

FORMACIÓN PROFESIONAL · INVESTIGACIÓN COMERCIAL · UNIDAD 8

Tamaño, precisión y no respuesta

Lunes 12 de abril de 2027. Quique ofrece a Nora «quinientas encuestas en una semana, con un error del ±4,4 % al 95 % de confianza». ¿Cuántas respuestas completas necesita el Estudio de kits, a cuántos hogares hay que llamar y qué garantiza de verdad un «±4,4 %»?

PARA EMPEZAR

En esta unidad aprenderás a...



Explicar cómo influyen la confianza, el error, la variabilidad, la población y el diseño en el tamaño de la muestra, y calcular el error de muestreo de un tamaño dado.



Aplicar la corrección para población finita y saber cuándo basta la fórmula de población grande.



Calcular el tamaño, redondearlo y distinguir respuestas completas de personas que hay que contactar.



Repartir la muestra entre estratos y comprobar las sumas y sus consecuencias.



Planificar los contactos con una tasa de respuesta prevista y explicar lo que el tamaño no corrige.

SITUACIÓN DE PARTIDA

«Quinientas encuestas en una semana»

Lunes 12 de abril de 2027. Quique, comercial de una empresa de encuestas en línea, enseña a Nora su oferta: «Quinientas encuestas en una semana, con un error del **±4,4 %** al 95 % de confianza». Nora está tentada: «Son muchas más de las que podemos hacer nosotros por teléfono».

Tu equipo de análisis tiene **hasta el viernes 23** para entregar el tamaño y el plan de contactos del Estudio de kits de **La Despensa de Nora** (caso ficticio): cuántas respuestas completas hacen falta para la precisión que Nora necesita, y con qué supuestos; a cuántos hogares hay que llamar para conseguirlas y cómo se reparten por zonas; y qué garantiza y qué no garantiza un «±4,4 %».

❓ **¿Qué debe decidir o comprobar Nora antes de comparar la oferta de Quique con un estudio propio?**

❓ **¿Qué sabemos de esas 500 encuestas y qué estamos suponiendo?**

❓ **Si responde el 8 % de los invitados, ¿cuántas respuestas dan 500 invitaciones? ¿Y 6.250?**

❓ **¿Qué prueba permitiría revisar la primera respuesta de Nora?**

MAPA DE LA UNIDAD

Lo que vamos a ver

8.1 Confianza y error

8.2 Variabilidad y población finita

8.3 Cálculo y redondeo

8.4 Asignación por estratos

8.5 No respuesta y limitaciones

¿Basta con llamar a 278?



Necesitamos 278 respuestas.

¿Basta con contactar a 278 personas?





No todas responderán; lo planificaré.

CLAVE

El tamaño n son **cuestionarios completos y válidos**. Como no todos contestan, hay que seleccionar y contactar a más: contactos = n ÷ tasa de respuesta prevista, al alza.

De qué depende cuántas respuestas



Error admitido (e)

La precisión, en tanto por uno: e = 0,05 es ±5 puntos porcentuales. Cuanto más pequeño, más muestra.



Nivel de confianza (z)

Proporción de veces que acertaría el procedimiento si se repitiera muchas veces. Se traduce en un coeficiente z de la normal.



Variabilidad (p · q)

Cuánto se espera que se repartan las respuestas: p es la proporción esperada y q = 1 − p.



Población (N) y diseño

N pesa poco salvo que la muestra sea una parte apreciable (8.2). Fórmulas del aleatorio simple sin reposición; un estratificado puede necesitar algo menos y un conglomerado, más.

CLAVE

Tabla 1 · Coeficiente z: **90 % → 1,645**; **95 % → 1,96**; **99 % → 2,576**. El número no sale de la costumbre ni del presupuesto.

La cuenta de Quique

$$e = z \cdot \sqrt{(p \cdot q / n)}$$

Oferta: n = 500, 95 %, p = q = 0,5	1,96 · √(0,25 / 500) = 1,96 · 0,02236	**0,0438 → ±4,4**
Con 200 respuestas	1,96 · √(0,25 / 200)	**±6,9 puntos**
Con 100 respuestas	1,96 · √(0,25 / 100)	**±9,8 puntos**
En la hoja de cálculo	=1,96*RAIZ(0,25/500)	**0,0438**

Para reducir el error a la mitad hay que multiplicar el tamaño por cuatro. El z sale de =DISTR.NORM.ESTAND.INV(0,975) para el 95 % (0,95 para el 90 % y 0,995 para el 99 %).

CLAVE

«±4,4 %» es un resultado condicionado: el cálculo es correcto **solo si** las 500 respuestas fueran un aleatorio simple de la población que interesa. Pregunta siempre cómo se seleccionó y cuántos contestaron.

8.1 · TABLA 2

Relación entre error, confianza y tamaño

Población grande, p = q = 0,5; respuestas completas, redondeadas al alza.

ERROR ADMITIDO (PUNTOS)	90 %	95 %	99 %
±3	752	1.068	1.844
±4	423	601	1.037
±5	271	385	664

El error de la fórmula mide solo la variación debida a observar una muestra. No incluye los sesgos de cobertura, selección, no respuesta y medición, que no bajan al aumentar la muestra (unidad 7).

CLAVE

Más confianza o menos error, **más muestra**: de 95 % a 99 % con ±5 puntos, de 385 a 664; de ±5 a ±4 puntos con el 95 %, de 385 a 601.

🙋 PREGUNTA PARA LA CLASE

Un estudio da su resultado «con ±5 puntos al 95 % de confianza». ¿Qué significa «al 95 %»?

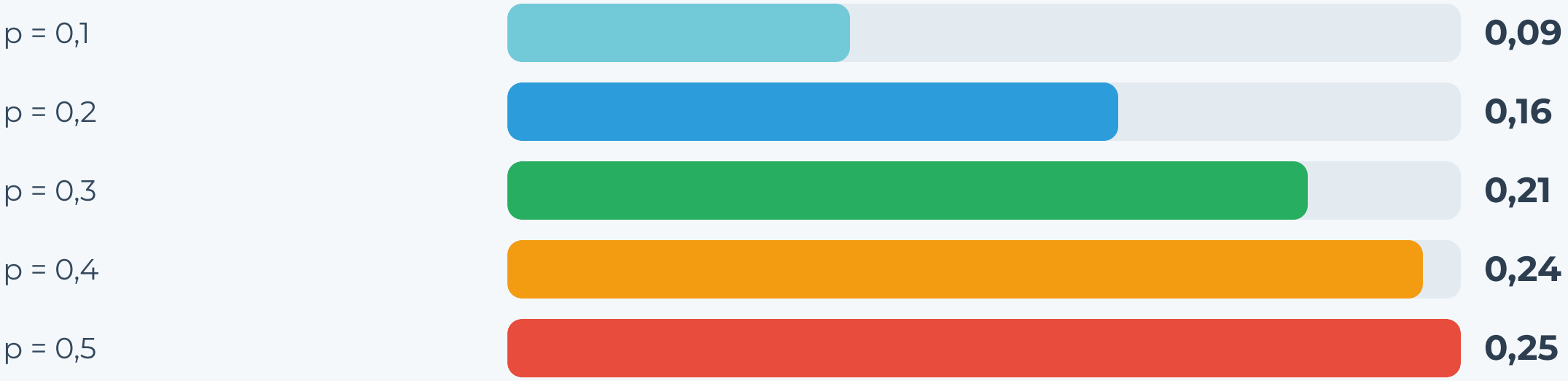
A Que el 95 % de los hogares piensa lo mismo

B Que, repitiendo el procedimiento con muchas muestras aleatorias del mismo marco, en unas 95 de cada 100 el resultado quedaría a menos de e del valor verdadero

C Que el resultado acierta con probabilidad 0,95 aunque la muestra sea de voluntarios

✔ Es una propiedad del procedimiento; por eso exige que la selección sea aleatoria (8.1).

p · q es máximo con p = 0,5



Con p = 0,2, la oferta de 500 respuestas tendría un error de 1,96 · √(0,16/500) = 0,035: baja porque hay menos variabilidad. Con N = 1.000 y p = 0,3 (dato de un estudio anterior), 245 respuestas en vez de 278.

CLAVE

Sin un dato fiable de un estudio anterior se toma **p = q = 0,5**, el supuesto más prudente: da el tamaño mayor.

El tamaño con N = 1.000

$$n = N \cdot z^2 \cdot p \cdot q \,/\, [e^2 \cdot (N - 1) + z^2 \cdot p \cdot q]$$

Numerador	1.000 · 3,8416 · 0,25	**960,4**
Denominador	0,0025 · 999 + 0,9604	**3,4579**
Tamaño	960,4 / 3,4579 = 277,74	**278**
Población grande	n_0 = 3,8416 · 0,25 / 0,0025 = 384,16	**385**
Relación entre las dos	$n = n_0 / [1 + (n_0 - 1)/N]$ = 384,16 / 1,38316	**277,74**

Sale de multiplicar p · q / n por el factor (N − n)/(N − 1). Con el listado de contactos del estudio (N = 1.000) y n = 278: 722/999 = 0,7227 y su raíz, 0,850: el error baja un 15 % frente al de N grande.

CLAVE

La corrección ahorra **385 − 278 = 107 respuestas**: con N = 1.000, 385 sería el 38,5 % de la población, y hay que corregir.

8.2 · TABLA 3

El tamaño según la población

95 %, ±5 %, p = q = 0,5, aleatorio simple sin reposición.

POBLACIÓN N	300	1.000	3.000	6.000	10.000	100.000	GRANDE
n exacto	168,70	277,74	340,65	361,10	369,98	382,69	384,16
n redondeado	169	278	341	362	370	383	385

Regla práctica: si n_0 es menos del 5 % de N, la corrección apenas cambia el resultado. Con N = 10.000, 385 es el 3,85 %: la fórmula de población grande da 385 y la finita, 370.

CLAVE

En tu sector. Poblaciones pequeñas: a veces, censo. Las 30 empresas del polígono: n = 30 · 0,9604 / (0,0025 · 29 + 0,9604) = 27,89 → **28**, casi todas. Se suele preguntar a todas y trabajar la no respuesta.

Supuestos, cálculo, redondeo y contactos

1

Supuestos

Población y marco (N), diseño, confianza (z), error (e) y p, por escrito.

→

2

Sustituir

Numerador y denominador por separado, con cuatro decimales.

→

3

Redondear al alza

Al entero superior, aunque el decimal sea pequeño: con una respuesta menos, el error supera el admitido.

→

4

Comprobar

El error con el tamaño redondeado no debe superar el admitido.

→

5

Contactos

n redondeado ÷ tasa de respuesta prevista, también al alza.

CLAVE

Ejemplo: N = 1.000, 95 %, **±4 %**: 960,4 / 2,5588 = 375,33 → **376**; al 70 %, 376 / 0,70 = 537,14 → **538 contactos** (537 · 0,70 = 375,9 no llega). Con 376, e = 0,0399.

8.3 · CUADRO 1

Cambios de parámetro con N = 1.000

SUPUESTO QUE CAMBIA	RESTO IGUAL (95 %, ±5 PUNTOS, P = 0,5)	RESPUESTAS COMPLETAS
Ninguno (caso base)	—	278
Confianza 99 %	z = 2,576	400
Confianza 90 %	z = 1,645	214
Error ±4 puntos	e = 0,04	376
p = 0,3 (de un estudio previo)	p · q = 0,21	245
Población grande (sin factor)	—	385

Errores frecuentes: el error en porcentaje (5 en lugar de 0,05); el z del 95 % para el 99 %; redondear al más próximo; calcular los contactos desde 277,74 o dividir por 1,70; contar como respuestas los hogares fuera del marco; comparar tamaños con supuestos distintos sin decirlo.

CLAVE

Para saber más. Efecto del diseño: con conglomerados de manzanas y un efecto de 1,5, las 278 respuestas pasarían a 278 · 1,5 ≈ 417 (ajuste conservador).

🙋 PREGUNTA PARA LA CLASE

Adrián redondea 375,33 a 375 «porque el decimal es pequeño». Con N = 1.000, 95 % y ±4 %, ¿qué pasa?

A Nada: 375 y 376 dan el mismo error

B Con 375 el error supera ligeramente 0,04; con 376 queda en 0,0399: se ha saltado el redondeo al entero superior

C Debería haber redondeado a 370

✔ Con una respuesta menos, el error supera el admitido. Se redondea siempre al alza y se comprueba el error con el tamaño redondeado (actividad 6).

8.4 · TABLA 4

Afijación proporcional del Estudio de kits

Marco de 1.000 contactos; n = 278; respuesta prevista 70 %. n_h = n · N_h / N.

ZONA	HOGARES N_H	PESO	CUOTA EXACTA	RESPUESTAS	CONTACTOS
Norte	420	42,0 %	116,76	117	168
Centro	350	35,0 %	97,30	97	139
Sur	230	23,0 %	63,94	64	92
Total	**1.000**	**100,0 %**	**278,00**	**278**	**399**

Partes enteras: 116 + 97 + 63 = 276; faltan 2, que van al Sur (resto 0,94) y al Norte (0,76). Contactos al alza en cada estrato: 167,14 → 168; 138,57 → 139; 91,43 → 92.

CLAVE

Suman **399**, uno más que los 398 del total, porque cada estrato redondea por su lado. Muestra casi autoponderada: cada respuesta representa a unos 3,6 hogares del marco.

Afijación igual y ponderación

Tasas hipotéticas de «sí» al kit: 50 % en el Norte, 40 % en el Centro y 30 % en el Sur.

ZONA	PESO EN EL MARCO	AFIJACIÓN IGUAL	HOGARES POR RESPUESTA	TASA HIPOTÉTICA
Norte	42 %	93	420/93 = 4,52	50 %
Centro	35 %	93	350/93 = 3,76	40 %
Sur	23 %	92	230/92 = 2,50	30 %

Media simple: (93 · 0,50 + 93 · 0,40 + 92 · 0,30) / 278 = 111,3 / 278 = 40,0 %, porque el Sur pesa en la muestra más que en el marco. Ponderada: 0,42 · 50 + 0,35 · 40 + 0,23 · 30 = **41,9 %**. El Sur gana precisión (con la corrección finita y N = 230): ±10,4 puntos con 64 respuestas; ±7,9 con 92.

CLAVE

Cada estrato tiene su propio error: el ±5 % es del total. La afijación igual conviene cuando comparar zonas importa tanto como el total; la óptima (de Neyman) da más a los estratos más variables o más baratos.

8.5 · TABLA 6

El embudo de respuesta

Simulación ficticia del equipo para preparar el plan.

ETAPA	HOGARES	SOBRE LA ETAPA ANTERIOR	PÉRDIDA
Seleccionados y contactados	399	—	—
Localizados (máximo 3 intentos)	336	84,2 %	63 no localizados
Aceptan la entrevista	284	84,5 %	52 rechazos
Cuestionarios completos	278	97,9 %	6 incompletos

Tasa de respuesta = completas ÷ seleccionados del marco: 278/399 = 69,7 %. Tasa de cooperación = aceptan ÷ localizados: 284/336 = 84,5 %. La no respuesta total es la del hogar que no contesta nada; la parcial, preguntas en blanco (unidad 9).

CLAVE

La cooperación separa el problema de **horario** (no se encuentra a nadie) del de **interés** (se encuentra y no quiere): piden remedios distintos.

8.5 · TABLA 5

Resultado posible del campo por zonas

ZONA	CONTACTOS	COMPLETAS	TASA DE RESPUESTA	PESO ENTRE LAS RESPUESTAS	PESO EN EL MARCO
Norte	168	101	60,1 %	36,3 %	42,0 %
Centro	139	104	74,8 %	37,4 %	35,0 %
Sur	92	73	79,3 %	26,3 %	23,0 %
Total	**399**	**278**	**69,7 %**	**100,0 %**	**100,0 %**

Se llega a 278 completas, pero no con el reparto previsto: al Norte le faltan 16 (101 de 117) y el Sur tiene 9 de más. Con la tasa del Norte, 117 / 0,601 = 194,7 → 195 contactos, no 168.

CLAVE

Si los hogares del Norte que no contestan son distintos, hay un **sesgo de no respuesta** que ningún aumento de llamadas corrige por sí solo.

Más llamadas no es menos sesgo

🔁

Reintentar

En horarios distintos antes de dar por perdido un hogar: es lo que más reduce la no respuesta.

➕

Ampliar contactos

En el estrato que falla, con la reserva sorteada: 16 / 0,601 = 26,6 → unos 27 contactos más en el Norte.

🔍

Comparar y ponderar

Zona, canal y antigüedad de quienes responden frente al marco; ponderar por zona corrige la composición por zona, no las diferencias dentro de cada zona.

📝

Informar

Tasa de respuesta, perfiles y limitaciones van en el informe.

CLAVE

Lo que no se hace: sustituir a quien no contesta por el vecino o por un voluntario, ni subir el tamaño «por si acaso» sin mirar quién falta.

Llamadas no deseadas



Los usuarios finales tienen derecho «a no recibir llamadas no deseadas con fines de comunicación comercial, salvo que exista consentimiento previo del propio usuario para recibir este tipo de comunicaciones comerciales o salvo que la comunicación pueda ampararse en otra base de legitimación de las previstas en el artículo 6.1 del Reglamento (UE) 2016/679».

Ley 11/2022, de 28 de junio, General de Telecomunicaciones, art. 66.1 b) · Consulta la versión vigente

SUJETO OBLIGADO

Quien llama: el listado del Estudio de kits son contactos que compraron y aceptaron llamadas para estudios.

CLAVE

Una llamada de estudio **no puede usarse para vender kits** (finalidad, RGPD, art. 5.1 b). El plan de contactos registra fecha, hora, intento y resultado de cada llamada.

Tamaño y contactos del Estudio de kits



El equipo de análisis, para Nora

Viernes 23/04/2027: listado de contactos autorizados 2027 (LC-v1.0), N = 1.000 hogares; respuesta prevista del 70 % (prueba de marzo: 21 de 30). Caso ficticio.

1. **Supuestos.** ±5 puntos al 95 % (z = 1,96); p = q = 0,5 sin dato previo; e = 0,05; aleatorio simple sin reposición. Alcance: los 1.000 contactos, no los unos 6.000 hogares del barrio.

2. **Cálculo.** Numerador 1.000 · 3,8416 · 0,25 = 960,4; denominador 2,4975 + 0,9604 = 3,4579; n = 277,74 → 278 respuestas completas.

3. **Contactos.** 278 / 0,70 = 397,14 → 398 (397 · 0,70 = 277,9 no llega; 398 · 0,70 = 278,6 sí).

4. **Comprobación.** Con 278, e = 0,0500; con 277, 0,0501. Ahorra 107 respuestas frente a 385. Zonas: 117, 97 y 64. El ±5 % no corrige la cobertura ni la no respuesta; revisión de tasas tras 100 llamadas.

✓ **278 respuestas completas** y **al menos 398 contactos**, con registro de cada contacto y comparación de perfiles con el marco: más contactos no eliminan el sesgo de no respuesta.

APRENDE CON LA IA

«Basta entrevistar 278 voluntarios»

Unai pregunta a un asistente de IA si pueden ahorrarse las llamadas publicando la encuesta en las redes. Respuesta: «Basta entrevistar 278 voluntarios para garantizar un error del 5 %». Pulsa cada tarjeta.

«278 respuestas dan ±5 puntos con N = 1.000»

REALIDAD

Solo con sus supuestos: aleatorio simple sin reposición del listado, 95 %, p = q = 0,5 y que todos los seleccionados respondan.

«Con 278 voluntarios de las redes el error es del 5 %»

BULO

La fórmula supone un sorteo; con voluntarios no hay probabilidad de inclusión conocida ni margen de error.

«Publicarla en las redes ahorra las 398 llamadas sin perder nada»

BULO

Responde quien quiere: sesgo de autoselección (unidad 7), que no baja con más respuestas.

«Con 278 voluntarios se puede afirmar algo»

REALIDAD

Lo que dicen quienes respondieron, como recuentos y sin margen de error, nombrando a quién deja fuera.

Tamaño y contactos

Para la investigación de tu negocio, con el marco de tu paso 7:

1. **Supuestos**
   N, diseño, confianza, error y p, con su justificación.
2. **Cálculo**
   En hoja de cálculo, con numerador, denominador y redondeo al alza; error comprobado con el tamaño redondeado.
3. **Estratos y contactos**
   Reparto por estratos con la suma comprobada; contactos con la tasa de respuesta prevista y su fuente; registro de contactos preparado.
4. **Limitaciones**
   Cobertura y no respuesta, responsable y fecha de revisión.

CLAVE

Comprueba tu paso: **cualquiera puede repetir tu cálculo** con tu archivo; ningún contacto de más se presenta como corrección del sesgo.

REPASO

Ideas clave

- El tamaño depende del error admitido, la confianza (z), la variabilidad (p · q), la población (N) y el diseño; no de la costumbre ni del presupuesto.
- Error dado n: $e = z \cdot \sqrt{(p \cdot q / n)}$. La oferta de Quique: $1{,}96 \cdot \sqrt{(0{,}25/500)} = \pm 4{,}4$, solo si fuera un aleatorio simple.
- «Al 95 %» es una propiedad del procedimiento: en unas 95 de cada 100 muestras aleatorias, el resultado quedaría a menos de e.
- p = q = 0,5 da el tamaño mayor (p · q = 0,25): es el supuesto prudente sin dato previo.
- Población finita: $n = N \cdot z^2 \cdot p \cdot q / [e^2 \cdot (N - 1) + z^2 \cdot p \cdot q]$. Con N = 1.000: 277,74 → 278 (385 sin corrección).
- Siempre al entero superior y comprobando el error; contactos = n ÷ tasa prevista, al alza: 278 / 0,70 → 398.
- Afijación proporcional por mayores restos: 117, 97 y 64; contactos por estrato, 399. La afijación igual obliga a ponderar.
- La tasa de respuesta (completas ÷ seleccionados) y la de cooperación (aceptan ÷ localizados) se calculan por zona.
- Más contactos aseguran el número, no que quienes responden se parezcan a quienes no: reintentar, comparar perfiles, ponderar e informar.



¡Ahora te toca a ti!

Repasa, practica y comprueba lo que has aprendido.

📝 **Resumen**

✅ **Test**

👆 **Interactivos**

Libro, presentaciones, resúmenes y tests gratis en **josesande.com**