Testing Poisson versus Poisson mixtures with application to neuroscience

Yunran Chen Surya Tapas Tokdar

Department of Statistical Science, Duke University
Durham, North Carolina, USA

AISC, 2018

Outline

Background
Non-Poisson behavior

- In neuroscience, spike counts are usually modeled as **Poisson distribution** for simplicity.
- **Non-Poisson behavior** is to be expected and has been documented under many situations. [1]
  - The stimuli or the internal state of the subject may **change over time and vary from trial to trial**.
  - "refractory period"
- **Poisson mixtures** attract increasing attention.
  - It can be seen as a generalized version of Poisson distribution.
  - It offers a rich class of alternatives to the Poisson distribution.

1. Kass R E, Ventura V, Brown E N. Statistical issues in the analysis of neuronal data[J]. Journal of neurophysiology, 2005, 94(1) : 8-25.

Background
Non-Poisson behavior

- In neuroscience, spike counts are usually modeled as **Poisson distribution** for simplicity.
- **Non-Poisson behavior** is to be expected and has been documented under many situations. [1]
  - The stimuli or the internal state of the subject may **change over time and vary from trial to trial**.
  - "refractory period"
- **Poisson mixtures** attract increasing attention.
  - It can be seen as a generalized version of Poisson distribution.
  - It offers a rich class of alternatives to the Poisson distribution.

1. Kass R E, Ventura V, Brown E N. Statistical issues in the analysis of neuronal data[J]. Journal of neurophysiology, 2005, 94(1) : 8-25.

Background

Unsuitable Poisson assumption

- **Testing Poisson versus Poisson mixtures**
  - Unsuitable model assumption may lead to distortion of inference.
  - Need to filter out non-Poisson behavior trials.
- Traditional testing procedure : χ^2 goodness of fit test
  - Whether χ^2 test can give us Poisson-like data ?
  - Is there better method for this ?
- Bayesian perspective : **Predictive recursion marginal likelihood (PRML) testing** [2]
  - Better performance as measured by ROC-AUC
  - Testing between different types of Poisson mixtures

2. Martin R, Tokdar S T. Semiparametric inference in mixture models with predictive recursion marginal likelihood[J]. Biometrika, 2011, 98(3) : 567-582.

Background

Unsuitable Poisson assumption

- **Testing Poisson versus Poisson mixtures**
  - Unsuitable model assumption may lead to distortion of inference.
  - Need to filter out non-Poisson behavior trials.
- Traditional testing procedure : χ^2 goodness of fit test
  - Whether χ^2 test can give us Poisson-like data ?
  - Is there better method for this ?
- Bayesian perspective : **Predictive recursion marginal likelihood (PRML) testing** [2]
  - Better performance as measured by ROC-AUC
  - Testing between different types of Poisson mixtures

2. Martin R, Tokdar S T. Semiparametric inference in mixture models with predictive recursion marginal likelihood[J]. Biometrika, 2011, 98(3) : 567-582.

Background

Unsuitable Poisson assumption

- **Testing Poisson versus Poisson mixtures**
  - Unsuitable model assumption may lead to distortion of inference.
  - Need to filter out non-Poisson behavior trials.
- Traditional testing procedure : χ^2 goodness of fit test
  - Whether χ^2 test can give us Poisson-like data ?
  - Is there better method for this ?
- Bayesian perspective : **Predictive recursion marginal likelihood (PRML) testing** [2]
  - Better performance as measured by ROC-AUC
  - Testing between different types of Poisson mixtures

2. Martin R, Tokdar S T. Semiparametric inference in mixture models with predictive recursion marginal likelihood[J]. Biometrika, 2011, 98(3) : 567-582.

Poisson and Poisson mixture

Poisson

$$Y_i \overset{i.i.d}{\sim} Poi(\mu), \mu \in (\mu_l, \mu_u)$$

Poisson mixture

$$Y_i \overset{i.i.d}{\sim} \int Poi(\mu) f(\mu) d\mu, support(f) = (\mu_l, \mu_u)$$

Or

$$Y_i \overset{i.i.d}{\sim} Poi(\mu_i), \mu_i \overset{i.i.d}{\sim} f$$

- A generalized version of Poisson distribution
- A rich class of alternatives to the Poisson distribution
- An overdispersion model

Hypothesis Testing

Consider the data Y_i for $i = 1, ..., n$,

- $H_0 : Y_i \overset{i.i.d}{\sim} Poi(\mu)$ for unknown $\mu \in (\mu_l, \mu_u)$
- $H_1 : Y_i \overset{i.i.d}{\sim} \int Poi(\mu) f(\mu) d\mu$ where $support(f) = (\mu_l, \mu_u)$

Methods

- Bayes Factor : $\frac{P(Y|M_0)}{P(Y|M_1)}$ - PRML algorithm
- p value : χ^2 goodness of fit test

Bayes Factor

Bayes Factor= $P(Y|M_0)/P(Y|M_1)$

Bayes' Factor : Ratio of marginal likelihood based on corresponding model assumption.

$$BF = \frac{P(Y|M_0)}{P(Y|M_1)}$$

The larger the Bayes' Factor, the stronger evidence showing Model 0 (Poisson) is better than Model 1 (Poisson mixture).

BF	Strength of evidence
1 to 3	not worth more than a bare mention
3 to 20	positive
20 to 150	strong
>150	very strong

Calculating marginal likelihood $P(Y|M_0)$, $P(Y|M_1)$

Bayes Factor

Bayes Factor$= P(Y|M_0)/P(Y|M_1)$

Bayes' Factor : Ratio of marginal likelihood based on corresponding model assumption.

$$BF = \frac{P(Y|M_0)}{P(Y|M_1)}$$

The larger the Bayes' Factor, the stronger evidence showing Model 0 (Poisson) is better than Model 1 (Poisson mixture).

BF	Strength of evidence
1 to 3	not worth more than a bare mention
3 to 20	positive
20 to 150	strong
$>$150	very strong

Calculating marginal likelihood $P(Y|M_0)$, $P(Y|M_1)$

Bayes Factor

Bayes Factor= $P(Y|M_0)/P(Y|M_1)$

Bayes' Factor : Ratio of marginal likelihood based on corresponding model assumption.

$$BF = \frac{P(Y|M_0)}{P(Y|M_1)}$$

The larger the Bayes' Factor, the stronger evidence showing Model 0 (Poisson) is better than Model 1 (Poisson mixture).

BF	Strength of evidence
1 to 3	not worth more than a bare mention
3 to 20	positive
20 to 150	strong
$>$150	very strong

Calculating marginal likelihood $P(Y|M_0), P(Y|M_1)$

Marginal likelihood approximation

$P(Y|M_0)$

- $H_0 : Y_i \overset{i.i.d}{\sim} Poi(\mu)$ for unknown $\mu \in (\mu_l, \mu_u)$
- Setting a prior and integrate out the parameter.
- If it is hard to get integral, we can apply Laplace approximation.

$P(Y|M_1)$

- $H_1 : Y_i \overset{i.i.d}{\sim} \int Poi(\mu)f(\mu)d\mu$ where $support(f) = (\mu_l, \mu_u)$
- Applying **predictive recursion marginal likelihood (PRML) algorithm.**

Marginal likelihood approximation

$P(Y|M_0)$

- $H_0 : Y_i \overset{i.i.d}{\sim} Poi(\mu)$ for unknown $\mu \in (\mu_l, \mu_u)$
- Setting a prior and integrate out the parameter.
- If it is hard to get integral, we can apply Laplace approximation.

$P(Y|M_1)$

- $H_1 : Y_i \overset{i.i.d}{\sim} \int Poi(\mu)f(\mu)d\mu$ where $support(f) = (\mu_l, \mu_u)$
- Applying **predictive recursion marginal likelihood (PRML) algorithm.**

PRML algorithm : Restate the problem

Estimate $P(Y|M_1)$

Calculate

$$p(Y|M_1) = \int p(Y|\mu)f(\mu)d\mu$$

Known :

- Likelihood Function : Poisson $p(Y|\mu)$
- support of $f(\cdot)$

Unknown :

- Mixture density $f(\mu)$

Mixture model density estimation : PRML

PRML algorithm

Estimate $P(Y|M_1)$

Calculate $p(Y|M_1) = \int p(Y|\mu)f(\mu)d\mu$
Predictive recursion (PR) is an accurate and computationally efficient algorithm for nonparametric estimation of mixing densities in mixture model.

Required information :

- $p(Y|\mu)$ known – Poisson ;
- support and continuity properties – Model assumptions. Pass the information via f_0 in initialization and $m_f(y)$ in integral.

Output

- Estimation on $f(\mu)$
- Estimation on marginal likelihood $p(Y|M_1)$

PRML Algorithm

Estimate $P(Y|M_1)$

$$p(Y|M_1) = \int p(Y|\mu)f(\mu)d\mu$$

Input : i.i.d observations $Y_1, ..., Y_n$
Output : $L = \prod_{i=1}^{n} m_i(y)$
Initialize : $f_0(\mu)$ – Usually uniformly distributed on the support.
$w_1, .., w_n \in (0, 1)$ – $w_i = \frac{1}{1+i}$; $\sum_{i=1}^{\infty} w_i = \infty, \sum_{i=1}^{\infty} w_i^2 < \infty$
For i = 1,...,n :

$$m_i(y) = \int p(Y_i|\mu)f_{i-1}(\mu)d\mu = \sum_{k=1}^{m} s_k p(Y_i|\mu_k)f_{i-1}(\mu_k)$$

$$f_i(\mu) = (1 - w_i)f_{i-1}(\mu) + w_i p(Y_i|\mu)f_{i-1}(\mu)/m_i(y)$$

PRML Algorithm

Estimate $P(Y|M_1)$

$$p(Y|M_1) = \int p(Y|\mu)f(\mu)d\mu$$

Input : i.i.d observations $Y_1, ..., Y_n$
Output : $L = \prod_{i=1}^{n} m_i(y)$
Initialize : $f_0(\mu)$– Usually uniformly distributed on the support.
$w_1, .., w_n \in (0,1)$ – $w_i = \frac{1}{1+i}$; $\sum_{i=1}^{\infty} w_i = \infty, \sum_{i=1}^{\infty} w_i^2 < \infty$
For i = 1,...,n :

$$m_i(y) = \int p(Y_i|\mu)f_{i-1}(\mu)d\mu = \sum_{k=1}^{m} s_k p(Y_i|\mu_k)f_{i-1}(\mu_k)$$

$$f_i(\mu) = (1 - w_i)f_{i-1}(\mu) + w_i p(Y_i|\mu)f_{i-1}(\mu)/m_i(y)$$

PRML Algorithm

Estimate $P(Y|M_1)$

$$p(Y|M_1) = \int p(Y|\mu)f(\mu)d\mu$$

Input : i.i.d observations $Y_1, ..., Y_n$
Output : $L = \prod_{i=1}^{n} m_i(y)$
Initialize : $f_0(\mu)$– Usually uniformly distributed on the support.
$w_1, .., w_n \in (0,1)$ – $w_i = \frac{1}{1+i}$; $\sum_{i=1}^{\infty} w_i = \infty, \sum_{i=1}^{\infty} w_i^2 < \infty$
For i = 1,...,n :

$$m_i(y) = \int p(Y_i|\mu)f_{i-1}(\mu)d\mu = \sum_{k=1}^{m} s_k p(Y_i|\mu_k)f_{i-1}(\mu_k)$$

$$f_i(\mu) = (1 - w_i)f_{i-1}(\mu) + w_i p(Y_i|\mu)f_{i-1}(\mu)/m_i(y)$$

PRML Algorithm : Permutation Version

Estimate $P(Y|M_1)$

- 1 dataset $\rightarrow$ 1 estimator L
- 1 dataset $\rightarrow$ shuffle $\rightarrow$ 10 datasets $\rightarrow$ 10 estimator $L_1, ..., L_{10}$ $\rightarrow$ average $\rightarrow$ L_p

Estimate $P(Y|M_0)$

Consider

$$p(Y|M_0) = \int p(Y|\mu)f(\mu)d\mu$$

- Setting a prior and integrate out the parameter.
  - Uniform prior : $Unif[\mu_l, \mu_u]$
  - $\int_{\mu_l}^{\mu_u} \prod_{i=1}^{n} dpoi(Y_i|\mu) \times \frac{1}{\mu_u - \mu_l} d\mu$
- For unknown μ_l, μ_u, use robust estimator
  - $\hat{\mu}_l = Y_{0.25} - \alpha \times IQR$
  - $\hat{\mu}_u = Y_{0.75} + \alpha \times IQR$
  - $IQR = Y_{0.75} - Y_{0.25}$
  - Simulation shows the performance of PRML is not sensitive to parameter α

Estimate $P(Y|M_0)$

Consider

$$p(Y|M_0) = \int p(Y|\mu)f(\mu)d\mu$$

- Setting a prior and integrate out the parameter.
  - Uniform prior : $Unif[\mu_l, \mu_u]$
  - $\int_{\mu_l}^{\mu_u} \prod_{i=1}^{n} dpoi(Y_i|\mu) \times \frac{1}{\mu_u - \mu_l} d\mu$
- For unknown μ_l, μ_u, use robust estimator
  - $\hat{\mu}_l = Y_{0.25} - \alpha \times IQR$
  - $\hat{\mu}_u = Y_{0.75} + \alpha \times IQR$
  - $IQR = Y_{0.75} - Y_{0.25}$
  - Simulation shows the performance of PRML is not sensitive to parameter α

Pearson χ^2 goodness of fit test

- $H_0 : Y_i \overset{i.i.d}{\sim} Poi(\mu)$ for $\mu \in (\mu_l, \mu_u)$
- $H_0 : Y_i \overset{i.i.d}{\sim} Poi(\hat{\mu})$
-
$$X = \sum_i \frac{(O_i - E_i)^2}{E_i}$$
- $X \sim \chi^2_{df}$
- approximate X using Monte Carlo p-value calculation

Simulation Result

Testing Poisson versus Poisson mixtures

Testing Poisson versus Poisson mixture

$$H_0 : Y_i \overset{i.i.d}{\sim} Poi(240)$$

$$H_1 : Y_i \overset{i.i.d}{\sim} \int_{150}^{300} Poi(\mu) gamma_{[150,300]}(\mu|480, 2) d\mu$$

- Generate datasets with size $N = 200$, half comes from Poisson and half comes from Poisson mixture.

Simulation Result

Testing Poisson versus Poisson mixtures

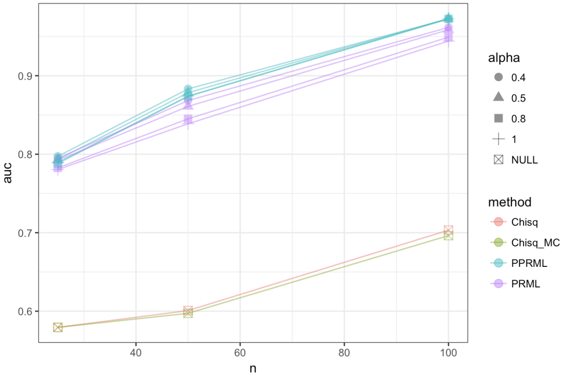

FIGURE – Plots of the AUC. x-axis indicates different sample size $n = 25, 50, 100$. Different colors indicate different methods. Different shapes of the point indicate different value for α.

Simulation Result

Testing Poisson versus Poisson mixtures

- PRML, PPRML testing perform much better than tradition χ^2 test
- As sample size increases, the performance improves.
- PPRML is much stable than PRML testing.

Comments

Traditional testing procedure based on p-value sets too general alternative hypothesis containing too large "model space", leading to a conservative decision, or we say a loss of power (or sensitivity).

Simulation Result

Poisson versus Poisson mixed with normal

Testing Poisson versus Poisson mixed with normal

$$H_0 : Y_i \overset{i.i.d}{\sim} Poi(240)$$

$$H_1 : Y_i \overset{i.i.d}{\sim} 0.9Poi(240) + 0.1N_{[0,\infty)}(240, \sigma^2)$$

- Generate datasets with size $N = 200$, half comes from Poisson and half comes from Poisson mixed with normal.

Simulation Result

Poisson versus Poisson mixed with normal

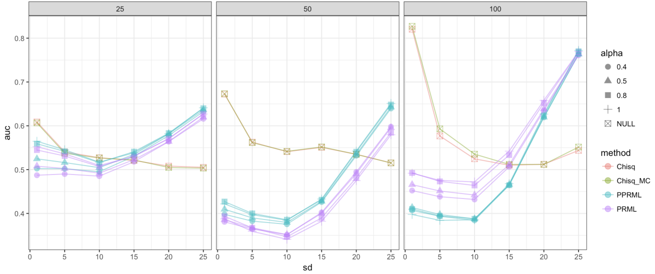

FIGURE – Plots of the AUC. Different panels indicate different sample size *n*. Different colors indicate different methods. Different shapes of the point indicate different value for α.

Simulation Result

Poisson versus Poisson mixed with normal

- For $\sigma \leq 15$, PRML, PPRML testing perform much worse than Pearson χ^2 testing. For $\sigma > 15$, PRML, PPRML testing perform better than Pearson χ^2 testing.
- As σ increase, the performances of PRML, PPRML testing improve.
- $\sigma < \sqrt{240} (\approx 15.5)$ – underdispersion model

Comments

When the alternative model is mis-specified (underdispersion model), PRML, PPRML testing on Poisson versus Poisson mixtures is not applicable.

Hypothesis Testing

- $Y_j^A \stackrel{i.i.d}{\sim} Poi(\mu^A)$, $Y_j^B \stackrel{i.i.d}{\sim} Poi(\mu^B)$ for unknown μ^A, μ^B
- $Y_j^{AB} \stackrel{i.i.d}{\sim} \int Poi(\mu) f(\mu) d\mu$ with four competing scenarios for the support of f :
  - M_1(Mixture) : for unknown $\mu \in \{\mu^A, \mu^B\}$
  - M_2(Intermediate) : for unknown $\mu \in (min(\mu^A, \mu^B), max(\mu^A, \mu^B))$
  - M_3(Outside) : for unknown $\mu \in [\mu_l, min(\mu^A, \mu^B))$ or $\mu \in (max(\mu^A, \mu^B), \mu_u]$, where known μ_l, μ_u indicate the lower bound and upper bound of μ respectively.
  - M_4(Single) : for $\mu = \mu^A$ or $\mu = \mu^B$
- Our goal : Choosing the best fit model.
- $$BF_{12} = \frac{P(Y^{AB}|M_i)}{P(Y^{AB}|M_j)}$$

Four competitive scenarios

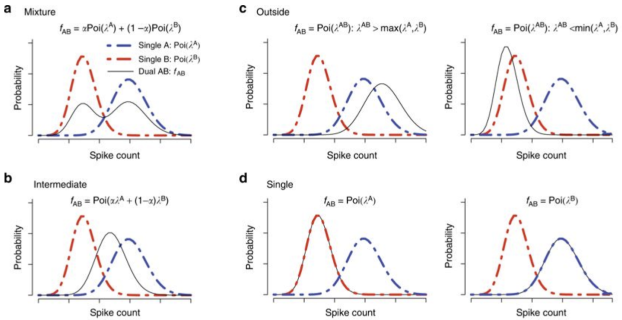

FIGURE – Four possible types of Poisson mixtures which spike counts may exhibit [3]

3. Caruso V C, Mohl J T, Glynn C, et al. Single neurons may encode simultaneous stimuli by switching between activity patterns[J]. Nature communications, 2018, 9(1) : 2715.

Known and Unknown

Known :

- Triplets Data : Y^A, Y^B, Y^{AB}
- Likelihood Function : Poisson
  $p(Y^A|\mu^A), p(Y^B|\mu^B), p(Y^{AB}|\mu^{AB})$
- Relationship on support between μ^A, μ^B, μ^{AB}
  - Mixture : $\mu^{AB} \in \{\mu^A, \mu^B\}$
  - Intermediate : $\mu^{AB} \in (\mu_{min}, \mu_{max})$
  - Outside : $\mu^{AB} \in [\mu_l, \mu_{min})$ or $\mu^{AB} \in (\mu_{max}, \mu_u]$
  - Single $\mu^{AB} = \mu^A$ or $\mu^{AB} = \mu^B$

Unknown :

- Parameters : μ^A, μ^B, μ^{AB}
- Mixture density $f(\mu^{AB}|\mu^A, \mu^B)$

Goal : Marginal likelihood

$$p(Y^{AB}|M, Y^A, Y^B) = \int_\Theta p(Y^{AB}|\theta, M)p(\theta|Y^A, Y^B)d\theta$$

Estimate $P(Y|M_i)$: Laplace Approximation

$$\begin{aligned} p(Y^{AB}|M, Y^A, Y^B) &= \int_\Theta p(Y^{AB}|\theta, M)p(\theta|Y^A, Y^B)d\theta \\ &= \int_\Theta \int p(Y^{AB}|\mu^{AB})f(\mu^{AB}|\theta)d\mu^{AB}p(\theta|Y^A, Y^B)d\theta \\ &\approx (2\pi)^{k/2}|\hat{\Sigma}|^{1/2}p(Y^{AB}|M, \hat{\theta})p(\hat{\theta}|Y^A, Y^B) \end{aligned}$$

With Laplace approximation, we have

$$p(Y^{AB}|M, Y^A, Y^B) \approx \frac{p(Y^{AB}|M, \hat{\theta})p(\hat{\theta}|Y^A, Y^B)}{N(\hat{\theta}|\hat{\theta}, \hat{\Sigma})}$$

where $k = dim(\theta)$

$$\hat{\theta} = \underset{\theta}{argmax} \, log p(Y^{AB}|M, \theta)p(\theta|Y^A, Y^B)$$

$$\hat{\Sigma} = \{-\nabla^2 log p(Y^{AB}|M, \theta)p(\theta|Y^A, Y^B)|_{\theta=\hat{\theta}}\}^{-1}$$

Estimate $P(Y|M_i)$: optimization problem

Object function : $l(\theta) = logp(Y^{AB}|M,\theta)p(\theta|Y^A, Y^B)$

Hessian Matrix : $H = \nabla^2 l(\theta)$

Marginal likelihood estimator :

$$p(Y^{AB}|M, Y^A, Y^B) \approx (2\pi)^{k/2}|-H|^{1/2}e^{l(\hat{\theta})}$$

- If we provide the gradient $\nabla l(\mu^A, \mu^B)$, the computation could be eased a lot.
- PRML gradient algorithm (PRMLG) : calculate gradient in each recursion without significant computation increase.

Estimate $P(Y|M_i)$:PRMLG algorithm

$p(Y^{AB}|\mu^A,\mu^B,M) = \int p(Y^{AB}|\mu^{AB})f(\mu^{AB}|\mu^A,\mu^B)d\mu^{AB}$

Input : i.i.d observations $Y_1, ..., Y_n$

Output : $logL = \sum_{i=1}^n logm_i(y)$, $\nabla logL = \sum_{i=1}^n \nabla logm_{i,\theta}(Y_i)$

Initialize : $f_0(\mu^{AB}|\mu^A,\mu^B)$, $\nabla f_{0,\theta}$, weights $w_1, .., w_n \in (0,1)$

For i = 1,...,n :

$$m_i(y) = \int p(Y_i^{AB}|\mu^{AB})f_{i-1}(\mu^{AB})d\mu^{AB}$$

$$f_i(\mu^{AB}) = (1-w_i)f_{i-1}(\mu^{AB}) + w_i p(Y_i^{AB}|\mu^{AB})f_{i-1}(\mu^{AB})/m_i(y)$$

$$\nabla logm_{i,\theta}(Y_i) = \int G(\theta,u)d\mu(u)/m_{i,\theta}(Y_i)$$

$$\nabla f_{i,\theta}(u) = (1-w_i)\nabla f_{i-1,\theta}(u) + w_i \frac{G(\theta,u) - p(Y_i|\theta,u)f_{i-1,\theta}(u)\nabla logm_{i,\theta}(}{m_{i,\theta}(Y_i)}$$

Estimate $P(Y|M_i)$: specify $p(Y^{AB}|M,\theta)$ – PRML

support and continuity properties – Model assumptions.

Model	Support	f_0	$m_f(y)$
Mixture	$\{\mu^A, \mu^B\}$	$(0.5, 0.5)$	$\sum_{A,B} p(Y_i^{AB}\|\mu^{AB}) f_{i-1}(\mu')$
Intermediate	(μ_{min}, μ_{max})	$Unif(\mu_{min}, \mu_{max})$	$\int_{\mu_{min}}^{\mu_{max}} p(Y_i^{AB}\|\mu') f_{i-1}(\mu') d\mu'$
OutsideA	(μ_l, μ_{min})	$Unif(\mu_l, \mu_{min})$	$\int_{\mu_l}^{\mu_{min}} p(Y_i^{AB}\|\mu') f_{i-1}(\mu') d\mu'$
OutsideB	(μ_{max}, μ_u)	$Unif(\mu_{min}, \mu_u)$	$\int_{\mu_{max}}^{\mu_u} p(Y_i^{AB}\|\mu') f_{i-1}(\mu') d\mu'$
SingleA	$\{\mu^A\}$	1	$p(Y_i^{AB}\|\mu^{AB}) f_{i-1}(\mu^A)$
SingleB	$\{\mu^B\}$	1	$p(Y_i^{AB}\|\mu^{AB}) f_{i-1}(\mu^B)$

TABLE – PRML setting under different model assumptions

Estimate $P(Y|M_i)$: Reparameterization

Model	Reparameterization	Support
Mixture	$\mu^{AB} = h(z) = \mu_{min} + z(\mu_{max} - \mu_{min})$	$\{0, 1\}$
Intermediate	$\mu^{AB} = h(z) = \mu_{min} + z(\mu_{max} - \mu_{min})$	$[0, 1]$
OutsideA	$\mu^{AB} = h(z) = \mu_l + z(\mu^A - \mu_l)$	$[0, 1]$
OutsideB	$\mu^{AB} = h(z) = \mu^B + z(\mu_u - \mu_B)$	$[0, 1]$

TABLE – Reparameterization for PRMLG algorithm

Model	Restriction	Reparameterization
Mixture	$\mu^A, \mu^B > 0$	$\theta = (log(\mu^A), log(\mu^B))$
Intermediate	$0 < \mu_{min} < \mu_{max}$	$\theta = (log(\mu_{min}), log(\mu_{max}))$
OutsideA	$0 < \mu_l < \mu^A$	$\theta = log(\mu^A - \mu_l)$
OutsideB	$0 < \mu^B < \mu_u$	$\theta = logit(\frac{\mu^B}{\mu_u})$
SingleA	$\mu^A > 0$	$\theta = log(\mu^A)$
SingleB	$\mu^B > 0$	$\theta = log(\mu^B)$

TABLE – Reparameterization for optimization

Simulation Setting

Generate $N = 100$ samples for each model. Set $\mu_A = 150$, $\mu_B = 300$, sample size $n = 25, 50$.

- Generate $Y^A \sim Poi(\mu^A)$ with $n^A = 1.5n$; $Y^B \sim Poi(\mu^B)$ with $n^B = 1.2n$.
- Generate Y^{AB} with sample size n.
  - Mixture : $Y^{AB} \sim \alpha Poi(\mu_A) + (1-\alpha)Poi(\mu_B)$ with $\alpha = 0.5$;
  - Intermediate : $Y^{AB} \sim \int Poi(\mu)Ga_{[180,270]}(\mu|144, 0.6)d\mu$;
  - Outside B : generate $Y^{AB} \sim Poi(400)$;
  - Outside A : generate $Y^{AB} \sim \int Poi(\mu)Ga_{[30,120]}(\mu|20.25, 0.225)d\mu$.
  - Single A : generate $Y^{AB} \sim Poi(150)$;
  - Single B : generate $Y^{AB} \sim Poi(300)$.

For estimation, consider $\mu_l = 30$, $\mu_u = 600$, $nGQ = 20$, $nP = 100$. For PRML-LP, set conjugate prior $r_A = 15$, $s_A = 0.1$, $r_B = 30$, $s_B = 0.1$; initial value $\mu_0^A = 120$, $\mu_0^B = 330$.

Testing between four Poisson mixtures

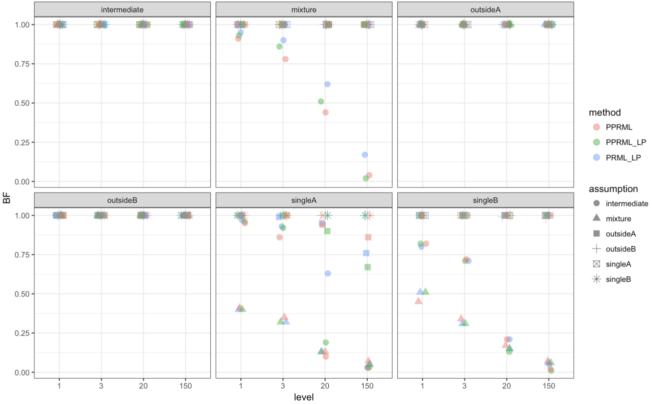

FIGURE – Bayes Factor with sample size $n = 25$.

Testing between four Poisson mixtures

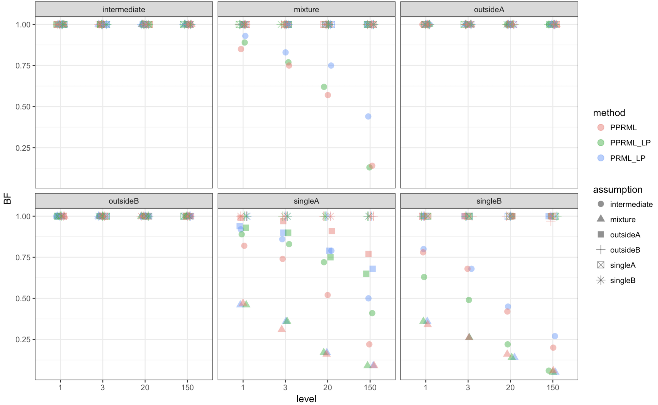

FIGURE – Bayes Factor with sample size $n = 50$.

Estimate $P(Y|M_i)$: Introduce parameter e

Model	Reparameterization	Support
Mixture	$\mu^{AB} = h(z) = \mu_{min} + z(\mu_{max} - \mu_{min})$	$\{0, 1\}$
Intermediate	$\mu^{AB} = h(z) = \mu_{min} + z(\mu_{max} - \mu_{min})$	$[0, 1]$

TABLE – Reparameterization for PRMLG algorithm under $e = 0$

Model	Reparameterization	Support
Mixture	$\mu^{AB} = h(z) = \mu_{min} + z(\mu_{max} - \mu_{min})$	$\{e, 1 - e\}$
Intermediate	$\mu^{AB} = h(z) = \mu_{min} + z(\mu_{max} - \mu_{min})$	$[e, 1 - e]$

TABLE – Reparameterization for PRMLG algorithm under e

Effect of parameter e

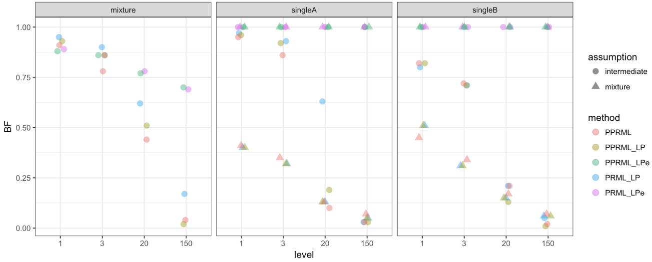

FIGURE – Bayes Factor sample size $n = 25$

Effect of parameter e

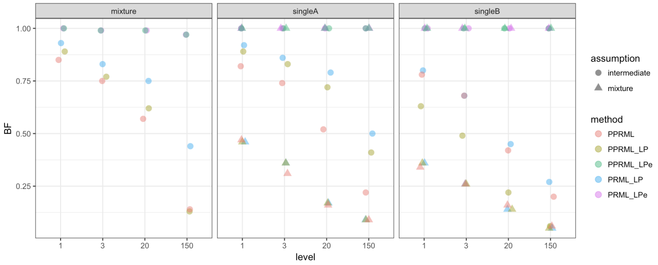

FIGURE – Bayes Factor sample size $n = 50$

Simulation result

- As **sample size** *n* **increase**, the performances of our proposed testing would get **improved.**
- PRML-LP and PPRML-LP perform **as well as the benchmark method.** They can identify these four Poisson mixtures **with strong evidence** except for distinguishing mixture, intermediate and single.
- This is due to the **specification of the continuity** of the domain of μ.
- With introduce parameter *e* to define mixture and single more clearly, our proposed testing (PRML-LP and PPRML-LP) can distinguish single, mixture and intermediate with **strong evidence** (especially when sample size is large).

Conclusion

Pros

- Testing Poisson versus Poisson mixture
- Testing between different Poisson mixtures

Cons

- Misspecification of model
- Choice of w_i remains openinng question
- Normality assumption

References I

Martin, Ryan and Tokdar, Surya T,
Semiparametric inference in mixture models with predictive recursion marginal likelihood.
Biometrika 98.3 (2011) : 567-582.

Caruso V C, Mohl J T, Glynn C, et al.
Single neurons may encode simultaneous stimuli by switching between activity patterns[J].
Nature communications, 2018, 9(1) : 2715.

Kass R E, Ventura V, Brown E N.
Statistical issues in the analysis of neuronal data[J].
Journal of neurophysiology, 2005, 94(1) : 8-25.