On the existence of maximum likelihood estimators in a Poisson-gamma HGLM and a negative binomial regression model

GNING Lucien Diégane
Laboratoire d'Etudes et de Recherches en Statistiques et Développement (LERSTAD), UFR Sciences Appliquées et Technologies, Université Gaston Berger.
Saint-Louis (B.P. 234), Sénégal
E-mail: luciendesgning@yahoo.fr

Pierre-Loti-Viaud, Daniel
Universit Pierre et Marie Curie, LSTA
4 place Jussieu
Paris (F-75252, Cedex 05), France
E-mail: daniel.pierre-loti-viaud@upmc.fr

Introduction

Poisson-gamma HGLM are members of the hierarchical generalized linear model family (see [8]), which is an extension of generalized linear models and generalized linear mixed models.
HGLM are applied particularly to longitudinal data that are generally correlated. An important motivation comes from the use of Poisson-gamma HGLM to describe numbers of claim in order to evaluate an a posteriori premium in non-life insurance (for more details see [2] or [3]). It is then important to establish the existence of MLE and possibly the uniqueness for a credible foundation for the method. In this paper, we are interested in the existence of MLE of parameters in a Poisson-gamma HGLM where the regressors are not dependent of time.
In this sense, our best achievement is the finding of a sufficient condition for the existence of MLE in this type of Poisson-gamma HGLM. Then, we remark that this model is linked to a negative binomial regression model in which the index parameter is unknown. By doing so, the condition obtained is also sufficient for the existence of MLE of parameters in the former as well as in the latter model. A necessary and sufficient condition for the existence and uniqueness of MLE of the index parameter in a negative binomial sample has been established in [9] , we show that our condition is a natural generalization of this latter.
The remainder of this paper is organized as follows. In the first section, we introduce Poisson-gamma HGLM and define notation that will be useful later. In the second section, we give our main result and preliminary results that we will be used to prove the main result. The paper concludes by proposing a conjecture. Indeed, it seems quite likely that our condition is also a necessary and sufficient condition for the existence and the uniqueness of MLE in both models considered, the Poisson-gamma HGLM where the regressors are not dependent of time and the negative binomial regression in the case of an unkwown index parameter.

Poisson-gamma HGLM

We may work here with an insurance portfolio composed of n policyholders or insureds, each insured being observed during T periods. Let Y_{kt}, be the number of reported claims for insured k during period t, and $x_1, \ldots, x_J$, be the covariates which represent the observable non random characteristics of insureds, taking x_{ktj}, $k = 1, \ldots, n$, $t = 1, \ldots, T$, $j = 1, \ldots, J$, as the observed value of the j^{th} covariate for the insured k at the period t. Furthermore, let for each insured k, Θ_k be a non-observed real random characteristic. We suppose that:

$(H1)$ each insured is represented by the random vector $(\Theta_k \ Y_{k1} \dots Y_{kT})'$, and these vectors are independent;
$(H2)$ given Θ_k, the numbers of claims $Y_{k1}, \dots, Y_{kT}$ are independent, and satisfy:

$$(1) \quad \mathcal{L}\big(Y_{kt}/\Theta_k{=}\theta_k\big) = \mathcal{P}(\lambda_{kt}\theta_k), \ \ k = 1, \dots, n, \ \ t = 1, \dots, T,$$

where $\log \lambda_{kt} = \boldsymbol{x}'_{\boldsymbol{kt}}\boldsymbol{\beta}$, and $\boldsymbol{\beta} = (\beta_1 \dots \beta_J)'$ is the regression parameters vector for explanatory variables $\boldsymbol{x}'_{\boldsymbol{kt}} = (x_{kt1} \dots x_{ktJ})$.
Moreover, we suppose that:
$(H3)$ $\mathcal{L}(\Theta_k) = \gamma(a, a)$, $k = 1, \dots, n$.
At the following of this work, we take $\boldsymbol{x_{kt}}{=}\boldsymbol{x_k}$ and $\lambda_{kt}{=}\lambda_k$ for $k = 1, \dots, n$, $t = 1, \dots, T$.
Now, let $X_{_R}$, be the $n \times J$ matrix defined by:

$$X_{_R} = \begin{pmatrix} \boldsymbol{x}'_{\boldsymbol{1}} \\ \boldsymbol{x}'_{\boldsymbol{2}} \\ \vdots \\ \boldsymbol{x}'_{\boldsymbol{n}} \end{pmatrix},$$

and $\mathcal{F}_{_R}$ the subspace $\mathrm{Im}(X_{_R})$. We suppose that $X_{_R}$ is a full rank matrix. After that, we note that the defined model satisfies to following constraint:

$$(2) \quad \log \boldsymbol{\lambda} \in \mathcal{F}_{_R},$$

where $\boldsymbol{\lambda} = (\lambda_1 \dots \lambda_n)'$.

Existence results for MLE

In this section, we state a sufficient condition for the existence of MLE of the parameters in a Poisson-gamma HGLM where $\lambda_{kt} = \lambda_k$. As we will prove it, this condition is also sufficient for the existence of MLE of the parameters in a negative binomial model in which the index parameter is unknown. First, we give the expression of the log-likelihood function of the model. Next, we give preliminary results that we will use to prove the already mentioned condition.

Log-likelihood function

Given Θ_k, the periodical numbers of claims $Y_{k1}, \dots, Y_{kT}$ are independent. The joint probability function of $Y_{k1}, \dots, Y_{kT}$ is thus given by:

$$\begin{aligned} \mathbb{P}(Y_{k1}{=}y_{k1}, \dots, Y_{kT}{=}y_{kT}) &= \int_0^{+\infty} \mathbb{P}(Y_{k1}{=}y_{k1}, \dots, Y_{kT}{=}y_{kT}/\Theta_k = \theta_k) f(\theta_k) d\theta_k \\ &= \int_0^{+\infty} \prod_{t=1}^{T} \big(\mathbb{P}(Y_{kt}{=}y_{kt}/\Theta_k = \theta_k)\big) f(\theta_k) d\theta_k, \end{aligned}$$

where f is the density function of the gamma distribution. With straightforward calculation, we deduce from this expression that the likelihood function of the model is given by:

$$(3) \quad L_{\boldsymbol{y}}(a, \boldsymbol{\lambda}) = \prod_{k=1}^{n} \left(\frac{\Gamma(a{+}s_k)}{\Gamma(a) \prod_{t=1}^{T} y_{kt}!} \left(\frac{a}{a{+}T\lambda_k}\right)^{a} \left(\frac{\lambda_k}{a{+}T\lambda_k}\right)^{s_k} \right),$$

where $s_k = \sum_{t=1}^{T} y_{kt}, k = 1, \ldots, n$. As a result, the log-likelihood function of the model minus some constant is given by:

$$\ell_{\boldsymbol{y}}(a, \boldsymbol{\beta}) = \sum_{k=1}^{n} \left(s_k \boldsymbol{x_k}' \boldsymbol{\beta} - (a + s_k) \log \left(a + T e^{\boldsymbol{x_k}' \boldsymbol{\beta}} \right) + a \log a \right) + \sum_{k=1}^{n} \log \left(\frac{\Gamma(a + s_k)}{\Gamma(a)} \right).$$

Therefore, using the properties of the function Γ (see [1]), we rewrite this log-likelihood function as:

$$\ell_{\boldsymbol{y}}(a, \boldsymbol{\beta}) = \sum_{k=1}^{n} \left(s_k \boldsymbol{x_k}' \boldsymbol{\beta} - (a + s_k) \log \left(a + T e^{\boldsymbol{x_k}' \boldsymbol{\beta}} \right) + a \log a \right) + \sum_{k=1}^{n} \sum_{j=0}^{s_k - 1} \log(a + j),$$

where, for $k = 1, \ldots, n$, $\sum_{j=0}^{s_k - 1} \log(a + j) = 0$, if $s_k = 0$. Let now $r = \frac{1}{a}$, and Φ the function defined on $]0, +\infty[$ by:

$$(4) \quad \Phi(r, \boldsymbol{\beta}) = \ell_{\boldsymbol{y}}\left(\frac{1}{r}, \boldsymbol{\beta}\right) = \sum_{k=1}^{n} \left(s_k \boldsymbol{x_k}' \boldsymbol{\beta} - \left(\frac{1}{r} + s_k\right) \log \left(1 + r T e^{\boldsymbol{x_k}' \boldsymbol{\beta}} \right) \right) + \sum_{k=1}^{n} \sum_{j=0}^{s_k - 1} \log \left(1 + r j \right).$$

In the following of this paper we take $S_k = \sum_{t=1}^{T} Y_{kt},\ \ k = 1, \ldots, n$, $\boldsymbol{S} = (S_1 \ldots S_n)'$ and we denote $\mathrm{diag}(u_k)$, the diagonal matrix with elements the u_k, $\ k = 1, \ldots, n$.

Preliminary results

We obtain the following results.

Lemma 1 *Let K be a compact space in $\mathbb{R}^J$, then $\Phi(r, \boldsymbol{\beta})$ converges uniformly on K to the function $\boldsymbol{\beta} \to \Phi(0, \boldsymbol{\beta})$ defined by:*

$$(5) \quad \Phi(0, \boldsymbol{\beta}) = \sum_{k=1}^{n} \left(s_k \boldsymbol{x_k}' \boldsymbol{\beta} - T e^{\boldsymbol{x_k}' \boldsymbol{\beta}} \right),$$

as $r \searrow 0$. Furthermore, the function $\boldsymbol{\beta} \to \Phi(0, \boldsymbol{\beta})$ is the log-likelihood function minus some constant of a Poisson regression model for clustered data.

Lemma 2 *The function $\boldsymbol{\beta} \to \Phi(0, \boldsymbol{\beta})$ is strictly concave on $\mathbb{R}^J$, and if there exists $\delta \in \mathcal{F}_R^{\perp}$ such that $S_k + \delta_k > 0$ for each $k = 1, \ldots, n$, then there exists a unique $\hat{\boldsymbol{\beta}}(0) \in \mathbb{R}^J$ satisfying:*

$$\Phi(0, \hat{\boldsymbol{\beta}}(0)) = \max_{\boldsymbol{\beta} \in \mathbb{R}^J} \Phi(0, \boldsymbol{\beta}).$$

Lemma 3 *For $r > 0$, the function $\boldsymbol{\beta} \to \Phi(r, \boldsymbol{\beta})$ is strictly concave on $\mathbb{R}^J$, and if there exists $\delta \in \mathcal{F}_R^{\perp}$ such that $S_k + \delta_k > 0$ for each $k = 1, \ldots, n$, then there exists a unique $\hat{\boldsymbol{\beta}}(r) \in \mathbb{R}^J$ satisfying:*

$$\Phi(r, \hat{\boldsymbol{\beta}}(r)) = \max_{\boldsymbol{\beta} \in \mathbb{R}^J} \Phi(r, \boldsymbol{\beta}).$$

Lemma 4 *If there exists $\delta \in \mathcal{F}_R^{\perp}$ such that $S_k + \delta_k > 0$ for each $k = 1, \ldots, n$, then for all neighborhood $\mathcal{V}$ of $\hat{\boldsymbol{\beta}}(0)$, and for enough small r, the function $\boldsymbol{\beta} \to \Phi(r, \boldsymbol{\beta})$ has a local maximum on $\mathcal{V}$ which is $\hat{\boldsymbol{\beta}}(r)$.*

Theorem 1 *If there exists $\delta \in \mathcal{F}_R^{\perp}$ such that $S_k + \delta_k > 0$ for each $k = 1, \ldots, n$, then:*

$$\lim_{r \searrow 0} \hat{\boldsymbol{\beta}}(r) = \hat{\boldsymbol{\beta}}(0).$$

Theorem 2 *If there exists* $\delta \in \mathcal{F}_R^{\perp}$ *such that* $S_k + \delta_k > 0$ *for each* $k = 1, \ldots, n$, *then for* $r > 0$ *in a neighborhood of* 0, *we have:*

$$\hat{\boldsymbol{\beta}}(r) = \hat{\boldsymbol{\beta}}(0) + rD\hat{\boldsymbol{\beta}}(0) + o(r),$$

where:

$$D\hat{\boldsymbol{\beta}}(0) = \left(X_R^{'} diag(Te^{\boldsymbol{x_k}'\hat{\boldsymbol{\beta}}(0)})X_R\right)^{-1}\left(X_R^{'} diag(T^2 e^{2\boldsymbol{x_k}'\hat{\boldsymbol{\beta}}(0)})\mathbb{1}_n - X_R^{'} diag(Te^{\boldsymbol{x_k}'\hat{\boldsymbol{\beta}}(0)})\boldsymbol{S}\right),$$

and $\mathbb{1}_n = (1 \ldots 1)' \in \mathbb{R}^n$.

Main result

Now, we state in the theorem below the sufficient condition for the existence of MLE of the model. We denote by:

$$\hat{\lambda}_k(0) = e^{\boldsymbol{x}_k^{'}\hat{\boldsymbol{\beta}}(0)}, \ \ k = 1, \ldots, n,$$

when $\hat{\boldsymbol{\beta}}(0)$ exists.

Theorem 3 *If the following conditions* $(C1)$ *and* $(C2)$ *are satisfied, then the MLE of the parameters* a *and* $\boldsymbol{\beta}$ *exists.*
$(C1)$: *If there exists* $\delta \in \mathcal{F}_R^{\perp}$ *such that* $S_k + \delta_k > 0$ *for each* $k = 1, \ldots, n$.
$(C2)$: $\frac{1}{n}\sum_{k=1}^{n}\left(\frac{1}{T}\sum_{t=1}^{T} Y_{kt} - \hat{\lambda}_k(0)\right)^2 - \frac{1}{T}\left(\frac{1}{n}\sum_{k=1}^{n}\left(\frac{1}{T}\sum_{t=1}^{T} Y_{kt}\right)\right) > 0.$

Remark 1 *If* $T = 1$, *then the random variables are* Y_{kt} *i.i.d. In this case* $(C2)$ *become "the empirical variance is strictly greater than the empirical mean" which is a necessary and sufficient condition for the existence and the uniqueness of the MLE of the index parameter in a negative binomial sample(see [9]).*

Remark 2 *The condition (C2) is interpreted as the difference between the residual variance of the Poisson regression model (which is not hierarchical) associated with our model, and the empirical mean divided by the number of periods.*

Link to the negative binomial regression model

We complete the presentation of our results by presenting the link between a Poisson-gamma HGLM where $\lambda_{kt} = \lambda_k$ and a negative binomial regression model.

Proposition 1 *The Poisson-gamma HGLM with* $\lambda_{kt} = \lambda_k$, *corresponds to a negative binomial regression model associated at the variables* S_k, $k = 1, \ldots, n$ *with a logarithmic link and an unknown index parameter .*

Corollary 1 *It comes from the proposition* 1 *that, if* $(C1)$ *and* $(C2)$ *are satisfied, then MLEs of all parameters (including the index parameter) of a negative binomial regression model with logarithmic link exists.*

Conclusion

The Remark 1 and the Corollary 1 encourage us to declare the following conjecture.

Conjecture. The MLE's of the parameters a and $\boldsymbol{\beta}$ of the Poisson-gamma HGLM model with$\lambda_{kt} = \lambda_k$, and that of the negative binomial regression, exists and are unique, if and only if $(C1)$ and $(C2)$ are satisfied.

References

[1] Abramowitz M., Stegun I. (1972). *Handbook of Mathematical Functions.* Dover Publications, New York.

[2] Boucher J., Denuit M., Guilln M.(2008). Models of Insurance claim counts with time dependance based on generalization of Poisson and negative binomial distributions. *Casuality Actuarial Society,* **2**(1)*,* $136-161$.

[3] Dionne G., Vanasse Ch. (1989). A generalization of automobile insurance rating models: the negative binomial binomial distribution with a regression component. *Astin Bulletin* **19**(2), $199-212$.

[4] Gning L., *Utilisation des mlanges de lois de Poisson et de lois binomiales ngatives pour tablir des tarifs a priori et a posteriori en assurance non vie.* Thse de mathmatiques en prparation.

[5] Gning L., Pierre-Loti-Viaud, D. (2010). Sur les estimateurs du maximum de vraisemblance dans les modles multiplicatifs de Poisson et binomiale ngative. *Afrika Statistika* **5**(13)*,* $297-305$.

[6] Hilbe J.M. (2007). *Negative Binomial Regression.* Cambridge University Press, Cambridge.

[7] Johnson N., Kotz S., Balakrishnan N. (1996). *Discrete Multivariate Distributions (*2^e *dition).*Wiley, New York.

[8] Lee Y., Nelder J. (1996). Hierarchical Generalized Linear Models. *Journal of Royal Statistical Society, Series B, (Methodological)* **58**, $619-678$.

[9] Levin B., Reeds J. (1977). Compound multinomial likelihoods functions are unimodal: proof a conjecture of I.J.Good. *The Annals of Statistics.* **5**, $79-87$.

[10] Santner J., Duffy E. (1989). *The Statistical Analysis of Discrete Data.* Springer-Verlag, New York.

[11] Spivak M.(1979). *A comprehensive introduction to differential geometry.* Publish or Perish deuxime dition, Houston .

[12] Zeger S., Liang K. (1986). Longitudinal Data Analysis using generalized linear models. *Biometrika,* **73**, $13-22$.

ABSTRACT

We are interested in the existence of maximum likelihood estimators (MLE) for all parameters in a hierarchical generalized linear model (HGLM) of Poisson-gamma type, and in a negative binomial regression model. These models are widely used in pricing non-life insurance (see [2], [3], [6]). However, for these models, it seems that the existence of MLE has not been studied in literature. In this paper, we address this issue for first time for a particular sub-model, indeed, we give a sufficient condition for the existence of MLE of parameters in the Poisson-gamma HGLM where the regressors are not dependent of time. We also show that the condition obtained is sufficient for the existence of MLE of parameters in the negative binomial regression model in the case of an unknown index parameter. In this latter case, our condition appears as a natural extension of the necessary and sufficient condition known for the existence and uniqueness of MLE of the index parameter for a sample of negative binomial distirbution.