

El muestreo: técnica al servicio de la valoración documental¹

Juan E. Piccini²

Resumen. La técnica de muestreo nació y se desarrolló con el objetivo de proporcionar muestras representativas de una población.

La imposibilidad de obtener información en muchos casos con base en la medición de todas las unidades que componen la población (censo) conlleva a seleccionar sólo una parte de ella (muestra) por métodos que permitan considerarla representativa y así argüir acerca de esas características poblacionales de interés.

El muestreo es por lo tanto, una herramienta de la investigación científica, cuya función básica es determinar qué parte de una población debe examinarse, con la finalidad de hacer inferencias sobre dicha población.

Ahora bien, estas técnicas, aplicadas a los archivos son en general una respuesta y/o medida que se toma frente al ritmo continuo de crecimiento vertiginoso de las series y subseries documentales conservadas tanto en archivos privados como públicos.

La finalidad entonces, no es averiguar algo sobre la población en cuestión, sino extraer individuos o unidades que serán preservadas como testimonio en función de la información en ellos contenida, para luego eliminar el resto de la población.

Es pues importante tener en cuenta esta diferencia a la hora de considerar la Selección por muestreo.

La propuesta de este trabajo es mostrar el alcance de una herramienta cuyo origen está en otra ciencia pero que el archivólogo debe saber cómo y cuándo utilizar. Importa además incluir la técnica de muestreo como objeto de estudio en la formación de la evaluación documental conciliando la

¹ Material preparado para el dictado de los cursos “Vigencia y actualización de la valoración documental” y “Valoración documental y Muestreo. Una nueva visión”, en el marco del Programa de Educación Permanente de la Escuela Universitaria de Bibliotecología y Ciencias Afines de la Universidad de la República con la colaboración archivística de las Profs. Lourdes Ramos y Alejandra Villar, docentes responsables de la asignatura “Selección Documental” de la Carrera de Archivología.

² Máster en Ingeniería Matemática, Licenciado en Matemáticas. Docente G°3, Profesor del Instituto de Matemáticas y Estadística de la Facultad de Ingeniería, Universidad de la República. Correo electrónico: jepiccini@gmail.com

archivística y la estadística como dos voces y un camino conjunto por recorrer y explorar.

Palabras clave: Archivos; Muestreo; Valoración documental

Title: Sampling: technical Assessment serving documentary.

Abstract. The sampling technique was born and developed in order to provide representative samples of a population. The impossibility of obtaining information in many cases based on the measurements of all units that compose the population (census), leads to select only a part of it (sample) by methods that allow consider it representative and so argue about these population characteristics of interest. Sampling is, therefore, a tool of scientific research, whose basic functions to determine which part of a population must be examined in order to make inferences about this population. However, these techniques are, applied to the archives, often an answer and/or a measure taken against the continued rate of exponential growth of documentary series and sub series preserved both in private and public archives. The aim then is not to find out something about the population in question, but to extract individuals or units to be preserved as testimony in light of the information contained in them, then eliminate the rest of the population. Therefore it's important to note this difference when considering the sample selection. The purpose of this paper is to show the scope of a tool whose origin is in another science but the archivist must know how and when to use. It's important, also, to include the sampling technique as a research subject on the formation of the documentary assessment reconciling the archives and statistics as two voices and a way to go and explore joint.

Keywords: Archives; Sampling; Documentary Assessment

1. Introducción

El muestreo es una herramienta matemática que tiene su origen en las ciencias naturales, en donde ha sido fundamental para estudiar grandes volúmenes de datos a partir de una pequeña fracción de los mismos. Esta técnica nació y se desarrolló con el objetivo de proporcionar muestras representativas de una población, de modo que las conclusiones que pudieran extraerse de dichas muestras pudieran generalizarse a la población de origen.

La archivística recurre al muestreo ante la imposibilidad y/o inconveniencia de guardar todos los documentos producidos por una institución. Este instrumento permite preservar una muestra representativa de la población de documentos, como forma de reflejar las funciones y actividades que en ellos se registran.

La finalidad no es averiguar algo sobre la población en cuestión, sino extraer individuos o unidades que serán preservadas en función de la información en ellos contenida, como testimonio para luego eliminar el resto de la población. Es conveniente tener presente esta significativa diferencia a la hora de considerar la Selección por muestreo.

En nuestra opinión se entienden como distintas, técnicas que son casos particulares de otras, lo que obstaculiza una visión más clara del abanico de las herramientas disponibles así como el entendimiento dentro del colectivo archivístico. El espíritu de este documento es pues intentar unificar definiciones y conceptos y promover la discusión acerca del uso acrítico del muestreo en la archivística.

2. ¿Por qué muestreo?

Nuestros conocimientos, actitudes y acciones están basados en gran parte sobre muestras. Emitimos opiniones y juicios basados solamente en una parte del panorama total. Hay por tanto una pérdida de información que condiciona las extrapolaciones y juicios que podamos hacer, pues se basan en la visión de un fragmento del total.

Es entonces fundamental que ese fragmento sobre el que construimos nuestros juicios y opiniones sea “representativo” de ese total, en el sentido de minimizar la pérdida de información. La teoría del muestreo es una herramienta para minimizar dicha pérdida.

3. Conceptos - definiciones – terminología

La palabra *población* se emplea para referirse al conjunto dentro del cual se elige el subconjunto llamado *muestra*. Su definición puede no presentar problema, por ejemplo cuando se muestrean las piezas documentales de un fondo documental a fin de determinar su valor promedio. Es necesario que se encuentre bien definida, de modo que pueda decidirse si un caso dudoso pertenece o no a la misma. La población de la que se elige la muestra (la *población muestreable*) debería coincidir con la población sobre la cual se desea información (la *población objetivo, target population*).

Suele suceder que la población muestreable es más restringida que la población objetivo. En este caso debe tenerse presente que las conclusiones extraídas de la muestra serán aplicables a la población muestreable, y hará falta información extra para decidir hasta qué grado son aplicables dichas conclusiones en la población objetivo.

A modo de ejemplo, si de los documentos generados por cierta institución extraemos una muestra de aquellos situados en los anaqueles más accesibles (correspondientes al último año, p.ej.), sería aventurado extrapolar las conclusiones al total de documentos generados por dicha institución, a menos que tengamos información adicional (p.ej. que el último año ha sido igual que los anteriores).

El muestreo puede tener un fin *descriptivo* o *analítico*. En el primer caso el objetivo es obtener cierta información respecto a una población. Por ejemplo la cantidad de profesionales que existen en la serie documental “legajos de personal” de una Institución.

En el segundo caso se comparan subgrupos de la población para averiguar si existen diferencias y formular o contrastar hipótesis sobre sus causas. (Por ejemplo, comparar la proporción mujer/hombre entre los profesionales de la serie documental “legajos de personal” anteriormente citada). La distinción entre descriptivo y analítico rara vez es tajante. Muchas muestras brindan datos que sirven a ambos propósitos.

Antes de seleccionar la muestra, la población muestreable debe partitionarse en *unidades de muestreo* (los individuos que forman la población) Los resultados pueden ser bien distintos según la unidad de muestreo que se elija. Por ejemplo se tiene la serie documental “tarjetas de control de horario” de una cierta empresa (ésta es la población muestreable), y se desea guardar una muestra, eliminando el resto.

Si se toman las distintas secciones de la empresa como unidades de muestreo, dicha muestra estará formada por todas las tarjetas de control de horario de la o las secciones elegidas. Si en cambio, la unidad de muestreo es el mes, la muestra estará formada por todas las tarjetas de control de horario correspondientes al mes o meses elegidos. Este *marco* (lista de las unidades que serán muestreadas) debe ser claro, dado que si

la unidad es la sección entonces las conclusiones que puedan surgir del análisis de la muestra serán conclusiones sobre secciones y no sobre meses.

Observaciones

Conviene remarcar nuevamente que en general el muestreo se utiliza para extraer muestras sobre las cuales se miden ciertos rasgos o atributos, los que luego serán extrapolados a la población muestreable, sin embargo en la archivística el muestreo se utiliza no para averiguar algo sobre la población en cuestión, sino para extraer individuos o unidades que serán preservadas según su valor, eliminándose al resto de la población.

Dicho valor está íntimamente relacionado con el contenido de la unidad, es decir, la información que la misma contiene. Esto no excluye que el propio soporte o tipo documental pueda tener valor en sí mismo (pergaminos, papiros, cédulas, etc.). Pero dejando estos casos de lado, en general, el valor de una unidad estriba en la información que contiene.

Es importante observar que las herramientas matemáticas involucradas en el muestreo aleatorio descansan en resultados que presuponen poblaciones grandes. El tamaño que la muestra debe tener para garantizar un nivel de representatividad dado no es el mismo para una población pequeña que para una grande. De este modo, una muestra de mil individuos de una población de un millón será representativa, mientras que una muestra de un individuo de una población de mil no lo será, pese a que la relación muestra/población es la misma en ambos casos. *Esto es especialmente relevante en la archivística, donde se suelen tener pocas series que totalizan muchos documentos.*

4. Tipos de muestreo

Existen dos métodos para seleccionar muestras de poblaciones: el *muestreo aleatorio* (que incorpora el azar como recurso en el proceso de selección), y el *muestreo no aleatorio (intencional o de juicio)*.

Cuando el primero cumple con la condición de que todos los elementos de la población tienen alguna oportunidad de ser escogidos en la muestra, si la probabilidad correspondiente a cada una de las posibles muestras es conocida de antemano (y en esto es fundamental que el marco esté bien definido) recibe el nombre de *muestreo probabilístico*.

En este caso se habla de *muestras probabilísticas*, pues no es en rigor correcto hablar de *muestras representativas* dado que al no conocer las características de la población, no es posible tener certeza de que tales

características se haya conseguido plasmar en la muestra. Esto es especialmente importante en el caso de la archivística, donde las muestras no se extraen para medir rasgos o características que puedan ser extrapoladas luego al total de la población, sino que las muestras se utilizan para seleccionar aquellos individuos que serán preservados (sea como testimonio o para futuros estudios), eliminándose al resto de la población.

En lo que sigue:

N = tamaño de la población (número de individuos o unidades que la forman)

n = tamaño de la muestra (número de individuos o unidades que serán elegidos de entre los N de la población)

4.1. Muestreo aleatorio

a) Muestreo aleatorio simple

Supongamos que la población muestreable se divide en N unidades (por ejemplo 1000 unidades), entonces elegimos por ejemplo 50 de dichas unidades de modo que cada una de las muestras posibles de 50 unidades elegidas de entre las 1000 tenga la misma probabilidad de salir. Esto puede lograrse mediante el muestreo con y sin reposición, entendiéndose por esto cuando la unidad elegida se vuelve a incorporar a la población, pudiendo ser elegida nuevamente. Cuando N (el tamaño de la población) es grande y n (el tamaño de la muestra) es muy pequeño comparado con N , las predicciones matemáticas de ambas opciones son muy similares, pero el desarrollo matemático del muestreo con reposición es más sencillo. En la archivística los muestreos son sin reposición, por los motivos ya mencionados.

Con el muestreo aleatorio simple toda unidad de la población tiene las mismas posibilidades de formar parte de la muestra. Es un método sencillo y de fácil comprensión.

La muestra se elige sin tomar en consideración ninguno de los atributos o características de interés que las unidades puedan poseer.

Por eso, cuando se trabaja con una población heterogénea es posible que la muestra no represente adecuadamente a dicha población.

b) Muestreo aleatorio sistemático

En este muestreo, las N unidades que conforman la población se enumeran de 1 a N según un cierto criterio. Buscamos elegir una muestra que represente una fracción $1/k$ del total. Entonces el tamaño de la muestra deberá ser $n=N/k$.

Estas n unidades se extraen equidistantes (siendo k la distancia entre unidades). Primero elegimos una unidad al azar entre las k primeras y el resto de las $n-1$ unidades se eligen consecutivamente, cada una a distancia k de la elegida anteriormente. Es como tener un peine con n dientes, cada diente a distancia k de sus dos dientes vecinos.

A modo de ejemplo, supongamos que tenemos $N = 300$ expedientes ordenados por antigüedad, de modo que el expediente más viejo tiene el número 1 y el más nuevo el número 300.

Deseamos elegir una muestra de razón $1/k = 1/5$, de donde tendremos que la muestra estará formada por $n = 300/5 = 60$ expedientes. Elegimos un expediente (por ejemplo al azar) entre los 5 primeros, digamos el 3 (aquí irá el primer diente del peine). El segundo expediente de la muestra será entonces el 8 (el tercero de los 5 segundos), luego 13 (el tercero de los 5 terceros), el 18 (tercero de los 5 cuartos), etc. hasta completar los $n = 60$ expedientes.

c) Muestreo aleatorio estratificado

Consiste en la división previa de la población bajo estudio en grupos o clases (llamadas estratos) que se suponen homogéneos respecto a alguna característica. Se decide primero cuántas unidades de cada estrato compondrán la muestra (o sea se asigna una cuota a cada estrato). Habitualmente se hace de dos formas:

Por asignación proporcional: el tamaño de la muestra dentro de cada estrato es proporcional al tamaño del estrato dentro de la población. Por asignación óptima: la muestra recogerá más individuos de aquellos estratos que tengan más variabilidad. Para ello es necesario un conocimiento previo de la población.

Una vez decidido cuantas unidades se elegirán en cada estrato, resta ver cómo elegir las. Suelen usarse dos métodos: a) Dentro de cada estrato se puede usar la técnica de muestreo sistemático (se llama entonces muestreo estratificado sistemático) y b) también puede usarse la técnica de muestreo aleatorio simple dentro de cada estrato.

El muestreo aleatorio estratificado se recomienda cuando la población tiene marcadas diferencias. Al agrupar en estratos más homogéneos se gana en precisión.

Este muestreo pretende asegurar la representación de cada grupo en la muestra, tiende a asegurar que la muestra represente adecuadamente a la población en función de unas variables seleccionadas, pero se ha de conocer la distribución en la población de las variables utilizadas para la estratificación.

Para ilustrar veamos un ejemplo de muestreo aleatorio estratificado por asignación proporcional: supongamos una serie documental de una Universidad con 1500 documentos de los cuales el 25% son de naturaleza contable, el 40% de naturaleza administrativa no contable y el 35% restante son de naturaleza académica. Si se desea una muestra de por ejemplo 100 documentos, 25 de ellos se elegirán de entre los documentos contables, 40 de entre los administrativos no contables y 35 de entre los documentos académicos.

Ahora un ejemplo pero con asignación óptima: volviendo al ejemplo anterior, supongamos que los documentos administrativos no contables presentan una mayor variabilidad o diversidad que los académicos y los administrativos contables, de donde se asignará una mayor cuota dentro de los 100 documentos que formarán la muestra, por ejemplo eligiendo 50 documentos administrativos no contables en vez de 40 (en detrimento de los otros estratos).

Observación

El muestreo sistemático puede verse como un caso particular del muestreo estratificado. Revisemos nuestro ejemplo de $N = 300$ documentos ordenados cronológicamente, de los que deseamos extraer $n = 60$. Al elegir $k = 5$, estamos partiendo la población en $n = 60$ estratos de 5 unidades cada uno (las primeras 5 unidades, las segundas 5, las terceras 5, etc.), y luego elegimos una unidad en cada estrato, equiespaciada de las unidades elegidas en los otros estratos por cuanto siempre elegimos p.ej. la tercera unidad de cada estrato.

El pensar al muestreo sistemático como un caso del muestreo estratificado, muestra su principal debilidad: el imponer una estratificación que no necesariamente responde a alguna homogeneidad intrínseca de los datos. Funcionará bien si la característica que se mide no está relacionada con el orden impuesto a las unidades.

Si los datos presentan algún tipo de periodicidad, existe el peligro de perder dicha característica al capturar siempre la misma posición relativa. Por ejemplo registros climatológicos anuales ordenados por mes. Si guardamos siempre los registros correspondientes al mismo mes, tendremos una idea equivocada sobre el clima anual si nos basamos en esa muestra.

d) Muestreo aleatorio por estadios múltiples

Esta técnica es la única opción cuando no se dispone de la lista completa de la población de referencia o bien cuando por medio de la técnica de muestreo simple o estratificado se obtiene una muestra con unidades distribuidas de tal forma que resultan de difícil acceso. Ilustraremos con un ejemplo: supongamos que deseamos una muestra de la población de empleados públicos que tienen título terciario (sus legajos). Si tuviésemos una lista donde constan todos los N empleados públicos, podríamos asignar un número a cada empleado y luego elegir n números al azar de entre esos N. Ahora bien, puede suceder que no tengamos tal lista, o que la misma no nos diga si tienen o no título terciario, o que los legajos de algunos de los empleados sorteados se encuentren en una región de difícil acceso, resultando caro viajar hasta allá por un solo legajo.

En el muestreo por estadios múltiples se subdivide la población en varios niveles ordenados que se extraen sucesivamente por medio de un procedimiento de embudo. El muestreo se desarrolla en varias fases o extracciones sucesivas para cada nivel, no por motivos estadísticos sino por motivos económicos y de eficiencia, buscando facilitar la recogida de la muestra.

Volviendo a nuestro ejemplo, se desea una muestra de empleados públicos con título terciario. Éstos pueden subdividirse en unidades primarias (los organismos públicos) y unidades secundarias (los propios empleados con título terciario).

Primero se extrae una muestra de las unidades primarias (esto presupone que contamos con una lista completa de tales unidades, en este caso organismos públicos), y luego extraemos una muestra de unidades secundarias (funcionarios con título terciario) de cada una de las unidades primarias seleccionadas en la primera extracción.

e) Muestreo por conglomerados

Se utiliza cuando la población se encuentra dividida de manera natural en grupos que se supone contienen toda la variabilidad de la población, es decir, la representan fielmente respecto a la característica elegida. Pueden entonces seleccionarse sólo algunos de estos grupos o *conglomerados* para la realización del estudio.

A modo de ejemplo supongamos que queremos conservar una muestra representativa de los escalafones existentes en la Administración Pública. Los funcionarios públicos se agrupan de manera natural en ministerios públicos, y en cada uno de ellos existe todo el abanico de los escalafones (administrativos, técnicos, profesionales, etc.). Entonces en

una primera etapa elegimos algunos Ministerios de los cuales, en una segunda etapa elegiremos algunos funcionarios como representantes de los escalafones existentes.

Este método, similar al de estadios múltiples, tiene la ventaja de simplificar la recogida de información muestral.

Observaciones

i) Las ideas de *estratos* y *conglomerados* son en cierto sentido, opuestas. Los primeros funcionan mejor cuanto más homogénea es la población del estrato y más diferentes son éstos entre sí (poca disimilaridad interna, mucha disimilaridad externa). En los segundos ocurre lo contrario, los conglomerados deben presentar toda la variabilidad de la población y ser muy parecidos entre sí (mucha disimilaridad interna, poca disimilaridad externa).

ii) *Homogéneo* significa en el contexto de la estratificación, que no hay mucha variabilidad. Los estratos funcionan mejor cuanto más homogéneos son cada uno de ellos respecto a la característica a medir.

iii) Por el contrario, la heterogeneidad hace inútil la división en estratos. Si se dan las mismas diferencias dentro del estrato que en toda la población, no hay por qué usar este método de muestreo. En los casos en los que existan grupos que contengan toda la variabilidad de la población, lo que se construyen son conglomerados, que ahorran algo del trabajo que supondría analizar toda la población.

En resumen, estratos y conglomerados funcionan bajo principios opuestos: los estratos son mejores cuanto más homogéneo es el grupo respecto a la característica a estudiar y los conglomerados lo son si representan fielmente a la población, esto es, contienen toda su variabilidad, son *heterogéneos*.

4.2. Muestreo no aleatorio

Aquél para el que no puede calcularse la probabilidad de extracción de una determinada muestra. Se busca seleccionar n individuos que se juzga de antemano tienen una información vital o sobresalen del resto según algún criterio preestablecido, individuos con alto valor según dicho criterio.

a) Muestreo no aleatorio por cuotas

En primer lugar es necesario dividir la población de referencia en varios estratos definidos por algunas variables de distribución conocida. Posteriormente se calcula el peso proporcional de cada estrato, es decir, la proporción de población que representan. Finalmente se multiplica cada peso por el tamaño de n de la muestra para determinar la cuota precisa en cada estrato. Es como el muestreo estratificado con asignación proporcional, solo que ahora en cada estrato se elige en forma no aleatoria.

Volvamos al ejemplo de los documentos de una Universidad con 1500 documentos de los cuales el 25% son de naturaleza contable, el 40% de naturaleza administrativa no contable y el 35% restante son de naturaleza académica. Deseamos conservar una muestra de 100 documentos, pero ahora entre los 25 documentos de naturaleza contable priorizaremos los relacionados con inversiones en obras, de los 40 documentos de naturaleza administrativa no contable daremos prioridad a los legajos de personal, y de los 35 documentos de naturaleza académica consideraremos especialmente las actas de resolución de los Claustros.

b) Muestreo no aleatorio subjetivo por decisión razonada

En este caso las unidades de la muestra se eligen en función de algunas de sus características de manera racional y no casual. Así, dada una colección personal de documentos pertenecientes a un personaje célebre, se conservarán el certificado de nacimiento, fotografías, correspondencia personal, etc. pero no la factura de compra de una licuadora por ejemplo.

5. ¿Otros tipos de muestreo?

En la comunidad de archivistas se confunde muestreo con muestreo aleatorio, y se menciona la selección cualitativa intrínseca y extrínseca como métodos separados de los métodos de muestreo. Consideramos que éstos son casos particulares del muestreo no aleatorio y aleatorio respectivamente, como intentaremos mostrar a continuación. Asimismo se menciona la valoración como algo separado del muestreo.

Las definiciones de selección cualitativa que siguen fueron tomadas del material correspondiente al Tema 8 del curso Archivística de la Diplomatura de Biblioteconomía y Documentación de la Universidad

de Granada.³ Cabe mencionar que si consultamos cualquier material de los usados en archivística, encontraremos contenidos similares.

5.1. Selección cualitativa intrínseca

“Establece a priori un criterio cuya presencia o ausencia determinará la conservación o la eliminación de cada documento o expediente”. Por ejemplo para expedientes judiciales se puede establecer como criterio conservar los relacionados con los casos más sonados de la época. *“Este modo de selección busca eliminar series voluminosas de documentación sin interés, conservándose sólo lo que se considera relevante”.*

Pensamos que estamos ante un muestreo no aleatorio, en particular el muestreo subjetivo por decisión razonada, donde la muestra se compone de todas las unidades (documentos, expedientes, etc.) que presentan el atributo o criterio definido a priori y que los caracteriza como valiosos o de interés.

5.2. Selección cualitativa extrínseca

Este tipo de selección, por el contrario, no define a priori un criterio o atributo cuya presencia o ausencia en las unidades decidirán su preservación o eliminación. Se presenta dividida en *alfabética, cronológica, topográfica y combinada.*

a) Selección Cualitativa Extrínseca Alfabética:

“Se emplea en el caso de los expedientes que contienen datos de carácter personal. Consiste en elegir aquellos expedientes cuyos apellidos comiencen por una o más iniciales”. Vemos que si las letras se eligen al azar, estaremos ante un caso de muestreo aleatorio sobre la población de letras, dado que se eligen letras del alfabeto y se preservan los expedientes de personas cuyos apellidos comienzan con alguna de las letras elegidas. El que dichas letras se elijan mediante muestreo aleatorio simple, sistemático u otro es irrelevante para el caso. Si la elección de las letras no es aleatoria, estaremos ante un caso de muestreo no aleatorio.

³ <http://www.ugr.es/~josemdc/Recursos/Tema8.pdf>

b) Selección Cualitativa Extrínseca Cronológica:

“Consiste en seleccionar la documentación de determinados años (uno cada cinco años, diez,...) o tomados al azar o los considerados significativos o más importantes por los acontecimientos desarrollados”. Nuevamente en este caso vemos que se hace un muestreo aleatorio sistemático o aleatorio simple en el conjunto de años (la población son los años, se eligen algunos años ya sea en forma sistemática o aleatoria simple), o un muestreo subjetivo por decisión razonada cuando elegimos los años considerados significativos o importantes.

“Otra posibilidad consiste en conservar una representación de la documentación de cada año, bien sea un mes determinado, un mes rotatorio o uno elegido al azar”. En este caso se estratifica primero dividiendo cada año en meses y luego se muestrea en la población de meses, ya sea con o sin aleatoriedad.

c) Selección Cualitativa Extrínseca Topográfica:

“Se basa en la elección de áreas geográficas cuya documentación es conservada íntegramente o la de determinados lugares de cada área”. Aquí la unidad de muestreo es el área geográfica o la subárea, la población es el conjunto de tales áreas (subáreas). Se eligen algunas unidades (presumiblemente mediante muestreo aleatorio simple o por estadios múltiples, a juzgar por el comentario que sigue) de dicha población, cuya documentación será la que se preserve íntegramente.

“El problema es de nuevo la representatividad, ya que ¿puede una región o varias localidades representar al conjunto de las regiones en las que se enclavan?”

Recordemos que una de las desventajas del muestreo aleatorio simple era justamente que para poblaciones heterogéneas, la muestra elegida podía no ser representativa. Asimismo el muestreo por estadios múltiples es un método pensado desde lo logístico, no desde lo estadístico. Para lidiar con esta falta de representatividad pensamos que mejor sería implementar un muestreo por conglomerados o un muestreo estratificado (no necesariamente aleatorio), suponiendo que se tenga el necesario conocimiento previo de la población a muestrear.

d) Selección Cualitativa Extrínseca Combinada

“Los métodos expuestos no suelen aplicarse individualmente, sino que con frecuencia se combinan varios de ellos para obviar en la medida de lo posible los inconvenientes de cada uno”. Es decir que se combinarán el muestreo no aleatorio con el aleatorio. Por ejemplo eligiendo algunas regiones y/o años de eminente importancia (ya sea mediante muestreo subjetivo por decisión razonada o mediante muestreo no aleatorio por cuotas), o por ejemplo estratificando, o por conglomerados o por muestreo aleatorio simple.

5.3. Valoración

El hecho mismo de valorar implica efectuar un muestreo, ya que cuando se evalúa y se decide luego en función de dicha evaluación, el efecto final es extraer una muestra no aleatoria de una población, eliminándose al resto (recordemos que el valor asignado a una unidad está íntimamente asociado con el contenido informativo de la misma).

En el proceso de valoración se hace primero un muestreo no aleatorio (la llamada selección cualitativa intrínseca, dado que se fija un criterio, plasmado en el valor, según el cual se conservará o no la unidad en cuestión) el cual produce grupos claramente diferenciados: unidades de evidente valor mínimo que serán eliminadas en su totalidad y el otro grupo, el cual a su vez suele dividirse en dos subgrupos, el de las unidades de alto valor (las que se conservarán en su totalidad), y las unidades de valor intermedio o zona gris, las cuales serán eliminadas salvo una muestra.

El muestreo no aleatorio se agota en los grupos de bajo y alto valor, aplicándose alguna variedad de muestreo aleatorio en la zona gris. Esto suele llamarse selección cualitativa extrínseca en la comunidad de archivistas, siendo en realidad casos particulares de los métodos aleatorios como ya hemos visto.

En esta zona gris, donde todos los individuos valen lo mismo y solamente se pueden conservar algunos, resulta importante que la muestra que se va a conservar sea representativa y capture del mejor modo posible los rasgos característicos (generalmente desconocidos) de la población que será eliminada.

5.4. Conclusiones

- Se incorpora a la archivística una herramienta concebida en principio para otros fines.
- Al hacerse en forma acrítica se genera confusión en lo conceptual:
 - Se confunde el concepto de tomar una muestra (el qué) con el cómo (si interviene o no el azar), llegándose a reducir el muestreo al caso particular del muestreo aleatorio, como si fueran sinónimos.
 - Se percibe la valoración documental como algo separado del muestreo, cuando en realidad se está aplicando primero un muestreo no aleatorio seguido de otro aleatorio en uno de los subgrupos producidos por el primer muestreo.
 - Se definen como métodos distintos (incluso de otro orden) lo que en realidad son casos particulares. Así, se presentan las selecciones cualitativas como métodos aparte de los métodos de muestreo, y dentro de éstas las cualitativas extrínsecas como métodos aparte de los de muestreo aleatorio.
 - Esta visión fragmentada dificulta la comprensión de los distintos métodos y por tanto su aplicación y eventual mejoría (véase selección cualitativa extrínseca topográfica).

Referencias bibliográficas

COCHRAN, William G. *Técnicas de Muestreo*. México : Ed. CECSA, 1996.

DUVERGER, Maurice. *Métodos de las Ciencias Sociales*. Barcelona : Ariel, 1962. (Demos - Colección de Ciencias Políticas).

HULL, Félix. “Utilización de técnicas de muestreo en la conservación de registros”. *Estudio del RAMP y directrices*. Paris, Programa General de Información y UNISIST. Paris: UNESCO, 1981, 67 p. (PGI-81/WS/26).

MORALES DEL CASTILLO, José Manuel. *Tema 8 de la materia Archivística, 2º curso de la Diplomatura de Biblioteconomía y Documentación de la Universidad de Granada* página web:

<http://www.ugr.es/~josemdc/Recursos/Tema8.pdf> (página consultada el 1/7/2010)

SLONIM, MORRIS, James. *Muestreo : guía ágil y precisa de estadística práctica*. Buenos Aires : Americana, 1974.

Wikipedia.org.

http://es.wikipedia.org/wiki/Muestreo_en_estad%C3%ADstica *Muestreo en estadística*. (Página consultada el 30/6/2010)

Nota: El presente artículo ha sido corregido, en relación a su estilo de escritura, por la Lic. Yanet Fuster, con el consentimiento de su autor.