Pruebas de Hipotesis

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.

Paso 1: Hipotesis

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

Como se dsitribuye $\bar{X}$

X se distribuye como una normal n $<$ 30 No conocemos sigma

Entonces: la distribucion muestral es una t de Student

Paso 2: Regla de decision

Nivel de confianza es de 0.98 Nivel de significancia es 0.02

Necesito encontrar a cuantas desviaciones estandar esta lejos el el valor frontera.

```
datos = 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,

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

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

Rechazo H_0 si:

$|t_e| > 2.53$ valorp < 0.02

Paso 3: Analisis de resultado

t_e: Numero de desvaciones estandar al que $\bar{X}$ se encuentra lejos de $\mu = 11.7$ Valor p: Probabilidad de obtener loq ue obtuve en la muestra o un valor mas extremo

Estadistico de prueba

```r
mu = 11.7
xbar = mean(datos)
s = sd(datos)

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

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

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

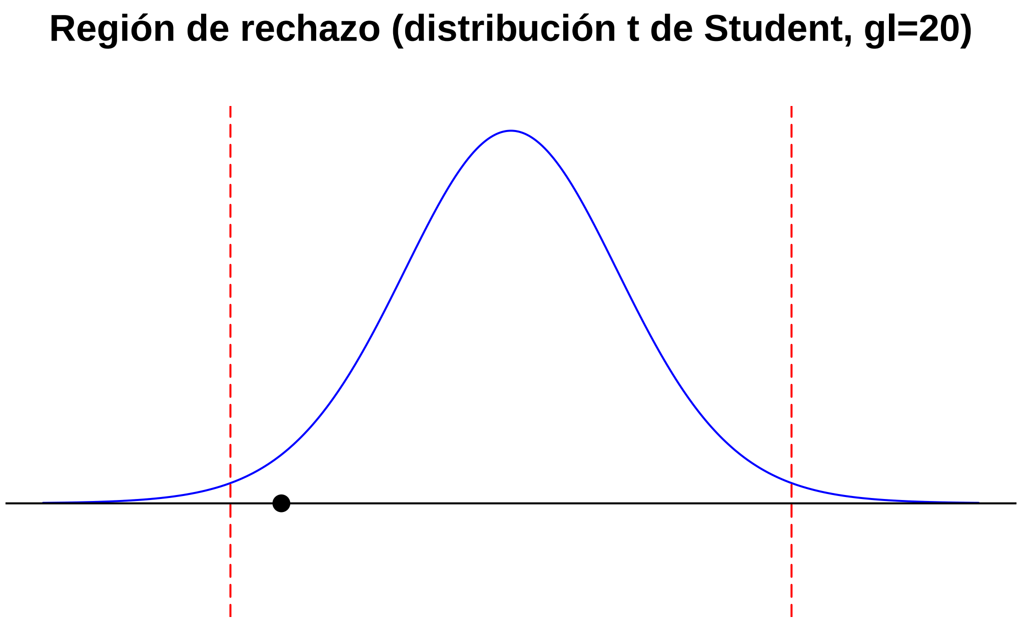
```
## El valor p = 0.0517299
```

Manera mas facil

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

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

```r
sigma = sqrt((n-1)/(n-3))
x=seq(-4*sigma,4*sigma,0.01)
y=dt(x,n-1)
plot(x,y,type="l",col="blue",xlab="",ylab="",ylim=c(-0.1,0.4),frame.plot=FALSE,xaxt="n",yaxt="n",main="
abline(v=t_f,col="red",lty=5)
abline(v=-1*t_f,col="red",lty=5)
abline(h=0)
points(mu,0,col="blue",pch=19)
points(te, 0, pch=19, cex=1.1)
```



El valorp es mayor a α entonces no se tiene la suficiente evidencia para rechazar

Paso 4: Conclusion

Comparar: Regla de decision vs analisis del resultado

Rechazo H_0 si:

$|t_e| = 2.07 < 2.53$ -> No RH0 valor p $= 0.05 > 0.02$ -> No RH0

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

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

```
datos = c(17, 11, 12, 23, 20, 23, 15, 16, 23, 22, 18, 23, 25, 14, 12, 12, 20, 18, 12, 19, 11, 11, 20, 2

n = length(datos)
alfa = 0.07
z_f = abs(qnorm(alfa))
cat("z_f =", z_f)
```

```
## z_f = 1.475791
```

```
mu = 15
sigma = 4
xbar = mean(datos)
```

```r
z_e = (xbar-mu)/(sigma/sqrt(n))
cat("z_e =", z_e, "\n")
```

```
## z_e = 2.95804
```

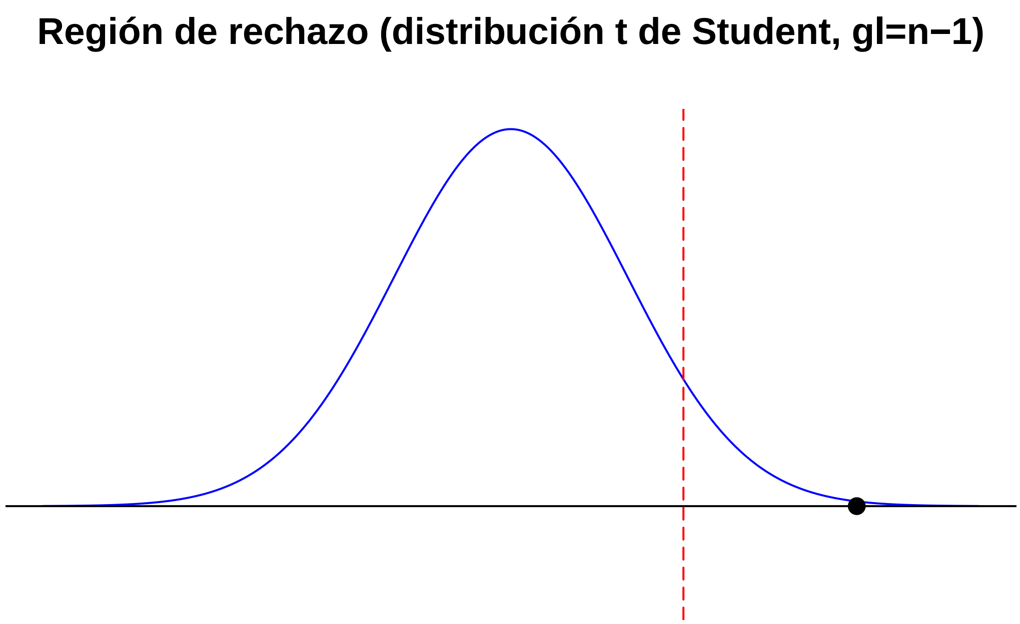
```r
valorp = 1-pnorm(z_e)
cat("El valor p =", valorp, "\n")
```

```
## El valor p = 0.00154801
```

```r
inf <- xbar - z_f * (sigma / sqrt(n))
cat("El valor Inf =", inf)
```

```
## El valor Inf = 16.00218
```

```r
sigma =1
x=seq(-4*sigma,4*sigma,0.01)
y=dnorm(x)
plot(x,y,type="l",col="blue",xlab="",ylab="",ylim=c(-0.1,0.4),frame.plot=FALSE,xaxt="n",yaxt="n",main="
abline(v=z_f,col="red",lty=5)
abline(h=0)
points(mu,0,col="blue",pch=19)
points(z_e, 0, pch=19, cex=1.1)
```



Viendo estos resultados vemos que la llamada promedio es mayor a 15 minutos y la tarifa adicional es justificada.