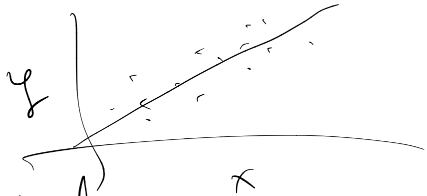
# Linear Regression




A model for real $y$ with conditional mean expressed as linear function of predictor variables (i.e., "covariates").

For a single covariate $x_i$, we have

Data model: $y_i \sim N(\beta_0 + \beta_1 x_i, \sigma^2)$, $i = 1, \ldots, n$

Parameter model(s): $\left.\begin{array}{l} \beta_0 \sim N(\mu_0, \sigma_0^2) \\ \beta_1 \sim N(\mu_1, \sigma_1^2) \\ \sigma^2 \sim IG(q, r) \end{array}\right]$ a priori independent

Posterior:

$$[\beta_0, \beta_1, \sigma^2 \mid y] \propto \left(\prod_{i=1}^{n} [y_i \mid \beta_0, \beta_1, \sigma^2]\right)[\beta_0][\beta_1][\sigma^2]$$

Full-conditional Distributions:

$$[\beta_0 \mid \cdot] \propto \prod_{i=1}^{n} [y_i \mid \beta_0, \beta_1, \sigma^2][\beta_0]$$

Note: $y - \beta_0 - \beta_1 x = (y - \beta_1 x) - \beta_0$

$$\propto \exp\left\{-\frac{1}{2} \frac{\sum (y_i - \beta_0 - \beta_1 x_i)^2}{\sigma^2}\right\} \exp\left\{-\frac{1}{2} \frac{(\beta_0 - \mu_0)^2}{\sigma_0^2}\right\}$$

$$\propto \exp\left\{-\frac{1}{2}\left(\frac{-2\sum (y_i - \beta_1 x_i)\beta_0}{\sigma^2} + \frac{n\beta_0^2}{\sigma^2} - \frac{2\mu_0 \beta_0}{\sigma_0^2} + \frac{\beta_0^2}{\sigma_0^2}\right)\right\}$$

$$\propto \exp\left\{-\frac{1}{2}\left(-2\underbrace{\left(\frac{\sum (y_i - \beta_1 x_i)}{\sigma^2} + \frac{\mu_0}{\sigma_0^2}\right)}_{b}\beta_0 + \underbrace{\left(\frac{n}{\sigma^2} + \frac{1}{\sigma_0^2}\right)}_{a}\beta_0^2\right)\right\}$$

$$= N(a^{-1}b, a^{-1})$$

Note: $y - \beta_0 - \beta_1 X = (y - \beta_0) - \beta_1 X$

$$[\beta_1 | \cdot] \propto \prod [y_i | \beta_0, \beta_1, \sigma^2] [\beta_1]$$

$$\propto \exp\left\{-\frac{1}{2} \frac{\sum_i (y_i - \beta_0 - \beta_1 x_i)^2}{\sigma^2}\right\} \exp\left\{-\frac{1}{2} \frac{(\beta_1 - \mu_1)^2}{\sigma_1^2}\right\}$$

$$\propto \exp\left\{-\frac{1}{2}\left(-2 \frac{\sum_i (y_i - \beta_0) X_i \beta_1}{\sigma^2} + \frac{(\sum_i x_i^2) \beta_1^2}{\sigma^2} - 2 \frac{\mu_1 \beta_1}{\sigma_1^2} + \frac{\beta_1^2}{\sigma_1^2}\right)\right\}$$

$$\propto \exp\left\{-\frac{1}{2}\left(-2 \underbrace{\left(\frac{\sum_i (y_i - \beta_0) x_i}{\sigma^2} + \frac{\mu_1}{\sigma_1^2}\right)}_{b} \beta_1 + \underbrace{\left(\frac{\sum_i x_i^2}{\sigma^2} + \frac{1}{\sigma_1^2}\right)}_{a} \beta_1^2\right)\right\}$$

$$= N(a^{-1} b, a^{-1})$$

$$[\sigma^2 | \cdot] \propto \left[\prod_{i=1}^{n} [y_i | \beta_0, \beta_1, \sigma^2]\right] [\sigma^2]$$

$$\propto (\sigma^2)^{-\frac{n}{2}} \exp\left\{-\frac{1}{2\sigma^2} \sum_i (y_i - \beta_0 - \beta_1 x_i)^2\right\} (\sigma^2)^{-(q+1)} \exp\left\{-\frac{1}{r\sigma^2}\right\}$$

$$\propto (\sigma^2)^{-(\underbrace{\frac{n}{2} + q}_{\tilde{q}} + 1)} \exp\left\{-\frac{1}{\sigma^2} \underbrace{\left(\frac{\sum_i (y_i - \beta_0 - \beta_1 x_i)^2}{2} + \frac{1}{r}\right)}_{\tilde{r}^{-1}}\right\}$$

$$= IG(\tilde{q}, \tilde{r})$$

---

MCMC Algorithm:

1.) Init $\beta_0^{(0)}$, $\beta_1^{(0)}$, set $k = 1$
2.) sample $\sigma^{2(k)} \sim [\sigma^2 | \beta_0^{(k-1)}, \beta_1^{(k-1)}, y]$
3.) Sample $\beta_0^{(k)} \sim [\beta_0 | \sigma^{2(k)}, \beta_1^{(k-1)}, y]$
4.) sample $\beta_1^{(k)} \sim [\beta_1 | \sigma^{2(k)}, \beta_0^{(k)}, y]$
5.) Let $k = k+1$, Goto 2 for $k \leq K$

---

Fitted values: $E(y_i | \beta_0, \beta_1, \sigma^2) = \beta_0 + \beta_1 X_i$