OVERDISPERSION AND POISSON-TWEEDIE EXPONENTIAL DISPERSION MODELS

C.C. Kokonendji, C.G.B. Demétrio and S. Dossou-Gbété

Abstract. We investigate two sets of overdispersed models when Poisson distribution does not fit to count data: a class of Poisson mixture with Tweedie mixing distributions and a class of exponential dispersion models which have a unit variance function of the form $\mu + \mu^p$, where p is a real number. These two classes generalize the negative binomial distribution which is classically used in the framework of regression models for count data when overdispersion results in a lack of fit of the Poisson regression model. Some properties are then studied and discussed.

Keywords: Natural exponential family, Poisson mixture, negative binomial distribution, stable distribution, unit variance function.

AMS classification: 62E15, 62F10, 62J02.

§1. Introduction

When Poisson distribution does not fit to count data and the observed dispersion is greater than the predicted dispersion, the negative binomial distribution is often used to overcome this lack of fit, the so-called overdipersion. It is well known that negative binomial distribution can be understood as a Poisson mixture with gamma mixing distribution. Hougaard *et al.* [5] have considered a large family of mixture distributions, including the inverse Gaussian mixture distribution, to improve significantly the fitness to certain data. We will call the *Poisson-Tweedie* class a completed set of these distrbutions.

It is common use to handle overdispersion in regression models for count data by replacing the initial Poisson regression model by a model where the mean-variance function has a more general form. Hinde and Demétrio [3, page 14] propose for overdispersed count data the use of the variance function (or variance as non constant function of the mean)

$$V_{\phi,p}(m) = m + \phi m^p, \tag{1}$$

where $\phi > 0$ and $p \in \mathbb{R}$ fixed, which is also a generalization of negative binomial variance function with $p = 2$. To make short that the *Hinde-Demétrio* class, we here call the *Hinde* class of "exponential dispersion models" (a term to be made precise) all distributions corresponding to variance function (1). Note that overdispersion can be measured by the variance inflation factor $V_{\phi,p}(m)/m = 1 + \phi m^{p-1}$, which is 1 for the Poisson distribution. Obviously, when ϕ goes to 0 in (1) the corresponding limit distribution must be a Poisson.

The aim of this work is to provide a complete identification of both the Poisson-Tweedie and the Hinde exponential dispersion models from their variance functions. In section 2, we briefly review some basic properties of exponential dispersion models and, in particular, of the *Tweedie* class with variance function ϕm^p. In section 3, we describe the possible Poisson mixture distributions with a Tweedie for obtaining the Poisson-Tweedie class. In section 4, we first characterize the Hinde class and we then compare it to the Poisson-Tweedie class. The last section is devoted to some discussion about statistical inference and some concluding remarks on the Hinde class.

§2. Exponential dispersion models

Exponential dispersion models are important statistical models, particularly for the treatment of generalized linear models (McCullagh and Nelder [15]). They have a number of important mathematical properties, which are relevant in pratice, and they include several well-known families of distributions as special cases, giving a convenient general framework. The reader can be referred to Jørgensen [7] for more details or to the contribution of Muriel Casalis in Chapter 54 of Kotz *et al.* [11] for a recent panorama of the subject.

Let $\theta \in \Theta$ and $\lambda \in \Lambda$, where Θ is generally an interval with interior $(\mathrm{int}\Theta)$ non-empty of $\mathbb{R}$ and Λ a subset of $(0,\infty)$. A random variable X from an *exponential dispersion* distribution with parameters θ and λ, denoted $X \sim ED(\theta,\lambda)$, if its density or mass function can be written as

$$c(x;\lambda)\exp\{\theta x - \lambda K(\theta)\}, \quad x \in S \subseteq \mathbb{R}. \tag{2}$$

The associated (additive) *exponential dispersion model* (EDM) is the set of probabilities $ED(\theta,\lambda)$ with $\theta \in \Theta$ and $\lambda \in \Lambda$. [Note here that the reproductive version of $X \sim ED(\theta,\lambda)$ is given by $Z = X/\lambda \sim ED^*(\mu;\sigma^2)$, where $\mu = K'(\theta)$ and $\sigma^2 = 1/\lambda$.]

For fixed $\lambda > 0$, the EDM is a *natural exponential family* (NEF). Therefore θ is the canonical parameter and λ the index parameter. These parameters satisfy the following convolution formula: $ED(\theta,\lambda_1) * ED(\theta,\lambda_2) = ED(\theta,\lambda_1+\lambda_2)$. So the EDM is closed under convolution and $\{1,2,\cdots\} \subseteq \Lambda$. Also, the model is infinitely divisible if and only if $\Lambda = (0,\infty)$.

The cumulant function K in (2) is such that, if ν is this reference measure (e.g., Lebesgue or counting or σ-finite and positive) on $\mathbb{R}$ then

$$K(\theta) = \ln \int_{\mathbb{R}} e^{\theta x} c(x;1)\, d\nu(x).$$

Therefore K is strictly convex on $\mathrm{int}\Theta$ and, for a random variable $X \sim ED(\theta,\lambda)$, one has its expectation and variance:

$$\mathbb{E}(X) = \lambda K'(\theta) \quad \text{and} \quad \mathrm{Var}(X) = \lambda K''(\theta), \tag{3}$$

where $K'(\theta)$ and $K''(\theta)$ are, respectively, the first and second derivatives of K at the point θ. From (3) with $\lambda = 1$, the function V defined on the domain $M = K'(\mathrm{int}\Theta)$ such that

$$K''(\theta) = V\{K'(\theta)\}$$

is called *unit variance function*. As for the NEFs, the unit variance function characterizes the EDM. Numerous properties have been established in the literature and, for many cases, the unit variance function presents a simpler expression than the density (e.g., Letac and Mora [13]). The role of the unit variance function in data fitting should be to identify the class of adequate distributions.

The reparametrization $\mu = K'(\theta)$ (unit mean) permits us to write the EDM as follows: $\{ED(\mu,\lambda); \mu \in M, \lambda \in \Lambda\}$. It is sometimes considered the reparametrization of the EDM by the mean $m = \mathbb{E}(X) = \lambda\, K'(\theta)$. From (3) the unit variance function V provides the variance $V_\lambda = \mathrm{Var}(X)$ of $X \sim ED(\theta,\lambda)$ in terms of m, called *variance function* and expressed as follows: $V_\lambda(m) = \lambda V(m/\lambda)$, for all $m \in \lambda M$. Before giving a basic example of EDMs with power variance functions, let us recall that the discrete overdispersed EDM compared to the Poisson distribution must satisfy

$$V(\mu) > \mu, \tag{4}$$

for all $\mu > 0$, where $V(\mu) = \mu$ is the unit variance function of the Poisson model (e.g., Jourdan and Kokonendji [9]).

2.1. Tweedie EDMs

A complete description of the EDMs with power unit variance functions

$$V(\mu) = \mu^p, \quad p \in (-\infty, 0] \cup [1, \infty), \tag{5}$$

is given by Jørgensen [7] where, for $p = \infty$ the corresponding unit variance function takes the exponential form $V(\mu) = \exp(\beta\mu)\,,\, \beta \neq 0$. This class called the *Tweedie class* is introduced by Tweedie [16]; see also Bar-Lev and Enis [1]. Instead of p, it is also convenient to introduce the index parameter α of stable distribution, defined by

$$(p-1)(1-\alpha) = 1. \tag{6}$$

According to the above notations, we can denote by $Tw_p(\theta,\lambda)$ any distribution of this class where p and α are connected by (6), $\lambda \in (0,\infty) = \Lambda$ for all p of (5), and $\theta \in \Theta_p$ with

$$\Theta_p = \begin{cases} \mathbb{R} & \text{for} \quad p = 0, 1 \\ [0,\infty) & \text{for} \quad p < 0 \;\text{ or }\; 0 < p < 1 \\ (-\infty, 0) & \text{for} \quad 1 < p \leq 2 \\ (-\infty, 0] & \text{for} \quad 2 < p < \infty. \end{cases} \tag{7}$$

Thus, for $s \in \Theta_p - \theta$, the Laplace transform $\mathbb{E}(e^{sX})$ of $X \sim Tw_p(\theta,\lambda)$ is given by

$$G_p(s;\theta,\lambda) = \begin{cases} \exp\{\frac{\lambda(1-p)^\alpha\theta^\alpha}{(2-p)}[(1+s/\theta)^\alpha - 1]\} & \text{for} \quad p \neq 1, 2 \\ (1+s/\theta)^{-\lambda} & \text{for} \quad p = 2 \\ \exp\{\lambda e^\theta(e^s - 1)\} & \text{for} \quad p = 1. \end{cases} \tag{8}$$

As shown in Table 1, the Tweedie class includes several well-known families of distributions amongst which one may the inverse-Gaussian distributions with $p = 3$ or $\alpha = 1/2$. Observe that the extreme stable distributions ($p < 0$) are not “steept’t’ and only one distribution, named Poisson ($p = 1$), is discrete.

Let us conclude this section by precising the notion of steepness. If $M = \Omega$, where Ω denotes the interior of the convex hull of the support S of EDM, the model is then said to be *steep*. From here to the end, an EDM is always assumed to be steep.

Table 1: Summary of Tweedie EDMs (Jørgensen [7])

Distribution	p	α	M	S
Extreme stable	$p < 0$	$1 < \alpha < 2$	$(0, \infty)$	$\mathbb{R}$
Normal	$p = 0$	$\alpha = 2$	$\mathbb{R}$	$\mathbb{R}$
[Do not exist]	$0 < p < 1$	$2 < \alpha < \infty$		
Poisson	$p = 1$	$\alpha = -\infty$	$(0, \infty)$	$\mathbb{N}$
Compound Poisson	$1 < p < 2$	$\alpha < 0$	$(0, \infty)$	$(0, \infty)$
Gamma	$p = 2$	$\alpha = 0$	$(0, \infty)$	$(0, \infty)$
Positive stable	$p > 2$	$0 < \alpha < 1$	$(0, \infty)$	$[0, \infty)$

§3. Poisson-Tweedie EDMs

Before showing the form of its (unit) variance function, let us precise what we call the Poisson-Tweedie class of EDMs.

Let X be a non-negative random variable following $Tw_p(\theta, \lambda)$ an element of Tweedie class. If a discrete random variable Y is such that the conditional distribution of Y given X is Poisson with mean X, then the EDM generated by the distribution of Y is of the Poisson-Tweedie class; and we can denote by $PTw_p(\theta, \lambda)$ any distribution of this class. That means that the support S of $X \sim Tw_p(\theta, \lambda)$ must be positive for defining the following individual probabilities of Y:

$$\Pr(Y = y) = \int_0^\infty \frac{e^{-x} x^y}{y!} Tw_p(\theta, \lambda)(dx), \quad y = 0, 1, \cdots. \tag{9}$$

Hence, from Table 1, these Poisson mixtures are possible only for $p \geq 1$ for which some basic properties are described by Hougaard *et al.* [5, Theorem 1] except for $p = 1$ or $\alpha = -\infty$. We here present the main results of Poisson-Tweedie distributions (9).

Theorem 1. *Let $Y \sim PTw_p(\theta, \lambda)$ defined by (9), where the parameter set is $\lambda > 0$, $\theta \in \Theta_p$ given by (7), and $p \geq 1$ or $\alpha \in [-\infty, 1)$ from (6). (i) If $Y_1, \cdots, Y_n$ are independent, with $Y_i \sim PTw_p(\theta, \lambda_i)$, $i = 1, \cdots, n$, then $Y_1 + \cdots + Y_n$ follows $PTw_p(\theta, \lambda_1 + \cdots + \lambda_n)$. The distribution $PTw_p(\theta, \lambda)$ is infinitely divisible. (ii) The Laplace transform of Y is*

$$\mathbf{E}(e^{sY}) = \begin{cases} \exp\{\frac{\lambda(1-p)^\alpha}{(2-p)}[(e^s - 1 + \theta)^\alpha - \theta^\alpha]\} & \text{for} \quad p \neq 1, 2 \\ [(e^s - 1 + \theta)/\theta]^{-\lambda} & \text{for} \quad p = 2 \\ \exp\{\lambda[\exp(e^s - 1 + \theta) - e^\theta]\} & \text{for} \quad p = 1, \end{cases} \tag{10}$$

for $s \in \Theta_p - \theta$. For $p = 1$, it is a Neyman type A distribution; for $p = 2$, it is a negative binomial distribution; and, for $p = 3$, it is the Sichel or Poisson-inverse Gaussian distribution. (iii) The distribution $PTw_p(\theta, \lambda)$ is overdispersed with respect to the Poisson distribution for all $p \geq 1$. (iv) The distribution $PTw_p(\theta, \lambda)$ is unimodal for $p \geq 2$. (v) The unit variance function of the EDM generated by $Y \sim PTw_p(\theta, 1)$ is exactly

$$V_{PT}(\mu) = \mu + \mu^p \exp\{(2 - p)\Phi(\mu)\}, \quad \mu > 0, \tag{11}$$

where $\Phi(\mu)$ denotes the inverse of the increasing function $s \mapsto d\{\ln \mathbb{E}(e^{sY})\}/ds$.

Proof. (i) It is deduced from the convolution formula of EDM.

(ii) Since the Laplace transform of a Poisson with mean X is $\exp\{X(e^s-1)\}$, from (9) and (8) we have

$$\mathbb{E}(e^{sY}) = \mathbb{E}(\mathbb{E}(e^{sY}|X)) = G_p(e^s-1;\theta,\lambda);$$

and the remainder becomes trivial.

(iii) Using (ii) for the cumulant generating function $K(s) = \ln \mathbb{E}(e^{sY})$, the rth cumulant of $Y \sim PTw_p(\theta,\lambda)$ is given by $\kappa_r = d^rK(0)/ds^r$. Therefore, the first two cumulants of Y, respectively, are given by

$$\kappa_1 = \mathbb{E}(Y) = \begin{cases} \frac{\lambda\alpha(1-p)^\alpha}{(2-p)}\theta^{\alpha-1} & \text{for} \quad p \neq 1,2 \\ -\lambda\theta^{-1} & \text{for} \quad p = 2 \\ \lambda e^\theta & \text{for} \quad p = 1, \end{cases}$$

$$\kappa_2 = \mathrm{Var}(Y) = \begin{cases} \frac{\lambda\alpha(1-p)^\alpha}{(2-p)}\theta^{\alpha-1}[1+(1-p)^{-1}\theta^{-1}] & \text{for} \quad p \neq 1,2 \\ \lambda(1-\theta)\theta^{-2} & \text{for} \quad p = 2 \\ 2\lambda e^\theta & \text{for} \quad p = 1. \end{cases}$$

It follows that the index of dispersion (ID $= \kappa_2/\kappa_1$) of Y verifies ID > 1 for all $\theta \in \Theta_p$, $\lambda > 0$ and $p \geq 1$, proving (iii).

(iv) See Theorem 1 (f) of Hougaard *et al.* [5] with $\alpha \in [0,1]$.

(v) Let $K(s) = \ln \mathbb{E}(e^{sY})$ for $Y \sim PTw_p(\theta, 1)$. From (ii) with $\lambda = 1$ and using (6) to simplify, the first derivative of $K(s)$ is

$$\mu = K'(s) = \begin{cases} e^s[(1-p)(e^s-1+\theta)]^{\alpha-1} & \text{for} \quad p \neq 1,2 \\ -e^s(e^s-1+\theta)^{-1} & \text{for} \quad p = 2 \\ e^s\exp\{e^s-1+\theta\} & \text{for} \quad p = 1, \end{cases}$$

and the second derivative of $K(s)$ may be expressed as follows:

$$V_{PT}(\mu) = K''(s) = \begin{cases} K'(s) + e^{2s}[(1-p)(e^s-1+\theta)]^{\alpha-2} & \text{for} \quad p \neq 1,2 \\ K'(s) + [K'(s)]^2 & \text{for} \quad p = 2 \\ K'(s) + e^sK'(s) & \text{for} \quad p = 1. \end{cases}$$

For $p \neq 1, 2$ we can also write

$$K''(s) = K'(s) + e^{2s}\left[\frac{K'(s)}{e^s}\right]^{\frac{\alpha-2}{\alpha-1}}.$$

Since (6) implies $(\alpha-2)/(\alpha-1) = p$, the expression given in (11) is easily obtained. □

The probability function of $Y \sim PTw_p(\theta,\lambda)$ is given by Hougaard *et al.* [5] which is the key difficulty in applying this family of models. Some other properties are discussed in the same paper such that the skweness and the bimodality, indeed trimodality. For $p = 1$ we can refer to Johnson *et al.* [6, pages 368-] for obtaining some properties on the Neyman type A distribution, which is therefore both a Poisson mixture of Poisson distributions, and also a Poisson-stopped sum of Poisson distributions. The Poisson-Tweedie EDMs can be summarized as in Table 2.

Note however that the conditional distribution of Y given X by (9) is not the same of that proposed by Jørgensen [7, page 166] which is given by $Y|X = x$ follows Poisson with mean xe^η, where $\eta \in \mathbb{R}$ is an additional parameter. Thus, the result in terms of unit variance function is slightly different.

Table 2: Summary of Poisson-Tweedie EDMs

Distribution	p	α	M	S
[Do not define]	$p < 1$	$1 < \alpha < \infty$		
Neyman type A	$p = 1$	$\alpha = -\infty$	$(0, \infty)$	$\mathbb{N}$
Poisson-compound Poisson	$1 < p < 2$	$\alpha < 0$	$(0, \infty)$	$\mathbb{N}$
Negative binomial	$p = 2$	$\alpha = 0$	$(0, \infty)$	$\mathbb{N}$
Poisson-positive stable	$p > 2$	$0 < \alpha < 1$	$(0, \infty)$	$\mathbb{N}$
Poisson-inverse Gaussian	$p = 3$	$\alpha = 1/2$	$(0, \infty)$	$\mathbb{N}$

§4. Hinde EDMs

In this section, we characterize the Hinde class which is the set of EDMs with unit variance function of the "simple" form

$$V_H(\mu) = \mu + \mu^p, \quad \mu \in M_H, \tag{12}$$

where $p \in \mathbb{R}$ and, then, we compare it to the Poisson-Tweedie class (11). The EDM corresponding to (12), if it exists for a given p, is denoted $Y \sim H_p(\mu, \lambda)$. For all integers $p \in \{1, 2, \cdots\}$, the models $H_p(\mu, \lambda)$ exist and they are infinitely divisible by using the Bar-Lev criterion described in Letac and Mora [13, Corollary 3.3], because we can write $V_H(\mu) = \mu\Delta(\mu)$ on $M_H = (0, r)$, where Δ is a polynomial with non-negative coefficients and $r \in (0, \infty]$.

We now state the result of characterization.

Theorem 2. *Let $p \in \mathbb{R}$. The function (12): $\mu \mapsto V_H(\mu) = \mu + \mu^p$, defined on a suitable domain $M_p = M_H$ corresponds to a unit variance function of a discrete (steep) EDM when*

$$p \in \{0\} \cup [1, \infty), \tag{13}$$

with $M_0 = (-1, \infty)$ and $M_p = (0, \infty)$ for $p \geq 1$; and the domain Θ_p of the canonical parameter is given by (7). In particular, if $p = 0$ the model $H_0(\mu, \lambda)$ is a positive-translated Poisson; if $p = 1$ the model $H_1(\mu, \lambda)$ is a scaled Poisson; if $p = 2$ the model $H_2(\mu, \lambda)$ is negative binomial; if $p = 3$ the model $H_3(\mu, \lambda)$ is strict arcsine (Kokonendji and Khoudar, 2004).

Before embarking on the proof, let us recall that, from (4), the Hinde class (12) is the set of overdispersed EDMs compared with the Poisson distribution, as well as the Poisson-Tweedie class (11). As consequence to previous results, we have the following comparison result. It means that only negative binomial $H_2(\mu, \lambda)$ of the Hinde class is interpreted as $PTw_2(\mu, \lambda)$ of the Poisson-Tweedie class and, for fixed $p \geq 1$ and $\lambda > 0$, each $H_p(\mu, \lambda)$ can be approximated by $PTw_p(\mu, \lambda)$ when μ goes to ∞. For this reason we can call the "Hinde-compound Poisson" (resp. "Hinde-positive stable") EDMs for $1 < p < 2$ (resp. $p > 2$) in (12).

Proposition 3. *Let $H = \{H_p(\mu, \lambda); p \in \mathbb{R}\}$ be the Hinde class and $PT = \{PTw_p(\mu, \lambda); p \in \mathbb{R}\}$ the Poisson-Tweedie class. Then: (a) $H \cap PT = \{H_2(\mu, \lambda) = PTw_2(\mu, \lambda)\}$. (b) For fixed $p \geq 1$, $V_{PT}(\mu) \sim V_H(\mu)$ as $\mu \to \infty$.*

Proof. Since the unit variance function characterizes the EDM, letting $p = 2$ in (11) we easily deduce the result (a). For the part (b) we can assume $p \neq 2$ from (a) and, then, we have from (11) that $\mu \to \infty$ implies $s = \Phi(\mu) \to 0$. From the Taylor expansion of the function $s \mapsto \exp\{(2-p)s\}$ in a neighborhood of $s = 0$, the result is easily obtained. □

For the proof of Theorem 2 we need the two following lemmas. The first is called an "impossibility criterion" to exclude case $0 < p < 1$, and the second is related to the steepness.

Lemma 4. *There are no EDM with $M = (0, \infty)$ and unit variance function $V(\mu) \sim \mu^\gamma$ as $\mu \to 0$ for $\gamma \in (0, 1)$.*

Proof. If $V(\mu) \sim \mu^\gamma$ as $\mu \to 0$, then

$$\theta = \psi(\mu) = \theta_0 + \int_0^\mu \frac{dt}{V(t)}$$

is left-bounded. Now, $V(\mu) \to 0$ as $\mu \to 0$ implies that the generating measure ν is concentrated on $[0, \infty)$ (see, e.g., Letac and Mora [12, Theorem 2.3 (2)]). Hence, the canonical parameter domain $\Theta(\nu)$ is not left-bounded, which yields a contradiction. □

Lemma 5 (Jørgensen et al. [8]). *Let $\mathcal{P} = \{\, ED(\theta, 1)\,; \theta \in \Theta \,\}$ be a NEF with variance function V on M and support S. If $\inf S = 0$, then: (i) $\inf M = 0$; (ii) $\lim_{\mu \to 0} V(\mu) = 0$; (iii) $\lim_{\mu \to 0} V(\mu)/\mu = c$, where $c = \inf\{S \setminus \{0\}\}$.*

Note that $c = 0$ for continuous distributions, and $c > 0$ for distributions which have an atom at zero; in particular, $c = 1$ for discrete integer valued distributions.

Proof of Theorem 2. We first make an observation based on Theorem 2.3 of Letac and Mora [12]. Since V_H must be an analytic positive function on the domain $M_p = (a, \infty)$, we have that both V_H has no zero in (a, ∞) and $V_H(a) = 0$. Thus, we have

$$M_p = \begin{cases} (0, \infty) & \text{for} \quad p \neq 0 \\ (-1, \infty) & \text{for} \quad p = 0, \end{cases} \tag{14}$$

and, in solving $\psi'(\mu) = 1/V_H(\mu) = 1/(\mu + \mu^p)$ that we ignore the arbitrary constants in the solutions,

$$\psi(\mu) = \begin{cases} \ln(\mu) - (p-1)^{-1} \ln(1 + \mu^{p-1}) & \text{for} \quad p \neq 0, 1 \\ \ln \sqrt{\mu} & \text{for} \quad p = 1 \\ \ln(1 + \mu) & \text{for} \quad p = 0. \end{cases} \tag{15}$$

We now examine the different situations of $p \in \mathbb{R}$ in (12) from (14).

- Consider case $p \in \{0\} \cup [1, \infty)$. Let $\theta = \psi(\mu)$ in (15), then we find $\mu = \mu(\theta) = K'(\theta)$ as follows:

$$K'(\theta) = \begin{cases} e^\theta (1 - e^{\theta(p-1)})^{-1/(p-1)} & \text{for} \quad p \neq 0, 1 \\ e^{2\theta} & \text{for} \quad p = 1 \\ e^\theta - 1 & \text{for} \quad p = 0, \end{cases}$$

and, hence (using Maple for $p \neq 0, 1, 2, 3$),

$$K(\theta) = \begin{cases} e^{\theta} - \theta & \text{for} \quad p = 0 \\ e^{2\theta}/2 & \text{for} \quad p = 1 \\ -\ln(1 - e^{\theta}) & \text{for} \quad p = 2 \\ \arcsin e^{\theta} & \text{for} \quad p = 3 \\ \frac{(-1)^{-1/(p-1)}}{(p-1)\Gamma[1/(p-1)]} \sum_{k=0}^{\infty} (1 - e^{k\theta(p-1)}) \frac{\Gamma[k+1/(p-1)]}{k\Gamma(k+1)} & \text{for} \quad p \neq 0, 1, 2, 3, \end{cases} \tag{16}$$

for $\theta \in \Theta$, where the interior of Θ is obtained by using (14) and (15):

$$\text{int}\Theta = \begin{cases} \mathbb{R} & \text{for} \quad p = 0, 1 \\ (0, \infty) & \text{for} \quad p < 0 \ \text{ or } \ 0 < p < 1 \\ (-\infty, 0) & \text{for} \quad 1 < p < \infty. \end{cases} \tag{17}$$

Since the cumulant generating function (16) is analytic, the domain Θ defined from its interior (17) coincides to Θ_p given by (7). Thus, for each $p \in \{0\} \cup [1, \infty)$, we define a unique discrete positive measure generating the corresponding (steep) EDM with unit variance function (12).

- Case $0 < p < 1$ is excluded by Lemma 4.

- Finally, let us exclude case $p < 0$ by steepness criterion. Indeed, by Lemma 5, it suffices to observe that $M = (0, \infty)$ from (14) and $\lim_{\mu \to 0} V(\mu)/\mu = \lim_{\mu \to 0}(1 + \mu^{p-1}) = \infty$. The proof of Theorem 2 is now complete. □

§5. Final remarks

From the unit variance function (12) of the Hinde class, it is easy to point out (graphically) different situations of overdispersion with respect to the Poisson distribution. The point $\mu = 1$ plays an interesting role for the degree or level of overdispersion which depends on $\mu < 1$ or $\mu > 1$.

The first behaviour of variance functions of Hinde distributions compared to Poisson provides a real view to the phenomena of overdispersion. Thus we can find, in order when $\mu < 1$, positive-translated Poisson ($p = 0$), scaled Poisson ($p = 1$), Hinde-compound Poisson ($1 < p < 2$), negative binomial ($p = 2$) and Hinde-positive stable ($p > 2$) distributions. The order changes when $\mu > 1$. Consequently, it is an indicator to choose the adequate class of distributions. See also Jourdan and Kokonendji [9] for the overdispersed generalized negative binomial distributions with respect to the negative binomial and, then, to the Poisson distributions. Let us mention here that the well-known case of Poisson-positive stable ($p > 2$) families is the Sichel or Poisson-inverse Gaussian ($p = 3$) distribution (e.g., Holla [4]; Willmot [17]); and the known case of Hinde-positive stable ($p > 2$) families is the strict arcsine ($p = 3$) distribution, which is not always unimodal (see Kokonendji and Khoudar [10]).

With respect to the limit cases which are Poisson, scaled Poisson and positive-translated Poisson, one can illustrate situations that there are no distributions in this class corresponding to the phenomena. Indeed, for $\mu < 1$ we have the following conclusions: between scaled Poisson and positive-translated Poisson, there are no EDM corresponding to these variance functions with $0 < p < 1$; above positive-translated Poisson which correspond to $p < 0$, there are no steep EDM. The positions change when $\mu > 1$. We must here observe that cases $p < 0$

can exist with a support S of distributions containing some negative point or being $\mathbb{Z} = -\mathbb{N} \cup \mathbb{N}$. Furthermore, when μ goes to ∞ the possible limit distribution could be a Poisson.

When the Poisson-Tweedie models $PTw_p(\mu, \lambda)$ or the Hinde models $H_p(\mu, \lambda)$ are used in the case of data set (univariate or regression), the real problem of statistical inference is the parameter p. If $\underline{x} = (x_1, \cdots, x_n)$ is an n-independent identically distributed observation from $PTw_p(\mu, \lambda)$ or $H_p(\mu, \lambda)$, it is recommended to use the moment estimate because it is simple (e.g., Kokonendji and Khoudar [10]) and it can be an initial estimate, for example, in the search for maximum likelihood estimate (e.g., Hougaard *et al.* [5]). For example, from (1) and when ϕ is fixed or known, we have by the moment method $p^* = \ln[(s_n^2 - \overline{x}_n)/\phi]/\ln(\overline{x}_n) \in \mathbb{R}$ under the overdispersion condition $s_n^2 - \overline{x}_n > 0$, where $\overline{x}_n$ and s_n^2 are, respectively, the mean and the variance from the count data $\underline{x}$. However, a profile estimate of p is necessary in this situation. Finally, as shown by Hougaard *et al.* [5] in the univariate case, the use of these models could be interesting in the regression case to unify the negative binomial and the Poisson-inverse Gaussian regression models (Dean *et al.* [2]) or the strict arcsine regression model (Marque and Kokonendji [14]) for a suitable model.

References

[1] BAR-LEV, S.K., AND ENIS, P. Reproducitibility and natural exponential families with power variance functions. *Ann. Statist.*, 14 (1986), 1507-1522.

[2] DEAN, C., LAWLESS, J.F., AND WILLMOT, G.E. A mixed Poisson-inverse Gaussian regression model. *Canadian J. Statist.*, 17 (1989), 171-181.

[3] HINDE, J., AND DEMÉTRIO, C.G.B. *Overdispersion: Models and Estimation*, ABE, São Paulo, 1998.

[4] HOLLA, M.S. On a Poisson-inverse Gaussian distribution. *Metrika*, 11 (1966), 115-121.

[5] HOUGAARD, P., LEE, M-L.T., AND WHITMORE, G.A. Analysis of overdispersed count data by mixtures of Poisson variables and Poisson processes. *Biometrics*, 53 (1997), 1225-1238.

[6] JOHNSON, N.L., KOTZ, S., AND KEMP, A.W. *Univariate Discrete Distributions, Second Edition*. John Wiley & Sons, New York, 1992.

[7] JØRGENSEN, B. *The Theory of Dispersion Models.* Chapman & Hall, London, 1997.

[8] JØRGENSEN, B., MARTÍNEZ, J.R., AND TSAO, M. Asymptotic behaviour of the variance function. *Scandinavian J. Statist.*, 21 (1994), 223-243.

[9] JOURDAN, A., AND KOKONENDJI, C.C. Surdispersion et modèle binomial négatif généralisé. *Rev. Statistique Appliquée*, L (3) (2002), 73-86.

[10] KOKONENDJI, C.C., AND KHOUDAR, M. On strict arcsine distribution. *Commun. Statist.-Theor. Meth.*, 33 (5) (2004), to appear.

[11] KOTZ, S., BALAKRISHNAN, N., AND JOHNSON, N.L. *Continuous Multivariate Distributions, Vol. 1: Models and Applications, Second edition*. Wiley, New York, 2000.

[12] LETAC, G., AND MORA, M. Sur les fonctions-variance des familles exponentielles naturelles sur $\mathbb{R}$. *C. R. Acad. Sc. Paris, Série I,* 302 (1986), 551-554.

[13] LETAC, G., AND MORA, M. Natural real exponential families with cubic variance functions. *Ann. Statist.*,18 (1990), 1-37.

[14] MARQUE, S., AND KOKONENDJI, C.C. A strict arcsine regression model. *Preprint LMA - University of Pau* (submitted), 2004.

[15] MCCULLAGH, P., AND NELDER, J.A. *Generalized Linear Models.* Chapman & Hall, London, 2nd edition, 1989.

[16] TWEEDIE, M.C.K. An index which distinguishes between some important exponential families. In *Statistics: Applications and new directions. Proceedings of the Indian Statistical Institute Golden Jubilee International Conference* (editors J.K. Ghosh and J. Roy), 1984, 579-604. Indian Statistical Institue, Calcutta.

[17] WILLMOT, G.E. The Poisson-inverse Gaussian distribution as an alternative to the negative binomial. *Scandinavian Actuarial J.,* (2) (1987), 113-127.

Célestin C. Kokonendji and
Simplice Dossou-Gbété
Laboratoire de Mathématiques Appliquées,
FRE 2570,
Université de Pau et des Pays de l'Adour
IPRA, BP1155, 64013 Pau Cedex, France
celestin.kokonendji@univ-pau.fr
and
simplice.dossou-gbete@univ-pau.fr

Clarice G.B. Demétrio
Departamento de Ciencias Exatas
Ecola Superior de Agricultura Luiz de Quieroz
University of São Paulo - Caixa Postal 9
13418-900 Piracicaba, SP - Brazil
clarice@carpa.ciagri.usp.br