Mario Alberto Rodríguez Morales 1860043 Grupo: 002

# Laboratorio 1: Muestreo Aleatorio Simple

1. En la auditoría contable de un hospital que tiene 1021 registros de pacientes se selecciona una muestra aleatoria simple de 200 registros y estos datos arrojan que la suma de las cuentas es de $1,844,400 y varianza muestral de $445.21. Estimar el promedio real de la deuda y establecer un límite para el error de estimación.

$N = 1021$

$n = 2000$

$\sum y_i = 1844400$

$s^2 = 445.21$

$$\hat{\mu} = \bar{y} = \frac{\sum y_i}{n} = \frac{1844400}{200} = 9222$$

$$\hat{Var}(\hat{\mu}) = \frac{s^2}{n}\left(\frac{N-n}{N}\right) = \frac{445.21}{200}\left(\frac{1021-200}{1021}\right)$$

$$Var(\hat{\mu}) = \frac{445.21}{200}\left(\frac{821}{1021}\right) = 1.7899$$

$$\hat{B} = 2\sqrt{Var(\hat{\mu})}$$

$$\hat{B} = 2\sqrt{1.7899}$$

$$\hat{B} = 2(1.3379) = 2.6758$$

Estableciendo un intervalo de confianza al 95%

$$\hat{\mu} - \hat{B} < \bar{y} < \hat{\mu} + \hat{B}$$

$$9222 - 2.6758 < \bar{y} < 9222 + 2.6758$$

$$9219.3242 < \bar{y} < 9224.6758$$

Con un margen de error

$$M.E: \frac{\hat{B}}{\hat{\mu}} * 100\% = \frac{2.6758}{9222} * 100\% = 0.029\%$$

∴ La deuda promedio real es de $9222 con una variabilidad de 1.7899. Este promedio se encuentra entre el rango de los $9219 y $9225. El límite para el error de estimación es de 2.6758.

2. Una muestra aleatoria simple de 9 registros del hospital es seleccionado para estimar la cantidad promedio de la deuda (en millones de pesos) sobre un total de 484 cuentas abiertas. Los valores de la muestra están a continuación:

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 34 | 43 | 45 | 32 | 40 | 43 | 52 | 41 | 39 |

Estimar la deuda promedio.

$N = 484$

$n = 9$

$$\bar{y} = \frac{\sum_{i=1}^{n} y_i}{n} = \frac{34+43+45+32+40+43+52+41+39}{9} = \frac{369}{9} = 41$$

$\hat{t} = N\bar{y} = (484)(41) = 19844$, deuda promedio

Con una variabilidad de:

$$Var(\hat{t}) = \frac{s^2 N}{n}(N-n)$$

$$s^2 = \frac{\sum_{i=1}^{n}(y_i - 41)^2}{9-1} = \frac{\sum (y_i-41)^2}{8} = \frac{280}{8} = 35$$

$$Var(\hat{t}) = \frac{(35)(484)}{9}(475) = \frac{8046500}{9} = 894055.5556$$

La cota para el error de estimación

$$B = 2\sqrt{Var(\hat{t})} = 2(945.5451) = 1891.09$$

Estableciendo un intervalo de confianza al 95%

$$\hat{t} - B < \hat{t} < \hat{t} + B$$

$$19844 - 1891.09 < \hat{t} < 19844 + 1891.09$$

$$17952.9097 < \hat{t} < 21735.0902$$

El margen de error es:

$$ME = \frac{B}{\hat{t}} * 100\% = \frac{1891.09}{19844} * 100\% = 9.5297\%$$

∴ La deuda promedio es de $19,844, aunque es un mal tamaño de muestra ya que la deuda está en el rango entre $17953 y $21735. Además se obtuvo un error mayor de 5%, se recomienda una muestra tamaño mayor para estimar la deuda con mayor exactitud.

3. Una empresa industrial está interesada en el tiempo por semana que los científicos emplean para ciertas tareas triviales. Las hojas de control del tiempo de una muestra aleatoria simple de 50 empleados muestran que la cantidad promedio de tiempo utilizado para esas tareas es de 10:31 hrs, con una varianza muestral de 2.25hrs. La compañía emplea 750 científicos. Estimar el número total de horas-hombre que se pierde por semana en las tareas insignificantes.

$N = 750$

$n = 50$

$\bar{y} = 10{:}31$ horas $= 631$ minutos

$s^2 = 2.25$ horas $= 135$ minutos

$$\hat{t} = N\bar{y} = (750)(631) = 473250 \text{ mins}$$

$$Var(\hat{t}) = \frac{s^2 N}{n}(N-n) = \frac{(135)(750)(700)}{50}$$

$$Var(\hat{t}) = \frac{70875000}{50} = 1417500 \text{ mins}^2$$

La cota para el error de estimación

$$B = 2\sqrt{Var(\hat{t})} = 2\sqrt{1417500} = 2(1190.58809) = 2381.17618 \text{ mins}$$

Convirtiendo a horas

$$\hat{t} = 473250 \cancel{\text{mins}} * \frac{\text{hrs}}{60 \cancel{\text{mins}}} = 7887.5 \text{hrs} = 7887 \text{ hrs y } 30 \text{ mins}$$

$$Var(\hat{t}) = 1417500 \cancel{\text{mins}^2} * \frac{\text{hrs}^2}{3600 \cancel{\text{mins}^2}} = 393.75 \text{hrs}^2$$

$$B = 2381.17618 \cancel{\text{mins}} * \frac{\text{hrs}}{60 \cancel{\text{mins}}} = 39.6862 \text{hrs} = 39 \text{hrs y } 41 \text{ mins}$$

Un intervalo de confianza al 95%.

$$\hat{t} - B < \hat{t} < \hat{t} + B$$

$$7887.5 - 39.6862 < \hat{t} < 7887.5 + 39.6862$$

$$7847.81373 < \hat{t} < 7927.18627$$

Y un margen de error de:

$$ME = \frac{39.6862}{7887.5} * 100\% = 0.5031\%$$

$\therefore$ Las horas-hombre totales que se pierden por semana en tareas insignificantes son 7887 horas y 30 minutos. Se encuentran en un rango de 7848hrs y 7927hrs con una confianza del 95% y un margen de error del 0.5031%.

4. Se requiere estimar la cantidad promedio de dinero para las cuentas por cobrar del hospital. Aunque no se cuenta con datos anteriores para estimar la varianza poblacional, se sabe que la mayoría de las cuentas cae dentro de un rango de variación de \$1000. Existen 1021 cuentas abiertas. Encuentre el tamaño de muestra necesario para estimar la deuda promedio por cliente con un límite para el error de estimación de \$30.00

$R = 1000$

$N = 1021$

$B = 30$

$$D = \frac{B^2}{4} = \frac{(30)^2}{4} = \frac{900}{4} = 225$$

$$\sigma^2 = \frac{R^2}{16} = \frac{(1000)^2}{16} = \frac{1000000}{16} = 62500$$

$$n = \frac{N\sigma^2}{(N-1)D+\sigma^2} = \frac{(1021)(62500)}{(1020)(225)+62500} = \frac{63812500}{229500+62500}$$

$$n = \frac{63812500}{292000} = 218.5359588 \approx 219$$

$\therefore$ La muestra necesaria para un buen muestreo debe ser de tamaño 219

5. Un investigador está interesado en estimar la ganancia en peso total en un periodo de 0 a 4 semanas de 2000 polluelos alimentados con una nueva fórmula. Obviamente pesar cada ave sería tedioso y tardado; por lo tanto se decide seleccionar una muestra de polluelos y con la información que se obtenga, estimar el peso total ganado por los polluelos. Determinar el número de polluelos que serán seleccionados en este estudio para estimar el peso total límite para el error de estimación igual a 100 gramos. Muchos estudios similares sobre nutrición de polluelos se han llevado a cabo en el pasado y usando los datos de esos estudios, el investigador encontró que la varianza poblacional es de 36. Determinar el tamaño de muestra requerido.

$N = 2000$

$B = 100$

$\sigma^2 = 36$

$$D = \frac{B^2}{4N^2} = \frac{(100)^2}{4(2000)^2} = \frac{10000}{16000000} = 0.000625$$

$$n = \frac{N\sigma^2}{(N-1)D + \sigma^2} = \frac{(2000)(36)}{(1999)(0.000625) + 36}$$

$$n = \frac{72000}{1.249375 + 36} = \frac{72000}{37.249375}$$

$$n = 1932.918338 \approx 1933$$

$\therefore$ Se deben seleccionar y analizar 1933 polluelos para hacen una estimación correcta en la ganancia del peso total.

6. Una muestra aleatoria simple de 100 estudiantes del último año de preparatoria fue seleccionada para estimar: (1) la proporción de estudiantes del último año que asistirán a una universidad de un total de 300 y (2) la proporción de estudiantes que han tenido trabajos de tiempo parcial durante su estancia en la preparatoria. Estimar $p_1$, la proporción de estudiantes del último año que planea asistir a una universidad y $p_2$, la proporción de estudiantes del último año que han tenido un trabajo de tiempo parcial durante sus cursos en el colegio (incluyendo los veranos) si de la muestra, 15 estudiantes planean asistir a una institución superior y 65 han tenido algún trabajo de tiempo parcial durante su estancia en el colegio.

$n = 100$
$X_1 = 15$
$X_2 = 65$

Proporción 1: Estudiantes del último año que asistirán a la universidad

$$\hat{p} = \frac{15}{100} = 0.15$$

$$E[\hat{p}] = 0.15$$

$$Var[\hat{p}] = \frac{\hat{p}\hat{q}}{n}\left[\frac{N-n}{N-1}\right] = \frac{(0.15)(0.85)}{100}\left[\frac{300-100}{300-1}\right]$$

$$Var[\hat{p}] = (0.001275)(0.668896321) = 0.000852842$$

$$B = 2\sqrt{Var(\hat{p})} = 2(0.0292) = 0.05860327$$

Con un intervalo de confianza al 95%

$$\hat{p} - B < \hat{p} < \hat{p} + B$$

$$0.15 - 0.0586 < \hat{p} < 0.15 + 0.0586$$

$$0.0913 < \hat{p} < 0.2086$$

Y un margen de error de

$$ME = \frac{0.0586}{0.15} * 100\% = 39.0688\%$$

$\therefore$ La proporción de estudiantes que asistirán a la universidad es el 0.15. Está en un rango entre el 0.0913 y 0.2086 con un 95% de confianza. Es un número muy bajo de estudiantes, ya que se tiene un margen de error del casi 40%.

Proporción 2: Estudiantes del último año que han tenido trabajos de tiempo parcial durante su estancia en el colegio.

$$\hat{p} = \frac{65}{100} = 0.65$$

$E[\hat{p}] = 0.65$

$$Var[\hat{p}] = \frac{(0.65)(0.35)}{100}\left[\frac{300-100}{300-1}\right] = \frac{0.2275}{100}\left[\frac{200}{299}\right]$$

$Var[\hat{p}] = (0.002275)(0.668896321) = 0.001521739$

$B = 2\sqrt{Var(\hat{p})} = 2(0.039) = 0.078281199$

Un intervalo de confianza al 95%

$$\hat{p} - B < \hat{p} < \hat{p} + B$$

$$0.65 - 0.0783 < \hat{p} < 0.65 + 0.0783$$

$$0.5717 < \hat{p} < 0.7283$$

El margen de error de

$$ME = \frac{0.0783}{0.65} * 100\% = 12.046\%$$

$\therefore$ La proporción de estudiantes del último año que han tenido trabajos de tiempo parcial es 0.65, la cual con un 95% de confianza se encuentra entre el rango de 0.5717 y 0.7283. Se obtuvo un margen de error del 12.046% por lo cual se recomienda aumentar la muestra ya que a pesar de ser un alto número de estudiantes (éxitos) se tiene un margen de error muy alto.