Estadística e inferencia I
Segundo Cuatrimestre — 2024

Práctica 2: Test de hipótesis

1. Sea X una variable aleatoria normal de varianza conocida $\sigma^2 = 1$ y esperanza desconocida μ. Se realiza un test que decide si rechazar o preservar la hipótesis $H_0 : \mu \leq 0$ a partir de $n = 100$ muestras. Para cierto $c > 0$, el test dice "Dadas muestras $x_1, \ldots, x_n$, rechazar H_0 si $\bar{x} > c$".

(a) Encontrar el valor de c para que el test tenga un nivel de 5%.

(b) Describir la región de rechazo, que es un subconjunto de $\mathbb{R}^n$.

(c) Calcular la potencia que tiene el test si el verdadero valor de μ es $\mu = 0.1$ y si es $\mu = 1$.

2. Según una encuesta, la cantidad de minutos que le lleva a un alumno de la facultad llegar al campus en tren es una variable aleatoria normal de media $\mu = 45$ y varianza $\sigma^2 = 144$. Se quiere saber si quienes vienen en colectivo tardan más.

(a) Plantear el test de hipótesis adecuado. En base a una muestra aleatoria de 100 datos, dar las regiones de rechazo de nivel $\alpha_1 = 0.01$ y de nivel $\alpha_2 = 0.05$

(b) Se sabe que, para una muestra de 100 datos, se rechazó la hipótesis nula a nivel 0.05, pero no a nivel 0.01. Indicar qué valores puede tomar el promedio muestra.

(c) En un test de nivel 0.05 basado en una muestra de 100 datos,¿cuál es la probabilidad de no rechazar H_0 si en realidad $\mu = 48$ minutos?

(d) Calcular el tamaño de la muestra que se debe considerar si se quiere rechazar la hipótesis nula a nivel 0.01 en el caso en que el promedio muestral sea 47 minutos.

3. En 1000 tiradas de moneda aparecen 560 caras y 440 cecas. ¿Es razonable suponer que la moneda es justa, es decir que la probabilidad de cara y de ceca son iguales?

4. Se desea determinar si un dado de seis caras está cargado o no. Para ello se arroja el dado 1000 veces obteniéndose los siguientes resultados:

Resultado del dado	1	2	3	4	5	6
Frecuencia	160	180	150	170	190	150

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 valor p aproximado?, ¿cuál es la decisión?

5. Históricamente, el 20% del mercado prefiere el jabón de la marca A. Para incrementar las ventas, la empresa A realiza una intensa campaña de publicidad. Al finalizar la misma se entrevistan $n = 400$ individuos y se les pregunta si prefieren o no la marca A, procurando demostrar que la campaña fue exitosa.

(*a*) Expresar H_0 y H_1 en términos de p, la probabilidad de que un cliente prefiera el jabón de la marca A al finalizar la campaña.

(*b*) El gerente de la empresa decide concluir que la campaña de publicidad es exitosa si al menos 92 de los 400 clientes entrevistados prefieren la marca A. Especificar cuál es la región de rechazo de la hipótesis nula H_0 e indicar, de manera aproximada, cuál es el nivel del criterio propuesto.

(*c*) El dueño de la empresa quiere establecer otro criterio de modo que, con probabilidad 0.05, la campaña se declare exitosa cuando en realidad no lo fue. Construya una región de rechazo para satisfacer al dueño. Si al realizar la encuesta se observa que 92 de los 400 encuestados prefirieron la marca A, calcular de manera aproximada el p-valor correspondiente a los datos obtenidos. Con estos datos, ¿se rechaza la hipótesis nula a nivel 0.05?

(*d*) Hallar de manera aproximada la probabilidad de cometer un error de tipo II con el criterio propuesto por el dueño si en realidad $p = 0.24$.