Gamma Family of Dynamic Models

Thiago Rezende dos Santos[a*] Dani Gamerman[b] Glaura C. Franco[c]

[a]*Departament of Mathematics, UFOP, MG, Brazil.*
[b]*Departament of Statistical Methods, UFRJ, RJ, Brazil.*
[c]*Departament of Statistics, UFMG, MG, Brazil.*

Abstract

The non-gaussian dynamic models have been used in modeling of count, proportions time series. In this article, a new family of dynamic models - Gamma Family of Dynamic Models is introduced, as well as particulares cases, generalizing previous proposals which were done in the literature. Besides sequential inference, are presented filtering and smoothing results. Through Monte Carlo Experiments, the behaviour of interval and point estimators of model parameters are investigated and compared. The results showed that the MLE and Bayesian estimators have a small and similar mean square error. Already interval estimators presented coverage rate near to the nominal level supposed and credibility intervals have, in general, a width slightly larger. As illustration of presented methodology, the Poisson, Gama models and Normal model of sthocastic volatility were fitted to two real time series and the results were satisfactories.

Keywords: State Space Models; Local Level Model; Classical Inference; Bayesian Inference; MCMC.

1 Introduction

In the literature, there are several models which are built based on normality, homocesdacity and independence assumptions of the errors, however, in some cases, it is not possible to satisfy these assumptions. Under time series context, the independence assumption of the error rarely is satisfied,

*Correspondence author: Thiago Rezende dos Santos. Address: Universidade Federal de Ouro Preto, Morro do Cruzeiro, Ouro Preto, MG, 35400-000, Brazil. E-mail: thiagors@iceb.ufop.br

while normality assumption has been considering in the mainly approaches of modeling time series.

The modeling by state space, that is the approached subject in this work, possesses one great amount of works and results, based on the assumption of normality. In this article, some possibilities will be presented for the treatment of time series that surpass this restriction.

The starting point for this extension is the article of Nelder & Wedderburn (1972), that proposed the family for them considered called of *generalized linear models* (GLM), unifying some existing models then of isolated form in a class. The basic idea of these models consists of opening the range of options for the response-variable distribution, allowing that the same one belongs to the exponential family of distributions, what also it brings a profit in the question of interpretation of the model. The function of linking of the data plays the role to relate the average of the data to the linear preditor, according to Nelder & Wedderburn (1972) and Dobson (2002).

In the context of time series, the correlation structure of observations can not be rejected. In this direction, one more general structure, called for *Dynamic Generalized Linear Models* (DGLM), proposed by West, Harrison & Migon (1985), attracting an immense interest due to great applicability of the same ones in diverse areas of the knowledge. Proof of this is the great number of works published on these models. Amongst which it can be cited the works of Gamerman & West (1987), Grunwald, Raftery & Guttorp (1993), Fahrmeir (1987), Fruhwirth-Schnatter (1994), Lindsey & Lambert (1995), Gamerman (1991a, 1998), Chiogna & Gaetan (2002), Hemming & Dhaw (2002) and Godolphin & Triantafyllopoulos (2006). Works that deal with non-gaussian time series, not necessarily the DGLM, include Smith (1979), Smith (1981), Cox (1981), Kaufmann (1987), Kitagawa (1987), Smith & Miller (1986), Harvey & Fernandes (1989), Shephard & Pitt (1997), Jørgensen et al. (1999) and Durbin & Koopman (2000), among others.

The problem with this models class (GDLM) is that the analytical form is easily lost, even using very simple components. Thus, the predictive distribution, that is basic for inference process, can only be gotten in an aproximated way. A particular case of these models assumes that only the trend floats and the effect of the covariables are fixed to long of the time. For these cases, a sufficiently wide models class exists that allows the exact computation of the predictive distribution, the Gamma Family of Dynamic Models (GFDM).

Thus, the main objective of this article is to consider in its second section

Gamma Family of Dynamic Models, which allows analytical computation of the predictive distributions. This family is gotten from a generalization of a Smith & Miller (1986)'s result. They had considered an exact evolution equation, thus making possible the analytical integration of the states and attainment of the predictive distributions that compose the likelihood function.

Thus, the contributions of the article are to consider and to characterize the GFDM, to present particular cases that belong to this family, as the Gamma, Pareto and Beta models and to consider a smoothing form of the component of level of the model. Besides, a study about the behavior of the estimators for a variety of possibilities of this family.

Gamma Family of Dynamic Models will be presented in Section 2. Its main theoretical results will be supplied as well as the form to make inference will be described. Classical approach, using the maximum likelihood estimator (MLE), and the Bayesian approach, using MCMC methods to get the Bayesian estimators (BE), are considered in the inference process. Besides, credibility and confidence intervals are built for the parameters. Next, in Section 3, particular cases will be presented inside of this family. Section 4 deals with the comparison of Classical and Bayesian points of view through exercises of simulation. In Section 5, the real data are made applications, adjusting Poisson and Gama models. Finally, Section 6 presents the main conclusions and final remarks.

2 Gamma Family of Dynamic Models

Smith & Miller (1986) and Harvey & Fernandes (1989) had presented particular cases of non-gaussian dynamic models. In this work, from these cases a generalization is made, that is, a wide family is introduced, called *Gamma Family of Dynamic Models*. An advantage of these models compared to the DGLM is that evolution equation is exact. On the other hand, it is not one trivial task to insert other sthocastic components, for example, the components of trend or/and sazonality. Although, the effect of these components can be caught in model through the covariables. In this section, the GFDM definition, the procedures of inference (classical and Bayesian) are considered, one way of making the forecast and the smoothing, as well as the model adequacy.

2.1 Definition

In a general way, it is defined that the time series $\{y_t\}$ possesses one distribution in the GFDM, if its distribution is written in the form:

$$p(y_t|\mu_t, \boldsymbol{\varphi}) = a(y_t, \boldsymbol{\varphi})\mu_t^{b(y_t,\boldsymbol{\varphi})} \exp(-\mu_t c(y_t, \boldsymbol{\varphi})), \quad (1)$$

where $y_t \in H(\boldsymbol{\varphi}) \subset \Re$ and $p(y_t|\mu_t, \boldsymbol{\varphi}) = 0$, otherwise. The functions $a(\cdot)$, $b(\cdot)$, $c(\cdot)$ and $H(\cdot)$ are such as $p(y_t|\mu_t, \boldsymbol{\varphi}) \geq 0$ and the Lebesgue-Stieljes integral $\int dF(y_t|\mu_t, \boldsymbol{\varphi}) = 1$ in which $F(\cdot)$ denotes the cumulative distribution function of y_t.

The GFDM is defined in the following way:

1. Se x_t is a covariate vector, the link function g relate the preditor to parameter μ_t through the relation $\mu_t = \lambda_t g(x_t, \boldsymbol{\beta})$, where $\boldsymbol{\beta}$ has regression coeficients (one of components of $\boldsymbol{\varphi}$) and λ_t is the parameter relate to the description of the dynamic level. If preditor is linear then, $g(x_t, \boldsymbol{\beta}) = g(x_t'\boldsymbol{\beta})$.

2. Level dynamic λ_t is given by evolution equation $\lambda_t = w^{-1}\lambda_{t-1}\varsigma_t$, where $\varsigma_t \sim Beta\,(wa_{t-1}, (1-w)a_{t-1})$, that is,

$$w\frac{\lambda_t}{\lambda_{t-1}} \mid \lambda_{t-1} \sim Beta\,(wa_{t-1}, (1-w)a_{t-1})\,.$$

3. The level dynamic λ_t is initialized with the specification *a priori* $\lambda_0|Y_0 \sim Gama(a_0, b_0)$. Therefore, using the scale propriety of the Gamma distribution, $\mu_0|Y_0 \sim Gama(a_0, b_0[g(x_t, \beta)]^{-1})$.

One of the specifications most usual for the link function g is the logarithmic function. It is interesting to highlight that, in this case, the evolution equation is gotten in the following way $\ln(\lambda_t) = \ln(\lambda_{t-1}) + \varsigma_t^*$, onde $\varsigma_t^* = \ln(\varsigma_t/w) \in \Re$. This equation is similar to the usual evolution equation given by a random walk, as the local level model. The parameter w varies between 0 and 1 and also composes $\boldsymbol{\varphi}$. As it will be seen to follow, w fulfills the function to increase the variance due to passing of the time. Thus, it plays a similar role to the system variances and it plays identical role to the discounting factors, used in the Bayesian approach for substituting these variances.

Case $b(y_t, \boldsymbol{\varphi}) = b(y)$ or $c(y_t, \boldsymbol{\varphi}) = c(y)$ and $H(\boldsymbol{\varphi})$ is a constant function (it does not depend of $\boldsymbol{\varphi}$), the gamma family of the dynamic models becomes a special case of exponential family of the distributions.

For one better characterization of the GFDM, some results are presented in Theorem 1 such as the distribution *a posteriori* of the level λ_t and the predictive density distributions. These results are the foundations in the inferencial process of the GFDM.

Theorem 1.

If the model is defined in the form of the Equation (1), the following results can be obtained:

1. the prior distribution $\lambda_t|\boldsymbol{Y}_{t-1}$ follows a Gamma$(a_{t|t-1},\, b_{t|t-1})$ distribution such as

$$a_{t|t-1} = wa_{t-1}, \tag{2}$$

$$b_{t|t-1} = wb_{t-1}, \tag{3}$$

and $0 < w \leq 1$.

2. $(\mu_t = \lambda_t g(x_t, \boldsymbol{\beta}))|\mathbf{Y}_{t-1}$, which is Gamma$(a^*_{t|t-1},\, b^*_{t|t-1})$, where

$$a^*_{t|t-1} = wa_{t-1}, \tag{4}$$
$$b^*_{t|t-1} = wb_{t-1}[g(x_t^{'}\beta)]^{-1}. \tag{5}$$

3. The posterior distribution of $\mu_t|\boldsymbol{Y}_t$ is Gamma (a^*_t, b^*_t), where

$$a^*_t = a^*_{t|t-1} + b(y_t, \boldsymbol{\varphi}), \tag{6}$$

$$b^*_t = b^*_{t|t-1} + c(y_t, \boldsymbol{\varphi}). \tag{7}$$

4. $\lambda_t = (\mu_t[g(x_t^{'}\boldsymbol{\beta})]^{-1})|\mathbf{Y}_t$ has Gamma(a_t, b_t) distribution in which

$$a_t = a_{t|t-1} + b(y_t, \boldsymbol{\varphi}), \tag{8}$$

$$b_t = b_{t|t-1} + c(y_t, \boldsymbol{\varphi})g(x_t, \boldsymbol{\beta}). \tag{9}$$

5. The predictive density function one step ahead is given by

$$p(y_t|\boldsymbol{Y}_{t-1}, \boldsymbol{\varphi}) = \frac{\Gamma(b(y_t, \boldsymbol{\varphi}) + a^*_{t|t-1})a(y_t, \boldsymbol{\varphi})(b^*_{t|t-1})^{a^*_{t|t-1}}}{\Gamma(a^*_{t|t-1})[c(y_t, \boldsymbol{\varphi}) + b^*_{t|t-1}]^{b(y_t, \boldsymbol{\varphi}) + a^*_{t|t-1}}}, \quad y_t \in H(\boldsymbol{\varphi}). \tag{10}$$

$\forall t \in N$; $t \leq n$ where n is the time series length and $\Gamma(\cdot)$ is gamma function. The proof of Theorem 1 is found in the Appendix 1. It is easy to see starting from (2)- (3) what $Var(\lambda_t|\boldsymbol{Y}_{t-1}) = w^{-1}Var(\lambda_{t-1}|\boldsymbol{Y}_{t-1})$. Thinking in terms of the precision (Inverse of the variance) like information measure, It has that for $t-1 \rightarrow t$ implies that just $100w\%$ of the information is preserved. It is exactly that the discount factors use in West & Harrison (1997). These factors measure the information quantity (measure by the system precision) preserved in the course of time.

The predictive density function of the observations, given $\boldsymbol{Y}_t$, h ($h > 0$) steps ahead is given by

$$p(y_{t+h}|\boldsymbol{Y}_t, \boldsymbol{\varphi}) = \frac{\Gamma(b(y_{t+h}, \boldsymbol{\varphi}) + a^*_{t+h|t})a(y_{t+h}, \boldsymbol{\varphi})(b^*_{t+h|t})^{a^*_{t+h|t}}}{\Gamma(a^*_{t+h|t})[c(y_{t+h}, \boldsymbol{\varphi}) + b^*_{t+h|t}]^{b(y_{t+h},\boldsymbol{\varphi})+a^*_{t+h|t}}}, \quad y_{t+h} \in H(\boldsymbol{\varphi}). \tag{11}$$

Corollary. According to result 1 of the Theorem 1, equations it (2) and (3), given a $h > 0$, it can be obtained, h steps ahead, the distribution of λ_{t+h}, given all information avaliable until to instant t, whose form is given by:

$$\lambda_{t+h}|\mathrm{Y}_t \sim \text{Gamma}(w^h a_t, w^h b_t). \tag{12}$$

As the result,

$$[\mu_{t+h}]\,|\mathbf{Y}_t \sim \text{Gamma}(w^h a_t, w^h b_t[g(x_{t+h}, \boldsymbol{\beta})]^{-1}). \tag{13}$$

The distributions in (12) and (13) describe the uncertain associated to level forecasts and based on them can be extrate resume measures of distribution, such as mean, percentiles.

2.2 Inference Procedures

The model parameters can be divided into latent state λ_t and fixed parameters, usually denominated hyparameters(φ). The *on-line* inference the state parameters λ_t was trated in the Section 2.1 and the smoothing inference will be treated in the Section 2.4. In this subsection, It will be discussed the inference on the hyperparameters.

2.2.1 Classical Inference

One way of making classical inference about the parameter vector $\boldsymbol{\varphi}$ is through marginal likelihood function whose form is given by

$$L(\boldsymbol{\varphi};\boldsymbol{Y}_n) = \prod_{t=1}^{n} p(y_t|\boldsymbol{Y}_{t-1},\boldsymbol{\varphi}) = \prod_{t=1}^{n} \frac{\Gamma(b(y_t,\boldsymbol{\varphi})+a^*_{t|t-1})a(y_t,\boldsymbol{\varphi})(b^*_{t|t-1})^{a^*_{t|t-1}}}{\Gamma(a^*_{t|t-1})[c(y_t,\boldsymbol{\varphi})+b^*_{t|t-1}]^{b(y_t,\boldsymbol{\varphi})+a^*_{t|t-1}}},$$
$$y_t \in H(\boldsymbol{\varphi}), \tag{14}$$

where $\boldsymbol{\varphi}$ is composed by ω, β and by specific parameters of the model; $\boldsymbol{Y}_n = (y_1, \ldots, y_n)'$.

In Equation (14), definition of τ, as presented above, it is due to the following fact: Gama prior distribution, that is, the initial distribution μ_t tends to turn be become non-informative when $a_0, b_0 \rightarrow 0$, although it is improper when $a_0 = b_0 = 0$. Note that if $a_0, b_0 \rightarrow 0$ e $y_1 = 0$, the posterior distribution $p(\mu_t|\boldsymbol{Y}_t)$ can be improper, then the predictive densities functions can not be defined.

A proper distribution of μ_t can be obtained in time $t = \tau$, in which τ is the index of the forst observation different from zero. Although, if $a_0 > 0$ e $b_0 > 0$, it not necesary the use of τ. From this moment, It will be assumed $a_0 > 0$ e $b_0 > 0$.

By asymptotic properties of the MLE (Harvey, 1989), under some regularity conditions, it leads

$$\hat{\varphi} \xrightarrow{\mathbb{D}} \mathrm{N}\left[\varphi, I^{-1}(\varphi)\right], \tag{15}$$

when $n \longrightarrow \infty$. If $\hat{\varphi}$ is obtained, maximizing marginal likelihood function, what is obtained analytically as the product of predictive densities, the result above is restricted to some following conditions (Harvey, 1989):

1. φ is a inside point of the parametric space;
2. the derivatives of the log-likelihood until to order 3 with respect to φ, exist and are continuous in the neighborhood of the true parameter value;
3. φ is identificable.

Condition (1): When $w = 1$ and/or some parameter of the model definied in the positive semi-straight is equal to zero, φ will not be a inside point to parametric space and the distribution limit can be affected.

Condition (2): The derivatives, for an arbitrary instant t with $t = 1, \ldots, n$, exist and can be computed derivativing the predictive density function in (10) with respect to φ and its continuity is result of the continuity of $a^*_{t|t-1}$ and $b^*_{t|t-1}$. In the case in which one of model parameters depends the suport, the derivatives of this respective parameter can not exist, for example, the Pareto model.

Condition (3): Like the likelihood function is obtained analytically by the product of predictives density functions in (10) and being φ an interior point to the parametric, given two points φ^1 and φ^2 arbitrary and different, the family of join density
$\mathcal{F} = \{p(y_1, \ldots, y_n; \varphi); \varphi \in \Phi \text{ (parametric space)}\}$ will produce different models and, as a result, different values from likelihood function.

The asymptotic confidence interval for $\boldsymbol{\varphi}$ is built based on a numerical approximation for Fisher information matrix, using $I(\boldsymbol{\varphi}) \cong -H(\boldsymbol{\varphi})$ in which $-H(\boldsymbol{\varphi})$ is the observed information matrix, computed through second derivatives of the log-likelihood function with respect to the parameters. As the compute of the derivatives is not an easy task, numerical derivatives are utilized.

Be φ_i, $i = 1, \ldots, p$, any parameter in vector $\boldsymbol{\varphi}$. Then, an asymptotic confidence interval of $100(1-\kappa)\%$ for φ_i is given by:

$$\hat{\varphi}_i \pm z_{\kappa/2}\sqrt{\widehat{Var}(\hat{\varphi}_i)},$$

where $z_{\kappa/2}$ is the $\kappa/2$ percentile of standard normal distribution and $\widehat{Var}(\hat{\varphi}_i)$ is obtained of diagonal elements of Fisher information matrix.

The observed information matrix is asymptotically equivalent to expected information matrix- result known (Migon & Gamerman, 1999) and Apparently corroborated for state models space by the led simulations in (Cavanaugh & Shumway, 1996). This practice of approaching the expected information matrix by the observed information matrix is relatively common and suggested in lots of texts as (Cavanaugh & Shumway, 1996) and (Sallas & Harville, 1988), mainly for large samples.

In some problems, the main interest is to calculate a function of the parameter estimator. In these cases, the method Delta (Casella & Berger , 2002) is used, which is defined soon below. Be $g(\cdot)$ an one-to-one function whose the first derivative exists and is different from zero. Using EMV's asymptotic property in (15), by Delta method, it has that

$$g(\hat{\varphi}_i) \xrightarrow{\mathbb{D}} \mathrm{N}\left[g(\varphi_i), I^{-1}_{ii}(\varphi)(g^{'}(\varphi_i))^2\right], \tag{16}$$

when $n \longrightarrow \infty$ and for $i = 1, \ldots, p$. $I_{ii}^{-1}(\varphi)$ is the i-th diagonal element of Fisher informatiom matrix $I^{-1}(\varphi)$.

Under the classical approach, asymptotic confidence intervals are built for the parameters, but these intervals can present border problems, that is, the intervals limits overtake the borders of the parametric space. In these cases, it applies the Delta method to correct this problem of the following way:

1. The asymptotic confidence interval for $g(\varphi_i)$ is built with $i = 1, \ldots, p$, using the Delta method;
2. To follow, the inverse transformation $g^{-1}(\cdot)$ is applied in the interval limits, obtaining asymptotic interval for $\boldsymbol{\varphi}$.

2.2.2 Bayesian Inference

Already for making Bayesian inference for the model parameters, like the posterior distribution of the parameters is not analytically tractable, it is used the method MCMC, algorithm of Metropolis-Hastings (Gamerman & Lopes , 2006) - so that of obtaining a marginal posterior distribution sample of parameter vector $\boldsymbol{\varphi}$ whose form is given by:

$$\pi(\boldsymbol{\varphi}|\boldsymbol{Y}_n) \propto L(\boldsymbol{\varphi}; \boldsymbol{Y}_n)\pi(\boldsymbol{\varphi}), \tag{17}$$

where$L(\boldsymbol{\varphi}; \boldsymbol{Y}_n)$ is the likelihood function obtained in (14) and $\pi(\boldsymbol{\varphi})$ is prior distrbution of $\boldsymbol{\varphi}$. In this work, a proper Uniform distribution is utilized, $\pi(\boldsymbol{\varphi}) = c$ for all possible values of $\boldsymbol{\varphi}$ in a pre-fixed interval and 0, otherwise.

Credibility intervals for φ_i, $i = 1, ..., p$ are built as following. Given a value $0 < \kappa < 1$, all interval $(t_1, t_2)'$ satisfying

$$\int_{t_1}^{t_2} \pi(\varphi_i \mid \boldsymbol{Y}_n)\, d\varphi_i = 1 - \kappa$$

is a credibility interval for φ_i with level $100(1-\kappa)\%$.

2.3 Model Adequacy and fit

Harvey & Fernandes (1989) suggest some diagnostic methods, based on the standard residuals (Pearson). These residuals are defined by

$$\nu_t = \frac{y_t - \mathbb{E}(y_t|\boldsymbol{Y}_{t-1}, \varphi)}{DP(y_t|\boldsymbol{Y}_{t-1}, \varphi)}, \tag{18}$$

where $DP(y_t|\boldsymbol{Y}_{t-1},\varphi)$ is the standard deviation of distribution of $y_t|\boldsymbol{Y}_{t-1},\varphi$.

Some diagnostic methods:

1. Examine residuals graph versus the time and versus an estimative of level component.

2. Check if the sample variance of standard residuals is close to 1. A value greater than 1 indicate superdispersation of the model.

For more details about these and other diagnostic methods, it consult Harvey & Fernandes (1989).

3 Smoothing

The forecast of the future values of time series is an important topic in time series analysis, due to the great interest in extrapolate the fit model results and project the future values. The forecast for the series observations can be obtained through the predictive density distributions. Already, the forecast for the level, it is shown below as I built them, based on GFDM.

If the interest is to find level component estimate λ, based on all avaliable information ($\boldsymbol{Y}_n$), then the smoothing techniques should be used. Harvey & Fernandes (1989) present a level component estimate of the process in an application to a real series, obtained by the application of the smoothing algorithm of the fixed interval (Harvey, 1989) to a model of random walk plus a noise. However, there is not a theoretical base to build a smoothing. They refer to this procedure as a "quasi-smoothing". Because of this, in this subsection, is proposed one way of making the smoothing of this component in the models of GFDM.

The goal is to obtain a sample distribution of $\lambda|\boldsymbol{Y}_n$, in which $\lambda = (\lambda_1,\ldots,\lambda_n)'$. For so much, it considers

$$\begin{aligned} p(\lambda,\varphi|\boldsymbol{Y}_n) &= p(\lambda|\varphi,\boldsymbol{Y}_n)p(\varphi|\boldsymbol{Y}_n) \\ &= p(\lambda_n|\varphi,\boldsymbol{Y}_n)\prod_{t=1}^{n-1} p(\lambda_t|\lambda_{t+1},\ldots,\lambda_n,\varphi,\boldsymbol{Y}_n)p(\varphi|\boldsymbol{Y}_n) \\ &= p(\lambda_n|\varphi,\boldsymbol{Y}_n)\prod_{t=1}^{n-1} p(\lambda_t|\lambda_{t+1},\varphi,\boldsymbol{Y}_t)p(\varphi|\boldsymbol{Y}_n) \end{aligned}$$

So, the smoothing problem can be solved via MCMC. Once it has a posterior distribution sample of φ, the problem is to find a sample distribution of

$\lambda|\varphi, \boldsymbol{Y}_n$ - what is proposed to follow. Firstly, it finds the distribution of $\lambda_{t-1}|\lambda_t, \varphi, \boldsymbol{Y}_t$. Can be written

$$p(\lambda_{t-1}|\lambda_t, \varphi, \boldsymbol{Y}_t) = p(\lambda_{t-1}|\lambda_t, \varphi, \boldsymbol{Y}_{t-1}) \quad (19)$$

$$= \frac{p(\lambda_t|\lambda_{t-1}, \varphi, \boldsymbol{Y}_{t-1})p(\lambda_{t-1}|\varphi, \boldsymbol{Y}_{t-1})}{p(\lambda_t|\varphi, \boldsymbol{Y}_{t-1})}. \quad (20)$$

Note that the three distributions of fraction above are known.

Theorem 2.

Agreing the three distributions in (19), the following result is obtained (the demonstration can be found in the Appendix 2):

$$\lambda_{t-1} - w\lambda_t|\lambda_t, \varphi, \boldsymbol{Y}_t \sim \text{Gamma}\left((1-w)a_{t-1}, b_{t-1}\right), t = 1, \ldots, n. \quad (21)$$

The result of Theorem 2 is important, because
$p(\lambda_1, \ldots, \lambda_n|\varphi, \boldsymbol{Y}_n) = p(\lambda_n|\varphi, \boldsymbol{Y}_t) \prod_{t=1}^{n-1} p(\lambda_t|\lambda_{t+1}, \varphi, \boldsymbol{Y}_t)$. Of this way, it can be obtained a sample smoothed distribution of $(\lambda|\varphi, \boldsymbol{Y}_n)$, following the steps of algorithm below:

1. Set $t = n$ and sample $p(\lambda_n|\varphi, \boldsymbol{Y}_n)$;
2. Set $t = t - 1$ and sample $p(\lambda_t|\lambda_{t+1}, \varphi, \boldsymbol{Y}_t)$;
3. if $t > 1$, go back to (2); otherwise, the sample of $(\lambda_1, \ldots, \lambda_n|\varphi, \boldsymbol{Y}_n)$ is complete.

The result presented (21), utilized in smoothing algorithm above, is exact, however aproximations of smoothed distribution of $\lambda_{t-1}|\varphi, \boldsymbol{Y}_n$ may be obtained, which are showed below.

Theorem 3. Supposing that $\lambda_t|\varphi, \boldsymbol{Y}_n \sim \text{Gama}(a_t^n, b_t^n)$ and using the aproximation
$\mathbb{E}(\exp(-qZ)) \doteq \exp(-q\mathbb{E}(Z))$ in which q is a constant and Z is a random variable, it is possible to obtain

$$\lambda_{t-1}|\varphi, \boldsymbol{Y}_n \dot{\sim} \text{Gama}(a_{t-1}^n, b_{t-1}^n);$$

where

$$a_{t-1}^n = a_t^n + (1-w)a_{t-1} \quad (22)$$

$$b_{t-1}^n = b_t^n + \frac{(1-w)a_{t-1}(b_{t-1} - b_t^n)}{a_{t-1}^n}, \quad (23)$$

for $t = 1, \ldots, n$. The recursions are initialized considering $a_n^n = a_n$ and $b_n^n = b_n$. The Theorem 3 demonstration is in the Appendix 3.

4 Special cases of Gamma family of dynamic models

In the next subsections, some particular cases of GFDM will be discussed such as the Poisson, Gamma, Weibull, Beta and Normal models.

4.1 The Poisson Model

Suppose that an observation in instant t is drawn of the Poisson distribution with mean μ_t,

$$p(y_t|\mu_t, \boldsymbol{\varphi}) = \mu_t^{y_t} \exp(-\mu_t)/y_t!, \tag{24}$$

where $y_t = 0, 1, \ldots$, $\mu_t = \lambda_t g(x_t, \boldsymbol{\beta})$. This model belong to Gamma family of dynamic models in which $a(y_t, \boldsymbol{\varphi}) = (y_t!)^{-1}$, $b(y_t, \boldsymbol{\varphi}) = y_t$ and $c(y_t, \varphi) = 1$. Then, $\boldsymbol{\varphi} = (w, \beta)^{'}$.

The prior distribution is the same of Theorem 1. With the functions $b(\cdot,\cdot)$ and $c(\cdot,\cdot)$, using the Theorem 1, the posterior distribution of $\mu_t|\boldsymbol{Y}_t$ is given by Gamma distribution with parameters

$$\begin{aligned} a_t^* &= a_{t|t-1}^* + y_t, \\ b_t^* &= b_{t|t-1}^* + 1. \end{aligned}$$

Therefore, it follows that $\lambda_t = \mu_t[g(x_t^{'}\boldsymbol{\beta})]^{-1}|\boldsymbol{Y}_t$ has Gamma distribution with parameters (update equations)

$$\begin{aligned} a_t &= w a_{t-1} + y_t, \\ b_t &= w b_{t-1} + g(x_t, \boldsymbol{\beta}). \end{aligned}$$

Replacing the functions $a(\cdot,\cdot)$, $b(\cdot,\cdot)$, $c(\cdot,\cdot)$ and using Theorem 1, it obtains the predictive distribution, what is Negative binomial, given by

$$p(y_t|\boldsymbol{Y}_{t-1}, \boldsymbol{\varphi}) = \begin{pmatrix} a_{t|t-1}^* + y_t + 1 \\ y_t \end{pmatrix} (b_{t|t-1}^*)^{a_{t|t-1}^*} (1 + b_{t|t-1}^*)^{-(a_{t|t-1}^* + y_t)} \quad ,$$

in which $y_t = 0, 1, 2, \ldots$

and

$$\begin{pmatrix} a_{t|t-1}^* + y_t + 1 \\ y_t \end{pmatrix} = \frac{\Gamma(a_{t|t-1}^* + y_t)}{\Gamma(y_t + 1)\Gamma(a_{t|t-1}^*)}.$$

The likelihood function has the following form

$$\ln L(\boldsymbol{\varphi};\boldsymbol{Y}_n)=\sum_{t=1}^{n}\ln\Gamma(a^*_{t|t-1}+y_t)-\ln y_t!-\ln\Gamma(a^*_{t|t-1})+ \\ a^*_{t|t-1}\ln b^*_{t|t-1}-(a^*_{|t-1}+y_t)\ln(1+b^*_{t|t-1}), \tag{25}$$

where $\boldsymbol{\varphi}=(w,\beta)^{'}$.

Then, by properties of Negative binomial distribution, the mean and variance of predictive distribution of $y_{t+1}|\boldsymbol{Y}_t,\boldsymbol{\varphi}$ are, respectively,

$$y_{t+1}=\mathbb{E}(y_{t+1}|\boldsymbol{Y}_t,\boldsymbol{\varphi})=a^*_{t+1|t}/b^*_{t+1|t}$$

and

$$var(y_{t+1}|\boldsymbol{Y}_t,\boldsymbol{\varphi})=a^*_{t+1|t}(1+b^*_{t+1|t})/(b^*_{t+1|t})^2.$$

4.2 The Gamma model

Suppose that the time series $\{y_t\}$ is generated Gamma distribution with unknown shape parameter χ and scale parameter $\chi\mu_t$, then:

$$p(y_t|\mu_t,\boldsymbol{\varphi})=\frac{y_t^{\chi-1}\exp(-\mu_t\chi y_t)}{\Gamma(\chi)(\mu_t\chi)^{-\chi}},y_t>0 \tag{26}$$

where $\mu_t=\lambda_t g(x_t,\boldsymbol{\beta})$ and $\forall t\leq n$. The expected value $(y_t|\mu_t,\boldsymbol{\varphi})$ is $1/\mu_t$. If $\chi=1$, $(y_t|\mu_t,\boldsymbol{\varphi})$ has exponential distribution with mean $1/\mu_t$.

The model Gamma can be written in the form of Gamma family of dynamic models in which $a(y_t,\boldsymbol{\varphi})=\frac{y_t^{\chi-1}\chi^{\chi}}{\Gamma(\chi)}$, $b(y_t,\boldsymbol{\varphi})=\chi$ and $c(y_t,\boldsymbol{\varphi})=\chi y_t$.

By Theorem 1, given the t-th observation and the functions $b(\cdot,\cdot)$ and $c(\cdot,\cdot)$, the posterior distribution of $\mu_t|\boldsymbol{Y}_t$ is Gamma with parameters

$$\begin{aligned} a^*_t &= a^*_{t|t-1}+\chi \\ b^*_t &= b^*_{t|t-1}+\chi y_t. \end{aligned}$$

So, it follows that $(\lambda_t=\mu_t[g(x_t,\boldsymbol{\beta})]^{-1})|\boldsymbol{Y}_t\sim$ Gamma(a_t,b_t) and update equations are given by:

$$\begin{aligned} a_t &= a_{t|t-1}+\chi \\ b_t &= b_{t|t-1}+\chi y_t g(x_t,\boldsymbol{\beta}). \end{aligned}$$

Replacing the functions $a(\cdot,\cdot)$, $b(\cdot,\cdot)$ and $c(\cdot,\cdot)$ in (10), the predictive density function $y_t|\boldsymbol{Y}_{t-1}, \boldsymbol{\varphi} \sim$ Gamma-gamma$(a^*_{t|t-1}, b^*_{t|t-1}/\chi, \chi)$ whose form is given by:

$$p(y_t|Y_{t-1}, \boldsymbol{\varphi}) = \frac{\Gamma(\chi + a^*_{t|t-1}) y_t^{\chi-1}}{\Gamma(a^*_{t|t-1})\Gamma(\chi)(b^*_{t|t-1}/\chi)^{-a^*_{t|t-1}}(y_t + b^*_{t|t-1}/\chi)^{\chi + a^*_{t|t-1}}},$$

if $y_t > 0$ and 0, otherwise.

The likelihood function is the product of predictive density functions given by:

$$\begin{aligned}
\ln L(\boldsymbol{\varphi}; \boldsymbol{Y}_n) &= \ln \prod_{t=1}^{n} p(y_t|\boldsymbol{Y}_{t-1}, \boldsymbol{\varphi}) \\
&= \sum_{t=1}^{n} \ln p(y_t|\boldsymbol{Y}_{t-1}, \boldsymbol{\varphi}) \\
&= \begin{array}{l} \sum_{t=1}^{n} \ln \Gamma(\chi + a^*_{t|t-1}) - \ln(\Gamma(\chi)\Gamma(a^*_{t|t-1})) + \\ a^*_{t|t-1} \ln(b^*_{t|t-1}/\chi) + \ln y_t^{(\chi-1)} - (\chi + a^*_{t|t-1}) \ln(y_t + b^*_{t|t-1}/\chi), \end{array}
\end{aligned}$$

where $\boldsymbol{\varphi} = (w, \beta, \chi)'$.

By properties of Gamma-gamma distribution, the mean and the variance of y_{t+1}, conditioned the information until time t, are

$$y_{t+1|t} = \mathbb{E}(y_{t+1}|\boldsymbol{Y}_t, \boldsymbol{\varphi}) = \frac{b^*_{t|t-1}}{a^*_{t|t-1} - 1}$$

and

$$var(y_{t+1}|\boldsymbol{Y}_t, \boldsymbol{\varphi}) = \frac{(b^*_{t|t-1})^2[\chi^2 + \chi(a^*_{t|t-1} - 1)]}{(a^*_{t|t-1} - 1)^2(a^*_{t|t-1} - 2)}.$$

4.3 The Weibull model

If the observations at instant t are generated of triparametric Weibull distribution (Ross , 2002) and parameters $\nu_t = \nu$ and $\rho_t = \rho = 0$ are invariants in time and unknown, so:

$$p(y_t|\mu_t, \boldsymbol{\varphi}) = \nu\mu_t(y_t)^{\nu-1} \exp[-\mu_t(y_t)^{\nu}], \qquad (27)$$

where $y_t > 0$, μ_t, $\nu > 0$ and $\mu_t = \lambda_t g(x_t, \boldsymbol{\beta})$.

The Weibull model can be written in the Gamma family of dynamic models form in which $a(y_t,\boldsymbol{\varphi}) = \nu(y_t)^{\nu-1}$, $b(y_t,\boldsymbol{\varphi}) = 1$ and $c(y_t,\boldsymbol{\varphi}) = (y_t)^{\nu}$.

By Theorem 1, the posterior distribution of $\mu_t|\boldsymbol{Y}_t$ is Gamma with parameters

$$\begin{aligned} a_t^* &= a_{t|t-1}^* + 1, \\ b_t^* &= b_{t|t-1}^* + (y_t)^{\nu}. \end{aligned}$$

Then, according to Theorem 1, it has that $\lambda_t = \mu_t[g(x_t,\boldsymbol{\beta})]^{-1}|\boldsymbol{Y}_t \sim \text{Gamma}(a_t, b_t)$ and update equations are given by:

$$a_t = a_{t|t-1} + 1, \tag{28}$$

$$b_t = b_{t|t-1} + (y_t)^{\nu} g(x_t^{'}\boldsymbol{\beta}). \tag{29}$$

Until this momment, was not done any remark for data with censoring, which are very communs in survival analysis. Now, can be assumed that y_t is observed if $\delta_t = 1$ or right censoring if $\delta_t = 0$. In this way, via Bayes' theorem, the update equation a_t in (28) become $a_t = a_{t|t-1} + \delta_t$, where δ_t is a indicator of right censoring.

Knowing $a(\cdot,\cdot)$, $b(\cdot,\cdot)$ and $c(\cdot,\cdot)$ and using the Teorema 1, the predictive density function $y_t|\boldsymbol{Y}_{t-1},\varphi$ is given by:

$$\frac{\Gamma(1 + a_{t|t-1}^*)\nu(y_t)^{\nu-1}}{\Gamma(a_{t|t-1}^*)(b_{t|t-1}^*)^{-a_{t|t-1}^*}[(y_t)^{\nu} + b_{t|t-1}^*]^{1+a_{t|t-1}^*}};$$

where $y_t > 0$, $a_{t|t-1}^*$ and $b_{t|t-1}^*$ are given by Theorem 1.

The likelhood function, what is the product of predictive density functions, is given by:

$$\begin{aligned} \ln L(\boldsymbol{\varphi};\boldsymbol{Y}_n) &= \ln \prod_{t=1}^{n} p(y_t|\boldsymbol{Y}_{t-1},\boldsymbol{\varphi}) \\ &= \sum_{t=1}^{n} \ln\Gamma(1 + a_{t|t-1}^*) + \ln \nu(y_t)^{\nu-1} \\ &\quad - \ln\Gamma(a_{t|t-1}^*) + a_{t|t-1}^* \ln b_{t|t-1}^* - (1 + a_{t|t-1}^*) \ln[(y_t)^{\nu} + b_{t|t-1}^*], \end{aligned}$$

in which $\boldsymbol{\varphi} = (\omega, \beta, \nu)^{'}$.

4.4 The Pareto model

The Pareto distribution has several applications in economic, social and geophysical problems (Johnson, Kotz & Balakrishnan , 1997). If the observations at time t are generated of Pareto distribution with parameters

$\rho > 0$, unknown and invariant in time, and $\mu_t > 0$, so:

$$p(y_t|\mu_t, \boldsymbol{\varphi}) = \mu_t \rho^{\mu_t} y_t^{-\mu_t - 1}, \tag{30}$$

where $y_t > \rho$ and $\mu_t = \lambda_t g(x_t, \boldsymbol{\beta})$.

Also, the Pareto model can be written in Gamma family of dynamic models form in which $a(y_t, \boldsymbol{\varphi}) = y_t^{-1}$, $b(y_t, \boldsymbol{\varphi}) = 1$, $c(y_t, \boldsymbol{\varphi}) = \ln y_t - \ln \rho$ and $H(\boldsymbol{\varphi}) = \rho$.

The prior distribution are the same of Gamma family of dynamic models. When the t-th observation is obtained, the posterior distribution of $\mu_t|Y_t$, by Theorem 1, is Gamma with parameters

$$\begin{aligned} a_t^* &= a_{t|t-1}^* + 1, \\ b_t^* &= b_{t|t-1}^* - \ln \rho + \ln y_t. \end{aligned}$$

Then, making the inverse transformation, it follows that $\lambda_t = \mu_t [g(x_t, \boldsymbol{\beta})]^{-1} | \boldsymbol{Y}_{t-1} \sim Gamma(a_t, b_t)$ where the update equation are given by:

$$\begin{aligned} a_t &= a_{t|t-1} + 1, \\ b_t &= b_{t|t-1} + (-\ln \rho + \ln y_t) g(x_t, \boldsymbol{\beta}). \end{aligned}$$

By Theorem 1, with the functions $a(\cdot, \cdot)$, $b(\cdot, \cdot)$ and $c(\cdot, \cdot)$, the predictive density function $y_t | \boldsymbol{Y}_{t-1}, \boldsymbol{\varphi}$ has the following form:

$$\frac{a_{t|t-1}^* y_t^{-1} [b_{t|t-1}^*]^{a_{t|t-1}^*}}{[b_{t|t-1}^* - \ln \rho + \ln y_t]^{a_{t|t-1}^* + 1}},$$

where $y_t > \rho$.

With the predictive density functions, it is possible to determine the likelihood function, which is:

$$\begin{aligned} \ln L(\boldsymbol{\varphi}; \boldsymbol{Y}_n) &= \ln \prod_{t=1}^{n} p(y_t | \boldsymbol{Y}_{t-1}, \boldsymbol{\varphi}) \\ &= \sum_{t=1}^{n} \ln p(y_t | Y_{t-1}, \boldsymbol{\varphi}) \\ &= \sum_{t=1}^{n} \ln(a_{t|t-1}^* y_t^{-1} (b_{t|t-1}^*)^{a_{t|t-1}^*}) - (a_{t|t-1}^* + 1) \ln[b_{t|t-1}^* - \ln \rho + \ln y_t], \end{aligned}$$

where $\boldsymbol{\varphi} = (w, \beta, \rho)'$.

4.5 The Beta model

When a parameter of Beta distribution is equal to 1, it can be written in GFDM form. This model is very useful for modeling time series that are proportions and probabilities. Suppose that the time series $\{y_t\}$ is generated of the Beta distribution with parameters μ_t and 1 whose density function is given by

$$p(y_t|\mu_t, \boldsymbol{\varphi}) = \mu_t y_t^{\mu_t - 1}, \tag{31}$$

where $0 < y_t < 1$.

This model belong to the GFDM in which $a(y_t, \boldsymbol{\varphi}) = y_t^{-1}$, $b(y_t, \boldsymbol{\varphi}) = 1$ and $c(y_t, \boldsymbol{\varphi}) = -\ln(y_t)$. The update equations and the predictive density function can be found similarly as the previous models, using Theorem 1. In this case, $\boldsymbol{\varphi} = (w, \beta)^{'}$.

4.6 The Normal model

The methodology developed in this article, can be applied, also, to gaussian models. If the observations at time t are generated of Normal distribution with mean $z_t^{'}\phi$ (z_t is a covariate vector with respct to time series mean) and precision parameter (the inverse of variance) $\mu_t > 0$, then:

$$p(y_t|\mu_t, \boldsymbol{\varphi}) = \frac{\mu_t^{1/2}}{\sqrt{2\pi}} \exp\left(\frac{-\mu_t(y_t - z_t^{'}\phi)^2}{2}\right), \tag{32}$$

where $-\infty < y_t < \infty$ e $\mu_t = \lambda_t g(x_t, \boldsymbol{\beta})$.

The Normal model may be written in GFDM form that $a(y_t, \boldsymbol{\varphi}) = (2\pi)^{-1/2}$, $b(y_t, \boldsymbol{\varphi}) = 1/2$ and $c(y_t, \boldsymbol{\varphi}) = (y_t - z_t^{'}\phi)^2/2$.

The prior distribution $\lambda_t|\boldsymbol{Y}_{t-1}$ is given by item 2 of Theorem 1. As the Theorem 1, the posterior distribution of $\mu_t|\boldsymbol{Y}_t$ also can be obtained, being Gamma with parameters

$$\begin{aligned} a_t^* &= a_{t|t-1}^* + 1/2, \\ b_t^* &= b_{t|t-1}^* + (y_t - z_t^{'}\phi)^2/2. \end{aligned}$$

Therefore, using again the scale property of Gamma distribution, it follows that $\lambda_t = \mu_t[g(x_t, \boldsymbol{\beta})]^{-1}|\boldsymbol{Y}_t \sim \text{Gamma}(a_t, b_t)$ where the update equations are given by:

$$\begin{aligned} a_t &= a_{t|t-1} + 1/2, \\ b_t &= b_{t|t-1} + ((y_t - z_t^{'}\phi)^2/2)g(x_t, \beta). \end{aligned}$$

The predictive density function $y_t|\boldsymbol{Y}_{t-1}, \boldsymbol{\varphi}$, by Theorem 1, has the following form:

$$\frac{\Gamma(a^*_{t|t-1} + 1/2)(2\pi)^{-1/2}(b^*_{t|t-1})^{a^*_{t|t-1}}}{\Gamma(a^*_{t|t-1})[(y_t - z'_t\phi)^2/2 + b^*_{t|t-1}]^{a^*_{t|t-1}+1/2}},$$

where $-\infty < y_t < \infty$.

With the predictive density functions, it possible to build the likelihood function:

$$\begin{aligned}
\ln L(\boldsymbol{\varphi}; \boldsymbol{Y}_n) &= \ln \prod_{t=1}^{n} p(y_t|Y_{t-1}, \boldsymbol{\varphi}) \\
&= \sum_{t=1}^{n} \ln p(y_t|Y_{t-1}, \boldsymbol{\varphi}) \\
&= \begin{array}{l} \sum_{t=1}^{n} \ln \Gamma(a^*_{t|t-1} + 1/2)(2\pi)^{-1/2}(b^*_{t|t-1})^{a^*_{t|t-1}} - \\ \ln \Gamma(a^*_{t|t-1})[(y_t - z'_t\phi)^2/2 + b^*_{t|t-1}]^{a^*_{t|t-1}+1/2}, \end{array}
\end{aligned}$$

in which $\boldsymbol{\varphi} = (w, \phi, \beta)'$.

Following the same idea of Normal model with evolution equation in the variance, other models can be built. For example, the case of the Lognormal model whose density function is:

$$p(y_t|\mu_t) = \frac{\mu_t^{1/2}}{\sqrt{2\pi}} \exp\left(\frac{-\mu_t}{2}(\ln y_t - z'_t\phi)^2\right), \text{onde}$$

$y_t > 0$; $\mu_t > 0$; z'_t and ϕ are covariate and parameter vectors, respectively, such as $-\infty < z'_t\phi < +\infty$.

This distribution belong to GFDM with functions $a(y_t, \varphi) = \frac{1}{\sqrt{2\pi}}$, $b(y_t, \varphi) = 1/2$ and $c(y_t, \varphi) = \frac{-(\ln y_t - z'_t\phi)^2}{2}$.

5 Simulations study

In this section, Monte Carlo simulations are done for two models of the GFDM: Poisson and Gamma. The MLE and the Bayesian estimators are compared with respect to the bias and the MSE, as well as the credibility and confidence intervals are assessed with respect to the width and coverage rate. All results showed were obtained through programs developed in the software Ox (Doornik , 1999).

5.1 The Poisson model

Through Monte Carlo simulations, the performances of the maximum likelihood estimator(MLE) and the Bayesian estimators - BE-mean and BE-median - were investigated for time series with length $n = 100$, generated under the Poisson model with a covariable $x_t = \cos(2\pi t/12)$, for $t = 1, \ldots, n$, and parameters $w = 0.90$ and $\beta = 1$. Two chains with 5000 samples were generated of which the 3000 first were excluded. The number of Monte Caro replications was fixed in 500. The level of confidence and the credibility probability of the confidence and credibility intervals, respectively, were fixed in 0.95. The state initial condition was $\lambda_0|Y_0 \sim Gama(0.01, 0.01)$. The proper prior distributions Uniform were adopted for w and β.

In the Figure 1, is a time series simulated under Poisson model with the same parameters described above. Observe that time series osciltates around a mean level equal to 3. Note that, also, dashed and dotted lines which represent the mean of predictive distribution of the classical and Bayesian fits, respectively, from Poisson model follow well the behaviour of time series (full line).

From Figure 6, note that the values of Gelman and Rubin's method for assessing the convergence of generated chains via MCMC of w and β for each Monte Carlo are less than 1.07 and 1.04 (less than 1.10, reference value), respectively. Therefore, it has evidences of chains convergence (Gelman , 1996, ver).

In the Tabel 1, the MLE and the Bayesian estimators are compared with respect to the bias and MSE. For β, all estimators have bias and MSE values very close. For w, the BE-mean and EB-median present smaller MSE values than MLE.

In Table 2, the credibility and confidence intervals are compared through coverage rate and width. For β, the coverage rate of confidence interval is closer to nominal level of 0.95 than the credibility intervals. The intervals have coverage rate equals for w.

5.2 The Gamma model

The criteria of MC simulation from the Gamma model are similar to Poisson model, which was shown in the previous section. The performances of MLE and Bayesian estimators - BE-Mean and BE-Median - were investigated by Monte Carlo experiments for time series of length $n = 100$, generated under Gamma model with a covariate $x_t = \cos(2\pi t/12)$, $t = 1, \ldots, n$, and parameters $\omega = 0.90$, $\chi = 5.00$ and $\beta = 0.50$. Two chains of 5000 samples

Table 1: MLE and BE for the Poisson model.

	MLE	**BE-Median**	**BE-Mean**
φ	**estimate** **Bias** **(MSE)**	**estimate** **Bias** **(MSE)**	**estimate** **Bias** **(MSE)**
$w = 0.90$	0.917	0.899	0.893
	0.017	-0.001	-0.007
	(0.003)	(0.002)	(0.002)
$\beta = 1.00$	1.003	1.001	1.003
	0,003	0.001	0.003
	(0.011)	(0.011)	(0.011)

Table 2: Credibility and confidence intervals for the Poisson model with noimal level of 95%.

	Cred. Int.	**Conf. Int.**
φ	**mean limits** **width** (**coverage**)	**mean limits** **width** (**coverage**)
$w = 0.90$	[0.785; 0.966]	[0.651; 0.964]
	0.181	0.313
	(0.98)	(0.98)
$\beta = 1.00$	[0.801; 1.214]	[0.806; 1.199]
	0.413	0.393
	(0.98)	(0.97)

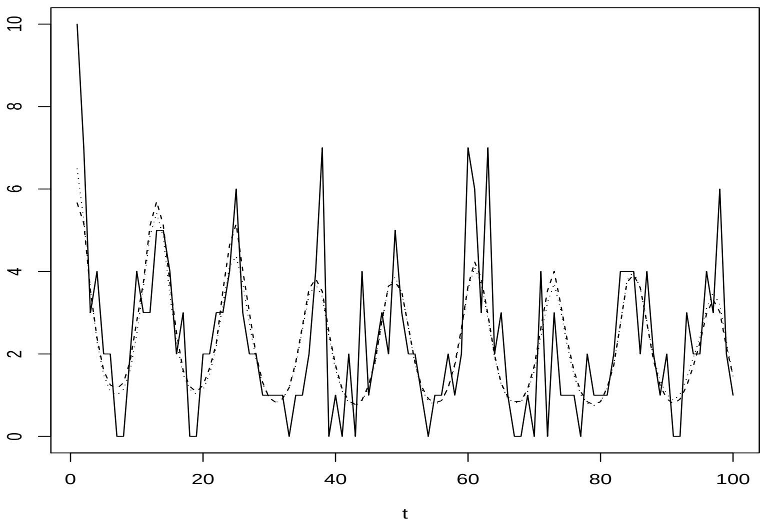

Figure 1: The full line represents the simulated time series under the Poisson model, the dashed and dotted lines indicate the smoothing mean of classical and Bayesian fits, using the exact smoothing method, respectively.

were obtained of which the 3000 first were excluded. The number of Monte Caro replications was fixed in 500. The level of confidence and the credibility probability of the confidence and credibility intervals, respectively, were fixed in 0.95. The state initial condition was $\lambda_0|Y_0 \sim Gama(0.01, 0.01)$. The proper prior distributions Uniform were adopted for w, β and χ.

The Figure 2 presents an example of simulated time series as this model. The dashed and dotted lines represent the smoothed mean of classical and Bayesian fits, respectively, and have a similar behaviour.

From Figure 7, observe that the values of Gelman and Rubin's method of generated chains by MCMC for w, β and χ are less than 1.07, 1.03 ad 1.03, respectively. so, it has evidences of the chains convergence.

From Tabel 5 - what presents the results of MC study of classical and Bayesian point estimators -, note that the estimators have the same MSE except the parameter χ whose MLE has MSE slightly small. The MLE possesses smaller bias than the Bayesian estimators (BE-Mean and BE-Median) for all parameters of model.

From Table 6, observe that the intervals possess a coverage rate of 0.98

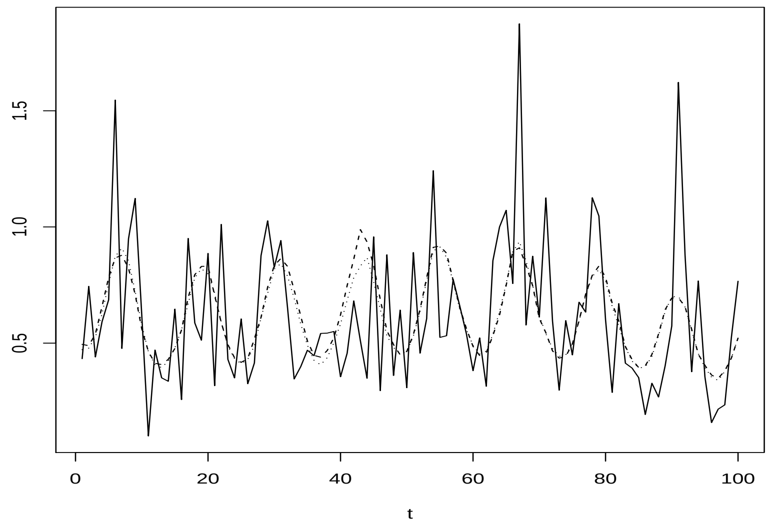

Figure 2: The full line represents the time series simulated under Gamma model, the dashed and dotted lines indicate the smoothed mean of the clas-sial and Bayesian fitso, using the exact smooth method, respectively.

Table 3: MLE and BE for Gamma model.

	MLE **Bias** **(MSE)**	**BE-Mean** **Bias** **(MSE)**	**BE-Median** **Bias** **(MSE)**
$\omega = 0.90$	0.905	0.876	0.883
	0.005	-0.024	0.017)
	(0.003)	(0.003)	(0.003)
$\beta = 0.50$	0.486	0.483	0.483
	-0.014	-0.017	-0.017
	(0.004)	(0.004)	(0.004)
$\chi = 5.00$	5.174	5.349	5.295
	0.174	0.349	0.295
	(0.488)	(0.597)	(0.549)

Table 4: Confidence and credibility intervals for Gamma model with nominal level of 95%.

	Conf. Int. mean limits width (coverage)	Cred. Int. mean limits width (coverage)
$\omega = 0.90$	[0.666; 0.960] 0.294 (0.98)	[0.750; 0.959] 0.209 (0.98)
$\beta = 0.50$	[0.357; 0.612] 0.255 (0.96)	[0.347; 0.619] 0.272 (0.98)
$\chi = 5.00$	[3.701; 6.647] 2.946 (0.98)	[3.894; 7.096] 3.202 (0.98)

for all parameters of model, except the confidence interval for β whose coverage rate is 0.96. In general, the confidence interval has a width slightly lesser than the credibility interval.

6 Application to real time series

6.1 ASIAV time series

In this subsection, Poisson model is adjust to the monthly data of the patient number with **affection** of the superior and inferior aerial vias - ASIAV - from São Paulo city , from 1997 to 1999, which is composed of 48 observations. A time series of sulfur dioxide SO_2 (x_t) is considered, $t = 1, \ldots, 48$. Os dados of the patient number with affection of the superior and inferior aerial vias and the pollutant SO_2 were obtained by Health ministery (http://www.datasus.gov.br) and by Technology and **saneamento** environmental company from São Paulo (CETESB), respectively.

The Figure 3 shows the graphs of series ASIAV and SO_2. It does not obseve discrepant values of ASIAV in years assessed. The pollutant concentration SO_2 presents the highiest mean value in 1997, as well as o number with affection of the superior and inferior aerial vias (see Figure 3). It seems that exists a relation between these series. To follow, is showed the results of Poisson model fit to ASIAI series.

In the Table 5, are the MLE and the Bayesian estimates of the parameters w and β from Poisson model. The last is associated with a covariable x_t (SO_2) and its estiamte is equal among all methods (0.023). Already, the estimates of w is about 0.69.

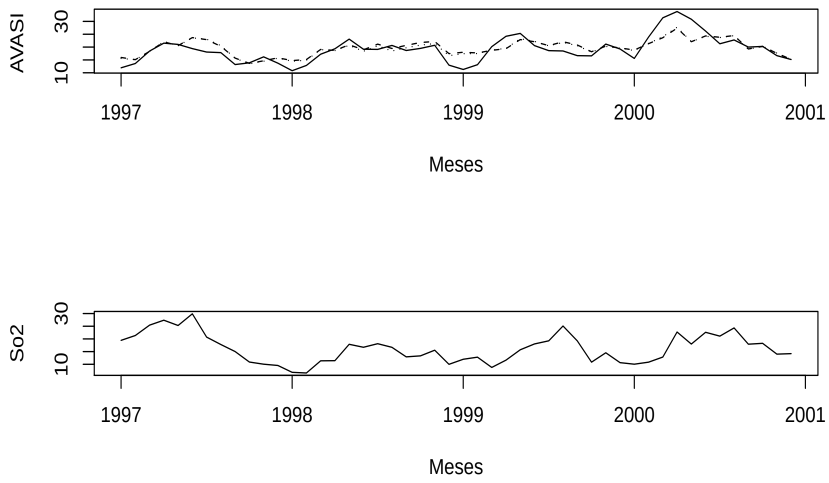

Figure 3: Graph 1: The full, dashed and dotted represent ASIAV time series, the smoothed mean of the classical and Bayesian, respectively. Graph 2: time series of poluent SO_2 in the years from 1997 to 2000.

Table 5: MLE and BE for the parameters of Poisson model fitted to the ASIAV time series.

	MLE	**BE-Median**	**BE-Mean**
w	0.698	0.684	0.683
β	0.023	0.023	0.023

From Tabel 6, it concludes that the parameters β and w are significants at confidence and credibility level of 0.95. Note that the lower limit of both intervals for β is close to zero.

It Tried include in the model **sine** and **cosine** covariables, computed at the Fourrier's frenquecies, for capturing some sazonality structure, however they were not significants. It is important to emphasize that the ASIAV series has a few observations ($n = 48$), so that the model can capture some sazonality structure from it. The model fit does not present inadequacy evidences.

The values Gelman and Rubin's criterion (Gelman , 1996) for the two chains of parameters w and β are 1.00 and 1.00, respectively. All values are

Table 6: Confidence and credibility intervals for the parameters of Poisson model fitted to the ASIAV time series with nomial level of 95%.

	Int. Assint.	**Int. Cred.**
w	[0.458; 0.849]	[0.492; 0.864]
β	[0.001; 0.038]	[0.007; 0.039]

close to 1, indicating the chains convergence. The Figure 8 shows the graph of two chains of each parameter. Observe that they superimpose, which is one evidence of convegence of the chains.

6.2 Petrobrás log-returns time series

The Normal model, described in the Subsection 3.3.1, is fitted to the daily data of Petrobrás stock market returns (Petro) in the period from 01/03/1995 a 12/27/2000, totalizing 1498 observations. It is important to emphasize that researches have been developed in order to incorporate a stationary evolution equation, from this way, allowing that the returns are stationaries.

The Figure 4 presents the times series plot of logarithm of Petrobrás returns. Once it presents a correlation structure with respect to the mean, first is fitted a Local Lvel Model (LLM) (Harvey, 1989) or a linear dynamic model with autoregressive structure. The residuals of LLM fiited also is shoed in the Figure 4 and is used for the fit from Normal model (Subsection 3.6) with mean zero, known. So after the fit of LLM, assumes that the residuals of fit $e_t|\sigma_t^2 \sim N(0, \sigma_t^2)$, so it has that Normal model (Subsection 3.6) with zero mean and precision $\lambda_t = \frac{1}{\sigma_t^2}$. Finally, it has the fit of LLM-Normal model.

Already, the Normal model with unknown mean is fitted considering the mean ϕy_{t-1} with autoregressive structure and precision $\lambda_t = \frac{1}{\sigma_t^2}$ (see, Subsection 3.6), named Normal model in tables and graphs, which it will be showed below. To treat as the mean as the variance of time series is done only one model fit, while in other proposed is done a combination of models, that is, the treatment is performed in two steps: one model for mean and one model for variance.

The Table 7 is composed of the log-likelihood, AIC and BIC for some models fiited. According to Harvey (1989), it adopts AIC and BIC criterion being $AIC = -2l(\hat{\varphi}) + 2p$ and $BIC = -2l(\hat{\varphi}) + n\ln(p)$, in which $l(\cdot)$ is the log-likelihood value, p the number os model parameters and n the number of observations. The DIC criterion (*Deviance Information Criterion*)

(Spiegelhalter et al. , 2002, see) is using for comparing the Bayesian models.

Normal model with unkonwn mean possesses the biggest log-likelihood values and the smallest AIC and BIC values, comparing to LLM-Normal model (mean know).

Table 7: Values of log-likelihood, AIC, BIC and DIC for the models fitted to the Petrobrás returns.

Models	log-likelihood	AIC	BIC	DIC
LLM + Normal[1] classical	3368.61	-6729.22	-4660.55	-
LLM + Normal[1] Bayesian	-	-	-	-6732.10
Normal[2] classical	**3371.92**	**-6739.84**	**-5705.51**	-
Normal[2] Bayesian	-	-	-	**-6739.60**

[1]zero mean. [2]unkonwn mean.

The results of the estimation from the models LLM-Normal and Normal model. The MLE from LLM with autoregressive structure are $\hat{\sigma}_{\eta}^2 = 0.001$, $\hat{\sigma}_{\epsilon}^2 = 0.000$ and $\hat{\rho} = 0.106$. From the Normal model with zero mean, w is the only parameter to be estimated. The MLE, the BE-Median and the BE-Mean of w are 0.798, 0.798 and 0.797, respectively. While the credibility and confidence intervals at level of 95% are $[0,769; 0,827]$ and $[0,768; 0,828]$, respectively.

Already, for the fit of Normal model with unknown mean, the MLE, the BE-Median and the BE-Mean of ϕ (w) are 0.126 (0.799), 0.126 (0.798) amd 0.126 (0.798), respectively. The credibility and confidence intervals at level of 95% for parameters ϕ (w) are $[0.074; 0,178]$ ($[0,770; 0,828]$) and $[0.071; 0,181]$ ($[0,768; 0,828]$), respectively.

The values of Gelman and Rubin's method (Gelman , 1996) for two chains is equal to 1.00 for the parameter w and observe that, in the Figure 9, the two chains for each parameter superimpose, indicating the convergence. The same can be observed for the Bayesian fit of Normal model with unknown mean. Figure 10 shows the generated chains of w and ϕ. In Figure 5, it finds the graph of volatility obtained by Normal model with unkonwn mean under classical and Bayesian approaches in several instants in time. It possessses discreptants values (pulses) which are explained by crisis period, hnown in the literature. It highlights that the fit captures the Nasdaq's drop in April/2000.

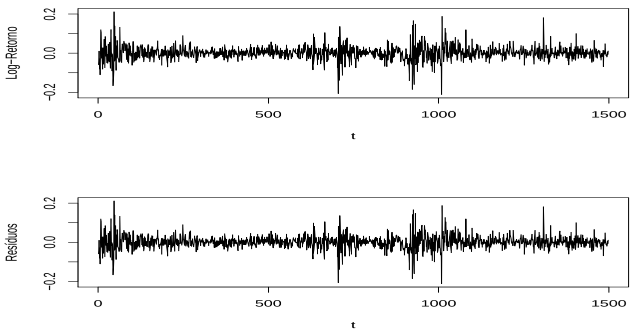

Figure 4: Graph of the Petrobrás log-return and residuals of the LLM fit, respectivamente.

7 Conclusions and final remarks

In this work, was presented a new family of dynamic models (the GFDM), as well as particular cases from it and a new way of making smooth of the level component. Through Monte Carlo experiments, was checked the performance of the point and intervalar estimators (classical and Bayesian) from the Poisson and Gamma models in finite samples. The results shown that both estimators possess MSE relatively small. Already, the intervalar estimators have a behaviour similar with respect to the coverage rate, although the confidence interval has the smallest width.

With respect to the non-gaussian models, several works can be developed. A study exploring the properties of Gamma family would be very relevant. Specific cases of the Gamma family of dynamic models can be found. An evolution equation with a autoregressive structure can be proposed. The piecewise exponential model is very used and it has an ample application, mainly, in studies in reliability and it fits into the molds of the non-gaussian models, however the observation dependence not be considered in the piecewise exponential model (Gamerman , 1991b), but it is can be applied to the models that take in count the autocorrelation of observations. Hyphotesis tests can be explored, as well as the use and application of the bootstrap in the GFDM, so that the inference about parameters model can be done. Other work interesting that can be developed is the comparation between the non-gaussian models and dynamic generalized linear models (West & Harrison , 1997) either via Monte Carlo experimentos or using real time series.

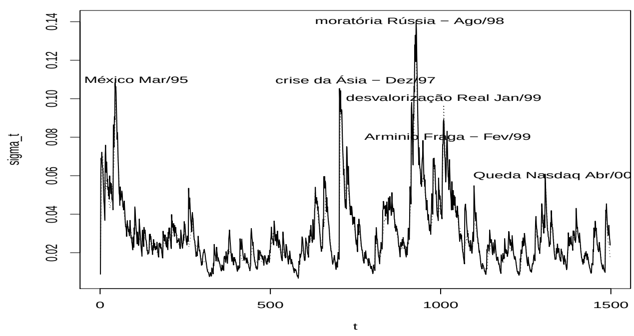

Figure 5: The full and dashed lines represent an smoothed estimative of sthocastic volatility, obtained by the fit of Normal model with unkonown mean under the classical and Bayesian approaches, respectively.

Acknowledgements

T. R. Santos was support by CNPq-Brazil and CAPES-Brazil. D. Gamerman was support by CNPq-Brazil and *Fundação de Amparo à Pesquisa no Estado do Rio de Janeiro* (FAPERJ foundation). G. C. Franco was partially support by CNPq-Brazil and *Fundação de Amparo à Pesquisa no Estado de Minas Gerais* (FAPEMIG foundation).

Appendix 1

Theorem 1 demonstration.

Assuming the model definition in the Section 2.1, The update equations can be obtained. It will be done the prove of the items 1, 2, 3, 4 and 5 from Theorem 1.
PS.: To facilitate the notation, the vector $\boldsymbol{\varphi}$ will be omitted in the distribution below.

Proof:

- If $t = 1$, $\lambda_0|Y_0 \sim Gamma(a_0, b_0)$ and $\mu_0|Y_0 \sim Gamma(a_0, b_0[g(x_t, \beta)]^{-1})$ - which is truth by assumption 4 of the model;

  By induction hyphotesis, supposing that $\lambda_{t-1}|\boldsymbol{Y}_{t-1} \sim Gamma(a_{t-1}, b_{t-1})$ s valid at t and, as result, $\mu_{t-1}|\boldsymbol{Y}_{t-1} \sim Gamma(a^*_{t-1}, b^*_{t-1})$ is valid at t:

- Now, it will prove that assumption is valid at $t + 1$.

  The distributions $\lambda_{t-1}|\boldsymbol{Y}_{t-1}$ and $\lambda_t|\lambda_{t-1}, \boldsymbol{Y}_{t-1}$ are known. The first by induction hyphotesis and the last by Lemma 1 below.

  1. Integrating out in λ_{t-1}, by Lemma II below, it conclude that $\lambda_t|\boldsymbol{Y}_{t-1} \sim Gamma\left(a_{t|t-1}, b_{t|t-1}\right)$ where $a_{t|t-1} = wa_{t-1}$ and $b_{t|t-1} = wb_{t-1}$.

  2. Therefore, from item (1), $(\mu_t = \lambda_t g(x_t, \boldsymbol{\beta}))|\mathbf{Y}_{t-1} \sim Gamma\left(a^*_{t|t-1}, b^*_{t|t-1}\right)$ onde $a^*_{t|t-1} = a_{t|t-1}$ e $b^*_{t|t-1} = b_{t|t-1} g(x_t, \boldsymbol{\beta})^{-1}$.
  3. By Bayes' theorem,
     $p(\mu_t|\boldsymbol{Y}_t) \propto p(y_t|\mu_t) p(\mu_t|\boldsymbol{Y}_{t-1}) \propto \mu_t^{(a^*_{t|t-1} + b(y_t, \boldsymbol{\varphi})) - 1} \exp[-\mu_t(b^*_{t|t-1} + c(y_t, \boldsymbol{\varphi}))]$.
     Then, it follows thar $\mu_t|\boldsymbol{Y}_t \sim Gamma\,(a^*_t, b^*_t)$, where $a^*_t = a^*_{t|t-1} + b(y_t, \boldsymbol{\varphi})$ and $b^*_t = b^*_{t|t-1} + c(y_t, \boldsymbol{\varphi})$.
  4. Using the item (3),
     $(\lambda_t = \mu_t g(x_t, \boldsymbol{\beta})^{-1})|\boldsymbol{Y}_t \sim Gamma\,(a_t, b_t)$, in which $a_t = a_{t|t-1} + b(y_t, \boldsymbol{\varphi})$ and $b_t = b_{t|t-1} + c(y_t, \boldsymbol{\varphi}) g(x_t, \boldsymbol{\beta})$; $\forall t \in N$, $t \leq n$ where n is the time series length.

The inductive hyphotesis is verified.

5. Demonstration of predictive distribution a step ahead:

$$
\begin{aligned}
p(y_t|\boldsymbol{Y}_{t-1},\boldsymbol{\varphi}) &= \int_0^\infty p(y_t|\mu_t,\boldsymbol{\varphi})p(\mu_t|\boldsymbol{Y}_{t-1},\boldsymbol{\varphi})d\mu_t \\
&= \frac{a(y_t,\boldsymbol{\varphi})}{\Gamma(a^*_{t|t-1})(b^*_{t|t-1})^{-a^*_{t|t-1}}} \int_0^\infty \left[\mu_t^{b(y_t,\boldsymbol{\varphi})+a^*_{t|t-1}-1} \exp\left(-\mu_t(c(y_t,\boldsymbol{\varphi})+b^*_{t|t-1})\right)\right] d\mu_t \\
&= \frac{\Gamma\left(b(y_t,\boldsymbol{\varphi})+a^*_{t|t-1}\right) a(y_t,\boldsymbol{\varphi})(b^*_{t|t-1})^{a^*_{t|t-1}}}{\Gamma(a^*_{t|t-1})\left(c(y_t,\boldsymbol{\varphi})+b^*_{t|t-1}\right)^{a^*_{t|t-1}+b(y_t,\boldsymbol{\varphi})}}; \qquad \text{where}
\end{aligned}
$$

$$a^*_{t|t-1} = wa_{t-1},\ b^*_{t|t-1} = wb_{t-1}g(x_t,\boldsymbol{\beta})^{-1} \text{ e } y_t \in H(\boldsymbol{\varphi}).$$

□

Lemma I. Knowing that $\varsigma_t \sim B\left(wa_{t-1},(1-w)a_{t-1}\right)$ by item 2 of GFDM definition, the distribution of $\lambda_t = w^{-1}\lambda_{t-1}\varsigma_t$ is expressed by equation (33).

Proof:

The, using the Jacobian's method, it has that

$$
p_{\lambda_t|\lambda_{t-1},\mathrm{Y}_{t-1}}(\lambda_t) = \begin{cases} \dfrac{\Gamma(wa_{t-1})\Gamma((1-w)a_{t-1})}{\Gamma(a_{t-1})} \dfrac{w}{\lambda_{t-1}} \left(\dfrac{w\lambda_t}{\lambda_{t-1}}\right)^{wa_{t-1}-1} \left(1-\dfrac{w\lambda_t}{\lambda_{t-1}}\right)^{(1-w)a_{t-1}-1}; \\ if \quad 0<\lambda_t<w^{-1}\lambda_{t-1}, \\ 0; \quad \text{otherwise.} \end{cases} \tag{33}
$$

□

Lemma II.

If $\lambda_{t-1}|Y_{t-1} \sim Gama(a_{t-1},b_{t-1})$ and the distribution of $\lambda_t = w^{-1}\lambda_{t-1}\varsigma_t$ is given by equation (33), then $\lambda_t|Y_{t-1} \sim Gama(wa_{t-1},wb_{t-1})$.

Proof:

$$
\begin{aligned}
p(\lambda_t|\boldsymbol{Y}_{t-1},\varphi) &= \int p(\lambda_{t-1}|\boldsymbol{Y}_{t-1},\varphi)p(\lambda_t|\lambda_{t-1},\boldsymbol{Y}_{t-1},\varphi)d\lambda_{t-1} \\
&= \int_{w\lambda_t}^{\infty} \left[\frac{\lambda_{t-1}^{a_{t-1}-1}\exp(-b_{t-1}\lambda_{t-1})}{\Gamma(a_{t-1})b_{t-1}^{-a_{t-1}}}\right] \left[\frac{w\lambda_{t-1}^{-1}(\frac{w\lambda_t}{\lambda_{t-1}})^{wa_{t-1}-1}(1-\frac{w\lambda_t}{\lambda_{t-1}})^{(1-w)a_{t-1}-1}}{\frac{\Gamma(wa_{t-1})\Gamma((1-w)a_{t-1})}{\Gamma(a_{t-1})}}\right] d_{\lambda_{t-1}}
\end{aligned}
$$

Be $c = \dfrac{w(w\lambda_t)^{wa_{t-1}-1}}{b_{t-1}^{-a_{t-1}}\Gamma(wa_{t-1})\Gamma((1-w)a_{t-1})}$, so,

$$
\begin{aligned}
p(\lambda_t|\boldsymbol{Y}_{t-1},\varphi) &= c\int_{w\lambda_t}^{\infty} \left[\lambda_{t-1}^{a_{t-1}-1-wa_{t-1}+1-1}\exp(-b_{t-1}\lambda_{t-1})\right] \left[(1-\frac{w\lambda_t}{\lambda_{t-1}})^{(1-w)a_{t-1}-1}\right] d_{\lambda_{t-1}} \\
&= c\int_{w\lambda_t}^{\infty} \exp(-b_{t-1}\lambda_{t-1})(\lambda_{t-1}-w\lambda_t)^{(1-w)a_{t-1}-1} d_{\lambda_{t-1}}
\end{aligned}
$$

Be $z = \lambda_{t-1} - w\lambda_t$, then

$$
\begin{aligned}
p(\lambda_t|\boldsymbol{Y}_{t-1},\varphi) &= c\int_{0}^{\infty} \exp\left[-b_{t-1}(z+w\lambda_t)\right] z^{(1-w)a_{t-1}-1}dz \\
&= \frac{w^{wa_{t-1}-1+1}(\lambda_t)^{wa_{t-1}-1}\Gamma((1-w)a_{t-1})}{\Gamma(a_{t-1})b_{t-1}^{-wa_{t-1}}\Gamma(wa_{t-1})\Gamma\left((1-w)a_{t-1}\right)\left[\Gamma(a_{t-1})\right]^{-1}}\exp(-wb_{t-1}\lambda_t) \\
&= \frac{\lambda_t^{wa_{t-1}-1}\exp(-wb_{t-1}\lambda_t)}{(wb_{t-1})^{-wa_{t-1}}\Gamma(wa_{t-1})}, \lambda_t > 0.
\end{aligned}
$$

□

Appendix 2

Theorem 2 demonstration:

PS.: To facilitate the notation, the vector $\boldsymbol{\varphi}$ will be omitted in the distribution below.

Proof:

$$\begin{aligned}
p(\lambda_{t-1}|\lambda_t, \boldsymbol{Y}_t) &= p(\lambda_{t-1}|\lambda_t, \boldsymbol{Y}_{t-1}), \text{ by Markovian structure of the model} \\
&= \frac{p(\lambda_t|\lambda_{t-1}, \boldsymbol{Y}_{t-1})p(\lambda_{t-1}|\boldsymbol{Y}_{t-1})}{p(\lambda_t|\boldsymbol{Y}_{t-1})} \\
&= \frac{w}{\lambda_{t-1}}\left(\frac{w\lambda_t}{\lambda_{t-1}}\right)^{wa_{t-1}-1}\left(1-\frac{w\lambda_t}{\lambda_{t-1}}\right)^{(1-w)a_{t-1}-1} \times \frac{\Gamma(wa_{t-1})\Gamma((1-w)a_{t-1})}{\Gamma(a_{t-1})} \times \\
&\quad \frac{b_{t-1}^{a_{t-1}}}{\Gamma(a_{t-1})}\lambda_{t-1}^{a_{t-1}-1}\exp(-\lambda_{t-1}b_{t-1})\frac{\frac{\Gamma(wa_{t-1})}{(wb_{t-1})^{wa_{t-1}}}}{\lambda_t^{wa_{t-1}-1}\exp(-\lambda_t w b_{t-1})} \\
&\propto \frac{(\lambda_t)^{wa_{t-1}}(\lambda_{t-1}-w\lambda_t)^{(1-w)a_{t-1}-1}\lambda_{t-1}^{a_{t-1}-1}\exp(-b_{t-1}(\lambda_{t-1}-w\lambda_t))}{\lambda_t^{a_{t-1}+wa_{t-1}-1-1}} \\
&\propto (\lambda_{t-1}-w\lambda_t)^{(1-w)a_{t-1}-1}\exp\left(-b_{t-1}(\lambda_{t-1}-w\lambda_t)\right)
\end{aligned}$$

Then, $\lambda_{t-1} - w\lambda_t|\lambda_t, \boldsymbol{Y}_{t-1} \sim \text{Gama}\left((1-w)a_{t-1}, b_{t-1}\right)$.

□

Appendix 3

Theorem 3 demonstration:

PS.: To facilitate the notation, the vector $\boldsymbol{\varphi}$ will be omitted in the distribution below.

Proof:
Note that $p(\lambda_{t-1}|\boldsymbol{Y}_n) = \int p(\lambda_{t-1}|\lambda_t, \boldsymbol{Y}_t)p(\lambda_t|\boldsymbol{Y}_n)d\lambda_t$. Supposing $\lambda_t|\boldsymbol{Y}_n \sim$

$\text{Gamma}\left(a_t^n, b_t^n\right)$, it follows that

$$\begin{aligned}
p(\lambda_{t-1}|\boldsymbol{Y}_n) &\propto \int (\lambda_{t-1}-w\lambda_t)^{(1-w)a_{t-1}}\,\lambda_t^{a_t^n-1}\exp\left(-b_{t-1}(\lambda_{t-1}-w\lambda_t)-b_t^n\lambda_t\right)d\lambda_t \\
&\propto \lambda_{t-1}^{(1-w)a_{t-1}}\exp(-b_{t-1}\lambda_{t-1}) \times \\
&\quad \int\left(1-\frac{w\lambda_t}{\lambda_{t-1}}\right)^{(1-w)a_{t-1}}\lambda_t^{a_t^n-1}\exp\left(-\lambda_t(b_t^n-wb_{t-1})\right)d\lambda_t
\end{aligned}$$

Making the change variable $Z = w\lambda_t/\lambda_{t-1}$, then

$$\begin{aligned}
&\propto \lambda_{t-1}^{(1-w)a_{t-1}} \exp(-b_{t-1}\lambda_{t-1}) \times \\
&\int (1-Z)^{(1-w)a_{t-1}} \left(\frac{\lambda_{t-1}Z}{w}\right)^{a_t^n-1} \exp\left(-\frac{\lambda_{t-1}Z}{w}(b_t^n - wb_{t-1})\right) \frac{\lambda_{t-1}}{w} dZ \\
&\propto \frac{\lambda_{t-1}^{(1-w)a_{t-1}+a_t^n}}{w^{a_t^n}} \exp(-b_{t-1}\lambda_{t-1}) \times \\
&\int_0^1 Z^{a_t^n-1} (1-Z)^{(1-w)a_{t-1}} \exp\left(-\frac{\lambda_{t-1}Z}{w}(b_t^n - wb_{t-1})\right) dZ
\end{aligned}$$

Using $\mathbb{E}(\exp(-qZ)) \doteq \exp(-q\mathbb{E}(Z))$, where

$$q = -\lambda_{t-1}w^{-1}(b_t^n - wb_{t-1}) \text{ e } \mathbb{E}(Z) = \tfrac{a_t^n}{a_t^n + (1-w)a_{t-1}},$$

$$\begin{aligned}
p(\lambda_{t-1}|\boldsymbol{Y}_n) &\propto \frac{\lambda_{t-1}^{(1-w)a_{t-1}+a_t^n}}{w^{a_t^n}} \exp(-b_{t-1}\lambda_{t-1}) \mathbb{E}\left[\exp\left(-\frac{\lambda_{t-1}Z}{w}(b_t^n - wb_{t-1})\right)\right] \\
&\doteq \frac{\lambda_{t-1}^{(1-w)a_{t-1}+a_t^n-1}}{w^{a_t^n}} \exp\left[-b_{t-1}\lambda_{t-1} - \frac{\lambda_{t-1}\mathbb{E}(Z)}{w}(b_t^n - wb_{t-1})\right] \\
&\propto \lambda_{t-1}^{(1-w)a_{t-1}+a_t^n-1} \exp\left[-b_{t-1}\lambda_{t-1} - \frac{\lambda_{t-1}\frac{a_t^n}{a_t^n+(1-w)a_{t-1}}}{w}(b_t^n - wb_{t-1})\right] \\
&= \lambda_{t-1}^{(1-w)a_{t-1}+a_t^n-1} \exp\left[-\lambda_{t-1}\left(b_{t-1} + \frac{a_t^n(b_t^n - wb_{t-1})}{w(a_t^n + (1-w)a_{t-1})}\right)\right]
\end{aligned}$$

Note that the expression above is the nuclues of Gamma distribution. Therefore,

$$p(\lambda_{t-1}|\boldsymbol{Y}_n) \dot{\sim} \mathrm{G}\left(a_{t-1}^n, b_{t-1}^n\right), \text{ onde}$$

$$a_{t-1}^n = a_t^n + (1-w)a_{t-1} \text{ e } b_{t-1}^n = b_t^n + \tfrac{(1-w)a_{t-1}(b_{t-1} - b_t^n)}{a_{t-1}^n}.$$

□

Appendix 4

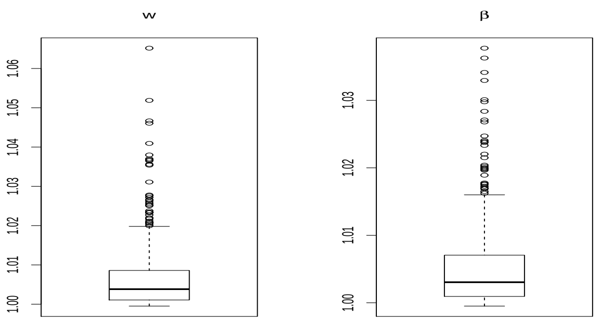

Figure 6: Boxplots for the Gelman and Rubin's values for diagnostic the chains convergence for parameters ω and β for the 500 Monte Carlo generated under the Poisson model, respectively.

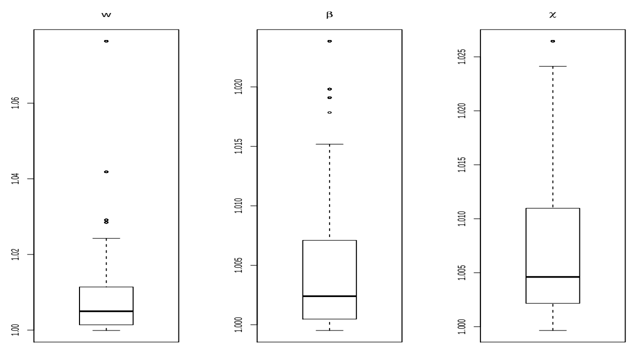

Figure 7: Boxplots of Gelman and Rubin's values for diagnostic of chain convergence for ω, β and χ for 500 Monte Carlo genearated under Gamma model, respectively.

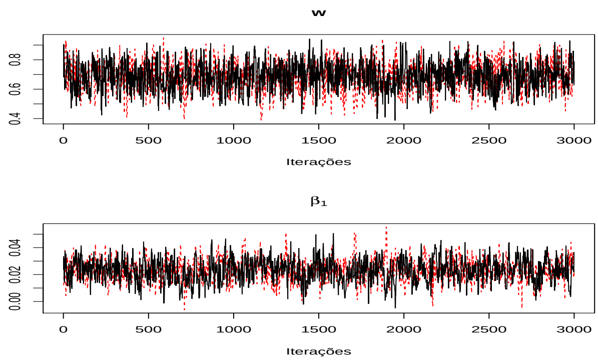

Figure 8: Graphs of the two chain generated by MCMC for the parameters w and β of the Bayesian model fit to the time series ASIAV, respectively.

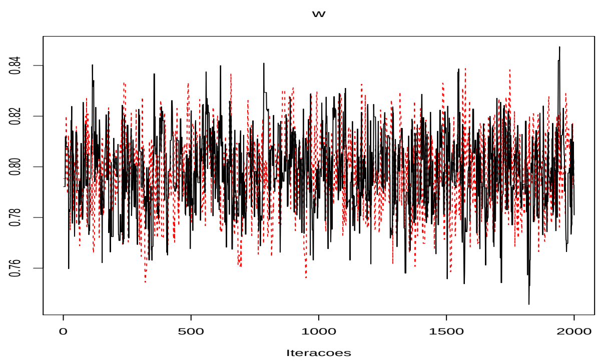

Figure 9: Graphs the two chains generated by MCMC for parameter w of the fit of Bayesian LLM-Normal model to the Petrobrás log-return time series, respectively.

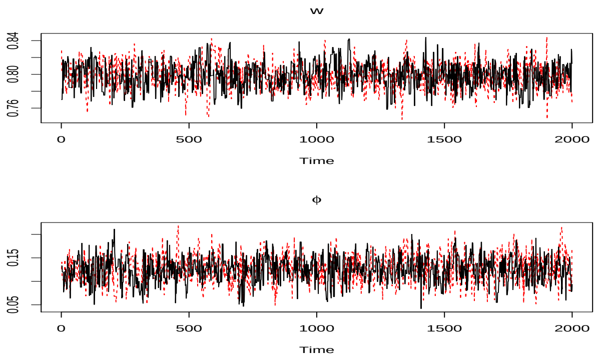

Figure 10: Graphs the two chains generated by MCMC for parameter w of the fit of Bayesian Normal model (GFDM, unknown mean) to the Petrobrás log-return time series, respectively.

References

Casella, G. and Berger, R. L. (2002). *Statistical Inference*, Second Edition, Duxbury, USA.

Cavanaugh, J.E. and Shumway, R.H. (1996). On computing the expected Fisher information matrix for state-space model parameters. *Statistics and Probability Letters*, Vol. **26**, p. 347-355.

Chiogna, M. and Gaetan, C. (2002). Dynamic generalized linear models with application to environmental epidemiology. *Applied Statistics*, **51**, 453-468.

Cox, D. R. (1981). Statistical analysis of time-series: some recent developments. *Scandanavian Journal of Statistics*, **8**, 93-115.

Dobson, A. J. (2002). *An Introduction to Generalized Linear Models*. 2nd edition, Chapman and Hall, New York.

Doornik, J. A. (1999). *Ox: An Object-Oriented Matrix Language*. 3. ed. London: Timberlake Consultants Press.

Durbin, J. and Koopman, S. J. (2000). Time series analysis of non-Gaussian observations based on state space models from both classical and Bayesian perspectives (with discussion). *Journal of the Royal Statistical Society Series B*, **62**, 3-56.

Engle, R.F. (1982). Autoregressive conditional heterocedasticity with estimates of the variance of U. K. inflation. *Econometrica*, **50**, 987-1008.

Fahrmeir, L. (1987). Regression models for nonstationary categorical time series. *Journal of Time Series Analysis*, **8**, 147-160.

Fruhwirth-Schnatter, S. (1994). Applied state space modelling of non-Gaussian time series using integration-based Kalman filtering. *Statistics and Computing*, **4**, 259-269.

Gamerman, D. and West, M. (1987). An application of dynamic survival models in unemployment studies. *The Statistician*, **36**, 269-274.

Gamerman, D. (1991). Dynamic Bayesian models for survival data. *Applied Statistics*, **40**, 63-79.

Gamerman, D. (1991). Bayes estimation of the piece-wise exponential distribution. *IEEE Transactions on Reliability*, **43**, 128-131.

Gamerman, D. (1998). Markov chain Monte Carlo for dynamic generalised linear models. *Biometrika*, **85**, 215-227.

Gamerman, D. and Lopes, H. F. (2006). *Markov Chain Monte Carlo. Stochastic simulation for Bayesian inference* (2nd edition). Chapman and Hall, London.

Gelman, A. (1996). Inference and monitoring convergence, in *Markov Chain Monte Carlo in Practice*, W.R. Gilks et al., eds., Chapman and Hall, London, pp. 131-140.

Godolphin, E. J. and Triantafyllopoulos, K. (2006). Decomposition of time series models in state-space form. *Computational Statistics and Data Analysis*, **50**, 2232-2246.

Grunwald, G. K., Raftery, A. E. and Guttorp, P. (1993).Time series of continuos proportions. *Journal of the Royal Statistical Society* , Ser. B, **55**, 1, pp. 103-116.

Harvey, A.C. (1989). *Forecasting, structural Time Series Models and the Kalman Filter*, University Press, Cambridge.

Harvey, A. C. and Fernandes, C. (1989). Time series models for count or qualitative observations. *Journal of Business and Economic Statistics*, **7**, 4, pp. 407-417.

Hemming, K. and Shaw, J. E. H. (2002). A parametric dynamic survival model applied to breast cancer survival times. *Applied Statistics*, **51**, 421-435.

Johnson, N.L., Kotz, S. & Balakrishnan, N. (1997). *Continuous Univariate Distrbutions*,vol. 1, Second Edition, John Wiley & Sons.

Jørgensen, B., Lundbye-Christensen, S., Song, P.X.K. and Sun, L. (1999). A state space models for multivariate longitudinal count data. *Biometrika*, **86**, 169-181.

Lindsey, J. K. and Lambert, P. (1995). Dynamic generalized linear models and repeated measurements. *Journal of Statistical Planning and Inference*, **47**, 129-139.

Kaufmann, H. (1987). Regression models for nonstationary categorical time series: asymptotic estimation theory. *Annals of Statistics*, 15, 79-98.

Kitagawa, G. (1987). Non-Gaussian state-space modelling of nonstationary time series. *Journal of the American Statistical Association*, **82**, 1032-1063.

Migon, H. S. & Gamerman, D. (1999). *Statistical Inference: an Integrated Approach*. London: Arnold.

Nelder. J.A. and Wedderburn, R.W.M. (1972). Generalized linear models. *Journal of the Royal Statistical Society A*, 135, 370-384.

R Development Core Team. (2005). *R: A language and environment for statistical computing*, R Foundation for Statistical Computing, Vienna.

Ross, S. (2002).*A First Course in Probability*, 6th ed., Prentice Hall.

Sallas, W.M. and Harville, D.A. (1988). Noninformative priors and restrited maximum likelihood estimation in the kalman filter. In: James C. Spall. Bayesian Analysis of Time Series and Dynamic Models. New York, Marcel Dekker Inc.

Shephard, N. and Pitt, M. K. (1997). Likelihood analysis of non-Gaussian measurement time series. *Biometrika*, **84**, 653-667.

Smith, R. L. and Miller, J. E. (1986). A non-gaussian state space model and application to prediction of records. *Journal of the Royal Statistical Society*, Ser. B, **48**, 79-88.

Smith, J. Q. (1979). A generalization of the Bayesian steady forecasting model. *Journal of the Royal Statistical Society*, Ser. B, **41**, 375-387.

Smith, J. Q. (1981). The multiparameter steady model. *Journal of the Royal Statistical Society*, Ser. B, **43**, 256-260.

Spall, J. C. (1988). *Bayesian Analysis of Time Series and Dynamic Models.* New York, Marcel Dekker Inc.

Spiegelhalter, D. J., Best, N. G., Carlin, B.P. and van der Linde, A. (2002). Bayesian measures of model complexity and fit (with discussion and rejoinder). *Journal of the Royal Statistical Society*, Series B, **64**, 583-639.

West, M., Harrison, P. J. and Migon, H. S.(1985). Dynamic generalized linear models and Bayesian forecasting (with discussion). *Journal of the American Statistical Association*, **81**, 741-750.

West, M. and Harrison, J. (1997). *Bayesian Forecasting and Dynamic Models*, Springer, New York.