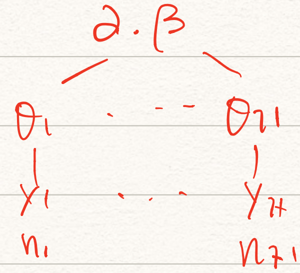
Hierarchical model

$\theta_j \sim \text{Beta}(\alpha, \beta)$

$y_j \sim \text{Binomial}(n_j, \theta_j)$



Today: Gaussian Hierarchical Model

consider J experiments with params $(\theta_1, \cdots, \theta_J)$

assume that $y_{ij} \sim N(\theta_j, \sigma^2)$, $i = 1, 2, \cdots, n_j$

Choice of discrete / continuous model

count: Binomial, Poisson...

score / height / income level: Gaussian...

→ For scores it has range $[a, b]$ (boundary)

consider the truncation. consider whether truncation will affect

$$\bar{y}_{\cdot j} = \sum_{i=1}^{n_j} y_{ij} / n_j \sim N(\theta_j, \sigma_j^2 = \sigma^2 / n_j),$$

(group mean)

we assume σ^2, n_j are known

plug in

$$\bar{y}_{\cdot\cdot} = \sum_{j=1}^{J} \bar{y}_{\cdot j} / \sigma_j^2$$

$$\frac{\cdots}{\sum_{j=1}^{J} \frac{1}{\sigma_j^2}} = \sum_{i,j} y_{ij} / \sum_{j=1} n_j \quad \text{(pooled)}$$

(overall mean)

Assume $n_1 = n_2 = \cdots = n_J$ ANOVA:

F-test: $\theta_1 = \theta_2 = \cdots = \theta_J$

(P.P. construct ANOVA table

→ calculate F-statistics → Hypothesis Testing

objective: have a model that's between "complete pooling" ($\bar{y}_{..}$) v.s. "no pooling at all" ($\bar{y}_{\cdot j}$)

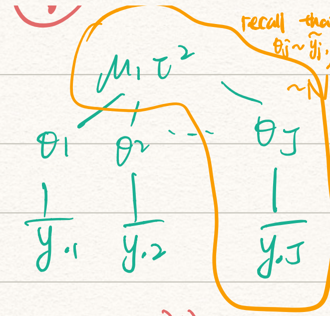
$$\hat{\theta}_j = \lambda_j \underline{\bar{y}_{\cdot j}} + (1-\lambda_j) \underline{\bar{y}_{..}}, \quad \lambda_j \in (0,1)$$

$$\begin{cases} \theta_j \sim N(\underline{\mu, \tau^2}) \\ \bar{y}_{\cdot j} \sim N(\theta_j, \sigma_j^2) \end{cases}$$

Now, set up priors for (μ, τ^2), $p(\mu, \tau^2) = p(\mu) p(\tau^2)$

$\propto p(\tau^2)$? $\left[(p(\mu) \propto 1) \text{ flat prior} \right]$

Bayes Formula :



recall that $\theta_j \sim \bar{y}_j, \mu, \tau^2$ $\sim N\left(\frac{\bar{y}_j/\sigma_j^2 + \mu/\tau^2}{1/\sigma_j^2 + 1/\tau^2}, \frac{1}{1/\sigma_j^2 + 1/\tau^2}\right)$ (Single Param)

$$p(\theta_{1:J}, \mu, \tau^2 \mid y_{1:J})$$

$$\propto p(y_{1:J} \mid \theta_{1:J}, \mu, \tau^2)\, p(\theta_{1:J}, \mu, \tau^2)$$

Conditional independence

$$= p(y_{1:J} \mid \theta_{1:J})\, p(\theta_{1:J} \mid \mu, \tau^2)\, p(\mu, \tau^2)$$

Factorization from C.I.

$$= \prod_{j=1}^{J} p(\bar{y}_{\cdot j} \mid \theta_j) \prod_{j=1}^{J} p(\theta_j \mid \mu, \tau^2) \quad p(\mu, \tau^2)$$

quanties of interest:

$\mathbb{E}[\theta_j \mid \bar{y}_{1:J}]$, $\mathrm{Var}[\theta_j \mid \bar{y}_{1:J}]$, quantiles of θ_j

posterior predictive quantities

Computation : Monte-Carlo Methods

obtain samples $\{\theta_j^{(i)}, \mu^{(i)}, \tau^{(i)}\}_{i=1 \cdots M}$

for $p(\mu, \tau^2, \theta_{1:J})$

$\propto p(\mu, \tau^2) \prod_{j=1}^{J} p(\bar{y}_j \mid \theta_j) \prod_{j=1}^{J} p(\theta_j \mid \mu, \tau^2)$

① $$p(\mu, \tau^2 \mid \bar{y}_{1:J}) = \int \underline{p(\mu, \tau^2, \theta_{1:J} \mid \bar{y}_{1:J})}\, d\theta_{1:J}$$

$$\propto \prod_{j=1}^{J} \left[\int p(\bar{y}_j \mid \theta_j)\, p(\theta_j \mid \mu, \tau^2)\, d\theta_j \right] p(\mu, \tau^2)$$

$$\propto \prod_{j=1} \left[\int \underbrace{p(\bar{y}_j \mid \theta_j)}_{e^{-\frac{(\bar{y}_j - \theta_j)^2}{2\sigma_j^2}}} \, \underbrace{p(\theta_j \mid \mu, \tau)}_{\frac{1}{\tau} e^{-\frac{(\theta_j - \mu)^2}{2\tau^2}}} \, d\theta_j \right] p(\mu, \tau)$$

$$\propto \prod_{j=1}^{J} \int \left(\frac{1}{\tau}\right) e^{-\frac{(\bar{y}_j - \theta_j)^2}{2\sigma_j^2} - \frac{(\theta_j - \mu)^2}{2\tau^2}} \, d\theta_j \; p(\mu,\tau^2)$$

Complete squares for θ_j

Use Gaussian Density

already derived

$$② \; p(\mu, \tau^2 \mid \bar{y}_{1:J}) = \frac{p(\mu, \tau^2, \theta_{1:J} \mid \bar{y}_{1:J})}{p(\theta_{1:J} \mid \bar{y}_{1:J}, \tau^2, \mu)}$$

?

$$= \prod_{j=1}^{J} p(\theta_j \mid \bar{y}_j, \mu, \tau^2)$$

(Gaussian - Gaussian model in single param)

if conjugate :

easy to simplify

if not, not easier than integration.

$X \sim N(1,1)$, $Y \sim N(2,1) \rightarrow X+Y \sim N$

$X \sim N(1,1)$, $Y \sim N(X,1) \longrightarrow Y$ normal.

$\downarrow$ $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $Y \sim N(1,2)$

($\underbrace{Y = X+Z}_{N(1,1)}$, $Z \perp\!\!\!\perp X$, $Z \sim N(0,1)$

③ Simpler: $\begin{cases} \theta_j \sim N(\mu, \tau^2) \\ \bar{y}_{\cdot j} \sim N(\theta_j, \sigma_j^2) \end{cases}$ $\boxed{\text{Use property of Gaussian}}$

$\Rightarrow \bar{y}_{\cdot j} \mid \mu, \tau^2 \sim N(\mu, \tau^2 + \sigma_j^2)$

$\underline{p(\mu,\tau^2)\, p(\bar{y}_{1:J} \mid \mu, \tau^2) \longrightarrow p(\mu, \tau^2 \mid \bar{y}_{1:J})}$

$\downarrow$

$\{\mu^{(i)}, \tau^{2(i)}\}$

$\downarrow$

$p(\theta_j \mid \mu^{(i)}, \tau^{2(i)}, \bar{y}_j)$

① ② ③ to get $p(\mu, \tau^2 \mid \bar{y}_{1:J})$

marginal prior.