Pontificia Universidad Católica de Chile
Facultad de Matematicas
Departamento De Estadística
Segundo Semestre 2020
Docente: Fernando Quintana

Interrogación 2 : Parte II

Nicolás Sumonte

Parte A

Notemos que:

$$p(\mu_1, \mu_2, \sigma^2|y_1, y_2) = \frac{p(y_1, y_2|\sigma^2, \mu_1, \mu_2)\ p(\mu_1, \mu_2, \sigma^2)}{p(y_1, y_2)}$$

con $y_1 = (y_{1,1}, ..., y_{1,n1})$ e $y_2 = (y_{2,1}, ..., y_{2,n1})$, y por axiomas de probabilidad sabemos que:

$$p(\mu_1, \mu_2, \sigma^2) = p(\mu_1, \mu_2|\sigma^2)p(\sigma^2) = p(\mu_1|\sigma^2)p(\mu_2|\sigma^2)p(\sigma^2)$$

Por lo que nos queda:

$$p(\mu_1, \mu_2, \sigma^2|y_1, y_2) = p(\ y_1, y_2|\mu_1, \mu_2, \sigma^2)p(\mu_1|\sigma^2)p(\mu_2|\sigma^2)p(\sigma^2)$$

Como ambas muestras son independientes entre sí, remplazamos y tenemos que:

$$p(\mu_1, \mu_2, \sigma^2|y_1, y_2) = \prod_{i=1}^{n_1} p(y_{1,i}|\mu_1, \sigma^2)\ \prod_{i=1}^{n_2} p(y_{2,i}|\mu_2, \sigma^2)\ p(\mu_1|\sigma^2)\ p(\mu_2|\sigma^2)\ p(\sigma^2)$$

Agrupando términos nos queda:

$$p(\mu_1, \mu_2, \sigma^2|y_1, y_2) = \prod_{i=1}^{n_1} p(y_{1,i}|\mu_1, \sigma^2)\ p(\mu_1|\sigma^2)\ \prod_{i=1}^{n_2} p(y_{2,i}|\mu_2, \sigma^2)\ p(\mu_2|\sigma^2)\ p(\sigma^2)$$

y se puede observar que:

$$\prod_{i=1}^{n_1} p(y_{1,i}|\mu_1, \sigma^2)\ p(\mu_1|\sigma^2)\ \&\ \prod_{i=1}^{n_2} p(y_{2,i}|\mu_2, \sigma^2)\ p(\mu_2|\sigma^2)$$

tienen la forma de posterioris de normales con σ^2 conocido y μ_k con $k \in \{1, 2\}$ como parámetro. Calcularé esta posteriori para μ_1 con perjuicio que para μ_2 será lo mismo, solo se debe cambiar el índice.

$$p(\mu_1|y_{1,1}, ..., y_{1,n_1}) = p(y_{1,1}, ..., y_{1,n_1|mu_1}, \sigma^2)\ p(\mu_1|\sigma^2)$$

donde

$$p(\mu_1|\sigma^2) \approx N(\mu_{1,0}, \sigma^2/k_{1,0})$$

Si consideramos $\tau_0^2 = \sigma^2/k_{1,0}$ ó $\sigma^2/k_{1,0}$ (dependiendo del caso) como la precisión a priori nos queda:

$$p(\mu_1|y_{1,1}, ..., y_{1,n_1}) \propto \{e^{\frac{-1}{2\tau_0^2}(\mu_1-\mu_{1,0})^2} \cdot e^{\frac{-1}{2\sigma^2}\sum_{i=1}^{n}(y_1-\mu_1)^2}\}$$

Si de momento ignoramos los $\frac{-1}{2}$ nos quedará:

$$\{e^{\frac{1}{\tau_0^2}(\mu_1^2-2\mu_1\mu_{1,0}+\mu_{1,0}^2)+\frac{1}{\sigma^2}(\sum_{i=1}^{n_1} y_1^2-\sum_{i=1}^{n_1} y_1\mu_1+n_1\cdot\mu_1^2)}\}$$

La cual tiene forma cuadrática: $a\mu_1^2 - 2b\mu_1 + c$, con c constante que no depende del parámetro y a y b:

$$a = \frac{1}{\tau_0} + \frac{n}{\tau_0^2} \ ; \ b = \frac{\mu_{1,0}}{\tau_0^2} + \frac{\sum_{i=1}^{n_1} y_i}{\sigma^2}$$

luego si completamos el cuadrado de:

$$p(\mu_1|\sigma^2, y_{1,1}, ..., y_{1,n_1}) \ e^{\frac{-1}{2}(a\mu_1^2-2b\mu_1)}$$

tendríamos:

$$e^{\frac{-1}{2}a(\mu_1^2-\frac{2b\mu_1}{a}+\frac{b^2}{a})+\frac{1}{2}\frac{b^2}{a}} \to e^{\frac{-1}{2}a(\mu_1-\frac{b}{a})^2+\frac{1}{2}\frac{b^2}{a}} \to e^{\frac{-1}{2}((\mu_1-\frac{b}{a})/\frac{1}{\sqrt{a}})^2+\frac{1}{2}\frac{b^2}{a}}$$

luego $p(\mu_1|\sigma^2, y_{1,1}, ..., y_{1,n_1})$ distribuye normal con media $\frac{b}{1}$ y varianza $\frac{1}{\sqrt{a}}$ lo que nos queda:

$$\sim N(\mu_{1,n_1}, \frac{\sigma^2}{k_1, n_1})$$

donde:

$$\mu_{1,n_1} = \frac{k_0}{k_0+n_1}mu_{1,0} + \frac{n_1}{k_0+n_1}\bar{y_1} = \frac{k_0\mu_{1,0}+\bar{y_1}n_1}{k_0+n_1}$$

y $k_{1,n1} = k_0 + n_1$ esto es debido a que k_0 representa el tamaño muestral de un conjunto de observaciones a priori y debemos sumarle las observaciones de la verosimilitud. Luego, haciendo lo propio para la μ_2 nos queda: $p(\mu_2|\sigma^2, y_{2,1}, ..., y_{2,n_2}) \sim N(\mu_{2,n_2}, \frac{\sigma^2}{k_{2,n_2}})$ donde:

$$\mu_{2,n_2} = \frac{k_0}{k_0+n_2}mu_{2,0} + \frac{n_2}{k_0+n_2}\bar{y_2} = \frac{k_0\mu_{2,0}+\bar{y_2}n_2}{k_0+n_2}$$

y $k_{2,n2} = k_0 + n2$.

Ahora la distribución a posteriori deσ^2 puede obtenerse integrando sobre los valores de μ_1 y μ_2. Esto es,

$$p(\sigma^2|y_1, y_2) \propto p(\sigma^2) \ p(y_1, y_2|\sigma^2)$$
$$= \sigma^2[\int p(y_{1,1}, ..., y_{1,n1}|\sigma^2, \mu_1)p(\mu_1, \sigma^2 du + \int p(y_{2,1}, ..., y_{2,n2}|\sigma^2, \mu_2)p(\mu_2, \sigma^2 du]$$

resolviendo esta integral para uno de los dos parametros (considerando que para el otro será lo mismo) nos queda

$$p(\sigma^2|y_{1,1},...,y_{1,n1}) \propto \chi^-2(\nu_0,\sigma_0^2)\int(2\pi\sigma^2)^{-n/2}\cdot e^{\frac{-1}{2}\sum_{i=1}^{n1}(y_i-\mu_1)^2}\cdot\frac{1}{\sqrt{2\pi\frac{\sigma^2}{k_0}}}\cdot e^{\frac{-1}{2}\frac{(\mu_1-\mu_{1,0})^2}{\frac{\sigma}{k_0}}}$$

lo cual sabemos que es

$$\chi^-2\sim(\nu_n1,\sigma_{n1}^2)$$

en donde

$$\nu_{n1}=n1+\nu_0$$
$$\nu_n\sigma_{n1}^2=\nu_0\sigma_0^2+(n_1-1)\,S_1^2+\frac{k_{1,0}n_1}{k_{1,0}+n_1}\,(\bar{y}_1-\mu_{1,0})^2$$

pero como es para ambos parametros finalmente nos queda:

$$p(\sigma^2|y_1,y_2)\sim\chi^{-2}(\nu_n,\sigma_n^2)$$

con

$$\nu_n=n1+n2+\nu_0$$
$$\nu_n\sigma_n^2=\nu_0\sigma_0^2+(n_1-1)\,S_1^2+(n_2-1)\,S_2^2+\frac{\kappa_{1,0}n_1}{\kappa_{1,0}+n_1}\,(\bar{y}_1-\mu_{1,0})^2+\frac{\kappa_{2,0}n_2}{\kappa_{2,0}+n_2}\,(\bar{y}_2-\mu_{2,0})^2$$ y con ello queda demostrado lo pedido.

Parte B

Nos piden $p(\Delta|y_1,y_2)$ con $\Delta=y_1-y_2$ luego,

$$\propto p(y_1,y_2|\mu_1-\mu_2,\sigma^2)p(\mu_1-\mu_2|\sigma^2)$$

pero claramente esto corresponde a una distribucion normal univariada de parametro Δ debido a que la distribución condicional para $\mu_1-\mu_2$ no deja de ser normal, esto es, debido a que la distribucion de σ^2 no cambia, luego,

$$\frac{(\mu_1-\mu_2)-\mu_{\Delta n}}{\sigma_n\cdot\sqrt{\frac{1}{k_{1,n_1}+k_{2,n2}}}}\sim t_{1-\alpha/2}$$

con $\nu_0+n_1+n_2$ grados de libertad y,

$$\mu_{\Delta n}=\mathbb{E}(\mu_1-\mu_2)=\frac{k_0\mu_{1,0}+\bar{y_1}n_1}{k_0+n_1}-\frac{k_0\mu_{2,0}+\bar{y_2}n_2}{k_0+n_2}$$

y σ_n como fue definido anteriormente. Luego nos queda,

$$(\mu_1-\mu_2\in\mu_{\Delta n}\pm t_{\nu_0+n_1+n_2,1-\alpha/2}\cdot\sigma_n\cdot\sqrt{\frac{1}{k_{1,n_1}+k_{2,n2}}}$$

Parte C

Si tomamos una priori impropia, es claro ver que k_0 y $\nu_0 \to 0$ porque no estamos usando ïnformación a priori"luego,

$$\mu_{1,n1} \to \bar{y_1}$$

$$\mu_{2,n2} \to \bar{y_2}$$

$$\sigma_n^2 \to \frac{1}{n_1}(\sum_{i=1}^{n_1}(y_{1,i} - \bar{y_1})^2) + \frac{1}{n_2}(\sum_{i=1}^{n_2}(y_{2,i} - \bar{y_2})^2)$$

y si usamos la distribución impropia $\hat{p}(\mu_1, \mu_2, \sigma^2) = \frac{1}{\sigma^2}$ nos queda que

$$p(\mu_1, \mu_2, \sigma^2 | y_1, y_2) \propto p(y_1, y_2 | \mu_1, \mu_2, \sigma^2) \cdot \hat{p}(\mu_1, \mu_2, \sigma^2)$$

y obtendremos la misma distribución marginal para μ_1 y μ_2 pero no para σ^2 la cual será $\chi^{-2}(n-1, \sigma_n^2)$ con σ_n^2 definido antriormente. Luego aplicando lo mismo para b), nos queda

$$\frac{(\mu_1 - \mu_2) - (\bar{y_1} - \bar{y_2})}{(s_1 + s2)\sqrt{\frac{1}{n_1+n_2}}} \sim t_{n1+n2-2}$$

Si comparamos con el intervalo clásico el cual es

$$\frac{(\bar{y_1}^* - \bar{y_2}^*) - (\mu_1 - \mu_2)}{(s_1 + s2)\sqrt{\frac{1}{n_1+n_2}}} \sim t_{n1+n2-2}$$

Es interesante notar los ordenes debido a que esta ultima nos dice que antes de muestrear los datos la incertidumbre acerca de σ^2 se representa con una distribución t-student con $n-2$ grados de libertad y el primer intervalo que encontramos nos dice que que después de obtener un muestreo nuestra incertidumbre distribuye t_{n-2}. La diferencia está en que antes de hacer el muestreo $\bar{y_1}^* - \bar{y_2}^*$ y μ_1, μ_2 son desconocidos, pero cuando sabemos el valor de $\bar{y_1}^* = \bar{y_1}, \bar{y_2}^* = \bar{y_2}$ esto nos proporciona información sobre $\mu_1 - \mu_2$.

1. Parte D y E

Estas partes estan explicadas en el script de R, para el cual se uso una sola libreria llamada "distributions3".