

Práctica 2

Los ejercicios señalados con (*) son opcionales para los alumnos de la Carrera de Especialización en Estadística y de la Maestría en Estadística Matemática

Ejercicio 1: En el ejemplo dado en clase se consideraban los ángulos de error en días de sol. Los siguientes datos se tomaron sobre 13 pájaros en días nublados.

8	10	38	43	45	57	73	76	83	105	112	126	141
---	----	----	----	----	----	----	----	----	-----	-----	-----	-----

a) Construya un test de nivel aproximado para

$$H_0: \theta = 90^\circ \quad \text{vs} \quad H_1: \theta \neq 90^\circ$$

¿Cuál es el p-valor?

b) Obtenga un estimador puntual y un intervalo de confianza de nivel 90% para θ .

Ejercicio 2: Seis estudiantes fueron sometidos a una dieta para perder peso con los siguientes resultados. El peso se registró en libras.

Estudiante	Antes	Después
A	174	165
E	191	186
J	188	183
M	182	178
P	201	203
R	188	181

¿Es la dieta una forma efectiva de bajar de peso?.

Ejercicio 3: En un laboratorio, insectos de cierto tipo son liberados en el centro de un círculo dibujado en una superficie plana. Una esencia que atrae a tal tipo de insectos es localizada en un extremo de la mesa. Cada insecto es liberado individualmente y es observado hasta que cruza el límite del círculo. En ese momento se registra si el insecto cruzó por la mitad del círculo en dirección al objeto o en la mitad opuesta. Al final del experimento 33 insectos fueron hacia la esencia, 16 en contra y 12 no cruzaron el círculo dentro de un tiempo razonable. ¿Hay evidencia de que la esencia atrae a los insectos?

Ejercicio 4 (*): Sean X_1, \dots, X_n v.a.i.i.d. con distribución $G(x)=F(x - \theta)$, $F \in \Omega_0$ donde

$$\Omega_0 = \{F / F \text{ es continua con única mediana igual a } 0, \text{ o sea } F(0) = 1/2\}$$

Se desea testear

$$H_0: \theta = 0 \quad \text{vs} \quad H_1: \theta > 0$$

o equivalentemente,

$$H_0: G(0) = 1/2 \quad \text{vs} \quad H_1: G(0) < 1/2$$

usando el test basado en $S = \# \{ i / x_i > 0 \}$.

a) Muestre que bajo H_0 , S tiene distribución libre.

b) ¿Es cierto el resultado bajo la alternativa?.

c) Sea $S^* = S/n$. El test del signo también rechaza la hipótesis nula si S^* es grande. Muestre que

$$S^* \rightarrow \begin{cases} 1/2 & \forall F \in \Omega_o, \theta = 0 \\ > 1/2 & \forall F \in \Omega_o, \theta > 0 \end{cases}$$

¿Qué conclusión saca?

d) En base a los resultados anteriores, muestre que si F tiene densidad f , la eficacia de S es

$$c_S = 2 f(0)$$

Muestre que el test paramétrico usual, el test t , tiene eficacia

$$c_t = \frac{1}{\sigma_f^2} \quad \text{donde} \quad \sigma_f^2 = \int x^2 f(x) dx$$

Calcule la eficiencia del test del signo respecto del test t .

Ejercicio 5: Para $n = 10$ y $\alpha = 0.0547$.

a) Construya el test del signo de nivel α para

$$H_0: \theta = 0 \quad \text{vs} \quad H_1: \theta > 0$$

b) Halle la potencia del test para
i) la distribución $N(1,1)$.

ii) la distribución con densidad $f(x) = \frac{1}{2} e^{-|x-1|}$.

c) Halle la potencia aproximada en la distribución $N(1,1)$ usando la aproximación normal con corrección por continuidad.

Ejercicio 6 (*): Construya el gráfico de $S(\theta)$ para n impar y deduzca el estimador y el intervalo de confianza para θ .

Ejercicio 7 (*): a) Muestre que el test t tiene tolerancia 0 de aceptación.

b) Halle la tolerancia de rechazo del test del signo y muestre que converge a la tolerancia asintótica de la mediana.

Ejercicio 8: Fueron seleccionados al azar 135 ciudadanos y se les pidió su opinión sobre la política exterior de los Estados Unidos, de ellos 43 estaban en contra. Después de varias semanas durante las que recibieron una carta informativa fueron nuevamente consultados, 37 estaban en contra, y 30 de las 37 eran personas que originalmente no estaban opuestas a la política exterior.

a) ¿Es el cambio en el número de personas opuestas significativo?

b) Suponiendo ahora que las 37 personas opuestas a la política exterior después del experimento fueran las mismas que se opusieron antes, ¿es el cambio en el número de personas opuestas significativo?

Ejercicio 9: En cierta ciudad la mortalidad por cada 100000 ciudadanos debida a accidentes automovilísticos fue en los últimos 15 años la siguiente:

17.3	17.9	18.4	18.1	18.3	19.6	18.6	19.2	17.7	20.0	19.0	18.8	19.3	20.2	19.9
------	------	------	------	------	------	------	------	------	------	------	------	------	------	------

¿Hay base para la afirmación de que la tasa de mortalidad por accidentes automovilísticos está aumentando?

Ejercicio 10: Durante los últimos 34 años un colegio registró las alturas promedio de los varones. Los promedios fueron

68.3	68.6	68.4	68.1	68.4	68.2	68.7	68.9	69.0	68.8	69.0	68.6	69.2	69.2	68.9	68.6	68.6
68.8	69.2	68.8	68.7	69.5	68.7	68.8	69.4	69.3	69.3	69.5	69.5	69.0	69.2	69.2	69.1	69.9

¿Son indicativos estos datos de una tendencia creciente en la alturas?.

Ejercicio 11: En un experimento para determinar la influencia de la sugestión, 20 líneas rectas de longitud variable fueron mostradas de a una por vez a dos sujetos A y B, y los sujetos estimaban la longitud de cada línea. El sujeto A establecía su preferencia en primer lugar y, sin saberlo el sujeto B, tenía instrucciones de sobreestimar las 10 primeras líneas y subestimar las 10 últimas. Después de oír al sujeto A, el sujeto B daba su estimación. Los errores de las estimaciones, medidos restando la verdadera longitud de cada línea de la longitud estimada fue:

Línea	Sujeto A	Sujeto B	Línea	Sujeto A	Sujeto B
1	0.3	-0.1	11	-1.3	-0.6
2	1.1	0.6	12	-1.1	-1.2
3	0.9	1.0	13	-1.3	-1.0
4	0.6	0.7	14	-0.7	-0.7
5	1.0	0.2	15	-1.4	-0.7
6	1.3	0.9	16	-1.4	-0.1
7	0.8	-0.1	17	-0.8	-0.5
8	1.6	0.2	18	-0.5	0.0
9	1.2	0.0	19	-1.2	-0.4
10	0.8	0.5	20	-1.0	-0.3

¿Existe correlación positiva entre los errores correspondientes a los dos sujetos?.

Ejercicio 12: Los datos para el test de Mc Nemar pueden ser escritos como observaciones bivariadas (X_i, Y_i) donde cada observación es 0 ó 1 *antes* y 0 ó 1 *después*. Supongamos que se usase el test t para datos apareados en base a la diferencia $D_i = X_i - Y_i$ con $i = 1, \dots, n$. El estadístico del test es

$$t = \frac{\bar{D} \sqrt{n-1}}{s}$$

siendo \bar{D} la media muestral y s la desviación estandar muestral de las diferencias. Este estadístico se compara con los cuantiles de la distribución t con $n-1$ grados de libertad. Muestre que la siguiente relación se cumple entre t y el estadístico $T_1 = (b-c)^2 / (b+c)$, correspondiente al test de Mc Nemar:

$$t^2 = \frac{(n-1)T_1}{n-T_1} \quad \text{ó} \quad T_1 = \frac{nt^2}{n-1-t^2}$$

O sea, a medida que T_1 se hace grande, t^2 también se hace grande, por lo que los dos tests rechazan H_0 para valores grandes de T_1 o t^2 , con las regiones críticas correspondientes.

Ejercicio 13 (*): Sean X_1, \dots, X_{2n} observaciones independientes tales que $X_i \sim F(x-\theta_i)$, con $F \in \Omega_0$ simétrica. Para testear

$$H_0: \theta_1 = \dots = \theta_{2n} \quad \text{vs} \quad H_1: \theta_1 \leq \dots \leq \theta_{2n}$$

con por lo menos una desigualdad estricta consideremos el test de Cox-Stuart, basado en

$$S = \sum_{i=1}^n s(X_{n+i} - X_i) \quad \text{donde} \quad s(x) = \begin{cases} 1 & \text{si } x > 0 \\ 0 & \text{si } x < 0 \end{cases}$$

a) Sean

$$p_i = P(X_{n+i} > X_i) \quad q_i = 1 - p_i \quad Z_i = s(X_{n+i} - X_i) - p_i$$

Pruebe que

$$\frac{S - \sum_{i=1}^n p_i}{\left[\sum_{i=1}^n p_i q_i \right]^{1/2}} \xrightarrow{d} N(0,1) \quad \text{si} \quad \sum_{i=1}^{\infty} p_i q_i \text{ diverge}$$

- b) ¿Cuál es la distribución de S bajo la hipótesis nula y la distribución asintótica?
 c) Muestre que S/n provee un test consistente siempre que

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n (p_i - 1/2) \rightarrow \infty$$

En particular el test es consistente si $\theta_{n+i} - \theta_i = \Delta > 0, \forall i = 1, \dots, n$.

(Sugerencia: Use la distribución dada en a))

Ejercicio 14: La tabla muestra las concentraciones de anticuerpos anti-Estreptococos Grupo B Tipo III en 20 voluntarios antes de la inmunización y 4 semanas después (Baker et al. N. Engl. J. Med. (1980), **303**:173-8).

- a) Los autores presentan en la publicación la tabla con los niveles de anticuerpos antes y después de la inmunización, y resumen la comparación de los niveles de anticuerpos como ' $t = 1.8; p > 0.05$ '. Comente la validez de este resultado.
 b) ¿Qué método hubiera sido el más apropiado para analizar estos datos? ¿A qué conclusión arriba utilizando este método?.

Voluntario	Antes	4 semanas después	Diferencia
1	0.4	0.4	0.0
2	0.4	0.5	0,1
3	0.4	0.5	0,1
4	0.4	0.9	0,5
5	0.5	0.5	0.0
6	0.5	0.5	0.0
7	0.5	0.5	0.0
8	0.5	0.5	0.0
9	0.5	0.5	0.0
10	0.6	0.6	0.0
11	0.6	12.2	11,6
12	0.7	1.1	0,4
13	0.7	1.2	0,5
14	0.8	0.8	0.0
15	0.9	1.2	0,3
16	0.9	1.9	1
17	1.0	0.9	-0,1
18	1.0	2.0	1
19	1.6	8.1	6,5
20	2.0	3.7	1,7

Ejercicio 15: Para los ejercicios 2, 3, 8, 9, 10 y 11, discuta la posibilidad de aplicar una técnica paramétrica. ¿Qué supuestos se necesitan?. ¿Qué supuestos se violan?.