Comparison between two bivariate Poisson distributions through the phi-divergence.

Prévot C. Batsindila Nganga[1], Rufin Bidounga[1], Dominique Mizère[1,*] and Célestin C. Kokonendji[2,*]

[1]Marien Ngouabi University, BP 69, Brazzaville, Congo.
[2]Laboratoire de Mathématiques de Besançon, University Bourgogne Franche-Comté, Besançon, France.

Abstract. In the literature, there are two probabilistic models of bivariate Poisson : the model according to Holgate and the model according to Berkhout and Plug. These two models express themselves by their probability mass function. The model of Holgate puts in evidence a strictly positive correlation, which is not always realistic. To remedy this problem, Berkhout and Plug proposed a bivariate Poisson distribution accepting the correlation as well negative, equal to zero, that positive. In this paper, we show that these models are nearly everywhere asymptotically equal. From this survey that the ϕ-divergence converges toward zero, both models are therefore nearly everywhere equal. Also, the model of Holgate converges toward the one of Berkhout and Plug. Some graphs will be presented for illustrating this comparison.

*Corresponding author : Dominique Mizère, domizere@gmail.com
Rufin Budounga, rufbid@yahoo.fr
Célestin C. Kokonendji, celestin.kokonendji@univ-fcomte.fr

Résumé. Dans la littérature, il y a deux modèles probabilistes de Poisson bivariés: le modèle selon Holgate puis le modèle selon Berkhout et Plug. Ces deux modèles sont définis par leur fonction de masse. Le modèle de Holgate met en évidence une corrélation strictement positive, ce qui n'est pas toujours réaliste. Pour remédier à ce problème, Berkhout et Plug ont proposé une distribution de Poisson bivariée acceptant la corrélation aussi bien négative, nulle que positive. Dans ce papier, nous montrons que ces modèles sont presque partout asymptotiquement égaux car la $\phi-$divergence de ces deux modèles converge vers zéro. Aussi, le modèle de Holgate converge vers celui de Berkhout et Plug. Les graphes seront présentés pour illustrer cette comparaison.

1. Introduction

The bivariate Poisson distribution has been introduced by Campbell (1934) who considered the limit of the distribution of a contingency table with two dimensions. Practically at the same time, Guldberg (1934) obtains the independent binomial distributions. The explicit form of the distribution is due a few years later in 1944 (Morin (1993)) to Aitken (1994). It is however necessary to await Holgate (1934) in 1964 to obtain a bivariate variables starting from three univariate variables of Poisson independent, that is with a matrix of variance-covariance not diagonal. Contrary to what is said in the paper of Elion *et al.* (2016), one cannot assign this find to Johnson and Kotz (1969) whose research are belated. Besides, the literature informs that several bivariate Poisson distributions have been put in evidence. The applied statistics makes use of the bivariate Poisson distribution more according to Holgate (1934) and a lot of scientist works (Kawamura (1973), Kocherlakota *et al.* (1992)) rest on this distribution. It permits to have a log-likelihood function with a difficult expression to manipulate of which the maximum of likelihood is estimated by means of the Hessian matrix. But, outside of the difficult matrix manipulations that it requires to estimate the parameters, the model of Holgate puts in evidence a strictly positive correlation. To remedy this problem, Berkhout *et al.* (2004) proposed a bivariate Poisson distribution accepting the correlation as well negative, equal to zero, that positive. Otherwise, in their paper, Berkhout and Plug reviewed the model of Holgate without comparing it to their model. (Elion *et al.* (2016)) passed in magazine the two models without saying a word on the relation that exists between the two. We wondered if these two models are identical or not. The model of Holgate (1934) and the one of Berkhout and Plug express themselves by their probability mass function (*pmf*).

Cuenin *et al.* (2016) proposed also a model of bivariate Poisson based on the distribution of Tweedie of which one knows that the *pmf* is not easy to calculate. In the literature, we raised an instrument permitting to compare two densities of probabilities to know: the ϕ-divergence (Toma (2009)).

We show, in this work, that the ϕ-divergence from model of Berkhout *et al.* (2004) to model of Holgate (1934) converge toward zero: therefore the model of Hogate is nearly everywhere asymtotically equal to the the model of Berkhout and Plug. We also shows that the model of Berkhout *et al.* (2004) is the asymptotic distribution of the model of Holgate (1934). Some graphs will be presented at the end of this paper to illustrate the comparison of these two models.

2. Generality

2.1. The bivariate Poisson distribution according to Holgate.

Let us consider three independent variables V_1, V_2 and U which follow the univariate Poisson distribution of respective parameters λ_1, λ_2 and λ_3. With these three variables, one builds two new variables dependent Y_1 and Y_2 such as:

$$Y_j = V_j + U, \quad j = 1, 2. \tag{1}$$

Then the *pmf* of the bivariate Poisson distribution is written:

$$\mathbb{P}\left(Y_1 = y_1, Y_2 = y_2\right) = e^{-\lambda_1 - \lambda_2 - \lambda_3} . \sum_{l=0}^{min(y_1, y_2)} \frac{\lambda_3^l}{l!} . \frac{\lambda_1^{y_1 - l}}{(y_1 - l)!} . \frac{\lambda_2^{y_2 - l}}{(y_2 - l)!}.$$

The disribution of bivariate Poisson according to Holgate will be noted by

$$f_H\left(y_1, y_2; \lambda_1, \lambda_2, \lambda_3\right).$$

We have the following the following class of properties.

Class of Property (I).

(a) The covariance between Y_1 and Y_2 gives:

$$cov\left(Y_1, Y_2\right) = cov\left(V_1 + U, V_2 + U\right) = \lambda_3.$$

(b) The correlation between Y_1 and Y_2 is equal to:

$$cor\left(Y_1, Y_2\right) = \frac{\lambda_3}{\sqrt{\left(\lambda_1 + \lambda_3\right)\left(\lambda_2 + \lambda_3\right)}}.$$

(c) The marginal distributions follow:

$$\mathbb{P}\left(Y_1 = y_1\right) = e^{-\lambda_1 - \lambda_3} \frac{\left(\lambda_1 + \lambda_3\right)^{y_1}}{y_1!} \quad \text{and} \quad \mathbb{P}\left(Y_2 = y_2\right) = e^{-\lambda_2 - \lambda_3} \frac{\left(\lambda_2 + \lambda_3\right)^{y_2}}{y_2!}.$$

(d) The conditional distribution is equal to:

$$\mathbb{P}\left(Y_1 = y_1 | Y_2 = y_2\right) = \mathbb{P}\left(Y_1 = y_1, Y_2 = y_2\right) \times \left[\frac{\left(\lambda_2 + \lambda_3\right)^{y_2}}{y_2!} e^{-\lambda_2 - \lambda_3}\right]^{-1}$$

$$= e^{-\lambda_1} \times \sum_{l=0}^{min(y_1, y_2)} \binom{l}{y_2} \left(\frac{\lambda_3}{\lambda_3 + \lambda_2}\right)^l \left(\frac{\lambda_2}{\lambda_3 + \lambda_2}\right)^{y_2 - l} \frac{\lambda_2^{y_1 - l}}{(y_1 - l)!},$$

and has characteristics:

$$\mathbb{E}\left(Y_1|Y_2\right) = \lambda_1 + y_2 . \frac{\lambda_3}{\lambda_3 + \lambda_2} var\left(Y_1|Y_2\right)$$

$$= \lambda_1 + \frac{y_2 \lambda_2 \lambda_3}{\left(\lambda_3 + \lambda_2\right)^2}.$$

To finish, let us retain that the conditional law of Y_1 knowing Y_2 is the convolution of the variable of Poisson U of parameter λ_1 and of the variable binomial V of parameters y_2 and $\lambda_3 / \left(\lambda_3 + \lambda_2\right)$.

2.2. The bivariate Poisson distribution according to Berkhout and Plug.

Let us consider a couple of dependent positive integer whole random variables (Y_1, Y_2). According to theory of conditional probabilities, their joined density can be written like the product of a distribution marginal and of a conditional distribution, such as:

$$\mathbb{P}\left(Y_1 = y_1, Y_2 = y_2\right) = f\left(y_1, y_2\right) = g_1\left(y_1\right) . g_2\left(y_2|y_1\right)$$
$$\mathbb{P}\left(Y_1 = y_1, Y_2 = y_2\right) = f\left(y_1, y_2\right) = g_1^*\left(y_1\right) . g_2^*\left(y_1|y_2\right)$$

This decomposition of the joined density can take several forms compared to permutations of Y_1 and Y_2. In this bivariate case, we have two permutations:

That is to say π an indicator of permutation taking value 0 and 1; z_1 and z_2 can be written like permutations realized by y_1 and y_2, such as:

$$z_1 = (1 - \pi) y_1 + \pi y_2 \text{ and } z_2 = \pi y_1 + (1 - \pi) y_2$$

où $\pi \in \{0, 1\}$.
Let us suppose g_1^π and g_2^π two Poisson distribution, for any joined distribution f^π. For the choice of the permutation, the existing data enable us to fix the density $f^\pi\left(z_1, z_2\right)$. Then the joined density is written:

$$f\left(y_1, y_2\right) = f^\pi\left(z_1, z_2\right) = g_1^\pi\left(z_1\right) . g_2^\pi\left(z_2|z_1\right) \quad \text{for} \quad \pi \in \{0, 1\}.$$

When the marginal one of z_1 is equal to:

$$g_1^\pi\left(z_1\right) = \frac{\mu_1^{z_1}}{z_1!} e^{-\mu_1}, \tag{2}$$

and the conditional distribution $z_2|z_1$ which:

$$g_2^\pi (z_2|z_1) = \frac{\mu_2^{z_2}}{z_2!} e^{-\mu_2} \tag{3}$$

where $\mu_1 = exp(x'\beta_1)$ and $\mu_2 = exp(x'\beta_2 + \eta z_1)$, with β_1, β_2 and η are parameters of g_1^π and g_2^π and $x' = (x_1, x_2, ..., x_p)$ the vector of the explanatory variables or factors. Then joined density of the permutations z^π:

$$f^\pi (z_1, z_2) = \frac{e^{z_1(x'\beta_1) - \exp(x'\beta_1) + z_2(x'\beta_2 + \eta z_1) - \exp(x'\beta_2 + \eta z_1)}}{z_1! z_2!}$$

is called *pmf* of the bivariate Poisson distribution according to Berkhout and Plug, and will be denoted

$$f_{BP}(z_1, z_2; \mu_1, \mu_2).$$

Likely, it is difficult to handle such an expression is a direct way, but that is possible to see that by using the factorial moment $(r, s)^{th}$ of the joint distribution, we get this second class of properties.

Class of Property (II).

(a) The factorial moment (r, s) of the joined distribution can be written:

$$\mathbb{E}[Z_1(Z_1 - 1)\cdots(Z_1 - r + 1)Z_2(Z_2 - 1)\cdots(Z_2 - s + 1)] \tag{4}$$
$$= \sum_{z_1=r}^{+\infty} \sum_{z_2=s}^{+\infty} \frac{e^{z_1(x'\beta_1) - \exp(x'\beta_1) + z_2(x'\beta_2 + \eta z_1) - \exp(x'\beta_2 + \eta z_1)}}{z_1! z_2!}$$
$$= e^{-\exp(x'\beta_1) + \exp(x'\beta_1 + s\eta) + rx'\beta_1 + sx'\beta_2 + rs\eta}$$
$$= \mu_1 e^{\mu_1(\exp(s\eta) - 1) + sx'\beta_2 + rs\eta} \tag{5}$$

(b) From the expression (5), we can now calculate every moment of the distribution joined just by the choice of r and s, one has what follows:
(b1) For $s = 1$ and $r = 0$,

$$\mathbb{E}(Z_2) = \exp[x'\beta_2 + \mu_1(\exp(\eta) - 1)].$$

(b2) For $s = 2$ and $r = 0$, one has for the variable Z_2:

$$\mathbb{E}(Z_2^2) = \mathbb{E}(Z_2) + e^{2x'\beta_2 + \mu_1(\exp(2\eta) - 1)},$$
$$var(Z_2) = \mathbb{E}(Z_2) + (\mathbb{E}(Z_2))^2 \left(e^{\mu_1(\exp(\eta) - 1)^2} - 1\right).$$

We note that for all $\eta \neq 0$, $var(Z_2) - \mathbb{E}(Z_2) > 0$, which implies that Z_2 is overdispersed (Castillo *et al.* (1998)).

(c) Let us pose $s = 0$ and $r = 1$, $r = 2$, $r = 0$ successively, we have for the Z_1 variable:

$$\mathbb{E}(Z_1) = \mu_1 \quad , var(Z_1) = \mu_1, \quad \mathbb{E}(Z_1 Z_2) = \mu_1 e^{\mu_1(e^\eta - 1) + x'\beta_2 + \eta}.$$

We have

$$\mathbb{E}(Z_1) \times \mathbb{E}(Z_2) = \mu_1 e^{\mu_1(e^\eta - 1) + x'\beta_2},$$

and then

$$cov(Z_1, Z_2) = \mathbb{E}(Z_1 Z_2) - \mathbb{E}(Z_1)\mathbb{E}(Z_2) = \mu_1 \mathbb{E}(Z_2)(\exp(\eta) - 1).$$

(d) Thus the correlation, between Z_1 and Z_2, can be obtained as follows

$$cor(Z_1, Z_2) = \frac{\mu_1 \mathbb{E}(Z_2)(\exp(\eta) - 1)}{\sqrt{\mu_1 [\mathbb{E}(Z_2) + \mathbb{E}^2(Z_2)](e^{\mu_1(\exp(\eta) - 1)} - 1)}}.$$

We note that the correlation is positive (resp. negative) if $\eta > 0$ (resp. $\eta < 0$). However for $\eta = 0$, we have

$$\mathbb{E}(Z_2) = var(Z_2) = \mu_2$$

and $cov(Z_1, Z_2) = 0$. Then, to respect with the interdependence of the variables Z_1 and Z_2, it is necessary that $\eta \neq 0$. $\diamond$

Remark 1. We will suppose in all that follows that the permutation is equal to $\pi = 0$, which provides that $z_1 = y_1$ and $z_2 = y_2$.

3. Comparison of the densities of probabilities f_H and f_{BP}.

We are going to show here that the densities f_H and f_{BP} are nearly everywhere equal asymptotycally.

3.1. functional relation between f_H and f_{BP}.

Proposition 1. *While taking $\mu_1 = \lambda_1 + \lambda_3$ and $\mu_2 = \lambda_2 + \lambda_3$ the pmf f_H decompose himself in a product of factors:*

$$f_H(y_1, y_2; \mu_1, \mu_2, \lambda_3) = \left(\frac{\mu_1^{y_1}}{(y_1)!} e^{-\mu_1}\right)\left(\frac{\mu_2^{y_2}}{(y_2)!} e^{-\mu_2}\right) \times b(y_1, y_2; \mu_1, \mu_2, \lambda_3) \tag{6}$$

with

$$b(y_1, y_2; \mu_1, \mu_2, \lambda_3) = e^{\lambda_3}\left(1 - \frac{\lambda_3}{\mu_1}\right)^{y_1}\left(1 - \frac{\lambda_3}{\mu_2}\right)^{y_2} \sum_{l=0}^{min(y_1, y_2)} (-y_1)_l (-y_2)_l \frac{z^l}{l!} \tag{7}$$

and

$$z = \frac{\lambda_3}{(\mu_1 - \lambda_3)(\mu_2 - \lambda_3)}. \tag{8}$$

Proof of Proposition 1.
Let us put $\mu_1 = \lambda_1 + \lambda_3$, $\mu_2 = \lambda_2 + \lambda_3$, we have:

$$f_H(y_1, y_2; \mu_1, \mu_2, \lambda_3) = e^{-\mu_1 - \mu_2 + \lambda_3} \sum_{l=0}^{min(y_1,y_2)} \frac{(\mu_1 - \lambda_3)^{y_1 - l}}{(y_1 - l)!} \frac{(\mu_2 - \lambda_3)^{y_2 - l}}{(y_2 - l)!} \frac{\lambda_3^l}{l!}$$

$$= e^{-\mu_1 - \mu_2} e^{\lambda_3} \sum_{l=0}^{min(y_1,y_2)} \frac{(\mu_1 - \lambda_3)^{y_1}}{(\mu_1 - \lambda_3)^l} \frac{(\mu_2 - \lambda_3)^{y_2}}{(\mu_2 - \lambda_3)^l} \frac{\lambda_3^l}{l!} \frac{1}{(y_1 - l)!(y_2 - l)!}$$

$$= \frac{\mu_1^{y_1}}{(y_1)!} \frac{\mu_2^{y_2}}{(y_2)!} e^{-\mu_1 - \mu_2} \frac{(y_1)!}{\mu_1^{y_1}} \frac{(y_2)!}{\mu_2^{y_2}} e^{\lambda_3} \sum_{l=0}^{min(y_1,y_2)} \frac{(\mu_1 - \lambda_3)^{y_1} (\mu_1 - \lambda_3)^{y_2}}{(y_1 - l)!(y_2 - l)!} \frac{z^l}{l!},$$

with

$$z = \frac{\lambda_3}{(\mu_1 - \lambda_3)(\mu_2 - \lambda_3)}.$$

We have:

$$f_H(y_1, y_2; \mu_1, \mu_2, \lambda_3) = \frac{\mu_1^{y_1}}{(y_1)!} \frac{\mu_2^{y_2}}{(y_2)!} e^{-\mu_1 - \mu_2} e^{\lambda_3} \sum_{l=0}^{min(y_1,y_2)} \left(1 - \frac{\lambda_3}{\mu_1}\right)^{y_1} \left(1 - \frac{\lambda_3}{\mu_2}\right)^{y_2} \frac{(-1)^l y_1}{(y_1 - l)!} \frac{(-1)^l y_2}{(y_2 - l)!} \frac{z^l}{l!}$$

$$f_H(y_1, y_2; \mu_1, \mu_2, \lambda_3) = \left(e^{-\mu_1} \frac{\mu_1^{y_1}}{y_1!}\right) \left(e^{-\mu_2} \frac{\mu_2^{y_2}}{y_2!}\right) \left(1 - \frac{\lambda_3}{\mu_1}\right)^{y_1} \left(1 - \frac{\lambda_3}{\mu_2}\right)^{y_2} e^{\lambda_3} \sum_{l=0}^{min(y_1,y_2)} (-y_1)_l (-y_2)_l \frac{z^l}{l!}.$$

While putting

$$b(y_1, y_2; \mu_1, \mu_2, \lambda_3) = \left(1 - \frac{\lambda_3}{\mu_1}\right)^{y_1} \left(1 - \frac{\lambda_3}{\mu_2}\right)^{y_2} e^{\lambda_3} \sum_{l=0}^{min(y_1,y_2)} (-y_1)_l (-y_2)_l \frac{z^l}{l!},$$

with $(-y_1)_l$ represent Pochhammer's symbol (Johnson *et al.* (1993)). The proof is finished. $\square$

The following results are easily derived for previous ones.

Corollary 1. *In the expression (6), while taking (see expressions (2) and (3)):*

$$\frac{\mu_1^{y_1}}{(y_1)!} e^{-\mu_1} = \mathbb{P}[Y_1 = y_1],$$

the density marginal of Y_1 with $\ln \mu_1 = x'\beta_1$ and

$$\frac{\mu_2^{y_2}}{(y_2)!} e^{-\mu_2} = \mathbb{P}[Y_2 = y_2/Y_1 = y_1],$$

the conditional distribution of Y_2 when one considers $Y_1 = y_1$, with $\ln \mu_2 = x'\beta_2 + \eta y_1$.

Indeed, we have

$$\mathbb{P}[Y_1 = y_1, Y_2 = y_2] = \mathbb{P}[Y_1 = y_1]\,\mathbb{P}[Y_2 = y_2/Y_1 = y_1] = f_{BP}(y_1, y_2; \mu_1, \mu_2),$$

and therefore:

$$f_H(y_1, y_2; \mu_1, \mu_2, \lambda_3) = f_{BP}(y_1, y_2; \mu_1, \mu_2) \times b(y_1, y_2; \mu_1, \mu_2, \lambda_3). \tag{9}$$

Proposition 2. *Let us consider two generalized linear models* $\ln \mu_1 = x'\beta_1$ *and* $\ln \mu_2 = x'\beta_2 + \eta y_1.$ *The response variables* Y_1 *and* Y_2 *which follows the densities of Poisson of parameters* μ_1 *and* $\mu_2.$
Let us put $\lambda_3 = \frac{1}{n}$ *with* $n \in \mathbb{N}^*$; *we construct a family of pmf* $\{f_{H,n}/n \in \mathbb{N}^*\}$ *of Holgate, with* $f_{H,n}(y_1, y_2; \mu_1, \mu_2) = f_H\left(y_1, y_2; \mu_1, \mu_2, \frac{1}{n}\right).$

Then we have the following result:

$$\lim_{n \longrightarrow +\infty} b\left(y_1, y_2; \mu_1, \mu_2, \frac{1}{n}\right) = 1. \tag{10}$$

Proof of Proposition 2.

According to the expressions (7) and (8), we have:

$$b\left(y_1, y_2; \mu_1, \mu_2, \frac{1}{n}\right) = e^{\frac{1}{n}}\left(1 - \frac{\frac{1}{n}}{\mu_1}\right)^{y_1}\left(1 - \frac{\frac{1}{n}}{\mu_2}\right)^{y_2} \sum_{l=0}^{min(y_1,y_2)} (-y_1)_l\,(-y_2)_l\,\frac{z^l}{l!}$$

with

$$z = \frac{\frac{1}{n}}{\left(\mu_1 - \frac{1}{n}\right)\left(\mu_2 - \frac{1}{n}\right)}.$$

Let us $z_n = z$. We have

$$b\left(y_1, y_2; \mu_1, \mu_2, \frac{1}{n}\right) = e^{\frac{1}{n}}\left(1 - \frac{\frac{1}{n}}{\mu_1}\right)^{y_1}\left(1 - \frac{\frac{1}{n}}{\mu_2}\right)^{y_2}\left[(-y_1)_0\,(-y_2)_0 + \sum_{l=1}^{min(y_1,y_2)} (-y_1)_l\,(-y_2)_l\,\frac{(z_n)^l}{l!}\right]$$

$$= e^{\frac{1}{n}}\left(1 - \frac{\frac{1}{n}}{\mu_1}\right)^{y_1}\left(1 - \frac{\frac{1}{n}}{\mu_2}\right)^{y_2}\left[1 + \sum_{l=1}^{min(y_1,y_2)} (-y_1)_l\,(-y_2)_l\,\frac{(z_n)^l}{l!}\right]$$

as $\lim_{n \longrightarrow +\infty} z_n = 0$, then

$$\lim_{n \longrightarrow +\infty} b\left(y_1, y_2; \mu_1, \mu_2, \frac{1}{n}\right) = 1.$$

The proof is complete. $\square$

The two following corollaries are easy to derive from the above results.

Corollary 2. *While taking into account the expressions (9) and (10), we have the result*

$$\lim_{n \longrightarrow +\infty} f_{H,n}(y_1, y_2; \mu_1, \mu_2) = f_{BP}(y_1, y_2; \mu_1, \mu_2).$$

Corollary 3. *Let us consider $F_{H,n}$ and F_{BP} the cumulative distributions functions associated to $f_{H,n}$ and f_{BP}. We have the following result.*

$$\forall (x_1, x_2) \in \mathbb{R}^2, \lim_{n \longrightarrow +\infty} F_{H,n}(x_1, x_2) = F_{BP}(x_1, x_2).$$

3.2. ϕ-divergence from f_{BP} to f_H

We have

Proposition 3. *For the distances of the Kullback-Leibler, χ^2, Hellinger and the variational distance (L^1-distance) not belonging to the class introduced by Cressie and Read (Csiszár (1967)),*

$$\lim_{n \longrightarrow +\infty} D_\phi(f_{H,n}, f_{BP}) = 0.$$

Therefore, the pmf $f_{H,n}$ and f_{BP} are nearly everywhere equal asymptotically (Csiszár (1967)).

Proof Proposition 3.

By definition, while considering the distance associated to the divergence of Kullback-Leibler (Csiszár (1967))),

$$
\begin{aligned}
D_\phi(f_{H,n}, f_{BP}) &= \sum_{y_1=0}^{+\infty} \sum_{y_2=0}^{+\infty} f_{H,n} \ln\left(\frac{f_{H,n}}{f_{BP}}\right) \\
&= \sum_{y_1, y_2} \left(\frac{\mu_1^{y_1}}{(y_1)!} e^{-\mu_1}\right) \left(\frac{\mu_2^{y_2}}{(y_2)!} e^{-\mu_2}\right) \times b\left(y_1, y_2; \mu_1, \mu_2, \frac{1}{n}\right) \times \\
&\quad \times \ln b\left(y_1, y_2; \mu_1, \mu_2, \frac{1}{n}\right).
\end{aligned}
$$

Knowing that $\lim_{n \longrightarrow +\infty} b\left(y_1, y_2; \mu_1, \mu_2, \frac{1}{n}\right) \times \ln b\left(y_1, y_2; \mu_1, \mu_2, \frac{1}{n}\right) = 0$. We are assured of the answer.

Besides, while considering the distance of the χ^2 (Csiszár (1967)),

$$D_\phi(f_{H,n}, f_{BP}) = \sum_{y_1=0}^{+\infty} \sum_{y_2=0}^{+\infty} \left(\frac{f_{H,n}}{f_{BP}} - 1\right)^2 f_{BP}.$$

However, we have

$$\lim_{n \longrightarrow +\infty} \frac{f_{H,n}}{f_{BP}} = \lim_{n \longrightarrow +\infty} b\left(y_1, y_2; \mu_1, \mu_2, \frac{1}{n}\right) = 1,$$

and therefore

$$\lim_{n \longrightarrow +\infty} D_\phi\left(f_{H,n}, f_{BP}\right) = 0.$$

Otherwise, while considering the distance of Hellinger (Csiszár (1967)), we have

$$D_\phi\left(f_{H,n}, f_{BP}\right) = \sum_{y_1=0}^{+\infty} \sum_{y_2=0}^{+\infty} \left(\sqrt{\frac{f_{H,n}}{f_{BP}}} - 1\right)^2 f_{BP}.$$

Therefore, the result is proved.

Finally, as considering the L^1 distance, we have:

$$D_\phi\left(f_{H,n}, f_{BP}\right) = \sum_{y_1=0}^{+\infty} \sum_{y_2=0}^{+\infty} f_{BP} \left|\frac{f_{H,n}}{f_{BP}} - 1\right|.$$

It follow that:
$$\lim_{n \longrightarrow +\infty} D_\phi\left(f_{H,n}, f_{BP}\right) = 0.$$

While being based on the distance of Kolmogorov-Smirnov, we propose a new divergence that we call divergence of Kolmogorov-Smirnov below:

Definition 1.

$$D_{KS}\left(F_{H,n}, F_{BP}\right) = \sum_{y_{1i}=0}^{+\infty} \sum_{y_{2i}=0}^{+\infty} Sup_{\substack{y_{1i} \leqslant x_1 \\ y_{2i} \leqslant x_2}} \left|F_{H,n}\left(x_1, x_2\right) - F_{BP}\left(x_1, x_2\right)\right|, \quad \forall\left(x_1, x_2\right) \in \mathbb{R}^2.$$

With y_{1i} (y_{2i}) the ith realization of Y_1 (Y_2).

Here is the result based on that divergence measure.

Proposition 4. *We have the following results :*

$$\lim_{n \longrightarrow +\infty} D_{KS}\left(F_{H,n}, F_{BP}\right) = 0$$

and

$$D_{KS}\left(F_{H,n}, F_{BP}\right) = D_{L^1}\left(f_{H,n}, f_{BP}\right).$$

Proof of Proposition 4.

The first point of the proposition is obvious and then, omited.

As to the second, we have, by definition: $\forall (x_1, x_2) \in \mathbb{R}^2$, $F_{H,n}(x_1, x_2) = \sum\limits_{\substack{y_{1i} \leqslant x_1 \\ y_{2i} \leqslant x_2}} f_{H,n}(y_{1i}, y_{2i})$,

$F_{BP}(x_1, x_2) = \sum\limits_{\substack{y_{1i} \leqslant x_1 \\ y_{2i} \leqslant x_2}} f_{BP}(y_{1i}, y_{2i})$,

and:

$$
\begin{aligned}
|F_{H,n}(x_1, x_2) - F_{BP}(x_1, x_2)| &= \left| \sum\limits_{\substack{y_{1i} \leqslant x_1 \\ y_{2i} \leqslant x_2}} f_{H,n}(y_{1i}, y_{2i}) - f_{BP}(y_{1i}, y_{2i}) \right| \\
&\leqslant \sum\limits_{\substack{y_{1i} \leqslant x_1 \\ y_{2i} \leqslant x_2}} |f_{H,n}(y_{1i}, y_{2i}) - f_{BP}(y_{1i}, y_{2i})|,
\end{aligned}
$$

and therefore,

$$
Sup_{\substack{y_{1i} \leqslant x_1 \\ y_{2i} \leqslant x_2}} |F_{H,n}(y_1, y_2) - F_{BP}(y_1, y_2)| = |f_{H,n}(y_{1i}, y_{2i}) - f_{BP}(y_{1i}, y_{2i})|.
$$

Finally, we have

$$
\begin{aligned}
D_{KS}(F_{H,n}, F_{BP}) &= \sum_{y_{1i}=0}^{+\infty} \sum_{y_{2i}=0}^{+\infty} |f_{H,n}(y_{1i}, y_{2i}) - f_{BP}(y_{1i}, y_{2i})| \\
&= \sum_{y_{1i}=0}^{+\infty} \sum_{y_{2i}=0}^{+\infty} f_{BP}(y_{1i}, y_{2i}) \left| \frac{f_{H,n}(y_{1i}, y_{2i})}{f_{BP}(y_{1i}, y_{2i})} - 1 \right| \\
&= D_{L^1}(f_{H,n}, f_{BP}).
\end{aligned}
$$

This closes the proof. $\square$

We present the following graph to illustrate this comparison (see figures 1 and 2).

Concluding remarks

If we note D_H, D_{KL}, D_{χ^2} and D_{L^1} the divergences according to the respective distances of Hellinger, Kullback-Leibler, Khi-2 and L^1; it takes out again of the following figures that $D_H \leqslant D_{KL} \leqslant D_{\chi^2} \leqslant D_{L^1}$. The graphs of the function (or sequences) $n \longmapsto D_\phi(f_{H,n}, f_{BP})$, $\phi \in I$, with $I = \{H, KL, \chi^2, L^1\}$ are monotonic decreasing to the lower bound zero :
$\forall \varepsilon > 0$, $\exists N_\phi(\varepsilon) \in \mathbb{N}$, such as $\forall n \geqslant N_\phi(\varepsilon) \Rightarrow D_\phi(f_{H,n}, f_{BP}) \leqslant \varepsilon$. The integers $N_\phi(\varepsilon)$, $\phi \in I$, inform us on the speed of convergence of the sequences.

For ε (arbitrarily small) given, a graphic resolution allows us to determine the $N_\phi(\varepsilon)$, $\phi \in I$. We found, for $\varepsilon = 10^{-4}$, $N_H(\varepsilon) = 100$, $N_{KL}(\varepsilon) = 150$, $N_{\chi^2}(\varepsilon) = 200$ et $N_{L^1}(\varepsilon) > 2000$. For the smallest integer $N_H(\varepsilon) = 100$, we cannot say that D_H converge very quickly toward zero. On the other hand, one can compare the speeds of convergence as follows:

D_H converge more quickly than D_{KL} toward zero ; D_{KL} converges more quickly than D_{χ^2} toward zero and that D_{χ^2} converges more quickly than D_{L^1} toward zero. The divergences tend to zero. It attests that the two distributions are nearly everywhere equal asymptotically.

The *pmf* of the bivariate Poisson distribution according to Berkhout and Plug is the asymptotic density of the couple (Y_1, Y_2), while the *pmf* of the bivariate distribution according to Holgate is her exact density.

References

Aitken, A.C. (1994). Statistical Mathematics, *Ed. Oliver and Boyd.*

Berkhout, P., and Plug, E. (2004). A bivariate Poisson count data model using conditional probabilities. *Statist. Neerlandica*, 58(3), 349-364.

Cuenin, J., Jørgensen, J., and Kokonendji, C.C. (2016). Simulations of full multivariate Tweedie with flexible dependence structure. *Comput Stat.*, 31, 1477-1492.

Campbell, J.T. (1934). The Poisson correlation function.*Proc. Edinburgh Math. Soc.*, 4, 18-26.

Castillo, J., and Pérez-Castillo, M. (1998). Weighted Poisson distributions for overdispersion and underdispersion situations. *Ann. Inst. Statist. Math.*, 20, 567-585.

Csiszár, I. (1967). On topology properties of f-divergences. *Studia Sci. Math. Hungar.]*, 2, 329-339,

Elion, L.-L., Koukouatikissa, D.M., Bidounga, R., Louzayadio, C.G., Mizère, D., Makany, R., and Kissita, G. (2016). The bivariate weighted Poisson distribution: Statistical study of the arterial hypertension data according to the weight and blood sugar level. *Far East Journal of Theorical Statistics*, 52, 365-393.

Evren, A., and Tuna, E. (2012). On some properties of goodness of measures based on statistical entropy. *IJRRAS* , 13(1).

Guldberg, A.(1934). On discontinuous frequency functions of 2 variables. *Skand. Aktuar.*, 17, 89-117.

Holgate, P. (1934). Estimation for the bivariate Poisson distribution. *Biometrika* , 51, 241-245.

Johnson, N.L., Kotz, S., and Kemp, A.W. (1993). *Univariate discrete distributions*. Second Edition. Wiley Series in Probability and Mathematical Statistics, 17, 307-318.

Johnson, N.L., and Kotz, S. (1969). Distributions in Statistcs: Discrete Distributions. *Johnson Wiley & Sons,* New York.

Kawamura, K. (1973). The structure of bivariate Poisson distribution. *Kodai Math. Sem. J.*, 25, 246-256.

Kocherlakota, S., and K. Kocherlakota, K.(1992). Bivariate Discrete Distributions. *Marcel Dekker,* New York.

Morin, A. (1993). Loi de Poisson multivariée. *Journal de la société statistique de Paris*, Tome 134 (2), 3-13.

Toma, A. (2009). Optimal robust M-estimators using divergences. *Statistics and Probability Letters*, 79, 1-5.
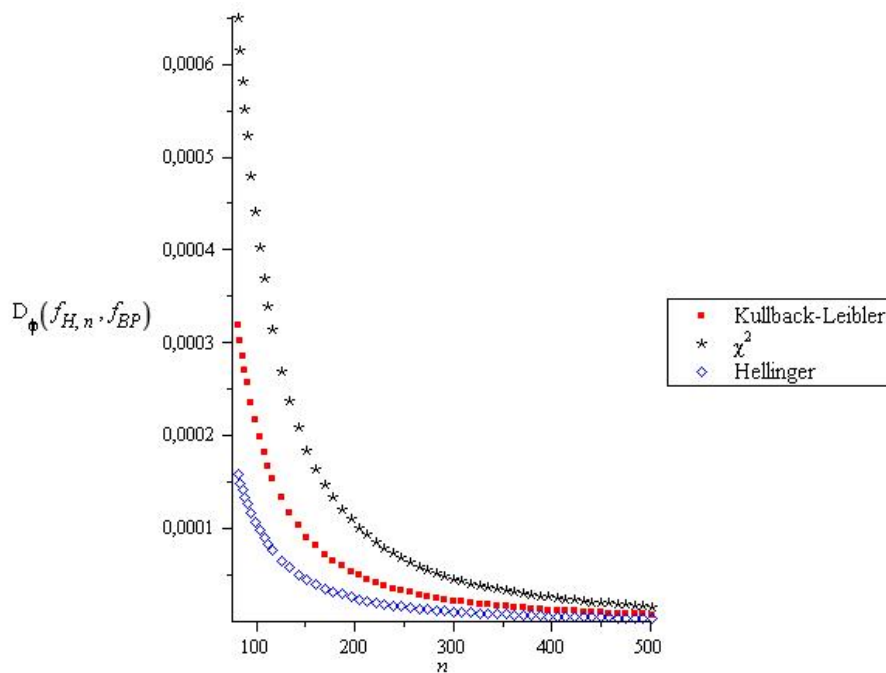
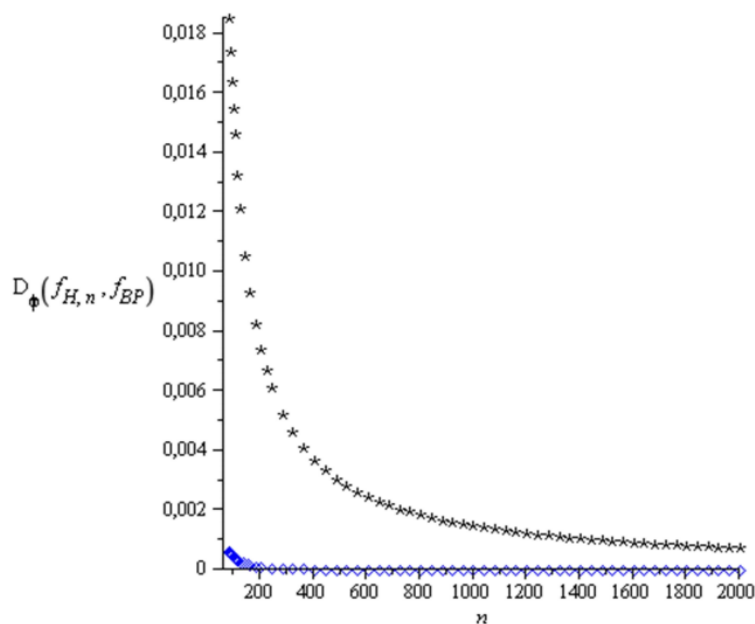**Fig. 1.** $\mu_1 = 0.4$, $\mu_2 = 0.6$, $y_1 = y_2 \in [0, 32]$ and $n \in [80, 500]$

Fig. 2. $\mu_1 = 0.4$, $\mu_2 = 0.6$, $y_1 = y_2 \in [0, 32]$ and $n \in [80, 2000]$