Pruebas de hipótesis

Diego Alberto Baños Lopez
A01275100

23-08-2023

Primer Problema - Enlatados

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.

Paso 1: Definir la hipotesis

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

Estadistico: $\bar{x}$

Distribucion del estadistico: t de Student

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

Paso 2: Regla de Decision

Nivel de Confianza $= 0.98$ $\alpha = 0.02$

```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
```

Paso 3: Calculos

t*: es el numero de desviaciones estandar al que $\bar{x}$ esta lejos de μ

H_0 se rechaza si:

- |t*| >2.53
- valor p <0.02

Tenemos que calcular:

- t* (que tan lejos esta $\bar{x}$ de μ)
- Valor p (la probabilidad de que $\bar{x}$ este en las colas de la distribucion)

*Calculo de t**

```
m <- mean(X)
s <- sd(X)
sm <- s / sqrt(n)

te <- (m - 11.7) / sm
cat("t* =", te)
```

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

Calculo de valor p

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

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

Paso 4: Conclusiones

- Como valor p (0.05173) es mayor que 0.02, entonces no RH_0
- Como |t*| (2.07) es menor que 2.53, entonces no RH_0

En el contexto del problema esto significa que la hipotesis se rechaza.

Mas facil:

```
t.test(X, alternative = "two.sided", mu = 11.7, 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
```

Segundo Problema - 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?

Paso 1: Definir la hipótesis

$H_0 : \mu = 15$ $H_1 : \mu \neq 15$

Estadistico: $\bar{x}$

Distribucion del estadistico: Normal

$\mu_{\bar{x}} = 15$, $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$

Paso 2: Regla de decisión

Nivel de significancia: $\alpha = 0.07$

```r
alfa <- 0.07
z0 <- qnorm(1 - alfa)
cat("z0 =", z0)
```

```
## z0 = 1.475791
```

Paso 3: Calculos

Necesitamos calcular el valor del estadistico de prueba z* en donde $z^* = \frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}}$

x es la media de la muestra.

μ es la media hipotética (en este caso, 15 minutos).

σ es la desviación estándar de la población (dada como 4 minutos).

n es el tamaño de la muestra.

```r
X2 <- 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
)
sigma <- 4
n2 <- length(X2)
m2 <- mean(X2)

z_estrella <- (m2 - 15) / (sigma / sqrt(n2))
cat("z* =", z_estrella)
```

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

Paso 4: Conclusiones

Aqui comparamos los resultados de z* con z_0 si z* es mayor se rechaza la hipotesis nula H_0

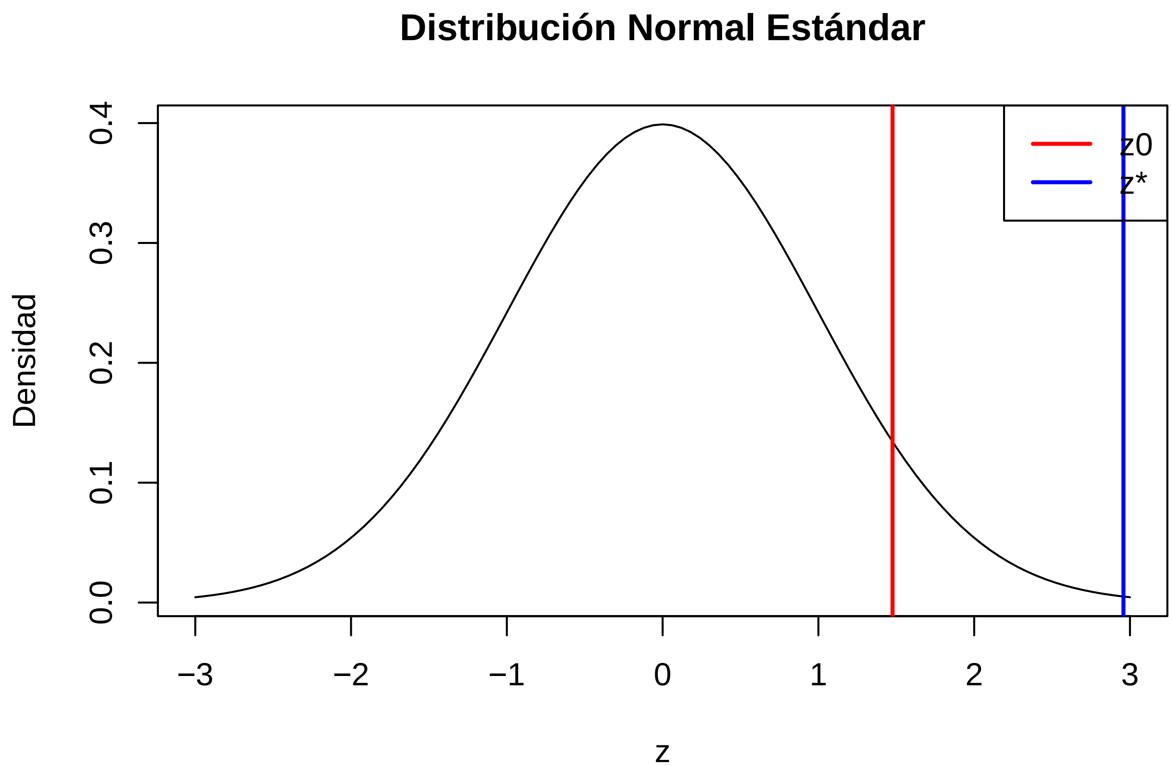
- Si se rechaza H_0 esi indica que hay evidencia para afirmar que el tiempo promedio de las llamadas es mayor a 15 min. justificando la tarifa adicional
- De lo contrario no hay evidencia para que el tiempo promedio sea mayor a 15 min. injustificando la tarifa adicional

```r
decision <- ifelse(z_estrella > z0, "Rechazar", "No Rechazar")
cat("Decisión:", decision)
```

```
## Decisión: Rechazar
```

Graficas

```r
curve(dnorm(x), -3, 3,
    main = "Distribución Normal Estándar",
    xlab = "z", ylab = "Densidad"
)
abline(v = z0, col = "red", lwd = 2)
abline(v = z_estrella, col = "blue", lwd = 2)
legend("topright", legend = c("z0", "z*"), col = c("red", "blue"), lwd = 2)
```



Distribución Normal Estándar
z0
z*
Densidad
0.0
0.1
0.2
0.3
0.4
−3
−2
−1
0
1
2
3
z