

El proceso para el diseño de muestras en el entorno de las bibliotecas universitarias

José Luis Sandoval Dávila

Introducción

En la administración de las bibliotecas universitarias es recurrente la necesidad de disponer de información que formalmente evidencie la efectividad de su gestión, procesos y operaciones, así como la detección de áreas de oportunidad y análisis de sus causas. Un factor determinante al respecto es el momento en que debe tenerse la información, ya que a partir de ella se sustentará la toma de decisiones. Es insoslayable considerar que cada biblioteca es diferente de todas aquellas que conforman su universo; sin embargo, la normatividad, planeación y decisiones estratégicas se realizan en forma global, pero debe corresponder a cada entidad su administración operativa. Al determinar el objeto de estudio por entidad, cada una de las bibliotecas universitarias puede ser considerada como un conjunto independiente a observar, o todas ellas formar un universo de análisis.

Sea cual fuere el contexto o delimitación del universo o población en estudio, es menester considerar que en la búsqueda de la información que permita tomar las mejores decisiones, la administración correspondiente deberá partir del análisis de la situación y condiciones actuales de operación, lo que le permitirá obtener un diagnóstico que evidenciará las áreas de oportunidad y sus alternativas de intervención. Para ello y ante la gran diversidad de elementos de incidencia como colecciones, servicios, usuarios y variables involucradas, es necesario minimizar los recursos sin lesionar la calidad de la información que habrá de obtenerse. Esto se simplificaría si se hace uso de las herramientas que proporcionan los métodos cuantitativos, tales como el diseño de muestras y la explicación del compor-

tamiento entre las variables por medio de la realización de pruebas estadísticas diversas, entre otras.

Este capítulo presenta al estudiante y profesional del ámbito bibliotecológico las etapas, consideraciones y criterios para el diseño óptimo de muestras, que le permitirán conocer la realidad de su contexto, orientándolo hacia la explicación del comportamiento de las variables incidentes en su entorno operacional, a partir de la realización de las pruebas estadísticas básicas que se recomiendan en la última parte de esta contribución. Para esto deberá considerarse el tipo de población, universo de estudio, unidades de muestreo, objetivo del estudio o investigación y costo, factor este último que además implica tiempo, recursos materiales, humanos y financieros, así como el impacto social de la decisión que habrá de tomarse y que en muchos de los casos es el que determina las acciones a seguir. Sin embargo, al considerar que el profesional medio de la bibliotecología no está plenamente identificado o familiarizado con los métodos estadísticos, se debe procurar utilizar un lenguaje sencillo y simple que, sin deterioro de la calidad del contenido, permita la fácil comprensión de conceptos y términos.

La naturaleza de las poblaciones (usuarios, colecciones, servicios, personal) puede simplificar el proceso, debido a que es posible observar su forma real a partir de tomar como referencia sólo una parte de ella, que se concretará en una muestra poblacional. Pero, ¿Qué confianza puede tenerse al observar sólo a una parte de la población y a partir de ella emitir afirmaciones en relación con las características generales, así como los diversos aspectos y condiciones en que se encuentra el universo en estudio? ¿Cuántos elementos deben ser observados para ese propósito, de tal manera que posean la credibilidad necesaria para tomar las medidas pertinentes al respecto?

La respuesta a esas preguntas es simple, ya que para estar seguro de las condiciones en que se encuentra una población, sólo bastará con observar detalladamente una parte de ella para llegar a la conclusión de que todos los elementos que la conforman tienen características y actitudes comunes, entre otras, que permiten su generalización a un menor costo; por supuesto que la observación detallada referida implica un proceso metodológico a seguir, lo que garantiza la efectividad de los resultados de la investigación.

Por tal virtud, se requiere que el interesado posea un amplio dominio de la metodología para el diseño de muestras, donde cada una de las etapas que la conforman resulte ser la más importante al

momento de su aplicación. Soslayar este requisito ha contribuido a que los responsables de la administración bibliotecaria recurran a diversas instancias para tal propósito, las cuales comúnmente carecen del grado de participación requerido en el entorno bibliotecario para lograr la efectividad del estudio.

Como se ha señalado, cada una de las etapas del diseño de muestras es la más sustancial al momento en que se utiliza; sin embargo, la muestra no lo es todo en el estudio, pero sí es fundamental, por lo que no hay que perder de vista que una vez recogida la muestra se tiene que seguir un conjunto de métodos, técnicas y procedimientos para la obtención de la información requerida por la investigación. Surgirá una variedad de supuestos (Cochran, 1996) que tratarán de explicar el porqué del comportamiento de las variables, analizar la relación entre ellas, conocer sus rasgos importantes, su magnitud, así como diversos aspectos de interés, tanto para la investigación como para el mismo investigador.

Deberá considerarse que después de la etapa del diseño de la muestra, los datos tendrán que ser analizados a través de su descripción gráfica y numérica, así como mediante la utilización de pruebas para la comprobación de hipótesis, como mediciones estadísticas y pruebas de correlación entre variables, explicadas ampliamente en las referencias que se presentan al final de este capítulo.

Conceptos básicos para el muestreo

El objetivo de este apartado es conocer de manera sencilla conceptos esenciales relacionados con el diseño de muestras:

- **Población:** conjunto de elementos bien determinados y con características comunes, que los hace formar parte del mismo. Ejemplo: estudiantes del turno matutino que asisten a la biblioteca por las tardes. Si algún elemento del conjunto no contiene todas las características referidas, entonces no formará parte de él. Ejemplo: si algún estudiante asiste a la biblioteca por las tardes pero no es del turno matutino, no formará parte de ese conjunto.
- **Población finita:** es aquel conjunto en el cual su número de elementos está definido o determinado. Ejemplo: los empleados

de la biblioteca que recibieron estímulos económicos por su rendimiento en el trabajo.

- *Población infinita*: es aquel conjunto en el que el número de sus elementos es indefinido o indeterminado. Ejemplo: el número de préstamos interbibliotecarios que se realizarán los próximos diez años.
- *Muestra*: fracción o parte de la población. Ésta debe ser representativa del conjunto de donde ha sido extraída. Para ostentar el carácter de representatividad, la muestra deberá evidenciar tanto las coincidencias como las discrepancias existentes en la población.
- *Parámetro*: es todo valor o característica que ha sido calculado, observado o extraído de la población.
- *Estimador*: también llamado factor estadístico, es todo valor o característica que ha sido calculado, observado o extraído de una muestra. El propósito del estimador es dar a conocer los parámetros sin tener que observar la totalidad de elementos que conforman una población.

Etapas para el diseño de muestras

La experiencia indica que una proporción significativa de los estudiantes y profesionales del ámbito bibliotecario manifiesta cierta aversión por todo aquello relacionado con números, ya no se diga con matemáticas, como es el caso de la estadística y su amplia gama de recursos; sin embargo, resulta paradójico saber que para aplicar la estadística en ese contexto, se requiere de ingredientes tan naturales en el ser humano como el sentido común y la intuición, así como no perder de vista que las técnicas y los métodos son importantes, pero no determinantes, aunque resulte difícil de creer; he aquí la explicación.

Cuando se posee el atributo que confiere a una persona la habilidad para analizar; la dificultad para el entendimiento, comprensión y aplicación de la metodología estadística se reduce notablemente y da la oportunidad para una mayor dedicación al análisis de variables, a la vez que disminuye el tiempo de operación, lo que garantizará información más oportuna y mejores decisiones.

Por tal virtud, debe enfatizarse la importancia que el proceso de diseño de muestras tiene en el desarrollo de estudios e investiga-

ciones, consistente en tres etapas que ofrecen respuesta a las preguntas ¿cómo, quién, cuántos? Es oportuno subrayar que este proceso es más efectivo y garantiza mejores resultados cuando es aplicable a poblaciones finitas.

Métodos de muestreo

Consiste en determinar cómo serán seleccionados los elementos que habrán de integrar la muestra y se clasifica en muestreo con reemplazo y sin reemplazo (Busha, 1990). El primero indica que los elementos de la población pueden pertenecer más de una vez a la muestra; es decir, que si un individuo ya fue seleccionado se le extraen los datos requeridos por la investigación y se integra nuevamente a su universo, sin perder la oportunidad de volver a ser seleccionado para su estudio en la misma muestra. Por su parte, el muestreo sin reemplazo indica que si un elemento de la población ha sido extraído para la muestra, no tendrá la oportunidad de pertenecer nuevamente a ella.

¿En qué casos o cómo se determina el método de muestreo a seguir en un proceso de recolección de datos para un estudio? Como ejemplo puede ser un estudio donde se pretende medir la satisfacción de los usuarios que pertenecen a la red de bibliotecas universitarias. Para este caso, una respuesta sencilla sería que las características de ese estudio requieren de encuestar a un usuario en más de una ocasión y analizar las mismas variables, sólo que en tiempos diferentes, pues la opinión del usuario puede variar.

Tipos de muestreo

La etapa para determinar el tipo de muestreo a utilizar en un estudio como el referido, reviste singular importancia, debido a que en ella se determina quiénes serán los usuarios que formarán parte de la muestra, de tal manera que esta última conserve las características de su población, según lo comentado en párrafos anteriores.

El tipo de muestreo se clasifica en probabilístico y no probabilístico. El primero de ellos es aquel en el que todos los elementos de la población tienen una probabilidad conocida para formar parte de la muestra (Sandoval, 2004a). En el segundo se desconoce o es in-

determinable esa probabilidad, restándole ciertos efectos de representatividad.

El muestreo probabilístico ofrece seguridad y control en la forma de la muestra (Sandoval, 2004a), lo que da mayor garantía a las pruebas estadísticas que se realizan para explicar el comportamiento de las variables o sus niveles de relación entre sí.

En nuestro caso, para realizar la extracción de una muestra probabilística es indispensable: conocer el total de usuarios (población finita), que cada uno de ellos se encuentre identificado (por número de credencial, número de usuario, número de boleta) y localizable, así como poseer una herramienta que permita realizar la selección de la manera más objetiva posible, sin que intervengan emociones, sentimientos, gustos y otros aspectos de carácter personal en el investigador, para que sea otorgada la absoluta responsabilidad de selección a la metodología empleada; esta herramienta toma la forma de tabla de números aleatorios, calculadora electrónica o programa de cómputo.

Básicamente, el muestreo probabilístico se clasifica de la siguiente forma:

- Muestreo aleatorio o al azar
- Muestreo sistemático
- Muestreo estratificado

La utilización y seguimiento de ellos en un proceso de investigación, dependerá de las características poblacionales, así como del objetivo y alcances del estudio.

Muestreo aleatorio o al azar

En esta modalidad, todos los elementos de la población tienen la misma probabilidad de ser seleccionados para la muestra (Raj, 1979). Para la selección se sugiere observar que la población sea finita y que cada uno de los elementos que la conforman estén identificados, habitualmente por un número. Finalmente, se requiere disponer de algún instrumento que facilite la selección, como puede ser una calculadora científica, una computadora o una tabla de números aleatorios.

Muestreo sistemático

En este tipo de muestreo, la oportunidad de que los elementos de la población tengan la misma probabilidad de ser seleccionados resulta relativa, debido a que una vez seleccionado el primer elemento para la muestra, se podría afirmar que automáticamente se conoce cuáles son los elementos que la constituirán.

Este tipo de muestreo consiste en dividir el tamaño de la población entre el tamaño de la muestra, lo que significa que la población es dividida en tantas partes como el tamaño de la muestra sea y de cada una de esas partes se extraerá un elemento que habrá de integrar a la misma (Cochran, 1996). Por ejemplo, si se tiene una población de 500 usuarios y de ella se determinó seleccionar a 50 elementos (tamaño de la muestra), entonces se divide 500 entre 50 dando como resultado $k=10$, lo que indica que la población ha sido fragmentada en 50 partes y de cada una de ellas se tomará un usuario para la muestra, con lo que al final se tendrá el total de los usuarios requeridos para el análisis muestral.

La desventaja en este tipo de muestreo está en los casos en que se dan periodicidades en la distribución de la población, ya que al elegir a los miembros de la muestra con una periodicidad constante (k) se puede introducir una homogeneidad que no se da realmente en la población (Blalock, 1998). Supóngase que se está seleccionando una muestra sobre listas de 10 individuos, en los que los 5 primeros son varones y los 5 últimos mujeres, si se utiliza el muestreo aleatorio sistemático con $k=10$ siempre serían seleccionados o sólo hombres o sólo mujeres y, por lo tanto, no podría haber una representación de los dos sexos.

Muestreo estratificado

Considera características heterogéneas en la población, llamadas estratos o subconjuntos, y que al mismo tiempo son distinguidas por la homogeneidad respecto a alguna característica, como la carrera que cursa el usuario, el sexo, el estado civil, la unidad académica a la que pertenece, entre otras (Blalock, 1998). Lo que se pretende con este tipo de muestreo es asegurar que todos los estratos de interés estén representados adecuadamente en la muestra (Raj, 1979). Cada estrato actúa de forma independiente y puede aplicarse dentro

de ellos el muestreo aleatorio simple o el sistemático, para elegir los elementos concretos que formarán parte de la muestra.

Este tipo de muestreo tiene la ventaja de que permite establecer controles cualitativos para garantizar la representatividad proporcional, por lo menos en forma, de la muestra según la distribución poblacional.

Determinación del tamaño de la muestra

Hasta este momento, el lector se habrá dado cuenta del detalle que implica diseñar muestras para una investigación o estudio. No obstante, ese detalle es superado cuando existe involucramiento en el área de oportunidad en estudio, así como el conocimiento de lo que se pretende realizar y el objetivo que se pretende alcanzar (Blalock, 1998). Ciertamente, las etapas abordadas anteriormente ofrecen al investigador las herramientas necesarias para la aplicación de diversos criterios en la selección de los individuos que formarán parte de la muestra; sin embargo, la importancia de esas etapas se minimiza si no se considera el número de sujetos que habrán de ser observados; por ello surge una etapa más para el diseño de la muestra: la determinación de su tamaño.

Antes de determinar el tamaño de una muestra, por razones pragmáticas se debe considerar que si la población es grande, naturalmente la muestra tiende a ser relativamente pequeña y si la población es pequeña, entonces la tendencia de la muestra es a ser proporcionalmente grande. A partir de dicho criterio de observación, se presentan los tres elementos básicos para la determinación del tamaño de una muestra. Esos tres elementos son el nivel de confianza, el margen de error y la dispersión en la población.

Para el cálculo del tamaño de una muestra deberá considerarse el tipo de población de donde será extraída y así seguir alguna de las siguientes fórmulas básicas (figura 1):

$$\text{Población finita: } n = \frac{N p q Z^2}{E^2 (N - 1) + p q Z^2} \quad (\text{Fórmula 1})$$

$$\text{Población infinita: } n = \frac{Z^2 p q}{E^2} \quad (\text{Fórmula 2})$$

Donde: Z representa el nivel de confianza.

E indica el margen de error o nivel de precisión.

p es la proporción o porcentaje de elementos en la población que tienen la característica en estudio.

q indica el complemento de p: $q = 1 - p$

N Tamaño de la población.

Figura 1. Fórmulas básicas para calcular el tamaño de una muestra.

Nivel de confianza

Llamamos nivel de confianza a la probabilidad de que el intervalo construido en torno a un estimador implique el verdadero valor distintivo del parámetro (Cochran, 1996). De manera sencilla, mediante el nivel de confianza se desea determinar cuántos, de cada cien elementos que se tomen como muestra, deberán poseer las características representativas de la población. En otros términos, es la probabilidad de que la estimación efectuada se ajuste a la realidad.

Con la aclaración de que se ha obviado el detalle de acceso a la tabla de áreas bajo la curva normal, en la cual se adquieren los coeficientes que representan el nivel de confianza para el cálculo del tamaño de la muestra y al considerar que comúnmente en los estudios de bibliotecas los niveles de confianza mayormente utilizados son del 90%, 95% y 99%, se obtiene $Z=1.64$, $Z=1.96$ y $Z=2.57$, respectivamente.

Margen de error

También llamado error muestral o nivel de precisión, es la diferencia máxima que el investigador está dispuesto a aceptar entre un estimador y su parámetro (Blalock, 1998). En una investigación siempre se considera la comisión de errores en el proceso, pero la naturaleza de la investigación indicará hasta qué grado es posible su presencia y tolerancia. Un estimador será más preciso en la medida en que su error sea más pequeño o su precisión sea mayor.

Dispersión en la población o variabilidad

Antes de iniciar una investigación, se deberá tener claro qué tipo de población será estudiada, qué características sobresalen, cuál es la variable principal u objeto de estudio y otros aspectos que son determinantes para el involucramiento en el tema y consecuentemente para realizar diagnósticos y análisis con mayor confiabilidad y credibilidad (Raj, 1979).

Un factor importante para determinar el tamaño de una muestra es conocer cómo se encuentra distribuida la población en relación con la variable principal del estudio. Este factor puede ser observado en términos relativos o porcentuales y aplicarse de cualquier manera en la fórmula para su cálculo, siempre y cuando se aplique en el mismo sentido: relativo o porcentual: siempre proporciones o siempre porcentajes.

Si la característica principal del estudio (por ejemplo usuarios de alto desempeño académico) la posee el 70% de la población, a este valor le llamaremos variabilidad o dispersión en la población (Blacklock, 1998) y al porcentaje de sujetos que no tienen esa característica (usuarios que no tienen alto desempeño académico), 30%, la identificaremos como el complemento a la variabilidad, identificados como p y q respectivamente. En conjunto deben sumar 100%. En caso de que se desconozca esa característica, se recomienda asignar la máxima variabilidad, esto es, 50% para cada valor.

Es menester puntualizar que, en la medida en que los requerimientos de la muestra tengan un mayor nivel de exigencia, su tamaño tenderá a incrementarse; esto es, si se desea un nivel de confianza más alto o mayor precisión, el número de elementos que exigirá la muestra naturalmente tenderá a incrementarse.

¿Quién decide o cómo se determinan los requerimientos para el cálculo del tamaño de la muestra? Éste es un cuestionamiento frecuente, sin que la respuesta sea una regla general; por lo tanto considérense algunos de los siguientes criterios: necesidades del estudio o investigación, experiencia del responsable o líder del proyecto, intuición, estudios similares, pruebas de ensayo y error, entre otros. Sólo con la práctica y experiencia en la materia, es seguro obtener el mejor criterio al respecto.

Con el conocimiento adquirido hasta el momento, ya es posible determinar el tamaño de una muestra. A continuación se presentan algunos ejemplos.

Ejemplo 1. Calcular el tamaño de la muestra para una población de 500 usuarios como objetos de estudio, que requiere un nivel de confianza del 90%, con un margen muestral del 8 %, y se sabe que el 70 % de la población posee la característica de un alto nivel de desempeño académico, contra el 30% con un bajo nivel de desempeño académico.

Los valores a sustituir en la fórmula 1 son: $Z = 1.64$ para el 90% de confianza, $E = 8$ como el porcentaje del error muestral, $p = 70$ como valor de la variabilidad, $q = 30$ como complemento de la variabilidad y $N = 500$ como tamaño de la población de usuarios.

Al sustituir los datos en la fórmula 1 y realizar las operaciones se obtiene lo siguiente (figura 2):

$$n = \frac{(500)(70)(30)(1.64)^2}{(8)^2(500-1) + (70)(30)(1.64)^2}$$

$$n = 75$$

Figura 2. Resultados de la aplicación de la fórmula 1.

Ejemplo 2. Para realizar una encuesta de opinión entre jóvenes universitarios, usuarios de los servicios bibliotecarios en la ciudad de México, se requiere una muestra con el 95% de confianza, un margen de error muestral del 5% y dadas las características de la población, considérese para este caso como infinita. En este ejemplo se tomará la máxima variabilidad.

Bajo el mismo principio de asignación de valores que en el ejemplo 1, al sustituir los datos en la fórmula 2 y realizar las operaciones correspondientes, deberá tomarse una muestra de (figura 3):

$$n = \frac{(1.96)^2 (50)(50)}{(5)^2} \quad (\text{Fórmula 2})$$

$$n = 384$$

Figura 3. Resultados de la aplicación de la fórmula 2.

Ahora bien, para confirmar las afirmaciones de que las poblaciones grandes requieren muestras relativamente pequeñas y que las poblaciones pequeñas presentan muestras relativamente grandes,

se proponen los siguientes ejercicios, en los cuales se tomará como referencia la primera de las dos últimas fórmulas, la cual considera la determinación del tamaño de una muestra para poblaciones finitas:

Con los mismos requerimientos de nivel de confianza (90%), error muestral (8%) y variabilidad ($p=70\%$ y $q=30\%$), calcule el tamaño de una muestra para cada una de las poblaciones que se presentan a continuación y compárelas; finalmente se comprobará la afirmación del párrafo anterior, con lo que se tendrán mayores elementos para garantizar la representatividad de una muestra a partir del tamaño de su población:

- Población 1: $N = 100$ elementos
- Población 2: $N = 1\,000$ elementos
- Población 3: $N = 10\,000$ elementos
- Población 4: $N = 100\,000$ elementos

Después de realizar estos ejercicios, se logra apreciar la sencillez con que puede determinarse el tamaño de una muestra. Para realizarlos se fijaron las características requeridas para la credibilidad de la muestra a partir de su tamaño.

Posterior al diseño de la muestra, que comprende las etapas del método, tipo y determinación de su tamaño, se procede al tratamiento estadístico de sus datos, del cual se deriva la información concerniente a la población. Es menester subrayar que la muestra no es mágica y por lo tanto no ofrece información automática; se requiere de un proceso detallado y metódico de análisis, así como pruebas estadísticas que permitan evidenciar la naturaleza y características de la población respecto al problema en estudio. Asimismo, se deberá seguir la metodología de la estadística para la descripción e inferencia de los resultados, sin olvidar la importancia implícita del responsable en el análisis y tratamiento de los datos, al estar debidamente involucrado en el problema a estudiar.

Las técnicas para el tratamiento y descripción de los datos, así como las pruebas estadísticas que explican el comportamiento de las variables se podrán consultar en las referencias presentadas al final de este capítulo.

Finalmente, es oportuno destacar y reiterar la importancia que el proceso para el diseño de muestras tiene en todo ámbito del conocimiento, ya que a partir de su efectividad podrán evidenciarse, a bajo costo, las características de una población a partir de una porción

de ella. Asimismo, es necesario recordar que cada etapa del proceso es determinante para el logro de los objetivos planteados en el proyecto de investigación; sin embargo, se requiere de la integración de múltiples factores para alcanzar los propósitos de la misma. Entre otros, podemos destacar la capacitación del equipo de trabajo, el diseño de los instrumentos para la recolección de los datos, el tipo de variables, las pruebas estadísticas a realizar, el análisis de los resultados y la presentación del informe final.

Referencias

- BUSHA, C. y HARTER, S. (1990). *Métodos de investigación en bibliotecología*. México: UNAM, Centro Universitario de Investigaciones Bibliotecológicas.
- COCHRAN, W. (1996). *Técnicas de muestreo*. México: CECSA
- RAJ, D. (1979). *La estructura de las encuestas por muestreo*. México: Fondo de Cultura Económica.
- SANDOVAL, J. L. (2004). *Estadística aplicada a la investigación social I*. México: UNAM, Escuela Nacional de Trabajo Social.
- SANDOVAL, J. L. (2004a). *Estadística aplicada a la investigación social II*. México: UNAM, Escuela Nacional de Trabajo Social.
- SPSS base 12.0: *manual del usuario* (2003). Dublín: SPSS Inc.