

Métodos de diseño y Análisis de Experimentos

Tarea 02

Rivera Torres Francisco de Jesús

Rodríguez Maya Jorge Daniel

Samayoa Donado Víctor Augusto

Trujillo Barrios Georgina

Marzo 05, 2019

1 Ejercicio 1

Se sabe que la pagina web de una famosa tienda departamental tiene un tiempo de carga en segundos que se distribuye Normal ($\mu = 5, \sigma^2 = 4$)

- Calcula y grafica la probabilidad de que la página tarde entre 4 y 6 segundos en cargar.
- Calcula y grafica la probabilidad de que la página cargue máximo en 4 segundos.
- Calcula y grafica la probabilidad de que la página tarde 6 segundos o más en cargar.
- Calcula y grafica el mínimo que tarda en cargar la página el 5% de las veces que carga mas lenta.
- Calcula y grafica cuál es el máximo que tarda en cargar la página el 10% de las veces que lo hace más rápido.

2 Ejercicio 2

Teniendo T una variable aleatoria t-student centrada en 0 y con 13gl.

- Calcula y grafica la probabilidad de que T este en el intervalo $[-1, 1]$
- Calcula y grafica la probabilidad de que T sea menor o igual a 2.16
- Calcula y grafica la probabilidad de que T sea igual o mayor a 2.16
- Calcula y grafica la probabilidad de que T sea menor o igual a -2.16
- Calcula y grafica la probabilidad de que T sea igual o mayor a -2.16

3 Ejercicio 3

Teniendo una variable aleatoria Q que se distribuye F con parámetros (4, 20)

- Calcula y grafica la probabilidad de que Q este en el intervalo $[3, 4]$
- Calcula y grafica la probabilidad de que Q sea 3 o menor
- Calcula y grafica la probabilidad de que Q sea 4 o mayor
- Calcula y grafica el percentil 0.95 de Q
- Calcula y grafica el percentil 0.5 de Q

4 Ejercicio 4

Una franquicia quiere determinar si existe diferencia entre la satisfacción de los clientes en los establecimientos de dos de sus franquiciatarios y para ello recolecta datos en 15 establecimientos de cada uno:

Tabla 1: Promedio de satisfacción de los clientes por establecimiento.

Franquiciatario 1	Franquiciatario 2
6.721351	8.3162646
6.323979	2.8591867
4.128115	12.9495849
9.593806	5.5420510
11.176376	3.8361638
5.460104	1.1963828
2.517744	4.8126178
9.186292	1.8920791
4.235253	6.1332265
8.824826	10.1599013
5.568107	5.9033151
6.794284	0.5051285
5.670497	0.4907579
8.418545	4.6517146
5.995717	7.0236920

Nota:

Donde 0 es completamente insatisfecho y 15 completamente satisfecho

- Escribe la hipótesis nula y la hipótesis alternativa

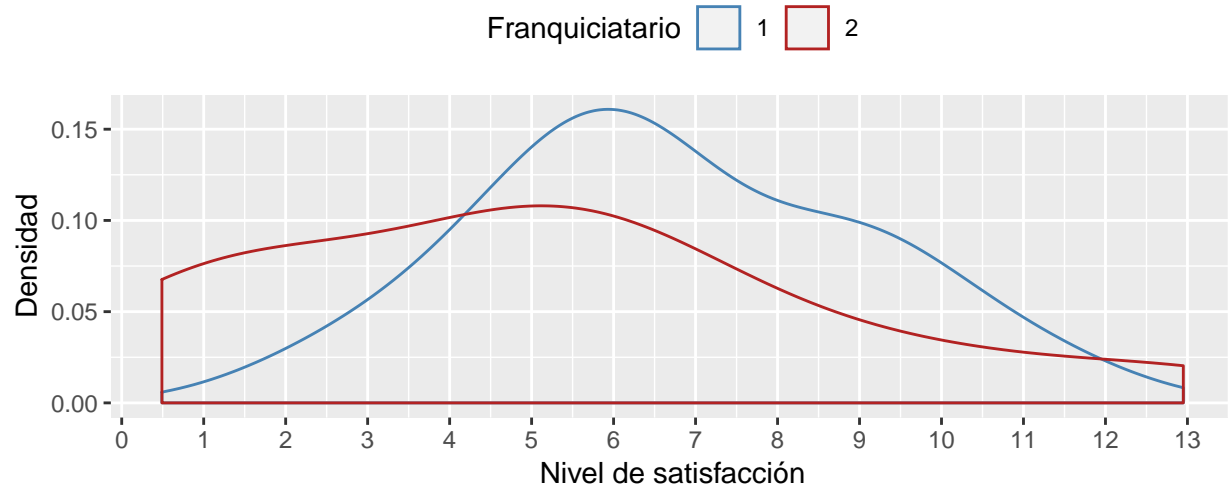
Consideremos como μ_i la media de satisfacción de los clientes en los establecimientos del franquiciatario i , $i = 1, 2$. Entonces la hipótesis nula (H_0) y la hipótesis alternativa (H_a) quedan como sigue:

$$H_0 : \mu_1 - \mu_2 = 0 \quad \text{v.s.} \quad H_a : \mu_1 - \mu_2 \neq 0$$

- Realiza la prueba de hipótesis correspondiente

Procedemos a graficar las distribuciones muestrales de los datos para tener una noción sobre que tan “diferentes” pueden llegar a ser.

Distribución muestral de niveles de satisfacción en los establecimientos



Primero, se procede a realizar una prueba de hipótesis para la igualdad de varianzas (ya que necesitamos saber si la comparación de las medias se realizará con varianzas desconocidas iguales o distintas),

El planteamiento de hipótesis para igualdad de varianzas está dado por:

$$H_0 : \frac{\sigma_y^2}{\sigma_x^2} = 1 \quad \text{v.s.} \quad H_a : \frac{\sigma_y^2}{\sigma_x^2} \neq 1$$

donde la región de rechazo está dada por

$$C = \left\{ x \in X \left| \left(\frac{\frac{1}{n_1} \sum_{i=1}^{n_1} (y_i - \mu_y)^2}{\frac{1}{n_2} \sum_{i=1}^{n_2} (x_i - \mu_x)^2} \right) > F_{(n_1-1, n_2-1)}^{1-\alpha/2} \right. \right\}$$

donde $\alpha = 0.1$ para una confianza del 90% y $n_1 = n_2 = 15$.

Calculando el estadístico se tiene que .

```
alpha <- 0.1
```

```
n1 <- nrow(datos)
```

```
n2 <- nrow(datos)
```

```
est.f <- qf(1 - alpha, df1 = n1 - 1, df2 = n2 - 1)
```

obteniendo así un valor de $F_{(n_1-1, n_2-1)}^{1-\alpha/2} = 2.0224339$.

Realizando los cálculos para la región de rechazo, se obtiene que

```
library(tidyverse)

x <- datos %>%
  pull(f1)
y <- datos %>%
  pull(f2)

var_f1 <- sum((y - mean(y))^2)/n1
var_f2 <- sum((x - mean(x))^2)/n2

f <- var_f1/var_f2
```

$$\left(\frac{\frac{1}{n_1} \sum_{i=1}^{n_1} (y_i - \mu_y)^2}{\frac{1}{n_2} \sum_{i=1}^{n_2} (x_i - \mu_x)^2} \right) = 2.3181465$$

En este caso observamos que:

$$2.3181465 = \left(\frac{\frac{1}{n_1} \sum_{i=1}^{n_1} (y_i - \mu_y)^2}{\frac{1}{n_2} \sum_{i=1}^{n_2} (x_i - \mu_x)^2} \right) > F_{(n_1-1, n_2-1)}^{1-\alpha/2} = 2.0224339$$

por lo tanto se rechaza la hipótesis nula de que ambas poblaciones tienen varianza igual. Es decir, consideramos que las poblaciones (establecimientos de franquiciatario 1 y establecimientos de franquiciatario 2) tienen varianza distinta.

Con lo anterior, se procede a realizar una prueba de hipótesis bilateral (de dos colas), para las medias de ambas poblaciones (establecimientos de franquiciatario 1 y establecimientos de franquiciatario 2).

```
test4 <- t.test(x, y, alternative = "two.sided", paired = FALSE,
  conf.level = 0.9, var.equal = FALSE, conf.int = TRUE)
```

```
test4

##
## Welch Two Sample t-test
##
## data: x and y
## t = 1.4735, df = 24.184, p-value = 0.1535
## alternative hypothesis: true difference in means is not equal to 0
## 90 percent confidence interval:
## -0.2608457 3.5065696
## sample estimates:
## mean of x mean of y
## 6.707666 5.084804
```

En este caso observamos que $p\text{-value} > 0.1 = \alpha$

- Construye el intervalo de confianza para la diferencia de medias usando un nivel de confianza de 90%

El intervalo de confianza está dado por:

```
test4$conf.int
```

```
## [1] -0.2608457  3.5065696  
## attr(,"conf.level")  
## [1] 0.9
```

En este caso observamos que $0 \in (-0.26, 3.51)$

- Concluye

Lo anterior nos indica que, con una cofianza del 90%, podemos afirmar que el nivel de satisfacción, en promedio, que proporcionan los establecimientos de ambos franquiciatarios es el mismo.