Parameter Estimation and Multiple Changepoint Detection for Count Data

Şirag Erkol, Berk Görgülü, Mehmetcan Tütüncü

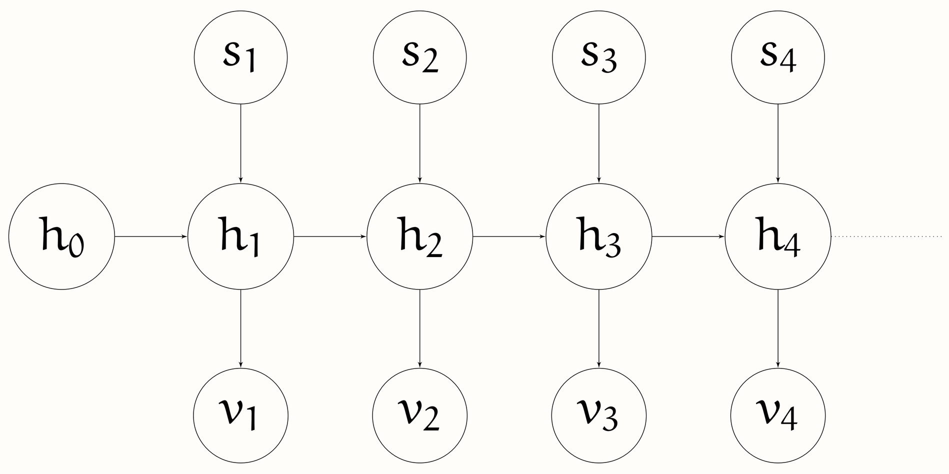
Boğaziçi University, Istanbul

Introduction

Changepoint models are common tools to determine the time points where mean shifts occur. In this project our aim is to discover multiple changepoints and finding posterior probability of change with associated intensity levels in a given count data without specified parameters using a switching state space model in a Hidden Markov Model form.

Since our data is discrete non-negative count data and we are aiming to detect changepoints on this data, we use Gamma-Poisson model in this approach and assuming that observation v_t is a Poisson random variable with intensity parameter h_t.

Model



Generative Model

$$h_0 \sim \Omega(h_0; w)$$
$$s_t \sim BE(s_t; \pi)$$
$$h_t | s_t, h_{t-1} \sim [s_t = 0]\delta(h_t - h_{t-1}) + [s_t = 1]\Omega(h_t; w)$$
$$v_t | h_t \sim \Theta(v_t; h_t)$$

Gamma-Poisson Reset Model

$$\Omega(h; w) = G(h; a, b) = \frac{b^a}{\Gamma(a)} h^{a-1} e^{-bh}$$
$$\Theta(v; h) = P(v; h) = \frac{h^v e^{-h}}{\Gamma(v+1)}$$

Approach

Our aim in this model requires two step application.

1. Parameter estimation step where we utilize Expectation Maximization(EM) algorithm to estimate hyper-parameters of the underlying distribution.
2. Inference step where we determine posterior changepoint probabilities and associated intensity levels using Forward-Backward algorithm.

Parameter Estimation(EM)

$$L(\theta) = \log\, p(s_{1:T}, h_{0:T}, v_{1:T} | \theta)$$
$$= \log\, \Omega(h_0; w) + \sum_{t=1}^{T} \Big[[s_t = 0](\log(1-\pi)) + \log\theta(h_t - h_{t-1})$$
$$+ [s_t = 1](\log\pi + \log\Omega(h_t; w) + \log\Gamma(v_t; h_t)) \Big]$$

E-Step

$$E_{p(s_{1:T}, h_{0:T} | v_{1:T}, \theta^{\tau})}[L(\theta)] =$$
$$\sum_{t=0}^{T} E_{p(s_{1:T}, h_{0:T} | v_{1:T}, \theta^{\tau})} \Big[[s_t = 0](\log(1-\pi)) \Big]$$
$$+ \sum_{t=0}^{T} E_{p(s_{1:T}, h_{0:T} | v_{1:T}, \theta^{\tau})} \Big[[s_t = 1](\log\pi + \log\Omega(h_t; w)) \Big]$$
$$+ \sum_{t=0}^{T} E_{p(s_{1:T}, h_{0:T} | v_{1:T}, \theta^{\tau})} \Big[\log\Theta(v_t; h_t) \Big]$$
$$= \log(1-\pi) \sum_{t=0}^{T} p(s_t = 0 | v_{1:T}, \theta^{(\tau)})$$
$$+ \log\pi \sum_{t=0}^{T} p(s_t = 1 | v_{1:T}, \theta^{(\tau)})$$
$$+ \sum_{t=0}^{T} p(s_t = 1 | v_{1:T}, \theta^{(\tau)}) E_{p(h_t | v_{1:T}, \theta^{(\tau)})} \Big[\log\Omega(h_t; w) \Big]$$
$$+ \sum_{t=1}^{T} E_{p(s_{1:T}, h_{0:T} | v_{1:T}, \theta^{\tau})} \Big[\log\Theta(v_t; h_t) \Big]$$

M-Step

$$0 = \frac{\partial E_{p(s_{1:T}, h_{0:T} | v_{1:T}, \theta^{(\tau)})}}{\partial w}$$
$$= \sum_{t=0}^{T} p(s_t = 1 | v_{1:T}, \theta^{(\tau)}) \frac{\partial}{\partial w} E_{p(h_t | v_{1:T}, \theta^{(\tau)})}$$

Update Equations for GP Model

Updating π

$$0 = \frac{\partial E_{p(s_{1:T}, h_{0:T} | v_{1:T}, \theta^{(\tau)})}}{\partial \pi}$$
$$= \frac{1}{\pi} \sum_{t=0}^{T} p(s_t = 1 | v_{1:T}, \theta^{(\tau)}) - \frac{1}{1-\pi} \sum_{t=0}^{T} p(s_t = 0 | v_{1:T}, \theta^{(\tau)})$$
$$\pi = \frac{1}{T} \sum_{t=0}^{T} p(s_t = 1 | v_{1:T}, \theta^{(\tau)})$$

Updating a

$$0 = \frac{\partial E_{p(s_{1:T}, h_{0:T} | v_{1:T}, \theta^{(\tau)})}}{\partial a}$$
$$\psi(a) = \frac{\sum_{t=0}^{T} p(s_t = 1 | v_{1:T}, \theta^{(\tau)})(\log b + E_{p(h_t | v_{1:T}, \theta^{(\tau)})}[\log h}{\sum_{t=0}^{T} p(s_t = 1 | v_{1:T}, \theta^{(\tau)})}$$
$$= \log b + \frac{\sum_{t=0}^{T} p(s_t = 1 | v_{1:T}, \theta^{(\tau)}) E_{p(h_t | v_{1:T}, \theta^{(\tau)})}[\log h_t]}{\sum_{t=0}^{T} p(s_t = 1 | v_{1:T}, \theta^{(\tau)})}$$
$$a^{new} = \psi^{-1}(\log b + \frac{\sum_{t=0}^{T} p(s_t = 1 | v_{1:T}, \theta^{(\tau)}) E_{p(h_t | v_{1:T}, \theta^{(\tau)})}[\log}{\sum_{t=0}^{T} p(s_t = 1 | v_{1:T}, \theta^{(\tau)})}$$

Updating b

$$0 = \frac{\partial E_{p(s_{1:T}, h_{0:T} | v_{1:T}, \theta^{(\tau)})}}{\partial b}$$
$$b^{new} = \frac{a \sum_{t=0}^{T} p(s_t = 1 | v_{1:T}, \theta^{(\tau)})}{\sum_{t=0}^{T} p(s_t = 1 | v_{1:T}, \theta^{(\tau)}) E_{p(h_t | v_{1:T}, \theta^{(\tau)})}[h_t]}$$

Inference

Forward-Backward Algorithm

Forward Messages:

$$\alpha_{0|0}(s_0, h_0) = p(h_0)$$
$$t = 1 \ldots T$$
$$\alpha_{t|t-1}(s_t, h_t) = p(s_t, h_t, v_{1:t-1})$$
$$\alpha_{t|t}(s_t, h_t) = p(s_t, h_t, v_{1:t})$$

Backward Messages:

$$\beta_{T|T} = p(v_{t+1:T | s_t, h_t})$$
$$t = M - 1 \ldots 1$$
$$\beta_{t|t+1}(s_t, h_t) = p(v_{t+1:T} | s_t, h_t)$$
$$\beta_{t|t}(s_t, h_t) = p(v_{t:T} | s_t, h_t)$$

Recursions:

$$\alpha_{t|t-1}(s_t, h_t) = \sum_{s_{t-1}} \int dh_{t-1} p(s_t, h_t | s_{t-1}, h_{t-1}) \alpha_{t-1|t-1}(s_{t-1}, h_{t-1})$$

$$\alpha_{t|t}(s_t, h_t) = p(v_t | s_t, h_t) \alpha_{t|t-1}(s_t, h_t)$$
$$\beta_{t|t+1}(s_t, h_t) = p(v_{t+1:M} | s_t, h_t)$$
$$= \sum_{s_{t+1}} \int dh_{t+1} p(s_{t+1}, h_{t+1} | s_t, h_t) \beta_{t+1|t+1}(s_{t+1}, h_{t+1})$$

$$\beta_{t|t}(s_t, h_t) = p(s_t, h_t, v_{t:M})$$
$$p(v_t | s_t, h_t) \beta_{t|t+1}(s_t, h_t)$$

Smoothing

$$p(s_t, h_t | v_{1:T}) \propto p(v_{1:T}, s_t, h_t)$$
$$= p(v_{1:t-1}, s_t, h_t) p(v_{1:T} | s_t, h_t, v_{1:t-1})$$
$$= p(v_{1:t-1}, s_t, h_t) p(v_{1:T} | s_t, h_t)$$
$$= \alpha_{t|t-1} \beta_{t|t}$$

Inference using GP Model

Results

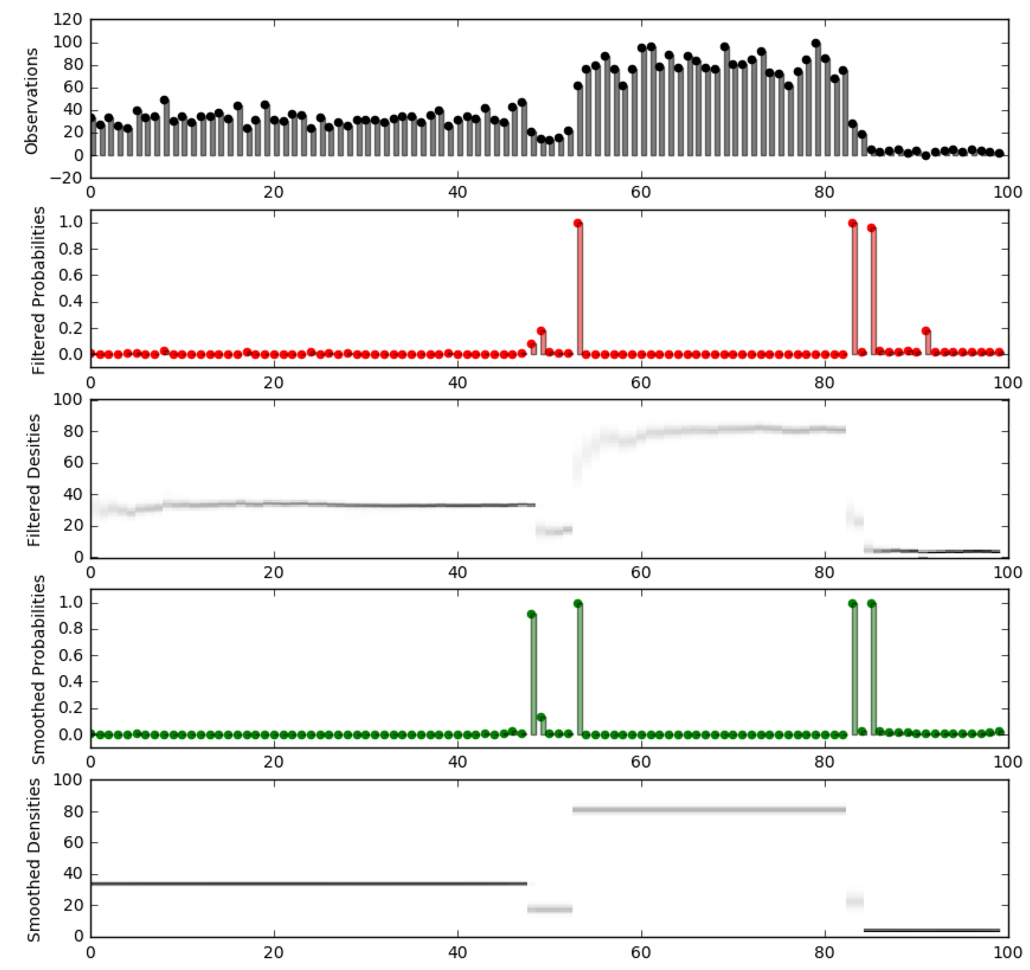


Figure 1: Example output 1

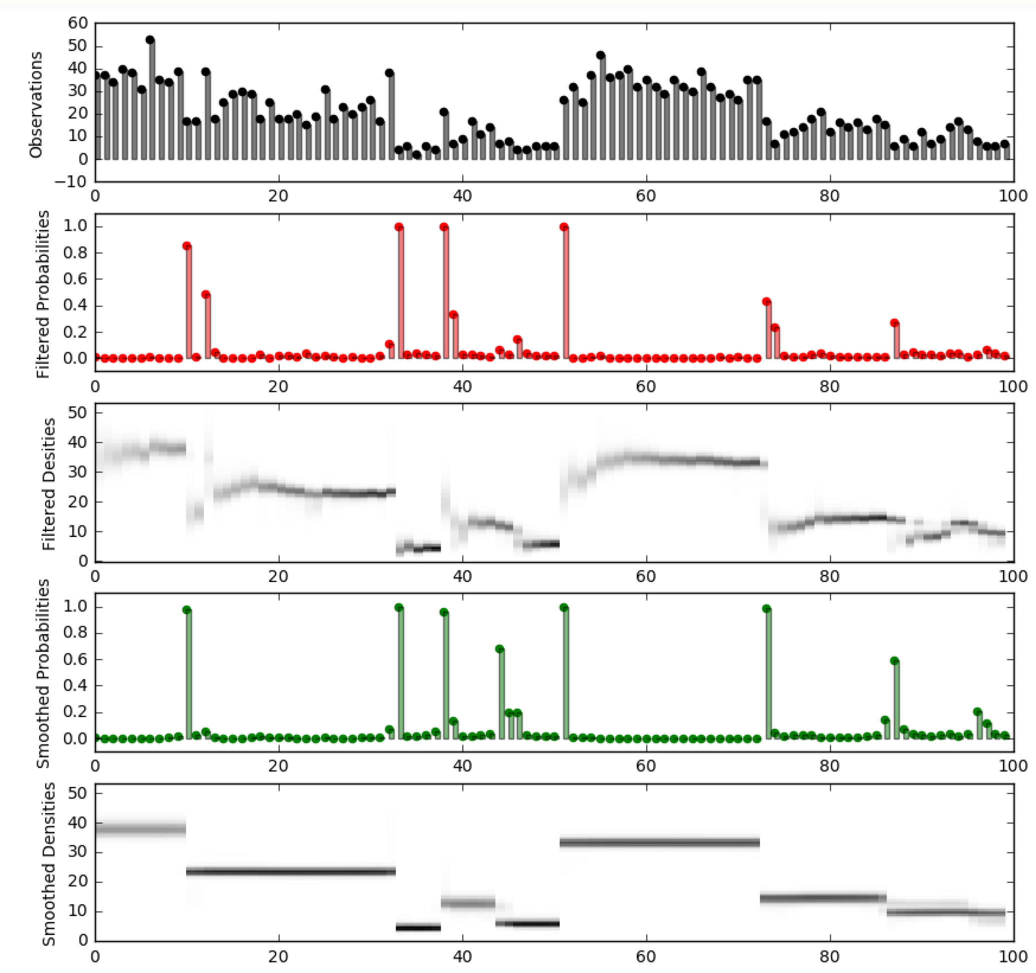


Figure 2: Example output 2

References

[1] A.T.Cemgil, A Multiple Changepoint mode for Count Data, *2009*

[2] D.Barber, Bayesian Reasoning and Machine Learning, *2017*

[3] C.Yıldız Inference and Parameter Estimation In Bayesian Change Point Models *2017*

Bayesian Statistics and Machine Learning, 6 June 2017, Istanbul, TURKEY