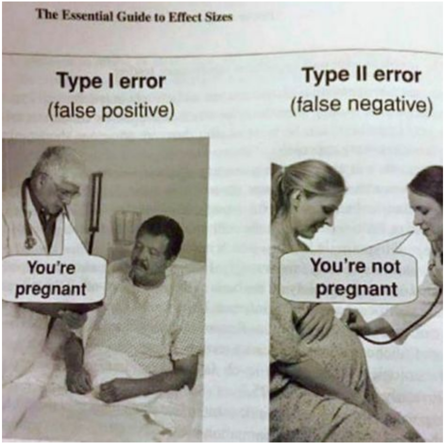


Figura: Hoy tenemos dos errores posibles

Clase 18 de Bioestadística

Tests de hipótesis

Ernesto Mordecki

CMAT, Facultad de Ciencias, Universidad de la República.

Uruguay

Contenidos de la clase

Test de hipótesis

Prueba de hipótesis para la esperanza

Construcción de la región crítica

Test de hipótesis

- Supongamos que tiramos una moneda 1000 veces y obtenemos 550 caras.
- ¿está equilibrada esa moneda?
- Es decir: las 50 caras de exceso de las esperadas, ¿obedecen al azar o a un desequilibrio?

- Esa pregunta se formula matemáticamente mediante un test de hipótesis.
- Para eso se formula una hipótesis (la hipótesis nula), que se denota H_0 (se lee “hache cero”)
- y se formula una hipótesis alternativa hache uno: H_1
  - $H_0 : \quad p = p_0 \ (1/2)$
  - $H_1 : \quad p \neq p_0 \ (1/2)$
- La idea es la siguiente: en principio pensamos que H_0 es verdadera, y la rechazamos si hay evidencia suficiente en los datos.

El procedimiento es el siguiente:

- Construímos un intervalo I asumiendo que H_0 es cierta de nivel α.
- Ese intervalo tiene centro en p_0 y amplitud ε con $\hat{p}$ estimado.
- El complemento $RC = [0,1] \setminus I$ se llama **región crítica**
- Calculamos $\hat{p}$ el estimador de p:
  - Si $\hat{p} \in RC$ **rechazamos la hipótesis nula** (el estimador pertenece a la región crítica)
  - Si $\hat{p} \notin RC$ **no rechazamos la hipótesis nula** (el estimador no pertenece a la región crítica)

Ejemplo

Tomamos como ejemplo las 550 caras de la tirada de la moneda. Asumimos $\alpha = 0{,}05$.

- El intervalo es $(\frac{1}{2} - \varepsilon, \frac{1}{2} + \varepsilon)$
- Tenemos
$$\varepsilon = 1{,}96 \frac{\sqrt{1/4}}{\sqrt{n}} = 1{,}96 \frac{1}{2\sqrt{1000}} = 0{,}0309.$$
- El intervalo es $(0{,}5 - 0{,}0309, 0{,}5 + 0{,}0309) = (0{,}4691, 0{,}5309)$
- ¿Cuál es la conclusión del test?

- **Rechazamos** la hipótesis nula, porque $0{,}55 \notin (0{,}4691, 0{,}5309)$.
- Decimos que hay evidencia estadística suficiente para decidir que la moneda no está equilibrada.
- ¿Que hubiera pasado si teníamos 5 500 caras en 10 000 tiradas?
- ¿Y si tenemos 55 caras en 100 tiradas?

- Si tenemos el mismo $\hat{p}$ con un n mayor, el intervalo se reduce y la conclusión es la misma.

- Si tenemos $n = 100$, la amplitud es

$$\varepsilon = 1{,}96 \frac{\sqrt{1/2}}{\sqrt{n}} = 1{,}96 \frac{1}{2\sqrt{100}} = 0{,}098.$$

- El intervalo es $(0{,}5 - 0{,}098, 0{,}5 + 0{,}0975) = (0{,}402, 0{,}598)$.

- Como $\hat{p} \in I$ no hay evidencia suficiente y **no rechazamos** la hipótesis nula.

Prueba de hipótesis para la esperanza

Formulamos ahora el mismo problema para la esperanza $\mu = E(X_1)$ de una m.a.s. $X_1, \ldots, X_n$

- Queremos testear si la esperanza de la muestra es un valor dado μ_0

- Formulamos nuestra hipótesis nula:

$$H_0 : \quad \mu = \mu_0$$

La hipótesis alternativa puede tomar diversas formas

- $H_1 : \mu = \mu_1$ (hipótesis simple)
- $H_1 : \mu \neq \mu_0$ (test bilateral)[1]
- $H_1 : \mu > \mu_0$ (test unilateral)
- $H_1 : \mu < \mu_0$ (test unilateral)

Los tests unilaterales se utilizan cuando hay información adicional:

- Queremos testear si un fármaco mejora un valor clínico
- Queremos testear si un método de enseñanza mejora el rendimiento

Si no tenemos esa información hacemos un test bilateral

[1] Si la hipótesis alternativa tiene más de un punto se llama compuesta.

Construcción de la región crítica

- Consideramos el promedio de los datos $\overline{X}_n$ como estimador de μ
- Suponemos que H_0 es cierta, es decir, que conocemos la esperanza y vale μ_0
- Nuestro problema es saber si está cerca o lejos de μ_0
- Fijamos un error de tipo I (α) que estamos dispuestos a tolerar
- Construímos un intervalo a partir de μ_0 con probabilidad $1 - \alpha$
- Eso lo hacemos aplicando el TCL

Construcción de la región crítica bilateral

- Como la hipótesis alternativa es $\mu \neq \mu_0$, nuestro intervalo es simétrico y centrado en μ_0
- Es decir, determinamos ε tal que

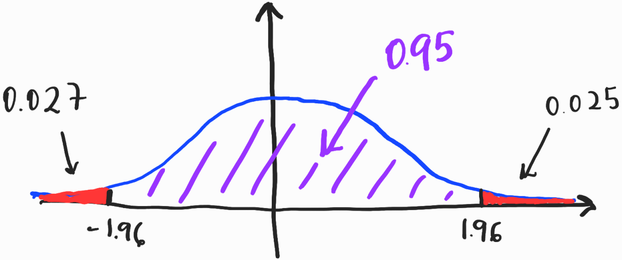
$$P_{H_0}(\overline{X}_n \in (\mu_0 - \varepsilon, \mu_0 + \varepsilon)) = 1 - \alpha$$

$$P_{H_0}(-\varepsilon \leq \overline{X}_n - \mu_0 \leq \varepsilon) = 1 - \alpha$$

$$P_{H_0}\left(-\varepsilon\frac{\sqrt{n}}{\sigma} \leq \frac{\sqrt{n}}{\sigma}\left(\overline{X}_n - \mu_0\right) \leq \frac{\sqrt{n}}{\sigma}\varepsilon\right) = 1 - \alpha$$

$$P_{H_0}\left(-z_{1-\alpha/2} \leq Z \leq z_{1-\alpha/2}\right) = 1 - \alpha$$

Es la misma situación que en los intervalos de confianza:



$\alpha = 0.05$

$\alpha/2 = 0.025$

$1 - \alpha/2 = 0.975$

$\text{qnorm}(0.975) = 1{,}96$

$z_{1-\alpha/2} = z_{0.955} = 1{,}96$

- Igualando

$$z_{1-\alpha/2} = \varepsilon \frac{\sqrt{n}}{\sigma}$$

- Y despejando

$$\varepsilon = z_{1-\alpha/2} \frac{\sigma}{\sqrt{n}}$$

- $z_{1-\alpha/2}$ es el cuantil $1 - \alpha/2$ de la normal
- Por ejemplo, si $\alpha = 0{,}05$, tenemos $1 - \alpha/2 = 0{,}975$ y $\texttt{qnorm}(0{,}975) = 1{,}96$
- Si no conocemos σ lo estimamos mediante

$$s_n = \sqrt{\frac{1}{n-1} \sum_{k=1}^{n} \left(X_k - \overline{X}_n\right)^2}$$

Conclusión

- Tenemos entonces nuestra región crítica que es el complemento de $(\mu_0 - \varepsilon, \mu_0 + \varepsilon)$
- Es decir $RC = (-\infty, \mu_0 - \varepsilon] \cup [\mu_0 + \varepsilon, \infty)$
- Si $\overline{X}_n \in RC$, rechazamos la hipótesis nula
- En caso contrario, no la rechazamos

Construcción de la región crítica en el caso unilateral

- Ahora la hipótesis alternativa es de la forma $\mu > \mu_0$
- Como sospechamos que el μ verdadero sea mayor que μ_0, rechazamos la hipótesis si es mucho mayor
- Es decir, construímos una región crítica de la forma

$$[\mu_0 + \varepsilon, \infty)$$

- Es un solo intervalo.
- Es decir, determinamos ε tal que

$$P_{H_0}(\overline{X}_n \geq \mu_0 + \varepsilon)) = 1 - \alpha$$

El razonamiento con el TCL es similar, pero aparece $z_{1-\alpha}$

$$P_{H_0}(\overline{X}_n \geq \mu_0 + \varepsilon)) = 1 - \alpha$$

$$P_{H_0}(\overline{X}_n - \mu_0 \geq \varepsilon) = 1 - \alpha$$

$$P_{H_0}\left(\frac{\sqrt{n}}{\sigma}\left(\overline{X}_n - \mu_0\right) \geq \frac{\sqrt{n}}{\sigma}\varepsilon\right) = 1 - \alpha$$

$$P_{H_0}\left(Z \geq z_{1-\alpha}\right) = 1 - \alpha$$

- Igualando

$$z_{1-\alpha} = \varepsilon \frac{\sqrt{n}}{\sigma}$$

- Y despejando

$$\varepsilon = z_{1-\alpha} \frac{\sigma}{\sqrt{n}}$$

- $z_{1-\alpha/2}$ es el cuantil $1-\alpha$ de la normal
- Por ejemplo, si $\alpha = 0{,}05$, tenemos $1-\alpha/2 = 0{,}95$ y $\mathtt{qnorm}(0{,}95) = 1{,}64$
- Si no conocemos σ lo estimamos mediante

$$s_n = \sqrt{\frac{1}{n-1} \sum_{k=1}^{n} \left(X_k - \overline{X}_n\right)^2}$$

Conclusión

- Tenemos entonces nuestra región crítica que es

$$RC = [\mu_0 + \varepsilon, \infty)$$

- Si $\overline{X}_n \in RC$, rechazamos la hipótesis nula
- En caso contrario, no la rechazamos

Potencia de un test

Cuando la hipótesis alternativa es simple, podemos calcular cuan bien se desempeña el test con el siguiente argumento.

- Suponemos que la hipótesis verdadera es H_1.
- Calculamos entonces cual es la probabilidad, siendo cierta H_1, de que el estimador **caiga** en la región crítica (como debería de ser).
- Esa probabilidad, que se espera sea alta si el test se desempeña correctamente, se llama potencia del test, y se denota mediante π.
- Su complemento, $\beta = 1 - \pi$ es el error de tipo 2 (falso negativo: aceptar un falso)
- (α se llama error del tipo 1, y es el error de tener un falso postivo (rechazar un verdadero)