EXTENDED TRUNCATED TWEEDIE-POISSON MODEL

JORDI VALERO, JOSEP GINEBRA, AND MARTA PÉREZ-CASANY

Abstract. It has been argued that by truncating the sample space of the negative binomial and of the inverse Gaussian-Poisson mixture models at zero, one is allowed to extend the parameter space of the model. Here that is proved to be the case for the more general three parameter Tweedie-Poisson mixture model. It is also proved that the distributions in the extended part of the parameter space are not the zero truncation of mixed Poisson distributions and that, other than for the negative binomial, they are not mixtures of zero truncated Poisson distributions either. By extending the parameter space one can improve the fit when the frequency of one is larger and the right tail is heavier than is allowed by the unextended model. Considering the extended model also allows one to use the basic maximum likelihood based inference tools when parameter estimates fall in the extended part of the parameter space, and hence when the m.l.e. does not exist under the unextended model. This extended truncated Tweedie-Poisson model is proved to be useful in the analysis of words and species frequency count data.

1. Introduction

The Poisson model is the usual one for describing non-negative integer data. Most often though, count data has larger variability than the one expected from the Poisson model. In many of those instances, the data generating process can be modelled through a two stage process in which the distribution of each count would be Poisson but with an expected value that changes from count to count randomly. By modelling the distribution of the Poisson expectation one is naturally lead to the use of Poisson mixture models.

The most common choice for a mixing distribution is the gamma distribution, which leads to the negative binomial model, and one popular alternative is the inverse Gaussian distribution, which leads to the inverse Gaussian-Poisson model (IG-Poisson from now on) first considered by Holla (1966) to model highly skewed non-negative integer data. Both the gamma and the inverse Gaussian models are special cases of the three parameter Tweedie model first considered in Tweedie (1984). By considering the model that results from using as mixing model the family of Tweedie distributions with non-negative support, one is lead to the Tweedie-Poisson mixture model first considered in Hougaard (1987).

Key words and phrases. Negative binomial; Overdispersion; Poisson mixture; Species frequency; Truncated distribution; Tweedie models; Word frequency.

In applications like the analysis of species (words) frequency count data typically the zeros can not be observed, because the total number of species (words) in the population (vocabulary) is unknown, and the data generating process can be well approximated through zero truncated Poisson mixture models. These models have the advantage that estimates of the model mixing distribution serve as estimates of the the distribution of the species (words) frequencies in the population (vocabulary) and can be used to estimate the size and diversity of the population (vocabulary). In particular, Tweedie models are typically a good approximation to species frequencies and as a consequence the Tweedie-Poisson mixture model is a good candidate for the analysis of species frequency count data, as illustrated in El-Shaarawi et al. (2010).

Engen (1974, 1978) argues that by truncating the negative binomial model at zero one is entitled to extend its parameter space, which leads to the extended truncated negative binomial model used in ecology and in stylometry (see, e.g., Sichel, 1997). An analogous phenomenon is documented to happen with the IG-Poisson mixture model in Puig et al. (2009). Here it is proved that this is the case under these particular models as well as under the more general Tweedie-Poisson mixture model. It is also proved that the extended part of this newly defined extended truncated Tweedie-Poisson model can not be posed as a truncated version of a Poisson mixture model and that, other than for the extended truncated negative binomial special case, it can not be posed as a mixture of zero truncated Poisson distributions either.

Note that the term 'extended model' here is not being used in the sense of generalizing a model by adding parameters to it, but in the sense of making it more flexible by expanding its parameter space in a natural way. Extending the parameter space of a statistical model is also useful in that it lets one to use all the basic maximum likelihood based inference tools when the m.l.e. estimate falls in the extended part of the parameter space and therefore, when it would either fall in the boundary or it would not exist if the unextended model was considered.

The paper is organized as follows. Section 2 describes and motivates the Tweedie-Poisson mixture model and its zero truncated version. Section 3 proves that one can extend the parameter space of the zero truncated Tweedie-Poisson model and, in particular, that that is also the case for the negative binomial and IG-Poisson mixture models. The extended part of this model is useful when the frequency of ones is larger and the upper tail is heavier than is allowed by the unextended model. Section 4 proves that in the extended part of the parameter space the model is not the zero truncation of a Poisson mixture distribution. It also proves that other than for the negative binomial special case, the extended part of the truncated Tweedie-Poisson model is not a mixture of zero truncated Poisson distributions either.

Section 5 uses the extended truncated Tweedie-Poisson model on 456 sets of word frequency count data, and on 89 sets of species frequency count data and

finds that it often fits data significantly better than the extended truncated negative binomial and IG-Poisson submodels. It also finds many instances where the maximum likelihood estimate of the parameters fall in the extended part of the parameter space. Even though these are instances where the fit under the extended model is not significantly better than the fits obtained under the unextended model, they correspond to situations where the m.l.e. under the unextended model do not exist and one can not resort to the usual confidence intervals and hypothesis tests.

2. The zero truncated Tweedie-Poisson mixture model

2.1. **The three parameter Tweedie model.** Exponential dispersion models, first systematically studied in Jorgensen (1987, 1997), are two parameter models formed by linear exponential distributions with an additional dispersion parameter that generalize exponential family models and serve as distributions for generalized linear models. Tweedie models are exponential dispersion models with a power mean-variance relationship, and hence such that the variance of the corresponding random variable is $V(Y) = \psi\mu^p$, where μ is its mean and ψ is its dispersion parameter (see Tweedie, 1984, Bar-Lev and Enis, 1986, Hougaard, 1986a,b, Aalen, 1992).

The index p uniquely determines a model in the Tweedie family. When $p < 0$, models have positive means but are supported on the whole real line and they are rarely used in practice. When $p = 0$ one obtains the normal model. For p in $(0, 1)$ the Tweedie models do not exist.

All random variables with a Tweedie model with $p \geq 1$ are non-negative. When $p = 1$ one obtains the (scaled) Poisson model. The set of Tweedie models obtained for p in $(1, 2)$ are stopped Poisson sums of Gamma distributions (see Aalen, 1988), and hence have distributions with a point mass at zero and a continuous density function on the positive real line, which makes them useful for modelling non-negative real valued data when exact zeros are possible. All Tweedie models with $p \geq 2$ are formed by stable continuous distributions with strictly positive support, and they include as special cases the gamma model, with $p = 2$, and the inverse Gaussian model, with $p = 3$. Tweedie models with $p > 1$ are the natural candidates for modelling non-negative continuous data with an arbitrary measurement scale, because they are the only exponential dispersion models closed under re-scaling.

The three parameter statistical model obtained by considering p to be an additional parameter with a range in $[1, \infty)$, encompassing all the exponential dispersion distributions with power mean-variance relationship and non-negative support, will be denoted by the Tweedie model. Instead of p one often uses the parameter $\beta = (2 - p)/(1 - p)$, which ranges in $(-\infty, 1)$.

In general, the density functions of the Tweedie model, $\psi_{\alpha,\beta,\theta}(\pi)$, can not be written in a closed form expression, but they can be evaluated through the Fourier

inversion of its characteristic functions as described in Dunn and Smyth (2008). Applications of Tweedie models can be found for example in Davidian (1990), Hougaard et al. (1992) and Smyth and Jorgensen (2002).

2.2. **The Tweedie-Poisson mixture model.** Non-negative integer data are often modelled through a Poisson model with a random effect for the mean. In applications one typically assumes either a finitely supported mixing distribution for that mean, which leads to a finite mixture of Poisson distributions, or a two-parameter continuous distribution with non-negative support, like the gamma or the inverse gaussian distributions leading to the negative binomial and the IG-Poisson models.

Hougaard (1987), Hougaard et al. (1997) and Jorgensen (1997, pp. 165-170) consider the distribution of the Poisson mean to be the three-parameter Tweedie model described above. That leads to the model with probability mass function,

$$p_r^{tp}(\alpha,\beta,\theta)=\int_{\mathbb{R}^+}\frac{\lambda^r e^{-\lambda}}{r!}\psi_{\alpha,\beta,\theta}(\lambda)d\lambda,\ \text{for } r=0,1,\ldots, \tag{1}$$

which Johnson et al. (2005) names as the Tweedie-Poisson mixture model, and Kokonendji et al. (2004) names as the Poisson-Tweedie model.

Willmot (1989) advocates for the Tweedie-Poisson mixture model for being the Poisson stopped sum of the extended truncated negative binomial model of Engen (1974). Puig and Valero (2006) prove that the set of statistical models obtained from (1) after fixing the value of β, and hence of p, are the only two parameter mixed Poisson statistical models that are "partially closed under addition", (i.e. such that the sum of their independent replicates share the same statistical model), and at the same time are such that the maximum likelihood estimate of the population mean is the sample mean. The Tweedie-Poisson mixture model has also been independently proposed by Hoffman (1955), by Gerber (1991), and by Zhu and Joe (2009).

There is not a simple closed form expression for (1), and the Tweedie-Poisson model is most conveniently characterized through its probability generating functions (pgfs)

$$G_{\alpha,\beta,\theta}^{tp}(t)=E_{\alpha,\beta,\theta}(t^R)=e^{\frac{\alpha(1-\beta)}{\beta}((1-\theta)^\beta-(1-\theta t)^\beta)}, \tag{2}$$

where R is distributed as in (1), $\alpha\in(0,\infty)$, $\beta\in(-\infty,1)$ and $\theta\in(0,1)$. Note that β determines the index p through $p=(2-\beta)/(1-\beta)$, and (α,θ) determine the mean, μ, and the dispersion index, δ, through

$$\mu=E_{\alpha,\beta,\theta}^{tp}(R)=\frac{\alpha\theta(1-\beta)}{(1-\theta)^{1-\beta}},$$

and

$$\delta=\frac{Var_{\alpha,\beta,\theta}^{tp}(R)}{E_{\alpha,\beta,\theta}^{tp}(R)}=\frac{1-\theta\beta}{1-\theta},$$

with $\mu \in (0, \infty)$ and $\delta \in (1, \infty)$. The derivatives of (2) at $t = 1$ yield the factorial moments for R, and the probabilities of the Tweedie-Poisson distribution in (1) can be calculated through the derivatives of (2) at $t = 0$ by making use of $p_r = G^{(r)}(0)/r!$. These probabilities can be computed recursively through

$$p_0^{tp}(\alpha, \beta, \theta) = e^{\frac{\alpha(1-\beta)}{\beta}((1-\theta)^{\beta}-1)}, \tag{3}$$

and

$$p_r^{tp}(\alpha, \beta, \theta) = \frac{1}{r} \sum_{i=1}^{r} i\, a_i\, p_{r-i}^{tp}(\alpha, \beta, \theta), \text{ for } r = 1, 2, \ldots,$$

where $a_1 = \alpha\, \theta\, (1 - \beta)$ and $a_{i+1}/a_i = \theta\, (i - \beta)/(i + 1)$.

When $\beta \in (-\infty, 0)$, $p \in (1, 2)$, one obtains the Pólya-Aeppli models that are useful for zero-heavy data and can be posed as the Poisson stopped sum of the negative binomial model (Katti and Gurland, 1961). The limit of (2) when β tends to 0, p tends to 2, is

$$G_{\alpha,\beta=0,\theta}^{tp}(t) = e^{\alpha(\ln(1-\theta)-\ln(1-\theta t))} = \Big(\frac{1-\theta}{1-\theta t}\Big)^{\alpha},$$

that are the pgf's of the negative binomial model with $\mu = \alpha\, \theta/(1 - \theta)$ and $\delta = 1/(1-\theta)$. When $\beta = .5$, $p = 3$, one obtains the pgf's of the inverse Gaussian-Poisson mixture model,

$$G_{\alpha,\beta=.5,\theta}^{tp}(t) = e^{\alpha(\sqrt{1-\theta}-\sqrt{1-\theta t})}.$$

The limit of (2) when β tends to $-\infty$, p tends to 1, while keeping μ and δ fixed, is the set of pgf's of the Neyman A model, and the limit of (2), when β tends to 1, p tends to ∞, and hence δ tends to 1 while μ stays fixed, is the set of pgf's of the Poisson model.

2.3. **The zero truncated Tweedie-Poisson mixture model.** When modelling count data sometimes one faces instances where the zero value can not be observed. That is the case for example in the analysis of word or species frequency count data in linguistics and ecology considered in Section 5. There are also instances where one observes the zeros but one needs to model their frequency apart from the rest of the distribution through zero-inflated or zero-deflated models (see Johnson et al. 2005, pp. 351-357).

In particular, if one intends to use Tweedie-Poisson mixture models in any of these contexts one needs to resort to its zero truncated version, with probability mass function:

$$p_r^{ttp}(\alpha, \beta, \theta) = \frac{p_r^{tp}(\alpha, \beta, \theta)}{1 - p_0^{tp}(\alpha, \beta, \theta)} \text{ for } r = 1, 2, \ldots, \tag{4}$$

denoted here as the TTP model. From the definition of pgf it follows that the pgf of any strictly positive integer valued random variable has to take the value

zero at zero. Furthermore, the pgf's of (4) can be posed in terms of the pgf's of the untruncated model in (2) through:

$$G^{ttp}_{\alpha,\beta,\theta}(t) = \frac{G^{tp}_{\alpha,\beta,\theta}(t) - G^{tp}_{\alpha,\beta,\theta}(0)}{1 - G^{tp}_{\alpha,\beta,\theta}(0)} = \frac{e^{\frac{\alpha(1-\beta)}{\beta}((1-\theta)^\beta - (1-\theta t)^\beta)} - e^{\frac{\alpha(1-\beta)}{\beta}((1-\theta)^\beta - 1)}}{1 - e^{\frac{\alpha(1-\beta)}{\beta}((1-\theta)^\beta - 1)}}. \tag{5}$$

The parameter space of any truncated version of a distribution can always be made to be the same as the one for the untruncated distribution. Hence, it is reasonable to assume that the parameter space of (4) is the same as the one for the untruncated Tweedie-Poisson model and therefore that θ is in $(0, 1)$ and (α, β) is in $(0, \infty) \times (-\infty, 1)$. Nevertheless, it turns that these are not the only values for (α, β, θ) that make (5) into a pgf, and hence that one can extend the zero truncated Tweedie-Poisson model by enlarging its parameter space.

3. Extended truncated Tweedie-Poisson model

3.1. **Definition of the extended truncated Tweedie-Poisson model.** The function in (5) is the pgf of a positive integer random variable for any θ in $(0, 1)$ and any (α, β) in $(0, \infty) \times (-\infty, 1)$ because it is the pgf of a zero truncated Tweedie-Poisson mixture distribution. The next result, proved in the Appendix, states that when θ is in $(0, 1)$ and (α, β) is either in $\{0\} \times (-\infty, 1)$ or in $[-1, 0) \times [0, 1)$, then (5) is also the pgf of a random variable.

As a consequence, one will be able to use as a parameter space for the model defined through (5) a set that is larger than the parameter space for the untruncated Tweedie-Poisson model by either letting α be 0 or letting α be in $[-1, 0)$ whenever β is in $[0, 1)$.

Theorem 3.1. *Consider the real function*

$$G_{\alpha,\beta,\theta}(t) = \frac{e^{\frac{\alpha(1-\beta)}{\beta}((1-\theta)^\beta - (1-\theta t)^\beta)} - e^{\frac{\alpha(1-\beta)}{\beta}((1-\theta)^\beta - 1)}}{1 - e^{\frac{\alpha(1-\beta)}{\beta}((1-\theta)^\beta - 1)}}, \tag{6}$$

and let

$$G_{\alpha=0,\beta,\theta}(t) = \frac{(1-\theta t)^\beta - 1}{(1-\theta)^\beta - 1}, \tag{7}$$

$$G_{\alpha,\beta=0,\theta}(t) = \frac{(1-\theta t)^{-\alpha} - 1}{(1-\theta)^{-\alpha} - 1}, \tag{8}$$

and

$$G_{\alpha=0,\beta=0,\theta}(t) = \frac{\log(1-\theta t)}{\log(1-\theta)}, \tag{9}$$

which are the limits of (6) *when* α *or* β *or both tend to* 0. *If* $\theta \in (0, 1)$, *it holds that:*

(1) *when* $\beta \in [0, 1)$*, then* $G_{\alpha,\beta,\theta}(t)$ *is the pgf of a positive integer valued random variable if, and only if,* $\alpha \in [-1, \infty)$*, and*
(2) *when* $\beta \in (-\infty, 0)$*, then* $G_{\alpha,\beta,\theta}(t)$ *is the pgf of a positive integer valued random variable if, and only if,* $\alpha \in [0, \infty)$*.*

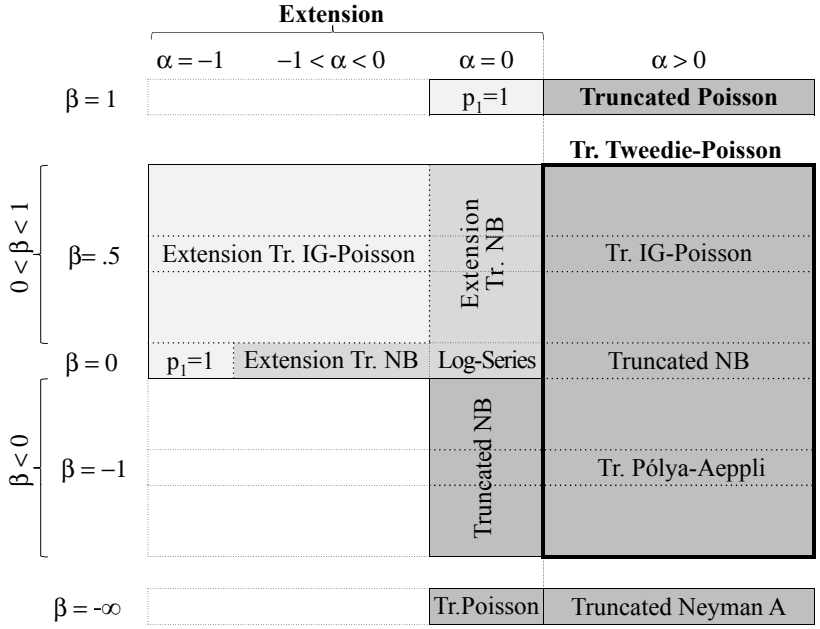

FIGURE 1. Map of the ETTP model and of two of its limiting models. The darkest grey shade indicates when the model is both the truncation of a Poisson mixture as well as a mixture of the truncated Poisson, the lightest grey indicates when it is neither one of these two kinds of models, and the medium grey when it is a mixture of the truncated Poisson but not the truncation of a Poisson mixture. The unshaded (α, β) combinations are not feasible.

That this result holds for $\beta = 0$ and for $\beta = .5$ had been documented in Engen (1974) and in Puig et al. (2009) respectively. Here the result is stated and proved in its full generality. Note that negative values of α lead to feasible probability models only after one truncates at zero the sample space of the Tweedie-Poisson mixture model, because under the untruncated model they would make the probability of zero, in (3), larger than 1.

Furthermore, in the truncated case one can neither extend the parameter space beyond θ in $(0,1)$ nor allow $\alpha < -1$ because that would make the second derivative of (6) at $t = 0$ in (12) (and hence $p_2^{ttp}(\alpha, \beta, \theta)$) negative. Hence we are lead to the next definition.

Definition 3.1. *The extended truncated Tweedie-Poisson model is the statistical model defined through the pgfs in* (6), (7), (8) *and* (9) *with* (α, β, θ) *either in* $[-1, \infty) \times [0, 1) \times (0, 1)$ *or in* $[0, \infty) \times (-\infty, 0) \times (0, 1)$. *It will be denoted as the ETTP model, its pgfs will be denoted by* $G^{ettp}_{\alpha,\beta,\theta}(t)$, *and its probability mass functions by* $p^{ettp}_r(\alpha, \beta, \theta)$.

Even though we do not provide the details, under the Poisson limiting case, when β tends to 1, and under the Neyman A limiting case, when β tends to $-\infty$, the zero truncated model can be extended to $\alpha = 0$ but not to negative α's. When $\alpha = 0$ the truncated Poisson limiting model becomes a degenerated one-point distribution concentrated at one, with $p_1 = 1$, and the zero truncated Neyman A limiting model becomes a zero truncated Poisson distribution.

Figure 1 maps the ETTP model and these two limiting models; it includes as particular submodels the extended truncated negative binomial model proposed in Engen (1974), obtained when $\beta = 0$ and denoted by ETNB, and the extended truncated inverse Gaussian-Poisson model proposed in Puig et al. (2009), obtained when $\beta = .5$. When $\alpha = 0$ and $\beta = 0$ the ETTP model becomes the logarithmic series model advocated for by Fisher et al. (1943).

The submodel that results from restricting α to be 0, at the boundary of the TTP model, coincides with the one that results from restricting β to be 0 and hence it is the ETNB model.

3.2. **When is the extended part of the ETTP model needed?** To explore the usefulness of the extended part of the ETTP model the next result, proved in the Appendix, describes how its probability at one and two and its first two moments relate to α. That will help understand what is gained by allowing α to take values in $[-1, 0]$.

Proposition 3.1. *Let* R *be a random variable with a distribution in the ETTP model. If one keeps* θ *and* β *fixed then:*

(1) $p^{ettp}_1(\alpha, \beta, \theta)$ *is decreasing with* α,
(2) $p^{ettp}_2(\alpha, \beta, \theta)/p^{ettp}_1(\alpha, \beta, \theta)$ *is increasing with* α,
(3) *its expectation,* $E^{ettp}_{\alpha,\beta,\theta}(R)$, *is increasing with* α, *and*
(4) *its dispersion index,* $V^{ettp}_{\alpha,\beta,\theta}(R)/E^{ettp}_{\alpha,\beta,\theta}(R)$, *is increasing with* α.

Figure 2 presents the contour plots for p^{ettp}_1, for p^{ettp}_2/p^{ettp}_1 and for the logarithm of the variance,

$$V^{ettp}_{\alpha,\beta,\theta}(R) = \frac{\alpha\,\theta(1-\beta)}{(1-\theta)^{2-\beta}\left(1-e^{\frac{\alpha(1-\beta)}{\beta}((1-\theta)^{\beta}-1)}\right)}\left(1-\theta\beta-\frac{\alpha\,\theta(1-\theta)^{\beta}(1-\beta)}{e^{\frac{\alpha(1-\beta)}{\beta}(1-(1-\theta)^{\beta})}-1}\right),$$

all for a fixed value of β equal to $.75, .5, .25, 0.$ and $-.5$, which correspond to p being equal to $5, 3, 2.33, 2$ and 1.66. That figure indicates that the variance is mainly determined by θ and it is increasing with θ, and that p_1^{ettp} is decreasing and p_2^{ettp}/p_1^{ettp} is increasing with θ. It can also be checked that if one keeps α and β fixed, the variance of R is mainly determined by θ and it is increasing with θ, and that p_1^{ettp} is decreasing and p_2^{ettp}/p_1^{ettp} is increasing with θ.

As a consequence of Proposition 3.1, when θ and β are fixed the expectation, the dispersion index and p_2/p_1 for the unextended TTP model all take their minimum value at $\alpha = 0$, and its probability at one takes its maximum at $\alpha = 0$, when it is equal to

$$p_1^{ettp}(\alpha = 0, \beta, \theta) = \frac{\beta\theta}{1 - (1-\theta)^{\beta}},$$

and:

$$p_1^{ettp}(\alpha = 0, \beta = 0, \theta) = \frac{-\theta}{\log(1-\theta)}.$$

When $\beta \geq 0$ and θ is near 1, which is most often the case when the data has a large dispersion, this maximum of p_1^{ttp} (under the unextended TTP model) is close to β and the minimum of p_2/p_1 is close to $(1-\beta)/2$. By extending the parameter space to include α in $[-1, 0]$, this statistical model gains flexibility by accommodating for count data sets with larger $\hat{p}_1$ and smaller $\hat{p}_2/\hat{p}_1$ than allowed by the unextended TTP model. By simultaneously decreasing α and increasing θ while keeping β fixed, the frequency of ones and the tail probabilities can simultaneously increase and lead to more dispersed distributions.

The maximum likelihood estimate of the mean of the ETTP model is the sample mean, because its pgf's can be posed as $G^{ettp}_{\alpha,\beta,\theta}(t) = h_{\alpha,\beta}(\theta t)/h_{\alpha,\beta}(\theta)$ for a given function $h_{\alpha,\beta}(\cdot)$ which is a sufficient condition for that property to hold (see Bondesson, 1997). Figure 3 presents the contour plots of the coefficient of variation, CV, of p_1, and of $(1-p_1)*CV$ for the ETTP model when one keeps its mean fixed. The contours of CV indicate that if the sample variability increases while keeping the sample mean fixed, either the estimate of α will decrease and eventually become negative or the estimate of β will increase towards 1; the smaller the mean the easier it will be to end up with negative estimates of α. Analogously, Figure 3 also indicates that if the probability of one is inflated while keeping the sample mean fixed, either the estimate of α will decrease and eventually become negative or the estimate of β will increase.

The quantity $(1-p_1)*CV$ captures the trade off between increasing the variability and inflating the probability of one. If the sample mean is fixed and the estimate of α is larger than 0, the larger this quantity the larger the estimate of β. Hence, if CV and p_1 increase but $(1-p_1)*CV$ decreases the estimate of β can not increase and hence the estimate of α will become negative, which will not happen if $(1-p_1)*CV$ increases.

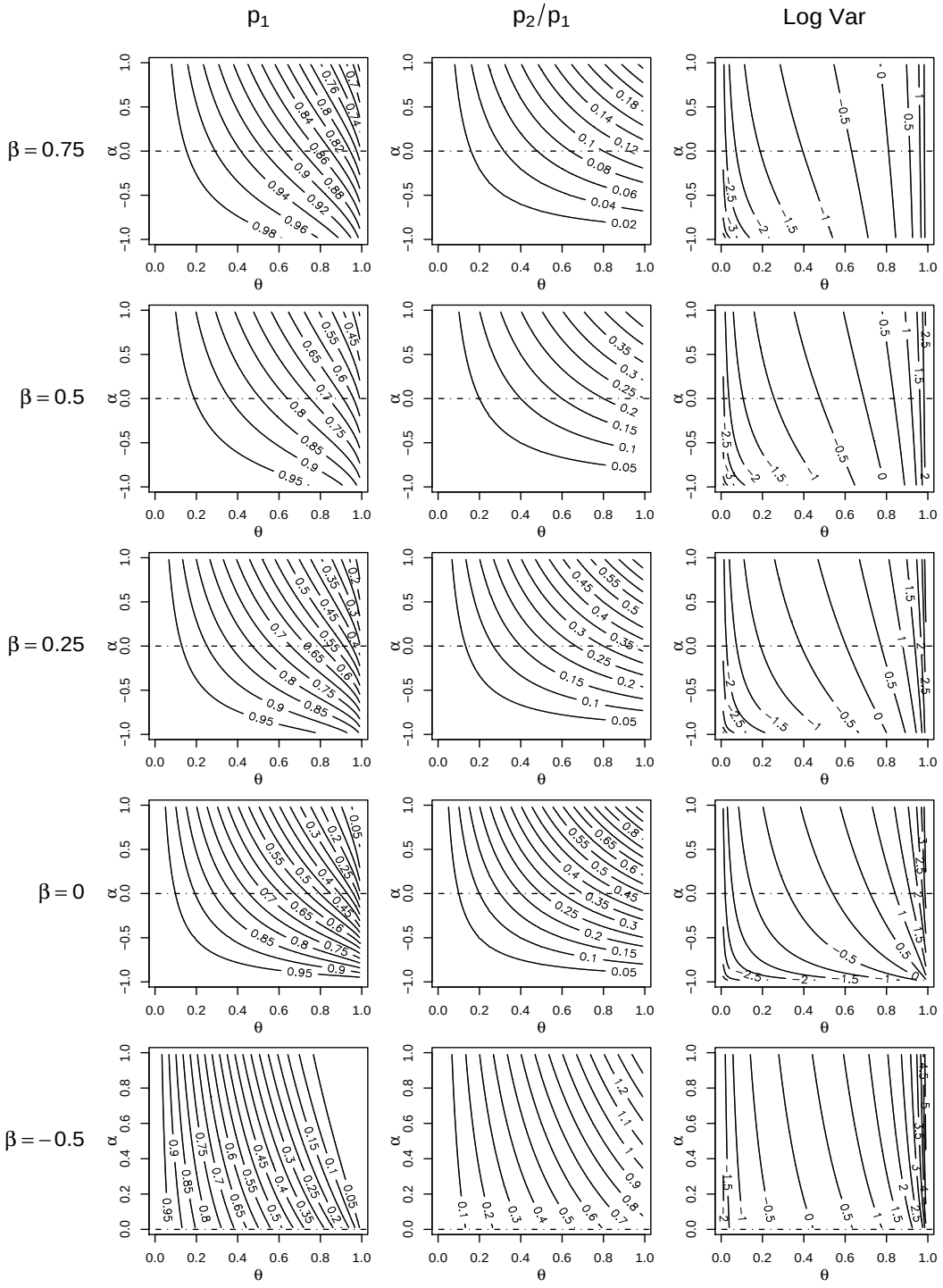

FIGURE 2. Contour plots of p_1^{ettp}, of p_2^{ettp}/p_1^{ettp} and of the base 10 logarithm of the variance of the ETTP model for five given values of β. The case $\beta = 0$ corresponds to the extended truncated negative binomial, and $\beta = .5$ to the extended truncated IG-Poisson.

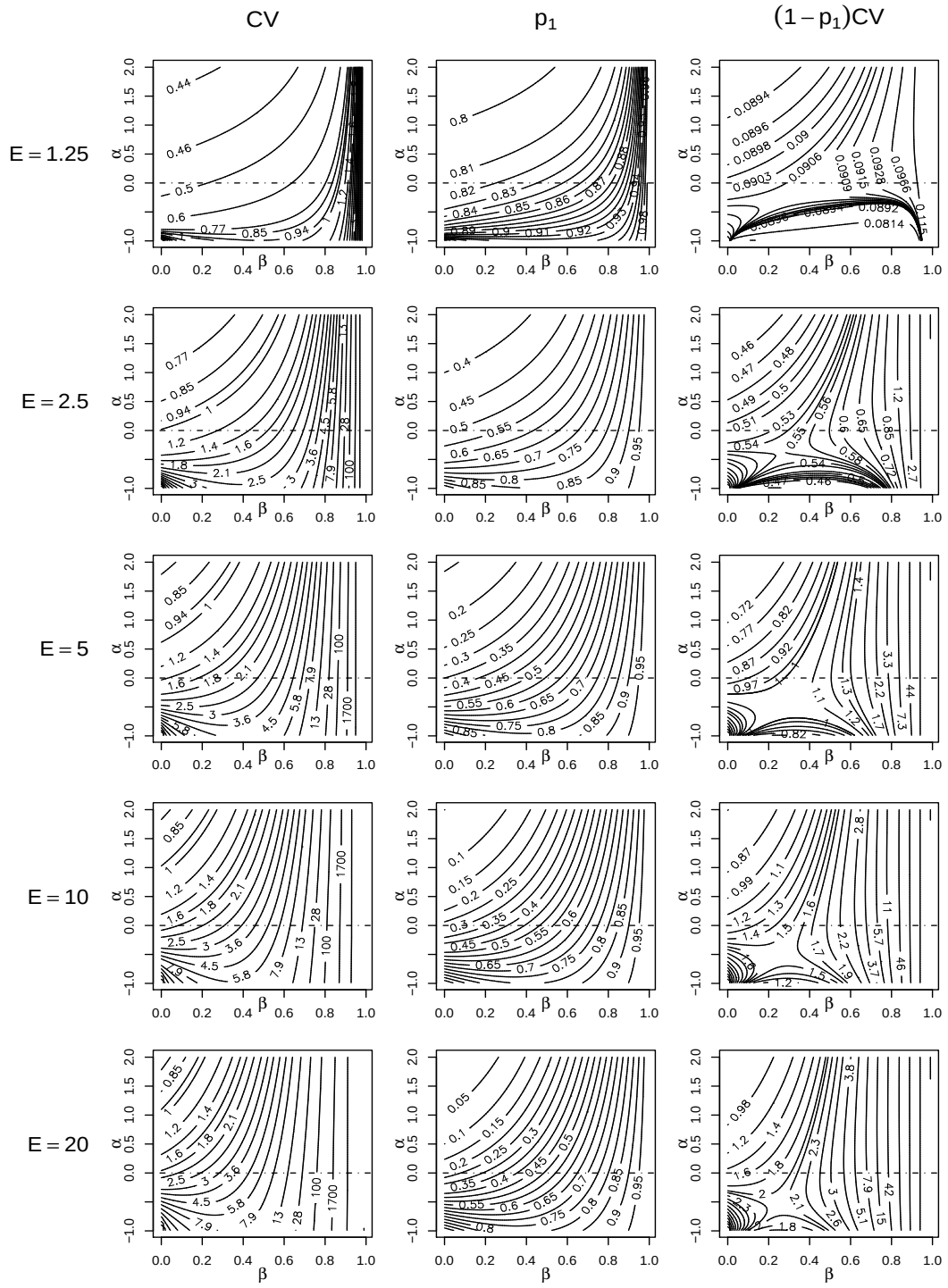

FIGURE 3. Contour plots of the coefficient of variation, CV, of p_1^{ettp} and of $(1 - p_1^{ettp})CV$ of the ETTP model for five given expected values, E. Note that the contour levels of CV and of p_1^{ettp} are set to be the same ones for all the five expectations considered.

4. Characterization of the ETTP model

When $\alpha > 0$ the ETTP model coincides with the TTP model and hence it is the truncation of Poisson mixture distributions. This is very useful when one is interested in estimating the mixing distribution or the probability of zero of the corresponding untruncated Poisson mixture model, like for example in the analysis of word or species frequency count data covered in Section 5. Hence, the importance of finding if the extended part of the model can also be posed as the truncation of Poison mixture distributions or not.

When $\alpha > 0$ the model is also a mixture of zero truncated Poisson distributions, as it is always the case for zero truncated Poisson mixtures (see Valero et al. 2009).

The next result, proved in the Appendix, states that when $\alpha < 0$ the ETTP model can not be obtained as the zero truncation of any Poisson mixture distribution and that, except for the special case obtained when $\beta = 0$, it can neither be obtained as a mixture of zero truncated Poisson distributions. The case $\alpha = 0$ coincides with the case $\beta = 0$.

Theorem 4.1. *Consider the ETTP model defined through the pgf's $G^{ettp}_{\alpha,\beta,\theta}(t)$ in (6), (7), (8) and (9) with $\theta \in (0,1)$. It holds that:*

1. *when $(\alpha,\beta) \in (0,\infty) \times (-\infty,1)$ the model is both the zero truncation of a Poisson mixture distribution as well as a mixture of zero truncated Poisson distributions,*
2. *when $(\alpha,\beta) \in (-1,0] \times \{0\}$ the model is a mixture of zero truncated Poisson distributions, but it is not the zero truncation of any Poisson mixture distribution,*
3. *when $(\alpha,\beta) \in [-1,0) \times (0,1)$ the model is neither the zero truncation of any Poisson mixture distribution, nor any mixture of zero truncated Poisson distributions,*
4. *when $(\alpha,\beta) \in \{0\} \times (-\infty,0)$ the model is both the zero truncation of a Poisson mixture distribution as well as a mixture of zero truncated Poisson distributions,*
5. *when $(\alpha,\beta) \in \{0\} \times [0,1)$ the model is a mixture of zero truncated Poisson distributions but it is not the zero truncation of any Poisson mixture distribution,*
6. *when $(\alpha,\beta) = (0,1)$ or $(\alpha,\beta) = (-1,0)$ the model is a degenerated distribution concentrated at one, with $p_1 = 1$.*

Note that the extended part of the ETTP submodel that results from imposing $\beta = 0$, when $\alpha \in [-1,0]$, which coincides with the submodel obtained with $\alpha = 0$ and $\beta \in [0,1]$, is an example of a model that can be obtained as a mixture of zero truncated Poisson distributions but not as the zero truncation of a mixture of Poisson distributions. That is in contrast with what happens under finite mixture models, because every model that results from truncating a finite mixture of

Poisson distributions can be interpreted as a finite mixture of truncated Poisson distributions and viceversa, as it is proved in Böhning and Kuhnert (2006).

The grey shades in Figure 1 capture the three possibilities considered in Theorem 4.1. As a consequence of this result, when (α, β) is in $[-1, 0] \times [0, 1)$ the ETTP model is not the zero truncation of a mixed Poisson distribution. Nevertheless, the extension will still be useful in that it will allow one to use maximum likelihood based tools when the m.l.e. under the unextended TTP model does not even exist, as exemplified next.

5. Use of the ETTP model on frequency count data

Typical word or species frequency count data has a reverse J-shaped distribution with an extraordinarily long upper tail. Modelling this type of data through Poisson mixture models is useful because the model mixing distribution can be interpreted as the distribution of the word or species frequency of the vocabulary of the author or the population. Hence estimates of their mixing density serve as estimates of the density of word or species frequencies and can be used as fingerprints of the style of the author or the ecosystem. That allows one to use the inverse of the expectation of the estimated mixing distribution to estimate the size (number of different words or species) of the vocabulary or population, and to use measures of the variability of the estimated mixing distribution as estimates of measures of their lack of diversity. For a discussion of the use of Poisson mixture models in the estimation of the frequency distribution, size and diversity of a population, see Ginebra and Puig (2010).

Given that the size of the vocabulary of an author and the total number of species are usually unknown, typically one can not count unobserved words or species and to model frequency count data one often needs to resort to the zero truncated version of Poisson mixture models. Engen (1974), Ord and Whitmore (1986), Holmes (1992) and Baayen (2001), among many others, fit this type of data through the truncated versions of either the negative binomial or the IG-Poisson model. Sichel (1975, 97) and Puig et al. (2010), following Good (1953), use a three parameter truncated generalized inverse gaussian-Poisson mixture model (GIG-Poisson), which includes the negative binomial and IG-Poisson models as special cases.

A natural alternative to the GIG-Poisson model is the Tweedie-Poisson model. Both the GIG and Tweedie mixing distributions serve as good approximations to the words or species frequencies distributions in typical vocabularies or populations, and the corresponding Poisson mixture models fit frequency count data extremely well. El-Shaarawi et al. (2010) is a first example of the use of the untruncated Tweedie-Poisson mixture model on species frequency count data where zeros are observed. When the unextended TTP model is used on zero truncated data, one can use:

$$size = \frac{N}{\mu} = \frac{N(1-\theta)^{1-\beta}}{\alpha\theta(1-\beta)}, \tag{10}$$

to estimate the total number of different words or species in the vocabulary of the author or population. Next, the usefulness of the extended truncated Tweedie-Poisson model is explored by fitting it to 456 sets of word frequency data and to 89 sets of species frequency data.

5.1. **Use of the ETTP model on words frequency count data.** Here the ETTP model is fitted to the word frequency count data of 456 chapters of *Tirant lo Blanc* presented in Table 1 and described in Riba and Ginebra (2006). The maximum likelihood estimate of α is negative for 84 out of the 456 chapters considered, and hence for them the ETTP model improves the fit of the TTP model that restricts α to be positive. Note that for these chapters the m.l.e. of α under the TTP model do not exist because the closer $\hat{\alpha}$ is to zero the larger its likelihood function and $\alpha = 0$ is not in the parameter space of the TTP model.

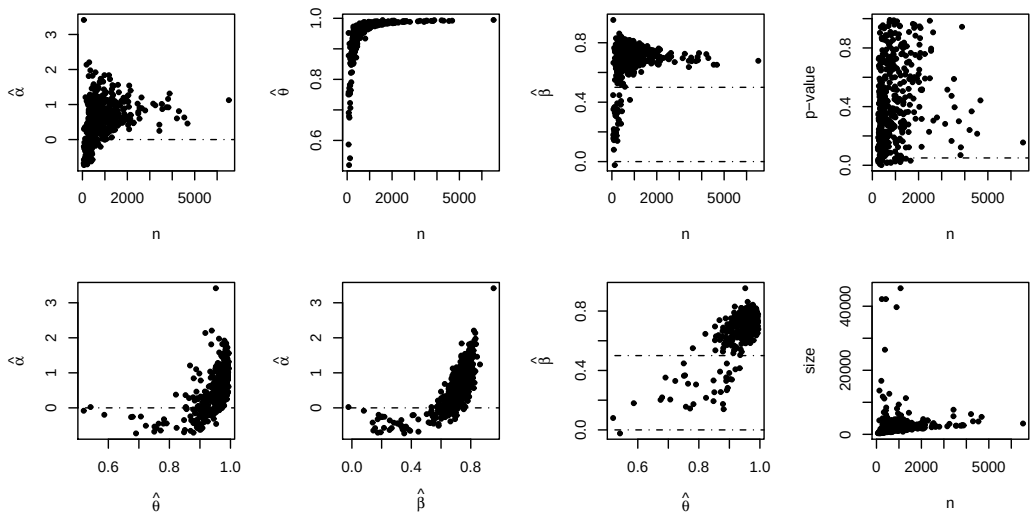

FIGURE 4. Maximum likelihood estimates of (α, β, θ) and of the size of vocabulary in (10), and χ^2 goodness of fit test p-value as a function of chapter size, n, and relationship between parameter estimates, all for the word frequency count data of the 456 chapters of *Tirant lo Blanc* with more than 40 different words.

Figure 4 shows that the chapters benefiting from the extended part of the model, with $\hat{\alpha} < 0$, tend to be the short ones. The χ^2 goodness of fit test p-values presented as a function of n indicate that the ETTP model fits this word frequency count data very well over all the range of n, even though some of the chapters are quite long.

Theorem 4.1 states that when (α, β) is in $[-1, 0] \times [0, 1)$ the ETTP model is not a zero truncated mixed Poisson distribution (see also Figure 1). Nevertheless, note that even when $\hat{\alpha}$ is negative, the true value of α could still be positive.

In fact, for all the 84 chapters where $\hat{\alpha}$ is negative but one, the (classical) 95% confidence intervals for α include positive values and therefore their word frequency counts are still compatible with Tweedie-Poisson mixture distributions. As a consequence, even when $\hat{\alpha}$ is negative, one can use the upper end of its confidence interval to provide confidence intervals for features associated with the Tweedie model mixing distribution, like for the number of different words or species estimated through (10).

Figure 4 also shows that, other than for very short chapters, most of the estimates of θ are close to 1, in line with the fact that the larger the text the more likely it is that words are repeated, the larger the variance of the distribution, and according to Figure 2 the larger θ. It is our experience that when $\hat{\alpha}$ becomes negative $\hat{\theta}$ increases relative to the estimate obtained under the unextended TTP model, which is another indication that the extension of the model is most useful when data is more overdispersed than allowed by the unextended model.

There is only one instance where $\hat{\beta}$ is negative, and most of its estimates are concentrated around $\hat{\beta} = .7$ which correspond to a mixture model with a mixing distribution with a mean-variance relationship of power $\hat{p}$ close to 5. Remember that the inverse Gaussian distribution has $p = 3$ and the negative binomial distribution has $p = 2$.

When one restricts β to be .5 and fits the extended IG-Poisson model proposed in Puig et al. (2009), 424 out of the 456 chapters have negative $\hat{\alpha}$. When one restricts β to be 0. and fits the ETNB model proposed in Engen (1974), all but one of the 456 chapters have negative $\hat{\alpha}$. The fits with the extended Pólya-Aepply model, with $\beta = -1$, are considerably worse, and 233 chapters lead to the extended part with $\hat{\alpha} = 0$, while 223 chapters have $\hat{\alpha} > 0$.

Figure 5 compares the performance of the three parameter ETTP model with the two parameter extended truncated IG-Poisson and ETNB submodels by plotting the corresponding log-likelihood ratios as a function of n. Applying the likelihood ratio tests with a .05 level of significance, one rejects the extended truncated IG-Poisson submodel in favor of the ETTP model in 45% of the chapters, and one rejects the ETNB submodel in 28% of the chapters.

Figure 5 also compares the ETTP model with the model at the boundary of the unextended TTP model, with $\alpha = 0$, in the 84 chapters where the fit under extended and unextended models differ. Under the corresponding likelihood ratio test one would only (barely) reject the $\alpha = 0$ submodel in one of these chapters. Hence, the usefulness of the ETTP model is not in that it significantly improves the fits that can be attained through the unextended TTP model. Its main advantage is in that by extending the parameter space one is allowed to compute the usual likelihood based confidence intervals and hypothesis tests for α. Remember that under the unextended TTP model the m.l.e. of α for these 84 chapters does not even exist, and it is not clear how one would construct confidence intervals for α under that model.

	$v_{1:n}$	$v_{2:n}$	$v_{3:n}$	$v_{4:n}$	$v_{5:n}$	$v_{6:n}$	$v_{7:n}$	$v_{8:n}$	$v_{9:n}$	$v_{10:n}$	$v_{11:n}$	...	n
Chapter 1	107	16	6	2	2	2	2	1	1	1	0	...	255
Chapter 2	172	26	19	7	2	2	2	2	1	1	1	...	476
Chapter 3	299	70	32	16	10	5	4	2	5	1	2	...	1174
...	...	...	...	...	...	...	...	...	...	...	...	...	...
Chapter 487	129	29	10	6	1	1	0	2	2	2	0	...	348

TABLE 1. Part of the word frequency counts in the 456 chapters of *Tirant lo Blanc* with more than 40 different words. By $v_{r:n}$ one denotes the number of words that are repeated exactly r times in a text with a total of n word counts.

5.2. **Use of the ETTP model on species frequency count data.** Here the ETTP model is fitted to the 89 samples of beetles and bugs frequency count data in Janzen (1973) with a total of more than $n = 40$ counts, partially presented in Table 2 and partially analysed in Huillet and Paroissin (2009). Figure 6 presents the m.l.e. of (α, β, θ) as a function of n. The estimates of α are negative for 45 out of these 89 samples, and hence here the ETTP model improves the fit of the TTP model in more than half the samples. The estimates of θ are even closer to one than for the previous word frequency examples, indicating that these species frequency data sets are even more overdispersed. The estimates of β and of p tend to be a bit smaller.

If one fits the extended truncated IG-Poisson model, with $\beta = .5$, 25 samples have negative $\hat{\alpha}$, and if one fits the ETNB model, with $\beta = 0.$, 82 samples have negative $\hat{\alpha}$. According to the likelihood ratio test with a .05 level of significance, one would only reject these two submodels in favor of the ETTP model for 2 out of the 89 samples, in line with the sample sizes here being considerably smaller than in the word frequency examples. On the other hand, the ETTP model fits 87 of the samples significantly better than the truncated Neyman A model, and it fits 83 of them significantly better than the truncated Pólya-Aepply model.

When one compares the fit of the ETTP model with the fit of the submodel with $\alpha = 0$ on the 45 samples with negative $\hat{\alpha}$, one does not reject the $\alpha = 0$ submodel for any sample. That indicates again that the fit of the ETTP model is not significantly better than fits attained through the TTP model, but unless one resorts to the ETTP model, the m.l.e. and the methods based on them would not be available for these 45 samples.

6. Final comments

It has been proved that by truncating the sample space of the Tweedie-Poisson mixture model at zero one is entitled to use a parameter space that is larger than the one for the un-truncated Tweedie-Poisson model. This phenomena had been

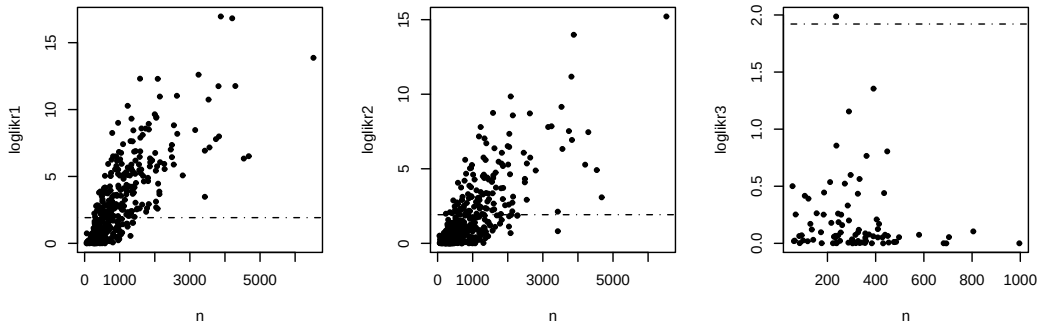

FIGURE 5. Logarithm of the ratios between the maximum of the likelihood of the ETTP model and the one of the extended truncated IG-Poisson (lglkr1), and the one of the ETNB (lglkr2), and the one of the boundary model (with $\alpha = 0$) of the unextended TTP model (lglkr3), as a function of n for the chapters of *Tirant lo Blanc*. For the third ratio only the chapters where the extended model improves the unextended one are included. The horizontal line indicates the likelihood ratio test level .05 critical region.

	$v_{1:n}$	$v_{2:n}$	$v_{3:n}$	$v_{4:n}$	$v_{5:n}$	$v_{6:n}$	$v_{7:n}$	$v_{8:n}$	$v_{9:n}$	$v_{10:n}$	$v_{11:n}$	...	n
Osa secondary	70	17	4	5	5	5	5	3	1	2	3	...	996
Osa primary flat	38	6	2	2	1	0	0	0	0	0	0	...	110
Osa primary hill	59	9	3	2	2	2	0	0	0	0	1	...	127
...	...	...	...	...	...	...	...	...	...	...	...	...	...
St. Johns	7	1	1	2	1	0	0	0	0	0	0	...	51

TABLE 2. Part of the 89 samples of beetles frequency counts in Janzen (1973) with a total of more than 40 beetles. By $v_{r:n}$ one denotes the number beetles species repeated exactly r times in a sample with a total of n beetle counts.

documented for the negative binomial and the IG-Poisson, even though a rigorous proof was lacking even for these two submodels.

The kind of extension described is analogous to the one reported in Griffiths (1973) for the truncated beta binomial model. It is not unlikely that similar

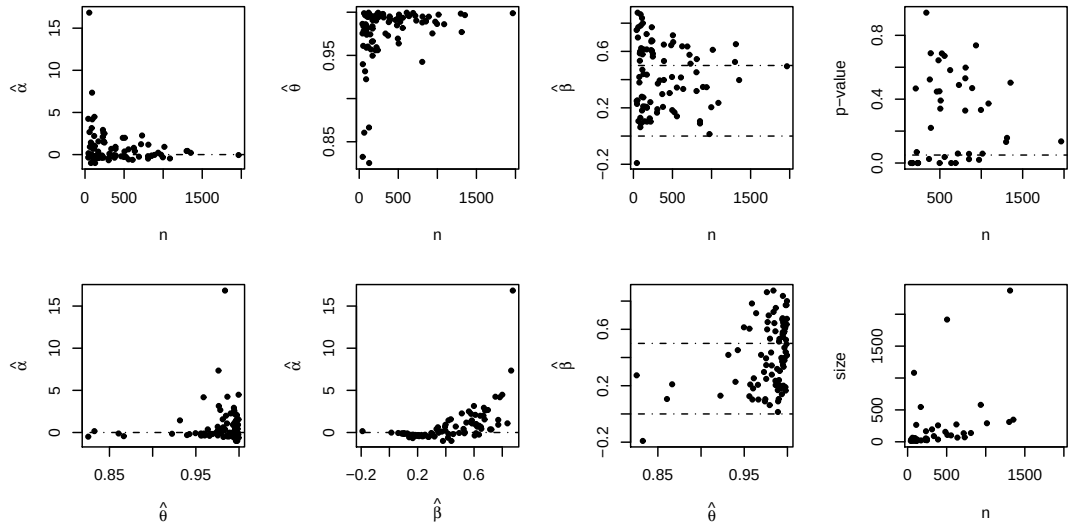

FIGURE 6. Maximum likelihood estimates of (α, β, θ) and of the size in (10) and χ^2 goodness of fit test p-value as a function of the beetle count total, n, and relation between parameter estimates, all for the 89 samples of beetles and bugs frequency count data of Janzen (1973) with a total of more than 40 counts.

phenomena occur for other zero truncated Poisson mixture models or the Hinde-Demetrio model in Kokonendji et al. (2007) and Kokonendji and Malouche (2008), and for multinomial mixture models. It is also natural to wonder what happens when one further truncates the Tweedie-Poisson model to exclude 0 and 1 or the tail of the distribution, even though these situations are less frequent in practice.

APPENDIX: PROOFS

Proof of Theorem 3.1. In order to prove this result one needs to use the fact that a real function $G(t)$ is a pgf of a non-negative integer valued random variable if, and only if, it verifies that:

a) $G(1) = 1$, that
b) it is analytical in a neighborhood of any $t_0 \in [0, 1)$, and hence that it can be represented by a power series at any such t_0, and that
c) all the coefficients of its Taylor series expansion at $t = 0$ are non-negative, and hence that $G(0)$ and all its derivatives at $t = 0$ are non-negative.

When a function $G(t)$ satisfies b) and c), then all the coefficients of its series expansion at any $t_0 \in (0, 1)$ are also non-negative. Furthermore, $G(t)$ is a pgf

of a strictly positive integer valued random variable if, and only if, on top of conditions a), b) and c) it also holds that $G(0) = 0$.

Proof of Theorem 3.1: The function $G_{\alpha,\beta,\theta}(t)$ is analytical at any $t \neq 1/\theta$ and hence, it is analytical at any t in $[0, 1)$ for all the range of (α, β, θ) considered. Furthermore, $G_{\alpha,\beta,\theta}(1) = 1$ and $G_{\alpha,\beta,\theta}(0) = 0$. Hence, to complete the proof one only needs to determine the values of α under which $G_{\alpha,\beta,\theta}(t)$ satisfies the condition c) above, when $\theta \in (0, 1)$ and $\beta \in (-\infty, 1)$.

The first derivative of (6) evaluated at $t = 0$ is:

$$G'_{\alpha,\beta,\theta}(0) = \frac{\alpha\theta(1-\beta)}{e^{\frac{\alpha(1-\beta)}{\beta}(1-(1-\theta)^{\beta})} - 1}, \tag{11}$$

and therefore $G'_{\alpha,\beta,\theta}(0)$ is non-negative for any α. Furthermore,

$$\frac{G''_{\alpha,\beta,\theta}(0)}{G'_{\alpha,\beta,\theta}(0)} = \theta(1-\beta)(\alpha+1), \tag{12}$$

and therefore $G''_{\alpha,\beta,\theta}(0)$ is non-negative when $\alpha \geq -1$. To check the sign of its higher derivatives at $t = 0$, the cases β in $[0, 1)$ and $\beta < 0$ will be considered apart.

1) Case $\beta \in [0, 1)$: To prove that the rest of derivatives of $G_{\alpha,\beta,\theta}(t)$ at $t = 0$ are non-negative for $\alpha \geq -1$, one needs to use the fact that if all the derivatives of the logarithm of a function at a given point are non-negative, then all the derivatives of the original function at that same point are also non-negative. Hence, let us check the sign of the derivatives of $h_{\alpha,\beta,\theta}(t) = \log G'_{\alpha,\beta,\theta}(t)$ at $t = 0$. The first derivative of $h_{\alpha,\beta,\theta}(t)$ is

$$h'_{\alpha,\beta,\theta}(t) = \frac{G''_{\alpha,\beta,\theta}(t)}{G'_{\alpha,\beta,\theta}(t)} = \theta(1-\beta)\Big(\frac{1}{1-\theta t} + \alpha(1-\theta t)^{\beta-1}\Big),$$

and hence $h'_{\alpha,\beta,\theta}(0)$ is equal to (12), and the m-th derivative of $h_{\alpha,\beta,\theta}(t)$ for $m \geq 2$ is

$$h^{(m)}_{\alpha,\beta,\theta}(t) = \theta^m(1-\beta)\Big(\frac{(m-1)!}{(1-\theta t)^m} + \alpha(\Pi_{i=1}^{m-1}(i-\beta))(1-\theta t)^{\beta-m}\Big),$$

and hence the m-th derivative of $h_{\alpha,\beta,\theta}(t)$ evaluated at $t = 0$ for $m \geq 2$ is:

$$h^{(m)}_{\alpha,\beta,\theta}(0) = \theta^m(1-\beta)((m-1)! + \alpha\Pi_{i=1}^{m-1}(i-\beta)). \tag{13}$$

Therefore, when $\beta \in [0, 1)$ and $\alpha \geq -1$ all the derivatives of $h_{\alpha,\beta,\theta}(t)$ at $t = 0$ are non-negative, and so are all the derivatives of $G'_{\alpha,\beta,\theta}(t)$ at $t = 0$, and as a consequence, $G_{\alpha,\beta,\theta}(t)$ satisfies all the conditions to be the pgf of a positive integer valued random variable.

2) Case $\beta \in (-\infty, 0)$: Note that (13) being negative when $\alpha < 0$ does not prove the result here, because the logarithm of a function having derivatives at

0 that are negative does not imply that the derivatives of that function at 0 are not all positive. Instead, note that:

$$G''_{\alpha,\beta,\theta}(t) = \frac{\alpha(1-\beta)^2\theta^2}{1-e^{\frac{\alpha(1-\beta)}{\beta}((1-\theta)^\beta-1)}} \frac{1+\alpha(1-\theta t)^\beta}{(1-\theta t)^{2-\beta}} e^{\frac{\alpha(1-\beta)}{\beta}((1-\theta)^\beta-(1-\theta t)^\beta)}, \tag{14}$$

and hence that $G''_{\alpha,\beta,\theta}(t_0) = 0$ for

$$t_0 = \frac{1-(-\alpha)^{-1/\beta}}{\theta}.$$

If $G_{\alpha,\beta,\theta}(t)$ was a pgf and α was in $(-1,0)$ then $G''_{\alpha,\beta,\theta}(t_0) = 2p_2 + 6p_3 t_0 + \ldots = 0$ for some t_0 in $(0, 1/\theta)$, and $G_{\alpha,\beta,\theta}(t)$ would have to be the pgf of a one-point distribution concentrated at one, which is not the case. Hence when β is negative, $G_{\alpha,\beta,\theta}(t)$ is a pgf if, and only if, $\alpha \geq 0$. □

Proof of Proposition 3.1. Note that the term

$$a = \alpha \frac{1-\beta}{\beta}(1-(1-\theta)^\beta)$$

is equal to α times a factor that is always positive and hence it is increasing in α. Given that the probability of one, in (11), can be written as:

$$p_1^{ettp}(\alpha,\beta,\theta) = \frac{a}{(e^a-1)} \frac{\beta\theta}{(1-(1-\theta)^\beta)},$$

which is a decreasing function of a times a positive term, it is also decreasing with α. Furthermore, it follows from (12) that

$$\frac{p_2^{ettp}(\alpha,\beta,\theta)}{p_1^{ettp}(\alpha,\beta,\theta)} = \frac{1}{2}\theta(1-\beta)(\alpha+1),$$

which is increasing in α. The expectation of R is the derivative of $G^{ettp}_{\alpha,\beta,\theta}(t)$ at $t=1$,

$$E^{ettp}_{\alpha,\beta,\theta}(R) = \frac{a}{1-e^{-a}} \frac{\beta\theta(1-\theta)^{\beta-1}}{1-(1-\theta)^\beta},$$

which is an increasing function of a times a positive constant and hence it is increasing with α. Finally the second derivative of $G^{ettp}_{\alpha,\beta,\theta}(t)$ at $t=1$ is the second factorial moment of R, and

$$\frac{V^{ettp}_{\alpha,\beta,\theta}(R)}{E^{ettp}_{\alpha,\beta,\theta}(R)} = \frac{-a}{(e^a-1)} \frac{\beta\theta(1-\theta)^{\beta-1}}{1-(1-\theta)^\beta} + 1 + (1-\beta)\frac{\theta}{1-\theta},$$

which is an increasing function of a times a positive constant plus a constant term, and hence it is increasing with α.

Proof of Theorem 4.1. To prove this result we use the following characterization of zero-truncated mixed Poisson distributions and of mixtures of zero-truncated Poisson distributions in terms of their pgf's due to Valero et al. (2010). A function $G(t)$ is the pgf of a mixture of zero truncated Poisson distributions with finite mean if, and only if,

a) $G(1) = 1$ and $G(0) = 0$,
b) it is analytical in $(-\infty, 1)$, and
c) all the coefficients of the series expansion of $G(t)$ at any $t_0 \in (-\infty, 1)$ are strictly positive except maybe the constant term.

A function $G(t)$ is the pgf of the zero truncation of a Poisson mixture with finite mean if, and only if, it satisfies conditions a), b) and c) and if its limit when t tends to $-\infty$ is finite.

We will also use the fact that all the coefficients of the series expansion at t_0 in $(-\infty, 1)$ for the pgf of an untruncated mixed Poisson distribution are positive (see Puri and Goldie, 1979), and as a consequence the limit of such a pgf when t tends to $-\infty$ is larger than or equal to zero.

Proof of Theorem 4.1: As already argued, in the range of (α, β, θ) considered $G^{ettp}_{\alpha,\beta,\theta}(1) = 1$, $G^{ettp}_{\alpha,\beta,\theta}(0) = 0$ and $G^{ettp}_{\alpha,\beta,\theta}(t)$ is analytical at any t in $(-\infty, 1)$. Hence, to prove the theorem one only needs to check the sign of the derivatives of $G^{ettp}_{\alpha,\beta,\theta}(t)$ in $(-\infty, 1)$, and find the limit when t tends to $-\infty$, which will be done case by case, keeping in mind that θ is always in $(0, 1)$.

1) Case $(\alpha, \beta) \in (0, \infty) \times (-\infty, 1)$: That (6) is the pgf of a zero truncated Poisson mixture follows from the way this function is constructed in Section 2. That it is also a mixture of zero truncated Poisson distributions follows from all zero truncated Poisson mixture distributions being mixtures of zero truncated Poisson distributions; in particular (4) can be written as:

$$p^{ttp}_r(\alpha, \beta, \theta) = \int_{R^+} \frac{\lambda^r e^{-\lambda}}{r!(1 - e^{-\lambda})} \Psi_{\alpha,\beta,\theta}(\lambda) d\lambda,$$

where:

$$\Psi_{\alpha,\beta,\theta}(\lambda) = \frac{1 - e^{-\lambda}}{1 - \int_{R^+} e^{-\lambda} \psi_{\alpha,\beta,\theta}(\lambda) d\lambda} \psi_{\alpha,\beta,\theta}(\lambda).$$

2) Case $\alpha \in (-1, 0]$ and $\beta = 0$: The first derivative of $G^{ettp}_{\alpha,\beta=0,\theta}(t)$ is

$$G'_{\alpha,\beta=0,\theta}(t) = \frac{\alpha\theta}{(1 - \theta) - (1 - \theta)^{\alpha+1}} \Big(\frac{1 - \theta}{1 - \theta t}\Big)^{\alpha+1}, \tag{15}$$

where the first term is a positive constant, and the second term is the pgf of an untruncated negative binomial$(\alpha + 1, \theta)$, which is a mixed Poisson distribution. Hence, when $\alpha > -1$ all the coefficients of the series expansion of (15) at any t_0 in $(-\infty, 1)$ are positive, and all the coefficients of the series expansion of $G^{ettp}_{\alpha,\beta=0,\theta}(t)$ are also positive, except maybe its constant term, which proves that it is the pgf of a mixture of zero truncated Poisson distributions.

When $\alpha > 0$, $G^{ettp}_{\alpha,\beta=0,\theta}(t)$ is also the pgf of the zero truncation of a (gamma-)Poisson mixture, as argued under case 1), but when $\alpha \in (-1, 0)$ it is not the truncation of any Poisson mixture distribution because the limit of (8) when t goes to $-\infty$ is $-\infty$. The result when $\alpha = 0$ and $\beta = 0$ follows from the limit of (9) when t tends to $-\infty$ being $-\infty$ coupled with the fact that

$$G^{(n)}_{\alpha=0,\beta=0,\theta}(t) = \frac{-(n-1)\,\theta^n}{\log(1-\theta)} \frac{1}{(1-\theta t)^n},$$

which is positive for any t_0 in $(-\infty, 1)$.

3) Case $(\alpha, \beta) \in [-1, 0) \times (0, 1)$: The second derivative of $G^{ettp}_{\alpha,\beta,\theta}(t)$ is (14), which is zero for

$$t_0 = \frac{1 - (-\alpha)^{-1/\beta}}{\theta},$$

with $t_0 < 0$. Hence, $G^{ettp}_{\alpha,\beta,\theta}(t)$ can neither be the pgf of the zero truncation of a Poisson mixture nor the pgf of a mixture of zero truncated Poisson distributions because all the derivatives of these distributions at any such t_0 have to be strictly positive.

4) Case $\alpha = 0$ and $\beta \in (-\infty, 0)$. The result follows from (7) with $\beta < 0$ being equivalent to (8) with $\alpha > 0$ and the result in case 1.

5) Case $\alpha = 0$ and $\beta \in [0, 1)$. The result follows from (7) with β in $[0, 1)$ being equivalent to (8) with α in $(-1, 0]$ and the result in case 2.

6) Case $(\alpha, \beta) = (0, 1)$ or $(-1, 0)$. It follows from $p_1 = G'_{\alpha=0,\beta=1,\theta}(0) = G'_{\alpha=-1,\beta=0,\theta}(0) = 1$. □

Acknowledgements: This work was funded in part by grants No. TIN2009-14560-C03-03, MTM2006-09920 and MTM2010-14887 of the Ministerio de Ciencia y Tecnologia of Spain and by Grant No. SGR-1187 of the Generalitat de Catalunya. The authors are grateful to Xavier Puig for his comments and help in preparing the figures. M. Pérez-Casany wishes to thank the Centre de Recerca Matemàtica, Bellaterra, Spain, for hosting her during part of her sabbatical in 2010.

References

Aalen OO. 1988. Heterogeneity in survival analysis. *Statistics in Medicine* **7**: 1121–1137.

Aalen OO. 1992. Modelling heterogeneity in survival analysis by the compound Poisson distribution. *The Annals of Applied Probability* **4**: 951–972.

Baayen H. 2001. Word frequency distributions. Dordretch: Kluwer.

Bar-Lev SK, Enis P. 1986. Reproducibility and natural exponential families with power variance functions. *The Annals of Statistics* **4**: 1507–1522.

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

Bondesson L. 1997. A generalization of Poincare's characterization of exponential families. *Journal of Statistical Planning and Inference* **63**: 147–155.

Davidian M. 1990. Estimation of variance functions in assays with possible unequal replication and nonnormal data. *Biometrika* **77**: 43–54.
Dunn PK, Smyth GK. 2008. Evaluation of Tweedie exponential dispersion model densities by Fourier inversion. *Statistical Computing* **18**: 73–86.
El-Shaarawi AH, Zhu R, Joe H. 2010. Modelling species abundance using the Poisson-Tweedie family. *Environmetrics*. DOI 10.1002/env.1036
Engen S. 1974. On species frequency models. *Biometrika* **61**: 263–270.
Engen S. 1978. Stochastic Abundance Models. London: Chapman Hall.
Fisher RJ, Corbet AS, Williams CB. 1943. The relation between the number of species and the number of individuals in a random sample from an animal population. *Journal of Animal Ecology* **12**: 42–58.
Gerber HU. 1991. From the generalized gamma to the generalized negative binomial distribution. *Insurance: Mathematics and Economics* **10**: 303–309.
Ginebra J, Puig X. 2010. On the measure and the estimation of evenness and diversity. *Computational Statistics and Data Analysis* **54**: 2187–2201.
Good IJ. 1953. The population frequencies of species and the estimation of population parameters. *Biometrika* **40**: 237–264.
Griffiths DA. 1973. Maximum likelihood estimation for the beta-binomial distribuiton and an application to the household distribution of the total of cases of a disease. *Biometrics* **29**: 637–648.
Hofmann M. 1955. Uber zusammengesetzte Poisson-Prozesse und ihre Anwendungen in der Unfallvewrsicherung. *Bulletin of the Swiss Association of Actuaries*, 499–575.
Holla MS. 1966. On a Poisson-inverse Gaussian distribution. *Metrika* **11**: 115–121.
Holmes DI. 1992. A stylometric analysis of mormon scripture and related texts. *Journal of the Royal Statistical Society, Series A* **155**: 91–120.
Hougaard P. 1986a. Survival models for heterogeneous populations derived from stable distributions. *Biometrika* **73**: 387–396.
Hougaard P. 1986b. A class of multivariate failure time distributions. *Biometrika* **73**: 671–678.
Hougaard P. 1987. Modelling multivariate survival. *Scandinavian Journal of Statistics* **14**: 291–304.
Hougaard P, Harvald B, Holm NV. 1992. Measuring the similarities between the lifetimes of adult Danish twins born between 1881-1930. *Journal of the American Statistical Association* **87**: 17–24.
Hougaard P, Lee M-LT, Whitmore GA. 1997. Analysis of overdispersed count data by mixtures of Poisson variables and Poisson processes. *Biomerics* **53**: 1225–1238.
Huillet T, Paroissin C. 2009. Sampling from Dirichlet partitions: estimating the number of species. *Environmetrics*. DOI: 10.1002/env.977
Janzen DH. 1973. Sweet samples of tropical foliage insects: Description of study sites, with data on species abundances and size distributions. Ecology **54**: 659–686.
Johnson NL, Kemp AW, Kotz S. 2005. Univariate Discrete Distributions, 3rd Ed. New York: Wiley.
Jorgensen B. 1987. Exponential dispersion models. *Journal of the Royal Statistical Society, Ser B* **49**: 127–162.
Jorgensen B. 1997. The Theory of Dispersion Models. London: Chapman Hall.
Katti SK, Gurland J. 1961. The Poisson Pascal distribution, *Biometrics* **17**: 527–538.
Kokonendji CC, Dossou-Gbete S, Demétrio CGB. 2004. Some discrete exponential dispersion models: Poisson-Tweedie and Hinde-Demetrio classes. *SORT* **28**: 201–214.
Kokonendji CC, Demétrio DGB, Zocchi SS. 2007. On Hinde-Demétrio regression models for overdispersed count data. *Statistical Methodology* **4**: 271–291.

Kokonendji CC, Malouche D. 2008. A property of count distributions in the Hinde-Demétrio family. *Communications in Statistics. Theory and Methods* **37**: 1823–1834.
Ord JK, Whitmore G. 1986. The Poisson-inverse Gaussian distribution as a model for species abundance. *Communications in Statistics, Theory and Methods* **15**: 853–871.
Puig P, Valero J. 2006. Count data distributions: Some characterizations with applications. *Journal of the American Statistical Association* **101**: 332–340.
Puig X, Ginebra J, Font M. 2010. The Sichel model and the mixing and truncation order. To appear in the *Journal of Applied Statistics*.
Puig X, Ginebra J, Pérez-Casany M. 2009. Extended truncated inverse Gaussian-Poisson model. *Statistical Modelling* **9**: 151–171.
Puri PS, Goldie CM. 1979. Poisson mixtures and quasi-infinite divisibility of distributions. *Journal of Applied Probability* **16**: 138–153.
Riba A, Ginebra J. 2006. Diversity of vocabulary and homogeneity of literary style. *Journal of Applied Statistics* **33**: 729–741.
Sichel HS. 1975. On a distribution law for words frequencies. *Journal of the American Statistical Association* **70**: 542–547.
Sichel HS. 1997. Modelling species-abundance frequencies and species-individual functions with the generalized inverse Gaussian-Poisson distribution. *South African Statistical Journal* **31**: 13–37.
Smyth GK, Jorgensen B. 2002. Fitting Tweedie's compound Poisson model to insurance claims data: dispersion modelling. *Astin Bulletin* **32**: 143–157.
Tweedie MCK. 1984. An index which distinguishes between some important exonential families. In *Statistics: Applications and New Directions, Proceedings of the Indian Statistical Institute Golden Jubilee International Conference*, JK Ghosh and J Roy (eds.), 579-604. Calcutta: Indian Statistical Institute.
Valero J, Pérez-Casany M, Ginebra J. 2010. On zero trucating and mixing Poisson distributions. Accepted for publication in *Advances in Applied Probability*.
Willmot GE. 1989. A remark on the Poisson-Pascal and some other contagious distributions. *Statistics and Probability Letters* **7**: 217–220.
Zhu R, Joe H. 2009. Modelling heavy-tailed count data using a generalized Poisson-inverse Gaussian family. *Statistics and Probability Letters* **70**: 1695–1703.

Jordi Valero
Departament de Matemàtica Aplicada 3
ESAB, Edifici ESAB D4 064
Avda del Canal Olímpic, s/n
Universitat Politècnica de Catalunya
08860 Castelldefels
E-mail address: jordi.valero@upc.edu

Josep Ginebra
Departament d'Estadística i Investigació Operativa
E.T.S.E.I.B., Avda Diagonal 647, 6a. planta
Universitat Politècnica de Catalunya
08028 Barcelona, Spain
E-mail address: josep.ginebra@upc.edu

Marta Pérez-Casany
Departament de Matemàtica Aplicada 2 i DAMA-UPC
Edifici Omega, Jordi Girona 1-3
Universitat Politècnica de Catalunya
08034 Barcelona, Spain
E-mail address: marta.perez@upc.edu