Bayesian Changepoint Detection

Mine Ogretir
mine.ogretir@aalto.fi

Onur Poyraz
onur.poyraz@aalto.fi

Aalto University
Department of Computer Science

February 21, 2020

OUTLINE

Aalto University
School of Science

1. Introduction
2. Model Descriptions
3. Analysis and Results
4. Discussion

Outline

Aalto University
School of Science

Motivating Problem

Aalto University
School of Science

- Given a sequential data

We might interest in;

- Changes in the streamed data

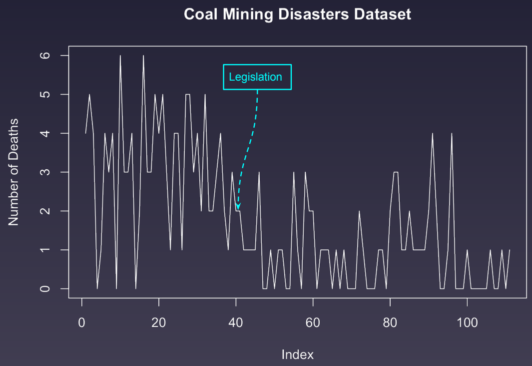
Assume;

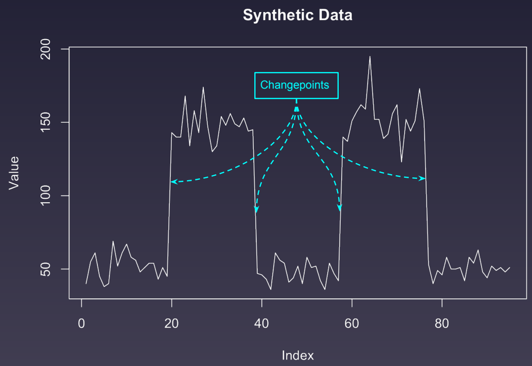
- Data is a result of such a generative process
- Generative process is changing in the changepoints

Goal;

- Find the changes in the parameters of the generative process

Dataset


Coal Mining Disasters Dataset
Legislation
Number of Deaths
Index
0
1
2
3
4
5
6
0
20
40
60
80
100


Synthetic Data
Changepoints
Value
Index
50
100
150
200
0
20
40
60
80

OUTLINE

MODELS

Aalto University
School of Science

Comparable Implemented Models

- Poisson-Gamma Single Changepoint Model (SCM) [2, 3]
- Poisson-Gamma Double Changepoint Model (DCM) [2, 3]

Other Implemented Model

- Bayesian Online Changepoint Detection Algorithm (BOCM)[1]

POISSON-GAMMA CHANGEPOINT DETECTION MODELS

Aalto University
School of Science

Single Changepoint Model

Means of the Intervals

$$e \sim Gamma\,(1,1)$$

$$l \sim Gamma\,(1,1)$$

Changepoint

$$s \sim Uniform(1,T)$$

$$X_t \sim Poisson(\lambda)$$

$$\lambda = \begin{cases} e & \text{if } t < s_1 \\ l & \text{if } s_1 <= t < s_2 \end{cases}$$

Double Changepoint Model

Means of the Intervals

$$e \sim Gamma\,(1,1)$$

$$l \sim Gamma\,(1,1)$$

$$m \sim Gamma\,(1,1)$$

Changepoints

$$s_1 \sim Uniform(1,T)$$

$$s_2 \sim Uniform(1,T)$$

$$X_t \sim Poisson(\lambda)$$

$$\lambda = \begin{cases} e & \text{if } t < s_1 \\ l & \text{if } s_1 <= t < s_2 \\ m & \text{else} \end{cases}$$

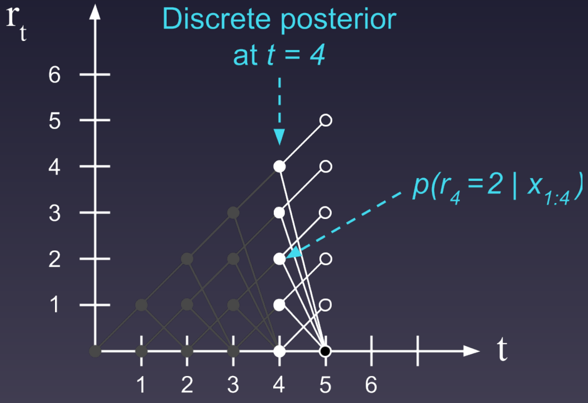


Figure 1: Run Length[1] [4]

Model

$$\alpha_{1:T} \sim Gamma(1, 1)$$
$$X_t \mid r_t \sim Poisson(\alpha_{t-r_t})$$

Changepoint Prior

$$P(r_t | r_{t-1}) = \begin{cases} H(r_{t-1} + 1) & \text{if } r_t = 0 \\ 1 - H(r_{t-1} + 1) & \text{if } r_t = r_{t-1} + 1 \\ 0 & \text{otherwise} \end{cases}$$

$$H(\tau) = 0.01$$

[1] Gregory Gundersen - Tutorial on Bayesian Online Changepoint Detection

OUTLINE

Convergence of Algorithm

Convergence Diagnostics

- All $\hat{R}$ values are 1.
- $n_{eff}/n_{transitions} > 0.001$ for all parameters.
- All of the
  - tree depth
  - E-BFMI
  - divergences

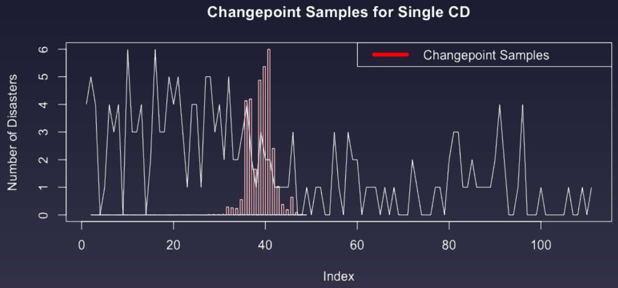
  are looking good.

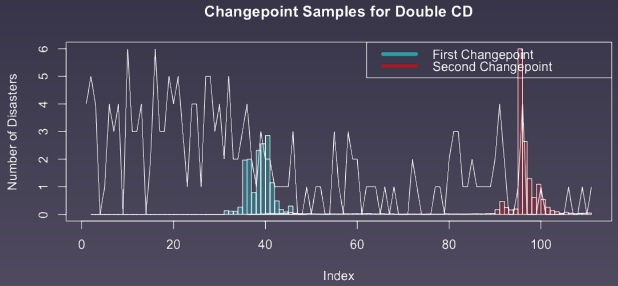
loo comparison	elpd diff	se diff
sigle model	0.0	0.0
double model	-3.2	5.7
hierarchical model	-5.0	6.4

Table 1: Result of loo comparison

Posterior Inference - Changepoint Estimations

Aalto University
School of Science

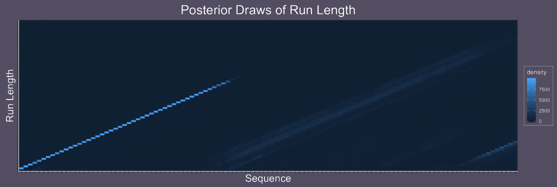




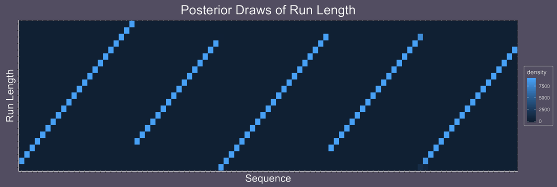
POSTERIOR INFERENCE - RUN LENGTH ESTIMATIONS

Aalto University
School of Science

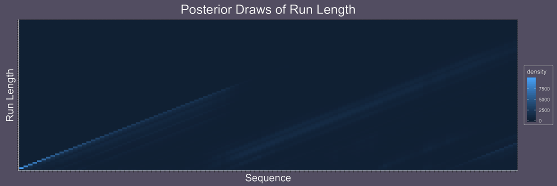
Coal Mining Dataset - Analytical solution



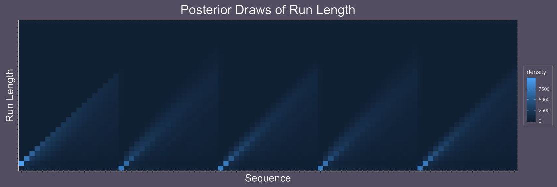
Synthetic Data - Analytical solution



Coal Mining Dataset - MCMC posterior samples



Synthetic Data - MCMC posterior samples



Outline

DISCUSSION

Aalto University
School of Science

Conclusion and Future Work

- The second changepoint is controversial
- Overall, Bayesian Online Changepoint Model is superior to other models.
- More generic versions of the MCMC solutions for Online Bayesian CD Model

More results at;

`https://github.com/onurpoyraz/bcpm-stan/`

References I

R. P. Adams and D. J. MacKay.
Bayesian online changepoint detection.
arXiv preprint arXiv:0710.3742, 2007.

C. Fonnesbeck, A. Patil, D. Huard, and J. Salvatier.
Pymc user's guide.

S. U. Guide.
Change point models.

G. Gundersen.
Bayesian online changepoint model.

Appendix

Aalto University
School of Science

Marginal Predictive Distribution

$$P\left(x_{t+1} \mid \boldsymbol{x}_{1:t}\right) = \sum_{r_t} P\left(x_{t+1} \mid r_t, \boldsymbol{x}_t^{(r)}\right) P\left(r_t \mid \boldsymbol{x}_{1:t}\right)$$

Recursive Run Length Estimation

$$P\left(r_t, \boldsymbol{x}_{1:t}\right) = \sum_{r_{t-1}} P\left(r_t \mid r_{t-1}\right) P\left(x_t \mid r_{t-1}, \boldsymbol{x}_t^{(r)}\right) P\left(r_{t-1}, \boldsymbol{x}_{1:t-1}\right)$$