Pruebas de hipótesis

Catherine Rojas

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 sabe 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

1. Formulación de hipótesis y datos

Hipótesis nula (H_0):

$$H_0: \mu = 11.7 \quad \text{(La media del peso es 11.7)}$$

Hipótesis alternativa (H_1):

$$H_1: \mu \neq 11.7 \quad \text{(La media del peso es diferente de 11.7)}$$

$\bar{X}$ se distribuye como una normal $ n < 30$ No conocemos sigma Entonces: la distribución muestral es una t de Student

```
pesos <- 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)
```

2. Regla de decisión: Nivel de significancia Se requiere encontrar a cuántas desviaciones estándar está lejos el valor frontera.

```
n = 21
alpha <- 0.02       # El nivel de confiaza es 98% o 0.98
t_f =qt(alpha/2, n-1)
cat("t_f = ", t_f)  # Valor frontera
```

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

Se rechaza H_0 si: * $|t_e > 2.53|$ * valor p < 0.02

3. Análisis del resultado

- t_e: Número de desviaciones al que $\bar{X}$ se encuentra lejos de $\mu = 11.7$
- Valor p: Probabiliad de obtener lo que obtuve en la muestra o un valor más extremo.

```
xb = mean(pesos)
s = sd(pesos)
mu = 11.7

t_e= (xb-mu)/ (s/sqrt(n))
cat("t_e = ", t_e)

## t_e =  -2.068884

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

## Valor p=  0.0517299
```

4. Conclusión

Comparar: Regla de decisión vs Análisis del resultado

Entonces: * $|t_e| = 2.07 < 2.53$ -> No RH_0 * Valor p = 0.05 < 0.02 -> No RH_0

En el contexto: Las latas de durazno tienen el peso requerido.

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

```
# Estadístico de prueba (t-test para una muestra)
t_test <- t.test(pesos, mu= mu, conf.level = 1 - alpha)

# Resultado de la prueba
t_test

##
##  One Sample t-test
##
## data:  pesos
## 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:
```

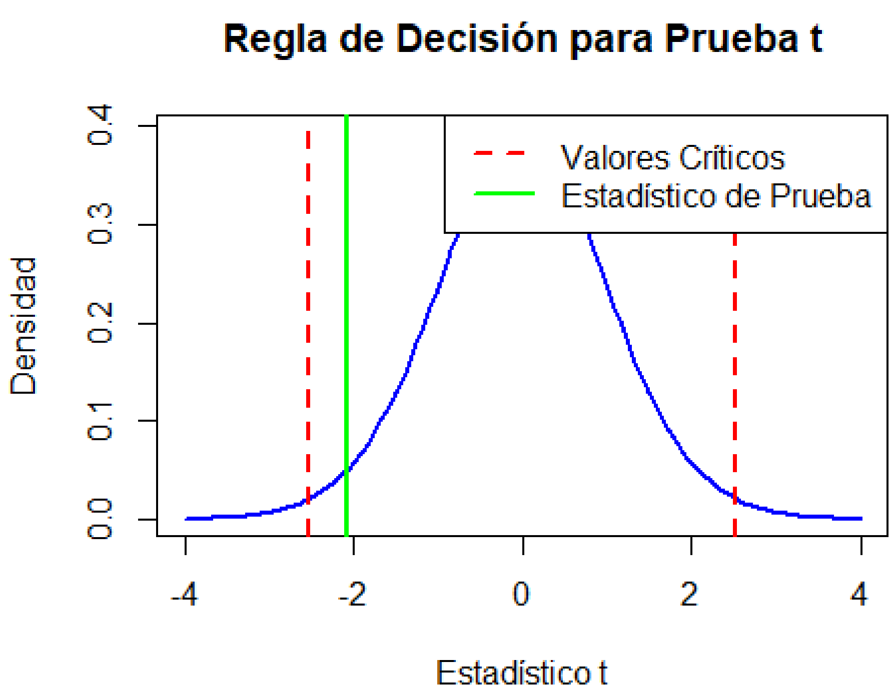
```
## mean of x
##  11.48571
```

Gráfico de la regla de decisión y el estadístico de prueba

```
# Extraemos los valores críticos para una prueba bilateral
gl <- length(pesos) - 1
valor_critico <- qt(1 - alpha / 2, gl)

# Generamos la distribución t
x <- seq(-4, 4, length = 100)
y <- dt(x, df = gl)

# Creamos el gráfico
plot(x, y, type = "l", lwd = 2, col = "blue", main = "Regla de Decisión
para Prueba t",
     xlab = "Estadístico t", ylab = "Densidad")
abline(v = c(-valor_critico, valor_critico), col = "red", lwd = 2, lty =
2)
abline(v = t_test$statistic, col = "green", lwd = 2)
legend("topright", legend = c("Valores Críticos", "Estadístico de
Prueba"),
       col = c("red", "green"), lwd = 2, lty = 2:1)
```




```
# Imprimimos los valores críticos y el valor del estadístico t
cat("Valores críticos:", -valor_critico, "y", valor_critico, "\n")
```

```
## Valores críticos: -2.527977 y 2.527977
cat("Estadístico t:", t_test$statistic, "\n")
## Estadístico t: -2.068884
```

Concluye en el contexto del problema.

El valor absoluto del estadístico t (-2.0689) no supera los valores críticos (-2.527977, 2.527977), lo que significa que no rechazamos la hipótesis nula al nivel de significancia del 2%, $\alpha = 0.02$. Además, el valor p (0.05173) es mayor que $\alpha = 0.02$, lo que refuerza esta decisión. Es decir, no podemos concluir que el peso medio es diferente de 11.7

Dado que la media de las latas podría estar dentro de este rango de pesos, los dueños podrían considerar que el peso está en un nivel aceptable, aunque el peso promedio observado es ligeramente inferior a 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 significación de 0.07, ¿está justificada la tarifa adicional?

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

1. Formulación de hipótesis y datos

Hipótesis nula (H_0):

$$H_0: \mu \leq 15 \quad \text{(El tiempo medio es 15 minutos o menos.)}$$

Hipótesis nula (H_1):

$$H_1: \mu > 15 \quad \text{(El tiempo medio es mayor a 15 minutos.)}$$

```
tiempos <- c(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)
```

2. Regla de decisión: Nivel de significancia

Regla de Decisión

Rechazamos H_0 si el valor del estadístico z es mayor que el valor crítico de $z_{\text{crítico}}$.

No rechazamos H_0 si el valor del estadístico z es menor o igual que el valor crítico.

- Si $z > 1.48$, rechazamos H_0.
- Si $z \leq 1.48$, no rechazamos H_0.

```r
# Media bajo la hipótesis nula
mu <- 15
sigma <- 4

# Nivel de significancia
alpha <- 0.07

n <- length(tiempos)
media_muestra <- mean(tiempos)

# Estadístico z
# Prueba t para una muestra
z <- (media_muestra - mu) / (sigma / sqrt(n))

# Regla de decisón
z_critico <- qnorm(1 - alpha)
decision <- ifelse(z > z_critico, "Rechazar H0", "No rechazar H0")

# Resultados
cat("La prueba t para una muestra es:")

## La prueba t para una muestra es:

z

## [1] 2.95804

cat("La regla de decisión es:")

## La regla de decisión es:

z_critico

## [1] 1.475791

decision

## [1] "Rechazar H0"
```

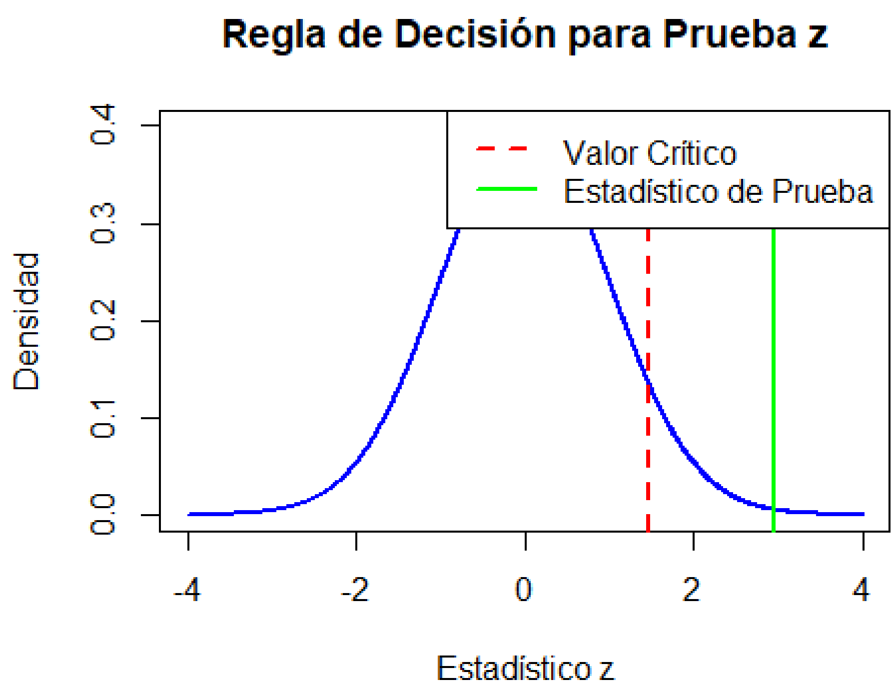
Para este caso, el valor crítico $z_{\text{crítico}}$ a un nivel de significancia de 0.07 es 1.48.
Rechazamos H_0

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

```r
# Gráfico de la regla de decisión
x <- seq(-4, 4, length = 1000)
y <- dnorm(x)

plot(x, y, type = "l", lwd = 2, col = "blue", xlab = "Estadístico z",
ylab = "Densidad", main = "Regla de Decisión para Prueba z")
abline(v = z_critico, col = "red", lwd = 2, lty = 2)
abline(v = z, col = "green", lwd = 2)

legend("topright", legend = c("Valor Crítico", "Estadístico de Prueba"),
col = c("red", "green"), lwd = 2, lty = 2:1)
```



4. Concluye en el contexto del problema.

Dado que el estadístico de prueba $z = 2.96$ es mayor que el valor crítico $z_{\text{crítico}} = 1.48$, rechazamos la hipótesis nula H_0. Esto significa que hay suficiente evidencia estadística para concluir que el tiempo promedio de las encuestas telefónicas es mayor a 15 minutos.

¿Está justificada la tarifa adicional?

Sí, con un nivel de significancia de 0.07, está justificada la tarifa adicional de Fowle Marketing Research, Inc., ya que los resultados sugieren que el tiempo promedio de las encuestas telefónicas excede los 15 minutos especificados. Por lo tanto, la tarifa adicional puede aplicarse.