Research Group Presentation

Nicolas Escobar, Annie Yu, Roger Zoh

Overview of Bayesian analysis

Motivating example: Likelihood

- Suppose we have Data

$$P(\Theta|Data) \propto \ell(Data|\Theta)P(\Theta)$$

- Advantages - Full joint distribution for Θ - You can do virtually everything - Pooling (full/Partial/and no)

Overview of Bayesian analysis

Motivating example: Likelihood

- Suppose we have Data
- Choose our model $\rightarrow$ likelihood $\rightarrow$ $\ell(Data|\Theta)$

$$P(\Theta|Data) \propto \ell(Data|\Theta)P(\Theta)$$

- Advantages - Full joint distribution for Θ - You can do virtually everything - Pooling (full/Partial/and no)

Overview of Bayesian analysis

Motivating example: Likelihood

- Suppose we have Data
- Choose our model $\rightarrow$ likelihood $\rightarrow \ell(Data|\Theta)$
- Find point estimates of Θ and Extra-step for uncertainty (Asymptotic theory/Bootstrap/etc.)

$$P(\Theta|Data) \propto \ell(Data|\Theta)P(\Theta)$$

- Advantages - Full joint distribution for Θ - You can do virtually everything - Pooling (full/Partial/and no)

Overview of Bayesian analysis

Motivating example: Likelihood

- Suppose we have Data
- Choose our model $\rightarrow$ likelihood $\rightarrow \ell(Data|\Theta)$
- Find point estimates of Θ and Extra-step for uncertainty (Asymptotic theory/Bootstrap/etc.)
- Go Bayesian

$$P(\Theta|Data) \propto \ell(Data|\Theta)P(\Theta)$$

- Advantages - Full joint distribution for Θ - You can do virtually everything - Pooling (full/Partial/and no)

Parameter Estimation

Overview of Stan

- Likelihood above can be very complicated, so sampling from the posterior distribution of parameters becomes an issue

Parameter Estimation

Overview of Stan

- Likelihood above can be very complicated, so sampling from the posterior distribution of parameters becomes an issue
- Usually, we deal with this using Markov Chain Monte Carlo (MCMC) techniques

Parameter Estimation

Overview of Stan

- Likelihood above can be very complicated, so sampling from the posterior distribution of parameters becomes an issue
- Usually, we deal with this using Markov Chain Monte Carlo (MCMC) techniques
- Coding MCMCs by hand is cumbersome. It can also lead to convergence issues

Parameter Estimation

Overview of Stan

- Likelihood above can be very complicated, so sampling from the posterior distribution of parameters becomes an issue
- Usually, we deal with this using Markov Chain Monte Carlo (MCMC) techniques
- Coding MCMCs by hand is cumbersome. It can also lead to convergence issues
- Software packages like BUGS and JAGS can automatically code MCMCs based on model structure

Parameter Estimation

Overview of Stan

- Likelihood above can be very complicated, so sampling from the posterior distribution of parameters becomes an issue
- Usually, we deal with this using Markov Chain Monte Carlo (MCMC) techniques
- Coding MCMCs by hand is cumbersome. It can also lead to convergence issues
- Software packages like BUGS and JAGS can automatically code MCMCs based on model structure
- Stan [Stan development team, 2012] improves on those packages by implementing HMC and NUTS

Parameter Estimation

Overview of Stan

- Likelihood above can be very complicated, so sampling from the posterior distribution of parameters becomes an issue
- Usually, we deal with this using Markov Chain Monte Carlo (MCMC) techniques
- Coding MCMCs by hand is cumbersome. It can also lead to convergence issues
- Software packages like BUGS and JAGS can automatically code MCMCs based on model structure
- Stan [Stan development team, 2012] improves on those packages by implementing HMC and NUTS
- Still need to becareful about your prior and even your model structure

Overview of BRMS

- Stan is low level: it's a programming language on it's own, built on top of C++

Overview of BRMS

- Stan is low level: it's a programming language on it's own, built on top of C++
- BRMS [Buckner, 2017] is a high level interface for Stan on R

Overview of BRMS

- Stan is low level: it's a programming language on it's own, built on top of C++
- BRMS [Buckner, 2017] is a high level interface for Stan on R
- It uses the same formula syntax as `lme4`

Overview of BRMS

- Stan is low level: it's a programming language on it's own, built on top of C++
- BRMS [Buckner, 2017] is a high level interface for Stan on R
- It uses the same formula syntax as `lme4`
- Under the hood, it automatically generates and compiles Stan code, which we can retrieve

Quantile regression

General Framework

- OLS regression: $E[Y|\mathbf{X}] = \mathbf{X}^T \boldsymbol{\beta}$

Quantile regression

General Framework

- OLS regression: $E[Y|\mathbf{X}] = \mathbf{X}^T\boldsymbol{\beta}$
- Quantile regression: $q_p(Y|\mathbf{X}) = \mathbf{X}^T\boldsymbol{\beta}(p)$

Quantile regression

General Framework

- OLS regression: $E[Y|\mathbf{X}] = \mathbf{X}^T\boldsymbol{\beta}$
- Quantile regression: $q_p(Y|\mathbf{X}) = \mathbf{X}^T\boldsymbol{\beta}(p)$
- Frequentist approach:

$$\widehat{\boldsymbol{\beta}(p)} = \mathrm{argmin}_{\boldsymbol{\beta}} \sum_i \rho_p(y_i - \mathbf{x}_i^T\boldsymbol{\beta})$$

where $\rho_p(u) = u(p - \mathbf{1}_{\{u<0\}})$

Bayesian quantile regression with AL

- It can be shown that the previous minimization problem is equivalent to maximizing the likelihood associated to the distribution

$$f_p^{\mathrm{AL}}(u) = p(1-p)\exp(-\rho_p(u))$$

Bayesian quantile regression with AL

- It can be shown that the previous minimization problem is equivalent to maximizing the likelihood associated to the distribution

$$f_p^{\mathrm{AL}}(u) = p(1-p)\exp(-\rho_p(u))$$

- More generally, we can add a location and a scale parameter:

$$f_p^{\mathrm{AL}}(u|\mu,\sigma) = p(1-p)\exp(-\rho_p((u-\mu)/\sigma))$$

this distribution is called *asymmetric Laplace* [Yu and Moyeed, 2001]

Bayesian quantile regression with AL

- It can be shown that the previous minimization problem is equivalent to maximizing the likelihood associated to the distribution

$$f_p^{\mathrm{AL}}(u) = p(1-p)\exp(-\rho_p(u))$$

- More generally, we can add a location and a scale parameter:

$$f_p^{\mathrm{AL}}(u|\mu,\sigma) = p(1-p)\exp(-\rho_p((u-\mu)/\sigma))$$

  this distribution is called *asymmetric Laplace* [Yu and Moyeed, 2001]

- Prior: $p(\boldsymbol{\beta}) \propto 1$ (improper)

Bayesian quantile regression with AL

- It can be shown that the previous minimization problem is equivalent to maximizing the likelihood associated to the distribution

$$f_p^{\mathrm{AL}}(u) = p(1-p)\exp(-\rho_p(u))$$

- More generally, we can add a location and a scale parameter:

$$f_p^{\mathrm{AL}}(u|\mu,\sigma) = p(1-p)\exp(-\rho_p((u-\mu)/\sigma))$$

  this distribution is called *asymmetric Laplace* [Yu and Moyeed, 2001]
- Prior: $p(\boldsymbol{\beta}) \propto 1$ (improper)
- This leads to a proper posterior distribution (only after some additional post-analysis steps)

Flexible quantile regression with GAL

- Problems: AL is very rigid

Flexible quantile regression with GAL

- Problems: AL is very rigid
  - Skewness is determined by p, in particular it is symmetric at $p = .5$

Flexible quantile regression with GAL

- Problems: AL is very rigid
  - Skewness is determined by p, in particular it is symmetric at $p = .5$
  - Mode is always at 0

Flexible quantile regression with GAL

- Problems: AL is very rigid
  - Skewness is determined by p, in particular it is symmetric at $p = .5$
  - Mode is always at 0
- Solution [Yan and Kottas, 2017]:

Flexible quantile regression with GAL

- Problems: AL is very rigid
  - Skewness is determined by p, in particular it is symmetric at $p = .5$
  - Mode is always at 0
- Solution [Yan and Kottas, 2017]:
  - We can write

$$f_p^{\mathrm{AL}}(y|\mu,\sigma) = \int_{\mathbb{R}} N(y|\mu + \sigma A(p)z, \sigma^2 B(p)z) Exp(z|1) dz$$

Flexible quantile regression with GAL

- Problems: AL is very rigid
  - Skewness is determined by p, in particular it is symmetric at $p = .5$
  - Mode is always at 0
- Solution [Yan and Kottas, 2017]:
  - We can write

$$f_p^{\mathrm{AL}}(y|\mu,\sigma) = \int_{\mathbb{R}} N(y|\mu + \sigma A(p)z, \sigma^2 B(p)z) Exp(z|1) dz$$

  - So, let us add a parameter

$$f_p^{\mathrm{GAL}}(y|\mu,\sigma)$$

$$= \int_{\mathbb{R}} N(y|\mu + \sigma\alpha s + \sigma A(p)z, \sigma^2 B(p)z) Exp(z|1) N^+(s|0,1) dz ds$$

Flexible quantile regression with GAL

- Problems: AL is very rigid
  - Skewness is determined by p, in particular it is symmetric at $p = .5$
  - Mode is always at 0
- Solution [Yan and Kottas, 2017]:
  - We can write

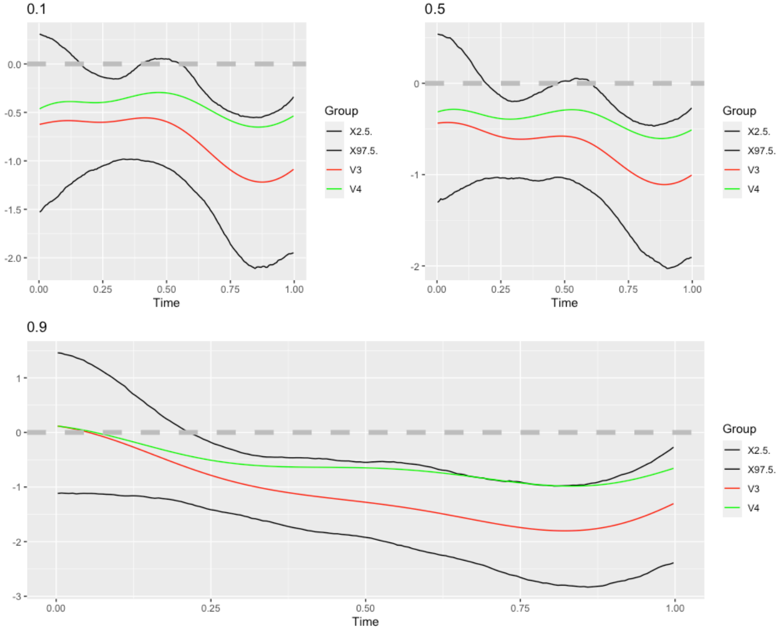
$$f_p^{\mathrm{AL}}(y|\mu,\sigma) = \int_{\mathbb{R}} N(y|\mu + \sigma A(p)z, \sigma^2 B(p)z) Exp(z|1) dz$$

  - So, let us add a parameter

$$f_p^{\mathrm{GAL}}(y|\mu,\sigma)$$

$$= \int_{\mathbb{R}} N(y|\mu+\sigma\alpha s+\sigma A(p)z, \sigma^2 B(p)z) Exp(z|1) N^+(s|0,1) dz ds$$

  - We even allow for a mixture of GAL as a likelihood (Dirichlet process mixture)


0.1
0.0
-0.5
-1.0
-1.5
-2.0
0.00
0.25
0.50
0.75
1.00
Time
Group
X2.5.
X97.5.
V3
V4
0.5
0
-1
-2
0.00
0.25
0.50
0.75
1.00
Time
Group
X2.5.
X97.5.
V3
V4
0.9
1
0
-1
-2
-3
0.00
0.25
0.50
0.75
1.00
Time
Group
X2.5.
X97.5.
V3
V4

Project we are working on

- Bayesian quantile regression with functional covariates with ME using GAL

Project we are working on

- Bayesian quantile regression with functional covariates with ME using GAL
- Bayesian quantile regression with functional covariates with ME using GAL with variance function

Project we are working on

- Bayesian quantile regression with functional covariates with ME using GAL
- Bayesian quantile regression with functional covariates with ME using GAL with variance function
- Bayesian quantile regression with censored outcomes

Project we are working on

- Bayesian quantile regression with functional covariates with ME using GAL
- Bayesian quantile regression with functional covariates with ME using GAL with variance function
- Bayesian quantile regression with censored outcomes
- A Bayesian medation model with multiple outcome, latent mediator and a function covariate measured with error