

Mesa redonda

Ética y Estadística: el caso del INDEC

14 de Diciembre de 2009 – Universidad de Belgrano

Auditorio Lavalle – Lavalle 485 – Buenos Aires

Herbert George Wells, escritor inglés, notable novelista y filósofo, famoso por sus novelas de ciencia ficción, de la que es considerado, junto a Julio Verne, uno de sus precursores, quien vivió entre 1866 y 1946, dijo: *“El pensamiento Estadístico llegará a ser tan necesario para ser un ciudadano responsable como el saber leer y escribir”*.

Por otra parte, Francis Galton, quien nació en 1822 y murió 1911. Fue explorador y científico británico con un amplio espectro de intereses: sicología, biología, tecnología, geografía, Estadística, meteorología. Primo de Charles Darwin. Realizó la mayoría de las investigaciones por su cuenta y recién a la edad de 87 años recibió un reconocimiento formal al concedérsele el título de Sir. Sus contribuciones más importantes en Estadística son: uso de la regresión lineal; pionero en la utilización de la distribución normal; inventó la máquina Quincunx, un instrumento para demostrar la ley del error y la distribución normal; descubrió las propiedades de la distribución normal bivariada y su relación con el análisis de regresión; en 1888 introdujo el concepto de correlación, dijo: *“Tengo que escribir acerca de un gran tema (las Estadísticas), pero siento vivamente mi incapacidad literaria de hacerlo de forma inteligible sin sacrificar su precisión y su veracidad”*.

Estos y muchos otros ejemplos nos muestran que personas de alta formación intelectual consideraron a la Estadística, por un lado, como una parte imprescindible dentro de la vida cívica y democrática de toda sociedad. Lo que lleva a la necesidad de que todos, como miembros de la sociedad civil, tengamos una “cultura Estadística” adecuada para usufructuar convenientemente la información suministrada. Por otro lado, se la considera también un área cuyo contenido técnico debería ser desarrollado por aquellos que hubieran logrado una formación apropiada a fin de que la información que llega a la sociedad civil sea accesible en su interpretación, precisa en sus apreciaciones y veraz en sus números.

El informar mal, utilizando material estadístico, podría llamarse manipulación estadística, y resumiéndolo en una sola palabra (aunque no sea muy buena), **estadisticulación**.

Muchas veces el proceso de mentir con estadísticas indica que esas operaciones son el producto del deseo de alguien de movernos al engaño. Ahora bien, un miembro de la American Statistical Association (ASA) dijo que la mayor parte de las veces ello no se trata de fraudes, sino de incompetencia. Quizás haya algo de verdad en esa afirmación, pero no creo que este supuesto (el de la incompetencia) sea menos ofensivo que el de fraude para los expertos en Estadística. Posiblemente sea de mucha importancia tener en cuenta que la tergiversación de datos estadísticos y su manipulación para conseguir una finalidad determinada, no son siempre obra de los profesionales en Estadísticas. Lo que sale lleno de virtudes de la mesa del experto puede verse cambiado, exagerado, demasiado simplificado y falseado al ser tomado por otras personas.

Sea quien sea el culpable en un caso determinado, se hace difícil achacarlo a su incapacidad inocente. Los gráficos falsos que aparecen en las revistas y diarios resultan sensacionalistas por su exageración, y raras veces minimizan nada. Quienes presentan argumentos estadísticos a favor del crecimiento industrial, de la falta de inflación, etc., pocas veces, según he experimentado, dan a los trabajadores o al público consumidor una información que mejore los hechos en sus sectores; en general los presentan aún peor para ellos.

Mientras los errores estén todos de una parte, no es fácil atribuirlos a la falta de conocimiento o a la casualidad. Por el contrario, es evidente que tienen un propósito que no es el de dar la información verídica a quienes deban recibirla.

El hecho es que, a pesar de su base matemática, las Estadísticas son tanto un arte como una ciencia. Muchas manipulaciones e incluso tergiversaciones son posibles dentro de los límites de su jurisdicción. A menudo, el experto en Estadísticas debe escoger entre distintos métodos y hallar el que debe utilizar para representar los hechos, lo que no deja de ser un proceso subjetivo. Sin embargo, en situaciones específicas, como el caso de las estadísticas oficiales, hay métodos que han sido tomados como los adecuados por una serie de convenciones (de expertos, internacionales, etc.) y que aseguran la uniformidad de los resultados, permitiendo comparaciones intertemporal y espacial. Todo cambio debería realizarse de acuerdo a metodologías técnica y

científicamente aceptadas y nunca siguiendo los caprichos circunstanciales de algunas personas o sectores.

Ahora sería interesante explicar cómo debe mirarse una Estadística falseada, y desenmascararla; y aun más importante, cómo reconocer los datos útiles y ciertos ante la marejada de fraudes a los cuales nos enfrentamos.

No toda la información Estadística que llegue a nuestras manos puede comprobarse con la seguridad del análisis químico o como se hace en un laboratorio para valorar metales. Pero podemos tamizar la información con el cedazo representado por cinco preguntas muy sencillas. Hallando las respuestas, evitaremos dar por verídicas muchas cosas que no lo son.

Pregunta 1: ¿QUIEN LO DICE?

Lo primero que se debe mirar es en que sentido puede estar influida la información: El laboratorio que necesita demostrar algo para beneficio de una teoría, una reputación o unos honorarios; el periódico cuya finalidad es publicar un buen artículo; los trabajadores y los patrones que discuten un aumento de salarios. Debemos buscar la influencia consciente. El método tal vez sea la desfiguración de un hecho en forma directa, o una afirmación ambigua que sirva igualmente para la finalidad que se desee y cuya culpabilidad no puede probarse. Puede tratarse de una selección de los datos favorables con supresión de los desfavorables. Pueden ser engañosas las unidades de medida, como cuando se toma un año para una comparación y se pasa a un año más favorable para otra. Quizás se utilice una medida inadecuada; una media cuando la mediana sea mucho más informativa (quizás demasiada informativa!), cubriendo el truco con la palabra “promedio” sin calificativos.

Debemos buscar con atención la influencia consciente o inconsciente que tiende a cambiar los números y la información. En los gráficos y predicciones de muchos economistas y expertos se suelen filtrar factores de influencia que llevan a producir resultados curiosos. Por ejemplo, se desestiman alegremente los trastornos de la estructura económica y se dan toda clase de pruebas justificadas estadísticamente para demostrar que podemos estar en una corriente de prosperidad.

Tal vez haga falta una segunda inspección para enterarse de quién lo dice.

Pregunta 2: ¿COMO LO SABE?

En una oportunidad se realizó un estudio a empresas para saber si habían aumentado los precios de los productos o no. De 169 empresas que contestaron, 108 dijeron que no había aumentado los precios y 61 que sí habían aumentado. Puesto así, tenemos que aproximadamente dos tercios de las empresas no están aumentando los precios y solamente un tercio lo hacen. Y sucede que en muchos casos eso es lo que se informa. Pero indagando un poco más, resultó que el cuestionario había sido enviado a 1.200 compañías importantes. Solamente contestaron 169, lo que es aproximadamente un 14%. El 86% restante no tenía ningún interés en informar a la opinión pública si subía los precios o no.

Quien informaba veía con muy buenos ojos los resultados obtenidos, pero de hecho había muy poca cosa de qué alardear. En realidad, de 1.200 compañías que estaban en la muestra y que debían tomar parte en la encuesta, el 9% dijo que no habían elevado los precios, el 5% que sí los habían elevado y el 86% no dio contestación alguna. Quienes contestaron constituyen una proporción muy pequeña de la muestra y es de esperar un factor de influencia. En estos casos debemos buscar los indicios indicativos de que se trata de una muestra influenciada: la que se ha escogido de forma inadecuada, o como ésta, que se ha seleccionado a sí misma. Hagámonos la pregunta ¿es suficientemente grande la muestra para llegar a alguna conclusión válida?

Lo mismo podríamos decir cuando se informa una correlación: ¿es suficientemente grande la muestra como para significar algo? ¿Hay bastantes casos que le den significado?

Pregunta 3: ¿QUE FALTA?

No siempre le darán el número de casos. La ausencia de esta cifra, en particular cuando la fuente de información es parte interesada, es suficiente para sospechar del conjunto. Tampoco debe tomarse demasiado en serio una correlación que se presente sin medida de confiabilidad (error probable, desviación estándar). Vigilemos los promedios cuya variedad no se especifica, en temas para los cuales quepa esperar que la

media y la mediana vayan a diferir sustancialmente. Cuidémonos de los porcentajes cuando no se nos informa el total sobre el cual fueron tomados.

A veces se mencionan los porcentajes, omitiendo el material numérico de base. Por ejemplo, una empresa informó que el 33 y 1/3% de las mujeres de su planta de trabajadoras se habían casado con miembros de la plana gerencial. Dadas las cifras brutas el cuadro queda más claro: al momento del informe había solamente tres mujeres en la planta de trabajadores y una de ellas se había casado con un gerente!!

Si nos enfrentamos con un índice, preguntemos qué falta. Tal vez sea la base, escogida con el objeto de tergiversar el resultado. A veces falta el factor determinante del cambio, por ejemplo cuando se informa de un aumento de la actividad comercial debido a incrementos en las ventas del mes de Abril con respecto a igual mes del año anterior. Pero lo que realmente se está omitiendo es el detalle que en el año considerado Pascua fue en Abril y en el año anterior fue en Marzo.

Pregunta 4: ¿DIO ALGUIEN CIERTO GIRO A LA INFORMACION?

Cuando compruebe una Estadística, busque la posible tendencia que alguien haya introducido en las cifras totales o en las conclusiones. Con frecuencia se da a conocer una cosa en lugar de otra.

Al comprobar un aumento en el número de casos de una cierta enfermedad no significa siempre una mayor virulencia de la misma. La victoria de un candidato en un distrito pequeño no tiene valor para una encuesta pública de mayores alcances. Que un número de lectores exprese preferencia por los artículos de información internacional, no quiere decir que lean estos artículo si se publican.

En cierta oportunidad, en un país, se informó que se había producido una vuelta al campo porque en el último censo se habían detectado medio millón más de granjas que en el anterior. Lo que realmente sucedió fue que la definición de granja había cambiado de un censo a otro.

La población de un área de China era de 29 millones de habitantes. Cinco años más tardes alcanzó los 105 millones. Muy poco de este aumento era real. El primer censo era para impuestos y el segundo para la lucha contra el hambre.

Pregunta 5: ¿TIENE SENTIDO?

Esta pregunta tiene mucha importancia cuando la aseveración se basa en un supuesto no probado. Existe una fórmula de amenidad en la lectura desarrollada por Rudolf Flesch. Se trata de medir la facilidad de lectura de un pasaje en prosa, por medio de criterios tan simples y objetivos como la longitud de las palabras y de las frases. Como todas las ideas que tienden a reducir a números lo imponderable y sustituyen el juicio por las matemáticas, esta es llamativa para decirlo de alguna manera. Al menos llamó la atención de quienes la usaron ya que, al aplicarla en algunos textos resultó que el “Cuento de la Bella Durmiente” era dos veces más difícil que la “República” de Platón. Aquí tenemos la historia de siempre: las Estadísticas se falsean en las propias narices de los usuarios. Se publican solamente porque la magia de los números anula el sentido común.

Las extrapolaciones son útiles, particularmente para la forma de predicción llamada previsión de tendencia. Pero al examinar las cifras o los gráficos derivados de las mismas, debe tenerse en presente una cosa: la tendencia actual tal vez sea un hecho, pero la tendencia futura sólo puede predecirse mediante una hipótesis razonable como la que lleva implícita la sentencia de que “siempre que lo demás siga siendo igual...” y que “las tendencias actuales continúen...”. De un modo u otro nada permanece igual, de lo contrario, la vida sería demasiado monótona.

Con esto he querido dar un breve recorrido por algunas de las tergiversaciones más frecuentes en Estadística y presentar un conjunto de preguntas básicas que debemos hacernos para iniciar un proceso de aceptación de la información Estadística que puede llegar a nuestras manos. Creo importante remarcar que el libro *Como Mentir con Estadísticas* de Darrell Huff sigue siendo hoy en día una obra muy esclarecedora en estas circunstancias.

Es evidente que en cada una de las consideraciones que siguieron a las cinco preguntas anteriores se podía identificar a una o más dificultades, presiones y vicisitudes por las que atraviesa el INDEC, el Sistema Estadístico Nacional y las Estadísticas Oficiales Argentinas. Como Sociedad Argentina de Estadística, no podemos permanecer indiferentes ante la mentira, el ocultamiento y el engaño a que nos vemos sometidos en cada informe referido a alguna cuestión de Estadísticas Oficiales.

Mucho menos podemos desviar la mirada ante el magnicidio que significa negar que grandes masas de personas permaneces por debajo de las líneas de indigencia y pobreza ya que ello es pura y simplemente condenarlos a la desaparición por falta de recursos mínimos, lo que en última instancia estaría violando el más básico de los Derechos Humanos: el de la Vida. Ya Su Santidad el Papa Benedicto XVI declaró a la pobreza en la Argentina como “un escándalo”, y el vocero del Episcopado Nacional afirmó que los obispos tratarán el tema y harán un informe sobre la pobreza, que para ellos está cerca del 40% y que la diferencia entre los más pobres y los más ricos “es inaceptable, un escándalo”. Similar situación de magnicidio ocurre cuando en algunas provincias argentinas se informa que “la mortalidad infantil ha disminuido”, cuando lo que realmente sucedió fue un simple cambio de definición: pasan a muchas muertes infantiles a considerarlas como muertes fetales. También preocupa que el falseamiento de la información lleve a que la incertidumbre desaliente la inversión deteriorando toda la actividad económica y llevando a que estemos cada día peor en cuanto a calidad de vida y, porqué no, a estabilidad y credibilidad política.

Adaptando algunos versos del conjunto Calle 13, raperos de Puerto Rico, podemos decir que “las Estadísticas del INDEC son sucias, pero lo que pasa es que la mente de ellos es flexible como gimnasta de Rusia”.

Es lamentable el manejo caprichoso que se está haciendo con la información estadística de nuestro Sistema Estadístico Nacional y la falta de un compromiso de querer convocar a expertos Estadísticos de Profesión a fin de que brinden sus aportes para mejorar la credibilidad de lo informado. Efectivamente, asistimos a la convocatoria de economistas, sociólogos, e inclusive veterinarios para que “brinden su opinión técnica”, pero prácticamente a casi ningún Estadístico de Profesión. Todos ellos conforman, en el mejor de los casos, un conjunto de usuarios de Estadísticas, pero bajo ningún aspecto los podemos considerar como expertos en nuestra área del conocimiento. Esperemos que esto se revierta. En estos casos debemos señalar que la Sociedad Argentina de Estadística está y estará siempre dispuesta a hacer los aportes técnicos que le sean requeridos en un ámbito de rigurosidad y transparencia.