Quantile Functional Regression by Quantlets

Jeffrey S. Morris
Quantile Functional Regression by Quantlets

Jeffrey S. Morris, Hojin Yang, Veera Baladandayuthapani

Department of Biostatistics
The University of Texas MD Anderson Cancer Center
Houston, TX 77030

July 24, 2017
Outline

1 Introduction
   - Motivation
   - Quantile Functions

2 Methods
   - Quantile Functional Regression
   - Quantlets
   - Model Setup
   - Estimation

3 Application
   - Simulation
   - GBM Data Analysis

4 Summary
Glioblastoma Multiforme (GBM)

- Most common and aggressive form of brain cancer
- No current prevention approaches, and poor outcomes
  - Median survival 12mo, 3-5% 5yr survival
- Exhibits heterogeneous physiological and morphological features as it proliferates
- Investigating these heterogeneities and relating them to clinical/genetic outcomes can lead to the development of personalized treatment strategies.
Glioblastoma Multiforme (GBM)

- Most common and aggressive form of brain cancer
- No current prevention approaches, and poor outcomes
  - Median survival 12mo, 3-5% 5yr survival
- Exhibits heterogeneous physiological and morphological features as it proliferates
- Investigating these heterogeneities and relating them to clinical/genetic outcomes can lead to the development of personalized treatment strategies.

Our Goal:
Assess how variability in tumor image intensities is associated with demographic, clinical, and genetic factors
Glioblastoma Images

- Presurgical T1-weighted post-contrast MRI images from GBM patients
- **Radiomics**: compute features summarizing tumor image characteristics and relate to clinical outcomes (100s of different features)
- **Histogram features**: Summaries computed from pixel intensity distributions (e.g. mean, variance, skewness, Q05, Q95)
The typical approach is to fit separate regression analyses to each radiomic feature, which has some major drawbacks:

- Multiple testing problems
- May miss distributional differences not contained in pre-chosen summaries.
The typical approach is to fit separate regression analyses to each radiomic feature, which has some major drawbacks:

- Multiple testing problems
- May miss distributional differences not contained in pre-chosen summaries.

**Alternative Approach**

Instead of just modeling the extracted summaries, model the entire distribution of pixel intensities (as functional data).
Various choices to represent pixel intensity distributions: density, cumulative distribution, or quantile functions.

**Definition of the quantile function**

\[
Q_Y(p) = F_Y^{-1}(p) = \inf \{ y : F_Y(y) \geq p \},
\]

where \( p = F_Y(y) \) is the proportion less than or equal to \( y \).
Various choices to represent pixel intensity distributions: density, cumulative distribution, or quantile functions.

We choose to use the quantile function. The quantile function of $Y$ on $p \in (0, 1)$, is defined as

**Definition of the quantile function**

$$Q_Y(p) = F_Y^{-1}(p) = \inf (y : F_Y(y) \geq p),$$

where $p = F_Y(y)$ is the proportion less than or equal to $y$. 

A. Various distributions with different parameters

B. Cumulative distribution functions

C. Quantile functions
Quantile functions have properties that make them useful here:

- Defined on a fixed domain, $p \in \mathcal{P} = (0, 1)$
Properties of Quantile Functions

Quantile functions have properties that make them useful here:

- Defined on a fixed domain, \( p \in \mathcal{P} = (0, 1) \)
- Straightforward to compute empirical estimates without choice of smoothing parameters

**eDF**

Let \( Y_{(1)} \leq \cdots \leq Y_{(m)} \) be order statistics from a sample of size \( m \). For \( p \in [1/(m + 1), m/(m + 1)] \), the eQF is given by

\[
\hat{Q}_Y(p) = (1 - w)Y_{([(m+1)p])} + wY_{([(m+1)p]+1)},
\]

where \( w \) is a weight such that \((m + 1)p = [(m + 1)p] + w\).
Properties of Quantile Functions

Quantile functions have properties that make them useful here:

- Defined on a fixed domain, \( p \in \mathcal{P} = (0, 1) \)
- Straightforward to compute empirical estimates without choice of smoothing parameters
- Straightforward formulas to calculate distributional moments

**Distributional Moments**

\[
\begin{align*}
\mu_Y &= \mathbb{E}(Y) = \int_0^1 Q_Y(p) \, dp \\
\sigma_Y^2 &= \text{Var}(Y) = \int_0^1 (Q_Y(p) - \mu_Y)^2 \, dp \\
\xi_Y &= \text{Skew}(Y) = \int_0^1 (Q_Y(p) - \mu_Y)^3 / \sigma_Y^3 \, dp
\end{align*}
\]
Quantile functional regression

**Approach:** Regress eQF as functional response on covariates.
Quantile functional regression

**Approach:** Regress eQF as functional response on covariates.

1. For each subject $i = 1, \ldots, n$, construct the eQF $Q_i(p)$ from the order statistics of $Y_{ij}, j = 1, \ldots, m_i$. 
Quantile functional regression

**Approach:** Regress eQF as functional response on covariates.

1. For each subject $i = 1, \ldots, n$, construct the eQF $Q_i(p)$ from the order statistics of $Y_{ij}, j = 1, \ldots, m_i$.

2. Regress $Q_i(p)$ on covariates $x_{ia}, a = 1, \ldots, A$, each with regression coefficients $\beta_a(p)$ defined on $p \in \mathcal{P} = (0, 1)$.

**Quantile Functional Regression Model**

$$Q_i(p) = \beta_0(p) + \sum_{a=1}^{A} x_{ia} \beta_a(p) + E_i(p)$$
**Approach:** Regress eQF as functional response on covariates.

1. For each subject $i = 1, \ldots, n$, construct the eQF $Q_i(p)$ from the order statistics of $Y_{ij}, j = 1, \ldots, m_i$.

2. Regress $Q_i(p)$ on covariates $x_{ia}, a = 1, \ldots, A$, each with regression coefficients $\beta_a(p)$ defined on $p \in \mathcal{P} = (0, 1)$.

**Quantile Functional Regression Model**

$$Q_i(p) = \beta_0(p) + \sum_{a=1}^{A} x_{ia} \beta_a(p) + E_i(p)$$

3. Test for significantly associated covariates: $H_0: \beta_a(p) \equiv 0$. 
**Approach:** Regress eQF as functional response on covariates.

1. For each subject \( i = 1, \ldots, n \), construct the eQF \( Q_i(p) \) from the order statistics of \( Y_{ij}, j = 1, \ldots, m_i \).
2. Regress \( Q_i(p) \) on covariates \( x_{ia}, a = 1, \ldots, A \), each with regression coefficients \( \beta_a(p) \) defined on \( p \in \mathcal{P} = (0, 1) \).

**Quantile Functional Regression Model**

\[
Q_i(p) = \beta_0(p) + \sum_{a=1}^{A} x_{ia} \beta_a(p) + E_i(p)
\]

3. Test for significantly associated covariates: \( H_0 : \beta_a(p) \equiv 0 \).
4. Characterize the significant distributional differences e.g. range of \( p \), mean, variance, skewness, Gausianness.
Types of Quantile and Functional Regression

| Response (·)          | Objective function $E(·|X)$                                      | Objective function $F^{-1}(·)(p|X)$                                      |
|-----------------------|------------------------------------------------------------------|---------------------------------------------------------------------|
| scalar $Y$ function $Y(t)$ | classic regression functional regression quantile functional regression* | quantile functional quantile regression |
| quantile function $F^{-1}(p)$ |                                                                                     |

- **Classic regression:** $E(Y|X)$
- **Quantile regression:** $F^{-1}_Y(p|X)$
  
e.g. He and Liang 2000; Koenker 2005
Types of Quantile and Functional Regression

| Response (·) | Objective function $E((·)|X)$ | Objective function $F_{(·)}^{-1}(p|X)$ |
|--------------|--------------------------------|----------------------------------------|
| scalar $Y$   | classic regression             | quantile regression                    |
| function $Y(t)$ | functional regression          | functional quantile regression         |
| quantile function $F^{-1}(p)$ | quantile functional regression* | quantile functional functional quantile regression |

- **Classic regression:**
  $E(Y|X)$

- **Quantile regression:**
  $F_{Y}^{-1}(p|X)$
  e.g. He and Liang 2000; Koenker 2005

- **Functional regression:**
  $E\{Y(t)|X\}$
  See review article by Morris (2015)

- **Functional quantile regression:**
  $F_{Y(t)}^{-1}(p|X)$
  e.g. Brockhaus et al. (2015)
### Types of Quantile and Functional Regression

| Response (\( \cdot \)) | Objective function \( E((\cdot)|X) \) | Objective function \( F_{(\cdot)}^{-1}(p|X) \) |
|-------------------------|----------------------------------------|----------------------------------------|
| scalar \( Y \)          | classic regression                     | quantile regression                    |
| function \( Y(t) \)     | functional regression                   | functional quantile regression         |
| quantile function \( F_{(\cdot)}^{-1}(p) \) | quantile functional regression*        | quantile functional quantile regression |

- **Classic regression**: \( E(Y|X) \)
- **Quantile regression**: \( F_Y^{-1}(p|X) \)
  - e.g. He and Liang 2000; Koenker 2005
- **Functional regression**: \( E\{Y(t)|X\} \)
  - See review article by Morris (2015)
- **Functional quantile regression**: \( F_Y^{-1}(t|X) \)
  - e.g. Brockhaus et al. (2015)
- **Quantile functional regression**: \( E\{F_Y^{-1}(p)|X\} \)
  - Expected quantile function given covariates
Quantile Functional Regression Model

\[ Q_i(p) = \beta_0(p) + \sum_{a=1}^{A} x_{ia}\beta_a(p) + E_i(p) \]

**Naive approach:** compute independent regressions for each \( p \)
- fail to borrow strength over \( p \) → wiggly, inefficient \( \hat{\beta}_a(p) \).
- ignore correlation over \( p \) in \( E_i(p) \) → loss of inferential power.
Quantile Functional Regression

Quantile Functional Regression Model

\[ Q_i(p) = \beta_0(p) + \sum_{a=1}^{A} x_{ia} \beta_a(p) + E_i(p) \]

**Naive approach:** compute independent regressions for each \( p \)
- fail to borrow strength over \( p \rightarrow \) wiggly, inefficient \( \hat{\beta}_a(p) \).
- ignore correlation over \( p \) in \( E_i(p) \rightarrow \) loss of inferential power.

**Functional regression approach:** Use *basis function* representations to account for correlation.
- \( \beta_a(p) \) regularized via L1/L2 penalization of basis coefficients.
- Basis functions induce correlation across \( p \) in \( \text{Cov}\{E_i(p)\} \).
- Common bases: splines, PC, Fourier bases, wavelets

Here, we introduce new custom basis functions *quantlets*. 
Multi-step process to derive custom quantlet basis functions:

1. Construct overcomplete dictionary

**Details of Step**

- **Gaussian bases:** $\psi_0(p) = 1$ for $p \in (0, 1)$, $\psi_1(p) = \Phi^{-1}(p)$.
- **Beta CDF bases:** $\psi_k(p) = F_{\theta_k}(p)$ for $k = 2, \ldots, K_0$.
- **Overcomplete dictionary:** $\mathcal{D}^0 = \{\psi_k, k = 0, \ldots, K_0\}$.
Construction of Quantlet Basis Functions

Multi-step process to derive custom quantlet basis functions:

1. Construct overcomplete dictionary
2. Choose sparse set of dictionary elements for each subject.

Details of Step

For each subject, use penalized regression (e.g. lasso) to find a sparse subset of dictionary elements.

\[ |Q_i(p) - \sum_{k \in \mathcal{D}_0} \psi_k(p) Q_{ik}^O|_2^2 + \lambda_i \sum_{k \in \mathcal{D}_0} |Q_{ik}^O|_1 \]

Obtain \( \mathcal{D}_i = \{\psi_k(p) \in \mathcal{D}_0 : Q_{ik}^O \neq 0\} \).
Construction of Quantlet Basis Functions

Multi-step process to derive custom quantlet basis functions:

1. Construct overcomplete dictionary
2. Choose sparse set of dictionary elements for each subject.
3. Take union set, and then find subset that is near-lossless.

Details of Step

Union set: $\mathcal{D}^U = \bigcup_{i=1}^n \mathcal{D}_i$

Cardinality $C$ set: $\mathcal{D}^C = \{\psi_k(p), k : \sum_{i=1}^n I(Q_{ik}^0 \neq 0) \geq C\}$

Lossless measure: Cross-validated concordance coefficient:

$$\rho^C_i = \text{Concordance}\{Q_i(p), \hat{Q}_i^C(p)\} \in (0, 1)$$

Plot $\rho^C_0 = \min_i \{\rho^C_i\}$ vs. $C$ and choose $C : \rho^C_0 < \epsilon$

Near-lossless set: $\mathcal{D}^\epsilon = \{\mathcal{D}^C \text{ with } C = \min(C : \rho^C_0 < \epsilon)\}$
Construction of Quantlet Basis Functions

Multi-step process to derive custom quantlet basis functions:

1. Construct overcomplete dictionary
2. Choose sparse set of dictionary elements for each subject.
3. Take union set, and then find subset that is near-lossless.
4. Orthogonalize this subset, regularize, and re-standardize.

Details of Step

Orthogonal set: \( \mathcal{D}^\perp = \{ \psi_k^\perp, k = 0, \ldots, K \} = \text{Gram-Schmidt}(\mathcal{D}^\epsilon) \)

Regularize \( \psi^\perp \) via wavelet denoising and then renormalize.

Resulting bases are called quantlets: \( \mathcal{D} = \{ \xi_k(p), k = 0, \ldots, K \} \)
First 16 Quantlets for GBM Data
Properties of Quantlets

- **Empirically defined**: adapts to characteristics of given data.
Properties of Quantlets

- **Empirically defined**: adapts to characteristics of given data.
- **Near-lossless**: rich enough to capture structure in each eQF.

![Graph showing quantitative properties of Quantlets](image)
Properties of Quantlets

- **Empirically defined:** adapts to characteristics of given data.
- **Near-lossless:** rich enough to capture structure in each eQF.
- **Regularity:** denoising removes wiggles $\rightarrow$ smooth quantlets.
Properties of Quantlets

- **Empirically defined:** adapts to characteristics of given data.
- **Near-lossless:** rich enough to capture structure in each eQF.
- **Regularity:** denoising removes wiggles $\rightarrow$ smooth quantlets.
- **Sparsity:** tends to produce low dimensional basis.
Properties of Quantlets

- **Empirically defined:** adapt to characteristics of given data.
- **Near-lossless:** rich enough to capture structure in each eQF.
- **Regularity:** denoising removes wiggles $\rightarrow$ smooth quantlets.
- **Sparsity:** tends to produce low dimensional basis.
- **Interpretability:** first two bases measure Gaussianity
**Basis Transform Modeling Approach**

**Data Space Model**

\[ Q_i(p) = X_i^T B(p) + E_i(p), \]

where \( B(p) = (\beta_1(p), \ldots, \beta_A(p))^T \) and \( E_i(p) \) is a noise process.

1. Compute quantlet basis coefficients

**Computing Quantlet Coefficients**

Let \( Q_i = [Q_i(p_1), \ldots, Q_i(p_{m_i})] \) with \( p_j = j/(m_i + 1) \)

Let \( \Psi_i \) be \( K \times m_i \) matrix with elements \( \psi_{i}(k, j) = \psi_k(p_j) \)

Quantlet coefficients: \( Q_i^* = Q_i \Psi_i^{-} \) where \( \Psi_i^{-} = \Psi_i^T (\Psi_i \Psi_i^T)^{-1} \).
Basis Transform Modeling Approach

### Data Space Model

\[ Q_i(p) = X_i^T B(p) + E_i(p), \]

where \( B(p) = (\beta_1(p), \ldots, \beta_A(p))^T \) and \( E_i(p) \) is a noise process.

1. Compute quantlet basis coefficients
2. Fit quantlet space model

### Quantlet Space Model

\[ Q^* = XB^* + E^* \]

where \( Q_i(p_j) = \sum_{k=1}^K Q_{ik}^* \psi_k(p_j) \) and \( \beta_a(p) = \sum_{k=1}^K B_{ak}^* \psi_k(p) \),
\[ E_i(p) = \sum_{k=1}^K E_{ik}^* \psi_k(p), \] and \( \{p_1, \ldots, p_J\} \in (0, 1). \)
\[ E_i^* \sim \text{MVN}(0, \Sigma^*) \] where \( \Sigma^* \) is \( K \times K \) covariance matrix.
Basis Transform Modeling Approach

Data Space Model

\[ Q_i(p) = X_i^T B(p) + E_i(p), \]
where \( B(p) = (\beta_1(p), \ldots, \beta_A(p))^T \) and \( E_i(p) \) is a noise process.

1. Compute quantlet basis coefficients
2. Fit quantlet space model
3. Transform results back to data space for inference

Transform Results to Data Space

\[ \beta_a(p) = \sum_{k=1}^{K} B_{ak}^* \psi_k(p), \]
and then perform desired inference.
We use a Bayesian modeling approach to fit this model.
We use a Bayesian modeling approach to fit this model.

- Sparsity prior on $B_{ak}^*$ to regularize $\beta_a(p)$. (spike Gaussian-slab)
We use a Bayesian modeling approach to fit this model.

- Sparsity prior on $B_{ak}^*$ to regularize $\beta_a(p)$. (spike Gaussian-slab)
- Vague proper prior on covariance parameters.

EBayes or hyperpriors on sparsity hyperparameters.

MCMC used to update parameters in the quantlet space model.

Complete conditional for $B_{ak}^*$ is mixture of $\delta_0$ and Gaussian.

Covariance parameters have conjugate complete conditionals.

Posterior samples transformed back to original data space to get posterior samples of $\beta_a(p)$ on desired grid of $p$. 

We use a Bayesian modeling approach to fit this model.

- Sparsity prior on $B_{ak}^*$ to regularize $\beta_a(p)$. (spike Gaussian-slab)
- Vague proper prior on covariance parameters.
- EBayes or hyperpriors on sparsity hyperparameters.
Bayesian Modeling

- We use a Bayesian modeling approach to fit this model.
  - Sparsity prior on $B_{ak}^*$ to regularize $\beta_a(p)$. (spike Gaussian-slab)
  - Vague proper prior on covariance parameters.
  - EBayes or hyperpriors on sparsity hyperparameters.
- MCMC used to update parameters in the quantlet space model.
We use a Bayesian modeling approach to fit this model.

- Sparsity prior on $B_{ak}^*$ to regularize $\beta_a(p)$. (spike Gaussian-slab)
- Vague proper prior on covariance parameters.
- EBayes or hyperpriors on sparsity hyperparameters.

MCMC used to update parameters in the quantlet space model.

- Complete conditional for $B_{ak}^*$ is mixture of $\delta_0$ and Gaussian.
Bayesian Modeling

- We use a Bayesian modeling approach to fit this model.
  - Sparsity prior on $B_{a_k}^*$ to regularize $\beta_a(p)$. (spike Gaussian-slab)
  - Vague proper prior on covariance parameters.
  - EBayes or hyperpriors on sparsity hyperparameters.

- MCMC used to update parameters in the quantlet space model.
  - Complete conditional for $B_{a_k}^*$ is mixture of $\delta_0$ and Gaussian.
  - Covariance parameters have conjugate complete conditionals.
We use a Bayesian modeling approach to fit this model.

- Sparsity prior on $B_{ak}^*$ to regularize $\beta_a(p)$. (spike Gaussian-slab)
- Vague proper prior on covariance parameters.
- EBayes or hyperpriors on sparsity hyperparameters.

MCMC used to update parameters in the quantlet space model.

- Complete conditional for $B_{ak}^*$ is mixture of $\delta_0$ and Gaussian.
- Covariance parameters have conjugate complete conditionals.

Posterior samples transformed back to original data space to get posterior samples of $\beta_a(p)$ on desired grid of $p$. 
Recommended Sequence of Bayesian Inference

1. Construct 95\% joint credible bands for each predictor.

100(1 − \alpha)\% Joint Credible Band (Ruppert/Wand/Carroll 2003)

\[
J_{a,\alpha}(p) = \hat{\beta}_a(p) \pm q_{1-\alpha} \left[ \text{StDev}\{\hat{\beta}_a(p)\} \right]
\]

where \(q_{1-\alpha}\) is \((1 - \alpha)\) quantile of:

\[
Z_a^{(m)} = \max_{p \in \mathcal{P}} \left| \frac{\beta^{(m)}_a(p) - \hat{\beta}_a(p)}{\text{St.Dev}\{\hat{\beta}_a(p)\}} \right|
\]
Recommended Sequence of Bayesian Inference

1. Construct 95% joint credible bands for each predictor.
2. Calculate global Bayesian p-value for each predictor.

Global Bayesian P-value (Meyer et al. 2015)

To assess \( H_0 : \beta_a(p) \equiv 0 \), we compute:

\[
P_{a,\text{Bayes}} = \min \{ \alpha : 0 \notin J_{a,\alpha}(p) \text{ for some } p \in \mathcal{P} \},
\]

and conclude \( \beta_a(p) \) differs from 0 whenever \( P_{a,\text{Bayes}} < \alpha \).
Recommended Sequence of Bayesian Inference

1. Construct 95% joint credible bands for each predictor.
2. Calculate global Bayesian p-value for each predictor.
3. For significant predictors, flag \( \{ p : P_{a,\text{SimBaS}} < \alpha \} \).

Simultaneous Band Scores (SimBaS, Meyer et al. 2015)

\[
P_{a,\text{SimBas}}(p) = \min\{ \alpha : 0 \notin J_{a,\alpha}(p) \}
\]

\[
= M^{-1} \sum_{m=1}^{M} I \left\{ \left| \frac{\hat{\beta}_a(p)}{\text{StDev}\{\hat{\beta}_a(p)\}} \right| \leq Z_{a}^{(m)} \right\}
\]
Recommended Sequence of Bayesian Inference

1. Construct 95% joint credible bands for each predictor.
2. Calculate global Bayesian p-value for each predictor.
3. For significant predictors, flag \( \{ p : P_{a,\text{SimBaS}} < \alpha \} \).
4. For significant predictors, assess which moments differ.

Probability scores for moments

\[
\mu^{(m)}_X = \int_0^1 X^T \beta^{(m)}(p) dp
\]

\[
P_{\mu_1 - \mu_2} = 2 \cdot \min \left\{ M^{-1} \sum_{m=1}^M I \left( \mu^{(m)}_{X_1} - \mu^{(m)}_{X_2} > 0 \right), \right. \\
\left. M^{-1} \sum_{m=1}^M I \left( \mu^{(m)}_{X_1} - \mu^{(m)}_{X_2} < 0 \right) \right\}
\]
Figure: 4 groups: mean distributions are \( N(1,5) \), \( N(3,5) \), \( N(1,6.5) \), and skewed normal with mean 1 and variance 5.
Figure: Simulated Data. $\beta_a(p)$ are location, scale, and skewness shifts.

- $Y_{ij}(p) = Q_{ij}(p) + \epsilon_{ij}(p)$ on 1,024 grid points $\{p_1, \ldots, p_{1024}\}$.
- $\epsilon_{ij}(p)$ follows AR(1) process to approximate biological variability within groups.
Simulation Results

Figure: Results of the simulation: estimations and 95% joint CI (A=Naive *one-p-at-a-time* method; D=*quantlets* with regularization)
## Simulation Results

**Table:** Area and coverage for the joint 95% credible intervals.

<table>
<thead>
<tr>
<th>Type</th>
<th>A (naive)</th>
<th>B (PCA)</th>
<th>C (no reg.)</th>
<th>D (regularized)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_1(p)$</td>
<td>1.603 (1.000)</td>
<td>1.092 (0.999)</td>
<td>1.186 (1.000)</td>
<td>1.069 (1.000)</td>
</tr>
<tr>
<td>$\beta_2(p)$</td>
<td>2.246 (1.000)</td>
<td>1.551 (1.000)</td>
<td>1.706 (1.000)</td>
<td>1.465 (1.000)</td>
</tr>
<tr>
<td>$\beta_3(p)$</td>
<td>2.242 (1.000)</td>
<td>1.599 (1.000)</td>
<td>1.717 (1.000)</td>
<td>1.457 (1.000)</td>
</tr>
<tr>
<td>$\beta_4(p)$</td>
<td>2.281 (1.000)</td>
<td>1.583 (1.000)</td>
<td>1.651 (1.000)</td>
<td>1.499 (1.000)</td>
</tr>
</tbody>
</table>

**Table:** Probability scores for differences in mean, variance, and skewness.

<table>
<thead>
<tr>
<th>$H_0$</th>
<th>True</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E (feature)</th>
<th>F (Gauss)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu_1 = \mu_3$</td>
<td>$\mu_1 = \mu_3$</td>
<td>0.001</td>
<td>0.193</td>
<td>0.211</td>
<td>0.217</td>
<td>0.205</td>
<td>0.212</td>
</tr>
<tr>
<td>$\mu_2 = \mu_4$</td>
<td>$\mu_2 = \mu_4$</td>
<td>0.001</td>
<td>0.447</td>
<td>0.465</td>
<td>0.445</td>
<td>0.438</td>
<td>0.462</td>
</tr>
<tr>
<td>$\sigma_1 = \sigma_3$</td>
<td>$\sigma_1 \neq \sigma_3$</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td>$\sigma_2 = \sigma_4$</td>
<td>$\sigma_2 = \sigma_4$</td>
<td>0.002</td>
<td>0.420</td>
<td>0.334</td>
<td>0.331</td>
<td>0.187</td>
<td>0.016</td>
</tr>
<tr>
<td>$\xi_1 = \xi_3$</td>
<td>$\xi_1 = \xi_3$</td>
<td>0.374</td>
<td>0.498</td>
<td>0.488</td>
<td>0.479</td>
<td>0.389</td>
<td>0.493</td>
</tr>
<tr>
<td>$\xi_2 = \xi_4$</td>
<td>$\xi_2 \neq \xi_4$</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.505</td>
</tr>
</tbody>
</table>
GBM Data Analysis

**Response:** T1 MRI images from 64 patients in glioblastoma (GBM) study, $Y_{ij} =$ intensity of pixel $j$ from subject $i$, $i = 1, \ldots, n$ and $j = 1, \ldots, m_i$, with $m_i$ ranging from 371 to 3421.

**Covariates:**
- **Demographic variables:** sex (21 F/43 M) & age (56.5 yr)
- **GBM subtype:** mesenchymal (30 mes./34 other)
- **Clinical outcome:** survival ($>12$ m/$<12$ m)
- **Genetic alterations:** DDIT3 (6 m/58 wt) & EGFR (24 m/58 wt)

**Model**

$$Q_i(p|X_i) = \beta_0(p) + x_{sex,i}\beta_{sex}(p) + x_{age,i}\beta_{age}(p) + x_{surv,i}\beta_{surv}(p) + x_{Mes,i}\beta_{Mes}(p) + x_{DDIT3,i}\beta_{DDIT3}(p) + x_{EGFR,i}\beta_{EGFR}(p) + E_i(p).$$
GBM Results

\[ P_{sex,\mu} = 0.004, \quad P_{sex,\sigma^2} = 0.121, \quad P_{sex,\xi} = 0.51 \]
$P_{DDIT3,\mu} = 0.008$, $P_{DDIT3,\sigma^2} = 0.023$, $P_{DDIT3,\xi} = 0.468$
Summary

- General approach to regress distributions on covariates
- Useful in many settings (e.g. activity data, climate data)
- Introduce quantlets basis functions that are sparse, regularized, near-lossless, empirically determined, and interpretable and lead to efficient regression.
- Bayesian framework yields global and local tests that adjust for multiple testing.
  - Greater power than naive one-p-at-a-time approach
  - No power loss compared with feature extraction.
Joining with BayesFMM framework allows us to extend to more complex settings with correlated observations across subjects, nonparametric effects (in $X$), functional predictors, and to perform robust regression when outliers are present.
Future Work

- Joining with BayesFMM framework allows us to extend to more complex settings with correlated observations across subjects, nonparametric effects (in $X$), functional predictors, and to perform robust regression when outliers are present.
- Applications of interest
Future Work

- Joining with BayesFMM framework allows us to extend to more complex settings with correlated observations across subjects, nonparametric effects (in $X$), functional predictors, and to perform robust regression when outliers are present.

- Applications of interest
  - Climate change data
Future Work

- Joining with BayesFMM framework allows us to extend to more complex settings with correlated observations across subjects, nonparametric effects (in $X$), functional predictors, and to perform robust regression when outliers are present.
- Applications of interest
  - Climate change data
  - Activity data/wearable computing
Future Work

- Joining with BayesFMM framework allows us to extend to more complex settings with correlated observations across subjects, nonparametric effects (in $X$), functional predictors, and to perform robust regression when outliers are present.
- Applications of interest
  - Climate change data
  - Activity data/wearable computing
  - Biomechanical optical data (250Hz for 6mo)
Future Work

- Joining with BayesFMM framework allows us to extend to more complex settings with correlated observations across subjects, nonparametric effects (in $X$), functional predictors, and to perform robust regression when outliers are present.

- Applications of interest
  - Climate change data
  - Activity data/wearable computing
  - Biomechanical optical data (250Hz for 6mo)

- Extensions
Joining with BayesFMM framework allows us to extend to more complex settings with correlated observations across subjects, nonparametric effects (in $X$), functional predictors, and to perform robust regression when outliers are present.

Applications of interest
- Climate change data
- Activity data/wearable computing
- Biomechanical optical data (250Hz for 6mo)

Extensions
- Extreme data (large $N \approx 100k$ and large $P > 10^{10}$)
Future Work

- Joining with BayesFMM framework allows us to extend to more complex settings with correlated observations across subjects, nonparametric effects (in $X$), functional predictors, and to perform robust regression when outliers are present.

- Applications of interest
  - Climate change data
  - Activity data/wearable computing
  - Biomechanical optical data (250Hz for 6mo)

- Extensions
  - Extreme data (large $N \approx 100k$ and large $P > 10^{10}$)
  - Low sample size ($P < 50$)
Joining with BayesFMM framework allows us to extend to more complex settings with correlated observations across subjects, nonparametric effects (in $X$), functional predictors, and to perform robust regression when outliers are present.

Applications of interest
- Climate change data
- Activity data/wearable computing
- Biomechanical optical data (250Hz for 6mo)

Extensions
- Extreme data ($N \approx 100k$ and large $P > 10^{10}$)
- Low sample size ($P < 50$)
- Spatially/temporally correlated distributions
Thank you.