

Research Article

Use of Non-Homogeneous Poisson Processes in the Modeling of Hospital over Admissions in Ribeirao Preto and Region, Brazil: An Application to Respiratory Diseases

Jorge Alberto Achcar¹, Edilberto Cepeda-Cuervo² and Edson Zangiacomi Martinez^{1*}

¹Universidade de Sao Paulo, FMRP Ribeirao Preto, SP, Brazil

²Universidad Nacional de Colombia-Bogotá-Colombia

Abstract

The daily number of hospital admission due to respiratory diseases can have a great variability. This variability could be explained by different factors as year seasons, temperature, pollution levels among many others [1,2]. In this paper, we have been using non-homogeneous Poisson processes with different intensity functions under the Bayesian paradigm and using standard existing MCMC (Markov Chain Monte Carlo) methods to simulate samples for the joint posterior distribution of interest. An application is given considering the daily number of hospital admission in Ribeirao Preto, Brazil in the period ranging from January 01, 1998 to December 31, 2007. The proposed model showed a good fit for the seasonality of the disease with simple interpretation in the framework of epidemiology.

Keywords: Hospital admissions; Respiratory diseases; Non-homogeneous Poisson process; Bayesian inference; MCMC methods

Introduction

The daily counting of hospital admissions, has large variations due to several factors: year season, climatic variation, variation in levels of different pollutants, among many others. A great interest to public health administrators is related to the modeling of these daily counts, especially for cases of over admissions, which can result in many problems in the hospitals such as lack of beds and equipment, or lack of medicines and lack of health professionals. Counting of hospital admissions due to different causes has been modeled in different ways [3-7].

Among the various diseases that lead to hospitalizations, one stands out among all others: the respiratory diseases. According to data obtained in DATASUS (a Brazilian health data center), respiratory diseases, classified in Chapter X of the Tenth International Classification of Diseases (ICD-10) of the World Health organization (WHO), were the sixth leading cause of hospital admissions in Ribeirao Preto, in hospitals of the National Health System (SUS) in 2008.

The data for this study were provided by the Data Processing Center Hospital (PCHR, 2010) of the Department of Social Medicine, School of Medicine of Ribeirao Preto, University of Sao Paulo. We analyzed the period ranging from January 01, 1998 to December 31, 2007. This database was composed of variables characterizing the patient's sex, age, occupation and town of residence, and variables characterizing the hospitalization as the time of entry and exit, exit condition, the main city of the diagnostic of the disease and where the patient was hospitalized.

A total of 25 municipalities were considered in this data set consisting of approximately 80,967 hospitalizations due to respiratory diseases with an average daily admissions equals to 22.4 and median equals to 22. We also observed that 75% of the observations had a value up to 28 admissions a day. Therefore we consider this point as the threshold for over admissions, that is, if there are 28 hospital admissions or more in a day, this will be considered an over admission. In figure 1 we observe the distribution of all daily admissions between January 01,1998 to November 31, 2007. In figure 2 we have the plot of the accumulated number of hospitalizations due respiratory diseases against time (days) when this occurs.

This paper has as main goal the counting modeling of hospital over admissions due to respiratory diseases using non-homogeneous Poisson processes with different intensity functions. Different intensity functions are considered and inferences for the proposed models are obtained under the Bayesian paradigm and using standard MCMC (Markov Chain Monte Carlo) methods.

The paper is organized as follows: in section 2 we introduce the modeling of hospital admissions; in section 3 the analysis of the respiratory data of Riberao Preto, and finally, in section 4, some conclusions.

Model and Methods

To model the number of times that hospital over admissions occur in Ribeirao Preto City, we consider a point process to count these violations. Let $N = \{N_t : t \in [0,T]\}$ be the process that registers the cumulative number of daily over admissions that are observed during the interval (0,T), i.e., for each $t \in [0,T]$, N_t is the number of hospital admission peaks that are observed during the time interval (0, t). Assume that N is modelled by a non-homogeneous Poisson process (NHPP) with intensity function,

$$\lambda(t) = \frac{d}{dt}m(t) = \frac{d}{dt}E(N_t)$$
(1)

where m(t) is the mean value function.

*Corresponding author: Edson Zangiacomi Martinez, Universidade de Sao Paulo, FMRP Ribeirao Preto, SP, Brazil, Email: edson@fmrp.usp.br

Received November 03, 2011; Accepted November 29, 2011; Published December 01, 2011

Citation: Achcar JA, Cepeda-Cuervo E, Martinez EZ (2012) Use of Non-Homogeneous Poisson Processes in the Modeling of Hospital over Admissions in Ribeirao Preto and Region, Brazil: An Application to Respiratory Diseases. J Biomet Biostat 3:131. doi:10.4172/2155-6180.1000131

Copyright: © 2012 Achcar JA. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

From the plot of figure 1, we observe a cyclic and small decreasing behavior in the number of daily hospital admissions during the observed period. Therefore, as first modeling for the counting data, it is interesting to have an intensity function $\lambda(t)$, $t \ge 0$, that is a monotonic decreasing function of *t*.

Different parametrical forms for monotonic intensity functions are very popular within the framework of software reliability studies [8,9,10]. A popular form for the monotonic intensity functions is given by power law processes (PLP) or the Weibull intensity function.

Intensity functions assumed to model hospital over admissions

Let $N^{(\Theta)} = \{N_t^{(\Theta)} : t \in [0, T]\}$ be a non-homogeneous Poisson process with mean value function $m(t|\theta)$ where θ is a vector of parameters. The function $m(t|\theta)$ represents the expected number of events registered by $N^{(\Theta)}$ up to time t (the hospital over admissions when the threshold is 28 admissions).

Equivalently, the process can be specified by its intensity function $\lambda(t \mid \Theta) = \frac{d}{m} m(t \mid \Theta)$ (2)

$$\frac{dt}{dt} = \frac{dt}{dt}$$

In this paper we explore some especial cases of NHPP to analyse





J Biomet Biostat

the hospital admissions data due to respiratory diseases of Ribeirao Preto: the power law (PLP); the Musa-Okumoto (MOP) [9]; the Goel-Okumoto (GOP) [11] and a generalized form of a Goel-Okumoto (GGOP) processes defined, respectively by the following mean value functions,

$$\begin{split} m^{(PLP)}\left(t \mid \theta\right) &= (t \mid \beta)^{\alpha}, \beta, \alpha > 0 \\ m^{(MOP)}\left(t \mid \theta\right) &= \beta \log(1 + t/\alpha), \beta, \alpha > 0 \\ m^{(GOP)}\left(t \mid \theta\right) &= \alpha(1 - \exp(-\beta t)), \beta, \alpha > 0 \\ m^{(GGOP)}\left(t \mid \theta\right) &= \alpha(1 - \exp(-\beta t^{\gamma})), \beta, \alpha, \gamma > 0 \end{split}$$
(3)

where $\theta = (\alpha, \beta)$ for the PLP, MOP, and GOP models and $\theta = (\alpha, \beta, \gamma)$ for the GGOP model. The intensity functions associated with those processes are given, respectively, by

$$\lambda^{(PLP)} (t \mid \theta) = (\alpha/\beta)(t \mid \beta)^{\alpha \cdot 1}, \beta, \alpha > 0$$

$$\lambda^{(MOP)} (t \mid \theta) = \beta/(t+\alpha), \beta, \alpha > 0$$

$$\lambda^{(GOP)} (t \mid \theta) = \alpha\beta \exp(-\beta t)), \beta, \alpha > 0$$

$$\lambda^{(GGOP)} (t \mid \theta) = \alpha\beta\gamma t^{\gamma \cdot 1} (1 - \exp(-\beta t^{\gamma})), \beta, \alpha, \gamma > 0$$

(4)

We observe that:

- 1. The intensity function given by (2) defines the hazard rate of the time between occurrence of events in the respective models.
- 2. We observe from (4) that the intensity function $\lambda^{(PLP)}(t|\Theta)$ gives a different behavior for the PLP depending on the value of α . We have that as a function of time that intensity function can be constant, decreasing or increasing depending on whether $\alpha = 1$, $\alpha < 1$ or $\alpha > 1$, respectively. The intensities $\lambda^{(MOP)}(t|\Theta)$ and $\lambda^{(GOP)}(t|\Theta)$ present a decreasing behavior as functions of t; and $\lambda^{(GGOP)}(t|\Theta)$ describes the situation where the intensity increases slightly at the beginning and then begins to decrease with t.

The intensity functions given by (4) also were considered by Achcar et al. [22,23,24] to analyze ozone pollution data of Mexico City in presence or not of one or more change points.

Since we have a cyclic behavior for the hospital admission counting (Figure 1), we propose new intensity functions adding a cyclic term for the intensity functions (4). This additional term is given by

$$a(t) = \delta \cos(\theta t) \tag{5}$$

As a special case, a new MOP intensity function is given by

$\lambda_{I}^{MOP}(t) = \frac{\beta}{t+\alpha} - \delta \sin(\theta t)$				
Model	Parameter	Mean	S. D.	95% Credible interval
	α	0.9232	0.03081	(0.867, 0.9848)
PLP	β	2.244	0.5621	(1.311, 3.504)
MOP	α	6964.0	1511.0	(4185.0,9657.0)
	β	2238.0	399.1	(1505.0, 2921.0)
GOP	α	2818.0	719.0	(1943.0, 4624.0)
	β	0.0001205	0.0000298	(0.000064, 0.000182)
GGOP	α	2229.0	655.6	(1502.0, 4248.0)
	β	0.000104	0.000028	(0.000061, 0.00017)
	γ	1.058	0.05368	(0.9554, 1.159)
MOP with	α	4421.0	1071.0	(2862.0,7115.0)
cyclic factor	β	1580.0	288.7	(1156.0, 2298.0)
	δ	7.694	0.6225	(6.455, 8.840)
	θ	0.01782	0.000031	(0.0178, 0.01789)

Tabel 1: Posterior sumaries.

Page 2 of 5



The corresponding mean value function associated to the intensity function (6) is given by

$$m_{L}^{MOP}(t) = \beta \log(1 + t/\alpha) + \delta \cos(\theta t)$$
⁽⁷⁾

In this model, δ is interpreted as the amplitude of oscillation and $1/\theta$ as the period of the oscillation. In this model, the positivity of the $\lambda_i^{MOP}(t)$ needed to be guaranteed with appropriate selection of the prior distribution to have $\beta/(T+\alpha) > \delta\theta$. Other choices for the cyclic term can be made, as for example, with the *sine* function. These choices can be explored depending on the nature of the data set.

The likelihood function

Denoting the data set by $D_T = \{n; t_1, t_2,...,t_n; T\}$ where *n* is the number of observed occurrence times which are such that $0 < t_1 < t_2 < ... < t_n < T$, and where these values are the epochs of the occurrence of hospital over admissions up to time *T*, the likelihood function for θ considering the time truncated model is [8] given by,

$$L(\theta|D_t) = [\prod_{i=1}^n \lambda(t_i | \theta)] \exp[-m(T | \theta)]$$
(8)

We get inferences for the parameters of the model under a Bayesian approach. Bayesian inferences for NHPP has been discussed by many authors [12,13,14].

Posterior summaries of interest are obtained using standard MCMC (Markov Chain Monte Carlo) methods [15,16,17].

Analysis of the Hospital Over Admissions due to Respiratory Diseases

In this section, we analyze the daily hospital admissions count due

to respiratory diseases in Riberao Preto region for the period ranging from January 01, 1998 to December 30, 2007. From this data set, we observe that in 948 days we have al least 28 hospital admissions due to respiratory diseases in the period of T = 3670 days.

To analyze this count data, we first assume a NHPP, considering the intensity functions (4) not in the presence of the cyclic term (5). For each proposed model, we assume the following prior distributions: $\alpha \sim U(0,2)$, $\beta \sim U(0,100)$ for the PLP model; $\alpha \sim U(0,1000)$, $\beta \sim U(0,3000)$ for the MOP model; $\alpha \sim U(1,10000)$, $\beta \sim U(0,1)$, for the GOP model, and $\alpha \sim U(0,5000)$, $\beta \sim U(0,0.001)$ and $\gamma \sim U(0.5,1.5)$ for the GGOP model, (*a*, *b*). We further assume prior independence among the parameters.

Observe that we are using approximately non-informative priors for the parameters of each model. Considering these priors, we obtained a proper joint posterior distribution of interest considering our data set (we have used standard uniform priors limited to a large finite interval) and in our case, we have an informative likelihood function. It is important to point out that identifiability issues and property of the posterior distribution of interest has been discussed by many authors, especially within the analysis of hierarchical models considering improper Jeffreys or Berger and Bernardo's reference priors for the parameters of the model [18,19,20]. We also observed that with small modifications in the hyperparameter values of the prior distributions, we obtained similar posterior summaries of interest. In some cases, the choice of the hyperparameter values was made to have convergence of the Gibbs Sampling algorithm using the WinBugs software [21], possibly a numerical limitation of the software.

Page 3 of 5

For each proposed model, we simulated 3000 initial Gibbs Samples considered as "burn-in-sample" to eliminate the effects of the initial values; after this "burn-in-sample" period we simulated another 30,000 Gibbs samples taking every 30th sample, to have approximately uncorrelated samples, which totalizes 1000 Gibbs samples.

In this simulation procedure, we have used the WinBugs software [21]. Convergence of the Gibbs Sampler algorithm was monitored using standard existing methods as the traceplots of the simulated samples.

In Table 1, we have the posterior means, posterior standard deviations and 95% credible intervals for all parameters.

In Figure 3 we have the plots of the empirical accumulated counts and the fitted mean value functions assuming each proposed model (panels a, b, c and d) versus time for occurrence of over admissions. From these plots, we observe that the MOP model gives better fit, but this could be improved considering the cyclic factor (6) in the MOP intensity function (intensity function 6).

In this way, we assume as fifth model, the MOP model in presence of the cyclic factor (5) and assuming the following priors: $\alpha \sim U(2000,10000), \beta \sim U(1000,3000), \delta \sim U(3,10), \beta \sim U(0.01,0.5)$. We also assume prior independence between the parameters. In this case, we have used some information of the previous Bayesian analysis not considering the presence of the cyclic factor (5) to choose the hyperparameter values for the prior distributions.

From the WinBugs output considering a "burn-in-sample" of size 5000 and 20000 simulated samples taking every 20th samples we also

have in Table 1, the posterior summaries of interest for the MOP with intensity function (6). The convergence was monitored using traceplots of the Gibbs samples simulated for each parameter (Figure 4).

In Figure 5, we have the plots of the empirical accumulated counts and the fitted mean values assuming the MOP model in the presence of a cyclic factor (6) versus time of occurrence of over admissions. We observe a good fit of this model for the hospital over admissions counting due to respiratory diseases in Ribeirao Preto. In Figure 5, we also have 95% confidence bands for the fitted mean values.

In the MOP model, the intensity function is a positive decreasing function of time. In the proposed model the intensity is positive, but it is not decreasing as it was derived in the application. Figure 6 shows the standard behavior of the fitted intensity obtained in applications. We observe that the oscillation function decreases in each periodic point.

Conclussion

In this paper we introduced new modeling approaches to analyze count data due to hospital over admission. In our case, we considered hospital over admission due to respiratory diseases in Riberao Preto, Brazil, but this class of model could be applied to any disease. The use of non-homogeneous Poisson processes assuming different intensity functions gives a great flexibility of fit for the count data. We also introduced a new model considering the introduction of the cyclic term (6) to capture the seasonality of the disease. This new modeling approach could be used to analyze data from many different seasonal diseases.







The use of Bayesian methods considering standard existing MCMC simulation methods to generate samples for the joint posterior distribution of interest, especially using the WinBugs software gives a great simplification to get accurate inference results and accurate predictions.

These results could be of great interest in epidemiology and medical research.

References

- Schwartz J (1996) Air pollution and hospital admissions for respiratory disease. Epidemiology 7: 20-28.
- Panagiotakos DB, Chrysohoou C, Pitsavos C, Nastos P, Anadiotis A, et al. (2004) Climatological variations in daily hospital admissions for acute coronary syndromes. International Journal of Cardiology 94: 229-233.
- Muller-Pebody B, Crowcroft NS, Zambon MC, Edmunds WJ (2006) Modeling hospital admissions for lower respiratory tract infections in the elderly in England. Epidemiology and Infection 134: 1150-1157.
- Morina D, Puig P, Rios J, Vilella A, Trilla A (2011) A statistical model for hospital admissions caused by seasonal diseases. Statistics in Medicine 30: 3125-3136.
- Burnett R, Bartlett S, Krewski D, Robert G, Raad-Young M (1994) Air pollution effects on hospital admissions: a statistical analysis of parallel time series. Environmental and Ecological Statistics 1: 325-332.

 Griffiths P, Murrells T, Dawoud D, Jones S (2010) Hospital admissions for asthma, diabetes, and COPD: is there an association with practice nurse staffing? A cross sectional study using routinely collected data. BMC Health Services Research 10: 276-287.

Page 5 of 5

- Holgate ST, Samet JM, Koren HS, Maynard RL (1999) Air Pollution and Health. Academic Press, Boston, USA.
- 8. Cox DR, Lewis PA (1966) Statistical Analysis of Series of Events, Methuen, UK.
- Musa JD, Okumoto K (1984) A logarithmic Poisson execution time model for software reliability measurement. Proceedings of Seventh International Conference on Software Engineering, Orlando, USA, pp: 230-238.
- Musa JD, Iannino A, Okumoto K (1987) Software reliability: measurement, prediction, application. McGraw Hill, New York, USA.
- Goel AL, Okumoto K (1978) An analysis of recurrent software failures on a real-time control system. Proceedings of ACM Conference, Washington, DC, USA, pp: 496-500.
- Kuo L, Yang TY (1996) Bayesian computation for nonhomogeneous Poisson process in software reliability. Journal of the American Statistical Association 91: 763-773.
- Pievatolo A, Ruggeri F (2004) Bayesian reliability analysis of complex repairable systems. Applied Stochastic Models in Business and Industry 20: 253-264.
- Ramirez-Cid JE, Achcar JA (1999) Bayesian inference for non-homogeneous Poisson process in software reliability models. Computational Statistics and Data Analysis 32: 147-159.
- Chib S, Greenberg E (1995) Understanding the Metropolis-Hastings algorithm. The American Statistician 49: 327-335.
- Gelfand AE, Smith AFM (1990) Sampling-based approaches to calculating marginal densities. Journal of the American Statistical Association 85: 398-409.
- Smith AFM, Roberts GO (1993) Bayesian computation via the Gibbs Sampler and related Markov Chain Monte Carlo methods. Journal of the Royal Statistical Society Series B 55: 3-23.
- Berger JO, Strawderman WE (1996) Choice of hierarchical priors: admissibility in estimation of normal means. Ann Statist 24: 931-951.
- Hobert JP, Casella G (1996) The effect of improper priors on Gibbs sampling in hierarchical linear mixed models. Journal of the American Statistical Association 91: 1461-1473.
- Gelfand A, Sahu S (1999) Identifiability, Improper Priors, and Gibbs Sampling for Generalized Linear Models. Journal of the American Statistical Association 94: 247–253.
- Spiegelhalter DJ, Thomas A, Best NG (2003) WinBugs: Bayesian inference using Gibbs sampling, MRC Biostatistics Unit, Cambridge, UK.
- Achcar JA, Fernandez-Bremauntz AA, Rodrigues ER, Tzintzun G (2008) Estimating the number of ozone peaks in Mexico City using a non-homogeneous Poisson model. Environmetrics 19: 469-485.
- Achcar JA, Rodrigues ER, Paulino CD, Soares P (2010) Non-homogeneous Poisson models with a change-point: an application to ozone peaks in Mexico city. Environmental and Ecological Statistics 17: 521-541.
- 24. Achcar JA, Rodrigues ER, Tzintzun G (2011) Using non-homogeneous Poisson models with multiple chance-points to estimate the number of ozone descendance in Mexico city. Environmetrics 22: 1-12.