

1. From page 26 in theory bayes15.pdf:

$$y_{ij} | \alpha, \beta_i, \sigma^2 \sim N(x_{ij}' \alpha + z_{ij}' \beta_i, \sigma^2)$$

$$\beta_i | D \sim N(0, D)$$

$$\alpha \sim N(\alpha_0, V_0)$$

$$\sigma^2 \sim \text{InverseGamma}(a/2, b/2)$$

$$D \sim \text{InverseWishart}_q(\nu, D_0)$$

assume:

$$y_j | \alpha, \beta, \sigma^2 \sim N(x' \alpha + \beta, \sigma^2) \quad \begin{matrix} q=1 \\ z_{ij}=1 \end{matrix}$$

$(D_0, \nu)$  possible mistake

$\beta$  a scalar  
 $\alpha$  a vector

$$f(\beta | D) = \frac{1}{(2\pi D)^{n/2}} e^{-\frac{1}{2} \left( \frac{\beta^2}{D} \right)}$$

$$= \frac{1}{(2\pi D)^{n/2}} e^{-\frac{\beta^2}{2D}}$$

$$f(D | \beta) \propto f(\beta | D) f(D)$$

$$\propto \left[ \frac{1}{(2\pi D)^{n/2}} e^{-\frac{\beta^2}{2D}} \right] \left[ \frac{|D_0|^{\nu/2}}{2^{\nu/2} \Gamma(\frac{\nu}{2})} |D|^{-(\nu+1)/2} e^{-\frac{1}{2} \text{tr}(D_0 D^{-1})} \right]$$

$$\propto \frac{|D_0|^{\nu/2}}{D^n 2^{\nu/2} \Gamma(\frac{\nu}{2})} e^{-\frac{\beta^2}{2D} - \frac{|D_0|}{2D}} D^{(\nu+2)/2}$$

$$\propto D^{-\frac{1}{2}(1+\nu+2)} e^{-\frac{1}{2}(\beta^2 + |D_0|)/D}$$

$$D | \beta \sim W^{-1}(\beta^2 + D_0, \nu + 1)$$

$$f(\sigma^2 | \alpha, \beta, y) \propto L(y | \alpha, \beta, \sigma^2) f(\alpha, \beta, \sigma^2)$$

$$\propto f(y | \alpha, \beta, \sigma^2) f(\sigma^2)$$

$$f(y | \alpha, \beta, \sigma^2) = \prod_{j=1}^n f(y_j | \alpha, \beta, \sigma^2)$$

$$= \frac{1}{(2\pi \sigma^2)^{n/2}} e^{-\frac{1}{2\sigma^2} \sum (y_j - x_j' \alpha - \beta)^2}$$

$$f(\sigma^2 | \alpha, \beta, y) \propto \frac{1}{(2\pi \sigma^2)^{n/2}} e^{-\frac{1}{2\sigma^2} \sum (y_j - x_j' \alpha - \beta)^2} \frac{(\frac{b}{2})^{\alpha/2}}{\Gamma(\frac{a}{2})} (\sigma^2)^{(\frac{a}{2}-1)} e^{-\frac{b/2}{\sigma^2}}$$

$$\propto \frac{1}{(2\pi)^{n/2}} \frac{(\frac{b}{2})^{\alpha/2}}{\Gamma(\frac{a}{2})} (\sigma^2)^{(-\frac{a}{2} - \frac{n}{2} - 1)} e^{-\frac{1}{2\sigma^2} (\sum (y_j - x_j' \alpha - \beta) + b)}$$

$$\sigma^2 | \alpha, \beta, y \sim IG\left(\frac{a+n}{2}, \frac{\sum (y_j - x_j' \alpha - \beta) + b}{2}\right)$$

observations  $y_{ij}$  on subject  $i$

$j = 1, \dots, n_i$

known covariate vectors  $\vec{x}_i$  and  $\vec{z}_i$

subject specific regression coeffs  $\beta_i$

population regression coeffs  $\alpha$

conditionally conjugate priors

InverseWishart  $\rightarrow$  distr on positive definite matrices

convenient prior for covariance matrices

positive definite: symmetric w/ all positive eigenvalues  
nonsingular

$$\text{Unknown } \theta = (\alpha, \{\beta_i\}_{i=1}^n, \sigma^2, D)$$

$W^{-1}(\bar{\alpha}, \nu)$   
pdf of InverseWishart:

$$\frac{|\bar{\alpha}|^{\nu/2}}{2^{\nu p/2} \Gamma_p(\frac{\nu}{2})} \frac{|X|}{\det(X)} e^{-\frac{1}{2} \text{tr}(\bar{\alpha} X^{-1})}$$

$\det(\bar{\alpha})$  multivariate gamma function

$$f(\beta_i | \alpha, \sigma^2, D, Y_i) \propto L(Y_i | \alpha, \beta, \sigma^2) f(\beta_i | D) \quad i=1$$

$$\propto \frac{1}{(2\pi\sigma^2)^{n/2}} e^{-\frac{1}{2\sigma^2} \sum (y_j - x_j' \alpha - \beta)^2} \frac{1}{(2\pi D)^{1/2}} e^{-\frac{\beta^2}{2D}}$$

$$\propto \exp\left(-\frac{1}{2} \left( \frac{1}{\sigma^2} \sum (y_j - x_j' \alpha - \beta)^2 + \frac{\beta^2}{D} \right)\right)$$

$$\propto \exp\left(-\frac{1}{2} \left( \frac{1}{\sigma^2} \sum (y_j^2 - y_j x_j' \alpha - y_j \beta - x_j' \alpha y_j + (x_j' \alpha)^2 + x_j' \alpha \beta - \beta y_j + \beta x_j' \alpha + \beta^2) + \frac{\beta^2}{D} \right)\right)$$

$$\propto \exp\left(-\frac{1}{2} \left( \frac{1}{\sigma^2} \sum (-y_j \beta + x_j' \alpha \beta - \beta y_j + \beta x_j' \alpha + \beta^2) + \frac{\beta^2}{D} \right)\right)$$

$$\propto \exp\left(-\frac{1}{2} \left( \frac{1}{\sigma^2} \sum (\beta^2 + 2\beta x_j' \alpha - 2\beta y_j) + \frac{\beta^2}{D} \right)\right)$$

$$\propto \exp\left(-\frac{1}{2} \left( \frac{1}{\sigma^2} \left( n\beta^2 + \sum (2\beta x_j' \alpha - 2\beta y_j) \right) + \frac{\beta^2}{D} \right)\right)$$

$$\propto \exp\left(-\frac{1}{2} \left( \frac{n\beta^2}{\sigma^2} + \frac{1}{\sigma^2} \sum (2\beta x_j' \alpha - 2\beta y_j) + \frac{\beta^2}{D} \right)\right)$$

$$\propto \exp\left(-\frac{1}{2} \left( \beta^2 \left( \frac{n}{\sigma^2} + \frac{1}{D} \right) - \frac{2\beta}{\sigma^2} \sum (y_j - x_j' \alpha) \right)\right)$$

$$\boxed{\beta | \alpha, \sigma^2, D, Y_i \sim N\left(\frac{\sum y_j - x_j' \alpha}{\sigma^2 \left( \frac{n}{\sigma^2} + \frac{1}{D} \right)}, \frac{1}{\left( \frac{n}{\sigma^2} + \frac{1}{D} \right)}\right)} \quad i=1$$

$$f(\alpha | \beta, y, \sigma^2) \propto L(y | \alpha, \beta, \sigma^2) f(\alpha)$$

$$\propto \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left(-\frac{1}{2\sigma^2} \sum (y_j - x_j' \alpha - \beta)^2\right) \frac{1}{(2\pi V_0)^{1/2}} \exp\left(-\frac{1}{2V_0} (\alpha - \alpha_0)^2\right)$$

$$\propto \exp\left(-\frac{1}{2} \left[ \frac{1}{\sigma^2} \sum (y_j - x_j' \alpha - \beta)^2 + \frac{1}{V_0} (\alpha - \alpha_0)^2 \right]\right)$$

$$\propto \exp\left(-\frac{1}{2} \left[ \frac{1}{\sigma^2} \sum (y_j - y_j x_j' \alpha - y_j \beta - x_j' \alpha y_j + (x_j' \alpha)^2 + x_j' \alpha \beta - \beta y_j + \beta x_j' \alpha + \beta^2) + \frac{1}{V_0} (\alpha' \alpha - 2\alpha' \alpha_0 + \alpha_0^2) \right]\right)$$

$$\propto \exp\left(-\frac{1}{2} \left[ \frac{1}{\sigma^2} \sum ((x_j' \alpha)^2 - 2x_j' \alpha y_j + 2x_j' \alpha \beta) + \frac{1}{V_0} (\alpha' \alpha - 2\alpha' \alpha_0) \right]\right)$$

$$\propto \exp\left(-\frac{1}{2} \left[ \frac{1}{\sigma^2} \sum (\alpha' x_j x_j' \alpha - 2x_j' \alpha y_j + 2x_j' \alpha \beta) + \frac{1}{V_0} (\alpha' \alpha - 2\alpha' \alpha_0) \right]\right)$$

$$\propto \exp\left(-\frac{1}{2} \left[ \frac{\sum \alpha' x_j x_j' \alpha}{\sigma^2} + \frac{\alpha' \alpha}{V_0} - \frac{2}{\sigma^2} \sum \alpha' x_j (y_j - \beta) - \frac{1}{V_0} (2\alpha_0' \alpha) \right]\right)$$

$$\alpha \exp\left(-\frac{1}{2}\left[\alpha'\left[\frac{\sum x_j x_j'}{\sigma^2} + \frac{1}{V_0}\right]\alpha - 2\left[\frac{\sum y_j x_j' - \beta x_j'}{\sigma^2} - \frac{\alpha'_0}{V_0}\right]\alpha\right]\right)$$

$$\alpha \exp\left(-\frac{1}{2}\left[\frac{\sum x_j x_j'}{\sigma^2} + \frac{1}{V_0}\right]\left[\alpha - \frac{\frac{\sum y_j x_j' - \beta x_j'}{\sigma^2} - \frac{\alpha'_0}{V_0}}{\frac{\sum x_j x_j'}{\sigma^2} + \frac{1}{V_0}}\right]^2\right)$$

$$\alpha | \beta, y_i, \sigma^2 \sim N\left(\frac{\sum y_j x_j' - \beta x_j'}{\sigma^2} - \frac{\alpha'_0}{V_0}, \left[\frac{\sum x_j x_j'}{\sigma^2} + \frac{1}{V_0}\right]^{-1}\right)$$

$$2. a) f(y_i | \alpha, \beta, \sigma^2, D) = L(y_i | \alpha, \beta, \sigma^2) f(\beta | D) f(D)$$

$$= (2\pi\sigma^2)^{-n/2} \exp\left(-\frac{1}{2\sigma^2} \sum_{j=1}^n (y_j - x_j' \alpha - \beta)^2\right) (2\pi D)^{-1/2} \exp\left(-\frac{\beta^2}{2D}\right)$$

$$= (2\pi)^{-\frac{n+1}{2}} (\sigma^2)^{-n/2} (D)^{-1/2} \exp\left(-\frac{1}{2} \left[ \frac{1}{\sigma^2} \sum_{j=1}^n (y_j - x_j' \alpha - \beta)^2 + \frac{\beta^2}{D} \right]\right)$$

$$= (2\pi)^{-\frac{n+1}{2}} (\sigma^2)^{-n/2} (D)^{-1/2} \exp\left(-\frac{1}{2} \left[ \frac{1}{\sigma^2} \sum_{j=1}^n (y_j^2 - y_j x_j' \alpha - y_j \beta - x_j' \alpha y_j + (x_j' \alpha)^2 + x_j' \alpha \beta - \beta y_j + \beta x_j' \alpha + \beta^2) + \frac{\beta^2}{D} \right]\right)$$

$$= (2\pi)^{-\frac{n+1}{2}} (\sigma^2)^{-n/2} (D)^{-1/2} \exp\left(-\frac{1}{2} \left[ \frac{1}{\sigma^2} \sum_{j=1}^n (y_j^2 + (x_j' \alpha)^2 - 2y_j x_j' \alpha - 2\beta y_j + 2\beta x_j' \alpha + \beta^2) + \frac{\beta^2}{D} \right]\right)$$

b) Integrate out  $\beta$  to get  $y_i | \alpha, \sigma^2, D$

$$f(y_i | \alpha, \sigma^2, D) = \int_{-\infty}^{\infty} f(y_i | \alpha, \beta, \sigma^2, D) d\beta$$

$$= (2\pi)^{-\frac{n+1}{2}} (\sigma^2)^{-n/2} (D)^{-1/2} \exp\left(-\frac{1}{2} \left( \frac{1}{\sigma^2} \sum_{j=1}^n (y_j^2 + (x_j' \alpha)^2 - 2y_j x_j' \alpha) \right)\right) \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2} \left[ \sum_{j=1}^n \frac{\beta^2 - 2\beta y_j + 2\beta x_j' \alpha}{\sigma^2} + \frac{\beta^2}{D} \right]\right) d\beta$$

integrand is proportional to

$$\beta | \alpha, \sigma^2, D, y_i \sim N\left(\frac{\sum y_j - x_j' \alpha}{\sigma^2 \left(\frac{1}{\sigma^2} + \frac{1}{D}\right)}, \frac{1}{\left(\frac{1}{\sigma^2} + \frac{1}{D}\right)}\right) \quad i=1$$

$$= (2\pi)^{\frac{1}{2} \left(\frac{n+1}{\sigma^2} + \frac{1}{D}\right)} (2\pi)^{-\frac{n+1}{2}} (\sigma^2)^{-\frac{n}{2}} (D)^{-\frac{1}{2}} \exp\left(-\frac{1}{2\sigma^2} \sum_{j=1}^n (y_j^2 + (x_j' \alpha)^2 - 2y_j x_j' \alpha)\right) \int_{-\infty}^{\infty} (2\pi)^{-1/2} \left(\frac{n}{\sigma^2} + \frac{1}{D}\right)^{1/2} \exp\left(-\frac{1}{2} \left(\frac{n}{\sigma^2} + \frac{1}{D}\right) \left(\beta - \frac{\sum y_j - x_j' \alpha}{\sigma^2 \left(\frac{1}{\sigma^2} + \frac{1}{D}\right)}\right)^2 - \frac{1}{2} \left(\frac{n}{\sigma^2} + \frac{1}{D}\right) \frac{(\sum y_j - x_j' \alpha)^2}{\sigma^2 \left(\frac{1}{\sigma^2} + \frac{1}{D}\right)}\right) d\beta$$

$$= (2\pi)^{\frac{1}{2} \left(\frac{n+1}{\sigma^2} + \frac{1}{D}\right)} (2\pi)^{-\frac{n+1}{2}} (\sigma^2)^{-\frac{n}{2}} D^{-\frac{1}{2}} \exp\left(-\frac{1}{2\sigma^2} \sum_{j=1}^n (y_j^2 + (x_j' \alpha)^2 - 2y_j x_j' \alpha)\right) \exp\left(-\frac{1}{2} \left(\frac{n}{\sigma^2} + \frac{1}{D}\right) \frac{(\sum y_j - x_j' \alpha)^2}{\sigma^2 \left(\frac{1}{\sigma^2} + \frac{1}{D}\right)}\right) \int_{-\infty}^{\infty} (2\pi)^{-1/2} \left(\frac{n}{\sigma^2} + \frac{1}{D}\right)^{1/2} \exp\left(-\frac{1}{2} \left(\frac{n}{\sigma^2} + \frac{1}{D}\right) \left(\beta - \frac{\sum y_j - x_j' \alpha}{\sigma^2 \left(\frac{1}{\sigma^2} + \frac{1}{D}\right)}\right)^2\right) d\beta$$

$$\alpha \exp\left(-\frac{1}{2\sigma^2} \left(1 + \frac{1}{\sigma^2 \left(\frac{1}{\sigma^2} + \frac{1}{D}\right)}\right)\right) \sum_{j=1}^n (y_j - x_j' \alpha)^2$$

Integrate to 1

$$y_i | \alpha, \sigma^2, D \sim N\left(x_j' \alpha, \sigma^2 \left(1 + \frac{1}{\sigma^2 \left(\frac{1}{\sigma^2} + \frac{1}{D}\right)}\right)^{-1}\right)$$

3. a) Design 1:

$\frac{n}{2}$  clinics intervention

$$x_{ij} = [0 \ 1]$$

$$y_{ij} = x'_{ij}\alpha + \beta_i + \epsilon_{ij}$$

$$\beta_i | D \sim N(0, D)$$

$$\epsilon_{ij} | \sigma^2 \sim N(0, \sigma^2)$$

$\frac{n}{2}$  clinics control

$$(\bar{Y}_1 - \bar{Y}_0) = \frac{\left( \sum_{i=1}^{n/2} \sum_{j=1}^J y_{ij} - \sum_{i=\frac{n}{2}+1}^n \sum_{j=1}^J y_{ij} \right)}{\frac{nJ}{2}} = \frac{2}{nJ} \left[ \sum_{i=1}^{n/2} \sum_{j=1}^J y_{ij} - \sum_{i=\frac{n}{2}+1}^n \sum_{j=1}^J y_{ij} \right]$$

$$= \frac{2}{nJ} \left[ \sum_{i=1}^{n/2} \sum_{j=1}^J (x'_{ij}\alpha + \beta_i + \epsilon_{ij}) - \sum_{i=\frac{n}{2}+1}^n \sum_{j=1}^J (x'_{ij}\alpha + \beta_i + \epsilon_{ij}) \right]$$

$$\begin{aligned} \text{Var}(\bar{Y}_1 - \bar{Y}_0) &= \left(\frac{2}{nJ}\right)^2 \text{Var} \left[ \sum_{i=1}^{n/2} \sum_{j=1}^J (x'_{ij}\alpha + \beta_i + \epsilon_{ij}) - \sum_{i=\frac{n}{2}+1}^n \sum_{j=1}^J (x'_{ij}\alpha + \beta_i + \epsilon_{ij}) \right] \\ &= \left(\frac{2}{nJ}\right)^2 \text{Var} \left[ \cancel{\sum_{i=1}^{n/2} \sum_{j=1}^J x'_{ij}\alpha} + \sum_{i=1}^{n/2} \sum_{j=1}^J \beta_i + \sum_{i=1}^{n/2} \sum_{j=1}^J \epsilon_{ij} - \cancel{\sum_{i=\frac{n}{2}+1}^n \sum_{j=1}^J x'_{ij}\alpha} - \sum_{i=\frac{n}{2}+1}^n \sum_{j=1}^J \beta_i - \sum_{i=\frac{n}{2}+1}^n \sum_{j=1}^J \epsilon_{ij} \right] \\ &= \left(\frac{2}{nJ}\right)^2 \text{Var} \left[ \sum_{i=1}^{n/2} J\beta_i + \sum_{i=1}^{n/2} \sum_{j=1}^J \epsilon_{ij} - \sum_{i=\frac{n}{2}+1}^n J\beta_i - \sum_{i=\frac{n}{2}+1}^n \sum_{j=1}^J \epsilon_{ij} \right] \end{aligned}$$

$$\beta_i | D \sim N(0, D) \rightarrow \text{Var}(\beta_i | D) = D$$

$$\text{Var}(X - Y) = \text{Var}(X) + \text{Var}(Y)$$

$$\epsilon_{ij} | \sigma^2 \sim N(0, \sigma^2) \rightarrow \text{Var}(\epsilon_{ij} | \sigma^2) = \sigma^2$$

$$= \left(\frac{2}{nJ}\right)^2 \left( \frac{n}{2} J^2 D + \frac{n}{2} J \sigma^2 + \left(\frac{n}{2}\right) J^2 D + \left(\frac{n}{2}\right) J \sigma^2 \right)$$

$$= \left(\frac{2}{nJ}\right) \left( n J^2 D + n J \sigma^2 \right)$$

$$= \left(\frac{4}{n^2 J^2}\right) \left( n J^2 D + n J \sigma^2 \right)$$

$$= \frac{4n J^2 D}{n^2 J^2} + \frac{4n J \sigma^2}{n^2 J^2}$$

$$= \frac{4}{n} D + \frac{4}{nJ} \sigma^2$$

$$= \frac{4}{n} \left( D + \frac{\sigma^2}{J} \right)$$

Design 2:

for each clinic,  $\frac{J}{2}$  patients intervention  
 $\frac{J}{2}$  patients control

$$\begin{aligned}
 (\bar{Y}_1 - \bar{Y}_0) &= \frac{\left( \sum_{i=1}^n \sum_{j=1}^{J/2} y_{ij} - \sum_{i=1}^n \sum_{j=\frac{J}{2}+1}^J y_{ij} \right)}{\frac{Jn}{2}} = \frac{2}{nJ} \left[ \sum_{i=1}^n \sum_{j=1}^{J/2} y_{ij} - \sum_{i=1}^n \sum_{j=\frac{J}{2}+1}^J y_{ij} \right] \\
 &= \frac{2}{nJ} \left[ \sum_{i=1}^n \sum_{j=1}^{J/2} (x'_{ij}\alpha + \beta_i + \epsilon_{ij}) - \sum_{i=1}^n \sum_{j=\frac{J}{2}+1}^J (x'_{ij}\alpha + \beta_i + \epsilon_{ij}) \right] \\
 \text{Var}(\bar{Y}_1 - \bar{Y}_0) &= \left( \frac{2}{nJ} \right)^2 \text{Var} \left[ \sum_{i=1}^n \sum_{j=1}^{J/2} (x'_{ij}\alpha + \beta_i + \epsilon_{ij}) - \sum_{i=1}^n \sum_{j=\frac{J}{2}+1}^J (x'_{ij}\alpha + \beta_i + \epsilon_{ij}) \right] \\
 &= \left( \frac{2}{nJ} \right)^2 \text{Var} \left[ \sum_{i=1}^n \sum_{j=1}^{J/2} \cancel{x'_{ij}\alpha} + \sum_{i=1}^n \sum_{j=1}^{J/2} \beta_i + \sum_{i=1}^n \sum_{j=1}^{J/2} \epsilon_{ij} - \sum_{i=1}^n \sum_{j=\frac{J}{2}+1}^J \cancel{x'_{ij}\alpha} - \sum_{i=1}^n \sum_{j=\frac{J}{2}+1}^J \beta_i - \sum_{i=1}^n \sum_{j=\frac{J}{2}+1}^J \epsilon_{ij} \right] \\
 &= \left( \frac{2}{nJ} \right)^2 \text{Var} \left[ \sum_{i=1}^n \sum_{j=1}^{J/2} \beta_i + \sum_{i=1}^n \sum_{j=1}^{J/2} \epsilon_{ij} - \sum_{i=1}^n \sum_{j=\frac{J}{2}+1}^J \beta_i - \sum_{i=1}^n \sum_{j=\frac{J}{2}+1}^J \epsilon_{ij} \right] \\
 &= \left( \frac{2}{nJ} \right)^2 \text{Var} \left[ \sum_{i=1}^n \cancel{\beta_i} + \sum_{i=1}^n \sum_{j=1}^{J/2} \epsilon_{ij} - \sum_{i=1}^n \cancel{\beta_i} - \sum_{i=1}^n \sum_{j=\frac{J}{2}+1}^J \epsilon_{ij} \right] \\
 &= \left( \frac{2}{nJ} \right)^2 \left( n \left( \frac{J}{2} \right) \sigma^2 + n \left( \frac{J}{2} \right) \sigma^2 \right) \\
 &= \left( \frac{4}{n^2 J^2} \right) \left( \frac{2nJ}{2} \sigma^2 \right) \\
 &= \frac{4nJ\sigma^2}{n^2 J^2} \\
 &= \frac{4}{nJ} \sigma^2
 \end{aligned}$$

$$\begin{aligned}
 b) \text{Var}_1(\bar{Y}_1 - \bar{Y}_0) - \text{Var}_2(\bar{Y}_1 - \bar{Y}_0) &= \frac{4}{n} \left( 0 + \frac{\sigma^2}{J} \right) - \frac{4}{nJ} \sigma^2 \\
 &= \frac{4}{n} D + \frac{4\sigma^2}{nJ} - \frac{4\sigma^2}{nJ} \\
 &= \frac{4}{n} D \geq 0 \quad \text{with } J, n > 0
 \end{aligned}$$

$\text{Var}(\bar{Y}_1 - \bar{Y}_0)$  for Design 1 is larger

$\text{Var}(\bar{Y}_1 - \bar{Y}_0)$  for Design 2 is smaller since we know  $D \geq 0$

- c) Yes, sometimes randomized block design will be more effective than randomized patient design if there is a possibility of cross contamination on patients affecting trial results — randomized block design is effective for grouping according to known/suspected variation isolated by block design