Práctico sobre Tests de Hipótesis

1. Los tiempos de secado de una marca de pintura “A ”siguen (según el fabricante) una distribución normal de parámetros $\mu = 5$ horas y desvío $\sigma = 45$ minutos. La competencia dice que su pintura tiene tiempos de secado también normales con el mismo desvío y media = 4 horas y media. Se toman 20 piezas iguales de madera (misma madera, mismas dimensiones, mismo tenor de humedad, etc.) se eligen al azar 10 de ellas y se les aplica una mano de pintura A, a las otras 10 se les aplica una mano de pintura B y si miden los tiempos de secado de cada pieza, obteniendo los siguientes datos:

A					
	4.21	5.09	4.05	6.08	5.17
	5.27	5.12	5.26	5.25	6.30

B					
	4.29	3.10	4.28	5.21	3.37
	3.15	5.07	4.03	4.47	4.50

   a) Hacer los diagramas de caja de cada conjunto de datos.
   b) Plantee un T.H. bilateral para ver si existe una diferencia significativa entre los tiempos de secado, hallando el p-valor del estadístico elegido. (Sugerencia: Si $X \sim N(\mu_x, \sigma_X^2)$ e $Y \sim N(\mu_Y, \sigma_Y^2)$, entonces $X - Y \sim N(\mu_X - \mu_Y, \sigma_X^2 + \sigma_Y^2)$).
   c) Resuelva dicho test “a mano ”y compare el resultado con el del test del paquete statistics de octave.
   d) Resuelva el test mediante simulaciones: genere muchas veces 10 datos con distribución normal de parámetros $\mu = 0h$, $\sigma = 6/4h$, para cada conuunto de 10 datos calcule el estadístico del test y decida si el valor observado en los datos de la letra tiene una probabilidad de ocurrencia razonable o si es muy poco probable haber observado bajo H_0 lo efectivamente observado.

2. Las distancias de frenado de un vehículo a 60 km/h cuando se usan frenos de la marca “A ”siguen (según el fabricante) una distribución normal de parámetros $\mu = 25$ mt. y desvío $\sigma = 5$ mt. La competencia dice que la

distancia de frenado al usar su sistema de frenos “B ”también tiene una distribución normal con el mismo desvío y media = 22mt.

Se toman 20 vehículos iguales, se eligen al azar 10 de ellos y se les equipa con los frenos A, al resto se les equipa con los frenos B y si miden las distancias de frenado, obteniendo los siguientes datos:

19.60	32.70	12.25	26.28	28.06	20.96	29.62	30.88	29.82	28.95
23.84	28.03	22.78	22.59	20.67	28.10	29.40	25.26	23.34	26.30

Tab. 1: Distancias de frenado con el sistema A

	1	2	3	4	5	6	7	8	9	10
1	30.99	17.55	17.50	19.16	12.28	24.09	25.19	31.98	25.23	16.93
2	13.78	20.08	22.41	22.30	19.63	32.00	20.06	18.67	22.56	21.69

Tab. 2: Distancias de frenado con el sistema B

a) Hacer los diagramas de caja de cada conjunto de datos.

b) Igual que en el problema anterior.

c) Igual que en el problema anterior.

e) Resuelva el test mediante simulaciones siguiendo los pasos del problema anterior.

3. En una empresa se introduce un cambio en el proceso de fabricación de un cierto producto. Para investigar si dicho proceso supone una mejora en la calidad, se hace un control de calidad: se seleccionan al azar 50 cajas (cada una contiene 50 unidades de un cierto producto), se examina cada una de las 50 unidades y se anota la cantidad de unidades defectuosas de la caja. El resultado es :

0	0	0	2	0	0	0	1	1	0
0	0	1	1	0	0	0	0	0	0
0	0	0	0	1	2	0	0	0	1
1	1	1	0	0	0	0	0	0	0
0	0	1	0	2	3	1	1	0	1

Tab. 3: Cantidad de defectuosos por caja

a) Elabore el diagrama de tallo y hojas de los datos.

b) Es razonable pensar que la distribución de los datos es $Bin(50, p)$. Se considera que el nuevo proceso representa una mejora sobre el anterior si $p \leq 0.01$ (0.01 era la proporción de defectuosos con el proceso anterior). Basado en estos datos, el nuevo proceso representa una mejora? Plantee el T.H. que usaría para tomar una decisión.

c) Encuentre las regiones de aceptación para $\alpha = 0.1$ y $\alpha = 0.05$.

d) Realice el test mediante simulación: simule muchos juegos de 50 datos $\sim Bin(50, 0.01)$, para cada uno calcule el valor del estadístico del test, y vea si la probabilidad de obtener dicho valor es aceptable o es muy pequeña.

4. Se tiene los siguientes datos sobre duración de lámparas LED (en horas):

869.39	177.11	735.20	721.46	850.72
636.72	624.48	539.24	4388.68	1376.25
1656.88	1397.89	523.88	254.30	304.92
263.11	584.39	360.07	139.33	88.86
567.65	948.83	918.23	1328.33	3112.35

Tab. 4: Duración en horas

a) Confeccione el diagrama de tallo y hojas y diagrama de caja para estos datos.

b) Estime la media y la varianza por el método de máxima verosimilitud.

c) Se presume que la distribución de los datos es $\mathcal{E}(\lambda)$ y se desea contrastar la hipótesis $\begin{cases} H_0: & \lambda = 1000 \\ H_1: & \lambda \neq 1000 \end{cases}$

c1) Repita lo anterior pero ahora con hipótesis alternativa unilateral: $\begin{cases} H_0: & \lambda = 1000 \\ H_1: & \lambda > 1000 \end{cases}$

c2) Realice dicho test mediante simulación: Genere muchas muestras de 25 datos $\sim \mathcal{E}(1000)$, calcule para cada vez el estimador de máxima verosimilitud para λ y evalúe si la probabilidad del valor obtenido a partir de la tabla (5) es aceptable o no.

d) Se mide la vida media de 100 lámparas de la misma marca que las anteriores, obteniéndose un valor de 850 horas. Asumiendo que la distribución es exponencial, contrastar la hipótesis de que la vida media de la población de la que se extrajeron esas 100 lámparas es de 1000 horas. Cuál es el p-valor del test?

5. En el armado de motores de combustión interna, es importante que los pistones encajen bien en las camisas de los cilindros.

Si el diámetro es demasiado pequeño, se perderá compresión, mientras que si es demasiado grande habrá que rectificarlos para que encajen.

Quienes hacen el control de calidad establecen como criterio de aceptación que los diámetros provengan de una distribución normal con una media de 8 cm. y un desvío de 1mm.

Para comprobar la calidad de un lote, se eligen 25 pistones al azar y se mide el diámetro, obteniendo:

7.57	8.17	8.06	7.95	7.96
8.28	8.12	8.14	7.82	8.09
7.79	8.02	8.34	8.08	8.06
7.90	8.09	7.80	8.05	8.05
8.05	7.76	7.87	8.14	7.98

Tab. 5: Diámetros (en cm.) de la muestra de pistones

a) Confeccione el diagrama de tallo y hojas y diagrama de caja para estos datos.

b) Estime la media y la varianza por el método de máxima verosimilitud.

c) Se desea contrastar $\begin{cases} H_0: & \mu = 8 \\ H_1: & \mu \neq 8 \end{cases}$ (σ conocido).

c1) Realize dicho test “a mano ”. Esto es, escriba el estadístico que usaría, su distribución bajo H_0, el intervalo de confianza al 95% y el p-valor del test (puede usarse octave para las cuentas).

c2) Realice dicho test usando la función que octave trae implementada, compare los resultados.

d) Realize el test mediante simulación, simulando 10000 veces 25 datos $\sim N(8, 0.1)$.

e) El lote es aceptable para $\alpha = 0.05$?

6. Se tienen los siguientes datos de rendimiento por hectárea (en quintales) de un cierto cultivo:

5.89	4.37	5.40	5.28	6.41
6.21	6.88	6.11	7.25	7.72
5.08	3.26	5.83	8.28	3.41
4.22	6.75	3.93	7.15	7.73
4.12	7.29	5.35	4.45	7.07

Tab. 6: Rendimientos por hectárea

a) Confeccione el diagrama de tallo y hojas y diagrama de caja para estos datos.

b) Se presume que la distribución de los datos es $N(\mu, \sigma^2)$ y se desea contrastar $\begin{cases} H_0: & \mu = 6 \\ H_1: & \mu \neq 6 \end{cases}$ (σ desconocido).

b1) Realize dicho test "a mano ". Esto es, escriba el estadístico que usaría, su distribución bajo H_0, el intervalo de confianza al 95% y el p-valor del test (puede usarse octave para las cuentas).

b2) Realice dicho test usando la función que octave trae implementada, compare los resultados.

b3) Repita lo anterior pero ahora con hipótesis alternativa unilateral:
$\begin{cases} H_0: & \mu = 6 \\ H_1: & \mu < 6 \end{cases}$

c) Realice ambos tests mediante simulación generando 10000 juegos de 25 datos c/u.