Afrika Statistika
Vol. 8, 2013, pages 583–594.
DOI: http://dx.doi.org/10.4314/afst.v8i1.7

Afrika Statistika
ISSN 2316-090X

Décomposition d'une loi de poisson pondérée en une combinaison convexe de lois duales

Dominique Mizere, Gélin Chedly Louzayadio, Rufin Bidounga and Gabriel Kissita

Université Marien NGOUABI, BP.69, Brazzaville. Congo

Received: December 24, 2012; Accepted: November 15, 2013

Copyright © 2013, Afrika Statistika. All rights reserved

Abstract. The probability distribution of a set of observation is most often defined as a convex combination of probability laws. To highlight this mixture of the laws, MCMC (Monte Carlo Markov Chain) which is an algorithm that generates a stationary Markov chain is often used; laws being considered as normal laws. In this paper, the observations are positive integer, so it is assumed that the mixture law is a Poisson weighted law and Blending laws are dual. The purpose of this work is to determine the dual laws by simple algebraic properties.

Résumé. La loi de distribution d'un ensemble d'observation est le plus souvent définie comme une combinaison convexe de plusieurs lois de probabilités. Pour mettre en évidence ce mélange de lois on utilise souvent la méthode MCMC (Monte Carlo par Chaîne de Markov) qui est un algorithme qui génère une chaîne de Markov stationnaire; les lois étant considérées comme des lois normales. Dans le présent papier, les observations sont entières positives; on suppose donc que la loi mélange est une loi de Poisson pondérée et les lois mélangeantes sont duales . L'objet de ce travail est de pouvoir déterminer ces lois duales par des propriétés algébriques simples.

Key words: Count Data; Exponetial Family; Weighted Poisson Distribution; Fisher Index; Overdispersion; Underdispersion; Dual Distribution; Convex Combination.
AMS 2010 Mathematics Subject Classification : 62F10; 62H30.

*Corresponding author Dominique Mizere: domizere@gmail.com
Gélin Chedly Louzayadio : gelinlouz@gmail.com
Rufin Bidounga : rufbid@yahoo.fr
Gabriel Kissita: gakissita@yahoo.fr

1. Introduction

Le modèle de Poisson est un modèle de repère pour l'analyse statistique des données de dénombrement. Lorsque ce modèle n'est pas approprié, principalement parce que les résultats issus du traitement des données ne permettent pas de le valider via le test d'adéquation du chi-deux; il est courant de chercher les familles de lois alternatives en se référant à l'indice de Fisher ou indice de dispersion de Fisher, à savoir le rapport de la variance sur la moyenne

$$I(Y) = \frac{\sigma^2}{\mu}$$

relativement à 1.

Lorsque $I(Y) = 1,\ I(Y) > 1,\ (I(Y) < 1)$, on dit que la variable Y est équidispersée, surdispersée (sous-dispersée). La loi de Poisson est équidispersée. Les distributions alternatives sont définies par au moins deux paramètres et appartiennent à des classes de lois de Poisson pondérée et de Katz.

Ce concept de loi pondérée est attribué à Fisher (1934) dans une étude de l'effet des méthodes de régulation basées sur les fréquences empiriques. Généralisant ces idées de base de Fisher, Patil et Rao (1978) ont compris l'intérêt d'unifier ces notions ou d'identifier les nombreuses situations où l'échantillonnage peut être modélisé par les lois pondérées. Supposons (Patil, 2002; Johnson *et al.*,1993) que la réalisation y de la variable aléatoire Y de densité de probabilité p soit enregistrée avec une probabilité proportionnelle à une fonction w(y); l'enregistrement y est la réalisation d'une variable aléatoire Y^w appelée version pondérée de Y (Patil, 2002) et qui a pour distribution de probabilité:

$$p_w(y,\theta) = \frac{w(y)}{E_\theta[w(Y)]} p(y,\theta) = P(Y^w = y),\ \forall\ y \in \mathbb{N}$$

avec $p(y,\theta) = \dfrac{\theta^y}{y!} e^{-\theta},\ \forall y \in \mathbb{N},\ \theta \in \mathbb{R}_+^*$, appelée loi de Poisson pondérée où $w(y)$ est une fonction positive, appelée fonction poids et $E_\theta[w(Y)]$ la constante de normalisation qui est la fonction poids moyenne de Y dépendant de θ, le paramètre de Poisson telle que: $0 < E_\theta[w(Y)] < \infty$.

La fonction poids $w(y) = w(y;\phi)$ peut dépendre d'un paramètre ϕ qui représente le mécanisme d'enregistrement des données.

Notons que: $w(y) = w(y;\phi,\theta)$ peut aussi dépendre du paramètre canonique θ.

Le souhait légitime de mettre en évidence d'avantage de lois de probabilité pour modéliser diverses situations nous a conduit également à définir la notion de lois Duales:dualité ponctuelle et dualité en moyenne.

Soit w_1 et w_2 deux fonctions poids strictement positives générant deux lois de Poisson pondérées. Ces deux lois sont dites (ponctuellement) duales si et seulement si leurs fonctions poids vérifient la relation

$$w_1(y) \times w_2(y) = 1, \forall y \in \mathbb{N}. \tag{1}$$

La notion de dualité ponctuelle peut être étendue à la notion de dualité en moyenne comme suite : on remplace $w_i > 0$ par $E_\theta[w_i(Y)] > 0$ et la condition (1) par:

$$E_\theta[w_1(Y)] \times E_\theta[w_2(Y)] = 1, \forall y \in \mathbb{N}$$

avec w_i non constant (Cf. Mizere, 2009).

Le problème que nous allons aborder dans cet article est celui du mélange de lois de probabilité dont la densité f se décompose en une combinaison convexe finie de fonction f_j

$$f(y, \theta) = \sum_{j=1}^{K} \alpha_j f_j(y, \theta_j)$$

avec $\alpha_j > 0 \; \forall j$, les coefficients de convexités et $\sum_{j=1}^{K} \alpha_j = 1$; θ est le paramètre défini dans la population mère et θ_j le paramètre définit dans la sous-population j.

Certains auteurs (cf. Robert *et al.*, 2000; Saint-Pierre, 2003) font l'hypothèse suivante: les lois f et f_j sont normales; ils procèdent alors à l'estimation des paramètres θ et θ_j par la méthode MCMC.

Pour notre part, nous allons supposer que :

1. Les lois f et f_j sont des lois de Poisson pondérées où f est connue et les f_j sont des lois duales inconnues. Pour les déterminer, nous allons utiliser les propriétés algébriques simples
2. $\theta_j = \theta, \forall j$.

Dans la suite de ce travail, nous allons d'abord rappeler les propriétés et définitions sur les lois Duales (Cf. Kokonendji *et al.*, 2008), ensuite nous procéderons à la décomposition de la loi de Poisson pondérée en une combinaison convexe des lois duales. Deux exemples d'illustration seront donnés selon que les lois f_j soient duales ponctuelles ou duales en moyenne.

2. Propriétés

Théorème 1. *(Cf. Mizere, 2009) Soient Y une variable aléatoire de Poisson de moyenne $\theta > 0$ et de fonction poids $w(y) = w(y;\phi)$, $y \in \mathbb{N}$, ne dépendant pas de θ. La moyenne de la fonction poids $\theta \mapsto E_\theta[w(Y)]$ est log-convexe (ou log- concave) si et seulement si la version pondérée Y^w de est surdispersée (ou sous-dispersée).*

Théorème 2. *Soient Y une variable aléatoire de Poisson de paramètre $\theta > 0$ et de fonction poids $w(y) = w(y;\phi)$, $\forall y \in \mathbb{N}$, dépendant du paramètre θ. Alors la version pondérée Y^w de Y est surdispersée (ou sous-dispersée) si la fonction poids $y \mapsto w(y;\phi)$ est log-convexe (ou log-concave).*

Remarque 1. 1. La stricte positivité des w_i dans la défnition 1 implique que les lois de Poisson pondérées correspondantes ont pour support l'ensemble $\mathbb{N}$ tout entier.
2. Etant donnée une loi de Poisson pondérée de fonction poids $w(y)$, sa loi duale existe si et seulement si sa constante de normalisation $E_\theta[\frac{1}{w(y)}]$ est finie.

Théorème 3. *Soit (Y^{w_1}, Y^{w_2}) une paire de distribution ponctuelle duale des variables aléatoires de Poisson pondérée. Si l'une des fonctions w_i $(i = 1, 2)$ est log-convexe (ou log-concave), alors cette paire ponctuellement duale est constituée des variables surdispersées et sous-dispersées.*

Cette définition reformulée par Mizère exclue la distribution de Poisson (cf. Remarque 1) comme une distribution duale en moyenne, car elle est ni surdispersée, ni sous-dispersée. Cette définition implique que la fonction poids ne peut pas être unique, car la valeur moyenne de la fonction $E_\theta[\cdot]$ n'est pas injective.

Théorème 4. *Soit (Y^{w_1}, Y^{w_2}) une paire des variables aléatoires de Poisson pondérée duale en moyenne telles que les fonctions poids $w_1(y) = w_1\ (y;\phi_1)$ et $w_2(y) = w_2(y;\phi_2)$ ne dépendent pas de θ et une des constantes de normalisation $E_\theta[w_i(Y, \phi_i)]$ $(i = 1, 2)$ est log-convexe (ou log-concave). Alors cette paire est constituée des variables surdispersées et sous-dispersées.*

Théorème 5. *Si la fonction poids $w_2(y) = w_2(y; \phi, \theta)$, dépend de θ et si $E_\theta[w_1(Y, \phi_1)]$ est log-convexe (ou log-concave), alors la paire duale en moyenne (Y^{w_1}, Y^{w_2}) est formée des variables surdispersées et sous-dispersées quand la fonction poids $w_2(y)$ est log-convexe (ou log-concave).*

Théorème 6. *Si n fonctions poids $w_1, w_2, \ldots, w_n$ génèrent la même distribution de Poisson pondérée, alors les suites $(E_\theta[w_i(Y)])_i$ et $(w_i)_i$ $(i = 1, \ldots, n)$ sont proportionnelles telles que:*

$$\frac{w_1(y)}{E_\theta[w_1(Y)]} = \frac{w_2(y)}{E_\theta[w_2(Y)]} = \cdots = \frac{w_n(y)}{E_\theta[w_n(Y)]} = \frac{\sum_{i=1}^n \alpha_i w_i(y)}{\sum_{i=1}^n \alpha_i E_\theta[w_i(Y)]}$$

Où les coefficients $\alpha_i \in \mathbb{R}$ sont choisis de tels sorte que les combinaisons linéaires soit toujours positive. Le coefficient de proportionnalité

$$q(y) = \frac{w_1(y)}{E_\theta[w_1(Y)]} = \frac{w_2(y)}{E_\theta[w_2(Y)]} = \cdots = \frac{w_n(y)}{E_\theta[w_n(Y)]} = \frac{\sum_{i=1}^n \alpha_i w_i(y)}{\sum_{i=1}^n \alpha_i E_\theta[w_i(Y)]}$$

est appelé fonction poids normalisée.

Corollaire 1. *Soit $\frac{w_3(y)}{E_\theta[w_3(Y)]}$ une fonction poids normalisée et $E_\theta[w_2(Y)]$ une constante de normalisation. Nous pouvons déterminer la fonction poids $w_2(y)$ qui génère la même distribution de Poisson pondérée que $w_3(y)$ par la quatrième proportionnelle c'est-à- dire*

$$w_2(y) = \alpha \times w_3(y)$$

avec $\alpha = \frac{w_2(y)}{E_\theta[w_3(Y)]} \in \mathbb{R}_+^*$.

Proposition 1. *Soit $w_1(y)$ et $w_2(y)$ deux fonctions poids respectivement de deux distributions duales en moyenne, où $w_2(y)$ est inconnue. Si $E_\theta[w_2(Y)] = \frac{1}{E_\theta[w_1(Y)]} = \varphi(\theta)$ admet une décomposition en série de Taylor au voisinage de θ_0 et de coefficient a_y, alors la loi duale en moyenne de la loi générée par $w_1(y)$ a pour fonction poids:*

$$w_2(y) = a_y \left(\frac{\theta - \theta_0}{\theta}\right)^y y! e^\theta;\ y = 0, 1, 2, \ldots.$$

3. Décomposition d'une loi de Poisson pondérée en une combinaison convexe de deux lois duales

3.1. Dualité ponctuelle

Théorème 7. *Soient $w_0(y,\phi)$, $w_1(y)$ et $w_2(y)$ des fonctions poids des lois de Poisson pondérées telles que $w_0(y,\phi)$ soit connue avec $w_0(y,\phi) > 2$, $\forall\ y \in \mathbb{N}$. Supposons que les lois de Poisson pondérées générées par $w_1(y)$ et $w_2(y)$ soient duales ponctuelles avec $w_0(y,\phi) = w_1(y) + w_2(y)$, $\forall\ y \in \mathbb{N}$. Alors nous avons les résultats suivants:*

$$w_1(y) = \frac{w_0(y,\phi) - \sqrt{w_0^2(y,\phi) - 4}}{2} \tag{2}$$

et

$$w_2(y) = \frac{w_0(y,\phi) + \sqrt{w_0^2(y,\phi) - 4}}{2} \tag{3}$$

Démonstration.

Il suffit de résoudre le système suivant:

$$\begin{cases} w_1(y) + w_2(y) = w_0(y,\phi) \\ w_1(y) \times w_2(y) = 1 \end{cases}$$

qui conduit à une équation du second degré de discriminant $w_0^2(y,\phi) - 4$ strictement positif.

Les fonctions poids $w_1(y) = w_1(y,\phi)$ et $w_2(y) = w_2(y,\phi)$ sont les racines de cette équation et dépendent de ϕ. □

Corollaire 2. *Soient $w_0(y,\phi)$, $w_1(y)$ et $w_2(y)$ des fonctions poids des lois de Poisson pondérées telles que $w_0(y,\phi)$ soit connue, $\forall\ y \in \mathbb{N}$. Supposons que $w_0(y,\phi) > 2$, et que les lois de Poisson pondérées générées par $w_1(y)$ et $w_2(y)$ soient duales ponctuelles avec $w_0(y,\phi) = w_1(y) + w_2(y)$, $y \in \mathbb{N}$. Alors, la loi de Poisson pondérée générée par $w_0(y,\phi)$ se décompose en une combinaison convexe des lois de Poisson pondérées générées par $w_1(y)$ et $w_2(y)$.*

Démonstration.

$$\begin{aligned} p^{w_0}(Y=y) &= \frac{w_0(y,\phi)}{E_\theta[w_0(Y,\phi)]}\frac{\theta^y}{y!}e^{-\theta} \quad ,y \in \mathbb{N} \\ &= \frac{w_1(y)+w_2(y)}{E_\theta[w_0(Y,\phi)]}\frac{\theta^y}{y!}e^{-\theta} \\ &= \frac{E_\theta[w_1(Y)]}{E_\theta[w_0(Y,\phi)]}\frac{w_1(y)}{E_\theta[w_1(Y,\phi)]}\frac{\theta^y}{y!}e^{-\theta} + \frac{E_\theta[w_2(Y)]}{E_\theta[w_0(Y,\phi)]}\frac{w_2(y)}{E_\theta[w_2(Y,\phi)]}\frac{\theta^y}{y!}e^{-\theta} \\ &= \alpha_1 p^{w_1}(Y=y) + \alpha_2 p^{w_2}(Y=y) \end{aligned}$$

avec $\alpha_1 = \dfrac{E_\theta[w_1(Y)]}{E_\theta[w_0(Y,\phi)]} > 0$ et $\alpha_2 = \dfrac{E_\theta[w_2(Y)]}{E_\theta[w_0(Y,\phi)]} > 0$.

Comme $w_0(y,\phi) = w_1(y) + w_2(y)$ $y \in \mathbb{N}$, alors; $E_\theta[w_0(Y,\phi)] = E_\theta[w_1(Y)] + E_\theta[w_2(Y)]$.
On a: $\alpha_1 + \alpha_2 = 1$. □

3.2. Dualité en moyenne

Théorème 8. *Soient $E_\theta[w_0(Y,\phi)]$, $E_\theta[w_1(Y)]$ et $E_\theta[w_2(Y)]$ des fonctions poids moyenne telles que $E_\theta[w_0(Y,\phi)]$ soit connue avec $E_\theta[w_0(Y,\phi)] > 2$. Supposons que les lois de Poisson pondérées générées par $w_1(y)$ et $w_2(y)$ soient duales en moyenne avec $E_\theta[w_0(Y,\phi)] = E_\theta[w_1(Y)] + E_\theta[w_2(Y)]$. Alors nous avons les résultats suivants:*

$$E_\theta[w_1(Y)] = \frac{E_\theta[w_0(Y,\phi)] - \sqrt{(E_\theta[w_0(Y,\phi)])^2 - 4}}{2}$$

et

$$E_\theta[w_2(Y)] = \frac{E_\theta[w_0(Y,\phi)] + \sqrt{(E_\theta[w_0(Y,\phi)])^2 - 4}}{2}$$

Démonstration.

Il suffit de résoudre le système suivant:

$$\begin{cases} E_\theta[w_1(Y)] + E_\theta[w_2(Y)] = E_\theta[w_0(Y,\phi)] \\ E_\theta[w_1(Y)] \times E_\theta[w_0(Y,\phi)] = 1 \end{cases}$$

qui conduit à une équation du second degré et on a le même type des résultats que ceux du Théorème 7. □

4. Décomposition en une combinaison convexe de plusieurs lois duales

4.1. Cas de la dualité ponctuelle

En généralisant les expressions (2) et (3), $\forall\, n \in \mathbb{N}^*$ et $w_{n-1}(y) > 2$, $\forall\, y \in \mathbb{N}$ on a:

$$w_{2n-1}(y) = \frac{w_{n-1}(y,\phi) - \sqrt{w_{n-1}^2(y,\phi) - 4}}{2}$$

et

$$w_{2n}(y) = \frac{w_{n-1}(y,\phi) + \sqrt{w_{n-1}^2(y,\phi) - 4}}{2}$$

qui impliquent

$$w_{2n-1}(y) + w_{2n}(y) = w_{n-1}(y,\phi) \tag{4}$$

et

$$w_{2n-1}(y) \times w_{2n}(y) = 1 \tag{5}$$

Remarque 2. L'expression (5) signifie que les lois de Poisson pondérées générées par les fonctions poids (à indices impair et pair successifs) $w_{2n-1}(y)$ et $w_{2n}(y)$ sont duales ponctuelles.

Théorème 9. *Soient $w_0(y,\phi)$,*

$$w_{2n-1}(y) = \frac{w_{n-1}(y,\phi) - \sqrt{w_{n-1}^2(y,\phi) - 4}}{2}$$

et

$$w_{2n}(y) = \frac{w_{n-1}(y,\phi) + \sqrt{w_{n-1}^2(y,\phi) - 4}}{2}$$

avec $w_{n-1}(y,\phi) > 2,\ y \in \mathbb{N},\ \forall n \in \mathbb{N}^*$ *des fonctions poids des lois de Poisson pondérées telles que* $w_{n-1}(y,\phi)$ *soit connue. Alors*

$$w_0(y,\phi) = \sum_{n=2^{n_0}-1}^{2(2^{n_0}-1)} w_n(y,\phi),\ \ y \in \mathbb{N};\ \ n_0 \in \mathbb{N}^*. \tag{6}$$

Démonstration.

En se donnant plusieurs valeurs de n $(n = 1, 2, 3, \cdots, n_0)$ dans l'expression (4), on obtient un système de plusieurs équations à plusieurs inconnues dont la sommation membre à membre jusqu'à l'étape n_0 donne la relation (6).

Les suites récurrentes définies dans le Théorème 9 permettent de construire l'arborescence de la Figure 1. □

4.2. Cas de la dualité en moyenne

Théorème 10. *Soient* $E_\theta[w_0(Y,\phi)]$

$$E_\theta[w_{2n-1}(Y)] = \frac{E_\theta[w_{n-1}(Y,\phi)] - \sqrt{(E_\theta[w_{n-1}(Y,\phi)])^2 - 4}}{2}$$

et

$$E_\theta[w_{2n}(Y)] = \frac{E_\theta[w_{n-1}(Y,\phi)] + \sqrt{(E_\theta[w_{n-1}(Y,\phi)])^2 - 4}}{2}$$

avec $E_\theta[w_{n-1}(Y,\phi)] > 2$, *des fonctions poids des lois de Poisson pondérées telles que* $E_\theta[w_0(Y,\phi)]$ *soit connue. Alors*

$$E_\theta[w_0(Y,\phi)] = \sum_{n=2^{n_0}}^{2(2^{n_0}-1)} E_\theta[w_n(Y,\phi)],\ \ y \in \mathbb{N};\ \ n_0 \in \mathbb{N}^*.$$

Ce théorème se démontre de la même façon que le Théorème 9.

4.3. Interprétation

Suivant les valeurs de n_0, de l'expression (4) on peut représenter l'arborescence entre les fonctions poids que l'on identifie à un arbre de la division cellulaire: la fonction poids $w_0(y)$ s'identifie à une cellule mère qui se divise en deux cellules filles $w_1(y)$ et $w_2(y)$ qui à leurs tour se divisent chacune en deux cellules et ainsi de suite. Ce modèle biologique est celui de la mitose car les deux cellules filles portent la même information que la cellule mère.

n_0 est le nombre de divisions.

Proposition 2. *Si* $w_0(y) > 2, \forall\ y \in \mathbb{N}$*; la loi de Poisson pondérée générée par* $w_0(y)$ *se décompose en une combinaison convexe de* 2^{n_0} *lois duales deux à deux.*

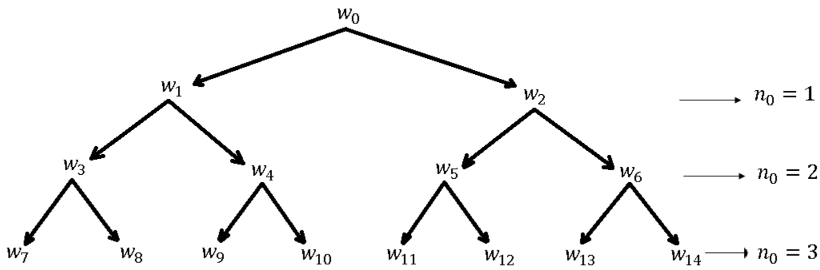


Figure 1. Arborescence de la décomposition d'une fonction poids

Démonstration.

$$\begin{aligned} p^{w_0}(Y=y) &= \frac{w_0(y)}{E_\theta[w_0(Y)]}\frac{\theta^y}{y!}e^{-\theta} \\ &= \sum_{n=2^{n_0}-1}^{2(2^{n_0}-1)} \frac{w_n(y)}{E_\theta[w_0(Y)]}\frac{\theta^y}{y!}e^{-\theta} \\ &= \sum_{n=2^{n_0}-1}^{2(2^{n_0}-1)} \frac{E_\theta[w_n(Y)]}{E_\theta[w_0(Y)]}\frac{w_n(y)}{E_\theta[w_n(Y)]}\frac{\theta^y}{y!}e^{-\theta}; \quad \forall n_0 \in \mathbb{IN}. \end{aligned}$$

En posant:

$$\alpha_n = \frac{E_\theta[w_n(Y)]}{E_\theta[w_0(Y,\phi)]}$$

on a: $\alpha_n > 0$ et $\sum_{n=2^{n_0}-1}^{2(2^{n_0}-1)} \alpha_n = 1$, alors $p^{w_0}(Y=y) = \sum_{n=2^{n_0}-1}^{2(2^{n_0}-1)} \alpha_n p^{w_n}(Y=y)$. □

Remarque 3. 1. L'entier n_0 ne peut "prendre" la valeur infinie sinon l'expression (4) n'aura pas de sens.
2. De l'expression (6), on obtient : $w_0(y,\phi) = w_0(y;\phi,n_0)$, $y \in \mathbb{N}$, $n_0 \in \mathbb{N}^*$; où n_0 est un paramètre latent.

5. Exemples d'illustration

5.1. Cas de la dualité ponctuelle

On considère la famille de lois de Poisson pondérée de Castillo et Pérez-Casany (1998) définie par:

$$p^{w_0}(Y=y) = \frac{(y+a)^r}{\sum_{y=0}^{+\infty}(y+a)^r\frac{\theta^y}{y!}}\frac{\theta^y}{y!}e^{-\theta}, \ y \in \mathbb{N}$$

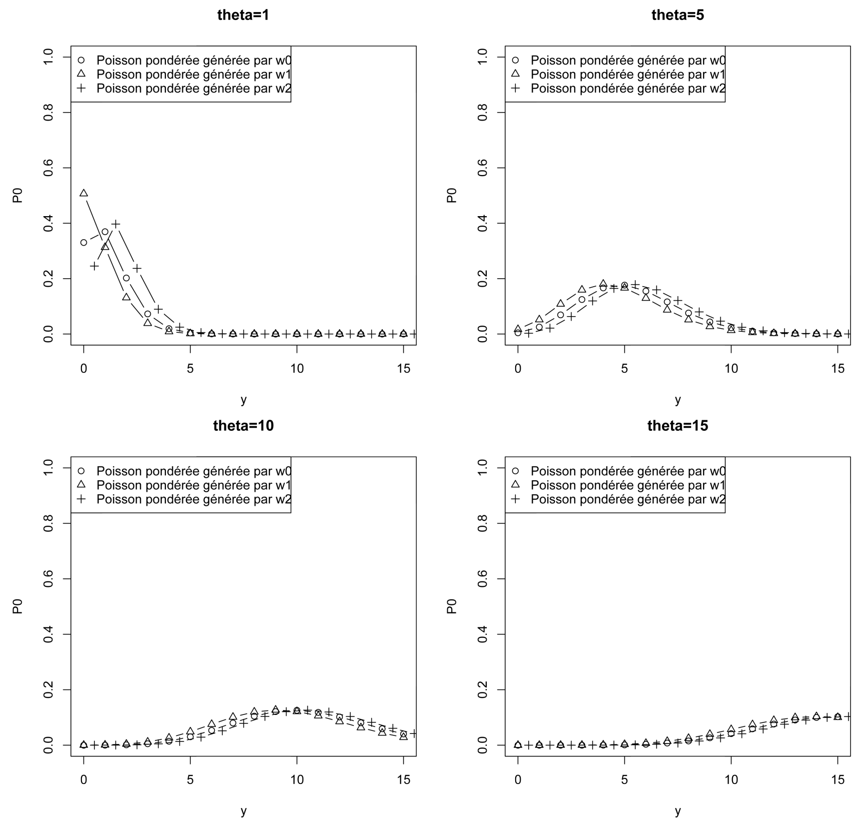
de fonction poids

$$w_0(y, \phi) = (y + a)^r \text{ et de constante de normalisation } E_\theta[w(Y, \phi)] = \sum_{y=0}^{+\infty}(y + a)^r \frac{\theta^y}{y!}e^{-\theta}$$

avec $a > 0$, $r > 0$ et $\phi = (a, r)$.

On prendra $a = 4$ et $r = 0.5$ pour toutes les valeurs de θ [H] On constate que les lois générées

Figure 2. Représentation des polygones de fréquences de la décomposition d'une loi de Poisson pondérée en deux lois duales ponctuelles
.



par w_1 et w_2 ont la même forme que celle générée par w_0 Les coefficients α_2 sont proche de 1, ceci voudrait dire que les courbes de lois générées par w_0 et w_2 sont plus proche.

Tableau 1. coefficients de convexité

θ	α_1	α_2
1	0.3262946	0.6737054
5	0.1353265	0.8646735
10	0.0802712	0.9197287
15	0.0572305	0.9427694

5.2. Cas de la dualité en moyenne

On considère la loi de Poisson biaisée par la taille (Patil, 2002) définie par:

$$p^{w_0}(Y=y) = \frac{y}{\theta}\frac{\theta^y}{y!}e^{-\theta}, \quad y \in \mathbb{N}^*$$

de fonction poids $w_0(y) = y$, $y \in \mathbb{N}^*$ et de constante de normalisation $E_\theta[w_0(Y)] = \theta$, où θ est le paramètre de Poisson.

Cette loi ne peut admettre de loi duale ponctuelle (cf. Remarque 1) car $w_0(0) = 0$. On a

$$E_\theta[w_1(Y)] = \frac{E_\theta[w_0(Y,\phi)] - \sqrt{(E_\theta[w_0(Y,\phi)])^2 - 4}}{2} = \frac{\theta - \sqrt{\theta^2 - 4}}{2}$$

et

$$E_\theta[w_2(Y)] = \frac{E_\theta[w_0(Y,\phi)] + \sqrt{(E_\theta[w_0(Y,\phi)])^2 - 4}}{2} = \frac{\theta + \sqrt{\theta^2 - 4}}{2}.$$

En application de la proposition 1, $E_\theta[w_1(Y)] = \varphi_1(\theta) + \varphi_2(\theta)$ avec $\varphi_1(\theta) = \theta$ et $\varphi_2(\theta) = \sqrt{\theta^2 - 4}$.

Les développements de Taylor au voisinage de 0 de φ_1 et φ_2 sont égaux à

$$\varphi_1(\theta) = \sum_{y=0}^{+\infty} a_y \times \theta^y$$

avec

$$a_y = \begin{cases} 1 & y = 1 \\ 0 & y \neq 1 \end{cases}$$

$$\varphi_2(\theta) = \sum_{y=0}^{+\infty} \binom{\frac{1}{2}}{y} (\theta^2 - 5)^y \text{ pour } 2 < \theta < \sqrt{6}$$

et

$$\binom{\alpha}{y} = \frac{\alpha(\alpha-1)(\alpha-2)\dots(\alpha-y+1)}{y!}; \quad \alpha = \frac{1}{2}, \ y \in \mathbb{N}.$$

Ce qui implique

$$w_1(y) = \frac{1}{2}\left[a_y - \binom{\alpha}{y}(\frac{\theta^2-5}{\theta})^y\right] y!e^\theta$$

et

$$w_2(y) = \frac{1}{2}\left[a_y + \binom{\alpha}{y}(\frac{\theta^2-5}{\theta})^y\right] y!e^\theta.$$

Ce tableau explique la même chose que celui de la dualité ponctuelle.

Figure 3. Représentation de la décomposition d'une loi de Poisson pondérée biaisée par la taille en deux lois duales en moyenne

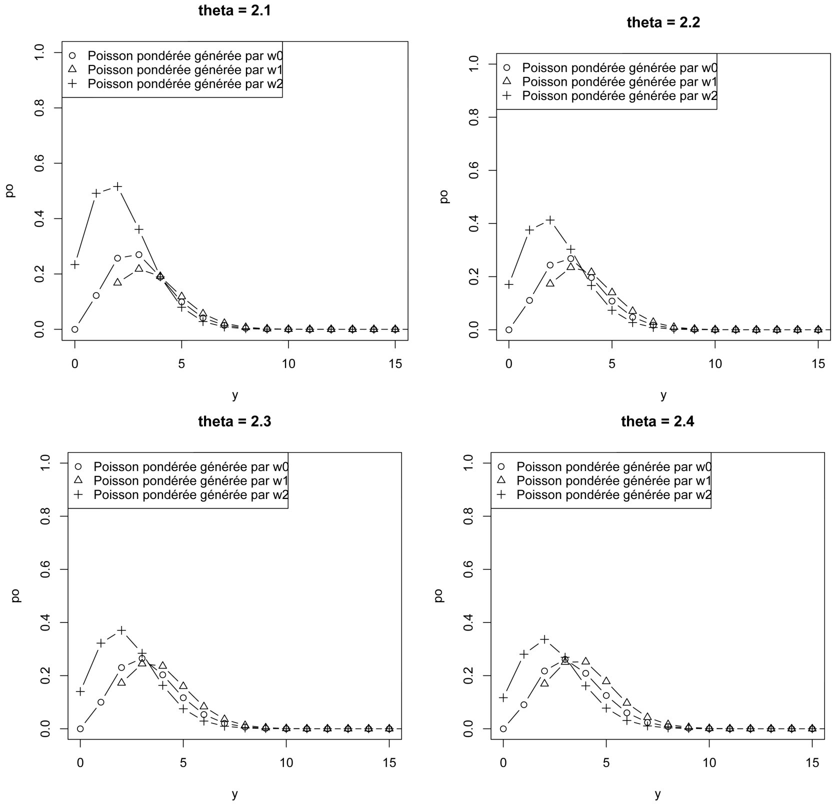


Tableau 2. coefficients de convexité

θ	α_1	α_2
2.1	0.3475447	0.6524553
3	0.127322	0.872678
4	0.0669873	0.9330127
5	0.04174243	0.9582576

6. Conclusion

Dans cet article nous avons décomposé une loi de Poisson pondérée en une combinaison convexe des lois duales à partir de sa fonction poids, du fait qu'une fonction poids génère une loi de Poisson pondérée. Cette décomposition n'est possible que lorsque la fonction poids est plus grande que 2. Ce problème de décomposition est une illustration d'un arbre d'une division cellulaire.

Références

Castillo J.D. and Pérez-Casany, M., 1998. Weighted Poisson distributions for overdispersion and underdispersion situations. *Ann. Inst. Statist. Math.* vol. 50, no. 3, 567-585.

Fisher R.A., 1934. The effects of methodes of ascertainment upon the estimation of frequencies. *Eugenics Ann.* **6**, 13-25.

Johnson N.L, Kotz S. and Kemp A.W., 1993, Univariate discrete distributions. Second Edition, Wiley Series in Probability and Mathematical Statistics..

Kokonendji, C.C., Mizere, D. and Balakrishnan, N., 2008. Connections of the Poisson weight function to overdispersion and underdispersion. *J. Statist. Plann. Inférence.* **138**, 1287-1296.

Mizere, D., Makany, R., Kissita, G. and Nkasa, Y., 2009. determination of the mean dual distribution through the Poisson weight function. *Far. East. Theor. Stat.* 29, no. 1. 53-63.

Patil G.P., (2002), Weighted distributions. In Encyclopedia of Environmetrics(eds A.H. El-Shaarawi and W.W. Piegorrsch), vol. **4**, 2369-2377. *Chichester: John Wiley and Sons.*

Patil G.P. and Rao C.R., (1978), Weighted distributions and size-biased sampling with applications to widlife populations and human families. *Biometrics.* **34**, 179-189.

Robert C.P., Ryden T., Titterington D.M., (2000), Bayesian inference in Hidden Markov Models through reversible jump Markov Chain Monte Carlo. J.Royal Stat. Society, Series B, Vol.62, pp 57-75.

Saint-Pierre, G., 2003. Identification du nombre de composants d'un mélange Gaussien par chaînes de Markov à sauts dans le cas multivarié ou par maximum de vraisemblance dans le cas univarié. Thèse présentée à l'université Toulouse 3-Paul Sabatier.