



Revista Electrónica de Metodología Aplicada

1998, Vol. 3 n° 1, pp. 1-15

URL:<http://www3.uniovi.es/~Psi/REMA/v3n1/a1v3n1.wp5>

---

**TEORÍA DEL MUESTREO:  
PARTICULARIDADES DEL DISEÑO  
MUESTRAL EN ESTUDIOS DE LA  
CONDUCTA SOCIAL.**

**Fernando Mandujano Bustamante**  
**Universidad de Playa Ancha**  
**Valparaiso, Chile**  
**e-mail:fm@uplaced.upa.cl**

**ABSTRACT.**

Sampling design for social studies have very specific requirements, that are generally ignored in the scarce bibliography about matter. The common assumption is that mathematical statistical theory applied in the testing in industrial control quality, pharmacology, medical therapy, etc. is adequate for social populations studies. The problem of the multivariation, the bringing up to date of sampling frame, the variance estimation for sample sizing, in social studies have exigencies no solved in the Sampling Theory. We confront the un-recognized sampling paradoxes: the more used probabilistic samples make use of non aleatory procedures for random selection of units; for estimate standard deviation needed to design a probabilistic sample, you can use a non-probabilistic sample; you adapt the sample size considering only one standard deviation of only one variable; however, you denominate "probabilistic" the whole sample for all other variables and its interrelations.

Reviewing the principles of sampling – trying avoid to use mathematical formulas- we analyze the troubles for truly probabilistic designs, and, the basis of the validity of non-probabilistic procedures.

**Key word:** Sampling Theory, Sample Sizing, Statistical Methods, Social Research.

**1.- Introducción**

El muestreo es el área teóricamente más débil de la metodología estadística para la investigación social; el diseño muestral en el estudio de fenómenos sociales tiene características y exigencias absolutamente específicas y que han sido subestimadas o simplemente ignoradas en la literatura. Esto ha contribuido a convertir el diseño muestral en los estudios sociales en una especie de zona extraña donde las cosas se hacen con oscuros rituales matemáticos que sólo escasos gurúes pueden entender . y que el público - como en el cuento del traje del emperador -, tiene que encontrar bueno aunque no lo comprenda. En este trabajo abordamos exactamente este problema, qué especificidades tiene, los alcances y limitaciones de la teoría muestral aplicada a poblaciones humanas.

Se ha procurado prescindir – en lo posible - de la utilización de fórmulas y expresiones matemáticas por dos razones: la primera es que el razonamiento estadístico es sólo una formalización de la teoría, no la reemplaza (aunque indudablemente permite precisar los enunciados); y, segundo, en la aún relativamente escasa bibliografía sobre el tema, no sólo los árboles numéricos han dificultado ver el bosque, sino que han generado un espacio muy propicio para la extrapolación impune de la teoría muestral estadística matemática y la aplicable al estudio de fenómenos no sociales (ciencias naturales, control de calidad industrial, etc.), a la conducta social. Estas dos consideraciones, explican abundantes errores en el razonamiento del muestreo, que pretendemos dilucidar en este documento.

El texto presentado asume que una adecuada comprensión de los principios teóricos del muestreo permitirá, un mejor manejo de los aspectos formales a la hora de decidir diseños muestrales para estudios sociales, y, de paso, hace visible cómo una perspectiva epistemológica mas "constructivista" contribuye a su esclarecimiento.

Para alcanzar las ventajas que el muestreo permite (Briones, 1990) necesitamos manejar los recursos teóricos y prácticos que permitan decidir diseños muestrales adecuados; en ese orden, revisamos enseguida los fundamentos teóricos del muestreo para luego tratar las condiciones de aplicación en los estudios sociales.

## 2.- Los fundamentos del muestreo

Todo el muestreo se basa en cuatro enunciados o principios: dos axiomas, un teorema y una ley:

**2.2.a) - Axioma 1:** Todo trabajo con muestras requiere un conocimiento previo del universo del cual se extraerán dichas muestras.

El tomar muestras se basa en la posibilidad de generalizar lo observado en una parte, hacia el todo; se basa en la idea de la similitud entre la composición de la muestra y el universo del cual se extrajo. Esto parece obvio, pero requiere un conocimiento anterior del universo. Si no se dispone de antecedentes que avalen ello, no podemos hablar de una muestra: para conocer algunas características de una población usando muestras. Necesitamos tener algunos conocimientos de dicha población en lo referente a la distribución de las características que nos importa estudiar. Esto, según Azorín y Crespo(1986, p.69), “el diseño óptimo de la muestra, en particular la determinación previa de su tamaño óptimo, sólo podría conseguirse a partir del conocimiento de la población. Esto es lo que algunos autores han llamado ‘paradoja de Friedman’”.

Por ejemplo: nosotros sabemos que para conocer cómo los seres humanos asimilan el oxígeno en los tejidos de su cuerpo, nos basta con tomar una muestra con  $n = 1$ . Un observador que no fuese humano ni conociera a ninguno, no podría saber que con la muestra mínima basta.

Por lo tanto, **muestra** es un conjunto de eventos, casos o unidades que son tomadas o extraídas de una población, que de acuerdo a nuestro conocimiento de dicha población, posee

algunas características similares al resto de conjunto del cual se extrae.

Esta distinción esta en la base de la propia identificación del universo. Y será un principio fundamental en todo diseño muestral posible.

**2.2.b) Axioma 2:** La similitud entre los componentes del universo no son propiedades dadas sino distinciones hechas por el observador.

La o las características presentes en los elementos de un universo que permiten identificarlo como tal y hacen posible el muestreo son elegidas por el investigador; a modo de ejemplo, desde el punto de vista de la identidad del yo, el universo de todos los seres humanos vivientes es absolutamente heterogéneo; en cambio, en cuanto a la asimilación celular del oxígeno, es absolutamente homogéneo. En el primer caso, la "muestra" sería la máxima posible ya que ningún elemento puede representar a otro; en el segundo caso, como vimos, con un elemento basta.

En el ejemplo indicado hay implícitas dos consecuencias; primero: hablamos de universo de acuerdo a la característica que definimos como observable en él. Así el universo de todos los seres humanos vivientes, no es el mismo que el de todos los seres humanos que hablan sólo español; el hecho de que uno puede ser subconjunto del otro puede ser relevante, sí, y sólo sí importa para el conocimiento que se busca alcanzar mediante el muestreo; cuando decimos la población de Chile, entendemos la suma total de los habitantes de Chile sólo por la fuerza del hábito, por omisión; pero en sentido estricto, Chile tiene tantas poblaciones como poblaciones se delimiten. En síntesis la propiedad de *ser* universo, no esta dada, es declarada por el investigador-observador.

Y, segundo, cualquier distinción de un universo requiere al menos una característica constante en todas las unidades que debe fijar el observador. Esta delimita, esto es, permite afirmar si un evento o elemento es o no parte del universo; hecha dicha distinción el observador fija nuevas características -las variables que estudiará- que se presentan en los elementos con diferente magnitud, cada una de estas características o variables tiene una distribución propia y distinta de las otras, es decir poseen una homogeneidad u heterogeneidad distinta, que estadísticamente se denomina varianza (o su raíz cuadrada, la desviación típica).

Ejemplo: El universo de los lectores del diario "X" (característica constante fijada para delimitar, en este caso sólo una) tiene elementos que varían poco en nivel socioeconómico, mucho en religiosidad, y mucho más en gustos musicales (para indicar sólo tres). Esto implica que el universo es perfectamente homogéneo con relación a la primera característica (la elegida para delimitar), y se hace cada vez más heterogéneo con cada una de las tres características, variables, siguientes. Por lo tanto el universo es homogéneo o heterogéneo según la característica el observador determine.

Además, aunque entre los lectores del diario X, exista la propiedad de creer o no en OVNIS, si el investigador no la considera, no es una característica del universo (no cumplirá ningún papel, aunque no se la niegue).

## 2.2. c)- Teorema del Límite Central

Una formulación de este (Daniel; 1985, p.108) (<sup>1</sup>):

*“Sin tener en cuenta la forma funcional de la población de donde se extrae la muestra, la distribución de las medias muestrales, calculadas con muestras de tamaño  $n$  extraídas de una población con media  $\mu$  y varianza finita  $\sigma^2$ , se aproxima a una distribución normal con media  $\mu$  y varianza  $\sigma^2/n$ , cuando  $n$  aumenta. Si  $n$  es grande, la distribución de medias muestrales puede aproximarse mucho a una distribución normal”.*

Con un ejemplo muy simple: si tomamos una muestra aleatoria de 20 habitantes de Valparaíso, y calculamos su promedio de edad, es mayor la probabilidad que el promedio así obtenido se acerque al promedio real de edad de los porteños, y viceversa, la probabilidad de que el promedio así obtenido se aleje del promedio real es más baja - si agregamos una segunda muestra, o tomamos una mas grande, aumentaremos la probabilidad de acierto.

## 2.2.d) La Ley de Los Grandes Números

En la formulación de Robinson (1981, p.16):

*“Si un experimento es repetido más y más veces, entonces la frecuencia relativa del evento tiende a acercarse a la probabilidad del evento”.*

Mientras más veces se lance un dado más clara será la tendencia a que cada una de sus caras cumpla con aparecer 1/6 del total de lanzamientos.

De este modo la tendencia de muchas variables sociales de distribuirse de acuerdo a la curva normal como la edad, la inteligencia, la estatura, etc., se hace visible sólo a partir de grandes números de casos, a partir de 30 casos tomados al azar.

Estos cuatro enunciados o principios, como se ha indicado, fundamentan el muestreo, hacen posible conocer totalidades a partir de partes. Los mas conocidos - por ser justamente más “matemáticos” - son los dos últimos, el teorema y la ley; los dos primeros se manejan sólo tácitamente, salvo la única mención encontrada para este estudio, con relación al axioma 1, la referencia que hace Azorín a la ‘paradoja de Friedman’.

## 3.- Muestras probabilísticas en la investigación social

Como se indicó en la fundamentación del Axioma 1, hablar de muestra implica alguna forma de representación del universo a través de una parte tomada de él. Clásicamente, uno de los primeros problemas que debe resolverse cuando se ha decidido utilizar muestras, es, si se trabajará con muestras con probabilidad de error muestral conocido, es decir **muestras probabilísticas**, o muestras sin conocimiento de esa probabilidad de error muestral, las denominadas **muestras no probabilísticas** (<sup>2</sup>).

Las primeras gozan, en la investigación social, de un prestigio sobrevalorado pese que

presentan dificultades habitualmente no resueltas, dificultades que, como hemos indicado, quedan en la zona oscura de la extrapolación simple de los enunciados de la estadística matemática o pura a los hechos sociales.

Veremos enseguida las dificultades que tiene el diseño de muestras probabilísticas en la conducta social, y cómo frecuentemente se toman ciertas licencias para su aplicación.

Para diseñar una muestra probabilística, es decir una muestra con probabilidad conocida de error muestral (diferencia entre el estadístico obtenido y el parámetro real en la población), se necesita cumplir cuatro requisitos:

3.a) Un esquema muestral completo y actualizado que permita numerar las unidades del universo; operacionalmente este esquema o marco muestral (un listado, un mapa), *es el universo o población*.

3.b) Un procedimiento de selección aleatorio, entendiéndose por tal uno que garantice una probabilidad conocida, igual e independiente para cada uno de los elementos que conforman la población o universo.

3.c) La fijación de una probabilidad de error muestral, denominada estadísticamente como nivel de significación o probabilidad de error  $\alpha$  (alfa). Este valor es fijado por el investigador y en estudios sociales se emplean valores que oscilan entre el 1 y el 5%.

3.d) Una estimación de la heterogeneidad del universo, que en términos estadísticos es su varianza (o la raíz de la varianza, la desviación típica), cuando se trata de niveles de medición de intervalo y razón.

La varianza es un estadístico propio de un alto nivel de medición; sin embargo en la investigación social la mayoría de las variables son de bajo nivel, es decir son nominales u ordinales. En tales casos, para poder determinar la heterogeneidad de la variable se sustituye la varianza por la dicotomización de la "heterogeneidad" utilizando la distribución binomial; así una categoría relevante se convierte en "p", o probabilidad de éxito, y la otra, o las otras, se convierten en "q", teniendo como único requisito (Spiegel, 1978) el que N sea grande, de al menos 30 casos, y  $Np$  y  $Nq$  sean superiores a 5, para que haya aproximación a la curva normal (Véase nota final "5").

Para disminuir los costos del muestreo, se acepta un quinto factor para determinar el tamaño, un margen de error de estimación, una renuncia parcial a la precisión de las cifras estimadas.

Los dos primeros requisitos determinan la forma de seleccionar los casos, y los dos últimos determinan la **adecuación** del tamaño muestral, esto es el número de casos ( <sup>3</sup> ).

### 3.1.- Dificultades para la muestra probabilística en estudios sociales

Las muestras son de tal naturaleza que si no cumplen cualquiera de los cuatro requisitos

descritos, no son probabilísticas. Pero en los estudios sociales, de esos cuatro, sólo la fijación del nivel de significación resulta poco problemática, simplemente porque se determina a priori; su impacto tiene que ver más con el costo de trabajar con muestras de mayor o menor tamaño. Vemos, entonces, las dificultades de los tres restantes, respecto al esquema o marco muestral, el procedimiento de selección y la estimación de la varianza.

### **3.1.1.- El esquema muestral completo y actualizado**

El esquema muestral completo y actualizado no siempre es posible de conseguir para estudios sociales. Cuando no se tiene el listado completo de personas, funcionarios de la empresa, hogares, tenemos una limitante. En estricto sentido, el uso de mapas donde se detallan manzanas y calles, sin una identificación precisa de las viviendas y/u hogares, por ejemplo, no permite cumplir el requisito de la probabilidad conocida de selección o igual e independiente para todos los casos (Robinson:1981); simplemente, no se sabe el tamaño exacto de la población (N); Aun con la realización habitual de algunos resguardos, sólo podemos suponer un buen nivel de aproximación.

Enseguida, vemos algunos ejemplos de muestreo donde se adopta decisiones que afectan en una magnitud desconocida el carácter aleatorio de la selección de las unidades, estos procedimientos son mucho más frecuentes que lo que parece en muestras donde no se dispone de un listado actualizado de los casos.

Nadie puede negar la importancia y calidad de la Encuesta CASEN en Chile (<sup>4</sup>), sin embargo, para graficar la dificultad de este punto, en la CASEN III, 1994, se usó como información básica el Censo de 1982 (Schkolnik, 1995); para la CASEN IV, 1996, en cambio, se usó como esquema muestral el Censo 1992, y se actualizó en las 88 comunas con “crecimiento post-censal significativo de acuerdo a antecedentes proporcionados por el MINVU (Ministerio de la Vivienda y Urbanismo) (MIDEPLAN,1997); suponiendo con ello la adecuada actualidad de los datos del MINVU, sin embargo la actualización final, el empadronamiento en los sectores urbanos, se “limitó a un recorrido “exterior” del sector seleccionado” (Ibid, p.19). Tenemos, entonces, en el primer caso 12 años de rezago en con respecto al censo, en el segundo caso, tenemos 4 años, información que se procura actualizar con los datos del Ministerio, pero que como a la vez no ofrece suficientes garantías, se actualiza finalmente a ojo.

El segundo ejemplo: para un estudio de opinión de jóvenes de Valparaíso en 1992 (FLACSO,1992), se empleó como muestra base una muestra “confeccionada en 1987, por un equipo de FLACSO para una encuesta de opinión pública en las ciudades de Viña del Mar y Valparaíso...El procedimiento utilizado para actualizar la muestra consistió en dar instrucciones específicas a los encuestadores para los casos en que se encontraran en terreno con cambios significativos. En efecto se empadronaron las manzanas inmediatamente contiguas, siempre y cuando cumplieran con las mismas características socioeconómicas de la original” (Ibid, p.3), no se indica cuales fueron las instrucciones específicas para discriminar cambios “significativos” de los “no significativos”, ni los indicadores para dictaminar “cumplimiento de las mismas características socioeconómicas”. En el trabajo no se señala, tampoco, ningún indicador respecto a la congruencia de la muestra final usada con la muestra tomada como

referencia que se diseñó 5 años antes y para estudiar otras variables.

Estas limitantes no son muy distintas en otros estudios basados en muestras de áreas. En las consultas periódicas, más conocidas por su figuración en la prensa de actualidad, utilizando la misma muestra, no resuelve el punto, se cumple la afirmación de von Foester, no puede verse que no se ve lo que no se ve.

Corroborando la idea de la dificultad respecto de la aleatoriedad del muestreo de áreas, que corresponde a una de las aplicaciones más usadas del muestreo de clusters o conglomerados en los estudios sociales, Blua y Cavallini (CEPAL, 1977:10) afirman: “Para tamaños muestrales iguales los estimadores del muestreo por conglomerados son en la mayoría de los casos menos eficientes que los estimadores de la *msa* (muestra simple aleatoria). Esto se debe a que las unidades elementales dentro de los conglomerados tienden a parecerse más entre sí, con respecto a la variable en estudio, que con el conjunto de unidades que forman la población completa. Los principales efectos positivos del uso de conglomerados son el que permiten bajar los costos y el tiempo empleado en viajes, se pueden identificar mejor las unidades elementales, y a veces se pueden reducir los errores no muestrales debido a que se logra ejercer una inspección y un control más riguroso”.

No agregaremos las distorsiones derivadas de los casos de rechazo, aun manteniendo un número adicional de casos para sustituir.

### 3.1.2.- La estimación de la varianza

El problema mayor que tiene el diseño probabilístico deriva de la estimación de la varianza o heterogeneidad del universo. Con relación a la estimación de la varianza surgen dos dificultades relacionadas con los axiomas formulados. Como señala el Axioma 1, necesitamos conocer el universo para poder estudiarlo a través de muestras, y ese conocimiento está vinculado con las variables que distinguimos en él (Axioma 2), incluyendo la o las delimitantes.

Aquí hay dos problemas, no uno, como se da por supuesto en la bibliografía: cómo estimamos la varianza, y qué varianza estimamos.

Esto implica que, para diseñar cualquier muestra, incluyendo una probabilística necesitamos tener una medición anterior del universo (axioma 1), pues bien esa estimación de la heterogeneidad o varianza tiene que estar disponible para determinar el tamaño de la muestra que necesitamos. Pregunta: ¿cómo se obtuvo ese dato?: en algunos casos mediante un censo, una investigación anterior. Si la determinación o estimación del parámetro necesitado se obtuvo, está resuelto parte del problema. ¿Y si la investigación anterior no usó muestra probabilística?, si fue así, implica que es un dato sesgado, con margen de error desconocido. ¿Podemos con un parámetro sesgado y con sesgo desconocido asegurar que nuestra determinación del tamaño es estadísticamente válida?. Mientras, revisemos como lo resuelve uno de los escasos autores que intenta profundizar en el tema.

Cochran (1971, p.112-113) señala: “en la práctica hay 4 formas de estimar la varianza

de la población para la determinación del tamaño de la muestra: (1º) tomando la muestra en dos pasos, el primero a través de una muestra simple aleatoria de tamaño  $n$ , de la cual se obtienen los valores de  $S^2$  o  $P$  y el valor de  $n$  requerido, (2º) a través de los resultados de una encuesta piloto; (3º) por muestreo previo de la misma población o de una similar; y (4º) por conjetura sobre la estructura de la población, secundada por algunos resultados matemáticos". La dificultad está en que para la primera forma, la muestra simple aleatoria se necesita cumplir los cuatro requisitos indicados, e igualmente, las tres formas restantes son no-probabilísticas; además que en ninguna parte explica como hacer la "conjetura estructural" secundada por qué "resultados matemáticos" que sugiere para la cuarta forma ( 5).

Si no hay investigación anterior que ofrezca la estimación del parámetro, simplemente no hay muestra probabilística.

Pero aquí no terminan las dificultades, viene la segunda, justamente la menos mencionada. La homogeneidad del universo, no es una propiedad "dura" o evidente, depende de la variable que se considere, por lo tanto cuando hablamos de la varianza que necesitamos para estimar el tamaño de la muestra probabilística, tenemos que precisar varianza de cuál variable. Y, como en la conducta social normalmente son muchas - a diferencia del control de calidad de las tuercas en las plantas industriales donde puede importar solo su diámetro, o en el estudio de la aceleración de partículas en el laboratorio de física, donde sólo importa su masa y velocidad, por ejemplo - tenemos que (axioma 2) el universo tiene tantas varianzas como variables tomemos en cuenta en él.

Si tomamos cualquier fórmula de determinación del tamaño muestral veremos que en todos los casos se considera **una** medida de la heterogeneidad, normalmente  $\sigma$  ( 6 ) (sigma, es decir la desviación estándar de una distribución muestral de medias ) (Spiegel, 1978; Galtung, 1966; Padua, 1982; Blalock, 1994; Naciones Unidas; Manual de Encuestas sobre Hogares, 1964; Emory, 1980; Conway, 1967; Yates, 1960; Kish, 1965), es decir la heterogeneidad del universo respecto a una sola variable.

El que matemáticamente sea imposible incorporar una segunda o más varianzas no obvia la dificultad. El problema que esto representa, es que la muestra probabilística que resulta es probabilística y "representativa" *sólo con relación a esa variable y no a las otras*.

Repasemos esto último: aquí estamos de frente a uno de los puntos más críticos. Como se deduce del Axioma 2, el universo o población tiene tantas variables como se determine estudiar en él, y cada variable tiene su propia distribución, su propia varianza, y, como hemos dicho, en los estudios sobre o en poblaciones humanas, prácticamente nunca se estudia una sola variable.

¿Es posible denominar probabilística a una muestra que sólo lo es con relación a una de las variables que se estudia y que no lo es respecto de las otras?. La respuesta es no, aunque se haga así hasta la saciedad: las muestras probabilísticas, sólo son probabilísticas respecto a la variable que se uso para estimar la heterogeneidad del universo, y no lo son respecto de las otras variables. Cuando alguien dice que uso una muestra con un margen de error (muestral) de  $x\%$ , debemos entender que una de las variables que estudió gozó, según su declaración, del

privilegio de contar con muestra probabilística.

¿Puede la varianza de una sola variable asegurar el carácter probabilístico de las mediciones o estimaciones que se hagan de las otras o con las otras variables?. La respuesta es no, salvo que se elija para hacer la estimación aquella variable con varianza conocida que tenga el valor absoluto en dígitos significativos más alto. Es decir aquella variable del total de variables en estudio, que posee una varianza tal que arroja el mayor tamaño muestral.

Por ejemplo, la varianza del nivel de ingresos es \$7.500, en cambio la del número de hijos nacidos vivos por mujer en edad fértil es 4,3; pues, si usamos el número 7500 en vez de 4,3 (o la raíz, es decir  $\sigma$ ), haremos que nuestra muestra sea probabilística para el ingreso, y de paso cubra la probabilidad de la segunda variable, aunque aumentando significativamente el  $n$  muestral. Esto obligaría a tener una estimación aceptable de las varianzas de todas las variables que se incluirán en el estudio, o al menos de las más relevantes. No sólo haría mucho más difícil, sino también mucho mayores los costos del muestreo.

Si, como se indicó, no es posible incorporar matemáticamente más de una varianza en la determinación del tamaño ¿Cómo puede elegirse la única variable que se considerará para el diseño probabilístico?. Veamos algunos casos de como se ha resuelto esto, cómo se han elegido dichas variables.

La bibliografía teórica y técnica sobre muestreo no tiene una respuesta; como afirmamos en la presentación de este trabajo la omisión de estas dificultades que son propias de la especificidad del estudio de la conducta social, es casi generalizada en la literatura metodológica y estadística para la Educación, la Sociología, la Psicología y las ciencias sociales en general.

Una de las pocas referencias encontradas aparece en Cochran (1971, p.107-108): “Usualmente es medida más de una característica en una encuesta por muestreo: algunas veces el número de características es grande. Si se especifica el grado de precisión deseado para cada característica, los cálculos llevan a una serie de valores conflictivos de  $n$ . uno para cada característica. Se debe encontrar algún método para reconciliar estos valores”, el autor no sólo no sugiere ninguno, sino que no vuelve a referirse el punto.

En los trabajos empíricos la situación se corrobora. En las encuestas de opinión y sondeos electorales se entrega muy poca información sobre el diseño muestral utilizado, en algunos casos en que se acompaña una ficha técnica se indica las ciudades o regiones incluidas, se insiste en que se trata de muestra probabilística y se afirma un margen de error de X%, representativa de un Y% de la población, sin precisar más .

Veamos algunos pocos casos donde es posible acceder a esta información.

Para la CASEN III se empleó “una muestra probabilística” y se estimó la varianza sobre la base de P, probabilidad binomial (donde  $\sigma^2 = Np(1-p)$ ), sin especificar en modo alguno de qué variable se habla ni cuando se presenta la fórmula usada para determinar  $n$ , ni cuando “se realizó un análisis de la heterogeneidad de los universos que componían el total”

(Schkolnik, 1995). En la CASEN IV (MIDEPLAN, 1997) se afirma también la utilización de una muestra probabilística estratificada, por conglomerados y polietápica (p.12), más adelante se declara que “se procuró estimar el tamaño de la muestra, tratando de calcular varianzas..” sin especificar, tampoco, de qué variable se habla, y cuando se menciona una variable, en el acápite “18 Estimadores Propuestos para la CASEN (1996)”, se afirma: “Para hacer más concreta la exposición, se ha tomado como ejemplo el calculo del ingreso medio estimado por vivienda” (p. 24). No hay en el texto otra especificación respecto de este punto.

Des Raj en su Teoría del Muestreo (1992), anuncia en las primeras páginas que, dadas la incertidumbre e imprecisión de otras formas de muestreo. “a partir de este momento, en este libro sólo tendremos en cuenta procedimientos de muestreo probabilísticos..”, y en toda su exposición las variables serán tratadas como “y” o “x”, sin hacer ninguna indicación respecto de cómo debe determinarse cual será la variable “y” o “x” ( 7 ). En el Manual de Encuestas sobre Hogares de las Naciones Unidas (1964), sin intentar alguna fundamentación teórica, propone - a través de un ejemplo - como variable para la estimación del tamaño muestral, el tamaño medio de los hogares. Se agrega como único argumento que justifique dicha elección el que en “la mayoría de los países se conoce el tamaño medio del hogar”, luego agrega menciones de encuestas a hogares en la antigua Rhodesia del Sur, y Suecia, donde por distintas razones no se utilizó el procedimiento estadístico que propone.

En un estudio de casos el Bureau Of Census (1968, p.4), realiza unas “consideraciones matemáticas del diseño de una muestra” y opta por una simplificación extrema, convirtiendo en discreta y binomial, p y q, la variable mas simple que encontró, por supuesto sin ninguna justificación teórica: “consideramos la estimación de una característica que posee alrededor de la mitad de la población; por ejemplo, el hecho de tener 12 años de edad o más”, entonces se funciona con la p máxima 0,5 (ver nota final N° 5).

Hasta aquí hemos visto las dificultades para trabajar con muestras probabilísticas en sentido estricto, aquellas que cumplan realmente los requisitos enunciados; hemos visto las dificultades para ello en el diseño muestral de la encuesta CASEN, por ejemplo; **pero no hemos afirmado que los datos obtenidos en ese mismo estudio carezcan de representatividad y validez.**

¿Cómo? ¿”Sin embargo se mueve”?; similar a un juego de cálculo silogístico, la consecuencia es verdadera aunque la premisa sea falsa. La explicación de esta aparente paradoja esta en las consecuencias de los cuatro principios del muestro con que iniciamos este escrito, y que se encuentran implícitos en el razonamiento que sustenta el muestro no-probabilístico.

#### **4.- Muestras no- probabilísticas en la investigación social**

En la literatura se acostumbra a describir dos tipos de muestras (Padua, 1982; Blalock, 1994; Raj, 1992), las probabilísticas: al azar simple, estratificada, sistemática, y de conglomerados o clusters. Y las no probabilísticas: intencional, casual, y por cuotas.

Como hemos señalado el carácter probabilístico de una muestra depende directamente

del cumplimiento de cada uno de los cuatro requisitos descritos en el acápite "2", por lo tanto estos "tipos" de muestras corresponden más propiamente a formas de selección de los casos.

De los tipos probabilísticos, el único que cumple estrictamente los requisitos descritos, es la muestra al azar simple, los otros corresponden a procedimientos que procuran acercarse a la selección aleatoria de las unidades, y justamente se emplean cuando un procedimiento aleatorio puro no es posible por falta de información o alto costo.

A modo de ejemplo, la muestra estratificada se justifica con la necesidad de asegurar la representatividad de una variable, en rigor una segunda variable que actúa como control. Esta estratificación además de suponer una estimación insesgada del peso de sus categorías en el universo, constituye una intervención no aleatoria por parte del investigador sobre la selección de los casos. No obstante, adicionalmente, si la varianza de la variable en base a la cual se estratifica determina un tamaño muestral menor que el que se establece con la primera variable, la primera asegura su representatividad; si fuese al revés, la variable para estratificar asegura a la primera; por lo tanto en ninguno de los dos casos - si efectivamente la selección de los casos fuese aleatoria - sería necesaria la estratificación.

Hasta aquí nos hemos encontrado con varias paradojas: las más usadas muestras probabilísticas se valen de procedimientos no aleatorios para seleccionar al azar sus unidades; para estimar la varianza que nos permita la adecuación de muestras probabilísticas podemos usar muestras no-probabilísticas; adecuamos el tamaño a la varianza de una sola variable sin embargo, denominamos probabilística a toda la muestra independientemente de la cantidad de variables que estudiemos.

La respuesta está justamente en los cuatro principios del muestreo con que iniciamos este trabajo; y son ellos justamente los que directamente nos permiten fundamentar la validez no solo de la intromisión de estos procedimientos no estocásticos aceptados en los diseños muestrales denominados probabilísticos, sino también de la validez de los diseños no probabilísticos.

Así, procurando abreviar ya esta exposición, en la muestra intencional es el investigador el que en base a su conocimiento del universo selecciona aquellos casos que cumplen una o más condiciones que necesita: líderes de opinión, sujetos o informantes claves, vecinos según localidades, clientes, etc. Si el conocimiento que el investigador tiene del universo es adecuado (Axioma 1), si los casos reúnen las características que ha definido previamente (Axioma 2), si toma un número grande de casos (Ley), los resultados tendrán una mayor probabilidad de acercarse al parámetro que de alejarse (Teorema); y podrá legítimamente tomar decisiones con los resultados aunque los casos no hayan sido seleccionados al azar, aunque no este en condiciones de indicar cuanto error muestral o nivel de significación contienen sus datos, cumplirá todos los requisitos para generalizar los resultados.

De lo indicado podemos deducir que en los estudios sociales hay más simulación de trabajo con muestras probabilísticas, que trabajos que realmente se ajusten a lo que estadísticamente se denomina con propiedad así. Los niveles de significación con que

habitualmente se presentan los datos procesados computacionalmente tienen muy poca validez<sup>(8)</sup> .

Pasado el siglo XIX y cerrando ya las puertas del XX, el fantasma fiscalista del viejo Comte aun consigue hacernos olvidar que entre la conducta humana y las partículas atómicas hay más de una diferencia.

## 5.- Referencias.

Azorín, F. y Sánchez-Crespo, J.L. (1986). **Métodos y Aplicaciones del Muestreo**. Madrid: Alianza Editorial.

Blalock, H. (1994). **Estadística Social**. México: Fondo de Cultura Económica.

Briones, G. (1990). **Métodos y Técnicas de Investigación para las Ciencias Sociales**. México: Editorial Trillas.

Bureau of the Census (1968). **Atlántida, un estudio de casos en encuestas de hogares por muestra. Unidad IV. Diseño de la Muestra**. Washington, DC: Agencia para el Desarrollo Internacional. Depto. de Estado U.S. Department of Commerce.

CEPAL(1977). **Diseño Muestral y Trabajo Operativo de una Investigación Piloto para Evaluar 3 Cuestionarios sobre Ingresos**. Elsa Blua y Carlos Cavallini. CEPAL/BORRADOR/EST/158.

Cochran, W. (1971). **Técnicas de Muestreo**. México: CECSA (Compañía Editorial Continental, S.A.).

Cohen, J. (1988). **Statistical power analysis for the behavioral sciences**. (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Conway, F. (1967). **Sampling. An introduction for social scientists**. London: George Allen & Unwin Ltd.

Daniel, W.W. (1985). **Estadística aplicada a las Ciencias Sociales y la Educación**. México: McGraw-Hill.

Emory, W.C. (1980). **Business Research Methods**. Homewood, Illinois: Richard D. Irwin, Inc.

FLACSO (1992). **Los Jóvenes en la Comuna de Valparaíso. Percepciones y opiniones. Informe de Encuesta**. FLACSO-U. DE PLAYA ANCHA.

Galtung, J. (1966). **Teoría y Métodos de Investigación Social**. Buenos Aires: EUDEBA.

**Internet Glossary of Statistical Terms**, The Animated Software Company: (March 27th,

1997): <<http://www.animatedsoftware.com>>.

Kish, L. (1965). **Survey Sampling**. New York: John Wiley & Sons, Inc.

Naciones Unidas (1964). **Manual de Encuestas sobre Hogares. Estudios de Métodos**. Nueva York.

MIDEPLAN (1997). **Casen 1996. Diseño de la Muestra. 4ta. versión**. Santiago de Chile: MIDEPLAN. UNIVERSIDAD DE CHILE. CIENES/OEA.

Padua, J. (1982). **Técnicas de Investigación Aplicadas a las Ciencias Sociales**. México: Fondo de Cultura Económica.

Raj, D. (1992). **Teoría del Muestreo**. México: Fondo de Cultura Económica.

Robinson, E.A.(1981). **Statistical Reasoning and Decision Making**. Houston, TX: Goose Pond Press.

Schkolnik, M. (1995). Chile: Un instrumento para la evaluación del impacto de los programas sociales. Encuesta de Caracterización Socioeconómica Nacional (CASEN). Anexo IV: Diseño de la muestra utilizada en la encuesta sobre políticas sociales en Chile. En: CEPAL: **Focalización de la Pobreza**. Cuadernos de la CEPAL, Naciones Unidas; Santiago de Chile.

Spiegel, M.R. (1978). **Teoría y Problemas de Estadística**. Cali: Libros McGraw-Hill.

Yates, F. (1960). **Sampling Methods for Censuses and Surveys**. London: Charles Griffin & Company Limited.

## 6.- Notas.

1) Formulaciones matemáticas de la Ley y el Teorema aparecen en: Raj, Des: Teoría del Muestreo, pág.32. Otra formulación disponible en la WWW, del Internet Glossary of Statistical Terms, se encuentra en:

<http://www.animatedsoftware.com/statglos/sgcltheo.htm>

2) No emplearemos aquí la denominación de muestras representativas y no representativas, simplemente por que, como vemos, la representatividad de la muestra no depende sólo del conocimiento de la probabilidad de error muestral, lo cual implica que muestras no probabilísticas pueden ser muy representativas. P.e: "con un muestreo carente de probabilidad, podemos acaso tener una muestra efectivamente muy representativa, pero no estaremos en condición de apreciar los riesgos de error implicados", Blalock; 1994:531.

3) El análisis de potencia no modifica la lógica de lo señalado. El análisis de potencia y el efecto del tamaño (power analysis y size effects (Cohen;1988)), que usados en test de hipótesis estadísticas, permiten determinar la probabilidad de rechazar la hipótesis nula cuando la alternativa es verdadera. La potencia de una muestra es así función de cuatro factores: función

directa con el tamaño muestral, la probabilidad de error tipo I o alpha, y función inversa con las diferencias esperadas en el parámetro que se estima y la varianza.

4) Encuesta Nacional de Caracterización Socio-económica, que permite evaluar políticas fiscales, distribución del ingreso etc. se aplica cada dos años en Chile.

5) Cuando se trabaja con variables no métricas, como señalamos, se utilizan las proporciones o porcentajes, con ello este problema no se reduce: la necesidad de dicotomizar (para convertir en binomiales las categorías (p y 1-p), no sólo hace una reducción forzada de las categorías (p.e: cristianos, budistas, moros, bahais, etc. se convierten en cristianos y no-cristianos), si no que además, a mayor diferencia entre p y q, menor será el tamaño muestral por lo tanto menor será la probabilidad de representar adecuadamente la, o las, categorías más bajas. Esto explica, por ejemplo, el que en las encuestas electorales, la votación de los partidos pequeños sea más difícil de predecir.

6) A modo de ejemplo, la fórmula para la determinación del tamaño en la muestra al azar simple (Spiegel:1978:161), es:

$$\frac{z \sigma^2}{\sqrt{N}} \leq e$$

donde  $z = \alpha$  (el puntaje z correspondiente al nivel de significación)

$\sigma$  = varianza de la variable elegida para determinar el tamaño

$N$  = el tamaño muestral, la incógnita

$e$  = error tolerado de la estimación

7) Véase en la obra citada, en el “Apéndice 1 Informe Sobre una Encuesta por Muestreo Real” el autor demuestra por omisión, o, contrariando su propia profesión de fe probabilística (pág. 41), que en las situaciones reales de muestreo en estudios sociales se deben tomar tantas decisiones no matemáticas que el concepto de muestra probabilística se hace, al menos, discutible.

8) En los paquetes estadísticos para procesamiento computacional de datos (SPSS, SAS, SYSTAT), los coeficientes de asociación o correlación obtenidos, presentan normalmente el nivel de significación para el cual son representativos, independientemente que la muestra sea o no sea probabilística, incluso independientemente del número de casos.

En los algoritmos de procesamiento de estos programas se asume que la muestra es probabilística, que los casos se seleccionaron al azar, y que la distribución de las variables es normal en el universo.

Con SPSS, a modo de ejemplo, basta con correlacionar consigo misma la serie de números del 1 al 5, para obtener r, r de Pearson, CC, etc. máximos con significancia de 1/1000 y más.

No podría incorporarse a la estructura de programas de este tipo la capacidad de discriminar

al respecto; teóricamente para que eso sea posible se necesitaría al menos que el software de proceso estadístico esté vinculado a la base de datos conteniendo el marco muestral completo y actualizado y la distribución de las variables que se estudiará. Es decir el modelo sería isomórfico con el sistema, no sería menester muestrear nada.

Evidentemente, lo indicado le resta validez a los valores de significación obtenidos con estos programas.