1. Determine el valor de verdad de las siguientes afirmaciones

a) Un censo consiste en seleccionar un subconjunto de la población, medir la característica de interés e inferir con esta información acerca de toda la población.
b) La probabilidad de que una unidad cualesquiera esté presente en la muestra es $\frac{n}{N}$.
c) Para N muy grande se tiene que $\mathbb{E}(S^2) \approx \sigma^2$.
d) Suponga que se tiene un grupo de 11 personas númeradas del 1 al 11, de este grupo se desea extraer una muestra de tamaño 6. Para esto se decide tirar un par de dados 6 veces y en cada lanzamiento se resta una unidad al resultado de los dados, siendo dicho resultado final el número del individuo que va a ser incluido en la muestra. En caso de que se repita algún resultado los dados se arrojan nuevamente, por tanto como se quiere extraer una muestra y el resultado es aleatorio, se está ante el diseño muestral de muestreo aleatorio simple sin reemplazo.

a) Falso, censo es sobre toda la población.

b) n = tamaño muestra; N = tamaño población

$$\underline{\quad}_1 \quad \underline{\quad}_2 \quad \underline{\quad}_3 \quad \cdots \quad \underline{\quad}_{n-1} \quad \underline{\quad}_n$$

una unidad cualesquiera tiene n casos favorables de ser seleccionada en la muestra.

Los casos posibles $= N$.

$\therefore$ Probabilidad de que cualquier unidad esté en la muestra $= \frac{n}{N}$

c) $$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2$$

Para una población finita

$$E[S^2] = \frac{N}{N-1}\sigma^2$$

Ahora bien

$$\lim_{N \to \infty} \frac{N}{N-1}\sigma^2 = \lim_{N \to \infty} \frac{\frac{d}{dN}(N)}{\frac{d}{dN}(N-1)}\sigma^2 = \sigma^2$$

$\therefore$ Para N grande $E[S^2] \approx \sigma^2$

d) En muestreo probabilístico, cada unidad elementar tiene la misma probabilidad de ser seleccionada en la muestra, sin embargo nótese para el ejercicio propuesto:

$$P(\text{Persona } 1 \text{ seleccionada}) = P(\underbrace{\textstyle\sum \text{dados} = 2}_{\text{Ambos dados caen en } 1})$$

$$\therefore = \frac{1}{36}$$

$$P(\text{persona } 6 \text{ seleccionada}) = P(\underbrace{\textstyle\sum \text{dados} = 7}_{(1,6),(6,1),(4,3),(3,4),(2,5),(5,2)})$$

$$\therefore = \frac{6}{36}$$

Como $P(\#1) \neq P(\#6)$, no es muestreo probabilístico y por tanto no es un m.a.s.

2. Una m.a.s de 100 contadores de agua es controlada dentro de una comunidad para estimar el promedio de consumo de agua diario por casa durante un periodo seco. Realizado el estudio, se halló que la media y varianza muestrales fueron 12.5 y 1.252 galones respectivamente. Haga una estimación del promedio de consumo diario y calcule su respectivo intervalo de confianza, realice lo propio para el consumo total diario de la comunidad. Suponga que en la comunidad existe un total de 10000 casas. Determine cuantas unidades son necesarias para obtener un límite para el error de estimación de la media del consumo diario de 1 galón.

$n = 100$; $N = 10000$

$$\underbrace{\bar{Y} = 12{,}5}_{\text{estimador insesgado del promedio}} \; ; \; s^2 = 1{,}252$$

Para el IC, $\alpha = 0{,}05$

Parámetro a estimar	Varianza estimada del estimador	Intervalo de confianza del $(1-\alpha)100\%$
μ	$\hat{V}\left(\bar{Y}\right) = \left(\frac{N-n}{N}\right)\frac{S^2}{n}$	$\bar{Y} \pm t_{\alpha/2,n-1}\sqrt{\hat{V}\left(\bar{Y}\right)}$

$$\hat{V}(\bar{Y}) = \left(\frac{10000-100}{10000}\right)\frac{1{,}252}{100} = 0.1113319$$

$$t_{\frac{\alpha}{2}, n-1} = 1.984217$$

$\therefore$ $(12.27909 \quad 12.72091)$

Para consumo total basta con multiplicar $\bar{Y}$ por N

$\therefore$ $\hat{\tau} = N\bar{Y} = 125000$

Para $\hat{var}(\tilde{\tau}) = \hat{var}(N\bar{Y}) = N^2 \hat{var}(\bar{Y})$

$\simeq$ 11133190

IC también basta con multiplicar por N, sin embargo, SÓLO CUANDO LA SIGNIFICANCIA COINCIDE, en este caso $\alpha = 0{,}05$

$\therefore$ (122790.9 127209.1)

Parámetro	Tamaño de muestra usando σ^2
μ	$n = \frac{N\sigma^2}{(N-1)D+\sigma^2}$, $D = \frac{\delta^2}{Z^2_{\alpha/2}}$. Para N muy grande: $n_0 = \frac{\sigma^2}{D}$

$\delta^2 = \ell^2$

$$n = \frac{10000\left(\frac{N-1}{N}\right)s^2}{(N-1)\cdot\frac{1}{Z^2_{\frac{\alpha}{2}}} + \left(\frac{N-1}{N}\right)s^2} = 4.807194$$

min

$\therefore \boxed{n = 5}$

3. París es una ciudad que recibe diariamente 1500 turistas. Se desea realizar un estudio y se ha seleccionado una muestra aleatoria simple de turistas, donde se les preguntó cuanto gastan diariamente y si eran extranjeros.

Tabla 1: Datos de los turistas

ID	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Gastos	48	41	34	25	32	25	36	31	30	38	31	19	26	27	22
Nacionalidad	0	1	1	0	1	0	1	1	0	1	1	0	1	1	1

a. Estime los gastos totales en consumo que realizan los turistas en París.
b. Estime el total de turistas que són extranjeros. Estime un intervalo de confianza.
c. Determine el tamaño de muestra mínimo necesario para estimar la propoción de extranjeros que visitan París en un día con un límite para el error de estimación de 2 % y una confianza de 95 %

a) $\hat{\tau} = N\bar{Y}$ donde $N = 1500$, $n = 15$

$$\bar{Y} = \frac{1}{n}\sum_{i=1}^{15} Y_i = 31$$

$$\therefore \hat{\tau} = (1500)(31) = 46500$$

b) Nos piden A (1 para extranjeros y 0 para nacionales) y su IC al 95%

Hallemos primero $\hat{p} = \frac{a}{n} = \frac{10}{15}$

IC con $\alpha = 0,05$

$$\hat{p} \pm t_{\frac{\alpha}{2}, n-1} \cdot \hat{se}(\hat{p})$$ donde

$$\hat{se}(\hat{p}) = \sqrt{\left(\frac{N-n}{N}\right)\frac{\hat{p}(1-\hat{p})}{n-1}} \approx 7.615773$$

$\therefore$ (0.3978034 0.9355299)

Para A sólo hace falta multiplicar por N, luego

$A = N\hat{p} = 1000$

IC al 95%

$(596.7051 \quad 1403.2949)$

c)

Parámetro	Tamaño de muestra
p y A	$n = \frac{Np(1-p)}{(N-1)Dp^2+p(1-p)} = \frac{N\,CV^2}{(N-1)D+CV^2} \approx \frac{CV^2}{D+\frac{1}{N}CV^2}$, con $D = \frac{\varepsilon^2}{Z^2_{\alpha/2}}$ y $CV = \sqrt{\frac{1-p}{p}}$. Para N muy grande: $n_0 = \frac{CV^2}{D}$

$$\varepsilon = \frac{\delta}{p} \rightarrow \varepsilon^2 = \frac{\delta^2}{p^2} \qquad \delta = 0{,}02$$

$$n = \frac{1500 \cdot \frac{10}{15}\left(1 - \frac{10}{15}\right)}{(1499) \cdot \frac{0{,}02^2}{Z^2_{\alpha/2}} + \frac{10}{15}\left(1 - \frac{10}{15}\right)} = 881.1145$$

min

$$\therefore \boxed{n = 882}$$

Una empresa industrial está interesada en estudiar el tiempo por semana que sus científicos emplean en ciertas tareas triviales. Las hojas de control del tiempo de una muestra aleatoria de $n = 50$ científicos muestran que la cantidad promedio de tiempo empleado en esas tareas es de 10.31 horas, con una varianza muestral $s^2 = 2.25$. La compañía emplea $N = 750$ científicos.

4. De realizar un intervalo del 95 % de confianza para µ se puede concluir:

a. Con una confianza del 95 % el tiempo medio que un científico de la empresa emplea en tareas triviales está entre 7.848 y 12.772 horas a la semana.
b. Con una confianza del 95 % se puede considerar que el tiempo medio que un científico de la empresa emplea en tareas triviales es inferior a 9.6 horas a la semana.
c. Con una confianza del 95 % el tiempo medio que un científico de la empresa emplea en tareas triviales está entre 9.898 y 10.722 horas a la semana.
d. Ninguna de las anteriores.

$\bar{Y} = 10,31$; $Var(\bar{Y}) = \left(\frac{N-1}{N}\right) \frac{s^2}{n}$

$= 0.04494$

Si $\alpha = 0,05$,

$t_{\frac{\alpha}{2}, n-1} = 2.009575$

IC del 95%

$\bar{Y} \pm t_{\frac{\alpha}{2}, n-1} * \sqrt{Var(\bar{Y})}$

$\therefore$ (9.883989 10.736011)

Respuesta c.

5. Estime el número total de horas que se pierden a la semana en las tareas insignificantes.

$\hat{\tau} = N * \bar{Y} = 7732.5$ h

6. Construya un intervalo de confianza del 95 % para total de horas que se pierden a la semana en las tareas insignificantes. Según los resultados se puede concluir:

a. Con una confianza del 95 % el número de total de horas que se pierden en la empresa en tareas insignificantes está entre 7711.90 y 7753.10 horas a la semana.
b. Con una confianza del 95 % el número de total de horas que se pierden en la empresa en tareas insignificantes está entre 5886.20 y 9578.80 horas a la semana.
c. Con una confianza del 95 % el número de total de horas que se pierden en la empresa en tareas insignificantes está entre 7609.41 y 7855.59 horas a la semana.
d. Con una confianza del 95 % el número de total de horas que se pierden en la empresa en tareas insignificantes está entre 7423.55 y 8041.45 horas a la semana.

Como ya se construyó IC para $\bar{Y}$ y además se hizo con el mismo nivel de confianza, basta con multiplicar por N el intervalo, así:

$$N\left(\bar{Y} \pm t_{\frac{\alpha}{2}, n-1} * \sqrt{var(\bar{Y})}\right)$$

(7412.992 8052.008)