

Distribución marginal de la media

Para calcular la distribución marginal de la media μ debemos integrar la f.d.p. conjunta con respecto a τ :

$$\xi_3(\mu) = \int_0^{\infty} \xi_1(\mu|\tau) \cdot \xi_2(\tau) d\tau$$

$$\Rightarrow \xi_3 \propto \int_0^{\infty} \tau^{\alpha_0-1/2} e^{-[\beta_0+1/2\lambda_0(\mu-\mu_0)^2]\tau} d\tau$$

Recordar que:

$$\int_0^{\infty} x^{\alpha-1} e^{-\beta x} dx = \frac{\Gamma(\alpha)}{\beta^\alpha}$$

$$\Rightarrow \xi_3(\mu) \propto [\beta_0 + 1/2\lambda_0(\mu - \mu_0)^2]^{-(\alpha_0+1/2)}$$

$$\Rightarrow \xi_3(\mu) \propto \left[1 + \frac{1}{2\alpha_0} \frac{\lambda_0\alpha_0}{\beta_0} (\mu - \mu_0)^2\right]^{-\frac{(2\alpha_0+1)}{2}}$$

Se define la variable aleatoria:

$$Y = \left(\frac{\lambda_0\alpha_0}{\beta_0}\right)^{1/2} (\mu - \mu_0)$$

$$\Rightarrow \mu = \left(\frac{\beta_0}{\lambda_0\alpha_0}\right)^{1/2} \cdot Y + \mu_0$$

Recordar (Pag. 167. Sección 3.8 del DeGroot & Schervish):

Supongamos que X y Y son dos variables aleatorias tales que $Y = r(X)$ y $X = s(Y)$.

$$G(y) = Pr(Y \leq y) = Pr(r(X) \leq y) = Pr(X \leq s(y)) = F(s(y))$$

$$\Rightarrow g(y) = \frac{dG(y)}{dy} = \frac{dF(s(y))}{dy} = f(s(y)) \cdot \frac{ds(y)}{dy}.$$

$$g(y) = \xi_3 \left[\left(\frac{\beta_0}{\lambda_0 \alpha_0} \right)^{1/2} y + \mu_0 \right] \frac{\beta_0^{1/2}}{(\lambda_0 \alpha_0)^{1/2}}$$

(donde aparece μ sustituimos $\frac{\beta_0^{1/2}}{(\lambda_0 \alpha_0)^{1/2}} y + \mu_0$)

$$\Rightarrow g(y) \propto \left(1 + \frac{y^2}{2\alpha_0} \right)^{-\frac{(2\alpha_0+1)}{2}}$$

$g(y)$ es proporcional a una t-student con $2\alpha_0$ grados de libertad.

La distribución marginal de μ es una t-student trasladada μ_0 unidades y con distinto factor de escala

La media y la varianza de la distribución marginal de μ se obtienen así:

Dado que

$$E[Y] = 0 \quad \text{si } \alpha_0 > 1/2$$

y

$$V[Y] = \frac{2\alpha_0}{2\alpha_0 - 2} = \frac{\alpha_0}{\alpha_0 - 1} \quad \text{si } \alpha_0 > 1$$

\Rightarrow

$$\begin{aligned} E[\mu] &= \mu_0 \\ \text{Var}[\mu] &= \frac{\beta_0}{\lambda_0 \alpha_0} V[Y] \\ &= \frac{\beta_0}{\lambda_0 (\alpha_0 - 1)} \end{aligned}$$

Notar que:

- La probabilidad de que μ esté en cualquier intervalo específico se puede obtener de una tabla t -student aunque $2\alpha_0$ no sea necesariamente entero.
- La distribución final marginal de μ también es una t -student. Por lo tanto la media y la varianza de esta distribución final marginal se pueden obtener de la distribución t -student correspondiente. En este caso debemos hablar de regiones o intervalos de probabilidad a posterior máxima (HPD)

Ejemplo:

- Encontrar el intervalo de probabilidad inicial para la media μ de una distribución Normal-Gamma tal que:
 $E[\mu] = 10, V[\mu] = 8, E[\tau] = 2, V[\tau] = 2.$
- Si se observa una muestra de tamaño 20 tal que $\bar{x}_n = 7,5$ y $\sum_{i=1}^n (x_i - \bar{x}_n)^2 = 28$, encontrar el intervalo de probabilidad final para μ .

(Ejemplo desarrollado en clase)

Ejemplo: Casas Hogares en New Mexico

- Datos de casas hogares con licencia analizados por Smith, Piland y Fisher (1992).
- Se considera la variable: Total anual de días de hospitalización (X), medida en cientos.
- Antes de observar los datos se modelan los valores de X como una normal para cada casa hogar, con media μ y precisión τ .
- Se usa la siguiente información adicional para estimar μ y τ :
 - Hay en promedio 111 camas con desviación estándar de 43.5 camas. Se asume a priori un 50% de ocupación. Se puede estimar una media (en cientos de pacientes por a no) como $0,5 \times 365 \times 1,11 \approx 200$ y una desviación estándar de $0,5 \times 365 \times 0,435 \approx 6300^{1/2}$.
 - Se atribuye la mitad de la varianza de 6300 a las casas hogares, y la otra mitad es la varianza de μ . Entonces $Var(\mu) = 3150$ y $E(\tau) = 1/3150$. Si se escoge $\alpha_0 = 2$, como $E(\tau) = \alpha_0/\beta_0$ entonces $\beta_0 = 6300$.
 - Usando que $E(\mu) = \mu_0$ y $Var(\mu) = \frac{\beta_0}{\lambda_0(\alpha_0-1)}$, se obtiene $\mu_0 = 200$ y $\lambda_0 = 2$.

- La variable aleatoria Y definida anteriormente tiene una distribución t-student con $2\alpha_0$ grados de libertad. Por lo tanto la variable $0.025(\mu - \mu_0)$ tiene una distribución t-student con cuatro grados de libertad.
- El cuantil de la t-student con dos grados de libertad, correspondiente a una probabilidad de 0.975 es 2.776. Por lo tanto $Pr[-2,776 < 0,025(\mu - 200) < 2,776] = 0,95$, lo cual es equivalente a: $Pr(89 < \mu < 311) = 0,95$. El intervalo (89,311) es el intervalo de probabilidad del 95 % para μ .
- Al observar la muestra:
128, 281, 291, 238, 155, 148, 154, 232, 316, 96, 146, 151, 100, 213, 208, 157, 48, 217,
se obtiene: $\bar{x}_n = 182,17$ y $s_n^2 = 88678,5$.
- Los parámetros de la normal-gamma final son: $\mu_1 = 183,95$,
 $\lambda_1 = 20$, $\alpha_1 = 11$, $\beta_1 = 50925,37$.
- Usando el argumento anterior se llega a un intervalo de probabilidad a posterior del 95 % = (152.38,215.52), el cual es mucho más corto que el intervalo inicial.

Comparación con el Intervalo de Confianza

La variable aleatoria:

$$U = \frac{n^{1/2}(\bar{X}_n - \mu)}{\left(\frac{S_n^2}{n-1}\right)^{1/2}}$$

tiene una distribución t-student con 17 grados de libertad ($n = 18$).

El cuantil 0.975 de la t-student es 2.110. Por lo tanto

$$Pr(-2,110 < U < 2,110) = 0,95$$

En este caso el intervalo con coeficiente de confianza del 95 % para μ es (146.25, 218.09), el cual es cercano al intervalo (152.38, 215.52).

Se obtienen resultados similares aunque la interpretación de los intervalos es distinta!

Contraste de Hipótesis

Sea θ un parámetro desconocido perteneciente al espacio paramétrico Ω . Supongamos que Ω puede descomponerse en los conjuntos Ω_0 y Ω_1 disjuntos.

Sea:

- H_0 : Hipótesis de que $\theta \in \Omega_0$
- H_1 : Hipótesis de que $\theta \in \Omega_1$

Como $\Omega = \Omega_0 \cup \Omega_1$, una de las dos hipótesis debe ser verdadera. Debemos decidir entre si aceptar H_0 ó aceptar H_1 .

Problema de Contraste de Hipótesis

Sólo hay dos decisiones posibles. Para tomar la decisión se lleva a cabo un procedimiento de contraste.

En general H_0 y H_1 se tratan de una forma distinta. A H_0 se le llama la **hipótesis nula** y a H_1 se le llama la **hipótesis alternativa**.

Hipótesis Simples y Compuestas

Supongamos que X_1, \dots, X_n es una muestra aleatoria de una f.d.p. $f(x|\theta)$ tal que $\theta \in \Omega$ y $\Omega = \Omega_0 \cup \Omega_1$. (Ω_0 y Ω_1 son disjuntos)

Consideremos:

- $H_0 : \theta \in \Omega_0$
- $H_1 : \theta \in \Omega_1$
 - Si el conjunto Ω_i sólo contiene un valor de θ se dice que H_i es **simple**.
 - Si Ω_i contiene más de un valor se dice que H_i es **compuesta**.

En el caso $H_0 : \theta = \theta_0$ el tamaño del contraste es $\pi(\theta_0)$.

Nota:

Con una hipótesis simple la distribución de las observaciones queda completamente especificada. Con una hipótesis compuesta se dice que las observaciones pertenecen a una cierta clase de distribuciones.

Región Crítica

Consideremos el problema de contraste:

- $H_0 : \theta \in \Omega_0$

- $H_1 : \theta \in \Omega_1$

Vamos a definir el concepto de Región Crítica.

Región Crítica (Cont.)

Antes de tomar una decisión se observa una muestra X_1, \dots, X_n de una distribución de probabilidades con parámetro desconocido θ .

Sea S el espacio aleatorio n-dimensional de todos los posibles valores de $\mathbf{X} = X_1, \dots, X_n$ (Vector aleatorio n-dimensional).

Dividimos a S en dos subconjuntos S_0 y S_1 :

- S_0 : El que contiene todos los valores de \mathbf{X} para los cuales se acepta H_0 .
- S_1 : El que contienen todos los valores de \mathbf{X} para los cuales se rechaza H_0 (por lo tanto se acepta H_1).

El conjunto para el cuál H_0 es rechazada se denomina **región crítica del contraste**.

En conclusión:

Un procedimiento de contraste se determina especificando la región crítica del contraste (donde H_0 será rechazada). Por lo tanto, en el complemento, H_0 será aceptada.

Nota: Las divisiones del espacio paramétrico Ω_0 y Ω_1 y del espacio muestral S_0 y S_1 están relacionadas entre sí, pero no coinciden. Si la muestra aleatoria X cae en la región crítica S_1 rechazamos la hipótesis nula Ω_0 . Si $X \in S_0$, no rechazamos Ω_0 .