CPE (SEGUNDO CURSO)

PRÁCTICA 16 **SOLUCIONES** (**Curso 2023–2024**)

1.– El resultado de 22 ensayos de adherencia en barras de acero corrugadas fue el siguiente (en MPa):

19.8	18.5	17.6	16.7	15.8	15.4	14.1	13.6	11.9	11.4	11.4
8.8	7.5	15.4	15.4	19.5	14.9	12.7	11.9	11.4	10.1	7.9

A la vista de los resultados,

a) ¿Puede afirmarse que la adherencia media de esas barras es superior a 10 MPa con una confianza del 95 %?

b) ¿A qué nivel de significación puede aceptarse que la adherencia media es superior a 15 MPa?

c) ¿Cuál es la adherencia media mínima que podemos aceptar con una confianza del 95 %?

——————SOLUCIÓN——————

La media muestral y la varianza muestral son:

$$\bar{x} = 13.71364;\ S^* = 3.55358;\ S^{*2} = 12.6279;\ S^2 = 12.0539;\ S = 3.47187$$

a) Con $\alpha = 0.05$ nuestro contraste es

$H_o : m \geq 10$ MPa

$H_1 : m < 10$ MPa

Si hacemos la hipótesis de que la variable subyacente está normalmente distribuida

$$T = \frac{(\bar{x} - m)}{S^*/\sqrt{n}} \equiv T_{n-1}$$

donde T_{n-1} representa a una T de Student con $n-1$ grados de libertad. En nuestro caso la región de aceptación es $T \geq c$. De las tablas de la T de Student con 21 grados de libertad y $\alpha = 0.05$ se obtiene $c = -1.72124$. En nuestro caso, para nuestra muestra, $T = 4.90168$. $\boxed{\text{Luego como } T \geq c \text{ podemos aceptar la hipótesis primaria } m \geq 10 \text{ MPa}}$

b) En este caso nuestro contraste será

$H_o : m \geq 15$ MPa

$H_1 : m < 15$ MPa

y nuestro estadístico vale $T = -1.69788$. Si buscamos para este valor en la tabla de la T de Student con 21 grados de libertad obtenemos $1 - \alpha = 0.94784$. $\boxed{\text{Es decir } \alpha = 0.05216}$

c) Hemos visto en el apartado a) que para una confianza del 95 % el valor crítico para este intervalo vale $c = -1.72124$. Es decir

$$\frac{(\bar{x} - m)}{S^*/\sqrt{n}} = -1.72124 \quad \Longrightarrow \quad m = 15.018$$

$\boxed{\text{Por tanto podemos decir con una confianza del 95 \% que la carga media excede 15.02 MPa.}}$

ESTE PROBLEMA TAMBIÉN SE PODRÍA HABER HECHO DE LA SIGUIENTE FORMA:

Supongamos que el tamaño de la muestra es lo suficientemente grande como para suponer que $\sigma^2 = S^2$. En este caso:

a) Con $\alpha = 0.05$ nuestro contraste es

$H_o : m \geq 10$ MPa

$H_1 : m < 10$ MPa

Como suponemos que la muestra es lo suficientemente grande

$$U = \frac{(\bar{x} - m)}{\sigma/\sqrt{n}} \equiv N(0, 1)$$

En nuestro caso la región de aceptación es $U \geq c$. De las tablas de la normal con $\alpha = 0.05$ se obtiene $c = -1.645$. En nuestro caso, para nuestra muestra, $U = 5.01704$. $\boxed{\text{Luego como } U \geq c \text{ podemos aceptar la hipótesis primaria } m \geq 10 \text{ MPa}}$

b) En este caso nuestro contraste será

$H_o : m \geq 15$ MPa

$H_1 : m < 15$ MPa

y nuestro estadístico vale $U = -1.73784$. Si buscamos para este valor en la tabla de la normal obtenemos $1 - \alpha = 0.9589$. $\boxed{\text{Es decir } \alpha = 0.04109}$

c) Hemos visto en el apartado a) que para una confianza del 95 % el valor crítico para este intervalo vale $c = -1.645$. Es decir

$$\frac{(\bar{x} - m)}{\sigma/\sqrt{n}} = -1.645 \quad \Longrightarrow \quad m = 14.931$$

$\boxed{\text{Por tanto podemos decir con una confianza del 95 \% que la carga media excede 14.9 MPa.}}$

Podemos observar que la introducción de nuevo conocimiento ($\sigma^2 = S^2$) transforma los resultados en otros ligeramente menos conservadores.

Y DE OTRA FORMA MÁS:

Supongamos que el tamaño de la muestra es lo suficientemente grande como para suponer que $\sigma^2 = S^{*2}$. En este caso:

a) Con $\alpha = 0.05$ nuestro contraste es

$H_o : m \geq 10$ MPa

$H_1 : m < 10$ MPa

Como suponemos que la muestra es lo suficientemente grande

$$U = \frac{(\bar{x} - m)}{\sigma/\sqrt{n}} \equiv N(0,1)$$

En nuestro caso la región de aceptación es $U \geq c$. De las tablas de la normal con $\alpha = 0.05$ se obtiene $c = -1.645$. En nuestro caso, para nuestra muestra, $U = 4.90168$. Luego como $U \geq c$ podemos aceptar la hipótesis primaria $m \geq 10$ MPa

b) En este caso nuestro contraste será

$H_o : m \geq 15$ MPa

$H_1 : m < 15$ MPa

y nuestro estadístico vale $U = -1.6979$. Si buscamos para este valor en la tabla de la normal obtenemos $1 - \alpha = 0.9552$. Es decir $\alpha = 0.04479$

c) Hemos visto en el apartado a) que para una confianza del 95 % el valor crítico para este intervalo vale $c = -1.645$. Es decir

$$\frac{(\bar{x} - m)}{\sigma/\sqrt{n}} = -1.645 \quad \Longrightarrow \quad m = 14.950$$

Por tanto podemos decir con una confianza del 95 % que la carga media excede 14.950 MPa.

Es decir, resultados intermedios entre los dos anteriores.

2.– De una población aleatoria X, con función de densidad

$$f_x(x) = \theta^2 x e^{-x\theta}, \quad x \geq 0, \quad \theta > 0$$

se toma una muestra de tamaño 1, x_1.

a) Determinar el estimador de máxima verosimilitud de θ. Estudiar su sesgo.

b) Hallar un contraste de máxima potencia con $\alpha = 0.1$ para contrastar la hipótesis primaria

H$_0$: $\theta = 1$

contra la alternativa

H$_1$: $\theta = 2$

———————SOLUCIÓN———————

a) Al tomar una muestra de tamaño 1, la función de verosimilitud es

$$L(\theta) = \prod_{i=1}^{n} f_{x_i}(x_i) = \theta^2 x_1 e^{-\theta x_1}$$

Tomando logaritmos

$$\ln L(\theta) = 2 \ln \theta + \ln x_1 - \theta x_1$$

Derivando con respecto a θ para hallar el máximo de $L(\theta)$ obtenemos

$$\frac{\partial \ln L(\theta)}{\partial\theta} = \frac{2}{\hat{\theta}} - x_1 = 0 \quad \Rightarrow \quad \boxed{\hat{\theta} = \frac{2}{x_1}}$$

Veamos el sesgo de este estimador

$$E[\hat{\theta}] = E\big[\frac{2}{x_1}\big] = E\big[\frac{2}{x}\big] = \int_0^\infty \frac{2}{x} f_X(x)dx = \int_0^\infty \frac{2}{x}\theta^2 x e^{-x\theta}dx = 2\theta^2 \int_0^\infty e^{-x\theta}dx = 2\theta$$

que es claramente sesgado.

b) El contraste de hipótesis es

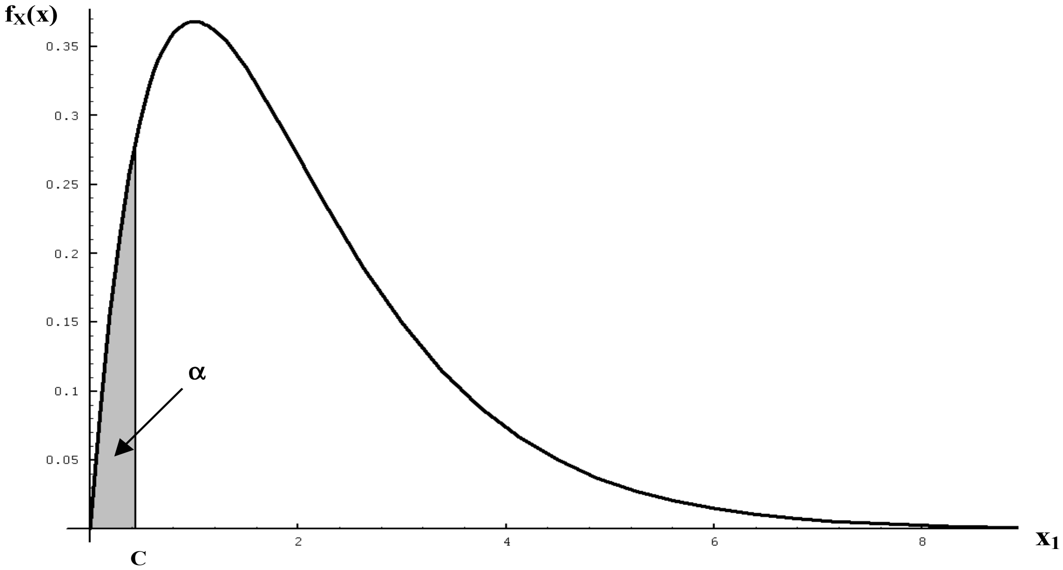
$H_o : \theta = 1$

$H_1 : \theta = 2$

Utilicemos el lema de Neyman-Pearson para evaluar la región crítica:

$$\frac{f_{x_1}(x_1|H_1)}{f_{x_1}(x_1|H_o)} = \frac{2^2 x_1 e^{-2x_1}}{1^2 x_1 e^{-1x_1}} = 4e^{-x_1} > K$$

Luego la región crítica debe contener los valores mínimos de x_1. Por lo tanto la región critica será la que indica la figura (recuérdese que la distribución de x_1 es la misma que la de X).



Ahora calcularemos el valor de c de forma que $F_X(c) = 0.1$. Pero en H_o se verifica $\theta = 1$. Luego

$$F_X(c) = 0.1 = \int_0^c xe^{-x}dx = 1 - (c+1)e^{-c} \quad \Rightarrow \quad \boxed{c \approx 0.53}$$

y rechazaremos la hipótesis primaria al 10 % de significación si $x_1 < 0.53$.

3.– Una compañía fabrica cables de acero para la construcción. En sus catálogos asegura que su modelo "MC300" tiene una tensión estándar de rotura media de 300 Kp, con una desviación típica de 24 Kp. La compañía adquiere una máquina nueva para fabricar cables y quiere comprobar si efectivamente produce cables con una tensión de rotura media mayor.

a) Diseñar una regla de decisión para contrastar esa hipótesis al nivel de significación 0.01, en base a una muestra de 64 cables "MC300" fabricados por la nueva máquina.

b) Una vez diseñada esa regla de decisión, calcular la probabilidad de aceptar que la nueva máquina fabrica según el estándar anterior cuando de hecho lo hace con una media de 310 Kp y la misma desviación típica.

c) Determinar la curva de potencia del contraste de hipótesis diseñado.

d) Diseñar una regla de decisión al nivel de significación 0.05, para contrastar las afirmaciones de que la nueva máquina fabrica según el estándar inicial, frente a que lo hace con una media mayor que 300 Kp. Calcular entonces la probabilidad de aceptar que la nueva máquina fabrica según el estándar inicial, cuando en realidad lo hace con media de 310 Kp y la misma desviación típica.

e) Considerando el test del apartado anterior, calcular el mínimo nivel de significación al cual se rechaza que la nueva máquina fabrica según el estándar inicial, si la media de una muestra resultó ser de 306 Kp.

——————SOLUCIÓN——————

a) Sea X la tensión de rotura de los cables fabricados con la máquina nueva. El contraste de hipótesis que tendremos que construir es

$H_0 : m_x \geq m_o = 300$ Kp

$H_1 : m_x = m_1 < m_o = 300$ Kp

Como la muestra que tenemos es de tamaño 64, podremos considerar que el estadístico de contraste $\bar{x} = \hat{m}_x$ es normal $N(m_o, \sigma_x^2/n)$ en H_o, y por lo tanto $U = \frac{\bar{x}-m_o}{\sigma_x/\sqrt{n}} \equiv N(0,1)$. El contraste será tal que la región crítica estará constituida por $RC = (-\infty,\ c]$ siendo c tal que $P[U \leq c] = F_U(c) = 0.01$ de donde $c = -2.3263$. Consecuentemente, la región crítica puede escribirse

$$RC = \Big(-\infty,\ m_o + c\frac{\sigma_x}{\sqrt{n}}\Big] = \Big(-\infty,\ 300 - 2.3263\frac{24}{8}\Big] = (-\infty,\ 293.0211]$$

Es decir, si $\bar{x} \geq 293.0211$ Kp aceptamos la hipótesis primaria.

b) La probabilidad de no aceptar la mejora en la tensión media, con los datos dados $(m_x = 310, \sigma_x = 24)$ es

$$P[\bar{x} \leq 293.0211 | m_x = 310, \sigma_x = 24] = P\Big[\frac{\bar{x} - 310}{\frac{\sigma_x}{\sqrt{n}}} \leq \frac{293.0211 - 310}{\frac{\sigma_x}{\sqrt{n}}}\Big] =$$

$$= F_U\Big(\frac{-16.9789}{3}\Big) = F_U(-5.6596) \approx 0$$

c) Sabemos que la potencia en un contraste de hipótesis es

$$\Pi = 1 - \beta = 1 - P[\text{Aceptar la hipótesis primaria}|H_1].$$

En nuestro caso

$$\Pi = 1 - \beta = 1 - P[\bar{x} \geq 293.0211|H_1] = 1 - P\Big[\frac{\bar{x} - m_1}{\frac{\sigma_x}{\sqrt{n}}} \geq \frac{293.0211 - m_1}{\frac{\sigma_x}{\sqrt{n}}}|H_1\Big] =$$

$$= 1 - \left(1 - F_U\Big(\frac{293.0211 - m_1}{3}\Big)\right) = F_U\Big(\frac{293.0211 - m_1}{3}\Big) = \Pi(m_1)$$

En la tabla siguiente pueden verse algunos valores de Π en función de H_1

m_1	300	299	298	297	296	295	294	293	292	291	290	289	288
Π	0,01	0,02	0,05	0,09	0,16	0,25	0,37	0,50	0,63	0,75	0,84	0,91	0,95

d) En este caso nuestro contraste de hipótesis sería

$H_0 : m_x = m_o = 300$ Kp

$H_1 : m_x = m_1 > m_o = 300$ Kp

Siguiendo un procedimiento similar al indicado en los apartados a) y b) obtendríamos:

$$RC = [c,\ \infty), \quad P[U \geq c] = 1 - F_U(c) = 0.05 \quad \Rightarrow \quad F_U(c) = 0.95 \quad \Rightarrow \quad c = 1.645$$

Consecuentemente, la región crítica, utilizando $\bar{x}$ como estadístico, puede escribirse como

$$RC = \Big[m_o + c\frac{\sigma_x}{\sqrt{n}},\ \infty) = \Big[300 + 1.645\frac{24}{8},\ \infty) = [304.935,\ \infty)$$

Es decir, si $\bar{x} < 305$ Kp aceptamos la hipótesis primaria.

Al igual que en el apartado b), la probabilidad de aceptar que la nueva máquina fabrica según el estándar inicial, cuando en realidad lo hace con media de 310 Kp y la misma desviación típica será

$$P[\bar{x} \leq 304.935 | m_x = 310, \sigma_x = 24] = P\Big[\frac{\bar{x} - 310}{\frac{\sigma_x}{\sqrt{n}}} \leq \frac{304.935 - 310}{\frac{\sigma_x}{\sqrt{n}}}\Big] =$$

$$= F_U\Big(\frac{-5.065}{3}\Big) = F_U(-1.6883) \approx 0.0457$$

e) El contraste de hipótesis es el mismo que el apartado anterior. Rechazamos la hipótesis primaria, a un nivel de significación de 0.05, si $\bar{x} \geq 305$ Kp. Tenemos una muestra (suponemos que sigue siendo de tamaño 64), con $\bar{x} = 306$. Tenemos que hallar el mínimo nivel de significación al que rechazaríamos la hipótesis primaria con la nueva media. En este caso nuestra región crítica sería $RC = [306,\ \infty)$. Entonces el valor crítico c sería

$$c = \frac{306 - 300}{\frac{\sigma_x}{\sqrt{n}}} = 2$$

y

$$1 - F_U(2) = 0.02275$$

que coincide con el nivel p del test para ese estadístico muestral.

4.– El responsable de la obra de construcción de una pasarela recibe remaches para la estructura metálica de dos suministradores diferentes (que llamaremos X e Y). El responsable sabe que la longitud media de los remaches fabricados por X es de 48 mm. y la de los remaches fabricados por Y es de 52 mm, con igual desviación típica $\sigma = 8$ mm. Acaba de recibir una partida de remaches, pero en el albarán no menciona la empresa de procedencia. Calcular el número de remaches que deberán medirse para que en los dos contrastes de hipótesis siguientes:

$$H_o : m = 48\ mm;\ H_1 : m = 52\ mm$$

$$H_o : m = 52\ mm;\ H_1 : m = 48\ mm$$

no puedan aceptarse simultáneamente ambas hipótesis primarias a un nivel de significación del 1 % (confianza del 99 %).

——————————SOLUCIÓN——————————

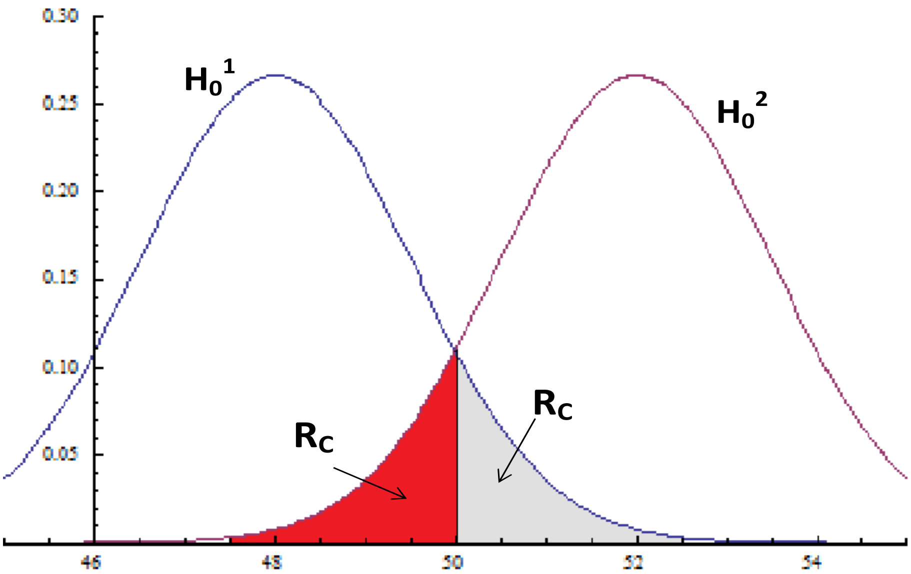
Para que no puedan aceptarse las dos hipótesis a la vez las regiones de aceptación de ambas no han de coincidir en ningún punto. En el límite (muestra menor) la región de aceptación del primer contraste ha de coincidir con la región crítica del segundo contraste. Por el tipo de test, la región de aceptación del primer contraste viene definida por $R_A^1 = (-\infty, a)$ y la región critica del segundo por $R_C^2 = (-\infty, b)$. Como la varianza es conocida en ambos casos, si llamamos $\overline{x}$ y $\overline{y}$ a los dos estadísticos correspondientes a las fábricas X y Y, respectivamente, y asumimos que o bien la muestra es grande (cosa que habremos de comprobar a posteriori) o bien las variables subyacentes son normales, se verificará en las correspondientes hipótesis primarias que $\overline{x} \equiv N(48, 8^2/n)$ y $\overline{y} \equiv N(52, 8^2/n)$. Por otra parte queremos que los límites de ambas regiones críticas sean el mismo, es decir $a = b$. Podemos por tanto escribir el siguiente sistema de ecuaciones:

$$F_{\overline{x}}(a) = 0.99 = F_U\left(\frac{a-48}{8/\sqrt{n}}\right) \implies \frac{a-48}{8/\sqrt{n}} = 2.326348$$

$$F_{\overline{y}}(a) = 0.01 = F_U\left(\frac{a-52}{8/\sqrt{n}}\right) \implies \frac{a-52}{8/\sqrt{n}} = -2.326348$$

de donde se obtiene $n = 86.59$. Es decir, necesitaríamos una muestra de tamaño mínimo de 87 remaches medidos para garantizar que no se pudieran aceptar ambas hipótesis primarias a la vez.

En la siguiente figura se muestra una explicación gráfica de lo anteriormente desarrollado.



5.– Una compañía que vende un cierto aditivo para hormigón indica que cada saco contiene 12 Kg. Supóngase que se sabe por experiencia que la población de pesos de los sacos tiene una desviación típica de 0.5 Kg. La compañía quiere asegurarse que una variación en el standard será detectada si la media se desplaza en 0.2 Kg. Supóngase que se pide una potencia de 0.99 para una media $\mu = 11.8$ Kg. ¿Cuál debe ser el tamaño de la muestra?

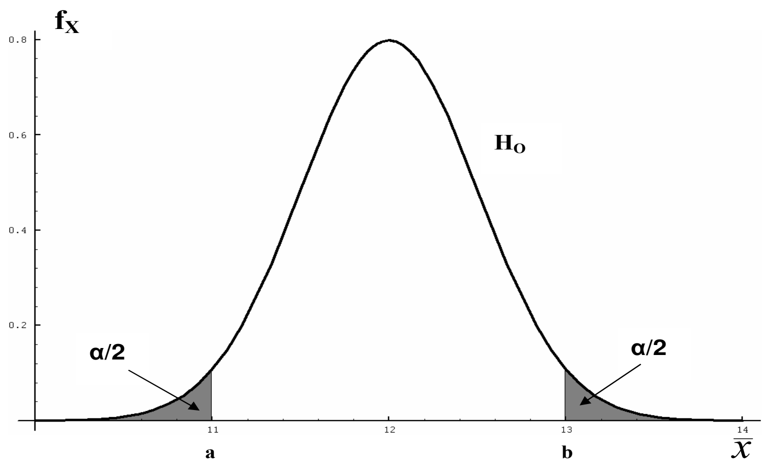
SOLUCIÓN

Sea X el peso de un saco. Sabemos que $\sigma_x = 0.5$ Kg. El contraste de hipótesis a emplear es

H_o: $\mathrm{m} = \mathrm{m}_o = 12$

H_1: $\mathrm{m} = \mathrm{m}_1 \neq 12$

Si $\bar{x}$ es la media muestral, sabemos que la región crítica está en ambas colas de la distribución de $\bar{x}$. Si suponemos $\bar{x}$ normal (ver Figura).



Si α es el nivel de significación

$$1-\alpha = P[a \leq \bar{x} \leq b] = P\left[\frac{a-m_o}{\frac{\sigma}{\sqrt{n}}} \leq \frac{\bar{x}-m_o}{\frac{\sigma}{\sqrt{n}}} \leq \frac{b-m_o}{\frac{\sigma}{\sqrt{n}}}\right] = P\Big[-c \leq \frac{\bar{x}-m_o}{\frac{\sigma}{\sqrt{n}}} \leq c\Big] =$$

$$= F_U(c) - F_U(-c) = 2F_U(c) - 1 = 1-\alpha \implies F_U(c) = 1 - \frac{\alpha}{2}$$

Para los valores usuales del nivel de significación

$\alpha = 0.1, \quad 1-\alpha = 0.9, \quad c = u_{0.95} = 1.645$

$\alpha = 0.05, \quad 1-\alpha = 0.95, \quad c = u_{0.975} = 1.96$

$\alpha = 0.01, \quad 1-\alpha = 0.99, \quad c = u_{0.995} = 2.576$

Entonces el error β es

$$\beta = P[a \leq \bar{x} \leq b|H_1] = P\Big[-c \leq \frac{\bar{x}-m_o}{\frac{\sigma}{\sqrt{n}}} \leq c|H_1\Big] = P\Big[\frac{m_o-m_1}{\frac{\sigma}{\sqrt{n}}} - c \leq \frac{\bar{x}-m_1}{\frac{\sigma}{\sqrt{n}}} \leq \frac{m_o-m_1}{\frac{\sigma}{\sqrt{n}}} + c|H_1\Big]$$

En nuestro caso $\mathrm{m}_o = 12$, $\mathrm{m}_1 = 11.8$, $\sigma = 0.5$. Por consiguiente

$$\beta = 1 - \Pi = P\Big[0.4\sqrt{n} - c \leq \frac{\bar{x}-m_1}{\frac{\sigma}{\sqrt{n}}} \leq 0.4\sqrt{n} + c|H_1\Big] = F_U(0.4\sqrt{n}+c) - F_U(0.4\sqrt{n}-c) = 0.01$$

Resolviendo en cada caso por tanteos

$\alpha = 0.1, \quad 1-\alpha = 0.9, \quad c = u_{0.95} = 1.645, \quad n = 99$

$\alpha = 0.05, \quad 1-\alpha = 0.95, \quad c = u_{0.975} = 1.96 \quad n = 115$

$\alpha = 0.01, \quad 1-\alpha = 0.99, \quad c = u_{0.995} = 2.576, \quad n = 151$

(Obsérvese que para cualquier nivel de significación la hipótesis $\bar{x} \equiv$ Normal ha sido correcta.)