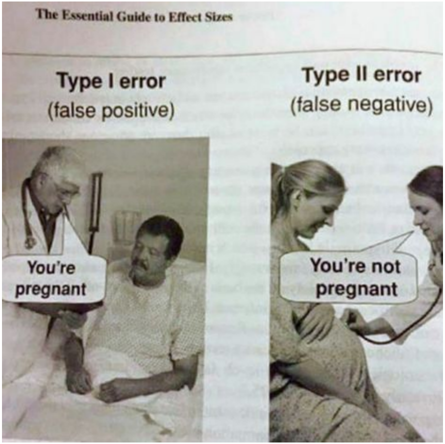


Figura: Hoy tenemos dos errores posibles

Clase 19 de Bioestadística

Tests de hipótesis (2): potencia y p-valor

Ernesto Mordecki

CMAT, Facultad de Ciencias, Universidad de la República.

Uruguay

Contenidos de la clase

Prueba de hipótesis para la esperanza

Construcción de la región crítica

Potencia de un test

p-valor

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

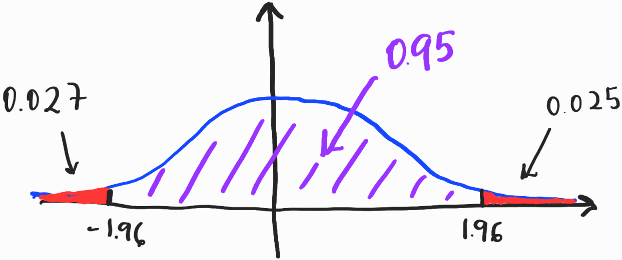
$$P_{H_0}(\overline{X}_n \in (\mu_0 - \varepsilon, \mu_0 + \varepsilon)) = 1 - \alpha$$

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

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

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

Es la misma situación que en los intervalos de confianza:



$\alpha = 0.05$

$\alpha/2 = 0.025$

$1 - \alpha/2 = 0.975$

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 `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

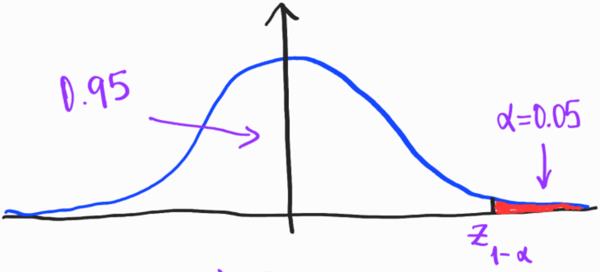
$$s_n = \sqrt{\frac{1}{n-1} \sum_{k=1}^{n} \left(X_k - \overline{X}_n \right)^2}$$

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



$$\text{qnorm}(1-\alpha) = z_{1-\alpha}$$

$$1-\alpha = 0.95$$

$$\text{qnorm}(0.95) =$$

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)

Ejemplo

- Para ser aprobada por las autoridades regulatorias, una vacuna tiene que dar resultado positivo por lo menos en el 90 % de los casos, en una determinada etapa de su producción.
- En la etapa anterior se confirmó su efectividad en un 80 % de los casos.
- Para determinar entonces su efectividad se aplica la vacuna 100 veces.
- Se obtiene efectividad en 89 casos.
- ¿Puede afirmarse con un 99 % de certeza que la vacuna supera la etapa?

- Formulamos un test de hipótesis con alternativa unilateral
- Tenemos $H_0: p = 0{,}8$ contra $H_1: p \geq 0{,}9$.
- Como $\alpha = 0{,}01$, tenemos `qnorm(0.99)=2.326`
- Entonces

$$\varepsilon = 2{,}326 \frac{\sqrt{p_0(1-p_0)}}{\sqrt{n}} = 2{,}326 \frac{\sqrt{0{,}8 \times 0{,}2}}{\sqrt{100}} = 0{,}09304$$

- El intervalo es $(0{,}8 + 0{,}09304,\ 1] = (0{,}893,\ 1]$
- ¿Cuál es la conclusión del test?

No se rechaza la alternativa, no hay evidencia suficiente para eso, y la vacuna no pasa la etapa, nos quedamos en H_0.

- Y nos vamos a dormir tranquilos, ¡es un 99 %!
- Pero, ¿que pasa si fuera cierta H_1? ¿Cuál es la probabilidad de equivocarme en ese caso?

Cálculo de la potencia:

Suponemos que H_1 es cierta:

$$\pi = P_{H_1}(\widehat{p} \in RC) = P_{H_1}(\widehat{p} \geq 0{,}893)$$

$$= P_{H_1}(\widehat{p} - p_1 \geq 0{,}893 - p_1)$$

$$= P_{H_1}\left(\frac{\sqrt{n}}{\sigma}(\widehat{p} - p_1) \geq \frac{\sqrt{n}}{\sigma}(0{,}893 - p_1)\right)$$

- Como $n = 100$, tenemos $\sqrt{n} = \sqrt{100} = 10$.
- Como suponemos H_1 cierta, tenemos $p_1 = 0{,}9$, y

$$\sigma = \sqrt{p_1(1 - p_1)} = \sqrt{0{,}9 \times 0{,}1} = 0{,}3$$

$$= P_{H_1}\left(\frac{\sqrt{n}}{\sigma}(\widehat{p} - p_1) \geq \frac{\sqrt{n}}{\sigma}(0{,}893 - p_1)\right)$$

$$= P_{H_1}\left(Z \geq \frac{10}{0{,}3}(0{,}893 - 0{,}9)\right)$$

$$= P_{H_1}\left(Z \geq -0{,}233\right)$$

$$= \int_{-2{,}33}^{\infty} \varphi(x)dx = 1 - \mathtt{pnorm}(-0{,}233) = \mathbf{0{,}59}$$

Y . . . ¿dormimos tranquilos?

p-valor

- Supongamos que tenemos un test de hipótesis al nivel α y rechazamos la hipótesis nula[2]
- Es decir, nuestro estadístico $\overline{X}_n$ está sospechosamente lejos de μ_0
- Pero, ¿si achicamos el α seguimos rechazando?
- El *p*-valor es el valor mas grande posible que me sigue dando rechazo.

[2] Puede ser unilateral o bilateral

Ejemplo

Obtenemos 550 caras al tirar 1000 veces una moneda. Asumimos $\alpha = 0{,}05$.

- El intervalo es

$$\left(\frac{1}{2} - \varepsilon, \frac{1}{2} + \varepsilon\right)$$

- Tenemos

$$\varepsilon = 1{,}96\frac{\sqrt{1/4}}{\sqrt{n}} = 1{,}96\frac{1}{2\sqrt{1000}} = 0{,}0309.$$

- El intervalo es

$$(0{,}5 - 0{,}0309, 0{,}5 + 0{,}0309) = (0{,}4691, 0{,}5309)$$

- Como $\widehat{p} = 0{,}55$, rechazamos la hipótesis nula

- Tomamos un $\alpha = 0{,}01$. En este caso, como el test es bilateral tengo $1 - \alpha/2 = 0{,}995$. Calculamos `qnorm(0.995)=2.58`
-
$$\varepsilon = 2{,}58 \frac{1}{2\sqrt{1000}} = 0{,}041.$$
- Entonces el intervalo es ahora
$$(0{,}5 - 0{,}041, 0{,}5 + 0{,}041) = (0{,}459, 0{,}541)$$
- Ahora $\hat{p} = 0{,}55$ está en el intervalo, sigo rechazando la hipótesis nula
- El *p*-valor es menor que 0,01

Cálculo del *p*-valor

- Hago ahora las cuentas así:

$$\varepsilon = z_{1-p/2}\frac{1}{2\sqrt{1000}} = 0{,}05.$$

- ¿Cuánto vale p?

$$z_{1-p/2} = 2\sqrt{1000} \times 0{,}05 = 3{,}16$$

- Calculamos `pnorm(3.16)=0.9992`.
- Entonces

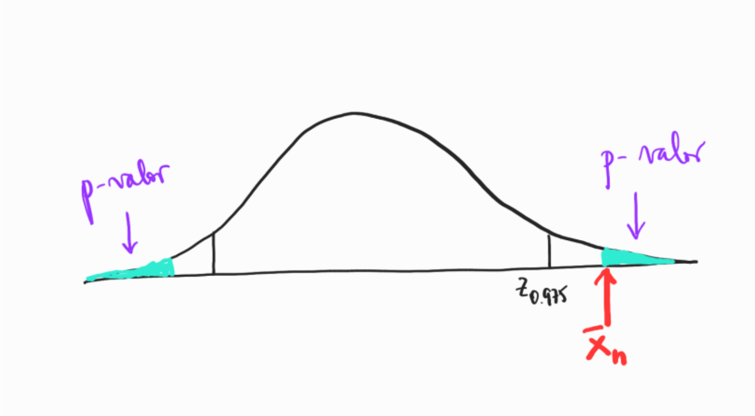
$$1 - \frac{p}{2} = 0{,}9992$$

- Despejando

$$p = 0{,}0016 = 1{,}6 \times 10^{-3}$$

  es el p-valor.

Gráfico del p-valor



- Empujo el gráfico de la región crítica hasta que el valor observado $\overline{X}_n$ queda en el borde.
- El área que me queda es el p-valor