

Laboratorio 10 - Inferencia Estadística

Pruebas de Hipótesis pt. 2 - Planteamiento de las hipótesis

Laboratorista: Héctor Lira Talancón

Ago-Dic 2017

1. Algunos analistas de mercado sostienen que últimamente el riesgo al invertir en el ramo de la construcción ha disminuido. Los resultados que se presentan a continuación corresponden a la variable W : rendimiento mensual (%), después de obtener una muestra aleatoria de este tipo de inversiones.

Datos: $n = 100$, $\min = 7.6125$, $\max = 28.3334$, $Q_1 = 13.8849$, $Q_2 = 16.2717$, $Q_3 = 19.5843$,
 $s = 4.3248$, $C_A = 0.8178$, $C_K = 0.2719$

a) Hasta el año pasado, el riesgo (desviación estándar) promedio había sido de 5%. ¿La información de la muestra confirma lo que sostienen los analistas? Plantear las hipótesis adecuadas y realizar la prueba correspondiente. Usar $\alpha = 0.025$.

b) Calcular la potencia de la prueba cuando $\sigma = 4$.

2. Un profesor de Probabilidad del ITAM asegura que sus alumnos obtienen mejores calificaciones en el segundo examen parcial en comparación con el primero. Él seleccionó una muestra aleatoria de 10 alumnos para los cuales sus calificaciones del primero y segundo examen parcial fueron: $\bar{X} = 58.1$, $\bar{Y} = 65.2$, $S_X = 24.5241$, $S_Y = 20.0322$, $\sum_{i=1}^{10} X_i Y_i = 40361$, $\sum_{i=1}^{10} (X_i - Y_i) = -71$, $\sum_{i=1}^{10} (X_i - Y_i)^2 = 4569$.

a) La información anterior presenta evidencia de lo que asegura el profesor? Usar el valor-p.

b) El profesor también asegura que los alumnos que obtienen altas (bajas) calificaciones en el primer examen parcial, suelen tener altas (bajas) calificaciones en el segundo examen. Comprobar estadísticamente lo anterior. Usar el valor-p.

3. Una muestra aleatoria de 40 observaciones de la variable $Y \sim N(\mu, 20)$ dio como resultado $\bar{Y} = 14, S_Y = 4.3$. Realice la prueba $H_0 : \mu = 15$ vs. $H_1 : \mu \neq 15$ usando exclusivamente un intervalo de confianza adecuado. El nivel de significancia deseado es 0.08.

4. Se desea probar $H_0 : \mu = 8$ vs. $H_1 : \mu = 3$, en donde μ es la media de una población $N(\mu, 9)$. Determinar el valor del tamaño de muestra n de tal forma que $P(ET1) = 0.02$ y $P(ET2) = 0.04$.

5. Establezca la verdad o falsedad de la siguiente afirmación: Si las pruebas de hipótesis se hicieran con un nivel de significancia de 0.07, la hipótesis nula generalmente no se rechazaría cuando ésta es verdadera.

6. Establezca la verdad o falsedad de la siguiente afirmación: Sean $H_0 : p \leq 0.45$ vs. $H_1 : p > 0.45$, donde $p = P(\text{una solicitud de crédito sea rechazada})$. Si para una muestra de tamaño 8 se va a rechazar H_0 cuando el número de solicitudes rechazadas es mayor o igual a 6, entonces el nivel de significancia de la prueba es menor a 0.02.

7. Sea X el ahorro mensual que hace una persona en la sucursal A, y Y el ahorro mensual que hace una persona en la sucursal B. La dirección del Banco del Ahorro sospecha que hay diferencia entre la variabilidad de X y Y . Se planea un estudio en donde se seleccionan aleatoriamente 13 y 16 cuentahabientes de las sucursales A y B, respectivamente. Los resultados fueron:

	N	Media	Desv.Est.
Suc. A (X)	13	2.488	1.6
Suc. B (Y)	16	2.926	1.1

¿La información muestra confirma la sospecha del Banco? Plantear las hipótesis adecuadas y realizar la prueba correspondiente. Usar un nivel de significancia igual a $\alpha = 0.02$. ¿Cuál es el valor-p? ¿Cuáles son los supuestos requeridos para esta prueba?

8. En cada una de dos zonas del país se selecciona una muestra aleatoria de Casas de Cambio. A continuación se presenta la cantidad de dólares (miles) que vendió cada Casa de Cambio en un día.

								\bar{X}	S
Zona A	15	13.6	12.1	11.7	11.5	15.1	13.3	13.19	1.49
Zona B	10.5	13.1	11.5	14.7	10.9	9.4	9.2	11.33	1.99

- a) ¿La variabilidad en la cantidad de dólares que se venden al día es igual en las dos zonas? Use $\alpha = 0.06$.
- b) Independientemente de la respuesta del inciso anterior, probar si las casas de cambio de la zona A venden más dólares (en promedio) que las de la zona B. Use $\alpha = 0.05$.
9. Los resultados de la encuesta antes del día de elecciones en cuanto a las personas a favor de un determinado candidato presidencial, afirman que existe mayor preferencia en el nivel socioeconómico alto que en el nivel socioeconómico bajo por dicho candidato. Si te contratan para evaluar dicha afirmación, ¿qué prueba de hipótesis debes plantear?
- Diferencia de medias para muestras independientes.
 - Diferencia de medias de datos apareados (muestras dependientes).
 - Diferencia de proporciones de una sola cola.
 - Diferencia de proporciones de dos colas.
10. La longitud de los pasos de una persona es una variable aleatoria con media 80 cm. y desviación estándar 10 cm. Demuestre que para estar 99% seguro que avanzará 200 metros la persona tiene que dar 255 pasos. Establezca claramente los supuestos que utilice.
11. Suponga que una mejor medida sobre la efectividad de una campaña publicitaria para determinado artículo es el aumento en las ventas de dicho artículo, es decir, $D_i = X_i - Y_i$, para $i = 1, \dots, 101$ donde Y_i es la venta del artículo en la i -ésima tienda medida antes de empezar la campaña y X_i es la venta del artículo después de la campaña. Si la media muestral del cambio en ventas es 5.3 (en millones de pesos) y la varianza muestral del cambio en ventas es 144.0, establezca paso a paso (planteamiento de la hipótesis, región de rechazo, estadístico de prueba, etc.) la prueba de hipótesis que propondría para validar la afirmación de que el cambio promedio en ventas es mayor a 3 (millones de pesos).
12. La gerencia de un banco desea comparar dos sistemas de cómputo para atención a clientes. El sistema 1 es el actual y el sistema 2 es el nuevo. En condiciones similares se obtienen dos muestras aleatorias independientes de tamaño 70 de los tiempos de servicio con cada uno de los dos sistemas.
- Describe verbalmente los parámetros poblacionales de interés y enuncia las hipótesis nula y alternativa apropiadas para ver si el nuevo sistema mejora la situación del servicio.
 - Describe la regla de decisión para concluir que hay evidencia de que el nuevo sistema es mejor que el sistema actual con 3% de significancia.

13. Se tiene un proceso de fabricación que se vigila en forma periódica tomando muestras de 20 manufacturas que se consideran como muestras aleatorias. Cada manufactura de la muestra se clasifica como defectuosa o no, y supondremos que esta clasificación se hace sin error. Si la muestra arroja evidencia de que la proporción de defectuosos en el proceso es mayor a 15% se detiene la producción y se revisa el proceso para encontrar el posible problema.
- Describe el parámetro poblacional de interés y enuncia las hipótesis nula y alternativa apropiadas.
 - Encuentra una regla para detener el proceso de forma que la probabilidad de que se paralice innecesariamente sea lo más cercana a el 5% de las veces.
 - Calcula la probabilidad de que se detecte correctamente que la proporción de defectuosos en el proceso sea del 25%.

14. En una universidad se tomó una muestra aleatoria de 50 alumnos del último semestre y resultó que 21 estaban haciendo su servicio social. Se desea ver si al menos el 30% de los alumnos del último semestre de la universidad están haciendo su servicio social. Define el parámetro de interés, formula las hipótesis nula y alternativa, aproxima el valor-p y úsalo para hacer la prueba con un 5% de nivel de significancia. Escribe tu conclusión en términos del contexto del problema.

15. Considera un proceso industrial de horneado de galletas. El personal de Investigación y Desarrollo está ensayando una nueva masa que produzca menor variación en el diámetro de la galleta, ya que una galleta demasiado grande presenta serios problemas de empaque. Se considera que el diámetro tiene distribución normal. Las muestras se obtuvieron independientemente una de la otra. Los resultados del ensayo fueron:

Masa	Tamaño de muestra	Promedio	Desv.Est.
Original	21	27.3mm	1.9mm
Nueva	20	28.5mm	1.3mm

¿Hay evidencia al 10% de significancia que la nueva masa reduce la variación en el diámetro de la galleta? Define los parámetros de interés, plantea las hipótesis nula y alternativa correspondientes y lleva a cabo la prueba.

16. Una empresa de bienes raíces desea comparar las características entre dos colonias. Para eso obtiene una muestra aleatoria de casas de cada colonia y entre otras cosas registra las siguientes variables: X = el precio de la casa en millones de pesos y Y = una variable indicadora que vale 1 si la casa tiene más de dos recámaras y 0 de otra forma. Un resumen de los resultados está en la siguiente tabla:

Colonia	n	\bar{X}	S_X	\bar{Y}
A	71	3.7	1.3	0.45
B	121	4.3	2.5	0.65

- a) Si se desea comparar los precios, define verbalmente los parámetros de interés para esta cuestión.
- b) Plantea las hipótesis nula y alternativa relevantes para este problema.
- c) Haz la prueba con un 5% de significancia.
- d) Obtén la conclusión en términos del contexto del problema.
- e) Si se desea probar que la proporción de casas que tienen más de dos recámaras en la colonia B es mayor a la correspondiente proporción en la colonia A, define verbalmente los parámetros de interés para este problema.
- f) Plantea las hipótesis nula y alternativa relevantes para este problema.
- g) Haz la prueba con un 5% de significancia.
- h) Si se desea comparar la variabilidad de los precios entre ambas colonias, define verbalmente los parámetros de interés para esta cuestión.
- i) Plantea las hipótesis nula y alternativa relevantes para este problema.
- j) Suponiendo normalidad de los precios de las casas haz la prueba con un 5% de significancia.

17. Se tiene un proceso productivo que se intenta mejorar. Un equipo de ingenieros se reunió y propuso cambios en la manera de llevar a cabo el proceso. Se obtuvo una muestra de 120 manufacturas antes de llevar a cabo los cambios y se detectaron 42 defectuosos. Después de los cambios se tomó una muestra de 190 manufacturas y se detectaron 48 defectuosos.

- a) Plantea las hipótesis nula y alternativa que consideres convenientes para evaluar si hubo una mejora en el proceso.
- b) Haz la prueba de hipótesis que consideres conveniente con un 5% de significancia y obtén una conclusión con los datos aportados por los ingenieros.

18. De acuerdo a la filosofía de calidad de E. Deming la reducción en variabilidad es sinónimo de mejor calidad. Un equipo de ingenieros se reunió para mejorar un proceso de ensamblaje desde este punto de vista. Se seleccionó una operación de pintura de carrocerías y se eligió como característica de interés el espesor promedio de un panel lateral. Se eligieron aleatoriamente 7 autos de la producción de un día antes de los cambios propuestos y 8 autos después de los cambios propuestos obteniéndose los siguientes resultados:

									\bar{X}	Desv. Est.	Q_1	Q_2	Q_3
Antes	6.6	6.5	2.3	1.5	6.8	11.7	3.2		5.5	3.52	2.75	6.5	6.7
Después	5.4	4.6	5.5	5.6	2.7	4.4	5.4	6.3	4.98	1.11	4.5	5.4	5.55

- a) Plantea las hipótesis nula y alternativa que consideres convenientes para evaluar si hubo un cambio significativo.

- b) Haz la prueba de hipótesis que consideres conveniente con un 10% de significancia y obtén una conclusión con los datos aportados por los ingenieros.
- c) Menciona tres supuestos implícitos en la elección que hiciste en el inciso c).

19. Se quiere comparar la compresión del suelo de dos tipos de suelo. Para dos muestras de suelo se mide X = compresión en toneladas por pie cuadrado. Se obtuvieron los datos de la siguiente tabla.

Tipo de suelo	n	\bar{X}	S_X
A	71	1.65	0.26
B	121	1.43	0.48

- a) ¿Hay evidencia de diferencia significativa entre la compresión promedio para los dos tipos de suelo? Plantea las hipótesis relevantes y haz la prueba con un nivel de significancia del 5%.
- b) Calcula el valor-p correspondiente y ahora haz la prueba de hipótesis con un nivel de significancia del 5% con este dato.

20. Se lleva a cabo una auditoría por muestreo aleatorio simple (MAS) a las 10350 operaciones llevadas a cabo por una empresa. Se tomó una muestra de 120 operaciones las cuales se analizaron detenidamente. De entre las variables que se registraron estuvieron:

X = Ingreso total de la operación; y

W = Variable indicadora que vale uno si en la operación se detectaron irregularidades.

\bar{X}	S_X	W
5300	795	0.15

- i) ¿Hay evidencia de que la proporción poblacional de las operaciones con irregularidades es mayor al 10%? Usa un nivel de significancia del 5%.

- ii) Considera las siguientes afirmaciones:

A. Para justificar el procedimiento del inciso i) se requiere que W tenga distribución normal.

B. El tamaño de muestra es uno de los elementos que justifican los cálculos hechos en i).

Entonces,

- a) A y B son apropiados.
- b) A es apropiado pero B no.
- c) A no es apropiado pero B sí.
- d) A y B no son apropiados.

21. Se sabe que el peso de los contenedores marítimos que se apegan a las prácticas internacionales se distribuye como una normal con desviación estándar de 5 toneladas. A partir de una muestra aleatoria grande de contenedores se quiere probar $H_0 : \mu = 30$ vs. $H_1 : \mu > 30$. ¿Qué tamaño de muestra garantiza $\alpha = \beta(32) = 0.01$ para esta prueba?
- a) $n = 27$
b) $n = 68$
c) $n = 136$
d) Ninguno de los anteriores.
22. Una empresa de tiendas de conveniencia desea ver el impacto de una campaña publicitaria en las ventas de una línea de productos. Tiene 417 establecimientos y se seleccionó aleatoriamente con reemplazo y probabilidades iguales una muestra de 10 establecimientos. De manera piloto hubo despliegue de la campaña solo en esas 10 tiendas. Se midieron las ventas del producto en el mes previo y en el posterior al despliegue de la campaña publicitaria. Se eligieron meses en los cuales las ventas del producto son similares.
- a) Usa Y para probar la hipótesis de si la campaña tiene un impacto positivo sobre las ventas con un nivel de significancia lo más cercano a 5%. Indica la hipótesis nula y alternativa, el estadístico de prueba, la regla de decisión y tu conclusión si se observó $Y = 8$.
b) Considera $p = P(X_i = 1)$. Calcula la potencia de la prueba del inciso a) cuando p toma los valores 0.75 y 0.95.
23. Un equipo de ingenieros se reunió para mejorar un proceso de producción desde el punto de vista de reducción de variabilidad. Se seleccionó una operación de llenado de botellas y se eligió como característica de interés la cantidad de líquido en el recipiente. Se propone elegir aleatoriamente 31 botellas un día antes de los cambios propuestos y 25 botellas después de los cambios. Alguien del equipo dice que es suficiente con que la desviación estándar muestral después de los cambios sea menor que la de antes para asegurar que hay mejora en el proceso. ¿Qué tan menor (en porcentaje) debe ser la desviación estándar muestral después de los cambios respecto a la de antes para que haya evidencia de mejora con un nivel de significancia del 5%?
24. Los siguientes datos provienen de una encuesta nacional de los EUA donde se estudió la relación que hay entre el peso del bebé al nacer (X) y los hábitos de fumar de la madre. Los datos recopilados fueron:

fumó	n	X	Desv. Est.
0	1176	120.1	20.3
1	212	111.1	19.2

Prueba la hipótesis de si la proporción de mujeres que no fuma es mayor al 80%. Usa un nivel de significancia del 5%.