Nombre:____ ***Curso***:_____	
Estadística – Muestreo, Inferencia e Hipótesis – 03/04/2014	

EJERCICIO 1: [2,5]

a) [1,25] Dada la población **X** = { 1 , 4 , 7 }, construya todas las muestras posibles de tamaño 2 que puedan formarse mediante muestreo aleatorio simple, y halle la varianza de las medias muestrales de todas esas muestras.

b) [1,25] Una población de 6000 personas se ha dividido en 3 estratos, uno con 1000 personas, otro con 3500 y otro con 1500. En esa población se ha realizado un muestreo estratificado con afijación proporcional, en el que se han elegido al azar 15 personas del tercer estrato. Determine el tamaño de la muestra total obtenida con este muestreo y su composición.

EJERCICIO 2:

Las naranjas de una cosecha tienen un peso medio de 210 gr. y una desviación típica de 20 gr. Las naranjas se ponen a la venta en cajas de 64 naranjas.

a) [1] ¿Cómo es la distribución de las medias muestrales?

b) [1,5] ¿Cuál es la probabilidad de que el peso medio de las naranjas de una caja no sea inferior a 217 gr.?

EJERCICIO 3:

Se conoce que la acidez de una solución es una variable aleatoria que sigue una distribución Normal con desviación típica 0.2. Se ha tomado una muestra aleatoria de cinco soluciones y se han obtenido las siguientes medidas de la acidez:

7.92 ; 7.95 ; 7.91 ; 7.9 ; 7.94

a) [1.25] Halle el intervalo de confianza, al 99%, para la media poblacional.

b) [0.5] ¿Qué error máximo se ha cometido en el intervalo anterior?

c) [0.75] Para el mismo nivel de confianza, calcule el tamaño mínimo muestral que permita reducir el error anterior a la mitad.

EJERCICIO 4: [2,5]

Según un estudio botánico, el 75% de los geranios atacados por el taladro muere. Se desea contrastar ese porcentaje, con un nivel de significación del 2%.

Para ello se hizo el seguimiento de 500 geranios atacados por el taladro, de los que 350 murieron.

a) Escribe la hipótesis nula e indica la hipótesis alternativa.

b) Halla el intervalo de aceptación y dibujalo junto con la región crítica.

c) ¿Hay evidencias estadísticas que nos permitan rechazar la hipótesis, o por el contrario la confirman?

EJERCICIO 1:

a) La variable aleatoria en la población es X = { 1 , 4 , 7 }

El conjunto de las muestras de tamaño $n = 2$, con reemplazamiento, y las medias muestrales son:

$$\mathfrak{M} = \left\{ \begin{array}{ccc} (1,1) & (1,4) & (1,7) \\ (4,1) & (4,4) & (4,7) \\ (7,1) & (7,4) & (7,7) \end{array} \right\} \quad \xrightarrow[medias]{} \overline{\mathbf{X}} = \left\{ \begin{array}{ccc} 1 & 2.5 & 4 \\ 2.5 & 4 & 5.5 \\ 4 & 5.5 & 7 \end{array} \right\}$$

Así, la media y la desviación típica de las medias muestrales son:

$$\overline{\mu} = \frac{36}{9} = 4$$

$$\overline{\sigma}^2 = \frac{171}{9} - 4^2 = 3$$

Observemos las barras que tienen los parámetros: fundamentales para no confundirlos con los parámetros de la población.

b) Mostremos esquemáticamente las composiciones de la población y de la muestra:

Población		
I	II	III
1000	3500	1500
N = 6000		

Muestra		
I	II	III
10	35	15
n = 60		

Como en el tercer estrato se eligen 15: $\frac{1500}{6000} \times n = 15 \;\rightarrow\; n = 60$

Para el segundo estrato: $\frac{3500}{6000} \times 60 = 35.$

Para el primer estrato: $60 - 35 - 15 = 10.$

Ahora elegimos en cada estrato, mediante muestreo aleatorio simple, el número de individuos señalados anteriormente.

EJERCICIO 2:

La variable **X** = "peso de las naranjas" tiene $\left\{ \begin{array}{ccc} \mu & = & 210 \\ \sigma & = & 20 \end{array} \right.$

Tamaño muestral: $n = 64$.

a) La distribución de las medias muestrales $\overline{\mathbf{X}}$ es normal (porque $n > 30$) con $\left\{ \begin{array}{l} \mu = \mu = 210 \\ \overline{\sigma} = \dfrac{\sigma}{\sqrt{n}} = \dfrac{20}{\sqrt{64}} = 2.5 \end{array} \right.$

b) La probabilidad pedida es:

$$\mathrm{p}\,(\overline{x} \geq 217) \overset{(*)}{=} \mathrm{p}\,(z \geq 2.8) = 1 - 0.9974 = 0.0026$$

$(*)\; z = \dfrac{\overline{x} - \overline{\mu}}{\overline{\sigma}} = \dfrac{217 - 210}{2.5} = 2.8$

EJERCICIO 3:

La v.a. **X** = "acidez de una disolución" es normal con $\begin{cases} \mu = ¿? \\ \sigma = 0.2 \end{cases}$

Nivel de confianza: $p = 1 - \alpha = 0.99 \quad \rightarrow \quad$ Valor crítico: $z_{\alpha/2} \approx 2.575$

a) Tamaño muestral: $n = 5$

Media muestral: $\overline{x} = 7.924.$

El intervalo de confianza, para la media de la población, es:

$$I = \left(\overline{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\,, \overline{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right) = (7.924 - 0.2303\,, 7.924 + 0.2303) = (7.6937\,, 8.1543)$$

b) El error máximo cometido es:

$$E = z_{\alpha/2}\frac{\sigma}{\sqrt{n}} = 0.2303$$

c) Ahora primero el error máximo:

$$E = \frac{0.2303}{2} = 0.11515$$

Del error sacamos el tamaño muestral:

$$E = z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \rightarrow 0.11515 = 2.575 \cdot \frac{0.2}{\sqrt{n}} \rightarrow \sqrt{n} = 2.575 \cdot \frac{0.2}{0.11515} \rightarrow \sqrt{n} = 4,4724 \rightarrow n \approx 20$$

EJERCICIO 4:

a) <u>Hipótesis</u>:

$\mathrm{H}_0 \,:\, p = 0.75$ hipótesis nula $\quad \rightarrow \quad \mathrm{H}_1 \,:\, p \neq 0.75$ hipótesis alternativa

Es unilateral sobre la proporción.

b) <u>Muestra y estadístico</u>:

Tamaño muestral: $n = 500$ ➡ Proporción muestral: $\widetilde{p} = \dfrac{350}{500} = 0.7.$

<u>Significación y valor crítico</u>:

Significación: $\alpha = 0.02$ ➡ Valor crítico: $z_{\alpha/2} \approx 2.33$

$\mathrm{p}\left(z > z_{\alpha/2}\right) = 0.01 \rightarrow \mathrm{p}\left(z < z_{\alpha/2}\right) = 0.99 \rightarrow z_{\alpha/2} \approx 2.33$

<u>Intervalo de aceptación</u>:

$$I = \left(p_0 - z_{\alpha/2} \cdot \sqrt{\frac{p_0 q_0}{n}}\,, p_0 + z_{\alpha/2} \cdot \sqrt{\frac{p_0 q_0}{n}}\right) = (0.75 - 0.0451\,, 0.75 + 0.0451) = (0.7049\,, 0.7951)$$

c) <u>Conclusión</u>:

$$\widetilde{p} \notin I \quad \rightarrow \quad \text{Rechazamos } \mathrm{H}_0$$

A la vista de los datos, con el nivel de significación dado, rechazamos que el 75% de los geranios atacados por el taladro mueran.