The Application Of Bayesian Change Point Detection On Homogeneous Poisson Process.

CERTARA Simcyp

Anthonia Afuape[1] and Golnaz Shahtahmassebi[2]

[1]Certara UK Limited, Simcyp Division, Level 2-Acero, 1 Concourse Way, Sheffield, S1 2BJ, United Kingdom

[2] School of Science and Technology, Nottingham Trent University, Clifton, NG11 8NS

NOTTINGHAM TRENT UNIVERSITY

Abstract

Observations with temporal effects such as the number of survivals within a time-frame arise in repeated count data, number of signalling molecules in a cellular environment or annual disasters at a coal mining follow a Poisson process[1,2,3]. The objective of this study is to present the Bayesian estimation approach of parameters of Univariate Poisson change-point processes. For this, we defined a class of Prior distributions and used it to obtain the Joint Posterior distribution using Markov Chain Monte Carlo (MCMC) sampling methods as discussed by Maria Rizzo (2007)[2]. All analysis were performed in R 3.3.4.

Background

A change-point (CP) splits a sequence of data over time into segments. One of the assumptions is that a dataset from the same segment comes from the same population (model). Let Y_i be an independent and identically distributed (IID) random variable (RV), representing the number of observations at times, $i = 1, \dots, n$. A simple form of a CP model can be described as shown in equation (1) below. k is the unknown parameter called the CP. When $k = n$ or $k = 1$, the model is interpreted to have no change.

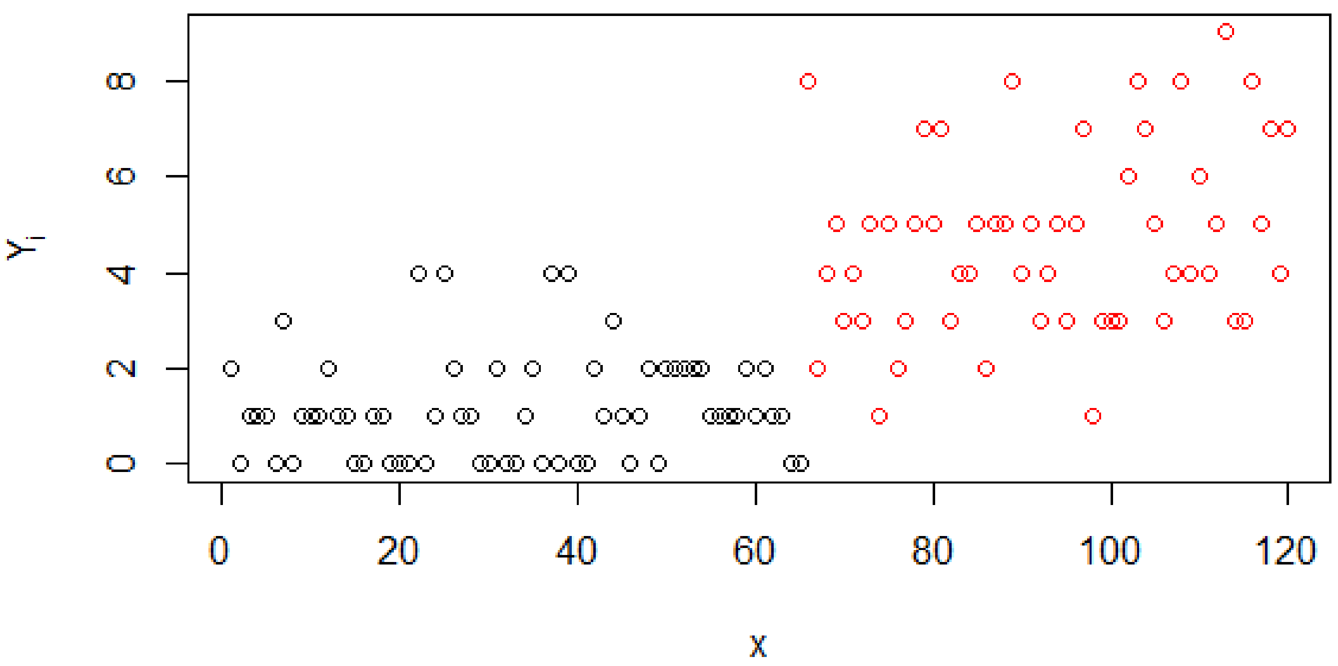
$$Y_i \sim Poisson(\lambda_1), \quad i = 1,\dots,k \qquad (1)$$

$$Y_i \sim Poisson(\lambda_2), \quad i = k+1,\dots,n$$

Methods

For Y_i we generated 120 random samples (Figure 1) from a system of two Poisson processes with $\lambda_1 = 1$ and $\lambda_2 = 5$. k was sampled randomly from a uniform distribution ($k = 65$). To estimate k, we could simply assume it lies between two time points i.e. k lies between x_{t-1} and x_t. However, this method is not as efficient as using CP detection models which are better able at isolating k and able to estimate λ_1 and λ_2.

Figure 1. 120 simulated samples from CP Poisson model.



Bayesian inference

Bayesian Inference is used to estimate the parameters of interest (k, λ_1 and λ_2) of the current model for CP. It is assumed the Joint Posterior distribution (2) is proportional to the product of the known Prior distribution and the Likelihood function:

$$P(\lambda_1,\lambda_2,k \mid y) \propto P(\lambda_1,\lambda_2,k)\, P(y \mid \lambda_1,\lambda_2,k) \qquad (2)$$

Here $P(y \mid \lambda_1,\lambda_2,k)$ is the likelihood function of n observations denoted by $L(Y_i)$ and

$$L(Y_i) \propto \prod_{i=1}^{k} e^{-\lambda_1} \lambda_1^{y_i} \prod_{i=k+1}^{n} e^{-\lambda_2} \lambda_2^{y_i} \qquad (3)$$

$$= e^{-k\lambda_1} \lambda_1^{\sum_{i=1}^{k} y_i} e^{-\lambda_2(n-k)} \lambda_2^{\sum_{i=k+1}^{n} y_i}$$

The following are the choice of prior distribution used:

$$\lambda_1 \sim Gamma\,(\alpha_1, \beta_1),$$
$$\lambda_2 \sim Gamma\,(\alpha_2, \beta_2),$$
$$k \sim Uniform\,[1, \dots, n]$$

Gibbs Sampler

Gibbs sampler (GS) is a special case of Metropolis-Hastings sampler that is used to generate RVs from marginal distributions. With this technique we are able to obtain characteristics such as mean or variance of marginal densities indirectly without having to calculate the actual posterior density. This is done by carrying out large enough iterations till the estimates converge[3]. Therefore, to estimate k, λ_1 and λ_2, we ran 10,000 iterations of GS algorithm using the following conditional distributions[2].

$$f(\lambda_1 | y, \lambda_2, k) \sim Gamma\left(\alpha_1 + \sum_{i=1}^{k} y_i\,, k + \beta_1\right),$$

$$f(\lambda_2 | y, \lambda_1, k) \sim Gamma\left(\alpha_2 + \sum_{i=k+1}^{n} y_i\,, n - k + \beta_2\right),$$

$$f(k \mid y, \lambda_1, \lambda_2) = \frac{L(y; k, \lambda_2, \lambda_1)}{\sum_{j=1}^{n} L(y; j, \lambda_2, \lambda_1)},$$

where,

$$L(y; k, \lambda_2, \lambda_1) = e^{k(\lambda_2 - \lambda_1)} \frac{\lambda_1}{\lambda_2}^{\sum_{i=1}^{k} y_i}.$$

Results

After discarding 5,000 burn-in samples, the mean estimates were calculated. The estimates suggests Y_i follows a CP model with $\widehat{\lambda_1} = 1.096$, $\widehat{\lambda_2}$ =5.177 and $\hat{k} = 64.909$, which are very close to the observed mean values. This suggests provided the choice of prior is accurate, the algorithm will perform effectively and accurately.

Table 1. Summary estimates of $\widehat{\lambda_1}, \widehat{\lambda_2}$ and $\hat{k}$ after 5,000 burn-in GS iterations.

Parameter	Mean	SD	2.5%	97.5%	Min	Max
$\widehat{\lambda_1}$	1.096	0.128	1.007	1.178	0.705	1.649
$\widehat{\lambda_2}$	5.177	0.305	4.971	5.373	4.133	6.420
$\hat{k}$	64.909	0.325	65.000	65.000	60.000	66.000

Conclusion

We performed our analysis on 10 CP models with different parameter values. We found despite the differences in the model parameters, GS algorithm still performed effectively and accurately. However, we noticed that when the change-point k is close to the upper and lower bounds of n, the algorithm struggled and sometimes failed. This can be explained by the "Random walk" effects which the algorithm is built on. At each iteration, a sample is generated depending on the sample form the previous step. So, if the sample for k is close to n, then the next estimate could fall on or beyond the maximum value of n causing the iteration to terminate. The Bayesian CP detection of Univariate Poisson processes has be extended to estimate parameter of Bivariate Poisson CP models[4]. Also, there is evidence of other estimation methods such as the stochastic version of EM algorithm (SAEM)[5] used to estimate the parameters of posterior distribution of change-point problems.

References

[1] EL Plan, CPT Pharmacometrics Syst. Pharmacol. (2014) 3, e129; doi:10.1038/psp.2014.27

[2] Maria L. Rizzo, 2007. Markov Chain Monte Carlo Methods. In: Statistical Computing With R. *New York: Chapman and Hall/CRC*, pp 245-279.

[3] West, W. R. and T. R. Ogden (1997). Continuous-time estimation of a change-point in a Poisson process. *Journal of Statistical Computation and Simulation* 56 (4), 293-302.

[4] Karlis, D. and P. Tsiamyrtzis (2008). Exact bayesian modeling for bivariate poisson data and extensions. Statistics and Computing 18 (1), 27-40.

[5] Lavielle, Marc. (2005). Using penalized contrasts for the change-point problem. Signal Processing. DOI: 10.1016/j.sigpro.2005.01.012

The Royal Statistical Society Conference 2018 , Cardiff City Hall. 23A