

Comparación de medias

Dr. Marco Aurelio González Tagle

08 /09/ 2021

Índice

Procedimiento general para la pruebas de t una muestra	2
Ejercicio 1	3
Ejercicio 2	3
Ejercicio 3	3

Una empresa de alimento para venados esta interesado en determinar si el peso neto medio del contenido de sus costales es de 80 kg como se anuncia en su producto. Digamos que un consumidor al paso del tiempo ha comprado y pesado de forma precisa el contenido neto de 44 costales de 80 kg seleccionados al azar. El consumidor reclama a la empresa que el contenido de sus costales **ha sido menor** porque sus datos no provienen de una distribución con media de 80 kg. Para investigar esta demanda, la empresa plantea una hipótesis usando el grado de confiabilidad de $\alpha = 0.05$.



La hipótesis nula será: No existen diferencias entre la media observada y la declarada en sus costales (80 kg).

La hipótesis alternativa se declara como: La media observada es menor a la declarada en los costales de 80 kg.

$$H_0 : \mu = 80 \tag{1}$$

$$H_1 : \mu < 80 \quad (2)$$

Procedimiento general para la pruebas de t una muestra

Ingresar los datos

```
costal <- c(87.7, 80.01, 77.28, 78.76, 81.52, 74.2, 80.71, 79.5, 77.87, 81.94, 80.7,
            82.32, 75.78, 80.19, 83.91, 79.4, 77.52, 77.62, 81.4, 74.89, 82.95,
            73.59, 77.92, 77.18, 79.83, 81.23, 79.28, 78.44, 79.01, 80.47, 76.23,
            78.89, 77.14, 69.94, 78.54, 79.7, 82.45, 77.29, 75.52, 77.21, 75.99,
            81.94, 80.41, 77.7)
```

Explorar los datos

La media y la desviación estándar deben ser estimados de la muestra

```
# Determinar el número de observaciones
n <- length(costal)
n
```

```
## [1] 44
```

```
# Determinar la media
costa.media <- mean(costal)
costa.media
```

```
## [1] 78.91068
```

```
# Desviación estándar
costa.sd <- sd(costal)
costa.sd
```

```
## [1] 3.056023
```

Necesitamos responder la siguiente pregunta: Dada la desviación estándar (3.056), ¿Cuál es la probabilidad de observar una media de la muestra (cuando $n = 44$) de 78.91 Kg o de menor cantidad si la media verdadera es de 80 kg.

La prueba de T de una muestra puede ser resulta con la siguiente fórmula:

$$T = \frac{\bar{x} - \mu_0}{(s/\sqrt{n})} \quad (3)$$

El denominador se le conoce como error estándar de la media (se).

```
costa.se <- costa.sd/ sqrt(n)
costa.se
```

```
## [1] 0.4607128
```

Entonces podemos calcular el valor de T

```
costa.T <- (costa.media - 80)/ costa.se
costa.T
```

```
## [1] -2.364419
```

Finalmente, el valor de p puede ser calculado.

```
pt(costa.T, df = n-1) # Arriba se declaro "n" como número de costales observados
```

```
## [1] 0.01132175
```

Ejercicio 1

Resuelva en R

El resultado de la prueba de t de una muestra del procedimiento anterior puede ser verificado con la fórmula ya establecida en R.

- ¿Cuál es el valor de p?
- ¿Cuántos grados de libertad tiene el experimento?
- ¿Cuál es la hipótesis aceptada?
- ¿Existe evidencia de que el valor medio promedio de los costales observados es menor (significativamente) a los que anuncia el producto?

Ejercicio 2

En el siguiente cuadro se registraron los datos de emisiones de óxido de azufre en toneladas al año de una planta industrial. De acuerdo con la reglamentación de la empresa, la cantidad máxima de toneladas por año emitida no debe superar las 17.5 Ton/año.

```
azufre <- c(15.8, 22.7, 26.8, 19.1, 18.5, 14.4, 8.3, 25.9, 26.4, 9.8,
            22.7, 15.2, 23.0, 29.6, 21.9, 10.5, 17.3, 6.2, 18.0, 22.9,
            24.6, 19.4, 12.3, 15.9, 11.2, 14.7, 20.5, 26.6, 20.1, 17.0,
            22.3, 27.5, 23.9, 17.5, 11.0, 20.4, 16.2, 20.8, 13.3, 18.1)
```

Resuelva:

Encuentre si el valor de emisiones de óxido de azufre registradas son significativamente mayores a la establecida por la empresa (17.5 Ton/ha).

- ¿Cuál es el valor de p?
- ¿Cuáles son los intervalos de confianza al 95 % ?
- ¿Cuántos grados de libertad tiene el experimento?
- ¿Cuál es la hipótesis aceptada?
- ¿Existe evidencia de que el valor medio promedio de las emisiones observadas es mayor (significativamente) a la declarada en los procedimientos de seguridad de la empresa?

Ejercicio 3

De acuerdo con la plataforma de datos del Gobierno de México. Descargar la Temperatura actual diaria de las estaciones climatológicas convencionales contenidas en el Sistema de Información Hidrológica (SIH) almacenados en GitHub: https://raw.githubusercontent.com/mgtagle/MCF-202_Agosto_2021/main/TEMPAIRE_DIA.csv

```
##   estacionID      fecha temp_media tmax tmin
## 1    AGSAG 08/09/21 0:00      20.5 26.0 15.0
## 2    ALMAG 08/09/21 0:00      15.5 21.0 10.0
## 3    ANVAG 08/09/21 0:00      19.5 25.0 14.0
## 4    CDRAG 08/09/21 0:00      19.0 23.5 14.5
## 5    CNSAG 08/09/21 0:00      20.0 25.5 14.5
## 6    CSOAG 08/09/21 0:00      14.0 20.0  8.0
```

Resuelva:

Encuentre si el valor promedio de la temperatura (`temp_media`) registradas es significativamente mayor a la establecida (24 Grados).

- ¿Cuál es el valor de p ?
- ¿Cuáles son los intervalos de confianza al 95 % ?
- ¿Cuántos grados de libertad tiene el experimento?
- ¿Cuál es la hipótesis aceptada?
- ¿Existe evidencia de que el valor medio promedio de las emisiones observadas es mayor (significativamente) a la declarada en los procedimientos de seguridad de la empresa?