

PROBABILIDADES Y ESTADÍSTICA (C)

PRÁCTICA 9

1. Un productor agropecuario cultiva una variedad de soja cuyo rendimiento por hectárea (ha.) puede considerarse una variable aleatoria con distribución $N(37, 25)$. Un vendedor de granos pone a la venta una nueva variedad de soja, diciendo que tiene un rendimiento mayor. El productor agropecuario realizó una compra de esta nueva variedad y quiere asegurarse de que tienen un rendimiento mayor que las semillas que compra habitualmente. Con ese fin decide cultivar 10 parcelas de 1 ha., observar sus rendimientos y aplicar un test de hipótesis de nivel 0.05. Suponiendo que el rendimiento por ha. de la nueva variedad de soja tiene una distribución normal con varianza 25, resolver los siguientes ítems.
 - (a) Definir las variables aleatorias y los parámetros involucrados en este problema y plantear el test de hipótesis que debe realizar el productor agropecuario.
 - (b) El productor cultiva las 10 parcelas y obtiene los siguientes rendimientos:
37 - 39.50 - 41.70 - 42 - 40 - 41.25 - 43 - 44.05 - 38 - 38.50
¿Cuál es la conclusión del test en base a estos datos observados? Calcular la probabilidad del error de tipo II suponiendo que el verdadero rendimiento medio de la nueva variedad de soja es de 40 toneladas por ha.
 - (c) Encontrar el test para el caso en que se cultiven n parcelas (con igual nivel).
 - (d) Determinar el número n de parcelas a cultivar para que la probabilidad del error de tipo II sea menor o igual que 0.05 suponiendo que el verdadero rendimiento medio de la nueva variedad de soja es de 40 toneladas por ha.
 - (e) Calcular la función de potencia $\pi(\mu)$ del test planteado en c), verificar que es creciente para $\mu \in \mathcal{R}$ y deducir que este test sirve también para testear

$$H_0 : \mu \leq 37 \quad H_1 : \mu > 37.$$

2. En la construcción de un edificio debe usarse un concreto que tenga una tensión media mayor a 300 psi. Para verificar si el concreto preparado a partir del cemento “Loma Blanca” cumple con este requerimiento, se realizan 15 mediciones en forma independiente de la tensión de este concreto. Se observa una media muestral de 304 psi y un desvío estándar muestral de 10 psi. El constructor está dispuesto a correr un riesgo del 5% de comprar cemento “Loma Blanca” cuando éste produce un concreto que no cumple con las especificaciones. Suponiendo que los datos están normalmente distribuidos:
 - (a) Plantear el test correspondiente. ¿Qué decisión se toma?
 - (b) Acotar el p -valor.

3. Se diseñó un nuevo sistema de riego de manera tal que el desvío del tiempo de activación sea menor que 6 segundos. Se lo prueba 11 veces, obteniéndose los siguientes tiempos de activación:
- 27 41 22 27 23 35 30 24 27 28 22
- Suponiendo que el tiempo de activación (en segundos) es una v.a. con distribución normal:
- ¿Usted decidiría, a nivel 0.05, que el sistema cumple la especificación?
 - Acotar el p -valor
 - Encontrar la función de potencia $\pi(\sigma^2)$, implementarla en R y graficarla. ¿Es creciente o decreciente? ¿Cuánto vale en $\sigma = 6$?
4. Se supone que 1 de cada 10 fumadores prefiere la marca A. Después de una campaña publicitaria en cierta región de ventas, se entrevistó a 200 fumadores para determinar la efectividad de la campaña. El resultado de esta encuesta mostró que 26 personas preferían la marca A.
- ¿Indican estos datos, a nivel aproximado 0.05, un aumento en la preferencia por la marca A?
 - Calcular el p -valor aproximado.
 - ¿Cuál es la probabilidad aproximada de decidir que la campaña publicitaria no fue efectiva, cuando en realidad la proporción de preferencia por la marca A después de la campaña es 0.15?
 - Implementar la función potencia y graficarla. ¿Qué características tiene esta función? Interpretar.
 - ¿Qué tamaño de muestra debería tomarse para que la probabilidad de c) fuese a lo sumo 0.05?
 - Decidir si las siguientes afirmaciones son V o F . Justificar.
 - Si la verdadera proporción de fumadores que prefieren la marca A es mayor que 0.15, la probabilidad de que este test rechace H_0 es mayor que 0.725.
 - Cuánto más pequeña es la verdadera proporción de fumadores que prefieren la marca A más improbable es que el test rechace H_0 .
 - $\pi(0.1) = 0.05$.
 - La probabilidad de error de tipo II es una función creciente de p , donde p es la verdadera proporción de fumadores que prefieren la marca A.
5. Se supone que el tiempo de duración de cierto tipo de lamparitas tiene distribución exponencial. Una fábrica garantiza que el tiempo medio de vida de las lamparitas que produce es mayor que 50 días, y la empresa vendedora quiere asegurarse que la producción satisface las especificaciones antes de sacarla a la venta. Para ello toma

al azar una muestra de 40 lamparitas y observa el tiempo de duración de las mismas, obteniendo un promedio de 53 días.

- (a) Proponer un test de nivel exacto con hipótesis nula simple para este problema de manera tal que la probabilidad de que la empresa decida no vender si no se satisfacen los requerimientos sea del 0.95.
 - (b) ¿Qué decisión se toma en base a los tiempos observados?
 - (c) Repetir a) usando un nivel aproximado.
 - (d) Calcular la función de potencia aproximada para el test del ítem anterior. ¿Este test conserva el nivel aproximado 5% si ampliamos la hipótesis nula, reemplazándola por una hipótesis compuesta?
 - (e) Utilizando el test aproximado, ¿qué probabilidad tiene la empresa de no sacar la producción a la venta, si el promedio de vida verdadero es 52 días?
 - (f) ¿Cuál debería ser el tamaño de la muestra para que la probabilidad de d) fuese 0.1?
6. Se desea determinar si un dado de seis caras está cargado o no. Para ello se arroja el dado 1000 veces obteniéndose los resultados contenidos en el archivo **dado.txt**, que se encuentra en la página web de la materia.

Estamos interesados en testear las hipótesis H_0 : El dado es equilibrado vs. H_1 : El dado está cargado. Para ello consideramos distintos tests:

- (a) Sea X = cantidad de veces que el resultado del dado es par. Plantear las hipótesis y determinar la región de rechazo para un test de nivel aproximado $\alpha = 0.05$ basado en X . Para el conjunto de datos obtenido, ¿cuál es el valor p aproximado?, ¿cuál es la decisión?.
- (b) Sea Y = cantidad de veces que el resultado del dado es menor o igual que 3. Plantear claramente las hipótesis y determinar la región de rechazo para un test de nivel aproximado $\alpha = 0.05$ basado en Y . Para el conjunto de datos obtenido, ¿cuál es el p -valor aproximado?, ¿cuál es la decisión?.
- (c) Sea W = cantidad de veces que el resultado del dado es exactamente 3. Plantear claramente las hipótesis y determinar la región de rechazo para un test de nivel aproximado $\alpha = 0.05$ basado en W . Para el conjunto de datos obtenido, ¿cuál es el p -valor aproximado?, ¿cuál es la decisión?.
- (d) Sea U = cantidad de veces que el resultado del dado es exactamente 2. Plantear claramente las hipótesis y determinar la región de rechazo para un test de nivel aproximado $\alpha = 0.05$ basado en U . Para el conjunto de datos obtenido, ¿cuál es el p -valor aproximado?, ¿cuál es la decisión?.
- (e) En base a los resultados obtenidos. ¿Qué decisión tomaría? ¿Está cargado el dado?

7. En cada caso indique si la afirmación es verdadera o falsa y justifique:

- (a) El nivel de significación de un test es igual a la probabilidad de que la hipótesis nula sea cierta.
- (b) Un Error de tipo II es más grave que un Error de tipo I.
- (c) Si el p -valor es 0.3, el test correspondiente rechazará al nivel 0.01.
- (d) Si un test rechaza al nivel de significación 0.06, entonces el p -valor es menor o igual a 0.06.
- (e) Si un intervalo de confianza de nivel 0.99 para la media μ de una distribución normal calculado a partir de una muestra da como resultado (2.1, 3.3), entonces el test para las hipótesis $H_0 : \mu = 3$ vs $H_1 : \mu \neq 3$ basado en los mismos datos rechazará la hipótesis nula al nivel 0.01.