était ce qu'il fallait démontrer. □

Pour valider la formule de variance asymptotique $\hat{\sigma}_{\hat{\theta}}^2$ donnée dans la proposition 3.1.2, 100 échantillons de taille n de la loi de Skellam bivariée de première espèce ont été engendrés en fixant

$$\theta = 0.5, \quad \lambda_{11} = 1, \quad \lambda_{12} = 2, \quad \lambda_{21} = 3, \quad \lambda_{22} = 4.$$

La variance expérimentale $S_{\hat{\theta}_n}^2$ de $\hat{\theta}_n$ a ensuite été calculée et comparée à $\sigma_{\hat{\theta}}^2/n$. La procédure a été répétée pour chaque valeur de $n \in \{100, 500, 1000, 1500, 2000\}$. Les résultats sont consignés dans le tableau 3.1, où l'on peut voir que la correspondance est excellente, surtout pour les grandes tailles d'échantillon.

n	$S_{\hat{\theta}_n}^2$	$\sigma_{\hat{\theta}}^2/n$	Écart
100	0.200	0.217	0.017
500	0.039	0.043	0.004
1000	0.020	0.021	0.001
1500	0.0148	0.0145	0.0003
2000	0.0104	0.0106	0.0002

TABLE 3.1 – Écart observé entre l'approximation asymptotique de la variance de $\hat{\theta}_n$ et sa valeur moyenne observée sur la base de 100 échantillons de taille n , où $n \in \{100, 500, 1000, 1500, 2000\}$.

3.2 Loi asymptotique des estimateurs des moments des paramètres du modèle $\mathcal{BS}_2$

L'objectif de cette section est d'établir la loi asymptotique des estimateurs des moments des paramètres de dépendance du modèle de Skellam bivarié $\mathcal{BS}_2$ donnés par (2.6) et (2.7). Pour ce faire, nous allons étudier le comportement asymptotique du vecteur des estimateurs des moments défini par $\hat{\Theta}_n = (\hat{\theta}_{n1}, \hat{\theta}_{n2})$.

Proposition 3.2.1. *Soit $(X_1, X_2), (X_{11}, X_{12}), (X_{21}, X_{22}), \dots$ une suite d'observations mutuellement indépendantes de loi $\mathcal{BS}_2$. Soit $\hat{\Theta}_n = (\hat{\theta}_{n1}, \hat{\theta}_{n2})^\top$ l'estimateur des moments du paramètre de dépendance $\Theta = (\theta_1, \theta_2)^\top$ de la loi, dont les composantes sont définies par les équations (2.6) et (2.7). Quand $n \rightarrow \infty$,*

$$\sqrt{n}(\hat{\Theta}_n - \Theta) \rightarrow \mathcal{N}_2(0, \Sigma_{\hat{\Theta}}),$$

où

$$\Sigma_{\hat{\Theta}} = \begin{pmatrix} \sigma_{\hat{\theta}_1}^2 & \sigma_{\hat{\theta}_1 \hat{\theta}_2} \\ \sigma_{\hat{\theta}_1 \hat{\theta}_2} & \sigma_{\hat{\theta}_2}^2 \end{pmatrix}$$

est une matrice 2×2 dont les termes, qui sont fonction des paramètres

$$\mu_0 = \theta_1 - \theta_2, \quad \mu_1 = \lambda_{11} - \lambda_{12}, \quad \mu_2 = \lambda_{21} - \lambda_{22},$$

$$\sigma_0^2 = \theta_1 + \theta_2, \quad \sigma_1^2 = \lambda_{11} + \lambda_{12}, \quad \sigma_2^2 = \lambda_{21} + \lambda_{22},$$

s'expriment comme suit :

$$\begin{aligned} \sigma_{\hat{\theta}_1}^2 = & \frac{1}{2} \mu_0 + \frac{13}{8} \mu_0^2 + \frac{1}{4} \mu_0 \mu_1 + \frac{1}{4} \mu_0 \mu_2 + \frac{1}{8} \mu_1 \mu_2 - \frac{1}{8} \mu_1^2 \mu_2^4 - \frac{1}{8} \mu_1^4 \mu_2^2 + \frac{1}{4} \mu_1^3 \mu_2^3 \\ & + \frac{1}{2} \sigma_0^2 + \frac{11}{8} \sigma_0^4 + \frac{1}{4} \sigma_0^6 + \frac{1}{2} \sigma_0^2 \sigma_1^2 + \frac{1}{2} \sigma_0^2 \sigma_2^2 + \frac{1}{2} \sigma_0^2 \sigma_1^2 \sigma_2^2 + \frac{3}{8} \sigma_1^2 \sigma_2^2 \\ & + \frac{1}{4} \sigma_0^4 \sigma_1^2 + \frac{1}{4} \sigma_0^4 \sigma_2^2 + \frac{1}{8} \sigma_1^4 \sigma_2^2 + \frac{1}{8} \sigma_1^2 \sigma_2^4 \\ & + \frac{3}{2} \mu_0 \sigma_0^2 + \frac{1}{2} \mu_0 \sigma_1^2 + \frac{1}{2} \mu_0 \sigma_2^2 + \frac{1}{4} \mu_1 \sigma_2^2 + \frac{1}{4} \mu_2 \sigma_1^2 \\ & + \frac{9}{4} \mu_1 \mu_2 \sigma_0^2 - \frac{9}{4} \mu_1 \mu_2 \sigma_0^4 + \frac{3}{16} \mu_1^2 \sigma_1^2 \sigma_2^2 \\ & + \frac{3}{16} \mu_2^2 \sigma_1^2 \sigma_2^2 + \frac{3}{16} \mu_1^2 \mu_2^2 \sigma_1^2 + \frac{3}{16} \mu_1^2 \mu_2^2 \sigma_2^2, \end{aligned}$$

$$\begin{aligned} \sigma_{\hat{\theta}_2}^2 = & -\frac{1}{2} \mu_0 + \frac{13}{8} \mu_0^2 + \frac{1}{4} \mu_0 \mu_1 + \frac{1}{4} \mu_0 \mu_2 + \frac{1}{8} \mu_1 \mu_2 - \frac{1}{8} \mu_1^2 \mu_2^4 - \frac{1}{8} \mu_1^4 \mu_2^2 + \frac{1}{4} \mu_1^3 \mu_2^3 \\ & + \frac{1}{2} \sigma_0^2 + \frac{11}{8} \sigma_0^4 + \frac{1}{4} \sigma_0^6 + \frac{1}{2} \sigma_0^2 \sigma_1^2 + \frac{1}{2} \sigma_0^2 \sigma_2^2 + \frac{1}{2} \sigma_0^2 \sigma_1^2 \sigma_2^2 + \frac{3}{8} \sigma_1^2 \sigma_2^2 \\ & + \frac{1}{4} \sigma_0^4 \sigma_1^2 + \frac{1}{4} \sigma_0^4 \sigma_2^2 + \frac{1}{8} \sigma_1^4 \sigma_2^2 + \frac{1}{8} \sigma_1^2 \sigma_2^4 \\ & - \frac{3}{2} \mu_0 \sigma_0^2 - \frac{1}{2} \mu_0 \sigma_1^2 - \frac{1}{2} \mu_0 \sigma_2^2 - \frac{1}{4} \mu_1 \sigma_2^2 - \frac{1}{4} \mu_2 \sigma_1^2 \\ & + \frac{9}{4} \mu_1 \mu_2 \sigma_0^2 - \frac{9}{4} \mu_1 \mu_2 \sigma_0^4 + \frac{3}{16} \mu_1^2 \sigma_1^2 \sigma_2^2 \\ & + \frac{3}{16} \mu_2^2 \sigma_1^2 \sigma_2^2 + \frac{3}{16} \mu_1^2 \mu_2^2 \sigma_1^2 + \frac{3}{16} \mu_1^2 \mu_2^2 \sigma_2^2, \end{aligned}$$

$$\begin{aligned} \sigma_{\hat{\theta}_1 \hat{\theta}_2} = & -\frac{13}{8} \mu_0^2 - \frac{1}{4} \mu_0 \mu_1 - \frac{1}{4} \mu_0 \mu_2 - \frac{1}{8} \mu_1 \mu_2 + \frac{1}{8} \mu_1^2 \mu_2^4 + \frac{1}{8} \mu_1^4 \mu_2^2 - \frac{1}{4} \mu_1^3 \mu_2^3 \\ & - \frac{7}{8} \sigma_0^4 - \frac{1}{4} \sigma_0^6 - \frac{1}{2} \sigma_0^2 \sigma_1^2 - \frac{1}{2} \sigma_0^2 \sigma_2^2 - \frac{1}{2} \sigma_0^2 \sigma_1^2 \sigma_2^2 + \frac{1}{8} \sigma_1^2 \sigma_2^2 \\ & - \frac{1}{4} \sigma_0^4 \sigma_1^2 - \frac{1}{4} \sigma_0^4 \sigma_2^2 - \frac{1}{8} \sigma_1^4 \sigma_2^2 - \frac{1}{8} \sigma_1^2 \sigma_2^4 \\ & - \frac{9}{4} \mu_1 \mu_2 \sigma_0^2 + \frac{9}{4} \mu_1 \mu_2 \sigma_0^4 - \frac{3}{16} \mu_1^2 \sigma_1^2 \sigma_2^2 \\ & - \frac{3}{16} \mu_2^2 \sigma_1^2 \sigma_2^2 - \frac{3}{16} \mu_1^2 \mu_2^2 \sigma_1^2 - \frac{3}{16} \mu_1^2 \mu_2^2 \sigma_2^2. \end{aligned}$$

Démonstration. Pour tout $i \in \{1, \dots, n\}$, posons

$$W_i = (X_{i1}, X_{i2}, X_{i1}^2, X_{i2}^2, X_{i1}X_{i2}, X_{i1}^2X_{i2}, X_{i1}X_{i2}^2)$$

et pour tout $n \in \mathbb{N}$, soit en outre $\bar{W}_n = (W_1 + \dots + W_n)/n$. Puisque $W_1, W_2, \dots$ forment une suite de vecteurs mutuellement indépendants de même loi dont tous les moments d'ordre 2 sont finis, il découle de la version multivariée du théorème central limite énoncée à la proposition 3.1.1 que, quand $n \rightarrow \infty$,

$$\sqrt{n}(\bar{W}_n - \Gamma) \rightarrow \mathcal{N}_7(0, \Sigma)$$

où Γ et Σ dénotent respectivement l'espérance et la variance du vecteur

$$W = (X_1, X_2, X_1^2, X_2^2, X_1X_2, X_1^2X_2, X_1X_2^2).$$

Considérons maintenant deux applications $h_1 : \mathbb{R}^7 \rightarrow \mathbb{R}$ and $h_2 : \mathbb{R}^7 \rightarrow \mathbb{R}$ définies, pour tout $(x_1, \dots, x_7) \in \mathbb{R}^7$, par

$$\begin{aligned} h_1(x_1, x_2, x_3, x_4, x_5, x_6, x_7) &= \frac{1}{2} \left\{ (x_5 - x_1x_2)(1 - x_1 - x_2) + \frac{x_7 + x_6 - x_2x_3 - x_1x_4}{2} \right\}, \\ h_2(x_1, x_2, x_3, x_4, x_5, x_6, x_7) &= \frac{1}{2} \left\{ (x_5 - x_1x_2)(1 + x_1 + x_2) - \frac{x_7 + x_6 - x_2x_3 - x_1x_4}{2} \right\}. \end{aligned}$$

On déduit alors immédiatement des équations (2.6) et (2.7) que

$$\hat{\theta}_{n1} = h_1(\bar{W}_n), \quad \hat{\theta}_{n2} = h_2(\bar{W}_n).$$

Si on traite le couple (h_1, h_2) comme une fonction $h : \mathbb{R}^{14} \rightarrow \mathbb{R}^2$, on a alors

$$h(\bar{W}_n) = (h_1(\bar{W}_n), h_2(\bar{W}_n))^T = \hat{\Theta}_n.$$

En faisant appel à la méthode Delta, on conclut que, quand $n \rightarrow \infty$,

$$\sqrt{n}\{h(\bar{W}_n) - h(\Gamma)\} \rightarrow \mathcal{N}_2(0, \Sigma_{\hat{\Theta}}),$$

où la matrice $\Sigma_{\hat{\Theta}} = \epsilon^T \Sigma \epsilon$ est déduite de Σ à l'aide d'une matrice ϵ de dimension 7×2 dont les deux colonnes sont données par

$$\begin{aligned} \epsilon_1 &= \nabla h_1(w) \big|_{w=\Gamma} = (\epsilon_{11}, \epsilon_{21}, \epsilon_{31}, \epsilon_{41}, \epsilon_{51}, \epsilon_{61}, \epsilon_{71}), \\ \epsilon_2 &= \nabla h_2(w) \big|_{w=\Gamma} = (\epsilon_{12}, \epsilon_{22}, \epsilon_{32}, \epsilon_{42}, \epsilon_{52}, \epsilon_{62}, \epsilon_{72}). \end{aligned}$$

Après quelques calculs élémentaires, on trouve

$$\begin{aligned}
 \epsilon_{11} &= \frac{1}{4} \left[2\{E(X_2)\}^2 + 4E(X_1)E(X_2) - 2E(X_2) - 2E(X_1X_2) - E(X_2^2) \right], \\
 \epsilon_{21} &= \frac{1}{4} \left[2\{E(X_1)\}^2 + 4E(X_1)E(X_2) - 2E(X_1) - 2E(X_1X_2) - E(X_1^2) \right], \\
 \epsilon_{31} &= -\frac{1}{4} E(X_2), \\
 \epsilon_{41} &= -\frac{1}{4} E(X_1), \\
 \epsilon_{51} &= \frac{1}{2} \{1 - E(X_1) - E(X_2)\}, \\
 \epsilon_{61} &= \epsilon_{71} = \frac{1}{4}.
 \end{aligned}$$

De plus,

$$\begin{aligned}
 \epsilon_{12} &= \frac{1}{4} [E(X_2^2) - 2E(X_2) - 2\{E(X_2)\}^2 + 2E(X_1X_2) - 4E(X_1)E(X_2)], \\
 \epsilon_{22} &= \frac{1}{4} [E(X_1^2) - 2E(X_1) - 2\{E(X_1)\}^2 + 2E(X_1X_2) - 4E(X_1)E(X_2)], \\
 \epsilon_{32} &= \frac{1}{4} E(X_2), \\
 \epsilon_{42} &= \frac{1}{4} E(X_1), \\
 \epsilon_{52} &= \frac{1}{2} \{1 + E(X_1) + E(X_2)\}, \\
 \epsilon_{62} &= \epsilon_{71} = -\frac{1}{4}.
 \end{aligned}$$

Il ne reste alors qu'à calculer les 49 éléments de la matrice de variance de W . Compte tenu de la symétrie de $\text{var}(W)$, il faut étudier 28 termes faisant intervenir des moments mixtes de la forme $E(X_1^j X_2^k)$, où $j, k \in \{0, 1, 2, 3\}$. Pour ce faire, on peut utiliser la fonction génératrice des moments de la loi $\mathcal{BS}_2(\theta_1, \theta_2; \lambda_{11}, \lambda_{21}; \lambda_{12}, \lambda_{22})$, dont l'expression est obtenue en posant $s_1 = e^{t_1}$ et $s_2 = e^{t_2}$ dans la fonction génératrice de probabilités énoncée dans la proposition 2.3.6. On a, pour tous $t_1, t_2 \in \mathbb{R}$,

$$\begin{aligned}
 \phi(t_1, t_2) &= \exp\{(\lambda_{11} - \theta_1)(e^{t_1} - 1) + (\lambda_{12} - \theta_2)(e^{-t_1} - 1)\} \\
 &\quad \times \exp\{(\lambda_{21} - \theta_1)(e^{t_2} - 1) + (\lambda_{22} - \theta_2)(e^{-t_2} - 1)\} \\
 &\quad \times \exp\{\theta_1(e^{t_1+t_2} - 1) + \theta_2(e^{-t_1-t_2} - 1)\}.
 \end{aligned}$$

Ainsi on a, pour tous $j, k \in \mathbb{N}$,

$$E(X_1^j X_2^k) = \frac{\partial^{j+k}}{\partial t_1^j \partial t_2^k} \phi(t_1, t_2) \Big|_{t_1=0, t_2=0}.$$

On voit facilement que

$$\frac{\partial}{\partial t_1} \phi(t_1, t_2) = H(t_1, t_2) \phi(t_1, t_2),$$

et

$$\frac{\partial}{\partial t_2} \phi(t_1, t_2) = K(t_1, t_2) \phi(t_1, t_2),$$

où

$$\begin{aligned} H(t_1, t_2) &= (\lambda_{11} - \theta_1)e^{t_1} - (\lambda_{12} - \theta_2)e^{-t_1} + \theta_1 e^{t_1+t_2} - \theta_2 e^{-t_1-t_2}, \\ K(t_1, t_2) &= (\lambda_{21} - \theta_1)e^{t_2} - (\lambda_{22} - \theta_2)e^{-t_2} + \theta_1 e^{t_1+t_2} - \theta_2 e^{-t_1-t_2}, \end{aligned}$$

et

$$R(t_1, t_2) = \theta_1 e^{t_1+t_2} - \theta_2 e^{-t_1-t_2}.$$

Posons en outre

$$\bar{H}(t_1, t_2) = (\lambda_{11} - \theta_1)e^{t_1} + (\lambda_{12} - \theta_2)e^{-t_1} + \theta_1 e^{t_1+t_2} + \theta_2 e^{-t_1-t_2},$$

$$\bar{K}(t_1, t_2) = (\lambda_{21} - \theta_1)e^{t_2} + (\lambda_{22} - \theta_2)e^{-t_2} + \theta_1 e^{t_1+t_2} + \theta_2 e^{-t_1-t_2}$$

et

$$\bar{R}(t_1, t_2) = \theta_1 e^{t_1+t_2} + \theta_2 e^{-t_1-t_2}.$$

Il vient alors

$$\frac{\partial}{\partial t_1} H(t_1, t_2) = \bar{H}(t_1, t_2), \quad \frac{\partial}{\partial t_1} \bar{H}(t_1, t_2) = H(t_1, t_2),$$

$$\frac{\partial}{\partial t_2} K(t_1, t_2) = \bar{K}(t_1, t_2), \quad \frac{\partial}{\partial t_2} \bar{K}(t_1, t_2) = K(t_1, t_2),$$

$$\frac{\partial}{\partial t_2} H(t_1, t_2) = \bar{R}(t_1, t_2), \quad \frac{\partial}{\partial t_1} K(t_1, t_2) = \bar{R}(t_1, t_2)$$

et

$$\frac{\partial}{\partial t_2} \bar{H}(t_1, t_2) = R(t_1, t_2), \quad \frac{\partial}{\partial t_1} \bar{K}(t_1, t_2) = R(t_1, t_2).$$

De plus on a, pour $j \in \{1, 2\}$,

$$\frac{\partial}{\partial t_j} R(t_1, t_2) = \bar{R}(t_1, t_2), \quad \frac{\partial}{\partial t_j} \bar{R}(t_1, t_2) = R(t_1, t_2).$$

Par conséquent, on peut aisément exprimer $\partial^{j+k} \phi(t_1, t_2) / \partial t_1^j \partial t_2^k$ en fonction de ϕ à travers les fonctions H , K , R , $\bar{H}$, $\bar{K}$ et $\bar{R}$. Il suffit en fait d'évaluer ces fonctions en $(0, 0)$. On trouve $\phi(0, 0) = 1$ et

$$H(0, 0) = \mu_1, \quad \bar{H}(0, 0) = \sigma_1^2,$$

$$K(0, 0) = \mu_2, \quad \bar{K}(0, 0) = \sigma_1^2$$

et

$$R(0,0) = \mu_0, \quad \bar{R}(0,0) = \sigma_0^2.$$

Ces informations permettent ensuite de calculer $E(X_1^j X_2^k)$ pour tous $j, k \in \{0, 1, 2, 3\}$ en fonction des paramètres $\mu_0, \mu_1, \mu_2, \sigma_0^2, \sigma_1^2$ et σ_2^2 . Au bout du compte, on obtient les expressions de $\sigma_{\hat{\theta}_1}^2, \sigma_{\hat{\theta}_2}^2$ et $\sigma_{\hat{\theta}_1 \hat{\theta}_2}$ annoncées dans la proposition. Le détail des calculs se trouve dans l'annexe. $\square$

Pour valider les variances asymptotiques $\sigma_{\hat{\theta}_1}^2$ et $\sigma_{\hat{\theta}_2}^2$ énoncées dans la proposition 3.2.1, 100 échantillons de taille n de la loi de Skellam bivariée de deuxième espèce ont été engendrés en fixant

$$\theta_1 = 0.5, \quad \theta_2 = 1.5, \quad \lambda_{11} = 1, \quad \lambda_{12} = 2, \quad \lambda_{21} = 3, \quad \lambda_{22} = 4.$$

Les variances expérimentales $S_{\hat{\theta}_{n1}}^2$ et $S_{\hat{\theta}_{n2}}^2$ ont été calculées et comparées respectivement à $\sigma_{\hat{\theta}_1}^2/n$ et $\sigma_{\hat{\theta}_2}^2/n$. La procédure a été répétée pour chaque valeur de $n \in \{100, 500, 1000, 1500, 2000\}$. Les résultats sont consignés dans les tableaux 3.2 et 3.3, où l'on peut voir que la correspondance s'améliore à mesure que croît la taille d'échantillon.

n	$S_{\hat{\theta}_{n1}}^2$	$\sigma_{\hat{\theta}_1}^2/n$	Écart
100	0.795	0.801	0.006
500	0.150	0.160	0.010
1000	0.075	0.080	0.005
1500	0.052	0.053	0.001
2000	0.038	0.040	0.002

TABLE 3.2 – Écart observé entre l'approximation asymptotique de la variance de $\hat{\theta}_{n1}$ et sa valeur moyenne observée sur la base de 100 échantillons de taille n , où $n \in \{100, 500, 1000, 1500, 2000\}$.

n	$S_{\hat{\theta}_{n2}}^2$	$\sigma_{\hat{\theta}_2}^2/n$	Écart
100	1.107	1.101	0.006
500	0.223	0.220	0.003
1000	0.111	0.110	0.001
1500	0.074	0.073	0.001
2000	0.054	0.055	0.001

TABLE 3.3 – Écart observé entre l'approximation asymptotique de la variance de $\hat{\theta}_{n2}$ et sa valeur moyenne observée sur la base de 100 échantillons de taille n , où $n \in \{100, 500, 1000, 1500, 2000\}$.

3.3 Estimateurs à vraisemblance maximale des paramètres du modèle $\mathcal{BS}_1$

Nous allons maintenant examiner comment il est possible d'estimer les paramètres de la loi de Skellam bivariée de première espèce au moyen de la méthode du maximum de vraisemblance. À cet effet, soit $(X_{11}, X_{12}), \dots, (X_{n1}, X_{n2})$ un échantillon aléatoire provenant de la loi $\mathcal{BS}_1(\theta; \lambda_{11}, \lambda_{21}; \lambda_{12}, \lambda_{22})$. Posons

$$\Lambda = (\lambda_{11}, \lambda_{21}, \lambda_{12}, \lambda_{22}, \theta), \quad \lambda_1 = (\lambda_{11}, \lambda_{21}), \quad \lambda_2 = (\lambda_{12}, \lambda_{22})$$

et dénotons par

$$\mathbf{x}_1 = (x_{11}, \dots, x_{n1}), \quad \mathbf{x}_2 = (x_{12}, \dots, x_{n2})$$

les vecteurs d'observations correspondant aux deux marges de la loi. La fonction de vraisemblance de ce modèle est donnée par

$$L(\Lambda, \mathbf{x}_1, \mathbf{x}_2) = e^{-n(\lambda_{11} + \lambda_{12} + \lambda_{21} + \lambda_{22} - \theta)} \prod_{i=1}^n h(\Lambda, x_{i1}, x_{i2}),$$

où, pour tout $(x_{i1}, x_{i2}) \in \mathbb{Z}^2$,

$$h(\Lambda, x_{i1}, x_{i2}) = \sum_{k=0}^{\infty} g_{\lambda_1, \theta}(x_{i1} - k) g_{\lambda_2, \theta}(x_{i2} - k) g_{\theta}(k),$$

avec $g_{\theta}(k) = \theta^k / k!$ et, pour $j \in \{1, 2\}$,

$$g_{\lambda_j, \theta}(x) = \sum_{r=\max(0, -x)}^{\infty} \frac{(\lambda_{j1} - \theta)^{x+r} \lambda_{j2}^r}{(x+r)! r!}. \quad (3.1)$$

On commence par énoncer deux résultats préliminaires.

Lemme 3.3.1. *Pour tous $x \in \mathbb{Z}$ et $j \in \{1, 2\}$, on a*

$$\begin{aligned} \frac{\partial}{\partial \lambda_{j1}} g_{\lambda_j, \theta}(x) &= \frac{x}{\lambda_{j1} - \theta} g_{\lambda_j, \theta}(x) + \frac{\lambda_{j2}}{\lambda_{j1} - \theta} g_{\lambda_j, \theta}(x+1), \\ \frac{\partial}{\partial \lambda_{j2}} g_{\lambda_j, \theta}(x) &= g_{\lambda_j, \theta}(x+1), \\ \frac{\partial}{\partial \theta} g_{\lambda_j, \theta}(x) &= -\frac{\partial}{\partial \lambda_{j1}} g_{\lambda_j, \theta}(x), \\ \frac{\partial}{\partial \theta} g_{\theta}(x) &= g_{\theta}(x-1). \end{aligned}$$

Démonstration. Partant de (3.1), on remarque que

$$\begin{aligned} \frac{\partial}{\partial \lambda_{j1}} g_{\lambda_j, \theta}(x) &= \sum_{r=\max(0, -x)}^{\infty} \frac{(x+r)(\lambda_{j1} - \theta)^{x+r-1} \lambda_{j2}^r}{(x+r)! r!} \\ &= \frac{x}{\lambda_{j1} - \theta} g_{\lambda_j, \theta}(x) + \frac{\lambda_{j2}}{\lambda_{j1} - \theta} g_{\lambda_j, \theta}(x+1), \end{aligned}$$

soit la première identité. Les autres relations se déduisent de manière semblable. $\square$

Lemme 3.3.2. *Pour tout $(x_1, x_2) \in \mathbb{Z}^2$, on a*

$$\begin{aligned} \frac{\partial}{\partial \lambda_{11}} h(\Lambda, x_1, x_2) &= \frac{x_1}{\lambda_{11} - \theta} h(\Lambda, x_1, x_2) + \frac{\lambda_{12}}{\lambda_{11} - \theta} h(\Lambda, x_1 + 1, x_2) \\ &\quad - \frac{\theta}{\lambda_{11} - \theta} h(\Lambda, x_1 - 1, x_2 - 1), \\ \frac{\partial}{\partial \lambda_{12}} h(\Lambda, x_1, x_2) &= h(\Lambda, x_1 + 1, x_2), \\ \frac{\partial}{\partial \lambda_{21}} h(\Lambda, x_1, x_2) &= \frac{x_2}{\lambda_{21} - \theta} h(\Lambda, x_1, x_2) + \frac{\lambda_{22}}{\lambda_{21} - \theta} h(\Lambda, x_1, x_2 + 1) \\ &\quad - \frac{\theta}{\lambda_{21} - \theta} h(\Lambda, x_1 - 1, x_2 - 1), \\ \frac{\partial}{\partial \lambda_{22}} h(\Lambda, x_1, x_2) &= h(\Lambda, x_1, x_2 + 1), \\ \frac{\partial}{\partial \theta} h(\Lambda, x_1, x_2) &= -\frac{\partial}{\partial \lambda_{11}} h(\Lambda, x_1, x_2) - \frac{\partial}{\partial \lambda_{21}} h(\Lambda, x_1, x_2) \\ &\quad + h(\Lambda, x_1 - 1, x_2 - 1). \end{aligned}$$

Démonstration. Ces identités sont de simples conséquences de la définition de h et du lemme 3.3.1. $\square$

Ces lemmes jouent un rôle dans la démonstration du prochain résultat, qui précise les équations à résoudre pour l'obtention de l'estimation à vraisemblance maximale $\hat{\Lambda}_n$ du paramètre Λ dans le modèle de Skellam bivarié de première espèce.

Proposition 3.3.3. *Les estimations à vraisemblance maximale des paramètres λ_{11} , λ_{21} , λ_{12} , λ_{22} et θ sont les solutions des équations suivantes*

$$\hat{\lambda}_{n11} - \hat{\lambda}_{n12} = \bar{X}_{n1}, \quad \hat{\lambda}_{n21} - \hat{\lambda}_{n22} = \bar{X}_{n2},$$

$$Q_{n1} = \sum_{i=1}^n \frac{h(\hat{\Lambda}_n, X_{i1} + 1, X_{i2})}{h(\hat{\Lambda}_n, X_{i1}, X_{i2})} = n, \quad (3.2)$$

$$Q_{n2} = \sum_{i=1}^n \frac{h(\hat{\Lambda}_n, X_{i1}, X_{i2} + 1)}{h(\hat{\Lambda}_n, X_{i1}, X_{i2})} = n, \quad (3.3)$$

$$Q_{n3} = \sum_{i=1}^n \frac{h(\hat{\Lambda}_n, X_{i1} - 1, X_{i2} - 1)}{h(\hat{\Lambda}_n, X_{i1}, X_{i2})} = n, \quad (3.4)$$

sous la contrainte $\hat{\theta}_n \leq \min(\hat{\lambda}_{n11}, \hat{\lambda}_{n12})$.

Démonstration. L'estimateur à vraisemblance maximale de Λ est obtenu en résolvant le système à cinq équations suivant :

$$\frac{\partial}{\partial \theta} \ln\{L(\Lambda, \mathbf{x}_1, \mathbf{x}_2)\} = 0, \quad \frac{\partial}{\partial \lambda_{jk}} \ln\{L(\Lambda, \mathbf{x}_1, \mathbf{x}_2)\} = 0 \quad j, k \in \{1, 2\}. \quad (3.5)$$

En exploitant les relations énoncées dans le lemme 3.3.2, on trouve que, pour $j \in \{1, 2\}$,

$$\frac{\partial}{\partial \lambda_{j2}} \ln\{L(\Lambda, \mathbf{x}_1, \mathbf{x}_2)\} = -n + \sum_{i=1}^n \frac{1}{h(\Lambda, x_{i1}, x_{i2})} \frac{\partial}{\partial \lambda_{j2}} h(\Lambda, x_{i1}, x_{i2}) = 0,$$

ce qui entraîne que $Q_{n1} = Q_{n2} = n$. On voit de manière semblable que $Q_{n3} = n$ et que

$$\begin{aligned} \frac{n\bar{X}_{n1}}{\hat{\lambda}_{n11} - \hat{\theta}_n} + \frac{\hat{\lambda}_{n12}Q_{n1}}{\hat{\lambda}_{n11} - \hat{\theta}_n} - \frac{\hat{\theta}_nQ_{n2}}{\hat{\lambda}_{n11} - \hat{\theta}_n} &= n, \\ \frac{n\bar{X}_{n2}}{\hat{\lambda}_{n21} - \hat{\theta}_n} + \frac{\hat{\lambda}_{n22}Q_{n1}}{\hat{\lambda}_{n21} - \hat{\theta}_n} - \frac{\hat{\theta}_nQ_{n2}}{\hat{\lambda}_{n21} - \hat{\theta}_n} &= n. \end{aligned}$$

Il s'ensuit que $\hat{\lambda}_{n11} - \hat{\lambda}_{n12} = \bar{X}_{n1}$ et $\hat{\lambda}_{n21} - \hat{\lambda}_{n22} = \bar{X}_{n2}$, ce qui permet de conclure. $\square$

En pratique, on commence par remplacer $\hat{\lambda}_{n11}$ par $\bar{X}_{n1} + \hat{\lambda}_{n12}$ et $\hat{\lambda}_{n21}$ par $\bar{X}_{n2} + \hat{\lambda}_{n22}$. Puis on résout le système non linéaire composé des équations (3.2), (3.3) et (3.4) en se servant de

$$\hat{\Lambda}_n = (\bar{X}_{n1} + \hat{\lambda}_{n12}, \hat{\lambda}_{n12}, \bar{X}_{n2} + \hat{\lambda}_{n22}, \hat{\lambda}_{n22}, \hat{\theta}_n)$$

ce qui permet d'obtenir $\hat{\lambda}_{n12}$, $\hat{\lambda}_{n22}$ et $\hat{\theta}_n$.

3.4 Estimateurs à vraisemblance maximale des paramètres du modèle $\mathcal{BS}_2$

Nous allons maintenant établir les équations normales permettant d'estimer les paramètres λ_{11} , λ_{21} , λ_{12} , λ_{22} , θ_1 et θ_2 par la méthode à vraisemblance maximale pour le

modèle $\mathcal{BS}_2$. À cet effet, soit $(X_{11}, X_{12}), \dots, (X_{n1}, X_{n2})$ un échantillon aléatoire provenant de la loi $\mathcal{BS}_2(\theta_1, \theta_2; \lambda_{11}, \lambda_{21}; \lambda_{12}, \lambda_{22})$. Posons

$$\Lambda = (\lambda_{11}, \lambda_{21}, \lambda_{12}, \lambda_{22}, \theta_1, \theta_2), \quad \lambda_1 = (\lambda_{11}, \lambda_{21}), \quad \lambda_2 = (\lambda_{12}, \lambda_{22}), \quad \Theta = (\theta_1, \theta_2)$$

et dénotons par

$$\mathbf{x}_1 = (x_{11}, \dots, x_{n1}), \quad \mathbf{x}_2 = (x_{12}, \dots, x_{n2})$$

les vecteurs d'observations correspondant aux deux marges de la loi. La fonction de vraisemblance de ce modèle, à savoir

$$L(\Lambda, \mathbf{x}_1, \mathbf{x}_2) = e^{-n(\lambda_{11} + \lambda_{12} + \lambda_{21} + \lambda_{22} - \theta_1 - \theta_2)} \prod_{i=1}^n h(\Lambda, x_{i1}, x_{i2}),$$

dépend de la fonction de masse de (X_1, X_2) définie en tout $(x_1, x_2) \in \mathbb{Z}^2$ par

$$h(\Lambda, x_1, x_2) = \sum_{k=-\infty}^{\infty} g_{\lambda_1, \Theta}(x_1 - k) g_{\lambda_2, \Theta}(x_2 - k) g_{\Theta}(k)$$

et où, pour tout $x \in \mathbb{Z}$,

$$g_{\Theta}(x) = \sum_{r=\max(0, -x)}^{\infty} \frac{\theta_1^{x+r} \theta_2^r}{(x+r)! r!}$$

et pour $j \in \{1, 2\}$,

$$g_{\lambda_j, \Theta}(x) = \sum_{r=\max(0, -x)}^{\infty} \frac{(\lambda_{j1} - \theta_1)^{x+r} (\lambda_{j2} - \theta_2)^r}{(x+r)! r!}. \quad (3.6)$$

Commençons par établir deux lemmes préliminaires.

Lemme 3.4.1. *Pour tout $x \in \mathbb{Z}$ et $j \in \{1, 2\}$, on a*

$$\begin{aligned} \frac{\partial}{\partial \lambda_{j1}} g_{\lambda_j, \Theta}(x) &= \frac{x}{\lambda_{j1} - \theta_1} g_{\lambda_j, \Theta}(x) + \frac{\lambda_{j2} - \theta_2}{\lambda_{j1} - \theta_1} g_{\lambda_j, \Theta}(x+1), \\ \frac{\partial}{\partial \lambda_{j2}} g_{\lambda_j, \Theta}(x) &= g_{\lambda_j, \Theta}(x+1), \\ \frac{\partial}{\partial \theta_1} g_{\lambda_j, \Theta}(x) &= -\frac{\partial}{\partial \lambda_{j1}} g_{\lambda_j, \Theta}(x), \\ \frac{\partial}{\partial \theta_2} g_{\lambda_j, \Theta}(x) &= -g_{\lambda_j, \Theta}(x+1), \end{aligned}$$

et

$$\frac{\partial}{\partial \theta_1} g_{\Theta}(x) = g_{\Theta}(x-1), \quad \frac{\partial}{\partial \theta_2} g_{\Theta}(x) = g_{\Theta}(x+1).$$

Démonstration. En partant de (3.6), on trouve, pour tout $x \in \mathbb{Z}$ et $j \in \{1, 2\}$,

$$\begin{aligned} \frac{\partial}{\partial \lambda_{j1}} g_{\lambda_j, \Theta}(x) &= \sum_{r=\max(0, -x)}^{\infty} \frac{(x+r)(\lambda_{j1} - \theta_1)^{x+r-1}(\lambda_{j2} - \theta_2)^r}{(x+r)!r!} \\ &= \frac{x}{\lambda_{j1} - \theta_1} g_{\lambda_j, \Theta}(x) + \frac{\lambda_{j2} - \theta_2}{\lambda_{j1} - \theta_1} g_{\lambda_j, \Theta}(x+1), \end{aligned}$$

ce qui conduit à la première identité. Les autres équations se déduisent de la même manière. $\square$

Lemme 3.4.2. *Pour tout $(x_1, x_2) \in \mathbb{Z}^2$, on a*

$$\begin{aligned} \frac{\partial}{\partial \lambda_{11}} h(\Lambda, x_1, x_2) &= \frac{x_1}{\lambda_{11} - \theta_1} h(\Lambda, x_1, x_2) + \frac{\lambda_{12} - \theta_2}{\lambda_{11} - \theta_1} h(\Lambda, x_1 + 1, x_2) \\ &\quad - \frac{\theta_1}{\lambda_{11} - \theta_1} h(\Lambda, x_1 - 1, x_2 - 1) + \frac{\theta_2}{\lambda_{11} - \theta_1} h(\Lambda, x_1 + 1, x_2 + 1), \\ \frac{\partial}{\partial \lambda_{12}} h(\Lambda, x_1, x_2) &= h(\Lambda, x_1 + 1, x_2), \\ \frac{\partial}{\partial \lambda_{21}} h(\Lambda, x_1, x_2) &= \frac{x_2}{\lambda_{21} - \theta_1} h(\Lambda, x_1, x_2) + \frac{\lambda_{22} - \theta_2}{\lambda_{21} - \theta_1} h(\Lambda, x_1, x_2 + 1) \\ &\quad - \frac{\theta_1}{\lambda_{21} - \theta_1} h(\Lambda, x_1 - 1, x_2 - 1) + \frac{\theta_2}{\lambda_{21} - \theta_1} h(\Lambda, x_1 + 1, x_2 + 1), \\ \frac{\partial}{\partial \lambda_{22}} h(\Lambda, x_1, x_2) &= h(\Lambda, x_1, x_2 + 1), \\ \frac{\partial}{\partial \theta_1} h(\Lambda, x_1, x_2) &= -\frac{\partial}{\partial \lambda_{11}} h(\Lambda, x_1, x_2) - \frac{\partial}{\partial \lambda_{21}} h(\Lambda, x_1, x_2) + h(\Lambda, x_1 + 1, x_2 + 1), \\ \frac{\partial}{\partial \theta_2} h(\Lambda, x_1, x_2) &= -h(\Lambda, x_1 + 1, x_2) - h(\Lambda, x_1, x_2 + 1) + h(\Lambda, x_1 - 1, x_2 - 1). \end{aligned}$$

Démonstration. Pour démontrer la première identité, notons que

$$\begin{aligned} \frac{\partial}{\partial \lambda_{11}} h(\Lambda, x_1, x_2) &= \sum_{k=-\infty}^{\infty} \frac{\partial}{\partial \lambda_{11}} g_{\lambda_1, \Theta}(x_1 - k) g_{\lambda_2, \Theta}(x_2 - k) g_{\Theta}(k) \\ &= \frac{x_1}{\lambda_{11} - \theta_1} h(\Lambda, x_1, x_2) + \frac{\lambda_{12} - \theta_2}{\lambda_{11} - \theta_1} h(\Lambda, x_1 + 1, x_2) \\ &\quad - \frac{1}{\lambda_{11} - \theta_1} \sum_{k=-\infty}^{\infty} g_{\lambda_1, \Theta}(x_1 - k) g_{\lambda_2, \Theta}(x_2 - k) k g_{\Theta}(k) \\ &= \frac{x_1}{\lambda_{11} - \theta_1} h(\Lambda, x_1, x_2) + \frac{\lambda_{12} - \theta_2}{\lambda_{11} - \theta_1} h(\Lambda, x_1 + 1, x_2) \\ &\quad - \frac{\theta_1}{\lambda_{11} - \theta_1} h(\Lambda, x_1 - 1, x_2 - 1) + \frac{\theta_2}{\lambda_{11} - \theta_1} h(\Lambda, x_1 + 1, x_2 + 1). \end{aligned}$$

Cette dernière égalité est vraie parce que $k g_{\Theta}(k) = \theta_1 g_{\Theta}(k-1) - \theta_2 g_{\Theta}(k+1)$. Toutes les autres formules énoncées dans le lemme s'obtiennent de manière semblable. $\square$

Armés de ces lemmes, on peut aisément déduire les équations qui conduisent à l'estimation à vraisemblance maximale de $\Lambda = (\lambda_{11}, \lambda_{21}, \lambda_{12}, \lambda_{22}, \theta_1, \theta_2)$. Le résultat est énoncé ci-dessous.

Proposition 3.4.3. *Les estimations à vraisemblance maximale des paramètres λ_{11} , λ_{21} , λ_{12} , λ_{22} , θ_1 et θ_2 sont les solutions des équations suivantes :*

$$\hat{\lambda}_{n11} - \hat{\lambda}_{n12} = \bar{X}_{n1}, \quad \hat{\lambda}_{n21} - \hat{\lambda}_{n22} = \bar{X}_{n2},$$

$$Q_{n1} = \sum_{i=1}^n \frac{h(\hat{\Lambda}_n, X_{i1} - 1, X_{i2} - 1)}{h(\hat{\Lambda}_n, X_{i1}, X_{i2})} = n,$$

$$Q_{n2} = \sum_{i=1}^n \frac{h(\hat{\Lambda}_n, X_{i1} + 1, X_{i2})}{h(\hat{\Lambda}_n, X_{i1}, X_{i2})} = n,$$

$$Q_{n3} = \sum_{i=1}^n \frac{h(\hat{\Lambda}_n, X_{i1}, X_{i2} + 1)}{h(\hat{\Lambda}_n, X_{i1}, X_{i2})} = n,$$

$$Q_{n4} = \sum_{i=1}^n \frac{h(\hat{\Lambda}_n, X_{i1} + 1, X_{i2} + 1)}{h(\hat{\Lambda}_n, X_{i1}, X_{i2})} = n,$$

sous les contraintes suivantes :

$$\hat{\theta}_{n1} \leq \min(\hat{\lambda}_{n11}, \hat{\lambda}_{n12}), \quad \hat{\theta}_{n2} \leq \min(\hat{\lambda}_{n21}, \hat{\lambda}_{n22}).$$

Démonstration. En invoquant le lemme 3.4.2, on voit que

$$\frac{\partial}{\partial \theta_2} \ln\{L(\Lambda, \mathbf{x}_1, \mathbf{x}_2)\} = \frac{\partial}{\partial \lambda_{12}} \ln\{L(\Lambda, \mathbf{x}_1, \mathbf{x}_2)\} = \frac{\partial}{\partial \lambda_{22}} \ln\{L(\Lambda, \mathbf{x}_1, \mathbf{x}_2)\} = 0.$$

Il s'ensuit que $Q_{n1} = Q_{n2} = Q_{n3} = Q_{n4} = n$. On déduit ensuite que

$$\frac{n\bar{X}_{n1}}{\hat{\lambda}_{n11} - \hat{\theta}_{n1}} + \frac{(\hat{\lambda}_{n12} - \hat{\theta}_{n2})Q_{n1}}{\hat{\lambda}_{n11} - \hat{\theta}_{n1}} - \frac{\hat{\theta}_{n1}Q_{n3}}{\hat{\lambda}_{n11} - \hat{\theta}_{n1}} + \frac{\hat{\theta}_{n2}Q_{n4}}{\hat{\lambda}_{n11} - \hat{\theta}_{n1}} = n,$$

$$\frac{n\bar{X}_{n2}}{\hat{\lambda}_{n21} - \hat{\theta}_{n1}} + \frac{(\hat{\lambda}_{n22} - \hat{\theta}_{n2})Q_{n2}}{\hat{\lambda}_{n21} - \hat{\theta}_{n1}} - \frac{\hat{\theta}_{n1}Q_{n3}}{\hat{\lambda}_{n21} - \hat{\theta}_{n1}} + \frac{\hat{\theta}_{n2}Q_{n4}}{\hat{\lambda}_{n21} - \hat{\theta}_{n1}} = n.$$

Après simplifications, on conclut que $\hat{\lambda}_{n11} - \hat{\lambda}_{n12} = \bar{X}_{n1}$ et $\hat{\lambda}_{n21} - \hat{\lambda}_{n22} = \bar{X}_{n2}$. Le résultat s'ensuit. $\square$

Dans la pratique, on pose d'abord $\hat{\lambda}_{n11} = \bar{X}_{n1} + \hat{\lambda}_{n12}$ et $\hat{\lambda}_{n21} = \bar{X}_{n2} + \hat{\lambda}_{n22}$, puis on résout le système non linéaire formé des équations $Q_{n1} = Q_{n2} = Q_{n3} = Q_{n4} = n$ en posant

$$\hat{\Lambda}_n = (\bar{X}_{n1} + \hat{\lambda}_{n12}, \hat{\lambda}_{n12}, \bar{X}_{n2} + \hat{\lambda}_{n22}, \hat{\lambda}_{n22}, \hat{\theta}_{n1}, \hat{\theta}_{n2}),$$

ce qui permet de calculer $\hat{\lambda}_{n12}$, $\hat{\lambda}_{n22}$, $\hat{\theta}_{n1}$ et $\hat{\theta}_{n2}$.

3.5 Loi asymptotique des estimateurs du maximum de vraisemblance des paramètres du modèle $\mathcal{BS}_1$

L'objectif de cette section est d'établir la loi asymptotique de l'estimateur du maximum de vraisemblance $\hat{\Lambda}_n$ décrit dans la section précédente. Commençons par rappeler que la fonction de masse d'une paire aléatoire (X_1, X_2) obéissant à une loi de Skellam $\mathcal{BS}_1$ de paramètre $\Lambda = (\lambda_{11}, \lambda_{21}, \lambda_{12}, \lambda_{22}, \theta)$ est donnée en tout $(x_1, x_2) \in \mathbb{Z}^2$ par

$$f(\Lambda, x_1, x_2) = \sum_{k=0}^{\infty} \tilde{g}_{\lambda_1, \theta}(x_1 - k) \tilde{g}_{\lambda_2, \theta}(x_2 - k) \tilde{g}_{\theta}(k),$$

où $\tilde{g}_{\theta}(k) = \theta^k e^{-\theta} / k!$ pour tout $k \in \mathbb{N}$ et, pour tous $x \in \mathbb{N}$ et $j \in \{1, 2\}$,

$$\tilde{g}_{\lambda_j, \theta}(x) = e^{-(\lambda_{j1} + \lambda_{j2} - \theta)} \sum_{r=\max(0, -x)}^{\infty} \frac{(\lambda_{j1} - \theta)^{x+r} \lambda_{j2}^r}{(x+r)! r!}.$$

Proposition 3.5.1. *Soit $\hat{\Lambda}_n$ l'estimateur à vraisemblance maximale du vecteur $\Lambda = (\lambda_{11}, \lambda_{21}, \lambda_{12}, \lambda_{22}, \theta)$ des paramètres de la loi de Skellam de première espèce. Quand $n \rightarrow \infty$,*

$$\sqrt{n} (\hat{\Lambda}_n - \Lambda) \rightarrow \mathcal{N}_5(0, I^{-1}(\Lambda)),$$

où $I(\Lambda)$ est la matrice d'information de Fisher de taille dont l'élément $I_{j,k}$ est donné, pour tous $j, k \in \{1, \dots, 5\}$, par

$$I_{j,k} = \mathbb{E}\{\ell_j(\Lambda, X_1, X_2) \ell_k(\Lambda, X_1, X_2)\},$$

où

$$\begin{aligned} \ell_1(\Lambda, X_1, X_2) &= \frac{f(\Lambda, X_1 - 1, X_2) - f(\Lambda, X_1, X_2)}{f(\Lambda, X_1, X_2)}, \\ \ell_2(\Lambda, X_1, X_2) &= \frac{f(\Lambda, X_1 + 1, X_2) - f(\Lambda, X_1, X_2)}{f(\Lambda, X_1, X_2)}, \\ \ell_3(\Lambda, X_1, X_2) &= \frac{f(\Lambda, X_1, X_2 - 1) - f(\Lambda, X_1, X_2)}{f(\Lambda, X_1, X_2)}, \\ \ell_4(\Lambda, X_1, X_2) &= \frac{f(\Lambda, X_1, X_2 + 1) - f(\Lambda, X_1, X_2)}{f(\Lambda, X_1, X_2)}, \end{aligned}$$

et

$$\begin{aligned} \ell_5(\Lambda, X_1, X_2) &= \frac{1}{f(\Lambda, X_1, X_2)} \{f(\Lambda, X_1 - 1, X_2 - 1) - f(\Lambda, X_1 + 1, X_2) \\ &\quad - f(\Lambda, X_1, X_2 + 1) + f(\Lambda, X_1, X_2)\}. \end{aligned}$$

Démonstration. Il découle de la théorie asymptotique classique de l'estimation à vraisemblance maximale que, pour tous $j, k \in \{1, \dots, 5\}$,

$$I_{j,k}(\Lambda) = E \left[\frac{\partial}{\partial \theta_j} \ln \{f(\Lambda, X_1, X_2)\} \frac{\partial}{\partial \theta_k} \ln \{f(\Lambda, X_1, X_2)\} \right],$$

où on pose par commodité $\theta_1 = \lambda_{11}$, $\theta_2 = \lambda_{12}$, $\theta_3 = \lambda_{21}$, $\theta_4 = \lambda_{22}$, $\theta_5 = \theta$. On vérifie facilement que

$$\frac{\partial}{\partial \theta_j} \ln \{f(\Lambda, X_1, X_2)\} = \ell_j(\Lambda, X_1, X_2)$$

pour tout $j \in \{1, \dots, 5\}$, d'où le résultat. $\square$

En particulier, la variance limite de l'estimateur du maximum de vraisemblance de θ est donnée par le terme (5,5) de l'inverse de $I(\Lambda)$. Pour vérifier la qualité de l'approximation asymptotique, 100 échantillons de taille n de la loi de Skellam bivariée de première espèce ont été engendrés en fixant

$$\theta = 0.5, \quad \lambda_{11} = 1, \quad \lambda_{12} = 2, \quad \lambda_{21} = 3, \quad \lambda_{22} = 4.$$

La variance expérimentale $S_{\hat{\theta}_n}^2$ de $\hat{\theta}_n$ a ensuite été calculée et comparée à $\sigma_{\hat{\theta}_{MV}}^2/n$. La procédure a été répétée pour chaque valeur de $n \in \{100, 500, 1000, 1500, 2000\}$. Les résultats sont consignés dans le tableau 3.4, où l'on peut constater que la correspondance est excellente, surtout pour les grandes tailles d'échantillon.

n	$S_{\hat{\theta}_n}^2$	$\sigma_{\hat{\theta}_{MV}}^2/n$	Écart
100	0.227	0.217	0.010
500	0.045	0.043	0.002
1000	0.022	0.021	0.001
1500	0.015	0.014	0.001
2000	0.011	0.010	0.001

TABLE 3.4 – Écart observé entre l'approximation asymptotique de la variance de $\hat{\theta}_n$ et sa valeur moyenne observée sur la base de 100 échantillons de taille n , où $n \in \{100, 500, 1000, 1500, 2000\}$.

3.6 Loi asymptotique des estimateurs du maximum de vraisemblance des paramètres du modèle $\mathcal{BS}_2$

Cette section a pour objectif d'établir la loi asymptotique de l'estimateur de maximum de vraisemblance $\hat{\Lambda}$ pour le modèle $\mathcal{BS}_2$. Rappelons que la fonction de masse de

$(X_1, X_2) \sim \mathcal{BS}_2(\theta_1, \theta_2; \lambda_{11}, \lambda_{21}; \lambda_{12}, \lambda_{22})$ est donnée en tout $(x_1, x_2) \in \mathbb{Z}^2$ par

$$f(\Lambda, x_1, x_2) = \sum_{k=-\infty}^{\infty} \tilde{g}_{\lambda_1, \Theta}(x_1 - k) \tilde{g}_{\lambda_2, \Theta}(x_2 - k) \tilde{g}_{\Theta}(k),$$

où pour tout $x \in \mathbb{Z}$,

$$\tilde{g}_{\Theta}(x) = e^{-(\theta_1 + \theta_2)} \sum_{r=\max(0, -x)}^{\infty} \frac{\theta_1^{x+r} \theta_2^r}{(x+r)! r!}.$$

et pour $j \in \{1, 2\}$,

$$\tilde{g}_{\lambda_j, \Theta}(x) = e^{-(\lambda_{j1} + \lambda_{j2} - \theta_1 - \theta_2)} \sum_{r=\max(0, -x)}^{\infty} \frac{(\lambda_{j1} - \theta_1)^{x+r} (\lambda_{j2} - \theta_2)^r}{(x+r)! r!}.$$

Proposition 3.6.1. *Soit $\hat{\Lambda}_n$ l'estimateur à vraisemblance maximale du vecteur $\Lambda = (\lambda_{11}, \lambda_{21}, \lambda_{12}, \lambda_{22}, \theta_1, \theta_2)$ des paramètres de la loi de Skellam de deuxième espèce. Quand $n \rightarrow \infty$,*

$$\sqrt{n}(\hat{\Lambda}_n - \Lambda) \rightarrow \mathcal{N}_6(0, I^{-1}(\Lambda)),$$

où $I(\Lambda)$ représente la matrice d'information de Fisher dont l'élément (j, k) est donné, pour tous $j, k \in \{1, \dots, 6\}$, par

$$I_{j,k} = E\{\ell_j(\Lambda, X_1, X_2) \ell_k(\Lambda, X_1, X_2)\}$$

où

$$\begin{aligned} \ell_1(\Lambda, X_1, X_2) &= \frac{f(\Lambda, X_1 - 1, X_2) - f(\Lambda, X_1, X_2)}{f(\Lambda, X_1, X_2)}, \\ \ell_2(\Lambda, X_1, X_2) &= \frac{f(\Lambda, X_1 + 1, X_2) - f(\Lambda, X_1, X_2)}{f(\Lambda, X_1, X_2)}, \\ \ell_3(\Lambda, X_1, X_2) &= \frac{f(\Lambda, X_1, X_2 - 1) - f(\Lambda, X_1, X_2)}{f(\Lambda, X_1, X_2)}, \\ \ell_4(\Lambda, X_1, X_2) &= \frac{f(\Lambda, X_1, X_2 + 1) - f(\Lambda, X_1, X_2)}{f(\Lambda, X_1, X_2)}, \end{aligned}$$

$$\begin{aligned} \ell_5(\Lambda, X_1, X_2) &= \frac{1}{f(\Lambda, X_1, X_2)} \{f(\Lambda, X_1 - 1, X_2 - 1) - f(\Lambda, X_1 + 1, X_2) \\ &\quad - f(\Lambda, X_1, X_2 + 1) + f(\Lambda, X_1, X_2)\} \end{aligned}$$

et

$$\begin{aligned} \ell_6(\Lambda, X_1, X_2) &= \frac{1}{f(\Lambda, X_1, X_2)} \{f(\Lambda, X_1 + 1, X_2 + 1) - f(\Lambda, X_1 - 1, X_2) \\ &\quad - f(\Lambda, X_1, X_2 - 1) + f(\Lambda, X_1, X_2)\}. \end{aligned}$$

n	$S_{\hat{\theta}_{n1}}^2$	$\sigma_{\hat{\theta}_{1MV}}^2/n$	Écart
100	0.600	0.514	0.086
500	0.120	0.102	0.018
1000	0.062	0.051	0.011
1500	0.042	0.034	0.008
2000	0.028	0.025	0.003

TABLE 3.5 – Écart observé entre l'approximation asymptotique de la variance de $\hat{\theta}_{n1}$ et sa valeur moyenne observée sur la base de 100 échantillons de taille n , où $n \in \{100, 500, 1000, 1500, 2000\}$.

Au vu de ce résultat, les variances asymptotiques des estimateurs de θ_1 et θ_2 sont donc données par les termes (5, 5) et (6, 6) de l'inverse de $I(\Lambda)$. Pour vérifier la qualité de l'approximation asymptotique, 100 échantillons de taille n de la loi de Skellam bivariée de première espèce ont été engendrés en fixant

$$\theta_1 = 0.5, \theta_2 = 1.5, \quad \lambda_{11} = 1, \quad \lambda_{12} = 2, \quad \lambda_{21} = 3, \quad \lambda_{22} = 4.$$

Les variances expérimentales $S_{\hat{\theta}_{n1}}^2$ et $S_{\hat{\theta}_{n2}}^2$ de $\hat{\theta}_{n1}$ et $\hat{\theta}_{n2}$ ont ensuite été calculées et comparées respectivement à $\sigma_{\hat{\theta}_{1MV}}^2/n$ et $\sigma_{\hat{\theta}_{2MV}}^2/n$.

Comme dans tous les cas précédents, la procédure a été répétée pour chaque valeur de $n \in \{100, 500, 1000, 1500, 2000\}$. Les résultats sont consignés dans les tableaux 3.5 et 3.6, où l'on voit que la correspondance s'améliore à mesure que $n \rightarrow \infty$.

n	$S_{\hat{\theta}_{n2}}^2$	$\sigma_{\hat{\theta}_{2MV}}^2/n$	Écart
100	0.996	0.905	0.091
500	0.197	0.180	0.017
1000	0.097	0.090	0.007
1500	0.065	0.060	0.005
2000	0.049	0.045	0.004

TABLE 3.6 – Écart observé entre l'approximation asymptotique de la variance de $\hat{\theta}_{n2}$ et sa valeur moyenne observée sur la base de 100 échantillons de taille n , où $n \in \{100, 500, 1000, 1500, 2000\}$.

3.7 Comparaison entre les deux approches

Cette dernière section explore brièvement la question de savoir si la variance asymptotique de l'estimateur du maximum de vraisemblance $\hat{\sigma}_{MV}^2$ du paramètre de dépendance de la loi de Skellam bivariée est plus petite ou non que la variance asymptotique $\hat{\sigma}_{MM}^2$ de l'estimateur des moments.

La réponse peut évidemment varier selon que l'on considère la loi de Skellam de première ou de deuxième espèce. Elle peut aussi dépendre des valeurs des paramètres marginaux. À des fins de comparaison, on se limitera ici au cas où

$$\lambda_{11} = 1, \quad \lambda_{12} = 2, \quad \lambda_{21} = 3, \quad \lambda_{22} = 4. \quad (3.7)$$

La figure 3.1 montre le comportement de $\hat{\sigma}_{\hat{\theta}_{MV}}^2$ et de $\hat{\sigma}_{\hat{\theta}_{MM}}^2$ lorsque θ varie dans l'intervalle $[0.05, 0.95]$ dans le cadre du modèle de Skellam bivarié de première espèce. Comme on peut le constater, $\hat{\sigma}_{\hat{\theta}_{MV}}^2 < \hat{\sigma}_{\hat{\theta}_{MM}}^2$ si $\theta \leq 0.28$ (environ) et le contraire se produit quand $\theta \geq 0.28$. Il semble donc que l'estimateur des moments soit plus efficace que celui du maximum de vraisemblance (pour cette combinaison de paramètres marginaux), ce qui est un peu étonnant a priori. La contrainte $\theta \leq \min(\lambda_{11}, \lambda_{12})$, qui fait en sorte que l'espace paramétrique n'est pas carré, invalide sans doute la borne de Cramér–Rao–Fréchet. Si tel est le cas, on ne s'étonne pas que l'EMV soit inefficace.

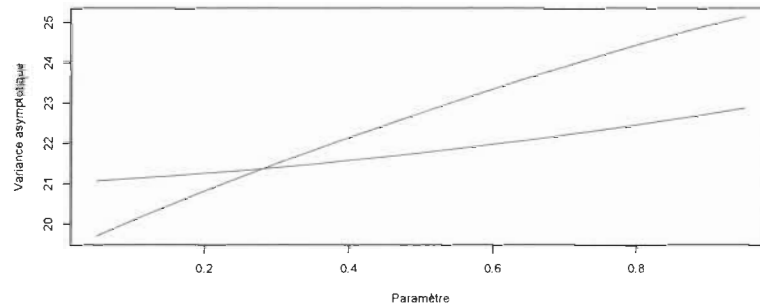


FIGURE 3.1 – Variance asymptotique de deux estimateurs du paramètre de dépendance θ dans le modèle de Skellam bivarié de première espèce lorsque les paramètres marginaux sont tels que donnés en (3.7). En rouge : estimateur à vraisemblance maximale ; en bleu : estimateur des moments.

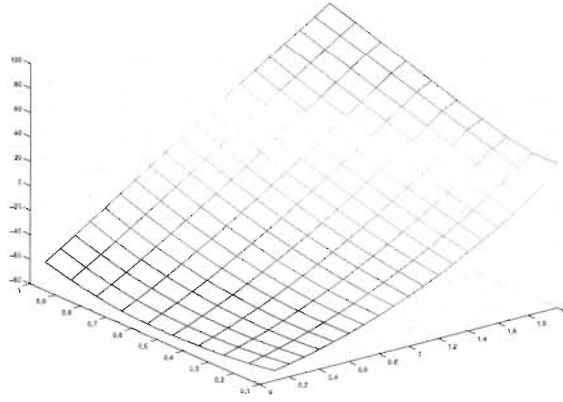


FIGURE 3.2 – Différence entre la variance asymptotique de deux estimateurs du paramètre de dépendance θ_1 dans le modèle de Skellam bivarié de deuxième espèce lorsque $\theta_1 \in [0, 1]$ et $\theta_2 \in [0, 2]$. En bleu : zone où l'estimateur des moments est moins efficace.

Le même phénomène se produit lorsque l'on compare l'efficacité asymptotique des estimateurs des moments et à vraisemblance maximale des paramètres de dépendance θ_1 et θ_2 dans le modèle de Skellam bivarié de deuxième espèce. La différence $\hat{\sigma}_{\hat{\theta}_{jMM}}^2 - \hat{\sigma}_{\hat{\theta}_{jMV}}^2$ est tracée en fonction de θ_1 et de θ_2 dans les figures 3.2 (pour $j = 1$) et 3.3 (pour $j = 2$). Cette différence peut être positive, négative ou nulle, selon les valeurs de θ_1 et θ_2 .

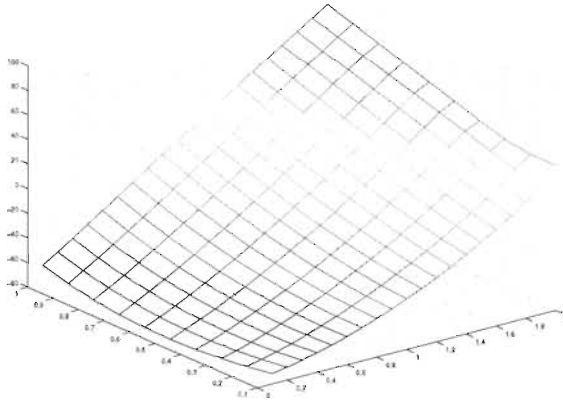


FIGURE 3.3 – Différence entre la variance asymptotique de deux estimateurs du paramètre de dépendance θ_2 dans le modèle de Skellam bivarié de deuxième espèce lorsque $\theta_1 \in [0, 1]$ et $\theta_2 \in [0, 2]$. En bleu : zone où l'estimateur des moments est moins efficace.

Conclusion

Ce mémoire est consacré à l'étude du comportement asymptotique de différents estimateurs des paramètres de dépendance des lois de Skellam bivariées de première et de deuxième espèce récemment proposées par Genest & Mesfioui (2014).

La première contribution originale de ce travail a consisté à montrer que les estimateurs des moments présentés par Genest & Mesfioui (2014) sont convergents et asymptotiquement gaussiens. Après avoir déterminé de façon explicite les variances asymptotiques de ces estimateurs, nous avons validé nos calculs par voie de simulation.

La seconde contribution originale de ce travail porte sur l'estimation des paramètres de dépendance des lois de Skellam bivariées de Genest & Mesfioui (2014) par la méthode du maximum de vraisemblance. En plus d'avoir montré comment calculer ces estimateurs au moyen des équations normales, nous avons établi leur convergence et leur normalité asymptotique.

Enfin, le calcul explicite des variances limites des estimateurs des moments et du maximum de vraisemblance nous a permis de réaliser une étude numérique de l'efficacité relative de ces estimateurs dans le contexte des lois de Skellam bivariées de première et de deuxième espèce.

Annexe

Pour calculer les moments de la forme $E(X_1^i X_2^j)$ pour tous $i, j \in \{0, 1, 2, 3\}$, on a recours à la fonction génératrice des moments bivariée, définie pour tous $t_1, t_2 \in \mathbb{R}$, par

$$\begin{aligned}\phi(t_1, t_2) &= \exp\{(\lambda_{11} - \theta_1)(e^{t_1} - 1) + (\lambda_{12} - \theta_2)(e^{-t_1} - 1)\} \\ &\times \exp\{(\lambda_{21} - \theta_1)(e^{t_2} - 1) + (\lambda_{22} - \theta_2)(e^{-t_2} - 1)\} \\ &\times \exp\{\theta_1(e^{t_1+t_2} - 1) + \theta_2(e^{-t_1-t_2} - 1)\}.\end{aligned}$$

En effet, on a

$$E(X_1^i X_2^j) = \frac{\partial^{i+j}}{\partial t_1^i \partial t_2^j} \phi(t_1, t_2) \Big|_{t_1=0, t_2=0}.$$

On a d'abord, pour tous $t_1, t_2 \in \mathbb{R}$,

$$\frac{\partial}{\partial t_1} \phi(t_1, t_2) = H(t_1, t_2) \phi(t_1, t_2), \quad \frac{\partial}{\partial t_2} \phi(t_1, t_2) = K(t_1, t_2) \phi(t_1, t_2),$$

où

$$\begin{aligned}H(t_1, t_2) &= (\lambda_{11} - \theta_1)e^{t_1} - (\lambda_{12} - \theta_2)e^{-t_1} + \theta_1 e^{t_1+t_2} - \theta_2 e^{-t_1-t_2}, \\ K(t_1, t_2) &= (\lambda_{21} - \theta_1)e^{t_2} - (\lambda_{22} - \theta_2)e^{-t_2} + \theta_1 e^{t_1+t_2} - \theta_2 e^{-t_1-t_2},\end{aligned}$$

et

$$R(t_1, t_2) = \theta_1 e^{t_1+t_2} - \theta_2 e^{-t_1-t_2}.$$

Définissons aussi, pour tous $t_1, t_2 \in \mathbb{R}$,

$$\begin{aligned}\bar{H}(t_1, t_2) &= (\lambda_{11} - \theta_1)e^{t_1} + (\lambda_{12} - \theta_2)e^{-t_1} + \theta_1 e^{t_1+t_2} + \theta_2 e^{-t_1-t_2}, \\ \bar{K}(t_1, t_2) &= (\lambda_{21} - \theta_1)e^{t_2} + (\lambda_{22} - \theta_2)e^{-t_2} + \theta_1 e^{t_1+t_2} + \theta_2 e^{-t_1-t_2}.\end{aligned}$$

et

$$\bar{R}(t_1, t_2) = \theta_1 e^{t_1+t_2} + \theta_2 e^{-t_1-t_2}.$$

On se convainc facilement que

$$\frac{\partial}{\partial t_1} H(t_1, t_2) = \bar{H}(t_1, t_2), \quad \frac{\partial}{\partial t_2} H(t_1, t_2) = \bar{R}(t_1, t_2),$$

$$\begin{aligned}\frac{\partial}{\partial t_1} \bar{H}(t_1, t_2) &= H(t_1, t_2), & \frac{\partial}{\partial t_2} \bar{H}(t_1, t_2) &= R(t_1, t_2), \\ \frac{\partial}{\partial t_1} K(t_1, t_2) &= \bar{R}(t_1, t_2), & \frac{\partial}{\partial t_2} K(t_1, t_2) &= \bar{K}(t_1, t_2),\end{aligned}$$

et

$$\frac{\partial}{\partial t_1} \bar{K}(t_1, t_2) = R(t_1, t_2), \quad \frac{\partial}{\partial t_2} \bar{K}(t_1, t_2) = K(t_1, t_2).$$

Pour $i \in \{1, 2\}$, on a en outre

$$\frac{\partial}{\partial t_i} R(t_1, t_2) = \bar{R}(t_1, t_2), \quad \frac{\partial}{\partial t_i} \bar{R}(t_1, t_2) = R(t_1, t_2).$$

Il est donc clair que $\partial^{i+j}\phi(t_1, t_2)/\partial t_1^i \partial t_2^j$ s'exprime en termes de ϕ et des fonctions H , $\bar{H}$, K , $\bar{K}$, R et $\bar{R}$. Il nous reste à évaluer ces fonctions au point $(0, 0)$. On trouve

$$\phi(0, 0) = 1, \quad H(0, 0) = \mu_1, \quad \bar{H}(0, 0) = \sigma_1^2$$

et

$$K(0, 0) = \mu_2, \quad \bar{K}(0, 0) = \sigma_2^2, \quad R(0, 0) = \mu_0, \quad \bar{R}(0, 0) = \sigma_0^2.$$

Il s'ensuit que

$$\begin{aligned}E(X_1) &= \left. \frac{\partial}{\partial t_1} \phi(t_1, t_2) \right|_{(t_1, t_2)=(0,0)} = \mu_1, \\ E(X_1^2) &= \left. \frac{\partial^2}{\partial t_1^2} \phi(t_1, t_2) \right|_{(t_1, t_2)=(0,0)} = \mu_1^2 + \sigma_1^2, \\ E(X_1^3) &= \left. \frac{\partial^3}{\partial t_1^3} \phi(t_1, t_2) \right|_{(t_1, t_2)=(0,0)} = \mu_1 + \mu_1^3 + 3\mu_1\sigma_1^2, \\ E(X_1^4) &= \left. \frac{\partial^4}{\partial t_1^4} \phi(t_1, t_2) \right|_{(t_1, t_2)=(0,0)} = 4\mu_1^2 + \mu_1^4 + \sigma_1^2 + 3\sigma_1^4 + 6\mu_1^2\sigma_1^2, \\ E(X_2) &= \left. \frac{\partial}{\partial t_2} \phi(t_1, t_2) \right|_{(t_1, t_2)=(0,0)} = \mu_2, \\ E(X_2^2) &= \left. \frac{\partial^2}{\partial t_2^2} \phi(t_1, t_2) \right|_{(t_1, t_2)=(0,0)} = \mu_2^2 + \sigma_2^2, \\ E(X_2^3) &= \left. \frac{\partial^3}{\partial t_2^3} \phi(t_1, t_2) \right|_{(t_1, t_2)=(0,0)} = \mu_2 + \mu_2^3 + 3\mu_2\sigma_2^2, \\ E(X_2^4) &= \left. \frac{\partial^4}{\partial t_2^4} \phi(t_1, t_2) \right|_{(t_1, t_2)=(0,0)} = 4\mu_2^2 + \mu_2^4 + \sigma_2^2 + 3\sigma_2^4 + 6\mu_2^2\sigma_2^2,\end{aligned}$$

et

$$\begin{aligned}
E(X_1 X_2) &= \frac{\partial^2}{\partial t_1 \partial t_2} \phi(t_1, t_2) \Big|_{(t_1, t_2)=(0,0)} = \mu_1 \mu_2 + \sigma_0^2, \\
E(X_1^2 X_2) &= \frac{\partial^3}{\partial t_1^2 \partial t_2} \phi(t_1, t_2) \Big|_{(t_1, t_2)=(0,0)} = \mu_0 + \mu_1^2 \mu_2 + 2\mu_1 \sigma_0^2 + \mu_2 \sigma_1^2, \\
E(X_1 X_2^2) &= \frac{\partial^3}{\partial t_1 \partial t_2^2} \phi(t_1, t_2) \Big|_{(t_1, t_2)=(0,0)} = \mu_0 + \mu_1 \mu_2^2 + 2\mu_2 \sigma_0^2 + \mu_1 \sigma_2^2, \\
E(X_1^2 X_2^2) &= \frac{\partial^4}{\partial t_1^2 \partial t_2^2} \phi(t_1, t_2) \Big|_{(t_1, t_2)=(0,0)} = 2\mu_0 \mu_1 + 2\mu_0 \mu_2 + \mu_1^2 \mu_2^2 + \sigma_0^2 + 2\sigma_0^4 + \mu_1^2 \sigma_2^2 \\
&\quad + \mu_2^2 \sigma_1^2 + 4\mu_1 \mu_2 \sigma_0^2, \\
E(X_1^3 X_2) &= \frac{\partial^4}{\partial t_1^3 \partial t_2} \phi(t_1, t_2) \Big|_{(t_1, t_2)=(0,0)} = 3\mu_0 \mu_1 + \mu_1 \mu_2 + \mu_1^3 \mu_2 \\
&\quad + \sigma_0^2 + 3\sigma_0^2 \sigma_1^2 + 3\mu_1^2 \sigma_0^2 + 3\mu_1 \mu_2 \sigma_1^2, \\
E(X_1 X_2^3) &= \frac{\partial^4}{\partial t_1 \partial t_2^3} \phi(t_1, t_2) \Big|_{(t_1, t_2)=(0,0)} = 3\mu_0 \mu_2 + \mu_1 \mu_2 + \mu_1 \mu_2^3 \\
&\quad + \sigma_0^2 + 3\sigma_0^2 \sigma_2^2 + 3\mu_2^2 \sigma_0^2 + 3\mu_1 \mu_2 \sigma_2^2.
\end{aligned}$$

On trouve aussi

$$\begin{aligned}
E(X_1^3 X_2^2) &= \frac{\partial^5}{\partial t_1^3 \partial t_2^2} \phi(t_1, t_2) \Big|_{(t_1, t_2)=(0,0)} \\
&= \mu_0 + 3\mu_0 \mu_1^2 + \mu_1 \mu_2^2 + \mu_1^3 \mu_2^2 + 6\mu_0 \sigma_0^2 + 3\mu_0 \sigma_1^2 + 2\mu_2 \sigma_0^2 \\
&\quad + 3\mu_1 \sigma_0^2 + 6\mu_1 \sigma_0^4 + \mu_1 \sigma_2^2 + \mu_1^3 \sigma_2^2 \\
&\quad + 6\mu_0 \mu_1 \mu_2 + 3\mu_1 \mu_2^2 \sigma_1^2 + 6\mu_1^2 \mu_2 \sigma_0^2 + 3\mu_1 \sigma_1^2 \sigma_2^2 + 6\mu_2 \sigma_0^2 \sigma_1^2, \\
E(X_1^2 X_2^3) &= \frac{\partial^5}{\partial t_1^2 \partial t_2^3} \phi(t_1, t_2) \Big|_{(t_1, t_2)=(0,0)} \\
&= \mu_0 + 3\mu_0 \mu_2^2 + \mu_1^2 \mu_2 + \mu_1^2 \mu_2^3 + 6\mu_0 \sigma_0^2 + 3\mu_0 \sigma_2^2 + 2\mu_1 \sigma_0^2 \\
&\quad + 3\mu_2 \sigma_0^2 + 6\mu_2 \sigma_0^4 + \mu_2 \sigma_1^2 + \mu_2^3 \sigma_1^2 \\
&\quad + 6\mu_0 \mu_1 \mu_2 + 3\mu_1^2 \mu_2 \sigma_2^2 + 6\mu_1 \mu_2^2 \sigma_0^2 + 3\mu_2 \sigma_1^2 \sigma_2^2 + 6\mu_1 \sigma_0^2 \sigma_2^2, \\
E(X_1^4 X_2) &= \frac{\partial^5}{\partial t_1^4 \partial t_2} \phi(t_1, t_2) \Big|_{(t_1, t_2)=(0,0)} \\
&= \mu_0 + 6\mu_0 \mu_1^2 + \mu_1^4 \mu_2 + 4\mu_1^2 \mu_2 + 6\mu_0 \sigma_1^2 + 8\mu_1 \sigma_0^2 \\
&\quad + \mu_2 \sigma_1^2 + 3\mu_2 \sigma_1^4 + 4\mu_1^3 \sigma_0^2 + 12\mu_1 \sigma_0^2 \sigma_1^2 + 6\mu_1^2 \mu_2 \sigma_1^2, \\
E(X_1 X_2^4) &= \frac{\partial^5}{\partial t_1 \partial t_2^4} \phi(t_1, t_2) \Big|_{(t_1, t_2)=(0,0)} \\
&= \mu_0 + 6\mu_0 \mu_2^2 + \mu_1 \mu_2^4 + 4\mu_1 \mu_2^2 + 6\mu_0 \sigma_2^2 + 8\mu_2 \sigma_0^2 \\
&\quad + \mu_1 \sigma_2^2 + 3\mu_1 \sigma_2^4 + 4\mu_2^3 \sigma_0^2 + 12\mu_2 \sigma_0^2 \sigma_2^2 + 6\mu_1 \mu_2^2 \sigma_2^2,
\end{aligned}$$

ainsi que

$$\begin{aligned}
E(X_1^3 X_2^3) &= \frac{\partial^6}{\partial t_1^3 \partial t_2^3} \phi(t_1, t_2) \Big|_{(t_1, t_2)=(0,0)} \\
&= 9\mu_0^2 + 3\mu_0\mu_1 + 3\mu_0\mu_2 + \mu_1\mu_2 + \mu_1\mu_2^3 + \mu_2\mu_1^3 + \mu_1^3\mu_2^3 + 9\mu_0\mu_1\mu_2^2 + 9\mu_0\mu_1^2\mu_2 \\
&\quad + \sigma_0^2 + 9\sigma_0^4 + 6\sigma_0^6 + 3\sigma_0^2\sigma_1^2 + 3\sigma_0^2\sigma_2^2 \\
&\quad + 3\mu_1^2\sigma_0^2 + 3\mu_2^2\sigma_0^2 + 18\mu_0\mu_1\sigma_0^2 + 9\mu_0\mu_1\sigma_2^2 + 9\mu_0\mu_2\sigma_1^2 \\
&\quad + 9\mu_1\mu_2\sigma_0^2 + 3\mu_1\mu_2\sigma_1^2 + 3\mu_1\mu_2\sigma_2^2 + 18\mu_1\mu_2\sigma_0^4 + 3\mu_1\mu_2^3\sigma_1^2 \\
&\quad + 3\sigma_2^2\mu_1^3\mu_2 + 18\mu_0\sigma_0^2\sigma_2^2 + 9\mu_1^2\sigma_0^2\sigma_2^2 + 9\mu_2^2\sigma_0^2\sigma_1^2 \\
&\quad + 9\mu_1^2\sigma_0^2\sigma_2^4 + 15\sigma_0^2\sigma_1^2\sigma_2^2 + 9\mu_1\mu_2\sigma_1^2\sigma_2^2, \\
E(X_1^4 X_2^2) &= \frac{\partial^6}{\partial t_1^4 \partial t_2^2} \phi(t_1, t_2) \Big|_{(t_1, t_2)=(0,0)} \\
&= 6\mu_0^2 + 8\mu_0\mu_1 + 2\mu_0\mu_2 + 4\mu_0\mu_1^3 + 4\mu_1^2\mu_2^2 + \mu_1^4\mu_2^2 + 12\mu_0\mu_1^2\mu_2 \\
&\quad + \sigma_0^2 + 8\sigma_0^4 + 6\sigma_0^2\sigma_1^2 + 12\sigma_0^4\sigma_1^2 + \sigma_1^2\sigma_2^2 + 3\sigma_1^4\sigma_2^2 \\
&\quad + 6\sigma_0^2\mu_1^2 + 12\sigma_0^4\mu_1^2 + 4\mu_1^2\sigma_2^2 + \mu_1^4\sigma_2^2 + \sigma_1^2\mu_2^2 + 3\mu_2^2\sigma_1^4 \\
&\quad + 24\mu_0\mu_1\sigma_0^2 + 12\mu_0\mu_1\sigma_1^2 + 12\mu_0\mu_2\sigma_1^2 + 16\mu_1\mu_2\sigma_0^2 \\
&\quad + 8\mu_1^3\mu_2\sigma_0^2 + 9\mu_1^2\mu_2^2\sigma_1^2 + 9\mu_1^2\sigma_1^2\sigma_2^2 + 24\mu_1\mu_2\sigma_0^2\sigma_1^2,
\end{aligned}$$

et

$$\begin{aligned}
E(X_1^2 X_2^4) &= \frac{\partial^6}{\partial t_1^2 \partial t_2^4} \phi(t_1, t_2) \Big|_{(t_1, t_2)=(0,0)} \\
&= 6\mu_0^2 + 8\mu_0\mu_2 + 2\mu_0\mu_1 + 4\mu_0\mu_2^3 + 4\mu_1^2\mu_2^2 + \mu_1^2\mu_2^4 + 12\mu_0\mu_1\mu_2^2 \\
&\quad + \sigma_0^2 + 8\sigma_0^4 + 6\sigma_0^2\sigma_2^2 + 12\sigma_0^4\sigma_2^2 + \sigma_1^2\sigma_2^2 + 3\sigma_1^2\sigma_2^4 \\
&\quad + 6\sigma_0^2\mu_2^2 + 12\sigma_0^4\mu_2^2 + 4\mu_2^2\sigma_1^2 + \mu_2^4\sigma_1^2 + \sigma_2^2\mu_1^2 + 3\mu_1^2\sigma_2^4 \\
&\quad + 24\mu_0\mu_2\sigma_0^2 + 12\mu_0\mu_2\sigma_2^2 + 12\mu_0\mu_1\sigma_2^2 + 16\mu_1\mu_2\sigma_0^2 \\
&\quad + 8\mu_1\mu_2^3\sigma_0^2 + 9\mu_1^2\mu_2^2\sigma_2^2 + 9\mu_2^2\sigma_1^2\sigma_2^2 + 24\mu_1\mu_2\sigma_0^2\sigma_2^2.
\end{aligned}$$

Bibliographie

- 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.
- Bulla, J., Chesneau, C. & Kachour, M. (2012). On the bivariate Skellam distribution. <http://hal.archives-ouvertes.fr/hal-00744355>
- Genest, C. & Mesfioui, M. (2014). Bivariate extensions of Skellam's distribution. *Probability in the Engineering and Informational Sciences* 28 : 401–417.
- Hwang, Y., Kim, J. & Kweon, I. (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.
- 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.
- Karlis, D. & Ntzoufras, I. (2003). Analysis of sports data by using bivariate Poisson models. *The Statistician* 52 : 381–393.
- Karlis, D. & Ntzoufras, I. (2006). Bayesian analysis of the differences of count data. *Statistics in Medicine* 25 : 1885–1905.
- Prékopa, A. (1952). On composed Poisson distributions. IV. Remarks on the theory of differential processes. *Acta Mathematica Academiae Scientiarum Hungaricae* 3 : 317–325.
- Romani, J. (1956). Distribución de la suma algebraica de variables de Poisson. *Trabajos de Estadística* 7 : 175–181.
- 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.
- Strackee, J. & van der Gon, J.J.D. (1962). The frequency distribution of the difference between two Poisson variates. *Statistica Neerlandica* 16 : 17–23.

Wilks, S.S. (1938). The large-sample distribution of the likelihood ratio for testing composite hypotheses. *The Annals of Mathematical Statistics* 9 : 60–62.