Métodos Bayesianos: Datos Normales

Sergio Marrero Marrero

Grado en Economía. Cuarto Curso

Asignatura: Métodos Bayesianos

Tutor: Francisco José Vázquez Polo

Universidad Las Palmas de Gran Canaria (ULPGC)

21 de abril de 2017

1. Ejercicio 1

Se resolverán las cuestiones para ambas poblaciones en paralelo.

Calcular el HDI al 95 % para la distribución a posteriori de la media. Estimadores puntuales. Se usan las siguientes expresiones para calcular la distribución normal de la media a posteriori.

$$\tau_1 = \tau_0 + n\tau \qquad E(\theta \mid \overline{x}) = \mu_1 = \mu_0 \cdot \frac{\tau_0}{\tau_0 + n\tau} + \overline{x} \cdot \frac{n\tau}{\tau_0 + n\tau}$$

Al haber simetría, los estimadores bayesianos puntuales coinciden, por ello se expresarán usando la media a posteriori.Utilizando los datos del problema, aplicando estas ecuaciones para ambas poblaciones y finalmente añadiendo los datos en el FirstBayes obtenemos:(Usando τ en lugar de σ^2)

$$\begin{array}{ll} \text{Poblacion 1:} & \mu \mid \overline{x} \sim \mathbf{N}(13.13, 6), HDI(95\,\%) = [12.331, 13.929], \quad \overline{\mu} = 13.13 \\ \text{Poblacion 2:} & \mu \mid \overline{x} \sim \mathbf{N}(16.75, 9.5), HDI(95\,\%) = [14.67, 18.8], \quad \overline{\mu} = 16.75 \end{array}$$

Calcular la posteriori para el caso no informativo El caso no informativo se modela haciendo que la precisión tienda a cero. Usando las ecuaciones anteriores es fácil ver que en este caso la posteriori tiene esta forma $\mathbf{N}(\overline{x}, n\tau)$, con lo cual se tiene:

$$\begin{array}{ll} \text{Población 1:} & \mu \mid \overline{x} \sim \mathbf{N}(12.95, 5), HDI(95\,\%) = [12.648, 13.827], \quad \overline{\mu} = 12.95 \\ \text{Población 2:} & \mu \mid \overline{x} \sim \mathbf{N}(17.21, 7.5), HDI(95\,\%) = [16.964, 17.456], \quad \overline{\mu} = 17.21 \end{array}$$

Realizar el test de hipótesis Bayesiano $H_0 : \mu = 15 \qquad H_0 : \mu = 16$.

En este caso el número de parámetros posibles se reduce a dos. Por esta razón la no información se modela con una probabilidad de 1/2 para cada alternativa, es decir:

$$\frac{p_0}{p_1} = \frac{\pi_0}{\pi_1} \cdot \frac{f(\underline{x} \mid H_0)}{f(\underline{x} \mid H_1)} = \frac{f(\underline{x} \mid H_0)}{f(\underline{x} \mid H_1)}) = exp\{\frac{-\tau}{2} \cdot \sum_{i=1}^{n} (x_i - \mu_{H_0})^2 - (x_i - \mu_{H_1})^2\}$$

Usando esta expresión para ambas expresiones se obtiene:

Poblacion 1: $\frac{p_0}{p_1} = 332701.177$ **La hipótesis nula es mucho más probable**

Poblacion 2: $\frac{p_0}{p_1} = 2.54X10^{-6}$ **La hipótesis alternativa es mucho más probable**

Considerando ahora el problema como uno de 2 muestras con varianza desconocida. ¿Puede considerarse que las dos medias son diferentes? Para este apartado, tanto la media como la varianza son desconocidas. Para averiguar esto, se analizará la probabilidad de que la media de la población dos sea superior a la de la poblacion 1 : $Prob\{\mu_2 - \mu_1 > 0\} = 1$. La distribución de probabilidad de esto viene en la siguiente imagen.

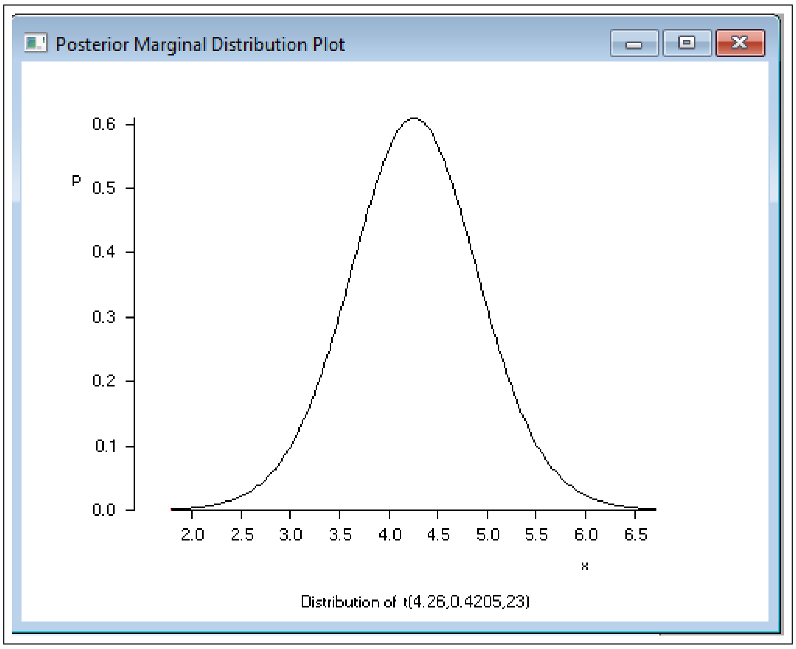

Figura 1: Distribución de la variable: $\mu_2 - \mu_1$

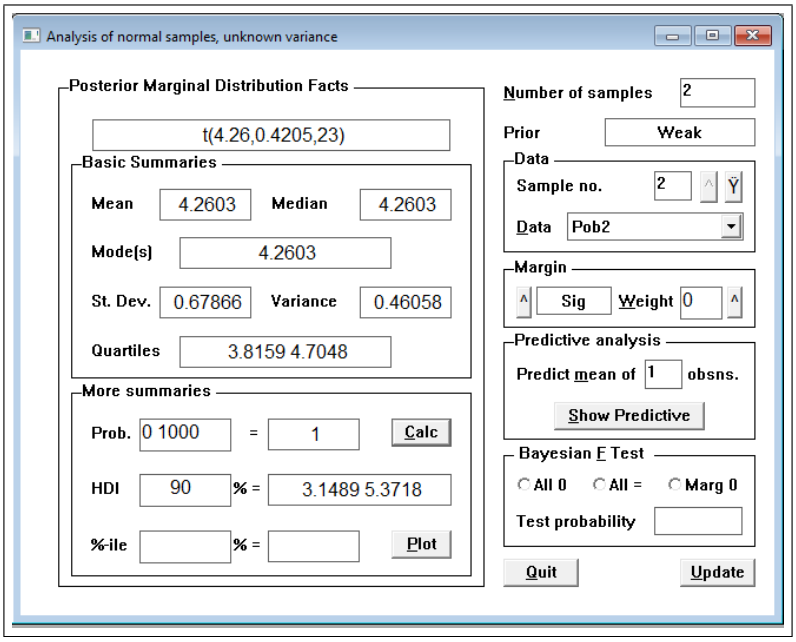

Figura 2: Datos obtenidos

2. Ejercicio 2

Comparar el éxito de la campaña en las dos áreas .

Para ello estudiaremos la probabilidad de que la media de incrementos de ventas en un área sea superior a la otra, de tal forma que : $Prob\{\mu_2 - \mu_1 > 0\}$. La distribución obtenida es la mostrada en la siguiente imagen.

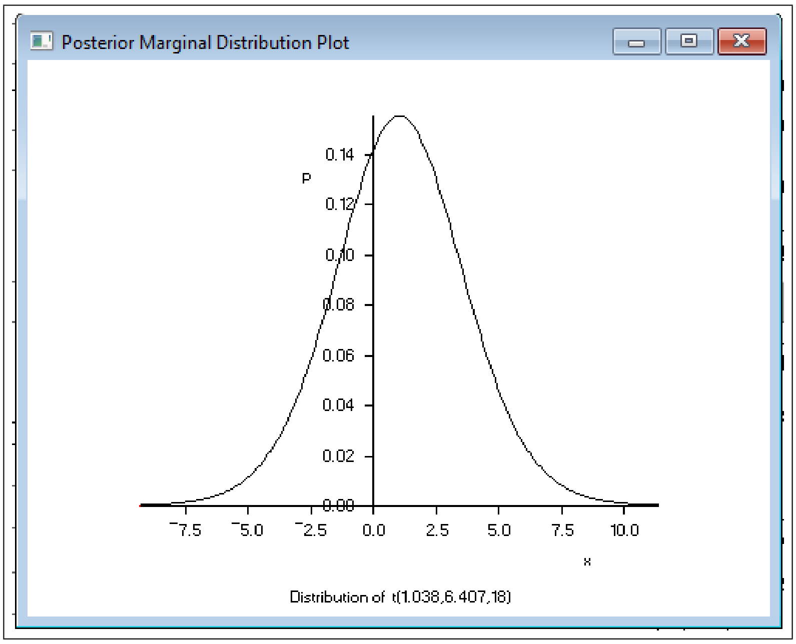

Figura 3: Distribución de la variable: $\mu_2 - \mu_1$. Áreas de ventas

Calculando probabilidades se tiene que : $Prob\{\mu_2 - \mu_1 > 0\} = 0.65$. Por otro lado, $HDI(95\,\%) = [-4.28, 6.25]$. Como regla general, si el intervalo bayesiano contiene al cero, entonces hay indicios para no aceptar que ambas medias provengan de distintas poblaciones. Por otro lado, el odd a posteriori $= \frac{0.65}{0.35} = 1.86$, para la interpretación de los odds también existe una norma general tal que si el odd no es superior a 3, entonces no hay evidencia para aceptar la hipótesis de que sean diferentes. Por todo esto, se concluye que en ambas áreas ha habido un incremento de ventas con igual éxito.

Obtener un HDI al 90 % para σ (la variabilidad de la respuesta entre las tiendas) El intervalo bayesiano de máxima densidad sobre la varianza al 90 % es:$HDI(90\,\%) = [15.52, 62.8]$. Debajo se muestra la distribución de la varianza.

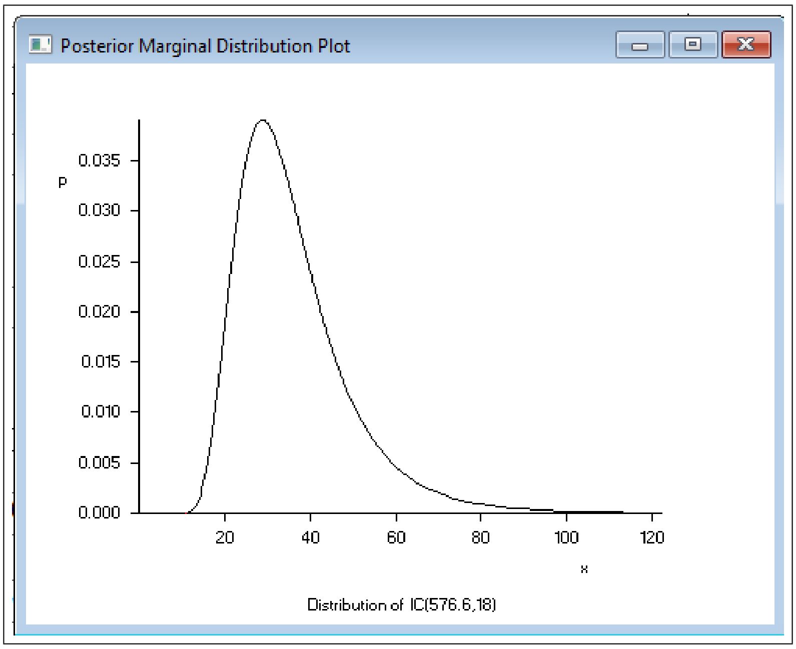

Figura 4: Distribución de la varianza. Áreas de ventas

Calcular las probabilidades de que σ sea menor que el extremo inferior del HDI anterior y de que sea superior al extremo superior. Calcular un intervalo simétrico al 90 %

$$Prob\left\{\sigma < HDI^{-}\right\} = 0.0052$$
$$Prob\left\{\sigma < HDI^{+}\right\} = 0.045$$

Un intervalo simétrico sería aquel que deja la misma probabilidad a ambos lados del HDI(90 %). Con lo cual lo calculamos para que deje un 5 % en ambas colas, siendo el resultado: $[19.963, 61.433]$