

Introducción a pruebas de hipótesis

ESTA 3042

enero 2013

Testing de Hipótesis

- Hemos visto como estimar un parámetro de una población.
- Ahora pasamos a como determinar si los datos proveen evidencia “a favor” de algún enunciado sobre la población.
- A dicho enunciado le llamamos hipótesis.
- Recuerde que desde la perspectiva matemática solo interesan enunciados con dos posibles valores de veracidad: cierto o falso.
- las hipótesis siempre se refieren a una población o a un modelo.

Testing de Hipótesis

Definition (Hipótesis nula (Null Hypothesis))

El enunciado a examinar en la prueba de hipótesis (test of significance) se llama **hipótesis nula** (denotada H_0). La prueba de hipótesis está concebida para determinar la fuerza de la evidencia **en contra** de la hipótesis nula.

- Usualmente H_0 se expresa como un enunciado de “no efecto”, “no cambio” o “status quo”.
- H_0 es un enunciado que se expresa en términos de algún parámetro(s) de la población.

Testing de Hipótesis

Definition (Hipótesis alterna (Alternate Hypothesis))

El enunciado que pasa a ser cierto en caso de que H_0 no lo sea es la **hipótesis alterna** (denotada H_a).

- H_a expresa el enunciado que se espera que la evidencia obtenida en las observaciones apoyen.
- H_a puede ser one-sided o two-sided. Por ejemplo:

$$H_0 : \mu = 0$$

$$H_a : \mu > 0 \quad (\text{one-sided})$$

$$H_a : \mu \neq 0 \quad (\text{two-sided}).$$

Ejemplo 1

Suponga que usted es el manager de producción de una compañía de cereal y que el proceso de llenar las cajas de cereal está bajo **control**, esto es, que las cajas se llenan con 368 g de cereal (esta es la especificación de la compañía). Usted quiere determinar si la cantidad media de cereal por caja es 368 g. Esto es:

- $H_0 : \mu = 368$
- Note que H_0 habla de la media de la pop. (μ) aunque lo que podremos medir será \bar{X} .
- H_0 se considera cierta y lo q. se considera es la evidencia en su contra. Esto es H_a :
- $H_a : \mu \neq 368$

Note que:

- 1 H_0 se refiere a un valor específico para un parámetro (por eso el signo de igualdad).
- 2 H_a no se refiere a **un** valor en particular.

Testing de Hipótesis

Observe que:

- Todo test está basado en la estadística que estima el parámetro que aparece en la hipótesis.
- Valores del estimado distantes del valor propuesto en H_0 proveen evidencia en contra de H_0 .
- H_a determina en que dirección se considera la evidencia en contra de H_0 .

Definition (Test statistic (estadística de prueba))

Un **test statistic** mide la compatibilidad entre H_0 y los datos.

Por ejemplo (z -statistic):

$$z = \frac{(\bar{X} - \mu)}{\sigma/\sqrt{n}}$$

Este es el z -test statistic para la media de la población (μ). Observe que este valor nos dice que tan lejos está \bar{X} de μ en términos del standard error .

Ejemplo 1b

Considere el ejemplo de las cajas de cereales postulado anteriormente:

$$H_0 : \mu = 368$$

$$H_a : \mu \neq 368$$

Y suponga que tomó un SRS con $n = 25$, que $\bar{X} = 372.5$, y que $\sigma = 15$ (especificación de la máquina que vierte el cereal). Luego

$$\begin{aligned} z &= \frac{(\bar{X} - \mu)}{\sigma/\sqrt{n}} \\ &= \frac{372.5 - 368}{15/\sqrt{25}} = 1.50 \end{aligned}$$

Testing de Hipótesis

- No todos los test statistics están asociados a z -scores.
- Los test statistics estarán asociados a probabilidades...

Definition (P -value)

La probabilidad de que el test statistic tomaría un valor tan extremo o más extremo de lo que es observado (en los datos) dado que H_0 es cierto se llama el **P -value** del test statistic. Matemáticamente:

$$P = Pr(\text{test-stat es tan extremo o más extremo que lo obs} | H_0 = \text{true})$$

- Mientras más pequeño el P -value más fuerte es la evidencia, provista por los datos, en contra de H_0 .
- El P -value es también llamado el valor de significancia observado, o el nivel de significancia más pequeño para el cual H_0 puede ser rechazado.
- Mire ejemplo 6.14.

Testing de Hipótesis

- El P -value determina la fuerza de la evidencia en contra de H_0 .
- En general, podemos cuantificar cuanta evidencia tenemos que acumular para rechazar H_0 .

Definition (Estadísticamente significativo (statistical significance))

Si el P -value es menor o igual que un valor α , entonces decimos que los datos son **estadísticamente significativos a un nivel α** .

- Si el P -value es mayor que α entonces H_0 no es rechazada.
- Si el P -value es menor que α entonces H_0 es rechazada.

Ejemplo 1c

Volvamos al ejemplo de las cajas de cereal.

- Sabemos que $z = 1.50$
- Usamos el z -statistic anterior para calcular el P -value. Esto es, la probabilidad de que Z sea tan o más extremo que 1.5 dado que H_0 es cierta ($Pr(Z \geq 1.5 | H_0 = \text{true})$).
- Como H_a era de dos lados ($\mu \neq 360$), el P -value incluye el otro extremo ($Pr(Z \leq -1.5 | H_0 = \text{true})$).
- El P -value es (recuerde que son prob. condicionales) :

$$\begin{aligned}P &= Pr(Z \geq 1.5) + Pr(Z \leq -1.5) \\&= (1 - Pr(Z < 1.5)) + Pr(Z \leq -1.5) \\&= 1 - 0.9332 + 0.0668 \\&= 0.0668 + 0.0668 = 2(0.0668) = 0.1336\end{aligned}$$

- Si queremos un nivel de significancia de un 5% ($\alpha = 0.05$). Entonces vemos que $P = 0.1336 > \alpha = 0.05$.
- Por lo tanto H_0 **no** puede ser rechazada.

Note que:

- no se supuso nada sobre la dist. de X , pero $Z \sim N(0, 1)$ por el TCL.
- El P -value nos da la probabilidad de que el test statistic z -test tenga el valor que se midió o más extremo dado que H_0 es **cierta**. (Mientras más peq. peor para H_0)
- α cuantifica cuan significativa es la evidencia (datos) en contra H_0 .
- Como H_a es two-sided α da el nivel de significancia considerando las dos colas.

Algorithm (z -test para μ)

Considere la hipótesis

$$H_0 : \mu = \mu_0$$

Tome un SRS de tamaño n de una población con media μ y desviación estandar σ . Compute la estadística

$$z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$$

$Z \sim N(0, 1)$, y el P -value para la prueba **en contra** H_0 es uno de los siguientes:

- $H_a : \mu > \mu_0$ $P = Pr(Z \geq z)$
- $H_a : \mu < \mu_0$ $P = Pr(Z \leq z)$
- $H_a : \mu \neq \mu_0$ $P = 2Pr(Z \geq |z|)$

Los P -values son exactos si la **población** se distribuye normalmente, y es aprox. correcta cuando n es grande.

Ejemplo 2

Suponga que administra una fábrica de aderezos de ensalada. Y que las embotelladoras funcionan bien cuando embotellan 8 oz de aderezo por botella. Sabemos que $\sigma = 0.15$. Se selecciona un SRS con $n = 50$ periódicamente y se detiene el proceso si hay evidencia de que $\mu \neq 8$. Suponga que para una muestra en particular $\bar{X} = 7.983$. Use un nivel de significancia de $\alpha = 0.05$ y determine si hay que parar el proceso.

$$H_0 : \mu = 8$$

$$H_a : \mu \neq 8$$

$$z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} = \frac{7.983 - 8}{0.15/\sqrt{50}} = \frac{-0.017}{0.0212} = -0.80139$$

$$\begin{aligned} P &= Pr(Z < -0.80139) + Pr(Z > 0.80139) = 2Pr(Z < -0.80139) \\ &= 2(0.2119) = 0.4238 \end{aligned}$$

$0.4238 = P > \alpha = 0.05$. Por tanto H_0 no debe ser rechazada, y la producción no debe detenerse.

Ejemplo 3

Suponga que usted es el encargado del control de calidad en una fábrica de bombillas. Se entiende que el proceso está en control cuando la media de la población es 375 hrs. Se conoce que $\sigma = 100$ hrs. Se toma un SRS de $n = 50$ y se calcula $\bar{X} = 350$ hrs. $\bar{X} = 350$ hrs, $\sigma = 100$ hrs, $n = 50$.

- 1 Enuncie la hipótesis nula y la alterna.

$$H_0 : \mu = 375 \quad \text{hrs}$$

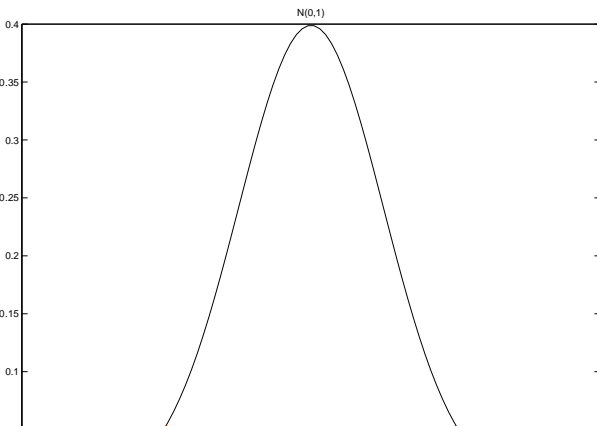
$$H_a : \mu \neq 375 \quad \text{hrs}$$

- 2 Use un nivel de significancia (α) de 0.05, ¿Qué debe concluir el encargado del control de calidad acerca del proceso basandose en los resultados de la muestra?

$$\begin{aligned} z &= \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} = \frac{350 - 375}{100/\sqrt{50}} \\ &= \frac{-25}{100/\sqrt{50}} = -1.76777 \end{aligned}$$

Intervalos de confianza (IC) y los two-sided tests de hipótesis

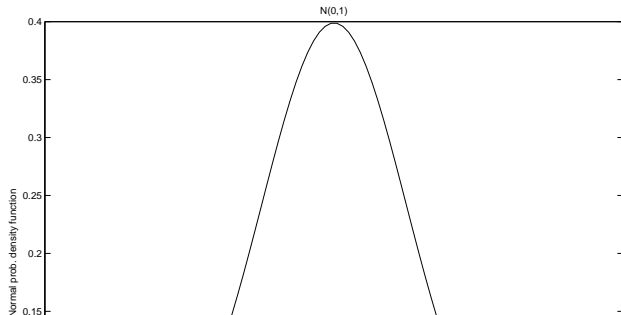
- Un two-sided test de significancia con nivel α rechaza la hipótesis nula ($H_0 : \mu = \mu_0$) exactamente cuando el valor de μ_0 está fuera del IC de nivel $1 - \alpha$ para μ .



- El área bajo curva y sobre $[-z, z]$ es igual $1 - \alpha$
- El área bajo curva y sobre $[-\infty, -z] \cup [z, \infty]$ es α

P-value versus α fijos

- Suponga que para un conjunto de datos se obtiene un $z = -5$.
- Por tanto los datos son significativos a un nivel $\alpha = 0.01$. ¿Por qué?
 - ▶ Note que -5 está a 5σ de μ en la $N(0, 1)$.
 - ▶ Si recuerda la regla 68-95-99.7 o ve la tabla, note que la prop. de obs. para $|z| > 5\sigma$ es *mucho* menor que el 0.03%
- Como vimos el z está más a la izq. del z requerido para un nivel de confianza de 1%. Gráficamente...



En la practica tenemos el siguiente procedimiento:

Algorithm

- 1 *Enuncie H_0 .*
- 2 *Enuncie H_a .*
- 3 *Escoja n .*
- 4 *Determine el test statistic apropiado.*
- 5 *Recolecte los datos, y calcule el test-statistic.*
- 6 *Calcule el P-value basado en el test statistic.*
- 7 *Compare el P-value con α .*
- 8 *Rechaze o no rechaze H_0*