Pruebas de hipótesis

Tenemos dos hipótesis $\longrightarrow$ H_0 la hipótesis nula
$\searrow$ H_1 la hipótesis alternativa

H_0 y H_1 tienen que ser incompatibles

Dada una MAS $X_1, \dots, X_n$ tenemos que decidir entre H_0 y H_1

→ definimos la región crítica RC

RC es la zona de rechazo de H_0 : son las muestras que nos convencen que H_1 es cierta.

→ tenemos dos tipos de errores

* error de tipo 1: rechazo H_0 cuando H_0 es cierta
* error de tipo 2: no rechazo H_0 cuando H_1 es cierta

⇒ queremos minimizar el error de tipo 1

$\alpha = P(\text{rechazar } H_0 \text{ cuando } H_0 \text{ es cierta})$

$= P_{H_0}(\text{rechazar } H_0)$

$= P_{H_0}((X_1, \dots, X_n) \in RC)$

α es el nivel de significación de la prueba

$\beta = P(\text{no rechazar } H_0 \text{ cuando } H_1 \text{ es cierta})$

$= P_{H_1}(\text{no rechazar } H_0)$

$= P_{H_1}((X_1, \dots, X_n) \notin RC)$

potencia de la prueba $= 1 - \beta$

$= 1 - P(\text{no rechazar } H_0 \text{ cuando } H_1 \text{ es cierta})$

$= P(\text{rechazar } H_0 \text{ cuando } H_1 \text{ es cierta})$

$= P_{H_1}(\text{rechazar } H_0)$

$= P_{H_1}((X_1, \dots, X_n) \in RC)$

Procedimiento para realizar una prueba de hipótesis:

① planteamos H_0 y H_1

② elegir el valor de α y construimos la región crítica de nivel α

③ tomamos una decisión: si la muestra está en RC rechazamos H_0 y si la muestra no está en RC no rechazamos H_0.

Ejercicio 1

La siguiente tabla registra los niveles de cloro en sangre de una muestra de 10 pacientes de una clínica, medido en milimoles por litro.

101,99	106,64	103,36	109,54	103,99
107,32	106,55	103,7	100,57	105,85

1. Asumiendo que los datos tienen distribución normal con media μ y desvío $\sigma = 2,5$, implemente la siguiente prueba de hipótesis:

$$\begin{cases} H_0: & \mu = 104 \text{ mg/dl} \\ H_1: & \mu > 104 \text{ mg/dl}. \end{cases}$$

Nota: Trabaje al nivel $\alpha = 0,05$

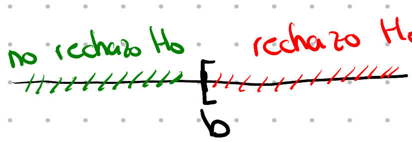
X = nivel de cloro en sangre

$X \sim \mathcal{N}(\mu, (2,5)^2)$

$$\begin{cases} H_0: \mu = 104 \\ H_1: \mu > 104 \end{cases}$$

$\alpha = 0,05$

la región crítica va a ser de la forma $RC = \{ \overline{X}_n \geq b \}$



buscamos b de forma que α valga 0,05 en una muestra de tamaño 10

$$\alpha = P(\text{rechazar } H_0 \text{ si } H_0 \text{ es cierta})$$

$$= P_{H_0}(\overline{X}_n \geq b)$$

como estamos suponiendo que H_0 es cierta:

$X \sim \mathcal{N}(104, 2,5^2)$

$\overline{X}_n \sim \mathcal{N}\left(104, \frac{2,5^2}{10}\right)$

$$= P\Bigg(\underbrace{\frac{\overline{X}_n - 104}{2,5}\sqrt{10}}_{\mathcal{N}(0,1)} \geq \frac{b-104}{2,5}\sqrt{10} \Bigg)$$

$$= 1 - \phi\left(\frac{b-104}{2,5}\sqrt{10}\right)$$

como $\alpha = 0{,}05$ tenemos

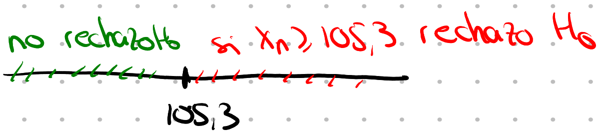
$$1-\phi\left(\frac{b-104}{2{,}5}\sqrt{10}\right) = 0{,}05$$

$$\phi\left(\frac{b-104}{2{,}5}\sqrt{10}\right) = 0{,}95$$

$$\frac{b-104}{2{,}5}\sqrt{10} = \phi^{-1}(0{,}95) = 1{,}645$$

$$\Rightarrow b = \frac{1{,}645 \cdot 2{,}5}{\sqrt{10}} + 104 = 105{,}3$$

Entonces $\quad RC = \{\overline{X}_n \geq 105{,}3\}$



el promedio muestral es : $\overline{X}_n = 104{,}951$

$\Rightarrow$ no rechazo H_0

2. Sabiendo que el verdadero valor de μ es 106 mg/dl, calcular la potencia de la prueba.

potencia = P(rechazar H_0 cuando H_1 es cierta)

$$= P_{H_1}(\overline{X}_n \geq 105{,}3)$$

como sabemos que $\mu = 106$:

$$X \sim \mathcal{N}(106, 2{,}5^2)$$

$$\overline{X}_n \sim \mathcal{N}\left(106, \frac{2{,}5^2}{10}\right)$$

$$= P\Big(\underbrace{\frac{\overline{X}_n - 106}{2{,}5}\sqrt{10}}_{\mathcal{N}(0,1)} \geq \frac{105{,}3-106}{2{,}5}\sqrt{10}\Big)$$

$$= 1-\phi\left(\frac{105{,}3-106}{2{,}5}\sqrt{10}\right)$$

$$= 1-\phi(-0{,}89)$$

$$= 1-(1-\phi(0{,}89))$$

$$= \phi(0{,}89)$$

$$= 0{,}8133$$

$\phi(0,89)$

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7703	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	$\boxed{0.8133}$
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389

$\phi^{-1}(0,95) = 1,645$

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.0
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7
0.7	0.7580	0.7611	0.7642	0.7673	0.7703	0.7734	0.7764	0.7794	0.7
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9
1.6	0.9452	0.9463	0.9474	0.9484	0.9495 $\boxed{0{,}95}$	0.9505	0.9515	0.9525	0.9
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9

Ejercicio 3

En años recientes los granjeros suecos fumigaron sus campos sembrados de cereales con un fungicida que contenía metil de mercurio. Para tener una estimación del grado de contaminación inducido a productos comestibles, se realizó un estudio sobre el nivel de mercurio de los huevos producidos en Suecia. Para tal fin se tomó una muestra aletoria de 1600 huevos y se observó que 940 de ellos estaban contaminados (esto es, tenían un nivel de metil de mercurio superior al máximo tolerado). En lo que sigue, denotamos por p a la proporción de huevos contaminados.

a) Construya un intervalo de confianza 90 % aproximado para p.

b) Realice una prueba de hipótesis al nivel $\alpha = 0,10$ para decidir entre las siguientes hipótesis:

$$\begin{cases} H_0: & p = 0,6 \\ H_1: & p < 0,6 \end{cases}$$

c) Sabiendo que el verdadero valor de p es 0,55, calcule la potencia de la prueba.

$X \sim Ber(p)$ $E(X) = p$, $Var(X) = p(1-p)$

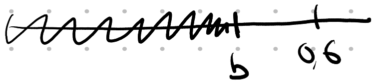
p es la proporción verdadera de huevos contaminados

proporción muestral de huevos contaminados es: $\dfrac{940}{1600} = \overline{X_n}$

$X_1, \dots, X_n$ MAS

$\overline{X_n}$ = proporción muestral

① $\begin{cases} H_0: p = 0{,}6 \\ H_1: p < 0{,}6 \end{cases}$ con $\alpha = 0{,}1$



la región crítica va a ser de la forma $\{ \overline{X}_n \leq b \}$

buscamos b para que α sea 0,1 en una muestra de tamaño 1600

$$\alpha = P(\text{rechazar } H_0 \text{ cuando } H_0 \text{ es cierta})$$

$$= P_{H_0}(\overline{X}_n \leq b)$$

$$= P\Big(\underbrace{\frac{\overline{X}_n - 0{,}6}{\sqrt{0{,}00015}}}_{\mathcal{N}(0,1)} \leq \frac{b - 0{,}6}{\sqrt{0{,}00015}}\Big)$$

$$= \phi\left(\frac{b - 0{,}6}{\sqrt{0{,}00015}}\right)$$

estamos suponiendo que H_0 es cierta

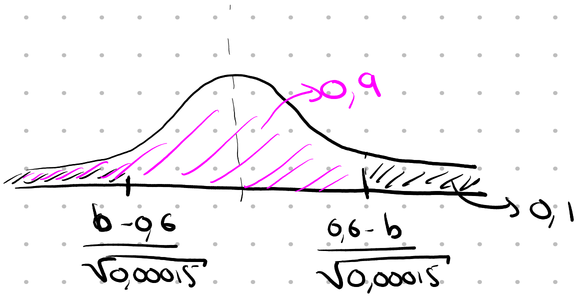
$X \sim Ber(0{,}6)$

$\overline{X}_n \sim \mathcal{N}\Big(0{,}6 \,;\, \underbrace{\frac{0{,}6(1-0{,}6)}{1600}}_{=\,0{,}00015}\Big)$

Como $\alpha = 0{,}1$

$$\phi\left(\frac{b - 0{,}6}{\sqrt{0{,}00015}}\right) = 0{,}1$$

$$\phi\left(\frac{0{,}6 - b}{\sqrt{0{,}00015}}\right) = 0{,}9$$



$$\Rightarrow \frac{0{,}6 - b}{\sqrt{0{,}00015}} = \phi^{-1}(0{,}9) = 1{,}285$$

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7703	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997 $\boxed{0{,}9}$	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319

$$\Rightarrow -b = 1{,}285 \cdot \sqrt{0{,}00015} - 0{,}6$$

$$b = 0{,}6 - 1{,}285\sqrt{0{,}00015} = 0{,}584$$

entonces $RC = \{ \overline{X_n} \leq 0{,}584 \}$

el promedio muestral es: $\overline{X_n} = \frac{940}{1600} = 0{,}5875$

$\overline{X_n} \notin RC$ por lo tanto no rechazamos H_0.