On Hinde-Demetrio Regression Models for Overdispersed Count Data

Célestin Kokonendji, Clarice G.B. Demétrio, Silvio S. Zocchi

On Hinde-Demétrio Regression Models for Overdispersed Count Data

Célestin C. Kokonendji [a,*],

[a] *University of Pau - LMA UMR 5142 CNRS - Pau, France*

Clarice G.B. Demétrio [b] and

[b] *University of São Paulo - ESALQ Piracicaba - SP, Brazil*

Silvio S. Zocchi [b].

Abstract

In this paper we introduce the Hinde-Demétrio (HD) regression models for analyzing overdispersed count data and, mainly, investigate the effect of dispersion parameter. The HD distributions are discrete additive exponential dispersion models (depending on canonical and dispersion parameters) with a third real index parameter p and have been characterized by its unit variance function $\mu + \mu^p$. For p equals to $2, 3, \cdots$, the corresponding distributions are concentrated on nonnegative integers, overdispersed and zero-inflated with respect to a Poisson distribution having the same mean. The negative binomial ($p = 2$), strict arcsine ($p = 3$) and Poisson ($p \to \infty$) distributions are particular count HD families. From generalized linear modelling framework, the effect of dispersion parameter in the HD regression models, among other things, is pointed out through the double mean parametrization: unit and standard means. In the particular additive model, this effect must be negligible within an adequate HD model for fixed integer p. The estimation of the integer p is also examined separately. The results are illustrated and discussed on a horticultural data set.

Key words: Additive exponential dispersion model, compound Poisson, generalized linear model, model selection, unit variance function, zero-inflation.
AMS classification: Primary 62J02; Secondary 62J12, 62F07.
Abbreviated title: Hinde-Demétrio regression models.

* *Address for correspondence:* C.C. Kokonendji. Université de Pau et des Pays de l'Adour. Laboratoire de Mathématiques Appliquées - UMR 5142 CNRS. Département STID. Avenue de l'Université. 64000 Pau, France. Tel +33(0)559 407 145; Fax +33(0)559 407 140.

Email addresses: celestin.kokonendji@univ-pau.fr (Célestin C. Kokonendji), clarice@carpa.ciagri.usp.br (Clarice G.B. Demétrio), sszocchi@esalq.usp.br (Silvio S. Zocchi).

1 INTRODUCTION

The regression models for count data are special because the structure of problem to solve and the inference techniques are not yet well developped for general families of discrete *exponential dispersion models* (EDMs), in particular for the Hinde-Demétrio families. Hence, fully flexible methods for analysis of count data are still not readily available (Cameron and Trivedi 1998, for econometric literature). The Poisson regression model provides a standard framework for the analysis of count data. One of the reasons is that Poisson distribution was historically considered as the "normal" distribution for count data. Because of its single parameter (having no dispersion parameter), many scenarios were necessary to construct suitable count distributions using some indexes as measures to detect departures from Poisson distribution. For instance, the most known, so frequent and well-explained of phenomena are overdispersion and zero-inflation (Hall and Berenhaut 2002; Mullahy 1997). Of course the opposite phenomena exist but is uncommon (Bosch and Ryan 1998; Castillo and Pérez-Casany 2005; Kokonendji and Mizère 2005). For this work our interest is on the *Hinde-Demétrio regression models* (HDRMs).

The Hinde-Demétrio distributions have been introduced by Kokonendji et al. (2004) as discrete *additive* EDMs (Jørgensen 1997; Vinogradov 2006, sec. 1-2) and are characterized by unit variance functions of the simple form:

$$V_p(\mu) = \mu + \mu^p, \quad p \in \{0\} \cup [1, \infty), \tag{1}$$

where $\mu > -1$ for $p = 0$ and $\mu > 0$ for $p \geq 1$. The index (or "power") parameter p given in (1) is associated to a particular additive EDM, which is a linear exponential family with a dispersion parameter $\phi > 0$. Recall that EDMs are the prototype response distributions for Generalized Linear Models (McCullagh and Nelder 1989; Jørgensen 2001). This third parameter p takes place on the support S_p of distributions as follows: $S_0 = \{-1, 0, 1, \cdots\} = \{-1\} \cup \mathbb{N}$, $S_1 = 2\mathbb{N}$ and $S_p = p\mathbb{N} \cup \mathbb{N}$ for $p > 1$. Consequently, we only need $p \in \{2, 3, \cdots\}$ for analyzing count data. As particular cases, we have the negative binomial for $p = 2$ and the strict arcsine for $p = 3$ (Kokonendji and Khoudar 2004; Kokonendji and Marque 2005). The limit case ($p \to \infty$) is associated to a Poisson distribution.

Note however that the origin of Hinde-Demétrio family could be an approximation (in terms of unit variance function) to the Poisson-Tweedie family, which is also the set of EDMs generated by Poisson mixture with positive stable mixing distribution (Hougaard *et al.* 1997). Short precisely, since positive stable distributions belong to the Tweedie family of EDMs with unit variance function μ^p, $p \in (-\infty, 0] \cup [1, \infty)$ (Jørgensen 1997, chap. 4; Dunn and Smyth 2005, for recent developments) the only reasonable mixtures of Poisson with Tweedie mixing distributions are produced for $p \geq 1$ and, then, the unit variance functions of the Poisson-Tweedie family are $\mu + \mu^p \exp\{(2-p)\Phi_p(\mu)\}$ for $\mu > 0$ and negative function $\Phi_p(\mu)$ generally implicit according to $p \geq 1$ (Kokonendji et al. 2004, Propositions 2 and 6). All Poisson-Tweedie distributions are concentrated on $\mathbb{N}$ with explicit density expressions for $p \geq 1$ (Kokonendji et al. 2004, Proposition 3). Their particular cases are Neyman type

A for $p = 1$ (Johnson et al. 1992, pp. 368-), Pólya-Aeppli for $p = 3/2$ (Vinogradov 2006), negative binomial for $p = 2$ (Lawless 1987) and Poisson-inverse Gaussian for $p = 3$ (Dean et al. 1989); the limit case ($p \to \infty$) is easily associated to a Poisson distribution like the Hinde-Demétrio family.

Let us also recall here the meaning of overdispersion and zero-inflation for any count distribution such that its mean is $m > 0$, its variance is $\sigma^2 > 0$ and its proportion of zeros is $p_0 > 0$ (Puig and Valero 2006). In fact, a count distribution is *overdispersed* or Poisson-overdispersed when its variance σ^2 is greater than the variance of a Poisson distribution having the same mean m; hence, its overdispersion index can be defined as $OD = (\sigma^2 - m)/m$. Similarly, a count distribution is said to be *zero-inflated* or Poisson-zero-inflated if its proportion of zeros p_0 exceeds the proportion of zeros of a Poisson distribution having the same mean m, that is $\exp(-m)$. Then the zero-inflation index of this distribution is $ZI = 1 + \log(p_0)/m$. Both indexes are null for a Poisson distribution and positive for any overdispersed and zero-inflated distribution. Then, both properties are closed under independence of effects (Puig 2003) and are explained by mixed Poisson distributions (Feller 1943) such Poisson-Tweedie distributions and by stopped Poisson distributions (Douglas 1980) such Hinde-Demétrio distributions. Only the negative binomial distribution is common to Poisson-Tweedie and Hinde-Demétrio families as for many families of count distributions (with three parameters) having at least one of those properties (Castillo and Pérez-Casany 2005; Kokonendji and Mizère 2005; Walhin and Paris 2006).

The aim of this paper is first to introduce HDRM for count data, for which we can estimate the index parameter p for the adequate additive EDM, and then to study the effect of dispersion parameter ϕ in HDRM when the index parameter p is fixed. The discussion will be essentially on modelling the unit and standard means. Hence, we organise the paper as follows. In Section 2 some main properties on the response distributions are recalled and the HDRMs are defined with the generalized linear modelling framework. Section 3 presents a test for choosing p in $\{2, 3, \cdots\} \cup \{\infty\}$ for the adequate additive EDM in the Hinde-Demétrio family. Section 4 briefly discusses the estimation methods which could be used for a given count additive EDM of the Hinde-Demétrio family when the index parameter p is fixed. Section 5 illustrates the methodology using a horticultural data set. Section 6 concludes.

2 HINDE-DEMÉTRIO REGRESSION MODELS

2.1 Properties of Hinde-Demétrio Models

Given $p \in \{0\} \cup [1, \infty)$ as in (1), the *probability mass function* (pmf) of any Hinde-Demétrio distribution $\mathcal{HD}_p(\theta, \phi)$ is written in form of the additive EDM as

$$P(y; p; \theta, \phi) = A_p(y; \phi) \exp\{\theta y - \phi K_p(\theta)\}, \quad y \in S_{p,} \tag{2}$$

where $\theta \in \Theta_p \subseteq \mathbb{R}$ is the canonical parameter, $\phi > 0$ is the dispersion (or scaling) parameter, $A_p(y;\phi)$ is the normalizing constant, $K_p(\theta)$ is the cumulant function checking $K_p''(\theta) = V_p(K_p'(\theta)) = K_p'(\theta) + (K_p'(\theta))^p$, and the support S_p is such that $S_0 = \{-1\} \cup \mathbb{N}$, $S_1 = 2\mathbb{N}$ and $S_p = p\mathbb{N} \cup \mathbb{N}$ for $p > 1$. Apart from the four distributions obtained for $p \in \{0, 1, 2, 3\}$, none of the Hinde-Demétrio models has explicit pmf even if the cumulant functions can be expressed as:

$$K_p(\theta) = e^{\theta}\, {}_2F_1\left(\frac{1}{p-1}, \frac{1}{p-1}; \frac{p}{p-1}; e^{\theta(p-1)}\right), \quad \theta < 0, \quad p > 1, \tag{3}$$

where ${}_2F_1\,(a, b; c; z) = 1 + (ab/c)(z/1!) + (a(a+1)b(b+1)/c(c+1))(z^2/2!) + \cdots$ is the Gaussian hypergeometric function (Johnson et al. 1992, pp. 17-19). When $p \to \infty$ the limit case $\mathcal{HD}_\infty(\theta, \phi)$ is the Poisson EDM with mean $m = \phi K_\infty'(\theta) = \phi e^\theta$ as described by Jørgensen (1997, pp. 90-92).

For $p > 1$, any Hinde-Demétrio distribution $\mathcal{HD}_p(\theta, \phi)$ is a stopped Poisson distribution. Indeed, let U be a discrete random variable taking its values on

$$U(\Omega) = \{1, p, 2p-1, 3p-2, \cdots\} \tag{4}$$

and such that its *probability generating function* (pgf) is

$$\mathbb{E}(z^U) = c(p, q)\; z\; {}_2F_1\left(\frac{1}{p-1}, \frac{1}{p-1}; \frac{p}{p-1}; (qz)^{p-1}\right), \tag{5}$$

where $q = q(\theta) \in (0, 1)$ is a reparametrization of θ given in (2) and $c(p, q)$ a normalizing constant. We denote $U \sim \mathcal{HD}_p^*(q) = \mathcal{HD}_p^*(\theta)$. Let N_t be a standard Poisson process on the interval $(0, t]$ ($N_0 = 0$) with intensity ϕ [that is $N_t \sim \mathcal{P}(\phi t)$] and supposed to be independent of U. From the pgf of

$$Y_t = \sum_{i=1}^{N_t} U_i = U_1 + \cdots + U_{N_t}, \tag{6}$$

where the U_i are independent and identically distributed as $U \sim \mathcal{HD}_p^*(q(\theta))$, one has $Y_1 \sim \mathcal{HD}_p(q, \phi) = \mathcal{HD}_p(\theta, \phi)$ by fixing the time to $t = 1$ (Kokonendji et al. 2004). From (4-6) one can appreciate different connection of parameters p, θ and ϕ with respect to auxiliary random variables U and N_1. For $p = 2$ we then have a new probabilistic interpretation of the negative binomial distribution $\mathcal{HD}_2(\theta, \phi)$. There the Poisson stopped-sum representation (6) can be also found under the popular name of compound Poisson (Feller 1971; Hinde 1982), but could be confuse (for example to certain mixed Poisson and to the Tweedie distributions with $1 < p < 2$). Another property of Hinde-Demétrio processes (6) is that their modified Lévy measures, which describes the probabilistic character of the jumps of Y_t, always are the negative binomial distributions (up to affinity) for all $p > 1$. See Kokonendji and Khoudar (2006) for more details.

For count Hinde-Demétrio $\mathcal{HD}_p(\theta, \phi)$ distributions, that is $p \in \{2, 3, \cdots\} \cup \{\infty\}$, Kokonendji and Malouche (2005) have shown the following original property: if we

denote $r_y = y\,P(y;p;\theta,\phi)/P(y-1;p;\theta,\phi) = r_y(p;\theta,\phi)$ for all $\theta < 0$ and $\phi > 0$ then

$$r_1 = r_2 = \cdots = r_{p-1} < r_p \neq r_{p+1} > r_1, \quad \forall p \in \{2,3,\cdots\}. \tag{7}$$

These relations hold for all θ and ϕ. The equality part of (7) means that the index parameter $p \in \{2,3,\cdots\}$ of any Hinde-Demétrio distribution is the first integer for which the recursive ratio r_y, $y \in \mathbb{N}^*$, is different from the previous. For the Poisson distribution with mean $m > 0$, we have $r_y = m$ for all $y \in \mathbb{N}^*$ and, therefore, it is regarded as limit of $\mathcal{HD}_p(\theta,\phi)$ when p tends to ∞; as we can also show from (6) with $U = 1$ almost surely. Hence, the two extremities of count Hinde-Demétrio distributions (corresponding to $p = 2$ as negative binomial and $p \to \infty$ as Poisson) belong to the so-called Katz family by Johnson et al. (1992, p. 78). The property (7) is characteristic of count Hinde-Demétrio distributions and constituted the point of departure for evaluating $p \in \{2,3,\cdots\}$ by statistical tests which do not depend on parameters θ and ϕ; see Section 3 below.

The last serie of properties is classical but most important for defining the regression models. The first is that any Hinde-Demétrio model is closed under convolution: $\mathcal{HD}_p(\theta,\phi_1) * \mathcal{HD}_p(\theta,\phi_2) = \mathcal{HD}_p(\theta,\phi_1+\phi_2)$ for $\phi_1,\phi_2 > 0$. Then, for fixed $p \in \{0\}\cup[1,\infty)$, the expectation and variance of $Y \sim \mathcal{HD}_p(\theta,\phi)$ are $\mathbb{E}_{p,\theta,\phi}(Y) = \phi K_p'(\theta)$ and $var_{p,\theta,\phi}(Y) = \phi K_p''(\theta) = \phi V_p(K_p'(\theta))$, respectively. Since $\theta \mapsto K_p(\theta)$ is strictly convex for $\theta < 0$ and from (1), we firstly reparametrize the additive $\mathcal{HD}_p(\theta,\phi)$ by its unit mean $\mu = K_p'(\theta)$ as $\mathcal{HD}_p(\mu,\phi)$ for which the unit variance function $V_p(\mu)$ is proportional to the dispersion parameter ϕ like a "reproductive" EDM:

$$\mu = K_p'(\theta) = \phi^{-1}\mathbb{E}_{p,\theta,\phi}(Y) \;\text{ and }\; var_{\mu,\phi}(Y) = \phi V_p(\mu) = (\mu+\mu^p)\phi, \tag{8}$$

where $\phi > 0$ and $\mu = \mu(p;\theta) > 0$ (not depending on ϕ). Secondly, we use the standard mean parametrization $\mathcal{HD}_p(m,\phi)$ of the additive $\mathcal{HD}_p(\theta,\phi)$ as follows:

$$m = \phi K_p'(\theta) = \mathbb{E}_{p,\theta,\phi}(Y) \;\text{ and }\; var_m(Y) = \phi V_p(m/\phi) = m + \phi^{1-p}m^p, \tag{9}$$

where $\phi > 0$ and $m = m(p;\theta,\phi) > 0$. A reason behind the two mean parametrizations (8) and (9) is essentially due to the pratical use of the variance-to-mean relationship, which must provide the same behaviour in presence of data as ϕ is near of 1 for given p. Finally, it is also known that all Hinde-Demétrio family has both overdispersed [e.g., from (1): $V_p(\mu) > \mu > 0$] and zero-inflated [e.g., from (6) and Douglas (1980)] distributions with respect to Poisson distribution. From (2) the characteristic indexes are $OD(p,\theta,\phi) = [K_p'(\theta)]^{p-1}$ and $ZI(p,\theta,\phi) = 1 - K_p(\theta)/K_p'(\theta)$ respectively and, they are positive and do not depend on ϕ.

2.2 Hinde-Demétrio Regression Models (HDRMs)

Let Y be a single count response variable and let $\mathbf{x}$ be an associated vector of covariates with a vector $\boldsymbol{\beta}$ of unknown regression coefficients. The first HDRM for Y on $\mathbf{x}$ is defined from (8) as $Y \sim \mathcal{HD}_p(\mu,\phi)$, where $\mu = \mu(\mathbf{x};\boldsymbol{\beta})$ is a positive-valued function related to $\mathbf{x}$ and to $\boldsymbol{\beta}$ by a link function (McCullagh and Nelder

1989; Jørgensen 2001), $\phi > 0$ and $p \in \{2, 3, \cdots\}$. For convenience we can write $Y \sim \mathcal{HD}_p(\mu(\mathbf{x}; \boldsymbol{\beta}), \phi)$ to denote this HDRM for which the dispersion parameter ϕ is not connected to the covariates $\mathbf{x}$ via the unit mean $\mu = \mu(\mathbf{x}; \boldsymbol{\beta})$ but through the variance $var(Y) = (\mu + \mu^p)\phi$ of Y. From (1) and Kokonendji et al. (2004, formula (16)), its canonical link function is given, for all $p \in \{2, 3, \cdots\}$, as:

$$\mu = \exp(\mathbf{x}^T\boldsymbol{\beta})[1 - \exp\{(p-1)\mathbf{x}^T\boldsymbol{\beta}\}]^{-1/(p-1)}. \tag{10}$$

In practice we can use the common log-linear link function of count data $\mu = \exp(\mathbf{x}^T\boldsymbol{\beta})$, which does not depend on p and is also obtained from (10) when $p \to \infty$.

Similarly we define the second HDRM for Y on $\mathbf{x}$ as $Y \sim \mathcal{HD}_p(m, \phi)$ such that ϕ is, theoretically, connected both to the standard mean $m = m(\mathbf{x}; \boldsymbol{\beta})$ and to the variance $var(Y) = m + \phi^{1-p}m^p$ of Y. In Kokonendji and Marque (2005), this model $Y \sim \mathcal{HD}_p(m(\mathbf{x}; \boldsymbol{\beta}), \phi)$ is used for $p \in \{2, 3\}$ with the usual log-linear link function $\log(m) = \mathbf{x}^T\boldsymbol{\beta}$ and the maximum likehood method.

For this work we do not consider the dispersion parameter modelling $\phi = \phi(\mathbf{z}; \boldsymbol{\gamma})$, where $\mathbf{z}$ is a second vector of covariates (not necessarily independent of the first $\mathbf{x}$ and through to affect the dispersion) and $\boldsymbol{\gamma}$ is the corresponding vector of unknown regression coefficients. That would lead to a kind of Double Generalized Linear Model. See, for example, Smyth and Jørgensen (2002) for a Tweedie regression model. Also, it is not envisaged to modelling the index parameter as $p = p(\mathbf{w}; \boldsymbol{\delta})$. Finally, a HDRM can be defined in different ways according to the modelling parameters as above and also to the parametrizations, namely $\mathcal{HD}_p(\mu, \phi)$ from (8), $\mathcal{HD}_p(m, \phi)$ from (9), $\mathcal{HD}_p(\theta, \phi)$ from (2) and $\mathcal{HD}_p(q, \phi)$ from (6). Of course these definition could be extended to the others, likely Generalized Linear Mixed Model (Hinde and Demétrio 1998, chap. 6, and references therein) or Double Hierarchical Generalized Linear Models by Lee and Nelder (2006) for more ideas.

3 CHOICE OF THE RESPONSE EDM IN HDRMs

Consider a HDRM with unknown index parameter $p \in \{2, 3, \cdots\} \cup \{\infty\}$. Without loss of generality, let $\mathbf{Y} = (Y_1, \cdots, Y_n)$ be a vector of random sample of count response with $Y_i \sim \mathcal{HD}_p(\mu_i(\mathbf{x}; \boldsymbol{\beta}), \phi)$, $i = 1, \cdots, n$. To evaluate the adequate $p \in \{2, 3, \cdots\}$ for this HDRM, we first adapt the statistical test procedure developed by Kokonendji and Malouche (2005) with no covariates and then, if it is necessary, we use a criterion of model selection (Akaike 1973; Schwarz 1978; Linhart and Zucchini 1986; Pan 2001) to determine the appropriated value of p.

More precisely, from (7) we investigate the $p(p-1)/2$ testing problems

$$H_{0p}^{ij} : r_i = r_j \quad \text{versus} \quad H_{1p}^{ij} : r_i < r_j \quad (1 \le i < j \le p), \tag{11}$$

which are tested individually at significance level $\alpha \in (0, 1)$. Starting with $p = 2$, we stop at the first $p \in \{2, 3, \cdots\}$, denoted by $\widehat{p} = \widehat{p}(\mathbf{Y}, \alpha)$, for which one of

alternative hypothesis H_{1p}^{ip} $(i = 1, \cdots, p-1)$ is accepted, i.e., the corresponding individual p-value $\widehat{\alpha}_{ip}$ is smaller than α. Thus, if we hesitate between two values ($\widehat{p}_1$ and $\widehat{p}_2 = \widehat{p}_1 + 1$) of p according to the choice of the level α, we can use an indicator of model selection (e.g., log-likelihood or deviance criterion) to choose between them. This last part depends on the methods used to estimate $\boldsymbol{\beta}$ and ϕ in the HDRM $Y_i \sim \mathcal{HD}_p(\mu_i(\mathbf{x};\boldsymbol{\beta}), \phi)$; see the next section. So, in this section we only present the theory necessary to the determination of $\widehat{p}$ based on (11). See Kokonendji and Malouche (2005) for details on simulation studies and applications without covariates, where the criterion of model selection is the adequation chi-squared test of Pearson.

Let $\mathbf{Y} = (Y_1, \cdots, Y_n)$ be a vector of random sample of count response. Denote, for all $y \in \mathbb{N}$, $F_y = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{\{Y_i = y\}} = F_y(\mathbf{Y})$ the random sample relative frequency. Following Kokonendji and Malouche (2005), we consider the test stastistics of (11)

$$T_{ij} = \frac{i\, F_i\, F_{j-1}}{j\, F_{i-1}\, F_j} = T_{ij}(\mathbf{Y}) \quad (1 \leq i < j \leq p) \tag{12}$$

such that their asymptotic normalities are obtained by the classical frequency substitution method (Bickel and Doksum 1977) as:

Proposition 1 *Suppose $p_0, p_1, \cdots, p_k$ are (population) proportions with $p_l = \Pr(Y = l), l = 0, 1, \cdots, k-1$, $p_k = \Pr(Y \geq k)$, and $k \geq p \in \{2, 3, \cdots\}$. Then the statistics $T_{ij} = T_{ij}(\mathbf{Y})$ defined in (12) are asymptotically normal:*

$$\sqrt{n}(T_{ij} - r_i/r_j)/\sigma_{ij} \to \mathcal{N}(0,1) \quad as \;\; n \to \infty \quad (1 \leq i < j \leq p),$$

where $\sigma_{ij}^2 := var(T_{ij}) = \sigma_{ij}^2(p_0, p_1, \cdots, p_k)$ is given by

$$\sigma_{ij}^2 = \begin{cases} (r_i/r_{i+1})^2 \,(1/p_{i-1} + 4/p_i + 1/p_{i+1}) & \text{for } j = i+1 \\ (r_i/r_j)^2 \,(1/p_{i-1} + 1/p_i + 1/p_{j-1} + 1/p_j) & \text{for } j > i+1. \end{cases} \tag{13}$$

Remark 2 *(i) A consistent estimator of σ_{ij}^2 is usually provided by $\widehat{\sigma}_{ij}^2 = \sigma_{ij}^2(F_0, F_1, \cdots, F_k)$ with $F_k = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{\{Y_i \geq k\}}$. (ii) When $\widehat{p}$ is large (or $\widehat{p} \approx k$), we can decide that $Y \sim \mathcal{HD}_\infty(\mu, \phi)$, which is the Poisson distribution of mean $m = \phi\mu$. (iii) The frequent situation $\widehat{p} = 2$ means two things: p really equals to 2 (e.g. $Y_i \sim \mathcal{HD}_2(\mu_i(\mathbf{x};\boldsymbol{\beta}), \phi)$ as negative binomial model) or not (i.e. Y_i follows a Poisson model or another model which is not a HDRM); thus, the test $p = 2$ has general interest for count data.*

Another way to estimate $p \in \{2, 3, \cdots\}$ is to use the moment methods. Indeed, if $Y \sim \mathcal{HD}_p(\mu, \phi)$ or $Y \sim \mathcal{HD}_p(m, \phi)$ with $\sigma^2 = var(Y)$ then we easily get p from (8) as $p_1(\phi) = \log(\sigma^2/\phi - \mu)/\log(\mu)$ and from (9) as $p_2(\phi) = \log((\sigma^2 - m)/\phi)/\log(m/\phi)$, respectively. For fixed $\phi = 1$ we have $\mu = m$ and, then,

$$p_1(1) = p_2(1) = \frac{\log(\sigma^2 - m)}{\log(m)} = p. \tag{14}$$

Thus we obtain a real estimate p^* which lies between two integers $p_1^* = \lfloor p^* \rfloor$ and $p_2^* = p_1^* + 1$, where $\lfloor a \rfloor$ denotes the integer part of $a \in \mathbb{R}$. As presented above, i.e. after applying the twice corresponding HDRM with p_1^* and p_2^*, we must use here an indicator of model selection to decide "the" adequate p^* between them. Kokonendji et al. (2004) used this idea without covariates and, then, one could also use the criterion of the adequation chi-squared test of Pearson to select $p \in \{2, 3, \cdots\}$. However, an estimator by moment methods (14) is generally known to be inefficient, but sometimes more useful as correct indicator of the parameter p when the first p^* from (14) belongs to $[2, \infty)$.

4 BACKGROUND FOR ESTIMATING IN HDRMs

Now we consider a HDRM such that the index parameter $p \in \{2, 3, \cdots\}$ is fixed or estimated following the previous section. According to the value of p one can use various approaches, namely maximum likelihood, (extended) quasi-likelihood, pseudo-likelihood and moment methods. See, for example, Hinde and Demétrio (1998, chap. 3) for a summary. For both models (8) and (9), and in view of application in the next section, we here present the basic materials of likelihood and deviance which are the *unit pmf* $P^*(y; p; \mu, \phi)$ from (2) and the *unit deviance function*

$$D_p^*(y; \mu) := -2 \int_y^\mu \frac{y - t}{V_p(t)} \, dt = -2 \int_y^\mu \frac{y - t}{t + t^p} \, dt, \tag{15}$$

respectively (Jørgensen 1997). In fact, if the considered HDRM is from (8), then the pmf of $Y \sim \mathcal{HD}_p(\mu, \phi)$ is $P^*(\phi y; p; \mu, \phi)$ and its deviance function is $\phi^{-1} D_p^*(y; \mu)$. Also, if the HDRM is defined from (9) like $Y \sim \mathcal{HD}_p(m, \phi)$, then its pmf is given by $P^*(y; p; m\phi^{-1}, \phi)$ and its deviance function is $\phi^{-1} D_p^*(y; m\phi^{-1})$. It is worth noting for correct use of the models in terms of the dispersion parameter ϕ.

Letting $\mu = K_p'(\theta)$ and, then, $B_p(\mu) = \exp(\theta)$. From (2-3) the unit pmf $P^*(y; p; \mu, \phi) := P(y; p; \log B_p(\mu), \phi)$ can be expressed, for all $y \in \mathbb{N}$ and $p \in \{2, 3, \cdots\}$, as:

$$\begin{aligned} P^*(y; p; \mu, \phi) = {} & A_p(y; \phi) \, [B_p(\mu)]^y \\ & \times \exp\left\{ -\phi B_p(\mu) \; {}_2F_1\left(\frac{1}{p-1}, \frac{1}{p-1}; \frac{p}{p-1}; [B_p(\mu)]^{p-1} \right) \right\}, \end{aligned} \tag{16}$$

where $A_p(y; \phi)$ remains as in (2) and, from Kokonendji et al. (2004, formula (16)),

$$B_p(\mu) = \mu \left(1 + \mu^{p-1} \right)^{-1/(p-1)}. \tag{17}$$

For instance, that gives *explicitly* (Johnson et al. 1992, p. 18) the negative binomial case with $p = 2$ as

$$P^*(y; 2; \mu, \phi) = \frac{\Gamma(y + \phi)}{y! \Gamma(\phi)} \left(\frac{\mu}{1 + \mu} \right)^y \left(\frac{1}{1 + \mu} \right)^\phi, \quad y \in \mathbb{N},$$

and the strict arcsine case with $p = 3$ as

$$P^*(y; 3; \mu, \phi) = \frac{A_3(y; \phi)}{y!} \left(\frac{\mu}{\sqrt{1+\mu^2}} \right)^y \exp\left\{ -\phi \arcsin \frac{\mu}{\sqrt{1+\mu^2}} \right\}, \quad y \in \mathbb{N},$$

where $A_3(y; \phi)$ is given in Letac and Mora (1990) by

$$A_3(y; \phi) = \begin{cases} \prod_{k=0}^{z-1} (\phi^2 + 4k^2) & \text{if } y = 2z, \text{ and } A_3(0; \phi) = 1 \\ \phi \prod_{k=0}^{z-1} [\phi^2 + (2k+1)^2] & \text{if } y = 2z+1, \text{ and } A_3(1; \phi) = \phi. \end{cases}$$

Concerning to the unit deviance function (15) which can be written as

$$\begin{aligned} D_p^*(y; \mu) &= 2 \left[y \log \left(\frac{B_p(y)}{B_p(\mu)} \right) + \int_y^\mu \frac{dt}{1 + t^{p-1}} \right] \\ &= 2 \left[y \log \left(\frac{B_p(y)}{B_p(\mu)} \right) + \frac{\mu \Phi(-\mu^{p-1}, 1, 1/(p-1)) - y \Phi(-y^{p-1}, 1, 1/(p-1))}{p-1} \right], \end{aligned} \tag{18}$$

where $\Phi(a, s, b) = \sum_{k \geq 0} a^k / (k+b)^s$ is the Lerch Φ function (Erdélyi et al. 1955), we obtain from (17) the two only simple cases:

$$D_p^*(y; \mu) = \begin{cases} 2 \left[y \log \left\{ y(1+\mu)[\mu(1+y)]^{-1} \right\} + \log \left\{ (1+\mu)/(1+y) \right\} \right] & \text{for } p = 2 \\ 2 \left[y \log \left\{ (y\sqrt{1+\mu^2})/(\mu\sqrt{1+y^2}) \right\} + \arctan \mu - \arctan y \right] & \text{for } p = 3. \end{cases}$$

A computer algebra program (e.g. Maple) can provide long expressions of $D_p^*(y; \mu)$ for $p = 4, 5, 6$. Both expressions (16) and (18) show the need of good approximation technics when we use a count HDRM with $p \in \{4, 5, \cdots\}$ and quasi- and likelihood methods. See, for instance, Dossou-Gbété et al. (2006).

5 ILLUSTRATIVE EXAMPLE

5.1 Data Set

We consider the data used by Ridout et al. (1998, tables 1 and 2); see also Ridout et al. (2001, table 2). However, we examine them on a different way for studing the behaviour of HDRM when the index parameter p belongs to $\{2, 3\} \cup \{\infty\}$, which are extremities of its domain. Table 1 gives the number of roots produced by 270 micropropagaged shoots of the columnar apple cultivar Trajan and some related statistics. During the rooting period, all shoots were maintained under identical conditions, but the shoots themselves had been produced under an 8- or 16-hour photoperiod in culture systems. For each of two photoperiods, it was used one of four different concentrations of the cytokinin BAP in the culture medium. There were 140 shoots produced under the 8-hour photoperiod, which are weakly overdipersed,

as also shown in Figs. 1 and 2, and quasi-none zero-inflated. However, the other 130 shoots produced under the 16-hour photoperiod are really overdispersed (see, e.g., Figs. 3 and 4) with an important excess of zeros. Ridout et al. (1998) analyzed these phenomena by fitting various models to these data as a whole, based on the Poisson and negative binomial distributions and their so-called zero-inflated contreparts. They had concluded on the different effect of these two photoperiods.

Table 1 about here
Figures 1, 2, 3 and 4 about here

5.2 Methodology and Results

Here we use two HDRMs separately for each one of the two photoperiods 8 and 16 for pointing out the effect of dispersion parameter through the double mean parametrization (8) and (9). All computations are done using the R software (R Development Core Team 2005; Kuhnert and Venables 2005; Venables and Ripley 2002). The corresponding p-value(t_{12}) in Table 1 suggest that we have $\widehat{p} = 2$ for the two HDRMs [see Remark 2 (*iii*)]. However, the corresponding values of p^* allow to consider mainly two models (in the extended HDRM) between, first, $p = 2$ and $p = \infty$ for the 8-hour photoperiod, and second, $p = 2$ and $p = 3$ for the 16-hour photoperiod, respectively. For all models $HD_p(\mu(\mathbf{x};\boldsymbol{\beta}), \phi)$ and $HD_p(m(\mathbf{x};\boldsymbol{\beta}), \phi)$, we consider the common log-linear link function

$$\mu = \exp(\mathbf{x}^T\boldsymbol{\beta}) = m \quad \text{with} \quad \mathbf{x}^T\boldsymbol{\beta} = \beta_0 + \beta_1\mathbf{x} + \beta_2\mathbf{x}^2 + \beta_3\mathbf{x}^3,$$

and only the variances change following (8) and (9). Hence, for fixed $p \in \{2, 3, \cdots\}$, the adequate model would produce $\widehat{\phi}$ near of 1. Note that, following Jørgensen (1997, pp. 90-92), the Poisson HDRM is $HD_\infty(m) = HD_\infty(\mu, 1) = HD_\infty(m, 1)$, without dispersion parameter ϕ.

For fixed $p \in \{2, 3\} \cup \{\infty\}$, full maximum likelihood estimation can be used because the corresponding HDRMs present a complete and explicit (unit) pmf for the response variable. For a future use and since the unit variance function (1) has a simple form, we here describe the quasi-likelihood and related methods (Hinde and Demétrio 1998, chap. 3). Hence, for quasi-likelihood method, only deviance criterion for comparing the performance of models must be computed and we could use a modified Akaike's Information Criterion (AIC) which would provide the same behaviour (Pan 2001; Lee and Nelder 2003).

The principle of the quasi-likelihood method for a model (8) with variance of the form $var(Y_i) = \phi V_p(\mu_i)$ is to estimate the regression parameters by maximizing the quasi-likelihood

$$Q_p = -\frac{1}{2}\sum_{i=1}^{n}\frac{D_p^*(y_i;\mu_i)}{\phi},$$

where D_p^* is the unit deviance function (15) and (18). The regression parameter estimates $\widehat{\boldsymbol{\beta}}$ are the same to those for the respective non-dispersed model (here

Poisson) and the (over)dispersion parameter ϕ is estimated by equating the Pearson X^2 statistic to the residual degrees of freedom $n - r$:

$$\widehat{\phi} = \frac{1}{n-r} \sum_{i=1}^{n} \frac{(y_i - \widehat{\mu}_i)^2}{\widehat{\mu}_i}.$$

The estimation of $\boldsymbol{\beta}$ and ϕ is asymptotically independent.

Concerning to the model (9) or (8) which can be described with a variance of the form $var(Y_i) = \phi_i(\phi; m_i; p)V(m_i)$, where both the scale parameter ϕ_i and the variance function $V(.)$ may depend upon additional parameters. One suggests estimating the unknown parameters in the mean ($\boldsymbol{\beta}$) and in the variance model (ϕ) by maximizing the *extended quasi-likelihood* (EQL) function

$$Q_p^+ = -\frac{1}{2} \sum_{i=1}^{n} \left\{ \frac{D(y_i; m_i)}{\phi_i} + \log\left[2\pi\phi_i V(y_i)\right] \right\},$$

where $D(y; m) := -2\int_y^m (y - t)dt/V(t)$ is a kind of unit deviance function. Using EQL for the HDRMs (8) and (9), like negative binomial ($p = 2$) and stirct arcsine ($p = 3$) variance functions, we have the following variance decompositions

$$\begin{array}{lll} \text{for (8):} & \phi_i = (1 + \mu_i^{p-1})\phi \quad \text{and} & V(\mu_i) = \mu_i; \\ \text{for (9):} & \phi_i = 1 + m_i^{p-1}/\phi^{p-1} \quad \text{and} & V(m_i) = m_i. \end{array}$$

The regression parameter estimates $\widehat{\boldsymbol{\beta}}$ are here identical to those for the respective non-dispersed model (here Poisson with $\phi_i = 1$) and are obtained by estimating equations for a weighted Poisson model with weights $1/\phi_i$. We can obtain an estimate for ϕ by fitting a gamma model using the Poisson deviance components as y-variable, an identity link and taking the linear model to have, for (8), a linear regression model without intercept with $1+\mu_i^{p-1}$ as explanatory variable, and, for (9), a fixed intercept (offset) of 1 and m_i^{p-1} as explanatory variables. An approximate standard error is obtained for ϕ by setting the scale to 2, corresponding to modelling χ_1^2.

Table 2 about here
Table 3 about here

The results presented in Table 2 and Table 3 were obtained by EQL method, which quickly converged. They are very interesting and suitable. Indeed, the regression coefficient estimates $\widehat{\beta}_j(8)$ and $\widehat{\beta}_j(16)$ are identical for the five proposed models with respect to the photoperiods 8 and 16, and only standard errors change (Table 3). The reason comes from the residual deviances and, therefore, the dispersion parameter $\widehat{\phi}$ in Table 2. In fact, with respect to each photoperiod and except the Poisson model $HD_\infty(m)$ which has poor fitting, we must compare the deviances or AIC *within* [i.e. $HD_2(\mu(\mathbf{x}; \boldsymbol{\beta}), \phi)$ to $HD_2(m(\mathbf{x}; \boldsymbol{\beta}), \phi)$, and $HD_3(\mu(\mathbf{x}; \boldsymbol{\beta}), \phi)$ to $HD_3(m(\mathbf{x}; \boldsymbol{\beta}), \phi)$] and also *between* [i.e. $HD_2(\mu(\mathbf{x}; \boldsymbol{\beta}), \phi)$ to $HD_3(\mu(\mathbf{x}; \boldsymbol{\beta}), \phi)$, and $HD_2(m(\mathbf{x}; \boldsymbol{\beta}), \phi)$ to $HD_3(m(\mathbf{x}; \boldsymbol{\beta}), \phi)$] the models, respectively. Hence we can observe different effects or estmated values of ϕ near 1 or not for the models (McCullagh and Nelder 1989, p. 400). Note in passing that, for the photoperiod 8, all regression coefficient estimates

$\widehat{\beta}_j(8)$ are significant but, for the photoperiod 16, only the intercept coefficient $\widehat{\beta}_0(16)$ is significant for the different models.

For the photoperiod 16, the effect (or value near of 1) of dispersion parameter ϕ is insignificant within the negative binomial ($p = 2$) models, compared to the strict arcsine ($p = 3$) models. This suggests that the negative binomial model can be "the" best model among the HDRMs for this data set. As for the photoperiod 8, the effect of dispersion parameter is sizeable both within and between the negative binomial ($p = 2$) and strict arcsine ($p = 3$) models. Hence, also from the end of Remark 2 (*iii*), the regression analysis of this photoperiod 8 could be improved by other models which do not belong to the HDRMs (Ridout et al. 1998, 2001; Walhin and Paris 2006; Kokonendji and Mizère 2005, for a future use of weighted Poisson models taking into account both over- and underdispersion situations).

6 CONCLUDING REMARKS

HDRMs cover a new broad family of the overdispersed count data regression models and, among other models, provide an alternative to the mixtures of Poisson models (Hougaard et al. 1997; Walhin and Paris 2006). They can be applied in various domains like agriculture, finance, epidemiology and ecology. For any index parameter p fixed in $\{2, 3, \cdots\} \cup \{\infty\}$, the EQL method is numerically efficient for estimating parameters in the HDRMs. The EQL could be generalized when p is unknown. For instance, the maximum likelihood and quasi-likelihood methods can be used for $p \in \{2, 3\} \cup \{\infty\}$. In this paper we have presented a statistical evaluation of the index parameter $p \in \{2, 3, \cdots\} \cup \{\infty\}$, which is arbitrarily chosen in similar classes of models (e.g. Walhin and Paris 2006; Dunn and Smyth 2005, for the Tweedie models). Note that another way than the moment method (14) to evaluate the index parameter p would be to find the p to minimize a distance between the empirical distribution and the Hinde-Demétrio distribution.

For given $p \in \{2, 3, \cdots\}$ and from the mean parametrizations (8) and (9), the effect of dispersion parameter $\phi > 0$ according to the unit and standard means modelling is more important within and between the models when the HDRMs are not appropriated to the (overdispersed) data set. Conversely, if the dispersion parameter effect is negligible within a reasonable HDRM for a given $p \in \{2, 3, \cdots\}$, that is ϕ near of 1, then this model is the best one. The same log-linear link function used in this application for both the unit and standard mean parametrizations, clearly, allows the within and between comparison of the models. In many count data regressions using an additive EDM for the response distribution, we must have the same behaviour as in HDRMs. The standard mean parametrization (9) is generally and commonly used in lieu of the unit mean parametrization (8), which theoretically appears appropriate to many situations and present some pratical advantages.

Acknowledgments. This research was essentially done while the first author was visiting USP/ESALQ that he has always appreaciated the hospitality. We are very

grateful to FAPESP, CNPq, CCInt/USP and LMA/CNRS for funding that visit.

BIBLIOGRAPHICAL REFERENCES

Akaike, H. (1973), "Information Theory and an Extension of the Maximum Likelihood Principle," by Petrov, B. N. and Csáki, F. (Eds), *Second International Symposium on Inference Theory*, pp. 267-281, Budapest: Akadémiai Kiadó. [Reprinted by Samuel Kotz and Norman L. Johnson (eds.), *Breakthroughs in Statistics* (Vol. I), New York: Springer-Verlag (1992 ed.), pp. 599-624 (with an introduction by J. deLeeuw)].

Bickel, P. J., and Doksum, K. A. (1977), *Mathematical Statistics*, California - Oakland: Holden - Day.

Bosch, R. J., and Ryan, L. M. (1998), "Generalized Poisson Models Arising From Markov Processes," *Statistics and Probability Letters,* 39, 205-212.

Cameron, A. C., and Trivedi, P. K. (1998), *Regression Analysis of Count Data,* Cambridge - UK: Cambridge University Press.

Castillo, J., and Pérez-Casany, M. (2005), "Overdispersed and Underdispersed Poisson Generalizations," *Journal of Statistical Planning and Inference,* 134, 486-500.

Cox, D. R., and Reid, N. (1987), "Parameter Orthogonality and Approximate Conditional Inference" (with discussion), *Journal of the Royal Statistical Society,* Ser. B, 49, 1-39.

Dean, C., Lawless, J. F., and Willmot, G. E. (1989), "A Mixed Poisson-Inverse Gaussian Regression Model," *The Canadian Journal of Statistics,* 17, 171-181.

Dossou-Gbété, S., Demétrio, C. G. B., and Kokonendji, C. C. (2006), "An MM-Algorithm for a Class of Overdispersed Regression Models," in *Proceedings of the 9th International Conference Zaragoza-Pau on Applied Mathematics and Statistics*, Jaca (Spain), September 19-21th 2005 (to appear).

Douglas, J. B. (1980), *Analysis With Standard Contagious Distributions,* Fairland - Md: International Cooperative Publishing House.

Dunn, P. K., and Smyth, G. K. (2005), "Series Evaluation of Tweedie Exponential Dispersion Model Densities," *Statistics and Computing,* 15, 267-280.

Erdélyi, A., Magnus, W., Oberhettinger, F., and Tricomi, F. G. (1955), *Higher Transcendental Functions* (vol. III), New York: McGraw-Hill.

Feller, W. (1943), "On a General Class of Contagious Distributions," *The Annals of Mathematical Statistics,* 14, 389-400.

——— (1971), *An Introduction to Probability Theory and its Applications* (vol. 2, 2nd ed.), New York: Wiley.

Hall, D. B., and Berenhaut, K. S. (2002), "Score Tests for Heterogeneity and Overdispersion in Zero-Inflated Poisson and Binomial Regression Models, *The Canadian Journal of Statistics,* 30, 1-16.

Hinde, J. (1982), "Compound Poisson Regression Models," by R. Gilchrist (Ed.), *GLIM82*, pp. 109-121, New York: Springer-Verlag.

Hinde, J., and Demétrio, C. G. B. (1998), *Overdispersion: Models and Estimation,* São Paulo: Associação Brasileira de Estatística.

Hougaard, P., Lee, M-L. T., and Whitmore, G. A. (1997), "Analysis of Overdispersed Count Data by Mixtures of Poisson Variables and Poisson Processes," *Biometrics*, 53, 1225-1238.

Johnson, N. L., Kotz, S., and Kemp, A. W. (1992), *Univariate Discrete Distributions* (2nd ed.), New York: John Wiley & Sons.

Jørgensen, B. (1997), *The Theory of Dispersion Models,* London: Chapman & Hall.

——— (2001), "Generalized Linear Models," Research Report No. 33, Odense University, Dept. of Statistics and Demography [Contribution to *Encyclopedia of Environmetrics*, to be published by Wiley, Chichester].

Kokonendji, C. C., Demétrio, C. G. B., and Dossou-Gbété, S. (2004), "Some Discrete Exponential Dispersion Models: Poisson-Tweedie and Hinde-Demétrio Classes," *Statistics and Operations Research Transactions,* 28, 201-214.

Kokonendji, C. C., and Khoudar, M. (2004), "On Strict Arcsine Distribution," *Communication in Statistics, Part A - Theory and Methods,* 33, 993-1006.

——— (2006), "On Lévy Measure for Infinitely Divisible Natural Exponential Families," *Statistics and Probability Letters*, in press.

Kokonendji, C. C., and Malouche, D. (2005), "A Property of Count Distributions in the Hinde-Demétrio Family," Technical Report No. 0421, University of Pau, Laboratory of Appl. Math.(submitted for publication).

Kokonendji, C. C., and Marque, S. (2005), "A Strict Arcsine Regression Model," *Advances in Mathematics - African Diaspora Journal of Mathematics,* 1, 85-92.

Kokonendji, C. C., and Mizère, D. (2005), "Overdispersion and Underdispersion Characterization of Weighted Poisson Distributions," Technical Report No. 0523, University of Pau, Laboratory of Appl. Math. (submitted for publication).

Kuhnert, P., and Venables, B. (2005), *An Introduction to R: Software for Statistical Modelling & Computing,* Piracicaba - SP: ESALQ/USP Silvio S. Zocchi (ed.).

Lawless, J. F. (1987), "Negative Binomial and Mixed Poisson Regression," *The Canadian Journal of Statistics*, 15, 203-225.

Lee, Y., and Nelder, J. A. (2003), "Extended-REML Estimators," *Journal of Applied Statistics*, 30, 845-856.

——— (2006), "Double Hierarchical Generalized Linear Models" (with discussion), *Journal of the Royal Statistical Society,* Ser. C, 55 (2), in press.

Letac, G., and Mora, M. (1990), "Natural Real Exponential Families With Cubic Variance Functions," *The Annals of Statistics*, 18, 1-37.

Linhart, H., and Zucchini, W. (1986), *Model Selection,* New York: Wiley.

McCullagh, P., and Nelder, J. A. (1989), *Generalized Linear Models* (2nd ed.), London: Chapman & Hall.

Mullahy, J. (1997), "Heterogeneity, Excess Zeros, and the Structure of Count Data Models," *Journal of Applied Econometrics*, 12, 337-350.

Pan, W. (2001), Akaike's Information Criterion in Generalized Estimating Equations, *Biometrics*, 57, 120-125.

Puig, P. (2003), "Characterizing Additively Closed Discrete Models by a Property of their Maximum Likelihood Estimators With Application to Generalized Hermite Distributions," *Journal of the American Statistical Association*, 98, 687-692.

Puig, P., and Valero, J. (2006), "Count Data Distributions: Some Characterizations With Applications," *Journal of the American Statistical Association*, 101, 332-340.

R Development Core Team (2005), *R: A Language and Environment for Statistical Computing,* R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Ridout, M., Demétrio, C. G. B., and Hinde, J. (1998), "Models for Count Data With

Many Zeros," in *Proceedings of the XIXth International Biometrics Conference*, Cape Town, Invited Papers, pp. 179-192.

Ridout, M., Hinde, J., and Demétrio, C. G. B. (2001), "A Score Test for Testing a Zero-Inflated Poisson Regression Model Against Zero-Inflated Negative Binomial Alternatives," *Biometrics,* 57, 219-223.

Schwarz, G. (1978), "Estimating the Dimension of a Model," *The Annals of Statistics*, 6, 461-464.

Smyth, G. K., and Jørgensen, B. (2002), "Fitting Tweedie's Compound Poisson Model to Insurance Claims Data: Dispersion Modelling," *Astin Bulletin*, 32, 143-157.

Venables, W. N., and Ripley, B. D. (2002), *Modern Applied Statistics With S* (4th ed.), New York: Springer-Verlag.

Vinogradov, V. (2006), "On Infinitely Divisible Exponential Dispersion Model Related to Poisson-Exponential Distritbution," *Communication in Statistics, Part A - Theory and Methods,* submitted for publication.

Walhin, J. F., and Paris, J. (2006), "A Mixed Poisson Model With Varying Element Sizes," *Statistical Methodology,* in press.

Table 1
Frequency distributions of the number of roots produced by 270 shoots of the apple cultivar Trajan, classified by experimental conditions (photoperiod and BAP concentration) under which the shoots were reared; shown are the numbers of shoots that produced $0, 1, \cdots, 12$ roots; counts that exceeded 12 are shown individually (Redout *et al.*, 1998, 2001). Some statistics are also calculated with $OD = (\widehat{\sigma}^2 - \widehat{m})/\widehat{m}$, $ZI = 1 + \log(f_0)/\widehat{m}$, $p^* = \log(\widehat{\sigma}^2 - \widehat{m})/\log(\widehat{m})$ from (14) and p-value(t_{12}) from (12).

Photoperiod	8	8	8	8		16	16	16	16	
Bap (μM)	2.2	4.4	8.8	17.6		2.2	4.4	8.8	17.6	
No. of roots					Total 8					Total 16
0	0	0	0	2	2	15	16	12	19	62
1	3	0	0	0	3	0	2	3	2	7
2	2	3	1	0	6	2	1	2	2	7
3	3	0	2	2	7	2	1	1	4	8
4	6	1	4	2	13	1	2	2	3	8
5	3	0	4	5	12	2	1	2	1	6
6	2	3	4	5	14	1	2	3	4	10
7	2	7	4	4	17	0	0	1	3	4
8	3	3	7	8	21	1	1	0	0	2
9	1	5	5	3	14	3	0	2	2	7
10	2	3	4	4	13	1	3	0	0	4
11	1	4	1	4	10	1	0	1	0	2
12	0	0	2	0	2	1	1	1	0	3
> 12	13, 17	13	14, 14	14	6					
No. of shoots	30	30	40	40	140	30	30	30	40	130
Mean: $\widehat{m}$	5.8	7.8	7.5	7.2	7.1	3.3	2.7	3.1	2.5	2.9
Variance: $\widehat{\sigma}^2$	14.1	7.6	8.5	8.8	9.8	16.6	14.8	13.5	8.5	12.8
OD	1.42	−0.03	0.13	0.22	0.39	4.06	4.40	3.31	2.47	3.46
ZI	–	–	–	1.10	1.10	1.82	2.03	1.80	2.18	2.44
p^*	1.2	–	0	0.24	0.52	2.17	2.51	2.07	1.96	2.18
p-value(t_{12})	–	–	–	–	8e-8	–	3e-3	6e-4	9e-5	0

Table 2
Results of AIC, deviance and dipersion parameter with its standard error (se) for fitting various HDRMs [(8) and (9)] to the data from Table 1.

Model	Photoperiod	AIC	df	Deviance	$\widehat{\phi}(se)$
$HD_{\infty}(m)$	8	729.31	136	206.88	–
$HD_2(\mu,\phi)$	8	485.40	136	140.00	0.189(0.023)
$HD_2(m,\phi)$	8	503.99	136	143.34	15.491(5.890)
$HD_3(\mu,\phi)$	8	471.64	136	140.00	0.032(0.004)
$HD_3(m,\phi)$	8	514.95	136	147.44	10.851(8.579)
$HD_{\infty}(m)$	16	843.60	126	606.16	–
$HD_2(\mu,\phi)$	16	187.43	126	130.00	1.206(0.150)
$HD_2(m,\phi)$	16	187.40	126	129.96	0.782(0.124)
$HD_3(\mu,\phi)$	16	187.79	126	130.00	0.513(0.064)
$HD_3(m,\phi)$	16	188.41	126	130.51	1.495(0.474)

Table 3
Results of standard errors of regression coefficient estimates $\widehat{\beta}_j(8)$ and $\widehat{\beta}_j(16)$ for the Table 2 according to the photoperiods 8 and 16.

Model	$\widehat{\beta}_0(8)=1.15810$	$\widehat{\beta}_1(8)=0.36104$	$\widehat{\beta}_2(8)=-0.04203$	$\widehat{\beta}_3(8)=0.00137$
$HD_{\infty}(m)$	0.291691	0.145364	0.018804	0.000655
$HD_2(\mu,\phi)$	0.348643	0.178646	0.023343	0.000817
$HD_2(m,\phi)$	0.347798	0.175108	0.022733	0.000793
$HD_3(\mu,\phi)$	0.347472	0.184144	0.024330	0.000855
$HD_3(m,\phi)$	0.341715	0.173359	0.022571	0.000789
	$\widehat{\beta}_0(16)=1.64502$	$\widehat{\beta}_1(16)=-0.28746$	$\widehat{\beta}_2(16)=0.03843$	$\widehat{\beta}_3(16)=-0.00139$
$HD_{\infty}(m)$	0.431564	0.226852	0.030012	0.001057
$HD_2(\mu,\phi)$	0.949686	0.492558	0.065009	0.002289
$HD_2(m,\phi)$	0.951093	0.492903	0.065046	0.002290
$HD_3(\mu,\phi)$	0.984797	0.501132	0.065899	0.002319
$HD_3(m,\phi)$	0.974558	0.497715	0.065496	0.002305

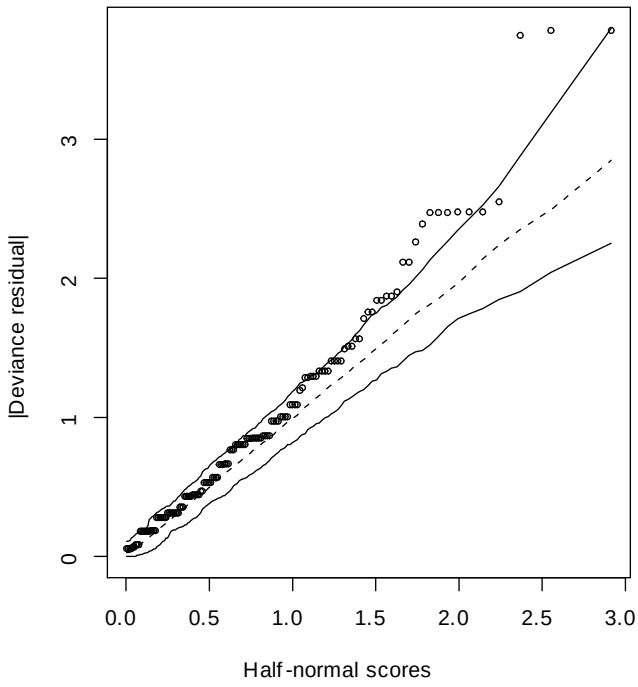

Fig. 1. Half-normal plot for photoperiod 8 with Poisson simulated envelope.

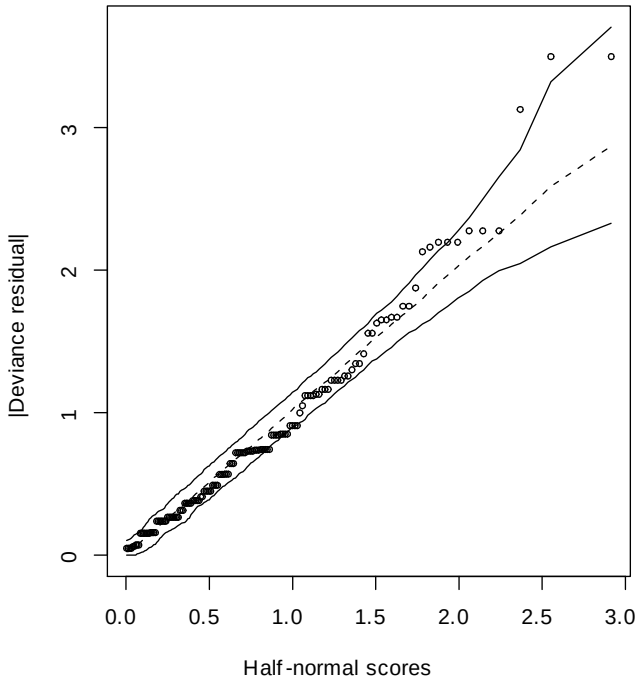

Fig. 2. Half-normal plot for photoperiod 8 with negative binomial simulated envelope.

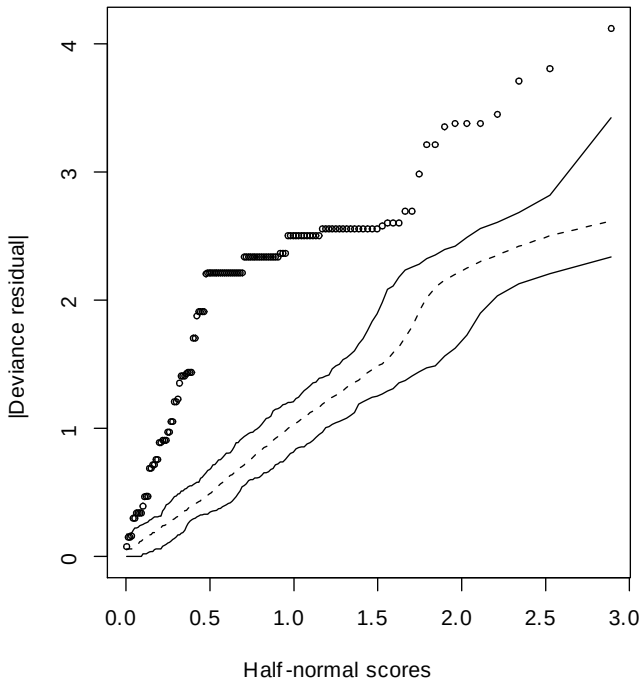

Fig. 3. Half-normal plot for photoperiod 16 with Poisson simulated envelope.

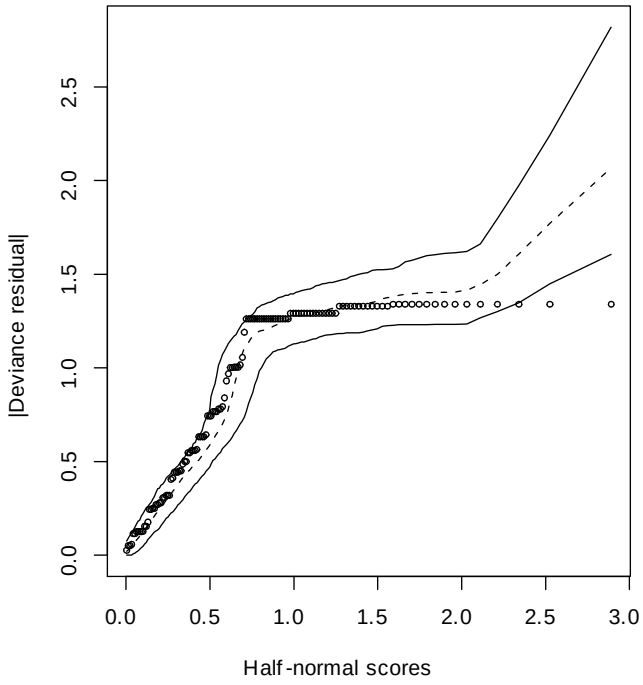

Fig. 4. Half-normal plot for photoperiod 16 with negative binomial simulated envelope.