

XS3310 Teoría Estadística

I Semestre 2021

Escuela de Estadística

2021-05-31

class: center, middle

¿Qué hemos visto hasta ahora?

Todo sobre estimadores puntuales + pivotes e intervalos de confianza. Bootstrap y contrastes de hipótesis (función de potencia, tamaño del contraste, el valor p).

¿Qué vamos a discutir hoy?

Contrastes de hipótesis: Dualidad en pruebas de hipótesis, contrastes más potentes y uniformemente más potentes. Cocientes de verosimilitud.

Dualidad en pruebas

Suponga que se tiene una prueba de hipótesis y se encuentra la región crítica (o región de rechazo) RC_δ de un contraste δ sobre θ con nivel α .

Entonces un intervalo de confianza con nivel α de este parámetro, está dado por RC_δ^c .

Ejemplo Supongamos que tenemos datos X_1, \dots, X_n con distribución Bernoulli con parámetro p .

Tenemos la hipótesis

$$H_0 : p < p_0 \text{ vs } H_1 : p \geq p_0.$$

El criterio de rechazo al nivel α es

$$Y = \sum_{i=1}^n X_i \geq c.$$

El valor c debe ser el menor valor que cumpla que

$$\Pr(Y \geq c \mid p = p_0) \leq \alpha_0$$

Suponga que $n = 10$, $p_0 = 0.3$, $Y = 6$ y $\alpha = 0.1$. Los valores de $P(Y > y)$ son

```
x <- matrix(pbinom(0:10, size = 10, prob = 0.3, lower.tail = FALSE), nrow = 1)
colnames(x) <- 0:10
knitr::kable(x, digits = 4)
```

0	1	2	3	4	5	6	7	8	9	10
0.9718	0.8507	0.6172	0.3504	0.1503	0.0473	0.0106	0.0016	1e-04	0	0

La regla decisión es “Rechazamos H_0 si $Y \geq c$ ”. En este caso rechazamos si $Y \geq 5$.

Entonces el intervalo de confianza sería la región donde $Y < 5$ o donde $p - \text{valor} > \alpha$.

En este caso tenemos

$$[0.01503, 1]$$

Ejemplo. $X = (X_1, \dots, X_n) \sim N(\mu, \sigma^2)$, $\theta = (\mu, \sigma^2)$ desconocido. Queremos probar

$$H_0 : \mu \leq \mu_0 \text{ vs } H_1 : \mu > \mu_0.$$

Por dualidad, basta con conocer un IC unilateral para μ :

$$\left(\bar{X}_n - t_{n-1, \alpha} \frac{s}{\sqrt{n}}, \infty \right).$$

Rechazamos H_0 si

$$\mu_0 \leq \bar{X}_n - t_{n-1, \alpha} \frac{s}{\sqrt{n}} \Leftrightarrow T = \frac{\bar{X}_n - \mu_0}{s/\sqrt{n}} \geq t_{n-1, \alpha}$$

(rechazando en la cola derecha de T).

Contrastes más potentes

- Es deseable encontrar un contraste donde las probabilidades $\alpha(\delta)$ y $\beta(\delta)$ sean lo más pequeñas posibles.
- Es posible que uno de los dos errores tenga probabilidad cero, por ejemplo tener un contraste donde H_0 se rechace siempre haciendo que la probabilidad de Error Tipo I siempre sea cero. No obstante, este contraste hace que la probabilidad de cometer un Error Tipo II sea siempre 1, lo que es indeseable.
- Es imposible encontrar un contraste que minimice ambas probabilidades de error por lo que nos enfocamos en minimizar una combinación lineal de estas.

Contrastes más potentes

- Una manera de hacer esto es fijar el tamaño del contraste y encontrar el contraste que tenga la *mayor potencia*.
- Para demostrar el procedimiento, vamos a enfocarnos en un tipo de contraste de hipótesis que contrasta una hipótesis simple contra otra hipótesis simple. Una hipótesis simple y una hipótesis compuesta se definen de la siguiente manera:

Hipótesis simple e hipótesis compuesta: Sea X_1, X_2, \dots, X_n una muestra aleatoria de una distribución con parámetro θ . Se dice que una hipótesis es **simple** si *especifica de manera única* la distribución de la población. Cualquier hipótesis que no sea simple se denomina **hipótesis compuesta**.

Contrastes más potentes

Supongamos que tenemos una muestra aleatoria donde $X_j \sim N(\mu, 1)$ y tenemos las hipótesis $H_0 : \mu = 0$ y $H_1 : \mu \neq 0$. En este caso la hipótesis nula nos dice que los datos pertenecen a una población $N(0, 1)$ mientras que la hipótesis alterna no especifica la distribución de manera única, por lo que H_0 es una *hipótesis simple* mientras que H_1 es una *hipótesis compuesta*.

Supongamos ahora que $X_j \sim N(\mu, \sigma^2)$ donde $\sigma^2 > 0$ y es desconocido. En este caso la hipótesis nula $H_0 : \mu = 0$ es una *hipótesis compuesta* ya que no está especificando la distribución de la población de una única manera (depende de σ^2 , el cual puede ser cualquier valor positivo).

Contrastes más potentes

Volviendo al objetivo de encontrar un buen contraste supongamos que queremos contrastar dos hipótesis simples y que fijamos el tamaño del contraste en un valor α_0 dado. De todos los posibles contrastes que existan con el mismo tamaño α_0 , el **contraste más potente** va a ser aquel que tenga la mayor potencia en Ω_1 . Este valor α_0 fijo usualmente se conoce como el **nivel de significancia**.

Contraste más potente: Suponga que se desean contrastar las hipótesis simples $H_0 : \theta = \theta_0$ contra $H_1 : \theta = \theta_1$ basados en una muestra aleatoria de una población con distribución que depende de θ . Para un valor $\alpha_0 \in]0, 1[$ sea $\mathbb{T}_{\alpha_0} = \{\delta | \alpha(\delta) = \alpha_0\}$ el conjunto de todos los contrastes δ que tienen tamaño α_0 . Entonces el contraste de tamaño α_0 **más potente**, denotado δ^* , satisface:

- $\delta^* \in \mathbb{T}_{\alpha_0}$
 - $\beta(\delta^*) \leq \beta(\delta), \forall \delta \in \mathbb{T}_{\alpha_0}$
-

Contrastes más potentes

El punto de este tema es que al enfocarse en solo dos valores para θ y fijar el tamaño del contraste en α_0 entonces solo tenemos que concentrarnos en la potencia en θ_1 . En general y en la práctica es raro fijar una hipótesis alterna que sea simple por lo que tendríamos que comparar toda la función potencia en Ω_1 entre distintos contrastes, lo que es considerablemente más complicado de hacer que solo comparar dos valores. No obstante, más adelante en el tema veremos una forma de lograr esto para ciertos contrastes.

Lema de Neyman-Pearson

Una manera de encontrar la región crítica para un contraste más potente δ^* se da por medio del Lema de Neyman-Pearson. Este dice lo siguiente:

Lema de Neyman-Pearson: Supongamos que se desea contrastar las hipótesis simples $H_0 : \theta = \theta_0$ contra $H_1 : \theta = \theta_1$ basados en una muestra aleatoria de una población con distribución que depende de θ . Además, sea $\mathcal{L}(\theta)$ la función de verosimilitud de la muestra para un valor de θ . Para un tamaño dado α_0 , el contraste que minimiza $\beta(\delta)$ tiene una región crítica determinada por: $RC_{\delta^*} = \left\{ \mathbf{X} \mid \frac{\mathcal{L}(\theta_0)}{\mathcal{L}(\theta_1)} < k \right\}$

De acuerdo con el teorema anterior, se rechaza H_0 si $\mathcal{L}(\theta_0) < k\mathcal{L}(\theta_1)$ y se acepta H_0 si $\mathcal{L}(\theta_0) > k\mathcal{L}(\theta_1)$. En el caso en que $\mathcal{L}(\theta_0) = k\mathcal{L}(\theta_1)$ se puede favorecer cualquiera de las dos hipótesis. El valor de k se selecciona de manera que el contraste tenga tamaño α_0 . Este contraste se denomina como **contraste más potente** para un tamaño α_0 , pues es la prueba con un valor máximo de Potencia(θ_1).

Contrastes más potentes

Podemos observar que un contraste con esta región crítica tiene sentido con la interpretación de la función de verosimilitud. Recordemos que entre más verosímil es un valor de θ para la muestra dada, mayor es su función de verosimilitud. Por lo tanto, si $\theta = \theta_0$ entonces se espera que $\mathcal{L}(\theta_0) > \mathcal{L}(\theta_1)$, favoreciendo a H_0 sobre H_1 y viceversa. De esta manera el contraste más potente consiste en escoger la hipótesis más verosímil para la muestra obtenida, fijando de antemano una máxima probabilidad de cometer Error Tipo I.

Ejemplo: Suponga que Y representa una sola observación con función de densidad dada por:

$$f_Y(y) = \begin{cases} \theta y^{\theta-1} & \text{si } 0 \leq y \leq 1 \\ 0 & \text{en otros casos} \end{cases}$$

Determine el contraste más potente con un tamaño $\alpha_0 = 0.05$, para contrastar $H_0 : \theta = 2$ contra $H_1 : \theta = 1$.

Contrastes más potentes

Solución: Por el Lema de Neyman-Pearson sabemos que se rechaza H_0 si $\frac{\mathcal{L}(2)}{\mathcal{L}(1)} < k$. Debemos encontrar esta expresión para el ejemplo que tenemos. En este caso la verosimilitud viene dada por: $\mathcal{L}(\theta) = \theta y^{\theta-1}$

Por lo tanto, al evaluar en Ω_0 y Ω_1 tenemos:

$$\mathcal{L}(2) = 2y, \quad \mathcal{L}(1) = 1, \quad \Rightarrow \frac{\mathcal{L}(2)}{\mathcal{L}(1)} = 2y$$

Por lo tanto, rechazamos H_0 si $2Y < k$, lo que equivale a decir que rechazamos si $Y < k'$, donde $k' = k/2$. Ahora debemos encontrar el valor de k' para que el contraste tenga un tamaño de 0.05.

$$P(Y < k' | \theta = 2) = F_Y(k' | \theta = 2) = k'^2 = 0.05, \quad \Rightarrow k' = \sqrt{0.05}$$

Podemos concluir que el contraste más potente consiste en rechazar H_0 si $Y < \sqrt{0.05}$, con un nivel de significancia de 0.05. Observen que no es necesario encontrar el valor de k exacto, ya que el tamaño del contraste va a seguir siendo igual. Solo queremos encontrar una regla de decisión para el contraste.

Contrastes más potentes

Ejemplo: Sea X_1, X_2, \dots, X_n una muestra aleatoria tal que $X_j \sim N(\mu, 1)$. Se desea contrastar las hipótesis $H_0 : \mu = \mu_0$ contra $H_1 : \mu = \mu_1$, donde $\mu_0 < \mu_1$. Obtenga el contraste más potente para un tamaño dado α_0 .

Solución: Recordemos que para esta muestra normal la verosimilitud viene dada por la siguiente expresión:

$$\mathcal{L}(\mu) = (2\pi)^{-\frac{n}{2}} e^{-\frac{\sum (x_j - \mu)^2}{2}} = (2\pi)^{-\frac{n}{2}} e^{-\frac{\sum x_j^2 - 2n\mu\bar{x} + n\mu^2}{2}}$$

Por lo tanto cuando obtenemos el cociente de las verosimilitudes para encontrar la región crítica según el Lema de Neyman-Pearson obtenemos:

$$\frac{\mathcal{L}(\mu_0)}{\mathcal{L}(\mu_1)} = \frac{e^{\frac{n\mu_0\bar{x} - \frac{n\mu_0^2}{2}}{2}}}{e^{\frac{n\mu_1\bar{x} - \frac{n\mu_1^2}{2}}{2}}} = e^{n(\bar{x}(\mu_0 - \mu_1) - \frac{1}{2}(\mu_0^2 - \mu_1^2))} = e^{n\bar{x}(\mu_0 - \mu_1)} \cdot e^{-\frac{n}{2}(\mu_0^2 - \mu_1^2)} < k$$

Contrastes más potentes

A partir de esta expresión debemos encontrar el valor crítico de la prueba, ya sabemos que rechazamos H_0 si:

$$e^{n\bar{x}(\mu_0 - \mu_1)} \cdot e^{-\frac{n}{2}(\mu_0^2 - \mu_1^2)} < k \Rightarrow e^{n\bar{x}(\mu_0 - \mu_1)} < k e^{\frac{n}{2}(\mu_0^2 - \mu_1^2)} \Rightarrow n\bar{x}(\mu_0 - \mu_1) < \ln \left(k e^{\frac{n}{2}(\mu_0^2 - \mu_1^2)} \right) \Rightarrow \bar{x} > \frac{\ln \left(k e^{\frac{n}{2}(\mu_0^2 - \mu_1^2)} \right)}{n(\mu_0 - \mu_1)} = k'$$

Por lo tanto, nuestro contraste consiste en rechazar H_0 si $\bar{X} > k'$. Ahora debemos encontrar el valor de k' para un tamaño dado α_0 :

$$\alpha_0 = P(\bar{X} > k' | \mu = \mu_0) = P(Z > \sqrt{n}(k' - \mu_0))$$

Usando el software adecuado podemos encontrar que $z_{1-\alpha_0} = \sqrt{n}(k' - \mu_0)$. Por lo tanto podemos despejar el valor de k' y encontrar que queda: $k' = \mu_0 + \frac{z_{1-\alpha_0}}{\sqrt{n}}$, y por lo tanto, nuestro contraste más potente consiste en rechazar H_0 si $\bar{X} > \mu_0 + \frac{z_{1-\alpha_0}}{\sqrt{n}}$ para un nivel de significancia α_0 .

Contrastes uniformemente más potentes

Una observación interesante que podemos realizar del ejemplo anterior es que este no depende realmente del valor de la hipótesis alterna, solo depende del valor de la hipótesis nula y de la relación entre ambos valores; esto es un patrón común con los contrastes obtenidos mediante el Lema de Neyman-Pearson. Por lo tanto, en ciertos casos, nos es posible generalizar el resultado de este tipo de contrastes para más de un valor en H_1 .

Supongamos que se desea probar las hipótesis $H_0 : \theta = \theta_0$ contra $H_1 : \theta > \theta_0$. En este caso parece que no podemos aplicar el Lema de Neyman-Pearson pues la hipótesis alterna es compuesta. Sin embargo, podemos redefinir la hipótesis alterna como $H_1 : \theta = \theta_1$, donde $\theta_1 > \theta_0$. De esta manera tenemos un contraste de hipótesis simples al que le podemos aplicar el Lema de Neyman-Pearson y que donde ya vimos que el valor de θ_1 no se utiliza en la definición de la región crítica. Podemos definir esto de la siguiente manera:

Contrastes uniformemente más potentes

Contraste uniformemente más potente: Un contraste de tamaño α_0 es **uniformemente más potente** (UMP) para contrastar $H_0 : \theta = \theta_0$ contra $H_1 : \theta \in \Omega_1$ si es más potente para $H_0 : \theta = \theta_0$ contra $H_1 : \theta = \theta_1$, para todo $\theta_1 \in \Omega_1$.

Bajo esta definición vamos a poder encontrar fácilmente contrastes UMP para las hipótesis alternas $H_1 : \theta > \theta_0$ o $H_1 : \theta < \theta_0$, siempre y cuando la distribución poblacional lo permite. Sin embargo, no es posible encontrar un contraste UMP para la hipótesis alterna $H_1 : \theta \neq \theta_0$ ya que será imposible encontrar el mismo contraste más potente para todo valor de θ_1 diferente de θ_0 .

Contrastes uniformemente más potentes

Ejemplo: Sea X_1, X_2, \dots, X_n una muestra aleatoria tal que $X_j \sim N(\mu, 1)$. Se desea contrastar las hipótesis $H_0 : \mu = \mu_0$ contra $H_1 : \mu > \mu_0$, con un nivel de significancia de α_0 .

Solución: Empecemos por definir un valor de μ_1 tal que $\mu_1 > \mu_0$. De esta forma podemos re-escribir la hipótesis alterna como $H_1 : \mu = \mu_1, \forall \mu_1 > \mu_0$. Una vez hecho esto el ejercicio se vuelve idéntico al anterior y por lo tanto podemos concluir que el contraste UMP consiste en rechazar H_0 si $\bar{X} > \mu_0 + \frac{z_{1-\alpha_0}}{\sqrt{n}}$ para un nivel de significancia α_0 .

Cuando se da el caso de que exista un contraste UMP (que será el caso para los ejemplos vistos en clases), existe una definición y un teorema que nos facilita la obtención del contraste.

Cociente de verosimilitudes monótono: Una muestra aleatoria con función de verosimilitud $\mathcal{L}(\theta)$ tiene un **cociente de verosimilitudes monótono** (CVM) en un estadístico $T = T(X_1, X_2, \dots, X_n)$ si, para $\theta_0 < \theta_1$, el cociente $\frac{\mathcal{L}(\theta_0)}{\mathcal{L}(\theta_1)}$ es una función monótona de T .

Contrastes uniformemente más potentes

Ejemplo: Sea X_1, X_2, \dots, X_n una muestra aleatoria tal que $X_j \sim N(\mu, 1)$. Demuestre que $\mathcal{L}(\mu)$ tiene un CVM en el estadístico \bar{X} .

Solución: Sea $\mu_0 < \mu_1$. Anteriormente habíamos trabajado una expresión para este cociente de verosimilitudes:

$$\frac{\mathcal{L}(\mu_0)}{\mathcal{L}(\mu_1)} = e^{n\bar{x}(\mu_0 - \mu_1)} \cdot e^{-\frac{1}{2}(\mu_0^2 - \mu_1^2)}$$

Lo que nos interesa es demostrar que esta es una función monótona de \bar{X} . Observemos que la cantidad $e^{-\frac{1}{2}(\mu_0^2 - \mu_1^2)}$ no depende de \bar{X} , por lo que puede ser tratada como una constante. Por otro lado la expresión $e^{n\bar{x}(\mu_0 - \mu_1)}$ es una función de \bar{X} , y sabemos que la exponencial es siempre creciente si su exponente es positivo o siempre decreciente si su exponente es negativo; no obstante, indiferentemente del signo del exponente, esta va a ser siempre una función monótona para \bar{X} .

Por lo tanto, concluimos que $\mathcal{L}(\mu)$ tiene un CVM en \bar{X} .

Contrastes uniformemente más potentes

Si la distribución tiene un CVM en el estadístico T , entonces el contraste UMP para $H_0 : \theta = \theta_0$ contra $H_1 : \theta > \theta_0$ o $H_1 : \theta < \theta_0$ existe y este se puede expresar en términos de T , en lugar del cociente $\frac{\mathcal{L}(\theta_0)}{\mathcal{L}(\theta_1)}$.

Este teorema nos facilita la obtención del contraste UMP pues nos garantiza que la regla de decisión va a depender de T . Además, al ser el cociente de verosimilitudes una función monótona de T entonces existe una relación directa entre la hipótesis alternativa y la regla de decisión. Básicamente si $\mathcal{L}(\theta)$ tiene CVM en T y tenemos $H_1 : \theta > \theta_0$ entonces el teorema anterior nos garantiza que rechazamos H_0 si $T > c$, mientras que si $H_1 : \theta < \theta_0$ entonces rechazamos H_0 si $T < c$. El valor de c se obtiene al igual que anteriormente, fijando el tamaño del contraste en α_0 .

Contrastes uniformemente más potentes

Ejemplo: Sea X_1, X_2, \dots, X_n una muestra aleatoria tal que $X_j \sim N(\mu, 1)$. Se desea contrastar las hipótesis $H_0 : \mu = \mu_0$ contra $H_1 : \mu < \mu_0$. Obtenga el contraste UMP para un tamaño dado α_0 .

Solución: Anteriormente se demostró que $\mathcal{L}(\mu)$ tiene un CVM en \bar{X} por lo que por el teorema anterior sabemos que vamos a rechazar H_0 si $\bar{X} < c$. Procedemos a encontrar el valor de c :

$$\alpha_0 = P(\bar{X} < c | \mu = \mu_0) = P(Z < \sqrt{n}(c - \mu_0))$$

Utilizando software podemos encontrar que $\sqrt{n}(c - \mu_0) = z_{\alpha_0} = -z_{1-\alpha_0}$. Despejando obtenemos que $c = \mu_0 - \frac{z_{1-\alpha_0}}{\sqrt{n}}$. Por lo tanto, el contraste UMP consiste en rechazar H_0 si $\bar{X} < \mu_0 - \frac{z_{1-\alpha_0}}{\sqrt{n}}$.

class: center, middle

¿Qué discutimos hoy?

Contrastes de hipótesis - contrastes más potentes, uniformemente más potentes, razón de verosimilitud.