

Nociones básicas para el cálculo del tamaño de la muestra

Análisis de Datos en Psicología I
Vicente Manzano

Decidir cuál es el mejor tamaño para una muestra es una de las preocupaciones principales relativas al muestreo. El primer principio es que no existe un tamaño bueno para todo. Según el tipo de muestreo que se vaya a realizar, los objetivos que se persigan, las características de la población y las condiciones en las que se van a realizar las estimaciones, serán aconsejables unos tamaños u otros. Podría parecer que una muestra es mejor cuanto más grande. Pues sí, podría parecerlo, pero no es cierto. Cuanto más grande las estimaciones serán más precisas y con menos riesgo de error. Pero también saldrán más caras y tal vez se reduzca el control en la recogida de datos, por lo que, repito, no existe un tamaño bueno para todo.

Este documento tampoco es bueno para toda decisión de tamaños de muestra. Se ha pensado para unas situaciones y no para cualesquiera. Pero no nos preocupemos en exceso. Mi experiencia es que buena parte de las investigaciones científicas publicadas en revistas de prestigio hacen muestreos que dejan bastante que desear. Intentaremos no ser peores sino mejorar un poco en panorama.

Principios generales

1. El tamaño de la muestra suele simbolizarse con la letra minúscula n .
2. n depende del tipo o modelo de muestro. Cada uno requiere sus propias expresiones de cálculo. Aquí partimos del muestreo aleatorio simple.
3. n depende del objetivo de análisis. Aquí vamos a suponer que estimamos una media aritmética o una proporción.
4. En una investigación se busca responder a varios objetivos. Cada objetivo requiere su propio cálculo de tamaño de muestra. Pero la muestra es única para todos. ¿Qué hacer? Hablaremos de ello más adelante (ver apartado "Tamaño y objetivos").
5. Aunque cada objetivo y modelo de muestreo cuentan con sus propias expresiones de cálculo, siempre existen cuatro variables básicas: tamaño de la población, variación de la población, precisión de la estimación y riesgo de error. En cuatro apartados del mismo nombre vamos a tratar estos conceptos.
6. n es una apuesta. Para calcularlo hemos de suponer cosas que no sabemos con exactitud. Podemos haber supuesto bien o menos bien o, incluso, mal. Por eso solemos ponernos en la peor de las situaciones, como verás en el ejemplo sobre la influencia de la variación de la población en el tamaño de la muestra.

Tamaño de la población

Parece lógico suponer que conforme aumenta el tamaño de la población, también debe hacerlo el de la muestra. En cierta medida es así. Pero la relación entre ambos no es lineal, sino logarítmica. Vale, es un palabrejo. Quiero decir que conforme la población es más grande también debe serlo la muestra, pero menos. El objetivo de la muestra es permitirnos realizar estimaciones sobre valores de la población. Es necesario partir de un tamaño mínimo. Estaremos de acuerdo en que una muestra de $n = 1$ no sirve para generalizar nada cuando medimos variables que tienen alguna variación en la población. Si lo que queremos medir es si la gente se muere al quitarle la cabeza, no hace falta obtener ninguna muestra, pues ya sabemos que la cabeza es fundamental para mantener la vida (aunque la utilicemos sólo un poco). Pero para averiguar qué piensan las personas sobre el precio de la remolacha o en qué medida están de acuerdo con la gestión de un personaje político, ya sabemos que las opiniones difieren y que no hay que jugársela con una muestra de $n = 1$. ¿Y $n = 2$? No hemos mejorado mucho.

Lo siento, no existe un tamaño mínimo. Pero sí un tamaño máximo.

Conforme aumenta el número de elementos de la población (personas, hogares, fábricas o lo que sea que estemos investigando), es lógico que deba ser mayor la muestra, pero cada nuevo elemento en la población añade menos cuanto ésta es más grande. Esto también es lógico. De 1 a 2 hemos duplicado el número, pero de 107 a 108 no hay mucha diferencia, mientras que 9999999995 ó 9999999996 son la misma cosa, y no admito discusiones tontas al respecto.

Por ejemplo, vamos a realizar una estimación por intervalo de una proporción. Suponemos que la varianza poblacional es 0,25. Utilizamos un error de precisión de 0,03. Y el riesgo de error es 0,05. Vamos a ir considerando tamaños de población como potencias de 10. Verás cómo la muestra no se multiplica del mismo modo:

N	10	100	1000	10000	100000	1000000	10000000	100000000	1000000000
n	10	92	517	965	1056	1066	1067	1068	1068

Esta circunstancia es trascendente. Imagina que queremos realizar estimaciones en una región con diez provincias. Cada provincia cuenta con 10 mil habitantes. Para hacer estimaciones con el error de precisión y el riesgo de error considerados, necesitamos una muestra de 1056 personas para toda la región. Pero si queremos hacer estimaciones por provincia, necesitamos 965 para cada una; es decir, 9650 (¡9 veces más!) en total.

Variación en la población

También es lógico pensar que conforme varíe más lo que queremos conocer, será necesario indagar en más unidades. Si la gente opina de forma muy similar con respecto a algo, bastará con preguntar a unas pocas. Si queremos hacer un estudio sobre el gasto de los hogares andaluces en alimentación, tal vez necesitemos una muestra grande porque sospechamos que hay muchas formas diferentes de organizar la economía familiar y muchas peculiaridades.

Vamos a ver esta influencia con otra tabla, esta vez partimos de un tamaño de población constante (diez mil unidades) para estimar una proporción con un error de precisión de 0,03; riesgo de equivocación de 0,05; y una varianza que vamos a considerar entre 0,10 y 0,25; utilizando intervalos de 0,03.

σ^2	0,10	0,13	0,16	0,19	0,22	0,25
n	410	526	640	751	859	965

He dicho que este documento se escribe suponiendo que vamos a realizar estimaciones de medias aritméticas o de proporciones. No he dicho nada de varianzas o desviaciones tipo. Se supone también que si no conocemos la media o la proporción poblacional ¿cómo vamos a conocer su desviación? ¿Qué hacemos entonces?

Cuando ocurre lo habitual, es decir, cuando no sabemos qué valor tiene la variación poblacional, adoptamos lo que se llama una "postura conservadora". Ya hemos visto que conforme es mayor la variación, también es mayor el tamaño de la muestra. Imagina que infravaloramos la variación, que es 0,22 en lugar de 0,13 como pensamos. La muestra va a ser más pequeña de lo que debería ser (calcularemos 526 en lugar de 859 que es lo que deberíamos haber hecho). Después, a la hora de hacer las estimaciones a partir de los datos de la muestra que hemos obtenido, nos tropezaremos con las consecuencias del error: por culpa de un tamaño de muestra demasiado pequeño, nuestras estimaciones serán más erróneas o más imprecisas o ambas cosas a un mismo tiempo.

Para evitar ese efecto indeseable, nos ponemos en la peor de las situaciones: la de la máxima varianza que pueda ocurrir desde la información con que contamos. A esto se le llama "postura conservadora". Si nos equivocamos será por exceso. Habremos implicado más dinero o más tiempo de lo necesario, pero conseguiremos estimaciones tanto o más precisas como lo deseado y tanto o más seguras de lo diseñado. Mejor eso que lo contrario.

¿Qué valor tiene la varianza entonces, desde una perspectiva conservadora? Hemos de pensar en una situación límite: la máxima varianza que se pueda obtener. En el caso de las proporciones es fácil. La varianza de proporciones es $p(1-p)$. Haz todas las pruebas que quieras, la varianza máxima ocurre cuando $p=0,5$, lo que lleva a $S^2=0,5(1-0,5)=0,25$. Así que si no sabemos qué valor tiene la varianza poblacional cuando estamos estimando proporciones, vale con la postura conservadora de que $S^2=0,25$.

En las medias aritméticas es más complicado. Es muy difícil que el valor de la desviación tipo llegue a igualar el valor de la media aritmética. Como la varianza es el cuadrado de la desviación tipo, lo que podemos hacer es imaginar un intervalo de valores esperables para la media de la población. Tomamos el límite superior del intervalo y lo elevamos al cuadrado. He ahí una estimación conservadora para la varianza poblacional. Por ejemplo, vamos a estimar el número de peines que compra una persona en una población, cada año. No sé nada sobre ello, pero dudaría que la media sea superior a 3. Entonces, podemos utilizar $3^2=9$ como estimación conservadora de la varianza.

Error de precisión

Llamamos error de precisión al radio del intervalo de estimación. Podemos pensar en ello con dos formas diferentes para referirse a lo mismo:

- La cantidad que se resta y se suma al estadístico o estimador para construir el intervalo de confianza en cuyo interior esperamos que se encuentre el parámetro.
- La máxima diferencia que cabe esperar que exista entre el estadístico y el parámetro.

Luego, a mayor error de precisión, menos precisión (más imprecisión). El ideal es buscar la máxima precisión posible (el mínimo valor para el error de precisión). Pero es lógico pensar que conforme hacemos el intervalo más pequeño, haya que exigir un tamaño de muestra más grande o correr más riesgo de equivocarse al suponer que el parámetro se encuentra dentro del intervalo. Como el riesgo es algo que preferimos no tocar, está claro que la aspiración de un error de precisión mínimo debe buscarse mediante un tamaño de muestra máximo, en las circunstancias concretas donde se defina tal cosa.

Conforme aumenta el tamaño de la muestra se obtienen valores de los estadísticos que deben parecerse más al valor del parámetro, si todo lo demás sigue igual. La tabla siguiente muestra esta relación, partiendo de una estimación por intervalo de una proporción, mediante un riesgo de error de 0,05 y una población de tamaño 10 mil y varianza de 0,25.

e_p	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09	0,10
n	4900	1937	965	567	370	260	193	148	118	96

Observa con qué velocidad se reduce el tamaño de la muestra al aumentar el error de precisión. Sin duda es una variable con gran sensibilidad. Entre otras razones, por esto mismo es la que suele utilizarse para variar las consecuencias de un tamaño de muestra diferente. Imagina que aunque deseamos un error de precisión de valor 0,06, no tenemos dinero para 260 entrevistas. De hecho, sólo podríamos costear 200. Pues no es muy grave si aceptamos igualmente un error de precisión de valor 0,07.

¿Cómo decidir un valor concreto para e_p ? En principio hay que considerar la escala en la que se mueve el parámetro. Si vamos a estimar una proporción y ésta creemos que ronda el 50%, podríamos considerar un buen intervalo, por ejemplo, de 45% a 55% (un error de precisión de 5%). Pero si la proporción ronda un 3%, no tiene sentido sumarle y restarle 5%. Algo admisible en este caso podría ser un error de precisión de valor 1%, que construye un intervalo del 2% al 4%, entre otras posibilidades.

Como eso de plantearse “la sensación positiva” de la amplitud de un intervalo es algo que roza lo asombroso para muchas mentes y muchos corazones, podemos pensar en estrategias aparentemente menos sentimentales. Una de ellas es considerar un error de precisión que no vaya más allá de un 10% del valor supuesto para el parámetro. Para una proporción del 50%, un error de precisión del 5%. Para una proporción del 3%, un error de precisión de 0,3%. Si estimamos una media aritmética, pongamos que el número medio de litros de bebida alcohólica que se consumen en una familia española durante las vacaciones de Navidad, y creemos que ese valor debe encontrarse entre los 15 y 25 litros, utilizamos el centro (20 litros) como referencia y un error de precisión de la décima parte, es decir, de 2 litros.

La sugerencia del párrafo anterior sólo tiene sentido en situaciones de pérdida total de referentes, que es lo que suele ocurrir, por ejemplo, cuando hay que hacer un trabajo para una asignatura. En una situación práctica se utilizan los referentes habituales. Así, por ejemplo, si vamos a estimar el número de escaños que va a ocupar un partido en el Congreso de los Diputados tras unas elecciones generales, lo habitual es procurar un error de precisión no superior a 2 escaños.

Riesgo de error

Como ya sabemos, lo que llamamos familiarmente con “riesgo de error” es la expresión abreviada de otra más larga: “riesgo de equivocarse al afirmar que el parámetro se encuentra en el intervalo de confianza que se propone”. Es un valor de probabilidad que, como todos los valores de probabilidad, debe encontrarse entre 0 y 1.

Como ocurre con el error de precisión, es el equipo investigador quien decide un valor para el riesgo de error. ¿Cómo escogerlo? Sabemos que a mayor riesgo, menor tamaño de muestra, puesto que somos menos exigentes con la situación. Lo ideal es no equivocarse, lo que es imposible de garantizar. Así que aspiramos al mínimo error en la práctica, lo que aconseja el máximo tamaño de muestra. Observa la relación en la siguiente tabla, para estimación de una proporción con una población de 10 mil unidades, una varianza poblacional de 0,25 y un error de precisión de 0,03.

riesgo	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09	0,10
n	1557	1307	1157	1049	965	895	836	785	740	700

Ya ves que su efecto no es tan drástico como el del error de precisión. Pero, volvamos a la pregunta ¿Cómo escoger un valor para el riesgo de error?

Lo suyo sería pensar en cuáles son las consecuencias que se derivan de equivocarse. Cuanto más graves o indeseables sean, menor tendrá que ser el riesgo de llegar a ello. Imagina cuál es el valor de riesgo que considerarías en las siguientes ocasiones: 1) si te equivocas mueres pues hablamos del riesgo de tener un accidente de tráfico, 2) si te equivocas no tendrás el yogurt gratis, pues hablamos de tomar una decisión sobre qué calle tomar para ir a la universidad sabiendo que en las inmediaciones están repartiendo yogurt.

En muchas ocasiones (como trabajos de asignaturas, como siempre), es difícil ponerse en situación y decidir un valor para el riesgo que sea adecuado a la conciencia de las consecuencias. En estos casos acudimos a Fisher. Como sabes, Fisher fue un agrónomo que le complicó la vida a

todos los estudiantes de estadística desde que irrumpió en escena en el primer tercio del siglo XX. Una de las ocurrencias que tuvo fue pensar cómo saber si una viejecita decía la verdad o no. Aquella viejecita afirmaba que era capaz de distinguir si en un vaso de té con leche se había echado antes el té o la leche. Fisher pensó que podría acertar por casualidad, pero que demasiados aciertos serían indicador de verdad en la afirmación. Finalmente decidió cuántas tazas debía poner ante la señora y le salió, en el cálculo, que con una probabilidad de equivocarse del 5% podría rechazar la hipótesis de que la viejecita había acertado por azar. Desde entonces, nadie ha abierto la boca para decir otra cosa. Se considera que un riesgo de 0,05 es válido por defecto. Si alguien osa tomar otra decisión se le preguntará por qué. Como no queremos dar explicaciones, todo el mundo coge 0,05. Se admiten dos excepciones más: 0,01 para mostrar que uno es muy exigente y ** , es decir, el número de asteriscos que suministra SPSS, un programa de ordenador que se está utilizando como referente desde hace tiempo. Sabemos que muchas personas deciden dejar de pensar cuando se les presenta la oportunidad para ello. SPSS, un programa tonto de ordenador (como todos los programas de ordenador, por definición, desprovistos de inteligencia), suministra esa oportunidad. Realmente no es cierto, pero mucha gente ha llegado a creérselo y si afirmas “me lo dijo SPSS”, suelen dejarte en paz.

Tamaño y objetivos

Los estudios tienen varios objetivos de análisis. En principio, habría que calcular un tamaño de muestra para cada objetivo. Imagina este caso: tenemos dos proporciones que estimar. Una consideramos que ronda el valor 40% (es decir, una varianza de $[0,4 \cdot (1-0,4)] = 0,24$; y un error de precisión de $[0,4/10] = 0,04$). En una población de tamaño prácticamente infinito y un riesgo de 0,05, el tamaño de muestra es $n = 577$. De la otra proporción suponemos que tiene un valor en torno al 10% (es decir, una varianza de $[0,1 \cdot (1-0,1)] = 0,09$; y un error de precisión de $[0,1/10] = 0,01$). En una población de tamaño prácticamente infinito y un riesgo de 0,05, el tamaño de muestra es $n = 3458$. ¿Qué hacemos?

Una postura conservadora aconseja tomar el tamaño más grande de entre todos los calculados. En este caso es 3458. No obstante, puede ser muy exagerado. En tales situaciones, podemos prescindir de los tamaños extremos, sabiendo que las estimaciones que realizaremos después para esos objetivos tendrán menos precisión o seguridad.

Como es temporalmente costoso calcular un tamaño de muestra para cada objetivo, es raro que se lleve a la práctica. Hay dos soluciones. Una es considerar el objetivo estrella o más importante. La otra es considerar que no tenemos ni idea de nada y adoptar una postura conservadora que llevará hacia un tamaño máximo de muestra.

Cálculo de n con MAS

Existen varias utilidades que permiten calcular el tamaño de una muestra. Hay de pago y gratis. Las hay de Internet y de puesto local. Las hay en Windows y en otros sistemas. Las hay más claras y más oscuras, con más o menos situaciones de consideración, etc.

Una de las utilidades de libre distribución, pensadas para cubrir todas las situaciones expuestas aquí, es MAS. Se puede bajar de

<http://www.personal.us.es/vmanzano/distribuye/programas.htm>

La siguiente figura muestra una pantalla típica para MAS. Como observarás, se encuentran todas las variables consideradas: varianza de la población, riesgo de error, error de precisión, tamaño de la población y tamaño de la muestra.

Hay dos aspectos interesantes para resaltar. El primero es que MAS sirve también para poner cualquiera de las variables en función de las otras. Observa el recuadro de la derecha de la

ventana. Al marcar sobre la variable, ésta se transforma en dependiente, su recuadro queda marcado de amarillo y recibirá los resultados de cálculo que se establezcan sobre todas las demás.

La segunda característica de interés a resaltar es que MAS genera un informe donde consta la situación de partida, los valores de las variables independientes, la fórmula de cálculo y el resultado. Se accede a esta generación pulsando el botón “Leer / Escribir” y escogiendo la opción “Generar”.

En el ejemplo que se muestra en la pantalla, el resultado de la generación del informe es el que sigue:

MAS: Información de contexto y resultados

* Contexto:

Muestreo Aleatorio Simple desde una población de tamaño infinito.
 Estimación de una media o proporción.
 Supuesta una distribución muestral normal.
 Tiempo de interacción: 01:43:14.

* Variables:

Variable dependiente: Tamaño de la muestra (n)

Variables independientes:

Varianza poblacional (V)	:	0,09
Distancia estandarizada (Z)	:	1,96
Probabilidad de error (p)	:	0,05
Error de precisión (E)	:	0,01

* Expresión de cálculo:

$$n = V \cdot Z^2 / E^2$$

* Resultado: 3458