Semiparametric Bayesian Latent Trajectory Models

Amy H. Herring, David B. Dunson, and Anna Maria Siega-Riz

August 8, 2007

Pregnancy weight gain

- Associated with both maternal and child health outcomes (National Research Council and Institute of Medicine (IOM), 2007)

Pregnancy weight gain

- Associated with both maternal and child health outcomes (National Research Council and Institute of Medicine (IOM), 2007)
- Maternal: postpartum weight retention (+ later obesity), gestational diabetes, duration of breastfeeding

Pregnancy weight gain

- Associated with both maternal and child health outcomes (National Research Council and Institute of Medicine (IOM), 2007)
- Maternal: postpartum weight retention (+ later obesity), gestational diabetes, duration of breastfeeding
- Child: birth weight, toddler at risk for overweight status, gestational age at delivery

Pregnancy weight gain

- Associated with both maternal and child health outcomes (National Research Council and Institute of Medicine (IOM), 2007)
- Maternal: postpartum weight retention (+ later obesity), gestational diabetes, duration of breastfeeding
- Child: birth weight, toddler at risk for overweight status, gestational age at delivery
  - We will concentrate on birth weight today

Pregnancy weight gain

- Associated with both maternal and child health outcomes (National Research Council and Institute of Medicine (IOM), 2007)
- Maternal: postpartum weight retention (+ later obesity), gestational diabetes, duration of breastfeeding
- Child: birth weight, toddler at risk for overweight status, gestational age at delivery
  - We will concentrate on birth weight today
- *Modifiable* risk factor

IOM recommendations for weight gain in pregancy

- Set to minimize risk of low birth weight

IOM recommendations for weight gain in pregancy

- Set to minimize risk of low birth weight
- Optimal for other outcomes?

IOM recommendations for weight gain in pregancy

- Set to minimize risk of low birth weight
- Optimal for other outcomes?
- Recommends normal weight women gain 25-35 pounds (more gain recommended for underweight women and less gain recommended for overweight and obese women)

IOM recommendations for weight gain in pregancy

- Set to minimize risk of low birth weight
- Optimal for other outcomes?
- Recommends normal weight women gain 25-35 pounds (more gain recommended for underweight women and less gain recommended for overweight and obese women)
  - Should pregnant woman who has gained 35 pounds 5 weeks before her due date give up the pleasure of eating ice cream at the local dairy farm?

PIN weight gain data

- PIN and PIN Postpartum studies (PI's: David Savitz and Anna Maria Siega-Riz) enrolled women at University of North Carolina Hospital in first and second trimesters of pregnancy

PIN weight gain data

- PIN and PIN Postpartum studies (PI's: David Savitz and Anna Maria Siega-Riz) enrolled women at University of North Carolina Hospital in first and second trimesters of pregnancy
- Pregnancy weight gain data: weight at each prenatal visit obtained from medical chart (mean of 11 pregnancy weights recorded); height measured; self-report of pregravid weight

PIN weight gain data

- PIN and PIN Postpartum studies (PI's: David Savitz and Anna Maria Siega-Riz) enrolled women at University of North Carolina Hospital in first and second trimesters of pregnancy
- Pregnancy weight gain data: weight at each prenatal visit obtained from medical chart (mean of 11 pregnancy weights recorded); height measured; self-report of pregravid weight
- Birth weight recorded

Goal of Analysis

Determine whether *pattern* and *amount* of pregnancy weight gain are related to birth weight (and in particular, risk of small-for-gestational age births).

Here, our outcome, birth weight, is univariate, and our predictor, weight gain, is longitudinal.

Complication: how to characterize pattern of pregnancy weight gain? Pregnancy weight measures are an example of a *functional predictor*: a random curve that varies over time, space, or some other domain, with observations at every point in the domain that may only be measured at a finite set of points (indexed by j).

General Problem & Data Structure

- Interest: relationship between functional predictor f_i & response y_i (*inference & prediction*)

General Problem & Data Structure

- Interest: relationship between functional predictor f_i & response y_i (*inference & prediction*)
- Predictor f_i takes value $f_i(t)$ at location $t \in \mathcal{T}$.

General Problem & Data Structure

- Interest: relationship between functional predictor f_i & response y_i (*inference & prediction*)
- Predictor f_i takes value $f_i(t)$ at location $t \in \mathcal{T}$.
- Data consist of $\{\mathbf{x}_i, y_i\}_{i=1}^n$, with $\mathbf{x}_i = (x_{i1}, \ldots, x_{i,n_i})'$.

General Problem & Data Structure

- Interest: relationship between functional predictor f_i & response y_i (*inference & prediction*)
- Predictor f_i takes value $f_i(t)$ at location $t \in \mathcal{T}$.
- Data consist of $\{\mathbf{x}_i, y_i\}_{i=1}^n$, with $\mathbf{x}_i = (x_{i1}, \ldots, x_{i,n_i})'$.
  - x_{ij} = error-prone measure of $f_i(t_{ij})$ (pregnancy weight gain at week t_{ij})

General Problem & Data Structure

- Interest: relationship between functional predictor f_i & response y_i (*inference & prediction*)
- Predictor f_i takes value $f_i(t)$ at location $t \in \mathcal{T}$.
- Data consist of $\{\mathbf{x}_i, y_i\}_{i=1}^n$, with $\mathbf{x}_i = (x_{i1}, \ldots, x_{i,n_i})'$.
  - x_{ij} = error-prone measure of $f_i(t_{ij})$ (pregnancy weight gain at week t_{ij})
  - t_{ij} = location (or time) of observation j

General Problem & Data Structure

- Interest: relationship between functional predictor f_i & response y_i (*inference & prediction*)
- Predictor f_i takes value $f_i(t)$ at location $t \in \mathcal{T}$.
- Data consist of $\{\mathbf{x}_i, y_i\}_{i=1}^n$, with $\mathbf{x}_i = (x_{i1}, \ldots, x_{i,n_i})'$.
  - x_{ij} = error-prone measure of $f_i(t_{ij})$ (pregnancy weight gain at week t_{ij})
  - t_{ij} = location (or time) of observation j
  - n_i = number of observations on subject i

General Problem & Data Structure

- Interest: relationship between functional predictor f_i & response y_i (*inference & prediction*)
- Predictor f_i takes value $f_i(t)$ at location $t \in \mathcal{T}$.
- Data consist of $\{\mathbf{x}_i, y_i\}_{i=1}^n$, with $\mathbf{x}_i = (x_{i1}, \ldots, x_{i,n_i})'$.
  - x_{ij} = error-prone measure of $f_i(t_{ij})$ (pregnancy weight gain at week t_{ij})
  - t_{ij} = location (or time) of observation j
  - n_i = number of observations on subject i
  - y_i = response variable (infant birth weight)

Latent class trajectory model

Group-based trajectory models are used to identify clusters of subjects following similar trajectories over time.

While we may not believe that each subject's weight gain measures in pregnancy exactly follow one of C curves, this may be a very useful summary of the data.

Issues

- Good choice of parametric form for latent trajectory curves often unclear (prefer a nonparametric form?)

Issues

- Good choice of parametric form for latent trajectory curves often unclear (prefer a nonparametric form?)
- iid $N(0, \sigma^2)$ residuals restrictive - may imply lots of latent classes

Issues

- Good choice of parametric form for latent trajectory curves often unclear (prefer a nonparametric form?)
- iid $N(0, \sigma^2)$ residuals restrictive - may imply lots of latent classes
- Number of latent classes unknown - BIC criteria may be poor

Our Interest

- Develop semiparametric Bayes joint modeling framework

Our Interest

- Develop semiparametric Bayes joint modeling framework
- For ease in interpretation, group individuals into functional predictor clusters (patterns of pregnancy weight gain), with number of clusters not specified in advance

Our Interest

- Develop semiparametric Bayes joint modeling framework
- For ease in interpretation, group individuals into functional predictor clusters (patterns of pregnancy weight gain), with number of clusters not specified in advance
- Allow response (birth weight) distribution to vary nonparametrically across clusters

Our Interest

- Develop semiparametric Bayes joint modeling framework
- For ease in interpretation, group individuals into functional predictor clusters (patterns of pregnancy weight gain), with number of clusters not specified in advance
- Allow response (birth weight) distribution to vary nonparametrically across clusters
- Conduct inferences on changes in quantiles of birth weight

Semiparametric Hierarchical Model

$$
\begin{aligned}
y_i &\sim \mathrm{N}(\mu_i, \tau_y^{-1}), \ i = 1, \ldots, n, \\
x_i(t_{ij}) &\sim \mathrm{N}\left(f_i(t_{ij}), \tau_x^{-1}\right), \ j = 1, \ldots, n_i, \\
(f_i, \mu_i) &\sim Q,
\end{aligned}
$$

- Dependence between functional predictor $f_i \in \Omega$ & response $y_i \in \Re$ characterized through Q
- Q=random probability measure

Hierarchical Basis Function Representation

- To simplify modeling, we express the function f_i as

$$f_i(t) = \sum_{m=1}^{p} \beta_{im} b_m(t), \quad \forall t \in \mathcal{T}.$$

Hierarchical Basis Function Representation

- To simplify modeling, we express the function f_i as

$$f_i(t) = \sum_{m=1}^{p} \beta_{im} b_m(t), \quad \forall t \in \mathcal{T}.$$

  - We use a cubic spline model with knots at the trimesters, motivated by recommendations for weight gain:

$$f_i(t) = \beta_{i1} t + \beta_{i2} t^2 + \beta_{i3} t^3 + \beta_{i4} (t-13)_+^{\ 3} + \beta_{i5} (t-26)_+^{\ 3}$$

Hierarchical Basis Function Representation

- To simplify modeling, we express the function f_i as

$$f_i(t) = \sum_{m=1}^{p} \beta_{im} b_m(t), \quad \forall t \in \mathcal{T}.$$

  - We use a cubic spline model with knots at the trimesters, motivated by recommendations for weight gain:

$$f_i(t) = \beta_{i1} t + \beta_{i2} t^2 + \beta_{i3} t^3 + \beta_{i4} (t-13)_+^{\ 3} + \beta_{i5} (t-26)_+^{\ 3}$$

- $\boldsymbol{\beta}_i = (\beta_{i1}, \ldots, \beta_{ip})'$ are subject-specific basis coefficients.

Hierarchical Basis Function Representation

- To simplify modeling, we express the function f_i as

$$f_i(t) = \sum_{m=1}^{p} \beta_{im} b_m(t), \quad \forall t \in \mathcal{T}.$$

  - We use a cubic spline model with knots at the trimesters, motivated by recommendations for weight gain:

$$f_i(t) = \beta_{i1} t + \beta_{i2} t^2 + \beta_{i3} t^3 + \beta_{i4} (t-13)_+^{\ 3} + \beta_{i5} (t-26)_+^{\ 3}$$

- $\boldsymbol{\beta}_i = (\beta_{i1}, \ldots, \beta_{ip})'$ are subject-specific basis coefficients.
- Let $\boldsymbol{\theta}_i = (\boldsymbol{\beta}_i', \mu_i) \sim P$, with P a probability measure on $(\Re^{p+1}, \mathcal{B})$ $[\pi(P) \rightarrow \pi(Q)]$

Hierarchical Basis Function Representation

- To simplify modeling, we express the function f_i as

$$f_i(t) = \sum_{m=1}^{p} \beta_{im} b_m(t), \quad \forall t \in \mathcal{T}.$$

  - We use a cubic spline model with knots at the trimesters, motivated by recommendations for weight gain:

$$f_i(t) = \beta_{i1} t + \beta_{i2} t^2 + \beta_{i3} t^3 + \beta_{i4} (t-13)_+^{\ 3} + \beta_{i5} (t-26)_+^{\ 3}$$

- $\boldsymbol{\beta}_i = (\beta_{i1}, \ldots, \beta_{ip})'$ are subject-specific basis coefficients.
- Let $\boldsymbol{\theta}_i = (\boldsymbol{\beta}_i', \mu_i) \sim P$, with P a probability measure on $(\Re^{p+1}, \mathcal{B})$ $[\pi(P) \rightarrow \pi(Q)]$
- Dependence between f_i and y_i is characterized through P

Dirichlet Process Joint Models

- A natural approach is to let P be unknown with $P \sim DP(\alpha P_0)$

Dirichlet Process Joint Models

- A natural approach is to let P be unknown with $P \sim DP(\alpha P_0)$
- $DP(\alpha P_0) =$ denotes the Dirichlet process with precision α & base probability measure P_0

Dirichlet Process Joint Models

- A natural approach is to let P be unknown with $P \sim DP(\alpha P_0)$
- $DP(\alpha P_0) =$ denotes the Dirichlet process with precision α & base probability measure P_0
- This DP prior induces the following Blackwell & MacQueen (1973) rule:

$$(\boldsymbol{\theta}_i \,|\, \boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_{i-1}) \sim \left(\frac{\alpha}{\alpha + i - 1}\right) P_0 + \sum_{j=1}^{i-1} \left(\frac{1}{\alpha + i - 1}\right) \delta_{\boldsymbol{\theta}_j},$$

$\delta_\theta =$ measure concentrated at θ.

Comments on DP Joint Models

- Subjects automatically grouped into an unknown number of functional trajectory clusters.

Comments on DP Joint Models

- Subjects automatically grouped into an unknown number of functional trajectory clusters.
- Cluster h has functional trajectory $f(t) = \beta_h^{*'}\mathbf{b}(t)$ & response density $N(\mu_h^*, \tau_x^{-1})$

Comments on DP Joint Models

- Subjects automatically grouped into an unknown number of functional trajectory clusters.
- Cluster h has functional trajectory $f(t) = \beta_h^{*'}\mathbf{b}(t)$ & response density $N(\mu_h^*, \tau_x^{-1})$
- Marginal density of y is a mixture of normals - flexible

Comments on DP Joint Models

- Subjects automatically grouped into an unknown number of functional trajectory clusters.
- Cluster h has functional trajectory $f(t) = \beta_h^{*'}\mathbf{b}(t)$ & response density $N(\mu_h^*, \tau_x^{-1})$
- Marginal density of y is a mixture of normals - flexible
- Within predictor cluster, density of y is a single normal - inflexible!

Comments on DP Joint Models

- Subjects automatically grouped into an unknown number of functional trajectory clusters.
- Cluster h has functional trajectory $f(t) = \beta_h^{*'}\mathbf{b}(t)$ & response density $N(\mu_h^*, \tau_x^{-1})$
- Marginal density of y is a mixture of normals - flexible
- Within predictor cluster, density of y is a single normal - inflexible!
- Does not allow flexible inferences on changes in risk of low and high birth weight (quantiles) over weight gain clusters (dictated by shape of normal distribution)

Comments on DP Joint Models

- The DP assumes identical clusters in the predictor & response

Comments on DP Joint Models

- The DP assumes identical clusters in the predictor & response
- It is appealing to consider alternatives that allow separate, but dependent clustering

Predictor Component Model

- First, we focus on the functional predictor portion of the model

Predictor Component Model

- First, we focus on the functional predictor portion of the model
- We need a joint distribution on the basis coefficients, characterized by the probability measure P

Predictor Component Model

- First, we focus on the functional predictor portion of the model
- We need a joint distribution on the basis coefficients, characterized by the probability measure P
  - We cluster individuals into unique groups, with each group h characterized by unique values of the basis coefficients ${\beta_h}^*$

Predictor Component Model

- We start with the following model for the basis coefficients:

$$\beta_i \sim P_1 = \sum_{h=1}^{H} \pi_h \delta_{\beta_h^*}, \ \beta_h^* \overset{iid}{\sim} P_{01},$$

$$\boldsymbol{\pi} = (\pi_1, \ldots, \pi_H)' \sim \text{Diri}\left(\frac{\alpha}{H}, \ldots, \frac{\alpha}{H}\right),$$

Predictor Component Model

- We start with the following model for the basis coefficients:

$$\beta_i \sim P_1 = \sum_{h=1}^{H} \pi_h \delta_{\beta_h^*}, \ \beta_h^* \overset{iid}{\sim} P_{01},$$

$$\boldsymbol{\pi} = (\pi_1, \ldots, \pi_H)' \sim \text{Diri}\left(\frac{\alpha}{H}, \ldots, \frac{\alpha}{H}\right),$$

- H is an upper bound on the number of predictor clusters

Predictor Component Model

- We start with the following model for the basis coefficients:

$$\beta_i \sim P_1 = \sum_{h=1}^{H} \pi_h \delta_{\beta_h^*}, \ \beta_h^* \overset{iid}{\sim} P_{01},$$

$$\boldsymbol{\pi} = (\pi_1, \ldots, \pi_H)' \sim \text{Diri}\left(\frac{\alpha}{H}, \ldots, \frac{\alpha}{H}\right),$$

- H is an upper bound on the number of predictor clusters
- We draw cluster specific coefficients from the base measure P_{01}

Predictor Component Model

- We start with the following model for the basis coefficients:

$$\beta_i \sim P_1 = \sum_{h=1}^{H} \pi_h \delta_{\beta_h^*}, \ \beta_h^* \overset{iid}{\sim} P_{01},$$

$$\boldsymbol{\pi} = (\pi_1, \ldots, \pi_H)' \sim \text{Diri}\left(\frac{\alpha}{H}, \ldots, \frac{\alpha}{H}\right),$$

- H is an upper bound on the number of predictor clusters
- We draw cluster specific coefficients from the base measure P_{01}
- Weights π have a Dirichlet prior, so that $\lim_{H \to \infty} P_1 \sim DP(\alpha P_{01})$.

Predictor Component Model

- Equivalently expressed as

$$\beta_i \sim \delta_{\beta^*_{\gamma_i}}, \quad \gamma_i \sim \sum_{h=1}^{H} \pi_h \delta_h.$$

Predictor Component Model

- Equivalently expressed as

$$\beta_i \sim \delta_{\beta^*_{\gamma_i}}, \quad \gamma_i \sim \sum_{h=1}^{H} \pi_h \delta_h.$$

- $\gamma_i \in \{1, \ldots, H\}$=predictor cluster index for subject i, so $\gamma_i = h$ with probability π_h

Response Component Model

- Now, consider the response component of the model

Response Component Model

- Now, consider the response component of the model
- Conditional on membership in weight gain cluster h, we could use the same structure we used for the weight gain trajectories, within each weight gain cluster

Response Component Model

- Now, consider the response component of the model
- Conditional on membership in weight gain cluster h, we could use the same structure we used for the weight gain trajectories, within each weight gain cluster
- The drawback is that we wouldn't be borrowing information about the distribution of birth weights across weight gain clusters (in some sense we would be forcing differences in the response distribution)

Response Component Model

- Now, consider the response component of the model
- Conditional on membership in weight gain cluster h, we could use the same structure we used for the weight gain trajectories, within each weight gain cluster
- The drawback is that we wouldn't be borrowing information about the distribution of birth weights across weight gain clusters (in some sense we would be forcing differences in the response distribution)
- In order to have cluster-specific response distributions while also borrowing information across clusters, we use a mixture specification with both global and local components

Response Component Model

- To allow response distribution to change flexibly across predictor clusters, instead of using a single Dirichlet prior, we use a mixture of local and global components

$$\begin{aligned}(\mu_i \,|\, \gamma_i = h) &\sim P_{2h} = \sum_{l=1}^{L} \nu_{hl}\delta_{\mu_l^*}, \quad \mu_l^* \overset{iid}{\sim} P_{02}, \\ \boldsymbol{\nu}_h &= (\nu_{h1}, \ldots, \nu_{hL})' \sim (1-\psi)\delta_{\boldsymbol{\nu}_0^*} + \psi\delta_{\boldsymbol{\nu}_h^*}, \\ \boldsymbol{\nu}_h^* &\overset{iid}{\sim} \text{Diri}\left(\frac{\lambda}{L}, \ldots, \frac{\lambda}{L}\right), \; h = 0, 1, \ldots, H,\end{aligned}$$

- Implies $\lim_{L\to\infty} P_{2h} \sim (1-\psi)P_{20}^* + \psi P_{2h}^*$, with $P_{2h}^* \sim DP(\lambda P_{02})$ independently for $h = 0, 1, \ldots, H$.
- P_{20}^* = *global* component (allows borrowing of information)
- P_{2h}^* = *local* component specific to predictor cluster h

Some Properties

- Marginalizing out the random weights and atoms, we can obtain a bivariate generalization of the DP prediction rule

Some Properties

- Marginalizing out the random weights and atoms, we can obtain a bivariate generalization of the DP prediction rule
- The conditional prior probability of response clustering for i and j is

$$\Pr(\mu_i = \mu_j \,|\, \gamma_i, \gamma_j) = \frac{\lambda/L + 1}{\lambda + 1}\{\psi^2 1_{(\gamma_i = \gamma_j)} + (1-\psi)^2\},$$

Some Properties

- Marginalizing out the random weights and atoms, we can obtain a bivariate generalization of the DP prediction rule
- The conditional prior probability of response clustering for i and j is

$$\Pr(\mu_i = \mu_j \,|\, \gamma_i, \gamma_j) = \frac{\lambda/L + 1}{\lambda + 1}\{\psi^2 1_{(\gamma_i = \gamma_j)} + (1-\psi)^2\},$$

- In limit as $\psi \to 0$, predictor and response clustering occur independently (global component dominant and no association between weight gain and birth weight)

Some Properties

- Marginalizing out the random weights and atoms, we can obtain a bivariate generalization of the DP prediction rule
- The conditional prior probability of response clustering for i and j is

$$\Pr(\mu_i = \mu_j \,|\, \gamma_i, \gamma_j) = \frac{\lambda/L + 1}{\lambda + 1}\{\psi^2 1_{(\gamma_i = \gamma_j)} + (1-\psi)^2\},$$

- In limit as $\psi \to 0$, predictor and response clustering occur independently (global component dominant and no association between weight gain and birth weight)
- In limit as $\psi \to 1$, $\Pr(\mu_i = \mu_j) = 0$ if i and j are not in same predictor cluster (local component dominant so that there are very strong differences in birth weight among weight gain clusters)

Posterior Computation

- Gibbs sampling is straightforward to implement, involving simple steps for sampling from standard distributions

Posterior Computation

- Gibbs sampling is straightforward to implement, involving simple steps for sampling from standard distributions
- Highly computationally intensive in our application ($22,467$ observations on $1,888$ subjects)

Posterior Computation

- Gibbs sampling is straightforward to implement, involving simple steps for sampling from standard distributions
- Highly computationally intensive in our application ($22,467$ observations on $1,888$ subjects)
- Label switching a pain! - can be addressed by post-processing (*Stephens, 2000*), but adds considerably to computational burden

Posterior Computation

- Gibbs sampling is straightforward to implement, involving simple steps for sampling from standard distributions
- Highly computationally intensive in our application ($22,467$ observations on $1,888$ subjects)
- Label switching a pain! - can be addressed by post-processing (*Stephens, 2000*), but adds considerably to computational burden
- Fast alternatives are needed

Combined SUGS and Gibbs sampling

- Wang & Dunson (2007) proposed a fast sequential updating & greedy search (SUGS) algorithm for posterior computation in DP mixture models (see also Zhang, Ghahramani, and Yang, 2005)

Combined SUGS and Gibbs sampling

- Wang & Dunson (2007) proposed a fast sequential updating & greedy search (SUGS) algorithm for posterior computation in DP mixture models (see also Zhang, Ghahramani, and Yang, 2005)
- Excellent performance for Bayes density estimation

Combined SUGS and Gibbs sampling

- Wang & Dunson (2007) proposed a fast sequential updating & greedy search (SUGS) algorithm for posterior computation in DP mixture models (see also Zhang, Ghahramani, and Yang, 2005)
- Excellent performance for Bayes density estimation
- Can be implemented within a few seconds even in moderately large data sets

Combined SUGS and Gibbs sampling

- Wang & Dunson (2007) proposed a fast sequential updating & greedy search (SUGS) algorithm for posterior computation in DP mixture models (see also Zhang, Ghahramani, and Yang, 2005)
- Excellent performance for Bayes density estimation
- Can be implemented within a few seconds even in moderately large data sets
- We use SUGS for clustering of the functional predictor data

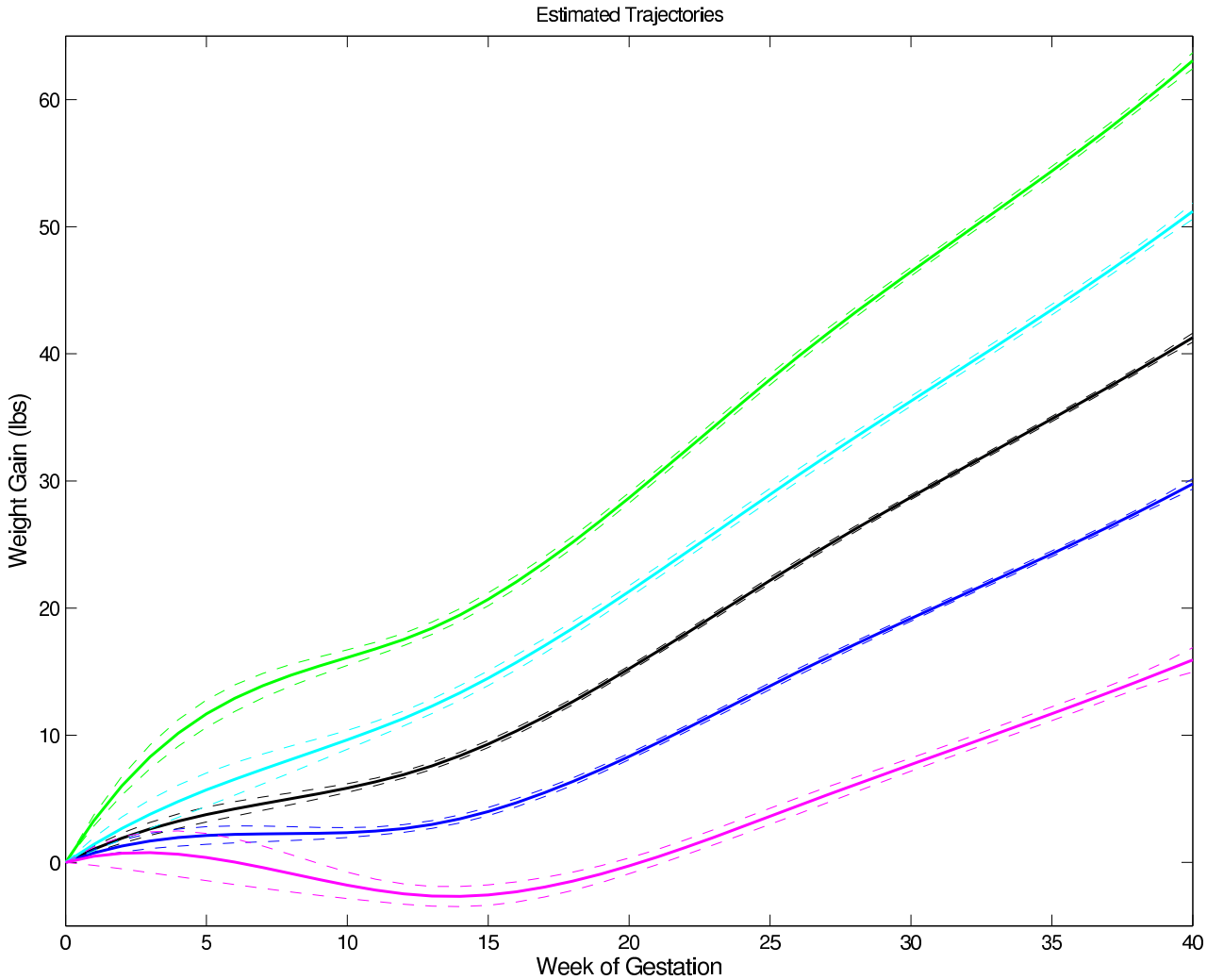
Combined SUGS and Gibbs sampling

- Wang & Dunson (2007) proposed a fast sequential updating & greedy search (SUGS) algorithm for posterior computation in DP mixture models (see also Zhang, Ghahramani, and Yang, 2005)
- Excellent performance for Bayes density estimation
- Can be implemented within a few seconds even in moderately large data sets
- We use SUGS for clustering of the functional predictor data
- Gibbs sampling is then used for posterior computation in the conditional response component model

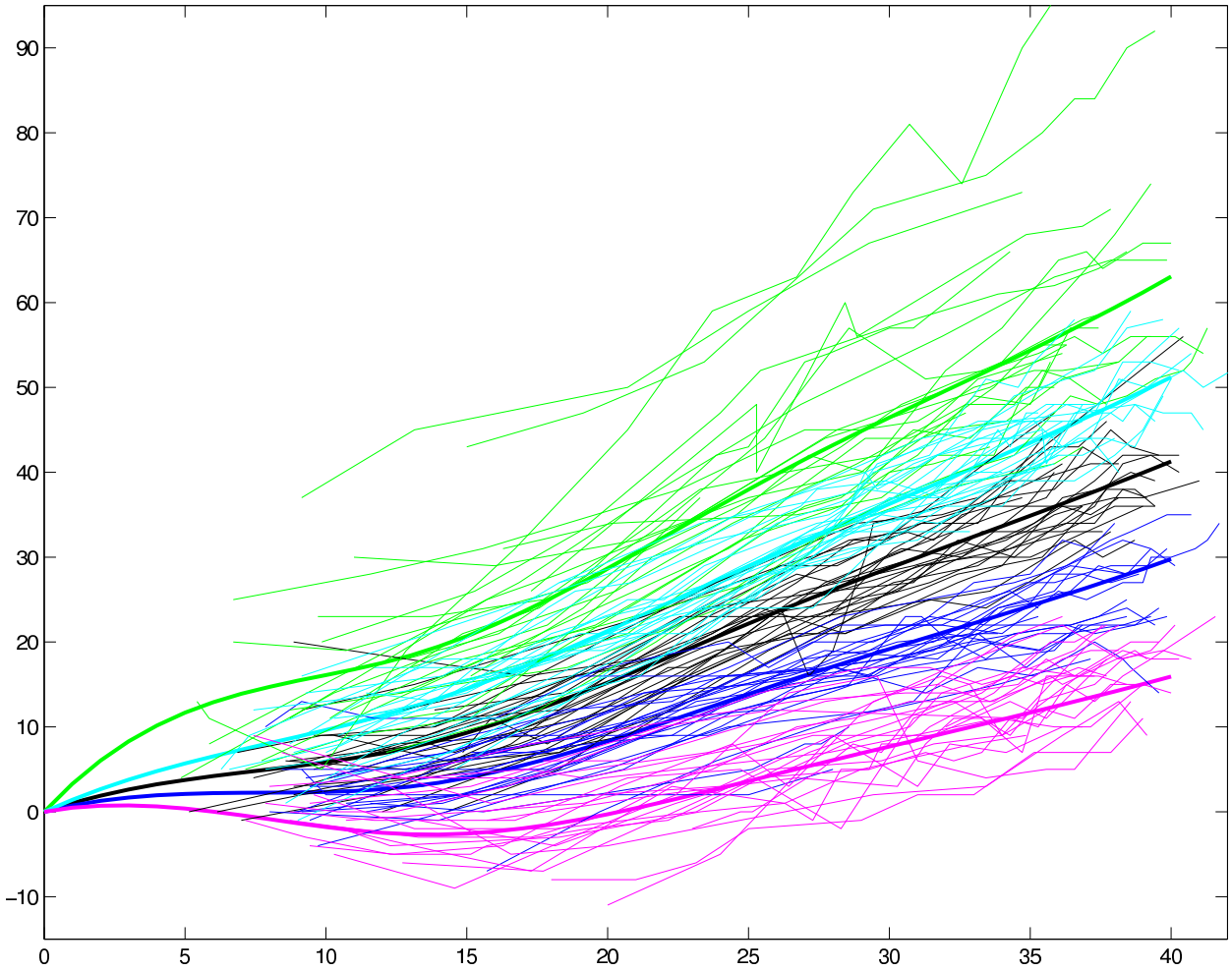
Combined SUGS and Gibbs sampling

- Wang & Dunson (2007) proposed a fast sequential updating & greedy search (SUGS) algorithm for posterior computation in DP mixture models (see also Zhang, Ghahramani, and Yang, 2005)
- Excellent performance for Bayes density estimation
- Can be implemented within a few seconds even in moderately large data sets
- We use SUGS for clustering of the functional predictor data
- Gibbs sampling is then used for posterior computation in the conditional response component model
- Avoids label ambiguity and is quite efficient

Estimated trajectories



Estimated and observed trajectories

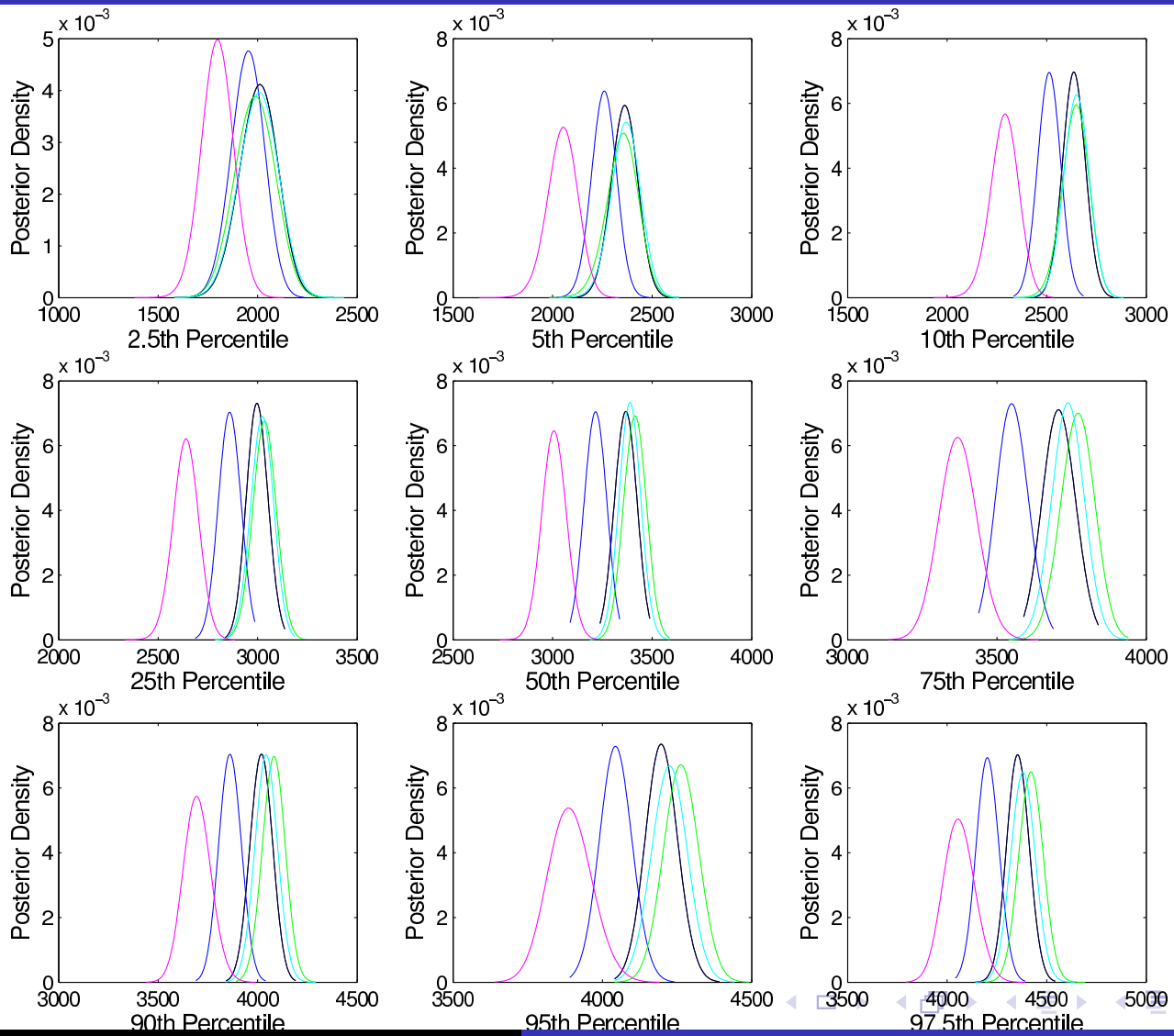


Birth weight results

Next we present plots of the quantiles of the birth weight distribution for each cluster. The posterior distribution of each quantile in the group $\{2.5, 5, 10, 25, 50, 75, 90, 95, 97.5\}$ is plotted by cluster (adjusted for gestational age at delivery).

Note that using normality within a cluster would imply that the 9 plots on the following page should be identical.

Quantile plots



Association between weight gain and low birth weight

Using 2500g as a cutoff for low birth weight (standard),

- Women in lowest weight gain cluster had 17.2% low birth weight babies

Association between weight gain and low birth weight

Using 2500g as a cutoff for low birth weight (standard),

- Women in lowest weight gain cluster had 17.2% low birth weight babies
- Women in higher weight gain clusters had 9.3, 6.8, 6.6, and 6.8% low birth weight babies, respectively

Summary

- Flexible Bayes nonparametric latent trajectory model

Summary

- Flexible Bayes nonparametric latent trajectory model
- Allows response distribution to vary nonparametrically with functional predictors

Summary

- Flexible Bayes nonparametric latent trajectory model
- Allows response distribution to vary nonparametrically with functional predictors
- Approach useful in other applications - e.g., joint modeling of images & disease status