



Métodos Estadísticos y Simulación

Unidad 6: Inferencia Estadística: Pruebas de hipótesis paramétricas

Mg. J. Eduardo Gamboa U.

2025-06-16

Conceptos básicos

Hipótesis estadística

Una hipótesis estadística es una afirmación sobre la distribución de probabilidad de una población o sobre el valor o valores de uno o más parámetros, como la media (μ), la variancia (σ^2) o la proporción (π).

Esta afirmación debe estar basada en la comprensión del fenómeno y sus variables. Una buena hipótesis permite hacer predicciones específicas y, si es rechazada, ayuda a revelar la complejidad del fenómeno.

Tipos de hipótesis estadísticas

Hipótesis nula (H_0 o H_p): Es la hipótesis que es aceptada provisionalmente como verdadera y cuya validez será sometida a verificación experimental. Los resultados experimentales nos permitirán seguir aceptándola como verdadera o si debemos rechazarla como tal.

Hipótesis alterna (H_1 o H_a): Es la hipótesis que se acepta en caso de que la hipótesis nula sea rechazada. La H_1 es la suposición contraria a H_0 .

Prueba de hipótesis

Una prueba de hipótesis es un proceso estructurado para tomar decisiones basadas en datos. Se fundamenta en el método hipotético-deductivo, donde las hipótesis se contrastan con la evidencia en lugar de verificarse directamente.

Una prueba de hipótesis estadística es el proceso mediante el cual se toma la decisión de aceptar o rechazar la hipótesis nula.

El proceso de prueba de hipótesis determina si se rechaza o no H_0 , pero **no se prueba su veracidad absoluta**.

Plantear hipótesis **antes del análisis de datos** (hipótesis a priori) mejora la solidez del estudio, enfocando la prueba en relaciones específicas y reduciendo sesgos.

Tipos de pruebas de hipótesis

En principio, se pueden formular hasta tres tipos de prueba, la cual dependerá de la forma de la hipótesis alterna que se plantee en el estudio:

Hipótesis unilateral con cola a la derecha	Hipótesis bilateral o de dos colas	Hipótesis unilateral con cola a la izquierda
$H_0 : \theta \leq \theta_0$ $H_1 : \theta > \theta_0$	$H_0 : \theta = \theta_0$ $H_1 : \theta \neq \theta_0$	$H_0 : \theta \geq \theta_0$ $H_1 : \theta < \theta_0$

donde θ es el parámetro de interés a probarse, pudiendo ser μ , σ^2 , π (o algún otro que, por cuestiones de tiempo y/o complejidad no es abordado en el curso), y θ_0 es el valor o los valores supuestos que puede tomar el parámetro.

Ejemplo

Enunciado	Formulación de la hipótesis	Tipo de prueba de hipótesis
El tiempo promedio de respuesta de una IA educativa no debe superar los 2.5 segundos.	$H_0 : \mu \leq 2.5$ $H_1 : \mu > 2.5$	Unilateral derecha
El porcentaje de estudiantes que aprueban un curso con clases híbridas no es 70%.	$H_0 : \pi = 0.70$ $H_1 : \pi \neq 0.70$	Bilateral
La varianza de los niveles de CO ₂ en aulas ventiladas debe ser menor que 400 ppm ² .	$H_0 : \sigma^2 \geq 400$ $H_1 : \sigma^2 < 400$	Unilateral izquierda
El peso promedio de una variedad mejorada de palta Hass es mayor a 280 gramos.	$H_0 : \mu \leq 280$ $H_1 : \mu > 280$	Unilateral derecha
La desviación en el consumo de energía entre dos modalidades de iluminación LED no difiere significativamente.	$H_0 : \sigma_1^2 = \sigma_2^2$ $H_1 : \sigma_1^2 \neq \sigma_2^2$	Bilateral

Tipo de Errores

Al tomarse una decisión respecto a una hipótesis nula (H_0), se puede presentar cuatro posibles casos que determinan si la decisión tomada es correcta o incorrecta, esto se presenta en la siguiente tabla:

Decisión	H_0 verdadera	H_0 falsa
Aceptar H_0	Decisión correcta con probabilidad $1 - \alpha$	Error tipo II con probabilidad β
Rechazar H_0	Error tipo I con probabilidad α	Decisión correcta con probabilidad $1 - \beta$

La probabilidad de cometer error tipo I se denota por α , conocido como nivel de significación. Determina el tamaño de la zona de rechazo de H_0

La probabilidad de cometer error tipo II se denota por β . Su complemento ($1 - \beta$) es conocido como potencia de prueba.

Pruebas de hipótesis para un parámetro (una población)

Prueba de hipótesis para una media

- ▶ Supuestos: Población normal o $n \geq 30$, e independencia de observaciones
- ▶ Hipótesis:
 - ▶ Unilateral a la derecha: $H_0 : \mu \leq \mu_0$ versus $H_1 : \mu > \mu_0$
 - ▶ Unilateral a la izquierda: $H_0 : \mu \geq \mu_0$ versus $H_1 : \mu < \mu_0$
 - ▶ Bilateral: $H_0 : \mu = \mu_0$ versus $H_1 : \mu \neq \mu_0$
- ▶ Estadístico de prueba:

$$t_{calc} = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}} \sim t_{(n-1)}$$

- ▶ Decisión: Comparar con valores críticos o usar p-valor

Ejemplo

Una universidad afirma que sus estudiantes dedican en promedio 4 horas diarias al estudio. Se sospecha que el promedio es realmente menor, por lo que se toma una muestra aleatoria de 35 estudiantes, quienes reportan las siguientes horas de estudio:

2.8, 3.9, 4.2, 2.1, 3.7, 4.5, 3.8, 4.1, 4.6, 3.6, 4.3, 4.0, 2.9, 1.4,
3.9, 4.5, 4.2, 4.0, 3.8, 1.0, 4.7, 3.5, 2.1, 2.2, 3.6, 4.3, 4.8, 4.2,
3.9, 3.7, 4.4, 3.0, 3.8, 4.6, 2.5

Verificar la afirmación con un nivel de significancia del 10%.

$$H_0 : \mu \geq 4 \quad H_1 : \mu < 4 \quad \alpha = 0.10$$

```
tiempo <- c(2.8, 3.9, 4.2, 2.1, 3.7, 4.5, 3.8, 4.1, 4.6, 3.6, 4.3, 4.0,  
           2.9, 1.4, 3.9, 4.5, 4.2, 4.0, 3.8, 1.0, 4.7, 3.5, 2.1, 2.2,  
           3.6, 4.3, 4.8, 4.2, 3.9, 3.7, 4.4, 3.0, 3.8, 4.6, 2.5)  
(tcalc <- (mean(tiempo) - 4)/(sd(tiempo)/sqrt(length(tiempo))))
```

[1] -2.383265

```
(tcrit <- qt(p = 0.10, df = length(tiempo)-1))
```

[1] -1.306952

$t_{calc} < t_{crit}$ y la hipótesis es unilateral a la izquierda, entonces se rechaza la hipótesis nula. En conclusión...

```
t.test(x = tiempo, mu = 4, alternative = "less")
```

One Sample t-test

data: tiempo

t = -2.3833, df = 34, p-value = 0.01144

alternative hypothesis: true mean is less than 4

95 percent confidence interval:

-Inf 3.88878

sample estimates:

mean of x

3.617143

$p\text{-value} < \alpha$, entonces se rechaza la hipótesis nula. En conclusión...

Si solo tuviésemos datos resumidos:

```
library(BSDA)
tsum.test(
  mean.x = 3.617,
  s.x = 0.95,
  n.x = 35,
  mu = 4,
  alternative = "less")
```

One-sample t-Test

```
data: Summarized x
t = -2.3851, df = 34, p-value = 0.0114
alternative hypothesis: true mean is less than 4
95 percent confidence interval:
    NA 3.888527
sample estimates:
mean of x
    3.617
```

Prueba de hipótesis para una varianza

- ▶ Supuestos: Población normal e independencia de las observaciones
- ▶ Hipótesis:
 - ▶ Unilateral a la derecha: $H_0 : \sigma^2 \leq \sigma_0^2$ versus $H_1 : \sigma^2 > \sigma_0^2$
 - ▶ Unilateral a la izquierda: $H_0 : \sigma^2 \geq \sigma_0^2$ versus $H_1 : \sigma^2 < \sigma_0^2$
 - ▶ Bilateral: $H_0 : \sigma^2 = \sigma_0^2$ versus $H_1 : \sigma^2 \neq \sigma_0^2$
- ▶ Estadístico de prueba:

$$\chi_{calc}^2 = \frac{(n-1)s^2}{\sigma_0^2} \sim \chi_{(n-1)}^2$$

- ▶ Decisión: Comparar con valores críticos o usar p-valor

Ejemplo

Para los mismos datos de tiempo de estudio, se sospecha que la varianza es mayor a 1 hora. ¿Se puede verificar dicha afirmación con un nivel de significancia del 10%?

$$H_0 : \sigma^2 \leq 1 \quad H_1 : \sigma^2 > 1 \quad \alpha = 0.10$$

```
(chicalc = (length(tiempo)-1)*var(tiempo)/1)
```

```
[1] 30.70971
```

```
(chicrit = qchisq(p = 0.90, df = length(tiempo)-1))
```

```
[1] 44.90316
```

```
library(EnvStats)
```

```
varTest(x = tiempo, sigma.squared = 1, alternative = "greater")
```

```
$statistic
```

```
Chi-Squared
```

```
30.70971
```

```
$parameters
```

```
df
```

```
34
```

```
$p.value
```

```
[1] 0.6296898
```

```
$estimate
```

```
variance
```

```
0.9032269
```

```
$null.value
```

```
variance
```

```
1
```

```
$alternative
```

```
[1] "greater"
```

```
$method
```

```
[1] "Chi-Squared Test on Variance"
```

Prueba de hipótesis para una proporción

- ▶ Supuestos: $n\pi_0 \geq 5$, $n(1 - \pi_0) \geq 5$ e independencia de las observaciones
- ▶ Hipótesis:
 - ▶ Unilateral a la derecha: $H_0 : \pi \leq \pi_0$ versus $H_1 : \pi > \pi_0$
 - ▶ Unilateral a la izquierda: $H_0 : \pi \geq \pi_0$ versus $H_1 : \pi < \pi_0$
 - ▶ Bilateral: $H_0 : \pi = \pi_0$ versus $H_1 : \pi \neq \pi_0$
- ▶ Estadístico de prueba:

$$Z_{calc} = \frac{p - \pi_0}{\sqrt{\frac{\pi_0(1-\pi_0)}{n}}} \sim N(0, 1)$$

- ▶ Decisión: Comparar con valores críticos o usar p-valor

Ejemplo

Una universidad sostiene que el 80% de los estudiantes están satisfechos con el servicio de la biblioteca. Se encuesta a 100 estudiantes al azar y 71 dicen estar satisfechos. Verificar si la proporción real difiere de 0.80, con un nivel de significancia del 5%.

$$H_0 : \pi = 0.80 \quad H_1 : \pi \neq 0.80 \quad \alpha = 0.05$$

```
p <- 71/100  
(Z_calc <- (p - 0.8) / sqrt(0.8 * (1 - 0.8) / 100))
```

```
[1] -2.25
```

```
(Z_crit1 <- qnorm(0.025))
```

```
[1] -1.959964
```

```
(Z_crit2 <- qnorm(0.975))
```

```
[1] 1.959964
```

```
prop.test(x=71, n=100, p=0.80, alternative = "two.sided", correct = F)
```

1-sample proportions test without continuity correction

data: 71 out of 100, null probability 0.8

X-squared = 5.0625, df = 1, p-value = 0.02445

alternative hypothesis: true p is not equal to 0.8

95 percent confidence interval:

0.6146111 0.7898516

sample estimates:

p

0.71

Pruebas de hipótesis para dos parámetros (dos poblaciones)

Prueba de hipótesis de homogeneidad de dos varianzas

- ▶ Supuestos: Ambas poblaciones normales e independientes entre ellas, así como independencia de las observaciones dentro de cada muestra
- ▶ Hipótesis: $H_0 : \frac{\sigma_1^2}{\sigma_2^2} = 1$ versus $H_1 : \frac{\sigma_1^2}{\sigma_2^2} \neq 1$
- ▶ Estadístico de prueba:

$$F_{calc} = \frac{s_1^2}{s_2^2} \sim F_{(n_1-1, n_2-1)}$$

- ▶ Decisión: Comparar con valores críticos o usar p-valor

Ejemplo

Un laboratorio desea determinar si la variabilidad de concentración (en mg/L) de un fármaco en la sangre es igual para dos fabricantes distintos, considerando un nivel de significancia del 10%. Los datos de concentración con el fabricante A son: 8.1, 7.9, 8.3, 7.8, 8.0, 8.2, 7.7, mientras que con el fabricante B: 7.5, 7.2, 7.1, 7.4, 7.3, 7.6, 7.2, 7.5.

$$H_0 : \frac{\sigma_A^2}{\sigma_B^2} = 1 \quad H_1 : \frac{\sigma_A^2}{\sigma_B^2} \neq 1 \quad \alpha = 0.10$$

```
A <- c(8.1, 7.9, 8.3, 7.8, 8.0, 8.2, 7.7)
B <- c(7.5, 7.2, 7.1, 7.4, 7.3, 7.6, 7.2, 7.5)
(Fcalc <- var(A)/var(B))
```

```
[1] 1.484848
```

```
(Fcrit1 <- qf(0.05, 6, 7))
```

```
[1] 0.2377184
```

```
(Fcrit2 <- qf(0.95, 6, 7))
```

```
[1] 3.865969
```

```
var.test(A, B, alternative = "two.sided", ratio = 1)
```

F test to compare two variances

data: A and B

F = 1.4848, num df = 6, denom df = 7, p-value = 0.6136

alternative hypothesis: true ratio of variances is not equal to 1

95 percent confidence interval:

0.290089 8.456911

sample estimates:

ratio of variances

1.484848

Prueba de hipótesis para dos medias independientes

- ▶ Supuestos: Ambas poblaciones normales (o con tamaño de muestra grande $n > 30$ cada una) e independientes entre ellas, así como independencia de las observaciones dentro de cada muestra
- ▶ Hipótesis:
 - ▶ Unilateral a la derecha: $H_0 : \mu_1 - \mu_2 \leq \mu_0$ versus $H_1 : \mu_1 - \mu_2 > \mu_0$
 - ▶ Unilateral a la izquierda: $H_0 : \mu_1 - \mu_2 \geq \mu_0$ versus $H_1 : \mu_1 - \mu_2 < \mu_0$
 - ▶ Bilateral: $H_0 : \mu_1 - \mu_2 = \mu_0$ versus $H_1 : \mu_1 - \mu_2 \neq \mu_0$
- ▶ Estadístico de prueba:
 - ▶ Varianzas iguales:

$$t = \frac{\bar{X}_1 - \bar{X}_2 - \mu_0}{\sqrt{s_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}} \sim t_{(n_1+n_2-2)} \quad \text{donde} \quad s_p^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1 + n_2 - 2}$$

- ▶ Varianzas distintas:

$$t = \frac{\bar{X}_1 - \bar{X}_2 - \mu_0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \sim t_{\nu-Welch}$$

- ▶ Decisión: Comparar con valores críticos o usar p-valor

Ejemplo

Queremos comparar los salarios mensuales de trabajadores de dos empresas diferentes a través de sus medias. No conocemos nada acerca de sus varianzas. Considerar $\alpha = 0.10$. Los datos recolectados en una muestra aleatoria de cada empresa es:

- ▶ Empresa 1: 2500, 2700, 2600, 2800, 2900
- ▶ Empresa 2: 3000, 2100, 3200, 2300, 3400

Primero, debemos verificar si las varianzas son homogéneas:

$$H_0 : \frac{\sigma_A^2}{\sigma_B^2} = 1 \quad H_1 : \frac{\sigma_A^2}{\sigma_B^2} \neq 1 \quad \alpha = 0.10$$

```
E1 = c(2500, 2700, 2600, 2800, 2900)
E2 = c(3000, 2100, 3200, 2300, 3400)
var.test(E1, E2, alternative = "two.sided", ratio = 1)
```

F test to compare two variances

data: E1 and E2

F = 0.076923, num df = 4, denom df = 4, p-value = 0.02915

alternative hypothesis: true ratio of variances is not equal to 1

95 percent confidence interval:

0.008009041 0.738809991

sample estimates:

ratio of variances

0.07692308

Ahora procederemos a comparar las medias. Dado que no se especifica una dirección particular en la comparación, se asumirá que el objetivo es determinar si las medias son iguales o diferentes.

$$H_0 : \mu_1 - \mu_2 = 0 \quad H_1 : \mu_1 - \mu_2 \neq 0 \quad \alpha = 0.10$$

```
t.test(E1, E2, alternative = "two.sided", mu = 0, var.equal = F, paired = F)
```

Welch Two Sample t-test

data: E1 and E2

t = -0.37796, df = 4.6118, p-value = 0.7222

alternative hypothesis: true difference in means is not equal to 0

95 percent confidence interval:

-797.699 597.699

sample estimates:

mean of x mean of y

2700 2800

Ejemplo

Un equipo de agrónomos desea evaluar si el tipo de fertilizante influye en el rendimiento de maíz (kg por parcela). Se aplicaron dos tipos de fertilizante (A y B) en parcelas similares bajo las mismas condiciones de riego y clima. Los investigadores quieren si fertilizante A incrementa el rendimiento del maíz en más de 20 kg/parcela en comparación con el fertilizante B. El estudio se realiza en condiciones controladas y se usa un nivel de significancia del 10%.

► Fertilizante A: 820, 830, 815, 860, 825, 835, 822

► Fertilizante B: 800, 805, 798, 810, 802, 799, 803, 777, 789, 815

$$H_0 : \frac{\sigma_A^2}{\sigma_B^2} = 1 \quad H_1 : \frac{\sigma_A^2}{\sigma_B^2} \neq 1 \quad \alpha = 0.10$$

```
A <- c(820, 830, 815, 860, 825, 835, 822)
B <- c(800, 805, 798, 810, 802, 799, 803, 777, 789, 815)
var.test(A, B, alternative = "two.sided", ratio = 1)
```

F test to compare two variances

data: A and B

F = 1.9719, num df = 6, denom df = 9, p-value = 0.3457

alternative hypothesis: true ratio of variances is not equal to 1

95 percent confidence interval:

0.45648 10.89142

sample estimates:

ratio of variances

1.971867

Probando la diferencia de medias:

$$H_0 : \mu_1 - \mu_2 \leq 20 \quad H_1 : \mu_1 - \mu_2 > 20 \quad \alpha = 0.10$$

```
A <- c(820, 830, 815, 860, 825, 835, 822)
B <- c(800, 805, 798, 810, 802, 799, 803, 777, 789, 815)
t.test(A, B, alternative = "two.sided", mu = 20, var.equal = T, paired = F)
```

Two Sample t-test

data: A and B

t = 1.5824, df = 15, p-value = 0.1344

alternative hypothesis: true difference in means is not equal to 20

95 percent confidence interval:

16.60920 42.93365

sample estimates:

mean of x mean of y

829.5714 799.8000

Prueba de hipótesis para dos medias pareadas

- ▶ Supuestos: Las poblaciones no son independientes. Las diferencias son normales (o con tamaño de muestra grande $n > 30$ cada una) e independientes entre ellas.
- ▶ Hipótesis:
 - ▶ Unilateral a la derecha: $H_0 : \mu_D \leq \mu_0$ versus $H_1 : \mu_D > \mu_0$
 - ▶ Unilateral a la izquierda: $H_0 : \mu_D \geq \mu_0$ versus $H_1 : \mu_D < \mu_0$
 - ▶ Bilateral: $H_0 : \mu_D = \mu_0$ versus $H_1 : \mu_D \neq \mu_0$
- ▶ Estadístico de prueba:

$$t_{calc} = \frac{\bar{d} - \mu_0}{\frac{s_d}{\sqrt{n}}} \sim t_{(n-1)}$$

- ▶ Decisión: Comparar con valores críticos o usar p-valor

Ejemplo

Un grupo de agrónomos desea evaluar el efecto de un regulador de crecimiento sobre la altura de plantas de tomate. Se midieron las alturas de las mismas plantas antes y después de aplicar el producto, por lo que se trata de un diseño pareado. Se quiere saber si el producto genera un aumento promedio mayor a 3 cm, utilizando un nivel de significancia del 5%.

Los datos recolectados antes y después de la aplicación del producto son: - Antes: 120, 122, 121, 119, 118, 123, 121, 120, 122, 119, 115, 123 - Después: 125, 127, 126, 124, 123, 129, 126, 125, 128, 123, 116, 129

$$H_0 : \mu_D \leq 3 \quad H_1 : \mu_D > 3 \quad \alpha = 0.05$$

```
antes    = c(120, 122, 121, 119, 118, 123, 121, 120, 122, 119, 115, 123)
despues  = c(125, 127, 126, 124, 123, 129, 126, 125, 128, 123, 116, 129)
t.test(despues, antes, mu = 3, alternative = "greater", paired = T)
```

Paired t-test

data: despues and antes

t = 4.7497, df = 11, p-value = 3e-04

alternative hypothesis: true mean difference is greater than 3

95 percent confidence interval:

4.140136 Inf

sample estimates:

mean difference

4.833333

Prueba de hipótesis para dos proporciones

- ▶ Supuestos: $n\pi \geq 5$ en ambos grupos, e independencia de las observaciones entre grupos y dentro de cada grupo.
- ▶ Hipótesis:
 - ▶ Unilateral a la derecha: $H_0 : \pi_1 - \pi_2 \leq \pi_0$ versus $H_1 : \pi_1 - \pi_2 > \pi_0$
 - ▶ Unilateral a la izquierda: $H_0 : \pi_1 - \pi_2 \geq \pi_0$ versus $H_1 : \pi_1 - \pi_2 < \pi_0$
 - ▶ Bilateral: $H_0 : \pi_1 - \pi_2 = \pi_0$ versus $H_1 : \pi_1 - \pi_2 \neq \pi_0$
- ▶ Estadístico de prueba:

$$Z_{calc} = \frac{p_1 - p_2 - \pi_0}{\sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}}$$

Caso especial: cuando $\pi_0 = 0$, se puede usar proporción combinada:

$$Z_{calc} = \frac{p_1 - p_2 - \pi_0}{\sqrt{\hat{p}(1-\hat{p}) \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}} \quad \hat{p} = \frac{x_1 + x_2}{n_1 + n_2}$$

- ▶ Decisión: Comparar con valores críticos o usar p-valor

Ejemplo

Se desea comparar la proporción de hogares que hierven el agua antes de consumirla. En zona urbana: 90 de 120 hogares lo hacen; en zona rural: 80 de 110. ¿Las proporciones son las mismas?

$$H_0 : \pi_{urbana} - \pi_{rural} = 0 \quad H_1 : \pi_{urbana} - \pi_{rural} \neq 0 \quad \alpha = 0.05$$

```
x1 <- 90; n1 <- 120; (p1 <- x1/n1)
```

```
[1] 0.75
```

```
x2 <- 80; n2 <- 110; (p2 <- x2/n2)
```

```
[1] 0.7272727
```

```
(p <- (x1+x2)/(n1+n2))
```

```
[1] 0.7391304
```

```
(zcalc <- (p1-p2)/sqrt(p*(1-p)*(1/n1+1/n2)))
```

```
[1] 0.3921012
```

```
(zcrit1 <- qnorm(0.025))
```

```
[1] -1.959964
```

```
(zcrit2 <- qnorm(0.975))
```

```
[1] 1.959964
```

```
prop.test(x = c(x1, x2), n = c(n1, n2), alternative = "two.sided", correct = F)
```

2-sample test for equality of proportions without continuity correction

data: c(x1, x2) out of c(n1, n2)

X-squared = 0.15374, df = 1, p-value = 0.695

alternative hypothesis: two.sided

95 percent confidence interval:

-0.09097861 0.13643315

sample estimates:

prop 1	prop 2
0.7500000	0.7272727

Ejemplo

Un equipo de especialistas en gestión ambiental desea evaluar si una campaña de sensibilización ambiental logra aumentar sustancialmente la proporción de hogares que clasifican adecuadamente sus residuos sólidos.

- ▶ En el barrio sin campaña, 45 de 100 hogares clasifican correctamente.
- ▶ En el barrio con campaña, 70 de 100 hogares lo hacen.

El equipo busca determinar si la proporción de hogares que clasifican adecuadamente sus residuos sólidos en el barrio con campaña supera en más de un 20% a la del barrio sin campaña, lo cual justificaría su implementación a mayor escala. Para ello, se emplea un nivel de significancia del 5%

$$H_0 : \pi_{con} - \pi_{sin} \leq 0.20 \quad H_1 : \pi_{con} - \pi_{sin} > 0.20 \quad \alpha = 0.05$$

```
x1 <- 70; n1 <- 100; (p1 <- x1/n1)
```

```
[1] 0.7
```

```
x2 <- 45; n2 <- 100; (p2 <- x2/n2)
```

```
[1] 0.45
```

```
(zcalc <- (p1-p2)/sqrt(p1*(1-p1)/n1+p2*(1-p2)/n2))
```

```
[1] 3.696106
```

```
(zcrit2 <- qnorm(0.95))
```

```
[1] 1.644854
```

```
(pv <- 1-pnorm(zcalc))
```

```
[1] 0.0001094656
```