$[a,b]$

$$P_i(\boldsymbol{x}_{\setminus i}, S_{\boldsymbol{x}_{\setminus i},i,a,b}) = \frac{\int_a^b f_k(x_1,\ldots,x_{i-1},z,x_{i+1},\ldots,x_n)\mathrm{d}z}{\int_{-\infty}^{\infty} f_k(x_1,\ldots,x_{i-1},z,x_{i+1},\ldots,x_n)\mathrm{d}z}$$

$$P_i(x_{\setminus i}, S_{x\setminus i}, \{y\})$$

$$\approx \frac{f_k(x_1,\ldots,x_{i-1},y,x_{i+1},\ldots,x_n)}{f(x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)}$$

$$= f_k(y \mid x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)$$

$$= f_k(y \mid x_{\setminus i})$$

$$\boxed{= f_k(y \mid x_{-i})}$$

$\pi(y)$

$f(y)$

$\pi(x_1,\ldots,x_{i-1},y_i,x_{i+1},\ldots,x_n)$

$f(x_i \mid y_i) = \pi(x_i) f(y_i \mid x_i)$

$$\ell(y_i) = \pi(x_1,\ldots,x_{i-1},y_i,x_{i+1},\ldots,x_n)$$

$$g(y_i \mid z_i) = f(y_i \mid X_{-i})$$

Modelo jerárquico simple basado en la distribución normal

Consideremos J experimentos independientes, con el experimento j estimando el parámetro θ_j de n_j muestras independientes distribuidas normal denotadas por y_{ij} con cada experimento con

varianza conocida σ^2

$$y_{ij} \mid \theta_j \sim N(\theta_j, \sigma^2) \qquad i = 1, \ldots, n_j \qquad j = 1, \ldots, J$$

$$\bar{y}_{\cdot j} = \frac{1}{n_j} \sum_{i=1}^{n_j} y_{ij}$$

$$\sigma_j^2 = \frac{\sigma^2}{n_j}$$

$$\bar{y}_{\cdot j} \mid \theta_j \sim N(\theta_j, \sigma_j^2)$$

$$\theta = (\theta_1, \ldots, \theta_J)$$

$$\bar{y}_{\cdot j} \rightsquigarrow \theta_j$$

$$\bar{y}_{\cdot\cdot} = \frac{\sum_{j=1}^{n} \frac{y_{\cdot j}}{\sigma_j^2}}{\sum_{j=1}^{n} \frac{1}{\sigma_j^2}} = \sum_{j=1}^{J} \sum_{i=1}^{n_j} \frac{y_{ij}}{n}$$

$\to A \qquad B$

	df	SS	MS	$E(MS \mid \sigma^2, \tau)$
Between groups	$J-1$	$\sum_i \sum_j (\bar{y}_{.j} - \bar{y}_{..})^2$	$SS/(J-1)$	$n\boxed{\tau^2} + \sigma^2$
Within groups	$J(n-1)$	$\sum_i \sum_j (y_{ij} - \bar{y}_{.j})^2$	$SS/(J(n-1))$	σ^2
Total	$Jn-1$	$\sum_i \sum_j (y_{ij} - \bar{y}_{..})^2$	$SS/(Jn-1)$	

$\to H_0 = \tau = 0$

$\to H_a = \tau \neq 0$

$\hat{\theta}_j = \bar{y}_{..}$

$\hat{\theta}_j = \bar{y}_{.j}$

$\frac{A}{B} > 1$

$$\rightarrow \hat{\theta}_j = \lambda_j \bar{y}_{\cdot j} + (1-\lambda_j)\bar{y}_{\cdot\cdot}$$

$$\lambda_j \in (0,1)$$

1) La estimación sin mezclar, $\hat{\theta}_j = \underline{\bar{y}_{\cdot j}}$, es la media a posterior si los σ valores θ_j tienen densidad a priori uniformes en $(-\infty, \infty)$

2) La estimación con mezcla $\hat{\theta}_j = \bar{y}_{\cdot\cdot}$ es la media a posterior si los $\overline{J}$ valores tienen la restricción de ser los mismos, con una distribución a priori unif. en el valor común θ.

3) La combinación con pesos es la media a posteriori si los $\bar{U}$ valores θ_j tiene función de densidad a priori (i.i.d. normales) normales independientes

Modelo jerárquico

$$p(\theta_1,\dots,\theta_J \mid \mu,\tau) = \prod_{j=1}^{J} N(\theta_j \mid \mu,\tau^2)$$

$$p(\theta_1,\dots,\theta_J) = \int \prod_{j=1}^{J} \left[N(\theta_j \mid \mu,\tau^2)\right] p(\mu,\tau)\, d(\mu,\tau)$$

$$p(\mu,\tau) = p(\mu \mid \tau)\, p(\tau) \propto 1$$

Distribución conjunta a posteriori

$\bar{y}_{\cdot j}$

$$p(\theta, \mu, \tau \mid y) \propto p(\mu, \tau)\, p(\theta \mid \mu, \tau)\, p(y \mid \theta)$$

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

$$p(\theta \mid \mu, \tau, y)$$

$$\theta_j \mid \mu, \tau, y \sim N(\hat{\theta}_j, V_j)$$

$$\hat{\theta}_j = \frac{\frac{1}{\sigma_j^2}\bar{y}_{\cdot j} + \frac{1}{\tau^2}\mu}{\frac{1}{\sigma_j^2} + \frac{1}{\tau^2}}$$

$$= \frac{\sigma_j^2 \tau^2 \left(\frac{1}{\sigma^2}\bar{y}_{\cdot j} + \frac{1}{\tau^2}\mu\right)}{\sigma_j^2 + \tau^2}$$

$$= \frac{1}{\sigma_j^2 + \tau^2}\left[\tau^2 \bar{y}_{\cdot j} + \sigma_j^2 \mu\right]$$

$$= \lambda_j \mu + (1-\lambda_j)\bar{y}_{\cdot j}$$

$$\lambda_j = \frac{\sigma_j^2}{\sigma_j^2 + \tau^2}$$

$$V_j = \frac{1}{\frac{1}{\sigma_j^2} + \frac{1}{\tau^2}}$$

Dificultades con una estimación no Bayesiana de los hiperpárametros

$$\hat{\mu} = \bar{y}_{\cdot\cdot}$$

$$\hat{\tau}^2 = (MS_B - MS_W)/n$$

$$<$$

School	Estimated treatment effect, y_j	Standard error of effect estimate, σ_j
A	28	15
B	8	10
C	−3	16
D	7	11
E	−1	9
F	1	11
G	18	10
H	12	18

Intento 1 (Estimaciones separadas)

-Difícil estadísticamente de distinguir con cualquiera de los experimentos (todos los intervalos se interpolan bastante)

Intento 2 (Pooled estimate)

Non-informative prior distribution

Media a Posteriori $\bar{y}_{..} = 7.7$

Varianza a posterior $\left(\sum_{j=1}^{8} \frac{1}{\sigma_j^2}\right)^{-1} = 16.6$

Error estándar $= \sqrt{16.6} = 4.1$

95% CI $= [-0.5, 15.9]$