$y \to$ data

$\theta \mapsto$ parameters

Model: $f(y|\theta) \mapsto$ likelihood

Prior: $f(\theta)$

posterior: $f(\theta|y) = \dfrac{f(y|\theta)f(\theta)}{\int f(y|\theta)f(\theta)d\theta} \propto f(y|\theta)f(\theta)$

(1) $\hat{\theta}$: mean median mode

$E\theta|X$ Median$(\theta|X)$ mode$(\theta|X)$

(2) Credible Interval $(1-\alpha)$ - equal-tail CI

highest Posterior density CI

(3) $H_0: \theta \in \Theta_0$

$H_a: \theta \in \Theta_0^c$

$P(\theta \in \Theta_0 | y) \geq \frac{1}{2}$, Accept the null

$< \frac{1}{2}$ Reject the null

8.4. Mulitparameter Models

	θ	y $n=1447$
George Bush	θ_1	y_1
Michael Dukakis	θ_2	y_2
Other	θ_3	y_3

$y_1+y_2+y_3 = 1447$

$\theta_1+\theta_2+\theta_3 = 1$

$\theta_1-\theta_2$

Solut: $y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \sim \text{Multinomial}\left(n, \begin{pmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \end{pmatrix}\right)$

Solut $y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \sim$ Multinomial $\left(n, \begin{pmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \end{pmatrix} \right)$

$$f(y|\theta) = \frac{n!}{y_1! y_2! y_3!} \theta_1^{y_1} \theta_2^{y_2} \theta_3^{y_3} \propto \theta_1^{y_1} \theta_2^{y_2} \theta_3^{y_3}$$

Conjugate prior: $\frac{\Gamma(\Sigma \alpha_i)}{\Gamma(\alpha_1)\Gamma(\alpha_2)\Gamma(\alpha_3)} \theta_1^{\alpha_1 - 1} \theta_2^{\alpha_2 - 1} \theta_3^{\alpha_3 - 1}$ Dirichlet Distribution (α)

$$f(\theta) \propto \theta_1^{\alpha_1 - 1} \theta_2^{\alpha_2 - 1} \theta_3^{\alpha_3 - 1}$$

$$f(\theta|y) \propto f(y, \theta) \propto \theta_1^{y_1 + \alpha_1 - 1} \theta_2^{y_2 + \alpha_2 - 1} \theta_3^{y_3 + \alpha_3 - 1}$$

$$f(\theta|y) \sim Dir\left((y_1 + \alpha_1, y_2 + \alpha_2, y_3 + \alpha_3) \right)$$

$\alpha_1 = \alpha_2 = \alpha_3 = 1$ $\quad y = (727, 583, 137)$

$$\Rightarrow \theta | y \sim Dir\left((728, 584, 138) \right)$$

$$f(\theta_1 - \theta_2 | y)$$

Normal - Normal model

$$Y \sim N(\theta, \sigma^2)$$

$$\theta \sim N(\mu, \tau^2)$$

$$\theta | Y \sim N(MY + (1-M)\mu, M\sigma^2) \quad \text{where } M = \frac{\tau^2}{\tau^2 + \sigma^2}$$

$$\tilde{\theta} = (\theta, \sigma^2) = (\theta_1, \theta_2)$$

Inference for θ_1

θ_1

$$f(\theta|y) \propto f(y|\theta_1, \theta_2) f(\theta_1, \theta_2)$$

$$f(\theta_1|y) = \int f(\theta_1, \theta_2 | y) \, d\theta_2$$

Example $\quad Y_i \Rightarrow$ i-th observation

$$Y_i \sim N(\mu, \sigma^2)$$

$$\begin{cases} Y_i \sim N(\mu, \sigma^2) \\ \mu | \sigma^2 \sim N(\mu_0, \frac{\sigma^2}{\kappa_0}) \\ (\sigma^2)^{-1} \sim \text{Gamma}(\frac{\nu_0}{2}, \underline{\frac{\nu_0 \sigma_0^2}{2}}) \end{cases}$$

$y \sim \text{Gamma}(\alpha, \beta)$

$\frac{\beta^\alpha}{\Gamma(\alpha)} y^{\alpha-1} e^{-\beta y}$

Normal - Inv - Gamma $(\mu_0, \kappa_0, \nu_0, \sigma_0^2)$

Normal - Inv - χ^2 ( )

$y^{-1} \sim \text{INGamma}(\alpha, \beta)$

$\propto y^{-\alpha-1} e^{-\frac{\beta}{y}}$

$f(\underline{\mu} | \tilde{y})$

$$f(y_i | \mu, \sigma^2) \propto \left(\frac{1}{\sigma^2}\right)^{\frac{1}{2}} \exp\left\{-\frac{(y_i-\mu)^2}{2\sigma^2}\right\}$$

$$\mu | \sigma^2 \propto \left(\frac{1}{\sigma^2}\right)^{-\frac{1}{2}} \exp\left\{-\frac{(\mu-\mu_0)^2}{2\sigma^2/\kappa_0}\right\}$$

$$\sigma^2 \propto \left(\frac{1}{\sigma^2}\right)^{\frac{\nu_0}{2}+1} \exp\left\{-\frac{\nu_0\sigma_0^2}{2}\frac{1}{\sigma^2}\right\}$$

$$f(\underline{\mu}, \sigma^2 | y) \propto \left(\frac{1}{\sigma^2}\right)^{\frac{n}{2}+\frac{\nu_0+1}{2}+1} \exp\left\{-\frac{1}{2\sigma^2}\left(\sum \underline{(y_i-\mu)^2} + \kappa_0(\mu-\mu_0)^2 + \frac{\nu_0\sigma_0^2}{2}\right)\right\}$$

$$\mu | \sigma^2, y \propto \exp\left\{-\frac{1}{2\sigma^2}\left(\sum(\mu-y_i)^2 + \kappa_0(\mu-\mu_0)^2\right)\right\}$$

$$\begin{aligned}
\sum(\mu-y_i)^2 + \kappa_0(\mu-\mu_0)^2 &= \sum(\mu^2 - 2\mu y_i + y_i^2) + \kappa_0(\mu^2 - 2\mu\mu_0 + \mu_0^2) \\
&= (n+\kappa_0)\mu^2 - 2\mu\left(\sum y_i + \kappa_0\mu_0\right) + \text{Constant} \\
&= (n+\kappa_0)\left(\mu^2 - 2\cdot\mu\cdot\frac{\sum y_i + \kappa_0\mu_0}{n+\kappa_0}\right) + \text{Constant} \\
&= (n+\kappa_0)\left(\mu - \frac{\sum y_i + \kappa_0\mu_0}{n+\kappa_0}\right)^2 + \text{Constant}
\end{aligned}$$

$$\mu | \sigma^2, y \sim N(\underline{M\bar{y} + (1-M)\mu_0}, M\sigma^2)$$

where $M = \frac{n}{n+\kappa_0}$

$$\sigma^2 | y \quad \cancel{\propto} \ \text{Inv-Gamma}\left(\frac{\nu_n}{2}, \frac{\nu_n\sigma_n^2}{2}\right)$$

where $\nu_n = \nu_0 + n$

$$\nu_n\sigma_n^2 = \underline{\nu_0\sigma_0^2} + (n-1)s^2 + \frac{\kappa_0 n}{\kappa_0+n}(\bar{y}-\mu_0)^2$$

$\underline{\underline{\mu \mid y}}$?

① generate $\underline{\underline{\sigma^2 \mid y}}$ according to Inv-Gamma($\frac{\cdot}{2}$, $\frac{\cdot}{2}$)

② generate μ according to $\underline{\underline{\mu \mid \sigma^2, y}}$.

Model Checking   posterior predictive checking.

Let y be the observed data θ be the parameters

$$f(\underline{\underline{y}}^{rep} \mid y) = \int f(y^{rep} \mid \theta)\, f(\theta \mid y)\, d\theta.$$

Test quantities $T(y, \theta)$ $T(y^{rep}, \theta)$

posterior predictive p-values

$$P\left(T(y^{rep}, \theta) \geq T(y; \theta)\right) \approx \frac{1}{G} \sum_g 1\left(T(y_g^{rep}, \theta) \geq T(y, \theta)\right)$$

Example 8.4.3

(a) Use the histogram

(b) Use the min(y)

(c) $T(y, \theta) = |y_{(61)} - \theta| - |y_{(6)} - \theta|$

8.5. Hierarchical Model

Cardiac Treatment

$y_j \mid \theta_j \sim f(y_j \mid \theta_j)$

$\theta_j \mapsto j$-th hospital. Survival Probability

$\theta_j \sim f(\theta \mid \alpha) \rightarrow$ distribution of the survival probability of all the hospital in a region

$\alpha \sim f(\alpha \mid \beta)$

hyperparameters. $\theta_j \mid Y$. hyperparameters.

Example. Baseball example (Efron and Morris).

Let Y_i be the batting average on the first 45 at bats.

θ_i be the batting average

$Y_i \mid \theta_i, \sigma^2 \sim N(\theta_i, \sigma^2)$

$$\begin{cases} \theta_i \mid \mu, \tau^2 \sim N(\mu, \tau^2) \\ \mu \sim N(0, 1000) \\ (\sigma^2)^{-1}, (\tau^2)^{-1} \sim \text{Gamma}(0.001, 0.001) \end{cases}$$

$\theta_i \mid y$?

Markov Chain Monte Carlo (MCMC) Gibbs Sampler.

Example. 8.5.2

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \sim \text{Bivariate Normal}\left(\begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix}, \begin{pmatrix} 1 & 0.5 \\ 0.5 & 1 \end{pmatrix} \right)$$

$$\begin{cases} \theta_1 \sim N(0, 1000) \\ \theta_2 \sim N(0, 1000) \end{cases} \qquad \begin{matrix} \pi(\theta_1) \propto 1 \\ \pi(\theta_2) \propto 1 \end{matrix}$$

$$\begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} \Bigg| \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

$$\begin{cases} \theta_1 \mid \theta_2, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \sim N(y_1 + 0.5(\theta_2 - y_2), 1 - 0.5^2) \\ \theta_2 \mid \theta_1, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \sim N(y_2 + 0.5(\theta_1 - y_1), 1 - 0.5^2) \end{cases}$$

Initial Value

Dependence

$$\begin{pmatrix} \theta_1^{k} \\ \theta_2^{k} \end{pmatrix} \qquad \begin{pmatrix} \theta_1^{k+1} \\ \theta_2^{k+1} \end{pmatrix}$$

burn-in : discard the first k - simulation

thin. Keeping every l-th simulation.

$$\theta^{10} \quad \theta^{11} \quad \theta^{12} \quad \theta^{13} \quad \ldots \quad \theta^{20}$$

11,000 throw $k = 1{,}000$

$l = 10$.

Autocorrelation
$$\rho_k = \frac{Cov(\theta_t, \theta_{t-k})}{\sqrt{V(\theta_t) \cdot V(\theta_{t-k})}}$$