MAT-032: Test de hipótesis

Felipe Osorio

fosorios.mat.utfsm.cl

Departamento de Matemática, UTFSM

Decidir lanzando una moneda[1]

[1] Harvey Dent, o “Dos Caras”, un enemigo de Batman.

Ejemplo de juguete: Problema de Monty Hall[2]

Puerta N° 1

Puerta N° 2

Puerta N° 3

[2] conocido también como "el Problema del presentador"

Problema de Monty Hall: ¿Debe cambiar su elección?[3]

Puerta N° 1

Puerta N° 2

Puerta N° 3

[3] ¿Que opinan? La respuesta NO es intuitiva...

Problema de Monty Hall: ¿Debe cambiar su elección?[4]

Puerta N° 1

Puerta N° 2

Puerta N° 3

[4] ¿Quiere ganar el auto? hagamos unos **(pocos)** cálculos...

- Para el problema original (3 puertas cerradas), considere:

$$C = \{\text{abrimos la puerta que tiene el auto}\}.$$

$$\overline{C} = \{\text{abrimos la puerta que tiene una cabra}\},$$

  es fácil notar que

$$\mathsf{P}(C) = \frac{1}{3}, \qquad \mathsf{P}(\overline{C}) = 1 - \frac{1}{3} = \frac{2}{3}.$$

- El presentador abre la puerta N° 3, y hay una cabra... ☹
- Una vez que se ha abierto la puerta, ¿Cuál es la probabilidad de ganar el auto?
- Pregunta: ¿Debemos cambiar nuestra elección inicial?

EX UMBRA IN SOLEM

Problema de Monty Hall: Sobre como ganar el auto...

- Aunque parezca extraño: ¡El presentador **nos está ayudando**!
- Abrir una puerta, modifica las probabilidades de ganar el auto... ☺ en nuestro beneficio.[5]
- En el problema modificado (presentador abre una puerta), considere:

$$A = \{\text{Ud. elige la puerta con el premio } \textbf{antes} \text{ de cambiar de opción}\},$$

$$B = \{\text{Ud. elige la puerta con el premio } \textbf{después} \text{ de cambiar de opción}\}.$$

- Usando el Teorema de probabilidad total,

$$\begin{aligned} \mathsf{P}(A) &= \mathsf{P}(B \cap A) + \mathsf{P}(B \cap \overline{A}) = \mathsf{P}(B|A)\,\mathsf{P}(A) + \mathsf{P}(B|\overline{A})\,\mathsf{P}(\overline{A}) \\ &= 0 \cdot \frac{1}{3} + 1 \cdot \frac{2}{3} = \frac{2}{3}, \end{aligned}$$

es decir, tenemos un 66.6% de chance de ganar un auto!

[5] Aunque no nos demos cuenta, esta información es **muy relevante**!

Problema de Monty Hall: Comentarios

- Este es un problema donde la intuición nos engaña.
- Tomamos una decisión (cambiar o no de puerta) en base al cálculo de probabilidades.
- Note que, aún podemos equivocarnos...
  (hay un 33.3% de chances de ganar una cabra!)

En Estadística:

Deseamos usar los datos disponibles para tomar mejores decisiones.

Un poco de controversia: Resultado de un juicio penal

Parecer de un juez:

¿La evidencia dada por el fiscal es suficiente para declarar culpable al acusado?

El juez tiene 2 opciones:

- Declarar al acusado culpable.
- Declarar al acusado inocente.

En términos científicos (o estadísticos)[6], debemos plantear las siguientes hipótesis:

$$H_0 : \text{el acusado es inocente.}$$
$$H_1 : \text{el acusado es culpable.}$$

Así, el fiscal debe probar que el acusado es culpable, *más allá de toda duda razonable*.

[6] Estadísticos suelen considerar $H_0 : \theta = \theta_0$ vs. $H_1 : \theta = \theta_1$

Parecer de un juez:

¿La evidencia dada por el fiscal es suficiente para declarar culpable al acusado?

El juez tiene 2 opciones:

- Declarar al acusado culpable.
- Declarar al acusado inocente.

En términos científicos (o estadísticos)[6], debemos plantear las siguientes hipótesis:

$$H_0 : \text{el acusado es inocente.}$$
$$H_1 : \text{el acusado es culpable.}$$

Así, el fiscal debe probar que el acusado es culpable, *más allá de toda duda razonable*.

[6] Estadísticos suelen considerar $H_0 : \theta = \theta_0$ vs. $H_1 : \theta = \theta_1$

Un poco de controversia: Resultado de un juicio penal

De este modo, en el juicio puede ocurrir lo siguiente:

Decisión	El acusado es	
	inocente	culpable
preso	falso positivo	OK
libre	OK	falso negativo

En la nomenclatura de test de hipótesis, tenemos:

Decisión	El acusado es	
	H_0 es verdadero	H_1 es verdadero
rechazar H_0	error tipo I	OK
aceptar H_0	OK	error tipo II

¿Cómo luce una regla de decisión?

Considere el **criterio** para aprobar una asignatura.

Suponga que usted ha obtenido las siguientes notas en (por ejemplo) MAT-032:

$$\boldsymbol{x} = \{x_1, x_2, x_3\} = \{68, 32, 70\}.$$

De este modo, el profesor calcula su promedio obteniendo: $\overline{x} = 56$. Por tanto,

$$\overline{x} \geq 55, \qquad \text{es decir, Ud. ha aprobado.}$$

Podemos reescribir lo anterior como la siguiente regla de decisión:

- Si $\overline{x} \in [55, 100]$, el alumno es aprobado.
- En caso contrario,[7] el alumno reprueba la asignatura.

[7] Es decir, $\overline{x} \notin [55, 100]$

Objetivo:

Usar la evidencia en los datos para concluir en favor de alguna hipótesis de interés.

Esencialmente, se debe crear una regla de decisión, tal que

$$\alpha = \mathsf{P}(\text{rechazar } H_0 | H_0 \text{ es verdadero}),$$

$$\beta = \mathsf{P}(\text{aceptar } H_0 | H_1 \text{ es verdadero}),$$

además $\pi = 1 - \beta$ es llamado potencia del test.

Observación:

Evidentemente, deseamos crear reglas de decisión tal que α y β sean lo más pequeños posible.[8]

Malas noticias... ☹

Lamentablemente, la minimización de ámbos, α y β es un problema infactible.

[8] y análogamente que tenga una alta potencia $\pi = 1 - \beta$.

- No todo son malas noticias, podemos fijar α (error tipo I) y escoger el test más potente (aquél con menor β)
- La regla de decisión será del tipo

$$\text{Rechazar } H_0 \text{ si:} \quad T(\boldsymbol{X}) \in C,$$

donde C representa la región de rechazo.

Recuerde que:

Test de hipótesis está basado en argumentos probabilísticos...

(es decir, aún podemos equivocarnos!)[9]

[9] Aunque los estadísticos solemos no equivocarmos tanto! ☺

Test de hipótesis

Sea $X_1, \dots, X_n$ muestra aleatoria desde $f(x; \boldsymbol{\theta})$ con $\boldsymbol{\theta} \in \Theta \subset \mathbb{R}^p$. Suponga que deseamos probar

$$H_0 : \boldsymbol{\theta} \in \Theta_0, \qquad \text{versus} \qquad H_1 : \boldsymbol{\theta} \in \Theta_1, \tag{1}$$

donde $\Theta = \Theta_0 \cup \Theta_1$ tal que $\Theta_0 \cap \Theta_1 = \varnothing$.[10]

Sea $0 < \alpha < 1$ y considere

$$L(\boldsymbol{\theta}) = \prod_{i=1}^{n} f(x_i; \boldsymbol{\theta}), \qquad \boldsymbol{\theta} \in \Theta.$$

El test más potente de tamaño α para probar $H_0 : \boldsymbol{\theta} \in \Theta_0$ es dado por el estadístico de razón de verosimilitudes (LR) dado por

$$\Lambda = \frac{\max_{\theta \in \Theta_0} L(\boldsymbol{\theta})}{\max_{\theta \in \Theta} L(\boldsymbol{\theta})} = \frac{L(\widetilde{\boldsymbol{\theta}})}{L(\widehat{\boldsymbol{\theta}})},$$

donde $\widetilde{\boldsymbol{\theta}}$ denota el MLE de $\boldsymbol{\theta}$ sujeto a la restricción $H_0 : \boldsymbol{\theta} \in \Theta_0$. Finalmente, rechazamos H_0 si y sólo si Λ es pequeño $(< k)$.

[10] Por ejemplo, podemos considerar $H_0 : \theta = \theta_0$, versus $H_1 : \theta \neq \theta_0$.

Observación:

Es posible notar que

$$0 < \Lambda < 1,$$

esto permite encontrar un k para rechazar H_0. Además, se suele considerar la estadística

$$\begin{aligned} LR &= 2\log\Lambda = 2\{\log L(\widetilde{\boldsymbol{\theta}}) - \log L(\widehat{\boldsymbol{\theta}})\} \\ &= 2\{\ell(\widetilde{\boldsymbol{\theta}}) - \ell(\widehat{\boldsymbol{\theta}})\}. \end{aligned}$$

Usando argumentos asintóticos, es posible mostrar que

$$LR \xrightarrow{\mathsf{D}} \chi^2(\nu).$$

Otros estadísticos (asintóticamente) equivalentemente al test LR son:

- Test de Wald.
- Test score o de multiplicadores de Lagrange.
- Test gradiente.

Suponga $X_1, \dots, X_n$ variables IID desde $\mathsf{N}(\mu, \sigma^2)$ donde $\mu \in \mathbb{R}$, $\sigma^2 > 0$ con σ^2 conocido. Considere

$$H_0 : \mu = \mu_0, \qquad H_1 : \mu \neq \mu_0,$$

con μ_0 fijo.

En este caso $\Theta_0 = \{\mu_0\}$ y $\Theta = \mathbb{R}$

$$\begin{aligned}\max_{\mu \in \Theta_0} L(\mu) &= (2\pi\sigma^2)^{-n/2} \exp\Big\{ -\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu_0)^2 \Big\} \\ &= L(\mu_0).\end{aligned}$$

Note que

$$\sum_{i=1}^n (x_i - \mu_0)^2 = \sum_{i=1}^n (x_i - \overline{x})^2 + n(\overline{x} - \mu_0)^2.$$

Por otro lado, bajo $H_1 : \mu \neq \mu_0$

$$\begin{aligned}\max_{\mu \in \Theta} L(\mu) &= (2\pi\sigma^2)^{-n/2} \exp\Big\{ -\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \overline{x})^2 \Big\} \\ &= L(\widehat{\mu}).\end{aligned}$$

De este modo, el estadístico LR adopta la forma

$$\begin{aligned}\Lambda = \frac{L(\mu_0)}{L(\widehat{\mu})} &= \frac{(2\pi\sigma^2)^{-n/2} \exp\Big\{ -\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu_0)^2 \Big\}}{(2\pi\sigma^2)^{-n/2} \exp\Big\{ -\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \overline{x})^2 \Big\}} \\ &= \frac{(2\pi\sigma^2)^{-n/2} \exp\Big\{ -\frac{1}{2\sigma^2} \Big[\sum_{i=1}^n (x_i - \overline{x})^2 + n(\overline{x} - \mu_0)^2 \Big] \Big\}}{(2\pi\sigma^2)^{-n/2} \exp\Big\{ -\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \overline{x})^2 \Big\}} \\ &= \exp\Big\{ -\frac{n}{2} \Big(\frac{\overline{x} - \mu_0}{\sigma} \Big)^2 \Big\},\end{aligned}$$

y rechazamos $H_0 : \mu = \mu_0$ si y solo si Λ es pequeño.

EX UMBRA IN SOLEM

Equivalentemente , podemos rechazar H_0 a un nivel α si

$$\frac{n(\overline{X} - \mu_0)^2}{\sigma^2} > k.$$

Es decir, rechazamos H_0 si y sólo si,

$$\left| \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}} \right| > z_{1-\alpha/2},$$

donde $z_{1-\alpha/2}$ es un valor cuantil $1 - \alpha/2$ de la distribución $\mathsf{N}(0,1)$.

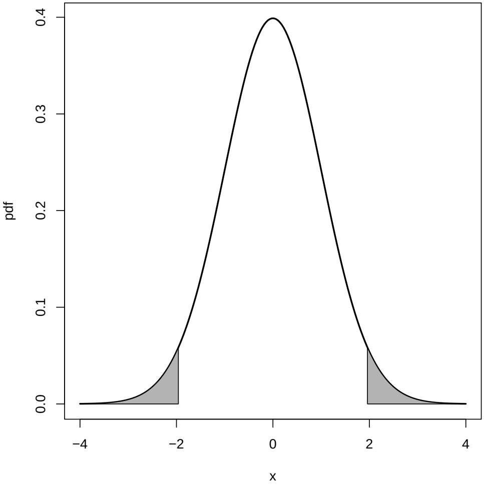
Observación:

Debido a la definición del estadístico de prueba:

$$Z = \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}} \sim \mathsf{N}(0,1).$$

en ocasiones este test de hipótesis es llamado test-Z

Test LR para la media



Test LR para la media

Suponga $X_1, \dots, X_n$ variables IID desde $\mathsf{N}(\mu, \sigma^2)$ donde $\mu \in \mathbb{R}$, $\sigma^2 > 0$ con σ^2 desconocido. Considere la hipótesis

$$H_0 : \mu = \mu_0, \qquad H_1 : \mu \neq \mu_0,$$

con μ_0 fijo.

En este caso

$$\Lambda = \Big(\frac{\widetilde{\sigma}^2}{\widehat{\sigma}^2}\Big)^{n/2} = \Big\{1 + \frac{n(\overline{x} - \mu_0)}{\sum_{i=1}^n (x_i - \overline{x})^2}\Big\}^{-n/2}$$

De este modo, rechazamos H_0 si y sólo si

$$\Big|\frac{\overline{X} - \mu_0}{S/\sqrt{n}}\Big| > t_{1-\alpha/2}(n-1),$$

donde

$$S^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \overline{X})^2,$$

y $t_{1-\alpha/2}(n-1)$ denota un valor cuantil $1-\alpha/2$ de la distribución t de Student con $n-1$ grados de libertad.

EX UMBRA IN SOLEM

Test LR para la varianza

Suponga $X_1, \dots, X_n$ variables IID desde N(μ, σ^2) donde ambos $\mu \in \mathbb{R}$ y $\sigma^2 > 0$ son desconocidos. Se desea probar la hipótesis

$$H_0 : \sigma = \sigma_0, \qquad H_1 : \sigma \neq \sigma_0,$$

donde σ_0 es fijo.

El test LR lleva al estadístico de prueba

$$Q = \frac{(n-1)S^2}{\sigma_0^2} \sim \chi^2(n-1).$$

De este modo, rechazamos H_0 si y sólo si

$$Q > \chi^2_{1-\alpha/2}(n-1), \qquad \text{o bien} \qquad Q < \chi^2_{\alpha/2}(n-1)$$

donde

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2,$$

y $\chi^2_{\alpha/2}(n-1)$, $\chi^2_{1-\alpha/2}(n-1)$ denota un valores cuantiles $\alpha/2$ y $1-\alpha/2$ de la distribución chi-cuadrado con $n-1$ grados de libertad.

Test LR para dos muestras

Sea $X_1, \dots, X_n$ y $Y_1, \dots, Y_m$ muestras aleatorias desde $\mathsf{N}(\mu_X, \sigma^2)$ y $\mathsf{N}(\mu_Y, \sigma^2)$, respectivamente y suponga que se desea probar la hipótesis

$$H_0 : \mu_X = \mu_Y, \qquad H_1 : \mu_X \neq \mu_Y,$$

El test LR es equivalente a rechazar H_0 cuando

$$\Big| \frac{\overline{X} - \overline{Y}}{s_p \sqrt{\frac{1}{n} + \frac{1}{m}}} \Big| > t_{1-\alpha/2}(n+m-2),$$

donde

$$S_p = \frac{1}{n+m-2} \big\{ (n-1)s_x^2 + (m-1)s_y^2 \big\}.$$

Test LR para dos muestras

Sea $X_1, \ldots, X_n$ y $Y_1, \ldots, Y_m$ muestras aleatorias desde $\mathsf{N}(\mu_X, \sigma_X^2)$ y $\mathsf{N}(\mu_Y, \sigma_Y^2)$, respectivamente. Se desea probar la siguiente hipótesis

$$H_0 : \sigma_X^2 = \sigma_Y^2, \qquad H_1 : \sigma_X^2 \neq \sigma_Y^2.$$

Sea

$$F = \frac{\sum_{i=1}^m (y_i - \overline{y})^2/(m-1)}{\sum_{i=1}^n (x_i - \overline{x})^2/(n-1)} \sim F(m-1, n-1)$$

De este modo, rechazamos H_0 si

$$F \geq F_{1-\alpha}(m-1, n-1),$$

donde $F_{1-\alpha}(m-1, n-1)$ denota un valor cuantil $1-\alpha$ desde la distribución F con $m-1$ y $n-1$ grados de libertad.

Test LR para dos muestras

Suponga $X_1, \ldots, X_n$ y $Y_1, \ldots, Y_m$ muestras aleatorias desde $\text{Ber}(\theta_1)$ y $\text{Ber}(\theta_2)$, respectivamente. En este caso,

$$\widehat{\theta}_1 = \frac{1}{n}\sum_{i=1}^{n} X_i, \qquad \widehat{\theta}_2 = \frac{1}{m}\sum_{i=1}^{m} Y_i.$$

Suponga que se desea probar la hipótesis

$$H_0 : \theta_1 = \theta_2, \qquad H_1 : \theta_1 \neq \theta_2.$$

Sea

$$Z = \frac{\widehat{\theta}_1 - \widehat{\theta}_2}{\sqrt{\widehat{\theta}(1-\widehat{\theta})}\sqrt{\frac{1}{n} + \frac{1}{m}}},$$

donde $\widehat{\theta} = (\sum_{i=1}^{n} X_i + \sum_{j=1}^{m} Y_j)/(m+n)$. De este modo, rechazamos H_0 si

$$|Z| > z_{1-\alpha/2}.$$

Suponga observaciones IID $(X_1, Y_1), \dots, (X_n, Y_n)$ desde $\mathsf{N}_2(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ y considere la hipótesis de interés

$$H_0 : \rho = 0, \qquad H_1 : \rho \neq 0.$$

Sea

$$r = \frac{\sum_{i=1}^n (x_i - \overline{x})(y_i - \overline{y})}{\big\{ \sum_{i=1}^n (x_i - \overline{x}) \sum_{i=1}^n (y_i - \overline{y}) \big\}^{1/2}}.$$

De este modo se rechaza H_0 si

$$\Big| \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} \Big| > t_{1-\alpha/2}(n-2).$$