Tarea 8 - Pruebas de Hipótesis

Héctor Hibran Tapia Fernández - A01661114

2024-08-23

Enlatados

Los pesos de 21 latas de duraznos empacados elegidas al azar fueron:

Peso de las latas: 11, 11.6, 11.6, 11.7, 10.9, 11.6, 12, 11.2, 11.5, 12, 12, 11.4, 11.2, 10.8, 10.5, 11.8, 12.2, 10.9, 11.8, 11.4, 12.1

Por estudios anteriores se saber que población del peso de las latas se distribuye normalmente.

Si a los dueños no les conviene que el peso sea menor, pero tampoco mayor a 11.7, prueba la afirmación de que el verdadero peso de las latas es de 11.7 con un nivel de confianza de 0.98 haciendo uso de los datos obtenidos en la muestra.

Muestra tu procedimiento siguiendo los 4 pasos de las pruebas de hipótesis.

Paso 1. Hipótesis (De dos colas)

- $H_0 : \mu = 11.7$
- $H_1 : \mu \neq 11.7$

¿Cómo se distribuye $\bar{x}$?

- X se distribuye como una Normal
- n < 30
- No conocemos sigma

Entonces usaremos la Prueba t de Student.

Paso 2. Regla de Decisión

Nivel de confianza es de 0.98 Nivel de significancía es de 0.02

Necesito encontrar a cuántas desviaciones estándar está lejos el valor frontera.

```
n = 21
alfa = 0.02
t_f = qt(alfa/2, n-1)
cat("t_f = ", t_f)
```

```
## t_f =  -2.527977
```

Rechazo H_0 sí:

- $|t_e| > 2.53$
- p-valor < 0.02

Paso 3. Análisis del Resultado

- t_e: Número de desviaciones al que $\bar{x}$ se encuentra lejos de $mu = 11.7$
- Valor P: Proababilidad de obtener lo que obtuve de muestra o un valor más extremo.

Estadístico de Prueba

```
X = c(11, 11.6, 11.6, 11.7, 10.9, 11.6, 12, 11.2, 11.5, 12, 12, 11.4, 11.2, 10.8, 10.5,
11.8, 12.2, 10.9, 11.8, 11.4, 12.1)

xb = mean(X)
s = sd(X)
mu = 11.7

te = (xb - mu) / (s/sqrt(n))
cat("te =", te, "\n")
```

```
## te = -2.068884
```

```
valorp = 2*pt(te, n-1)
cat("Valor p =", valorp)
```

```
## Valor p = 0.0517299
```

Un atajo

```
t.test(X, mu = 11.7, alternative = ("two.sided"), conf.level = 0.98)
```

```
##
##  One Sample t-test
##
## data:  X
## t = -2.0689, df = 20, p-value = 0.05173
## alternative hypothesis: true mean is not equal to 11.7
## 98 percent confidence interval:
##  11.22388 11.74755
## sample estimates:
## mean of x
##  11.48571
```

Paso 4. Conclusión

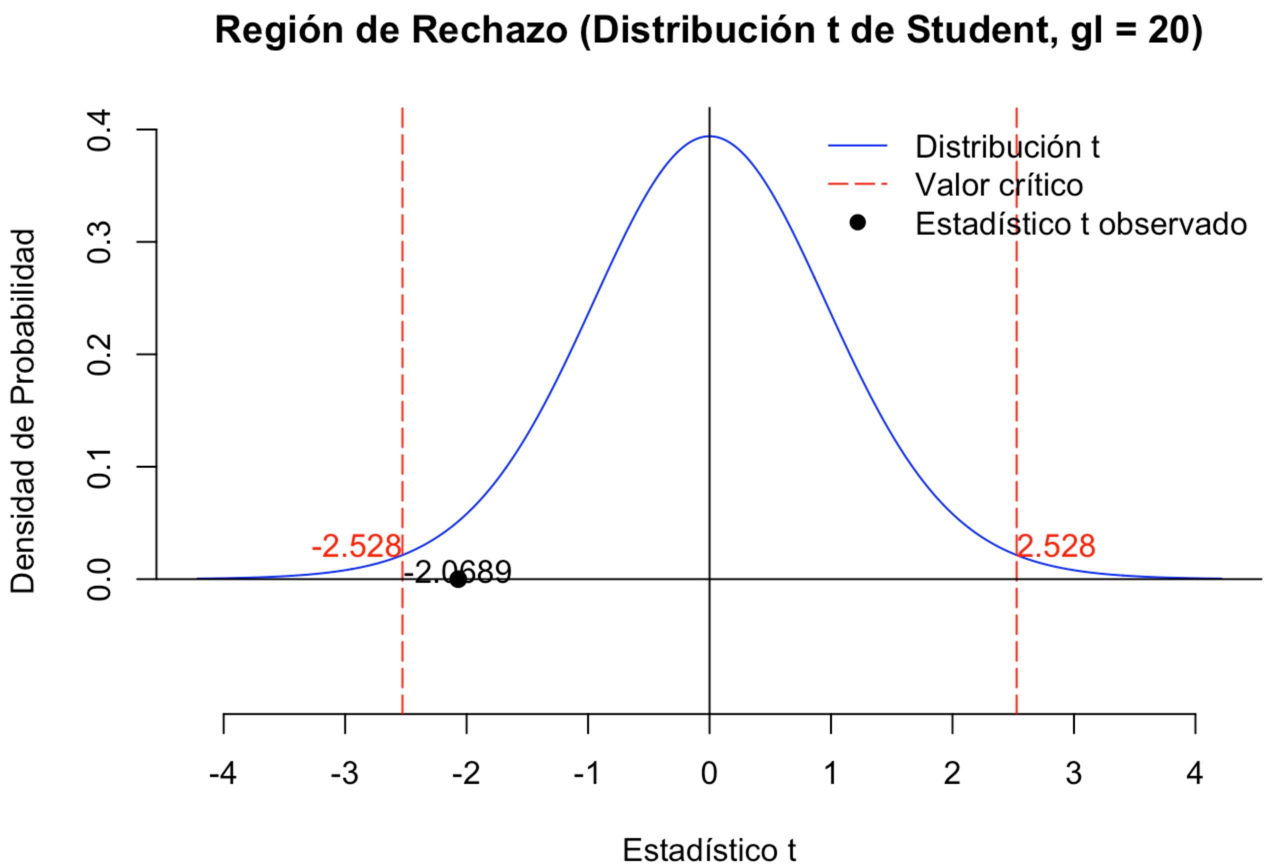
Comparamos: Regla de Decisión vs Análisis de Resultado.

$|t_e| = 2.07 < 2.53$ -> *No se rechaza la* H_0 $p\text{-valor} = 0.05 > 0.02$ -> No se rechaza la H_0

Elabora un gráfico que muestre la regla de decisión y el punto

donde queda el estadístico de prueba.

```
n <- 21  # tamaño de la muestra
gl <- n - 1  # grados de libertad

sigma <- sqrt((n - 1) / (n - 3)) # Factor de ajuste para la distribución t
x <- seq(-4 * sigma, 4 * sigma, 0.01) # Valores para el eje x
y <- dt(x, gl) # Densidad de la distribución t con gl grados de libertad
alpha <- 0.02 # Estadístico t crítico para la región de rechazo a un nivel de confianza
del 98%

t_f <- qt(1 - alpha / 2, gl)  # Valor crítico de t para dos colas
te <- -2.0689  # Este es el valor del estadístico t obtenido del test

plot(x, y, type = "l", col = "blue", xlab = "Estadístico t", ylab = "Densidad de Probabi
lidad",
     ylim = c(-0.1, 0.4), frame.plot = FALSE, xaxt = "n", yaxt = "n",
     main = "Región de Rechazo (Distribución t de Student, gl = 20)")

axis(1, at = seq(-4, 4, by = 1))
axis(2, at = seq(0, 0.4, by = 0.1))

abline(v = t_f, col = "red", lty = 5)
abline(v = -t_f, col = "red", lty = 5)

text(t_f, 0.03, paste0(round(t_f, 3)), col = "red", adj = 0)
text(-t_f, 0.03, paste0(round(-t_f, 3)), col = "red", adj = 1)

abline(h = 0)
abline(v = 0, col = "black")

points(te, 0, pch = 19, cex = 1.1)
text(te, -0.02, "-2.0689", pos = 3)

legend("topright", legend = c("Distribución t", "Valor crítico", "Estadístico t observad
o"),
       col = c("blue", "red", "black"), lty = c(1, 5, NA), pch = c(NA, NA, 19), bty =
"n")
```



Concluye en el contexto del problema.

- El valor del estadístico de prueba t es -2.069, que está dentro del rango de no rechazo de la hipótesis nula, ya que $|-2.069| < 2.528$, lo que significa que no rechazamos la hipótesis nula. Con un nivel de confianza del 98%, no hay suficiente evidencia para concluir que el verdadero peso promedio de las latas sea diferente de 11.7.

La decisión de Fowle Marketing Research, Inc.

Fowle Marketing Research, Inc., basa los cargos a un cliente bajo el supuesto de que las encuestas telefónicas (para recopilación de datos) pueden completarse en un tiempo medio de 15 minutos o menos. Si el tiempo es mayor a 15 minutos entonces se cobra una tarifa adicional. Compañías que contratan estos servicios piensan que el tiempo promedio es mayor a lo que especifica Fowle Marketing Research Inc. así que realizan su propio estudio en una muestra aleatoria de llamadas telefónicas y encuentran los siguientes datos:

Tiempo: 17, 11, 12, 23, 20, 23, 15, 16, 23, 22, 18, 23, 25, 14, 12, 12, 20, 18, 12, 19, 11, 11, 20, 21, 11, 18, 14, 13, 13, 19, 16, 10, 22, 18, 23

Por experiencias anteriores, se sabe que σ=4 minutos. Usando un nivel de significancia de 0.07, ¿está justificada la tarifa adicional?

Muestra tu procedimiento siguiendo los 4 pasos de las pruebas

de hipótesis.

Paso 1. Hipótesis (De una cola)

- $H_0 : \mu \leq 15$
- $H_1 : \mu > 15$

¿Cómo se distribuye $\bar{x}$?

- X se distribuye como una Normal
- n > 30
- Conocemos sigma

Entonces usaremos la Prueba Z.

Paso 2. Regla de Decisión

Nivel de confianza es de 0.93 Nivel de significancia es de 0.07

Ya sabemos que la desviación estándar de la población (sigma) es igual a 4.

Para α = 0.07 podemos, buscar el valor crítico z correspondiente a un test de una cola, usando tablas de la distribución normal estándar, o lo siguiente:

```
alfa = 0.07
z_critico = qnorm(1 - alfa)
cat("z_critico = ", z_critico)
```

```
## z_critico =  1.475791
```

Rechazo H_0 sí:

- $z > 1.48$
- $p\text{-valor} < 0.07$

Paso 3. Análisis del Resultado

- z: Indica cuántas desviaciones estándar está la media muestral por encima de la media poblacional supuesta bajo la hipótesis nula (μ = 15 minutos).
- Valor P: Proababilidad de obtener lo que obtuve de muestra o un valor más extremo.

Estadístico de Prueba

```
X = c(17, 11, 12, 23, 20, 23, 15, 16, 23, 22, 18, 23, 25, 14, 12, 12, 20, 18, 12, 19, 1
1, 11, 20, 21, 11, 18, 14, 13, 13, 19, 16, 10, 22, 18, 23)

n = length(X)
xb = mean(X)
mu = 15
sigma = 4

z = (xb - mu) / (sigma / sqrt(n))
cat("Estadístico de Prueba Z =", z)
```

```
## Estadístico de Prueba Z = 2.95804
```

```
p_value = 1 - pnorm(z)
cat("Valor p =", p_value, "\n")
```

```
## Valor p = 0.00154801
```

Un atajo

```
library(BSDA)
```

```
## Loading required package: lattice
```

```
## Warning: package 'lattice' was built under R version 4.2.3
```

```
##
## Attaching package: 'BSDA'
```

```
## The following object is masked from 'package:datasets':
##
##     Orange
```

```
sigma = 4
z_test_result = z.test(X, mu = 15, sigma.x = sigma, alternative = "greater")
z_test_result
```

```
##
##  One-sample z-Test
##
## data:  X
## z = 2.958, p-value = 0.001548
## alternative hypothesis: true mean is greater than 15
## 95 percent confidence interval:
##  15.88788       NA
## sample estimates:
## mean of x
##        17
```

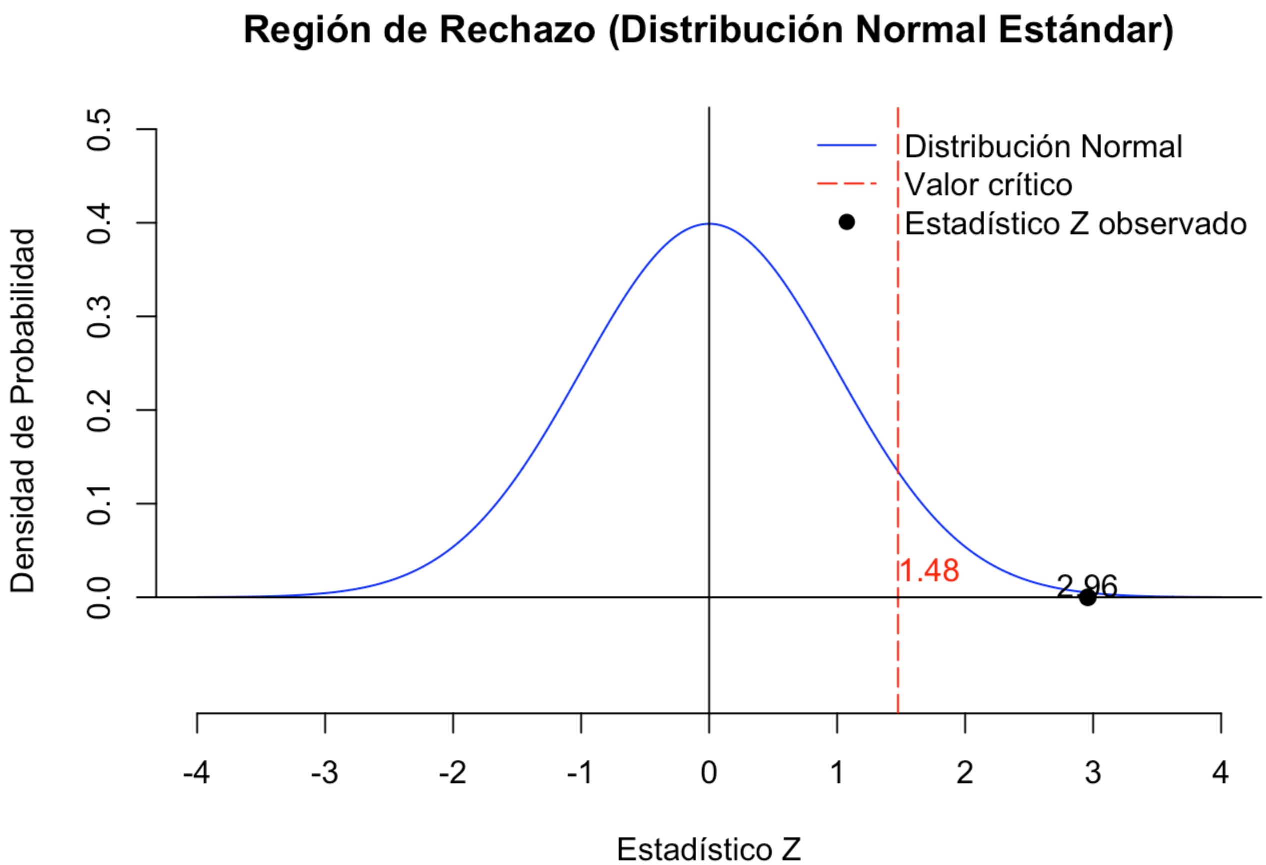
Paso 4. Conclusión

Comparamos: Regla de Decisión vs Análisis de Resultado.

- $z = 2.958 > 1.48$ -> Se rechaza la H_0
- p-valor $= 0.001548 < 0.07$ -> Se rechaza la H_0

Elabora un gráfico que muestre la regla de decisión y el punto donde queda el estadístico de prueba.

```
mu <- 15  # Media poblacional bajo la hipótesis nula
sigma <- 4  # Desviación estándar de la población
n = length(X) # Tamaño de la muestra
alpha <- 0.07  # Nivel de significación
z_observado <- 2.95804  # Estadístico Z observado
z_critico <- qnorm(1 - alpha)  # Valor crítico de Z para una cola

x <- seq(-4, 4, 0.01)
y <- dnorm(x)

plot(x, y, type = "l", col = "blue", xlab = "Estadístico Z", ylab = "Densidad de Probabi
lidad",
     ylim = c(-0.1, 0.5), frame.plot = FALSE, xaxt = "n", yaxt = "n",
     main = "Región de Rechazo (Distribución Normal Estándar)")

axis(1, at = seq(-4, 4, by = 1))
axis(2, at = seq(0, 0.5, by = 0.1))

abline(v = z_critico, col = "red", lty = 5)
text(z_critico, 0.03, paste0(round(z_critico, 2)), col = "red", adj = 0)

abline(h = 0)
abline(v = 0, col = "black")

points(z_observado, 0, pch = 19, cex = 1.1)
text(z_observado, -0.02, round(z_observado, 2), pos = 3)

legend("topright", legend = c("Distribución Normal", "Valor crítico", "Estadístico Z obs
ervado"),
       col = c("blue", "red", "black"), lty = c(1, 5, NA), pch = c(NA, NA, 19), bty =
"n")
```



Concluye en el contexto del problema.

-
    - El valor del estadístico de prueba z es 2.958, que está fuera dentro del rango de no rechazo de la hipótesis nula, ya que $z = 2.958 > 1.48$, lo que significa que rechazamos la hipótesis nula. Con un nivel de confianza del 97%, hay suficiente evidencia para concluir que el tiempo promedio de las encuestas telefónicas es mayor a 15 minutos. Por lo tanto, la tarifa adicional está justificada.