

Práctica 2 - AD137-Estadística

Julio César Ramirez-Pacheco

17/11/2020

Distribuciones muestrales de la media

En esta práctica abordaremos el tema de la distribución muestral de la media para diversas distribuciones de probabilidad y veremos como éstas se ajustan a una distribución de probabilidad normal o Gaussiana como la que se muestra en la ecuación,

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Antes de iniciar, es necesario que nos familiaricemos con los siguientes comandos:

```
#hist()  
#qqnorm()  
#qqline()
```

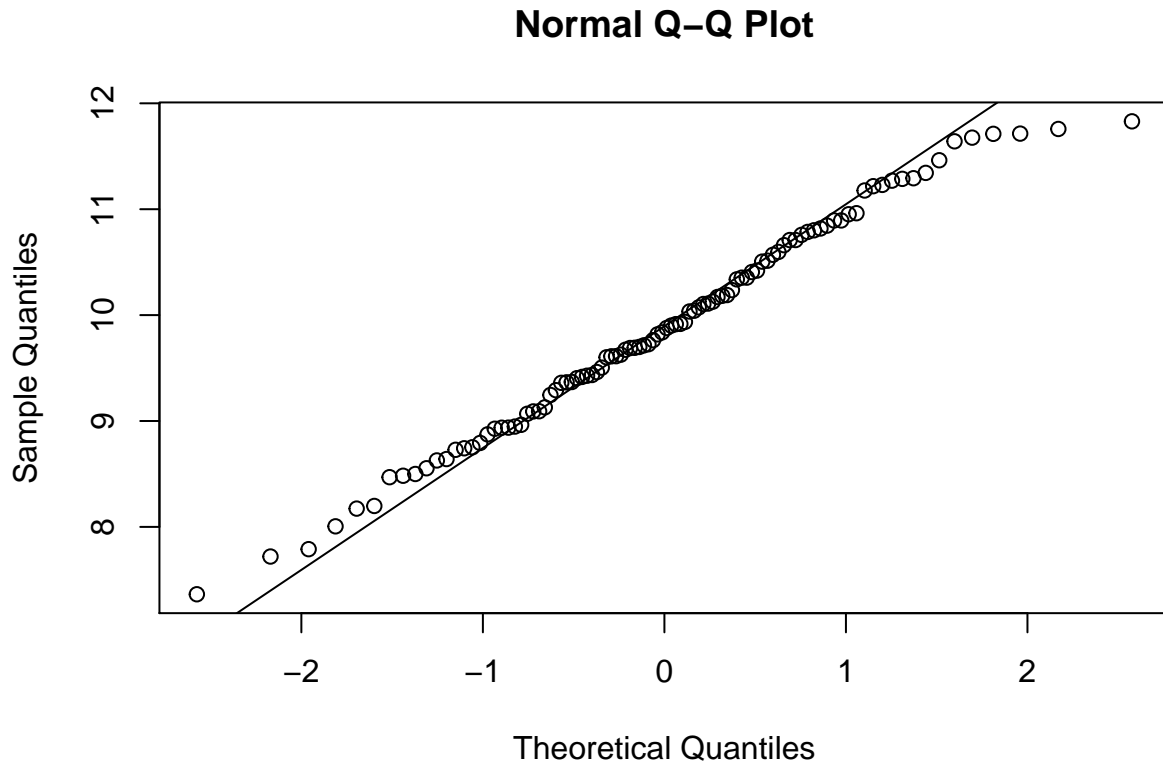
para ello podemos hacer `help(comando)` y obtener ayuda sobre cada uno de ellos. Ahora responda lo siguiente.

1. ¿Cuál es la función del comando `qqnorm`?
2. ¿Cuál es la función del comando `qqline`? ¿y de `hist`?

Aplicación de `qqnorm` y `qqline`

Veamos ahora la aplicación de `qqnorm` y `qqline` en un ejemplo practico. Supongamos que queremos verificar si los datos se ajustan a una distribución normal, entonces para hacerlos aplicamos,

```
datos <- rnorm(100, mean = 10)  
qqnorm(datos)  
qqline(datos)
```



Actividad.

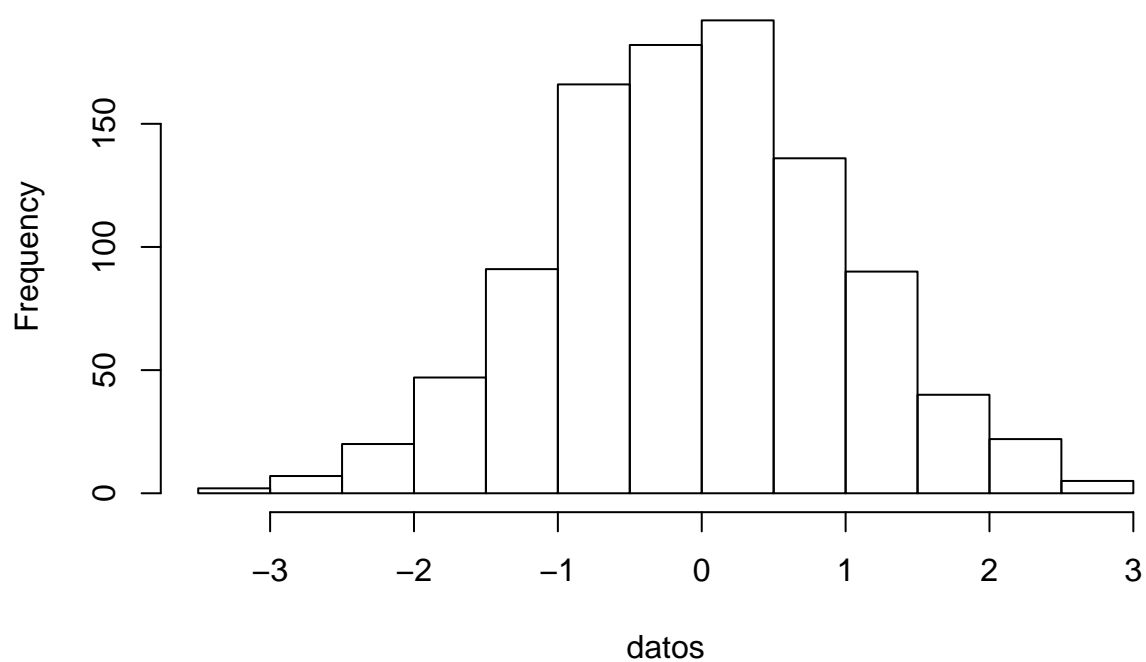
1. Ahora cambie los nombres de los ejes x , y y del título del gráfico anterior.
2. ¿Qué sucede si la longitud de los datos es pequeña, digamos 10, 20?
3. ¿Qué sucede si los datos tienen varianza grande? ¿En qué parte del gráfico Q-Q se puede ver la varianza grande?
4. ¿Y si cambiamos la media?
5. ¿Qué sucede si los datos son exponenciales?

Aplicación de `hist`

El comando `hist` permite obtener el histograma de una serie de datos. El histograma, de igual forma que el gráfico Q-Q permite ajustar una serie de datos a una distribución por medio de qué tanto se parece a la densidad. Veamos cómo se aplica.

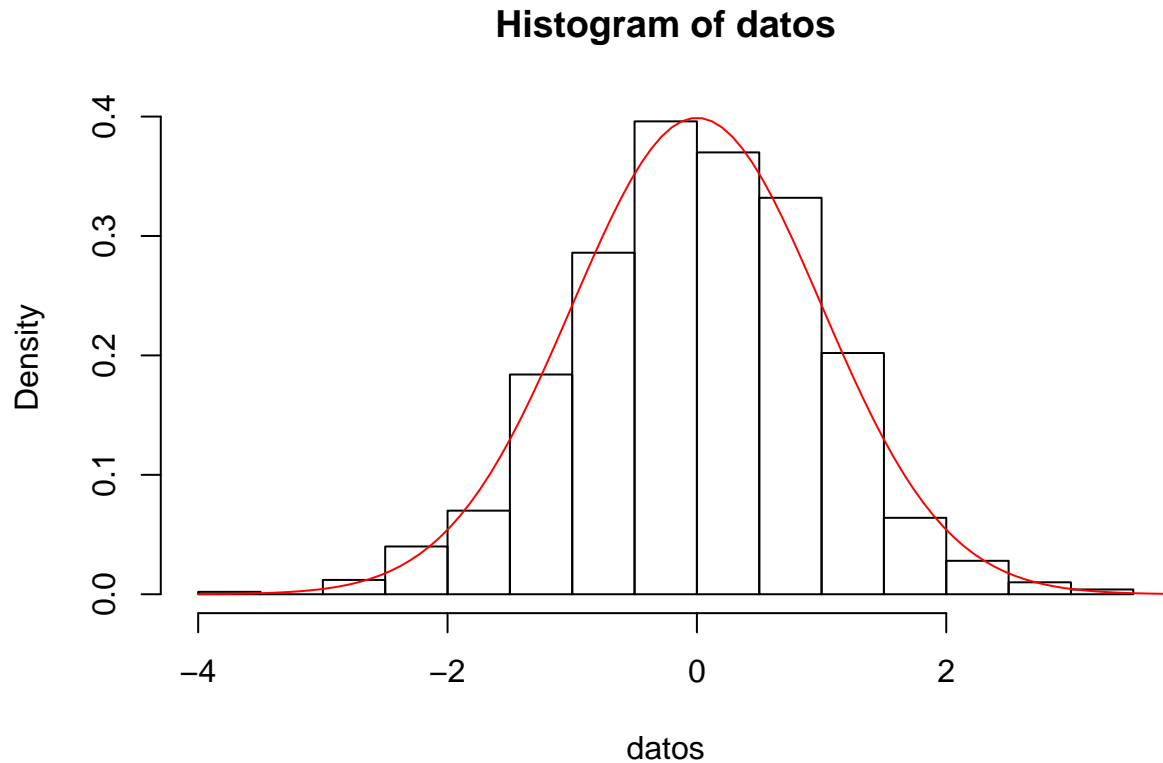
```
datos <- rnorm(1000) #1000 muestras de una densidad normal
hist(datos)
```

Histogram of datos



Podemos ver en el gráfico del histograma que éste es similar al gráfico de la densidad. El siguiente código lo muestra:

```
datos <- rnorm(1000) # muestras normales con media 0 y var 1.
teorica <- dnorm(seq(-4,4, length=100),mean = 0, sd = 1)
hist(datos,freq = FALSE)
lines(seq(-4,4, length=100),teorica, col="red")
```



Actividades.

1. ¿Qué sucede si el número de muestras es pequeño?
2. ¿Y si cambio la varianza? ¿Tiene algún efecto la varianza grande?
3. ¿Cuáles crees que son las ventajas y desventajas del histograma y el gráfico Q-Q?

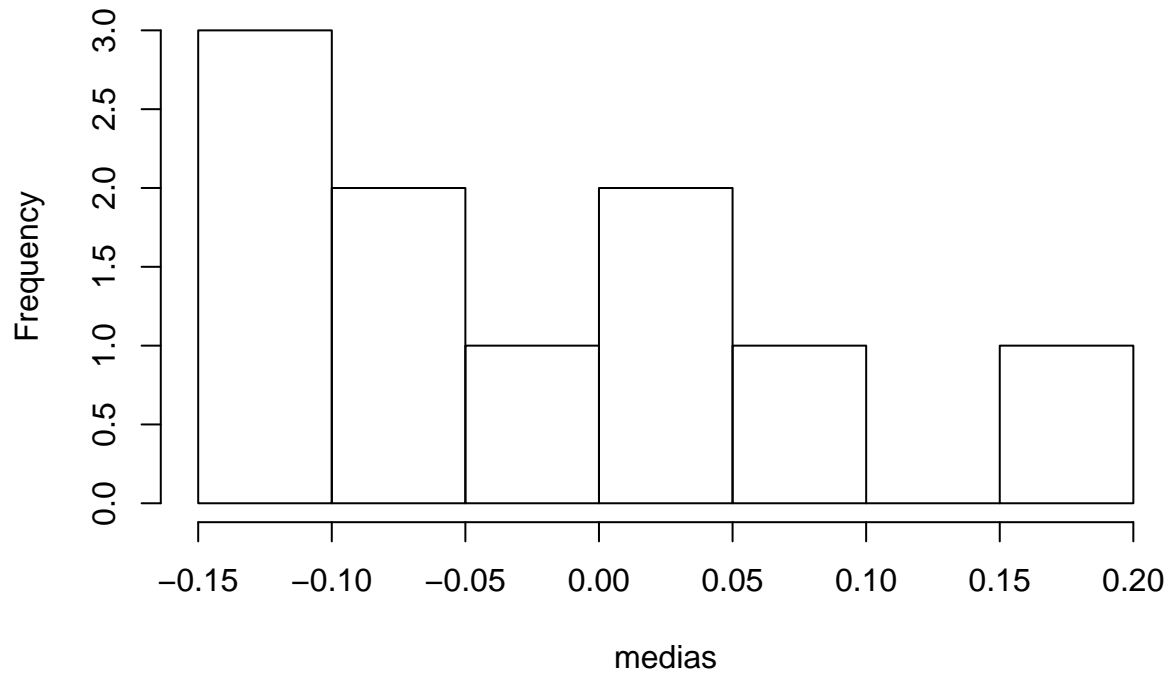
Aplicación de la distribución muestral de la media.

El término distribución muestral de la media se refiere a la distribución (ajustada mediante el gráfico Q-Q o el histograma) de la media muestral de un conjunto de datos. Por ejemplo

```
N <- 10 # número de medias (es decir número de muestras)
M <- 100 # número de la longitud de cada serie de datos
medias <- numeric(N) # Vector que guardará las medias

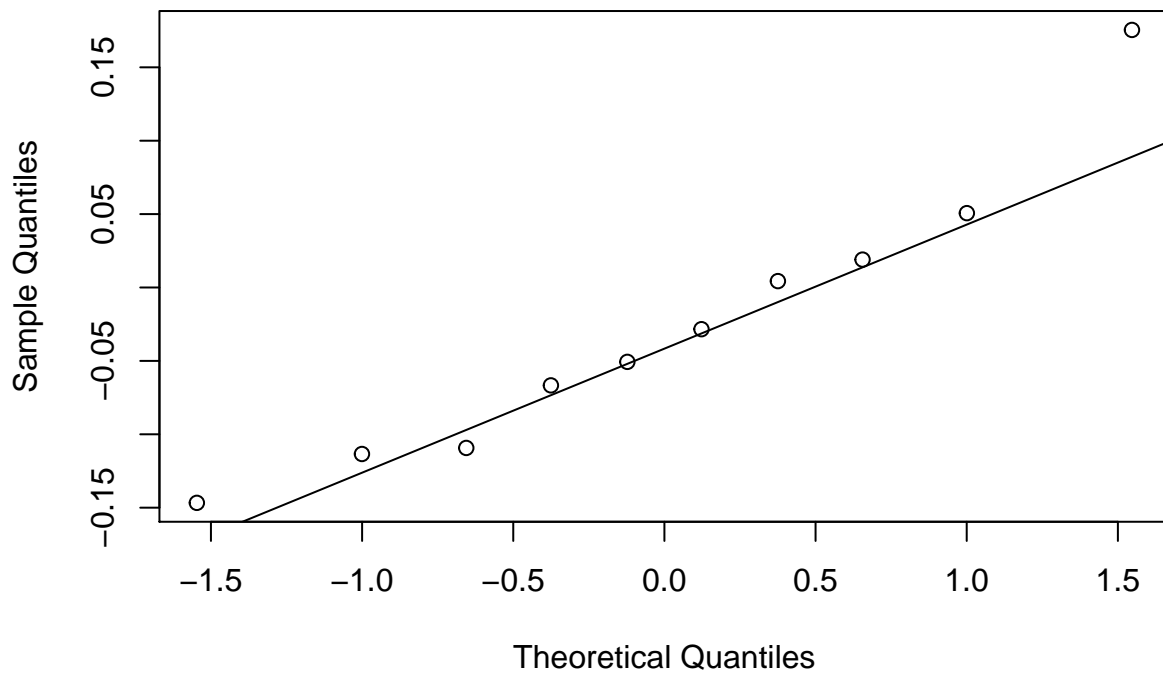
for(i in 1:N)
{
  datos <- rnorm(M)
  medias[i] <- mean(datos)
}
hist(medias)
```

Histogram of medias



```
qqnorm(medias)  
qqline(medias)
```

Normal Q-Q Plot



Actividad

1. ¿Qué sucede si hago N grande, por ejemplo 100, 1000, 10000?

2. ¿Qué sucede si en lugar de tomar una muestra Gaussian, tomo una muestra exponencial con N muy grande?

Fecha de entrega: Jueves 19 de Noviembre de 2020 a las 19:00