Noninformative Priors

Want data to be emphasized.

Idea: use invariance

Ex: Jeffrey's Prior is invariant to reparameterization

$\mathbb{P}(\theta) \propto I(\theta)^{1/2}$ $\quad I(\theta) = $ fisher information

$$= \mathbb{E}_{x|\theta}\left[\left(\frac{d\ell}{d\theta}\right)^2\right] \qquad \ell(\theta) = \log \mathbb{P}(x|\theta)$$

$$\doteq -\mathbb{E}_{x|\theta}\left[\frac{d^2\ell}{d\theta^2}\right]$$ (under regularity conditions)

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

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

$$= I(\lambda)\left(\frac{d\lambda}{d\theta}\right)^2$$

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

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

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

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

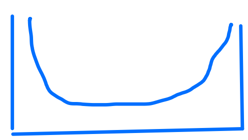
$$\frac{d\ell}{d\theta} = \frac{x}{\theta} - \frac{n-x}{1-\theta} \implies \frac{d\ell^2}{d\theta^2} = \frac{-x}{\theta^2} - \frac{n-x}{(1-\theta)^2}$$

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

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

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

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



arcsin distribution

Examples from Exp. Families

(i)

$$X_i \mid \mu \overset{iid}{\sim} N(\mu, \sigma^2)$$

σ^2 known

$$\mathbb{P}(x_1,\dots,x_n|\mu) \propto \prod_{i=1}^{n} \exp\left\{-\frac{1}{2\sigma^2}(x_i-\mu)^2\right\} = \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\left((x_i-\bar{x})+(\bar{x}-\mu)\right)^2\right)$$

But $\sum_i (x_i-\mu)^2 = \sum(x_i-\bar{x})^2 + 2\left(\cancelto{0}{\sum_i x_i - \bar{x}}\right)(\bar{x}-\mu) + n(\bar{x}-\mu)^2$

$$= \sum_i (x_i-\bar{x})^2 + n(\bar{x}-\mu)^2$$

So

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

data

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

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

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

Posterior Predictive: We can compute

$\mathbb{P}(\tilde{x}|x) = \int \mathbb{P}(\tilde{x}|\mu)\,\mathbb{P}(\mu|x)\,d\mu$ but there is an easier way

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

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 $e \perp\!\!\!\perp f$.

$\tilde{x} = \mu_1 + e + f \sim N(\mu, \sigma^2 + \tau_1^2)$ So $N(\mu_1, \sigma_1^2 + \tau_1^2)$

(ii) $X_i | \sigma^2 \overset{iid}{\sim} N(\underbrace{\mu}_{\text{known}}, \sigma^2)$

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

$$\mathbb{P}(x_1,..,x_n|\sigma^2) \propto \prod_{i=1}^{n} (\sigma^2)^{-1/2} \exp\left(\frac{1}{2\sigma^2}(x_i-\mu)^2\right)$$

$$\propto (\sigma^2)^{-\overset{\text{data}}{n}/2} \exp\left(\frac{1}{2\sigma^2} \underbrace{\sum_1 (x_i-\mu)^2}_{S(x) \text{ data}}\right)$$

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

$$\mathbb{P}(\sigma^2) \propto (\sigma^2)^{-\left(\frac{\nu_0}{2}+1\right)} \exp\left(\frac{-\nu_0\tau_0^2}{2\sigma^2}\right) \implies \sigma^2 \sim \text{Inv}\,\chi^2(\nu_0, \tau_0^2)$$

$$\mathbb{P}(\sigma^2|X) \propto \mathbb{P}(X|\sigma^2)\mathbb{P}(\sigma^2)$$

$$\propto (\sigma^2)^{-n/2} e^{-S(x)/2\sigma^2} \times (\sigma^2)^{-\left(\frac{\nu_0}{2}+1\right)} \exp\left(\frac{-\nu_0\tau_0^2}{2\sigma^2}\right)$$

$$\propto (\sigma^2)^{-\left(\frac{n+\nu_0}{2}+1\right)} \exp\left(\frac{-1}{2\sigma^2}\left(S(x)+\nu_0\tau_0^2\right)\right)$$

So $\sigma^2|X \sim \text{Inv}\,\chi^2\left(n+\nu_0, \frac{S(x)+\nu_0\tau_0^2}{n+\nu_0}\right)$

See text for post. predictive.

(iii) $Y_i|\theta \overset{ind}{\sim} Po(x_i\theta)$, x_i are exposures <u>ex</u>: $\theta = \frac{\#\text{ events}}{\text{hour}}$ $x_i = \text{hours}$

$$\mathbb{P}(Y_1,\dots,Y_n|\theta) = \prod_i \frac{(x_i\theta)^{y_i} e^{-x_i\theta}}{y_i!} \propto \theta^{\overset{\text{data}}{\sum y_i}} e^{-\theta\underbrace{\sum_i x_i}_{\text{data}}}$$

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

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

Post Pred: $\tilde{y}|\theta \sim Po(\tilde{x}\theta)$

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

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

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

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