Test de Hipótesis

Sebastián Decuadro

Facultad de Ingeniería
Universidad Católica del Uruguay

22 de junio de 2023

Test de Hipótesis

Un test de hipótesis es un procedimiento estadístico mediante el cual se investiga la verdad o falsedad de una hipótesis acerca de una característica de una población o de un conjunto de poblaciones.

Un test de hipótesis está sujeto a la incertidumbre (azar), por lo que podemos equivocarnos al obtener una conclusión acerca de la veracidad de una hipótesis.

Test de hipótesis

Los tests que veremos harán contraste entre dos hipótesis incompatibles:

- H_0 llamada hipótesis nula
- H_1 llamada hipótesis alternativa

Por ejemplo, tenemos un dado y nos gustaría saber si es equilibrado.

$$\begin{cases} H_0: & \text{el dado está equilibrado} \\ H_1: & \text{el dado no está equilibrado} \end{cases}$$

Test de hipótesis

Para realizar un test siempre se tendrá una muestra $X_1, X_2, \ldots, X_n$ iid, la idea es realizar algún procedimiento a partir de esta muestra para tomar la decisión.

En el ejemplo del dado, supongamos que tiramos el dado 10 veces y obtenemos 5 veces el número 1. ¿Deberíamos rechazar H_0?

Tipos de test de hipótesis

Un test de hipótesis es *paramétrico* si sus hipótesis dan información sobre los parámetros de la distribución de la muestra.

Por ejemplo: $X_1, X_2, \ldots, X_n \sim N(\mu, \sigma^2)$ iid con parámetros desconocidos. El siguiente test es paramétrico:

$$\begin{cases} H_0: & \mu = 5 \\ H_1: & \mu = 6 \end{cases}$$

Tipos de test de hipótesis

Un test de hipótesis es *no paramétrico* cuando sus hipótesis dan información que no tiene que ver con los parámetros de la distribución.

Por ejemplo: $X_1, X_2, \ldots, X_n$ iid con distribución desconocida. El siguiente test es no paramétrico:

$$\begin{cases} H_0 : \text{la muestra viene de una} \\ \text{distribución exponencial} \\ H_1 : \text{no } H_0 \end{cases}$$

Tipos de test de hipótesis

Otro ejemplo de test no paramétrico es el test de independencia.

$X_1, X_2, \ldots, X_n$ iid con distribución de una variable aleatoria X.

$Y_1, Y_2, \ldots, Y_n$ iid con distribución de una variable aleatoria Y.
El siguiente test no es paramétrico

$$\begin{cases} H_0 : X \text{ e } Y \text{ son independientes} \\ H_1 : \text{no } H_0 \end{cases}$$

Tipos de error

Al decidir en un test de hipótesis hay cuatro posibilidades:

- Rechazo H_0 pero H_0 es verdadera.

  (Error de tipo I)
- Rechazo H_0 y H_1 es verdadera.

  (Test exitoso)

Tipos de error

- No rechazo H_0 pero H_1 es verdadera.

  (Error de tipo II)

- No rechazo H_0 y H_0 es verdadera.

  (Test exitoso)

Observación: la suma de las probabilidades de los cuatro casos es 1.

Notación: $P(\text{Error tipo I}) = \alpha$ y $P(\text{Error tipo II}) = \beta$

Tipos de error

Como α y β son probabilidades de error nos gustaría que ambos sean chicos.

No es posible diseñar un test de hipótesis para el cual α y β sean chicos al mismo tiempo. En general si uno quiere que ambos sean chicos se necesita de una muestra muy grande.

Diseñaremos tests de hipótesis para los cuales α es chico.

Ejemplo

En un preparado alimenticio infantil se especifica que el contenido medio de proteínas es al menos $42\,\%$. Tratamos de comprobar esta especificación y para ello tomamos 10 preparados que analizamos para determinar su contenido en proteínas, se obtiene un promedio muestral de $40\,\%$ y una desviación estándar muestral de $3,5\,\%$.

Diseñar un test de hipótesis que decida si la especificación es correcta.

Ejemplo

Asumiremos que tenemos $X_1, X_2, \ldots, X_{10} \sim N(\mu, \sigma^2)$ (alguien hizo un test de hipótesis que determinó que la hipótesis de normalidad de la muestra es razonable). Plantearemos el siguiente test de hipótesis:

$$\begin{cases} H_0: & \mu = 0{,}42 \\ H_1: & \mu < 0{,}42 \end{cases}$$

Intentaremos construir un criterio de decisión cuya probabilidad de error tipo I sea igual a α prefijado.

Ejemplo

Nuestro criterio de decisión tendrá la siguiente forma: rechazaremos H_0 si

$$\overline{X}_n < 0,42 - \delta$$

y trataremos de hallar δ tal que $P_{H_0}(\text{Rechazar } H_0)$

$$= P(\overline{X}_n < 0,42 - \delta \mid \mu = 0,42) = \alpha$$

Ejemplo

Para poder calcular esa probabilidad necesito conocer la distribución del promedio. Recordemos que $\overline{X}_n \sim N(\mu, \sigma^2/n)$.

Ejemplo

Tomando $\mu = 0,42$ para el cálculo de la probabilidad de error tipo I se tiene entonces

$$\alpha = P(\overline{X}_n < 0,42 - \delta) =$$

$$P\left(\sqrt{10}\frac{(\overline{X}_{10} - 0,42)}{S_n} < \sqrt{n}\frac{(0,42 - \delta - 0,42)}{S_n}\right)$$

$$= P\left(T < \sqrt{n}\frac{-\delta}{S_n}\right)$$

Donde T tiene distribución T de Student con 9 grados de libertad.

Ejemplo

Ahora despejamos δ usando la función cuantil de la T de Student.

$$\sqrt{n}\frac{-\delta}{S_n} = t.ppf(\alpha, n-1)$$

Despejando obtenemos

$$\delta = -t.ppf(\alpha, n-1)\frac{S_n}{\sqrt{n}}$$

Elijamos el nivel de significación $\alpha = 0,05$.

Ejemplo

$t.ppf(0.05, 9) \approx -1,833113$

Usando los datos del problema, llegamos a que $\delta \approx 0,0203$, por lo tanto, nuestro criterio de decisión para rechazar H_0 queda de la siguiente forma:

$$\overline{X}_n < 0,42 - \delta \approx 0,3997$$

Como $0,4$ no es menor a $0,3997$ no rechazamos H_0.

Sobre la conclusión de un test

Cuando rechazamos la hipótesis nula, tenemos prueba estadística de que la hipótesis nula es falsa. En cambio, si no podemos rechazar la hipótesis nula, no tenemos prueba estadística de que la hipótesis nula sea verdadera.

Esto se debe a que en general no tenemos control sobre β, no establecimos la probabilidad β de aceptar equivocadamente la hipótesis nula para que sea pequeña. De hecho si n es fijo, achicar α en general aumenta β.

Terminología

- Al criterio para decidir el rechazo de H_0 se le llama región crítica del test.
- α se llama nivel de significación del test.
- $1-\beta$ se llama potencia del test.
- En el ejemplo, $\sqrt{n}\frac{(\overline{X}_n-\mu)}{S_n}$ se llama estadístico del test. Otros test usarán estadísticos diferentes, pero es importante que el estadístico tenga distribución conocida cuando H_0 es cierta para poder construir el test.

Propiedad

Por la simetría de la distribución T de student se tiene que

$$-qt(\alpha, k) = qt(1-\alpha, k)$$

En la otra notación se escribe de la siguiente forma

$$qt(1-\alpha, k) = t_{\alpha,k}$$

Regiones críticas

En general, si tenemos una muestra $X_1, \ldots, X_n$ iid con distribución normal y el test de hipótesis

$$\begin{cases} H_0: & \mu = \mu_0 \\ H_1: & \mu < \mu_0 \end{cases}$$

entonces usaremos la región crítica

$$RC = \{\overline{X}_n < \mu_0 - \delta\} = \left\{\overline{X}_n < \mu_0 - t_{\alpha,n-1}\frac{S_n}{\sqrt{n}}\right\}$$

Regiones críticas

La misma región crítica funciona para los siguientes tests:

$$\begin{cases} H_0: & \mu \geq \mu_0 \\ H_1: & \mu < \mu_0 \end{cases} \qquad \begin{cases} H_0: & \mu = \mu_0 \\ H_1: & \mu = \mu_1 \end{cases}$$

donde $\mu_1 < \mu_0$.

Regiones críticas

En general, si tenemos una muestra $X_1, \ldots, X_n$ iid con distribución normal y el test de hipótesis

$$\begin{cases} H_0: & \mu = \mu_0 \\ H_1: & \mu > \mu_0 \end{cases}$$

entonces usaremos la región crítica

$$RC = \{\overline{X}_n > \mu_0 + \delta\} = \left\{\overline{X}_n > \mu_0 + t_{\alpha,n-1}\frac{S_n}{\sqrt{n}}\right\}$$

Regiones críticas

La misma región crítica funciona para los siguientes tests:

$$\begin{cases} H_0: & \mu \leq \mu_0 \\ H_1: & \mu > \mu_0 \end{cases} \qquad \begin{cases} H_0: & \mu = \mu_0 \\ H_1: & \mu = \mu_1 \end{cases}$$

donde $\mu_0 < \mu_1$.

Regiones críticas

Todos estos son test unilaterales para la media. Veamos un test bilateral:

$$\begin{cases} H_0: & \mu = \mu_0 \\ H_1: & \mu \neq \mu_0 \end{cases}$$

La región crítica para este test es:

$$RC = \{|\overline{X}_n - \mu_0| > \delta\} = \left\{|\overline{X}_n - \mu_0| > t_{\frac{\alpha}{2}, n-1} \frac{S_n}{\sqrt{n}}\right\}$$

Generalización

Estas regiones críticas también funcionan de forma aproximada cuando n es grande para muestras que no tengan distribución normal. Se justifican por el teorema central del límite, cuando H_0 es cierta

$$\sqrt{n}\frac{(\overline{X}_n - \mu_0)}{S_n} \approx N(0,1)$$

Por lo tanto usamos z_α en lugar de $t_{\alpha,n-1}$ y $z_{\frac{\alpha}{2}}$ en lugar de $t_{\frac{\alpha}{2},n-1}$ según corresponda.

p-valor

Veamos un ejemplo del práctico donde tenemos que realizar el test

$$\begin{cases} H_0: & \mu \geq 6 \\ H_1: & \mu < 6 \end{cases}$$

para una muestra con distribución normal de parámetros desconocidos, $n = 15$, $\alpha = 0,05$, $\overline{x}_n = 5$ y $s_n^2 = 2,143$.

p-valor

Por lo visto antes sabemos que la región crítica que tenemos que usar es

$$RC = \{\overline{X}_n < \mu_0 - \delta\} = \left\{\overline{X}_n < \mu_0 - t_{\alpha,n-1}\frac{S_n}{\sqrt{n}}\right\}$$

$$RC = \left\{\overline{X}_n < 6 - t_{\alpha,n-1}\frac{\sqrt{2,143}}{\sqrt{15}}\right\}$$

$$RC = \{\overline{X}_n < 5,334\}$$

Como se pertenece a la región crítica, rechazamos H_0.

p-valor

Observemos que si cambiamos el valor de $\alpha = 0,05$ a $\alpha = 0,01$ la nueva región crítica queda

$$RC = \{\overline{X}_n < 5,008\}$$

pero como sigo en la región crítica, seguimos rechazando H_0.

Observemos que si usamos $\alpha = 0,005$ entonces la nueva región crítica queda

$$RC = \{\overline{X}_n < 4,8748\}$$

y ahora no rechazamos H_0.

p-valor

Resumiendo, en este test:

1. Si $\alpha = 0,05$ rechazamos H_0.
2. Si $\alpha = 0,01$ rechazamos H_0.
3. Si $\alpha = 0,005$ no rechazamos H_0.

Tenemos un cambio de estado, pasamos de rechazar H_0 a no rechazar H_0. ¿Existe un valor que separa las opciones de rechazo y no rechazo?

p-valor

Existe un valor de $\alpha = \alpha^*$, llamado p-valor, que separa las decisiones que tomamos. En este ejemplo, existe un valor de α^* que me da una región crítica de la forma

$$RC = \{\overline{X}_n < 5\}$$

p-valor

Si tomamos valores de α más grandes que el p-valor, la región crítica queda con extremo más grande que 5 y por lo tanto rechazamos H_0. Por otro lado si tomamos valores de α más chicos que el p-valor, la región crítica queda con extremo más chico que 5 y por lo tanto no rechazamos H_0.

p-valor

Para calcular el p-valor α^* tratemos de hallar α tal que

$$RC = \left\{ \overline{X}_n < \mu_0 - t_{\alpha,n-1} \frac{S_n}{\sqrt{n}} \right\} = \{ \overline{X}_n < 5 \}$$

Por lo tanto necesitamos hallar α tal que

$$\mu_0 - t_{\alpha,n-1} \frac{S_n}{\sqrt{n}} = 5$$

p-valor

Despejando obtenemos que

$$t_{\alpha,n-1} = (\mu_0 - 5)\frac{\sqrt{n}}{S_n} = (6-5)\frac{\sqrt{15}}{\sqrt{2,143}}$$

$$qt(1-\alpha, n-1) = (6-5)\frac{\sqrt{15}}{\sqrt{2,143}}$$

$$1-\alpha = \mathrm{t.cdf}\left(\frac{\sqrt{15}}{\sqrt{2,143}},\ n-1\right)$$

p-valor

Por lo tanto el p-valor en este caso es

$$1 - \mathrm{t.cdf}\left(\frac{\sqrt{15}}{\sqrt{2,143}}, 14\right) \approx 0,009595$$

Regla para tomar decisiones con el p-valor

Si $\alpha \leq \alpha^*$ entonces no rechazamos H_0.

Si $\alpha > \alpha^*$ entonces rechazamos H_0.

Este criterio de decisión sirve para todos los test de hipótesis.

Esto es importante recordarlo porque muchos software de estadística, incluyen el p-valor de un test cuando lo ejecutan.

Definición de p-valor

El p-valor de un test es un número entre $(0, 1)$ que depende del test y depende de la muestra.

Es la probabilidad bajo H_0 de que el estadístico del test tome un valor tanto o más extremo que el valor que se observó en la muestra. Donde extremo significa contradictorio con H_0.

p-valor

En nuestro ejemplo el p-valor es

$$P_{H_0}\left(\overline{X}_n < 5\right)$$

En general, los software de estadística estiman el p-valor generando nuevas muestras (cuando H_0 es cierta conocemos la distribución de los datos) y estimar la probabilidad de que el estadístico sea más extremo que el estadístico observado.