Lecture - 12 03/19/2020

$$X \sim \text{Poisson}(\theta) = \frac{e^{-\theta}\theta^{x}}{\underbrace{x!}_{\text{constant}}} = p(x;\theta) \propto \overbrace{e^{-\theta}\theta^{x}}^{\text{kernal}}$$
(Discrete)

$$y \sim \text{Gamma}(\alpha,\beta) = \frac{\beta^{\alpha}}{\Gamma(\alpha)} y^{\alpha-1} e^{-\beta y} \propto y^{\alpha-1} e^{-\beta y}$$
(Continuous)

F: iid poisson $x_1, \dots, x_n ; \theta \overset{iid}{\sim} \text{poisson}(\theta)$

$$p(x \mid \theta) = \prod_{i=1}^{n} \frac{e^{-\theta}\theta^{x_i}}{x_i!} = \frac{e^{-n\theta}\theta^{\Sigma x_i}}{\prod_{i=1}^{n} x_i!} \; ; \; \theta \sim \text{Gamma}(\alpha,\beta)$$

Put a prior on θ; last time we learnt that the gamma is conjugate for the poisson likelihood model.

$$P(\theta \mid x) = \frac{P(x \mid \theta)\,P(\theta)}{P(x)} \propto P(x \mid \theta)\,P(\theta)$$

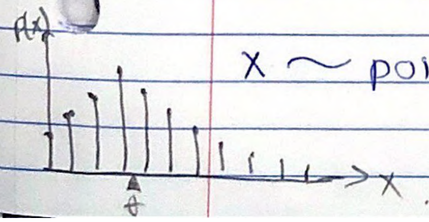
$$= \underbrace{\left(\frac{e^{-n\theta}\theta^{\Sigma x_i}}{\prod_{i=1}^{n} x_i!}\right)}_{\text{likelihood}} \underbrace{\left(\frac{\beta^{\alpha}}{\Gamma(\alpha)} \theta^{\alpha-1} e^{-\beta\theta}\right)}_{\text{Gamma}}$$

$$\propto \underbrace{e^{-n\theta}\theta^{\Sigma x_i}\theta^{\alpha-1}e^{-\beta\theta}}_{\text{(kernal)}} = \theta^{\Sigma x_i + \alpha - 1} e^{-(n+\beta)\theta}$$

$$\propto \text{Gamma}(\Sigma x_i + \alpha, n+\beta) = P(\theta \mid x)$$

$P(\theta) \xrightarrow{x} P(\theta \mid x)$ Gamma distribution

$$\text{Gamma}(\alpha,\beta) \xrightarrow{x} \text{Gamma}(\Sigma x_i + \alpha, n+\beta)$$ # of trials



$X \sim \text{poisson}(\theta)$ $\text{Supp}[x] = \mathbb{N}_0 \; (0, 1, \dots \infty)$

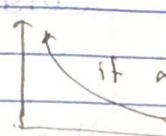
$\theta = np$ (# of trials, prob. of success)

parameter space $\theta \in (0, \infty)$

$E[x] = \theta$

$$E[\theta] = \frac{\alpha}{\beta}, \quad \text{mode}[\theta] = \frac{\alpha - 1}{\beta} \quad \text{if } \alpha \geq 1$$

$$\hat{\theta}_{MMSE} = E[\theta \mid x] = \frac{\sum x_i + \alpha}{n + \beta}$$



if $\alpha < 1$ (no mode in the function

$$\hat{\theta}_{MMAE} = \text{med}[\theta \mid x] = \text{qgamma}(0.5, \textstyle\sum x_i + \alpha, n + \beta)$$

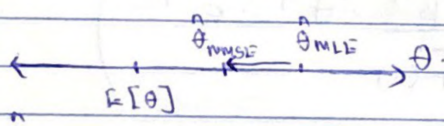
$$\hat{\theta}_{MAP} = \text{mode}[\theta \mid x] = \frac{\sum x_i + \alpha - 1}{n + \beta} \; ; \text{ if } \textstyle\sum x_i + \alpha \geq 1$$

Pseudo count interpretation.

$$\theta \mid x \sim \text{Gamma}(\underbrace{\textstyle\sum x_i}_{\text{\# of total success}} + \underbrace{\alpha}_{\text{\# of pseudo success}} \; ; \; \underbrace{n}_{\text{\# of trials}} + \underbrace{\beta}_{\text{\# of pseudo trials } (n_0)})$$

# of experiment you made; (If gamma be large ...) stronger data

Shrinkage.



$$\hat{\theta}_{MMSE} = \frac{\sum x_i + \alpha}{n + \beta} = \frac{\sum x_i}{n+\beta} \cdot \frac{n}{\underset{\bar{x}}{n}} + \frac{\alpha}{n+\beta} \cdot \frac{\beta}{\beta}$$

Prior expectation $= E[\theta]$

$$= \underbrace{\frac{n}{n+\beta}}_{1-\rho} \cdot \underset{\hat{\theta}_{MLE}}{\bar{x}} + \underbrace{\frac{\beta}{n+\beta}}_{\rho} E[\theta]$$

If you have more & more data, you will get more & more shrinkage.

$$\lim_{n \to \alpha} \rho = 0$$

Prove the MLE

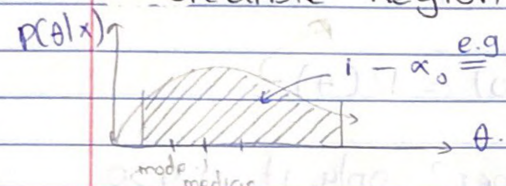
$$\mathcal{L}(\theta; x) = \frac{e^{-n\theta}\,\theta^{\Sigma x_i}}{\prod x_i!}, \quad \ell(\theta; x) = -n\theta + (\Sigma x_i)\ln(\theta) - \ln(\prod x_i!)$$

$$\ell'(\theta; x) = -n + \frac{\Sigma x_i}{\theta} \overset{set}{=} 0$$

$$\implies \frac{\Sigma x_i}{\theta} = n$$

$$\implies \Sigma x_i = n\theta. \quad \implies \hat{\theta}_{MLE} = \bar{X}$$

Credible Region



$$CR_{\theta, 1-\alpha} = [\,qgamma(\alpha_{0}/2, \Sigma x_i + \alpha, n+\beta),\ qgamma(1-\alpha_0/2, \Sigma x_i + \alpha, n+\beta)\,]$$

Baysian hypothesis testing (same as before)

Uninformative priors. (principal)

1. Laplace indifference prior Gamma (1,0)
2. Haldane ignorance prior Gamma (0,0)
3. Jeffrey's prior Gamma (½,0)

$\theta \sim Gamma(\alpha, \beta)$ $\quad Supp[\theta] = [0, \alpha)$

$\theta \sim U(0, \alpha)$; Answer $\implies$ NO

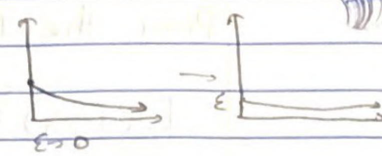
$P(\theta) = \frac{1}{\alpha} = 0$ not valid pdf!

$$P(\theta) = Gamma(1, \varepsilon) = \frac{\varepsilon}{\Gamma(1)}\,\theta^{1-1} e^{-\varepsilon\theta}$$

$$= \varepsilon e^{-\varepsilon\theta}$$

If ε is really small $\longrightarrow \approx 0$

$$P(0) = \varepsilon e^{-\varepsilon(0)} = \varepsilon$$



<u>Laplace</u>

$\theta \sim Gamma(1, 0)$ proper? No, b/c $\beta \not> 0$

n has to be greater than 1

$\Longrightarrow P(\theta|x) = Gamma(\sum x_i + 1, n)$

Proper? Yes (always)

shrinkage is undefined here

<u>Haldane</u>: $\beta = 0$, $\alpha = 0$

\# of pseudo success

\# of pseudo trials

$$p(\theta|x) = Gamma(\sum x_i + \alpha, n + \beta)$$

$$Gamma(0, 0) = P(\theta)$$

$\Longrightarrow Gamma(\sum x_i, n)$ proper? only if $\exists x_i > 0$.

$$\hat{\theta}_{MMSE} = E[\theta|x] = \frac{\sum x_i}{n} = \bar{x} = \hat{\theta}_{MLE}$$

<u>Jeffrey's prior</u>

$$P_J(\theta) \propto \sqrt{I(\theta)} = \sqrt{n/\theta} \propto \theta^{-1/2} = \theta^{1/2 - 1} \quad Gamma(1/2, 0)$$

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

$$I(\theta) = E_x[-l''(\theta; x)] = E_x\left[\frac{\sum x_i}{\theta^2}\right] = \frac{1}{\theta^2}\sum_{i=1}^{n} E[x_i]$$

$$= \frac{1}{\theta^2} n\theta = n \cdot 1/\theta$$

$P(\theta|x) = Gamma(\sum x_i + 1/2, n)$ Always proper? YES!

Prediction

X_* is the next observation that you want to predict

$X_* | X \sim ?$ $\quad \text{Supp}[X_* | X] = \text{Supp}[X] = \{0, 1, 2, \ldots\}$

$$p(x_* | x) = \int_{\Theta} \underbrace{p(x_* | \theta)}_{\text{likelihood}} \underbrace{p(\theta | x)}_{\text{posterior}} \, d\theta .$$

$$= \int_0^{\infty} \left(\frac{e^{-\theta} \theta^{x_*}}{x_*!} \right) \left(\frac{(n+\beta)^{\sum x_i + \alpha}}{\Gamma(\sum x_i + \alpha)} \cdot \theta^{\sum x_i + \alpha - 1} e^{-(n+\beta)\theta} \right) d\theta$$