Trabajo práctico 9: Intervalos de confianza

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.

PARA EL DOCENTE

Los ejercicios 1 a 4 son para el encuentro práctico, con las tablas de la t y de la normal; el 1 y el 3 son los centrales. El 5 necesita computadora: sirve para un encuentro virtual o como tarea, y se complementa con el simulador de cobertura.

El 6 conviene hacerlo en voz alta, con todo el curso: las interpretaciones erróneas que propone son las que aparecen en los parciales y en los diarios.

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.

RESOLUCIÓN

$$\bar{x} = \frac{135{,}4}{12} = 11{,}283 \qquad s = 1{,}322 \qquad EE = \frac{s}{\sqrt{n}} = \frac{1{,}322}{\sqrt{12}} = 0{,}3817$$

El desvío se calcula dividiendo por n − 1, porque estima el de la población. El error estándar usa ese desvío: σ no se conoce.

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?

> RESOLUCIÓN
>
> $$\bar{x} \pm t_{11;\,0{,}975}\,\frac{s}{\sqrt{n}} = 11{,}283 \pm 2{,}201 \cdot 0{,}3817 = 11{,}283 \pm 0{,}840 \;\Rightarrow\; (10{,}44;\ 12{,}12)$$
>
> Se usa la t porque el desvío poblacional no se conoce y se reemplaza por s, que también varía de muestra a muestra: la t tiene colas más pesadas que la normal y da un intervalo más ancho, que compensa esa incertidumbre extra. Los grados de libertad son n − 1 = 11. Con la normal, 1,96 en lugar de 2,201, el intervalo quedaría demasiado angosto para una muestra tan chica.

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

> RESOLUCIÓN
>
> «Con un 95 % de confianza, la hemoglobina media de las embarazadas en el tercer trimestre que se controlan en este centro está entre 10,44 y 12,12 g/dL.» La confianza se refiere al método: si se repitiera el estudio muchas veces, el 95 % de los intervalos construidos así contendría la media verdadera. No dice que el 95 % de las embarazadas tenga la hemoglobina en ese rango: eso describiría a las personas, y el intervalo describe a la media.

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

> RESOLUCIÓN
>
> $$11{,}283 \pm 3{,}106 \cdot 0{,}3817 = 11{,}283 \pm 1{,}186 \;\Rightarrow\; (10{,}10;\ 12{,}47)$$
>
> Más ancho. Para tener más confianza en que el intervalo atrape a la media, hay que abrirlo más: con los mismos datos, la confianza se paga con precisión. La única forma de tener las dos cosas es aumentar la muestra.

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

> RESOLUCIÓN
>
> $$t_{47;\,0{,}975} \cdot \frac{1{,}322}{\sqrt{48}} = 2{,}012 \cdot 0{,}1909 = 0{,}384$$
>
> Pasa de 0,840 a 0,384: menos de la mitad. Cuadruplicar la muestra divide el error estándar por dos (por la raíz de n), y además el cuantil t baja de 2,201 a 2,012, porque con más grados de libertad la t se parece más a la normal.

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?

> RESOLUCIÓN
>
> $$\hat{p} = \frac{253}{500} = 0{,}5060 \qquad EE(\hat{p}) = \sqrt{\frac{\hat{p}\,(1-\hat{p})}{n}} = \sqrt{\frac{0{,}5060 \cdot 0{,}4940}{500}} = 0{,}0224$$
>
> Sí: hay 253 respuestas «sí» y 247 «no», las dos muy por encima de 10. En el error estándar se usa p̂ en lugar de p, que no se conoce.

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

> RESOLUCIÓN
>
> $$\hat{p} \pm 1{,}96 \cdot EE = 0{,}5060 \pm 0{,}0438 \;\Rightarrow\; (0{,}4622;\ 0{,}5498)$$
>
> «El 50,6 % de los adultos usa el colectivo al menos una vez por semana, con un margen de error de 4,4 puntos porcentuales y un nivel de confianza del 95 %.» El margen de error es la mitad del ancho del intervalo; se informa en puntos porcentuales, no como porcentaje del resultado.

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

> RESOLUCIÓN
>
Confianza	z	Margen de error	Intervalo
90 %	1,645	0,0368	0,4692 a 0,5428
95 %	1,96	0,0438	0,4622 a 0,5498
99 %	2,576	0,0576	0,4484 a 0,5636
>
> El centro es siempre p̂; solo cambia cuánto se abre. Los tres intervalos salen de los mismos datos: elegir el nivel de confianza es una decisión previa, no algo que se ajusta después de ver los resultados.

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 %.

> RESOLUCIÓN
>
> El 46 % queda afuera del intervalo del 95 % (0,4622 a 0,5498), pero adentro del de 99 % (0,4484 a 0,5636). Con 95 % de confianza, los datos indican que la proporción aumentó; con 99 %, no alcanzan para descartar que siga igual.
>
> La conclusión depende del nivel elegido, y eso es lo que hay que informar: la evidencia de un cambio es moderada, no contundente. Esta lógica —un valor fuera del intervalo es poco compatible con los datos— es la misma de las pruebas de hipótesis, que se ven la semana que viene. Además, en estricto rigor el valor de hace cinco años también tenía su propio error de muestreo.

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?

> RESOLUCIÓN
>
> $$n_0 = \frac{z^2\, p\,(1-p)}{E^2} = \frac{1{,}96^2 \cdot 0{,}5 \cdot 0{,}5}{0{,}04^2} = 600{,}2 \;\Rightarrow\; 601$$
>
> Se redondea siempre para arriba: redondear para abajo daría un margen apenas mayor que el pedido. Con p = 0,5 el producto p(1 − p) es el máximo posible (0,25), así que es el caso más desfavorable: el tamaño alcanza sea cual sea la proporción verdadera.

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

> RESOLUCIÓN
>
> $$n = \frac{n_0}{1 + \frac{n_0 - 1}{N}} = \frac{600{,}2}{1 + \frac{599{,}2}{12000}} = 571{,}7 \;\Rightarrow\; 572$$
>
> Baja de 601 a 572: la muestra es el 5,0 % de la población, y por eso la corrección se nota algo. Con poblaciones grandes, la corrección casi no cambia nada: el tamaño de muestra depende del margen pedido, no del tamaño de la población. Una encuesta nacional y una municipal necesitan muestras parecidas para la misma precisión.

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)?

> RESOLUCIÓN
>
> $$n_0 = \frac{1{,}96^2 \cdot 0{,}25 \cdot 0{,}75}{0{,}04^2} = 450{,}2 \;\Rightarrow\; 451$$
>
> Bastante menos que 601, porque p(1 − p) es menor lejos de 0,5. El riesgo: si la proporción cambió y ahora está más cerca de 0,5, el margen real va a ser mayor que el planeado. Si la encuesta mide varias proporciones a la vez, se usa 0,5.

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?

> RESOLUCIÓN
>
> $$\frac{572}{0{,}8} = 715{,}0 \;\Rightarrow\; 715$$
>
> Agrandar la muestra compensa la cantidad de respuestas, pero no el sesgo: si los hogares que no responden son distintos de los que sí (por ejemplo, los que no están en casa porque trabajan todo el día), la estimación puede quedar sesgada aunque se llegue al tamaño buscado.

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?

> RESOLUCIÓN
>
> $$n = \left(\frac{z\,\sigma}{E}\right)^2 = \left(\frac{1{,}96 \cdot 40}{10}\right)^2 = 61{,}5 \;\Rightarrow\; 62$$
>
> Para un margen de 5,0 mil pesos hacen falta 246 hogares: unas cuatro veces más. Dividir el margen por dos multiplica la muestra por cuatro, y eso suele decidir cuánta precisión se puede pagar. El desvío sale de un piloto o de estudios anteriores, y si se subestima, el margen real será mayor.

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?

> RESOLUCIÓN
>
> $$E = \frac{53900 - 47900}{2} = 3.000 \qquad EE = \frac{E}{1{,}96} = \frac{3000}{1{,}96} = 1.530{,}6$$
>
> El centro del intervalo es la media muestral, y la mitad de su ancho es el margen de error. Con n tan grande, el cuantil t y el z son prácticamente iguales, y se puede usar 1,96.

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

> RESOLUCIÓN
>
> $$s = EE \cdot \sqrt{n} = 1.530{,}6 \cdot \sqrt{300} = 26.511$$
>
> El desvío es mucho más grande que el margen: los turistas gastan muy distinto unos de otros (un desvío de unos $26.511), pero el promedio de 300 se conoce con bastante precisión. Es la diferencia entre la variabilidad de los datos y la del estimador.

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

> RESOLUCIÓN
>
> No. El intervalo es para el gasto promedio, no para el gasto de cada turista. Con un desvío de unos $26.511, los gastos individuales están mucho más dispersos: una gran parte de los turistas gasta fuera de ese rango. Lo correcto: «Con 95 % de confianza, el gasto diario promedio por turista está entre $47.900 y $53.900».

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

RESOLUCIÓN

$50.900 \pm 1{,}645 \cdot 1.530{,}6 = 50.900 \pm 2.518 \;\Rightarrow\; (48.382;\; 53.418)$

Con 1200 turistas y el mismo desvío, el error estándar se divide por $\sqrt{4} = 2$, y el margen pasa a ser de unos $1.500. Para reproducir los cálculos de un informe, a veces alcanza con el intervalo y el tamaño de muestra: el resto se despeja.

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.

RESOLUCIÓN

$n = 594 \qquad \bar{x} = 881.992 \qquad s = 935.780 \qquad t_{593;\,0{,}975} = 1{,}964$

$$881.992 \pm 1{,}964 \cdot \frac{935.780}{\sqrt{594}} = 881.992 \pm 75.408$$

Con 95 % de confianza, el ingreso medio de la ocupación principal está entre $806.584 y $957.399. En R: x ← eph$P21[eph$ESTADO == 1 & eph$P21 > 0]; t.test(x)$conf.int. Aunque los ingresos son muy asimétricos, con 594 casos el teorema central del límite hace que el intervalo t funcione bien.

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

RESOLUCIÓN

$$\hat{p} = \frac{26}{657} = 0{,}0396 \qquad 0{,}0396 \pm 1{,}96\sqrt{\frac{0{,}0396 \cdot 0{,}9604}{657}} = 0{,}0396 \pm 0{,}0149$$

Entre 2,5 % y 5,4 %. Hay 26 desocupados en la muestra: por encima de 10, el mínimo habitual para la aproximación normal, pero no por mucho; prop.test usa el intervalo de Wilson, con corrección por continuidad, que da un intervalo algo asimétrico, más confiable con proporciones chicas. El INDEC publica la tasa con el ponderador y con márgenes que tienen en cuenta el diseño, más anchos que este.

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.

RESOLUCIÓN

En R: mu ← mean(x); cubre ← replicate(2000, { m ← sample(x, 10); ic ← t.test(m)$conf.int; ic[1] <= mu & mu <= ic[2] }); mean(cubre).

En una corrida, los intervalos con n = 10 contuvieron a la media en el 88,0 % de las muestras, y con n = 30 en el 88,8 %. Los dos quedan por debajo del 95 % prometido. El intervalo t supone que los datos son normales; con ingresos tan asimétricos y muestras chicas, la media y el desvío muestrales están correlacionados y el intervalo falla más de lo que dice. Pasar de 10 a 30 casos casi no lo arregla: con una cola tan larga, la cobertura recién ronda el 93 % con muestras de 100, y se acerca al 95 % con muestras de cientos de personas, como la de la consigna a.

d) ¿Qué conclusión práctica sacás de la simulación?

RESOLUCIÓN

Que el nivel de confianza es una promesa que se cumple solo si se cumplen los supuestos. Con muestras chicas de una variable muy asimétrica, conviene transformarla (por ejemplo, con el logaritmo), usar la mediana o métodos como el bootstrap. Con muestras grandes el problema desaparece. La simulación sirve justamente para eso: comprobar si un método hace lo que dice en un caso concreto.

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.»

RESOLUCIÓN

Incorrecta. El intervalo es para la media, no para las madres. Las edades individuales están mucho más dispersas: hay madres primerizas de 17 y de 38 años. Un rango que contenga al 95 % de las personas es otra cosa (un intervalo de predicción o de referencia), y es mucho más ancho.

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

RESOLUCIÓN

Correcta: es la interpretación rigurosa. El 95 % es una propiedad del método, no de este intervalo en particular. La edad media verdadera es fija; lo que varía de muestra a muestra es el intervalo.

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

RESOLUCIÓN

Es la forma en que casi todos lo dicen, pero en sentido estricto no es correcta en la estadística clásica: una vez calculado, el intervalo contiene o no a la media, y no hay probabilidad en juego. Por eso se dice «confianza». La estadística bayesiana sí permite una frase así, con sus intervalos de credibilidad. En la práctica, «con un 95 % de confianza, está entre…» es la redacción segura.

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

> RESOLUCIÓN
>
> Incorrecta. Con cualquier tamaño, el intervalo del 95 % acierta el 95 % de las veces (si se cumplen los supuestos). Lo que cambia con una muestra más grande es el ancho: el intervalo es más angosto, o sea, más preciso, con la misma confianza.

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.»

> RESOLUCIÓN
>
> Incorrecta en las dos direcciones. Que se superpongan no prueba que las medias sean iguales: la diferencia puede ser significativa igual, porque el error de una diferencia no es la suma de los dos márgenes. Y aunque no se pudiera afirmar que son distintas, eso no prueba que sean iguales: puede faltar precisión. Para comparar dos medias hay que construir el intervalo de la diferencia, o hacer una prueba.

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

> RESOLUCIÓN
>
> Ninguno es mejor: hay un compromiso. El del 99 % se equivoca menos veces pero dice menos; el del 95 % es más preciso pero falla una de cada veinte veces. La elección depende del costo de equivocarse, y se hace antes de mirar los datos.

---

estadistica.ar/docentes/trabajos-practicos/semana-9 · Las lecturas de cada tema y la hoja de fórmulas están en el sitio.