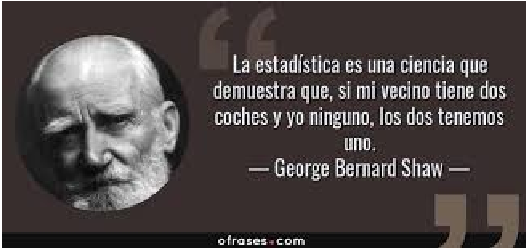
La estadística es una ciencia que demuestra que, si mi vecino tiene dos coches y yo ninguno, los dos tenemos uno.
— George Bernard Shaw —
ofrases.com

Probabilidad - Clase 13
Test de hipótesis, hipótesis unilaterales, potencia de un test.

Ernesto Mordecki

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

Curso de la Licenciatura en Matemática - 2020

Contenidos

Teorema Central del Límite: TCL

Del TCL, para $a > 0$:

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

Tenemos

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

La **amplitud** es:

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

Nos queda determinar a tal que

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

Entonces, tenemos que hallar a tal que

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

Utilizamos la función inversa de Φ en R :

$$\texttt{qnorm(1-}\alpha/2\texttt{)=}a\texttt{.}$$

La amplitud entonces es

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

Intervalos de probabilidad 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	$1.645\frac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}}$	$\mathbf{1.96}\frac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}}$	$2.58\frac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}}$

- Los intervalos de confianza se centran en $\hat{p}$, la estimación de p.
- Los intervalos para test de hipótesis se centran en el valor que da la hipótesis. Puede ser
  - p_0 (si suponemos verdadera la hipótesis nula H_0)
  - p_1 (si suponemos verdadera la hipótesis alternativa)
- Ambos intervalos tienen la misma amplitud ε.

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 de probabilidad $1-\alpha$ asumiendo que H_0. Llamamos a α error de tipo 1.
- 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 $\hat{p}$, 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

$$\varepsilon = 1.96 \frac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}} = 1.96 \frac{\sqrt{0.55 \times 0.45}}{\sqrt{1000}} = 0.031.$$

- El intervalo es $(0.5 - 0.031, 0.5 + 0.031) = (0.47, 0.53)$
- ¿Cuál es la conclusión del test?

- **Rechazamos** la hipótesis nula, porque $0.55 \notin (0.47, 0.53)$.
- 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 $n = 100$, la amplitud es

$$\varepsilon = 1.96 \frac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}} = 1.96 \frac{\sqrt{0.55 \times 0.45}}{\sqrt{100}} = 0.0975.$$

- El intervalo es $(0.5 - 0.0975, 0.5 + 0.0975) = (0.4, 0.6)$.
- Como $\hat{p} \in I$ no hay evidencia suficiente y **no rechazamos** la hipótesis nula.

Variantes de la hipótesis alternativa (nuevo)

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)

Hipótesis alternativa simple

- Tenemos dos puntos distinguidos distintos: p_0 y p_1.
- Sunpongamos $p_0 < p_1$ (el otro caso es análogo).
- Hipótesis nula simple: $H_0 : p = p_0$
- Hipótesis alternativa simple $H_1 : p = p_1 \quad (> p_0)$
- El intervalo es de la forma $[0, p_0 + \varepsilon]$
- La región crítica es $S = [p_0 + \varepsilon, 1]$
- ε es tal que
$$\mathbf{P}_{H_0}(\hat{p} < p_0 + \varepsilon) = 1 - \alpha.$$

Aplicamos el TCL

$$\begin{aligned}
\mathbf{P}_{H_0}(\hat{p} < p_0 + \varepsilon) &= \mathbf{P}_{H_0}\left(\frac{\hat{p} - p_0}{\sqrt{np_0q_0}} < \frac{\varepsilon}{\sqrt{np_0q_0}}\right) \\
&= \mathbf{P}_{H_0}\left(\frac{\mu - np_0}{\sqrt{np_0q_0}} < \frac{\sqrt{n}\varepsilon}{\sqrt{p_0q_0}}\right) \\
&\sim \int_{-\infty}^{a} \varphi(x)dx = 1 - \alpha
\end{aligned}$$

- Pusimos $a = \frac{\sqrt{n}\varepsilon}{\sqrt{p_0q_0}}$
- El intervalo es unilateral, de amplitud

$$\varepsilon = \frac{a\sqrt{p_0q_0}}{\sqrt{n}}.$$

Intervalos de confianza unilaterales:

Calculando `qnorm(1 − α)` obtenemos

```
> qnorm(0.9)
[1] 1.281552
> qnorm(0.95)
[1] 1.644854
> qnorm(0.99)
[1] 2.326348
```

confianza	0.90	**0.95**	0.99
error α	0.10	**0.05**	0.01
a (percentil)	1.28	**1.645**	2.33
amplitud ε	$1.28\frac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}}$	$\mathbf{1.645}\frac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}}$	$2.33\frac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}}$

El mismo razonamiento se aplica a un test con hipótesis unilateral de la forma

$$H_0 : p = p_0, \qquad H_1 : p > p_1, \qquad (p_0 < p_1).$$

Ejemplo

- Para ser aprobada por las autoridades regulatorias, una vacuna tiene que dar resultado positivo por lo menos en el 90% de los casos, en una determinada etapa de su producción.
- En la etapa anterior se confirmó su efectividad en un 80% de los casos.
- Para determinar entonces su efectividad se aplica la vacuna 100 veces.
- Se obtiene efectividad en 89 casos.
- ¿Puede afirmarse con un 99% de certeza que la vacuna supera la etapa?

- Formulamos un test de hipótesis con alternativa unilateral
- Tenemos $H_0: p = 0.8$ contra $H_1: p \geq 0.9$.
- De nuestra tabla (con $\alpha = 0.01$) tenemos amplitud

$$\varepsilon = 2.326 \frac{\sqrt{\hat{p}_0(1-\hat{p}_0)}}{\sqrt{n}} = 2.326 \frac{\sqrt{0.8 \times 0.2}}{\sqrt{100}} = 0.09304$$

- El intervalo es $(0.8 + 0.09304,\ 1] = (0.893,\ 1]$
- ¿Cuál es la conclusión del test?

No se rechaza la alternativa, no hay evidencia suficiente para eso, y la vacuna no pasa la etapa, nos quedamos en H_0.

- Y nos vamos a dormir tranquilos, ¡es un 99%!
- Pero, ¿que pasa si fuera cierta H_1? ¿Cuál es la probabilidad de equivocarme en ese caso?

Potencia de un test

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

- Suponemos que la verdadera hipótesis 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)

Vamos a calcular la potencia del ejemplo:

$$\pi = P_{H_1}(\hat{p} \in S) = P_{H_1}(\hat{p} \geq 0.893)$$

$$\begin{aligned}
&= P_{H_1}(\hat{p} - p_1 \geq 0.893 - p_1) \\
&= P_{H_1}\left(\frac{n\hat{p} - np_1}{\sqrt{p_1(1-p_1)}} \geq \frac{n0.893 - np_1}{\sqrt{p_1(1-p_1)}}\right) \\
&= P_{H_1}\left(\frac{n\hat{p} - np_1}{\sqrt{p_1(1-p_1)}} \geq \frac{89.3 - 90}{\sqrt{0.9 \times 0.1}}\right) \\
&= P_{H_1}\left(\frac{n\hat{p} - np_1}{\sqrt{p_1(1-p_1)}} \geq -2.33\right) \\
&= \int_{-2.33}^{\infty} \varphi(x)dx = 1 - \texttt{pnorm}(-2.33) = 0.9900969
\end{aligned}$$

Y . . . ¿dormimos tranquilos?

Probabilidades de que baje de peso antes de que acabe el año.