Pruebas_Hipotesis

Arturo Garza Campuzano

2023-08-23

Autor: Arturo Garza Campuzano

Matrícula: A00828096

Pruebas de hipótesis

1. Resuelve las dos partes del problema "Enlatados" que se encuentran al final de la presentación de Pruebas de hipótesis.

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

11.0, 11.6, 10.9, 12.0, 11.5, 12.0, 11.2, 10.5, 12.2, 11.8, 12.1, 11.6, 11.7, 11.6, 11.2, 12.0, 11.4, 10.8, 11.8, 10.9, 11.4

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.

1.1 Muestra tu procedimiento siguiendo los 4 pasos.

Paso 1. Definir las hipótesis

Se establecen la hipótesis nula y la hipótesis alternativa:

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

El estadístico designado para esta prueba de hipótesis es $\bar{x}$, el cual se supondrá que sigue una distribución t Student.

Por lo tanto, tenemos lo siguiente:

- $\mu_{\bar{x}} = 11.7$
- $\sigma_{\bar{x}} = \frac{s}{\sqrt{n}}$

Paso 2. Regla de decisión

El nivel de confianza con el que se trabajará es 0.98 y, por lo tanto: $\alpha = 0.02$.

```r
x = c(11.0, 11.6, 10.9, 12.0, 11.5, 12.0, 11.2, 10.5, 12.2, 11.8, 12.1,
11.6, 11.7, 11.6, 11.2, 12.0, 11.4, 10.8, 11.8, 10.9, 11.4)
alfa = 0.02
n  = length(x)
t0 = qt(alfa/2,n-1) # Valor frontera
cat("t0 =",t0)
```

```
## t0 = -2.527977
```

Como regla de decisión, H_0 se rechaza si:

- $|t^*| > 2.53$
- $p < 0.02$

Donde, t^* es el número de desviaciones estándar al que $\bar{x}$ está lejos de μ.

Paso 3. Análisis de resultado

Se tiene que calcular:

- t^* (qué tan lejos está $\bar{x}$ de μ).
- Valor de p (la probabilidad de que $\bar{x}$ esté en las colas de la distribución).

```
# Calculo de t*
m = mean(x)
s = sd(x)
sm = s/sqrt(n)
te = (m-11.7)/sm
cat("t =",te,"\n")
```

```
## t = -2.068884
```

```
# Calculo de valor p
valorp = 2*pt(te,n-1)
cat("Valor p =",valorp,"\n")
```

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

Paso 4. Conclusiones

Después de haber calculado t^* y el valor p, se pueden llegar a las siguientes observaciones:

- Como el valor p (0.05173) es mayor que 0.02, entonces H_0 no se rechaza.
- Como $|t^*|$ (2.07) es menor que 2.53, entonces H_0 no se rechaza.

Más fácil (paso 3):

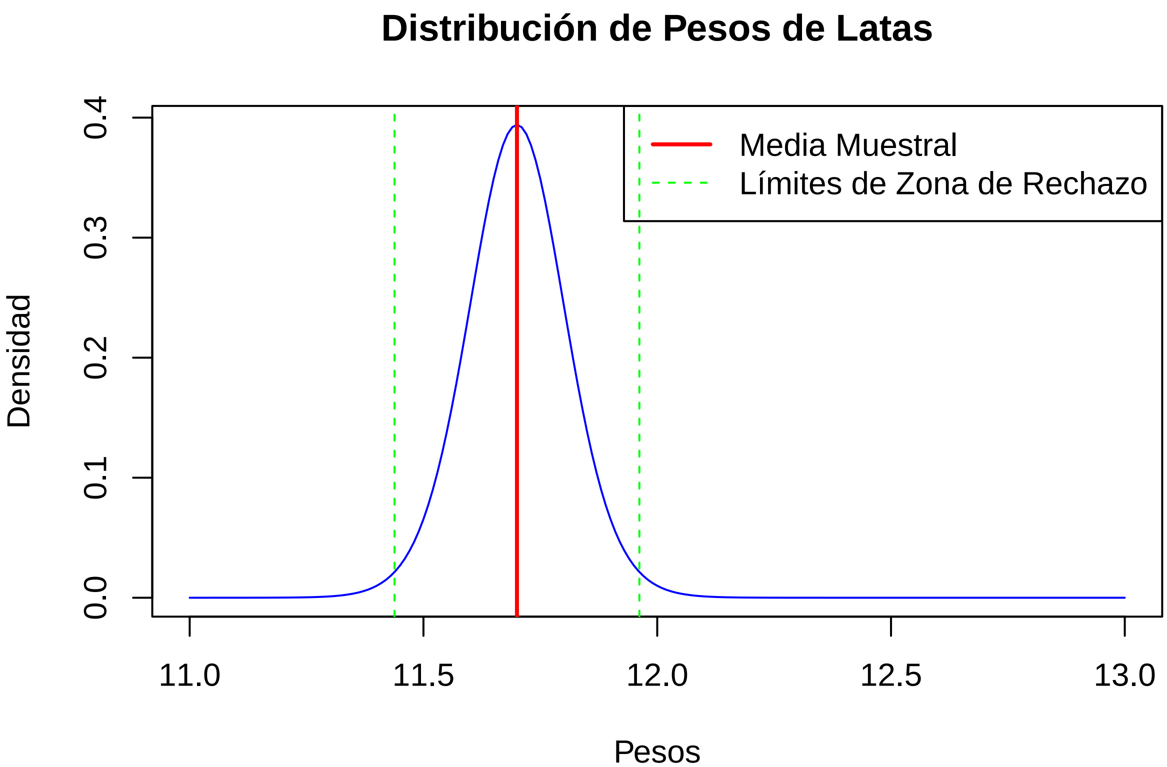
```
result <- t.test(x, alternative = "two.sided", mu = 11.7, conf.level = 0.98)
print(result$statistic)
```

```
##         t
## -2.068884
```

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

```
# Crear secuencia de valores para la gráfica
x <- seq(11, 13, by = 0.01)

# Calcular la densidad de la distribución t de Student
y <- dt((x - 11.7) / sm, df = n - 1)

# Graficar la distribución
plot(x, y, type = "l", col = "blue", xlab = "Pesos", ylab = "Densidad", main = "Distribución de Pesos d

# Agregar línea vertical en la media muestral
abline(v = 11.7, col = "red", lwd = 2)

# Agregar zonas de rechazo (derecha e izquierda)
abline(v = 11.7 - t0 * sm, col = "green", lty = 2)
abline(v = 11.7 + t0 * sm, col = "green", lty = 2)
```

```r
# Agregar leyenda
legend("topright", legend = c("Media Muestral", "Límites de Zona de Rechazo"), col = c("red", "green"),
```



1.3 Concluye en el contexto del problema.

En el contexto del problema esto significa que no hay evidencia estadística suficiente para afirmar que el peso promedio de las latas es diferente de 11.7. Es decir, no se puede concluir que el peso promedio sea mayor ni menor que 11.7.

2. Encuestas telefónicas

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 $\sigma = 4$ minutos. Usando un nivel de significación de 0.07, ¿está justificada la tarifa adicional?

2.1 Muestra tu procedimiento siguiendo los 4 pasos de solución.

Paso 1. Definir las hipótesis

Se establecen la hipótesis nula y la hipótesis alternativa:

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

El estadístico designado para esta prueba de hipótesis es $\bar{x}$, el cual se supondrá que sigue una distribución normal.

Por lo tanto, tenemos lo siguiente:

- $\mu_{\bar{x}} = 15$
- $\sigma_{\bar{x}} = 4$

Paso 2. Regla de decisión

El nivel de confianza con el que se trabajará es 0.93 y, por lo tanto: $\alpha = 0.07$.

```
x = 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, 1
alfa = 0.07
n  = length(x)
z0 = qnorm(alfa/2) # Valor frontera
cat("z0 =",z0)
```

```
## z0 = -1.811911
```

Como regla de decisión, H_0 se rechaza si:

- $|z^*| > 1.81$
- $p < 0.07$

Donde z^* es el número de desviaciones estándar al que $\bar{x}$ está lejos de μ.

Paso 3. Análisis de resultado

Se tiene que calcular:

- z^* (qué tan lejos está $\bar{x}$ de μ)
- Valor de p (la probabilidad de que $\bar{x}$ esté en las colas de la distribución)

```
# Calculo de z*
m = mean(x)
s = 4
sm = s/sqrt(n)
ze = (m-15)/sm
cat("z* =",ze,"\n")
```

```
## z* = 2.95804
```

```
# Calculo de valor p
valorp = 1 - pnorm(ze)
cat("Valor p =",valorp,"\n")
```

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

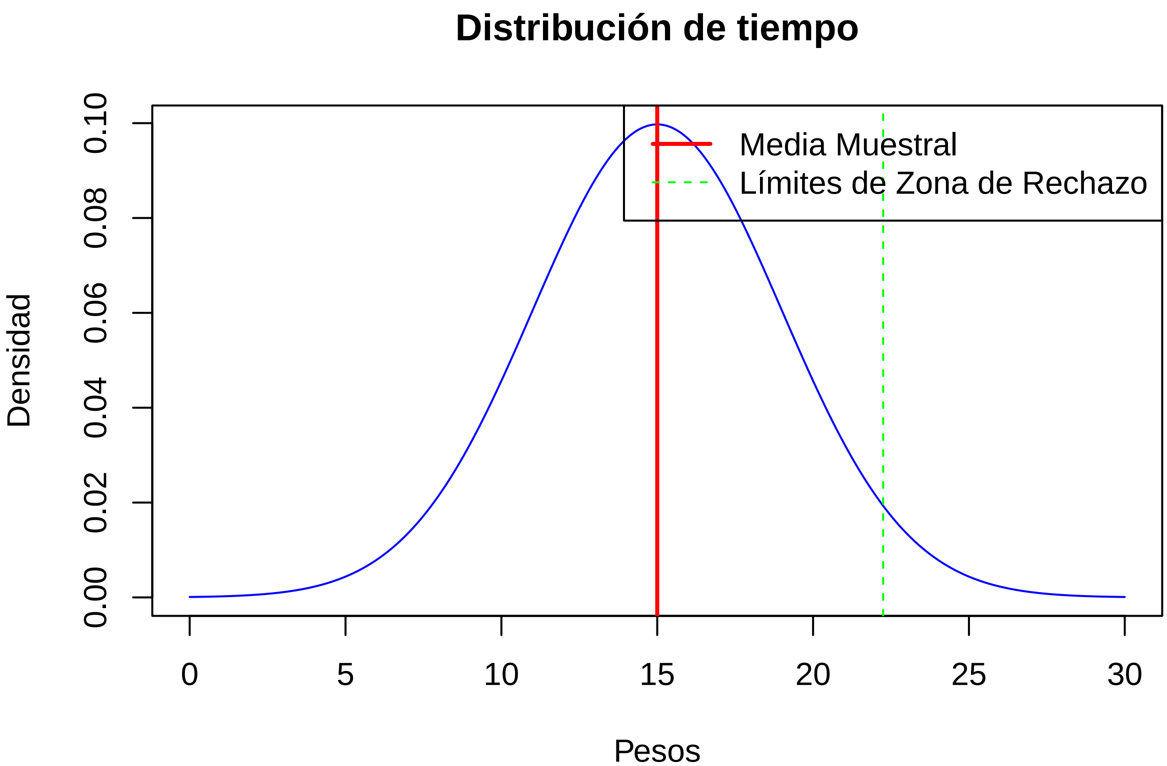
Paso 4. Conclusiones

Como $|z^*|$ (2.95804) es mayor que 1.81 y el valor p (0.00154801) es menor que 0.07, entonces H_0 se rechaza.

2.2 Grafica la regla de decisión y el valor del estadístico de prueba.

```
# Crear secuencia de valores para la gráfica
x <- seq(0, 30, by = 0.01)

# Calcular la densidad de la distribución normal
y <- dnorm(x, mean = 15, sd = 4)

# Graficar la distribución
plot(x, y, type = "l", col = "blue", xlab = "Pesos", ylab = "Densidad", main = "Distribución de tiempo")

# Agregar línea vertical en la media muestral
abline(v = 15, col = "red", lwd = 2)
```

```r
# Agregar zonas de rechazo (derecha e izquierda)
abline(v = 15 - z0 * 4, col = "green", lty = 2)

# Agregar leyenda
legend("topright", legend = c("Media Muestral", "Límites de Zona de Rechazo"), col = c("red", "green"),
```



2.3 Concluye en el contexto del problema.

En el contexto del problema esto significa que hay evidencia estadística suficiente para afirmar que el tiempo medio para completar las encuestas telefónicas es mayor a 15 minutos. Por lo tanto, no está justificada la tarifa adicional.