Apuntes - Inferencia estadística

Curso: Universidad 1

La inferencia estadística utiliza una muestra para sacar conclusiones sobre una población. Se estudian dos problemas básicos: estimar un parámetro (media, proporción) con un margen de confianza y contrastar hipótesis sobre él. Se apoya en la probabilidad del tema anterior, sobre todo en la distribución normal.

Actividad de introducción

Antes de unas elecciones, un instituto de opinión entrevista a unas mil personas y publica que un partido obtendrá entre el 35 % y el 41 % de los votos. ¿Cómo puede una muestra tan pequeña decir algo sobre millones de votantes? ¿Y qué significa exactamente ese intervalo? La inferencia estadística responde a estas preguntas: permite sacar conclusiones sobre una población a partir de una muestra, cuantificando la incertidumbre. Con ella se estiman parámetros mediante intervalos de confianza y se toman decisiones con contrastes de hipótesis, por ejemplo para aprobar un medicamento, controlar la calidad de una fábrica o comprobar si un método docente funciona.

Muestreo y distribución de los estimadores

Definición: muestreo y distribución de los estimadores

Una **muestra aleatoria simple** $X_1, \dots, X_n$ son variables independientes con la distribución de la población (media μ, desviación σ). Un **estimador** es una función de la muestra: la media muestral $\bar{X} = \frac{1}{n}\sum X_i$ estima μ y la cuasivarianza $s^2 = \frac{1}{n-1}\sum (X_i - \bar{X})^2$ estima σ^2. Un estimador es **insesgado** si su esperanza es el parámetro; $\bar{X}$ lo es y cumple $E[\bar{X}] = \mu$, $\operatorname{Var} \bar{X} = \sigma^2/n$. Si la población es normal (o n es grande, por el **teorema central del límite**), $\bar{X} \sim N(\mu, \sigma/\sqrt{n})$. Para una proporción, $\hat{p} \approx N\left(p, \sqrt{p(1-p)/n}\right)$.

Intervalos de confianza

Definición: intervalos de confianza

Un **intervalo de confianza** al nivel $1-\alpha$ para un parámetro es un intervalo calculado con la muestra que contiene al parámetro con probabilidad $1-\alpha$ en el muestreo repetido. Para la media con σ conocida: $\bar{x} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$ con $z_{0{,}025} = 1{,}96$ y $z_{0{,}005} = 2{,}576$. Si σ es desconocida y la población normal se usa la t de Student con $n-1$ grados de libertad: $\bar{x} \pm t_{n-1,\alpha/2}\frac{s}{\sqrt{n}}$. Para una proporción: $\hat{p} \pm z_{\alpha/2}\sqrt{\hat{p}(1-\hat{p})/n}$. El **margen de error** E disminuye al aumentar n (como $1/\sqrt{n}$) y aumenta con la

confianza. Tamaño muestral necesario: $n \geq \left(\frac{z_{\alpha/2}\sigma}{E}\right)^2$.

Contrastes de hipótesis

Propiedad

Se enfrentan la **hipótesis nula** H_0 (lo que se supone cierto) y la **alternativa** H_1. Se calcula un **estadístico** con la muestra; si cae en la **región crítica** se rechaza H_0. El **error de tipo I** es rechazar H_0 siendo cierta (probabilidad α, nivel de significación); el **de tipo II** es no rechazarla siendo falsa (probabilidad β); la **potencia** es $1-\beta$. El **p-valor** es la probabilidad, bajo H_0, de un resultado al menos tan extremo como el observado: se rechaza si p. No rechazar H_0 no prueba que sea cierta, solo que no hay evidencia suficiente en su contra.

Propiedad

Contraste de la media ($H_0 : \mu = \mu_0$): estadístico $z = \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}$ (o t con s). Bilateral al 5 %: se rechaza si $|z| > 1{,}96$. Unilateral al 5 %: $z > 1{,}645$. Para la proporción, $z = \frac{\hat{p}-p_0}{\sqrt{p_0(1-p_0)/n}}$. La **bondad de ajuste** usa $\chi^2 = \sum \frac{(O_i-E_i)^2}{E_i}$ con $k-1$ grados de libertad.

Método: intervalo de confianza para la media

1. Identifica media muestral, desviación y tamaño n.
2. Elige el valor crítico z para el nivel de confianza.
3. Calcula el margen z·σ/√n.
4. Escribe el intervalo (media − margen, media + margen).

Ejemplo 1

Con $\sigma = 15,\ n = 36,\ \bar{x} = 100$: margen $1{,}96 \cdot \frac{15}{6} = 4{,}9$.

El intervalo al 95 % es $(95{,}1;\ 104{,}9)$.

Ejemplo 2

Contraste $H_0 : \mu = 100$ frente a $H_1 : \mu \neq 100$ con $\bar{x} = 105,\ \sigma = 15,\ n = 36$: $z = \frac{5}{2{,}5} = 2$**;** $p = 2(1-\Phi(2)) = 0{,}0455$**.**

Se rechaza H_0 al 5 %.

Ejemplo 3

En $n = 400$ individuos, $\hat{p} = 0{,}3$: margen $1{,}96\sqrt{\frac{0{,}3 \cdot 0{,}7}{400}} \approx 0{,}045$.

Intervalo $(0{,}255;\ 0{,}345)$.

¡Atención!

Interpretar «95 % de confianza» como probabilidad de que el parámetro esté en el intervalo ya calculado; confundir el p-valor con la probabilidad de que H_0 sea cierta; decir que «se acepta H_0»; usar σ en lugar de $\sigma/\sqrt{n}$ en el error típico; olvidar que la proporción necesita n grande.

Cómo razonar en inferencia

Definición: muestreo aleatorio

Una muestra aleatoria simple de tamaño n se obtiene eligiendo los individuos de forma que todos tengan la misma probabilidad de ser seleccionados y que las elecciones sean independientes. En la práctica se usan también el muestreo estratificado, que divide la población en grupos homogéneos, como comunidades autónomas o franjas de edad, y toma una muestra de cada uno en proporción a su tamaño, y el muestreo por conglomerados, que selecciona al azar grupos completos, como colegios o barrios. Un buen diseño de muestreo es la condición previa de toda inferencia: si la muestra no representa a la población, las fórmulas de los intervalos y de los contrastes dan resultados precisos pero equivocados.

Observación: nivel de confianza y significación

El nivel de confianza $1 - \alpha$ de un intervalo y el nivel de significación α de un contraste están relacionados: un intervalo al 95 % corresponde a contrastes bilaterales al 5 %. Elegir α es una decisión que depende de las consecuencias de equivocarse. En física de partículas se exige una evidencia de cinco desviaciones típicas antes de anunciar un descubrimiento, mientras que en estudios exploratorios de ciencias sociales es habitual trabajar al 5 % o incluso al 10 %.

Método: contraste de hipótesis

1. Plantea la hipótesis nula H_0, que expresa «no hay efecto» o el valor de referencia, y la alternativa H_1, que puede ser bilateral o unilateral según la pregunta.
2. Fija el nivel de significación α, normalmente 0,05.
3. Calcula el estadístico de contraste suponiendo que H_0 es cierta.
4. Determina la región crítica o calcula el p-valor, la probabilidad de obtener un resultado tan extremo como el observado si H_0 fuera cierta.
5. Decide: se rechaza H_0 si el estadístico cae en la región crítica, es decir, si el p-valor es menor que α. Redacta la conclusión en el lenguaje del problema.

Definición: estimador insesgado y eficiente

Un estimador es una función de la muestra que se usa para aproximar un parámetro. Es insesgado si su esperanza coincide con el parámetro, de modo que no se equivoca sistemáticamente hacia arriba ni hacia abajo, y es más eficiente que otro si tiene menor varianza. La media muestral es un estimador insesgado de la media poblacional, con varianza $\frac{\sigma^2}{n}$, y la cuasivarianza es un estimador insesgado de la varianza.

Propiedad: amplitud de un intervalo

La amplitud de un intervalo de confianza para la media es $2z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$. Aumenta con el nivel de confianza y con la dispersión de la población, y disminuye con el tamaño de la muestra, pero solo como $\frac{1}{\sqrt{n}}$: para reducir el error a la mitad hay que cuadruplicar la muestra. Por eso las encuestas rara vez superan unos pocos miles de entrevistas.

Ejemplo 4

Una muestra de 49 lámparas da una duración media de 1200 h. Si $\sigma = 140$ h, construye el intervalo al 95 %.

El error estándar es $\frac{140}{7} = 20$ h y el margen $1{,}96 \cdot 20 = 39{,}2$ h. El intervalo es $[1160{,}8;\ 1239{,}2]$ h.

Ejemplo 5

Un fabricante asegura que el 90 % de sus envíos llegan a tiempo. De 400 envíos, llegan a tiempo 344. ¿Es creíble al 5 %?

$\hat{p} = 0{,}86$ y, bajo $H_0 : p = 0{,}9$, el error estándar es $\sqrt{0{,}9 \cdot 0{,}1/400} = 0{,}015$. Entonces $z = \frac{0{,}86 - 0{,}9}{0{,}015} \approx -2{,}67$: se rechaza la afirmación del fabricante.

Ejemplo 6

En una muestra de 16 cafés servidos en una cafetería de Salamanca, la temperatura media es 68 °C con $s = 4$ °C. Da un intervalo al 95 % ($t_{15;\,0{,}025} = 2{,}131$).

Margen $2{,}131 \cdot \frac{4}{4} = 2{,}131$ °C. Intervalo: $[65{,}87;\ 70{,}13]$ °C. Se usa la t porque la muestra es pequeña y σ es desconocida.

Interpretar bien los resultados

Los conceptos de la inferencia se malinterpretan con frecuencia, incluso en la prensa científica. Un intervalo de confianza al 95 % no significa que el parámetro tenga un 95 %

de probabilidad de estar dentro de él: el parámetro es un número fijo, y lo que es aleatorio es el intervalo. Significa que el método acierta en el 95 % de las muestras posibles. Del mismo modo, el p-valor no es la probabilidad de que la hipótesis nula sea cierta, sino la probabilidad de obtener datos tan extremos si lo fuera. No rechazar H_0 no demuestra que sea verdadera: puede que la muestra sea pequeña y el contraste tenga poca potencia. Y un resultado estadísticamente significativo no siempre es importante en la práctica: con muestras enormes, diferencias minúsculas resultan significativas. Por eso conviene acompañar cada contraste de un intervalo de confianza que muestre el tamaño del efecto. Además, todo el razonamiento depende de que la muestra sea aleatoria: una encuesta hecha solo por internet o a voluntarios puede estar sesgada, y ningún cálculo corrige un mal diseño del muestreo.

¡Atención!

En el intervalo para una proporción se usa $\hat{p}$ en el error estándar, pero en un contraste se usa el valor p_0 de la hipótesis nula. La hipótesis nula siempre lleva el signo igual. Y elegir la alternativa unilateral después de ver los datos para conseguir significación es una mala práctica.

¿Sabías que?

William Gosset, químico de una fábrica de cerveza de Dublín, publicó en 1908 la distribución t con el seudónimo Student, porque su empresa no permitía a los empleados firmar artículos. Necesitaba sacar conclusiones con muestras muy pequeñas de cebada. Ronald Fisher y Jerzy Neyman desarrollaron después la teoría de los contrastes que se usa hoy en todos los ensayos clínicos.

El consejo de Otto

escribe siempre antes de calcular: parámetro, hipótesis, nivel de significación y estadístico. Un esquema fijo evita casi todos los errores de planteamiento.