today: Poisson; bivariate uncertainty

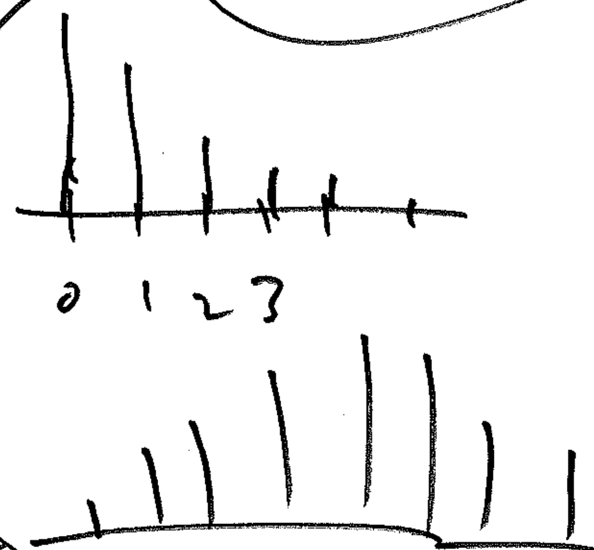
CS3: L05

AMS 206
19 Feb 19

① our uncertainty about $y = (y_1, \dots, y_n)$ is exchangeable → $n = 14$

$$\left\{ (Y_i \mid ?) \overset{IID}{\sim} ? \qquad ? \sim ? \right\}$$

y_i: non-negative integers



$$(Y_i \mid \lambda \, \underset{\text{poisson}}{\mathcal{P}} \, \mathcal{B}) \overset{IID}{\sim} \text{Poisson}(\lambda)$$

$(\lambda > 0)$

⓪ for integer $y_i \geq 0$ (marginal)

$$p_{Y_i}(y_i \mid \lambda \, \mathcal{P} \, \mathcal{B}) = \frac{\lambda^{y_i} e^{-\lambda}}{y_i!}$$

① (joint)

$$p_{Y_1, \dots, Y_n}(y_1, \dots, y_n \mid \lambda \, \mathcal{P} \, \mathcal{B}) = \prod_{i=1}^{n} \frac{\lambda^{y_i} e^{-\lambda}}{y_i!}$$

$$= \frac{\lambda^s e^{-n\lambda}}{\prod_{i=1}^{n} y_i!}$$

② $(\lambda > 0)$ $\left(s = \sum_{i=1}^{n} y_i \right)$ ②

$$\ell(\lambda \mid y\, \mathcal{P}\, \mathcal{B}) = c\, \lambda^s e^{-n\lambda}$$

③ $$\ell\ell(\lambda \mid y\, \mathcal{P}\, \mathcal{B}) = s \log \lambda - n\lambda$$

(y) s is sufficient for λ (n)

④ $$\frac{d}{d\lambda} \ell\ell(\lambda \mid y\, \mathcal{P}\, \mathcal{B}) = \frac{s}{\lambda} - n = 0$$

$$\rightarrow \hat{\lambda}_{MLE} = \frac{s}{n} = \bar{y} = \frac{1}{n} \sum_{i=1}^{n} y_i \doteq 2.07$$

⑤ $$\frac{d^2}{d\lambda^2} \ell\ell(\lambda \mid y\, \mathcal{P}\, \mathcal{B}) = -\frac{s}{\lambda^2}$$

$$\hat{I}(\hat{\lambda}_{MLE}) = \left[-\frac{d^2}{d\lambda^2} \ell\ell(\lambda \mid y\, \mathcal{P}\mathcal{B}) \right]_{\lambda = \hat{\lambda}_{MLE}}$$

3

$$= \left(\frac{s}{\lambda^2} \right)_{\lambda = \frac{s}{n}} = \frac{s}{\frac{s^2}{n^2}} = \frac{n^2}{s} = n \left(\frac{n}{s} \right) = \frac{n}{\bar{y}} = O(n)$$

$$\hat{V}_{RS}(\hat{\lambda}_{MLE}) = \hat{I}^{-1}_{RS}(\hat{\lambda}_{MLE}) = \frac{\bar{y}}{n}$$

$$\hat{SE}_{RS}(\hat{\lambda}_{MLE}) = \sqrt{\hat{V}_{RS}(\hat{\lambda}_{MLE})} = \sqrt{\frac{\bar{y}}{n}}$$

approx. 95% CI for λ

$(\alpha = 0.05)$ $\quad \Phi^{-1}(1 - \frac{\alpha}{2})$

$$\hat{\lambda}_{MLE} \pm 1.96 \, \hat{SE}(\hat{\lambda}_{MLE})$$

Bayesian algorithm

(B) (1) (2) (3) same as ML story (maximum likelihood)

$$\ell(\lambda \mid y \, P \, B) = c \, \lambda^s e^{-n\lambda}$$

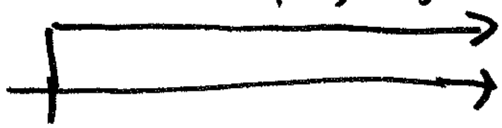
(4) is there a conjugate prior? yes: $\Gamma(\alpha, \beta)$ $(s > 0, n > 0)$

sidebar (familiarize yourself with the $\Gamma(\alpha, \beta)$ family of PDFs as ($\alpha > 0$, $\beta > 0$) vary.

5 conjugate updating

$$c\,\lambda^{(\alpha+s)-1} e^{-(\beta+n)\lambda} = \underbrace{\left[c\,\lambda^{\alpha-1} e^{-\beta\lambda}\right]}_{\text{prior}} \underbrace{\left[c\,\lambda^{s} e^{-n\lambda}\right]}_{\text{likelihood}}$$

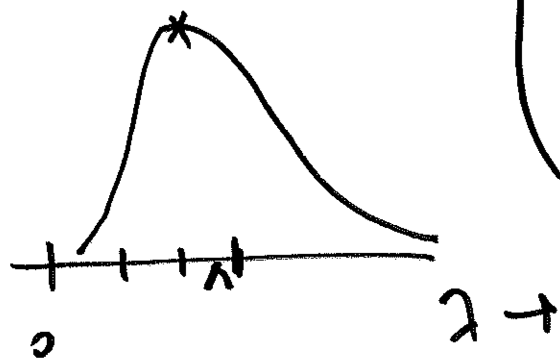
$U(0, \infty)$



little info about λ external to y

an improper prior for λ

$\int_0^{\infty} p(\lambda \mid \mathcal{B})\, d\lambda = +\infty$



$\ell(\lambda \mid y \, \mathcal{P} \, \mathcal{B})$

$p(\lambda \mid y \, \mathcal{P} \, \mathcal{B})$

(5)

$\Gamma(\alpha, \beta)$ prior ; prior sample size β

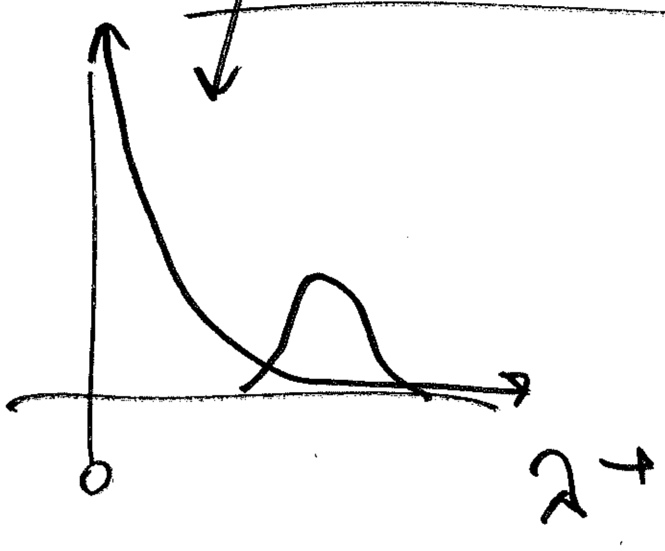
so just choose β small but > 0,

e.g. $\beta = \epsilon \doteq 0.01$

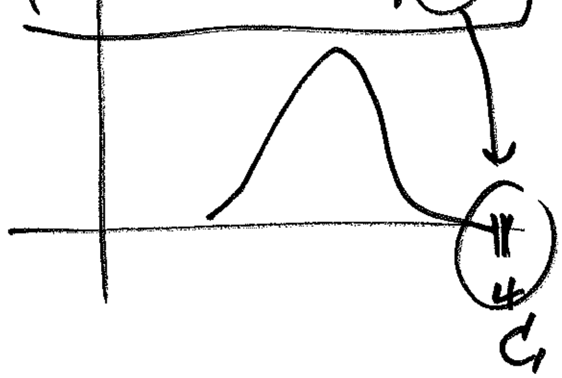
prior data set : prior mean of $\frac{\alpha}{\beta}$

common lazy (diffuse, flat) but proper ~~noninformative~~ prior for any parameter, e.g., $\lambda > 0$:

$\Gamma(\epsilon, \epsilon)$ prior (prior data set with ϵ votes & mean of 1)



$\boxed{\lambda \sim U(0, 4)}$ 74



Description ✓
Inference ✓
Prediction ✓
Decision

$$(\lambda \mid \alpha \beta \Gamma) \sim \Gamma(\alpha, \beta)$$ 6

$$(Y_i \mid \lambda \, \mathcal{P} \, \mathcal{B}) \overset{IID}{\sim} \text{Poisson}(\lambda)$$

$$(y_1, \dots, y_n) \quad (i = 1, \dots, n)$$

$$(\lambda \mid y \, \mathcal{P} \, \Gamma \, \mathcal{B}) \sim \Gamma(\alpha + s, \beta + n)$$

posterior predictive distribution for

$$p(y_{n+1} \mid y \, \mathcal{P} \, \Gamma \, \mathcal{B})$$ a new data value

Law of total prob. (LTP)

$$= \int_0^\infty p(y_{n+1}, \lambda \mid y \, \mathcal{P} \, \Gamma \, \mathcal{B}) \, d\lambda$$

$$= \int_0^\infty p(y_{n+1} \mid \lambda \, \cancel{y} \, \mathcal{P} \, \Gamma \, \mathcal{B}) \boxed{p(\lambda \mid y \, \mathcal{P} \, \Gamma \, \mathcal{B})} \, d\lambda$$

$$= \int \left(\text{Poisson sampling dist}\right) \left(\Gamma \text{ posterior}\right)$$