



Módulo 1

Introducción a la Estadística

Autor

Dres. José Tessler y Juan Gagliardi

Abril de 2007

ÍNDICE

1. INTRODUCCIÓN A LA ESTADÍSTICA.....	2
1.1. ¿QUÉ ES LA ESTADÍSTICA?(*), ¿POR QUÉ NECESITAMOS USARLA?(*), ¿CUÁNDO ES NECESARIO UTILIZARLA?(*).....	2
1.2. <i>¿Es posible estimar con un error del 0 %,(*), un intervalo dentro del cual se encuentre el porcentaje de la población, cuando solamente tenemos los resultados de una muestra?</i>	<i>6</i>

Grupo de Estudio, Docencia e Investigación Clínica
Av. Rivadavia 2358 P.B. "4"
1034 Buenos Aires – Argentina
Tel./Fax: (011) 4952-4112
E-mail: gedic@advancedsl.com.ar / gedic@fibertel.com.ar
www.gedic.com.ar



Existe una versión animada para Power Point de la figura 1. Si Ud. desea utilizarla, imprima el texto y baje de Medioteca el archivo **introd.ppt**. Cada vez que leyendo el texto impreso encuentre un asterisco (*), debe pulsar el botón izquierdo del mouse, la teclas **Enter** o las flechas para abajo o la derecha. Si hay más de un asterisco, debe pulsar la tecla tantas veces como asteriscos haya. Para retroceder, debe pulsar las flechas para arriba o a la izquierda, cada vez que la pulse retrocederá un paso.

1. Introducción a la estadística

1.1. ¿Qué es la estadística? (*) ¿Por qué necesitamos usarla? (*) ¿Cuándo es necesario utilizarla? (*)

Para responder a estas preguntas construiremos un escenario (fig. 1): un sistema de medicina prepaga (*) desea saber qué porcentaje de sus 100.000 afiliados tiene antecedentes de vacunación antitetánica en los últimos 10 años. (*)

Hay 2 maneras (*) de poder conocer ese porcentaje. Una de ellas (*) consiste en llamar por teléfono a los 100.000 afiliados (*) y preguntarle a cada uno (o a sus padres, en el caso de niños) si se vacunaron contra el tétanos en los últimos 10 años. Luego de efectuados los llamados, (*) el resultado fue 23.704 vacunados, (*) es decir, (*) exactamente el 23,704 %.

(*) Este dato nos da una **certeza** (*) absoluta, pues no hay duda que 23.704 afiliados dijeron haber sido vacunados y no fueron 23.703 ni 23.705, por lo que el porcentaje es 23,704 % y no 23,703 % ni 23,705 %.

Grupo de Estudio, Docencia e Investigación Clínica
Av. Rivadavia 2358 P.B. "4"
1034 Buenos Aires – Argentina
Tel./Fax: (011) 4952-4112
E-mail: gedic@advancedsl.com.ar / gedic@fibertel.com.ar
www.gedic.com.ar

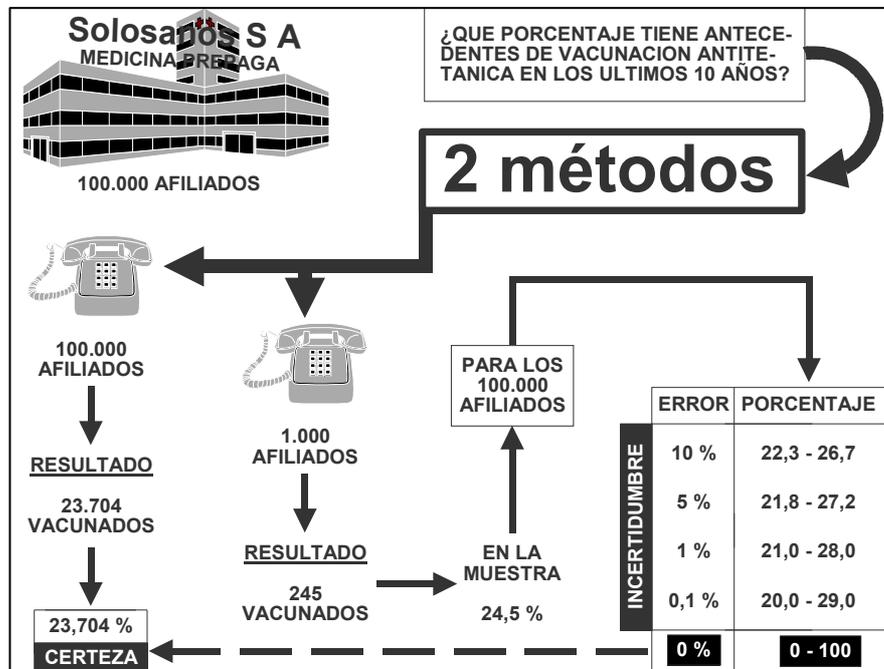


Fig. 1. Certeza e incertidumbre: Cuándo es necesario aplicar estadística.

El otro método(*) consiste en llamar solamente a una cantidad limitada de afiliados (muestra), por ejemplo, 1.000 afiliados(*) Es obvio que para que la muestra sea representativa tiene que haber sido tomada al azar. Una vez llamados esos 1.000 afiliados,(*) se estableció que 245 habían sido vacunados,(*) es decir,(*) el 24, 5 %.(*)

El problema está en que 24,5 % es el porcentaje entre los 1.000 individuos muestreados, pero aún no sabemos cuál es el porcentaje entre los 100.000 afiliados (población),(*) por lo que necesitamos proyectar el resultado hacia el conjunto de la población.(*) Al proyectar el resultado de la muestra a la población cometeremos siempre algún error(***) . Cuando aceptamos un error del 10 %,(*) tenemos una probabilidad del 90 % de que el porcentaje en la población sea un valor entre 22,3 y 26,7,(*) pero queda una probabilidad

Grupo de Estudio, Docencia e Investigación Clínica
Av. Rivadavia 2358 P.B. "4"
1034 Buenos Aires – Argentina
Tel./Fax: (011) 4952-4112
E-mail: gedic@advancedsl.com.ar / gedic@fibertel.com.ar
www.gedic.com.ar



del 10 % de que sea menor de 22,3 o mayor de 26,7. A medida que pretendemos disminuir el error, el intervalo se va ampliando.(****) Si pretendemos un error de solamente el 0,1 %, (*) tendremos una probabilidad del 99,9 % de que el porcentaje en la población sea un valor entre 20 y 29, (*) pero queda una probabilidad del 0,1 % de que sea menor de 20 o mayor de 29. En resumen, no obtenemos un número único como cuando llamamos a toda la población, sino un intervalo dentro del cual se

encontrará el porcentaje real en la población. Es decir, siempre tendremos **incertidumbre(*)** acerca del porcentaje real en la población.

Al llamar a los 100.000 afiliados obtuvimos con certeza que el porcentaje era 23,704 y no necesitamos hablar de probabilidades. En cambio cuando tomamos una muestra, tuvimos que utilizar probabilidades. Como vemos, las probabilidades son necesarias cuando no se dispone de la información de todos los individuos, pues en estos casos siempre existe incertidumbre. **Es importante remarcar que probabilidad e incertidumbre son dos conceptos indisolublemente ligados.**

La estadística es la parte de las matemáticas que estudia las probabilidades y las aplica a la solución de problemas. La estadística es, pues, necesaria solamente cuando no se dispone de los datos de todos los individuos, es decir, cuando hay incertidumbre.

Si nuestra hipótesis fuera que menos del 30 % de los afiliados han recibido alguna dosis de vacuna antitetánica en los últimos 10 años, al llamar a los 100.000 afiliados y encontrar que solamente el 23,704 % dice haber recibido la vacuna, hemos confirmado nuestra hipótesis sin necesidad de recurrir a la estadística. Si, en cambio, solamente llamamos a 1.000 afiliados, necesitamos de la estadística para aceptar o rechazar nuestra hipótesis, pero nunca podremos confirmarla pues siempre existe una incertidumbre, un margen de error.

Así como a la hipótesis la confirmamos al llamar a los 100.000 afiliados y solamente la aceptamos o rechazamos (sin seguridad absoluta) si nos limitamos a llamar a 1.000 afilia-

Grupo de Estudio, Docencia e Investigación Clínica
Av. Rivadavia 2358 P.B. "4"
1034 Buenos Aires – Argentina
Tel./Fax: (011) 4952-4112
E-mail: gedic@advancedsl.com.ar / gedic@fibertel.com.ar
www.gedic.com.ar



dos, también podemos decir que *calculamos* el porcentaje vacunado en el primer caso y *estimamos* un intervalo dentro del cual se encontrará el porcentaje, en el segundo caso. Es decir, con la estadística no calculamos los valores en la población sino que los estimamos. La estimación incluye un nivel de incertidumbre que no existe en el cálculo exacto efectuado sobre la población completa.

INDEC (Argentina) son las siglas de *Instituto Nacional De Estadísticas y Censos*. Esta institución efectúa 2 tipos distintos de relevamientos:

- La encuesta de hogares, una muestra sobre la que elabora **estadísticas** sobre empleo, pobreza, etc.
- **Censo** de población cada 10 años (además de otros censos). En este caso, los censistas visitan todos y cada uno de los hogares del país y recogen los datos de **todos** sus habitantes. Es decir, en este caso no produce estadísticas sino **datos censales**.

En nuestro ejemplo, cuando se llamó a los 100.000 afiliados se efectuó un censo y cuando se llamó solamente a 1.000 fue necesario aplicar estadística.

Está muy difundido el uso incorrecto del término “estadística” para referirse a datos censales. Incluso en algunos hospitales existe una oficina de Estadística que lo que hace en realidad es el censo de todas las actividad asistenciales del Hospital. Cuando hay

elecciones, las encuestas preelectorales o en boca de urna son estadísticas, pero los resultados de la elección son datos censales. Muchas veces los periodistas cometen el error de llamar “estadísticas” a estos resultados.

Grupo de Estudio, Docencia e Investigación Clínica
Av. Rivadavia 2358 P.B. “4”
1034 Buenos Aires – Argentina
Tel./Fax: (011) 4952-4112
E-mail: gedic@advancedsl.com.ar / gedic@fibertel.com.ar
www.gedic.com.ar



1.2. ¿Es posible estimar con un error del 0 %, (*) un intervalo dentro del cual se encuentre el porcentaje de la población, cuando solamente tenemos los resultados de una muestra?

Es decir, ¿es posible afirmar con absoluta certeza (*) entre qué valores se encuentra el porcentaje de vacunados en la población? La respuesta es **sí**: ese intervalo va del 0 al 100 % (fig. 1). (*) Es absolutamente seguro que el porcentaje de la población se encuentra entre el 0 y el 100 %. Obviamente, un intervalo tan amplio nos resulta totalmente inútil.

Ahora consideremos el primer caso, cuando llamamos a los 100.000 afiliados: ¿Es confiable el dato que el 23,704 % de los afiliados recibió alguna dosis de vacuna antitetánica en los últimos 10 años? Desde el punto de vista numérico sí, pero se plantea la

duda de si el llamar por teléfono a una persona y preguntarle si fue vacunada en los últimos 10 años es un método adecuado o no. Este ejemplo nos permite diferenciar entre el valor numérico de un dato y su validez metodológica. Es obvio que solamente deberían utilizarse datos válidos, para la validez de los datos no es un problema estadístico, sino metodológico y será tratado en un módulo posterior.

El ejemplo del párrafo anterior nos lleva a señalar que la estadística es solamente un aspecto de la planificación de un estudio y del análisis de los datos, pero los aspectos metodológicos son de fundamental importancia pues si éstos fallan, aplicaremos la estadística a datos no válidos.

Es importante señalar que los aspectos estadísticos son los más fáciles de corregir en un trabajo de investigación. Si al finalizar el trabajo resulta que la muestra fue insuficiente, es posible ampliarla (si ello estaba previsto). Si se objeta el método estadístico utilizado, es posible aplicar otro método a los mismos datos. Pero si el diseño no fue adecuado o si los métodos utilizados son objetables, no existe solución y se perdió todo el trabajo. Podemos afirmar que *la estadística es a la investigación lo que el electrocardiograma es al cardió-*

Grupo de Estudio, Docencia e Investigación Clínica
Av. Rivadavia 2358 P.B. "4"
1034 Buenos Aires – Argentina
Tel./Fax: (011) 4952-4112
E-mail: gedic@advancedsl.com.ar / gedic@fibertel.com.ar
www.gedic.com.ar



logo, es decir, un elemento necesario pero no suficiente. Es un serio error limitar la planificación de un trabajo y el análisis de los datos a los aspectos estadísticos.

Grupo de Estudio, Docencia e Investigación Clínica
Av. Rivadavia 2358 P.B. "4"
1034 Buenos Aires – Argentina
Tel./Fax: (011) 4952-4112
E-mail: gedic@advancedsl.com.ar / gedic@fibertel.com.ar
www.gedic.com.ar