Likelihood
Prior
Posterior } Normal – normal

3/28/17

$X_1, \dots, X_n | \theta, \sigma^2 \overset{iid}{\sim} N(\theta, \sigma^2)$

prior: $\theta | \sigma^2 \sim N(\mu_0, \tau^2)$

$$\theta | X_1, \dots, X_n, \sigma^2 \sim N\left(\underbrace{\frac{\frac{\bar{X}_n}{\sigma^2} + \frac{\mu_0}{\tau^2}}{\frac{n}{\sigma^2} + \frac{1}{\tau^2}}}_{\theta_p}, \underbrace{\frac{1}{\frac{n}{\sigma^2} + \frac{1}{\tau^2}}}_{\sigma_p^2} \right)$$

Normal – normal
conjugate model

$$\hat{\theta}_{MMSE} = E(\theta|X) = \hat{\theta}_{MAP} = \hat{\theta}_{MAE}$$

$$= \frac{\frac{\bar{X}_n}{\sigma^2} + \frac{\mu_0}{\tau^2}}{\frac{n}{\sigma^2} + \frac{1}{\tau^2}} = \underbrace{\frac{\sigma^2}{n\tau^2 + \sigma^2}}_{p} \underbrace{\mu_0}_{E(\theta)} + \underbrace{\frac{n\tau^2}{n\tau^2 + \sigma^2}}_{1-p} \underbrace{\bar{X}}_{\hat{\theta}_{MLE}}$$

μ_0: prior mean
τ^2: prior variance

pseudocount interpretation?

Imagine you see n_0 previous trials $Y_1, \dots, Y_n$

$$\mu_0 = \bar{Y} = \frac{1}{n_0} \sum_{i=1}^{n_0} y_i$$

Let $\tau^2 = \frac{\sigma^2}{n_0}$ since σ^2 known

$X_1, \dots, X_n | \theta, \sigma^2 \overset{iid}{\sim} N(\theta, \sigma^2)$

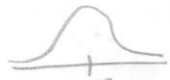
$\theta | \sigma^2 \sim N(\mu_0, \frac{\sigma^2}{n_0})$

$$\Rightarrow \theta | x_1, \dots, x_n, \sigma^2 \sim N\left(\frac{\bar{X}_n + \mu_0 n_0}{n + n_0}, \left(\frac{\sigma}{\sqrt{n + n_0}} \right)^2 \right)$$

$$\Rightarrow \theta_p = \frac{\frac{\bar{X}_n}{\sigma^2} + \frac{\mu_0}{\tau^2}}{\frac{n}{\sigma^2} + \frac{1}{\tau^2}} = \frac{\frac{\bar{X}_n}{\sigma^2} + \frac{\bar{Y} n_0}{\sigma^2}}{\frac{n}{\sigma^2} + \frac{n_0}{\sigma^2}} = \frac{\bar{X}_n + \bar{Y} n_0}{n + n_0}$$

For Poisson, gamma

Laplace: Equal θs
Haldane: ignorance
Jeffreys: Uninformative



not indiffent

$$= \frac{\sum_{i=1}^{n} X_i + \sum_{i=1}^{n_0} Y_i}{n + n_0}$$

Master average of prior data and observational data

$$\sigma^2_p = \frac{1}{\frac{n+n_0}{\sigma^2}} = \frac{\sigma^2}{n+n_0}$$

Laplace Prior for $\theta|\sigma^2$

$$P(\theta|\sigma^2) \propto 1$$

Improper

$$P(\theta|X,\sigma^2) \propto P(X|\theta,\sigma^2)\overbrace{P(\theta|\sigma^2)}^{\propto 1} \propto P(X|\theta,\sigma^2)$$

$$\propto e^{\frac{\bar{X}_n}{\sigma^2}\theta} e^{-\frac{n}{2\sigma^2}\theta^2} \propto N\left(\bar{X}, \left(\frac{\sigma}{\sqrt{n}}\right)^2\right)$$ always proper

$$N(c, v^2) \propto e^{-\frac{1}{2v^2}(\theta-c)^2}$$

$$= e^{-\frac{1}{2v^2}(\theta^2 - 2c\theta + c^2)}$$

$$= e^{-\frac{1}{2v^2}\theta^2} e^{\frac{c}{v^2}\theta} e^{-\frac{c^2}{2v^2}}$$

$$\propto e^{-\frac{1}{2v^2}\theta^2} e^{\frac{c}{v^2}\theta}$$

$$\frac{c}{v^2} = \frac{\bar{X}_n}{\sigma^2} \Rightarrow$$

$$c = \frac{\bar{X}_n}{\sigma^2} v^2 = \frac{\bar{X}_n}{\sigma^2} \cdot \frac{\sigma^2}{n} = \bar{X}$$

$$-\frac{1}{2v^2} \Rightarrow -\frac{1}{2v^2} = \frac{-n}{2\sigma^2}$$

$$\Rightarrow v^2 = \frac{\sigma^2}{n}$$

$$P(\theta|\sigma^2) \propto \sqrt{I(\theta)}$$

Jeffrey's method

$$I(\theta) = E\left(-\ell''(\theta; x, \sigma^2)\right)$$

$$\ell'(\theta; x, \sigma^2) = \frac{\bar{X}_n}{\sigma^2} - \frac{n\theta}{\sigma^2}$$

$$\ell''(\theta; x, \sigma^2) = \frac{-n}{\sigma^2}$$

$$I(\theta) = E\left(\frac{n}{\sigma^2}\right) = \frac{n}{\sigma^2} \Rightarrow P(\theta|\sigma^2) \propto \sqrt{\frac{n}{\sigma^2}} \propto 1$$

Bayesian point estimate $\Rightarrow$ $\hat{\theta}_{MMSE} = \bar{X} = \hat{\theta}_{MAP} = \hat{\theta}_{MME} = \hat{\theta}_{MLE}$

$X|\theta \sim Bin(n, \theta)$

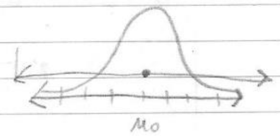
Any prior arbitrarily close to Haldane is proper

Convolution

$\theta \sim Beta(\alpha, \beta)$

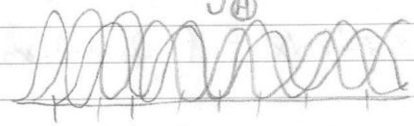
$\theta | X \sim Beta(\alpha + x, \beta + n - x)$

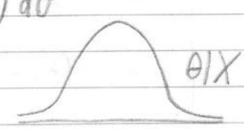
$\lim_{\substack{\alpha \to 0 \\ \beta \to 0}} P(\theta | x) = Beta(x, n-x)$



$\lim_{\tau^2 \to \infty} N(\mu_0, \tau^2) \propto 1$ $\qquad \bar{X}, \frac{\sigma^2}{n}$

$$P(X^* | X) = \int_{\Theta} P(X^* | \theta) P(\theta | X) \, d\theta$$



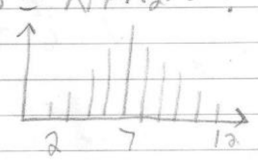


for $n^* = 1$

$$P(X^* | X, \sigma^2) = \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x^* - \theta)^2} \frac{1}{\sqrt{2\pi\sigma_p^2}} e^{-\frac{1}{2\sigma_p^2}(\theta - \theta_p)^2} d\theta$$

$X_1, X_2 \overset{iid}{\sim} U(\{1,2,3,4,5,6\})$

$S = X_1 + X_2 \sim ?$



$P(S=1) = 0 \qquad P(S=2) = P(X_1=1)P(X_2=1) = \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36}$

$P(S=3) = P(X_1=1)P(X_2=2) + P(X_1=2)P(X_2=1) = \frac{1}{36} + \frac{1}{36} = \frac{2}{36}$

$= \sum_{x \in supp[X_1]} P(X_1 = x) P(X_2 = 3 - x)$

$P(S=s) = \sum_{x \in supp(X_1)} P(X_1 = x) P(X_2 = s - x) = \sum_{x \in supp(X_2)} P(X_2 = x) P(X_1 = s - x)$

$X_1 \sim N(\mu_1, \sigma_1^2)$, $X_2 \sim N(\mu_2, \sigma_2^2)$ independently

$X_1 \sim f_{X_1}$, $X_2 \sim f_{X_2}$ indep. cont. R.V.-s

$$S = X_1 + X_2 \sim f_S(s) = \underset{\substack{\uparrow \\ \text{convolution operator}}}{f_{X_1} * f_{X_2}} = \int_{\text{Supp}(X_1)} f_{X_1}(x)\, f_{X_2}(s-x)\, dx$$

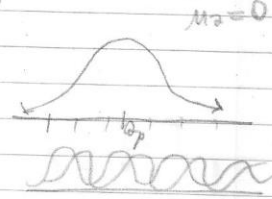
$$S = X_1 + X_2 \sim N(\mu_1 + \mu_2, \underbrace{\sigma_1^2 + \sigma_2^2}_{= \left(\sqrt{\sigma_1^2+\sigma_2^2}\right)^2})$$

$$\sim f_S(s) = f_{X_1}(x) * f_{X_2}(x) = \int_{\mathbb{R}} f_{X_1}(x)\, f_{X_2}(s-x)\, dx$$

$$= \int \frac{1}{\sqrt{2\pi\sigma_1^2}} e^{-\frac{1}{2\sigma_1^2}(x-\mu_1)^2} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(s-x-\mu_2)^2}\, dx$$

$$\underbrace{p(X^* \mid X, \sigma^2)}_{S = X^*} = \int_{\mathbb{R}} \underbrace{\frac{1}{\sqrt{2\pi\sigma_p^2}} e^{-\frac{1}{2\sigma_p^2}(\theta-\theta_p)^2}}_{\substack{\sigma_1^2 = \sigma_p^2 \\ \mu_1 = \theta_p}} \underbrace{\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(X^*-\theta-0)^2}}_{\substack{s = x^* \\ \sigma_2^2 = \sigma^2 \\ \mu_2 = 0}}\, d\theta$$

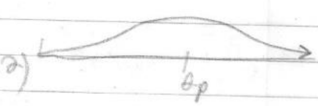
$$= N(\theta_p, \sigma_p^2 + \sigma^2)$$



If Jeffrey's prior,

what is post. pred. dist.?

$$p(X^* \mid X, \sigma^2) = N(\theta_p, \sigma_p^2 + \sigma^2)$$



$$= N\left(\bar{X}, \frac{\sigma^2}{n} + \sigma^2\right)$$

as $n \to \infty \Rightarrow \theta$ $\qquad = \sigma^2\left(\frac{1}{n}+1\right) = \sigma^2\left(\frac{n+1}{n}\right) = \left(\sigma\sqrt{\frac{n+1}{n}}\right)^2$

degenerates as $n \to \infty$

$$P(X^* | X) = \int_{\Theta} P(X^* | \theta) P(\theta | x) d\theta = P(X^* | \theta, \sigma^2)$$

Normal-normal w/ σ^2 known

$X_1, \dots, X_n | \theta, \sigma^2 \overset{iid}{\sim} N(\theta, \sigma^2)$

θ is known, σ^2 unknown

Let's find MLE for σ^2

$$\mathcal{L}(\sigma^2; X, \theta) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x_i - \theta)^2}$$

$$= \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \theta)^2}$$

$$\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n = \left(\frac{1}{\sqrt{2\pi}}\right)^n (\sigma^2)^{-\frac{n}{2}}$$

$$\ell(\sigma^2; x, \theta) = n \ln\left(\frac{1}{\sqrt{2\pi}}\right) - \frac{n}{2}\ln(\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \theta)^2$$

$$\ell'(\sigma^2; x, \theta) = \frac{-n}{2\sigma^2} + \frac{1}{2(\sigma^2)^2}\sum(x_i - \theta)^2 \overset{set}{=} 0 \; (2)(\sigma^2)$$

$$\Rightarrow -n + \frac{1}{\sigma^2}\sum(x_i - \theta)^2 = 0 \Rightarrow \hat{\sigma}^2_{MLE} = \underbrace{\frac{1}{n}\sum_{i=1}^{n}(x_i - \theta)^2}_{\text{sum of squared errors (SSE)}} = \frac{SSE}{n}$$

Frequentist point estimate

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

$Y = f(\theta) = \frac{1}{\theta} \sim ?$

$\theta = f^{-1}(y) = \frac{1}{y}$

$$f_Y(y) = f_\theta(f^{-1}(y)) \left| \frac{d}{dy}[f^{-1}(y)] \right|$$

$$\underbrace{y^{-\alpha+1}} \qquad |-y^{-2}| = y^{-2}$$

$$= \frac{\beta^\alpha}{\Gamma(\alpha)} e^{-\frac{\beta}{y}} \left(\frac{1}{y}\right)^{\alpha-1} \left|\frac{d}{dy}\left[\frac{1}{y}\right]\right|$$

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

$Y \sim \text{InvGamma}(\alpha, \beta)$

Formula $E(Y) = \frac{\beta}{\alpha-1}$ if $\alpha > 1$

$Med(Y) =$ qinvgamma $(0.5, \alpha, \beta)$

$Mode(Y) = \frac{\beta}{\alpha+1}$

Param. Space $\alpha, \beta > 0$

$Supp(Y) = (0, \infty)$

$$P(\sigma^2 | X, \theta) \propto P(X | \theta, \sigma^2) P(\sigma^2 | \theta)$$

$$= \left(\prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x_i-\theta)^2}\right) P(\sigma^2 | \theta)$$

$$= \left(\frac{1}{\sqrt{2\pi}}\right)^n (\sigma^2)^{-\frac{n}{2}} e^{-\frac{1}{2\sigma^2} \underbrace{\sum_{i=1}^{n}(x_i-\theta)^2}_{n\hat{\sigma}^2_{MLE}}} P(\sigma^2 | \theta)$$

$$\propto \underbrace{(\sigma^2)^{-\frac{n}{2}} e^{-\frac{n\hat{\sigma}^2_{MLE}}{2\sigma^2}}} \quad P(\sigma^2 | \theta)$$

kernel for $\text{InvGamma}\left(\frac{n}{2}-1, \frac{n\hat{\sigma}^2_{MLE}}{2}\right)$

$-\frac{n}{2} = -\alpha - 1$

$1 - \frac{n}{2} = -\alpha$

$\alpha = \frac{n}{2} - 1$