Journal of Statistical Theory and Applications
Volume 11, Number 4, 2012, pp. 337-352 ISSN 1538-7887

A Note on Weighted Count Distributions

Célestin C. Kokonendji* and Marta Pérez-Casany †

Abstract

As particular case of the Radon-Nikodym theorem we show that any count distribution is a weighted Poisson distribution (WPD) and, more generally, a weighted version of any other count distribution with suitable support. Several properties and consequences of this result are then given. With a condition on their supports, we deduce that two count distributions are connected by the weightening operator which commutes with some classical transformations. The notion of dual distribution, initially introduced for WPD, is also extended to any count distribution and a practical interpretation of its meaning is given. Finally, some concluding remarks and discussions are made for practical purposes.

Key Words and Phrases: Dual distribution, exponential dispersion model, left-truncation, mixed Poisson distribution, size biasing, weighted count distribution, zero-modification.

AMS 2010 Subject Classifications: 60E05, 62E10, 62E15.

*Université de Franche-Comté, UFR Sciences et Techniques, Laboratoire de Mathématiques de Besançon - UMR 6623 CNRS-UFC, 16 route de Gray, 25030 Besançon cedex, France. email: celestin.kokonendji@univ-fcomte.fr

†Technical University of Catalonia, Department of Applied Mathematics II and DAMA-UPC, Jordi Girona 1-3 08034 Barcelona, Spain. email: marta.perez@upc.edu

1 Introduction

The most general form of modifying an initial count distribution on its support is to multiply it by a weight function as follows. Let X be a count random variable (r.v.) with *probability mass function* (pmf) $\mathbb{P}(X = k) = p(k;\theta)$, $k \in \mathbb{N} := \{0, 1, 2, \cdots\}$, depending on parameters $\theta \in \Theta \subseteq \mathbb{R}^d$. Suppose that when the event $X = k$ occurs, the probability of ascertaining it is equal to $\omega(k)$; e.g. Fisher (1934). The recorded k is thus a realization of the count r.v. X^ω, which is said to be the *weighted version of* X. Its pmf is given by

$$\mathbb{P}(X^w = k) = p_\omega(k;\theta) = \frac{\omega(k)\ \mathbb{P}(X = k)}{\mathbb{E}\left[w\left(X\right)\right]}, \quad k \in \mathbb{N}, \tag{1}$$

where the denominator is the normalizing constant depending on θ. The discrete weight (or recording) function $\omega(k)$ is a nonnegative function on $\mathbb{N}$ and, from (1), it is clear that $0 < \mathbb{E}\left[\omega\left(X\right)\right] = \sum_{k\in\mathbb{N}} \omega(k)\ \mathbb{P}(X = k) < \infty$. The discrete weight function $\omega(k) \equiv \omega(k;\phi)$ can depend on a parameter ϕ representing the recording mechanism, and it may also be connected to the underlying initial parameter θ.

The most popular and useful of weighted count distributions are the *weighted Poisson distributions* (WPDs), obtained when the initial count r.v., X, follows a Poisson distribution. Such WPDs provide a unified approach to handle, among others, both dispersion and zero-proportion phenomena; see Kokonendji et al. (2008, 2009) and the references therein and, also Balakrishnan and Kozubowski (2008) and Borges et al. (2012) for the stochastic processes version. The concept of WPD was originally introduced by Rao (1965) and used widely as a tool in the selection of appropriate models for observed data drawn without a proper frame (e.g. Patil and Rao, 1978). Recently, several authors have proposed and analyzed different types of WPDs (e.g. Ridout and Besbeas, 2004; Shmueli et al., 2005; and, Castillo and Pérez-Casany, 2005); other authors have defined some forms of weight functions for investigating theoretical properties of the initial r.v. like probability weighted moment, moment generating function, proportional hazards, etc.; see, for example, Avval Riabi et al. (2010). Some authors have studied directly count distributions and their behaviours on real count data (e.g. Nikoloulopoulos and Karlis, 2008; Puig and Valero, 2006). Since it is easy to observe that any count distribution can be formulated as a WPD (Theorem 1), other authors have characterized the Poisson weight function with respect to a specific property of the initial count distribution like a uniform measure to detect departures from the reference Poisson distribution by Kokonendji et al. (2008); in particular, the connection between the logconvexity (logconcavity) of the Poisson weight function and the overdispersion (underdispersion) has been determined. Of course, this fact does not mean that their generating process is necessarily an imperfect recording of the classical Poisson distribution. It is also known that a discrete logconcave distribution is necessarily increasing a failure rate (IFR) and strongly unimodal, while a discrete logconvex distribution is necessarily decreasing failure rate (DFR) and infinitely divisible (e.g. Steutel, 1985). Further, it is known that if a count distribution is infinitely divisible, it is overdispersed, and if it is not infinitely divisible, it can be either overdispersed or underdispersed. Finally, it is well-known that the Poisson distribution is also a WPD, logconcave (and hence IFR and strongly unimodal), infinitely divisible, and equidispersed; see, e.g., Johnson et al. (2005). For a discussion on over- and underdispersion, one may refer to Xekalaki (2006).

The objectives of this work are of two-fold: to fix the representation of any count distribution as a WPD with their consequences and to know if the weightening operation commutes with some useful transformations applied to count distributions. The rest of this paper is organized as follows. In Section 2 we show several results related to the representation of any count distribution as a weighted version of another one. We also prove the commutativity of the weightening operation under an appropriate condition and we introduce a converse representation following the concept of dual distribution. Section 3 is devoted to some illustrations on two count models and their consequences with respect to three useful transformations of our main results. Finally, some concluding remarks and perspectives are briefly made for practical situations in Section 4.

2 Some results about weighted distributions

The Radon-Nikodym theorem states that given a measurable space $(\mathbb{T}, \mathcal{A})$, and given two σ-finite measures μ_1 and μ_2 on it, if μ_2 is absolutely continuous with respect to μ_1, then there exists a measurable positive real function f defined on $\mathbb{T}$, such that for any measurable set A,

$$\mu_2(A) = \int_A f d\mu_1.$$

In particular the theorem holds for probability measures. Thus, if one takes $(\mathbb{T}, \mathcal{A}) = (\mathbb{N}, \mathcal{F})$, being $\mathcal{F}$ the σ-algebra of events, and μ_1 equal to the Poisson measure, one has the following result:

Theorem 1 *Let Y be a count r.v, and let X be a Poisson r.v. with mean $\mu > 0$. Then:*

$$Y \stackrel{d}{=} X^w,$$

where "$\stackrel{d}{=}$" stands for "equality in distribution" and the Poisson weight function w is

$$w(k) = e^{\mu} \mu^{-k} (k!) \, \mathbb{P}(Y = k), \quad \forall k \in \mathbb{N}. \tag{2}$$

Remark 1 ***(i)*** *It is also possible to define the Poisson weight function (2) without the factor* $\exp(\mu)$*. In that case,* $\mathbb{E}[w(X)] = \exp(-\mu)$*.*
(ii) *If $\mathbb{E}(Y) = \mu_Y < \infty$, one can take $\mu = \mu_Y$ and thus the weight function only depends on the parameters of the Y probability distribution. This choice allows to interpret a given count r.v. as a weighted version of a Poisson r.v. with the same mean. This is particularly interesting, when one wants to compare the two probability distributions from the dispersion or the zero-proportion point of view.*
(iii) *If the distribution of Y depends on $\theta \in \Theta \subseteq \mathbb{R}^d$ and $\mathbb{E}(X) = \mu(\theta)$, i.e. it also depends on θ, then $w(\cdot) = w(\cdot; \theta, \mu)$ is proportional to $w(\cdot; \theta')$ with $\theta' \in \Theta' \subseteq \Theta$.*

The next proposition points out the connection between two representations of a count distribution as WPDs.

Proposition 1 *Let X_1 and X_2 be two Poisson r.v.'s with parameters μ_1 and μ_2 respectively.*

If a count r.v. Y is such that $Y \overset{d}{=} X_1^{w_1}$ with w_1 defined as in (2), then $Y \overset{d}{=} X_2^{w_2}$ with

$$w_2(k) = \left(\frac{\mu_1}{\mu_2}\right)^k w_1(k), \quad \forall k \in \mathbb{N}. \tag{3}$$

PROOF: Assume that $Y \overset{\mathrm{d}}{=} X_1^{w_1}$ with w_1 defined as in (2). Since

$$\mathbb{E}\left[w_2\left(X_2\right)\right] = \sum_{k\in\mathbb{N}} \left(\frac{\mu_1}{\mu_2}\right)^k w_1(k)\ \frac{e^{-\mu_2}\mu_2^k}{k!} = \exp\left(\mu_1 - \mu_2\right) < \infty,$$

the pmf of $X_2^{w_2}$ from (3)

$$\mathbb{P}(X_2^{w_2} = k) = \frac{e^{-\mu_2}\mu_2^k\ w_2(k)}{k!\ \mathbb{E}\left[w_2\left(X_2\right)\right]} = \frac{e^{-\mu_1}\mu_1^k\ w_1(k)}{k!} = \mathbb{P}(X_1^{w_1} = k) = \mathbb{P}(Y = k)$$

for all $k \in S_Y = S_{X_1^{w_1}}$, and $\mathbb{P}(X_2^{w_2} = k) = 0$ for $k \notin S_Y = S_{X_1^{w_1}}$. □

The following proposition is somehow the converse of Theorem 1, and shows how to obtain the Poisson distribution as a weighted version of a count distribution with support the nonnegative integers set.

Proposition 2 *Let Y be a count r.v. with support $S_Y = \mathbb{N}$, and let X be a Poisson r.v. with mean $\mu > 0$. Then:*

$$Y^{1/w} \overset{d}{=} X$$

where w is the positive Poisson weight function defined in (2).

In order to prove it, we need the following theorem which is more general and points out the commutativity of the weightening operation.

Theorem 2 *Let Y be a count r.v., and let w_1 and w_2 be two weight functions such that $0 < \mathbb{E}\left[w_j\left(Y\right)\right] < \infty$, $j = 1, 2$, and $0 < \mathbb{E}\left[w_1\left(Y\right) w_2\left(Y\right)\right] < \infty$. Then:*

$$\left(Y^{w_1}\right)^{w_2} \overset{d}{=} Y^{w_1 w_2} \overset{d}{=} \left(Y^{w_2}\right)^{w_1}.$$

Proof: Without loss of generality, we assume that the support of Y is $S_Y = \mathbb{N}$. Given that,

$$\mathbb{E}\left[w_2\left(Y^{w_1}\right)\right] = \frac{1}{\mathbb{E}\left[w_1\left(Y\right)\right]} \sum_{k\in\mathbb{N}} w_1(k)w_2(k)\mathbb{P}(Y=k) = \frac{\mathbb{E}\left[w_1\left(Y\right)w_2\left(Y\right)\right]}{\mathbb{E}\left[w_1\left(Y\right)\right]},$$

the pmf of $(Y^{w_1})^{w_2}$ is given by

$$\mathbb{P}\left((Y^{w_1})^{w_2} = k\right) = \frac{w_2(k)\mathbb{P}(Y^{w_1}=k)}{\mathbb{E}\left[w_2\left(Y^{w_1}\right)\right]} = \frac{w_1(k)w_2(k)\mathbb{P}(Y=k)}{\mathbb{E}\left[w_1\left(Y\right)\right]\times\mathbb{E}\left[w_2\left(Y^{w_1}\right)\right]} = \mathbb{P}\left(Y^{w_1w_2}=k\right).$$

And the same is true by interchanging w_1 and w_2. $\square$

Remark 2 *That $\mathbb{E}\left[w_j\left(Y\right)\right] < \infty$, $j = 1, 2$, do not necessarily implies that $\mathbb{E}\left[w_1\left(Y\right)w_2\left(Y\right)\right] < \infty$.*

It is enought to take Y a Poisson r.v. with mean $\mu \in (1/e, 1)$, $w_1(k) = k!$ and $w_2(k) = \exp(k)$ for all $k \in \mathbb{N}$. In that case, $\mathbb{E}\left[w_1\left(Y\right)\right] = \exp(-\mu)/\left(1-\mu\right) < \infty$, $\mathbb{E}\left[w_2\left(Y\right)\right] = \exp\left[\mu\left(e-1\right)\right] < \infty$, and $\mathbb{E}\left[w_1\left(Y\right)w_2\left(Y\right)\right] = \exp(-\mu)\sum_{k\geq 0}\left(\mu e\right)^k = \infty$ because $\mu e > 1$.

Proof of Proposition 2: Given that $S_Y = S_X = \mathbb{N}$, from Theorem 1 we have $Y \overset{\text{d}}{=} X^w$ where w defined as in (2) is positive (i.e. $w(k) > 0$ for all $k \in \mathbb{N}$) and $\mathbb{E}\left[w\left(X\right)\right] = 1 < \infty$. Thus, $1/w$ is also a positive weight function. Combining two properties in Kokonendji et al. (2008), which are Part (ii) of Remark (page 1292) and Part (i) of Theorem 6, one has that

$$\lim_{k\to\infty} \frac{\mu \times w(k-1)}{k \times w(k)} = \lim_{k\to\infty} \left(\frac{\mu}{k}\right)^2 \frac{\mathbb{P}(Y=k-1)}{\mathbb{P}(Y=k)} = 0$$

which implies that $\mathbb{E}\left[1/w\left(X\right)\right] < \infty$. Applying Theorem 2 (with $w_1 \equiv w$, $w_2 \equiv 1/w$ and $Y \equiv X$) we can write

$$\left(X^w\right)^{1/w} \overset{\text{d}}{=} X \iff Y^{1/w} \overset{\text{d}}{=} X.\square$$

To get an interpretation of the converse representation obtained in Proposition 2 one can use the concept of duality for WPDs introduced by Kokonendji et al. (2008, Section 3). The following definition extends this notion to any reference count distribution different from the Poisson distribution.

Definition 1 *Let Y be a count r.v. on $S_Y = \mathbb{N}$, and let w_1 and w_2 be two positive weight functions such that $\mathbb{E}[w_j(Y)] < \infty$, $j = 1, 2$. The two corresponding weighted versions $Y^{w_1} \equiv Y_1$ and $Y^{w_2} \equiv Y_2$ are said to be a dual pair with respect to Y if, and only if,*

$$w_1(k)w_2(k) = 1, \quad \forall k \in \mathbb{N}.$$

Next result generalizes Theorem 1, replacing the reference Poisson r.v. by another count r.v. It is also a particular situation of the Radon-Nykodym theorem.

Theorem 3 *Let Y and Z be two count r.v.'s with support S_Y and S_Z, respectively. If $S_Y \subseteq S_Z$ then*

$$Y \stackrel{d}{=} Z^w.$$

where the weight function w is

$$w(k) = \frac{\mathbb{P}(Y = k)}{\mathbb{P}(Z = k)}, \quad \forall k \in S_Z. \tag{4}$$

Remark 3 ***(i)*** *Theorem 1 can be considered as a corollary of Theorem 3. Nevertheless, it has been presented in that way because the Poisson distribution is the most used in practice, and the corresponding Poisson weight function provides some properties or characteristics on the original count r.v. (e.g. Kokonendji et al., 2008). Theorem 3 could be used to study the weighted version of binomial distribution; see, also, Chakraborty and Das (2006).*
(ii) *An alternative proof of Theorem 3 when $S_Z = \mathbb{N}$ is obtained from Theorem 1 using the dual notion (Definition 1) and Theorem 2 as follows: let X be a Poisson r.v. such that $Y \stackrel{d}{=} X^{w_Y}$ (Theorem 1) and $Z \stackrel{d}{=} X^{w_Z} \Leftrightarrow X \stackrel{d}{=} Z^{1/w_Z}$ (Proposition 2); then using therefore Theorem 2 and taking $w = w_Y/w_Z$, we can write:*

$$Y \stackrel{d}{=} X^{w_Y} \stackrel{d}{=} \left(Z^{1/w_Z}\right)^{w_Y} \stackrel{d}{=} Z^{w_Y/w_Z} \stackrel{d}{=} Z^w.$$

The following Proposition generalizes Proposition 1.

Proposition 3 *Let Z_1 and Z_2 be two count r.v. with the same supports $S_{Z_1} = S_{Z_2} = S$. If a count r.v. Y is such that $S_Y \subseteq S$ and $Y \stackrel{d}{=} Z_1^{w_1}$ with w_1 defined as in (4), then $Y \stackrel{d}{=} Z_2^{w_2}$ with*

$$w_2(k) = w_1(k)\frac{\mathbb{P}(Z_1 = k)}{\mathbb{P}(Z_2 = k)} \times \frac{\mathbb{E}\left[w_2\left(Z_2\right)\right]}{\mathbb{E}\left[w_1\left(Z_1\right)\right]}, \quad \forall k \in \mathbb{N}.$$

Finally note that we also have an equivalence to Proposition 2 with $Y^{1/w} \stackrel{\mathrm{d}}{=} Z$ if we assume $S_Y = S_Z$ which can be interpret from the dual point of view.

3 Illustrations and consequences

This section has two parts, in the first one we give a representation as a weighted Poisson model, of a general exponential dispersion model and a general mixed Poisson model. In the second part, we define three transformations very useful in practice, and for each one we stablish a consequence obtained from the commutativity of the weighted operator.

3.1 Two count models

A given r.v. Y follows an exponential dispersion distribution on $\mathbb{N}$ if, and only if, its pmf has the following form:

$$\mathbb{P}(Y = k) = \varphi(k;\phi)\exp\left[\theta k - K(\theta;\phi)\right], \qquad k \in \mathbb{N},$$

where $\phi > 0$, $\theta \in \Theta \subseteq \mathbb{R}$, $\varphi(\cdot;\phi) > 0$ and $K(\theta;\phi)$ is the normalizing constant; see, for example, Kokonendji et al. (2004). Thus, if a r.v. X follows the Poisson distribution with mean $\mu = \exp(\theta)$, from Theorem 1 we can write $Y \stackrel{\mathrm{d}}{=} X^w$ with

$$w(k;\theta,\phi) = \exp\left(e^\theta - \theta k\right)\ (k!)\,\mathbb{P}(Y = k) = (k!)\varphi(k;\phi)\exp\left[e^\theta - K(\theta;\phi)\right] \propto (k!)\varphi(k;\phi);$$

see also Kokonendji et al. (2008, Theorem 2). This illustrates Part (iii) of our Remak 1. As particular cases of exponential dispersion models, we have all distributions of natural exponential families concentrated on $\mathbb{N}$ (with fixed $\phi > 0$) like binomial and negative binomial distributions. Thus, the logconvexity (logconcavity) of $k \mapsto (k!)\varphi(k;\phi) =: \omega(k;\phi)$ provides the overdispersion (underdispersion) of the exponential dispersion r.v. Y with respect to the reference Poisson distribution; see Kokonendji et al. (2008, Corollary of Theorems 3 and 4).

Following the notation of Grandell (1997), a r.v. Y follows a mixed Poisson distribution with $\mathbb{E}(Y) = \mu > 0$ if, and only if, there exists a nonnegative random variable T with cumulative distribution function F_T such that

$$\mathbb{P}(Y = k) = \int_0^\infty \frac{e^{-\mu t}(\mu t)^k}{k!} dF_T(t), \qquad k \in \mathbb{N};$$

i.e., the distribution of Y given $T = t$ is the Poisson distribution with mean μt. Particular cases of mixed Poisson distributions consist of the negative binomial, the Delaporte, the Pólya-Aeppli, the lognormal-Poisson and, more globally, the Poisson-Tweedie distributions that one can refer to Kokonendji et al. (2004) and El-Shaarawi et al. (2010). Considering the reference Poisson r.v. X with mean $\mu > 0$ and assuming that the mixing distribution has moments of any order, from Theorem 1 we get that $Y \stackrel{\mathrm{d}}{=} X^w$ with a weight function proportional to the k-th moment of the mixing distribution, since

$$w(k;\mu) = e^\mu \int_0^\infty e^{-\mu t} t^k dF_T(t) = e^\mu (-1)^k \frac{d^k}{d\mu^k} \mathbb{E}\left(e^{-\mu T}\right) \propto \mathbb{E}\left(T^k\right).$$

3.2 Three useful transformations

Here we lay out some consequences of the commutativity result of the weightening operation (Theorem 2).

The first transformation in consideration is the *left-truncation* of an initial count r.v. Y at fixed $k_0 \in \mathbb{N}$, denoted by Y^{k_0-tr}, which is represented as Y^{w_1} with

$$w_1(k;k_0) = \begin{cases} 0 & \text{if} \quad 0 \le k \le k_0 \\ 1 & \text{if} \quad k \ge k_0 + 1. \end{cases}$$

Dealing with real data, quite often the researcher has to modify the distribution of the initial count r.v. in order to adapt it to the impossibility of observing the zero value. This fact makes the left-truncation at zero the most usual type of left-truncation. The following result is straightforward from Theorem 2:

Corollary 1 *Given a count r.v.* Y *and a weight function* w *such that* $\mathbb{E}[w(Y)] < \infty$*, then:*

$$\left(Y^{k_0-tr}\right)^w \stackrel{d}{=} (Y^w)^{k_0-tr}.$$

From a practical point of view, this result implies that if the researcher needs both to truncate and to weight with a given function an initial distribution in order to adapt it to the data ascertaining mecanism, he does not have to worry about the order in which the two operations are done. This does not hold if one considers the mixing operation instead of the weightening operation as pointed out in the following remark.

Remark 4 ***(i)*** *Böhning and Kuhnert (2006) proved that if one considers the set of mixed Poisson distributions with a mixing distribution with finite support, the set of distributions obtained by zero-truncating a mixed Poisson distribution and the set of distributions obtained by mixing zero-truncated Poisson distributions are equivalent. Nevertheless, the mixing distribution is not the same and thus, mixing with a finite mixture distribution and truncating do not commute.*
(ii) *Recently, Valero et al. (2010) proved that if one considers all plausible mixing distributions, i.e. with any kind of support, the set of distributions obtained by zero-truncating a mixed Poisson is striclty included in the one obtained by mixing zero-truncated Poisson distributions.*

The second transformation considered is the zero-modification of a count distribution for which the zero-truncation is a limiting case. That is useful when the data obtained by the experimenters exhibits a larger (smaller) proportion of zeros compared with an initial count distribution. Given a count r.v. Y, and given a value ϵ belonging to the interval $(-p_0/(1-p_0), 1)$ with $p_0 = \mathbb{P}(Y = 0)$, the *zero-modification* of Y with value ϵ, denoted by Y_ϵ^{ZM}, is defined as the finite mixture between the original distribution of Y and the degenerate distribution at zero with weights $1-\epsilon$ and ϵ, respectively. We can rewrite Y_{ϵ}^{ZM} as a weighted version of Y with weight function:

$$w_2(k;\epsilon) = \begin{cases} \frac{(1-\epsilon)p_0+\epsilon}{(1-\epsilon)\,p_0} & \text{for} \quad k=0 \\ 1 & \text{for} \quad k \geq 1. \end{cases}$$

Corollary 2 *Given a count r.v. Y and a weight function w such that $\mathbb{E}\left[w\left(Y\right)\right] < \infty$, it follows that:*

$$\left(Y_{\epsilon}^{ZM}\right)^{w} \stackrel{d}{=} \left(Y^{w}\right)_{\epsilon}^{ZM}.$$

The third transformation is the *size-biasing* of order α ($\alpha > 0$) of a given count r.v. Y. It is defined to be the weighted version of Y, $Y^{*\alpha}$, with weight function

$$w_3(k;\alpha) = k^{\alpha}, \qquad k \in \mathbb{N}.$$

The size-biased version of order one is the most used in practice, since it corresponds to the sample distribution when the probability to ascertain an observation is proportional to the value of the observation itself.

Corollary 3 *Given a count r.v. Y and a weight function w such that $\mathbb{E}\left[w\left(Y\right)\right] < \infty$. If $\left(Y^{*\alpha}\right)^{w}$ or $\left(Y^{w}\right)^{*\alpha}$ exists, it follows that*

$$\left(Y^{*\alpha}\right)^{w} \stackrel{d}{=} \left(Y^{w}\right)^{*\alpha}.$$

This result does not hold if one considers mixed Poisson distributions instead of weighted Poisson distributions. Indeed, Valero et al. (2010) proved that the size-biased version of a Poisson mixture is neither a zero-truncated mixed Poisson distribution nor a mixture of zero-truncated Poisson distributions.

4 Concluding remarks

In this work, we pointed out an efficient representation of any count distribution as a weighted version of any previously fixed count distribution. In particular this is true when the reference distribution is the well-known Poisson distribution. We have also established that under an appropriate condition, weighting an already weighted distribution is equivalent to weighting once the initial distribution with the product of weights. As a consequence, classical transformations like left-truncating, zero-modifying or size-biasing commute with the weightening operation. These results are important for the practitioners because in cases where it is requiered to apply two or more types of weightening, they do not have to worry about the order in which the weights are considered.

The main fact that any count distribution may be viewed as a WPD allows several echoes. Firstly, we hence have a new interpretation of classical count distributions in the literature. We can also deduce the converse interpretation through the concept of duality. Secondly, it allows to investigate or compare count distributions means of comparing their corresponding Poisson weight functions. For example, the overdispersion and underdispersion properties of a count distribution have been studied through the nature of their Poisson weight functions by Kokonendji et al. (2008); similar characterization can be done for the infinite divisibility property of a count distributions presented in Steutel and Van Harn (2003, Chapter 2).

Finally, in order to estimate any count function $g : \mathbb{N} \to \mathbb{R}$ under nonparametric assumption like pmf and regression functions on count explanatory variables we can henceforth consider a semiparametric approach by decomposing

$$g(k) = w(k;\theta)p(k;\theta) =: g_w(k;\theta), \qquad k \in \mathbb{N} \tag{5}$$

with $p(k;\theta)$ being the parametric part well specified and depending on θ, say Poisson or binomial, and $w(k;\theta)$ being the unknown discrete weight function or nonparametric part for fixed θ. Works in these directions are recently done, without theoretical justification, by Kokonendji et al. (2009) for pmf and by Abdous et al. (2012) for count regression function. In the previous papers, the semiparametric estimation of $g(\cdot) = g_w(\cdot;\theta)$ following the decomposition (5) is represented as $\widetilde{g}(\cdot) = g_{\widetilde{w}}(\cdot;\widehat{\theta}_P)$ where, firstly the estimation of parameter θ, $\widehat{\theta}_P$, is obtained by maximum likelihood or by least squared methods, and secondly, for the value of $\widehat{\theta}_P$ obtained, one obtains the discrete kernel estimation of $w(\cdot) = w(\cdot;\widehat{\theta}_P)$ namely $\widetilde{w}(\cdot) = \widetilde{w}(\cdot;\widehat{\theta}_P)$. This procedure can be pursued for investigating the convergence of the estimator.See also Kokonendji et al.

(2007), Kokonendji and Zocchi (2010), Kokonendji and Senga Kiessé (2011), and Zougab et al. (2012) for some new materials related to the discrete associated kernel method for smoothing discrete functions.

Note that another way to estimate $g(\cdot) = g_\omega(\cdot;\theta) = \omega(\cdot)p(\cdot;\theta)/\sum_k \omega(k)p(k;\theta)$ of (5) represented in the sense of (1) is to consider an *empirical weighted parametric* estimation

$$\widehat{g}(k) = g_{\widehat{\omega}}(k;\widehat{\theta}_W) = \widehat{\omega}_k p(k;\widehat{\theta}_W)/\sum_k \widehat{\omega}_k p(k;\widehat{\theta}_W)$$

for all k, where $\widehat{\omega}_k$ is the frequency of count k in a sample of size n and $\widehat{\theta}_W$ is an estimation of θ. Work in this direction must be done for investigating the convergence of $\widehat{g}(\cdot)$ to something that makes sense, method and consistency of $\widehat{\theta}_W$, and comparison of $\widetilde{g}(\cdot)$ with respect to $\widehat{g}(\cdot)$ and their perfomance through simulations and real count data.

Acknowledgements. Part of this work was done while the second author was at Laboratoire de Mathématiques de Besançon - UMR 6623 CNRS-UFC, with the partial supports of the Spanish Ministry of Education and Science and FEDER grants TIN2009-14560-C03-03, MTM2006-09920 and MTM2006-01477, and Generalitat de Catalunya through grant SGR-1187. The authors thank Editors and Referees for useful comments and suggestions on an earlier draft of the paper.

References

Abdous, B., Kokonendji, C.C. and Senga Kiessé, T. (2012), On semiparametric regression for count explanatory variables. *Journal of Statistical Planning and Inference* 142, 1537-1548.

Avval Riabi, M.Y., Borzadaran, G.R.M. and Yari, G.H. (2010), β-entropy for Pareto-type distributions and related weighted distributions. *Statistics and Probability Letters* 80, 1512-1519.

Balakrishnan, N. and Kozubowski, T.J. (2008), A class of weighted Poisson processes. *Statistics and Probability Letters* 78, 2346–2352.

Böhning, D. and Kuhnert, R. (2006), Equivalence of truncated count mixture distributions and mixtures of truncated distributions. *Biometrics* 62, 1207–1215.

Borges, P., Rodrigues, J. and Balakrishnan, N. (2012). A class of correlated weighted Poisson processes. *Journal of Statistical Planning and Inference* 142, 366-375.

Castillo, J. and Pérez-Casany, M. (2005), Overdispersed and underdispersed Poisson generalizations. *Journal of Statistical Planning and Inference* 134, 486–500.

Chakraborty, S. and Das, K.K. (2006), On some properties of a class of weighted quasi-binomial distributions. *Journal of Statistical Planning and Inference* 136, 156–182.

El-Shaarawi, A.H., Zhu, R. and Joe, H. (2010), Modelling species abundance using the Poisson–Tweedie family. *Environmetrics* 22, 152-164.

Fisher, R.A. (1934), The effects of methods of ascertainment upon the estimation of frequencies. *Annals of Eugenics* 6,13-25.

Grandell, J. (1997), *Mixed Poisson Processes*, Monographs on Statistics and Applied Probability, London: Chapman and Hall.

Johnson, N.L., Kemp, A.W. and Kotz, S. (2005), *Univariate Discrete Distributions*, 3rd edition. New Jersey: John Wiley and Sons, Hoboken.

Kokonendji, C.C. and Senga Kiessé, T. (2011), Discrete associated kernel methods and extensions. *Statistical Methodology* 8, 495-516.

Kokonendji, C.C. and Zocchi, S.S. (2010), Extensions of discrete triangular distribution and boundary bias in kernel estimation for discrete functions. *Statistics and Probability Letters* 80, 1655-1662.

Kokonendji, C.C., Dossou-Gbété, S. and Demétrio, C.G.B. (2004), Some discrete exponential dispersion models: Poisson-Tweedie and Hinde-Demétrio classes. *Statistics and Operations Research Transactions* 28, 201-214.

Kokonendji, C.C., Mizère, D. and Balakrishnan, N. (2008), Connections of the Poisson weight function to overdispersion and underdispersion. *Journal of Statistical Planning and Inference* 138, 1287–1296.

Kokonendji, C.C., Senga Kiessé, T. and Balakrishnan, N. (2009), Semiparametric estimation for count data through weighted distributions. *Journal of Statistical Planning and Inference* 139, 3625-3638.

Kokonendji, C.C., Senga Kiessé, T. and Zocchi, S.S. (2007), Discrete triangular distributions and non-parametric estimation for probability mass function. *Journal of Nonparamatric Statistics* 19, 241–254.

Nikoloulopoulos, A.K. and Karlis, D. (2008). On modeling count data: a comparison of some well-known discrete distributions. *Journal of Statistical Computation and Simulation* 78, 437-457.

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

Puig, P. and Valero, J. (2006), Count data distributions: some characterizations with applications. *Journal of the American Statistical Association* 101, 332–340.

Rao, C.R. (1965), Weighted distributions arising out of methods of ascertainment. In *Classical and Contagious Discrete Distributions*, G. P. Patil (Eds). Calcuta: pergamon Press and Statistical Publishing Society, 320-332.

Ridout, M.S. and Besbeas, P. (2004), An empirical model for underdispersed count data. *Statistical Modelling* 4, 77–89.

Shmueli, G., Minka, T.P., Kadane, J.P., Borle, S. and Boatwright, P. (2005), A useful distribution for fitting discrete data: revival of the Conway–Maxwell–Poisson distribution. *Journal of the Royal Statistical Society, Ser. C* 54, 127–142.

Steutel, F.W. (1985), Log-concave and log-convex distributions. In: Kotz, S., Johnson, N.L., Read, C.B. eds., *Encyclopedia of Statistical Sciences* vol. 5: 116-117. New York: Wiley.

Steutel, F.W. and Van Harn, K. (2003), *Infinite Divisibility of Probability Distributions on the Real Line*. New York: Marcel Dekker.

Valero, J., Pérez-Casany, M. and Ginebra, J. (2010), On zero-truncating and mixing Poisson distributions. *Advances in Applied Probability* 42, 1013-1027.

Xekalaki, E. (2006), Under- and overdispersion, In: *Encyclopedia of Actuarial Science,* vol. 3, pp.1700-1705, John Wiley & Sons, Hoboken, New Jersey.

Zougab, N., Adjabi, S. and Kokonendji, C.C. (2012), Binomial kernel and Bayes local bandwidth in discrete functions estimation. *Journal of Nonparametric Statistics* [In Press; DOI: 10.1080/10485252.2012.678847].