
GABRIEL DAVIDOVICS

Magister en Estadística Matemática. Profesor de Teoría del Muestreo, Universidad de Chile.

GUSTAVO MARTINEZ

Master en Sociología y Ciencia Política, Profesor de Metodología de la Investigación en la P. Universidad Católica de Chile y en la U. de Chile.

Muestreo Probabilístico y Muestreo de Cuotas en los Estudios Preelectorales y de Opinión Pública. *Argumentos para una Discusión Crítica y Necesaria en Chile* *

I. EL RESURGIMIENTO DE LAS ENCUESTAS POR MUESTREO, SU CALIDAD Y LAS CRITICAS QUE HAN RECIBIDO

Desde el restablecimiento democrático se ha producido en Chile un verdadero florecimiento de los estudios de opinión pública y de las encuestas preelectorales, las que han acaparado la atención de los medios de comunicación de masas. Hoy son muchos los institutos y empresas que ofrecen resultados periódicos de estudios de opinión pública, y cuando las elecciones se aproximan surgen esporádicamente otras entidades entregando sus propios resultados¹.

Tanto la prensa escrita como la televisiva y radial han hecho de tales estudios, particularmente de los que pronostican resultados de elecciones futuras, temas insustituibles de sus agendas de noticias.

* Agradecemos los comentarios y observaciones de las profesoras Erika Himmel (U. Católica) y María Elisa Osses (U. de Chile).

¹ Los Institutos que periódicamente difunden estudios de opinión pública en los medios de comunicación son el Centro de Estudios Públicos (CEP),

Desgraciadamente, también han formado parte de las agendas de los medios, artículos generalmente sarcásticos. En ellos, algunos periodistas —que desgraciadamente no son especialistas en estos temas— han hecho mofa de la mala calidad de las predicciones efectuadas por los investigadores que han difundido los resultados de sus estudios a través de los medios de comunicación². Tampoco han estado ausentes las críticas que se han dedicado a ciertos investigadores, los que, a través de cartas en los medios (diarios y revistas) o de crónicas en la prensa, han sido acusados de incompetencia profesional, de parcialidad política e incluso de deshonestidad³. Estos últimos calificativos son bastante deshonorosos para los investigadores, ya que implican acusarlos de falsear deliberadamente los resultados obtenidos de sus estudios.

Ciertamente, la generalidad de estas críticas es exagerada y, en muchos casos, bastante injustas, ya que es frecuente que quienes “afilan sus cuchillos” en contra de los investigadores no aportan las pruebas suficientes que respalden cargos tan graves. Quienes formulan estas ácidas acusaciones suelen tomar como expresiones indicativas de ambas censuras (deshonestidad e incompetencia profesional) el mero desajus-

el Centro de Estudios de la Realidad Contemporánea (CERC) y el Departamento de Estudios Sociológicos de la Universidad Católica (DESUC).

Entre las empresas comerciales que realizan similar labor se encuentran Adimark y Gémines. Con investigaciones de intencionalidad de voto y sólo en períodos preelectorales están Investmerc y Benchmark.

² Véase “Resultados cuestionan validez de las encuestas”. Diario “La Hora” (12-XII-1997).

³ Un ejemplo de ello es la carta de Andrés Tagle, publicada en “El Mercurio” del 23 de Diciembre de 1997, en que, junto con objetarle a Roberto Méndez una serie de vaticinios preelectorales erróneos que incluyen el exit-poll de 1993 para Canal 7, lo acusa de presentar en Notaría con ocasión de las elecciones del 11 de Diciembre de 1997, una pseudo encuesta para corregir sus equivocadas estimaciones anteriores. Tagle también le objeta a Méndez la intencionalidad de favorecer a los candidatos senatoriales de Renovación Nacional con sus cifras, en desmedro de los de la Unión Demócrata Independiente. Además, el mismo Andrés Tagle declaró que Roberto Méndez con la publicación de esta pseudo encuesta “sólo quería quedar bien” y que

te entre los porcentajes de votación que recibirían los candidatos en una elección, según las estimaciones entregadas por los estudios, y los resultados finales escrutados de la elección misma cuando ellos se entregan oficialmente⁴.

Indudablemente, los resultados de una elección constituyen un elemento de contrastación de cuán bueno o malo puede haber sido el trabajo realizado y, en tal sentido, si el estudio es muy próximo a la fecha de la elección podría otorgársele un valor predictivo⁵.

No obstante lo anterior, creemos que los críticos no apuntan a los verdaderos blancos hacia los cuales deberían ir dirigidos sus dardos.

Ellos no focalizan su atención en los aspectos metodológicos del trabajo, particularmente en el diseño muestral empleado, ni en las preguntas del instrumento de medición, ni en la cali-

ellas constituían una "manipulación de datos": ("El Mercurio", 14-XII-1997).

Véase también: "Ravinet invalida resultados de estudio" ("La Tercera", 7-I-1998). En entrevista periodística Jaime Ravinet descalifica los resultados del estudio del CERC que lo colocaban como el candidato presidenciable de la Democracia Cristiana percibido con menores opciones. En la entrevista, Ravinet señala que le resta importancia a la encuesta del CERC porque "en las predicciones que ha hecho este tipo de encuestas no le han apuntado a una. Así sucedió con las que publicó en Diciembre la socióloga Marta Lagos".

En otro pasaje de la entrevista, Ravinet expresa que a pesar de que quiere una elección al interior de su partido y que rechaza imposiciones cupulares para elegir el precandidato presidencial de la Democracia Cristiana: "un mecanismo para buscar consenso entre los implicados es a través de una encuesta seria, abierta, con unos dos mil encuestados. Estaría dispuesto a aceptar el resultado si el resto también se compromete".

⁴ Para la gran mayoría de los periodistas que cubren el sector de los estudios de opinión pública, un estudio estaría bien hecho si los resultados predictivos coinciden en el orden de llegada de los candidatos en la respectiva carrera electoral.

⁵ En ningún estudio de predicción electoral se puede descartar que parte de los electores modifique sus intenciones de voto en el período que media entre la recolección de los datos del estudio mismo y el día mismo de la elección.

dad de la ejecución de la recolección de los datos, ni en las tasas de no-respuesta de los potenciales entrevistados probabilísticos, ni en los procesos de inflatación de los resultados desde el nivel de la muestra al del universo de estudio y otros aspectos de similar naturaleza metodológica. Las críticas se formulan de un modo muy irresponsable, centrándolas en una pretendida deshonestidad por parte del investigador⁶.

Sin embargo, parte de la culpa de que ello acontezca así radica en los propios investigadores, los que omiten a la prensa la información metodológica relevante, sobre todo aquella que permitiría resituar o reubicar la evaluación de las predicciones electorales mismas. Esta omisión es, sin duda, la que permite que se generen las polémicas alrededor de las predicciones electorales las que, en último término, son las que levantan las pasiones⁷.

Normalmente no se saben las características de los diseños muestrales empleados, se desconocen los procedimientos de inflatación de los resultados, se ignoran las tasas de no-respuesta (atribuibles a rechazos, no habidos en el hogar, ausencia del seleccionado probabilístico, u otras causas). Igualmente, no se publican las medidas tomadas para tratar de corregir

⁶ Si los críticos, en lugar de formular acusaciones de deshonestidad intelectual o de parcialidad política, se concentraran en los aspectos técnicos de los trabajos se encontrarían con bastante paño para cortar.

⁷ Al parecer en Chile se cree desmesuradamente en un pretendido efecto de "bandwagon" que producirían las encuestas preelectorales. Tal hipótesis supondría al menos tres cosas. Primero, que los electores leen o escuchan las noticias referidas a las encuestas de predicciones electorales difundidas por los medios. Segundo, que les otorgan credibilidad en el sentido que ellas efectivamente pronostican lo que ocurrirá el día de las elecciones. Tercero, y esto es lo fundamental, que por efecto de esta información, asumida como veraz, ciertos electores modifican las intenciones de voto previamente adoptadas por ellos, o se deciden si es que eran electores indefinidos, a favor del presunto vencedor según las encuestas. Estas condiciones requieren obviamente de prueba empírica en la realidad chilena.

estas omisiones, y no se tiene la más mínima información sobre los procedimientos de imputación de los "missing data", si es que dichos métodos se han implementado⁸.

Este silenciamiento de cómo trabajan ellos mismos tiene un solo culpable: los institutos o empresas investigadoras. Ellos se reservan una información que debiera ser susceptible de ser evaluada analíticamente, si el estudio se difunde por los medios de comunicación de masas.

Tampoco cabe duda de que la investigación con técnicas de encuestas por muestreo en Chile está muy distante de los estándares de calidad de otras áreas geográficas científicamente más avanzadas, y ello se observa -entre otros aspectos- en el amplio uso que se hace en nuestro país de las muestras de cuotas o de las así llamadas "muestras semi-probabilísticas".

Podemos explicarnos que muchos investigadores trabajen con muestras de cuotas o semi-probabilísticas de cuotas. Total, si los financiadores o usuarios de los estudios ignoran la importancia del empleo del muestreo probabilístico, ¿para qué complicarse la vida con diseños muestrales sofisticados que exigen experiencia y sobre todo conocimientos profundos de muestreo, que además demandan un mayor tiempo de recolección de los datos y que requieren generalmente mayores costos de realización?

¿Para qué autoasignarse más trabajo y problemas si el destinatario de los estudios no es capaz de distinguir el rayo láser del cuchillo carnicero para operar?

¿Por qué motivos habrían de asumir el empleo de diseños muestrales probabilísticos si no se los exigen?

⁸ Lo cierto es que un velo de incógnitas rodea cada investigación, y la así llamada "ficha técnica" que acompaña a los estudios no pasa de ser una caricatura de lo que debería ser información metodológica relevante para los efectos de la evaluación del diseño muestral utilizado en el estudio.

¿Para qué autoflagelarse si nadie los censura desde un punto de vista científico?

¿Para qué intentar siquiera trabajar con muestras probabilísticas si los docentes que enseñan Teoría del Muestreo y Metodología de la Investigación guardan silencio?⁹

Podemos comprender entonces que muchos investigadores usen muestras de cuotas, aunque no podemos justificar tal conducta. Sin embargo, lo que se nos hace ininteligible es comprobar cómo algunos de esos investigadores intentan defender las muestras no-probabilísticas como si fueran científicas y modernas, aptas para ser utilizadas en los estudios de opinión pública, en los preelectorales, e incluso en los de marketing, donde en muchos casos se toman decisiones en que están involucradas grandes sumas de dinero y en los que la bondad de dichas decisiones depende en gran medida de la calidad y precisión de la información entregada por estos estudios.

La casi totalidad de los investigadores que en Chile emplean las técnicas de encuestas por muestreo y recurren a muestras no-probabilísticas defienden el uso de estos procedimientos, aunque sin llegar hasta el extremo de afirmar que tienen bases científicas. Una notable excepción ha estado constituida por la argumentación proporcionada –no en revistas científicas, sino en entrevistas periodísticas y declaraciones a los medios escritos– por Marta Lagos y Carlos Huneeus, altos ejecutivos de MORI y CERC, respectivamente. ¿Cuáles han sido los argumentos de ambos profesionales?, ¿son ellos susceptibles de crítica? Estimamos que sí.

⁹ La excepción está constituida por el trabajo de Víctor Correa: “Muestreo Probabilístico y por Cuotas. ¿Cuál es mejor?”, publicado en la revista “Estadística y Economía”. Instituto Nacional de Estadística, Vol. 15, Santiago, diciembre de 1997, págs. 63-75.

Tuvimos conocimiento del contenido de este artículo con posterioridad a la entrega de nuestro paper y pudimos constatar con satisfacción que no somos los únicos en esta cruzada en favor del muestreo probabilístico. Dejamos constancia que concordamos absolutamente con las ideas esenciales expresadas por Correa.

II. LAS MUESTRAS DE CUOTAS Y LAS PROBABILISTICAS EN LOS ESTUDIOS DE OPINION PUBLICA

Nos parece que los diseños muestrales correctos, desde el punto de vista científico, son los probabilísticos y no los de cuota y que los argumentos empleados por Marta Lagos y Carlos Huneeus para descalificar los diseños muestrales probabilísticos son insostenibles.

Esos argumentos aparecieron vertidos en la entrevista efectuada por Raquel Correa a Carlos Huneeus y Marta Lagos en "El Mercurio" del domingo 30 de abril de 1995, en el artículo aclaratorio de Marta Lagos publicado en "El Mercurio" del domingo 7 de Mayo de 1995 y en la crónica "Efecto Boomerang" de la revista "Qué Pasa" del sábado 6 de Mayo del mismo año¹⁰.

En general, podemos decir que hemos encontrado en ellos dos tipos de argumentaciones criticables, a saber:

- 1.- Algunas contradicciones entre sus afirmaciones respecto del muestreo probabilístico y su empleo.
- 2.- Algunos errores en sus planteamientos que niegan la científicidad y utilidad del muestreo probabilístico y, en cambio, valorizan injustificadamente el muestreo de cuotas.

¹⁰ Las afirmaciones en defensa del muestreo de cuotas hechas por Marta Lagos y Carlos Huneeus fueron cuestionadas por los autores de este trabajo en un artículo periodístico que con el título "Muestras de Cuotas y Probabilísticas", fue publicado por "El Mercurio" el domingo 23 de julio de 1995. Lo esencial de los argumentos de Marta Lagos y Carlos Huneeus y de nuestros contra argumentos constituyen la base de la sección II del artículo que el lector tiene entre manos. Por su relevancia en la discusión del tema, objeto de este paper, hemos considerado conveniente reproducir las ideas de ambas partes.

II.1. Contradicciones

- 1.- Huneus afirma: “.....El sistema probabilístico es muy caro y lento. Es un método del siglo XIX “, “... el sistema nuestro es el que se ocupa en Europa. Nosotros antiguamente hacíamos encuestas probabilísticas; los expertos europeos se mataban de la risa: un sistema obsoleto y engorroso”.

Si esto fuera así, cómo se sostiene esto con la afirmación de Marta Lagos el 7 de Mayo, en que refiriéndose al muestreo probabilístico y de cuotas, sostiene que “Quienes aplicamos ambas metodologías a diario...”. Surge la pregunta: ¿lo hacían antiguamente o lo hacen actualmente a diario? Y si esto último es cierto, ¿para qué lo utilizan si es un método obsoleto y engorroso?, ¿para qué emplean un método muestral que es del siglo XIX y que puede provocar la hilaridad de los expertos europeos?

- 2.- Huneus sostiene: “Decir hoy día que el sistema probabilístico es más científico, es desconocer el desarrollo de toda la industria de la encuesta en el mundo en los últimos 20 años”. Luego, cuando la periodista Raquel Correa le indica que es el que usan los norteamericanos, Huneus continúa: “ Porque ellos cuentan con un sistema de información perfecto. Las personas están listadas en un directorio que se actualiza mes a mes. Obviamente, si tuviéramos un listado de todos los habitantes de Chile no haríamos jamás una muestra por cuotas”.

Perfecto, si el sistema probabilístico es caro, lento, engorroso, obsoleto y además menos científico que el de cuotas, ¿ por qué si Huneus contara con ese famoso listado, que por lo demás no existe siquiera en los Estados Unidos, dejaría de lado las muestras de cuotas? ¿ Sólo se estarían topando como limitación con la falta de ese listado para abandonar de inmediato el muestreo de cuotas y trabajar, en consecuencia, con el muestreo probabilístico?

Afirmar que si ese listado de habitantes existiese “no haríamos jamás una muestra por cuotas” implica sostener que la muestra de cuotas es un sustituto a la falta del listado y no que es superior al método probabilístico.

- 3.- Huneus afirma peyorativamente: “...La única gracia que tiene la probabilística es que se puede calcular el error matemático...”. Si ésta es la única gracia que tiene, ¿por qué Marta Lagos agrega frecuentemente en la entrega de resultados el probable margen de error muestral que tendrían sus cifras, en el supuesto caso que su muestra fuera probabilística? Si ésta es la única gracia, ¿por qué adornar sus resultados con lo que tan despectivamente se desprecia? Hay que recordar que Marta Lagos agrega siempre que esos márgenes de error muestral se obtendrían sólo en el supuesto caso de que la muestra fuera probabilística. Pero como las muestras de cuotas no permiten calcular márgenes de error muestral, ello se transforma en un agregado innecesario. Es como decir que si un ciego tuviese visión, entonces podría ver lo maravilloso de los colores. Pero acontece que el ciego no puede ver, y las muestras de cuotas no posibilitan calcular los errores muestrales.
- 4.- Marta Lagos asevera que: “la muestra por cuotas es la mejor metodología aplicada a la realidad chilena. La probabilística no sirve para nuestra realidad. Es impracticable”.

¿Por qué hace esta afirmación?, ¿qué tendría de específico la realidad chilena para sostener que la muestra por cuotas es la mejor metodología aplicable a la realidad chilena? El “onus probandi” corre por cuenta de ella y mientras tal opinión no se fundamente aparece como un juicio gratuito y caprichoso.

Sin embargo, y al margen de lo anterior, en su artículo “¿Opinión Pública Chilena? aparecido el 31 de Diciembre de 1994 en “El Mercurio”, Marta Lagos acota, en relación a la contratación de dos técnicos norteamericanos en estudios de opinión, por parte de CEP: “La primera consecuencia es que con ello se afirma que no existe en el país gente suficientemente confiable

como para diseñar un estudio así. Sin desmerecer a mis colegas y amigos extranjeros, por cierto competentes, creo importante destacar que no es necesario salir del ámbito nacional para hacer una cosa similar. Las universidades chilenas tienen técnicos perfectamente capaces de diseñar muestras probabilísticas y nacionales, que no estén políticamente tildadas."

¿Por qué se defiende la competencia de los técnicos chilenos para diseñar muestras probabilísticas de cobertura nacional y simultáneamente se afirma que ellas no son la mejor metodología muestral aplicable a nuestra realidad?

II.2. Algunas afirmaciones erróneas

Primera afirmación: "Las encuestas políticas deben contemplar mil casos."

Esto no es así. Tal exigencia no existe para el muestreo probabilístico y para el de cuotas no existe criterio alguno. No existe un tamaño muestral específico que determine que una muestra se pueda considerar como científica o no científica.

En primer lugar, el tamaño muestral de una muestra probabilística debe adecuarse a los requerimientos de precisión del investigador o de los usuarios, de tal modo que si se desea aumentar la precisión respecto a la que se obtendría con determinado tamaño muestral, manteniendo constantes los restantes factores del diseño, se debe aumentar el tamaño muestral. En consecuencia, no existe tal deber de trabajar con un determinado tamaño muestral, ya sea éste de mil casos u otra cifra cualquiera.

El investigador tiene plena libertad de acción, dentro de los límites del presupuesto y del error muestral máximo deseado o permitido. En caso de poseer un presupuesto prefijado debe tratar de optimizar su diseño con el objeto de lograr la máxima precisión, minimizando también en lo posible, los sesgos y los errores no-muestrales. En cambio, si lo que se le exige al investigador es precisión preestablecida, debe tratar de lograr ese

objetivo minimizando los costos.

Claro que todo lo aquí expresado tiene sentido solamente para el caso de las muestras probabilísticas. En el caso de las muestras de cuotas, el tamaño muestral no tiene sentido como indicador de precisión. Aún más, con muestras de cuotas no existen ni fórmula matemático-estadística ni criterio posible para determinar tamaño muestral alguno en función de una precisión deseada.

Segunda afirmación: "Las encuestas políticas no pueden ser parciales sino nacionales."

En cuanto a la afirmación de que la encuesta debe ser nacional y no cubrir parcialmente el territorio nacional, nosotros afirmamos que la encuesta debe cubrir el área geográfica que el investigador, el cliente o el usuario, considere de interés. Así es que si sólo interesa conocer la opinión de las personas adultas que habitan en las ciudades sobre 50,000 habitantes, pues es legítimo realizar una encuesta que cubra sólo dichas ciudades, vale decir, una cobertura parcial del territorio nacional.

Lo que sí se le debe pedir al investigador, es que en dicho caso no trate de presentar sus resultados publicándolos como representativos de todos los chilenos.

En todo caso esta afirmación es curiosa dado que ni MORI ni el CERC utilizan muestras nacionales. Al menos, los estudios publicados por la prensa no cubren la población nacional en su totalidad.

Tercera afirmación: "El muestreo probabilístico sería un procedimiento del siglo XIX y estaría obsoleto".

No es así. Durante el siglo XIX no se utilizaron los diseños muestrales probabilísticos, sino los diseños basados en muestras "intencionales" (purposive samples), entre los que figura el muestreo de cuotas.

El muestreo probabilístico fue introducido y desarrollado por el gran estadístico polaco Jerzy Neyman en 1934, vale decir en pleno siglo XX. Él también tuvo que luchar contra la incomprensión de muchos, que no podían aceptar el hecho de que alguien propusiese trabajar en base a probabilidades.

Y como decía el propio Neyman más recientemente, en un simposio celebrado en la Universidad de Carolina del Norte: "...todavía seguimos presenciando a investigadores que le rinden homenaje al muestreo probabilístico, de la boca para afuera, pero que en la práctica siguen desechando su uso ya sea por ignorancia o por intención de conducir a engaño" ¹¹.

Cuarta afirmación: "El muestreo probabilístico sería obsoleto y engorroso".

Esta afirmación es errónea. Se hace "engorroso" cuando no se domina la teoría del muestreo, pero cualquier profesional de la especialidad (estadístico-muestrista) puede diseñar muestras probabilísticas, válidas y eficientes, si es que cuenta con el conocimiento y la experiencia correspondiente.

Obviamente, el diseñar muestras probabilísticas no es una tarea que cualquier investigador pueda hacer. Se trata de un conocimiento especializado que obliga a contar con un estadístico muestrista competente en la ejecución de un estudio realizado con técnicas de encuestas por muestreo (técnicas de "survey sampling"). Así también diseñar submarinos o aviones debe ser materia a ejecutar por ingenieros de la especialidad.

Quinta afirmación: "La única gracia del muestreo probabilístico consistiría en que se puede calcular el error muestral".

¡Y vaya qué gracia! Precisamente una de las principales bondades y utilidades científicas del muestreo probabilístico es

¹¹ Véase de Jerzy Neyman "Bias in surveys due to nonresponse" en Norman L. Johnson y Harry Smith Jr., "New developments in survey sampling", Wiley-Interscience, New York, 1969.

mirada con desdén por Carlos Huneeus.

Ciertamente, el muestreo probabilístico permite calcular los intervalos de confianza para niveles de confianza o seguridad estadística preestablecidos por el propio investigador, además de permitir calcular el tamaño muestral necesario para cumplir con ese objetivo.

La estimación de punto, como sería señalar, por ejemplo, que el 35% de cierta población apoya al Presidente Frei, no tiene un valor en sí misma, ya que no significa estadísticamente que ese mismo guarismo exista en el universo o población de la cual se extrajo la muestra (y que se hubiese obtenido si se hubiese realizado un censo o estudio de enumeración completa del universo).

La “estimación de punto” es sólo el inicio para calcular los límites superior e inferior de un intervalo de confianza (de la aprobación presidencial en nuestro ejemplo), sirviéndonos para ello de la estimación de punto y del error muestral (múltiplo del error standard) obtenidos a partir de la misma muestra, y todo esto con una probabilidad determinada por nosotros mismos. Esta característica es importantísima.

Nada de esto se puede realizar con las muestras de cuotas y, como consecuencia de ello, no podemos saber si sus resultados reflejan o no lo que ocurre en el universo o población.

Como ejemplo de lo anterior, podríamos decir que si como resultado de una encuesta efectuada con una muestra de cuotas, obtenemos un resultado hipotético de un 35% de las personas opinando en determinada forma, no podemos agregar ninguna medida de precisión a dicho resultado, y por lo tanto, no podemos decir cuán cerca o cuán lejos está del valor verdadero que habríamos obtenido de haber censado el universo.

En cambio, si la muestra fuera probabilística, con un error standard de 1% (por ejemplo), podríamos decir que hay un 95% de probabilidades que el valor verdadero esté compren-

dido entre 33% y 37% (éste es el intervalo de confianza, vale decir el 35% obtenido de la encuesta al cual le hemos restado y sumado simultáneamente 2%, que equivalen a dos veces el error standard del 1%).

El multiplicar por 2 es lo que permite afirmar que el nivel de confianza es del 95%. En una curva normal exacta, el valor real es de 1.96. En cambio, si queremos tener un nivel de confianza mayor de que el parámetro estaría ubicado al interior del intervalo de confianza, digamos de un 99.7%, entonces deberíamos multiplicar el error standard por 3.

En nuestro ejemplo hipotético, esto significaría que deberíamos restar y sumar simultáneamente 3% (vale decir tres veces el error standard) a la estimación del 35% obtenido de la encuesta. Esto nos daría un intervalo de confianza con un límite inferior del 32% y superior de 38%, y tendríamos una seguridad del 99.7% de que el parámetro que estamos tratando de estimar estaría ubicado en algún lugar entre estos dos valores.

Cabe señalar que el error standard de cada resultado, para cada una de las preguntas del cuestionario, se puede calcular a partir de la misma muestra, aplicando para ello la fórmula que corresponda al diseño muestral y al estimador utilizados.

Pero quisiéramos agregar a esta "gracia" de poder determinar la precisión de las estimaciones a partir de la misma muestra, el hecho de que para extraer una muestra probabilística no necesitamos contar con información sobre la composición del universo (como distribución por sexo, grupos etarios, niveles socio-económicos y otros), que sí son indispensables para "construir" una muestra de cuotas.

Esta característica del muestreo probabilístico puede parecer sorprendente para Marta Lagos y Carlos Huneeus, pero es así. Basta contar con material cartográfico suficientemente detallado y actualizado de las áreas geográficas que cubren el universo, para que se pueda extraer una muestra. Esta misma muestra nos proporcionará luego toda la información que de-

seemos del universo, incluyendo, por cierto, el perfil socio-demográfico.

Sexta afirmación: "El muestreo probabilístico es impracticable".

Otra afirmación errada. ¿Por qué sería impracticable?

El INE, en sus estudios sobre empleo y desempleo o del cálculo del Índice de Precios al Consumidor, utiliza el muestreo probabilístico. La serie de estudios del CONACE (Consejo Nacional de Control de Estupefacientes) sobre consumo de drogas también está basada en un diseño probabilístico. Los estudios del CEP (Centro de Estudios Públicos) sobre opinión pública también se basan en diseños muestrales probabilísticos¹². Si el INE, el CONACE y el CEP, entre otros institutos, no han encontrado las razones que los hacen impracticables, ¿cuáles serían éstas?

La información censal de que dispone el INE le permite extraer -si los investigadores lo desean- muestras probabilísticas multietápicas de áreas, asignando probabilidades a todas las unidades muestrales ya sea de primera, segunda o tercera etapa (por ej. comunas, distritos, zonas censales, manzanas).

La información existe y la cartografía necesaria también, como para poder emplear procedimientos científicamente adecuados.

Pero no sólo el INE puede disfrutar de estas posibilidades. Cualquiera que disponga del material cartográfico y estadístico pertinente, del INE, puede utilizar esta información (por ej. del último Censo de Población y Vivienda) para asignar probabilidades variables a las unidades muestrales del marco

¹² Adimark, Gémines y DESUC, señalan al dar cuenta de sus estudios que sus muestras son probabilísticas. Si tomamos en cuenta todos los institutos y empresas que habitualmente divulgan sus estudios de opinión pública, exclusivamente el CERC y MORI emplearían muestras de cuotas.

Respecto de los diseños muestrales utilizados por Benchmark y por Investmerc no se conocen públicamente sus características.

muestral, en el caso de trabajar con diseños muestrales conocidos como PPS (Probabilities proportionate to size). Los datos censales, en una muestra probabilística, no son necesarios para ponderar la muestra y así asemejarla al universo. La muestra, automáticamente, si está bien diseñada y apropiadamente inflactada al universo por los recíprocos de las fracciones muestrales de cada estrato (en el caso de muestreo estratificado), proporcionará todas las estimaciones necesarias correctamente ponderadas.

En el muestreo probabilístico, a diferencia de la muestra de cuotas, no hay que tratar de construir un universo en miniatura. La muestra probabilística lo construye automáticamente. Si es autoponderada, ya se asemeja en su composición al universo de estudio, y si no es así, se debe ponderar por los recíprocos de las fracciones muestrales utilizadas (y no por datos censales) y se obtiene la inferencia deseada.

Séptima afirmación: "El muestreo probabilístico es demoroso".

No es una limitación "per se" del muestreo probabilístico. Las visitas y revisitas de los entrevistadores en el muestreo probabilístico podrían hacerse en muy corto tiempo. Basta aumentar el número de entrevistadores y la frecuencia de sus visitas para acortar el tiempo de trabajo de terreno a los límites deseables, y todo esto sin aumentar los costos.

Octava afirmación: "El muestreo probabilístico ya no se practica en Europa".

Ello no es cierto. Hay quienes no lo practican y, en el caso de los que prefieren el muestreo de cuotas, los resultados -cuando ellos son contrastados con elementos externos como ocurre respecto a estudios sobre intenciones de voto y votos efectivos, en las proximidades de una elección- suelen ser desafortunados. Demasiado conocidas son las predicciones electorales del Literary Digest en 1936 y de George Gallup en 1948, como para reiterarlas aquí¹³.

¹³ En las elecciones presidenciales de 1936, una revista americana, el

Más frescos en la memoria están los fracasos predictivos de 54 estudios preelectorales, basados todos ellos en muestras de cuotas (usadas por su bajo costo y su rapidez en la recolección de datos), en las elecciones parlamentarias de Abril de 1992 en el Reino Unido. Ninguno de los Institutos de investigación logró predecir el triunfo de los Conservadores, a pesar de que ellos triunfaron por un margen superior al 7% sobre los Laboristas¹⁴.

Literary Digest, predijo el triunfo de Alf Landon sobre Roosevelt por 57% contra 43% del voto popular. Roosevelt ganó la reelección presidencial con el 61% de los votos populares.

El Literary Digest, que había pronosticado correctamente los ganadores en las elecciones presidenciales de 1920, 1924, 1928 y 1932, utilizaba una muestra intencional que además se sesgaba a favor de los votantes de mayor status socioeconómico, ya que los miembros de la muestra eran obtenidos de directorios telefónicos y de listados de propietarios de automóviles. La revista dejó de aparecer 2 años después de su último vaticinio electoral, cuyo fracaso fue atribuido por sus editorialistas a que los electores demócratas, a diferencia de los republicanos, no se daban el trabajo de contestar sus cupones de votación, por correo.

La muestra del Literary Digest en 1936 alcanzó a más de 2 millones de personas, hecho que prueba que lo importante no es el tamaño muestral si no el diseño muestral.

George Gallup, en 1948, estimulado por sus propios éxitos predictivos en las elecciones de 1936, 1940 y 1944, predijo erróneamente el triunfo de Thomas Dewey sobre Harry Truman. Para explicar su fracaso, Gallup, que utilizaba entonces muestras de cuotas, señaló que los datos censales de 1940 estaban desfasados por los rápidos cambios operados en la sociedad norteamericana. También adujo que su último estudio había sido realizado un mes antes de las elecciones y durante todo el curso de la campaña la intención de voto por Truman había crecido constantemente.

Muchos años después, Gallup, reconoció hidalgamente su error de 1948, atribuyéndoselo al empleo de una muestra no probabilística.

¹⁴ Véase de Roger Jowell et.al. : "The 1992 british election: the failure of the polls", Public Opinion Quarterly N° 57, 1993, págs. 238-263.

Incluso en Chile tenemos una mala experiencia con las predicciones de los exit-polls con muestras de cuotas de Adimark para el Canal 7 y Time para el Canal 13 en las elecciones de congresales de 1993.¹⁵

Novena afirmación: "Ningún diseño muestral sería en definitiva probabilístico, porque se reemplaza a los entrevistados probabilísticos que no se encuentran o rechazan cooperar".

¡Esto no es así! Estamos de acuerdo en que en el muestreo probabilístico no se debe reemplazar a quienes corresponda entrevistar aleatoriamente. Sin embargo, no debe atribuirse al diseño probabilístico, como si fuera una característica propia de él, la mala práctica de algunos investigadores que sí reemplazan los rechazos o los casos de "no habidos". Este diseño -el probabilístico- no permite los reemplazos¹⁶. Los institutos que trabajan correctamente y no reemplazan deben

¹⁵ Para un análisis de los exit-polls realizados por Adimark y Time en 1993, véase de Gustavo Martínez y Miguel Ángel López : " El exit-poll en las elecciones chilenas de 1993: las enseñanzas de un fracaso". Revista "Política", Instituto de Ciencia Política , Universidad de Chile, Santiago, 1996, Vol. 34.

Con posterioridad a la publicación de este artículo, Roberto Méndez reconoció honestamente en el Seminario sobre "Periodismo de Precisión" realizado en el Centro de Estudios de la Prensa de la Universidad Católica, que su error había consistido en haber trabajado con una muestra de cuotas.

¹⁶ Desgraciadamente, algunos institutos que trabajan con diseños probabilísticos, reemplazan los hogares deshabitados por otros de aparentemente similares características de la misma manzana o de alguna adyacente. Ciertamente esto es un error porque, al proceder de esa manera, se alteran las probabilidades de selección al nivel de hogares y se pueden introducir sesgos de insospechadas consecuencias. De partida, están reemplazando hogares deshabitados o en que no hay nadie, por hogares habitados o en los que hay alguien. ¿Quién asegura que, siendo disímiles en esta característica, sean sin embargo similares en las características bajo estudio?

Muchos investigadores creen que es una especie de pecado no completar el número de entrevistas contemplado en el diseño original de la muestra, en circunstancias que al reemplazar están cometiendo un daño mucho mayor al alterar las probabilidades de selección e introducir sesgos.

sí tener la honestidad de comunicar a sus clientes o usuarios, que la muestra no refleja la totalidad del universo de estudio original, sino sólo aquella proporción del universo que, en caso de haberse efectuado una "cobertura completa idéntica" -en lugar de haberse tomado una muestra- habría contestado¹⁷.

Décima afirmación: "Los expertos europeos se matan de la risa con los que trabajan en base a muestras probabilísticas".

¿Qué expertos europeos se matan de la risa por el empleo de muestras probabilísticas? Al menos, si se ríen, seguramente deben hacerlo en privado, pues estamos seguros que no se atreverían a hacerlo en público para no caer en el descrédito profesional.

No conocemos ningún "experto europeo" que haya publicado en los últimos dos decenios algún artículo defendiendo la fundamentación científica de las muestras de cuotas. Ni en revistas de estadística de prestigio como el *Journal of the American Statistical Association* o el *Journal of the Royal Statistical Society*. Tampoco en revistas de opinión pública importantes, como el *Public Opinion Quarterly*.

Creemos que sería importante nombrar los "expertos" que defienden el muestreo de cuotas y conocer sus fundamentos.

¹⁷ Véase de W.E. Deming: "Sample Design in Business Research", John Wiley and Sons, New York, 1960, pág. 50, para una muy precisa definición del concepto del "equal complete coverage" traducida por nosotros a "cobertura completa idéntica".

La práctica del muestreo de cuotas por parte de algunos investigadores, particularmente por los del Centro de Investigaciones Sociológicas de Madrid (CIS), sólo encuentran como argumento de justificación de su uso, el que su empleo es más barato y rápido, pero no se fundamentan en su calidad o en la precisión de sus resultados¹⁸.

Realmente son muy pobres los argumentos de aquellos que están dispuestos a sacrificar calidad y precisión por defender el uso de un método cuyo único mérito consiste en su bajo precio relativo y su presumiblemente mayor rapidez.

¿De qué sirven resultados obtenidos a un menor precio, pero carentes de precisión o al menos de indicaciones sobre su precisión?

¹⁸ Tal es la posición de los que defienden honestamente el uso del muestreo de cuotas. Fuera de la rapidez y menor costo, no hay realmente otro argumento. Al mismo tiempo se reconocen sus limitaciones en comparación al muestreo probabilístico.

Es por ello que nos sorprenden las palabras de Jacinto Rodríguez cuando, después de afirmar que con las muestras de cuotas "no se pueden hacer estimaciones rigurosas con medición de errores e intervalos de confianza", con lo cual estamos de acuerdo, añade lo siguiente: "Por ello, cuando se trata de encuestas estatales y, en general, de encuestas que se toman como base para decisiones importantes, hay que recurrir a muestras totalmente probabilísticas, a pesar de su costo y del tiempo que se necesita para la realización de una rigurosa recogida de la información en el trabajo de campo. Es la garantía para poder medir el grado de precisión de las estimaciones".

Véase de Jacinto Rodríguez : "Métodos de Muestreo", Cuadernos Metodológicos. Centro de Investigaciones Sociológicas (CIS), Madrid, 1991, pág. 45.

Si bien es consistente en sus afirmaciones, desconcierta el hecho que haga distinciones entre los estudios trascendentes (estatales y de información básica para decisiones importantes) y un complemento que no se especifica (¿estudios del sector privado?). Aparentemente, en estas últimas no interesaría la precisión de las estimaciones y se estaría aceptando en forma implícita, el uso de métodos muestrales no rigurosos. Curiosa distinción.

En el caso de algunos estudios de mercado, como por ejemplo en aquellos que se trata de estimar la demanda potencial de un nuevo producto o servicio, de cuyos resultados se deducen conclusiones y luego se toman decisiones de las cuales dependen generalmente inversiones considerables, ¿se atrevería un empresario sensato a ahorrar en el método de muestreo, para luego lamentarlo?

Décima primera afirmación: "El muestreo probabilístico requeriría tener un listado actualizado de las personas (con sus domicilios suponemos), para poder aplicarlo".

De ninguna manera ello es necesario para aplicar el muestreo probabilístico.

Nos parece que Marta Lagos y Carlos Huneeus confunden los requerimientos de elaboración del muestreo aleatorio simple para individualizar las unidades muestrales, con los requerimientos de marcos muestrales para muestras probabilísticas de áreas, como por ejemplo muestras multietápicas, estratificadas geográficamente, de conglomerados. No es necesario tener ningún listado. Basta la información del INE, (y ni siquiera necesita estar actualizada) y buenos mapas, idealmente actualizados para asegurar un adecuado cubrimiento del área geográfica que ocupa el universo de estudio. Esto es suficiente.

Décima segunda afirmación: "Las muestras probabilísticas deben asignar a cada entrevistado igual probabilidad de selección".

Esto es un error muy generalizado. En el muestreo probabilístico se trata de otorgarle una probabilidad positiva y conocida a cada unidad de muestreo.

El muestreo con equiprobabilidad de selección sólo es un caso particular. Es muy conocido el uso de probabilidades desiguales, como el muestreo con probabilidades proporcionales al tamaño (usando para estos efectos una variable altamente correlacionada con las variables de estudio más importantes).

Además, debe notarse que en los estudios probabilísticos de opinión pública, basados en diseños multietápicos, y en los que se selecciona aleatoriamente una persona por hogar, tampoco las personas contenidas en la muestra tienen igual probabilidad de ser seleccionadas, independientemente de si se recurre al reemplazo o no. Y esto no las invalida. El problema se resuelve utilizando el estimador adecuado para obtener estimaciones con el menor error total.

III. EL MUESTREO PROBABILISTICO Y LOS MALTENDIDOS EN TORNO A SU SIGNIFICADO

Desde un punto de vista estricto y riguroso, una muestra probabilística puede ser definida como aquel tipo de muestra en que cada una de las unidades o elementos que componen el universo de estudio o población objeto del estudio tiene *una probabilidad conocida o calculable, y positiva*, de aparecer en la muestra¹⁹.

¹⁹ Algunos estadísticos suelen diferenciar entre universo y población. Un universo sería un conjunto de objetos (personas, establecimientos comerciales, predios agrícolas, etc) de interés para un estudio. Por ejemplo, los chilenos de 18 años de edad y más, que viven en hogares particulares del Chile continental. Respecto de este universo podemos estar interesados en sus estaturas, sus pesos, sus preferencias partidarias, etc.

Cada una de estas variables, y su distribución a lo largo de todas las unidades del universo, generan lo que podemos llamar una "población estadística" la cual queremos analizar y de la cual queremos estimar diversos parámetros, como la media, una proporción, un total, la varianza, etc. Así estaríamos hablando de la población de las estaturas, de la población de ingresos, etc., que corresponden, en nuestro ejemplo, al universo ya definido de chilenos. En otras palabras, el universo no tiene parámetros, en cambio, las poblaciones estadísticas formadas por las variables de interés de ese universo, sí tienen. Sin embargo, es bastante común detectar el uso intercambiado de estos dos términos, como si fueran sinónimos.

Esta definición parte de la noción de probabilidad de un suceso, en este caso la extracción de las unidades que constituyen el universo o población, señalando que la probabilidad de selección tiene que ser conocida (a priori) al momento de selección de la muestra o conocible (calculable a posteriori) a tal proceso²⁰.

Ello conduce a la noción de marco muestral, que consiste en un medio material (listado o combinación de listados, o material cartográfico y antecedentes estadísticos relevantes) que permite llevar a la práctica la selección de la muestra. A los elementos constitutivos de esta muestra -que no es otra cosa que un subconjunto aleatorio extraído del universo- se les aplicará luego un instrumento de medición (el que normalmente será un cuestionario o cédula de entrevista) que permitirá al investigador obtener la información relevante para los efectos del estudio²¹.

²⁰ En los diseños aleatorios simples en que se tiene, a través de un listado completo y actualizado de todos los elementos de un universo -como podría serlo el caso de los alumnos de una universidad registrados en un año determinado- la probabilidad de selección es conocida a priori.

En cambio, en las muestras multietápicas de áreas geográficas, si bien se pueden controlar las probabilidades de selección hasta el nivel de manzanas, y en algunos diseños hasta al nivel de hogares, no se conoce a priori la cantidad de habitantes en cada uno de los hogares que van a componer la muestra. Por lo tanto, se desconoce a priori la probabilidad de selección del entrevistado aleatorio al interior de cada hogar, cuando la intención del estudio es de interrogar sólo a una persona seleccionada aleatoriamente dentro del hogar. Sin embargo, en estos casos la probabilidad final puede ser calculable a posteriori.

²¹ Cuando se dispone de un listado de elementos, como ocurre con las muestras de organizaciones o de individuos, el listado constituirá el marco muestral.

En las muestras de áreas geográficas el marco muestral estará compuesto por el conjunto de material estadístico disponible, generalmente censal, y el material cartográfico correspondiente

Se señala que la probabilidad de selección de los elementos del universo debe ser positiva, vale decir superior a cero e inferior o igual a 1, porque si fuere cero para alguna unidad muestral, ésta no podría aparecer en la muestra.

El valor de 1, o sea, certeza absoluta de aparecer en la muestra se acepta, puesto que hay numerosos diseños en que algunas unidades muestrales entran a la muestra con seguridad, por su importancia. Esto es especialmente cierto en diseños en que se utilizan probabilidades desiguales de selección, o en muestras estratificadas en que las unidades muestrales de algún estrato entran todas a la muestra con certeza, vale decir, con probabilidad 1.

En los diseños muestrales complejos, como lo son las muestras de áreas utilizadas generalmente en los estudios de opinión pública y preelectorales, se procede a obtener la muestra mediante selecciones aleatorias en múltiples etapas, en cada una de las cuales se seleccionan unidades muestrales de nivel jerárquico inferior a la de la etapa anterior, como serían en Chile, para las zonas urbanas: comunas, distritos, zonas censales, manzanas, hogares, personas.

Así por ejemplo, a nivel nacional, se puede seleccionar una muestra aleatoria de comunas, de cada una de las cuales se extrae a su vez una muestra aleatoria de distritos. De cada distrito seleccionado se podría extraer, a su vez, una muestra aleatoria de zonas censales²².

²² El INE define "zona censal" como un conjunto de manzanas que pueden ser censadas por un encuestador o grupo de encuestadores en un sólo día, cuando se levanta el Censo Nacional de Población y Viviendas.

Siguiendo con nuestro ejemplo, de cada zona censal se podría extraer una muestra aleatoria de manzanas y, de cada una de las manzanas seleccionadas, una muestra aleatoria de hogares. Éste es sólo un ejemplo de las muchas variantes y alternativas que existen en un así llamado muestreo multietápico de áreas.²³

De modo tal que, si bien al decir que en las muestras probabilísticas las unidades o elementos que componen el universo deben tener una probabilidad conocida o conocible de aparecer en la muestra resulta ser correcto para muchos diseños muestrales probabilísticos tales como el muestreo aleatorio simple, esto no es tan fácil de lograr en muestras multietápicas.

Por ello nos parece más apropiado, para estos últimos casos, una definición que califique como probabilística una muestra cuyas unidades muestrales finales (generalmente personas) tienen todas y cada una de ellas una probabilidad calculable de selección. Esta cualidad es indispensable para poder calcular luego los pesos o inflatores por los que se debe multiplicar los resultados de cada una de esas personas para efectuar la inferencia al universo de estudio.

Para lograr ese objetivo, es esencial conocer las probabilidades de selección de las unidades muestrales de cada etapa, en forma independiente para cada una de las etapas si el muestreo es multietápico. Luego el producto combinado de ellas nos proporcionará la así llamada fracción muestral general (o final).

Sin embargo, se debe destacar que en la última etapa de selección, en un diseño multietápico, la probabilidad de selección sólo se conoce en el momento mismo de empadronar a los ha-

²³ Cada país tiene su propia división geográfica. En el caso de Chile urbano, el INE divide las comunas en distritos censales, los que a su vez se subdividen – para facilitar el proceso censal – en zonas censales, luego éstas en manzanas y éstas en viviendas. En cuanto a las zonas rurales, la subdivisión jerárquica del INE consiste en: provincias, comunas, distritos censales, localidades pobladas y entidades.

bitantes del hogar que, por definición, deben integrar el universo de estudio. Normalmente esta probabilidad difiere de un hogar a otro, puesto que el número de habitantes que cumplen los requisitos para el estudio también varía de un hogar a otro.

Así, en un hogar de 4 personas que califican para ser entrevistados, la probabilidad de cada uno de ellos es de $1/4$, comparado con la probabilidad de 1 en los hogares en que sólo habita una persona que califica para el estudio y, por lo tanto, es segura su selección para la muestra (si el hogar fue seleccionado en la muestra de hogares).

En términos de inflatores utilizados en los estimadores estadísticos, y si pensamos en un caso hipotético en que llegamos a la etapa muestral a nivel de hogares y en el cual logramos la autoponderación, vale decir, que todos los hogares tengan igual probabilidad de pertenecer a la muestra, entonces el inflator de la persona seleccionada de un hogar de 4 personas que califican para el estudio es 4 veces superior al de la persona que fue seleccionada de un hogar en que sólo él o ella calificaba para el estudio.

Por lo tanto, al decir que la probabilidad de selección de cada unidad muestral final (personas en este ejemplo), tiene que ser conocida o calculable y positiva se prefiere usar, en el muestreo multietápico, el concepto de "calculable" en lugar de "conocida".

De acuerdo a muchos manuales de Metodología de la Investigación, lo que define una muestra probabilística sería la circunstancia de que todas las unidades o elementos del universo o población deberían tener la misma probabilidad de ser seleccionadas (equiprobabilidad de selección o autoponderación)²⁴, lo cual constituye una creencia generalizada de muchas personas.

²⁴ Éste es un concepto equivocado. Pueden extraerse muestras con probabilidades iguales de selección para todos los elementos del universo, como también se pueden extraer con probabilidades desiguales. Por ejemplo, en

Ello no constituye una idea correcta, porque si bien es cierto que toda muestra equiprobable es probabilística, como ocurre en el muestreo aleatorio simple, lo contrario no es cierto. Así tenemos que existen muestras en que sus unidades tienen probabilidades distintas de selección, incluso por ser diseñadas para que así ocurra, como es el caso del muestreo con probabilidades proporcionales al tamaño (de cada unidad), o en el caso de muestreo estratificado con afijación desproporcional a los estratos. Sin embargo, seguimos contando con una muestra probabilística y continuamos hablando -y con toda propiedad- de muestreo probabilístico.

Es precisamente el conocimiento que se tiene de las probabilidades de selección de cada unidad de la muestra lo que permite finalmente utilizar los inflatores correspondientes a cada uno de ellos y, en consecuencia, utilizar fórmulas de estimación ("estimadores") idealmente insesgados. Finalmente, es también este conocimiento de las probabilidades, lo que nos permite calcular los márgenes de error muestral y realizar inferencias científicamente fundadas.

También es corriente encontrar en algunos textos de Metodología de la Investigación, como igualmente en algunos medios de comunicación, que se usen como sinónimos los tres siguientes términos: "probabilístico", "aleatorio" y "al azar".

Es necesario clarificar que las muestras al azar no son probabilísticas, aun cuando en algunas circunstancias cumplan con cierta apariencia de ausencia de subjetividad en la selec-

el caso del muestreo de conglomerados monoetápico, la fracción de muestreo suele ser una constante. Es decir es idéntica para todos los conglomerados y los elementos que lo componen, pero existen situaciones en que el investigador desea extraer las unidades con probabilidades desiguales, como podría ser en el caso de una muestra de establecimientos industriales, en la que éstos son extraídos con probabilidades proporcionales a su tamaño. El tamaño podría ser definido como: facturación del año anterior, o número de trabajadores de la firma, u otra variable que podría ser considerada como indicativa de tamaño por parte del investigador.

ción de los entrevistados. Sin embargo, esta ausencia de subjetividad en la selección de la muestra, por sí sola, no es suficiente. Se requiere también que las probabilidades de selección sean conocidas o calculables, y esa posibilidad no existe en las muestras al azar.

Las muestras al azar, como aquellas por ejemplo que se efectúan en una calle concurrida, o en las plazas u otros sitios públicos no pueden ser consideradas como probabilísticas. Otro ejemplo lo constituyen las encuestas que se efectúan a la salida de las estaciones del Metro de Santiago, donde se entrevista a una de cada "n" personas que salen por algún punto al exterior. En estas encuestas, como en otras similares, no se puede determinar de qué universo se está hablando y no se puede conocer la probabilidad de selección de las personas. En consecuencia, no podemos inferir a un universo más amplio que el pequeño número de personas que salen de esas estaciones determinadas durante el tiempo específico que los encuestadores están trabajando en ella²⁵.

²⁵ En este tipo de muestras no existe ni un universo de estudio ni un marco muestral que permita la aleatoriedad de la selección. El caso más patético del uso de estas muestras son las mediciones coincidentales de sintonías radiales en automovilistas. Los datos que se estiman se obtienen de un número reducido de cruces de calles y el automovilista sólo es medido si la luz roja del semáforo lo detiene y el entrevistador alcanza a encuestarlo.

Todos los automovilistas que escuchan radio cuando se movilizan en sus automóviles tienen una probabilidad 0 (cero) de aparecer en la muestra en a lo menos alguna de las cuatro situaciones siguientes:

a) Cuando no atraviesan por los cruces donde están los entrevistadores. Éstos deben ser la mayoría de los conductores porque el número de puntos de entrevistas es limitado y, además, no elegidos aleatoriamente.

b) Cuando atravesando por esos "cruces" de calles los atraviesa porque dispone de luz verde y no se detiene.

c) Cuando atravesando los cruces de las calles que constituyen puntos de muestreo y deteniendo su coche por la presencia de la luz roja, el entrevistador no alcanza a formularle las preguntas correspondientes.

d) Cuando los automovilistas no salen de sus casas en los días de las entrevistas, y el conjunto de automovilistas que así se comportan no puede

Por lo tanto, no parece recomendable referirse a las muestras al azar como si este término fuera sinónimo de "muestras probabilísticas".

En cambio, los términos "probabilístico" y "aleatorio" sí son sinónimos y su uso puede intercambiarse libremente. Por eso no entendemos que a veces se lea en los medios escritos que un instituto informe que su muestra fue "probabilística y aleatoria". Es como si dijéramos que una reina de belleza es "bella y hermosa". Basta con uno de los dos términos. *Si la muestra es probabilística, entonces también es aleatoria.*

Existen aún en la literatura otros conceptos que se confunden. Por ejemplo, se suele identificar el "muestreo aleatorio" con el "muestreo aleatorio simple", siendo este último sólo un caso particular del primero, vale decir, es sólo uno más de los innumerables diseños que se pueden plantear. En consecuencia, no se debe utilizar el término "muestreo aleatorio simple" (o "MAS" como también se le conoce), para cualquier diseño probabilístico y menos aún utilizar las fórmulas del "MAS" para el cálculo del error muestral de las estimaciones obtenidas de la encuesta ²⁶.

ser considerado como una muestra aleatoria representativa del total de automovilistas.

¿Qué valor inferencial podrían tener estos datos para estimar adecuadamente las preferencias de sintonías radiales de los automovilistas? ¿Podrían servir para determinar opciones racionales de localización de comerciales publicitarios, por ejemplo?

No tenemos otra explicación para esta práctica que la del desconocimiento de las carencias de este tipo de muestras "al azar"

²⁶ No es descartable la explicación de que tales caracterizaciones de las muestras probabilísticas se hagan así por razones pedagógicas clarificadoras de la necesidad de selección aleatoria o bien como un reflejo de lo que es el muestreo aleatorio simple con reposición (muestreo irrestrictamente aleatorio).

Véase de Earl Babbie : "The practice of social research", Wadsworth Publishing Co., Belmont, California, 1983, pág. 145; de David Dooley "Social Research Methods", Prentice Hall, Englewood Cliffs, N. Jersey 1995, pág. 134.

Las fórmulas de estimación, como igualmente las fórmulas para el cálculo del error muestral varían de un diseño a otro y siempre deberían emplearse las fórmulas que correspondan al diseño que se utilice.

IV. EL MUESTREO NO-PROBABILISTICO: MUESTRAS DE CUOTAS Y MUESTRAS SEMIPROBABILISTICAS

Al contrario de lo que ocurre con las muestras probabilísticas, en las no-probabilísticas no es posible calcular (ni a priori ni a posteriori del proceso de recolección de datos) la probabilidad de selección de las unidades que componen el universo.

Para los practicantes del muestreo no-probabilístico, todo el proceso de asignación de probabilidades deja de ser un problema. Felizmente para ellos, pero desgraciadamente para los usuarios de la información que buscan obtener datos confiables, este proceso simplemente se elimina o se pasa por alto.

Las muestras no-probabilísticas reciben distintas denominaciones. Muestras disponibles, intencionales, de casos típicos y de juicio son nombres de uso corriente para designar sus variedades²⁷.

Pero, en definitiva, su común denominador lo constituye el hecho de que es el propio investigador o los entrevistadores que colaboran en el estudio con la recolección de los datos, los que determinan -aunque ello se haga en forma inconsciente- quiénes constituirán la muestra. Esto marca su diferencia con el muestreo probabilístico en el cual la selección queda en manos de un proceso aleatorio.

²⁷ Serían muestras disponibles aquellas en que el investigador compone la muestra con las personas que tiene a su alcance, como ocurriría con un profesor y sus alumnos en alguno de sus cursos. Muestra de juicio, intencional o de propósito es aquella en que el propio investigador elige y descarta -de acuerdo con sus propios criterios- quiénes compondrán la muestra.

Las muestras de casos típicos están constituidas por aquellas muestras donde se supone que los miembros elegidos representan el promedio respecto de una o más variables.

En cualquier caso, esto se debe a la falta de transparencia del método de selección, el cual carece de fundamentos matemáticos, lo que le impide además calcular los márgenes de error muestral. Esta limitación también imposibilita al investigador calcular a priori cuál sería el tamaño muestral adecuado para lograr determinada precisión en los resultados, y, a posteriori, también le hace imposible calcular los márgenes de error muestral de los resultados obtenidos.

Un cierto mayor refinamiento se encuentra en las llamadas muestras de cuotas, las que parten de una mala interpretación de la idea de la representación estadística.

En las muestras de cuotas, se procura componer una muestra que "represente" el universo en algunas de sus características esenciales. Entonces, si en el universo hay hombres y mujeres, se tratará que las cuotas de entrevistados de la muestra tengan hombres y mujeres, y en la medida de lo posible, en las mismas proporciones que en el universo.

Si el universo contiene personas de distintos tramos etarios, se intentará que esos diversos tramos etarios estén "representados" -como se suele decir- en la muestra y ojalá en las mismas proporciones que en el universo.

Ahora, ¿de dónde se obtienen los componentes individuales de la muestra? La respuesta es que de cualquier parte, con tal que los miembros seleccionados respondan a los criterios definidores de la combinación de variables socio-económicas y demográficas que caracterizan el universo y que deberían caracterizar la muestra (por ej. mujeres de nivel socioeconómico alto, entre los 18 y los 24 años de edad).

Cuando se procede así, se cree erróneamente que se tiene una muestra representativa del universo, porque "todos" los sectores estarían representados y, además, en las mismas proporciones que en el universo. Esta concepción de "representatividad" involucra el concepto de que la muestra, para ser representativa, debiera ser un universo en miniatura. La expresión

misma de "muestra representativa" es equívoca y bien haríamos, como lo sugiere Leslie Kish, en eliminarla de nuestro vocabulario estadístico por ser ambigua y carente de significado matemático²⁸.

Agreguemos que su uso lleva a un malentendido muy perjudicial porque hace pensar que lo que importa es que ciertas características del universo estén representadas proporcionalmente en la muestra, en lugar de pensar en la capacidad de inferir al universo que debe tener toda muestra útil y científicamente defendible.

Es mejor pensar, de no eliminarse el término, que una muestra será representativa o podría serlo si sus resultados logran ser inferidos al universo al mismo tiempo que lo están dentro de ciertos límites aceptables de precisión.

El máximo de sofisticación de las muestras de cuotas se logra con lo que Seymour Sudman ha llamado muestras probabilísticas de cuotas. Francisco Azorín las llama muestras semi-probabilísticas y otros las llaman muestras semi-probabilísticas de cuotas²⁹.

²⁸ Véase de Leslie Kish: "Survey Sampling", John Wiley and Sons. New York, 1965, cáp. 1 y del mismo autor "Selection of the sample", en Leon Festinger y Daniel Katz: "Research Methods in the behavioral sciences"; Holt, Rinehart and Winston, New York, 1966.

²⁹ Azorín Poch distingue entre los así llamados muestreo semi-probabilístico superior e inferior. Sería superior aquel donde es posible asignar -en el muestreo multietápico- probabilidades de selección a las unidades muestrales de mayor jerarquía (primeras etapas) pero no a las de inferior jerarquía (última(s) etapa(s)). Al revés, sería muestreo semi-probabilístico inferior aquel en que ocurre lo contrario. No vemos la importancia de esta diferenciación ni su utilidad práctica, porque las probabilidades de selección sólo tienen sentido si todas las unidades pertenecientes a todas las etapas de selección en el muestreo multietápico, tienen probabilidades conocidas o conocibles de selección.

Véase de Francisco Azorín y José Luis Sánchez-Crespo: "Métodos y aplicaciones del muestreo", Alianza Universidad Textos, Madrid 1986, pág. 20.

Bajo estas diferentes denominaciones se presenta un procedimiento de extracción de la muestra que tiene características similares de selección con las muestras probabilísticas –en los diseños muestrales multietápicas de áreas– hasta llegar al nivel de selección de manzanas. De ahí para adelante, ambos procedimientos muestrales divergen absolutamente.

En el muestreo probabilístico, seleccionada una manzana determinada, la selección de las viviendas también es aleatoria. Uno de los métodos más comunes de seleccionar una muestra aleatoria de viviendas dentro de una manzana, entre varios otros métodos, consiste en empadronarla y luego seleccionar aleatoriamente una vivienda como punto de partida para posteriormente seleccionar el resto de las viviendas aplicando un paso sistemático.

En las viviendas así seleccionadas, y en aquellos estudios de opinión o de mercado en los se requiera las respuestas de una sólo persona del hogar, se extrae también mediante un procedimiento aleatorio, a una sólo persona de entre aquéllas que cumplan con las definiciones requeridas para pertenecer al universo de estudio, luego de empadronarlas (estén o no, en ese momento, en la vivienda). Esta selección se efectúa normalmente a través del uso de una tabla de Kish.

En cambio, en el muestreo semi-probabilístico de cuotas, una vez seleccionada una manzana determinada, los entrevistadores siguen una ruta prefijada, casa a casa, buscando a personas que satisfagan las instrucciones exigidas por las cuotas. De no completarse las cuotas asignadas a esa manzana, el entrevistador continúa, siguiendo criterios de contigüidad espacial o “rutas”, a otras manzanas adyacentes.

Independientemente que los resultados de las muestras se sesguen a favor de las respuestas de aquellas personas que permanecen más tiempo en sus viviendas, por lo que tienen mayores probabilidades de ser encontradas en sus hogares y, en consecuencia, de pertenecer a las muestras obtenidas por el método de cuotas (dueñas de casa, jubilados, cesantes), no exis-

te ningún control sobre la probabilidad de selección de dicha persona dentro de hogares que normalmente tienen diferentes cantidades de habitantes entre los que se efectúa la selección. Sólo importa si cumple con las características que se necesitan para pertenecer a la cuota que se está procurando llenar.

Además, por el método de cuotas, se pierde el concepto de "conglomerados", el cual es típico en el muestreo de áreas y por ende, del muestreo probabilístico. En éste se definen los conglomerados a priori, es decir, al interior de oficinas utilizando el material cartográfico del que se dispone, aun en etapas del diseño muestral.

En el muestreo semi-probabilístico de cuotas, en cambio, una muestra de 10 viviendas, por ejemplo, puede ser extraída en algunos casos de una sólo manzana y, en cambio, en el mismo estudio pero en otros lugares de la ciudad, puede que sea necesario recorrer varias manzanas antes de lograr las 10 entrevistas. Este último caso es típico en los barrios de viviendas de alto nivel socioeconómico²⁹. Esto también contribuye a destruir cualquier asomo de similitud con las muestras probabilísticas.

Así es como los investigadores que trabajan en base a muestras semi-probabilísticas de cuotas, saben que para inflactar sus resultados a nivel del universo, deben utilizar estadísticas

²⁹ Con el método de cuotas se pierde el sentido estadístico del concepto de la conglomeración. En el muestreo probabilístico, en cambio, específicamente en el muestreo multietápico de áreas, el conglomerado es una unidad territorial estrictamente definida, sin ambigüedades de límites. Esto significa que cada vivienda de la ciudad está ubicada en un solo conglomerado, sea éste una sólo manzana o un conjunto predeterminado de manzanas. No existen indefiniciones de localización como es el caso en las muestras semi-probabilísticas de cuotas.

Para una descripción de uno de los métodos de selección en muestreo semi-probabilístico de cuotas, véase de Francisco Javier Barranco: "Técnicas de Marketing Político", Red Editorial Iberoamericana. México, 1994, Cap. 2, págs. 51-58.

secundarias en lugar de los recíprocos de la probabilidad final de selección.

Sin embargo, ellos no reparan en que estas estadísticas secundarias son de dudosa aplicación por estar normalmente desfasadas en el tiempo y pertenecer a áreas geográficas que no necesariamente coinciden con el área de estudio.

V. EL MUESTREO PROBABILISTICO CONSTITUYE EL UNICO TIPO DE MUESTREO DE FUNDAMENTACION CIENTIFICA

No negamos que el muestreo de cuotas es de utilidad en el trabajo científico. Por de pronto, la tiene en estudios exploratorios, en estudios piloto (pre-tests) que testean la comprensión de la redacción de los instrumentos de medición de una encuesta por muestreo, y en los estudios cualitativos del tipo focus-group o entrevistas focalizadas (focused interviews).

Lo que no es admisible es pretender que a partir de los resultados de un estudio en base a muestras de cuotas se pueda estimar el valor de los parámetros poblacionales, y menos aún que se puedan estimar los intervalos de confianza, vale decir, aquellos intervalos dentro de los cuales estarían los parámetros en cuestión con un cierto grado de probabilidad.

En tal sentido, la muestra de cuotas sólo se representa a sí misma y carece de propiedades inferenciales al universo, que es siempre lo que interesa cuando se trabaja con muestras.

Al respecto, conviene tener siempre en cuenta que una muestra tiene valor sólo como un medio para conseguir un fin, y éste no es otro que lograr estimar los parámetros poblacionales cuando el estudio es descriptivo, o lograr establecer correlaciones o relaciones de causa a efecto cuando el estudio es de carácter analítico.

Además, el muestreo probabilístico permite comparar los resultados de un estudio con el que le precede en el tiempo en

los diseños longitudinales. También nos permite realizar tests de hipótesis para comparar cambios a través del tiempo o diferencias entre estimaciones muestrales, a cierto nivel de significación estadística, análisis que sólo tienen sentido si las muestras sobre las cuales se basan las estimaciones en cuestión son probabilísticas.

Nada de esto es posible con muestras de cuotas o muestras semi-probabilísticas de cuotas, en que tales interpretaciones aparecen como conejos salidos del sombrero de un mago prestidigitador.

No tiene sentido hablar de muestras semi-probabilísticas porque no hay "semi-probabilidad".

Algunos investigadores, en su trabajo, se encuentran a veces enfrentados a un dilema similar al de Hamlet, pero esta vez trasladado al terreno de las estadísticas: trabajar con muestras probabilísticas o trabajar con muestras no-probabilísticas.

Algunos creen que pueden soslayar este dilema utilizando lo que eufemísticamente se suele llamar muestreo semi-probabilístico. Pero esto sólo es una ilusión. En realidad, no existe el llamado muestreo semi-probabilístico, aunque con ese nombre el lector incauto crea que está frente a un diseño casi tan bueno como un diseño probabilístico puro.

En matemáticas, si multiplicamos varios eslabones de una cadena multiplicativa y uno de estos eslabones es cero, entonces el producto entero de la multiplicación también es cero.

En muestreo probabilístico sucede algo similar. Para calcular las fórmulas de estimación o "estimadores" de los parámetros de interés se trabaja con la "fracción muestral final". Estas fracciones muestrales finales son el producto de la multiplicación de las fracciones muestrales de las varias etapas de un diseño multietápico, y, tal como ocurre con la cadena multiplicativa en matemáticas, en este caso basta con que una sólo de las fracciones muestrales de cualquier etapa se desconozca para que

el producto de todas ellas se pierda y no pueda ser utilizado por el estimador. Y justamente éste es el caso en las muestras semi-probabilísticas; se pierde el hilo de las probabilidades y la inferencia al universo nuevamente debe pasar por un proceso de ponderación utilizando estadísticas secundarias.

Una ventaja adicional en el caso del muestreo probabilístico es que la distribución de muestreo es conocida gracias al Teorema del Límite Central. Dicho conocimiento permite obtener la información requerida para calcular los márgenes de precisión de los resultados obtenidos, a determinados niveles de confianza. Éstos no son otra cosa que la probabilidad de que el parámetro buscado esté ubicado al interior del intervalo de confianza formado por el estimador al que se le suma y se le resta, simultáneamente, el margen de error muestral³⁰.

Por último, con muestras probabilísticas es posible comparar los resultados de dos categorías de respuesta de una misma pregunta y ver si las diferencias encontradas son o no estadísticamente significativas y, por lo tanto, dicha diferencia puede o no inferirse al universo.

Un buen ejemplo de ello es lo acontecido con los resultados del estudio de opinión pública de diciembre de 1997, efectuado por el Centro de Estudios Públicos.

En dicho estudio, las estimaciones de punto señalaban que el 37% desaprobaba la gestión del Gobierno de Eduardo Frei y el 34% lo aprobaba. Todos los periódicos difundieron la noticia de que por primera vez durante el mandato de Frei, eran más los que estaban en desacuerdo con su gestión que aquellos que la respaldaban.

³⁰ Véase una muy ilustrativa, simple y práctica explicación de esta distribución de muestreo en "Sample Survey Methods and Theory", Vol. I, de Hansen, Hurwitz and Madow, John Wiley and Sons, New York, Sexta edición, 1965, págs. 20-31.

Los datos no fueron acompañados de los resultados de los tests de hipótesis correspondientes, y la interpretación fue difundida con caracteres destacados para mostrar que Frei tenía más detractores que defensores.

Al efectuar el test de hipótesis correspondiente, nosotros constatamos que la diferencia entre las proporciones no era estadísticamente significativa al nivel de significación del 5%.

En consecuencia, la conclusión sustantiva del cambio en las proporciones en diciembre de 1997 entre los que aprobaban y los que reprobaban la gestión del gobierno de Eduardo Frei no puede sustentarse a nivel del universo de estudio conformado por todos los chilenos de 18 años y más de edad, inscritos y no inscritos en los Registros Electorales y que viven en hogares particulares del Chile continental.

El efectuar el test fue posible gracias a que la muestra del CEP era probabilística. Esto no se podría haber hecho con una muestra de cuotas o semi-probabilística de cuotas. Con tales muestras no hay interpretación posible.

No nos explicamos cómo aquellos que utilizan muestras no probabilísticas hacen en sus análisis de tendencias (trend studies), interpretaciones estadísticas sobre los aumentos o descensos que sufren las proporciones de ciertas variables entre un estudio y otro. Se nos asemeja a "palomitas" creadas por la varita mágica del mago en un show.

Como dice Deming -tal vez uno de los más destacados contribuyentes al desarrollo de la teoría del muestreo científico, junto con Mahalanobis- "el muestreo probabilístico y el de cuotas son dos monedas no intercambiables"³¹.

³¹ Véase de W.E. Deming: "Some Theory of Sampling", John Wiley and Sons, 1950, págs. 9-14. También de W.E. Deming: "Sample Design in Business Research", John Wiley and Sons, 1960, pág. 31.