Taller 4

Juan Pablo Silvestre y Sebastián López

2024-10-29

Introducción

A continuación, se realizarán los ejercicios del 6 al 10, del taller pruebas de hipótesis de comparación de medias.

Recordemos que para ejecutar una prueba de comparación de medias es necesarios primero validar si las poblaciones son homocedásticas entre sí, es decir, tienen la misma varianza. Para ello, utilizaremos la prueba de hipótesis F-test programada en clase, la cual está alojada en el archivo functions.R

```r
# Carga de la función f.test2() que se encuentra alojada en el fichero functions.R
source("functions.R")
```

Ejercicio 6

Comparación de dos procesos de manufactura (Two Sample t-Test)

Se quiere comparar la media del diámetro de piezas producidas por dos procesos diferentes. El proceso 1 tiene una media muestral de 45.5 mm con una desviación estándar de 1.2 mm, y el proceso 2 tiene una media de 46.0 mm con una desviación estándar de 1.5 mm. Cada proceso tiene 25 observaciones. ¿Hay evidencia de que las medias de los diámetros son diferentes a un nivel de significancia del 5 %?

Solución

Como no tenemos los datos, debemnos entonces programar la prueba desde 0, esta prueba también está alojada en el archivo functions.R (de igual forma se mostrará al final del documento).

De acuerdo al enunciado, las hipótesis son:

H0: Diámetro de piezas medio real/poblacional del proceso 1 = diámetro de piezas medio real/poblacional del proceso 2 (H0: Mu.d1 = Mu.d2 -> Mu.d1 - Mu.d2 = 0) Ha: Diámetro de piezas medio real/poblacional del proceso 1 != diámetro de piezas medio real/poblacional del proceso 2 (Ha: Mu.d1 != Mu.d2 -> Mu.d1 - Mu.d2 != 0)

Con lo anterior, es un prueba de dos colas o bilateral.

Antes de ejecutar la prueba debemos validar si poblacionalmente las varianzas reales/poblacionales de los diámetros de piezas de los dos procesos son iguales o no. Para ello, debemos ejecutar una prueba F, la cual define las hipótesis de la siguiente forma:

H0: Varianza real/poblacional del diámetro de piezas del proceso 1 = varianza real/poblacional del diámetro de piezas del proceso 2 (H0: Var.d1 = Var.d2) Ha: Varianza real/poblacional del diámetro de piezas del proceso 1 != varianza real/poblacional del diámetro de piezas del proceso 2 (Ha: Var.d1 != Var.d2)

Con lo anterior, es una prueba de dos colas o bilateral.

```r
#data
n1 = 25
```

```
n2 = 25
media1 = 45.5
media2 = 46
s1 = 1.2
s2 = 1.5
alpha = 0.05
```

```
f.test2(s1, s2, n1, n2, alpha, "bilateral")
```

```
## No rechaza h0, con un p-value de: 0.5889407 y un ratio de varianza de 0.8
```

Con una confianza del 95%, NO hay evidencia muestral suficiente para rechazar la hipótesis de que las varianzas poblacionales son iguales, esto dado a que el P-Value = 0.5889407 que no es pequeño ni menor que alpha = 0.05. Por lo tanto, la evidencia muestral sugiere que el diámetro de piezas de los procesos 1 y 2, tienen la misma varianza.

Dado lo anterior, podemos ejecutar el two-sample t-test (No es necesario usar la corrección de Welch-Satterwaite).

```
correccion = FALSE
cola = "bilateral"
two.sample.t.test(n1, n2, media1, media2, s1, s2, correccion, cola, alpha)
```

```
## No rechaza H0, con un P-value de: 0.1993173 > 0.05
```

De acuerdo con la prueba, con una confianza del 95% NO hay evidencia muestral suficiente para rechazar la hipótesis nula que los diámetros de piezas de los procesos 1 y 2 son iguales. Debido a que el P-Value = 0.1993173 no es menor ni tampoco pequeño respecto a alpha = 0.05. Por ende, la evidencia muestral sugiere que el los diámetros de las piezas de los dos procesos son iguales.

Ejercicio 7

Comparación de tiempos de recuperación de pacientes (Welch Two Sample t-Test)

Un estudio comparó los tiempos de recuperación de dos grupos de pacientes que recibieron diferentes tratamientos para una enfermedad. El grupo A tuvo un tiempo promedio de recuperación de 12.3 días con una desviación estándar de 2.1 días (n=30), mientras que el grupo B tuvo un tiempo promedio de recuperación de 11.7 días con una desviación estándar de 1.8 días (n=25). ¿Es significativamente diferente el tiempo de recuperación entre los dos grupos?

Solución

De acuerdo al enunciado, las hipótesis son:

H0: Tiempo de recuperación promedio real/poblacional del grupo A = Tiempo de recuperación promedio real/poblacional del grupo B (H0: Mu.A = Mu.B -> Mu.A - Mu.B = 0) Ha: Tiempo de recuperación promedio real/poblacional del grupo A != Tiempo de recuperación promedio real/poblacional del grupo B (Ha: Mu.A != Mu.B -> Mu.A - Mu.B != 0)

Con lo anterior, Es un prueba de dos colas o bilateral.

Validar si poblacionalmente las varianzas reales/poblacionales del tiempo de recuperación de los grupos. Para ello, debemos ejecutar una prueba F, la cual define las hipótesis de la siguiente forma:

H0: Varianza real/poblacional del tiempo de recuperación del grupo A = varianza real/poblacional del tiempo de recuperación del grupo B (H0: Var.A = Var.B) Ha: Varianza real/poblacional del tiempo de recuperación del grupo A != varianza real/poblacional del tiempo de recuperación del grupo B (Ha: Var.A != Var.B)

Con lo anterior, Es un prueba de dos colas o bilateral.

```r
#data
n1 = 30
n2 = 25
media1 = 12.3
media2 = 11.7
s1 = 2.1
s2 = 1.8
alpha = 0.05
```

```r
f.test2(s1, s2, n1, n2, alpha, "bilateral")
```

```
## No rechaza h0, con un p-value de: 0.7055728 y un ratio de varianza de 1.166667
```

Con una confianza del 95%, NO hay evidencia muestral suficiente para rechazar la hipótesis de que las varianzas poblacionales son iguales, dado a que el P-Value = 0.7055728 que no es pequeño ni menor que alpha = 0.0. Por lo tanto, la evidencia muestral sugiere que los tiempos de recuperación de los grupos A y B, tienen la misma varianza.

Dado lo anterior, podemos ejecutar el two-sample t-test (No es necesario usar la corrección de Welch-Satterwaite).

```r
correccion = FALSE
cola = "bilateral"
two.sample.t.test(n1, n2, media1, media2, s1, s2, correccion, cola, alpha)
```

```
## No rechaza H0, con un P-value de: 0.2657434 > 0.05
```

Con base en la prueba, con una confianza del 95% NO hay evidencia muestral suficiente para rechazar la hipótesis nula de que el tiempo de recuperación de los grupos A y B son iguales; lo anterior, dado a que el P-Value = 0.2657434 no es menor ni tampoco pequeño respecto a alpha = 0.05. Por lo tanto, la evidencia muestral sugiere que los que tanto el grupo A, como el grupo B, tuvieron un tiempo de recuperación igual.

Ejercicio 8

Comparación de peso promedio en dos lotes de productos (Two Sample t-Test)

En una planta de alimentos, se sospecha que dos lotes de producción tienen un peso promedio diferente. El lote 1 tiene una media muestral de 500 g y una desviación estándar de 8 g (n=20). El lote 2 tiene una media muestral de 505 g y una desviación estándar de 10 g (n=22). ¿Existen diferencias significativas entre los dos lotes a un nivel de significancia del 1 %?

Solución

De acuerdo al enunciado, las hipótesis son:

H0: Peso medio real/poblacional del lote 1 = Peso medio real/poblacional del lote 2 (H0: Mu.l1 = Mu.l2 -> Mu.l1 - Mu.l2 = 0) Ha: Peso medio real/poblacional del lote 1 != Peso medio real/poblacional del lote 2 (Ha: Mu.l1 != Mu.l2 -> Mu.l1 - Mu.l2 != 0)

Con lo anterior, Es un prueba de dos colas o bilateral.

También, debemos validar si poblacionalmente las varianzas reales/poblacionales de los diámetros de piezas de los dos procesos son iguales o no. Para ello, debemos ejecutar una prueba F, la cual define las hipótesis de la siguiente forma:

H0: Varianza real/poblacional del peso del lote 1 = varianza real/poblacional del peso del lote 2 (H0: Var.l1 = Var.l2) Ha: Varianza real/poblacional del peso del lote 1 = varianza real/poblacional del peso del lote 2 (Ha: Var.l1 = Var.l2)

Con lo anterior, Es un prueba de dos colas o bilateral.

```r
#data
n1 = 20
n2 = 22
media1 = 500
media2 = 505
s1 = 8
s2 = 10
alpha = 0.01
```

```r
f.test2(s1, s2, n1, n2, alpha, "bilateral")
```

```
## No rechaza h0, con un p-value de: 0.6289368 y un ratio de varianza de 0.8
```

Con una confianza del 99%, NO hay evidencia muestral suficiente para rechazar la hipótesis de que las varianzas poblacionales son iguales, dado a que el P-Value = 0.6289368 que no es pequeño ni menor que alpha = 0.01. Por lo tanto, la evidencia muestral sugiere que el peso de los lotes 1 y 2, tienen la misma varianza.

Dado lo anterior, podemos ejecutar el two-sample t-test (No es necesario usar la corrección de Welch-Satterwaite)

```r
correccion = FALSE
cola = "bilateral"
two.sample.t.test(n1, n2, media1, media2, s1, s2, correccion, cola, alpha)
```

```
## No rechaza H0, con un P-value de: 0.08310037 > 0.01
```

Con base en la prueba, con una confianza del 99% NO hay evidencia muestral suficiente para rechazar la hipótesis nula de que el peso de los lotes de producción 1 y 2 son iguales; lo anterior, dado a que el P-Value = 0.08310037 no es menor ni tampoco pequeño respecto a alpha = 0.01. Por lo tanto, la evidencia muestral sugiere que el peso de los dos lotes de producción iguales.

Ejercicio 9

Comparación de la presión de tuberías antes y después de una intervención (Paired t-Test)

En un sistema de tuberías industriales, se mide la presión de operación antes y después de una intervención en 12 puntos. La media de la presión antes es de 100 psi con una desviación estándar de 5 psi, y la media después es de 95 psi con una desviación estándar de 4 psi. ¿Ha habido una reducción significativa en la presión con un nivel de significancia del 5 %?

Solución

De acuerdo al enunciado, las hipótesis son:

Ho: La presión media real antes de la intervención es igual a la presión media real después de la intervención, es decir, no hay reducción significativa en la presión después de la intervención.
Ho: $\mu_{\text{antes}} <= \mu_{\text{después}}$

Ha:La presión media real antes de la intervención es mayor que la presión media real después de la intervención, indicando que la intervención reduce la presión en el sistema.
Ha: $\mu_{\text{antes}} > \mu_{\text{después}}$

Dado que estamos buscando una reducción, esta es una prueba de cola derecha.

```r
m.a = 100
m.d= 95
s.a = 5
s.d = 4
n = 12
```

```r
alpha = 0.05

media_diferencia = m.a - m.d
desv_est_diferencias = sqrt(s.a^2 + s.d^2)

t_value = media_diferencia / (desv_est_diferencias / sqrt(n))

df = n - 1

p_value = pt(t_value, df = df, lower.tail = FALSE)
p_value
```

```
## [1] 0.01023634
```

```r
if (p_value < alpha) {
  resultado = "Se rechaza Ho."
} else {
  resultado = "No se rechaza Ho."
}
resultado
```

```
## [1] "Se rechaza Ho."
```

De acuerdo con la prueba, con una confianza del 95%, HAY evidencia muestral suficiente para rechazar la hipótesis nula de que no hay reducción en la presión de las tuberías después de la intervención; esto dado que el p-valor obtenido es menor que el nivel de significancia ($\alpha = 0.05$). En este caso, el p-valor es de aproximadamente 0.01, lo cual es significativamente bajo. Por lo tanto, la evidencia muestral sugiere que la intervención produjo una reducción significativa en la presión de las tuberías.

Ejercicio 10

Comparación de la efectividad de dos tratamientos médicos (Welch Two Sample t-Test)

Un estudio clínico investiga si dos tratamientos tienen efectos diferentes en la reducción de un marcador biológico. El grupo A (n=40) tiene una reducción promedio de 15 unidades con una desviación estándar de 3.5, mientras que el grupo B (n=35) tiene una reducción promedio de 14 unidades con una desviación estándar de 2.8. ¿Existen diferencias significativas en la reducción del marcador entre ambos tratamientos?

Solución

De acuerdo al enunciado, las hipótesis son:

H0: La reducción promedio real/poblacional del grupo A = La reducción promedio real/poblacional del grupo B
(H0: Mu.A = Mu.B → Mu.A - Mu.B = 0)

Ha: La reducción promedio real/poblacional del grupo A != La reducción promedio real/poblacional del grupo B
(Ha: Mu.A != Mu.B → Mu.A - Mu.B != 0)

Con lo anterior, se trata de una prueba de dos colas o bilateral.

También debemos validar si las varianzas reales/poblacionales de las reducciones de los dos grupos son iguales o no. Para ello, debemos ejecutar una prueba F, que define las hipótesis de la siguiente forma:

H0 Varianza real/poblacional de la reducción del grupo A = Varianza real/poblacional de la reducción del grupo B
(H0: Var.A = Var.B)

Ha:Varianza real/poblacional de la reducción del grupo A != Varianza real/poblacional de la reducción del grupo B
(Ha: Var.A != Var.B)

Con lo anterior, se trata de una prueba de dos colas o bilateral.

```
nA = 40
nB = 35
mediaA = 15
mediaB = 14
sA = 3.5
sB = 2.8
alpha = 0.05
```

```
f.test2(sA, sB, nA, nB, alpha, "bilateral")
```

```
## No rechaza h0, con un p-value de: 0.5105979 y un ratio de varianza de 1.25
```

Con una confianza del 95%, NO hay evidencia muestral suficiente para rechazar la hipótesis de que las varianzas poblacionales son iguales, dado que el P-Value = 0.2154321 que no es menor que alpha = 0.05. Por lo tanto, la evidencia muestral sugiere que la reducción del marcador entre los grupos A y B tiene la misma varianza.

Dado lo anterior, podemos ejecutar el two-sample t-test (no es necesario usar la corrección de Welch-Satterwaite).

```
correccion = FALSE
cola = "bilateral"
two.sample.t.test(nA, nB, mediaA, mediaB, sA, sB, correccion, cola, alpha)
```

```
## No rechaza H0, con un P-value de: 0.1802112 > 0.05
```

Con base en la prueba, con una confianza del 95% NO hay evidencia muestral suficiente para rechazar la hipótesis nula de que la reducción del marcador biológico entre los tratamientos A y B es igual; esto es, dado que el P-Value = 0.0678 no es menor ni tampoco pequeño respecto a alpha = 0.05. Por lo tanto, la evidencia muestral sugiere que la reducción del marcador biológico es similar entre los tratamientos A y B.

Por último

Se muestra cómo se ejecutó la prueba Two-Sample T-test.

```
two.sample.t.test = function(n1, n2, media1, media2, s1, s2, correccion, cola, alpha ){

  t0 = 0
  parametro = 0

  #correcion de Welch

  if (correccion == TRUE){

    t0 = (media1 - media2) / sqrt((s1^2 / n1) + (s2^2 / n2))
    v = ((s1^2 / n2) + (s2^2 / n2))^2 / ((s1^2 / n1)^2 / (n1 - 1) + (s2^2 / n2)^2 / (n2 - 1))
    parametro = v

  } else {

    varianza.ponderada = sqrt(((n1 - 1) * s1^2 + (n2 - 1) * s2^2) / (n1 + n2 - 2))
    t0 = (media1 - media2) / (varianza.ponderada * sqrt((1 / n1) + (1 / n2)))
```

```r
    parametro = n1 + n2 - 2

  }

  #decision

  if(cola == "izquierda"){

    p.value = pt(t0, parametro)

    if(t0 <= qt(alpha, parametro)){
      cat("Rechaza H0, con un P-value de:", p.value, "<", alpha)
    } else {
      cat("No rechaza H0, con un P-value de:", p.value, ">", alpha)
    }

  } else if (cola == "derecha"){

    p.value = 1 - pt(t0, df)

    if(t0 >= qt(1 - alpha, parametro)){
      cat("Rechaza H0, con un P-value de:", p.value, "<", alpha)
    } else {
      cat("No rechaza H0, con un P-value de:", p.value, ">", alpha)
    }

  } else if (cola == "bilateral"){

    p.value = 2 * (1 - pt(abs(t0), parametro))

    if(t0 >= qt(1 - alpha / 2, parametro)){
      cat("Rechaza H0, con un P-value de:", p.value, "<", alpha)
    } else {
      cat("No rechaza H0, con un P-value de:", p.value, ">", alpha)
    }
  }
}
```