STT 465
Bayesian Multiple Linear Regression:

- Summary of model specification
- Conditional distribution of effects
- Model with unknown variances
- Gibbs sampler

Bayesian Multiple Linear Regression

- **Gaussian Linear Regression Model**

$$y_i = \sum_{j=1}^{p} x_{ij}\beta_j + \varepsilon_i$$

- **Matrix representation**

Stack equations 1-n to get $$y = X\beta + \varepsilon$$

- **Likelihood (assuming iid normal errors)** $$\varepsilon \sim MVN\left(0, \mathrm{I}\sigma_\varepsilon^2\right)$$

$$\left[y \mid \beta\right] \sim MVN\left(X\beta, R\right)$$

$$p\left(y \mid X, \beta, \sigma_\varepsilon^2\right) = \left(2\pi\right)^{-n/2} \left\| I\sigma_\varepsilon^2 \right\|^{-1/2} Exp\left\{-\frac{1}{2\sigma_\varepsilon^2}\left(y - X\beta\right)'\left(y - X\beta\right)\right\}$$

Prior with known variance components

Normal prior for reg. coefficients

$$\beta \sim N\left(\beta_0, \Sigma_0\right) = \left(2\pi\right)^{-n/2} \left\|\Sigma_0\right\|^{-1/2} Exp\left\{-\frac{1}{2}\left(\beta - \beta_0\right)' \Sigma_0^{-1}\left(\beta - \beta_0\right)\right\}$$

If we assume IID, zero-mean prior we have

$$\beta \sim N\left(0, I\sigma_\beta^2\right) = \left(2\pi\right)^{-n/2} \left\|I\sigma_\beta^2\right\|^{-1/2} Exp\left\{-\frac{\beta'\beta}{2\sigma_\beta^2}\right\}$$

Posterior Density

Likelihood

$$p(y \mid X,\beta)=(2\pi)^{-n/2}\left\|I\sigma_\varepsilon^2\right\|^{-1/2} Exp\left\{-\frac{1}{2}(y-X\beta)'\left[I\sigma_\varepsilon^2\right]^{-1}(y-X\beta)\right\}$$

Prior

$$\beta \sim N\left(0, I\sigma_\beta^2\right)=(2\pi)^{-n/2}\left\|I\sigma_\beta^2\right\|^{-1/2} Exp\left\{-\frac{1}{2}\beta'\left[I\sigma_\beta^2\right]^{-1}\beta\right\}$$

General for of the posterior density (derivation presented in class)

$$p(\beta \mid y) \sim N\left(C^{-1}rhs, C^{-1}\right)$$

$$C=\left[X'R^{-1}X+\Sigma_0^{-1}\right] \qquad \left[X'R^{-1}X+\Sigma_0^{-1}\right]\hat{\beta}=X'R^{-1}y+\Sigma_0^{-1}\beta_0$$

$$rhs=X'R^{-1}y+\Sigma_0^{-1}\beta_0$$

Special (most commonly used) case

Likelihood [iid residuals] $R = I\sigma_\varepsilon^2$

Prior [iid mean-zero effects] $R = I\sigma_\varepsilon^2$

$$\beta \sim N\left(0, I\sigma_0^2\right)$$

Posterior

$$p\left(\beta \mid y\right) \sim N\left(C^{-1}rhs, C^{-1}\right)$$

$$C = \left[X'X\sigma_\varepsilon^{-2} + I\sigma_0^2\right]$$

$$rhs = X'y\sigma_\varepsilon^{-2}$$

Discuss shrinkage and connection to Ridge Regression.

Model with unknown variances

Joint prior

$$p\left(\beta,\sigma_\varepsilon^2,\sigma_\beta^2\right)=p\left(\beta\mid\sigma_\beta^2\right)p\left(\sigma_\beta^2\right)p\left(\sigma_\varepsilon^2\right)$$

$$=N\left(\beta\middle|0,I\sigma_\beta^2\right)\chi^{-2}\left(\sigma_\beta^2\middle|df_\beta,S_\beta\right)\chi^{-2}\left(\sigma_\varepsilon^2\middle|df_\varepsilon,S_\varepsilon\right)$$

$$\propto\left[\left\|I\sigma_\beta^2\right\|^{-1/2}e^{-\frac{\beta'\beta}{2\sigma_\beta^2}}\right]\times\left[\left(\sigma_\beta^2\right)^{-(1+df_\beta/2)}e^{-\frac{S_\beta}{2\sigma_\beta^2}}\right]\times\left[\left(\sigma_\varepsilon^2\right)^{-(1+df_\varepsilon/2)}e^{-\frac{S_\varepsilon}{2\sigma_\varepsilon^2}}\right]$$

Model with unknown variances

Joint prior

$$p\left(\beta,\sigma_\varepsilon^2,\sigma_\beta^2\right)=p\left(\beta\mid\sigma_\beta^2\right)p\left(\sigma_\beta^2\right)p\left(\sigma_\varepsilon^2\right)$$

$$=N\left(\beta\middle|0,I\sigma_\beta^2\right)\chi^{-2}\left(\sigma_\beta^2\middle|df_\beta,S_\beta\right)\chi^{-2}\left(\sigma_\varepsilon^2\middle|df_\varepsilon,S_\varepsilon\right)$$

$$\propto\left[\left(\sigma_\beta^2\right)^{-p/2}e^{-\frac{\beta'\beta}{2\sigma_\beta^2}}\right]\times\left[\left(\sigma_\beta^2\right)^{-(1+df_\beta/2)}e^{-\frac{S_\beta}{2\sigma_\beta^2}}\right]\times\left[\left(\sigma_\varepsilon^2\right)^{-(1+df_\varepsilon/2)}e^{-\frac{S_\varepsilon}{2\sigma_\varepsilon^2}}\right]$$

Model with unknown variances

Joint posterior

$$p\left(\beta,\sigma_\varepsilon^2,\sigma_\beta^2 \mid y\right) = p\left(y \mid \beta,\sigma_\varepsilon^2\right) p\left(\beta \mid \sigma_\beta^2\right) p\left(\sigma_\beta^2\right) p\left(\sigma_\varepsilon^2\right)$$

$$p\left(\beta,\sigma_\varepsilon^2,\sigma_\beta^2 \mid y\right) \propto \left(\sigma_\varepsilon^2\right)^{-n/2} Exp\left\{-\frac{1}{2\sigma_\varepsilon^2}\left(y-X\beta\right)'\left(y-X\beta\right)\right\}$$

$$\times\left[\left(\sigma_\beta^2\right)^{-p/2} e^{-\frac{\beta'\beta}{2\sigma_\beta^2}}\right]\times\left[\left(\sigma_\beta^2\right)^{-(1+df_\beta/2)} e^{-\frac{S_\beta}{2\sigma_\beta^2}}\right]\times\left[\left(\sigma_\varepsilon^2\right)^{-(1+df_\varepsilon/2)} e^{-\frac{S_\varepsilon}{2\sigma_\varepsilon^2}}\right]$$

Gibbs Sampler

1. **Initialize parameters**

   - Any initial values with non-null prior prob. are valid.

2. Sample effects given variances

$$p(\beta \mid ELSE) \propto Exp\left\{-\frac{1}{2\sigma_\varepsilon^2}(y - X\beta)'(y - X\beta)\right\}\left[e^{-\frac{\beta'\beta}{2\sigma_\beta^2}}\right]$$

$$\text{[from previous results]} = N\left(C^{-1}rhs, C^{-1}\right) \qquad [1]$$

$$C = \left[X'X\sigma_\varepsilon^{-2} + I\sigma_0^2\right]$$

$$rhs = X'y\sigma_\varepsilon^{-2}$$

Gibbs Sampler

3. Sample error variance given effects

$$p\left(\sigma_\varepsilon^2 \mid ELSE\right) = \left(\sigma_\varepsilon^2\right)^{-n/2} Exp\left\{-\frac{1}{2\sigma_\varepsilon^2}\left(y - X\beta\right)'\left(y - X\beta\right)\right\}\left[\left(\sigma_\varepsilon^2\right)^{-(1+df_\varepsilon/2)} e^{-\frac{S_\varepsilon}{2\sigma_\varepsilon^2}}\right]$$

- Given effects the errors are known $\varepsilon = y - X\beta$
  - Therefore

$$p\left(\sigma_\varepsilon^2 \mid ELSE\right) \propto \left(\sigma_\varepsilon^2\right)^{-n/2} Exp\left\{-\frac{\varepsilon'\varepsilon}{2\sigma_\varepsilon^2}\right\}\left[\left(\sigma_\varepsilon^2\right)^{-(1+df_\varepsilon/2)} e^{-\frac{S_\varepsilon}{2\sigma_\varepsilon^2}}\right]$$

$$\propto \left(\sigma_\varepsilon^2\right)^{-[1+(n+df_\varepsilon)/2]} Exp\left\{-\frac{\varepsilon'\varepsilon + S_\varepsilon}{2\sigma_\varepsilon^2}\right\}$$

$$= \chi^{-2}\left(\sigma_\varepsilon^2 \middle| S = \varepsilon'\varepsilon + S_\varepsilon, df = n + df_\varepsilon\right) \qquad [2]$$

Gibbs Sampler

4. Sample the variance of effects

$$p\left(\sigma_\beta^2 \mid ELSE\right) \propto \left[\left(\sigma_\beta^2\right)^{-p/2} e^{-\frac{\beta'\beta}{2\sigma_\beta^2}}\right] \times \left[\left(\sigma_\beta^2\right)^{-(1+df_\beta/2)} e^{-\frac{S_\beta}{2\sigma_\beta^2}}\right]$$

$$\propto \left[\left(\sigma_\beta^2\right)^{-[1+(p+df_\beta)/2]} e^{-\frac{\beta'\beta+S_\beta}{2\sigma_\beta^2}}\right] \qquad \varepsilon = y - X\beta$$

$$==\chi^{-2}\left(\sigma_\beta^2 \middle| S = \beta'\beta + S_\beta, df = p + df_\beta\right) \qquad [3]$$