# Conjugate Analysis for the Linear Model

- If we have good prior knowledge that can help us specify priors for $\boldsymbol{\beta}$ and $\sigma^2$, we can use conjugate priors.
- Following the procedure in Christensen, Johnson, Branscum, and Hanson (2010), we will actually specify a prior for the error **precision** parameter $\tau = \frac{1}{\sigma^2}$:

$$\tau \sim \text{gamma}(a, b)$$

- This is analogous to placing an **inverse gamma** prior on $\sigma^2$.
- Then our prior on $\boldsymbol{\beta}$ will depend on $\tau$:

$$\boldsymbol{\beta}|\tau \sim MVN\Big(\boldsymbol{\delta}, \tau^{-1}[\tilde{\mathbf{X}}^{-1}\mathbf{D}(\tilde{\mathbf{X}}^{-1})']\Big)$$

(Note $\tau^{-1} = \sigma^2$)

# Conjugate Analysis for the Linear Model

- We will specify a set of $k$ *a priori* **reasonable** hypothetical observations having predictor vectors $\tilde{\mathbf{x}}_1, \ldots, \tilde{\mathbf{x}}_k$ (these — along with a column of 1's — will form the rows of $\tilde{\mathbf{X}}$) and prior expected response values $\tilde{\mathbf{y}}_1, \ldots, \tilde{\mathbf{y}}_k$.
- Our MVN prior on $\boldsymbol{\beta}$ is equivalent to a MVN prior on $\tilde{\mathbf{X}}\boldsymbol{\beta}$:

$$\tilde{\mathbf{X}}\boldsymbol{\beta}|\tau \sim MVN(\tilde{\mathbf{y}}, \tau^{-1}\mathbf{D})$$

- Hence prior mean of $\tilde{\mathbf{X}}\boldsymbol{\beta}$ is $\tilde{\mathbf{y}}$, implying that the prior mean $\boldsymbol{\delta}$ of $\boldsymbol{\beta}$ is $\tilde{\mathbf{X}}^{-1}\tilde{\mathbf{y}}$.
- $\mathbf{D}^{-1}$ is a diagonal matrix whose diagonal elements represent the weights of the "hypothetical" observations.
- Intuitively, the prior has the same "worth" as $\text{tr}(\mathbf{D}^{-1})$ observations.

# Conjugate Analysis for the Linear Model

- The joint density is

$$\pi(\boldsymbol{\beta}, \tau, \mathbf{X}, \mathbf{y}) \propto \tau^{n/2}\tau^{n/2}|\mathbf{D}|^{-1/2}\tau^{a-1}e^{-b\tau} \times \exp\Big\{-\tfrac{1}{2}(\mathbf{X}\beta - \mathbf{y})^{'}(\tau^{-1}\mathbf{I})^{-1}(\mathbf{X}\beta - \mathbf{y})\Big\} \times \exp\Big\{-\tfrac{1}{2}(\tilde{\mathbf{X}}\beta - \tilde{\mathbf{y}})^{'}(\tau^{-1}\mathbf{D})^{-1}(\tilde{\mathbf{X}}\beta - \tilde{\mathbf{y}})\Big\}$$

- It can be shown that the posterior for $\boldsymbol{\beta}|\tau$ is:

$$\boldsymbol{\beta}|\tau, \mathbf{X}, \mathbf{y} \sim MVN\big(\hat{\boldsymbol{\beta}}, \tau^{-1}(\mathbf{X}^{'}\mathbf{X} + \tilde{\mathbf{X}}^{'}\mathbf{D}^{-1}\tilde{\mathbf{X}})^{-1}\big)$$

where

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^{'}\mathbf{X} + \tilde{\mathbf{X}}^{'}\mathbf{D}^{-1}\tilde{\mathbf{X}})^{-1}[\mathbf{X}^{'}\mathbf{y} + \tilde{\mathbf{X}}^{'}\mathbf{D}^{-1}\tilde{\mathbf{y}}]$$

# Conjugate Analysis for the Linear Model

- And the posterior for $\tau$ is:

$$\tau|\mathbf{X},\mathbf{y} \sim \text{gamma}\left(\frac{n+2a}{2}, \frac{n+2a}{2}s^*\right)$$

where

$$s^* = \frac{(\mathbf{y}-\mathbf{X}\hat{\beta})'(\mathbf{y}-\mathbf{X}\hat{\beta}) + (\tilde{\mathbf{y}}-\tilde{\mathbf{X}}\hat{\beta})'\mathbf{D}^{-1}(\tilde{\mathbf{y}}-\tilde{\mathbf{X}}\hat{\beta}) + 2b}{n+2a}$$

- The subjective information is incorporated via $\hat{\beta}$ (a function of $\tilde{\mathbf{X}}$ and $\tilde{\mathbf{y}}$) and $s^*$ (a function of $\hat{\beta}$, $a$, and $b$).

# Prior Specification for the Conjugate Analysis

- We will specify a matrix $\tilde{\mathbf{X}}$ of hypothetical predictor values.
- We also specify (via expert opinion or previous knowledge) a corresponding vector $\tilde{\mathbf{y}}$ of reasonable response values for such predictors.
- The number of such "hypothetical observations" we specify must be one more than the number of predictor variables in the regression.
- Our prior mean for $\beta$ will be $\tilde{\mathbf{X}}^{-1}\tilde{\mathbf{y}}$.

# Prior Specification for the Conjugate Analysis

- We also must specify the shape parameter $a$ and the rate parameter $b$ for the gamma prior on $\tau$.
- One strategy is to choose $a$ first, based on the degree on confidence in our prior.
- For a given $a$, we can view the prior as being "worth" the same as $2a$ sample observations.
- A larger value of $a$ indicates we are more confident in our prior.

# Prior Specification for the Conjugate Analysis

- Here is one strategy for specifying $b$:
- Consider any of the "hypothetical observations" — take the first, for example.
- If $\tilde{\mathbf{y}}_1$ is the prior expected response for a hypothetical observation with predictors $\tilde{\mathbf{x}}_1$, then let $\tilde{\mathbf{y}}_{\max}$ be the *a priori* **maximum reasonable response** for a hypothetical observation with predictors $\tilde{\mathbf{x}}_1$.
- Then (based on the normal distribution) let a prior guess for $\sigma$ be $\frac{\tilde{\mathbf{y}}_{\max} - \tilde{\mathbf{y}}_1}{1.645}$.
- Since $\tau = \frac{1}{\sigma^2}$, this gives us a reasonable guess for $\tau$.
- Set this guess for $\tau$ equal to the mean $\frac{a}{b}$ of the gamma prior for $\tau$.
- Since we have already specified $a$, we can solve for $b$.

# Example of a Conjugate Analysis

- Example in `R` with Automobile Data Set
- We can get point and interval estimates for $\tau$ (and thus for $\sigma^2$).

- Given the estimate for $\tau$, we can get point and interval estimates for the elements of $\boldsymbol{\beta}$.