

# Sobre la Utilización de «Benchmarks» en Investigación de Mercados. Una Nota Metodológica



**Javier Alagón**  
*Estadística Aplicada*

**E**n la interpretación de estudios cuantitativos, la utilización de los famosos «*benchmarks*», o números de referencia, es frecuentemente uno de los valores agregados más buscados por los clientes usuarios dentro de nuestra industria. El poder afirmar categóricamente que en una prueba de concepto, por ejemplo, una intención de compra del 37% es «buena» o «mala», es «alta» o «baja», requiere, entre otras cosas, de **conocimiento profundo** del mercado y de una **extraordinaria tarea de síntesis**, basada en la lectura documentada y analítica de múltiples bases de datos.

Sin lo anterior, la calificación de distintas estadísticas de los estudios de mercado se convierte en un acto de audacia extrema o de arrogancia infinita, por muy comprensible que pueda ser la presión ejercida por el cliente o usuario. Ante la gran variabilidad (varianza, en términos estadísticos) de las respuestas que diferentes grupos sociodemográficos pueden tener ante una misma pregunta, las siguientes preguntas resultan sumamente pertinentes:

- ¿Es posible realmente establecer *benchmarks* en nuestra industria?
- ¿Qué tan válidos son los *benchmarks* obtenidos con distintos grupos sociodemográficos?
- ¿Es posible utilizar *benchmarks* obtenidos en otros países? ¿En otras categorías?

El **objetivo** de esta nota metodológica es tratar de dar respuesta a las preguntas anteriores con base en múltiples experiencias dentro del área; es decir, una respuesta fundamentada en argumentos empíricos, estadísticos y lógicos.

Para ilustrar la naturaleza de la variabilidad de las respuestas de los distintos grupos sociodemográficos, presentamos en la tabla 1 datos reales, estimaciones obtenidas en una evaluación de concepto promocional (medición previa) y una evaluación de desempeño promocional (medición posterior) para una marca muy conocida y cercana al grupo objeti-

**Tabla 1**

**Resultados Previos y Posteriores**

	Intención de Participación (Previa)	Nivel de Gusto (Previo)	Tasa de Participación (Posterior)
México	45%	72%	14%
Guadalajara	49%	75%	9%
Monterrey	68%	87%	15%
Niños de 8 a 12	64%	86%	13%
Jóvenes de 13 a 17 años	57%	80%	16%
Adultos jóvenes de 18 a 24 años	46%	71%	12%
Adultos de 24 a 35 años	49%	73%	11%
Nivel Alto	48%	76%	13%
Nivel Medio	54%	81%	15%
Nivel Bajo	52%	78%	12%

vo, con distintos grupos sociodemográficos. En la tabla, se presentan los datos de dos mediciones transversales con el mismo grupo objetivo en dos puntos en el tiempo: antes del lanzamiento de una promoción y después de terminada la promoción.

La tabla 1 muestra la forma en que responden los entrevistados ANTES de que la promoción hubiera sido lanzada (intención de participación, equivalente a intención de compra en pruebas de concepto) y lo que dicen haber hecho DESPUÉS de que terminó la promoción (en este caso, se tomó canje real del producto promocionado como variable de participación en la promoción). El propósito de haber incluido datos de ANTES y DESPUÉS es claro: en la contrastación de estadísticas se pueden hacer inferencias sobre las formas (¡y sesgos!) en las que responden los distintos grupos sociodemográficos.

Sin pretender que la información presentada en la tabla 1 represente la forma en que responden los distintos grupos sociodemográficos en todos los estudios de mercado (lo que sería equivalente a hacer inferencias estadísticas con una muestra de tamaño 1), sí podemos afirmar, con base en múltiples experiencias empíricas, que los comportamientos exhibidos, responden a patrones y a formas de respuesta que pudieran generalizarse de la siguiente manera:

- Los grupos de menor edad tienden a exagerar sus comportamientos y suelen ser más benévolo en el uso de escalas (es decir, usar los niveles superiores de la escala).
- Los entrevistados de ciertas ciudades (como es el caso de Monterrey en el ejemplo) tienden a exagerar sus comportamientos y suelen ser también más benévolo en el uso de escalas. (Este comportamiento suele ocurrir en ciudades como Monterrey y Mérida, así como en zonas rurales).
- Los entrevistados de la ciudad de México y los de nivel alto suelen ser menos benévolo en el uso de escalas.

La implicación inmediata de los puntos anteriores es muy clara: debemos tener mucho cuidado al mez-

clar la información de distintos grupos sociodemográficos y, para el caso específico que nos ocupa en este artículo, la utilización de los famosos *benchmarks* como referentes universales para hacer comparaciones, puede ser una actividad sumamente arriesgada si no se toma en cuenta qué grupos sociodemográficos se están analizando.

Si, por ejemplo, en la evaluación previa se hubiese tenido un *benchmark* de 55% (basado en experiencias anteriores) para intención de compra, se habría concluido erróneamente que la tasa de participación esperada fracasaría en México y en Guadalajara, o que solamente los niños y adolescentes estarían dispuestos a participar en la promoción. Como puede verse en el ejemplo, a lo largo de los distintos grupos sociodemográficos, la tasa de participación declarada una vez terminada la promoción, no estuvo directamente relacionada con la intención de participación declarada en la medición previa. Pero esto es debido a la forma en que contestan o usan las escalas los distintos grupos sociodemográficos, no a que no exista dicha relación.

El comportamiento exhibido por los datos de la tabla 1 es muy común. Por tanto, resulta sumamente arriesgado el tener *benchmarks* universales para establecer que una estadística específica es alta o baja, es mala o es buena. Y la variabilidad de las respuestas no tan sólo depende de cómo responden los distintos grupos sociodemográficos. La experiencia empírica nos indica que también puede depender de otras variables muy importantes en la construcción de estrategias de mercado, entre otras:

- Tipo de categoría.
- Nivel de involucramiento con la categoría (*heavy, medium y light*).
- Usuarios de distintas marcas.
- Nivel de emotividad inherente a conceptos presentados.

Resulta por tanto, sumamente cuestionable que, por ejemplo, en evaluaciones publicitarias previas o posteriores, o en pruebas de concepto y/o de pro-

**Tabla 2** **Importancia de Temas en Distintos Países**

	Familia	Trabajo	Amigos	Tiempo Libre	Religión	Política
Nigeria	94	94	53	68	85	39
Brasil	91	82	57	52	57	42
EUA	92	62	54	43	53	52
Argentina	91	76	51	40	40	31
Gran Bretaña	89	51	48	45	16	43
México	85	67	25	28	34	41
España	82	66	45	38	21	22
Japón	78	41	34	26	6	5 4
China	62	64	22	14	1	5 9
Portugal	6 5	3 5	2 0	16	1 7	22

ducto o en estudios de imagen y posicionamiento, por mencionar los más comunes, se utilicen *benchmarks* universales como si fueran la piedra filosofal de la vida.

Y aún más falaz puede ser la utilización de *benchmarks* obtenidos de estudios en otros países, puesto que ha sido ampliamente ilustrado, que los entrevistados de distintas nacionalidades, responden cuestionarios de maneras muy distintas, como se ilustra de manera elocuente en la tabla 2.

En la tabla 2, se muestra un extracto de los resultados de la Encuesta Mundial de Valores, recabada por la Universidad de Michigan, en donde se exhiben datos de 10 países (de un total de 42 incluidos en el estudio), correspondientes al nivel de importancia conferido, en una escala de 5 puntos, a 6 temas en la vida: familia, trabajo, amigos, tiempo libre, religión y política.

Sin pretender hacer un análisis exhaustivo de los datos anteriores, resulta claro que las estimaciones numéricas dependen de la forma en que responden los entrevistados de distintos países. El que para los estadounidenses resulte más importante la familia, los amigos o la religión, que para los mexicanos, pone de manifiesto el tipo de uso que se hace de las escalas en los dos países y no necesariamente las diferencias de importancias.

Hemos constatado en una multitud de estudios realizados en la agencia en diferentes países Latinoamericanos, que comparados con entrevistados de otros países, los mexicanos solemos encontrarnos a la mitad en el uso de escalas, muy diferente a lo que ocurre con los entrevistados brasileños, por ejemplo, en donde suele hacerse un uso muy generoso de las partes superiores de las escalas. Por tanto, es muy cuestionable el comentario frecuente que escuchamos respecto a la existencia de sesgos de respuesta muy fuertes en estudios en México. Es decir, que el multicitado «sesgo de cortesía» ocurre más frecuentemente en otros países que en el nuestro.

El mensaje final de esta nota metodológica es muy claro: si deseamos usar correctamente *benchmarks* en estudios de mercado, éstos deberán ser directamente comparables con el grupo objetivo que se está analizando en el estudio. La utilización ciega, como un acto de fe, de *benchmarks* externos (de otros grupos sociodemográficos, de otras categorías, de otros países) es equivalente a jugar la ruleta rusa.