

Capítulo 2: Intervalos de confianza

Un problema habitual es el de estimar parámetros que ayuden a caracterizar una variable. Por ejemplo el porcentaje de individuos que mejora ante un cierto tratamiento, o el tiempo medio que tarda un anestésico en hacer efecto.

Podríamos decir, tras realizar un estudio, *“el 75% de los pacientes tratados experimentó una mejoría”*. Una respuesta más sofisticada usando intervalos de confianza podría ser: *“Nuestro estudio muestra que en el 75% de los casos se experimenta una mejoría, siendo el margen de error del 6%. El nivel de confianza es del 95%”*.

El cálculo de intervalos de confianza para la estimación de parámetros son técnicas que nos permiten hacer declaraciones sobre qué valores podemos esperar para un parámetro. El intervalo calculado dependerá de:

- Lo estimado en la muestra (porcentaje, media,...) El intervalo de confianza esta formado por valores ligeramente menores y mayores que la aproximación ofrecida por la muestra.
- El tamaño muestral. Cuantos más datos hayan participado en el cálculo, más pequeño esperamos que sea la diferencia entre el valor estimado y el valor real desconocido.
- La probabilidad (nivel de confianza) con la que el método dará una respuesta correcta. Niveles de confianza habituales para los intervalos de confianza son el 95% y el 99%.

Puede parecer sorprendente que no busquemos respuestas con una confianza del 100%, pero ocurre que en ese caso, los intervalos serían tan grandes que no serían de gran provecho. La elección de un nivel de confianza como el 95% es un compromiso entre hacer declaraciones con una razonable probabilidad de acertar, y que además el intervalo declarado, sea lo suficientemente pequeño como para suscitar algún interés. El nivel de confianza hay que interpretarlo como que disponemos de un método de calcular intervalos que seguido con rigor, en cierto porcentaje de casos acierta (nivel de confianza) y en el resto falla.

2.1 Error típico o estándar

En multitud de ocasiones al utilizar un programa estadístico encontramos junto a las más diversas estimaciones como una media, una proporción, un coeficiente de regresión, un coeficiente de asimetría, etc., una cantidad misteriosa denominada **error estándar**² o también **error típico**.

El error estándar tiene mucho que ver con los intervalos de confianza. Para muchos parámetros, su intervalo de confianza es habitualmente la estimación obtenida sobre la

² El error estándar depende del parámetro que estemos calculando y de la distribución de la variable. Es una medida de la variabilidad del estimador. Su cálculo exacto es salvo excepciones y sin hacer simplificaciones, excesivamente complicado. Estas simplificaciones suelen ser del tipo que la variable en estudio tiene una distribución conocida, o que la muestra es lo suficientemente grande para considerar algunas aproximaciones adecuadas. No entraremos en más detalles sobre esta espinosa cuestión.

muestra (proporción, media,...), y un margen de error que nos es más que un múltiplo del error estándar. Un ejemplo muy común, consiste en elegir niveles de confianza del 95%. Para ello un margen de error de dos errores estándar es habitualmente la respuesta. Simple, ¿verdad?

Cuando un programa estadístico nos ofrece una estimación de una cantidad, junto a su error estándar, podemos estar seguros de que si se dan ciertas condiciones de validez, el estimador del parámetro tiene comportamiento normal cuya desviación típica es el error estándar, y por tanto una declaración como: “*La proporción de pacientes que mejoraron con el tratamiento es del 75%, con un error estándar del 3%*” puede enunciarse de forma más clara en forma de intervalo de confianza:

- *La proporción de pacientes que mejoró fue del 75%, con un margen de error de 2 EE=6%. El nivel de confianza es del 95%.*
- *La proporción de pacientes que mejoró fue del 75%, con un margen de error de 2.5 EE=7.5%. El nivel de confianza es del 99%.*

Lo anterior se deduce de que para distribuciones normales³ tenemos que alrededor del valor central y en un radio de 2 y 2.5 desviaciones típicas tenemos aproximadamente una masa de probabilidad del 95% y 99% respectivamente, que son los valores usados habitualmente como niveles de confianza.

2.2 Intervalo de confianza para una proporción

Raro será que veamos alguna vez un estudio donde se pregunta a los individuos si les gusta determinado producto y nos encontremos como respuesta simplemente el porcentaje de satisfechos. Como hemos mencionado anteriormente en la respuesta esperaríamos encontrar un margen de error, a ser posible pequeño (indicador de que la muestra es representativa) y con un nivel de confianza alto, por ejemplo 95%.

Para poder dar una respuesta es necesario calcular el error estándar. Si la muestra es suficientemente grande (más de 50 individuos por ejemplo), y el porcentaje a estimar, π , no es muy extremo (cercano al 0% o al 100%), la distribución del estimador de la proporción, p , es aproximadamente normal. El error estándar puede ser aproximado mediante:

$$EE = \frac{\sqrt{p \cdot (1-p)}}{\sqrt{n}}$$

Ejemplo: Se preguntó a 80 pacientes si habían sufrido algún trastorno tras seguir un tratamiento, de los cuales 60 ($p=60/80=3/4=75\%$) dijeron que no. La muestra es grande y no esperamos que el porcentaje real en caso de haber sido extendido a muchos más pacientes sea muy diferente. Por tanto el error estándar es:

$$EE = \sqrt{p \cdot (1-p)} / \sqrt{n} = \sqrt{0.75 \cdot 0.25} / \sqrt{80} = 0.048 = 4.8\%$$

Podemos decir, pues, que el 75% de los individuos no mencionaron haber sufrido trastornos, con un margen de error de $2 \cdot EE = 9.6\%$. La confianza es del 95%.

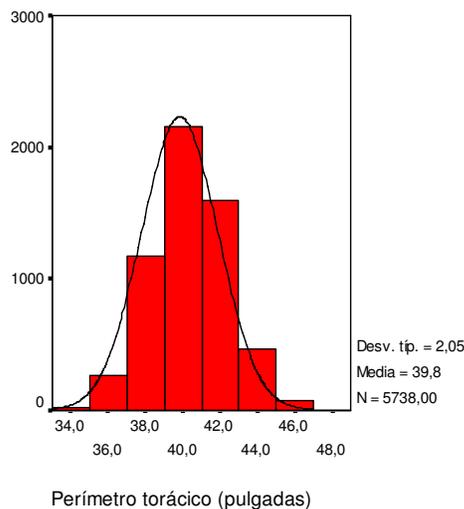
³ La distribución de los estimadores habitualmente es aproximadamente normal, aunque la distribución de la variable original no lo sea.

2.3 Intervalo de confianza para una media

Otra caso muy común es el de estimar el valor medio de una variable numérica. Como sabemos la media es interesante como medida de centralización cuando la distribución de la misma es más o menos normal. En este caso el error estándar se puede aproximar con:

$$EE = \frac{S}{\sqrt{n}}$$

En este caso, incluso podemos hacer estimaciones aunque las muestras sean pequeñas, con sólo verificarse que la distribución de la variable es normal, ya que la distribución del estimador es conocida de forma exacta como distribución t-student. En cualquier caso si las muestras son grandes, aunque la distribución de los datos no sea normal, la media como en el caso de las proporciones se distribuye de manera aproximadamente normal. Y el mismo método que usamos para proporciones sigue siendo válido.



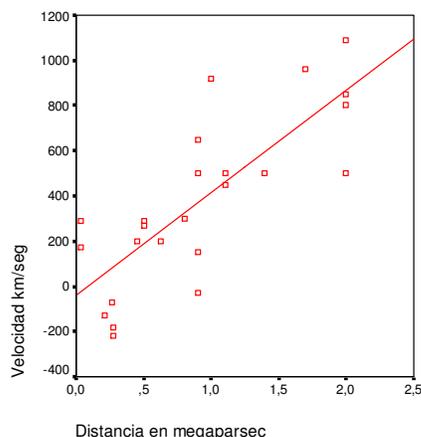
Ejemplo: En un trabajo de Quetelet se estudia la distribución del perímetro torácico medido en pulgadas de militares escoceses de principios del siglo XIX. Los resultados se muestran en la gráfica, y aparentan una distribución normal. La media es 39.8 y la desviación típica 2.05; El tamaño de la muestra es de 5738 individuos, por tanto el error estándar es:

$$EE = 2.05 / \sqrt{5738} = 0.027$$

Podemos decir que *el perímetro torácico medio es de 39.8 pulgadas con un margen de error de $2 \cdot EE = 0.054$ pulgadas. La confianza es del 95%.*

2.4 Intervalos de confianza para otros parámetros

Hay una técnica de estimación que usan casi todos los programas estadísticos que es la estimación por el método de máxima verosimilitud. Con ella es frecuente conseguir estimadores con muy buenas propiedades⁴, y la técnicas anteriores se extienden con facilidad. En especial, cuando junto a la estimación de un parámetro observamos un término denominado error estándar o error típico, podemos confiar en que si la muestra es suficientemente grande, las técnicas anteriores se extienden sin cambios.



Ejemplo: En 1929 Edwin Hubble estudió la relación existente entre la distancia de 29 nebulosas y la velocidad radial con respecto a la tierra. El análisis de los datos muestra que un modelo de regresión lineal es una buena aproximación, siendo proporcional la velocidad de separación a la distancia, lo que se le dio la idea de

son aproximadamente normales, de la menor variabilidad

que el universo estaba en expansión. En un análisis de regresión lineal se aprecia que el término constante es -40 , pero siendo el error típico de 83.4 , podemos decir que con una confianza del 95% (2 EE), el intervalo de confianza para dicha constante es $-40.78 \pm 2 \cdot 83.4$, que incluye al valor cero. Eso es una indicación de que el término constante puede ser eliminado del modelo por simplicidad. En cuanto al coeficiente que acompaña a la distancia se puede decir que es 454.2 con un error de $2 \cdot 75.2 = 150.4$, para un nivel de confianza del 95% . Como se advierte, dicho intervalo no contiene al cero, y sugiere como estimación de la velocidad radial en función de la distancia, el modelo carente de término constante:

$Velocidad = H_0 \cdot distancia$, donde $H_0 = 454.2$ es la estimación de la *constante de Hubble*, para la expansión del universo.

Modelo	Coeficientes no estandarizados		t	Sig.
	B	Error típ.		
1 (Constante)	-40,784	83,439	-,489	,630
Distancia en megaparsec	454,158	75,237	6,036	,000

2.5 Contrastes de hipótesis basados en intervalos de confianza

Aunque no hemos definido aún lo que es un contraste de hipótesis, en el último ejemplo se ha dejado entrever que es posible tomar ciertas decisiones como, por ejemplo, rechazar que la constante de Hubble sea cero, basándonos en que el intervalo de confianza al 95% no contiene a dicho valor. Hay una relación estrecha entre intervalos de confianza y contrastes de hipótesis. Buena parte de las hipótesis que se contrastan pueden rechazarse si dicha hipótesis establece un valor para el parámetro que no pertenece al intervalo de confianza. En un entorno clínico es preferible, con mucho, mostrar un intervalo de confianza que simplemente el resultado del contraste.