3/28/17

$$X_1, \dots, X_n \mid \theta, \sigma^2 \overset{iid}{\sim} N(\theta, \sigma^2)$$
$$\theta \mid \sigma^2 \sim N(\mu_0, \tau^2)$$
$$\theta \mid X_1, \dots, X_n, \sigma^2 \sim N\Bigg(\underbrace{\frac{\frac{\bar{x}n}{\sigma^2} + \frac{\mu_0}{\tau^2}}{\frac{n}{\sigma^2} + \frac{1}{\tau^2}}}_{\theta_p}, \underbrace{\frac{1}{\frac{n}{\sigma^2} + \frac{1}{\tau^2}}}_{\sigma^2_p}\Bigg)$$

What makes this conjugate normal?
- prior is N
- posterior is N
- likelihood is N
- posterior predictive is N

$$\hat{\theta}_{MMSE} = \hat{\theta}_{MAP} = \hat{\theta}_{MAE} = \frac{\frac{\bar{x}n}{\sigma^2} + \frac{\mu_0}{\tau^2}}{\frac{n}{\sigma^2} + \frac{1}{\tau^2}} = \underbrace{\frac{\sigma^2}{n\tau^2 + \sigma^2}}_{\rho} \underbrace{\mu_0}_{E[\theta]} + \underbrace{\frac{n\tau^2}{n\tau^2 + \sigma^2}}_{1-\rho} \cdot \underbrace{\bar{x}}_{\hat{\theta}_{MLE}}$$

μ_0: prior mean
τ^2: prior variance

pseudocount interpretation?

$$X_1, \dots, X_n \mid \theta, \sigma^2 \overset{iid}{\sim} N(\theta, \sigma^2)$$
$$\theta \mid \sigma^2 \sim N\left(\mu_0, \frac{\sigma^2}{n_0}\right)$$

Imagine you see n_0 previous trials: $\mu_0 = \bar{y} = \frac{1}{n_0} \sum_{i=1}^{n_0} y_i$

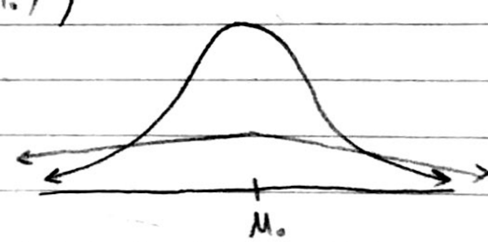
let $\tau^2 = \frac{\sigma^2}{n_0}$ since σ^2 known

$$\Rightarrow \theta_p = \frac{\frac{\bar{x}n}{\sigma^2} + \frac{\mu_0}{\tau^2}}{\frac{n}{\sigma^2} + \frac{1}{\tau^2}} = \frac{\frac{\bar{x}n}{\sigma^2} + \frac{\bar{y}n_0}{\sigma^2}}{\frac{n}{\sigma^2} + \frac{n_0}{\sigma^2}} = \frac{\bar{x}n + \bar{y}n_0}{n + n_0} = \frac{\sum_{i=1}^{n} x_i + \sum_{i=1}^{n_0} y_i}{n + n_0}$$

↳ master average of prior data and observed data

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

$$\Rightarrow \theta \mid X_1, \dots, X_n, \sigma^2 \sim N\left(\frac{\bar{x}n + \mu_0 n_0}{n + n_0}, \left(\frac{\sigma}{\sqrt{n + n_0}}\right)^2\right)$$



<u>Laplace</u> prior for $\theta \mid \sigma^2$

$$P(\theta \mid \sigma^2) \propto 1 \quad \Big\} \quad \lim_{\tau^2 \to \infty} N(\mu_0, \tau^2) \propto 1$$

improper

$$P(\theta \mid X, \sigma^2) \propto P(X \mid \theta, \sigma^2) \overbrace{P(\theta \mid \sigma^2)}^{\propto 1} \propto P(X \mid \theta, \sigma^2) \propto e^{\frac{\bar{x}n}{\sigma^2}\theta} e^{-\frac{n}{2\sigma^2}\theta^2}$$

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

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

$$\propto N\left(\bar{x}, \left(\frac{\sigma}{\sqrt{n}}\right)^2\right) \Rightarrow \text{always proper}$$

Jeffren's

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

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

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

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

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

$\hat{\theta}_{mmsE} = \hat{\theta}_{MAP} = \hat{\theta}_{MAE} = \bar{x} = \hat{\theta}_{MLE}$

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

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

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

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

↳ Improper priors can be thought of as the limit of proper priors.

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

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

$\theta | X_1, \ldots, X_n \sim N\left(\frac{\frac{\bar{x}n}{\sigma^2} + \frac{\mu_0}{\tau^2}}{\frac{n}{\sigma^2} + \frac{1}{\tau^2}}, \frac{1}{\frac{n}{\sigma^2} + \frac{1}{\tau^2}}\right) = N\left(\hat{\theta}_{mmsE}, \sigma_p^2\right)$

$P(x^*|x) = \int_\Theta P(x^*|\theta)P(\theta|x)\,d\theta = P(x^*|\theta, \sigma^2) \Rightarrow$ same thing as drawing ideas of what θ is

for $n^* = 1$

$$\underbrace{P(x^*|x, \sigma^2)}_{S = x^*} = \int_R \underbrace{\frac{1}{\sqrt{2\pi\sigma_p^2}} e^{-\frac{1}{2\sigma_p^2}(\theta - \theta_p)^2}}_{\substack{\sigma_1^2 = \sigma_p^2 \\ \mu_1 = \theta_p}} \underbrace{\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x^* - \theta - \theta)^2}}_{S = x^* \quad \sigma_2^2 = \sigma^2 \quad \mu_1 = 0} d\theta = N(\theta_p, \sigma_p^2 + \sigma^2)$$

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

$S = X_1 + X_2 \sim ?$

$P(S=1) = 0$

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

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

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

2>

3/28/17 ②

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

$X_1 \sim f_{X_1}, X_2 \sim f_{X_2}$ ⇒ independent continuous r.v.

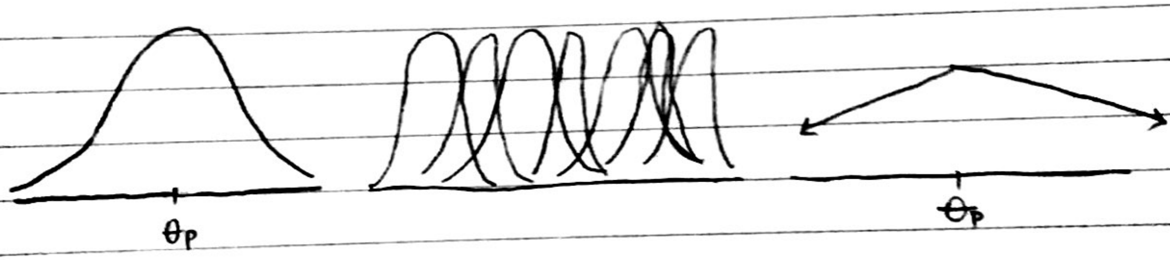
$S = X_1 + X_2 \sim f_{X_1} * f_{X_2} := \int_{Supp[X_1]} f_{X_1}(x) f_{X_2}(s-x)\, dx$

↳ convolusion

$X_1 \sim N(\mu_1, \sigma^2_1), X_2 \sim N(\mu_2, \sigma^2_2)$ ⇒ independent

$S = X_1 + X_2 \sim N(\mu_1 + \mu_2, \underbrace{\sigma^2_1 + \sigma^2_2}_{= (\sqrt{\sigma^2_1 + \sigma^2_2})^2})$

$$\sim f_S(s) = f_{X_1}(x) * f_{X_2}(x) = \int_{\mathbb{R}} f_{X_1}(x) f_{X_2}(s-x)\, dx$$
$$= \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi\sigma^2_1}} e^{-\frac{1}{2\sigma^2_1}(x-\mu_1)^2} \frac{1}{\sqrt{2\pi\sigma^2_2}} e^{-\frac{1}{2\sigma^2_2}(s-x-\mu_2)^2}\, dx$$



If Jeffrey's prior, what is posterior predictive distribution?

$$P(x^* | x, \sigma^2) = N(\theta_p, \sigma^2_p + \sigma^2)$$
$$= N(\bar{x}, \underbrace{\frac{\sigma^2}{n} + \sigma^2}_{\sigma^2(\frac{1}{n}+1)})$$
$$= \sigma^2 \frac{n+1}{n}$$
$$= \left(\sigma \sqrt{\frac{n+1}{n}}\right)^2$$

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

θ is known, σ^2 is unknown.

Let's find MLE for σ^2

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

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

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

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

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

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

$$Y = \frac{1}{\theta} \sim ? = t(\theta) \Rightarrow \theta = t^{-1}(y) = \frac{1}{y}$$

$$f_Y(y) = f_\theta(t^{-1}(y))\left|\frac{d}{dy}[t^{-1}(y)]\right| = \frac{\beta^\alpha}{\Gamma(\alpha)} e^{-\frac{\beta}{y}} \overbrace{\left(\frac{1}{y}\right)^{\alpha-1}}^{y^{-\alpha+1}} \overbrace{\left|\frac{d}{dy}\left[\frac{1}{y}\right]\right|}^{y^{-2} = |-y^{-2}|}$$

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

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

$E[Y] = \frac{\beta}{\alpha-1}$ if $\alpha > 1$

$\text{med}[Y] = \text{qinvgamma}(0.5, \alpha, \beta)$

$\text{mode}[Y] = \frac{\beta}{\alpha+1}$

parameter space $\alpha, \beta > 0$

$\text{supp}[Y] = (0, \infty)$

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

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

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

$$\propto \underbrace{(\sigma^2)^{-\frac{n}{2}} e^{-\frac{n\hat{\sigma}^2_{MLE}}{2\sigma^2}}}_{\text{Kernel of InvGamma}\left(\frac{n}{2}-1,\ \frac{n\hat{\sigma}^2_{MLE}}{2}\right)} P(\sigma^2|\theta)$$