

Noninformative Priors

Want data to be emphasized.

Idea: use invariance

Ex: Jeffreys' Prior is invariant to reparameterization

$$\begin{aligned} P(\theta) &\propto I(\theta)^{1/2} & I(\theta) &= \text{fisher information} \\ &= \mathbb{E}_{X|\theta} \left[\left(\frac{d\ell}{d\theta} \right)^2 \right] & \ell(\theta) &= \log P(X|\theta) \\ &= -\mathbb{E}_{X|\theta} \left[\frac{d^2 \ell}{d\theta^2} \right] & & \text{(under regularity conditions)} \end{aligned}$$

In this case, for $\lambda = h(\theta)$

$$\begin{aligned} I(\theta) &= \mathbb{E}_{X|\theta} \left[\left(\frac{d\ell}{d\theta} \right)^2 \right] = \mathbb{E}_{X|\lambda} \left[\left(\frac{d\ell}{d\lambda} \cdot \frac{d\lambda}{d\theta} \right)^2 \right] = \mathbb{E}_{X|\lambda} \left[\left(\frac{d\ell}{d\lambda} \right)^2 \right] \left(\frac{d\lambda}{d\theta} \right)^2 \\ &= I(\lambda) \left(\frac{d\lambda}{d\theta} \right)^2 \end{aligned}$$

$$\text{Then } P(\lambda) = \left| \frac{d\theta}{d\lambda} \right| P(\theta) \propto \left| \frac{d\theta}{d\lambda} \right| I(\theta)^{1/2} = \cancel{\left| \frac{d\theta}{d\lambda} \right|} I(\lambda)^{1/2} \cancel{\left| \frac{d\lambda}{d\theta} \right|} = I(\lambda)^{1/2}$$

So the prior is still prop. to square root of information

$$\text{Ex: } X|\theta \sim \text{Binom}(n, \theta) \quad I(\theta) = -\mathbb{E}_{X|\theta} \left[\frac{d^2 \ell}{d\theta^2} \right]$$

$$\ell(\theta) = \log \binom{n}{x} + x \log \theta + (n-x) \log(1-\theta)$$

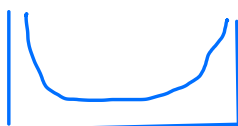
$$\frac{dl}{d\theta} = \frac{x}{\theta} - \frac{n-x}{1-\theta} \Rightarrow \frac{dl^2}{d\theta^2} = -\frac{x}{\theta^2} - \frac{n-x}{(1-\theta)^2}$$

$$-E_{X|\theta} \left[-\frac{x}{\theta^2} - \frac{n-x}{(1-\theta)^2} \right] = \frac{n\theta}{\theta^2} + \frac{n-n\theta}{(1-\theta)^2}$$

$$= n \left(\frac{1}{\theta} + \frac{1}{1-\theta} \right)$$

$$= \frac{n}{\theta(1-\theta)}$$

$$\text{Thus, } P(\theta) \propto I(\theta)^{1/2} \propto \theta^{-1/2} (1-\theta)^{-1/2} \sim \text{Beta}(1/2, 1/2)$$



arcsin distribution

Examples from Exp. Families

(i)

$$X_i | \mu \stackrel{iid}{\sim} N(\mu, \sigma^2)$$

\uparrow known

$$P(X_1, \dots, X_n | \mu) \propto \prod_{i=1}^n \exp \left\{ -\frac{1}{2\sigma^2} (x_i - \mu)^2 \right\} = \exp \left(-\frac{1}{2\sigma^2} \sum_{i=1}^n \left((x_i - \bar{x}) + (\bar{x} - \mu) \right)^2 \right)$$

$$\begin{aligned} \text{But } \sum_i (x_i - \mu)^2 &= \sum_i (x_i - \bar{x})^2 + 2 \cancel{\left(\sum_i x_i - \bar{x} \right)} (\bar{x} - \mu) + n(\bar{x} - \mu)^2 \\ &= \sum_i (x_i - \bar{x})^2 + n(\bar{x} - \mu)^2 \end{aligned}$$

So

$$P(X|\mu) \propto \exp \left(-\frac{1}{2\sigma^2} \sum_i (x_i - \bar{x})^2 \right) \exp \left(-\frac{1}{2\sigma^2} n(\bar{x} - \mu)^2 \right) \propto \exp \left(-\frac{1}{2\sigma^2 n} \underbrace{(\sum_i (x_i - \bar{x})^2)}_{data} \right)$$

This implies (in freq) $\bar{X} | \mu \sim N(\mu, \sigma^2/n)$

$$P(\mu) \propto \exp\left(-\frac{1}{2\tau_0^2} (\mu_0 - \mu)^2\right) \Rightarrow \mu \sim N(\mu_0, \tau_0^2)$$

$$\xRightarrow{\text{implies}} \mu | X \sim N(\mu_1, \tau_1^2) \text{ such that } \frac{1}{\tau_1^2} = \frac{1}{\tau_0^2} + \frac{n}{\sigma^2}, \mu_1 = \frac{\mu_0 \cdot \frac{1}{\tau_0^2} + \bar{X} \cdot \frac{n}{\sigma^2}}{\frac{1}{\tau_0^2} + \frac{n}{\sigma^2}}$$

Posterior Predictive: We can compute

$$P(\bar{X} | X) = \int P(\bar{X} | \mu) P(\mu | X) d\mu \text{ but there is an easier way}$$

$$\bar{X} | \mu \sim N(\mu, \sigma^2); \bar{X} = \mu + e, e \sim N(0, \sigma^2)$$

$$\text{On the other hand, } \mu | X \sim N(\mu_1, \tau_1^2) \mu = \mu_1 + f, f \sim N(0, \tau_1^2)$$

with $e \perp f$.

$$\bar{X} = \mu_1 + e + f \sim N(\mu_1, \sigma^2 + \tau_1^2) \text{ So } N(\mu_1, \sigma^2 + \tau_1^2)$$

$$(ic) X_i | \sigma^2 \stackrel{\text{known}}{\sim} N(\mu, \sigma^2)$$

$$P(X_i | \sigma^2) \propto \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2} (x_i - \mu)^2\right)$$

$$\begin{aligned} P(X_1, \dots, X_n | \sigma^2) &\propto \prod_{i=1}^n (\sigma^2)^{-1/2} \exp\left(-\frac{1}{2\sigma^2} (x_i - \mu)^2\right) \\ &\propto (\sigma^2)^{\underbrace{-n/2}_{\text{data}}} \exp\left(-\frac{1}{2\sigma^2} \underbrace{\sum_{i=1}^n (x_i - \mu)^2}_{S(x) \text{ data}}\right) \end{aligned}$$

$$P(\sigma^2) \propto (\sigma^2)^{-r/2} \exp\left(-\frac{1}{2\sigma^2} \tau^2\right)$$

$$P(\sigma^2) \propto (\sigma^2)^{-\left(\frac{v_0}{2}+1\right)} \exp\left(-\frac{r_0 z_0^2}{2\sigma^2}\right) \Rightarrow \sigma^2 \sim \text{Inv } \chi^2(r_0, z_0^2)$$

$$P(\sigma^2 | x) \propto P(x | \sigma^2) P(\sigma^2)$$

$$\propto (\sigma^2)^{-n/2} e^{-S(x)/2\sigma^2} \times (\sigma^2)^{-\left(\frac{v_0}{2}+1\right)} \exp\left(-\frac{r_0 z_0^2}{2\sigma^2}\right)$$

$$\propto (\sigma^2)^{-\left(\frac{n+r_0}{2}+1\right)} \exp\left(-\frac{1}{2\sigma^2} (S(x) + r_0 z_0^2)\right)$$

$$\text{So } \sigma^2 | x \sim \text{Inv } \chi^2\left(n+r_0, \frac{S(x) + r_0 z_0^2}{n+r_0}\right)$$

See text for post. predictive.

$$(iii) Y_i | \theta \stackrel{\text{ind}}{\sim} P_\theta(x_i; \theta), x_i \text{ are exposures ex: } \theta = \frac{\# \text{ events}}{\text{hour}} \quad x_i = \text{hours}$$

$$P(y_1, \dots, y_n | \theta) = \prod_i \frac{(x_i \theta)^{y_i} e^{-x_i \theta}}{y_i!} \propto \theta^{\sum_i y_i} e^{-\theta \sum_i x_i}$$

data data

$$P(\theta) \propto \theta^{\alpha-1} e^{-\beta\theta} \Rightarrow \theta \sim \text{Gamma}(\alpha, \beta)$$

$$P(\theta | y) \propto \theta^{\alpha+n\bar{y}-1} e^{-(\beta+n\bar{x})\theta} \Rightarrow \theta | y \sim \text{Gamma}\left(\underbrace{\alpha+n\bar{y}}_{\alpha_1}, \underbrace{\beta+n\bar{x}}_{\beta_1}\right)$$

$$\text{Post Pred: } \tilde{y} | \theta \sim P_\theta(\tilde{x}; \theta)$$

$$P(\tilde{y} | y) = \int P(\tilde{y} | \theta) P(\theta | y) d\theta = \int \frac{(\tilde{x}\theta)^{\tilde{y}} \exp(-\tilde{x}\theta)}{\tilde{y}!} \cdot \frac{\beta_1^{\alpha_1}}{\Gamma(\alpha_1)} \theta^{\alpha_1-1} e^{-\beta_1\theta} d\theta$$

$$= \frac{\tilde{x}^{\tilde{y}} \beta_1^{\alpha_1}}{\tilde{y}! \Gamma(\alpha_1)} \int \theta^{\tilde{y}+\alpha_1-1} e^{-(\tilde{x}+\beta_1)\theta} d\theta$$

$$= \frac{\bar{x}^{\tilde{y}} \beta^{\alpha_1}}{\tilde{y}! \Gamma(\alpha_1)} \cdot \frac{\Gamma(\tilde{y} + \alpha_1)}{(\bar{x} + \beta_1)^{\tilde{y} + \alpha_1}} = \frac{\Gamma(\tilde{y} + \alpha_1)}{\tilde{y}! \Gamma(\alpha_1)} \cdot \left(\frac{\bar{x}}{\bar{x} + \beta_1} \right)^{\tilde{y}} \left(\frac{\beta_1}{\bar{x} + \beta_1} \right)^{\alpha_1}$$

$$\tilde{y}|y \sim \text{Neg Bin} \left(\alpha_1, \frac{\bar{x}}{\bar{x} + \beta_1} \right)$$