Práctica 6: *Fundamentos de la Inferencia.*

Teorema Central del Límite.

1) Dada una población de media $\mu = 400$ y varianza $\sigma^2 = 1600$. Se obtiene una muestra aleatoria de tamaño 35. Detallar bajo qué condiciones se puede aplicar el TCL y responder:
   a) ¿Cuáles son la media y la varianza de la distribución de las medias muestrales en el muestreo?
   b) ¿Cuál es la probabilidad de que $\bar{x} > 412$?
   c) ¿Cuál es la probabilidad de que $393 \leq \bar{x} \leq 407$?
   d) ¿Cuál es la probabilidad de que $\bar{x} \leq 389$?
2) Completar y detallar bajo qué condiciones se puede aplicar el TCL:
   a) Se selecciona una muestra de tamaño $n = 50$ de una distribución con media $\mu = 53$ y $\sigma = 21$. La distribución muestral de $\bar{x}$ será aproximadamente ________________ con una media de _____________ y una desviación estándar (o error estándar) de ______________________.
   b) Se selecciona una muestra de tamaño $n = 40$ de una distribución con media $\mu = 100$ y $\sigma = 20$. La distribución muestral de $\bar{x}$ será aproximadamente ________________ con una media de _____________ y una desviación estándar (o error estándar) de ______________________.
3) La duración de las bombillas de un fabricante tiene una media de 1.200 horas y una desviación típica de 400 horas. La distribución poblacional es normal. Suponga que compra nueve bombillas, que puede considerarse que son una muestra aleatoria de la producción del fabricante.
   a) ¿Cuál es la media de la media muestral de la duración?
   b) ¿Cuál es la varianza de la media muestral?
   c) ¿Cuál es el error típico de la media muestral?
   d) ¿Cuál es la probabilidad de que esas nueve bombillas tengan, en promedio, una duración de menos de 1.050 horas?
4) El consumo de combustible, en kilómetros por litro, de todos los automóviles de un determinado modelo tiene una media de 25 y una desviación típica de 2. Se puede suponer que la distribución poblacional es normal. Se toma una muestra aleatoria de estos automóviles.
   a) Halle la probabilidad de que la media muestral del consumo de combustible sea inferior a 24 kilómetros por litro suponiendo que
      i) se toma una muestra de una observación.

   ii) se toma una muestra de cuatro observaciones.
   iii) se toma una muestra de 16 observaciones.

b) Explique por qué las tres respuestas del apartado (a) son diferentes. Trace un gráfico para explicar su razonamiento.

5) Se ha tomado una muestra aleatoria de 16 directivos de empresas de una gran ciudad para estimar el tiempo medio que tardan diariamente en desplazarse al trabajo. Suponga que los tiempos poblacionales siguen una distribución normal que tiene una media de 87 minutos y una desviación típica de 22 minutos.
   a) ¿Cuál es el error típico de la media muestral de los tiempos de desplazamiento?
   b) ¿Cuál es la probabilidad de que la media muestral sea de menos de 100 minutos?
   c) ¿Cuál es la probabilidad de que la media muestral sea de más de 80 minutos?
   d) ¿Cuál es la probabilidad de que la media muestral esté fuera del intervalo 85-90 minutos?
   e) Suponga que se toma una segunda muestra aleatoria (independiente) de 50 directivos. Indique, sin realizar los cálculos, si las probabilidades de los apartados (b), (c) y (d) serían mayores, menores o iguales en el caso de la segunda muestra. Ilustre sus respuestas gráficamente.

6) Una empresa decide realizar una prueba de aptitud a sus futuros empleados. Sabe, en base a estudios anteriores, que la cantidad media de aciertos es de 1000 con una desviación estándar de 125. Si se aplica la prueba a 100 individuos seleccionados al azar:
   a) ¿Cuál es la probabilidad de que la media muestral se encuentre entre 985 y 1015 aciertos?
   b) ¿Cuál es la probabilidad de que la media muestral sea de como mínimo 1020 aciertos?
   c) ¿Cuál es la probabilidad de que la media muestral se encuentre entre 960 y 1040 aciertos?

7) Mattel Corporation produce autos que funcionan con baterías AA. La vida media de las baterías posee una distribución normal con una media de 35 horas y una desviación estándar de 5,5 horas. Como parte de su programa se prueban 25 baterías:
   a) ¿Qué proporción de las muestras tendrá una media de vida útil de más de 36 hs?
   b) ¿Qué proporción de las muestras tendrá una media de vida útil entre 34,5 hs y 36 hs?

8) El tiempo que dedican los estudiantes a estudiar la semana antes de los exámenes finales sigue una distribución normal que tiene una desviación típica de 8 horas. Se toma una muestra aleatoria de cuatro estudiantes para estimar el tiempo medio de estudio de la población total de estudiantes.

a) ¿Cuál es la probabilidad de que la media muestral sea más de dos horas superior a la media poblacional?
b) ¿Cuál es la probabilidad de que la media muestral sea más de tres horas inferior a la media poblacional?
c) ¿Cuál es la probabilidad de que la media muestral difiera más de cuatro horas de la media poblacional?
d) Suponga que se toma una segunda muestra aleatoria (independiente) de diez estudiantes. Indique sin realizar los cálculos si las probabilidades de los apartados (a), (b) y (c) serán mayores, menores o iguales en el caso de la segunda muestra.

9) Un proceso industrial produce lotes de un producto químico cuyos niveles de impurezas siguen una distribución normal que tiene una desviación típica de 1,6 gramos por 100 gramos de producto químico. Se selecciona una muestra aleatoria de 100 lotes para estimar la media poblacional de los niveles de impurezas.
   a) La probabilidad de que la media muestral de los niveles de impurezas sea _________ gramos mayor que la media poblacional es de 0,05.
   b) La probabilidad de que la media muestral de los niveles de impurezas sea _________ gramos menor que la media poblacional es de 0,10.
   c) La probabilidad de que la media muestral de los niveles de impurezas difiera en _________ gramos de la media poblacional es de 0,15.

10) Las relaciones precio-beneficio de todas las empresas cuyas acciones cotizan en bolsa siguen una distribución normal que tiene una desviación típica de 3,8. Se selecciona una muestra aleatoria de estas empresas para estimar la media poblacional de las relaciones precio-beneficio.
   a) ¿Cuál debe ser el tamaño de la muestra para garantizar que la probabilidad de que la media muestral difiera más de 1,0 de la media poblacional es de menos de 0,10?
   b) Indique sin realizar los cálculos si sería necesaria una muestra mayor o menor que la del apartado (a) para garantizar que la probabilidad de que la media muestral difiera en más de 1,0 de la media poblacional es de menos de 0,05.
   c) Indique sin realizar los cálculos si sería necesaria una muestra mayor o menor que la del apartado (a) para garantizar que la probabilidad de que la media muestral difiera en más de 1,5 de la media poblacional es de menos de 0,10.

11) Una empresa metalúrgica produce cables de acero y se encuentra analizando la variable “resistencia a la rotura” de los mismos. Se conoce por estudios anteriores que esta variable posee una media de 15800 kg/m y un desvío estándar de 2600 kg/m. Se selecciona al azar una muestra de 36 cables de la misma longitud.

a) ¿Cuál es la probabilidad de que la media de la muestra sea superior a 17360 kg/m?
b) Determinar el valor de la media muestral superada por el 10% de los cables.
c) ¿Cuál es la probabilidad de que la media de la muestra se encuentre entre 16000 Kg/m y 18000 Kg/m?

12) La duración de la enfermedad de Alzheimer desde el inicio de los síntomas hasta la muerte varía de tres a veinte años. El promedio de duración de la enfermedad es de ocho años con una desviación estándar de cuatro años. El administrador de un centro médico selecciona al azar, de la base de datos los expedientes de 30 pacientes y registra la duración media de la enfermedad. Calcular:
a) probabilidad de que la duración media de la muestra sea inferior a los siete años.
b) probabilidad de que la duración media de la muestra este comprendida entre siete y nueve años.

13) Se ha determinado que el tiempo necesario para que los alumnos terminen un examen final es una variable aleatoria con distribución normal con media 84 minutos y desvío 18 minutos.
a) Se seleccionan 9 alumnos al azar ¿Cuál es la probabilidad de que el tiempo medio de finalización de esta muestra sea de al menos 89 minutos?
b) Si se seleccionan muestras de 5 alumnos ¿Cuál es el tiempo medio mínimo del 8% de dichas muestras?

14) Una empresa ha determinado que el monto de las ventas realizadas es una variable que tiene distribución normal con una media de $343.000 y un desvío de $48140. Si en el mes ha realizado 1800 ventas y se selecciona una muestra de 20 facturas.
a) ¿Cuál es la probabilidad de que la media muestral difiera de la media poblacional en más de $20000?
b) ¿Cuál es el valor de la media muestral que es superado con probabilidad 0,05?

15) Se selecciona una muestra aleatoria de tamaño $n = 50$ de una distribución binomial con una media de $p = 0{,}7$. La distribución muestral de $\hat{p}$ será aproximadamente ____________ con una media de ______________ y una desviación estándar (o error estándar) de _______. La probabilidad de que la proporción muestral sea inferior a 0,8 es___________________.

16) Se seleccionaron muestras aleatorias de tamaño $n = 500$ de una población binomial con $p = 0{,}1$:
a) ¿Es apropiado utilizar la distribución normal para aproximar la distribución de muestreo de $\hat{p}$? Justificar
b) De acuerdo al ítem anterior, calcular las siguientes probabilidades:

i) $P(\hat{p} > 0{,}12)$

ii) $P(\hat{p} < 0{,}08)$

iii) $P(0{,}09 < \hat{p} < 0{,}12)$

17) Suponga que tenemos una población con una proporción $p = 0{,}60$ y una muestra aleatoria de tamaño $n = 100$ extraída de la población.

a) ¿Cuál es la probabilidad de que la proporción muestral sea superior a 0,66?

b) ¿Cuál es la probabilidad de que la proporción muestral sea inferior a 0,48?

c) ¿Cuál es la probabilidad de que la proporción muestral esté comprendida entre 0,52 y 0,66?

18) En 1992, los canadienses votaron en un referéndum sobre una nueva constitución. En la provincia de Quebec, el 42,4% de los que votaron estaba a favor de la nueva constitución. Se extrajo una muestra aleatoria de 100 votantes de la provincia.

a) ¿Cuál es la media de la distribución de la proporción muestral a favor de una nueva constitución?

b) ¿Cuál es la varianza de la proporción muestral?

c) ¿Cuál es el error típico de la proporción muestral?

d) ¿Cuál es la probabilidad de que la proporción muestral sea superior a 0,5?

19) Una empresa de internet informa que el porcentaje de estudiantes que utilizan internet como su principal recurso para un proyecto escolar en el último año fue 66%. Suponiendo que se selecciona una muestra de 1000 estudiantes:

a) ¿Cuál es la probabilidad de que la proporción muestral sea superior a 0,68?

b) ¿Cuál es la probabilidad de que la proporción muestral se encuentre comprendida entre 0,64 y 0,68?

c) Determinar la proporción muestral determinada por Q3.

20) Una empresa está considerando la posibilidad de realizar una nueva emisión de bonos convertibles. La dirección cree que los términos de la oferta serán atractivos para el 20% de todos sus accionistas actuales. Suponga que está en lo cierto. Se toma una muestra aleatoria de 130 accionistas actuales.

a) ¿Cuál es el error típico de la proporción muestral que piensa que esta oferta es atractiva?

b) ¿Cuál es la probabilidad de que la proporción muestral sea superior a 0,15?

c) ¿Cuál es la probabilidad de que la proporción muestral esté comprendida entre 0,18 y 0,22?

d) Suponga que se hubiera tomado una muestra de 500 accionistas actuales. Indique, sin realizar los cálculos, si las probabilidades de los apartados (b) y (c) habrían sido mayores, menores o iguales que las obtenidas.

21) Se toma una muestra aleatoria de 100 votantes para estimar la proporción del electorado que está a favor de una subida del impuesto sobre la gasolina para obtener más ingresos para reparar las autopistas. ¿Cuál es el valor más alto que puede tomar el error típico de la proporción muestral que está a favor de esta medida?

22) Suponga en el ejercicio anterior que se decide que una muestra de 100 votantes es demasiado pequeña para obtener una estimación suficientemente fiable de la proporción poblacional. Se exige, por el contrario, que la probabilidad de que la proporción muestral difiera de la proporción poblacional (cualquiera sea su valor) en más de 0,03 no sea superior a 0,05. ¿De qué tamaño debe ser la muestra para que se cumpla este requisito?

23) Las tasas mensuales de rendimiento de las acciones de una empresa son independientes de las de otra y siguen una distribución normal que tiene una desviación típica de 1,7. Se toma una muestra de 12 meses.

a) Halle la probabilidad de que la desviación típica muestral sea inferior a 2,5.

b) Halle la probabilidad de que la desviación típica muestral sea superior a 1,0.

24) Se cree que los sueldos que perciben durante el primer año los contadores recién egresados siguen una distribución normal que tiene una desviación típica de \$2.500. Se toma una muestra aleatoria de 16 observaciones.

a) Halle la probabilidad de que la desviación típica muestral sea superior a \$3.000.

b) Halle la probabilidad de que la desviación típica muestral sea inferior a \$1.500.

25) En una gran ciudad se ha observado que durante el verano las facturas del consumo de electricidad siguen una distribución normal que tiene una desviación típica de \$100. Se ha tomado una muestra aleatoria de 25 facturas.

a) Halle la probabilidad de que la varianza muestral sea inferior a 5625.

b) Halle la probabilidad de que la desviación típica muestral sea superior a \$150.

26) Sea X una variable normal de media μ y desviación típica σ. Sea s^2 la varianza de muestras aleatorias de X de tamaño 41. Si se sabe que $P(s^2 > 46{,}763) = 0{,}10$, calcular el valor de σ.

27) Suponga que se tiene una muestra aleatoria de tamaño $2n$ tomada de una población X, con $E[X] = \mu \; y \; Var[X] = \sigma^2$. Sean:

$$\overline{X}_1 = \frac{1}{2n}\sum_{i=1}^{2n} X_i \qquad y \qquad \overline{X}_2 = \frac{1}{n}\sum_{i=1}^{n} X_i$$

dos estimadores de μ . ¿Cuál es el mejor estimador de μ ? Explique su elección.

28) Sean $X_1,\ldots,X_7$ una muestra aleatoria de una población que tiene media μ y varianza σ^2 . Considere los siguientes estimadores de μ :

$$\hat{\mu}_1 = \frac{X_1 + \cdots + X_7}{7} \qquad y \qquad \hat{\mu}_2 = \frac{2X_1 - X_6 + X_4}{2}$$

a) ¿Alguno de estos estimadores es insesgado?
b) ¿Cuál es mejor? ¿En qué sentido es mejor?

29) Suponga que $\hat{\Theta}_1\, y\, \hat{\Theta}_2$ son estimadores insesgados del parámetro θ. Se sabe que $Var(\hat{\Theta}_1) = 10 y\; Var(\hat{\Theta}_2) = 4$ ¿Cuál es mejor? ¿En qué sentido lo es?

30) Calcule la eficiencia relativa de los estimadores del ejercicio 27.

31) Calcule la eficiencia relativa de los estimadores del ejercicio 28.

Trabajo Práctico en RStudio Cloud

En el ejercicio 14 del trabajo práctico en R se utilizan datos del mercado inmobiliario de una ciudad de EEUU. Se toman muestras aleatorias de reducido tamaño y se analiza la distribución de las medias muestrales, intentando ver si se cumplen los resultados del Teorema Central del Límite.