



Berumen y Asociados, S.A. de C.V.
Altadena No. 15 Col. Nápoles,
Del. Benito Juárez Méx. D.F. CP 03810
Tel: 5093-8600 Fax: 5093-8615



VIII SEMINARIO DE ACTUALIZACION PROFESIONAL

Asociación Mexicana de Agencias de Investigación de Mercado y Opinión Pública

DE LA PARTE AL TODO VUELTA A LO BASICO

**Edmundo Berumen
17 de agosto 2001**

‘Buenos días, soy de y estamos realizando una encuesta sobre para conocer (Usted / Su hogar / Su empresa) ha sido seleccionada para representar a otras (personas / hogares / empresas) como usted en este estudio y quisiera hacerle unas preguntas. Sus respuestas son confidenciales y solo se usarán en forma agregada con otras.’.

A diario escuchamos en la radio, vemos en televisión, leemos en periódicos, revistas o informes de trabajo, los resultados de la encuesta más reciente sobre los temas más diversos. En base a ellos se compara, alaba, invierte, defiende, justifica, reprocha, ataca, modifica o, descontinúa, a productos, servicios, instituciones, organismos, administraciones, programas, proyectos, funcionarios, grupos, partidos políticos, personalidades, o sociedades enteras.

¿Pero qué son las encuestas? ¿Cómo se hacen? ¿Qué alcance tienen? La respuesta es tan amplia y diversa como el número de encuestadores, agencias de investigación, empresas o instituciones, nacionales y transnacionales, que las solicitan o realizan. Veamos.

PERTINENCIA DE LA INVESTIGACIÓN MEDIANTE UNA ENCUESTA

En general, quienes diseñan y levantan encuestas, los ‘productores’, ya sean agencias de investigación privadas u organismos públicos, usualmente lo hacen para uno o varios clientes, los ‘usuarios’, no para consumo propio. La relación **productor-usuario** para determinar el objetivo y alcance de la encuesta, siempre es difícil. El usuario tiene dificultad en expresar en términos

operativos lo que necesita y el grado de precisión y confianza que le es suficiente. Es responsabilidad del productor ayudar al usuario a definir sus requerimientos, y hacerle explícito desde el inicio las limitaciones y resultados posibles por medio de encuestas. En la ausencia de esta fase, ahí es donde nacen fallas serias de las encuestas y malas interpretaciones futuras de sus resultados. Al productor le toca hacer las preguntas difíciles al usuario (doblemente difícil al considerarlo como una fuente de ingresos importante), entre otras:

¿Por qué y para qué necesita información el usuario? El conocimiento cabal de los objetivos permitirá al productor serio y profesional señalarle al usuario alternativas a la encuesta, cuando las haya. En no pocas ocasiones las hay. Ya sea una fuente distinta, datos de registros administrativos o censales, o una encuesta reciente que investigó temas similares. Muchas variables de interés no son tan dinámicas como se piensa, y datos censales o de encuestas anteriores pueden ser suficientes.

Respecto a las encuestas **anteriores**, un patrón común es que son poco analizadas. Esto se debe por un lado a la dificultad de acceso a los datos básicos por parte de usuarios **distintos** a los que la originaron, y por otro a la facilidad con que se satisface al usuario original poco exigente. En particular, es común que a la agencia de investigación que se le está solicitando la nueva medición, no se le informe ni se le de acceso a los datos de investigaciones anteriores, lo que ocasiona desperdicio de recursos escasos. Hay una obligación ante la sociedad, el cliente y el informante, de aprovechar al máximo la información que nos proporciona la población investigada; para el mejor conocimiento del tema, la mejor toma de decisiones, y el respeto al tiempo valioso que ya nos dieron los informantes.

El argüir que no vale la pena el consultar los datos disponibles por que 'no son actuales' o 'podemos hacer la medición mucho mejor', es discutible. Es más reto para el productor el sacarle provecho a información existente, reconociendo las limitaciones que tenga, que el plantear una 'nueva y mejor' encuesta. Lo último es la salida fácil.

Finalmente, el dato de interés puede no ser medible vía una encuesta; si este es el caso, hay que hacérselo ver al usuario. Antes de volver a tocar la puerta del informante, hay que asegurarse que es necesario. Evitemos caer en el abuso con el encuestado.

EL UNIVERSO Y LAS POBLACIONES QUE LO INTEGRAN

La importancia de comprender cabalmente cuál es el universo de donde se derivan las distintas poblaciones de interés, luego tomar muestras de ellas para medir variables y parámetros asociados a éstas, y por último saber cómo realizar el viaje de vuelta, regresar de la parte al todo,

de los datos ‘muestrales’ a inferencias válidas para las poblaciones de interés, es lo cotidiano en toda investigación por muestreo. Para ello es necesario:

- Ir del todo a las partes que lo integran.
- Partir de todos ‘los mexicanos’, para reconocer a los distintos y variados grupos sociodemográficos, étnicos, regionales, urbanos-rurales, sociopolíticos, hasta llegar al elemento esencial, el individuo que conforma todo lo anterior.
- En el universo desmenuzado finalmente encontrar y rescatar en su definición la(s) ‘población(es) de interés’.

Se dice y lee pronto, pero este paso fundamental de toda investigación, por obvio se omite, y deja ya en su lugar los cimientos endebles del inicio de cualquier estudio.

Al inicio de cualquier investigación se requiere precisar los elementos sobre los que se desean obtener datos. Si bien al principio se pueden utilizar términos vagos como ‘los mexicanos’, ‘las mujeres’, ‘los niños’, ‘los adultos’, ‘los televidentes’, conforme se avanza es indispensable definir con precisión de quiénes queremos información.

Para ilustrar, tomemos un ejemplo cotidiano, al que todos estamos expuestos mes a mes. Desde hace varias décadas la Dirección General de Estadística del Instituto Nacional de Estadística, Geografía e Informática (INEGI), nos reporta estadísticas sobre “la fuerza laboral”, “el empleo”; en particular, para cada medición mensual entre otros parámetros difunde la “tasa de desocupación abierta”. Pero ¿a quién y a qué se refiere esta tasa?. Imaginemos el siguiente proceso para arribar a una definición precisa de la tasa:

1. Los ‘mexicanos que no trabajaron’, entre todos los mexicanos.
2. Los ‘mexicanos **de 12 años de edad cumplidos o mayores, residentes en el país, que no trabajaron**’, entre todos los mexicanos de **12 años de edad cumplidos o mayores, residentes en el país**.
3. Los ‘mexicanos de 12 años de edad cumplidos o mayores, residentes en **viviendas particulares** del país, que **no estando ocupados, buscaron activamente un trabajo**’, entre todos los mexicanos de 12 años de edad cumplidos o mayores, residentes en **viviendas particulares** del país, **que están desocupados u ocupados**.
4. Los ‘mexicanos de 12 años de edad cumplidos o mayores, residentes **permanentes** en viviendas particulares del país, que no estando ocupados, buscaron activamente un trabajo **en las 4 semanas previas a la semana de levantamiento**’, entre todos los mexicanos de 12 años de edad cumplidos o mayores, residentes **permanentes** en viviendas particulares del país, que están desocupados u ocupados.
5. Los ‘mexicanos de 12 años de edad cumplidos o mayores, residentes permanentes en viviendas particulares del país, que no estando ocupados, buscaron activamente un

- trabajo en las 4 semanas previas a la semana de levantamiento, **o hasta 8 semanas, siempre y cuando estén disponibles a incorporarse de inmediato**, entre todos los mexicanos de 12 años de edad cumplidos o mayores, residentes permanentes en viviendas particulares del país, que están desocupados o **trabajaron al menos una hora o un día en la semana de referencia.**
6. Los ‘mexicanos de 12 años de edad cumplidos o mayores, residentes permanentes en viviendas particulares del país, que no estando ocupados, buscaron activamente un trabajo en las 4 semanas previas a la semana de levantamiento, o hasta 8 semanas, siempre y cuando estén disponibles a incorporarse de inmediato’, entre todos los mexicanos de 12 años de edad cumplidos o mayores, residentes permanentes en viviendas particulares del país, que están desocupados o trabajaron al menos una hora o un día en la semana de referencia **a cambio de una remuneración monetaria o en especie.**
 7. Los ‘mexicanos de 12 años de edad cumplidos o mayores, residentes permanentes en viviendas particulares del país, que no estando ocupados, buscaron activamente un trabajo en las 4 semanas previas a la semana de levantamiento, o hasta 8 semanas, siempre y cuando estén disponibles a incorporarse de inmediato’, entre todos los mexicanos de 12 años de edad cumplidos o mayores, residentes permanentes en viviendas particulares del país, que están desocupados o trabajaron al menos una hora o un día en la semana de referencia a cambio de una remuneración monetaria o en especie, **o que tienen empleo pero no trabajaron por alguna causa, pero con retorno seguro a su trabajo en menos de 4 semanas.**
 8. Los ‘mexicanos de 12 años de edad cumplidos o mayores, residentes permanentes en viviendas particulares del país, que no estando ocupados, buscaron activamente un trabajo en las 4 semanas previas a la semana de levantamiento, o hasta 8 semanas, siempre y cuando estén disponibles a incorporarse de inmediato’, entre todos los mexicanos de 12 años de edad cumplidos o mayores, residentes permanentes en viviendas particulares del país, que están desocupados o trabajaron al menos una hora o un día en la semana de referencia a cambio de una remuneración monetaria o en especie, o que tienen empleo pero no trabajaron por alguna causa, pero con retorno seguro a su trabajo en menos de 4 semanas, **o que no tenían empleo pero iniciarán uno con seguridad en 4 semanas o menos.**
 9. Los ‘mexicanos de 12 años de edad cumplidos o mayores, residentes permanentes en viviendas particulares del país, que no estando ocupados, buscaron activamente un trabajo en las 4 semanas previas a la semana de levantamiento, o hasta 8 semanas, siempre y cuando estén disponibles a incorporarse de inmediato’, entre todos los mexicanos de 12 años de edad cumplidos o mayores, residentes permanentes en viviendas particulares del país, que están desocupados o trabajaron al menos una hora o un día en la semana de referencia a cambio de una remuneración monetaria o en

- especie, o que tienen empleo pero no trabajaron por alguna causa, pero con retorno seguro a su trabajo en menos de 4 semanas, o que no tenían empleo pero iniciarán uno con seguridad en 4 semanas o menos, **o que trabajaron al menos una hora o un día sin recibir pago alguno, ni monetario ni en especie, en un negocio propiedad de un familiar o no familiar.**
10. Los ‘mexicanos de 12 años de edad cumplidos o mayores, residentes permanentes en viviendas particulares **de las 45 localidades urbanas más importantes del país (que cubren alrededor del 90% de la población que habita en localidades de 100,000 habitantes)**, que no estando ocupados, buscaron activamente un trabajo en las 4 semanas previas a la semana de levantamiento, o hasta 8 semanas, siempre y cuando estén disponibles a incorporarse de inmediato’, entre todos los mexicanos de 12 años de edad cumplidos o mayores, residentes permanentes en viviendas particulares del país, que están desocupados o trabajaron al menos una hora o un día en la semana de referencia a cambio de una remuneración monetaria o en especie, o que tienen empleo pero no trabajaron por alguna causa, pero con retorno seguro a su trabajo en menos de 4 semanas, o que no tenían empleo pero iniciarán uno con seguridad en 4 semanas o menos, o que trabajaron al menos una hora o un día sin recibir pago alguno, ni monetario ni en especie, en un negocio propiedad de un familiar o no familiar.

En una simplificación excesiva, se ha sintetizado en 10 pasos las convenciones conceptuales que se derivan de 40 años de esfuerzo de la Organización Internacional del Trabajo (OIT), entre 1954 y 1993, a través de la 8ª, 11ª, 13ª, 14ª, y 15ª, Conferencia Internacional de Estadígrafos del Trabajo.

El esfuerzo anterior no sólo permite precisar a la población de interés, también proporciona guía para idear qué y a quién medir, cómo tomar la medición en la práctica; y no menos importante, cómo tomar una muestra de ella.

MUESTRAS, MUESTREO

El investigar un tema por medio de una encuesta, generalmente conduce a realizar ésta en base a una **muestra** de la población que se quiere estudiar (la excepción son los censos, en donde se *intenta* cubrir a toda la población de interés). El hacerlo así, reduce considerablemente los costos de la investigación, y aunque parezca paradójico, un diseño y tamaño adecuado de muestra permite un mejor control en la calidad, precisión y confiabilidad de los datos que se miden, que lo que puede proporcionar un censo.

La característica metodológica crucial de la encuesta por muestreo, es que el procedimiento que utilice para seleccionar a los encuestados, y el que use para procesar los datos recogidos, permita inferir conclusiones válidas sobre el tema investigado para **toda** la población de interés. Si el diseño de una encuesta **no tiene** esta característica, los resultados e interpretaciones que escuchamos, vemos y leemos, por lo general no son válidos, y en el mejor de los casos sólo se refieren a la muestra particular de encuestados.

Decir que una encuesta está basada en una "muestra representativa" o "aleatoria" de tamaño tal, con un grado de error o precisión tal, es un intento por asegurar el que se cumplió con los requisitos cruciales. Pero para quien escucha o lee lo anterior, el aceptar ésta declaración es más un acto de fe que el convencimiento razonado de que el estudio se apejó a una metodología adecuada.

Tamaño de muestra

Si la población de interés es 'pequeña', ¿para qué muestrear? En estos casos, es necesario dar seria consideración a la opción de censo.. Claro, hay situaciones en donde aún en estas circunstancias el muestreo es lo indicado, por ejemplo, cuando la medición es 'destruktiva'.

Si todos los elementos de la población son idénticos, ¿para qué medir a más de uno?

Si no estamos en ninguno de estos extremos, entonces el tamaño de la muestra es un parámetro relevante.

La diferencia entre una estimación particular y el 'valor real' (y desconocido) de lo que se está estimando, el que resulta de aplicar el mismo procedimiento de medición a todos los elementos de la población de interés, es el "error" de esa estimación particular, resultante de una muestra de las muchas posibles.

Si el tamaño de la muestra se hace cada vez más grande, hasta que se aproxima a un censo, entonces uno espera que el error se reduzca, hasta llegar a cero en el extremo de que el tamaño de muestra sea igual al tamaño de la población. Recordemos la conocida fórmula para determinar el tamaño de muestra **n** necesario para estimar el porcentaje **P** en una población de tamaño **N**, con un esquema de muestreo aleatorio simple:

$$(1) \quad n = \frac{\frac{k^2 \cdot PQ}{d^2}}{1 + \frac{k^2 \cdot PQ}{N d^2}} = \frac{n_0}{1 + \frac{n_0}{N}}, \quad \text{digamos}$$

donde:

- k : lo determina el nivel de confianza (1.96 para el 95%)
Q: 1-P
d: lo determina el nivel de precisión deseado (0.05 para el tan socorrido 5% de precisión)

Lo primero que observamos es que el cociente:

$$\frac{n_0}{N} = \frac{k^2 \cdot PQ}{d^2}$$

es despreciable a partir de que el tamaño de la población es ‘grande’, por lo que en estos casos podemos simplificar la fórmula a:

$$(2) \quad n = n_0 = \frac{k^2 \cdot PQ}{d^2}$$

En el **Cuadro 1** del **Anexo**, se presentan distintos tamaños de muestra requeridos para distintos órdenes de magnitud de **P**, niveles de precisión **d** y niveles de confianza **k**. Por ejemplo, el tamaño de muestra requerido para estimar un porcentaje del orden del **50%**, con una precisión de **±5%** a un nivel de confianza del **95%**, de acuerdo a la fórmula (2) está dado por:

$$n = \frac{1.96^2 \cdot 0.5 \cdot (1-0.5)}{0.05^2} = 384.16$$

al ‘redondear’ de ahí sale el tan conocido y recomendado tamaño de muestra de **400** casos. Si la población de interés tiene 10 millones, 1 millón, 100 mil, o 10 mil elementos, el tamaño con la fórmula (1) correcta es de:

$$n = 400 / \{1 + 400 / 10,000,000\} = 399.98$$

$$n = 400 / [1 + 400 / 1,000,000) = 399.84$$

$$n = 400 / \{1 + 400 / 100,000\} = 398.40$$

$$n = 400 / [1 + 400 / 10,000) = 384.61$$

De ahí que digamos que cuando de estimar proporciones, porcentajes o tasas se trata, el tamaño de la población es irrelevante para determinar el tamaño de la muestra a partir de poblaciones cuyo tamaño es de decenas de miles.

Son pocas las ocasiones prácticas en donde aquí termina el tema del tamaño de muestra, aquellas donde es pertinente y práctico el utilizar un esquema de muestreo aleatorio simple. Por ejemplo, las encuestas telefónicas es uno de los pocos casos en donde sí se puede utilizar de manera práctica un ilustraciones.

Ah, pero ninguna encuesta es de una sola variable, de una sola pregunta. Las consideraciones de control para las distintas fuentes de error que debe contemplar el diseño de una encuesta, se deben hacer para cada parámetro de interés. Declaraciones generales sobre la precisión y confianza que los resultados de una encuesta tienen, dicen todo y nada. ¿Para qué variables? En subpoblaciones de interés, por ejemplo, distintos grupos de edad y sexo, ¿cómo se afectan estas precisiones y confianzas para la misma variable?

Será necesario hacer compromisos y concentrarse en las variables principales, para los dominios de estudio (población o subpoblaciones para las que se requieren las estimaciones) prioritarios. Pero sobre todo, es importante calcular y presentar las estimaciones de error y confianza para las variables principales, en una forma sencilla y comprensible para el usuario. Desafortunadamente, una práctica poco común en la mayoría de las encuestas.

Dejemos inconcluso el tema, para luego retomarlo, y pasemos a diseños de muestra.

Diseños de muestra

Una de las bondades de las encuestas **probabilísticas**, las que le asignan una probabilidad conocida y distinta de cero a todos los elementos que pertenecen a la población de interés, es que permiten con los datos de la misma encuesta, esto es, sólo con las observaciones de la muestra particular que resultó seleccionada, una de las muchas posibles, estimar no sólo el parámetro de interés, si no también su margen de error, y a partir de ahí, construir un intervalo de confianza alrededor de la estimación del parámetro de interés.

Las encuestas que aseguran a **todos** los elementos de la población de interés una probabilidad, distinta de cero, de resultar seleccionados por el esquema de muestreo utilizado (sea cual sea), son

las que tienen el ingrediente principal para poder hacer inferencias estadísticamente válidas a toda la población de interés. Esto se dice fácil pero en la práctica se cumple poco.

Se requiere invertir bastante en construir y mantener un ‘marco de muestreo’ (un procedimiento que permitiría llegar a todos y cualquiera en particular de los elementos de la población de interés) que permita aseverar que el esquema de muestreo cumplió con lo anterior.

Una vez definida con precisión ‘operativa’ la población de interés, es necesario encontrar esquemas que nos permitan tomar muestras de ella de manera práctica. Continuemos con la familiar ilustración de la ‘tasa de desempleo abierto’. ¿Cómo tomar una muestra de los mexicanos de 12 años de edad o mayores, residentes permanentes en viviendas particulares del país? Las palabras clave son ‘residentes permanentes’ y ‘viviendas particulares’.

Resulta que todo elemento de la población que me interesa reside de manera permanente en alguna vivienda particular del país. Lo que entonces necesito es una manera de tomar una muestra probabilística de viviendas particulares del país, y luego dentro de las viviendas que resulten seleccionadas, tomar una muestra probabilística de los residentes permanentes de 12 años de edad o mayores. Empiezo a ver la luz.

El cómo llegar a una muestra probabilística de viviendas particulares del país, tiene tantas respuestas como muestristas, y su experiencia, ingenio y creatividad, existan en el país. Viajemos tan solo una de las posibles rutas.

1. Fácil, usemos una lista de las viviendas particulares del país, y de ahí tomamos la muestra. Ah caray! No existe tal lista, ni en el censo.
2. Bueno, las viviendas particulares de interés son estructuras físicas que están asentadas en algún terreno del territorio nacional. Tomemos entonces muestras de terrenos. Si, ¿pero cómo? Pues que tal si lo hacemos por partes, primero aprovechemos algo que particione al país en grandes divisiones territoriales, y tomemos una muestra probabilística de éstas; luego volvamos a particionar las divisiones seleccionadas en la primer etapa para crear subdivisiones de las que a su vez tomemos una muestra probabilística, y así sucesivamente hasta llegar a una partición fina de los espacios que contienen las viviendas particulares que nos interesan. Bien, ya llegamos a una idea, pongámosla en ejecución para ver si es práctica.
3. Que tal si primero seleccionamos municipios o delegaciones, son una división política-administrativa más o menos bien definida en sus límites, y toda parte del país está contenida dentro de algún municipio o delegación del Distrito Federal. Identificados los municipios seleccionados mediante cualquier esquema de muestreo probabilístico, los dividimos a su vez en localidades. Un momento ¿qué tan bien

- definidas en sus límites están las localidades? Hombre, no te preocupes, tomamos la definición que a efectos estadísticos use el mismo INEGI y ya está (conjunto de áreas geoestadísticas básicas, 'AGEB', bien delimitadas en cartografía censal, que a efectos estadísticos constituyen la localidad).
4. Resuelto, primero municipios; luego localidades urbanas y rurales; luego AGEB dentro de las localidades urbanas, y áreas que se delimiten en campo dentro de las localidades rurales; luego manzanas en AGEB urbanas; y finalmente viviendas particulares en manzanas urbanas y dentro de las áreas delimitadas en zonas rurales.
 5. ¿Cómo seleccionar las unidades de cada etapa? Pues aprovechando el que el INEGI produce información con la misma desagregación, entre ésta, el número de viviendas particulares que contiene cada unidad. Se puede entonces seleccionar con probabilidad proporcional al número de viviendas en cada una de las etapas, para en la última seleccionar un número fijo de viviendas tal que nos de el tamaño de muestra deseado.

Si por ejemplo para un estudio a nivel nacional se necesita una muestra de tamaño $n = 2,000$, y la logística de campo aconseja que al mandar a un entrevistador a una zona en muestra es conveniente que al menos levante **10** entrevistas para mantener bajo control los costos de campo, entonces podemos utilizar un esquema de muestreo que:

- a) Seleccione **10** viviendas de cada AGEB urbana o localidad rural en muestra, lo que implica seleccionar un total de **200** unidades de muestreo en la primer etapa, entre AGEB urbanas y localidades rurales.
- b) Seleccione **2** subdivisiones de las AGEB urbanas y localidades rurales que resulten seleccionadas en muestra, para a su vez seleccionar **5** viviendas de cada subdivisión.
- c) Seleccione las unidades muestrales con probabilidad proporcional a una medida de tamaño que se aproxime al número de viviendas que tienen; por ejemplo, las viviendas reportadas en el censo más reciente.

Con estos grandes lineamientos, podemos ya ejecutar un esquema de muestreo que los satisfaga, por ejemplo:

- i) Primero, ordenar en una lista a todas las entidades federativas del país, según su contigüidad geográfica.
- ii) Al interior de cada entidad federativa, ordenar todos los municipios de mayor a menor, según el número de viviendas que reportó el último censo.
- iii) Al interior de cada municipio, ordenar las AGEB urbanas por localidad, de mayor a menor según el número de viviendas, seguidas de todas las localidades rurales.

- iv) Aplicar un muestreo sistemático con probabilidad proporcional al número de viviendas (PPT) de cada AGEB o localidad rural en la lista ordenada.
- v) Seleccionar dos manzanas por AGEB urbana en muestra, con PPT, o dos áreas de las localidades rurales en muestra.
- vi) Seleccionar 5 viviendas con igual probabilidad de cada manzana o área en muestra.

Si V_k es el número de viviendas en la k -ésima AGEB urbana o localidad rural del país, y V_{kj} es el número de viviendas dentro de la j -ésima manzana, o área delimitada en zonas rurales, de la k -ésima AGEB o localidad rural, entonces podemos representar la probabilidad de selección de cualquier vivienda particular 'i' del país, ' p_i ', de la manera siguiente:

$$p_i = \frac{200 \cdot V_k}{\sum_{k=1} V_k} \cdot \frac{2 \cdot V_{kj}}{V_k} \cdot \frac{5}{V_{kj}} = \frac{2,000}{\sum_{k=1} V_k} = \frac{2,000}{V}$$

O sea que la muestra de 2,000 viviendas particulares es una muestra que da 'igual' probabilidad de selección a cualquier vivienda urbana particular del país, lo que en la jerga técnica se conoce como una muestra 'autoponderada'.

Las muestras autoponderadas facilitan considerablemente el proceso de estimación, pues todas las observaciones tomadas 'valen lo mismo', y las estimaciones de porcentajes o promedios son simples sumas de los datos pertinentes entre el número de informantes correspondientes.

Estimadores

Para la estimación de totales, todos recuerdan el estimador clásico cuando se trata de una muestra aleatoria simple de tamaño n de una población de tamaño N , que me permito describir varias veces:

$$\begin{aligned} \hat{Y} &= N * \bar{y} = N \frac{\sum_{i=1}^n y_i}{n} = \frac{1}{n} \sum_{i=1}^n y_i N \\ (3) \quad &= \frac{1}{n} \sum_{i=1}^n \frac{y_i}{\frac{1}{N}} \\ &= \frac{1}{n} \sum_{i=1}^n \frac{y_i}{p_i} \end{aligned}$$

Lo que este estimador hace es tomar una a una las observaciones ' y_i ' de la muestra y las multiplica por el 'factor de expansión', N , que equivale a dividir las por 1 entre el tamaño N de la población, que no es otra cosa que la probabilidad de selección del elemento i . Piensen por ejemplo en el ingreso mensual que tiene el informante ' i ', o la cantidad de leche que consumió el hogar ' i ' en el último mes; este valor se está multiplicando por el factor de expansión que no es otra cosa sino el inverso de la probabilidad de selección p_i . Pero como en la muestra se tiene ' n ' observaciones, se suman las n estimaciones del mismo total, y luego se promedian para obtener un promedio de n estimaciones de lo mismo.

Lo anterior se generaliza muy fácil a esquemas de muestreo más complicados: para estimar un total lo único que se necesita hacer es dividir cada observación entre su probabilidad de selección, y luego promediar el total de observaciones:

$$\hat{Y}_{ppt} = \frac{1}{n} \sum_{i=1}^n \frac{y_i}{p_i}$$

(4)

$$p_i = \frac{V_i}{V}, \text{ medida de tamaño}$$

La estimación de la varianza muestral de esta clase de estimadores está dada por:

$$(5) \quad v(\hat{Y}_{\text{Appt}}) = \frac{1}{n} \sum p_i \left(\frac{y_i}{p_i} - \hat{Y}_{\text{Appt}} \right)^2$$

Y la estimación de cualquier razón (proporción, porcentaje, media, tasa) es el simple cociente de la estimación del total de las variables correspondientes:

$$(6) \quad \hat{R}_{\text{ppt}} = \frac{\sum_{i=1}^n \frac{y_i}{p_i}}{\sum_{i=1}^n \frac{x_i}{p_i}} = \frac{\hat{Y}}{\hat{X}}$$

p_i = probabilidad de selección del elemento i

Al alejarse del muestreo aleatorio simple, se logran reducir considerablemente los costos de campo, pero el diseño utilizado puede tener una varianza mayor, igual o menor que la de un muestreo aleatorio simple del mismo tamaño.

El simple cociente de la varianza del parámetro estimado resultante para un muestreo distinto al aleatorio simple, V_{otro} , de tamaño " n ", entre la varianza del mismo parámetro resultante de un muestreo aleatorio simple, V_{mas} , del mismo tamaño " n ", es lo que en la jerga del muestreo se conoce como el 'efecto de diseño' o DEFF:

$$\text{DEFF} = V_{\text{otro}} / V_{\text{mas}}$$

Si DEFF es mayor a 1, el diseño usado es menos eficiente en términos de precisión que un muestreo aleatorio simple; si es igual a 1, son igual de eficientes; y si es menor a 1, además de ser más económico resulta más eficiente que el muestreo aleatorio simple.

En el **Cuadro 2** del **Anexo** se ilustran las consecuencias en el tamaño de muestra para algunos valores supuestos de DEFF.

PERIODICIDAD DE LA MEDICION Y GENESIS DEL CUESTIONARIO

¿La encuesta debe ser puntual, de una sola medición, o se requieren levantamientos periódicos? Los diferentes parámetros de interés en cualquier encuesta tienen dinamismo distinto. Hay variables que cambian significativamente con más rapidez que otras. Pero, ¿justifica esto el que las estemos midiendo con una periodicidad afín a su dinamismo? ¿O es más prudente el medirlas con una periodicidad acorde a **la capacidad del usuario** para reaccionar y tomar medidas que incidan sobre el fenómeno investigado? ¿De qué sirve medir de nuevo el parámetro si aún no se actúa con respecto a la medición anterior? Sólo para observar su evolución? Ciertamente, se pueden construir series de datos que permitirían modelar y analizar mejor la estadística, pero esto poco se hace, y ¿a qué costo-beneficio?

Vale la pena idear estrategias en donde la periodicidad de la encuesta se acorta conforme entendemos la medición y aprendemos a reaccionar ante ella. Esto permitirá un aprovechamiento más intensivo de cada encuesta, al poder dedicar los recursos especializados al análisis profundo, en lugar de continuamente estar preocupados sólo por la medición siguiente. Dará también tiempo a quienes tienen que tomar acción a tomar ésta con mayor comprensión del fenómeno, y dar tiempo a las acciones tomadas de poder incidir en la variable de interés antes de volver a medirla. Si a futuro se llega a un escenario donde la capacidad de acción es pareja al dinamismo de la variable, enhorabuena, midamos con más frecuencia, ahora sí se justifica.

Con la respuesta a estas y otras preguntas, el productor debe entonces traducir las necesidades del usuario a un **plan de análisis**, usualmente los tabulados que se desea obtener de la encuesta, y acordar con el usuario la relevancia del plan. **Después** de concertado el plan de análisis, el productor debe desarrollar un **cuestionario** sencillo, ágil, de corta duración en tiempo de entrevista, y con preguntas que no hieran ni la sensibilidad del informante ni su inteligencia. Las preguntas a su vez deben concentrarse en el tema y ser suficientes para obtener las estimaciones de los parámetros de interés contenidos en el plan de análisis.

Quien diseña el cuestionario tiene que encontrar el punto de equilibrio entre la tentación de elaborar preguntas con opciones de respuesta precodificadas, que si bien ahorran recursos pueden estar sesgadas dirigiendo al entrevistado a resultados preconcebidos, y la costosa opción de respuestas ‘abiertas’ que además conllevan el riesgo de errores de codificación. Tarea difícil pues entre más experto se es de un tema más fuerte la tentación a ‘dirigir’ los resultados desde la elaboración misma de las opciones de respuesta. Es necesario guardar un balance adecuado entre la medición objetiva y la eficiencia de un cuestionario precodificado.

Desafortunadamente, la práctica usual en el diseño de encuestas, define primero un cuestionario producto de la recopilación de ‘preguntas interesantes’ sobre el tema, sin mucha preocupación por su extensión, y luego, cuando ya es demasiado tarde, se preocupa del análisis que puede hacer con los datos que ya recogió.

El **rechazo** que sufren muchas de las encuestas se debe a la justa irritación del informante ante cuestionarios demasiado largos, con preguntas complejas, de respuesta dirigida, sin tacto para tratar temas sensibles, o simplemente irrelevantes.

Los problemas causados por el rechazo, aunados a las ‘sorpresas’ que recibe el usuario ante tabulados que no satisfacen sus necesidades de información, reflejan la poca previsión del productor. Aquí si, lo azaroso de la encuesta y sus resultados ya no son imputables al ‘azar técnico’, deseable y diseñado exprefeso, son producto de malos diseños y falta de previsión.

TRABAJO DE CAMPO

Para tener una buena encuesta basada en una muestra probabilística, no basta que el diseño de muestreo asegure a todos los elementos una probabilidad de selección distinta de cero. Seleccionada la muestra con un esquema así, es necesario que el personal de campo respete el diseño. Para ello debe identificar correctamente la muestra, insistir en lograr las entrevistas difíciles, y no sustituir elementos seleccionados por otros ‘parecidos’ que estaban más a la mano.

Con demasiada frecuencia, en lugar de lo anterior se utilizan métodos de muestreo que resultan mucho más económicos pero conllevan riesgos altos de caer en errores y sesgos de importancia. Los tan usados ‘muestreos de cuotas’, donde la cuota se cumple en forma ‘azarosa’ (en no pocas ocasiones mas bien desordenada en lugar de con un esquema aleatorio científico), sin asignar a los elementos de cada categoría una probabilidad de resultar seleccionados, si bien con frecuencia producen datos útiles, también son origen de múltiples anécdotas de resultados equívocos. Hay sin embargo usuarios que abiertamente aceptan el riesgo ante la economía de este tipo de encuesta. Lo grave es que hay otros que reciben información generada con estos procedimientos y que no lo saben, por lo que ignoran que los resultados que les están proporcionando tienen esos riesgos.

Igual de importante es el respeto a las instrucciones del cuestionario y/o manual de entrevista. El flujo correcto del cuestionario, la lectura textual de las preguntas, la observancia de ‘saltos’ según respuestas previas, la utilización de tarjetas con respuestas posibles, etc, deben ser observados al pie de la letra por cada encuestador.

La supervisión de lo anterior, y en general el control de la calidad del trabajo de campo, es crucial en cualquier encuesta seria. Los encuestadores se deben saber y sentir supervisados, la supervisión debe tener consecuencias explícitas, y éstas deben ser ampliamente difundidas. Las enseñanzas de la encuesta de hoy deben ponerse en práctica en la encuesta de mañana, y no perderse en la apatía y olvido.

Toda encuesta tiene casos de no-respuesta, el ignorarlos, es asumir que los que no respondieron se comportan como los que sí contestaron, y todos felices, lo único que sufrió fue la precisión (al tener menos casos) pero no hay sesgo en las inferencias. ¿De veras? ¿Y que tal si no?

La no-respuesta producto de deficiencias en el marco de muestreo no preocupa (viviendas deshabitadas, de uso temporal, sin residentes miembros de la población de interés, etc), la que es producto de deficiencias en el trabajo de campo, sí. Es necesario identificar la no-respuesta por causa, y examinar si se concentra o no en ciertos sectores de la población de interés, o en ciertos entrevistadores. Si lo se da cualquier tipo de concentración, es una llamada de atención que nos debe llevar a acciones explícitas y oportunas, desde repetir el trabajo de campo hasta utilizar factores de corrección en la etapa de procesamiento para compensar los casos de no-respuesta.

COMPARABILIDAD

La comparabilidad entre datos de encuestas continuas, entre encuestas continuas y encuestas ocasionales, entre datos de encuestas y otras fuentes de datos, y entre datos de encuestas de distinta naturaleza es un tema poco conocido aún por algunos usuarios asiduos de las mismas.

Encuestas continuas. Cuando una encuesta se repite en el tiempo para seguir la evolución de un tema sustantivo, es recomendable mantener el **mismo** cuestionario, el **mismo** diseño de muestra, la **misma** estrategia del trabajo de campo, el **mismo** proceso de estimación. El alterar cualquiera de estos elementos puede inducir diferencias en los resultados que se **confundirán** con cambios naturales en el nivel de las variables medidas. Esto complica considerablemente el análisis y empaña las conclusiones.

Por esta razón, algunos organismos responsables de las estadísticas oficiales en distintos países, mantienen los mismos conceptos, cuestionario, diseño y logística de campo de una encuesta durante largos periodos, a pesar de conocer y tener manera de mejorar alguno o varios aspectos de la metodología.

Cuando el cambio viene, es porque el beneficio de incorporar las mejoras excede el costo de 'romper' la serie. En estos casos, hay distintas estrategias para incorporar los cambios: se pueden

introducir todos de golpe; se pueden programar para introducirlos en forma paulatina en levantamientos sucesivos (primero el cuestionario, luego el diseño de muestra, etc); cuando el interés y el presupuesto lo permiten, se puede tener un periodo de levantamientos paralelos con la metodología 'vieja' y la 'nueva', y computar los resultados como un promedio ponderado que va desfasando la primera y reforzando la segunda hasta que solo queda ésta; etc.

Es difícil e innecesario para un usuario el estar interiorizado en todos estos aspectos de detalle. Sin embargo, cuando detecte cambios **sustantivos y sorpresivos** al comparar los resultados de distintas rondas de la misma encuesta, es recomendable que agregue a su análisis una indagación sobre posibles cambios realizados a la metodología de la encuesta. El que ésta conserve el mismo nombre y publique los mismos cuadros no es suficiente para asegurar que no ha sido alterada. Se podría estar nombrando de igual manera y comparando erróneamente, valores de conceptos que son distintos.

Encuestas ocasionales. Muchos temas donde las variables de interés no cambian con frecuencia o que solo requieren un estudio a profundidad en forma ocasional, se investigan mediante encuestas ad-hoc. No es raro el que estas encuestas capten datos sobre variables investigadas a través de sistemas continuos de encuesta especializados en medir éstas variables, y que en las ocasionales solo sean de interés colateral. La comparación de resultados para estas variables entre éstas encuestas y las continuas es inevitable y origen de controversias innecesarias. Con frecuencia se pretende invalidar los resultados del sistema especializado de encuesta en un tema, con resultados divergentes provenientes de estas encuestas ad-hoc.

El usuario precavido debe indagar al menos la definición operativa en ambas encuestas de conceptos que parecen idénticos, la población objetivo de ambas encuestas, sus diseños de muestra, y la estrategia para recoger los datos. Con esta información tendrá los elementos mínimos para el análisis comparativo de resultados. Si las estimaciones son comparables y los resultados divergentes en magnitud significativa, contará con más elementos para seleccionar la mejor estimación. (Que por cierto puede no ser ninguna de las dos en controversia, y basarse en un estimador que aproveche los datos de ambas encuestas.)

Encuestas y otras fuentes de datos. Es común y recomendable el comparar los resultados de las encuestas con los datos de registros administrativos, censos y estimaciones basadas en modelos. La precaución corriente es cuidarse de los periodos de referencia de las distintas fuentes. No lo es tanto el cuidar que se está hablando de los mismos conceptos y definiciones, para las mismas poblaciones de interés. Aquí el enemigo es la familiaridad con que usamos cotidianamente distintos términos económicos, sociales, demográficos o técnicos, que en su **definición operativa** pueden encubrir diferencias importantes entre lo que mide cada fuente.

En los casos raros en que hay coincidencia entre distintas fuentes, ¡enhorabuena! Pero lo más común es que exista divergencia; y entonces requerimos más elementos para formarnos un juicio sólido que la explique. Antes de calificar de acertada una fuente y errónea otra, es recomendable que el analista se cerciore que no está comparando ‘peras con manzanas’. Con demasiada facilidad se cae en este tipo de error.

Encuestas de distinta naturaleza. No todas las encuestas son de la misma índole. Las hay que registran **datos medibles en escalas comunes** sobre eventos pasados o presentes (por ejemplo: edad; ingreso; empleo; **si votó, cómo votó** en la elección pasada: etc.). Las hay que registran **actitudes asumidas** ante eventos conocidos pasados o presentes (por ejemplo: qué **hizo** cuando.....). Y finalmente las que registran **opiniones** ante escenarios reales o hipotéticos del pasado, presente o futuro (por ejemplo: qué **pensó/piensa/pensaría** si).

Los primeros dos tipos de encuestas, cuando están bien diseñadas y el informante responde con veracidad, pueden medir en forma objetiva los eventos de interés. Las estimaciones que se obtienen se pueden controlar en su precisión y confianza, y nos dan información sobre **hechos acontecidos** que ya no cambian con el tiempo.

La última clase es más difícil. Cuando se entra al terreno de la opinión, lo más que puede pretender una encuesta es medir el estado anímico provocado por la pregunta **en el momento y forma particular en que se le hizo al respondente**. Extrapolar de ahí comportamientos futuros es, por decirlo en términos simplistas, riesgoso. De ahí la necesidad de aplicar continuamente este tipo de encuestas y tratar de identificar patrones consistentes.

EPILOGO

Con la semblanza general del tema de encuestas que hemos intentado en ésta nota, queda una referencia global que invita al lector interesado a tomar cuidado en la discusión de los alcances y limitaciones de las múltiples encuestas particulares que en nuestro país realizan el sector privado, público, social y académico, así como organismos internacionales multilaterales y bilaterales. Si la ocasión lo permite, a futuro intentaremos compartir otras facetas metodológicas de algunas de ellas.

Muchas gracias.

ANEXO**Tamaño y Diseño de Muestra****A. Tamaño de muestra.**

1. En general, los parámetros a estimar en una investigación de mercado son distintos porcentajes y proporciones de variables referidas a una muestra de personas dentro de cierto rango de edad (por ejemplo, entre **18** y **54** años de edad).
2. Para determinar un tamaño de muestra inicial partimos del supuesto de un **muestreo aleatorio simple**, y de las exigencias de **precisión** y **confianza** requeridas para la investigación, expresados en términos generales por:

$$\text{Pr}[| p - P | \leq d] \geq 1 - \alpha$$

Donde :

P = valor del porcentaje o proporción **real** de interés,

p = porcentaje o proporción que **estima** la encuesta,

d = máxima diferencia aceptable entre el valor real, **P**, y su estimación, **p**, proveniente de la encuesta;

1 - α = confianza requerida en la precisión, o sea la probabilidad de que se cumpla la especificación expresada con **d**.

3. La fórmula (1) y las exigencias de precisión y confianza conducen a la siguiente expresión para la estimación **inicial** del tamaño de muestra :

$$n_0 \geq k^2 * \frac{P * (1 - P)}{d^2}$$



Berumen y Asociados, S.A. de C.V.
Altadena No. 15 Col. Nápoles,
Del. Benito Juárez Méx. D.F. CP 03810
Tel: 5093-8600 Fax: 5093-8615

Donde k se determina en función de la confianza requerida (el valor especificado para $1 - \alpha$, usualmente **90%** o **95%**).

4. El tamaño de muestra requerido para distintos valores de P , d , y k , se ilustran en el cuadro siguiente:

**TAMAÑO DE MUESTRA REQUERIDO PARA UN
MUESTREO ALEATORIO SIMPLE SIN REEMPLAZO**

PRECISION d	CONFIANZA 95% (k = 1.96)				
	P=.05	P=.10	P=.25	P=.40	P=.50
.01	1,825	3,457	7,203	9,220	9,604
.02	456	864	1,801	2,305	2,401
.03	203	384	800	1,024	1,067
.04	114	216	450	576	600
.05	73	138	288	369	384

PRECISION d	CONFIANZA 90% (k = 1.65)				
	P=.05	P=.10	P=.25	P=.40	P=.50
.01	1,293	2,450	5,105	6,534	6,806
.02	323	613	1,276	1,634	1,702
.03	144	272	567	726	756
.04	81	153	319	408	425
.05	52	98	204	261	272

PRECISION d	CONFIANZA 80% (k = 1.282)				
	P=.05	P=.10	P=.25	P=.40	P=.50
.01	781	1,479	3,082	3,944	4,109
.02	195	370	770	986	1,027
.03	87	164	342	438	457
.04	49	92	193	247	257
.05	31	59	123	158	164

5. El tamaño de muestra inicial del **Cuadro 1** se determinó bajo el supuesto de un muestreo aleatorio simple sin reemplazo. De hecho, este esquema de muestreo **nunca** se usa en la práctica en investigaciones de poblaciones como la de interés, pues resulta ineficiente y costoso.
6. En general, se proponen esquemas de muestreo estratificado con tres o más etapas de selección: por ejemplo, AGEB o secciones electorales en la primer etapa, manzanas en la segunda, y viviendas y ciudadanos en la tercera. Esto implica la necesidad de modificar el tamaño de muestra para reflejar la utilización de un diseño distinto al aleatorio simple sin reemplazo, o sea "corregir" los tamaños anteriores por el llamado "efecto de diseño" o **DEFF**.
7. El coeficiente **DEFF** es el simple cociente de la varianza del parámetro estimado resultante para un muestreo distinto al aleatorio simple, V_{otro} , de tamaño "**n**", entre la varianza del mismo parámetro resultante de un muestreo aleatorio simple, V_{mas} , del mismo tamaño "**n**", o sea:

$$\mathbf{DEFF} = V_{\text{otro}} / V_{\text{mas}}$$

8. Usualmente **DEFF** es mayor a **1**, por lo que el utilizar las fórmulas del muestreo aleatorio simple para estimar la precisión de los parámetros estimados **subestiman** la varianza de estos, y por tanto **sobre-estiman** la precisión de la estimación.
9. De acuerdo a la experiencias de **Berumen** para variables similares a las de interés, en donde se ha utilizado un diseño de muestra similar al del ejemplo, estratificado con tres etapas de selección, se han obtenido coeficientes de diseño **DEFF** que varían entre **1.5** y **2.8.**, para las distintos parámetros estimados. Habría entonces que multiplicar los tamaños de muestra dados en el cuadro por el factor **DEFF** correspondiente.
10. Para ilustrar una manera conservadora de corrección, propongamos un valor de **DEFF** igual a **2.25**. Los tamaños de muestra adecuados (observaciones exitosas) al diseño de muestra propuesto, estratificado con dos etapas de selección, son entonces:

CUADRO 2. TAMAÑO DE MUESTRA REQUERIDO PARA UNA

MUESTRA ESTRATIFICADA CON TRES ETAPAS DE SELECCION

PRECISION d	CONFIANZA 95% (k = 1.96)				
	P=.05	P=.10	P=.25	P=.40	P=.50
.01	4,106	7,778	16,207	20,745	21,609
.02	1,026	1,944	4,052	5,186	5,402
.03	457	864	1,800	2,304	2,401
.04	257	486	1,013	1,296	1,350
.05	164	311	648	830	864

d	CONFIANZA 90% (k = 1.65)				
	P=.05	P=.10	P=.25	P=.40	P=.50
.01	2,909	5,513	11,486	14,702	15,314
.02	727	1,379	2,871	3,677	3,830
.03	324	612	1,276	1,634	1,701
.04	182	344	718	918	956
.05	117	221	459	587	612

11. Si por ejemplo se seleccionan **2,500** personas entre **18-54** años de edad, de acuerdo al **Cuadro 2**, esto es suficiente para lograr estimaciones de porcentajes de cualquier orden de magnitud, con una precisión del **3%** y una confianza del **95%**. Para porcentajes de **10%** o menores mejora la precisión al **2%** o aún más, al mismo nivel del **95%** de confianza.