Trabajo práctico 9: Intervalos de confianza

Apellido y nombre: ..... Comisión: ..... Fecha: .....

AL TERMINAR ESTA GUÍA

- Construir intervalos de confianza para una media, con la distribución t, y para una proporción.
- Redactar la interpretación de un intervalo sin confundir la media con los datos individuales.
- Calcular el tamaño de muestra necesario para un margen de error dado, para una proporción y para una media.
- Comprender que el nivel de confianza es una propiedad del método, que depende de que se cumplan los supuestos.

EJERCICIO 1 La hemoglobina de las embarazadas

A mano

En un centro de atención primaria se midió la hemoglobina (en g/dL) de 12 embarazadas en el tercer trimestre, elegidas al azar entre las que se controlan ahí. En g/dL:

Embarazadas 1 a 6	11,2	12,4	9,1	12,4	12,2	12,9

Embarazadas 7 a 12	12,2	12,4	9,9	10,0	9,6	11,1

Se puede suponer que la hemoglobina tiene distribución aproximadamente normal en esa población.

a) Calculá la media y el desvío estándar de la muestra, y el error estándar de la media.

b) Construí el intervalo de confianza del 95 % para la hemoglobina media. ¿Por qué se usa la distribución t y no la normal? ¿Con cuántos grados de libertad?

c) Redactá la interpretación del intervalo en una oración que entienda alguien sin formación estadística.

d) Calculá el intervalo del 99 %. ¿Es más ancho o más angosto? ¿Por qué?

e) Si se hubieran medido 48 embarazadas y el desvío hubiera dado lo mismo, ¿cuánto valdría el margen de error del 95 %?

EJERCICIO 2 ¿Cuántos usan el colectivo?

A mano

Una encuesta a 500 personas adultas de una ciudad, elegidas por muestreo aleatorio simple, preguntó si usan el transporte público al menos una vez por semana. Respondieron que sí 253 personas.

a) Estimá la proporción de adultos que usan el transporte público cada semana, y su error estándar. ¿Se puede usar la aproximación normal?

b) Construí el intervalo de confianza del 95 % e informá el resultado como lo haría un diario, con el margen de error.

c) Calculá también los intervalos del 90 % y del 99 %.

d) En una encuesta de hace cinco años, la proporción era del 46 %. ¿Los datos nuevos son compatibles con que no cambió? Respondé con el intervalo del 95 % y con el del 99 %.

EJERCICIO 3 ¿A cuántos hogares hay que encuestar?

A mano

Un municipio de 12.000 hogares planea una encuesta para estimar qué proporción de hogares tiene acceso a internet fijo. Quiere un margen de error de 4 puntos porcentuales con un nivel de confianza del 95 %.

a) Si no se tiene ninguna idea previa de la proporción, ¿cuántos hogares hay que encuestar? ¿Por qué se usa p = 0,5?

b) La población es finita: 12.000 hogares. Aplicá la corrección por población finita. ¿Cambia mucho?

c) Un estudio anterior encontró que alrededor del 25 % de los hogares tenía internet fijo. Si se usa ese dato, ¿cuántos hogares hacen falta (sin la corrección por población finita)?

d) Se espera que responda solo el 80 % de los hogares contactados. ¿A cuántos hay que ir a buscar para conseguir las 572 respuestas de la consigna b?

e) La misma encuesta quiere estimar el gasto mensual en internet. Una prueba piloto dio un desvío estándar de 40 (en miles de pesos). ¿Cuántos hogares hacen falta para un margen de error de 10 mil pesos con 95 % de confianza? ¿Y para la mitad de ese margen?

EJERCICIO 4 Un intervalo en un informe

A mano

Un informe de una encuesta de turismo dice: «El gasto diario promedio por turista fue de $50.900. Intervalo de confianza del 95 %: $47.900 a $53.900. Muestra: 300 turistas.»

a) ¿Cuánto vale el margen de error? ¿Y el error estándar de la media?

b) ¿Cuánto vale aproximadamente el desvío estándar del gasto entre los turistas encuestados?

c) Un periodista escribe: «El 95 % de los turistas gasta entre $47.900 y $53.900 por día». ¿Está bien?

d) Calculá el intervalo del 90 %. ¿Cuál sería el margen del 95 % si la muestra hubiera sido cuatro veces más grande?

EJERCICIO 5 Intervalos con la EPH

Con software

Se usa la base «Ocupación e ingresos en Gran Catamarca», extracto de la EPH del primer trimestre de 2026 (estadistica.ar/datos/eph-catamarca-2026t1). Se puede usar R, Python, Excel, InfoStat o SPSS. En R, t.test(x)$conf.int da el intervalo t para la media y prop.test(k, n) el de una proporción.

En esta guía se trata a la base como una muestra aleatoria simple de personas, sin usar el ponderador. La EPH real tiene un diseño por conglomerados y ponderadores, y sus márgenes de error son más grandes que los que da esta cuenta.

a) Para las personas ocupadas que declararon ingreso (ESTADO = 1 y P21 > 0), calculá el intervalo de confianza del 95 % para el ingreso medio de la ocupación principal.

b) Calculá la tasa de desocupación (desocupados sobre activos, es decir ESTADO = 2 sobre ESTADO = 1 o 2) y su intervalo del 95 %.

c) Tomando a las personas ocupadas con ingreso como si fueran toda la población, simulá 2.000 muestras de 10 personas, calculá el intervalo t del 95 % de cada una y contá en qué proporción contiene a la media verdadera. Repetí con muestras de 30.

d) ¿Qué conclusión práctica sacás de la simulación?

EJERCICIO 6 ¿Qué dice y qué no dice un intervalo?

Para discutir

Un estudio informa que la edad media de las madres primerizas de una provincia está entre 24,1 y 25,3 años, con 95 % de confianza. Decidí si cada afirmación es correcta y, si no lo es, corregila.

a) «El 95 % de las madres primerizas tiene entre 24,1 y 25,3 años.»

b) «Si se repitiera el estudio muchas veces con muestras distintas, alrededor del 95 % de los intervalos calculados así contendría la edad media verdadera.»

c) «Hay un 95 % de probabilidad de que la edad media esté entre 24,1 y 25,3.»

d) «Si la muestra hubiera sido más grande, el intervalo del 95 % tendría más probabilidad de contener a la media.»

e) «En otra provincia, el intervalo dio de 24,9 a 26,0. Como los intervalos se superponen, la edad media es la misma en las dos provincias.»

f) «Como el intervalo del 99 % es más ancho, es peor que el del 95 %.»