Figura: ¡Mas estadística!

Probabilidad - Clase 12
Intervalos de confianza, test de hipótesis

Ernesto Mordecki

Facultad de Ciencias, Universidad de la República. Montevideo, Uruguay

Curso de la Licenciatura en Matemática - 2020

Contenidos

Intervalo de confianza

- Mediante máxima verosimilitud obtuvimos un valor aproximado de p, que es $\hat{p} = \mu(\omega)/n$, que se aproxima a p cuando n es grande.
- Queremos ahora dar un intervalo, de forma de tener una cierta certeza de no equivocarnos.
- Supongamos que esa certeza (o confianza) es de un 95%
- ¿Cómo hacemos?

- Los teoremas que vimos vienen en nuestra ayuda:
- Sabemos acotar y aproximar la probabilidad de un intervalo de la forma:

$$|\hat{p} - p| < \varepsilon$$

- Pero ahora lo que conocemos es $\hat{p}$, entonces el intervalo de confianza para p es de la forma

$$\hat{p} - \varepsilon < p < \hat{p} + \varepsilon.$$

- El problema es calcular ε para que el intervalo contenga 0.95 de probabilidad o más.

Teorema Central del Límite: TCL

Del teorema se obtiene, para $a > 0$, que

$$\mathbf{P}\left(-a < \frac{\mu - np}{\sqrt{npq}} \leq a\right) \sim \frac{1}{\sqrt{2\pi}} \int_{-a}^{a} e^{-x^2/2} dx.$$

Tenemos

$$\mathbf{P}\left(-a < \frac{\mu - np}{\sqrt{npq}} \leq a\right) = \mathbf{P}\left(-\frac{a}{\sqrt{n}}\sqrt{pq} < \frac{\mu}{n} - p \leq \frac{a}{\sqrt{n}}\sqrt{pq}\right)$$

Es decir

$$\varepsilon = \frac{a}{\sqrt{n}}\sqrt{pq} \sim \frac{a}{\sqrt{n}}\sqrt{\hat{p}(1-\hat{p})}$$

Nos queda determinar a tal que

$$\frac{1}{\sqrt{2\pi}} \int_{-a}^{a} e^{-x^2/2} dx = 0.95.$$

Calculamos con la distribución normal. Queremos a tal que

$$\frac{1}{\sqrt{2\pi}}\int_{-a}^{a} e^{-x^2/2}dx = 0.95.$$

Es decir, por simetría, a verifica

$$\frac{1}{\sqrt{2\pi}}\int_{0}^{a} e^{-x^2/2}dx = \frac{0.95}{2} = 0.475$$

Entonces,

$$\Phi(a) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{a} e^{-x^2/2}dx = 0.5 + 0.475 = 0.975.$$

Entonces, tenemos que hallar a tal que

$$\Phi(a) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a} e^{-x^2/2} dx = 0.975.$$

Utilizamos la función inversa de Φ en R :

```
> qnorm(0.975)
[1] 1.959964
```

Obtuvimos que $a = 1.96$. El intervalo era

$$\varepsilon = \frac{a}{\sqrt{n}}\sqrt{pq} \leq \frac{a}{2\sqrt{n}} = \frac{1.96}{2\sqrt{n}} = \frac{1.96}{2\sqrt{n}} = \frac{0.98}{\sqrt{n}}.$$

Ejemplo: un plebiscito

- Nos interesa conocer la opinión de una población respecto de un tema que se va a plebiscitar.
- El plebiscito se aprueba si recibe más de la mitad de los votos.
- Supongamos entonces que hacemos una encuesta a 1000 ciudadanos, y tenemos un porcentaje de afirmativos de 450 personas.
- Construir intervalos de confianza del 95 % y otro del 99% para la proporción verdadera en la población. Se usa decir que $\alpha = 0.05$, el error admisible del intervalo.

Solución (acotación de p)

- Tenemos $\hat{p} = \frac{450}{1000} = 0.45$.
- Por otra parte, acotando:

$$\varepsilon = \frac{1.96}{\sqrt{n}}\sqrt{\hat{p}(1-\hat{p})} \leq \frac{1.96}{2\sqrt{n}} \sim \frac{1}{\sqrt{1000}} = 0.032.$$

- El intervalo **estimado** para p es entonces de la forma

$$[0.45 - 0.032, 0.45 + 0.032] = [0.418, 0.482].$$

Solución (estimación de p)

- Tenemos $\hat{p} = \frac{450}{1000} = 0.45$.
- El intervalo para p tiene amplitud:

$$\varepsilon = \frac{1.96}{\sqrt{n}} \sqrt{\hat{p}(1-\hat{p})} = \frac{1.96}{\sqrt{n}} \sqrt{0.45 \times 0.55} = \frac{0.9751}{\sqrt{1000}} = 0.0308.$$

- Ambos intervalos son muy parecidos, pero este segundo es mas preciso.

Ejercicio

Calcular el intervalo de confianza para $\alpha = 0.10$. Repasemos:

$$\mathbf{P}\left(-a < \frac{\mu - np}{\sqrt{npq}} \leq a \right) \sim \frac{1}{\sqrt{2\pi}} \int_{-a}^{a} e^{-x^2/2} dx.$$

equivalente a

$$\mathbf{P}\left(-\frac{a}{\sqrt{n}}\sqrt{pq} < \frac{\mu}{n} - p \leq \frac{a}{\sqrt{n}}\sqrt{pq} \right)$$

Lo que nos da una amplitud de intervalo

$$\varepsilon = \frac{a}{\sqrt{n}}\sqrt{pq} \leq \frac{a}{2\sqrt{n}}.$$

Calculamos con la distribución normal. Queremos a tal que

$$\frac{1}{\sqrt{2\pi}}\int_{-a}^{a} e^{-x^2/2}dx = 0.90.$$

Es decir, por simetría, a verifica

$$\frac{1}{\sqrt{2\pi}}\int_{0}^{a} e^{-x^2/2}dx = \frac{0.90}{2} = 0.45$$

Entonces,

$$\Phi(a) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{a} e^{-x^2/2}dx = 0.5 + 0.45 = 0.95.$$

Entonces, tenemos que hallar a tal que

$$\Phi(a) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a} e^{-x^2/2} dx = 0.995.$$

Utilizamos la función inversa de Φ en R :

```
> qnorm(0.95)
[1] 1.644854
```

Obtenemos entonces que $a = 1.645$. El intervalo tiene amplitud

$$\varepsilon = \frac{a}{\sqrt{n}}\sqrt{pq} \leq \frac{1.645}{2\sqrt{n}} = \frac{0.82}{\sqrt{n}}.$$

0.99

Por último calculamos el intervalo de confianza para $\alpha = 0.01$ (confianza 0.99!). La amplitud es

$$\varepsilon = \frac{a}{\sqrt{n}}\sqrt{pq} \leq \frac{a}{2\sqrt{n}}.$$

donde *a* verifica

$$\Phi(a) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{a} e^{-x^2/2}dx = 0.5 + \frac{0.99}{2} = 0.995.$$

```
> qnorm(0.995)
[1] 2.575829
```

da

$$\varepsilon = \frac{a}{\sqrt{n}}\sqrt{pq} \leq \frac{2.576}{2\sqrt{n}} = \frac{1.29}{\sqrt{n}}.$$

Intervalos de confianza con **acotación** de p:

confianza	0.90	**0.95**	0.99
error α	0.10	**0.05**	0.01
amplitud ε	$0.82/\sqrt{n}$	$\mathbf{0.98/\sqrt{n}}$	$1.29/\sqrt{n}$

- Mayor confianza requiere mayor amplitud
- El denominador de la amplitud es siempre $\sqrt{n}$

Intervalos de confianza con **estimación** de p:

confianza	0.90	**0.95**	0.99
error α	0.10	**0.05**	0.01
a (percentil)	1.645	**1.96**	2.58
amplitud ε	$1.645\frac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}}$	$\mathbf{1.96}\frac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}}$	$2.58\frac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}}$

Table: En **negrita** los valores mas usuales de intervalos de confianza (95%)

Test de hipótesis

- Supongamos que tiramos una moneda y obtenemos 550 caras.
- ¿está equilibrada esa moneda?
- Es decir: las 50 caras de exceso de las esperadas, ¿obedecen al azar o a un desequilibrio?

- Esa pregunta se formula matemáticamente mediante un test de hipótesis.
- Para eso se formula una hipótesis (la hipótesis nula), que se denota H_0 (se lee "hache cero")
- y se formula una hipótesis alternativa hache uno: H_1
  - $H_0: \quad p = p_0 \; (1/2)$
  - $H_1: \quad p \neq p_0 \; (1/2)$
- La idea es la siguiente: en principio pensamos que H_0 es verdadera, y la rechazamos si hay evidencia suficiente en los datos.

El procedimiento es el siguiente:

- Construímos un intervalo I asumiendo que H_0 es cierta de nivel α.
- Ese intervalo tiene centro en p_0 y amplitud ε con $\hat{p}$ estimado.
- El complemento $S = [0, 1] \setminus I$ se llama **región crítica**
- Calculamos $\hat{p}$ el estimador de p
- Si $\hat{p} \notin S$ (el estimador no pertenece a la región crítica): **no rechazamos la hipótesis nula**
- Si $\hat{p} \in S$ (el estimador pertenece a la región crítica): **rechazamos la hipótesis nula**

Ejemplo

Tomamos como ejemplo las 550 caras de la tirada de la moneda. Asumimos $\alpha = 0.05$.

- El intervalo es $(\frac{1}{2} - \varepsilon, \frac{1}{2} + \varepsilon)$
- Tenemos [1]

$$\varepsilon = 1.96 \frac{\sqrt{1/4}}{\sqrt{n}} = 1.96 \frac{1}{2\sqrt{1000}} = 0.0309.$$

- El intervalo es $(0.5 - 0.0309, 0.5 + 0.0309) = (0.4691, 0.5309)$
- ¿Cuál es la conclusión del test?

[1] Acá corregimos usando $p_0 = 1/2$ para calcular la región crítica.

- **Rechazamos** la hipótesis nula, porque $0.55 \notin (0.4691, 0.5309)$.
- Decimos que hay evidencia estadística suficiente para decidir que la moneda no está equilibrada.
- ¿Que hubiera pasado si teníamos 5 500 caras en 10 000 tiradas?
- ¿Y si tenemos 55 caras en 100 tiradas?

- Si tenemos el mismo $\hat{p}$ con un n mayor, el intervalo se reduce y la conclusión es la misma.
- Si tenemos[2] $n = 100$, la amplitud es

$$\varepsilon = 1.96 \frac{\sqrt{1/2}}{\sqrt{n}} = 1.96 \frac{1}{2\sqrt{100}} = 0.098.$$

- El intervalo es $(0.5 - 0.098, 0.5 + 0.0975) = (0.402, 0.598)$.
- Como $\hat{p} \in I$ no hay evidencia suficiente y **no rechazamos** la hipótesis nula.

[2] Misma corrección

Variantes de la hipótesis alternativa

La hipótesis alternativa puede tener otras formas, dependiendo de la formulación de la pregunta o de información a prior que tengamos:

- Hipótesis alternativa simple: $H_1 : p = p_1$ (con $p_1 \neq p_0$)
- Hipótesis alternativa unilateral $H_1 : p > p_0$
- Hipótesis alternativa unilateral $H_1 : p < p_0$

En cada uno de estos tests se construyen intervalos y regiones críticas que son unilaterales (no las vamos a ver)

Potencia de un test

Cuando la hipótesis alternativa es simple, podemos calcular cuan bien se desempeña el test con el siguiente argumento.

- Suponemos que la verdadera hipótesis verdadera es H_1.
- Calculamos entonces cual es la probabilidad, siendo cierta H_1, de que el estimador **caiga** en la región crítica (como debería de ser).
- Esa probabilidad, que se espera sea alta si el test se desempeña correctamente, se llama **potencia** del test, y se denota mediante π.
- Su complemento, $\beta = 1 - \pi$ es el error de tipo 2 (falso negativo: aceptar un falso)
- (α se llama error del tipo 1, y es el error de tener un falso postivo (rechazar un verdadero)

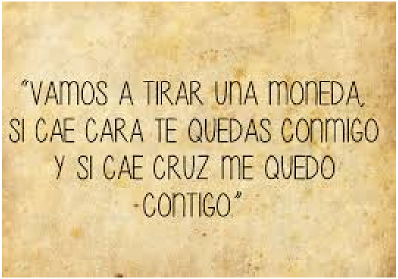

Figura: ¿Cara o número?