A Bayesian Comparative Analysis of Neuronal Point Processes

Sam Behseta
California State University Fullerton

June 7, 2012
Motivation: Action Potential

A Bayesian Comparative Analysis of Neuronal Point Processes
Raster Plot and Peri- Stimulus Time Histograms (PSTH)
Example 1: Two Conditions-

Random Mode

Response 4 → Target 2 on → Response 2

Repeating Mode (5-3-1)

Response 5 → Response 3 OR Target 3 on ...

Raster and PSTH plots for a neuron under repeating (left panel) and random (right panel) modes.
Example 2: Three Conditions-
Moorman and Olson (2012)

- Pre-Mark Delay: 400 msec
- Pre-Cue Delay: 500 msec
- Spatial Cue: 100 msec
- Post-Cue Delay: 800-900 msec
- Go Signal (Fix Spot Off)
- Saccade and Pre-Reward Fixation (200-300 msec)

A. Space
B. Ring
C. Dot

= fixation point
= current direction of gaze
= saccade

Sam Behseta  California State University Fullerton
A Bayesian Comparative Analysis of Neuronal Point Processes
Example 2: Three Conditions-
Moorman and Olson (2012)
A Bayesian Comparative Analysis of Neuronal Point Processes
Estimating The Firing Rate $f(t)$ With $\hat{f}(t)$
Multiple Curve-Fitting: Hierarchical Modeling of Firing Intensity Curves Using BARS
Part I: Bayesian Functional Data Analysis

- Multiple Curve-Fitting: Hierarchical Modeling of Firing Intensity Curves Using BARS
- Single Neuronal Analysis: Testing Equality of Two or More Curves
Part I: Bayesian Functional Data Analysis

- Multiple Curve-Fitting: Hierarchical Modeling of Firing Intensity Curves Using BARS
- Single Neuronal Analysis: Testing Equality of Two or More Curves
- Population-level Analysis: Testing Equality of Two Groups of Curves
Objects of Statistical Inference

Sam Behseta  California State University Fullerton  A Bayesian Comparative Analysis of Neuronal Point Processes
(Y_i|\beta, \xi_i, k_i, \lambda_i) \sim \text{Poisson}(\lambda_i)

\log \lambda(t_i) = \sum_{j=1}^{k+2} b_j(t_i) \times \beta_j

BARS (DiMatteo, Genovese and Kass, 2001): Free Knots, Fully Bayesian, Reversible Jump MCMC
Fitting BARS *simultaneously* by keeping same knots for all neurons, and forming hierarchical models
Fitting BARS **simultaneously** by keeping same knots for all neurons, and forming hierarchical models

Fitting BARS **separately**, and forming hierarchical models
Simultaneous Curve Fitting: Simultaneous-BARS

- Same \((\xi, k)\) for all curves
Simultaneous Curve Fitting: Simultaneous-BARS

- Same \((\xi, k)\) for all curves
- Random Coefficients Model:

\[
(f^i(u_1), ..., f^i(u_n))^T = X_\xi \beta^i_\xi
\]

\[
\beta^i_\xi | \xi, \alpha_\xi, D_\xi \sim N(\alpha_\xi, D_\xi)
\]
Simultaneous Curve Fitting: Simultaneous-BARS

- Same \((\xi, k)\) for all curves
- Random Coefficients Model:
  \[
  (f^i(u_1), ..., f^i(u_n))^T = X_\xi \beta^i_\xi
  \]
  \[
  \beta^i_\xi|\xi, \alpha_\xi, D_\xi \sim N(\alpha_\xi, D_\xi)
  \]

- We replace the first stage with the MLE and write:
  \[
  \hat{\beta}^i_\xi|\xi, k, \beta^i_\xi, R^i \sim N(\beta^i_\xi, R^i)
  \]
  \[
  \beta^i_\xi|\xi, k, \alpha_\xi, D_\xi \sim N(\alpha_\xi, D_\xi).
  \]
Simultaneous Curve Fitting: Simultaneous-BARS

- Same \((\xi, k)\) for all curves
- \textbf{Random Coefficients Model:}

\[
(f^i(u_1), ..., f^i(u_n))^T = X_\xi \beta^i_\xi
\]

\[
\beta^i_\xi | \xi, \alpha_\xi, D_\xi \sim N(\alpha_\xi, D_\xi)
\]

- We replace the first stage with the MLE and write:

\[
\hat{\beta}^i_\xi | \xi, k, \beta^i_\xi, R^i \sim N(\beta^i_\xi, R^i)
\]

\[
\beta^i_\xi | \xi, k, \alpha_\xi, D_\xi \sim N(\alpha_\xi, D_\xi).
\]

- Approximation has the accuracy of order \(O(n^{-1})\)
Fits are obtained separately
Independent Curve Fitting:
Hierarchical Gaussian Process Model (HGP)

- Fits are obtained separately
- For a time grid \((t_1, \ldots, t_p)\),

\[
\begin{aligned}
(\hat{f}_i(t_1), \ldots, \hat{f}_i(t_p)) &\sim N_p((f_i(t_1), \ldots, f_i(t_p)), S^i) \\
(f_i(t_1), \ldots, f_i(t_p)) &\sim N_p(\mu, V)
\end{aligned}
\]
Independent Curve Fitting:
Hierarchical Gaussian Process Model (HGP)

- Fits are obtained separately
- For a time grid \((t_1, \ldots, t_p)\),

\[
\begin{align*}
\left\{ \begin{array}{l}
(\hat{f}^i(t_1), \ldots, \hat{f}^i(t_p)) \sim N_p((f^i(t_1), \ldots, f^i(t_p)), S^i) \\
(f^i(t_1), \ldots, f^i(t_p)) \sim N_p(\mu, V)
\end{array} \right.
\end{align*}
\]

- Can be applied at any time resolution. So will consider \(f^i(t)\) as realizations of a Gaussian process.
Fits are obtained separately

For a time grid \((t_1, \ldots, t_p)\),

\[
\begin{align*}
(\hat{f}^i(t_1), \ldots, \hat{f}^i(t_p)) &\sim N_p((f^i(t_1), \ldots, f^i(t_p)), S^i) \\
(f^i(t_1), \ldots, f^i(t_p)) &\sim N_p(\mu, V)
\end{align*}
\]

Can be applied at any time resolution. So will consider \(f^i(t)\) as realizations of a Gaussian process.

This is general (not necessarily Bayesian) and direct.
Idea:

- Each Estimated Curve is a Gaussian Process:

\[ \hat{f}^i \sim GP(f^i, \Gamma_{\hat{f}_i}). \]
Multiple Curve Fitting: General Framework

Idea:

- Each Estimated Curve is a Gaussian Process:
  \[ \hat{f}^i \sim GP(f^i, \Gamma_{\hat{f}^i}). \]

- Underlying Functions are Gaussian Processes:
  \[ f^i \sim GP(\alpha, \Gamma_f). \]
Example: Variability Due to Curve Estimation

- Usual Method (Optican and Richmond 1997; Ramsey and Silverman, 1997): sample covariance between the estimated firing rates $\hat{f}^i(t)$

- First Proportion of Variance (first eigenvalue divided by sum of all eigenvalues) is substantially biased upward when usual FDA is applied

$$\text{Variance}_{\text{total}} = \text{Variance}_{\text{between}} + \text{Variance}_{\text{within}}$$
Example: Variability Due to Curve Estimation

- Usual Method (Optican and Richmond 1997; Ramsey and Silverman, 1997): sample covariance between the estimated firing rates $\hat{f}^i(t)$
- Our approach: Obtain the posterior distribution of the covariance among $f^i(t)$

First Proportion of Variance (first eigenvalue divided by sum of all eigenvalues) is substantially biased upward when usual FDA is applied

Variance total = Variance between + Variance within
Example: Variability Due to Curve Estimation

- Usual Method (Optican and Richmond 1997; Ramsey and Silverman, 1997): sample covariance between the estimated firing rates \( \hat{f}_i(t) \)

- Our approach: Obtain the posterior distribution of the covariance among \( f_i(t) \)

- First Proportion of Variance (first eigenvalue divided by sum of all eigenvalues) is substantially biased upward when usual FDA is applied
Example: Variability Due to Curve Estimation

- Usual Method (Optican and Richmond 1997; Ramsey and Silverman, 1997): sample covariance between the estimated firing rates $\hat{f}_i(t)$
- Our approach: Obtain the posterior distribution of the covariance among $f_i(t)$
- First Proportion of Variance (first eigenvalue divided by sum of all eigenvalues) is substantially biased upward when usual FDA is applied
- $\text{Variance}_{total} = \text{Variance}_{between} + \text{Variance}_{within}$
Data Analysis
Part II. Single Neuronal Analysis: Testing Equality of Two Functions Using BARS

- Want to test the hypothesis:
  \[ H_0 : f^1(t) = f^2(t) \]
  where \( f^i(t) \) are two functions.
Want to test the hypothesis:

\[ H_0 : f^1(t) = f^2(t) \]

where \( f^i(t) \) are two functions.

Of particular interest:

\[ H_0 : \lambda^1(t) = \lambda^2(t) \]

where \( \lambda^i(t) \) are Poisson process intensity function (neurophysiological applications).
Main Application:

i) Neurons in both conditions are not responsive to the stimulus or task.

ii) Screening: to eliminate those neurons that behave non-differentially.

Example: Repeating vs. Random conditions.
Main Application:
  i) Neurons in both conditions are not responsive to the stimulus or task.
  ii) Screening: to eliminate those neurons that behave non-differentially.
Example: Repeating vs. Random conditions.

Substantial literature on the general problem:
*Fan and Lin (1998)*
Limitations: pre-processing, equispaced time points.
*Neuymeyer and Dette (2003).*
Case I: Simultaneous Curve-Fitting:

i) Fit curves simultaneously with BARS and re-express the hypothesis as:

ii) \( H_0 : \beta_{\xi}^1 = \beta_{\xi}^2. \)

iii) Write Bayes the factor:

\[
B = \frac{\int p(y^1|\beta_{\xi}, \xi)p(y^2|\beta_{\xi}, \xi)\pi(\beta_{\xi}, \xi)d\beta d\xi}{\int p(y^1|\beta_{\xi}^1, \xi)p(y^2|\beta_{\xi}^2, \xi)\pi(\beta_{\xi}^1, \beta_{\xi}^2, \xi)d\beta_{\xi}^1d\beta_{\xi}^2d\xi}
\]
Case I: Simultaneous Curve-Fitting:

i) Fit curves simultaneously with BARS and re-express the hypothesis as:

\( H_0 : \beta_1^\xi = \beta_2^\xi. \)

ii) Write Bayes's theorem:

\[
B = \frac{\int p(y_1|\beta, \xi) p(y_2|\beta, \xi) \pi(\beta, \xi) d\beta d\xi}{\int p(y_1|\beta_1^\xi, \xi) p(y_2|\beta_2^\xi, \xi) \pi(\beta_1^\xi, \beta_2^\xi, \xi) d\beta_1^\xi d\beta_2^\xi d\xi}
\]

The theory in Kass and Wasserman (1995) applies and BIC may be used to approximate the posterior probability

\( P(\beta_1^\xi = \beta_2^\xi | \xi(g), y^1, y^2) \).
Case II: Fit curves separately

Consider a grid of points: \( \tilde{t}_1, \tilde{t}_2, \ldots, \tilde{t}_p \)

Evaluate functions on the grid:

\( U = (f(\tilde{t}_1), f(\tilde{t}_2), \ldots, f(\tilde{t}_p)) \)

Posterior means obtained by BARS: \( \hat{U} \)

We approximate \( \hat{U}_j \sim N(U_j, \Sigma_j) \), for \( j = 1, 2 \)

Null hypothesis becomes:

\( H_0: U_1 = U_2 \)

Modify \( T_2 \).

Suppose there are \( k \) positive eigenvalues and write

\[ T^k_2 = (\hat{U}_1 - \hat{U}_2)^T \Lambda_k^{-1} \Lambda_k \frac{1}{k} \hat{U}_1 \hat{U}_2 \]

\( T^k_2 \sim \chi^2_k \).
Testing Equality of Two Functions Using BARS

- Case II: Fit curves separately
  - Gaussian Process Test

Consider a grid of points: \( \tilde{t}_1, \tilde{t}_2, \ldots, \tilde{t}_p \)

Evaluate functions on the grid: \( U = (f(\tilde{t}_1), f(\tilde{t}_2), \ldots, f(\tilde{t}_p)) \).

Posterior means obtained by BARS: \( \hat{U} \).

We approximate \( \hat{U}_j \sim N(U_j, \Sigma_j) \), for \( j = 1, 2 \).

Null hypothesis becomes: \( H_0: U_1 = U_2 \)

Modify \( T_2 \).

Suppose there are \( k \) positive eigenvalues and write

\[
T_2^k = (\hat{U}_1 - \hat{U}_2)^T P_k \Lambda_k^{-1} P_k (\hat{U}_1 - \hat{U}_2)
\]

\( T_2^k \sim \chi^2_k \).
Case II: Fit curves separately

Gaussian Process Test

Consider a grid of points: $\tilde{t}_1, \tilde{t}_2, \ldots, \tilde{t}_p$

Evaluate functions on the grid: $\mathbf{U} = (f(\tilde{t}_1), f(\tilde{t}_2), \ldots, f(\tilde{t}_p))$.

Posterior means obtained by BARS: $\hat{\mathbf{U}}$.

We approximate $\hat{U}_j \sim \mathcal{N}(U_j, \Sigma_j)$, for $j = 1, 2$.

Null hypothesis becomes: $H_0: U_1 = U_2$

Modify $T_2$.

Suppose there are $k$ positive eigenvalues and write $T_2^k = (\hat{U}_1 - \hat{U}_2) \mathbf{P} \mathbf{\Lambda}^{-1} \mathbf{P}^T (\hat{U}_1 - \hat{U}_2)$. $T_2^k \sim \chi^2_k$. 
Case II: Fit curves separately

Gaussian Process Test

Consider a grid of points: \( \tilde{t}_1, \tilde{t}_2, \ldots, \tilde{t}_p \)

Evaluate functions on the grid: \( U = (f(\tilde{t}_1), f(\tilde{t}_2), \ldots, f(\tilde{t}_p)) \).
Case II: Fit curves separately
- Gaussian Process Test
- Consider a grid of points: $\tilde{t}_1, \tilde{t}_2, \ldots, \tilde{t}_p$
- Evaluate functions on the grid: $U = (f(\tilde{t}_1), f(\tilde{t}_2), \ldots, f(\tilde{t}_p))$.
- Posterior means obtained by BARS: $\hat{U}$. 
Case II: Fit curves separately

Gaussian Process Test

Consider a grid of points: \( \tilde{t}_1, \tilde{t}_2, \ldots, \tilde{t}_p \)

Evaluate functions on the grid: \( U = (f(\tilde{t}_1), f(\tilde{t}_2), \ldots, f(\tilde{t}_p)) \).

Posterior means obtained by BARS: \( \hat{U} \).

We approximate \( \hat{U}^j \sim N(U^j, \Sigma^j) \), for \( j = 1, 2 \).
Case II: Fit curves separately

Gaussian Process Test

Consider a grid of points: $\tilde{t}_1, \tilde{t}_2, \ldots, \tilde{t}_p$

Evaluate functions on the grid: $U = (f(\tilde{t}_1), f(\tilde{t}_2), \ldots, f(\tilde{t}_p))$.

Posterior means obtained by BARS: $\hat{U}$.

We approximate $\hat{U}^j \sim N(U^j, \Sigma^j)$, for $j = 1, 2$.

Null hypothesis becomes:

$H_0 : U^1 = U^2$
Testing Equality of Two Functions Using BARS

- **Case II: Fit curves separately**
- Gaussian Process Test
- Consider a grid of points: \( \tilde{t}_1, \tilde{t}_2, \ldots, \tilde{t}_p \)
- Evaluate functions on the grid: \( U = (f(\tilde{t}_1), f(\tilde{t}_2), \ldots, f(\tilde{t}_p)) \).
- Posterior means obtained by BARS: \( \hat{U} \).
- We approximate \( \hat{U}^j \sim N(U^j, \Sigma^j) \), for \( j = 1, 2 \).
- Null hypothesis becomes:
  \[ H_0 : U^1 = U^2 \]
- Modify \( T^2 \).
Case II: Fit curves separately

Gaussian Process Test

Consider a grid of points: $\tilde{t}_1, \tilde{t}_2, \ldots, \tilde{t}_p$

Evaluate functions on the grid: $U = (f(\tilde{t}_1), f(\tilde{t}_2), \ldots, f(\tilde{t}_p))$.

Posterior means obtained by BARS: $\hat{U}$.

We approximate $\hat{U}^j \sim N(U^j, \Sigma^j)$, for $j = 1, 2$.

Null hypothesis becomes:

$H_0 : U^1 = U^2$

Modify $T^2$.

Suppose there are $k$ positive eigenvalues and write

$T_k^2 = (\hat{U}^1 - \hat{U}^2)^T P_k^T \Lambda_k^{-1} P_k (\hat{U}^1 - \hat{U}^2)$. 
Case II: Fit curves separately

Gaussian Process Test

Consider a grid of points: \( \tilde{t}_1, \tilde{t}_2, \ldots, \tilde{t}_p \)

Evaluate functions on the grid: \( U = (f(\tilde{t}_1), f(\tilde{t}_2), \ldots, f(\tilde{t}_p)) \).

Posterior means obtained by BARS: \( \hat{U} \).

We approximate \( \hat{U}^j \sim N(U^j, \Sigma^j) \), for \( j = 1, 2 \).

Null hypothesis becomes:
\[ H_0 : U^1 = U^2 \]

Modify \( T^2 \).

Suppose there are \( k \) positive eigenvalues and write
\[ T_k^2 = (\hat{U}^1 - \hat{U}^2)^T P_k^T \Lambda_k^{-1} P_k (\hat{U}^1 - \hat{U}^2). \]

\[ T_k^2 \sim \chi_k^2 \]
Want to test the hypothesis:
\[ H_0 : f^1(t) = f^2(t) = \ldots = f^k(t) \]
where \( f^i(t) \) are two functions.
Want to test the hypothesis:
\[ H_0 : f^1(t) = f^2(t) = \ldots = f^k(t) \]
where \( f^i(t) \) are two functions.

Of particular interest:
\[ H_0 : \lambda^1(t) = \lambda^2(t) = \ldots = \lambda^k(t) \]
where \( \lambda^i(t) \) are Poisson process intensity function
Want to test the hypothesis:
\[ H_0 : f^1(t) = f^2(t) = \ldots = f^k(t) \]
where \( f^i(t) \) are two functions.

Of particular interest:
\[ H_0 : \lambda^1(t) = \lambda^2(t) = \ldots = \lambda^k(t) \]
where \( \lambda^i(t) \) are Poisson process intensity function

Hypothesis testing was conducted through likelihood ratio tests.
Part III: Bayesian Nonparametric Approach

- A fully inferential model-based approach for the comparative problem of part 1 for two or more conditions
Part III: Bayesian Nonparametric Approach

- A fully inferential model-based approach for the comparative problem of part 1 for two or more conditions
- Pointwise and global analysis
Part III: Bayesian Nonparametric Approach

- A fully inferential model-based approach for the comparative problem of part 1 for two or more conditions
- Pointwise and global analysis
- Computationally inexpensive
Comparing Two Conditions

- **Neuronal data:**
  \[ \{ y_{ij}^{(\ell)} : i = 1, \ldots, N^{(\ell)}; j = 1, \ldots, n_i^{(\ell)} \}, \]
  
  \( y_{ij}^{(\ell)} \) is the \( j \)-th firing time in the \( i \)-th trial under condition \( \ell \), with \( \ell = 1, 2 \).
Comparing Two Conditions

- **Neuronal data:**
  \[
  \{y_{ij}^{(\ell)} : i = 1, \ldots, N^{(\ell)}; j = 1, \ldots, n_i^{(\ell)} \},
  \]
  \(y_{ij}^{(\ell)}\) is the \(j\)-th firing time in the \(i\)-th trial under condition \(\ell\), with \(\ell = 1, 2\).

- **Probability model** is a NHPP with intensity function \(\lambda^{(\ell)}(\cdot)\), (i.e., \(\int_D \lambda^{(\ell)}(u)du < \infty\) for any bounded subset \(D\) of the positive real line).
Comparing Two Conditions

- **Neuronal data:**
  \[ \{ y_{ij}^{(\ell)} : i = 1, \ldots, N^{(\ell)}; j = 1, \ldots, n_i^{(\ell)} \}, \]
  \( y_{ij}^{(\ell)} \) is the \( j \)-th firing time in the \( i \)-th trial under condition \( \ell \), with \( \ell = 1, 2 \).

- **Probability model** is a NHPP with intensity function \( \lambda^{(\ell)}(\cdot) \),
  (i.e., \( \int_D \lambda^{(\ell)}(u)du < \infty \) for any bounded subset \( D \) of the positive real line).

- **Likelihood** for \( \lambda(\cdot) \):
  \[
  \exp\{- \int_0^1 \lambda(u)du\} \prod_{k=1}^K \lambda(t_k). 
  \]
Bayesian nonparametrics focuses mainly on the NHPP cumulative intensity function:
\[ \Lambda(t) = \int_0^t \lambda(u)du, \; t \in R^+. \]
Bayesian nonparametrics focuses mainly on the NHPP cumulative intensity function:
\[ \Lambda(t) = \int_0^t \lambda(u)du, \quad t \in R^+. \]

Modeling \( \lambda(t), \quad t \in (0, 1) \), through density function:
\[ f(t) = \frac{\lambda(t)}{\gamma}, \]
where
\[ t \in (0, 1), \quad \gamma = \int_0^1 \lambda(u)du. \]
Bayesian nonparametrics focuses mainly on the NHPP cumulative intensity function:
\[ \Lambda(t) = \int_0^t \lambda(u)du, \quad t \in R^+. \]

Modeling \( \lambda(t) \), \( t \in (0, 1) \), through density function:
\[ f(t) = \frac{\lambda(t)}{\gamma}, \quad t \in (0, 1), \quad \gamma = \int_0^1 \lambda(u)du. \]

Employ Dirichlet process (DP) mixture of Beta densities model for \( f(\cdot) \):
\[ f(t; G) = \int \text{be}(t; \nu, \tau)dG(\nu, \tau), \quad G \sim \text{DP}(\alpha, G_0). \]
Bayesian nonparametrics focuses mainly on the NHPP cumulative intensity function:
\[ \Lambda(t) = \int_0^t \lambda(u)du, \quad t \in R^+. \]

Modeling \( \lambda(t), \quad t \in (0,1) \), through density function:
\[ f(t) = \frac{\lambda(t)}{\gamma}, \quad t \in (0,1), \quad \gamma = \int_0^1 \lambda(u)du. \]

Employ Dirichlet process (DP) mixture of Beta densities model for \( f(\cdot) \):
\[
 f(t; G) = \int \text{be}(t; \nu, \tau)dG(\nu, \tau), \quad G \sim \text{DP}(\alpha, G_0).
\]

Write full Bayesian model:
\[
 \exp(-\gamma)\gamma^K \left\{ \prod_{k=1}^{K} \int \text{be}(t_k; \nu, \tau)dG(\nu, \tau) \right\} p(\gamma)p(G | \alpha, \beta)p(\alpha)p(\beta).
\]
Marginal likelihood for $\gamma$ is proportional to $\exp(-\gamma)\gamma^K$

$p(\gamma) \propto \gamma^{-1} 1(\gamma > 0)$ as the reference prior for $\gamma$. 
Comparing Two Conditions: Prior Specification

- Marginal likelihood for $\gamma$ is proportional to $\exp(-\gamma)\gamma^K$
  
  $p(\gamma) \propto \gamma^{-1}1(\gamma>0)$ as the reference prior for $\gamma$.

- $G \sim DP(\alpha, G_0)$
Marginal likelihood for $\gamma$ is proportional to $\exp(-\gamma)\gamma^K$

$p(\gamma) \propto \gamma^{-1}1(\gamma>0)$ as the reference prior for $\gamma$.

$G \sim \text{DP}(\alpha, G_0)$

gamma prior for $\alpha$
Comparing Two Conditions: Prior Specification

- Marginal likelihood for $\gamma$ is proportional to $\exp(-\gamma)\gamma^K$
  \[ p(\gamma) \propto \gamma^{-1}1(\gamma>0) \] as the reference prior for $\gamma$.
- $G \sim \text{DP}(\alpha, G_0)$
- Gamma prior for $\alpha$
- $G_0(\nu, \tau)$, with uniform(0,1) prior for $\nu$, and inverse gamma for $\tau$. 
Let $\theta = \{(\nu_k, \tau_k) : k = 1, \ldots, K\}$ be the vector that collects all the mixing parameters. Then:

$$p(\gamma, G, \theta, \alpha, \beta | t) = p(\gamma | t) p(G, \theta, \alpha, \beta | t)$$
Let $\theta = \{(\nu_k, \tau_k) : k = 1, \ldots, K\}$ be the vector that collects all the mixing parameters. Then:

$$p(\gamma, G, \theta, \alpha, \beta \mid t) = p(\gamma \mid t) \ p(G, \theta, \alpha, \beta \mid t)$$

$p(\gamma \mid t)$ is a gamma distribution.
Let $\theta = \{(\nu_k, \tau_k) : k = 1, \ldots, K\}$ be the vector that collects all the mixing parameters. Then:

$$p(\gamma, G, \theta, \alpha, \beta \mid t) = p(\gamma \mid t) \, p(G, \theta, \alpha, \beta \mid t)$$

- $p(\gamma \mid t)$ is a gamma distribution.
- To get $p(G, \theta, \alpha, \beta \mid t)$, we use

$$p(G, \theta, \alpha, \beta \mid t) = p(G \mid \theta, \alpha, \beta)p(\theta, \alpha, \beta \mid t).$$
Obtain posterior realizations over \((0, 1)\) for each of the two firing intensities to the full model for \((\gamma^{(1)}, f^{(1)}(\cdot))\) and \((\gamma^{(2)}, f^{(2)}(\cdot))\) given the data \(t^{(1)}\) and \(t^{(2)}\).
Obtain posterior realizations over $(0, 1)$ for each of the two firing intensities to the full model for $(\gamma^{(1)}, f^{(1)}(\cdot))$ and $(\gamma^{(2)}, f^{(2)}(\cdot))$ given the data $t^{(1)}$ and $t^{(2)}$.

Global Analysis: Consequently, we derive posterior point estimates and uncertainty bands for the function: $f^{(1)}(\cdot) - f^{(2)}(\cdot)$. 
Obtain posterior realizations over \((0, 1)\) for each of the two firing intensities to the full model for \((\gamma^{(1)}, f^{(1)}(\cdot))\) and \((\gamma^{(2)}, f^{(2)}(\cdot))\) given the data \(t^{(1)}\) and \(t^{(2)}\).

Global Analysis: Consequently, we derive posterior point estimates and uncertainty bands for the function: 
\[ f^{(1)}(\cdot) - f^{(2)}(\cdot). \]

Pointwise Analysis: Similarly, we derive the entire posterior \(p\{f^{(1)}(t_0; G^{(1)}) - f^{(2)}(t_0; G^{(2)}) \mid t^{(1)}, t^{(2)}\}\) for specific points \(t_0\).
Comparing Two Neurons: Data Analysis
Global Analysis for Neuron 1

Random condition

Repeating condition

Difference of densities

A Bayesian Comparative Analysis of Neuronal Point Processes
Global Analysis for Neuron 2

Random condition

Repeating condition

Difference of densities

Sam Behseta  California State University Fullerton  A Bayesian Comparative Analysis of Neuronal Point Processes
A Bayesian Comparative Analysis of Neuronal Point Processes
Neurons were recorded for a 4000ms (−2000ms, 2000ms) time interval for three conditions in which a memory-guided saccade was made to a visual target (space, ring, or dot).
Neurons were recorded for a 4000ms (–2000ms, 2000ms) time interval for three conditions in which a memory-guided saccade was made to a visual target (space, ring, or dot).

16 trials performed under each condition.
Neurons were recorded for a 4000ms (−2000ms, 2000ms) time interval for three conditions in which a memory-guided saccade was made to a visual target (space, ring, or dot).

16 trials performed under each condition.

\[
\text{Likelihood: } \left[ \prod_{i=1}^{3} \left( \exp \left\{ - \int_{0}^{1} \lambda_i(u) \, du \right\} \prod_{j=1}^{16} \prod_{k=1}^{n_j} \lambda_i(t_{ijk}) \right) \right]^{s_{ijk}}
\]
We transform the data to the scale $(0, 1)$: $t \rightarrow y \in (0, 1)$
We transform the data to the scale (0, 1): $t \rightarrow y \in (0, 1)$

model the density that defines the intensity function

$(f_i(y) = \lambda_i(y)/\gamma_i$ where $\gamma_i = \int_0^1 \lambda_i(u)du)$
We transform the data to the scale $(0, 1)$: $t \rightarrow y \in (0, 1)$

model the density that defines the intensity function

$(f_i(y) = \lambda_i(y)/\gamma_i$ where $\gamma_i = \int_0^1 \lambda_i(u)du$)

Use a logit-normal dependent Dirichlet process mixture model to borrow strength from other conditions:

$y_{ijk} | \sigma_i^2, G_i \sim \int \logit-N(y_{ijk}; \mu, \sigma_i^2) dG_i(\mu) = \sum_{l=1}^L p_l \logit-N(y_{ijk}; \mu_l, \sigma_i^2)$
We transform the data to the scale $(0, 1)$: $t \rightarrow y \in (0, 1)$

model the density that defines the intensity function

$f_i(y) = \lambda_i(y)/\gamma_i$ where $\gamma_i = \int_0^1 \lambda_i(u)du$

Use a logit-normal dependent Dirichlet process mixture model to borrow strength from other conditions:

$y_{ijk} | \sigma_i^2, G_i \sim \int \logit-N(y_{ijk}; \mu, \sigma_i^2)dG_i(\mu) = \sum_{l=1}^L p_l \logit-N(y_{ijk}; \mu_l, \sigma_i^2)$

Marginally, $G_i \sim DP(\alpha, G_0)$
We transform the data to the scale $(0, 1)$: $t \rightarrow y \in (0, 1)$

model the density that defines the intensity function $f_i(y) = \lambda_i(y)/\gamma_i$ where $\gamma_i = \int_0^1 \lambda_i(u)du$

Use a logit-normal dependent Dirichlet process mixture model to borrow strength from other conditions:

$y_{ijk} | \sigma_i^2, G_i \sim \int \text{logit-}N(y_{ijk}; \mu, \sigma_i^2) dG_i(\mu) = \sum_{l=1}^L p_l \text{logit-}N(y_{ijk}; \mu_l, \sigma_i^2)$

Marginally, $G_i \sim DP(\alpha, G_0)$

$p_l$'s are the weights obtained via DP stick-breaking construction corresponding to the component $\theta_{li} = (\mu_{li})$ and $L$ is the total number of components specified in the model.
We transform the data to the scale $(0, 1)$: $t \rightarrow y \in (0, 1)$

- model the density that defines the intensity function $(f_i(y) = \lambda_i(y)/\gamma_i$ where $\gamma_i = \int_0^1 \lambda_i(u)du$)

- Use a logit-normal dependent Dirichlet process mixture model to borrow strength from other conditions:

$$y_{ijk} | \sigma_i^2, G_i \sim \int \logit-N(y_{ijk}; \mu, \sigma_i^2) dG_i(\mu) = \sum_{l=1}^L p_l \logit-N(y_{ijk}; \mu_{il}, \sigma_i^2)$$

- Marginally, $G_i \sim DP(\alpha, G_0)$

- $p_l$’s are the weights obtained via DP stick-breaking construction corresponding to the component $\theta_{li} = (\mu_{li})$ and $L$ is the total number of components specified in the model

- Dependence across the conditions are implied through the common locations $\theta_l = (\theta_{l1}, \theta_{l2}, \theta_{l3})$
Data Analysis Per Condition

sp220b

sp259a

Sam Behseta  California State University Fullerton  A Bayesian Comparative Analysis of Neuronal Point Processes
Data Analysis– Pairwise Comparisons

- **sp220b, Space - Dot Conditions**
  - Difference in Densities:
    - Time (ms): -2000, -1000, 0, 1000, 2000
    - Difference Values: -5e-04, 5e-04

- **sp220b, Space - Ring Conditions**
  - Difference in Densities:
    - Time (ms): -2000, -1000, 0, 1000, 2000
    - Difference Values: -5e-04, 5e-04

- **sp259a, Space - Dot Conditions**
  - Difference in Densities:
    - Time (ms): -2000, -1000, 0, 1000, 2000
    - Difference Values: -5e-04, 0e+00, 5e-04

- **sp259a, Space - Ring Conditions**
  - Difference in Densities:
    - Time (ms): -2000, -1000, 0, 1000, 2000
    - Difference Values: -6e-04, 0e+00, 4e-04

- **sp220b, Dot - Ring Conditions**
  - Difference in Densities:
    - Time (ms): -2000, -1000, 0, 1000, 2000
    - Difference Values: -4e-04, 0e+00, 4e-04

- **sp259a, Dot - Ring Conditions**
  - Difference in Densities:
    - Time (ms): -2000, -1000, 0, 1000, 2000
    - Difference Values: -4e-04, 0e+00, 4e-04
Collaborators

Rob Kass, Carnegie Mellon, CNBC
Shoja Chenouri, Waterloo
Garrick Wallstrom, University of Pittsburgh
Peter Strick, CNBC
Thanasis Kottas, UCSC
Valerie Poynor, UCSC
Peter Strick, CNBC
Carl Olson, CNBC
David Moorman, CNBC, South Carolina Medical School
Thanks!

sbehseta@fullerton.edu