Last time

$$y|\theta \sim \text{binomial}(\underset{\text{fixed, known}}{n}, \theta)$$

$$\theta \sim \text{beta}(a, b)$$

Exercise : write $\underbrace{\mathbb{E}[\theta | y]}_{\text{"posterior mean"}}$ as $w\,\mathbb{E}[\theta] + (1-w)\bar{x}$

where $y = \sum_{i=1}^{n} x_i$

PROOF

$$\mathbb{E}[\theta | y] = \frac{a+y}{a+b+n} = \frac{a}{a+b+n} + \left(\frac{n}{a+b+n}\right)\bar{x}$$

idea : let $(1-w) = \dfrac{n}{a+b+n}$

$$\Rightarrow w = 1-(1-w) = \frac{a+b+n}{a+b+n} - \frac{n}{a+b+n} = \frac{a+b}{a+b+n}$$

$\Rightarrow$

$$\mathbb{E}[\theta | y] = w\,\mathbb{E}[\theta] + (1-w)\bar{x}$$

$$\frac{a+b}{a+b+n} \cdot \frac{a}{a+b} + \frac{n}{a+b+n} \cdot \bar{x}$$

$\square$

Lecture 3

1. Poisson-gamma model

2. exp. families

3. 1 line formula

1. Let y_i be # of assists made by a particular player in a particular game.

- Support of y? $y \in \{0, 1, 2, 3, 4, \ldots\}$

- $y_i | \lambda \sim \text{Poisson}(\lambda)$

Let $\lambda = \theta x$

x : # minutes a particular plays in a particular game.

θ : rate of assists per unit time

$$p(y_1, \ldots, y_n | \theta) = \prod_{i=1}^{n} p(y_i | \theta)$$

$$= \prod_{i=1}^{n} \frac{(\theta x_i)^{y_i} e^{-\theta x_i}}{y_i!}$$

2. Prior of unknowns

- What is unknown? answer: θ
- What is its support? answer: $\theta > 0$

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

We must choose a and b!

Ex 1: Choose $a = 9$, $b = 3$ s.t. $E(\theta) = 9/3 = 3$.

Is this reasonable? No!

Ex 2: Choose $a = 1$, $b = 10$ (better...)

3. Posterior

$$p(\theta \mid y_1, \dots, y_n) \propto \underbrace{p(y_1, \dots, y_n \mid \theta)}_{\text{likelihood}} \overbrace{p(\theta)}^{\text{prior}}$$

$$\propto \left(\theta^{\Sigma y_i} e^{-\theta \Sigma x_i}\right) \cdot \left(\theta^{a-1} e^{-b\theta}\right)$$

$$\propto \underbrace{\theta^{\Sigma y_i + a - 1} e^{-\theta(b + \Sigma x_i)}}$$

kernel of gamma(α, β)

$$\begin{cases} \alpha = \Sigma y_i + a \\ \beta = b + \Sigma x_i \end{cases}$$

sometimes posterior parameters written w/ subscript "n" (α_n, β_n) to emphasize dependence on $\underline{n}$ datapoints.

4. What is $E(\theta \mid y_1, \dots, y_n)$?

"posterior expectation of θ"

$$\frac{\alpha}{\beta} = \frac{\Sigma y_i + a}{b + \Sigma x_i}$$

How is this diff't from prior expectation?

$E\theta = a/b$

Σy_i = total assists

Σx_i = total min / played

prior says $\underline{a}$ assists in $\underline{b}$ minutes.

$var(\theta \mid y_1, \dots, y_n) = \alpha / \beta^2$

Poisson as a member of the exp. family

4.

$$p(y|\lambda) = \underbrace{\frac{1}{y!}}_{h(y)} \lambda^y e^{-\lambda}$$

Now λ^y needs to look like $e^{\phi t(y)}$

Let $\boxed{\phi = \log\lambda}$, then $e^{\phi} = \lambda$

$$p(y|\phi) = \underbrace{\boxed{\frac{1}{y!}}}_{h(y)} \underbrace{\boxed{e^{\phi y}}}_{e^{\phi t(y)}} \underbrace{\boxed{e^{-e^{\phi}}}}_{c(\phi)}$$

$t(y) = y$

therefore the conj. prior is

$$p(\phi|n_0,t_0) \propto c(\phi)^{n_0} e^{n_0 t_0 \phi}$$

$$= e^{-e^{\phi} \cdot n_0} e^{n_0 t_0 \phi}$$

how to go from $p(\phi|n_0,t_0) \to p(\lambda|n_0,t_0)$?

one-line formula:

$$p(\phi|n_0,t_0)\,d\phi = p(\lambda|n_0,t_0)\,d\lambda$$

$$p(\lambda|n_0,t_0) = p(\phi|n_0,t_0)\left|\frac{d\phi}{d\lambda}\right| \qquad \frac{d}{d\lambda}\log\lambda = \frac{1}{\lambda}$$

$$\underbrace{e^{-e^{\log\lambda} \cdot n_0} e^{n_0 t_0 \log\lambda}} \cdot \left|\frac{d\phi}{d\lambda}\right|$$

$$= e^{-\lambda n_0} \lambda^{n_0 t_0} \cdot \left|\frac{1}{\lambda}\right|$$

$$= \underbrace{e^{-\lambda n_0} \lambda^{n_0 t_0 - 1}}_{\text{gamma}(n_0 t_0,\ n_0)}$$

[prior parameters are sometimes written w/ subscript "0" e.g. n_0, t_0 to emphasize parameters depend on 0 data.]