

Lecture 11 3/21/17 Part 391

Review from last time...

$$X_1, \dots, X_n | \theta \sim \text{iid}(\text{Bern}(\theta))$$

$$\theta \sim \text{Beta}(\alpha, \beta)$$

$$\theta | X_1, \dots, X_n \sim \text{Beta}(n + \alpha, \sum X_i + \beta)$$

$$\Rightarrow \hat{\theta}_{MLE} = \dots = \frac{1}{1 + \bar{x}}$$

see previous lecture

$$\Rightarrow \hat{\theta}_{muse} = \frac{n + \alpha}{n + \alpha + \sum X_i + \beta}$$

α : # pseudocounts

β : # pseudocounts in reverse

Uniform priors

Haldane: $\theta \sim \text{Beta}(0, 0)$

$$\Rightarrow \hat{\theta}_{muse} = \frac{n}{n + \sum X_i} = \frac{1}{1 + \frac{\sum X_i}{n}} = \frac{1}{1 + \bar{x}} = \hat{\theta}_{MLE}$$

Laplace: $\theta \sim \text{Beta}(1, 1)$

$$\Rightarrow \hat{\theta}_{muse} = \frac{n+1}{n+1 + \sum X_i + 1} = \frac{1}{1 + \frac{\sum X_i + 1}{n+1}} \neq 1$$

Jeffreys: $\theta \sim \text{Beta}(0.5, 0.5)$

$$\Rightarrow \hat{\theta}_{muse} = \frac{n}{n + \sum X_i + 1}$$

$\frac{2n}{\bar{x}}$ but \bar{x} shrunk towards $E(\theta) = 0.5$

$$= \frac{1}{1 + \frac{\sum X_i + 1}{n}} \neq 1$$

same idea

$\theta \sim U(0,1)$

$$\hat{\theta}_{MSE} = \frac{n+1}{n+1+\sum x_i+1} = \frac{1}{1 + \frac{\sum x_i+1}{n+1}}$$

\bar{x} shrink toward $E(\theta)$

$\theta \sim \text{Beta}(\alpha, \beta)$

$$\hat{\theta}_{MSE} = \frac{n}{n+\sum x_i} = \frac{1}{1+\bar{x}} = \hat{\theta}_{MLE} \quad (\text{same as before})$$

Is there shrinkage?

$$\frac{1}{\hat{\theta}_{MSE}} = \frac{n+\alpha+\sum x_i+\beta}{n+\alpha} = \frac{\alpha+\beta}{n+\alpha} \cdot \frac{n}{\alpha} + \frac{n+\sum x_i}{n+\alpha} \cdot \frac{1}{n} = \frac{\alpha+\beta}{\alpha} \cdot \frac{\alpha}{n+\alpha} + \frac{n+\sum x_i}{n} \cdot \frac{n}{n+\alpha}$$

$$= p \frac{1}{E(\theta)} + (1-p) \frac{1}{\hat{\theta}_{MLE}}$$

this is shrinkage but a bit different from before...

Jeffrey's prior?

$$p(\theta) \propto \sqrt{J(\theta)} = \sqrt{\frac{1}{\theta^2(1-\theta)^2}} \propto \theta^{-1}(1-\theta)^{-1} \propto \text{Beta}(\theta, \frac{1}{2}) \quad \text{Improper!}$$

Post. Pred. Distrib.

$$p(x^* | x) = \int_0^1 p(x^* | \theta) p(\theta | x) d\theta = \int_0^1 (1-\theta)^{x^*} \theta \frac{1}{B(n+\alpha, \sum x_i + \beta)} \theta^{n+\alpha-1} (1-\theta)^{\sum x_i + \beta - 1} d\theta$$

for $n^*=1$

$$= \frac{1}{B(n+\alpha, \sum x_i + \beta)} \int_0^1 \theta^{n+\alpha+1-1} (1-\theta)^{x^* + \sum x_i + \beta - 1} d\theta = \frac{B(n+\alpha+1, x^* + \sum x_i + \beta)}{B(n+\alpha, \sum x_i + \beta)}$$

geometric
over-dispersed geometric

$$E(x^* | x) = \frac{\sum x_i + \beta}{n + \alpha - 1}$$

Beta-geometric Prior

$$\text{BetaGeo}(n+\alpha, \sum x_i + \beta) := \checkmark$$

Not covered in class

3

$$X_1, \dots, X_n | \theta \sim \text{Ngbin}(r, \theta) \quad \text{fixed}(r)$$

$$\theta \sim \text{Beta}(\alpha, \beta)$$

$$\theta | X_1, \dots, X_n \sim \text{Beta}(r+1 + \alpha, \sum X_i + \beta)$$

$$h^* = 1, \quad X^* | X \sim \text{Ngbin}$$

$$X \sim \text{Bin}(n, \theta) := \binom{n}{x} \theta^x (1-\theta)^{n-x} \quad \text{supp}(X) = \{0, \dots, n\}, \quad \theta \in \mathcal{H} = (0, 1)$$

if $n \rightarrow \infty, \theta \rightarrow 0$ s.t. $\lambda = n\theta$

$$\lim_{n \rightarrow \infty} \frac{n!}{x!(n-x)!} \left(\frac{\lambda}{n}\right)^x \left(1 - \frac{\lambda}{n}\right)^{n-x} = \frac{\lambda^x}{x!} \lim_{n \rightarrow \infty} \frac{1}{\frac{n \cdot (n-1) \cdots (n-x+1)}{n^x}} = \frac{\lambda^x}{x!} \lim_{n \rightarrow \infty} \left(1 - \frac{\lambda}{n}\right)^n \cdot \left(1 - \frac{\lambda}{n}\right)^{x-n} = \frac{\lambda^x e^{-\lambda}}{x!}$$

$$X \sim \text{Poisson}(\theta) := \frac{e^{-\theta} \theta^x}{x!}$$

↑
differs $\theta!!!$

$$\text{supp}(X) = \{0, \dots\} = \mathbb{N}_0, \quad \theta \in \mathcal{H} = (0, \infty)$$

$$E[X] = \theta$$

$$\text{Var}(X) = \theta$$

Final practice

$$P(\theta|x) \propto P(x|\theta) P(\theta) = \frac{e^{-\theta} \theta^x}{x!} P(\theta) \propto e^{-\theta} \theta^x P(\theta)$$

$P(\theta)$ should have the form...

$$P(\theta) \propto e^{-b\theta} \theta^a$$

Integrate and see

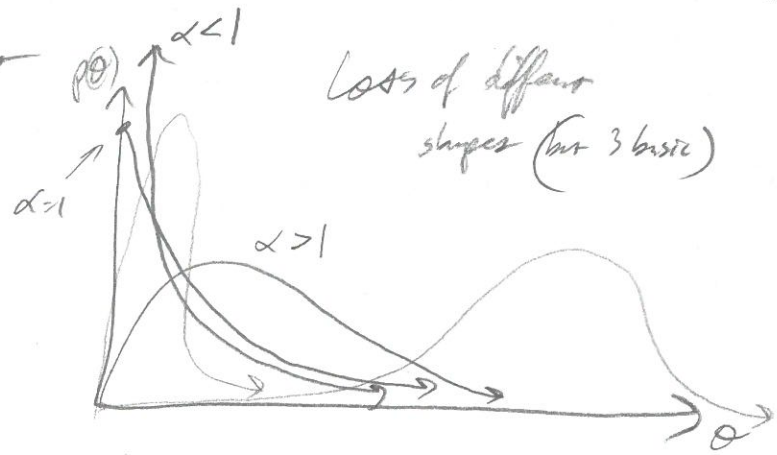
$$P(\theta) = \frac{b^{a+1}}{\Gamma(a+1)} e^{-b\theta} \theta^a$$

this is called the gamma distr.

Usually it's parameterized via

$$\theta \sim \text{Gamma}(\alpha, \beta) := \frac{\beta^\alpha}{\Gamma(\alpha)} e^{-\beta\theta} \theta^{\alpha-1}$$

Supp(θ) = $(0, \infty)$ exactly what we want
 param space $\alpha, \beta > 0$ just as before



lots of different shapes (but 3 basic)

$$\begin{aligned} E(\theta) &= \frac{\alpha}{\beta} \\ \text{Var}(\theta) &= \frac{\alpha}{\beta^2} \\ \text{Mode}(\theta) &= \frac{\alpha-1}{\beta} \text{ if } \alpha \geq 1 \end{aligned}$$

calculs exercises

$\text{Mod}(\theta) = \text{ggamm}(0.5, \alpha, \beta)$ (i.e., no closed form)

$$p(\theta|x) \propto p(x|\theta)p(\theta) = \frac{e^{-\theta}\theta^x}{x!} \frac{\beta^\alpha}{\Gamma(\alpha)} e^{-\beta\theta}\theta^{\alpha-1} \propto e^{-\theta}\theta^x e^{-\beta\theta}\theta^{\alpha-1} = e^{-(\beta+1)\theta}\theta^{x+\alpha-1}$$

$X|\theta \sim \text{Poisson}(\theta)$
 $\theta \sim \text{Gamma}(\alpha, \beta) \propto \text{Gamma}(x+\alpha, \beta+1)$

$\theta|x \sim \text{Gamma}(x+\alpha, \beta+1)$
 "Gamma is conjugate prior for Poisson likelihood"

Now lets say you have n iid Poissons

$$X_1, \dots, X_n | \theta \sim \text{Poisson}(\theta)$$

$\theta \sim \text{Gamma}(\alpha, \beta)$

$\theta | X_1, \dots, X_n \sim ?$

$$p(\theta|x) \propto p(x|\theta)p(\theta) = \left(\prod_{i=1}^n \frac{e^{-\theta}\theta^{x_i}}{x_i!} \right) \left(\frac{\beta^\alpha}{\Gamma(\alpha)} e^{-\beta\theta}\theta^{\alpha-1} \right) = \frac{e^{-\sum \theta_i} \theta^{\sum x_i}}{\prod x_i!} \frac{\beta^\alpha}{\Gamma(\alpha)} e^{-\beta\theta}\theta^{\alpha-1}$$

$$\propto e^{-n\theta} \theta^{\sum x_i} e^{-\beta\theta} \theta^{\alpha-1} \propto \text{Gamma}(\sum x_i + \alpha, n + \beta)$$

Interpretation

α = sum total of successes seen previously

β = # of pseudo trials performed

$$E(\theta) = \frac{\alpha}{\beta} = \text{avg \# of successes / trial}$$

$$\hat{\theta}_{\text{unbiased}} = \frac{\sum x_i + \alpha}{n + \beta}, \quad \hat{\theta}_{\text{prop}} = \text{gamma}(0.5, \sum x_i + \alpha, n + \beta), \quad \hat{\theta}_{\text{prop}} = \frac{\sum x_i + \alpha - 1}{n + \beta} \approx \hat{\theta}_{\text{unbiased}}$$

5

Can we say $\theta \sim U$? No..

But... what if $P(\theta) \propto 1$ Try proper since $\int_0^{\infty} c d\theta = \infty$

What happens?

$$P(\theta|x) \propto P(x|\theta) P(\theta) \propto e^{-n\theta} \theta^{\sum x_i + 1 - 1} \propto \text{gamma}(\sum x_i + 1, n) \quad (\text{always proper})$$

$$\theta \sim \text{Gamma}(1, 0) \propto 1$$

Laplace prior (improper)

What is MLE? Back to math stats...

Haldane prior is not considered for this model.

$$L(\theta; x) = \prod_{i=1}^n \frac{e^{-\theta} \theta^{x_i}}{x_i!} = \frac{e^{-n\theta} \theta^{\sum x_i}}{\prod x_i!}$$

$$l(\theta; x) = -n\theta + \sum x_i \ln(\theta) - \ln(\prod x_i!)$$

$$l'(\theta; x) = -n + \frac{\sum x_i}{\theta} \stackrel{\text{set}}{=} 0 \Rightarrow \frac{\sum x_i}{\theta} = n \Rightarrow \boxed{\hat{\theta} = \bar{x}}$$

Jeffrey's Prior?

$$l''(\theta; x) = -\frac{\sum x_i}{\theta^2}$$

$$I(\theta) = E[-l''(\theta; x)] = E\left(\frac{\sum x_i}{\theta^2}\right) = \frac{E(\sum x_i)}{\theta^2} = \frac{\sum E(x_i)}{\theta^2} = \frac{\sum \theta}{\theta^2} = \frac{n\theta}{\theta^2} = \frac{n}{\theta}$$

$$P(\theta) \propto \sqrt{I(\theta)} = \sqrt{\frac{n}{\theta}} \propto \sqrt{\frac{1}{\theta}} = \theta^{-\frac{1}{2}} \propto \text{Gamma}\left(\frac{1}{2}, 0\right)$$

Same idea as before... see I tried... but know here 0.5 successes somewhere!
 so two choices for uninformative prior: Haldane, Jeffrey's. Jeffrey's lead to proper post. always.

Shrinkage? Yes, should weighted ^{average} shrinkage

$$\hat{\theta}_{\text{shrink}} = \frac{\sum x_i + \alpha}{n + \beta} = \frac{1}{n} \cdot \frac{\sum x_i}{n + \beta} + \frac{\alpha}{n + \beta} \cdot \frac{\beta}{\beta} = \frac{1 - \rho}{n + \beta} \cdot \frac{\sum x_i}{n} + \frac{\rho}{n + \beta} \cdot \frac{\alpha}{\beta}$$

$$\lim_{n \rightarrow \infty} \rho = 0 \quad \checkmark$$

Poisson prob. distr.

$X|\theta \sim \text{Poisson}(\theta)$



Now θ is unknown... so this will be "dispersed"

the beta-binomial is the "dispersed binomial"

the beta-geometric is the "dispersed geometric"

What is a "dispersed Poisson"?

$$P(x^* | x) = \int_{\Theta} P(x^* | \theta) P(\theta | x) d\theta \quad \text{for } x^* = 1, \dots$$

$$= \int_0^{\infty} \left(\frac{e^{-\theta} \theta^{x^*}}{x^*!} \right) \left(\frac{\beta^{\alpha'}}{\Gamma(\alpha')} e^{-\beta\theta} \theta^{\alpha'-1} \right) d\theta$$

$$= \frac{\beta^{\alpha'}}{\Gamma(\alpha') x^*!} \int_0^{\infty} \underbrace{e^{-(\beta+1)\theta} \theta^{x^* + \alpha' - 1}}_{\text{kernel of Gamma}(x^* + \alpha', \beta + 1)} d\theta$$