

UNIDAD 3 –UNIDAD 4: Estimación de parámetros y Pruebas de Hipótesis

TRABAJO PRÁCTICO –Parte 2

1. A partir de un examen de habilidades matemáticas realizado por 12 estudiantes de primer año seleccionados al azar se obtuvieron las siguientes calificaciones:

6,71 5,90 7,12 7,20 6,83 6,05 7,26 7,45 6,66 7,16 5,64 7,05

Suponiendo que la población se distribuye normal calcule:

- Un intervalo del 90 % de confianza para la media poblacional. Interprete.
- Calcule un intervalo del 95% de confianza para la varianza poblacional y para el desvío estándar poblacional.

2. Una empresa automotriz está interesada en estimar la proporción de cigüeñales de motor defectuosos que le adquiere al proveedor A. Toma una muestra aleatoria de 800 cigüeñales (adquiridas a A), en la que resultaron 80 unidades defectuosas.

- Construya un intervalo del 95% de confianza para la proporción poblacional de cigüeñales defectuosos correspondientes al proveedor A. Interprete.
- Para comparar con otro proveedor, toma una muestra aleatoria de 200 unidades de este proveedor (B), en la que encuentra 30 cigüeñales defectuosos. A un nivel de significación del 5% ¿Puede concluir que los cigüeñales del proveedor A son mejores que los del B? Calcule el valor P.

3. En cada una de las siguientes situaciones indicar si se trata de una comparación de medias poblacionales de muestras independientes o no.

- Se eligieron diez adultos con sobrepeso para evaluar una dieta, cada persona se pesó antes de empezar, y nuevamente después de seguir la dieta durante 12 semanas. Se quiere probar el efecto de la dieta.
- Un granjero del Medio Oeste realiza un experimento para determinar si el uso de un aditivo químico especial en el fertilizante que ha estado usando para hacer crecer la soja acelera el crecimiento de las plantas. Se eligieron al azar 15 parcelas para el estudio. En cada una se trataron dos plantas de soja cercanas, una con el fertilizante acostumbrado y la otra con el que se le añadió el aditivo. El crecimiento de cada planta se midió en pulgadas después de un período de cuatro semanas.
- En un experimento para determinar si las personas afectadas con glaucoma tienen córneas anormalmente gruesas, se hizo el estudio a 16 personas. Ocho tenían glaucoma y ocho no. Se midió en micras el grosor de la córnea de cada sujeto.
- Se realiza un experimento para comparar las características de llenado del equipo de embotellado de dos diferentes fábricas vinícolas.

4. Suponga que se quiere comparar la enseñanza tradicional de la Geometría del plano con la enseñanza con soporte informático (por ejemplo, con GeoGebra). Para medir los resultados se usarán pruebas estándar.

- Describa cómo debería realizarse la experiencia para obtener dos muestras independientes.
- Describa cómo debería realizarse la experiencia para obtener dos muestras dependientes.

5. Se diseñó un estudio para comparar las actitudes de dos grupos de estudiantes de medicina hacia las computadoras. El grupo A había realizado un curso de estadística que implicaba interacción con la computadora a través de la utilización de un paquete de análisis de datos. El grupo B tomó un curso de estadística que no incluía el uso de computadoras. Las actitudes de los estudiantes se midieron mediante la administración del *Computer Anxiety Index* (CAIN) (menor puntaje indica actitud más positiva). Los datos obtenidos se resumen a continuación:

Grupo A (con computadora)	n=10	Media=60,3	s(x) = 5,5
Grupo B (sin computadora)	n=15	Media = 67,2	s(x) = 4,1

Suponiendo que las poblaciones se distribuyen en forma normal y que las varianzas se pueden considerar iguales: Construir un intervalo del 95% de confianza para la diferencia de medias poblacionales. Interprete.

6. Se realiza un experimento para comparar las características de llenado del equipo envasador de aceite refrigerador de dos destilerías diferentes. Para ello se seleccionan al azar 12 barriles de la destilería A y otros 16 barriles de la destilería B, resultando para A un promedio de 75,2 lt. y un desvío estándar de 0,36 lt. y para B un promedio de 75,5 lt. con un desvío estándar de 0,4 lt. Suponiendo que el volumen de llenado para ambas destilerías tiene distribución normal:

- ¿Puede concluir que existe diferencia entre las variancias? Aplique el test adecuado para corroborarlo.
- Se puede decir que la media de contenido poblacional de ambas destilerías son iguales? Utilice un nivel de significación del 5%. Calcule el valor P.
- Halle un intervalo de confianza del 95% para la diferencia verdadera entre los contenidos medio poblacional de ambas destilerías. Es compatible este intervalo con la respuesta del inciso b) Justifique.

7. a) ¿En base a qué supuestos se puede utilizar la distribución F para hacer inferencias sobre el cociente de variancias poblacionales?

b) El gerente de una empresa láctea piensa adquirir una nueva máquina embotelladora y considera los modelos fabricados por las compañías A y B. Si la durabilidad, el costo y la conveniencia son equiparables en los dos modelos, el factor decisivo es la variabilidad de los llenados (se prefiere un modelo cuyos llenados tengan menor variancia). Sean σ_A^2 y σ_B^2 las variancias de los llenados que producen A y B respectivamente. Analicemos diversas pruebas de la hipótesis nula $H_0: \sigma_A^2 = \sigma_B^2$. Al tomar muestras de llenado de las dos máquinas y utilizar el estadístico de prueba S_A^2 / S_B^2 , se podría utilizar como región de rechazo la región superior, inferior o dos colas de la distribución F dependiendo de los fines que se persigan. Identifique el tipo de región de rechazo que preferirán las siguientes personas y explique la razón:

Un vendedor de la compañía A.

Un vendedor de la compañía B.

El gerente de la empresa láctea.

8. A todos los estudiantes que se inscriben en un curso para ejercitar la memoria se les aplica una prueba antes del inicio del curso. A una muestra aleatoria de 10 estudiantes que terminaron este curso se les aplicó una post evaluación. Las calificaciones se muestran a continuación:

Estudiante	1	2	3	4	5	6	7	8	9	10
Antes	93	86	72	54	92	65	80	81	62	73
Después	98	92	80	62	91	78	89	78	71	80

Suponiendo que la calificación tiene una distribución normal, probar la hipótesis de que el curso fue beneficioso, a un nivel de significación del 1%. Calcule el valor P.

9. En un control de calidad de una fábrica se decide inspeccionar si el tiempo de ensamblaje (en minutos) tiene una distribución aproximadamente normal y para ello se seleccionaron al azar 200 registros de los mismos cuyos resultados se volcaron en la siguiente tabla:

i	X_i	f_i
1	(22-26]	6
2	(26-30]	28
3	(30-34]	48
4	(34-38]	65
5	(38-42]	38
6	(42-46]	13
7	(46-50]	2

Realice el test de hipótesis que le permita dar respuesta a la suposición de normalidad del tiempo de ensamblaje. Concluya.

10. En la producción de cierto material para soldar se ha establecido que la desviación estándar de la tensión a la ruptura es de 15 lb como máximo.

a) A partir de la información del archivo *tensionruptura.xlsx* correspondiente a una muestra aleatoria de 16 piezas de ese material ¿Qué puede decir de la hipótesis establecida?

b) Evalúe los supuestos necesarios para que la prueba anterior sea estadísticamente correcta.

c) Calcule un intervalo de confianza del 95% para la tensión media a la ruptura de ese material (tome en cuenta el resultado del inciso a) Interprete.

11. Los datos del archivo *resistenciaflexion.xlsx* corresponden a la resistencia a la flexión de dos muestras independientes de 20 juntas seleccionadas al azar sin recubrimiento y 25 juntas seleccionadas al azar con recubrimiento.

a) Verifique que la resistencia a la flexión tiene distribución normal según tipo de junta.

b) Calcule un intervalo de confianza del 95% para la resistencia promedio verdadera para cada tipo de junta.

c) Calcule un intervalo de confianza del 95% para la diferencia entre las resistencias promedio verdaderas de los dos tipos de juntas.

d) Se quiere probar que la resistencia de las juntas es mayor cuando no tienen recubrimiento lateral, plantear la prueba correspondiente con un nivel de significación del 0,05 y concluir al respecto.

e) Comentar acerca de la relación de los resultados encontrados en los 3 incisos anteriores.

12. Se quiere comparar los componentes electrónicos para las notebooks de dos proveedores. Para ello se eligen al azar 10 componentes del proveedor A y 15 componentes del proveedor B, midiéndose el tiempo hasta el fallo (en horas) La información está en el archivo *TiempoHastaFallo.xlsx*

a) Hacer dos gráficos de caja superpuestos y dos gráficos Q-Q plot del tiempo hasta el fallo según proveedor.

b) En base a estos gráficos puede suponer poblaciones normales?

c) Realice la prueba de Bondad de ajuste y concluya respecto a la normalidad del tiempo hasta el fallo de los componentes electrónicos de cada proveedor.

d) Con la información disponible, puede considerar homogeneidad de las variancias poblaciones de ambos proveedores? Aplique el test adecuado para corroborarlo.

e) ¿Presentan los datos suficiente evidencia para concluir que existe diferencia significativa entre los tiempos medios hasta el fallo de los componentes de los dos proveedores?

f) Construya un intervalo de confianza para la diferencia verdadera entre los tiempos medios hasta el fallo de ambos proveedores.

g) Halle un intervalo de confianza para el valor medio del tiempo hasta el fallo de los componentes de cada proveedor.

h) Compare los resultados obtenidos en los incisos d, e y g y diga si hay coherencia entre ellos y por qué.

13. Se realizó un estudio para estudiar la relación entre el hecho de que una gestante fume durante el embarazo y que el bebé presente bajo peso al nacer. Para ello, se hizo un seguimiento sobre una cohorte de 2000 gestantes, a las que se les interrogó sobre su hábito de fumar durante la gestación y se registró el peso de los recién nacidos, clasificándolos con bajo peso o no. Los resultados se volcaron en el archivo *Gestante.xlsx*

a) Aplique la prueba chi-cuadrado para resolver el problema planteado. Concluya.

b) ¿Se le ocurre alguna otra forma de solución? Sólo diga cuál y plantee la hipótesis correspondiente.

14. En un estudio para investigar el límite al que las personas perciben olores industriales en cierta región ("Annoyance and Health Reactions to Odor from Refineries and Other Industries in Carson, California", Environmental Research, 1978: 119-132), se obtuvo una muestra de personas de cada una de las tres regiones cercanas a instalaciones industriales: Región A, B y C.

Se le preguntó a cada persona seleccionada si percibía olores: Todos los días, al menos una vez a la semana, al menos una vez al mes, menos de una vez al mes, o nada en absoluto. Los datos se encuentran en el archivo *Ej14TP3.xlsx*.

- a) Exprese en forma simbólica y coloquial las hipótesis apropiadas para esta situación.
- b) Efectúe la prueba correspondiente y concluya al respecto.