Test de Hipótesis - Largada

Algunos ejemplos simpáticos para largar, del libro
Statistics: Unlocking the power of data. Lock, R.H., Lock, P.F., Morgan, K.L., Lock, E.F. and Lock, D.F., 2016.

¿Se parecen los perros a sus dueños?

© Kevin Klöpper/iStockphoto

¿Están los pollos contaminados?

Concentración media de arsénico superior a 80 ppb (partes por billón).

©Lauri Patterson/iStockphoto

How much arsenic is in this chicken?

¿Tienen hombre y mujeres diferente opinión sobre el divorcio?

Is divorce morally acceptable?

Test de hipótesis

Utilizar datos para determinar la validez de una aseveración hecha sobre un parámetro poblacional.

Determinar si los datos obtenidos resultan suficientemente convincentes para sacar alguna conclusión sobre el parámetro de interés.

Test – Jerga – Parte 1: H_0 y H_1: Hipótesis nula y alternativa

- Hipótesis (estadística): Aseveración sobre un parámetro poblacional (Gerardo Cueto)

- Hipótesis alternativa H_1: Es la hipótesis *del investigador*

- Hipótesis alternativa H_1: Escenario para el cual buscamos evidencia significativa.

- La hipótesis alternativa (H_1) se establece mediante la observación de evidencia (en los datos) que contradice la hipótesis nula (H_0) y apoya la hipótesis alternativa.

- Los datos son *raros* bajo la hipótesis nula (H_0) y ADEMÁS sugieren que sea rechazada en favor de la hipótesis alternativa
.

Hipótesis (del investigador): perros y dueños se parecen

© Kevin Klöpper/iStockphoto

Se muestra la foto de un dueño y dos mascotas, siendo una de ellas la propia. X vale 1 si acierta y 0 caso contrario. $X \sim \mathcal{B}(1, p)$.

$$H_0 : \qquad\qquad H_1 :$$

Hipótesis (del investigador): los pollos están contaminados

© Lauri Patterson/iStockphoto

How much arsenic is in this chicken?

X concentración de arsénico en un pollo elegido al azar. $\mu = \mathbb{E}(X)$

$$H_0 : \qquad\qquad H_1 :$$

Hipótesis (del investigador): hombre y mujeres tienen diferente opinión sobre el divorcio

© Sawayasu Tsuji/iStockphoto

Is divorce morally acceptable?

$$X \sim \mathcal{B}(1, p_1),\ Y \sim \mathcal{B}(1, p_2).$$

$$\delta = p_1 - p_2$$

$$H_0 : \qquad\qquad H_1 :$$

Test - Jerga - Parte II
Región de Rechazo (de H_0)

Vamos a indicar cómo deben ser los datos para DECIDIR rechazar la hipótesis nula en favor de la alternativa.

$\mathcal{R} :=$ Región de rechazo de H_0

$\mathcal{R} =:$ conjunto de resultados para los cuales H_0 es rechazada en favor de H_1

Test - Ingredientes

- H_0: Hipótesis nula. $H_0 : \theta \in \Theta_0$
- H_1: Hipótesis Alternativa. $H_1 : \theta \in \Theta_1$
- $\mathcal{R}$: Región que indica cómo deben ser las observaciones para rechazar H_0 en favor de H_1.

Test de Hipótesis - Ejemplo: De Pollos y Arsénico

p-valor - *con el vermouth - en la entrada*

- p-valor se calcula una vez realizado el experimento.
- Depende de los valores $(x_1, \ldots, x_n)$ observados.
- Indica *cuán probable es observar valores extremos como el obtenido, en la dirección de* H_1, *con* $(x_1, \ldots, x_n)$ *cuando* H_0 *es verdadero.*

  p-valor chico da evidencia contra H_0, *en favor de* H_1

Test de Hipótesis - Ejemplo

- $X_i \sim \mathcal{N}(\mu, 9)$

$$H_0 : \mu = 80 \qquad H_1 : \mu > 80$$

- Región de rechazo de H_0

$$\mathcal{R} = \left\{ \frac{\overline{X}_n - 80}{\sqrt{9/n}} \ \geq \ 1.64 \right\} = \left\{ \overline{X}_n \ \geq \ 80 + 1.64\sqrt{9/n} \right\}$$

Test de Hipótesis - Ejemplo

- Región de rechazo de H_0 (de nivel de significatividad α):

$$\mathcal{R} = \left\{ \frac{\overline{X}_n - 80}{\sqrt{9/n}} \ \geq \ 1.64 \right\} = \left\{ \overline{X}_n \ \geq \ 80 + 1.64\,\sqrt{9/n} \right\}$$

- Función de potencia:

$$\begin{aligned}
\pi(\mu) &= \mathbb{P}_\mu(\mathcal{R}) = \mathbb{P}_\mu(\overline{X}_n \ \geq \ 80 + 1.64\,\sqrt{9/n}) = \\
&= \mathbb{P}_\mu\left(\frac{\overline{X}_n - \mu}{\sqrt{9/n}} \ \geq \ \frac{80-\mu}{\sqrt{9/n}} + 1.64 \right) = \\
&= \mathbb{P}\left(Z \geq \frac{80-\mu}{\sqrt{9/n}} + 1.64 \right) = \\
&= 1 - \text{pnorm}\left(\frac{80-\mu}{\sqrt{9/n}} + 1.64 \right) = \\
&= 1 - \text{pnorm}\left(\frac{\sqrt{n}(80-\mu)}{\sqrt{9}} + 1.64 \right)
\end{aligned}$$

Test de Hipótesis - Ejemplo

- $X_i \sim \mathcal{N}(\mu, 9)$, $H_0 : \mu = 80$, $H_1 : \mu > 80$.
- Región de rechazo de H_0 (de nivel de significatividad α):

$$\mathcal{R} = \mathcal{R}_{n,\alpha} = \left\{ \frac{\overline{X}_n - 80}{\sqrt{9/n}} \;\geq\; z_\alpha \right\} = \left\{ \overline{X}_n \;\geq\; 80 + z_\alpha \sqrt{9/n} \right\}$$

- Función de potencia:

$$\pi(\mu) = \pi_n(\mu) = 1 - \text{pnorm}\left(\frac{\sqrt{n}(80 - \mu)}{\sqrt{9}} + z_\alpha \right)$$

- Propiedades:
  1. π es creciente: $\mu_a \leq \mu_b$, entonces $\pi(\mu_a) \leq \pi(\mu_b)$
  2. $\lim_{n\to\infty} \pi_n(\mu_1) = 1$, para todo $\mu_1 > 80$

Test

$$H_0 : \theta \in \Theta_0 \qquad H_1 = \theta \in \Theta_1$$

Un test es una regla de decisión que, en función de H_0 y H_1, determina cómo deben ser la muestra (o los datos) para que H_0 sea rechazada en favor de H_1.

La región de rechazo $\mathcal{R}$ indica en qué casos rechazamos H_0 en favor de H_1.

La región de rechazo $\mathcal{R}$ es el conjunto de valores que llevan a la decisión de rechazar H_0.

Test: Regla de decisión- posibles errores

$\mathcal{R}$: Región de rechazo. Determina cómo deben ser los datos para rechazar H_0 en favor de H_1.

	No Rechazamos H_0	Rechazamos H_0
H_0 es cierta	no hay error	**error Tipo I**
H_0 es falsa	**error Tipo II**	no hay error

- Error (es) tipo I: rechazar H_0 cuando es verdadera:
- Error (es) tipo II: NO rechazar H_0 cuando es falsa.

Test: Regla de decisión- posibles errores

$\mathcal{R}$: Región de rechazo. Determina cómo deben ser los datos para rechazar H_0 en favor de H_1.

	No Rechazamos H_0	Rechazamos H_0
H_0 es cierta	no hay error	**error Tipo I**
H_0 es falsa	**error Tipo II**	no hay error

- Error (es) tipo I: rechazar H_0 cuando es verdadera.

  Probabilidad (ERROR TIPO I) $= \mathbb{P}_\theta(\mathcal{R})$, con θ satisfaciendo H_0

- Error (es) tipo II: NO rechazar H_0 cuando es falsa.

$$\text{Probabilidad (ERROR TIPO II)} = \mathbb{P}_\theta(\mathcal{R}^c) = 1 - \mathbb{P}_\theta(\mathcal{R}),$$

con θ satisfaciendo H_1

$\pi(\theta)$: Función de Potencia (del Test)

$\pi(\theta)$ es la probabilidad de rechazar H_0 cuando el valor verdadero del parámetro es θ.

$$\pi(\theta) = \mathbb{P}_\theta(\mathcal{R})$$

- si θ satisface H_0, $\mathbb{P}_\theta(\mathcal{R}) = \pi(\theta)$ es la probabilidad de un error tipo I.
- si θ satisface H_1, $\mathbb{P}_\theta(\mathcal{R}^c) = 1 - \pi(\theta)$ es la probabilidad de un error tipo II.

$\pi(\theta)$: Función de Potencia (del Test) y Errores

- Función de potencia: probabilidad de rechazar H_0 cuando el valor verdadero del parámetro es θ:

$$\pi(\theta) = \mathbb{P}_\theta(\mathcal{R})$$

- si θ satisface H_0, $\mathbb{P}_\theta(\mathcal{R}) = \pi(\theta)$ es la probabilidad de un error tipo I.

  Queremos $\pi(\theta)$ chico cuando θ satisface H_0.

- si θ satisface H_1, $\mathbb{P}_\theta(\mathcal{R}^c) = 1 - \pi(\theta)$ es la probabilidad de un error tipo II.

  Queremos $\pi(\theta)$ GRANDE cuando θ satisface H_1. *Potencia grande en las alternativas*

Nivel de significatividad del Test cuando $H_0 : \theta = \theta_0$

Dados $H_0 : \theta = \theta_0$, $H_1 : \theta \in \Theta_1$ y $\mathcal{R}$, decimos que el test es de nivel α si

$$\mathbb{P}_{\theta_0}(\mathcal{R}) \leq \alpha$$

$$\mathbb{P}_{\theta_0}(\mathcal{R}) = \pi(\theta_0) \leq \alpha$$

Controlamos la probabilidad de Error Tipo I: está acotada por α (en nuestros ejemplos, conseguimos que sea igual a α)

Nivel de significatividad del Test - caso general

Dados $H_0 : \theta \in \Theta_0$, $H_1 : \theta \in \Theta_1$ y $\mathcal{R}$, decimos que el test es de nivel α si

$$\mathbb{P}_\theta(\mathcal{R}) \leq \alpha \quad \forall\, \theta \text{ satisfaciendo } H_0$$

Casos de probabilidad de Error tipo I: $\mathbb{P}_\theta(\mathcal{R}) = \pi(\theta) \leq \alpha\ , \quad \forall\, \theta \in \Theta_0$

Peor probabilidad de Error tipo I: $\sup_{\theta \in \Theta_0} \pi(\theta) \leq \alpha$

Controlamos la probabilidad de **todo posible** Error Tipo I: está acotada por α,

Test de Hipótesis - Ejemplo

- $X_i \sim \mathcal{N}(\mu, 9)$, $H_0 : \mu = 80$, $H_1 : \mu > 80$.
- Región de rechazo de H_0 (de nivel de significatividad α):

$$\mathcal{R} = \mathcal{R}_{n,\alpha} = \left\{ \frac{\overline{X}_n - 80}{\sqrt{9/n}} \ \geq \ z_\alpha \right\} = \left\{ \overline{X}_n \ \geq \ 80 + z_\alpha \sqrt{9/n} \right\}$$

Estadístico del test

$$Z = Z(X_1, \dots, X_n) := \frac{\overline{X}_n - 80}{\sqrt{\frac{9}{n}}} \quad \underset{\text{bajo } H_0}{\underset{\uparrow}{\sim}} \quad \mathcal{N}(0, 1)$$

- Región de rechazo de H_0 (de nivel de significatividad α):

$$\mathcal{R} = \{Z \geq z_\alpha\}$$

Test de Hipótesis - Ejemplo

- Región de rechazo de H_0 (de nivel de significatividad α):

$$\mathcal{R} = \mathcal{R}_{n,\alpha} = \left\{ \frac{\overline{X}_n - 80}{\sqrt{9/n}} \ \geq \ z_\alpha \right\} = \left\{ \overline{X}_n \ \geq \ 80 + z_\alpha \sqrt{9/n} \right\}$$

- Función de potencia:
  $\pi(\mu) = \pi_n(\mu) = 1 - \text{pnorm}\left(\frac{\sqrt{n}(80-\mu)}{\sqrt{9}} + z_\alpha \right)$
- Propiedades:
  1. π es creciente: $\mu_a \leq \mu_b$, entonces $\pi(\mu_a) \leq \pi(\mu_b)$

$$\sup_{\mu \leq 80} \pi(\mu) \leq \pi(80) = \alpha \quad \rightarrow \quad \text{Nivel } \alpha \text{ para } H_0 : \mu \leq 80$$

  2. $\lim_{n\to\infty} \pi_n(\mu_1) = 1$, para todo $\mu_1 > 80$

$$1 - \pi(\mu_1) \leq \beta \quad \equiv \quad n \geq \left\{ \sqrt{9}\ (\text{qnorm } (\beta) - z_\alpha)/(80 - \mu_1) \right\}^2$$

En síntesis, cuando $H_0 : \theta = \theta_0$

- Dado n, y α, se puede construir un test mediante una región de rechazo $\mathcal{R} = \mathcal{R}_{n,\alpha}$ de nivel α:

$$P_{\theta_0}(\mathcal{R}_{n,\alpha}) = \alpha$$

- Para armar la región de rechazo se usa un estadístico $T(X_1, \ldots, X_n)$ cuya distribución es CONOCIDA bajo H_0.
- La potencia del test está definida por

$$\pi(\theta) = \pi_{n,\alpha}(\theta) = \mathbb{P}_\theta(\mathcal{R}_{n,\alpha})$$

- Dado un valor θ_1 en H_1 y β, se puede encontrar n para que el error tipo II en θ_1 sea menor o igual a β.

$$\beta \geq P_{\theta_1}(\mathcal{R}^c) = 1 - \pi(\theta_1) \quad \equiv \quad \pi(\theta_1) \geq 1 - \beta$$

p- valores: Primer plato

- p-valor se calcula una vez realizado el experimento.
- Depende de los valores $(x_1, \ldots, x_n)$ observados.
- Indica *cuan probable es observar valores extremos como el obtenido con* $(x_1, \ldots, x_n)$ *cuando* H_0 *es verdadero.*

  p-valor chico da evidencia contra H_0*, en favor de* H_1

- Con los datos rechazo H_0 a nivel α si y solo si p-valor $\leq \alpha$.

Una muestra normal, varianza conocida

Sean $X_1, \ldots, X_n \sim N\left(\mu, \sigma_0^2\right)$ i.i.d., con σ_0^2 conocida.
Se quiere testear

$$H_0 : \mu = \mu_0 \text{ versus alguna de las alternativas siguientes}$$

$$H_1 : \mu > \mu_0 \qquad H_1 : \mu < \mu_0 \qquad H_1 : \mu \neq \mu_0$$

Estadístico del test

$$Z = Z(X_1, \ldots, X_n) := \frac{\overline{X}_n - \mu_0}{\sqrt{\frac{\sigma_0^2}{n}}} \quad \underset{\substack{\uparrow \\ \text{bajo } H_0}}{\sim} \quad \mathcal{N}\left(0, 1\right)$$

Región de rechazo de nivel α

Estadístico del test $Z := \frac{\overline{X}_n - \mu_0}{\sqrt{\frac{\sigma_0^2}{n}}} \sim \mathcal{N}(0,1)$, cuando $\mu = \mu_0$.

1. $H_0 : \mu = \mu_0$ vs. $H_1 : \mu > \mu_0$

$$\mathcal{R}_\alpha = \left\{ \frac{\overline{X}_n - \mu_0}{\sqrt{\frac{\sigma_0^2}{n}}} \geq z_\alpha \right\} = \{Z \geq z_\alpha\}$$

2. $H_0 : \mu = \mu_0$ vs. $H_1 : \mu < \mu_0$

$$\mathcal{R}_\alpha = \left\{ \frac{\overline{X}_n - \mu_0}{\sqrt{\frac{\sigma_0^2}{n}}} < -z_\alpha \right\} = \{Z \leq -z_\alpha\}$$

3. $H_0 : \mu = \mu_0$ vs. $H_1 : \mu \neq \mu_0$

$$\mathcal{R}_\alpha = \left\{ \left| \frac{\overline{X}_n - \mu_0}{\sqrt{\frac{\sigma_0^2}{n}}} \right| > z_{\alpha/2} \right\} = \{|Z| \geq z_{\alpha/2}\}$$

p-valor - Una muestra normal, varianza conocida

Estadístico del test $Z := \frac{\overline{X}_n - \mu_0}{\sqrt{\frac{\sigma_0^2}{n}}}$

Estadístico OBSERVADO:

$$z_{\textsf{obs}} = \frac{\overline{X}_{n,\textsf{obs}} - \mu_0}{\sqrt{\frac{\sigma_0^2}{n}}} = \frac{\overline{x}_n - \mu_0}{\sqrt{\frac{\sigma_0^2}{n}}}$$

1. $H_0 : \mu = \mu_0$ vs. $H_1 : \mu > \mu_0$

$$\text{p-valor} = \mathbb{P}(Z \geq z_{\textsf{obs}})$$

2. $H_0 : \mu = \mu_0$ vs. $H_1 : \mu < \mu_0$

$$\text{p-valor} = \mathbb{P}(Z \leq z_{\textsf{obs}})$$

3. $H_0 : \mu = \mu_0$ vs. $H_1 : \mu \neq \mu_0$

$$\text{p-valor} = 2\mathbb{P}(Z \geq |z_{\textsf{obs}}|)$$