

# Determinación de Tamaño de Muestra

## ANEXO: PROCEDIMIENTO DE MUESTREO Y DETERMINACIÓN DEL TAMAÑO DE LA MUESTRA

Extraído de Baca Urbina (2010)

### PROCEDIMIENTO DE MUESTREO Y DETERMINACIÓN DEL TAMAÑO DE LA MUESTRA

La teoría del muestreo es compleja y definitivamente este texto no pretende hacer un análisis exhaustivo de ella, por varias razones: desde un principio se mencionó que se deseaba elaborar una guía sencilla para la evaluación de proyectos, pero enfocada a la pequeña y mediana industria de capitales privados. Esto elimina en forma automática todos los proyectos del gobierno, ya sean regionales, rurales o estatales, y también elimina todos los grandes proyectos privados. Por tanto, solo se presenta lo que probablemente necesite conocer el evaluador de un proyecto al investigar los tipos de mercado mencionados.

Respecto del muestreo, selección de una pequeña parte estadísticamente determinada, para inferir el valor de una o varias características del conjunto, conviene señalar que existen dos tipos generales de muestreo: el probabilístico y el no probabilístico. En el primero, cada uno de los elementos de la muestra tiene la misma probabilidad de ser entrevistado, y en el muestreo no probabilístico, la probabilidad no es igual para todos los elementos del espacio muestral.

# Determinación de Tamaño de Muestra

Aunque pareciera que el muestreo probabilístico es el más usado en las investigaciones de mercado, esto no es así, vea por qué. Un estudio de mercado siempre está enfocado a investigar ciertas características de, por ejemplo, empresas, productos y usuarios; es decir, antes de iniciar la investigación siempre se hace una *estratificación*. Aunque se investiguen características que pueda tener toda la población, tales como usar calzado, fumar, hábitos de vestir, y otros, siempre se estratifica antes de encuestar. No se debe confundir, por ejemplo, con investigar el porcentaje de gente que fuma, porque esto no sería una investigación de mercado. Una investigación acerca de los fumadores tendría como primera pregunta si la persona fuma y seguiría una serie de cuestionamientos sobre sus gustos, preferencia de marcas, estrato social, etc. La estratificación implícita está en aplicar el cuestionario a quienes fuman, pues quien no fuma difícilmente opinaría con propiedad acerca de gustos o marcas preferidas. A cualquier persona se le puede preguntar si fuma, pero no a cualquiera se le aplicará el cuestionario, que es la verdadera investigación de mercado.

Si se examinan más casos de investigación de mercado con base en encuestas, se encontrará siempre una estratificación preliminar implícita, y esto es un muestreo no probabilístico. Por tanto, el probabilístico queda fuera de aplicación en la evaluación de proyectos. Esta teoría es muy interesante y de gran aplicación en control de calidad, donde el universo de la muestra es finito y conocido.

Para calcular el tamaño de la muestra se deben tomar en cuenta algunas de sus propiedades y el error máximo que se permitirá en los resultados. Para el cálculo de  $n$  (tamaño de la muestra) se puede emplear la siguiente fórmula:

$$n = \frac{\sigma^2 Z^2}{E^2}$$

Donde  $\sigma$  (sigma) es la desviación estándar, que puede calcularse por criterio, por referencia a otros estudios o mediante una prueba piloto. El nivel de confianza deseado se denota con  $Z$ , el cual se acepta que sea de 95% en la mayoría de las investigaciones. El valor de  $Z$  es entonces llamado número de errores estándar asociados con el nivel de confianza. Su valor se obtiene de la tabla de probabilidades de una distribución normal.

# Determinación de Tamaño de Muestra

Para un nivel de confianza de 95%,  $Z = 1.96$ , lo que significa que con una probabilidad total de 0.05 la media de la población caería fuera del intervalo a  $2\sigma$  ( $\sigma$  es la desviación estándar de la muestra). Finalmente,  $E$  es el error máximo permitido y se interpreta como la mayor diferencia permitida entre la media de la muestra y la media de la población ( $X \pm E$ ).

Suponga que se desea investigar la demanda actual del champú a partir de encuestas, dado que no existen estadísticas oficiales ni particulares para este producto. Así, se entiende que este procedimiento de cuantificación es conveniente para estudios de productos nuevos. El primer paso para la cuantificación es la determinación del *tamaño de la muestra* que se efectúa como sigue.

Se encuesta a una muestra piloto de al menos 30 personas; de acuerdo al teorema del límite central, la distribución de alguna de las propiedades de este tamaño de muestra tiende a una distribución normal, es decir, no está sesgada, que es lo que se busca en una investigación de mercado. En el ejemplo se intenta cuantificar el consumo promedio de champú por persona al año, la desviación estándar del consumo y los resultados que se obtengan no deben estar estadísticamente sesgados.

El objetivo de cuantificar la demanda de un producto mediante encuestas, obliga a que dentro de la misma se incluyan preguntas sobre el consumo del producto a través del tiempo. Suponga que la encuesta piloto arroja como resultado que una persona adulta consume un promedio de 5 litros de champú por año y que la desviación estándar del consumo es de 2.1 litros, esto supone que un hombre con poco pelo consumiría tan sólo  $5 - 2.1 = 2.9$  litros de champú al año, y que una mujer con una cabellera abundante consumiría hasta 7.1 litros del producto al año. El nivel de confianza generalmente aceptado es de 95%, que significa que en la cuantificación de la demanda se tiene 95% de certeza de que la media de consumo obtenida en la muestra (5 litros/año/ persona) estará dentro de tres desviaciones estándar de la media poblacional, la cual siempre es desconocida, esto es, nadie sabe con absoluta certeza cuál es el consumo promedio de champú por persona al año, pero la fórmula otorga al investigador 95% de certeza de que eso suceda. Por otro lado, la fórmula contiene el error permisible, que es el que el investigador de mercado quiere respecto a que la media poblacional esté alejada de la media muestral. En el ejemplo, el error se fijó en 40%.

# Determinación de Tamaño de Muestra

Hay que observar dos aspectos de la fórmula. Si la desviación estándar obtenida en el muestreo piloto es baja, el número de encuestas también será bajo. En términos prácticos esto significa que la gente consume este producto en límites muy estrechos, por lo cual, no harán falta tantas encuestas para tener 95% de certeza de que la media poblacional estará dentro de dos desviaciones estándar de la media muestral. Lo opuesto, es decir, que se obtuviera una desviación estándar muy elevada por medio de encuesta piloto, significa que la gente consume determinado producto de manera muy irregular, por lo que es necesario realizar gran cantidad de encuestas para tener más certeza en los resultados.

Por otro lado, si el error que se asigna es muy bajo, se elevará considerablemente el número de encuestas, es decir, si el investigador casi no se quiere equivocar al cuantificar la demanda, tendrá que hacer mucho más encuestas. Por tanto:

$$n = \frac{\sigma^2 Z^2}{E^2} = \frac{2.1^2 * 1.96^2}{0.4^2} = 106$$

Si el investigador quisiera un error en la cuantificación de la demanda de tan sólo 5%, la cantidad de encuestas se elevaría hasta 6,776 en el ejemplo.

## Referencia:

Baca Urbina, G. (2010). Evaluación de proyectos. Sexta Edición *McGrawHill*.