STAT40380/STAT40390/STAT40850
Bayesian Analysis

Dr Niamh Russell

School of Mathematics and Statistics
University College Dublin

niamh.russell@ucd.ie

February 2016

Simple linear regression models with Bayes

- We are used to fitting linear regression models of the form:

$$y_i \sim N(\alpha + \beta(x_i - \bar{x}), \phi)$$

  for paired data (x_i, y_i) with x_i as an explanatory variable and y_i as the response

- We often estimate the parameter set $(\alpha, \beta, \sigma^2($ or $\phi))$ through least squares (or maximum lilkelihood) to obtain:

$$\hat{\alpha} = \bar{y}, \hat{\beta} = \frac{\sum(y_i - \bar{y})(x_i - \bar{x})}{\sum(x_i - \bar{x})^2}, \hat{\phi} = \frac{1}{n-2}\sum(y_i - \bar{y})^2$$

- The model can be extended by considering extra explanatory variables, or by making the variance term depend on x_i, amongst others

Bayesian linear regression

- To perform Bayesian linear regression we need to specify the posterior:
$$p(\alpha, \beta, \phi | \mathbf{y}, \mathbf{x}) \propto p(\alpha, \beta, \phi) \prod_{i=1}^{n} p(y_i | x_i, \alpha, \beta, \phi)$$
- ...which depends on a prior distribution for $p(\alpha, \beta, \phi)$.
- Possible choices include:
  1. Independent priors, eg $\alpha \sim N(2, 3), \beta \sim N(1.5, 0.7), \phi \sim lognormal(4.2, 0.1)$, so $p(\alpha, \beta, \phi) \propto p(\alpha) \times p(\beta) \times p(\phi)$
  2. Dependent priors, eg
$$\begin{bmatrix} \alpha \\ \beta \\ \log(\phi) \end{bmatrix} \sim N\left(\begin{bmatrix} 2 \\ 3 \\ 4.2 \end{bmatrix}, \begin{bmatrix} 3 & 0 & -0.02 \\ 0 & 0.7 & -0.05 \\ -0.02 & -0.05 & 0.1 \end{bmatrix} \right)$$
- We will used (to start with) a reference prior, $p(\alpha, \beta, \phi) \propto \frac{1}{\phi}$
- Note that the probability distributions assume that the explanatory variables **x** are always fixed and thus have no probability distribution

Example 1

Example

Use the reference prior $p(\alpha, \beta, \phi) \propto \frac{1}{\phi}$ to find the joint posterior distribution of $\alpha, \beta, \phi | \mathbf{y}, \mathbf{x}$ in the simple linear regression model. Use the joint posterior distribution to find the marginal posterior distributions for $\alpha|\mathbf{x}$, $\beta|\mathbf{y}, \mathbf{x}$ and $\phi|\mathbf{y}, \mathbf{x}$

Example 2

Example

Meteorologists are interested in predicting the December rainfall (in mm) from the November rainfall at a particular site. Use the following summary statistics to compute 95% credible intervals for the posterior distributions of the parameters $\alpha, \beta, \phi | \mathbf{y}, \mathbf{x}$:

$$n = 10, \bar{x} = 57.8, \bar{y} = 40.8, S_{xx} = 13539, S_{yy} = 1889, S_{xy} = -2178$$

Multiple regression

- We ften want to fit models of the form

  $$y_i|\alpha, \beta_1, \ldots, \beta_p, \phi \mathbf{x}_1, \ldots, \mathbf{x}_p \sim N(\alpha + \beta_1(x_{1i} - \bar{x}_1) + \cdots + \beta_p(x_{pi} - \bar{x}_p), \phi)$$

  ie with p different explanatory variables
- (Side note: an even more interesting problem here is to let $\mathbf{x}_p = \mathbf{x}^p$ to create a polynomial regression. We could then treat p as a parameter and put a prior distribution on it)
- It is posssible to fit Bayesian models here, but we will require matrix notation:

  $$\mathbf{y}|\mathbf{x}, \boldsymbol{\beta} \sim N(\mathbf{X}\boldsymbol{\beta}, \phi\mathbf{I})$$

  where $\mathbf{X}$ is a design matrix, $\boldsymbol{\beta}$ a vector of regression coefficients (ie parameters), and $\mathbf{I}$ the identity matrix

Example 3

Example

Suppose that $\mathbf{y}|\mathbf{x}, \beta \sim N(\mathbf{X}\beta, \phi\mathbf{I})$ and that ϕ is known. Using the reference prior $p(\beta) \propto 1$, find the posterior distribution of $\beta|\mathbf{x}, \mathbf{y}$