PyMCon Web Series #11

PyMCon
WEBSERIES

THE ONLY CONSTANT IS CHANGE

Bespoke Changepoint Modelling in PyMC

Abuzar Mahmood

Neuroscientist
Brandeis University

LIVE Q&A
20th November

SPONSORS

ODSC

pymcon.com/sponsors

About Me

- Swartz Foundation Computational Neuroscience Postdoctoral Fellow
- Katz Lab – Dynamics of Neural Taste Processing
- Brandeis University, MA

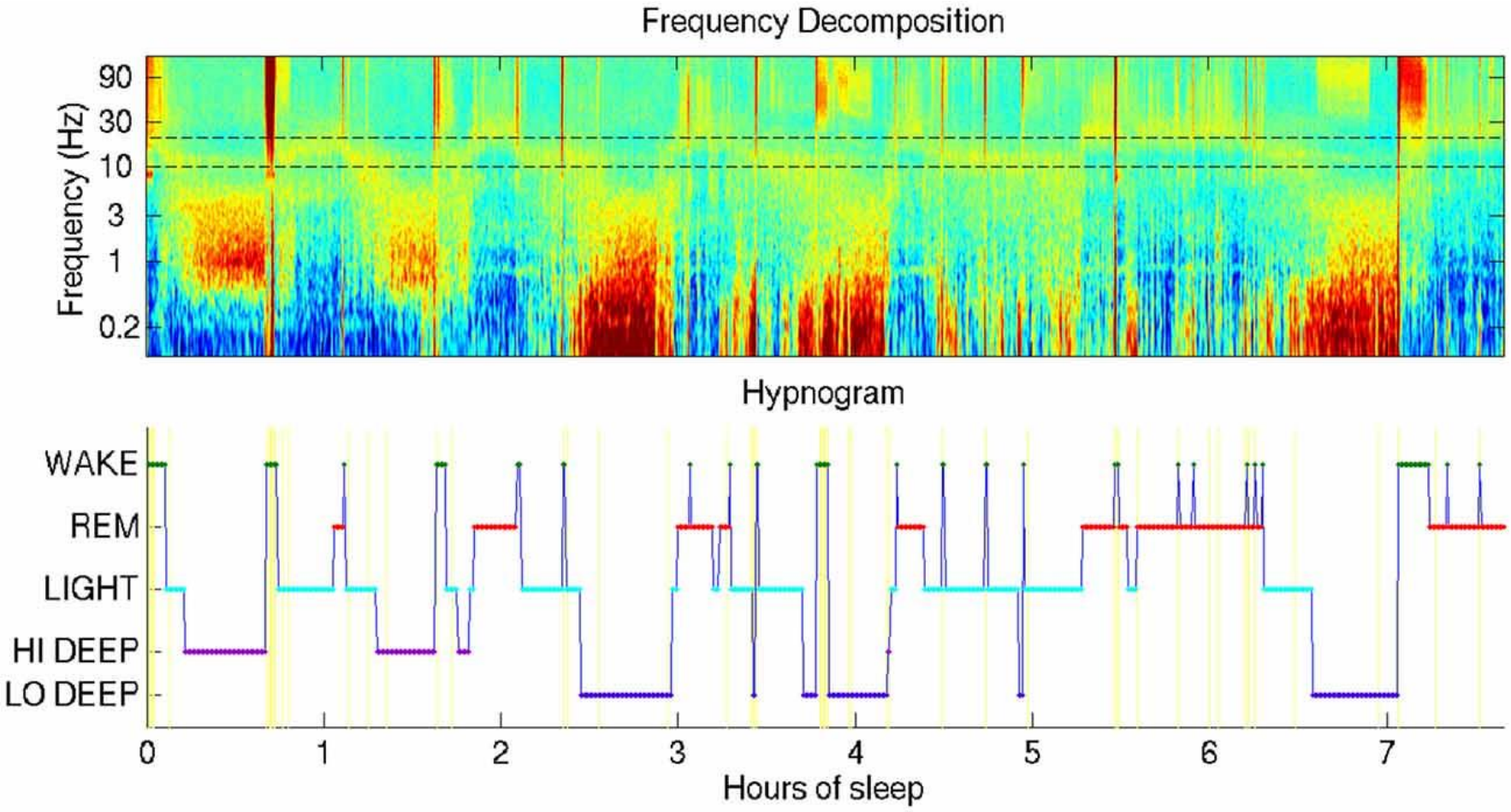
B.L.E.CH LAB

katzlab.squarespace.com

Layout of talk

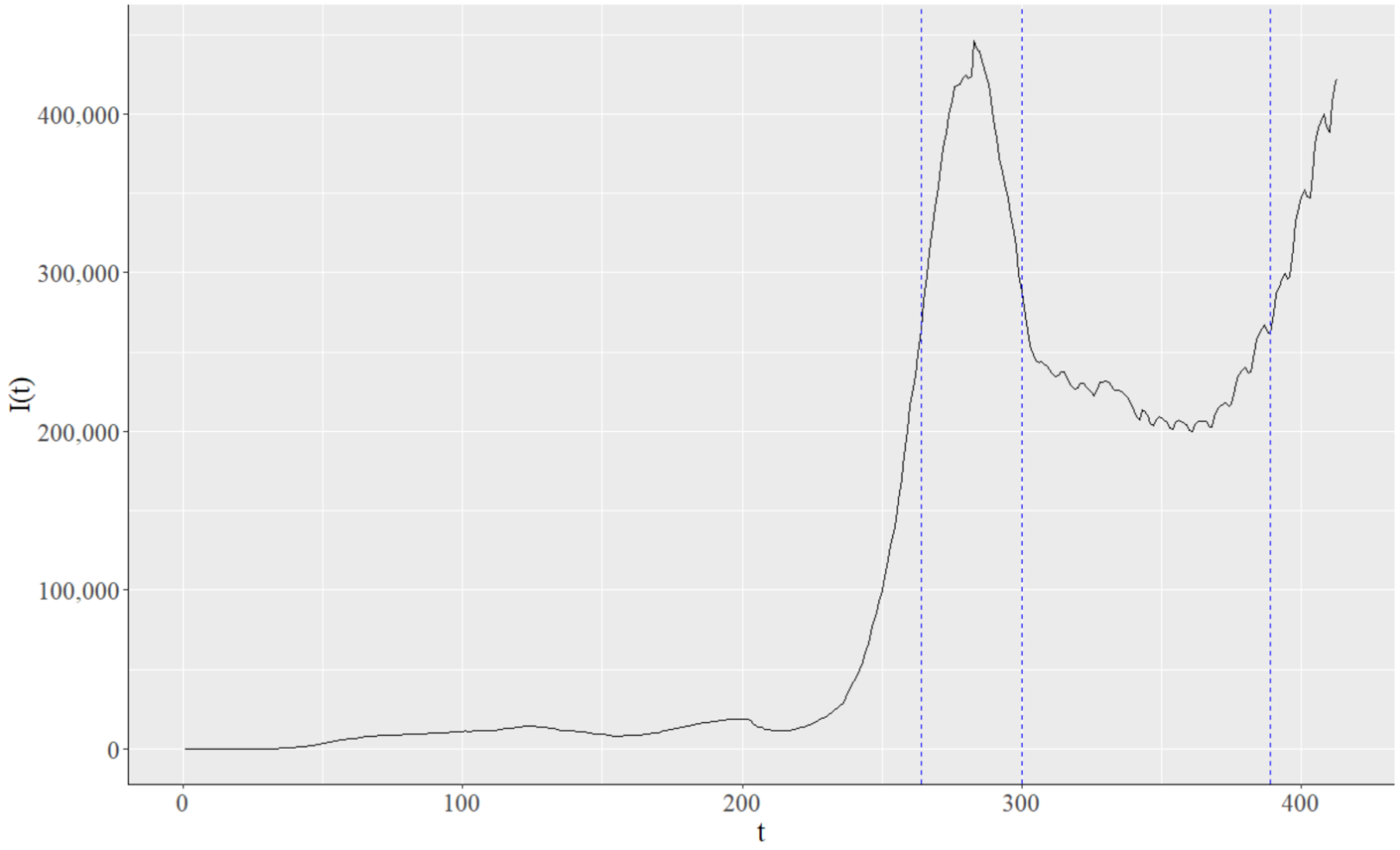
- Changepoint modelling use-cases:
  - Research (offline), and Industry (online) use-cases
  - My work:
    - What I study and how we use changepoint models
    - Why these are the appropriate types of models to use
- Finite and infinite mixtures
- Overview of changepoint models in shared code
  - Not going into detail on model construction

Offline example: Sleep-state segmentation from EEG



Onton, Julie A., et al. "Visualization of Whole-Night Sleep EEG From 2-Channel Mobile Recording Device Reveals Distinct Deep Sleep Stages with Differential Electrodermal Activity." *Frontiers in Human Neuroscience*, vol. 10, 2016. *Frontiers*, https://www.frontiersin.org/articles/10.3389/fnhum.2016.00605.

Offline example: Changes in number of COVID cases



Jegede, Segun Light, and Krzysztof J. Szajowski. “Change-Point Detection in Homogeneous Segments of COVID-19 Daily Infection.” *Axioms*, vol. 11, no. 5, May 2022, p. 213. *www.mdpi.com*, https://doi.org/10.3390/axioms11050213.

Offline example: Detecting changes in forest cover

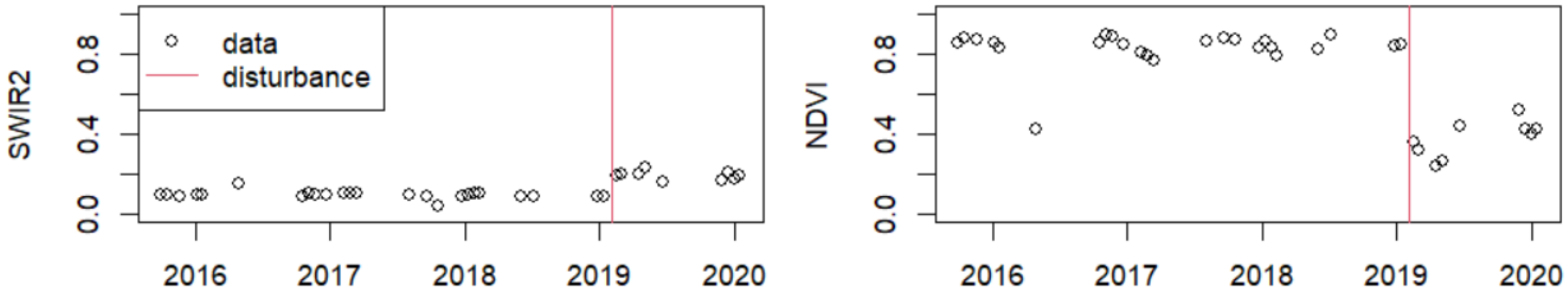
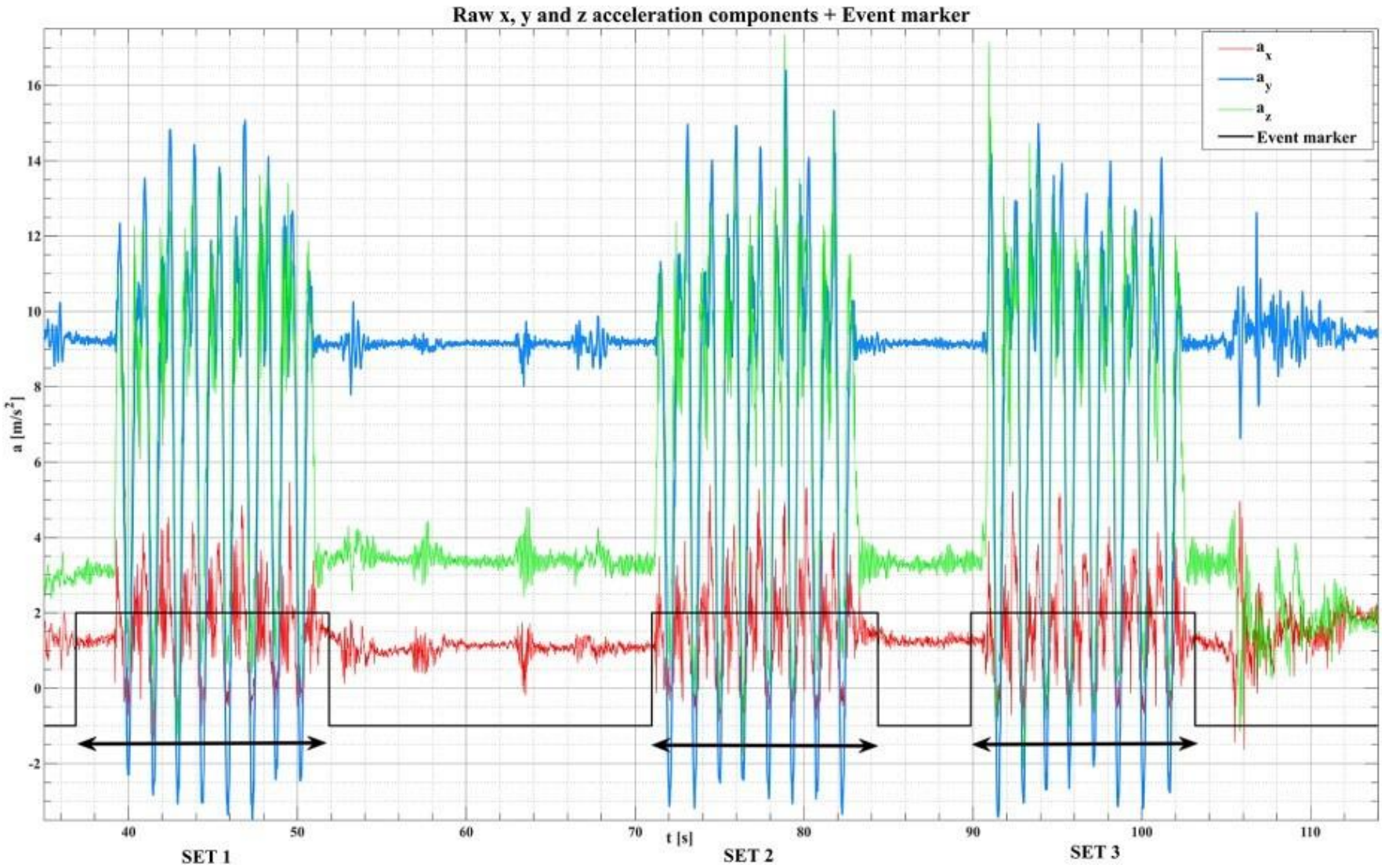


Figure 2: Observed SWIR2 and NDVI measurements of a 30 × 30m location in Myanmar with a deforestation event identified on Feb 2, 2019.

Wendelberger, Laura J., Josh M. Gray, Brian J. Reich and Alyson G. Wilson. “Monitoring Deforestation Using Multivariate Bayesian Online Changepoint Detection with Outliers.” (2021).

Online example: Detecting bouts of activity from accelerometer data



Džaja, D., et al. "Accelerometer-Based Algorithm for the Segmentation and Classification of Repetitive Human Movements during Workouts." Automatika, vol. 64, no. 2, Apr. 2023, pp. 211–24. DOI.org (Crossref), https://doi.org/10.1080/00051144.2022.2121247.

Online example: Detecting network attacks

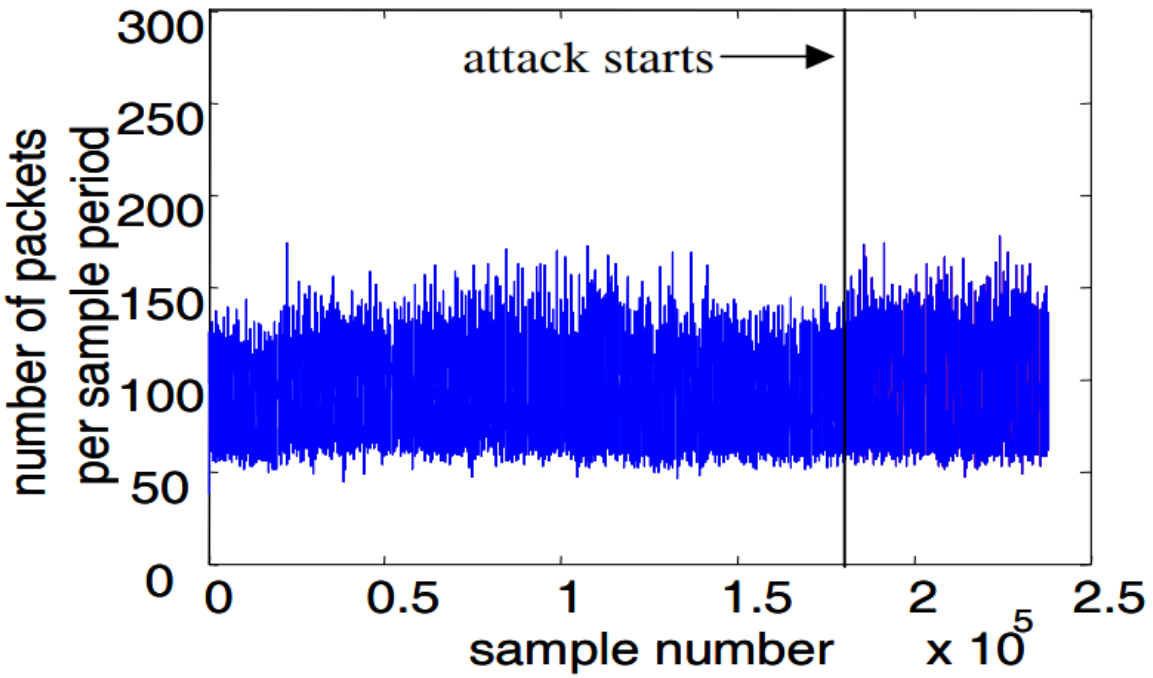
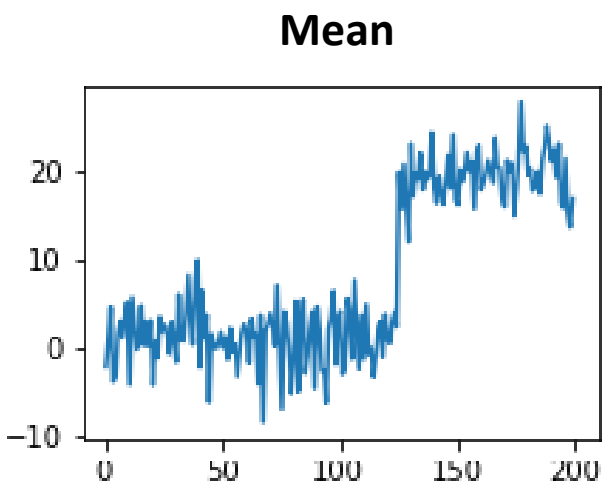


Figure 1: Number of packets in a sample period vs. time. Observe that the attack is not visible to the naked-eye.

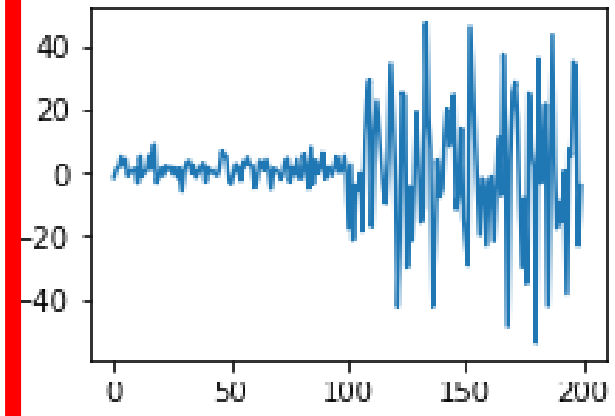
Tartakovsky, Alexander & Rozovskii, Boris & Shah, Khushboo. (2005). A nonparametric multichart CUSUM test for rapid intrusion detection.

Types of Changepoints

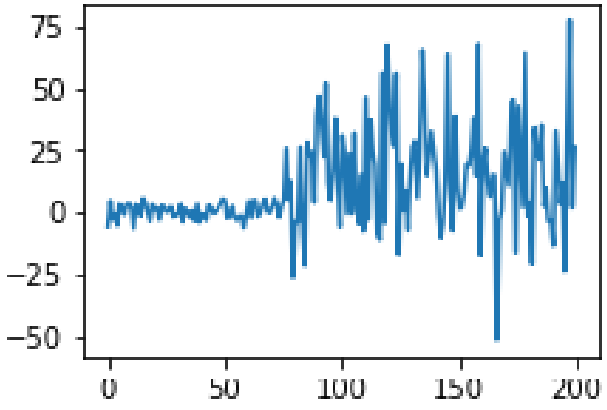
Mean



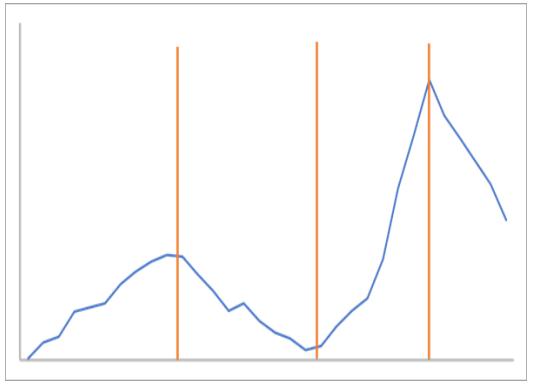
Variance



Mean + Variance

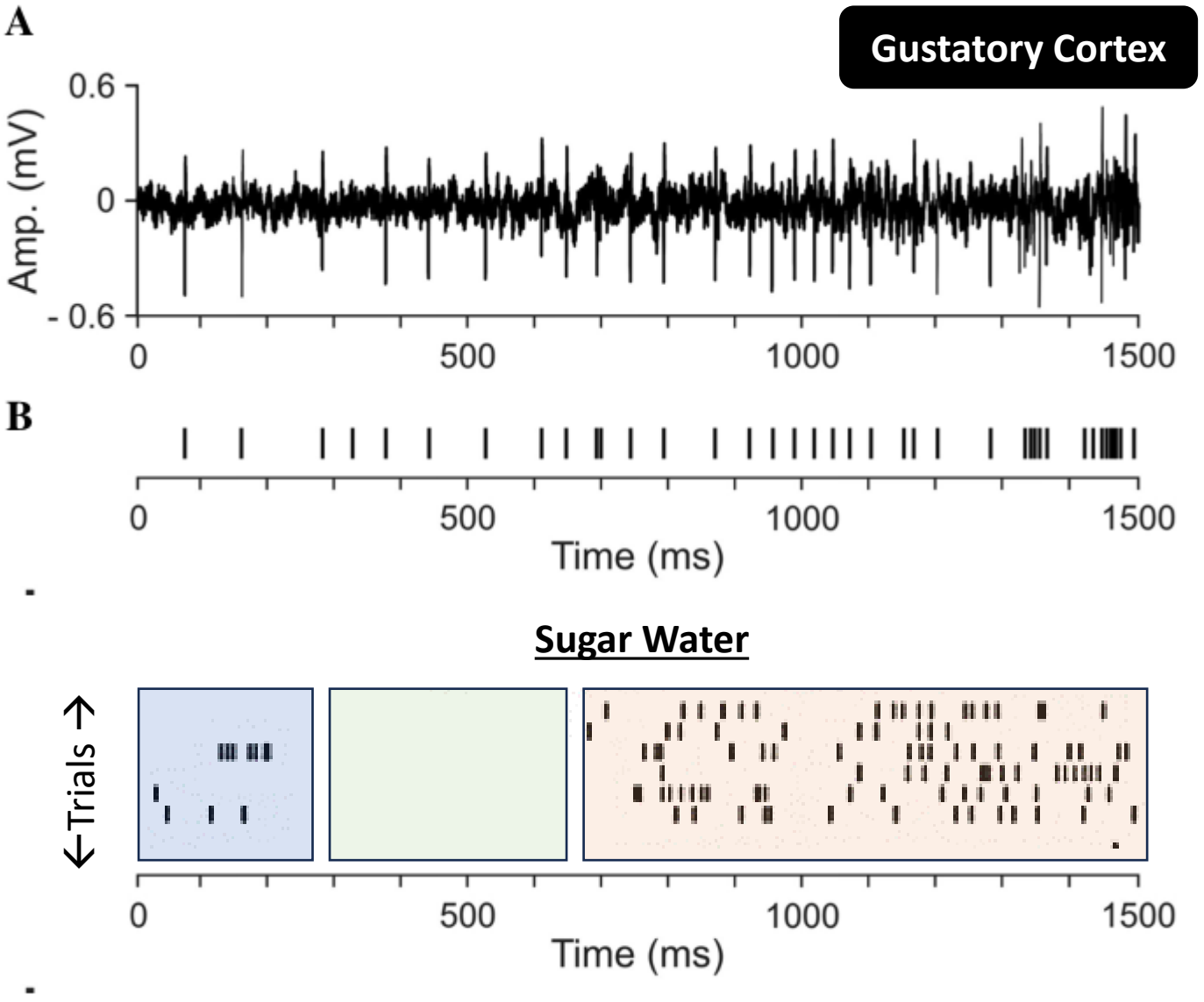


Slope



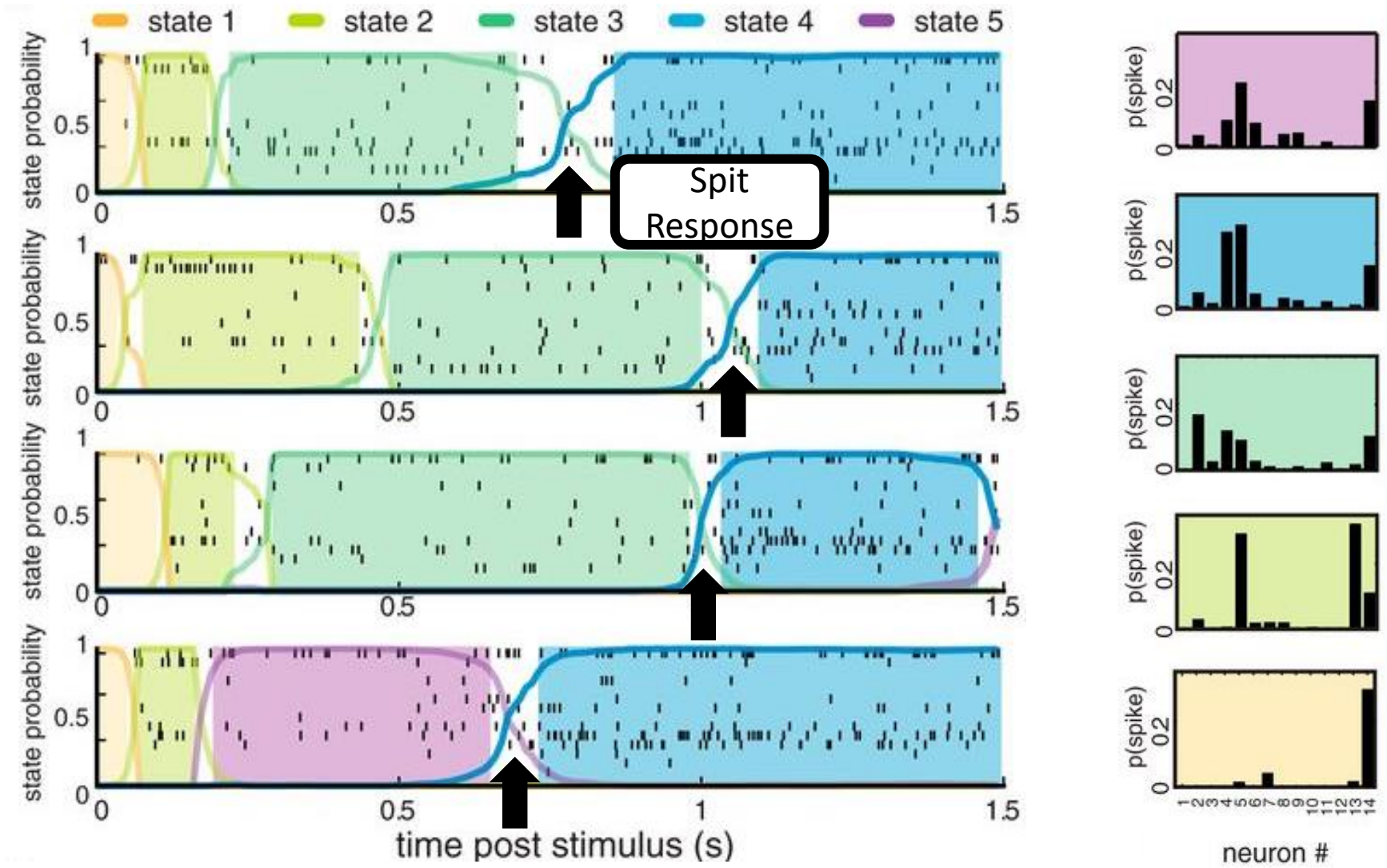
REFS:
Abuzar Mahmood (2022) Teachables [Source Code] https://github.com/abuzarmahmood/teachables/tree/main
https://pro.arcgis.com/en/pro-app/latest/tool-reference/space-time-pattern-mining/how-change-point-detection-works.htm

My Work: Neural Processing of Taste



Caro-Martín, C.R., Delgado-García, J.M., Gruart, A. et al. Spike sorting based on shape, phase, and distribution features, and K-TOPS clustering with validity and error indices. Sci Rep 8, 17796 (2018). https://doi.org/10.1038/s41598-018-35491-4

My Work: Neural Processing of Taste



Sadacca, Brian F., et al. “The Behavioral Relevance of Cortical Neural Ensemble Responses Emerges Suddenly.” Journal of Neuroscience, vol. 36, no. 3, Jan. 2016, pp. 655–69. www.jneurosci.org, https://doi.org/10.1523/JNEUROSCI.2265-15.2016.

My Work: Neural Processing of Taste

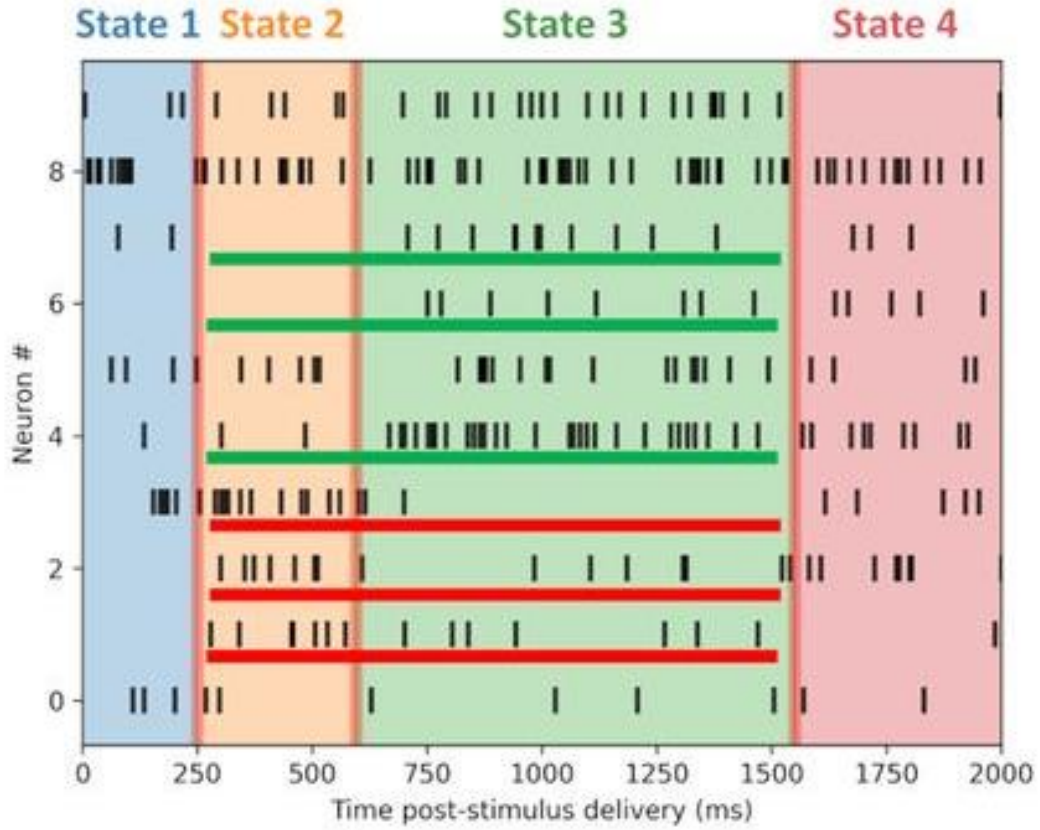
- Why use changepoint models:
  - Inspire and constrain theoretical models of processing in the brain.
  - Measure strength of coordination between brain regions over time.

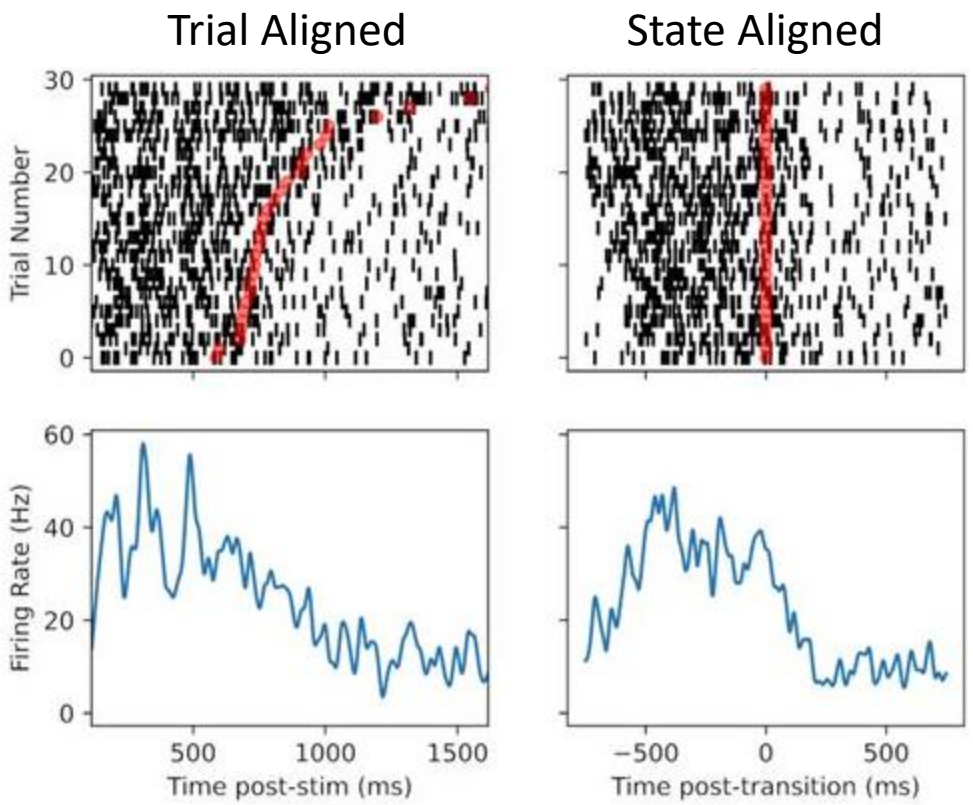
- Why Bayesian?
  - Because we like uncertainty estimates.

- Why PyMC?
  - Rapid iteration for model-building
  - Easy extension of models for specific use-cases/needs
  - Powerful inference

My Work: Neural Processing of Taste

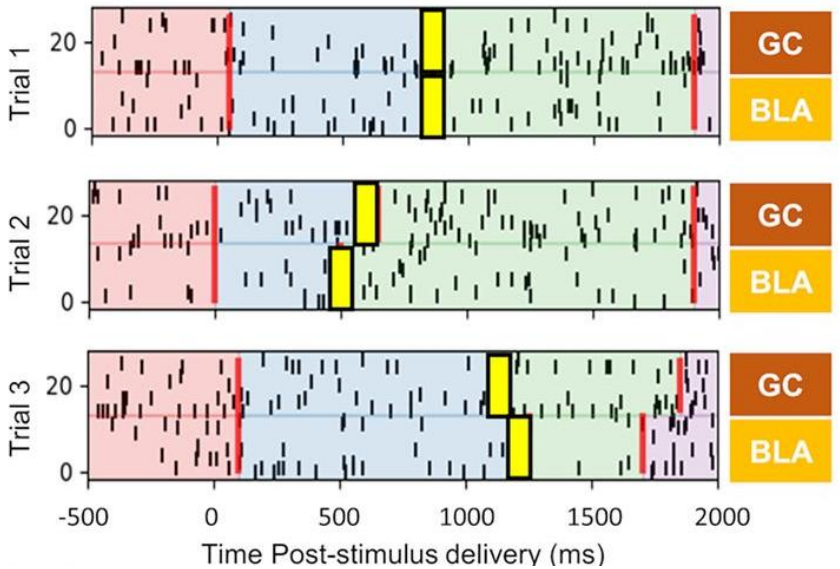
Basolateral Amygdala (BLA) Responses

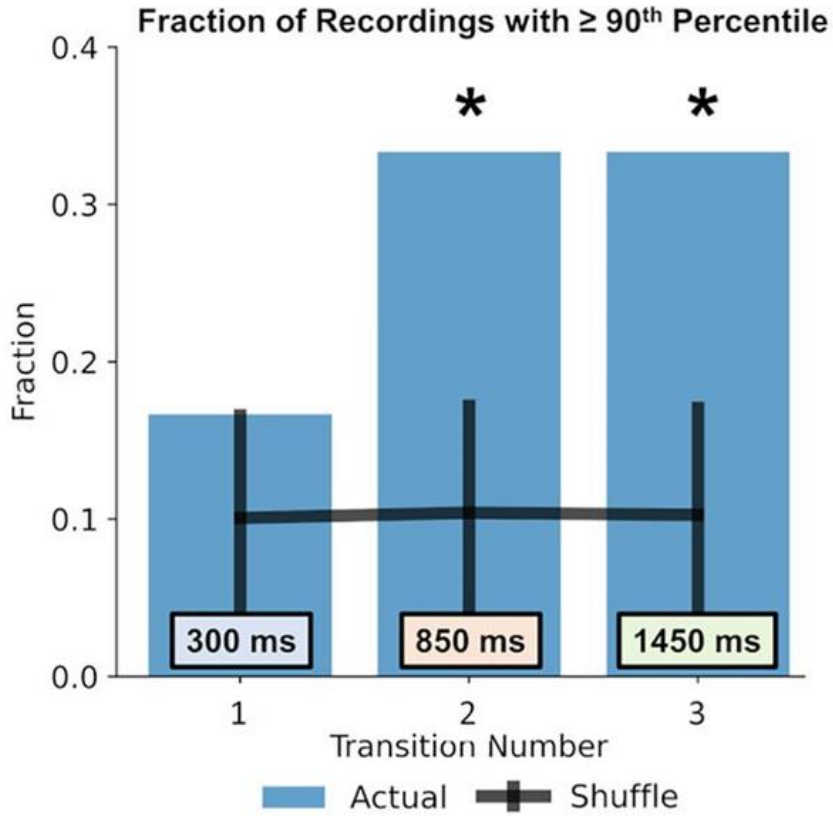


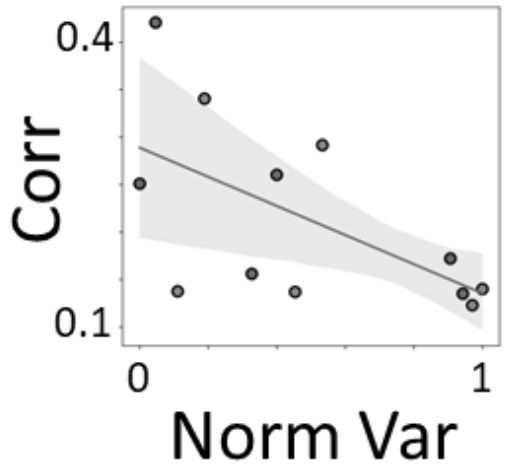


Mahmood, Abuzar, et al. “Coupled Dynamics of Stimulus-Evoked Gustatory Cortical and Basolateral Amygdalar Activity.” *Journal of Neuroscience*, vol. 43, no. 3, Jan. 2023, pp. 386–404. *www.jneurosci.org*, https://doi.org/10.1523/JNEUROSCI.1412-22.2022.

My Work: Neural Processing of Taste



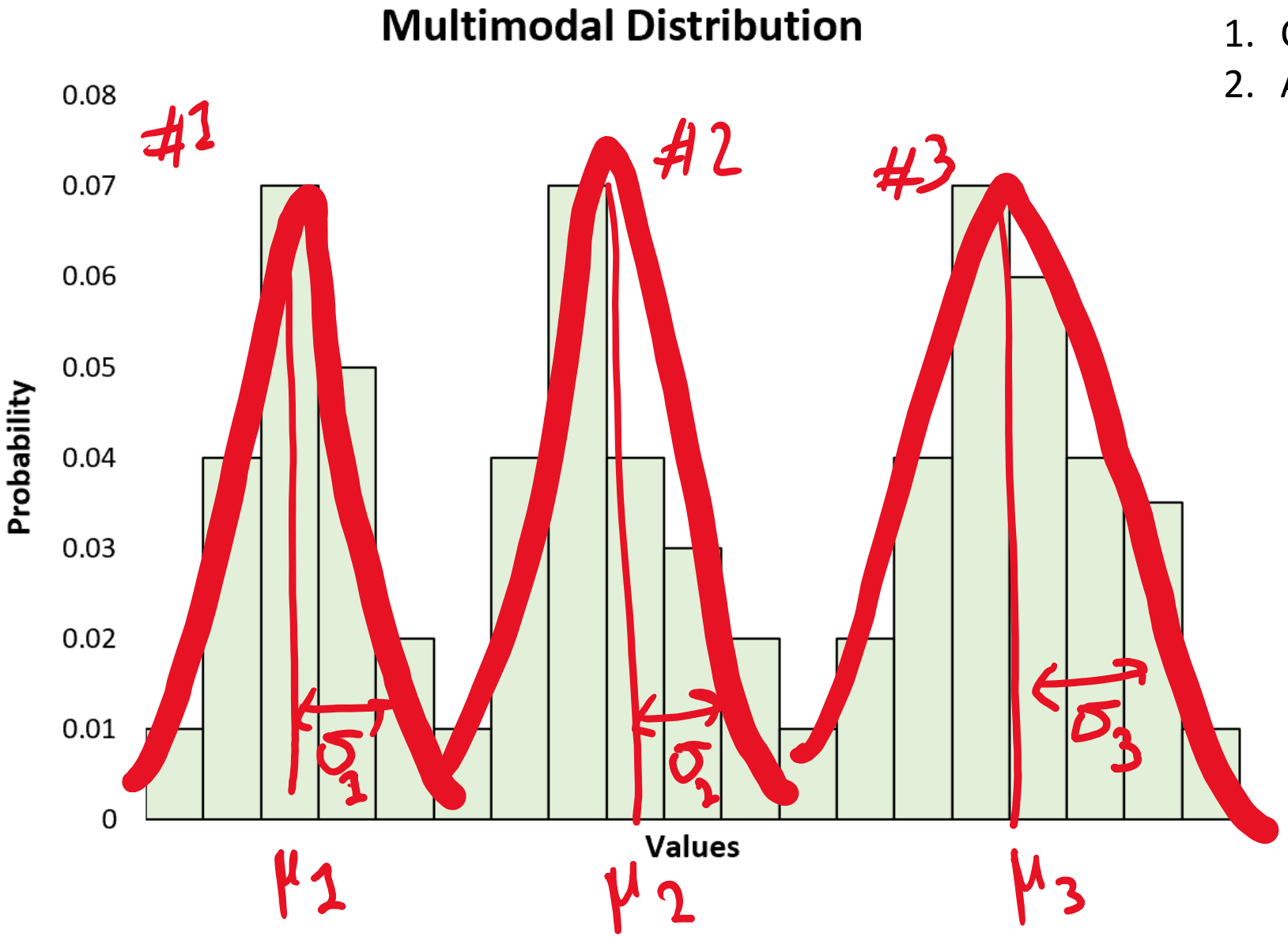




Layout of talk

- Changepoint modelling use-cases:
  - Research (offline), and Industry (online) use-cases
  - My work:
    - What I study and how we use changepoint models
    - Why these are the appropriate types of models to use
- Finite and infinite mixtures
- Overview of changepoint models in shared code

Finite Mixture Models: Intro

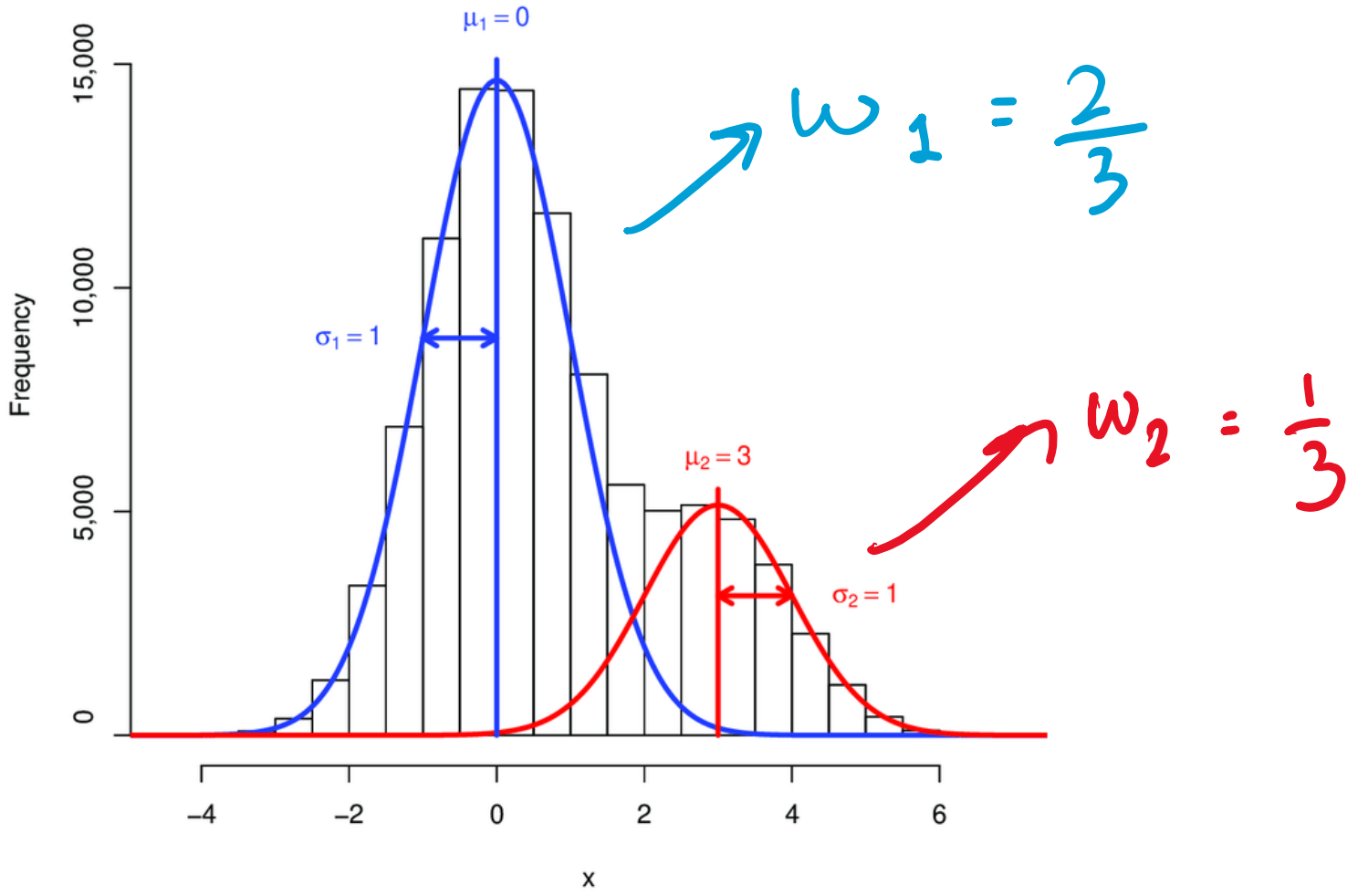


1. Clustering
2. Analytically describe the distribution

$$p \sim w_1 N(\mu_1, \sigma_1) + w_2 N(\mu_2, \sigma_2) + w_3 N(\mu_3, \sigma_3)$$

$$w = [\tfrac{1}{3}, \tfrac{1}{3}, \tfrac{1}{3}]$$

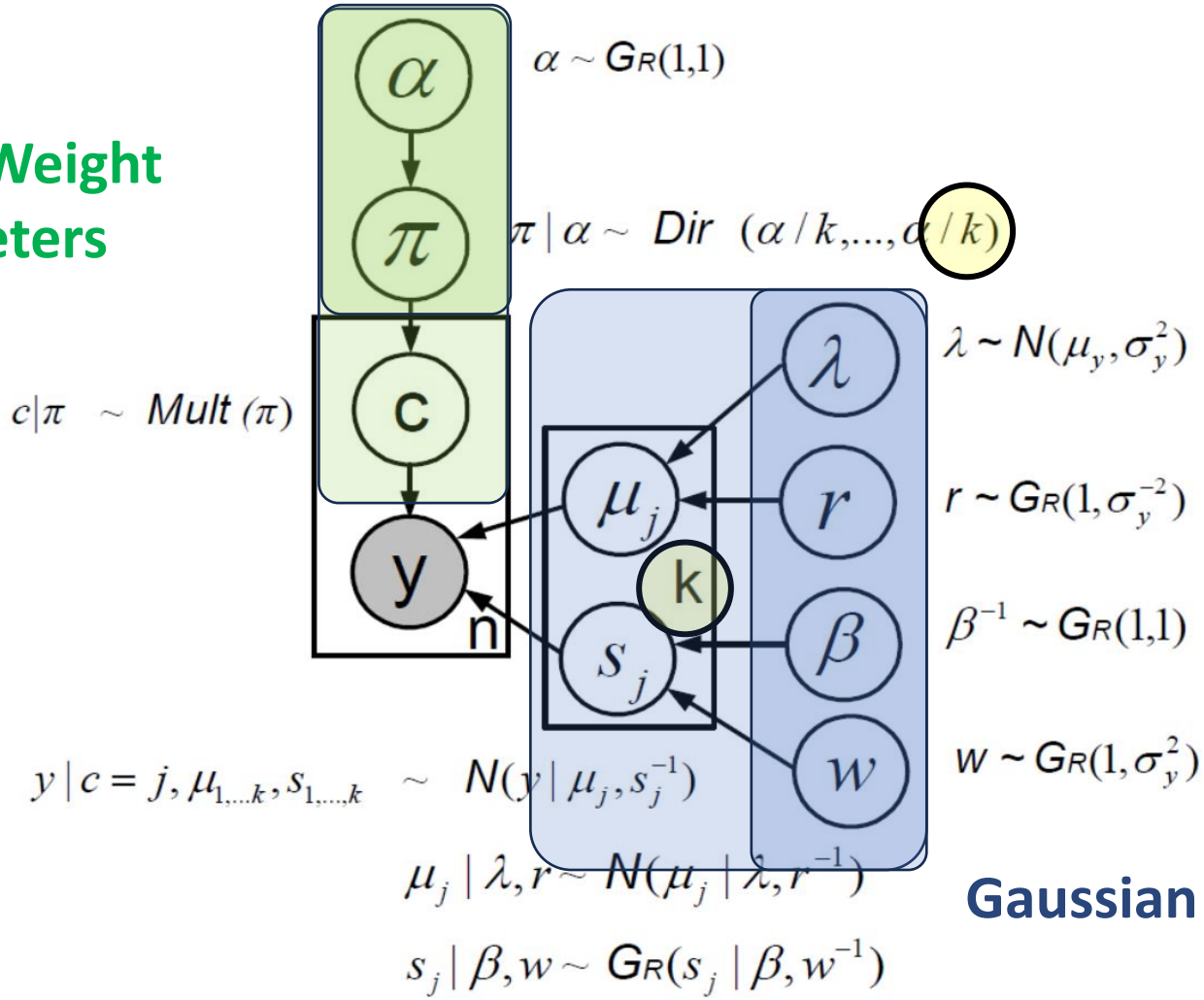
Finite Mixture Models: Intro



Beisemann, Marie, et al. “Comparison of Recent Acceleration Techniques for the EM Algorithm in One- and Two-Parameter Logistic IRT Models.” *Psych*, vol. 2, no. 4, Dec. 2020, pp. 209–52. *www.mdpi.com*, https://doi.org/10.3390/psych2040018.

Finite Mixture Models : Model Structure

Mixture Weight Parameters

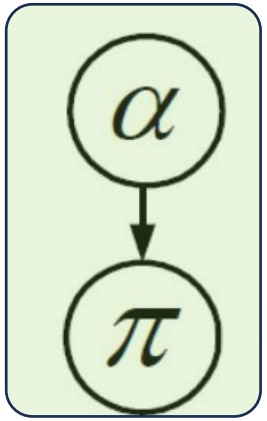


Gaussian Parameters

Figure 1: The graphical presentation of FGMM

Derivation of Gibbs Sampling for Finite Gaussian Mixture Model,
Xiaodong Yu,
University of Maryland, College Park,
2009,
https://yuxiaodong.files.wordpress.com/2009/09/derivation_of_gibbs_sampling_gmm1.pdf

Finite Mixture Models : Mixture Weight Parameters



Concentration parameter

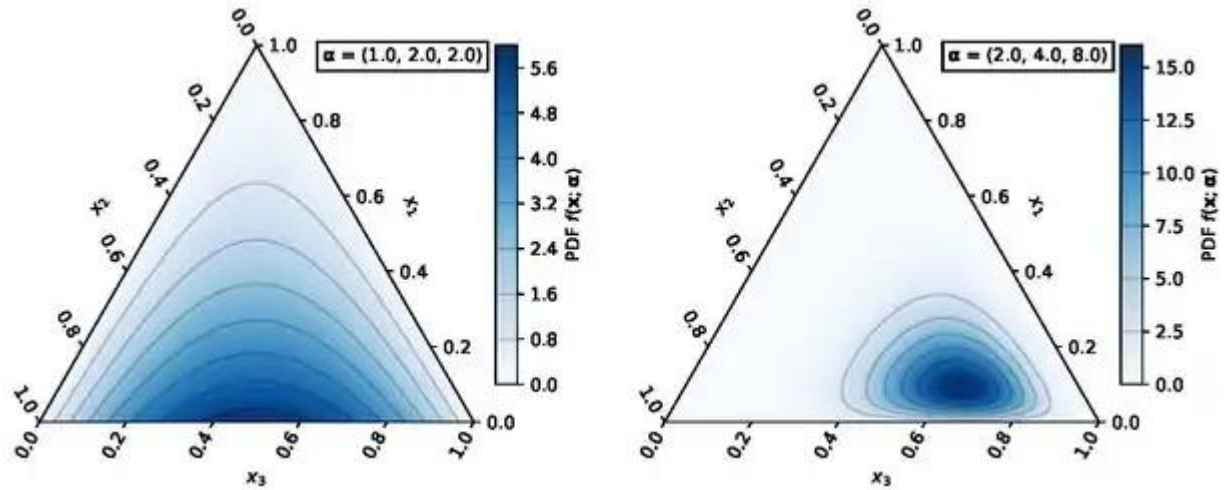
$$\alpha \sim G_R(1,1)$$

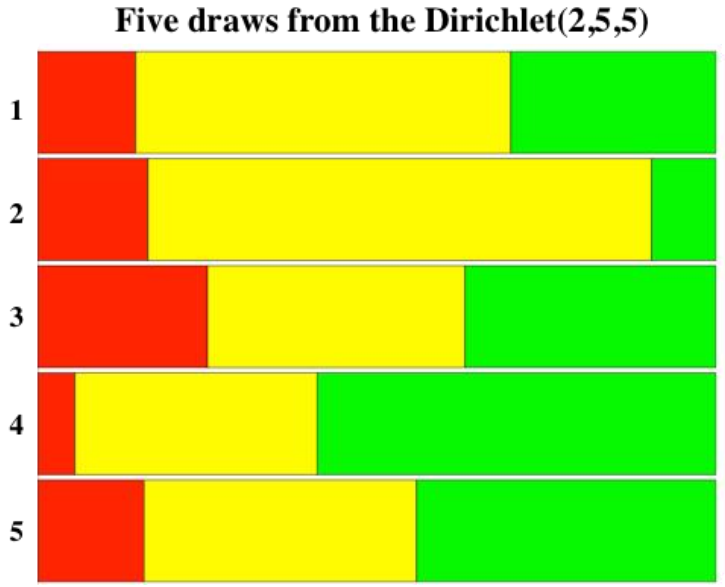
$$\pi \mid \alpha \sim Dir\ (\alpha / k, \ldots, \alpha / k)$$

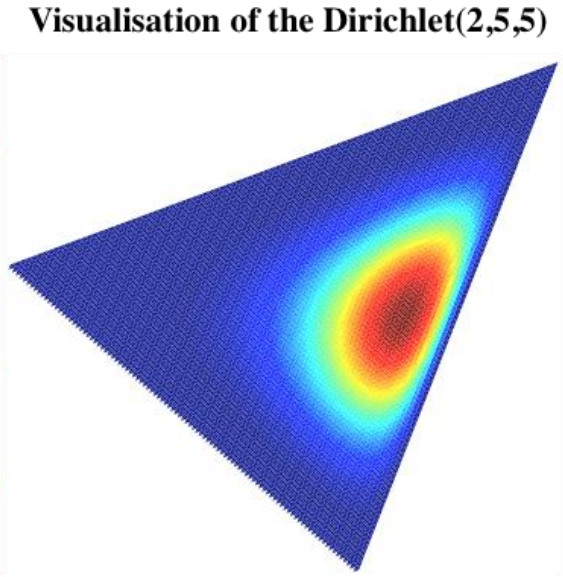
Mixture Weights

$k \to \infty$

Dirichlet Distribution







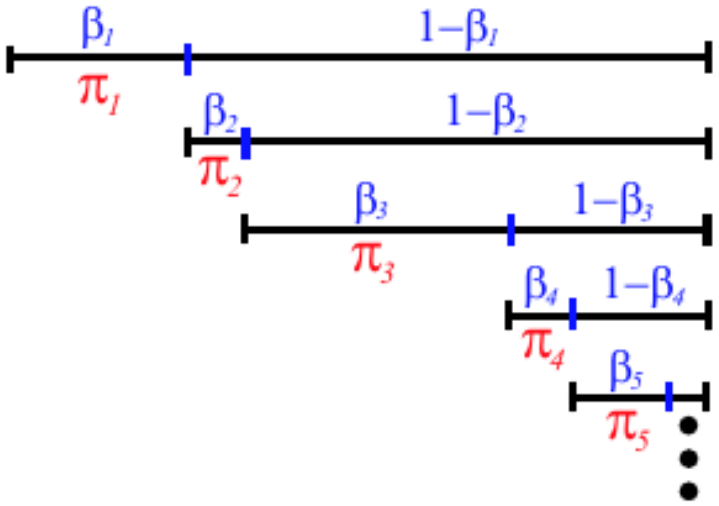
https://en.wikipedia.org/wiki/Dirichlet_distribution

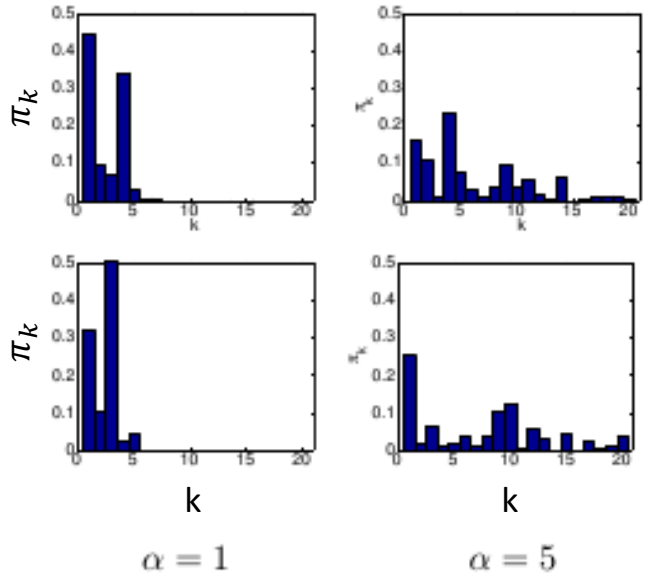
https://blog.shakirm.com/2015/12/machine-learning-trick-of-the-day-6-tricks-with-sticks/

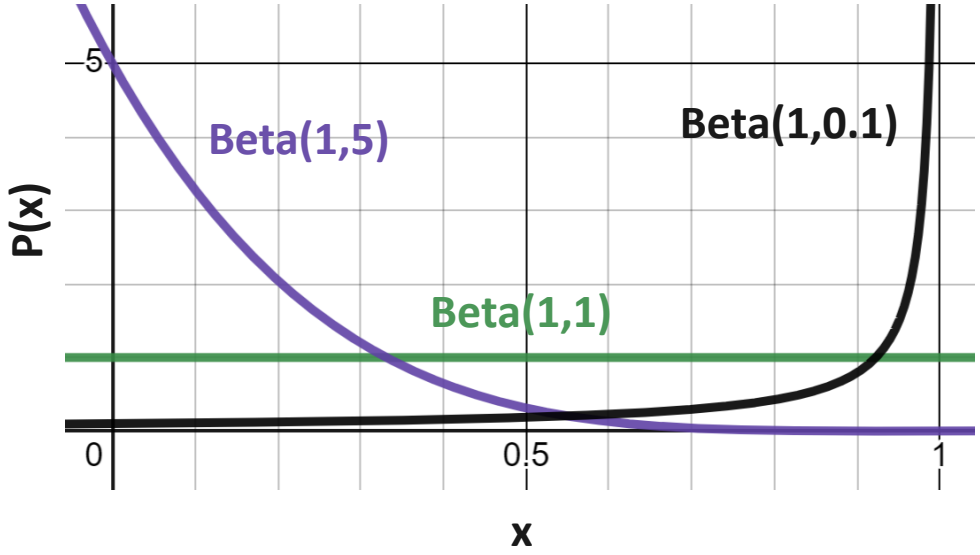
Infinite Mixture Models : Mixture Weight parameters

Dirichlet Process

$$\pi_k = \beta_k \prod_{l=1}^{k-1}(1-\beta_l) \qquad \beta_k \sim \mathrm{Beta}(1,\alpha)$$







Wiecki, Thomas V. Sequential Sampling Models in Computational Psychiatry: Bayesian Parameter Estimation, Model Selection and Classification. arXiv, 22 Mar. 2013. arXiv.org, https://doi.org/10.48550/arXiv.1303.5616.

Beta distribution plot generated using : desmos.com

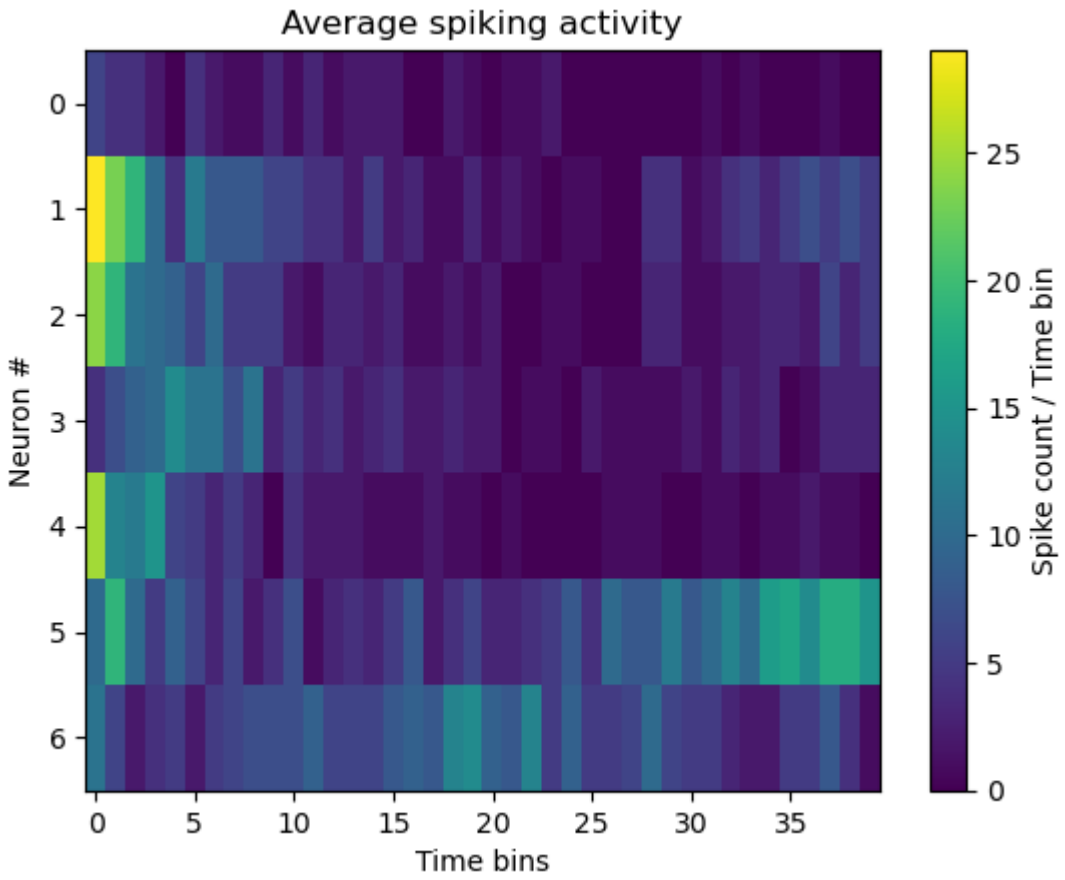
Clarifications:

- One model uses a finite mixture
- The other uses a Dirichlet Process (but not for an infinite mixture)

Layout of talk

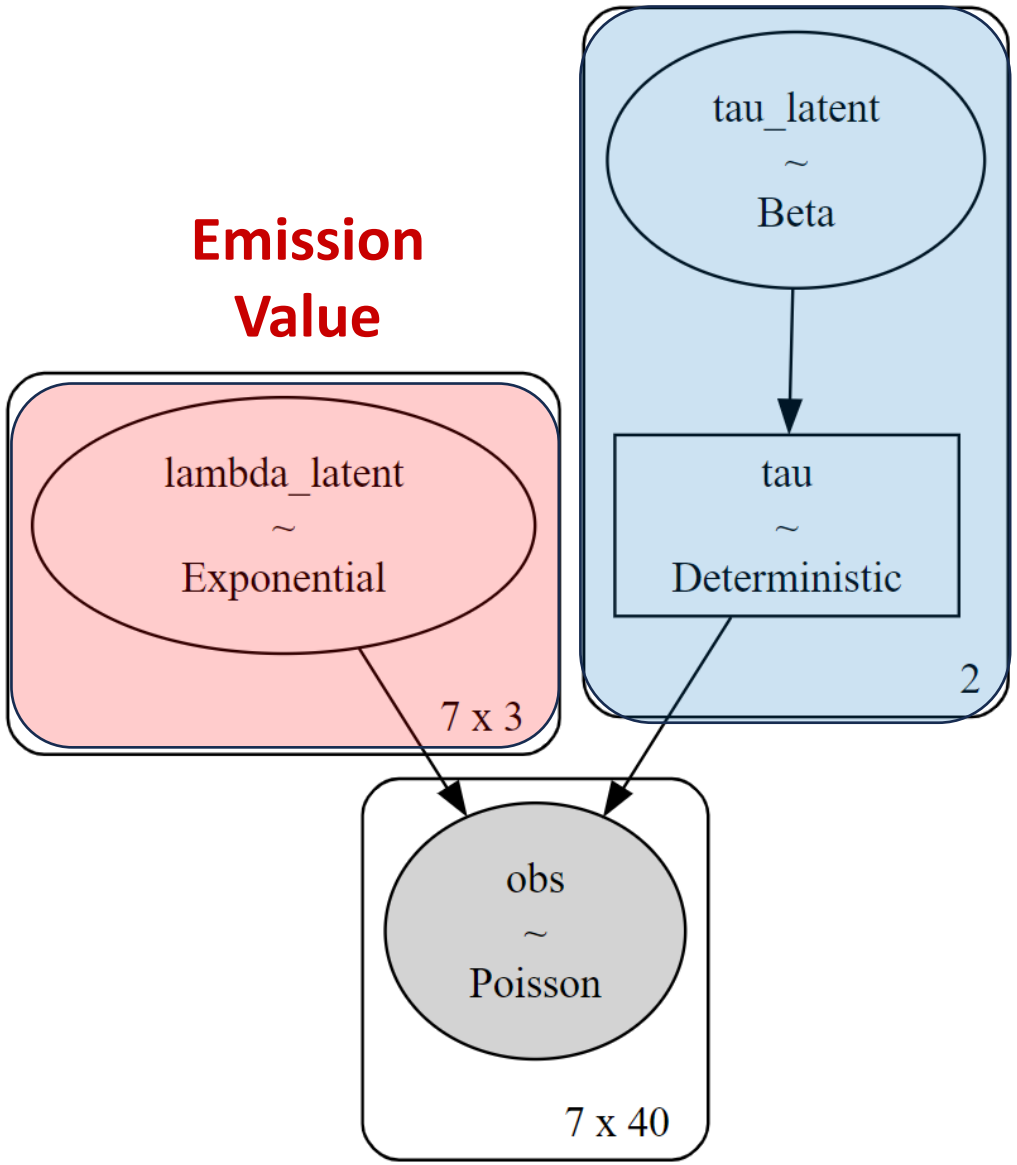
- Changepoint modelling use-cases:
  - Research (offline), and Industry (online) use-cases
  - My work:
    - What I study and how we use changepoint models
    - Why these are the appropriate types of models to use
- Finite and infinite mixtures
- Overview of changepoint models in shared code
  - Simple Model
  - Model with automatic state number determination (Dirichlet Process)
  - Model which allows drift across trials (Mixture Model)

“Simple” Poisson Changepoint

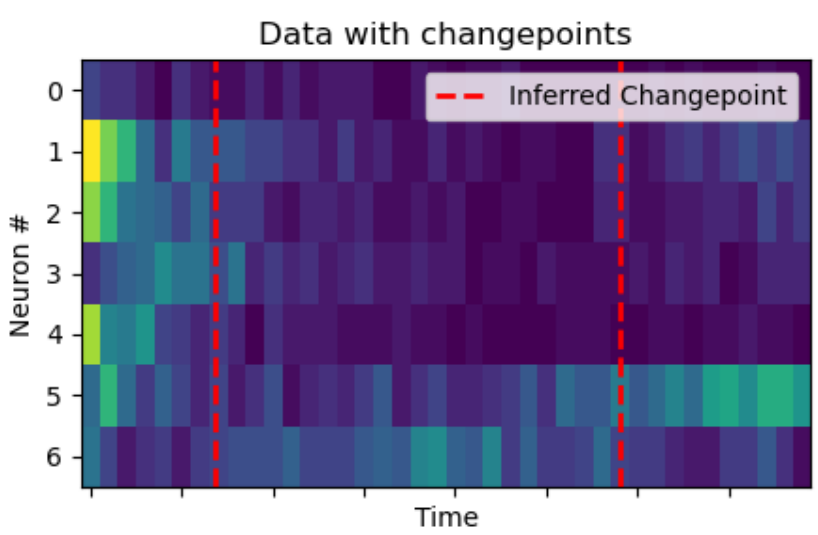


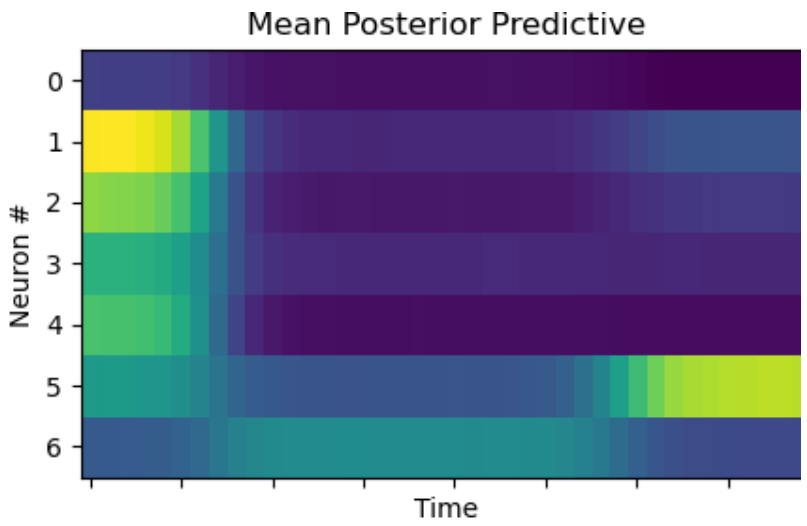
Changepoint Location

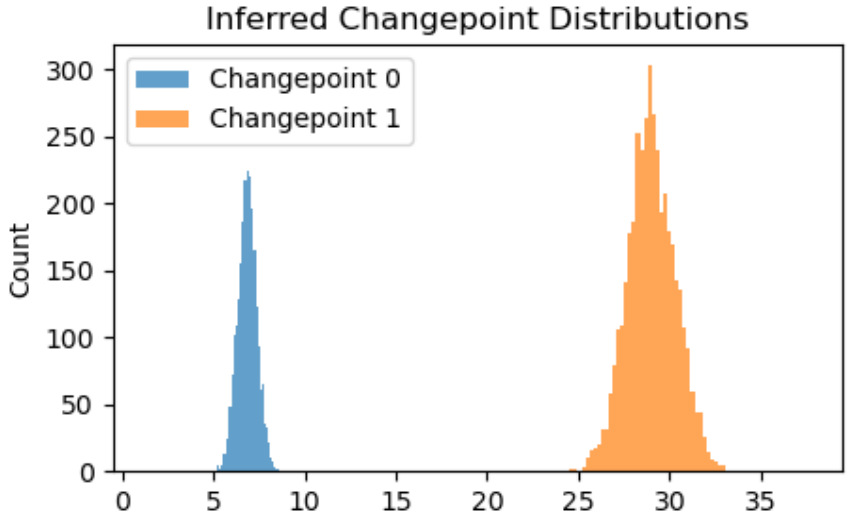
Emission Value



“Simple” Poisson Changepoint



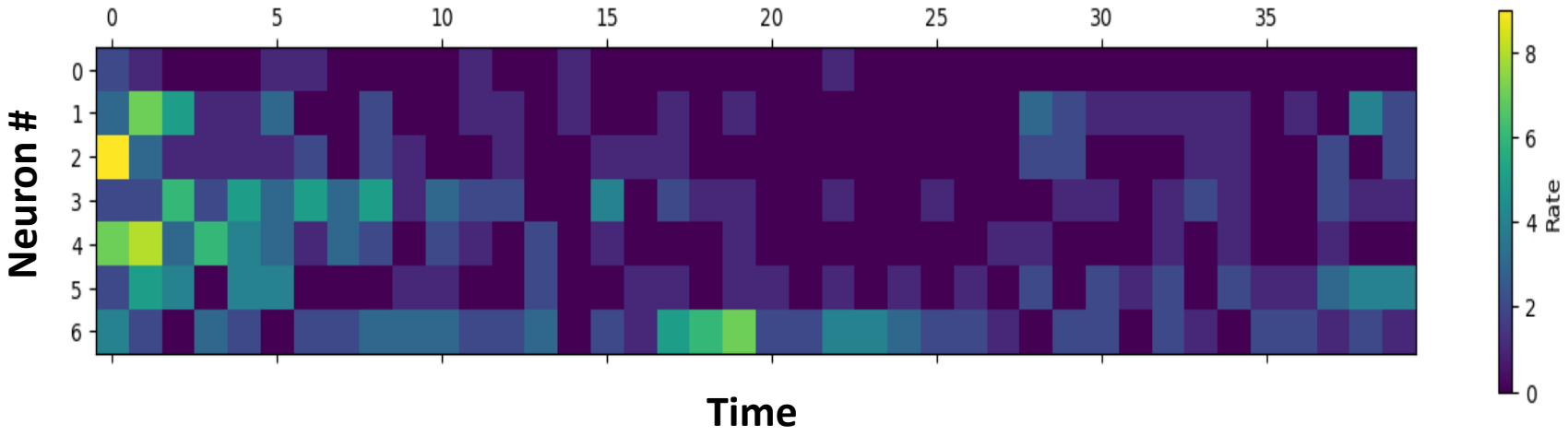




Layout of talk

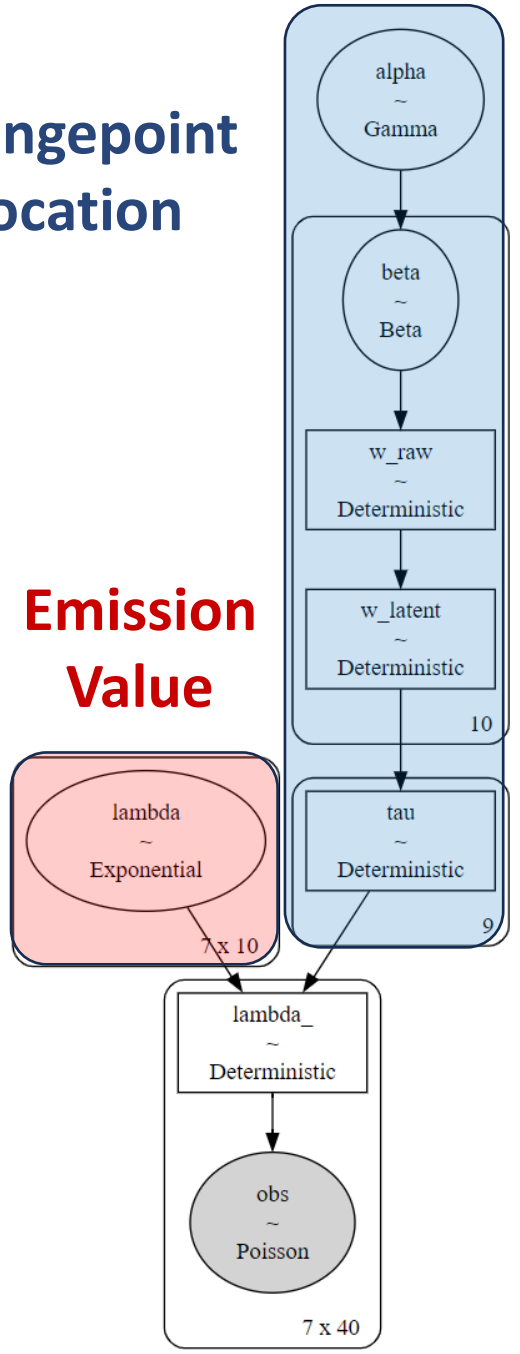
- Changepoint modelling use-cases:
  - Research (offline), and Industry (online) use-cases
  - My work:
    - What I study and how we use changepoint models
    - Why these are the appropriate types of models to use
- Finite and infinite mixtures
- Overview of changepoint models in shared code
  - Simple Model
  - Model with automatic state number determination (Dirichlet Process)
  - Model which allows drift across trials (Mixture Model)

Dirichlet Process Changepoint

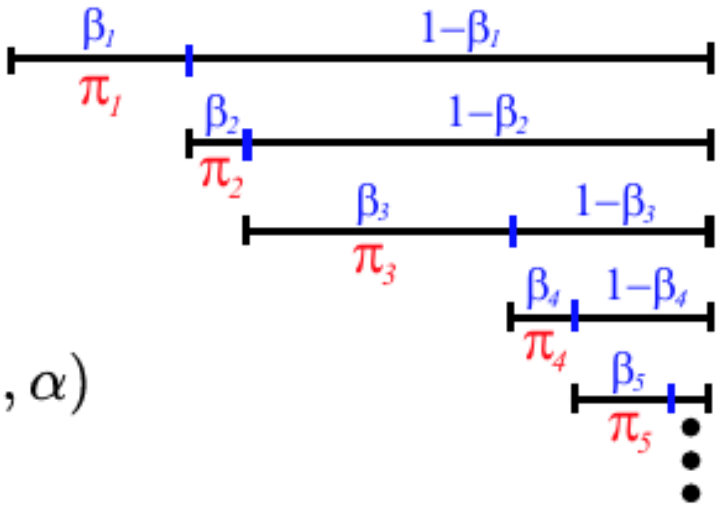


Changepoint Location

Emission Value



w_raw

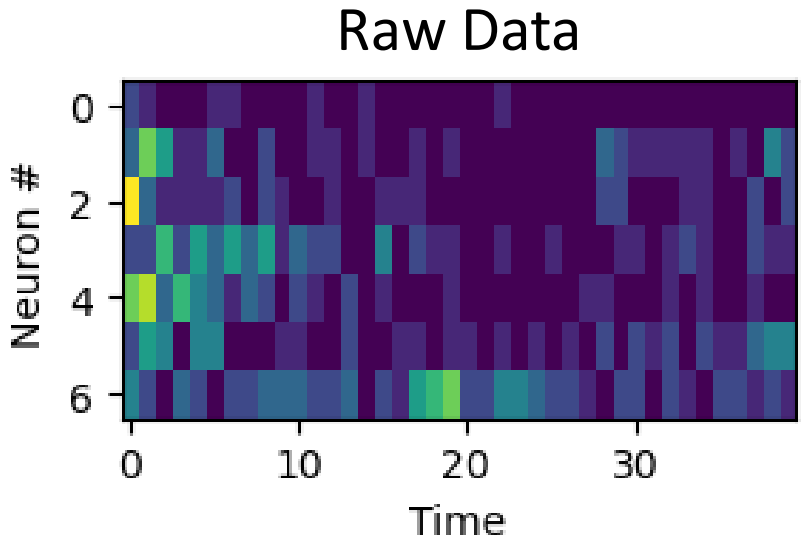


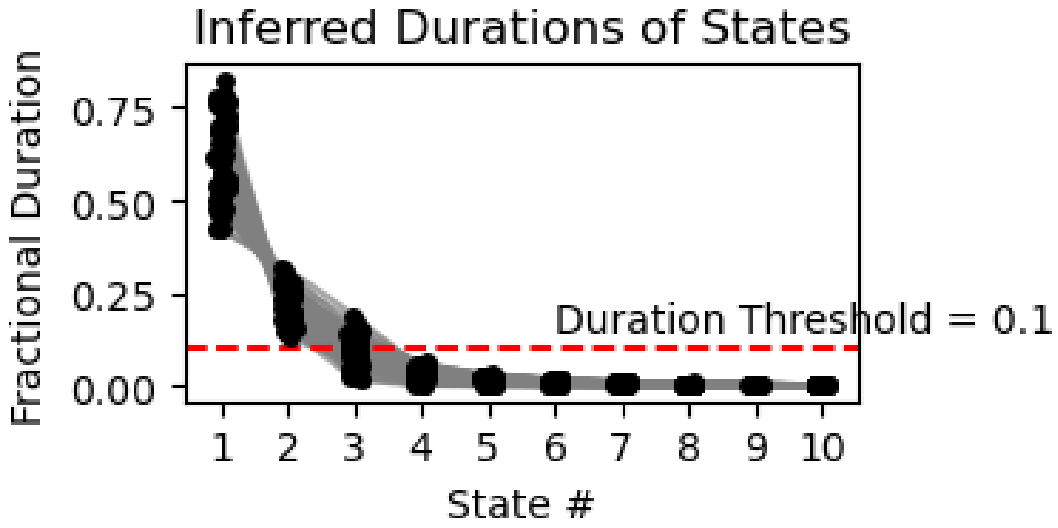
Dirichlet Process

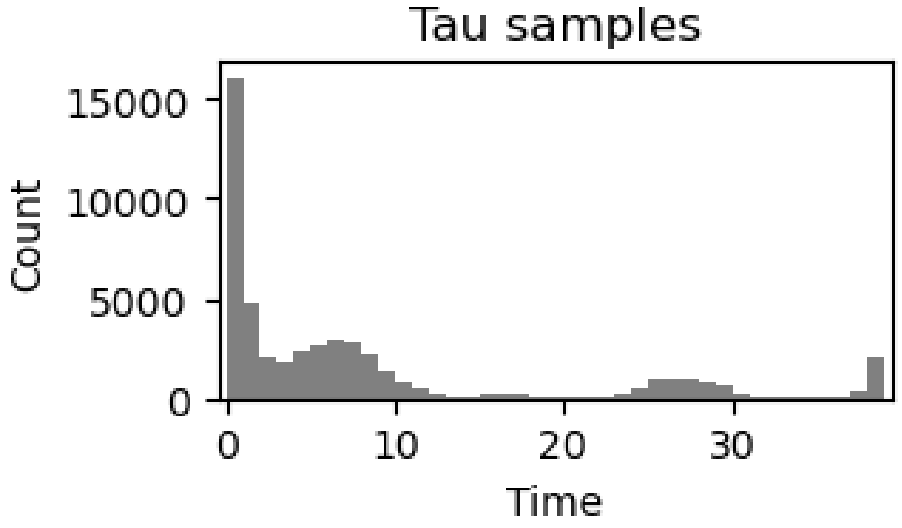
$$\pi_k = \beta_k \prod_{l=1}^{k-1}(1-\beta_l) \qquad \beta_k \sim \mathrm{Beta}(1,\alpha)$$

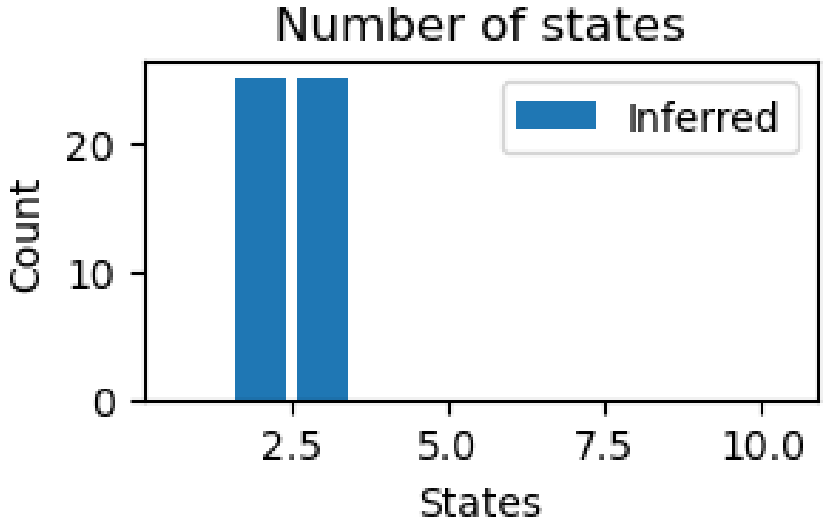
https://medium.com/@abuzar_mahmood/probabilistically-determining-the-number-of-components-for-unsupervised-learning-e1c3c4a82082

Dirichlet Process Changepoint



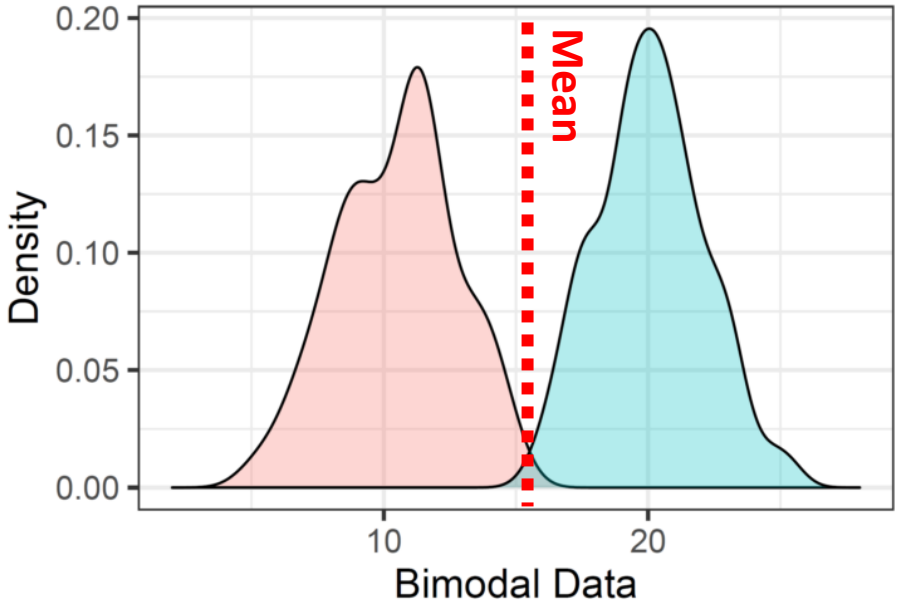






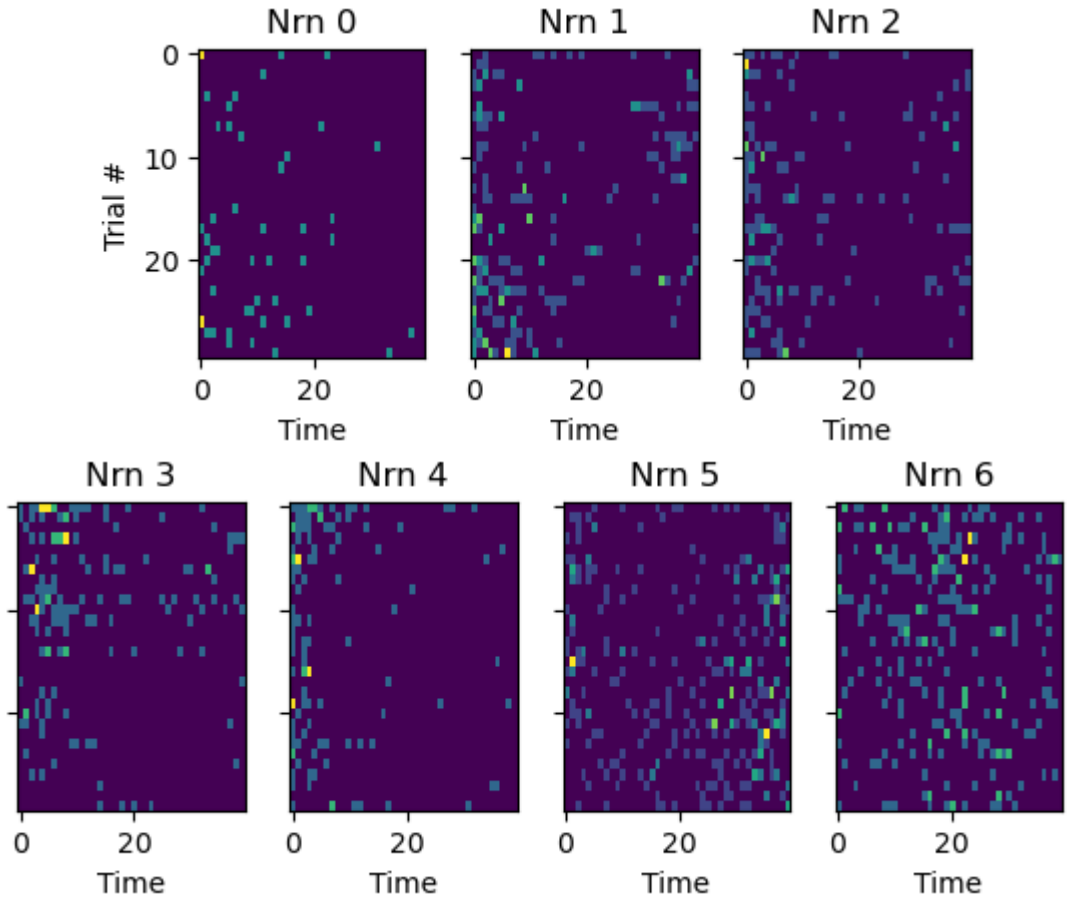
Layout of talk

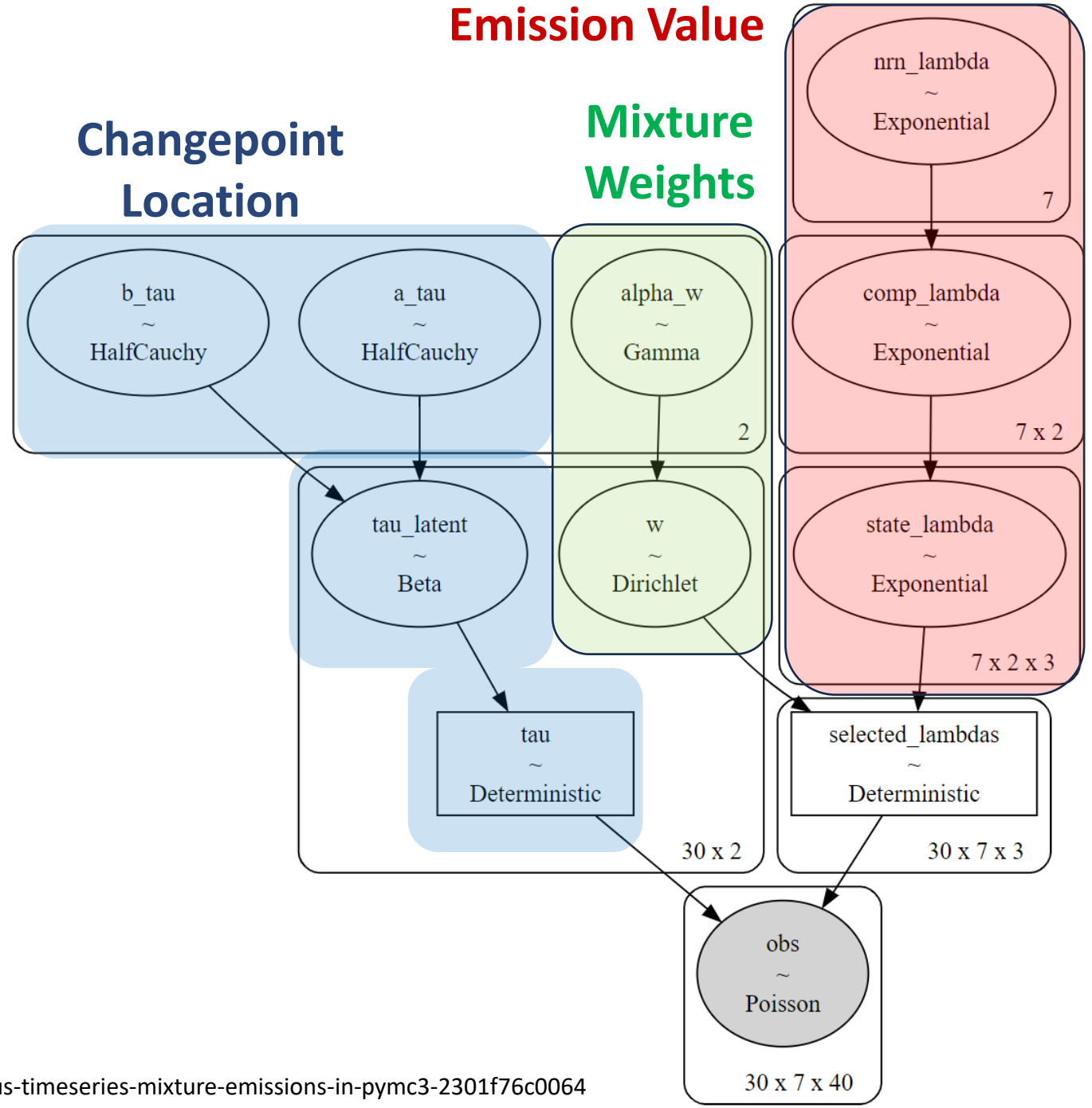
- Changepoint modelling use-cases:
  - Research (offline), and Industry (online) use-cases
  - My work:
    - What I study and how we use changepoint models
    - Why these are the appropriate types of models to use
- Finite and infinite mixtures
- Overview of changepoint models in shared code
  - Simple Model
  - Model with automatic state number determination (Dirichlet Process)
  - Model which allows drift across trials (Mixture Model)



https://universeofdatascience.com/how-to-determine-if-data-are-unimodal-or-multimodal-in-r/

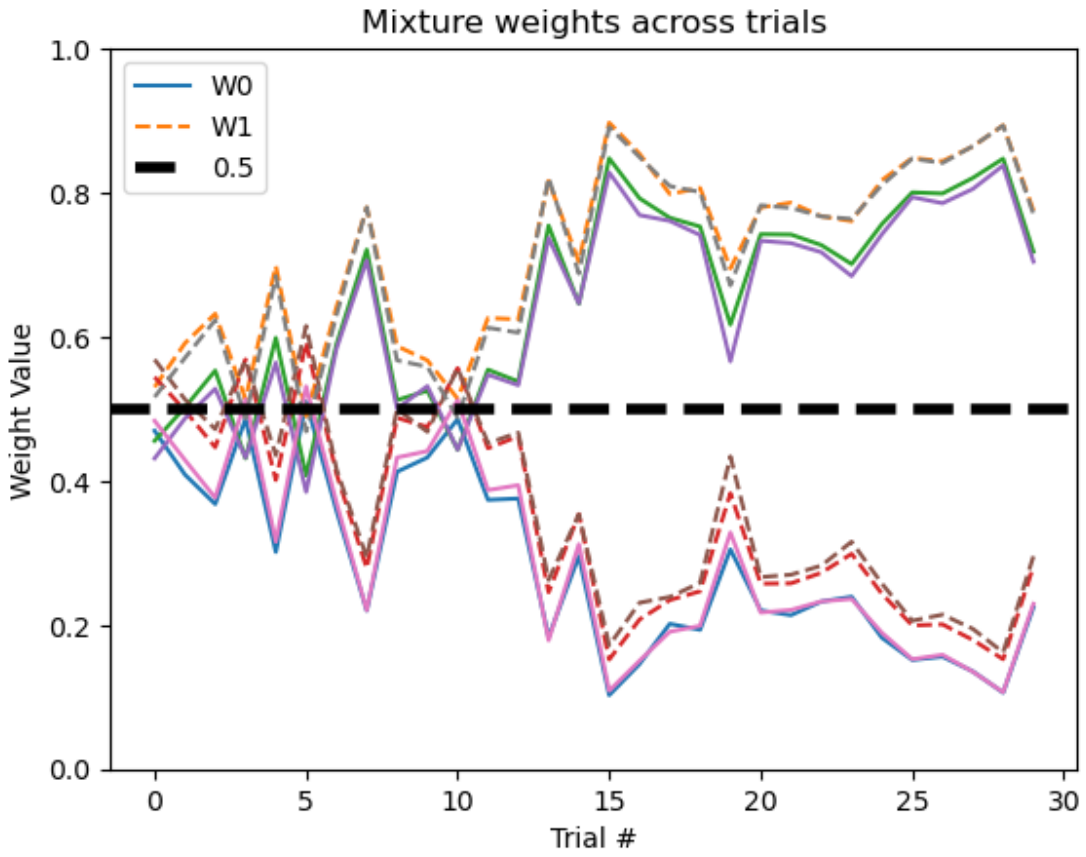
Mixture Changepoint

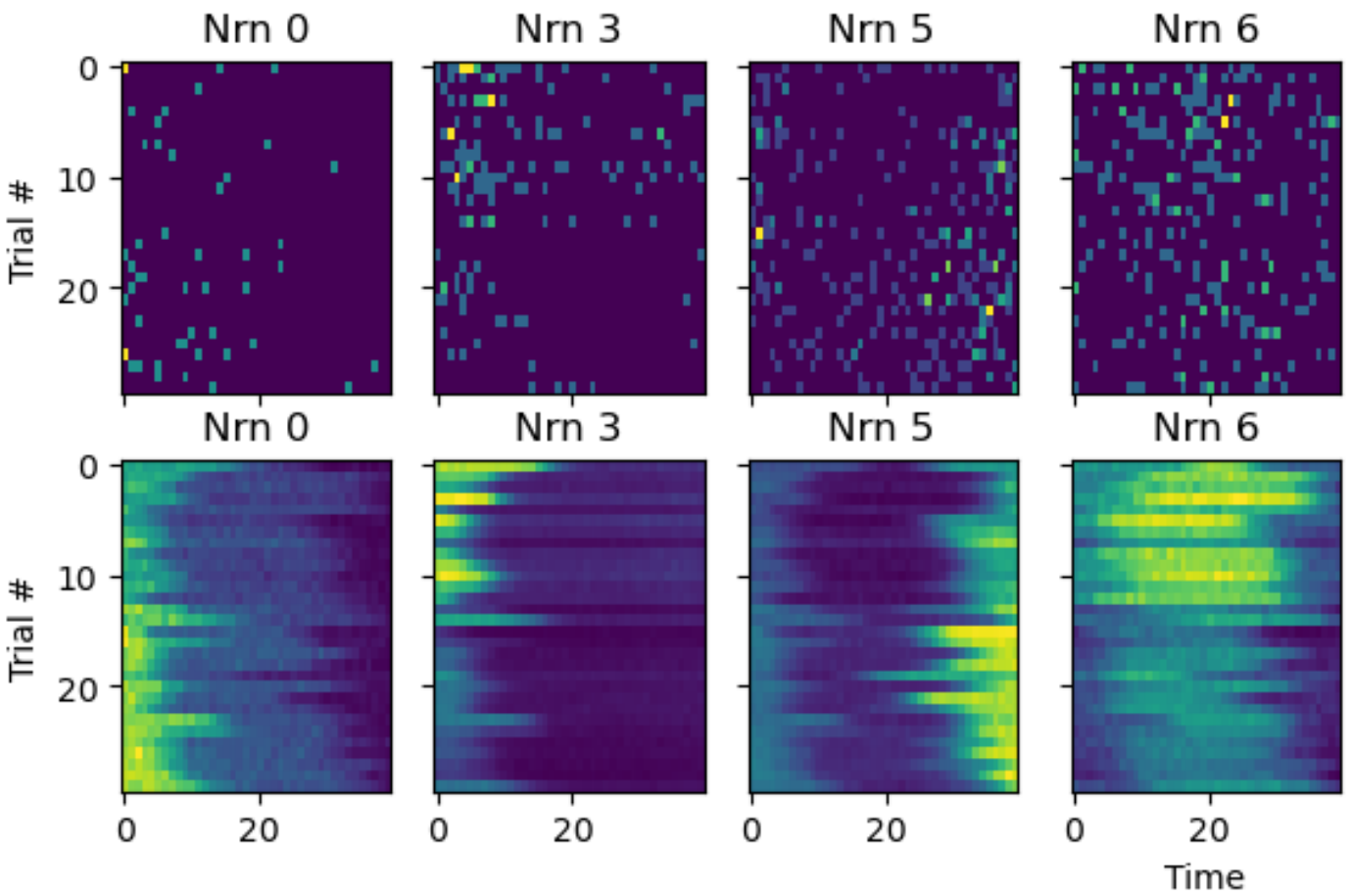




https://medium.com/@abuzar_mahmood/changepoint-detection-for-inhomogeneous-timeseries-mixture-emissions-in-pymc3-2301f76c0064

Mixture Changepoint





Wrap-up

1. Why changepoint models are useful
2. How they are used in the world
3. How PyMC allows easy iterative development of bespoke Bayesian models

Acknowledgements

PyMCon

SPONSORS | ODSC

- **Ravin Kumar**
- **Purna Chandra Mansingh**

NIH National Institutes of Health

THE SWARTZ FOUNDATION
www.theswartzfoundation.org

ACCESS | Advancing Innovation

- **Donald B. Katz**
- **Narendra Mukherjee**

$$P(\ Questions \mid Talk\ ) = \frac{P(Talk|\ Questions\ )\ P(Questions)}{P(Talk)}$$