PROBABILIDADES Y ESTADÍSTICA (C)

Práctica 9

1. Una empresa vende dos variedades de soja. La variedad 1 tiene un rendimiento por ha. que puede considerarse una variable aleatoria con distribución $N(37, 25)$, y la variedad 2 tiene un rendimiento por ha. que puede considerarse $N(40, 25)$. Un cliente realizó una compra de semillas de la variedad 2 y antes de continuar comprando a esta empresa, quiere asegurarse de que las semillas que le enviaron realmente pertenecen a esa variedad y no a la variedad 1.

   Con ese fin, cultiva 10 parcelas de 1 ha. y obtiene los siguientes rendimientos:

   37 - 39.50 - 41.70 - 42 - 40 - 41.25 - 43 - 44.05 - 38 - 38.50

   El cliente quiere que la probabilidad de seguir comprando a esta empresa cuando las semillas no son de la variedad pedida sea 0.05.

   a) Explicar por qué las hipótesis para este problema son
   $$H_0 : \mu = 37 \qquad H_1 : \mu = 40.$$

   b) Encontrar un test para estas hipótesis y decir qué decisión se toma en base a la muestra dada. Calcular la probabilidad del error de tipo II.

   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.

   e) Explicar por qué el test planteado en c) sirve también para las hipótesis
   $$H_0 : \mu = 37 \qquad H_1 : \mu > 37.$$

2. Según una estadística realizada por el Centro de Estudiantes la cantidad de minutos por semana que un alumno de Computación dedica a practicar alguna actividad deportiva es una variable aleatoria $N(45; 144)$. Se quiere saber si durante el receso de verano los estudiantes dedican más tiempo al deporte.

   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 muestral.

   c) En un test de nivel 0.05 basado en una muestra de 100 datos, ¿cuál es la probabilidad de equivocarse al sacar la conclusión 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 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 valor p si usa una tabla o calcularlo en forma exacta con el R.

4. 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:

a) ¿Usted decidiría, a nivel 0.05, que el sistema cumple la especificación?

b) Acotar el valor p o calcularlo en forma exacta con el R.

5. 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.

a) ¿Indican estos datos, a nivel aproximado 0.05, un aumento en la preferencia por la marca A?

b) Calcular el valor p.

c) ¿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?

d) ¿Qué tamaño de muestra debería tomarse para que la probabilidad de c) fuese a lo sumo 0.05?

6. 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.

La empresa quiere tener un 95 % de probabilidad de no vender si no se satisfacen los requerimientos.

a) Proponer un test de nivel exacto con hipótesis nula simple para este problema. ¿Qué decisión toma?
(SUGERENCIA: Mirar con atención la práctica 8.....)

b) Repetir a) usando un nivel aproximado.

d) 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?

e) ¿Cuál debería ser el tamaño de la muestra para que la probabilidad de d) fuese 0.1?

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 [-2.0, 3.0], 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.

1

1.

b) $H_0: \mu = 37$ vs $H_1: \mu > 37$

Bajo H_0

$$T = \sqrt{10}\,\frac{(\bar{X}_n - 37)}{\sqrt{25}} \geq Z_{0.05} = 1.64$$

$$\bar{x}_{10} = 40.5$$

$$\frac{\sqrt{10}}{5}(40.5 - 37) = 2.21 \geq 1.64$$

Rechazo H_0

P(Error tipo II) = P(No rechazar cuando H_0 es falsa)

$$40 = \mu_1 > \mu_0 = 37$$

$$P^{\mu=40}\left(\frac{\sqrt{10}(\bar{X}_n - 37)}{5} - \frac{40}{\frac{5}{\sqrt{10}}} < 1.64 - \frac{40}{\frac{5}{\sqrt{10}}}\right)$$

$$P^{\mu=40}\left(\frac{\bar{X}_n - 40}{\frac{5}{\sqrt{10}}} < 1.64 + \frac{37-40}{\frac{5}{\sqrt{10}}}\right)$$

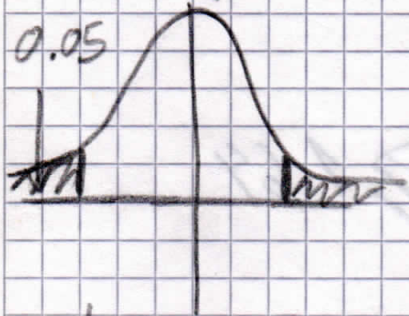
$$= \phi\left(1.64 + \frac{37-40}{\frac{5}{\sqrt{10}}}\right) = \phi(-0.26) = 1 - \phi(0.26)$$

$$= 1 - 0.5987 = 0.4013$$

c) Idem b) pero para n genérico

d) $P^{\mu=\mu_1}\left(\frac{\overline{X}_n - 40}{\frac{5}{\sqrt{n}}} < 1.64 + \frac{37-40}{\frac{5}{\sqrt{n}}} \right)$

$$\phi\left(1.64 + \frac{37-40}{\frac{5}{\sqrt{n}}}\right) \leq 0.05$$



$$\phi\left(-\left(1.64 + \frac{37-40}{\frac{5}{\sqrt{n}}}\right)\right) \geq 0.95$$

$$-1.64 - \frac{37-40}{\frac{5}{\sqrt{n}}} \geq 1.64$$

$$\frac{3}{5}\sqrt{n} \geq 3.28$$

$$\sqrt{n} \geq \frac{82}{15}$$

$$n \geq \frac{6724}{225}$$

$$n \geq 30$$

(2)

2.

a) $H_0: \mu = 45$ vs $H_1: \mu > 45$

$\alpha_1 = 0.01$

$$T = \sqrt{100}\,\frac{(\overline{X}_{100} - 45)}{12} > z_{0.01} = 2.33$$

$R = [2.33, +\infty)$

$\alpha_2 = 0.05$

$$R = [\underset{\substack{\| \\ z_{0.05}}}{1.64}, +\infty)$$

b) p-valor > 0.01

$$1.64 < \sqrt{100}\,\frac{(\overline{x}_{100} - 45)}{12} \leq 2.33$$

$$1.64 * \frac{12}{10} < \overline{x}_{100} - 45 \leq 2.33 * \frac{12}{10}$$

$$46.968 < \overline{x}_{100} \leq 47.796$$

c) $P(\text{Error tipo II}) = P^{\mu=48}\left(\sqrt{100}\frac{(\overline{X}_{100}-45)}{12} < 1.64\right)$

$= P^{\mu=48}\left(\frac{\overline{X}_{100}-48}{\frac{12}{10}} < 1.64 + \frac{45-48}{\frac{12}{10}}\right)$

$= P^{\mu=48}\left(\frac{\overline{X}_{100}-48}{\frac{6}{5}} < 1.64 - 2.5\right)$

$= \phi(-0.86) = 1 - \phi(0.86) = 1 - 0.8051$

$= 0.1949$

d) $\sqrt{n}\frac{(\overline{X}_n - 45)}{12} > 2.33$

$\overline{x}_n = 47$

$\sqrt{n}\frac{(\overline{x}_n - 45)}{12} > 2.33$

$\sqrt{n}\frac{(47-45)}{12} > 2.33$

$\sqrt{n} > 2.33 * 6$

$n > 195.4404$

$n = 196$

③

3. $n = 15$, $\bar{x}_n = 304$, $s = 10$, $\alpha = 0.05$

a) $H_0: \mu = 300$ vs $H_1: \mu > 300$

$$T = \frac{\bar{X}_n - 300}{\frac{S}{\sqrt{15}}} \sim t_{n-1}$$

$$\frac{\bar{X}_n - 300}{\frac{S}{\sqrt{n}}} > t_{n-1, 0.05}$$

$$\frac{\bar{x}_{15} - 300}{\frac{S}{\sqrt{15}}} > t_{14, 0.05} = 1.7613$$

$$\frac{304 - 300}{\frac{10}{\sqrt{15}}} = 1.55 > 1.7613$$

No se rechaza H_0

b) $0.05 < p\text{-valor} < 0.1$

4.

$n = 11 \quad \overline{x}_n = \frac{306}{11}$

a) $H_0: \sigma = 6$ vs $H_1: \sigma < 6$

$\alpha = 0.05$

$$(n-1)\frac{S^2}{\sigma_0} \sim \chi^2_{n-1}$$

$$S^2 = \frac{\sum_{i=1}^{n}(x_i - \overline{x}_n)^2}{n-1}$$

$$(n-1)\frac{S^2}{\sigma_0^2} < \chi^2_{n-1,\,1-\alpha}$$

$$\frac{\frac{3714}{11}}{6^2} < \chi^2_{10,\,0.95} = 3.9403$$

$$9.3788 \nless 3.9403$$

El sistema no cumple la especificación a nivel 0.05 (no rechazo H_0)

b)

$$\text{p-valor} = \alpha \;/\; 9.3788 = \chi^2_{10,1-\alpha}$$

$$0.1 < 1-\alpha < 0.9$$

$$0.1 < \alpha < 0.9$$

← p-valor

④

5.

$$X_i = \begin{cases} 1 & \text{el i-ésimo entrevistado prefiere A} \\ 0 & \text{"} \end{cases}$$

$$X_i \sim Be(P)$$

$$\sum_{i=1}^{200} X_i \sim Bi(200, P)$$

$$H_0: P = 0.1 \quad vs \quad H_1: P > 0.1$$

Por el TCL,

$$\frac{\overline{X}_n - P}{\sqrt{\frac{P(1-P)}{n}}} \xrightarrow{\mathcal{D}} N(0,1)$$

$$\overline{X}_{200} = \frac{26}{200} = 0.13$$

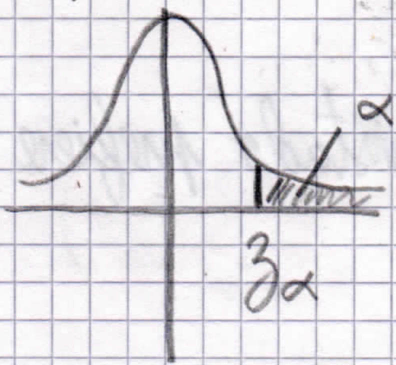
Bajo H_0

$$\sqrt{n}\frac{(\overline{X}_n - 0.1)}{\sqrt{0.1 * 0.9}} \geq z_{0.05} = 1.64$$

$$\sqrt{200}\frac{(0.13-0.1)}{\frac{3}{10}} = 0.4243 \geq 1.64$$

No puedo rechazar H_0 (los datos no indican un aumento en preferencia por A)

b) P-valor $= \alpha$ / $0.4243 = z_\alpha$



$$P(Z > z_\alpha) = \alpha$$

$$P(Z > 0.4243) = \alpha$$

$$1 - P(Z \leq 0.4243) = \alpha$$

$$1 - \phi(0.4243) = \alpha$$

$$1 - 0.6628 = \alpha$$

$$0.3372 = \alpha$$

c) $P(\text{Error Tipo II}) = P^{P=0.15}\left(\dfrac{\bar{X}_{200} - 0.1}{\sqrt{\frac{0.1 * 0.9}{200}}} < 1.64\right)$

$$P^{P=0.15}\left(\frac{\bar{X}_{200} - 0.15}{\sqrt{\frac{0.15 * 0.85}{200}}} < 1.64 + \frac{0.1}{\sqrt{\frac{9}{20000}}} - \frac{0.15}{\sqrt{\frac{51}{80000}}}\right)$$

$$\overset{TCL}{\simeq} \phi(0.41) = 0.6591$$

(5)

d) $P(\text{Error tipo II}) \leq 0.05$

$$P^{p=0.15}\left(\frac{\bar{X}_n - 0.1}{\frac{0.3}{\sqrt{n}}} < 1.64\right)$$