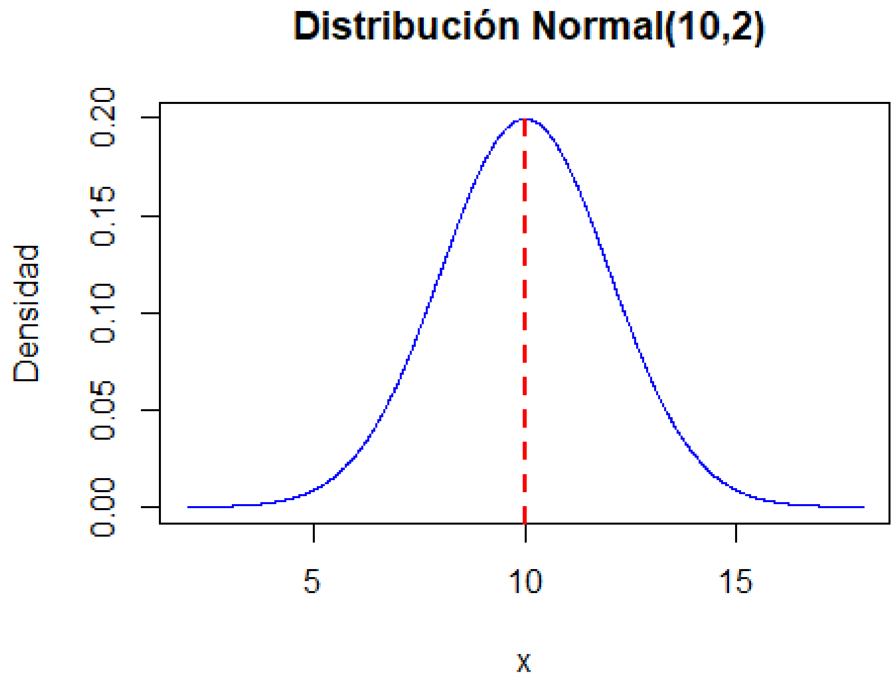
Estadistica

Catherine Rojas

2024-08-08

1. Graficar una distribución Normal con media μ = 10, y desviación estándar σ = 2

```
# Definir los parámetros
miu = 10
sigma = 2

# Generar una secuencia de valores de x
x = seq(miu - 4*sigma, miu +4*sigma, 0.01)

# Calcular la densidad de la distribución normal para cada valor
y = dnorm(x, miu, sigma)

# Graficar la distribución normal
plot(x, y, type = "l", col = "blue", main = " Distribución Normal(10,2)",
xlab = "x", ylab = "Densidad")

# Agregar una línea vertical en la media
abline(v = miu, col = "red", lwd =2, lty =2)
```



Observaciones:

Esta gráfica visualiza la distribución normal con una media de 10 y una desviación estandar de 2. Es una gráfica simétrica respecto a la línea vertical que indica la media.

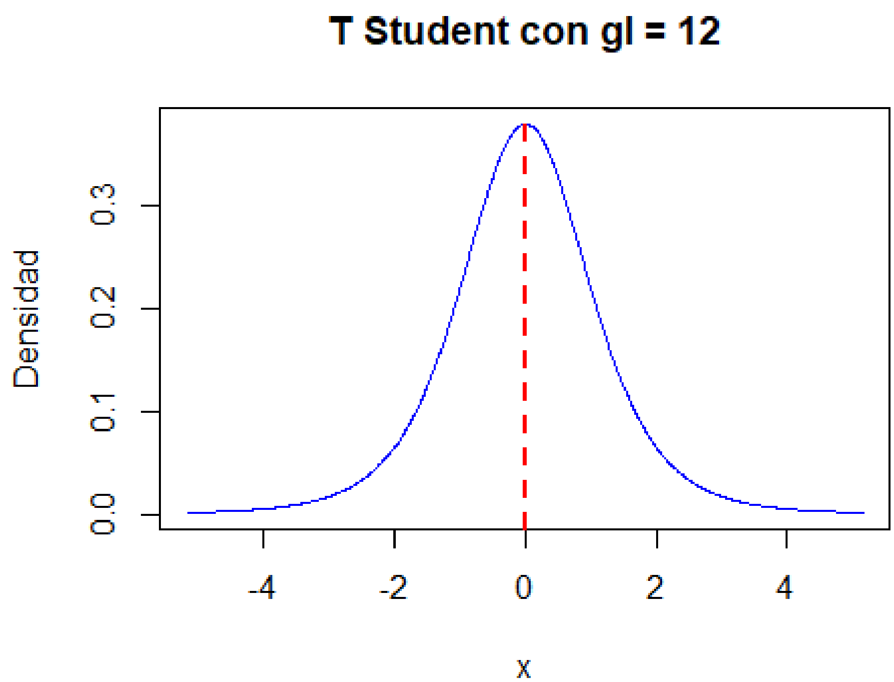
Cualquier área bajo la curva entre dos puntos específicos representa la probabilidad de que la variable aleatoria caiga dentro de ese rango.

Los valores que están muy alejados de la media tienen una probabilidad baja de ocurrir, ya que las colas de la distribución son muy bajas.

2. Graficar una distribución T Student con grados de libertad v = 12

```r
# Definir los grados de libertad
gl = 5

# Calcular el valor de sigma basado en los grados de libertad
sigma = sqrt(gl / (gl - 2))

# Generar una secuncia de valores x
x = seq(-4 *sigma, 4 * sigma, 0.01)

# Calcular la densidad de la distribución T-Student para cada valor de x

y = dt(x, gl)
```

```r
# Graficar la distribución T-Student
plot(x, y, type = "l", col = "blue",  main = "T Student con gl = 12",
xlab = "x", ylab = "Densidad")

# Agregar una línea vertical en la media (cero en T-Student)
abline(v = 0, col = "red", lwd = 2, lty =2)
```



Observaciones:

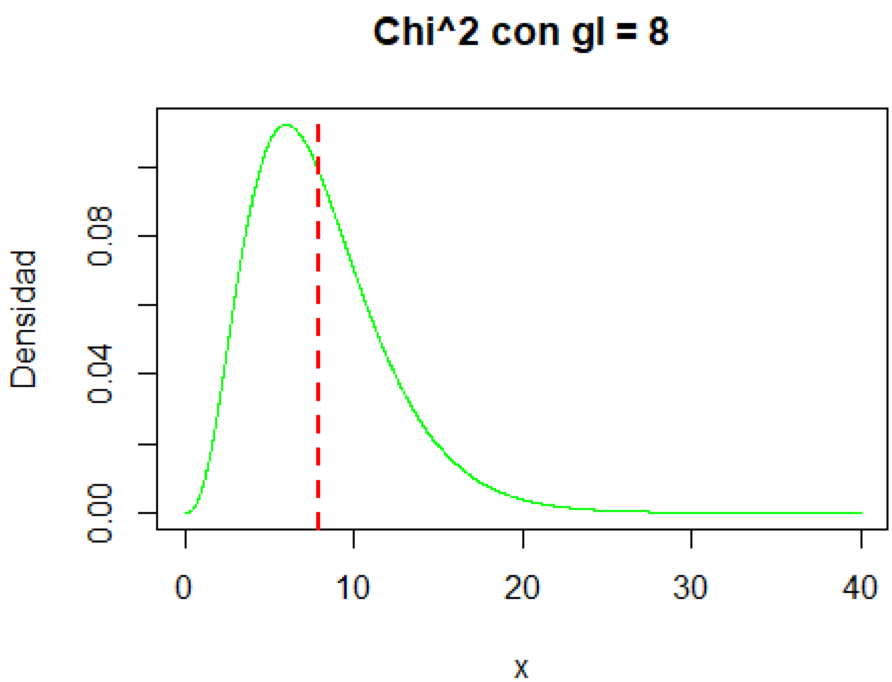
La gráfica representa una distribución t Student con 12 grados de libertad. La distribución t de Student es simétrica alrededor de su media, que es 0.

Los valores que están muy alejados de la media tienen una probabilidad baja de ocurrir, pero esta probabilidad es mayor en comparación con la distribución normal debido a que esta tiene colas más gruesas.

3. Gráfique la distribución Chi-cuadrada con 8 grados de libertad.

```r
# Definir los grados de libertad
gl = 8

# Calcular el valor sigma basado en los grados de libertad
sigma = sqrt(2 *gl)

#Generar una secuencia de valores de x
x =seq(0, gl +8 * sigma, 0.01)
```

```r
# Calcular la densidad de la distribución Cho-cuadrada para cada valor de x
y = dchisq(x,gl)

# Graficar la distribución Chi-cuadrada
plot(x, y, type = "l", col = "green", main = "Chi^2 con gl = 8", xlab = "x", ylab = "Densidad")

# Agregar una línea vertical en la media (gl en Chi-cuadrada)

abline(v= gl, col = "red", lwd = 2, lty =2)
```



Observaciones:

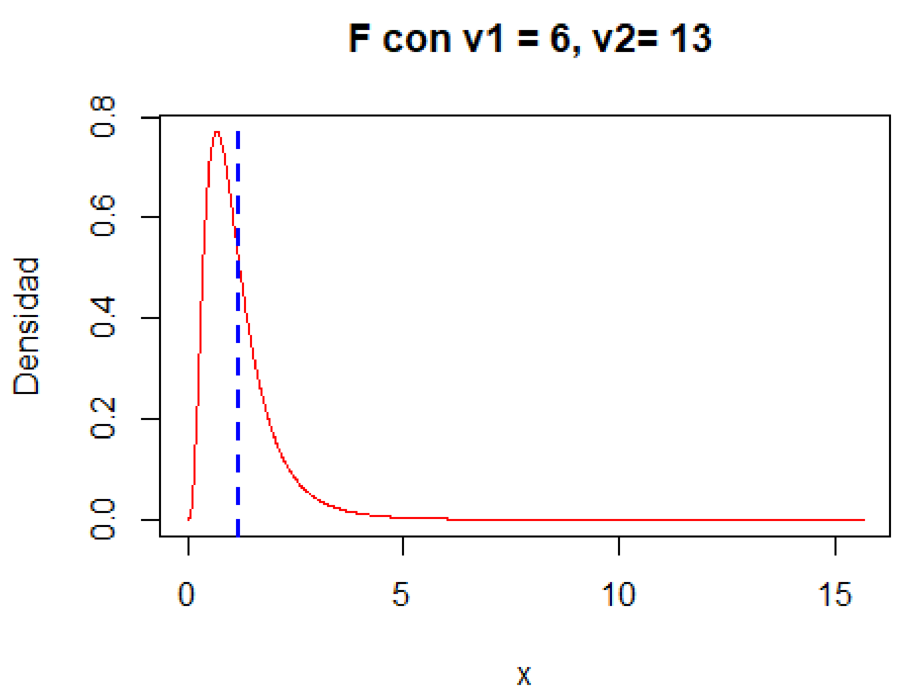
La gráfica representa la distribución chi-cuadrada con 8 grados de libertad. Esta distribución es asimétrica a la derecha, es decir, exiate una mayor concentración de probabilidades en los valores más bajos, y la cola se extiende hacia valores más altos.

La línea vertical roja indica el valor esperado o la media de la distribución chi-cuadrada, que es igual a los grados de libertad (df=8).

El pico de la gráfica se encuentra a la izquierda y representa el valor con la mayor densidad de probabilidad. La mayor densidad se encuentra cerca del valor de la media (8) y disminuye a medida que nos alejamos de este valor.

4. Graficar una distribución F con v1 = 9, v2 = 13

```r
# Definir los grados de libertad

v1 = 9
v2 = 13

# Calcular el valor sigma basado en los grados de la libertad

sigma = sqrt(2) * v2 * sqrt(v2 + v1 - 2) / (sqrt(v2 - 4) * (v2 -2) *
sqrt(v1))

# Generar un secuencia de valores x
x = seq(0, v1 +8 *sigma, 0.01)

# Calcular la densidad de la distribución F para cada valor de x
y = df(x, v1, v2)

# Gráficar la distribución de F
plot(x, y, type= "l", col= "red", main = " F con v1 = 6, v2= 13", xlab =
"x", ylab = "Densidad")

# Agregar una línea vertical en el valor esperado
abline(v = (v2 / (v2 - 2)), col = "blue", lwd = 2, lty = 2)
```



La distribución F es asimétrica y tiene una cola larga hacia la derecha. Esto significa que hay una mayor concentración de probabilidades en los valores más bajos, y la probabilidad disminuye lentamente a medida que los valores aumentan.

Los valores más alejados del pico tienen una menor probabilidad de ocurrir. Sin embargo, debido a la cola larga, hay una pequeña probabilidad de observar valores extremos.

5. Si Z es una variable aleatoria que se distribuye normalmente con media 0 y desviación estándar 1, hallar los procedimientos de:

a) P(Z > 0.7) = 0.2419637
b) P(Z < 0.7) = 0.7580363
c) P(Z = 0.7) = 0
d) Hallar el valor de Z que tiene al 45% de los demás valores inferiores a ese valor.

Utilice la función pnorm, por ejemplo P(Z < 2.1) = pnorm(2.1) Cuando lo que se quiere es hallar el valor de Z dada el área a la izquierda bajo la curva se usa qnorm(área izq).

a) P(Z > 0.7) = 0.2419637

```
# Calcular P(Z > 0.7)

prob_z_mayor_0.7 = 1 - pnorm(0.7)
prob_z_mayor_0.7

## [1] 0.2419637
```

b) P(Z < 0.7) = 0.7580363

```
# Calcular P(Z < 0.7)

prob_z_menor_0.7 = pnorm(0.7)
prob_z_menor_0.7

## [1] 0.7580363
```

c) P(Z = 0.7) = 0

```
# Calcular P(Z = 0.7)
prob_z_igual_0.7 = pnorm(0.7) - pnorm(0.7)
prob_z_igual_0.7

## [1] 0
```

d) Hallar el valor de Z que tiene al 45% de los demás valores inferiores a ese valor.

```r
# Calcular el valor de Z que tiene al 45% de los valores inferiores
z_45_percentil = qnorm(0.45)
z_45_percentil

## [1] -0.1256613
```

6. Hallar el procedimiento para verificar los siguientes resultados si se sabe que X se distribuye normalmente con una media de 100 y desviación estándar de 7.

a) P(X < 87) = 0.031645
b) P(X > 87) = 0.968354
c) P(87 < X < 110) = 0.89179 Utilice la función pnorm(x, miu, sigma) de R

a) P(X < 87) = 0.031645

```r
# Parámetros de la distribución
miu = 100
sigma = 7

# Calcular P(X < 87)
prob_x_menor_87 = pnorm(87, mean = miu, sd = sigma)
prob_x_menor_87

## [1] 0.03164542
```

b) P(X > 87) = 0.968354 Para calcular P(X > 87),se puede utilizar el complemento de la probabilidad de P(X < 87):

```r
# Calcular P(X > 87)
prob_x_mayor_87 = 1 - pnorm(87, mean = miu, sd = sigma)
prob_x_mayor_87

## [1] 0.9683546
```

c) P(87 < X < 110) = 0.89179 Para calcular P(87 < X < 110), se calcula la diferencia entre P(X < 110) y P(X < 87):

```r
# Calcular P(87 < X < 110)
prob_x_entre_87_110 = pnorm(110, mean = miu, sd = sigma) - pnorm(87, mean
= miu, sd = sigma)

prob_x_entre_87_110

## [1] 0.8917909
```

7. Hallar el procedimiento para verificar los siguientes resultados si se sabe que X se distribuye T Student con gl= 10, hallar:

a) P(X < 0.5) = 0.6860532
b) P(X > 1.5) = 0.082253
c) La t que sólo el 5% son inferiores a ella. (t = -1.812461) Utilice pt(x, gl) y qt(área izq, gl)

a) P(X <0.5) = 0.6860532

```
# Grados de libertad
gl = 10

# Calcular P(X < 0.5)
prob_x_menor_0.5 = pt(0.5, df =gl)
prob_x_menor_0.5

## [1] 0.6860532
```

b) P(X > 1.5) = 0.082253 Para calcular P(X > 1.5), se puede utilizar el complemento de la probabilidad de P(X < 1.5)

```
# Calcular P(X > 1.5)
prob_x_mayor_1.5 = 1 - pt(1.5, df = gl)
prob_x_mayor_1.5

## [1] 0.08225366
```

c) La t que sólo el 5% son inferiores a ella. (t = -1.812461) Para encontrar el valor de t tal que solo el 5% de los valores son inferiores a este valor, se utiliza la función qt:

```
# Calcular la t que deja el 5% de los valores a su izquierda
t_5_percentil = qt(0.05, df = gl)
t_5_percentil

## [1] -1.812461
```

8. Hallar el procedimiento para verificar los siguientes resultados si se sabe que X se distribuye Chi-cuadrada con gl = 6, hallar

a) P(X2 < 3) = 0.1911532
b) P(X2 > 2) = 0.9196986
c) El valor x de chi que sólo el 5% de los demás valores de x es mayor a ese valor = 12.59159 Utilice pchisq(x, gl) y qchisq(área izq., gl)

a) P(X2 < 3) = 0.1911532

```
# Grados de libertad
gl = 6
```

```r
# Calcular P(X^2 < 3)
prob_x2_menor_3 = pchisq(3, df = gl)
prob_x2_menor_3

## [1] 0.1911532
```

b) P(X2 > 2) = 0.9196986

```r
# Calcular P(X² > 2)
prob_x2_mayor_2 = 1 - pchisq(2, df = gl)
prob_x2_mayor_2

## [1] 0.9196986
```

c) El valor x de chi que sólo el 5% de los demás valores de x es mayor a ese valor = 12.59159 Para encontrar el valor de chi-cuadrada tal que solo el 5% de los valores son mayores a este valor, se utiliza la función qchisq con el complemento del área (1 - 0.05 = 0.95):

```r
# Calcular el valor de chi-cuadrada que deja el 5% de los valores a su derecha
chi_5_percentil = qchisq(0.95, df = gl)
chi_5_percentil

## [1] 12.59159
```

9. Hallar el procedimiento para verificar los siguientes resultados si se sabe que X se distribuye F con v1 = 8, v2 = 10, hallar

a) P(X < 2) = 0.8492264
b) P(X > 3) = 0.05351256
c) El valor de x que sólo el 25% de los demás valores es inferior a él. = 0.6131229)

a) P(X < 2) = 0.8492264

```r
# Grados de libertad
v1 = 8
v2 = 10

# Calcular P(X < 2)
prob_x_menor_2 = pf(2, df1 = v1, df2 = v2)
prob_x_menor_2

## [1] 0.8492264
```

b) P(X > 3) = 0.05351256 Para calcular P(X>3), se puede utilizar el complemento de la probabilidad de P(X<3):

```r
# Calcular P(X > 3)
prob_x_mayor_3 = 1 - pf(3, df1 = v1, df2 = v2)
prob_x_mayor_3

## [1] 0.05351256
```

c) El valor de x que sólo el 25% de los demás valores es inferior a él. = 0.6131229) Para encontrar el valor de x tal que solo el 25% de los valores son inferiores a este valor, se utiliza la función qf:

```r
# Calcular el valor de F que deja el 25% de los valores a su izquierda
f_25_percentil = qf(0.25, df1 = v1, df2 = v2)
f_25_percentil

## [1] 0.6131229
```

10. Resolver el siguiente problema:

Una compañía de reparación de fotocopiadoras encuentra, revisando sus expedientes, que el tiempo invertido en realizar un servicio, se comporta como una variable normal con media de 65 minutos y desviación estándar de 20 minutos. Calcula la proporción de servicios que se hacen en menos de 60 minutos. Resultado en porcentaje con dos decimales, ejemplo 91.32%.

En este ejercicio se aplica la función 'pnorm', pues se busca calcular la proporción de servicios que se hacen en menos de 60 minutos.

Debido que el tiempo invertido en realizar un servicio se distribuye normalmente con una media de 65 minutos y una desviación estándar de 20 minutos, se quiere encontrar P(X < 60 donde X es el tiempo de servicio.)

```r
# Parámetros de la distribución
miu = 65
sigma = 20

# Calcular la proporción de servicios que se hacen en menos de 60 minutos
prob_menos_60 = pnorm(60, mean = miu, sd = sigma)

# Convertir la probabilidad a porcentaje y redondear a dos decimales
porcentaje_menos_60 = round(prob_menos_60 * 100, 2)
porcentaje_menos_60

## [1] 40.13
```