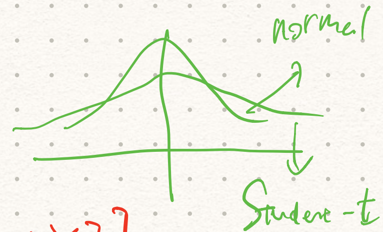
① Marginization $P(\mu|\text{observations}) = \int P(\mu, \sigma^2 | \text{obs.})\, d\sigma^2$

$\downarrow$ Student-t

$= \mathbb{E}_{P(\sigma^2|\text{obs})}[P(\mu | \sigma^2, \text{obs})]$

Gaussian distribution



$$\mathrm{Var}[X] = \mathrm{Var}[E[X|Y]] + E[\mathrm{Var}[X|Y]]$$

$$\geq \mathrm{Var}[E[X|Y]]$$

② Monte Carlo Approximation for posterior quantities

(mean, variance, intervals)

$$\mu^{(i)} \sim P(\mu|\text{obs}) \Rightarrow \mathbb{E}(\mu|\text{obs}) \approx \sum_{i=1}^{M} \mu^{(i)} / M$$

[unknown: μ, σ^2, observation: $y_1, \dots, y_n$ $\Rightarrow$

summarize all information about μ, σ^2 from

obs in $P(\mu, \sigma^2 | y_1, \dots, y_n)$ or $P(\mu | y_1, \dots, y_n)$

or $P(\sigma^2 | y_1, \dots, y_n)$

represent this distribution using random samples

$\{\mu^{(i)}, (\sigma^2)^{(i)}\}_{i=1,\dots,M}$

Model $\rightarrow P(y_{1,\dots,n} | \mu, \sigma^2)$

Pencil problem $y_i \overset{i.i.d.}{\sim} N(\mu, \sigma^2)$ with priors

$$\mu | \sigma^2 \sim N\left(\mu_0, \frac{\sigma^2}{K_0}\right), \quad \sigma^2 \sim \text{inverse-}\chi^2(\nu_0, \sigma_0^2)$$

① derive posterior distribution

$$p(\mu, \sigma^2 | y_1, \cdots, y_n), \quad p(\mu | y_1, y_2, \cdots, y_n) \sim t\text{-dis}$$

② derive a Gibbs sampler scheme

$$p(\mu | \sigma^2, y_1, y_2, \cdots, y_n)$$

$$p(\sigma^2 | \mu, y_1, y_2, \cdots, y_n)$$

Multinomial Model for Categorical Data ← generalization of Beta-Binomial model where there're only 2 categories $\{h, t\}$

e.g. election

Candidates: A, B, C, D, E

→ observation n_A n_B n_C n_d n_E (counts)

quantity of interest $P[\theta_A > \max\{\theta_B, \cdots, \theta_E \mid \text{observed counts}]$

(probability of A wins)

proportion of people ACTUALLY supports A

$$\boxed{\frac{n_A}{n_A+n_B+\cdots} \xrightarrow{\quad} \infty \longrightarrow \text{Consistent estimate}}$$

Let R be the ~~total~~ # classes.

$y_1, \dots, y_R$ be the observed counts $\sum_{j=1}^{R} y_j = n$

$\theta_1, \dots, \theta_R$ be the proportions $\sum_{j=1}^{R} \theta_j = 1$

model: $P(y|\theta) \propto \prod_{j=1}^{R} \theta_j^{y_j} \longleftarrow$ similar to $\theta^{\#head}(1-\theta)^{\#tail}$

prior: Dirichlet $p(\theta|\alpha) \propto \prod_{j=1}^{R} \theta_j^{\alpha_j - 1}$ (conjugate prior)

posterior

$P(\theta_1, \dots, \theta_R | \text{obs})$ is Dirichlet $(\alpha_j + y_j, \; j=1 \dots R)$

(X_i, n_i, y_i)

$y_i \longrightarrow$ positive response

$y_i | \theta_i \sim \text{Bin}(n_i, \theta_i)$

$$\text{logit}(\theta_i) = \alpha + \beta x_i$$

so likelihood

$$p(y_i|\alpha, \beta, n_i, x_i) \propto [\text{logit} \quad]^{y_i} [1 - \text{logit} \quad]^{n_i - y_i}$$

prior $\quad p(\alpha, \beta) \propto 1$

joint posterior

LD50

$$\log\left(\frac{\theta_i}{1-\theta_i}\right) = \alpha + \beta x_i \longrightarrow \text{plug in } \theta_i = 0.5. \quad x_i = -\frac{\alpha}{\beta}$$

s.t. $\theta_i = 0.5$

Multivariate normal:

$$\mu \sim N(\mu_0, \Lambda_0)$$