

NOCIONES DE MUESTREO

Cualquier investigación empírica versará sobre sistemas reales situados en el espacio y en el tiempo. El estudio de un objeto determinado es algo que puede demandar la definición de uno o varios universos. El fin que se persiga puede ser eminentemente práctico –describir qué piensan los obreros de una fábrica o cuáles son las intenciones de voto de los habitantes de una provincia, u obtener una estimación de la tasa de desempleo en una ciudad, etc.- o bien responder a una motivación más estrictamente científica –como cuando se trata de fundar hipótesis de un mayor grado de generalidad.

La generalidad es algo que se puede entender por lo menos en dos sentidos. Zetterberg distingue entre proposiciones teóricas y comunes basándose en cuál sea su “valor informativo”¹. Si se dice “A mayor educación, mayor ingreso monetario”, ésta puede ser una hipótesis de un nivel de generalidad muy elevado en el sentido de que, potencialmente, es aplicable a cualquier sociedad en la que se reconozca la existencia de algo llamado “educación” y en la que funcione alguna forma de moneda para regular el acceso de sus miembros a bienes y servicios. Sin embargo, no es ésta una hipótesis de alto valor informativo en cuanto a la variedad de fenómenos de los que da cuenta.

El problema del alcance de una hipótesis es entonces doble: por una parte, tiene que ver con sus referentes concretos –con la cantidad de sistemas o UUAA comprendidos dentro de su dominio-; por la otra, con el grado de abstracción de sus conceptos de propiedad respecto a conceptos más específicos –con la cantidad, por lo tanto, de conceptos de propiedad que subsumen. Así, cuando se dice “Toda persona tiende a realizar aquellas acciones que ayudan a mantener invariables las evaluaciones que recibe de sus iguales” se trata de una proposición de un alto nivel de generalidad en este segundo sentido. En efecto, como lo señala Zetterberg, de ésta pueden deducirse varias proposiciones más específicas; por ejemplo, los políticos intentarán mantener su popularidad, los investigadores elegirán temas que les permitan conservar su prestigio como miembros de una comunidad científica, etc. Un concepto tan abstracto como el de “evaluaciones” estará subsumiendo varios conceptos de menor nivel (“aprobación”, “estima”, “rango”); por lo que la mera sustitución de dicho concepto en la proposición más abstracta, permitirá generar toda una serie de hipótesis específicas.

Si se admite que el desarrollo del conocimiento científico supone tender hacia la formulación de hipótesis “generales abstractas”, se entenderá que tanto el grado de generalidad como el de “abstracción” son relevantes en cuanto a determinar el alcance de una hipótesis cualquiera. Así como era posible partiendo de la idea de “niveles de inclusión” determinar superunidades (“colectivos”), así también podría aplicarse la idea de inclusión a los conceptos de propiedad, obteniendo superconceptos.

El problema clásico de la inducción deviene de la imposibilidad lógica de sustentar la verdad de hipótesis generales a partir de un número de observaciones limitado y siempre finito. Ahora bien esta limitación del “número de observaciones” se refiere tanto a las UUAA como a las variables: ni estudiamos todos los elementos de un conjunto de UUAA, ni de todas las UUAA observadas consideramos todas las propiedades. Desde este punto de vista la técnica del muestreo podría considerarse como una alternativa estadística para superar el problema de la inducción. Empero esta técnica es más apta para resolver el problema en cuanto al número de las UUAA. Las razones por las que no se puede aplicar con el mismo éxito a las variables son varias; mencionemos simplemente la que deviene de la dificultad en delimitar un universo de variables. Por ende, nos

limitaremos a plantear el problema del muestreo en lo referido al problema de la selección de un cierto número de UUAA. Abordaremos luego la cuestión del tamaño de la muestra, y finalmente describiremos los diferentes tipos de muestras que se utilizan más corrientemente.

1. ELEMENTOS BÁSICOS DE MUESTREO

Ya hemos introducido el concepto de “universo” para referirnos a un conjunto de UUAA definidas por presentar valores constantes en un conjunto de variables. Indiferentemente puede utilizarse este término o el de “**población**” para denotar “el conjunto de todos los casos que concuerdan con una serie determinada de especificaciones”². Así, cuando se habla comúnmente de “la población de la Argentina”, se está haciendo referencia a un conjunto de UUAA –“personas”- que en un momento dado (por ejemplo, el 30 de setiembre de 1980), en la variable “país de residencia” presentan el valor “Argentina”. En la jerga estadística, una población se compone de “elementos”: así, cada persona habitante de la Argentina es un **elemento** de dicha población. Contrariamente al uso corriente del término “población”, los elementos no tienen por qué ser personas. Podrá tratarse también de grupos, acciones, productos o cualquier tipo de objetos o sistemas que se convenga en definir como tales: cada uno de los avisos transmitidos por un canal de TV durante una semana, de los informes incluidos en un archivo, de los desplazamientos turísticos realizados por los habitantes de región durante un período dado, de las relaciones existentes entre los miembros de una comunidad, etc., pueden ser considerados elementos de las correspondientes poblaciones. En cualquier caso una población incluye **todos** los elementos que satisfacen las propiedades que la definen: todos los obreros de una fábrica, todos los departamentos o partidos en que se dividen las provincias argentinas, todos los municipios de Misiones, todas las noticias sobre suicidios de jubilados publicadas en diarios de Buenos Aires durante 1992, etc.

Desde este punto de vista uno de los primeros pasos en cualquier investigación empírica consiste en la determinación del universo de la investigación. Así, por ejemplo, en un trabajo realizado para la Entidad Binacional Yacuyetá,³ el universo del estudio fue definido como constituido por unidades domésticas que presentaran las siguientes características: pertenecer a la ciudad de Posadas, estar ubicadas por debajo de la cota 85, dentro de la Zona 1 de relocalización y cuyos miembros no fueran propietarios del terreno en que se asentaban.

Así como distinguíamos sub-universos dentro de un universo, así es posible definir poblaciones que se encuentran incluidas dentro de poblaciones mayores. En el contexto de la teoría del muestreo tales sub-poblaciones reciben el nombre de “estratos de la población”, o simplemente “estratos”. Un **estrato** se define a partir de una o más variables que permiten dividir a una población en sus conjuntos mutuamente excluyentes. Dentro de la población de los electores de Posadas la variable “sexo” permitirá definir el estrato de los varones y el de las mujeres; dentro del universo de los establecimientos industriales de la Argentina se puede distinguir estratos de acuerdo con la variable “número de empleados”: los que ocupan hasta 5 empleados, de 6 a 50 empleados, y de más de 50 empleados, por ejemplo. Combinando las variables “sexo” y “condición de alfabetismo”, la población de los votantes de Posadas el 14 de mayo de 1989, podría dividirse en cuatro estratos, varones alfabetos, varones analfabetos, mujeres alfabetas y mujeres analfabetas, etc.

Supongamos que deseamos conocer qué proporción de una población presenta un determinado valor en una variable dada. Si la población es pequeña, ello no ofrece mayores problemas. Por ejemplo, para conocer el porcentaje de mujeres entre los

alumnos presentes en un aula, se puede rápidamente contar cuántos alumnos hay y cuántos son de sexo femenino. Se denomina **censo** a este procedimiento, que consiste en “un recuento de todos los elementos en una población y/o una especificación de las distribuciones de sus características, basados en la información obtenida para cada uno de los elementos”.⁴

Si por lo contrario, la población que nos interesa es de gran tamaño, por ejemplo, los habitantes de Buenos Aires en 1993, podría resultar sumamente trabajoso obtener información sobre todos sus elementos. En este caso convendrá utilizar una **muestra**, vale decir un subconjunto de elementos de una población seleccionados para averiguar algo sobre el conjunto total de elementos que constituye esa población.⁵ Cuando se utiliza una muestra, se lo hace en base a la creencia de que la exactitud de los resultados obtenidos de ese modo es lo más próxima posible a la que se hubiera obtenido realizando un censo desprovisto de todo tipo de error sobre la población. Por supuesto, tal creencia constituye una hipótesis cuya verdad sólo sería demostrable realizando paralelamente un censo sobre la misma población.⁶ Por una parte existen los **parámetros** poblacionales, es decir los valores “verdaderos” que caracterizan las distribuciones de variables en la población: por ejemplo, la media de ingresos de todos los jefes de unidades domésticas de Buenos Aires, o el porcentaje de todos los electores con intenciones de votar por el Justicialismo en las próximas elecciones; dichos parámetros son los valores que se obtendrían en una medición de todos los elementos desprovista de cualquier tipo de error. Por la otra, se obtienen **valores muestrales**, simples **estimadores** de aquellos parámetros. Lo cierto es que nunca resulta posible asegurar la coincidencia entre los valores muestrales y los parámetros; a lo sumo, la forma de selección de la muestra podrá maximizar nuestra creencia en dicha coincidencia.

El problema del muestreo consiste entonces en cómo seleccionar una muestra de modo tal de obtener la máxima aproximación a los parámetros poblacionales compatible con las restricciones de costo y de tiempo existentes. En efecto, si se dispusiera de una cantidad ilimitada de recursos, lo que cabría sería recurrir a un censo. Pero lo cierto es que las más de las veces no es ésta la situación en que se encuentra el investigador.

A menudo se habla de la “representatividad” de la muestra. La idea de representatividad tiene que ver con la posibilidad de que la muestra sea una réplica adecuada de la población en lo que hace a las variables relevantes. Al respecto, convendrá tener en cuenta las consideraciones siguientes.

1. Una muestra en particular no es representativa en abstracto; una muestra que sea representativa para determinados propósitos, puede no serlo para otros: no existe la muestra representativa en sí, para cualquier propósito.
2. Si bien existen procedimientos para evaluar la bondad de una muestra, no es la muestra en sí la que es representativa; más bien es nuestra creencia en su representatividad la que va a depender del plan de muestreo utilizado para seleccionarla. Por “**plan de muestreo**” se entiende el diseño de un conjunto ordenado de operaciones que conduce a la selección de un número determinado de unidades de análisis a partir de un universo determinado.

2. EL TAMAÑO DE LA MUESTRA

Sin duda, determinar el tamaño adecuado para una muestra es una elección crucial, por todas las consecuencias negativas que puede producir una equivocación. No obstante ello, no es infrecuente que el tamaño de la muestra se determine en función de los recursos disponibles. Está claro que si la muestra es demasiado grande, se habrá realizado una inversión inútil, un derroche; en cambio, si resulta exigua, ocurrirá que no servirá a los propósitos de la investigación.⁷ Lamentablemente, y contrariamente a

opiniones corrientes, no existe ninguna regla áurea del tipo “tomar un 10% de la población” para asegurarse el éxito. En algunos casos 10% resultará lastimosamente insuficiente, mientras que en otros será un total exceso.

Atendiendo a una distinción de Galtung, todas las muestras pueden clasificarse en dos grandes categorías según el tipo de hipótesis que se pretende poner a prueba. En efecto puede tratarse tanto de **hipótesis de generalización** –cuando a partir de los datos muestrales se pretende inferir, por ejemplo, el valor de un parámetro en una población dada-, o bien de **hipótesis sustantivas** –cuando se desea comprobar la existencia de determinada relación entre ciertas variables. En el primer caso los criterios pertinentes a tener en cuenta serán los habituales para cualquier tipo de inferencia estadística; en el segundo, pueden resumirse en unas breves reglas muy sencillas de comprender.

2.1. Muestras para verificar hipótesis de generalización.

Existen procedimientos estadísticos que nos permiten estimar la probabilidad de que un determinado valor muestral no difiera sustancialmente del parámetro que se hubiera obtenido de haber realizado un censo. Así, si se desea averiguar qué porcentaje de los elementos de una población presenta el atributo X, es posible adoptar un plan de muestreo tal que nos garantice, por ejemplo, con un **nivel de confianza** del 95% que el parámetro no se apartará más de un 2% (**margen de error**) del valor muestral que obtengamos. Si requiriéramos mayor seguridad y precisión elegiríamos un nivel de confianza mayor (99%, por ejemplo) y un margen de error menor (1%); empero, cuanto más alto sea nuestro nivel de aspiraciones, mayor deberá ser el tamaño de la muestra, lo que supondrá un costo más elevado, puesto que deberemos seleccionar y recolectar datos sobre mayor cantidad de unidades. A los efectos prácticos un plan de muestreo con un nivel de confianza del 99% significa que, siguiendo ese plan de muestreo, 99 veces de cada 100 no nos equivocaremos en nuestra estimación; empero, no existe nunca garantía alguna de que la muestra particular que haya resultado seleccionada no pertenezca a aquel conjunto del 1% de las muestras que nos conducirían a error.

Tabla A: Tamaños de muestras al azar simple requeridos para niveles de confianza del 99,7% y del 95,5%, según valores presumibles de p y q y límites de error.

Nivel de confianza de 99.7% (3 Ó)					
Límites de error (+/- %)	Valores presumibles de p y q en %				
	10/90	20/80	30/70	40/60	50/50
1,0	8.100	14.400	18.900	21.600	22.500
2,0	2.025	3.600	4.825	5.400	5.627
3,0	900	1.600	2.100	2.400	2.500
4,0	506	900	1.181	1.350	1.406
5,0	324	576	756	864	900
10,0	81	144	189	216	225
20,0	20	36	47	54	56

Nivel de confianza de 95,5% (2 Ó)					
Límites de error (+/- %)	Valores presumibles de p y q en %				
	10/90	20/80	30/70	40/60	50/50
1,0	3.600	6.400	8.400	9.600	10.000
2,0	900	1.600	2.100	2.400	2.500
3,0	400	711	933	1.067	1.111
4,0	225	400	525	600	625
5,0	144	256	336	384	400
10,0	36	64	83	96	100
20,0	9	16	21	24	25

Fuente: Sierra Bravo, 1979: 181–182.

La Tabla A ilustra cómo el tamaño requerido para una muestra es una función directa de tres factores.

1. El **nivel de confianza** requerido: la tabla está dividida en dos mitades. En las celdas del sector superior se presentan los tamaños de muestra requeridos para un nivel de confianza del 99,7%; en las del sector inferior los tamaños demandados por un menor nivel: 95,5%. Puede observarse cómo en el sector inferior los números son menores a sus análogos del sector superior.
2. El **margen de error**: en la primera columna se indican distintos márgenes de error para cada una de las hileras de la tabla. El menor límites de error incluido es de +/- 1%; vale decir que si con ese límite producimos una estimación de que un 40% de las unidades observadas en la muestra poseen el atributo, podremos inferir que la proporción en la población estará ubicada entre el 39 y el 41%. En ambos sectores de la tabla de abajo hacia arriba crecen los tamaños requeridos: cuanto menor sea el margen de error que deseemos, mayor deberá ser el tamaño de la muestra.
3. La **variabilidad** del atributo investigado: cada columna corresponde a diferentes distribuciones del atributo en el universo, que van desde la distribución 10/90 (10% poseen el atributo, 90% no lo poseen) hasta 50/50 (una mitad posee el atributo, la otra no). En cada sector los tamaños aumentan a medida que nos desplazamos de izquierda a derecha: cuanto mayor es la variabilidad de la característica,⁸ mayor es el tamaño de muestra necesario. Así por ejemplo, mientras que para un nivel de confianza del 95,5 y un margen de error de +/- 2%, 900 casos serían suficientes si el atributo se distribuyera en las proporciones 10/90, en el caso de que la distribución en el universo fuera 50/50 la muestra requerida sería de 2.500 unidades.⁹ Este último factor supone la necesidad de anticiparse a la distribución del atributo en el universo; y de no existir elementos que permitan sustentar otra alternativa, convendrá conservadoramente suponer la distribución más desfavorable: 50/50.

En síntesis, cuanto mayores sean nuestras exigencias respecto al grado de confiabilidad y de precisión de nuestros resultados, mayor habrá de ser el tamaño de la muestra. La tabla permite también apreciar cómo juegan aquí rendimientos decrecientes. Para un atributo cuya distribución en el universo es de 20/80, 256 casos alcanzan para determinar el valor del parámetro con un nivel de 95,5 y un margen de error de +/- 5%; si se lleva la muestra a un número de 711 unidades, el margen de error se reduce en dos puntos, pasando a ser de +/- 3%. Pero, para obtener una nueva reducción de dos puntos y alcanzar un margen de +/- 1%, el aumento en el tamaño deberá ser mucho mayor, requiriéndose ya 6.400 casos.

2.2. Muestras para someter a prueba hipótesis sustantivas.

Si el objetivo de nuestra investigación es primordialmente analítico antes que descriptivo, no estaremos tan interesados en la generalización como en la simple comprobación de la existencia de relaciones específicas entre las variables. Desde este punto de vista, el tamaño no se basará en los requerimientos demandados para producir inferencias estadísticas, sino que deberá ser tal que nos permita determinar la existencia de relaciones para el conjunto de las unidades incluidas en la muestra.

Si por ejemplo queremos estudiar la relación entre dos variables dicotómicas, ello demandará producir una tabla de contingencia en la que las unidades se clasifiquen simultáneamente en ambas variables.¹⁰ Una tabla tal consta de cuatro celdas, y para aplicar una simple diferencia porcentual difícilmente nos contentemos con un promedio de casos por celda inferior a 10. Vale decir que **con menos de 40 casos no es posible analizar ninguna relación entre variables**; incluso, para que la tendencia sea clara, preferiremos sin duda contar con un promedio de 20 casos por celda: aún así el simple cambio de un caso de una columna o de una hilera a otra producirá una diferencia porcentual del 2,5%.

En esta perspectiva, las preguntas a formularse para determinar el tamaño de la muestra tienen que ver con el número de variables que se pretende analizar simultáneamente y con el número de valores de cada una de esas variables. Galtung presenta la siguiente tabla:

Tabla B: Número mínimo de unidades de análisis para un promedio de 10 casos por celda (20 casos entre paréntesis)

Nº de variables por tabla	Nº de valores por variable		
	2	3	4
2	40 (80)	90 (180)	160 (320)
3	80 (160)	270 (540)	640 (1.280)
4	160 (320)	810 (1.620)	2.560 (5.120)

Así, por ejemplo, para estudiar la relación entre tres variables tricotómicas se requerirá un mínimo de 270 casos para contar con un promedio de 10 observaciones por celda (ó 540, si pretendemos tener 20 casos en cada combinación). Obviamente no es indispensable que todas las variables tengan el mismo número de valores, ni que deban limitarse a cuatro valores, lo que puede dar lugar a otros tamaños de muestra. Lo más práctico es entonces multiplicar el número de casos por celda por el número de valores de cada una de las variables que se desee analizar simultáneamente; si se quisiera investigar cómo se relacionan entre sí dos variables dicotómicas y una tercera de cinco valores, el número mínimo de casos requerido para contar con un promedio de 10 observaciones por celda sería $(10) \times 2 \times 2 \times 5 = 200$. Por supuesto, en cualquier investigación es posible pretender analizar la relación entre varios conjuntos de variables, por lo que la pregunta que debe responderse para determinar el tamaño de la muestra es cuál es la tabla más grande que se pretende generar.

En las muestras para hipótesis sustantivas es posible de todos modos aplicar determinados test de hipótesis como la de Student aplicada a la diferencia de medias o de proporciones, o la prueba de χ^2 . En estos casos, puede interpretarse que la hipótesis sustantiva pasa a funcionar también como una hipótesis de generalización, aunque no ya referida a una población finita y concretamente determinada, sino con relación a un universo teórico, sin límites definidos, que abarque las unidades de todo tiempo y lugar a las que sea posible aplicar las variables.

3. PRINCIPALES TIPOS DE MUESTRAS

No es nuestro propósito reseñar aquí todos los tipos posibles de muestras. Tampoco nos proponemos abordar este tema en una perspectiva estadística; ante cualquier investigación de envergadura, convendrá recurrir a algún especialista en la cuestión.¹¹ Nuestra intención es sólo describir las características de algunos diseños muestrales desde un punto de vista conceptual, para exponer algunos problemas típicos que se plantean con relación al uso de esta técnica. Desde el punto de vista de la teoría del muestreo, existen solamente dos grandes categorías de muestras. Por un lado están las muestras probabilísticas –o “al azar”–, que se caracterizan porque permiten especificar la probabilidad que tiene cada elemento de la población de resultar incluido en la muestra. El caso más sencillo es cuando todos los elementos tienen la misma probabilidad de integrar la muestra, pero no es éste el único tipo de muestreo de probabilidad. Por otra parte, hay muestras no-probabilísticas, en las que no existe forma de determinar la probabilidad que tiene cada elemento de ser seleccionado, y ni siquiera la seguridad de que cada elemento tenga alguna probabilidad de ser seleccionado.

Esta distinción es fundamental, puesto que solamente los planes de muestreo probabilísticos pueden ser considerados representativos en un sentido estadístico. El muestreo probabilístico es el único que permite basarse en la teoría de las probabilidades para estimar estadísticamente el grado en que los valores muestrales pueden tender a diferir de los parámetros; una muestra probabilística permite especificar el tamaño de muestra requerido para un nivel de confiabilidad y un margen de error determinados.

3.1. Muestras no-probabilísticas

En el muestreo de no-probabilidad no hay modo alguno de evaluar estadísticamente los resultados obtenidos a partir de la muestra. No obstante ello, algunas muestras de este tipo se utilizan mucho, en razón de las ventajas comparativas que exhiben en cuanto a su comodidad y bajo costo.¹²

Muestras accidentales (o “casuales”): consisten simplemente en tomar los casos que “caen bajo la mano” continuando el proceso hasta que se alcanza un cierto tamaño de la muestra. Por ejemplo, se incluirán en la muestra las primeras 100 personas que pasen por una esquina y que sean deseosas de ser entrevistadas. Si el universo es la población de una ciudad, es claro que no será ésta una muestra representativa: no todos los elementos de dicha población tienen la misma oportunidad de pasar por dicha esquina, ni menos aún coincidirán en la probabilidad de pasar por allí en el momento en que se realicen las entrevistas. Pero en este caso, no existe modo alguno de evaluar los sesgos introducidos por el modo de elección de las unidades; sólo cabrá esperar que la equivocación no sea demasiado grande.

Muestras por cuotas: en el muestreo por cuotas se busca garantizar la selección de elementos pertenecientes a los diferentes estratos que conforman la población y que dichos elementos puedan ser tenidos en cuenta en las mismas proporciones en que ocurren en esa población. En todos los casos, la técnica supone sostener alguna hipótesis acerca de cuáles son las variables relevantes, las que serán utilizadas para dividir la muestra en estratos. Si, por ejemplo, se quiere conocer las intenciones de voto de una población, y se piensa que éstas variarán en función del sexo, convendrá estratificar la muestra en dos estratos: “masculino y femenino”. Así, a cada entrevistador se le dará la consigna de cumplir con una “cuota” de entrevistados pertenecientes a cada estrato: tantos varones y tantas mujeres. El muestreo por cuotas requiere anticiparse a las diferencias que puedan presentar los elementos de los distintos estratos respecto al valor

que exhibirán en la variable a investigar, de modo tal de seleccionar una muestra que “represente” adecuadamente a la población, esto es que funcione como una réplica de la población en cuanto a las variables que se juzgan relevantes. En lo que respecta al estudio de preferencias, opiniones, actitudes, etc., se conoce que variables tales como el sexo, la edad, o el nivel educativo tienden a producir diferencias significativas en esas características.

Ahora bien, puede suceder que la proporción de elementos de un estrato en la muestra no sea idéntica a la existente en la población. Imaginemos que se encuesta a 100 personas obteniendo una cierta distribución¹³ por nivel educativo, la que podemos comparar con la existente en el universo:

Tabla 3.1: comparación de la distribución de la variable “nivel educativo” en la muestra y en la población

Nivel Educativo	% muestra	% población	Diferencia porcentual
Alto	30	15	+15
Medio	45	30	+15
Bajo	25	55	-30
Total	100	100	

La comparación muestra claramente que en la muestra los niveles alto y medio se encuentran sobrerrepresentados, mientras que el nivel bajo está subrepresentado con relación a la población.

Tabla 3.2.: Distribución en la muestra de la intención de voto según el nivel educativo.

Intención de voto	Nivel Educativo			
	Bajo	Medio	Alto	Total
Justicialista	14	21	10	45 (45%)
No justicialista	11	24	20	55 (55%)
Total	25	45	30	100 (100%)

Ahora bien, supongamos que se ha preguntado a los entrevistados por su intención de voto en las próximas elecciones para estimar qué porcentaje de votos obtendrá el Justicialismo, obteniendo estos resultados. Si se los toma sin tener en cuenta el peso distinto que tienen los diferentes niveles educativos en la muestra y en la población se concluye que hay un 45% de los entrevistados que piensan votar por el Justicialismo.

Pero puesto que el nivel bajo está subrepresentado en la muestra, y que la proporción de votos por el Justicialismo es sustancialmente mayor en ese estrato, es evidente que el porcentaje de votos con esa orientación debe ser sustancialmente mayor en la población,¹⁴ no hay inconveniente alguno en calcular cuál hubiera sido el porcentaje de votos total por el Justicialismo si los distintos estratos hubieran estado representados en la muestra en las proporciones correctas. Para ello se recurre a simples reglas de tres:

las frecuencias de la columna “bajo” se multiplican por 55/25, las de “medio” por 30/45 y las de “alto” por 15/30:

Tabla 3.3.: Distribución corregida de la intención de voto según nivel educativo

Intención de voto	Nivel educativo			Total
	Bajo	Medio	Alto	
Justicialista	31	14	5	50 (50%)
No justicialista	24	16	10	50 (50%)
Total	55	30	15	100 (100%)

El resultado es ahora sustancialmente distinto: el marginal inferior exhibe la misma distribución de los distintos niveles educativos que se da en la población, y el porcentaje resultante de votos por el Justicialismo se eleva al 50%.

Lo que se ha hecho mediante estas sencillas operaciones aritméticas es restablecer el “peso” de los distintos estratos en la muestra, **ponderando** los resultados de acuerdo con la distribución conocida de la variable en la población. Por lo tanto, no es una condición necesaria para el muestreo por cuotas que los distintos estratos se encuentren representados en la muestra en las mismas proporciones que en la población, sino que bastará con que se cuente con un número suficiente de casos en cada estrato como para poder estimar los valores para los estratos de la población. Se denominan “**muestras autoponderadas**” aquéllas en las cuales los estratos se encuentran representados en idénticas proporciones que en la población.

Las muestras por cuotas son mucho más eficaces que las muestras accidentales simples. De allí que su uso sea muy frecuente en sondeos electorales así como en la investigación de mercado en general. Sin embargo, el muestreo por cuotas sigue siendo en esencia un procedimiento accidental en el que cada estrato de la muestra es una muestra accidental del correspondiente estrato de la población.

En particular, la libertad otorgada a los entrevistadores puede dar lugar a **muestras sesgadas**. Éstos son proclives a encuestar a las personas que tengan más a mano –por ejemplo, a amigos o conocidos-, a descartar a las personas que no se encuentren en su domicilio cuando los visitan, a evitar realizar entrevistas en barrios apartados, etc. Ello puede redundar en un cierto **sesgo** de la muestra, en una desviación con respecto a determinadas características reales de la población que la torne poco representativa para un objetivo determinado. Por cierto, ello dependerá del grado en que estas desviaciones se encuentren asociadas a variables relevantes con respecto al propósito de la investigación. Así puede darse que estén subrepresentadas en la muestra las personas que habitan en barrios apartados y que este efecto carezca de importancia en sí, pero que se produzca por esta vía una subrepresentación de las personas de bajo nivel económico-social debido a que éstas tienden a residir en barrios alejados del centro urbano, por ejemplo. Si se pudiera garantizar que no se ha producido ningún efecto de esta naturaleza en la selección de las unidades, el problema estaría resuelto. Pero como se trata de una muestra accidental no existe ningún fundamento que nos permita asumir esa hipótesis.

3.2. Muestras probabilísticas

En la investigación científica, el azar no es algo que se pueda presuponer como naturalmente dado: antes bien, se busca garantizar el azar, lo cual requiere recurrir a procedimientos con frecuencia caros y laboriosos. El azar científico se persigue de manera sistemática, es algo que se construye. Así, en las muestras probabilísticas –o

“aleatorias”- se requerirá que todos los elementos de una población tengan una probabilidad **conocida** de ser seleccionados.

Muestra al azar simple: es el ejemplo más sencillo de un procedimiento de muestreo probabilístico; pero aún más que por su posibilidad de ser aplicado directamente, su interés consiste en que se encuentra a la base de todos los otros tipos de muestras aleatorias. En la muestra simple al azar, todos los elementos de la población tienen la **misma** probabilidad de resultar seleccionados, y todas las combinaciones de elementos para un tamaño dado de la muestra presentan también la misma probabilidad de selección. Así, si nuestro universo está constituido por 100 elementos, cada uno de éstos tendría la misma probabilidad de ser incluido en la muestra: $p = 1/100$.

La probabilidad de una muestra de un tamaño dado de ser extraída, por su parte es igual a:

$$\frac{1}{\frac{N!}{N!(N-n)!}}$$

En la que: N = tamaño de la población
N = tamaño de la muestra
! = factorial

Así, para cualquier muestra de tamaño 10, extraída de una población de 100 elementos, la probabilidad de ser seleccionada será:¹⁵

$$p = \frac{1}{\frac{100!}{10!(100-10)!}} = \frac{1}{124.303\ 388.140}$$

El requisito imprescindible para seleccionar una muestra al azar simple es disponer de un listado de todos los elementos que componen la población. A cada una de las unidades se le asigna un número desde 1 hasta N. Luego, recurriendo a una tabla de números al azar, a un bolillero, a una computadora, o a cualquier otro procedimiento que garantice la producción de una serie de números al azar, se extraen en la cantidad necesaria los números correspondientes a las unidades que integrarán la muestra. Así, si se tiene una población de $N = 1.000$ y se desea seleccionar una muestra de 100 unidades, se sortearán 100 números comprendidos entre el 1 y el 1.000, tomando los últimos tres dígitos de los números generados.

En la práctica, el procedimiento del muestreo al azar simple se torna por demás engorroso al trabajar con poblaciones grandes. Si además, por determinadas restricciones el número de casos en la muestra es pequeño, la muestra al azar simple puede conducir a una representación absolutamente insuficiente de los estratos de menor peso relativo en la población. Es por ello que en estas condiciones se suele optar por otros diseños muestrales libres de esas desventajas. No obstante, la importancia teórica de la muestra al azar simple es fundamental, dado que todos los otros diseños probabilísticos remiten a este modelo, y a que en todos esos otros diseños el muestreo al azar simple es utilizado al menos en alguna de sus etapas.

Muestra sistemática: se trata de un procedimiento de muestreo que, a la vez que es susceptible de simplificar notablemente la selección de las unidades, a todos los efectos prácticos presentará las mismas virtudes que el muestreo al azar simple. Aquí también se requiere como condición indispensable contar con el listado de todas las unidades, las que deberán ser numeradas correlativamente. Conociendo el número de

elementos que componen la población, y habiendo determinado el número de casos necesario para la muestra, se puede establecer la fracción de muestreo dividiendo el primer número por el segundo. Luego se selecciona un solo número al azar, que corresponderá a la primera unidad seleccionada; a este número se le suma la fracción de muestreo y se obtiene la segunda unidad; luego se suma nuevamente la fracción obteniendo la unidad siguiente, y se procede así sucesivamente hasta completar la muestra.

Supongamos que queremos extraer una muestra de 100 estudiantes de un universo constituido por una Facultad de 20.000 alumnos. Nuestra fracción de muestreo será: $20.000/100 = 200$. Seleccionamos un número al azar: el 4.387. Nuestra muestra quedará integrada por los estudiantes que lleven los números 4.387, 4.587, 4.787, 4.987, etc. Debe quedar claro que en este diseño ni todas las unidades ni todas las combinaciones de unidades tienen las mismas probabilidades de resultar seleccionadas. Así, en nuestro ejemplo, una muestra que incluya las unidades 4.387 y 4.388 tiene una probabilidad cero de ser seleccionada (dada nuestra fracción de muestreo de 200).

La condición para poder utilizar este procedimiento es que la numeración de las unidades en la lista no responda a ningún orden en particular, o por lo menos que de existir un orden éste no se base en ninguna propiedad susceptible de encontrarse relacionada con las variables que se desea estudiar. En determinadas circunstancias, sin embargo, la existencia de un orden puede resultar ventajoso pudiendo ponerse al servicio de la investigación. Imaginemos que queríamos estudiar las opiniones de los estudiantes de aquella Facultad respecto a cuestiones relacionadas con su organización académica. Como hipótesis, no resultaría irrazonable pensar que las opiniones podrían variar según la antigüedad de los alumnos en la institución. Ahora bien, si la numeración de la lista es tal que los números más bajos corresponden a los estudiantes más antiguos y los más altos a los más recientes, un procedimiento de selección de la muestra como el expuesto garantizaría la representación en la muestra de estudiantes con diferentes grados de antigüedad en la Facultad.

Muestra estratificada: siempre que hay razones para pensar que determinadas características de las UUA pueden encontrarse fuertemente relacionadas con las variables que tenemos interés en investigar, convendrá asegurar la representación en la muestra de los diferentes valores que pueden asumir las unidades en dichas características. Esto es, en base a determinadas variables que por hipótesis se asumen como pertinentes se divide la población en estratos. En el diseño más simple de muestra estratificada, la única diferencia con respecto a la muestra por cuotas es el hecho de que la selección de las unidades se realiza al azar. Por cierto, esta no es una diferencia menor, puesto que obtenemos una muestra apta para realizar cualquier tipo de inferencia estadística. Pero conceptualmente ambos tipos de muestra son muy similares: así como la muestra por cuotas consiste en una serie de muestras casuales tomadas cada una dentro de un estrato diferente, así una muestra estratificada se compone de varias muestras al azar simple seleccionadas dentro de otros tantos estratos.

Por supuesto, las mismas consideraciones que se hacían acerca de la ponderación de los estratos en la muestra por cuotas se aplican también al muestreo estratificado. Existen poderosas razones por las cuales puede resultar conveniente trabajar con un diseño de muestra que no sea autoponderado. En primer lugar, una muestra autoponderada puede llevar a seleccionar un número insuficiente de casos dentro de un estrato cuyo peso dentro de la población es pequeño. Así, por ejemplo, una muestra autoponderada estratificada por nivel económico social (NES), puede llevar a contar con un magro número de unidades dentro del estrato "Clase alta". Supongamos que la definición de "Clase alta" sea tal que comprenda a un 5% de la población.

Tomando una muestra de 400 casos, sólo se contaría con 20 elementos dentro de ese estrato; un número por cierto insuficiente para poder afirmar nada acerca de los integrantes de ese estrato. De ahí que pueda resultar interesante sobremuestrear voluntariamente ese estrato, de modo de contar, por ejemplo con 80 unidades con esa característica. Por supuesto, es factible hacer jugar esta consideración también en el muestreo por cuotas.

Pero en el caso del muestreo estratificado se agrega una razón más que hace a la eficiencia del procedimiento desde un punto de vista estadístico. En efecto, se ha visto que cuanto menor es la variabilidad de una característica en una población, menor será el número de casos necesario en la muestra para contar con una estimación adecuada del parámetro. Vale decir que para un número dado de observaciones muestrales, cuanto mayor sea la homogeneidad de la población, menor será el margen de error al estimar la distribución de la variable.

Dividir la población en estratos es útil en la medida en que los estratos sean internamente homogéneos y externamente heterogéneos en cuanto a las variables de nuestro interés. Si se cuenta con una estimación previa acerca de la variabilidad del parámetro dentro de cada estrato, entonces el modo más eficiente de muestrear¹⁶ es aquel en que el número de casos seleccionados sea proporcional a esa variabilidad. En la práctica, ello se resume en una consigna sencilla: sobremuestrear los estratos en que la variabilidad sea mayor, submuestrear los más homogéneos. Así, si en un sondeo sobre las intenciones de voto para las elecciones presidenciales se toman las provincias argentinas como estratos, se deberá sobre muestrear aquéllas en las que el resultado aparezca como más dudoso, por ejemplo. Lo expuesto debería bastar para comprender que, excepto por un pequeño ahorro de operaciones aritméticas, nada justifica muestrear a todos los estratos en la misma proporción y, por lo contrario, puede llegar a ser muy ventajoso no hacerlo.

Muestra por conglomerados y en etapas múltiples: si se quisiera hacer una muestra al azar simple —o sistemática, o aún estratificada— de la población de Buenos Aires, se tropezaría de inmediato con un obstáculo difícilmente salvable: la inexistencia de un listado de todos los habitantes de esta ciudad.¹⁷ Por cierto, nada impediría dedicar una etapa previa del muestreo a listar todos los elementos de esa población, a no ser pedrestres consideraciones de costos.

Una solución podría ser recurrir a una muestra por **conglomerados** (*clusters*). Se trata de conglomerados de UUAA que funcionan como unidades de muestreo definidas espacialmente. Este tipo de diseño se utiliza en general para reducir los costos de la recolección de datos. En efecto, cuando el universo de la investigación resulta geográficamente extenso, es muy ventajoso, en términos económicos, tratar de concentrar la tarea en el espacio. Se reducen de este modo los costos que suponen el traslado del entrevistador de una a otra unidad.

Idealmente, los conglomerados son unidades de un mismo tamaño. Como bien observa Blalock (1966: 441), ésta es en cierto modo una estrategia opuesta a la del muestreo estratificado, aunque en ambos casos se divide a la población en grupos. Si en la muestra estratificada se seleccionan los casos **dentro** de cada estrato, asegurando así que todos los estratos estarán representados, en este otro tipo de muestra se selecciona **entre** los conglomerados; correlativamente, así como se busca que los estratos sean lo más homogéneos posible internamente, para los conglomerados cuanto más heterogéneos sean mejor será el resultado. Dividida toda la población en conglomerados, se selecciona al azar un cierto número de éstos dentro de los cuales se entrevistará a todas las unidades de análisis.

A menudo la técnica por conglomerados se integra en diseños de **muestra en etapas múltiples**, esto es, muestras que suponen la definición de unidades de muestreo en diferentes niveles. Así, si se quisiera medir la tasa de desempleo en la población de Buenos Aires con 14 o más años de edad, se podría estratificar la ciudad en zonas, y dentro de cada zona o estrato seleccionar al azar un cierto número de manzanas; la gran ventaja es que todas estas operaciones se realizarán en la comodidad de un escritorio si se cuenta con buena cartografía. Luego para cada una de las manzanas seleccionadas, se puede proceder en campo a un rápido relevamiento del número de unidades domésticas existentes. A partir de los listados para cada manzana, es posible seleccionar al azar unidades domésticas.¹⁸ Finalmente, en cada unidad doméstica el entrevistador podrá sortear por algún procedimiento adecuado la persona a encuestar. En este diseño multietápico, las manzanas funcionarían como unidades de muestreo de 1era. Etapa, las UDD serían unidades de muestreo de 2da. etapa, y las UUA –personas- serían las unidades de muestreo de última etapa.

Sin duda, las muestras por conglomerados y en etapas múltiples son muy ventajosas desde el punto de vista económico, en el sentido que con igual cantidad de recursos es posible realizar mayor número de encuestas. Sin embargo, hay que tener en cuenta que con estas técnicas, contrariamente a la del muestreo estratificado, el grado de error aumenta: esto es para obtener el mismo grado de exactitud la muestra al azar simple se bastará con un menor número de casos.

Figura XIV: Resumen de los principales tipos de muestras

No-probabilísticas	Probabilísticas
Casual Por cuotas	Al azar simple
	Sistemática
	Estratificada
	Por conglomerados
	Por etapas múltiples

¹ Cf. Zetterberg; 1968: 66-67.

² Cf. Selltiz et al.; 1968:560.

³ “Estudio sobre los medios de subsistencia y la capacidad de pago de la población no propietaria de Posadas a relocalizar por Yacyretá” (Informe inédito producido por la FHCS, UnaM).

⁴ Cf. Selltiz et al.; 1968:561.

⁵ Nada impide combinar ambos procedimientos. De hecho el Censo Nacional de Población en la Argentina combina ahora un formulario corto que se aplica a todos los hogares (universo), y un formulario más extenso que sólo se aplica en una muestra de los hogares.

⁶ Aunque, obviamente, de ser este el caso ya no sería necesario recurrir a una muestra.

⁷ Existiría la posibilidad en ese caso de aumentar el número de unidades seleccionadas, pero ello entrañaría varias consecuencias negativas: a) el carácter de la muestra se vería desnaturalizado, en lo que hace a sus propiedades estadísticas; b) los costos de retornar al campo en busca de más datos pueden ser muy gravosos; c) se puede perder la comparabilidad entre los datos recolectados en dos períodos distintos, ya sea porque las variables objeto se hayan visto modificadas por el simple paso del tiempo, o porque el mismo proceso inicial de recolección haya conducido a una alteración en sus valores.

⁸ La variabilidad máxima es 50/50, vale decir cuando $p=q$ (siendo p la proporción de los que poseen el atributo y $q=1-p$). Intuitivamente, se comprende que si una población es absolutamente homogénea en cuanto a una característica dada, una muestra de cualquiera de los elementos de dicha población bastará para producir una estimación correcta del parámetro; en cambio si la población es heterogénea –se divide por mitades en varones y mujeres, por ejemplo- es claro que se necesitará un número mayor de observaciones para producir una estimación ajustada de la distribución de la característica en el universo.

⁹ Por ende, una muestra cuyo tamaño sea suficiente para estimar un determinado parámetro, podrá no serlo para otro parámetro, si la distribución de este último es menos favorable.

¹⁰ Cf. *Infra*: Cap. IV.

¹¹ O, al menos, remitirse a los capítulos pertinentes de los manuales de Blalock (1966), Padua (1979), Galtung (1966), Selltiz (1968), o Festinger y Katz (1975), o a textos específicos como el de Kalton (1987).

¹² Dada la naturaleza de este manual, nos concentramos en los tipos de muestras utilizados corrientemente con fines “cuantitativos”. En una perspectiva “cualitativa” –aunque también en la fase exploratoria de investigaciones que versen sobre un gran número de casos- cabe destacar la pertinencia de otros procedimientos de enorme valor heurístico. Así, “el **muestreo teórico** se efectúa para descubrir categorías y sus propiedades y para sugerir las interrelaciones dentro de una teoría. El muestreo estadístico se hace para obtener evidencia exacta sobre la distribución de la población entre categorías a ser usadas en descripciones o verificaciones. De este modo, en cada tipo de investigación la *muestra adecuada* que deberíamos buscar (como investigadores y lectores de investigaciones) es muy diferente.” (Cf. Glaser y Strauss; 1967:62.

¹³ Afortunadamente, se dispone con frecuencia de datos sobre la población a muestrear provenientes de otras fuentes. En este caso, los datos son imaginarios; pero si se tratara de muestrear la población de Posadas, por ejemplo, se podría recurrir fácilmente a los resultados de la Encuesta de Hogares para conocer la distribución de la variable en el universo.

¹⁴ Por ejemplo, se dispone de datos censales, o de la última onda de la Encuesta Permanente de Hogares realizada en la ciudad.

¹⁵ Cf. J. Padua (Ed.), 1979:66.

¹⁶ Vale decir, el que a igual cantidad de casos produzca mayor exactitud en la estimación del parámetro, o el que nos permita obtener al mínimo costo un resultado con un nivel de confianza y un margen de error dados.

¹⁷ Obsérvese que la utilización de la guía telefónica produciría una muestra sesgada, puesto que no todos los habitantes de Buenos Aires figuran en esa guía, y que la tenencia de teléfono, lejos de comportarse como una variable aleatoria, está asociada a otras características de las personas, como su nivel económico. A pesar de ello, las encuestas por vía telefónica son de uso frecuente (Cf. Lebart, 1992)

¹⁸ En la terminología del Censo de Población se habla tradicionalmente de “hogares” en vez de unidades domésticas.

