Tema 7

CONTRASTE DE HIPÓTESIS

7.1. Planteamiento del problema y conceptos básicos

Hasta ahora se ha estudiado, para un problema de inferencia paramétrica, la estimación puntual y la estimación por intervalos. Una tercera opción, que se puede plantear para este tipo de problemas, es el contraste de hipótesis.

Esta opción consiste en plantear una hipótesis sobre el parámetro y decidir según la información de la muestra si dicha hipótesis es rechazada por la muestra o no. Debe tenerse en cuenta que la distribución de la variable, de la que se tiene la muestra, depende de si la hipótesis planteada es cierta o no.

Ejemplo: Se tiene una moneda con probabilidad p de que salga cara y se desea saber si dicha probabilidad es $1/3$. Plantear el problema de contraste asociado.

Definición: Sea $X \rightsquigarrow F \in \{F_\theta : \theta \in \Theta \subseteq \mathbb{R}^k\}$. Una *hipótesis* (paramétrica) es una declaración (afirmación o conjetura) acerca del parámetro desconocido θ. Por ejemplo, $\theta \in \Theta_0$ con $\Theta_0 \subset \Theta$.

Cuando se formula una hipótesis sobre el parámetro, se establece una partición sobre el espacio paramétrico, entre los valores de la hipótesis que se formula y el resto de los valores: $\Theta = \Theta_0 \cup \Theta_1$, $\Theta_0 \cap \Theta_1 = \emptyset$.

Ejemplo: $X \rightsquigarrow \mathcal{N}(\mu, \sigma^2)$ Si la hipótesis que se desea plantear es $\mu = 3$, entonces separo $\Theta = \{3\} \cup \mathbb{R} - \{3\}$. Otra opción sería plantear la hipótesis $\mu > 0$, entonces separo $\Theta = \mathbb{R}^+ \cup \mathbb{R}^-$.

La hipótesis no sólo establece una partición sobre el espacio paramétrico, sino también sobre la familia de distribuciones: $\{F_\theta;\, \theta \in \Theta\}$=$\{F_\theta;\, \theta \in \Theta_0\} \cup \{F_\theta;\, \theta \in \Theta_1\}$.

Esto indica que toda hipótesis que se hace sobre el parámetro lleva implícita otra hipótesis que es la negación de la primera.

Usualmente, a la primera hipótesis que se hace se le llama *hipótesis nula*, y se denota por

$$H_0 : \theta \in \Theta_0 \subset \Theta$$

A la negación de dicha hipótesis, se le llama *hipótesis alternativa*, y se denota por

$$H_1 : \theta \in \Theta_1 = \Theta - \Theta_0$$

Un problema de contraste de hipótesis se leería: Se contrasta la hipótesis nula $H_0 : \theta \in \Theta_0$, frente a la hipótesis alternativa $H_1 : \theta \in \Theta_1$. Bajo el punto de vista clásico, la H_0 se formula como aquella hipótesis en la que se tiene mayor confianza y debe tenerse en cuenta que H_0 y H_1 no son intercambiables.

Si el subconjunto que define la hipótesis, Θ_0 (Θ_1), contiene un solo punto, entonces esa hipótesis, H_0 (H_1), se denomina *hipótesis simple*. En caso contrario, cuando conste de más de un punto, se denomina *hipótesis compuesta*.

Ejemplos:

1. Sea $X \rightsquigarrow B(1, p)$. Se pueden plantear, por ejemplo, los siguientes contrastes de hipótesis:

   - $\begin{cases} H_0 : & p = 1/2 \\ H_1 : & p \neq 1/2 \end{cases}$
   - $\begin{cases} H_0 : & p \leq 1/2 \\ H_1 : & p > 1/2 \end{cases}$
   - $\begin{cases} H_0 : & p = 1/2 \\ H_1 : & p = 1/3 \end{cases}$

   ¿Qué tipo de contraste son?. (El último contraste implica que $X \rightsquigarrow \{B(1, p); p \in \{1/2, 1/3\}\}$, ya que si el espacio paramétrico no se especifica de partida es la unión de las dos hipótesis.)

2. Si se tiene $X \rightsquigarrow \mathcal{N}(\mu, \sigma^2)$. ¿Qué tipo de contraste es el siguiente?

   $\begin{cases} H_0 : & \mu = \mu_0 \\ H_1 : & \mu \neq \mu_0 \end{cases}$

Una vez estudiado como se plantea un problema de contraste de hipótesis, el siguiente paso es buscar un procedimiento que permita llegar a una solución del problema. Dicho procedimiento debe estar basado en las observaciones de la variable aleatoria cuya distribución depende de que la hipótesis formulada sea cierta o no.

Definición: Un *test de hipótesis* es un estadístico, $\varphi(X_1, \ldots, X_n)$, con valores en $[0, 1]$, que especifica la probabilidad de rechazar H_0 para cada realización muestral.

Los test de hipótesis se pueden clasificar en dos tipos:

- Test No Aleatorizados: Un test no aleatorizado va a ser un procedimiento, mediante el cual, una vez observada la variable aleatoria, una o varias veces, según el tamaño de la muestra, se decide rechazar la hipótesis nula con probabilidad 1 ó se decide rechazarla con probabilidad 0.

  Matemáticamente hablando, se considera $\mathcal{X}^n$, el espacio muestral asociado a la muestra de tamaño n, $(X_1, \ldots, X_n)$, y el test consiste en elegir una *región crítica o de rechazo* $\mathcal{C} \subset \mathcal{X}^n$ tal que si:

  $(X_1, \ldots, X_n) \in \mathcal{C}$, entonces se rechaza la hipótesis nula H_0 con probabilidad 1.

  $(X_1, \ldots, X_n) \notin \mathcal{C}$, entonces se rechaza H_0 con probabilidad 0 o equivalentemente no se rechaza H_0 con probabilidad 1.

  Esto se puede escribir, de la siguiente forma, mediante una función $\varphi : \mathcal{X}^n \to \{0, 1\}$, donde $\varphi(X_1, \ldots, X_n) = \begin{cases} 1 & \text{si } (X_1, \ldots, X_n) \in \mathcal{C} \\ 0 & \text{si } (X_1, \ldots, X_n) \notin \mathcal{C} \end{cases}$

  Entonces, la región crítica $\mathcal{C} = \varphi^{-1}(1) = \{(x_1, \ldots, x_n) \in \mathcal{X}^n / \text{ se rechaza } H_0\}$. Al complementario de $\mathcal{C}$ se le llama región de aceptación del test, $\overline{\mathcal{C}}$.

  Este tipo de test sólo rechaza o no rechaza radicalmente ya que φ sólo puede tomar los valores 0 y 1.

- Test Aleatorizados: Un test aleatorizado va a ser un procedimiento, mediante el cual, una vez observada la variable aleatoria, una o varias veces, según el tamaño de la muestra, se decide rechazar la hipótesis nula con una cierta probabilidad que dependerá del valor observado.

  La diferencia con el test no aleatorizado es que φ no va a estar definida en $\{0, 1\}$, sino $\varphi : \mathcal{X}^n \to [0, 1]$, donde, como ya se ha indicado, $\varphi(X_1, \ldots, X_n)$ expresa la probabilidad de rechazar la hipótesis nula H_0 para cada realización muestral. Además, $\forall \theta \in \Theta$, $E_\theta \varphi(X_1, \ldots, X_n)$ va a representar la probabilidad media o esperanza de rechazar H_0 cuando θ es el verdadero valor del parámetro.

  La aplicación práctica de este tipo de test requiere de procedimientos auxiliares para tomar la decisión final. Sin embargo, son de gran interés teórico ya que permiten dar la solución óptima, (siempre desde el punto de vista clásico), a muchos problemas de contraste.

En relación con los problemas de diseño y elección de posibles test, se van a definir los siguientes conceptos: tipos de errores, nivel de significación, tamaño de un test y función de potencia.

- Tipos de errores: Los test llevan asociado dos tipos de errores:

  - Error de tipo I: Se rechaza la hipótesis nula cuando es cierta.
  - Error de tipo II: No se rechaza la hipótesis nula cuando es falsa.

	H_0 verdadera	H_0 falsa
No rechazar H_0	Correcto	Error Tipo II
Rechazar H_0	Error Tipo I	Correcto

  En test no aleatorizados, la probabilidad de cometer un error de tipo I sería $P_\theta[(X_1,\ldots,X_n)\in\mathcal{C}]=P_\theta(\mathcal{C})$ con $\theta\in\Theta_0$ y la probabilidad de cometer un error de tipo II sería $P_\theta[(X_1,\ldots,X_n)\in\overline{\mathcal{C}}]=P_\theta(\overline{\mathcal{C}})$ con $\theta\in\Theta_1$.

  Idealmente interesaría encontrar una región crítica, $\mathcal{C}$, que minimice los errores Tipo I y Tipo II. En la práctica, esto no será posible (salvo en situaciones muy específicas), por lo que el enfoque que se hace es el siguiente:

  1. Se acota o limita la probabilidad de error de tipo I a un nivel, α, prefijado. Evidentemente se van a consideran valores de α pequeños como 0.1, 0.05 ó 0.01.
  2. Se trata de minimizar entonces la probabilidad de error de tipo II, entre los test que son admisibles.

- Función de potencia: Para un test φ, se define la *función de potencia* como la función que asocia a cada parámetro, θ, la probabilidad media de rechazar H_0 cuando el (verdadero) valor del parámetro es θ.

$$\begin{array}{rcl} \beta_\varphi:\Theta & \to & [0,1] \\ \theta & \to & \beta_\varphi(\theta)=E_\theta[\varphi(X_1,\ldots,X_n)] \end{array}$$

  En el caso de un test φ no aleatorizado, esto se puede expresar,

$$\begin{array}{rcl} \beta_\varphi:\Theta & \to & [0,1] \\ \theta & \to & \beta_\varphi(\theta)=P_\theta[(X_1,\ldots,X_n)\in\mathcal{C}] \end{array}$$

- Tamaño de un test: Para un test φ, se define el *tamaño del test* como la máxima probabilidad media de cometer un error de tipo I con dicho test, es decir

$$\sup_{\theta\in\Theta_0} E_\theta[\varphi(X_1,\ldots,X_n)]=\sup_{\theta\in\Theta_0}\beta_\varphi(\theta)$$

  En el caso de un test φ no aleatorizado, esto se puede expresar,

$$\sup_{\theta\in\Theta_0} P_\theta[(X_1,\ldots,X_n)\in\mathcal{C}]$$

- Nivel de significación: Se dice que φ es un test de hipótesis con *nivel de significación* α ($\in [0,1]$) si su tamaño es menor o igual que que α (cota superior de las probabilidades medias de cometer error de tipo I), es decir:

$$\forall \theta \in \Theta_0,\ \beta_\varphi(\theta) = E_\theta[\varphi(X_1,\dots,X_n)] \leq \alpha.$$

  En el caso de un test φ no aleatorizado, esto se puede expresar,

$$\forall \theta \in \Theta_0\ P_\theta[(X_1,\dots,X_n) \in \mathcal{C}] \leq \alpha.$$

Planteado un problema de contraste de hipótesis, el objetivo será, dado un nivel de significación α prefijado, encontrar un test φ con dicho nivel de significación, es decir con error de tipo I menor o igual a α, que maximice la función potencia en todos los valores de la hipótesis alternativa, es decir, que minimice el error de tipo II.

En general puede ocurrir que existan más de un test que verifique dichas condiciones. En dicho caso se debe seleccionar el más potente, entendiendo que dados dos test φ y φ^*, ambos con nivel de significación α, se dice que φ es *más potente* que φ^*, para un valor $\theta \in \Theta_1$, si $\beta_\varphi(\theta) \geq \beta_{\varphi^*}(\theta)$.

Definición: Dado un nivel de significación α fijo, se dice que un test φ es *uniformemente más potente* (UMP), a ese nivel, si se cumplen dos cosas:

1. φ tiene nivel de significación α, es decir

$$E_\theta\varphi(X_1,\dots,X_n) \leq \alpha\ \forall \theta \in \Theta_0.$$

2. Para cualquier otro test φ^*, con nivel de significación α, se cumple que la función potencia en φ es mayor que en φ^*, $\forall \theta \in \Theta_1$,

$$\beta_\varphi(\theta) \geq \beta_{\varphi^*}(\theta), \forall \theta \in \Theta_1.$$

En general no tiene porque existir un test UMP para un problema dado (esta relación establece un orden parcial).

Ejemplos:

1. Sea $X \rightsquigarrow \mathcal{N}(\mu, 1)$. Formular el problema de contrastar $\mu = 15$ frente a $\mu = 18$ y resolverlo de forma intuitiva usando un test no aleatorizado basado en $\bar{X}$.

2. Sea $(X_1,\dots,X_5)$ una m.a.s. de $X \rightsquigarrow \{B(1,p);\ p \in [0,1]\}$. Construir un test para el contraste $H_0 : p = 1/2$ frente a $H_1 : p \neq 1/2$ con nivel de significación 0.1.

7.2. Test de Neyman-Pearson

Lema de Neyman-Pearson:

Sea $X \rightsquigarrow F \in \{F_\theta : \theta \in \Theta = \{\theta_0, \theta_1\}\}$, y sea f_θ la fmp o la fdp de X, según X sea discreta o continua. Por simplicidad, se denota $f_0 \equiv f_{\theta_0}$ y $f_1 \equiv f_{\theta_1}$. Sea $X_1, \ldots, X_n$, una m.a.s. de X y se platea el siguiente problema de contraste de hipótesis:

$$\begin{cases} H_0: & \theta = \theta_0, \\ H_1: & \theta = \theta_1. \end{cases}$$

a) Sea test φ un test de la forma

$$\varphi(X_1, \ldots, X_n) = \begin{cases} 1 & \text{si } f_1^n(X_1, \ldots, X_n) > k f_0^n(X_1, \ldots, X_n), \\ \gamma(X_1, \ldots, X_n) & \text{si } f_1^n(X_1, \ldots, X_n) = k f_0^n(X_1, \ldots, X_n), \\ 0 & \text{si } f_1^n(X_1, \ldots, X_n) < k f_0^n(X_1, \ldots, X_n), \end{cases}$$

para alguna constante $k \in \mathbb{R}^+ \cup \{0\}$, y $\gamma(x_1, \ldots, x_n) \in [0, 1]$. Si $\varphi(X_1, \ldots, X_n)$ tiene tamaño α, es de máxima potencia a nivel de significación α. Un test de esta forma se denomina test de Neyman-Pearson.

b) Para todo $\alpha \in (0, 1]$ existe un test de Neyman-Pearson de tamaño α, con $\gamma(X_1, \ldots, X_n) = \gamma$ constante.

c) Si $\varphi'(X_1, \ldots, X_n)$ es un test de tamaño $\alpha \in (0, 1]$ y es de máxima potencia a nivel de significación α, $\varphi'(X_1, \ldots, X_n)$ es un test de Neyman-Pearson.

d) El test de Neyman-Pearson de tamaño $\alpha \in (0, 1]$ con $\gamma(X_1, \ldots, X_n) = \gamma$ constante es único con probabilidad uno bajo P_{θ_0} y bajo P_{θ_1}.

e) El test de máxima potencia entre todos los de nivel de significación 0 (tamaño 0) es

$$\varphi_0(X_1, \ldots, X_n) = \begin{cases} 1 & \text{si } f_0(X_1, \ldots, X_n) = 0, \\ 0 & \text{si } f_0(X_1, \ldots, X_n) > 0. \end{cases}$$

Ejemplos:

1. Sea $(X_1, \ldots, X_5)$ una m.a.s. de $X \rightsquigarrow B(1, p)$. Encontrar el test más potente de tamaño α =0.05 para resolver el problema de contraste:

$$\begin{cases} H_0: & p = 1/2 \\ H_1: & p = 3/4 \end{cases}$$

2. Sea $(X_1, \ldots, X_n)$ una m.a.s. de X cuya función de densidad viene dada por $f_\theta(x) = e^{\theta - x}$ $x \geq \theta$. Encontrar el test más potente de tamaño α para resolver el problema de contraste:

$$\begin{cases} H_0: & \theta = \theta_0 \\ H_1: & \theta = \theta_1 \end{cases}$$

7.3. Test de la razón de verosimilitudes

Los test de la razón de verosimilitud están basados en el Principio de Máxima Verosimilitud. Sea $X \rightsquigarrow \{F_\theta : \theta \in \Theta \subseteq \mathbb{R}^k\}$ y se considera el siguiente contraste: $\left\{\begin{array}{ll} H_0: & \theta \in \Theta_0 \\ H_1: & \theta \in \Theta_1 \end{array}\right.$ con $\Theta = \Theta_0 \cup \Theta_1$. Sea $X_1, \ldots, X_n$ una m.a.s. de X. Se puede construir la siguiente función:

$$\lambda(x_1, \ldots, x_n) = \frac{\sup_{\theta \in \Theta_0} L_{x_1,\ldots,x_n}(\theta)}{\sup_{\theta \in \Theta} L_{x_1,\ldots,x_n}(\theta)} \quad \forall (x_1, \ldots, x_n) \in \chi^n$$

donde $L_{x_1,\ldots,x_n}(\theta) = f_\theta(x_1, \ldots, x_n)$ es la función de verosimilitud asociada a la muestra, $(x_1, \ldots, x_n)$. Es evidente que $0 \leq \lambda(x_1, \ldots, x_n) \leq 1$ puesto que el denominador es el supremo en todo el espacio paramétrico, luego será mayor o igual que el numerador.

El *test de la razón de verosimilitud* (TRV) será:

$$\varphi(X_1, \ldots, X_n) = \left\{\begin{array}{ll} 1 & \lambda(X_1, \ldots, X_n) < c \\ 0 & \lambda(X_1, \ldots, X_n) \geq c \end{array}\right.$$

siendo $c \in (0, 1]$ una contante que se determina imponiendo el tamaño o el nivel de significación.

El TRV es un test no aleatorizado, por tanto habrá veces que se pueda calcular el tamaño y otras en que se tomará el más próximo. Si se opta por aleatorizar el test, ya no será el TRV, sino la aleatorización de él.

Propiedades del TRV

Si $T(X_1, \ldots, X_n)$ es un estadístico suficiente para la familia $\{F_\theta : \theta \in \Theta\}$, entonces el TRV es función de $T(X_1, \ldots, X_n)$.

Ejemplo: Sea X la variable aleatoria que mide el número de accidentes por semana, $X \rightsquigarrow P(\theta)$, $\theta \leq 2.5$. Después de modificaciones en la vía se desea comprobar si la media de accidentes sigue sin superar 2.5 para lo cual se realiza un estudio del número de accidentes en 4 semanas obteniendo los valores 4, 3, 2 y 3. Resolver el problema de contraste planteado al nivel de significación de 0.05.

7.4. Contrastes sobre los parámetros de una población normal

Sea $X \rightsquigarrow \mathcal{N}(\mu, \sigma^2)$, y $X_1, \ldots, X_n$ una m.a.s. Se pueden plantear los siguientes test según sean conocidos o no los parámetros de la distribución:

- Test sobre la media de una distribución normal con varianza conocida:

  Fijado un nivel de significación, α, se calcula el valor del estadístico

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \rightsquigarrow \mathcal{N}(0,1) \Rightarrow z_{exp} = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$$

Contraste	Región Crítica del TRV de tamaño α Se rechaza H_0 a nivel α si
$H_0 : \mu = \mu_0$ $H_1 : \mu \neq \mu_0$	$z_{exp} < -z_{\alpha/2}$ o $z_{exp} > z_{\alpha/2} \quad \alpha \in [0,1]$ equivalentemente $\lvert z_{exp}\rvert > z_{\alpha/2}$
$H_0 : \mu \leq \mu_0$ $H_1 : \mu > \mu_0$	$z_{exp} > z_{\alpha} \quad \alpha \leq 1/2$
$H_0 : \mu \geq \mu_0$ $H_1 : \mu < \mu_0$	$z_{exp} < -z_{\alpha}(z_{1-\alpha}) \quad \alpha \leq 1/2$

  (Ver páginas 5 y 6 del resumen)

- Test sobre la media de una distribución normal con varianza desconocida:

  Fijado un nivel de significación, α, se calcula el valor del estadístico

$$T = \frac{\bar{X} - \mu}{S/\sqrt{n}} \rightsquigarrow t_{n-1} \Rightarrow t_{exp} = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

Contraste	Región Crítica del TRV de tamaño α Se rechaza H_0 a nivel α si
$H_0 : \mu = \mu_0$ $H_1 : \mu \neq \mu_0$	$t_{exp} \leq -t_{n-1;\alpha/2}$ o $t_{exp} \geq t_{n-1;\alpha/2} \quad \alpha \in [0,1]$ equivalentemente $\lvert t_{exp}\rvert > t_{n-1;\alpha/2}$
$H_0 : \mu \leq \mu_0$ $H_1 : \mu > \mu_0$	$t_{exp} > t_{n-1;\alpha} \quad \alpha \leq 1/2$
$H_0 : \mu \geq \mu_0$ $H_1 : \mu < \mu_0$	$t_{exp} < -t_{n-1;\alpha}(t_{n-1;1-\alpha}) \quad \alpha \leq 1/2$

  (Ver páginas 7 y 8 del resumen)

- Test sobre la varianza de una distribución normal con media conocida:

  Fijado un nivel de significación, α, se calcula el valor del estadístico

$$\chi^2 = \frac{\sum_{i=1}^{n}(X_i - \mu)^2}{\sigma^2} \rightsquigarrow \chi^2_n \Rightarrow \chi^2_{exp} = \frac{\sum_{i=1}^{n}(x_i - \mu)^2}{\sigma_0^2}$$

Contraste	Región Crítica del TRV de tamaño $\approx \alpha$ Se rechaza H_0 a nivel α si
$H_0 : \sigma^2 = \sigma_0^2$ $H_1 : \sigma^2 \neq \sigma_0^2$	$\chi^2_{exp} < \chi^2_{n;1-\alpha/2}$ o $\chi^2_{exp} > \chi^2_{n;\alpha/2}$ $\quad \alpha \in [0,1]$
$H_0 : \sigma^2 \leq \sigma_0^2$ $H_1 : \sigma^2 > \sigma_0^2$	$\chi^2_{exp} > \chi^2_{n;\alpha}$ $\quad \alpha \leq P(Y > n)$ $(Y \rightsquigarrow \chi^2(n))$
$H_0 : \sigma^2 \geq \sigma_0^2$ $H_1 : \sigma^2 < \sigma_0^2$	$\chi^2_{exp} < \chi^2_{n;1-\alpha}$ $\quad \alpha \leq P(Y \leq n)$ $(Y \rightsquigarrow \chi^2(n))$

(Ver páginas 9 y 10 del resumen)

- Test sobre la varianza de una distribución normal con media desconocida:

  Fijado un nivel de significación, α, se calcula el valor del estadístico

$$\chi^2 = \frac{(n-1)S^2}{\sigma^2} \rightsquigarrow \chi^2_{n-1} \Rightarrow \chi^2_{exp} = \frac{(n-1)s^2}{\sigma_0^2}$$

Contraste	Región Crítica del TRV de tamaño $\approx \alpha$ Se rechaza H_0 a nivel α si
$H_0 : \sigma^2 = \sigma_0^2$ $H_1 : \sigma^2 \neq \sigma_0^2$	$\chi^2_{exp} < \chi^2_{n-1;1-\alpha/2}$ o $\chi^2_{exp} > \chi^2_{n-1;\alpha/2}$ $\quad \alpha \in [0,1]$
$H_0 : \sigma^2 \leq \sigma_0^2$ $H_1 : \sigma^2 > \sigma_0^2$	$\chi^2_{exp} > \chi^2_{n-1;\alpha}$ $\quad \alpha \leq P(Y > n)$ $(Y \rightsquigarrow \chi^2(n-1))$
$H_0 : \sigma^2 \geq \sigma_0^2$ $H_1 : \sigma^2 < \sigma_0^2$	$\chi^2_{exp} < \chi^2_{n-1;1-\alpha}$ $\quad \alpha \leq P(Y \leq n)$ $(Y \rightsquigarrow \chi^2(n-1))$

(Ver páginas 11 y 12 del resumen)

7.5. Contrastes sobre los parámetros de dos poblaciones normales

Sean $X \rightsquigarrow \mathcal{N}(\mu_1, \sigma_1^2)$ e $Y \rightsquigarrow \mathcal{N}(\mu_2, \sigma_2^2)$, y sus muestras aleatorias $X_1, \ldots, X_{n_1}$ e $Y_1, \ldots, Y_{n_2}$. Se pueden plantear los siguientes test según sean conocidos o no los parámetros de las distribuciones:

- Test para comparar las medias de dos distribuciones normales con varianzas conocidas: Fijado un nivel de significación, α, se calcula el valor del estadístico

$$Z = \frac{\overline{X} - \overline{Y} - (\mu_1 - \mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} \rightsquigarrow \mathcal{N}(0,1) \Rightarrow z_{exp} = \frac{\overline{x} - \overline{y} - \mu_0}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}$$

Contraste	Región Crítica Se rechaza H_0 a nivel α si
$H_0 : \mu_1 - \mu_2 = \mu_0$ $H_1 : \mu_1 - \mu_2 \neq \mu_0$	$z_{exp} < -z_{\alpha/2}$ o $z_{exp} > z_{\alpha/2}$ equivalentemente $\lvert z_{exp}\rvert > z_{\alpha/2}$
$H_0 : \mu_1 - \mu_2 \leq \mu_0$ $H_1 : \mu_1 - \mu_2 > \mu_0$	$z_{exp} > z_\alpha$
$H_0 : \mu_1 - \mu_2 \geq \mu_0$ $H_1 : \mu_1 - \mu_2 < \mu_0$	$z_{exp} < -z_\alpha(z_{1-\alpha})$

- Test para comparar las medias de dos distribuciones normales con varianzas desconocidas pero iguales: Fijado un nivel de significación, α, se calcula el valor del estadístico

$$T = \frac{\overline{X} - \overline{Y} - (\mu_1 - \mu_2)}{S_p\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}} \rightsquigarrow t_{n_1+n_2-2} \Rightarrow t_{exp} = \frac{\overline{x} - \overline{y} - \mu_0}{s_p\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}}$$

con $S_p = \sqrt{\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}}$.

Contraste	Región Crítica Se rechaza H_0 a nivel α si
$H_0 : \mu_1 - \mu_2 = \mu_0$ $H_1 : \mu_x - \mu_y \neq \mu_0$	$t_{exp} < -t_{n_1+n_2-2;\alpha/2}$ o $t_{exp} > t_{n_1+n_2-2;\alpha/2}$ equivalentemente $\lvert t_{exp}\rvert > t_{n_1+n_2-2;\alpha/2}$
$H_0 : \mu_1 - \mu_2 \leq \mu_0$ $H_1 : \mu_1 - \mu_2 > \mu_0$	$t_{exp} > t_{n_1+n_2-2;\alpha}$
$H_0 : \mu_1 - \mu_2 \geq \mu_0$ $H_1 : \mu_1 - \mu_2 < \mu_0$	$t_{exp} < -t_{n_1+n_2-2;\alpha}(t_{n_1+n_2-2;1-\alpha})$

- Test para comparar las varianzas de dos distribuciones normales con medias conocidas: Fijado un nivel de significación, α, se calcula el valor del estadístico

$$F = \frac{\sigma_1^2}{\sigma_2^2}\frac{\sum_{i=1}^{n_2}(Y_i - \mu_2)^2/n_2}{\sum_{i=1}^{n_1}(X_i - \mu_1)^2/n_1} \rightsquigarrow F_{n_2,n_1} \Rightarrow F_{exp} = \frac{\sum_{i=1}^{n_2}(y_i - \mu_2)^2/n_2}{\sum_{i=1}^{n_1}(x_i - \mu_1)^2/n_1}$$

Contraste	Región Crítica Se rechaza H_0 a nivel α si
$H_0 : \sigma_1^2 = \sigma_2^2$ $H_1 : \sigma_1^2 \neq \sigma_2^2$	$F_{exp} < F_{n_2,n_1;1-\alpha/2}$ o $F_{exp} > F_{n_2,n_1;\alpha/2}$
$H_0 : \sigma_1^2 \leq \sigma_2^2$ $H_1 : \sigma_1^2 > \sigma_2^2$	$F_{exp} > F_{n_2,n_1;\alpha}$
$H_0 : \sigma_1^2 \geq \sigma_2^2$ $H_1 : \sigma_1^2 < \sigma_2^2$	$F_{exp} < F_{n_2,n_1;1-\alpha}$

- Test para comparar las varianzas de dos distribuciones normales con medias desconocidas: Fijado un nivel de significación, α, se calcula el valor del estadístico

$$F = \frac{\sigma_1^2}{\sigma_2^2} \cdot \frac{S_2^2}{S_1^2} \rightsquigarrow F_{n_2-1,n_1-1} \Rightarrow F_{exp} = \frac{s_2^2}{s_1^2}$$

Contraste	Región Crítica Se rechaza H_0 a nivel α si
$H_0 : \sigma_1^2 = \sigma_2^2$ $H_1 : \sigma_1^2 \neq \sigma_2^2$	$F_{exp} < F_{n_2-1,n_1-1;1-\alpha/2}$ o $F_{exp} > F_{n_2-1,n_1-1;\alpha/2}$
$H_0 : \sigma_1^2 \leq \sigma_2^2$ $H_1 : \sigma_1^2 > \sigma_2^2$	$F_{exp} > F_{n_2-1,n_1-1;\alpha}$
$H_0 : \sigma_1^2 \geq \sigma_2^2$ $H_1 : \sigma_1^2 < \sigma_2^2$	$F_{exp} < F_{n_2-1,n_1-1;1-\alpha}$

7.6. Dualidad entre estimación por intervalos y contraste de hipótesis

Se puede establecer una relación entre los intervalos de confianza para un problema y la región de aceptación si se resuelve el problema mediante un test.

Sea $X \rightsquigarrow \{F_\theta; \theta \in \Theta\}$ y $X_1, \ldots, X_n$ una m.a.s. de X. Según se ha visto en el tema anterior un problema de estimación puede resolverse mediante, un intervalo de confianza, ó de forma más general mediante una región de confianza, $(S(x_1, \ldots, x_n))$, al nivel de confianza $1 - \alpha$, que es un subconjunto del espacio paramétrico tal que verifica:

$$P_\theta[S(X_1, \ldots, X_n) \ni \theta] \geq 1 - \alpha \quad \forall \theta \in \Theta$$

Por otro lado, se ha visto en este tema que si se plantea el problema de estimación como un contraste de hipótesis del tipo $H_0 : \theta = \theta_0$ (frente a $H_1 : \theta \neq \theta_0$,), al nivel de significación α, el problema se resuelve dando un test ó, equivalentemente, dando una región de rechazo, $\mathcal{C}(\theta_0)$. Dicha región es un subconjunto del espacio de posibles valores de la muestra $\mathcal{X}^n$, de forma que:

$$P_{\theta_0}[(X_1, \ldots, X_n) \in \mathcal{C}(\theta_0)] = (\text{Error de tipo I}) \leq \alpha$$

El test también puede expresarse en función de la región complementaria a la de rechazo, es decir la región de aceptación que podemos denotar $A(\theta_0)$:

$$\varphi_{\theta_0}(x_1, \ldots, x_n) = \begin{cases} 1 & \text{si } (x_1, \ldots, x_n) \notin \mathcal{A}(\theta_0) \\ 0 & \text{si } (x_1, \ldots, x_n) \in \mathcal{A}(\theta_0) \end{cases}$$

$$P_{\theta_0}[(X_1, \ldots, X_n) \in A(\theta_0)] \geq 1 - \alpha.$$

Teorema: Cada uno de los test $\varphi_{\theta_0}(X_1,\ldots,X_n)$ aplicado al problema de contrastar $H_0: \theta=\theta_0$ (frente a $H_1: \theta\neq\theta_0$,) tiene nivel de significación α si y sólo si $S(X_1,\ldots,X_n)$ es una región de confianza para θ al nivel de confianza $1-\alpha$.

Es decir, existe una relación entre ambas formas de resolver el problema de estimación.

- Contraste de Hipótesis $\rightarrow$ Intervalo de Confianza

  Si se sabe resolver el problema de contraste de hipótesis, del tipo planteado anteriormente, para cada valor $\theta\in\Theta$, entonces se conocen conjuntos del tipo $A(\theta)$ verificado $P_{\theta_0}[(X_1,\ldots,X_n)\in A(\theta_0)]\geq 1-\alpha$. Se puede construir un intervalo de confianza, al nivel de confianza $1-\alpha$, como sigue:

  $$S(x_1,\ldots,x_n)=\{\theta\in\Theta:\ (x_1,\ldots,x_n)\in A(\theta)\}.$$

- Intervalo de Confianza $\rightarrow$ Contraste de Hipótesis

  Si se sabe resolver el problema mediante un intervalo de confianza, $S(x_1,\ldots,x_n)$, al nivel de confianza $1-\alpha$ y se desea resolver el problema de contraste, del tipo planteado anteriormente, se construye la región de aceptación, al nivel de significación α, de la siguiente forma:

  $$A(\theta_0)=\{(x_1,\ldots,x_n)\in\mathcal{X}^n:\ S(x_1,\ldots,x_n)\ni\theta_0\}$$

Ejemplo: Sea $X\rightsquigarrow\mathcal{N}(\mu,\sigma_0^2)$. Obtener el intervalo de confianza para μ al nivel de confianza $1-\alpha$ a partir del test de razón de verosimilitud para el contraste $\left\{\begin{array}{ll} H_0: & \mu=\mu_0 \\ H_1: & \mu\neq\mu_0\end{array}\right.$.

Siguiendo un razonamiento análogo al del ejemplo se puede estudiar la dualidad entre los intervalos de confianza para poblaciones normales, estudiados en el tema anterior, y los contrastes de hipótesis para esas mismas poblaciones, estudiados en este tema (ver pág. 13 del resumen).

Para los contrastes de igualdad de medias o de varianzas estudiados, resueltos por el test de razón de verosimilitud, se pueden encontrar también sus duales (ver pág. 14 del resumen). Por ejemplo, si se contrasta $\left\{\begin{array}{ll} H_0: & \mu_1=\mu_2 \\ H_1: & \mu_1\neq\mu_2\end{array}\right.$, este contraste es equivalente al contraste $\left\{\begin{array}{ll} H_0: & \mu_1-\mu_2=0 \\ H_1: & \mu_1-\mu_2\neq 0\end{array}\right.$, o de forma más general $\left\{\begin{array}{ll} H_0: & \mu_1-\mu_2=\delta \\ H_1: & \mu_1-\mu_2\neq\delta\end{array}\right.$, con $\delta\in\mathbb{R}$, el cual se pueden resolver rechazando la hipótesis nula si δ no pertenece al intervalo de confianza de menor longitud esperada para la diferencia de medias. De forma equivalente, si se contrasta $\left\{\begin{array}{ll} H_0: & \sigma_1^2=\sigma_2^2 \\ H_1: & \sigma_1^2\neq\sigma_2^2\end{array}\right.$, este contraste es equivalente a contrastar

$\left\{\begin{array}{ll} H_0: & \sigma_1^2/\sigma_2^2 = 1 \\ H_1: & \sigma_1^2/\sigma_2^2 \neq 1 \end{array}\right.$ ó $\left\{\begin{array}{ll} H_0: & \sigma_2^2/\sigma_1^2 = 1 \\ H_1: & \sigma_2^2/\sigma_1^2 \neq 1 \end{array}\right.$, los cuales pueden resolverse rechazando la hipótesis nula si 1 no pertenece al intervalo de confianza de menor longitud esperada para el cociente de varianzas adecuado.

Para los contrastes de hipótesis con alternativas unilaterales, es decir, para los contrastes del tipo $\left\{\begin{array}{ll} H_0: & \theta \geq \theta_0 \\ H_1: & \theta < \theta_0 \end{array}\right.$ ó $\left\{\begin{array}{ll} H_0: & \theta \leq \theta_0 \\ H_1: & \theta > \theta_0 \end{array}\right.$ también se pueden obtener los intervalos de confianza duales al problema de resolver el contraste de hipótesis planteado, sólo que en este caso los intervalos serán de tipo unilateral, $(a, +\infty)$ ó $(-\infty, b)$ (ver pág. 13 y 14 del resumen).

Ejemplo: Sea $X \rightsquigarrow \mathcal{N}(\mu, \sigma_0^2)$. Obtener el intervalo de confianza para μ al nivel de confianza $1-\alpha$ a partir del test de razón de verosimilitud para el contraste $\left\{\begin{array}{ll} H_0: & \mu \geq \mu_0 \\ H_1: & \mu < \mu_0 \end{array}\right.$.

7.7. Resolución de contrastes de hipótesis mediante el $p - valor$

Otra opción para resolver un problema de contraste hipótesis mediante un test no aleatorizado es usar el denominado $p - nivel$ o $p - valor$ asociado a los datos. Este método no depende del nivel de significación para el cual se desea resolver el contraste lo que lo diferencia de usar la región de rechazo ya que para determinar esta hace falta dicho valor.

Definición: El $p - valor$ es el mínimo nivel de significación para el cual el valor del estadístico en la muestra cae en la región de rechazo siendo la hipótesis nula verdadera.

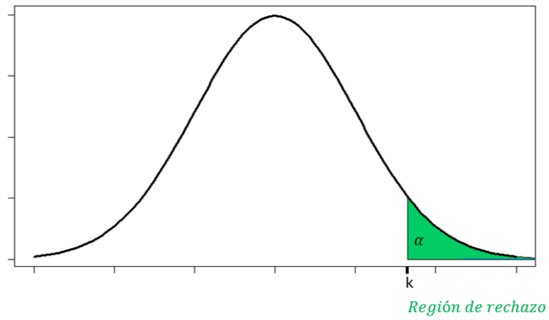
Si $T(X_1, \ldots, X_n)$ es el estadístico en función del cual está definido el test que resuelve el contraste y $T_{exp} = T(x_1, \ldots, x_n)$ es el valor del estadístico en la muestra observada, el $p - valor$ se determina mediante el siguiente cálculo:

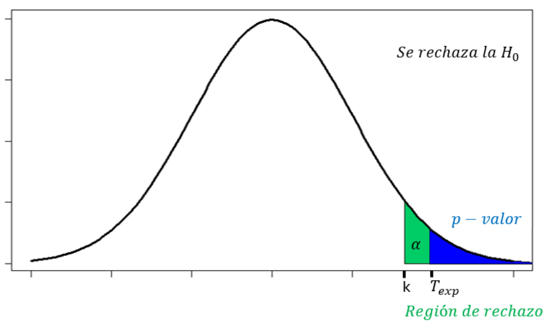
- Si la región de rechazo es de cola a la derecha,

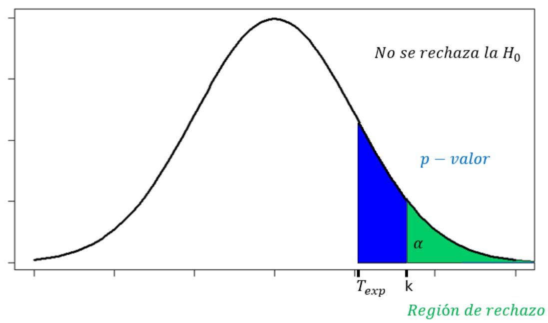
$$T_{exp} \geq k \text{ con } P[T(X_1, \ldots, X_n) \geq k] \leq \alpha,$$

se verifica

$$p - valor = P_{H_0}(T \geq T_{exp}).$$





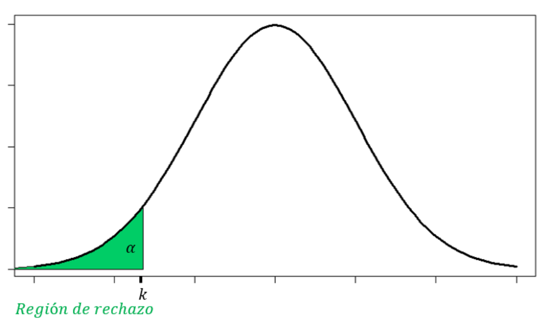


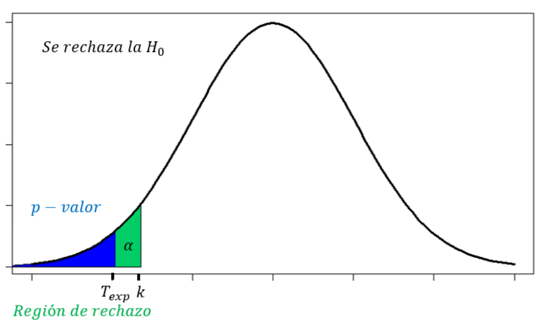
- Si la región de rechazo es de cola a la izquierda,

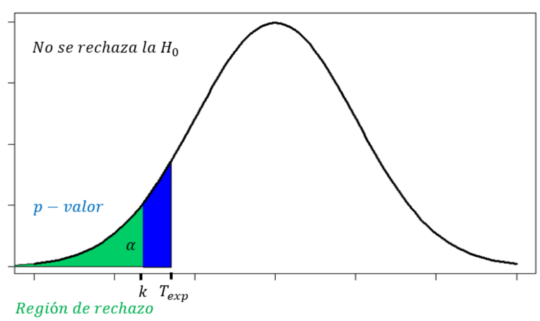
$$T_{exp} \leq k \text{ con } P[T(X_1, \ldots, X_n) \leq k] \leq \alpha,$$

se verifica

$$p - valor = P_{H_0}(T \leq T_{exp}).$$







- Si la región de rechazo es de dos colas,

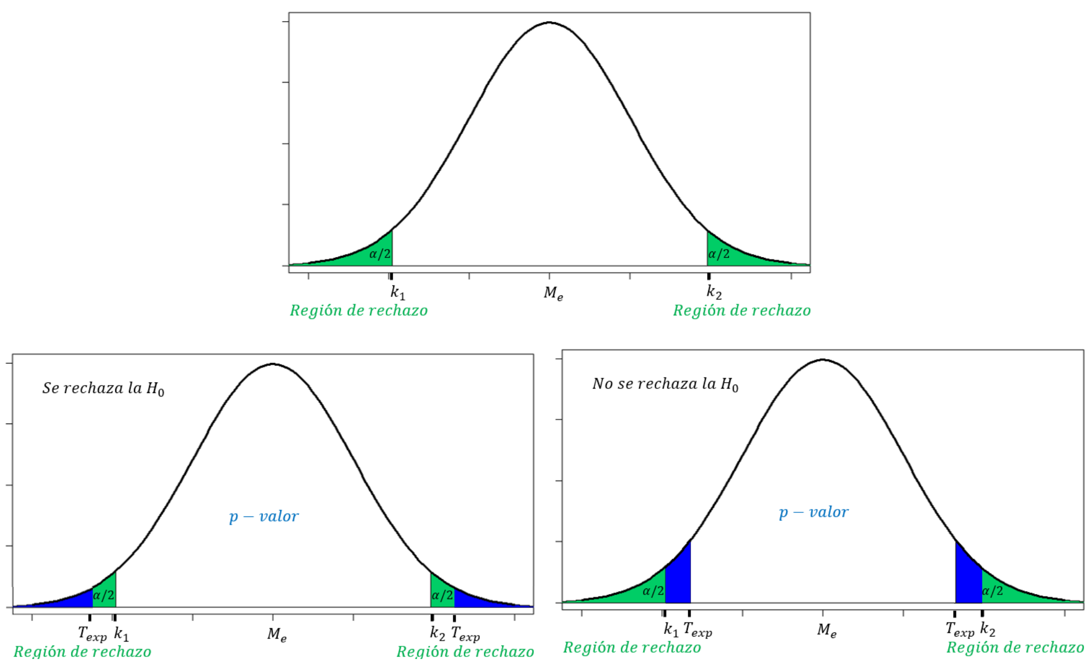
$$T_{exp} \leq k_1 \text{ o } T_{exp} \geq k_2, \text{ con } P[T(X_1,\ldots,X_n) \leq k_1] + P[T(X_1,\ldots,X_n) \geq k_2] \leq \alpha$$

se verifica

$$p - valor = \begin{cases} 2P_{H_0}[T(X_1,\ldots,X_n) \geq T_{exp}] & \text{si } T_{exp} \geq Me \\ 2P_{H_0}[T(X_1,\ldots,X_n) \leq T_{exp}] & \text{si } T_{exp} \leq Me \end{cases}$$

siendo Me la mediana de la distribución del estadístico bajo hipótesis nula

$$P_{H_0}[T(X_1,\ldots,X_n) \leq Me] = \frac{1}{2}.$$



En general la forma de interpretar el $p - valor$ es que para niveles de significación menores que el p-valor, el test no rechaza H_0 mientras que para niveles de significación mayores o iguales, el test rechaza H_0.

En la práctica, la conclusión del contraste en base al $p - valor$ sería la siguiente:

- Si el $p - valor$ es grande (mayor a 0.1), no se rechaza H_0, ya que para rechazar en base al él se deben tomar niveles de significación muy altos, es decir, asumir un error de tipo I grande. En ese caso se suele indicar que el contraste no es significativo.

- Si el $p - valor$ es pequeño (menor a 0.01), se rechaza H_0, ya que para los niveles de significación más estándar (0.01, 0.05 y 0.1) el $p - valor$ quedaría por debajo. En ese caso se suele indicar que el contraste es significativo para niveles de significación mayores o iguales a 0.01.

- Para valores intermedios, hay que tener en cuenta el nivel de significación con que se quiere trabajar.

  Si se toma un nivel de significación de 0.05, para rechazar H_0 se debe obtener un $p - valor$ inferior a dicho nivel, es decir si el $p - valor$ $<$0.05 el test rechazaría H_0. Si el $p - valor$ está entre 0.05 y 0.1, no se puede rechazar la H_0 pero como el $p - valor$ está cerca del valor de comparación, se suele indicar que el contraste está próximo a la significación. Si el $p - valor$ $>$0.1 el test no rechazaría H_0.

  Hay que tratar cada situación en particular aunque, normalmente, lo aconsejable en los casos donde se obtenga un $p - valor$ próximo a la significación por encima, sería poder tomar más datos y rehacer los cálculos para que el test pueda ser concluyente.