Trabajo Práctico Probabilidad y Estadística (c)

Luis Greco - Nicolas Hertzulis - Ruslan Sanmartin Sobol

20 de noviembre de 2017

Contents

Figure 1: Tendencia Probabilistica Revolucionaria

Lo primero que vamos a hacer es fijar la semilla de forma global.

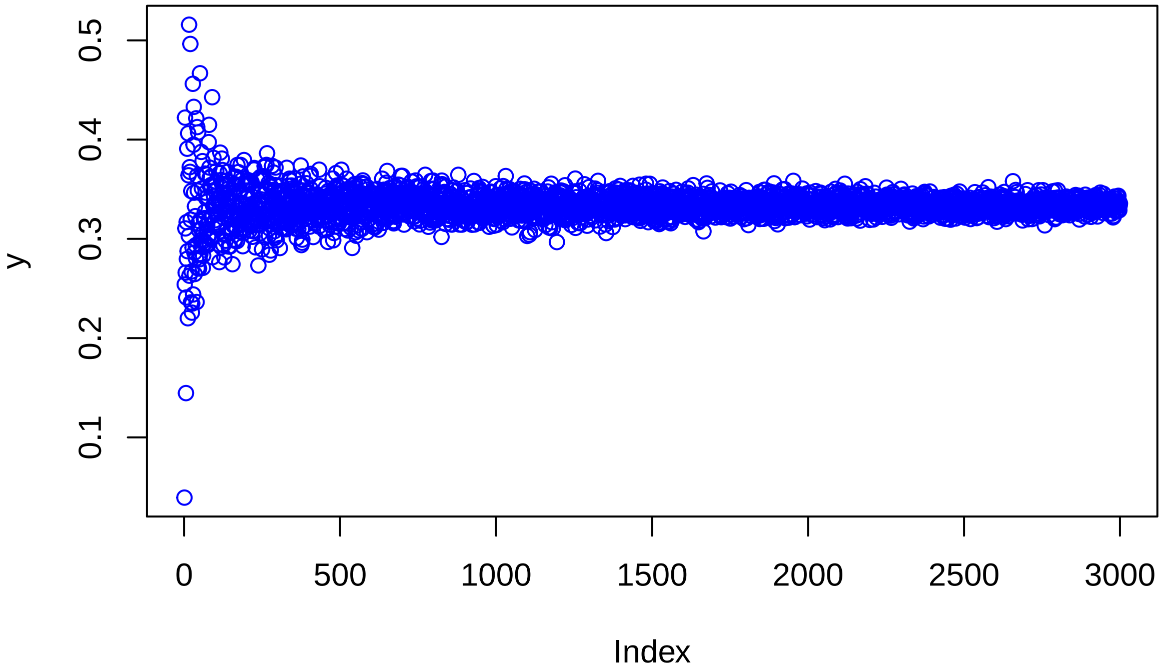
```
set.seed(1109)
```

Ejercicio 1

Acá lo que hicimos fue usar el mismo código visto en clase para generar la exponencial, solo que en este caso el valor de retorno es la media, por que es lo que necesitabamos. También podríamos haber usado la función "rexp()".

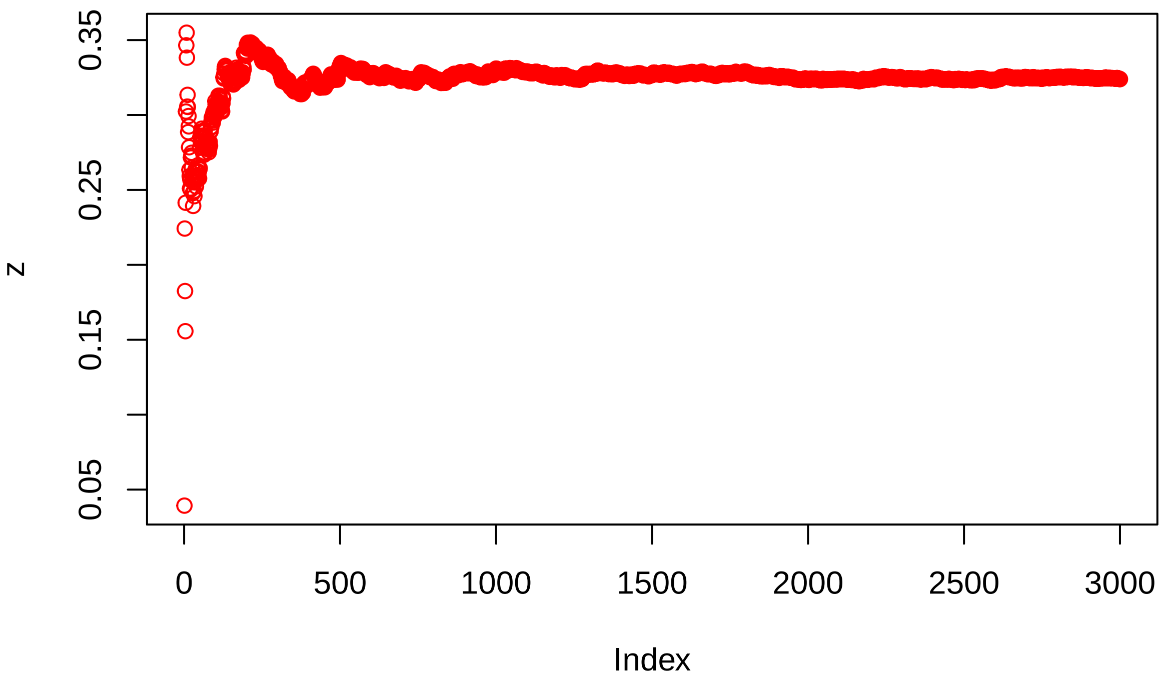
Ahora veamos los gráficos. Este primer gráfico es el que hicimos con el set.seed() "global".

```r
plot(y,col="blue")
```



Este otro gráfico es el que hicimos con el set.seed() dentro de la función.

```r
plot(z,col="red")
```

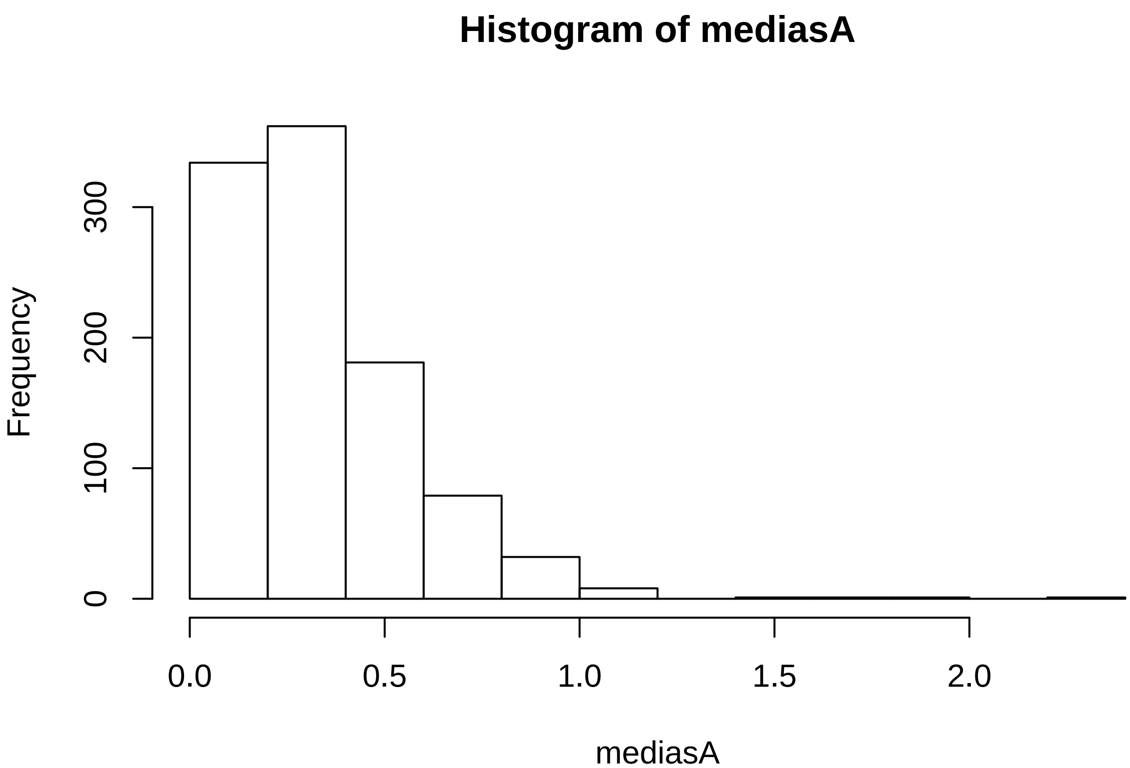


En cuanto a set.seed() nos encontramos con que un gráfico se encuentra notablemente más concentrado sobre ciertos valores que el otro. Creemos que esto se debe a que si fijamos set.seed() localmente dentro de una función, al ejecutarla varias veces, dará siempre el mismo resultado. Según lo que pudimos encontrar, cuando la función set.seed() es fijada globalmente, al generar números "aleatorios" se empieza usando una secuencia de números desde el lugar indicado, pero luego las simulaciones siguientes irán cambiando en función de la anterior, o sea, que en cada generación de números no tiene por qué dar resultados iguales.

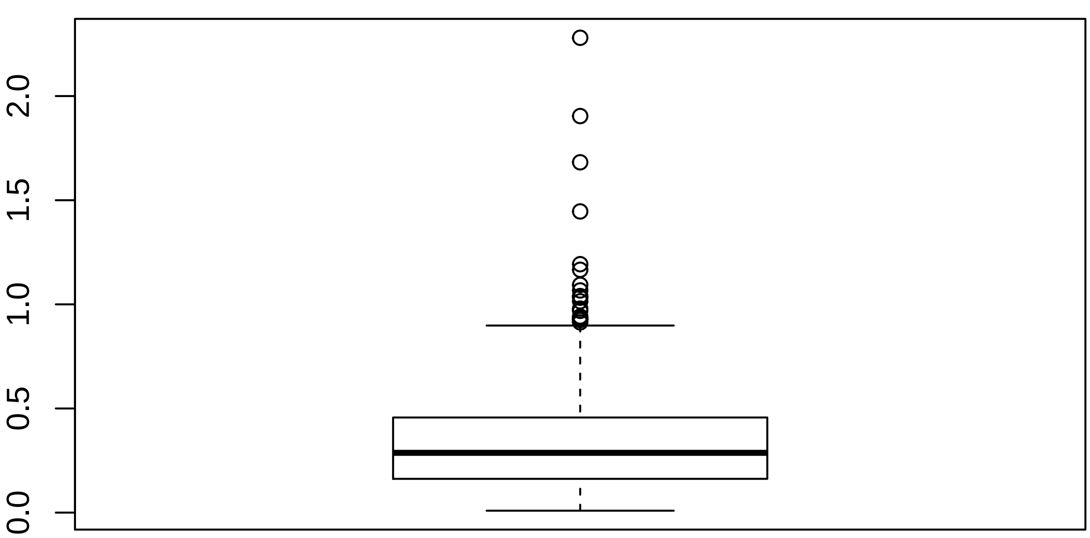
Se puede verificar mediante la Ley de Grandes Numeros, la media real y la media estimada, son casi exactas, con un posible error de (+0,05 o -0,05). En el plot de promedio, se observa como por la LGN, la misma converge a la media cuando su tamaño de muestras tiende a infinito. Se puede observar el comportamiento asintotico del promedio muestral.

Ejercicio 2

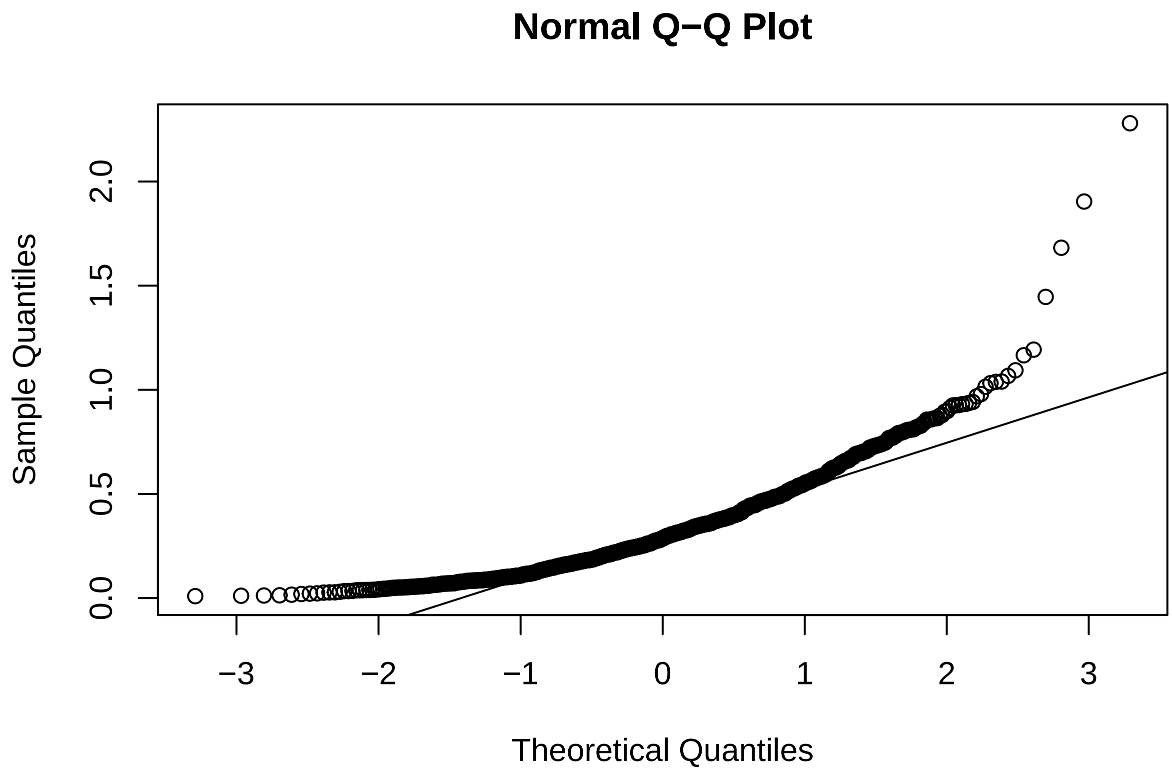
```r
hist(mediasA)
```




```r
boxplot(mediasA)
```



```
qqnorm(mediasA)
qqline(mediasA)#La cola del plot
```

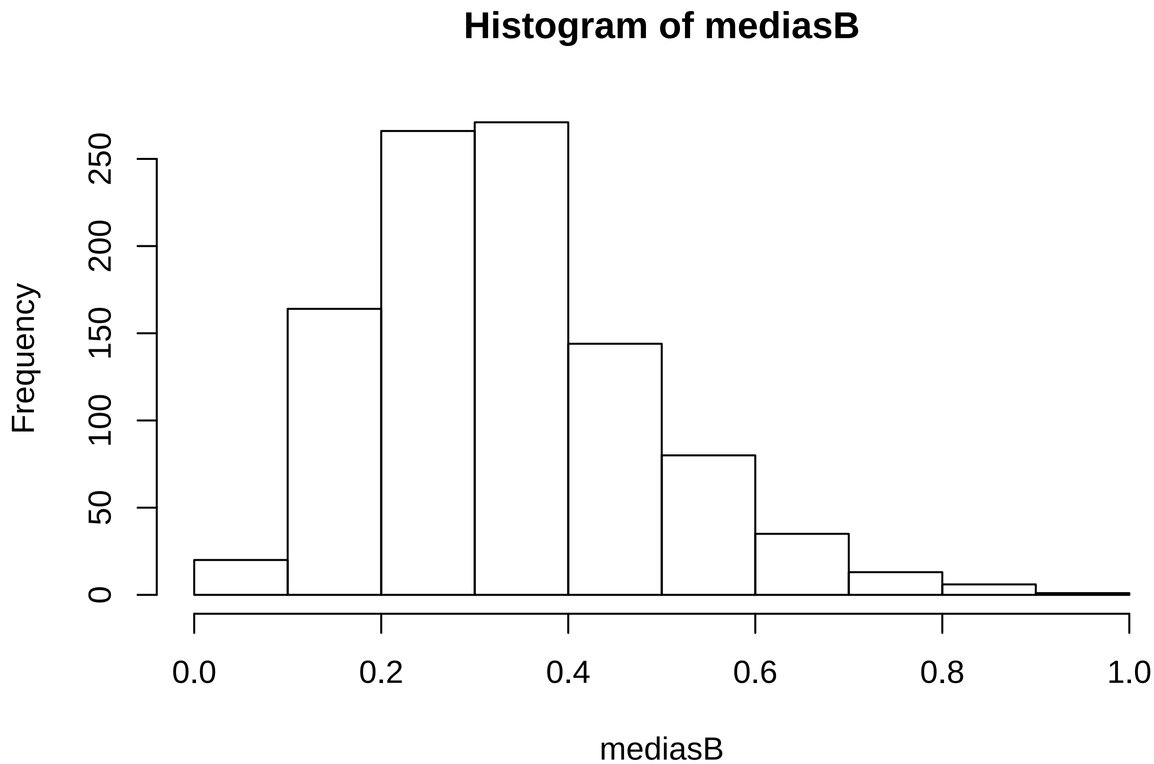


Se observa tanto en el histograma hay una asimetria a la derecha. En el Q-Q plot se observa una simetria de las colas livianas. En el boxplot se observa que tiene la cola superior pesada, ademas de los visibles outliers.

b)

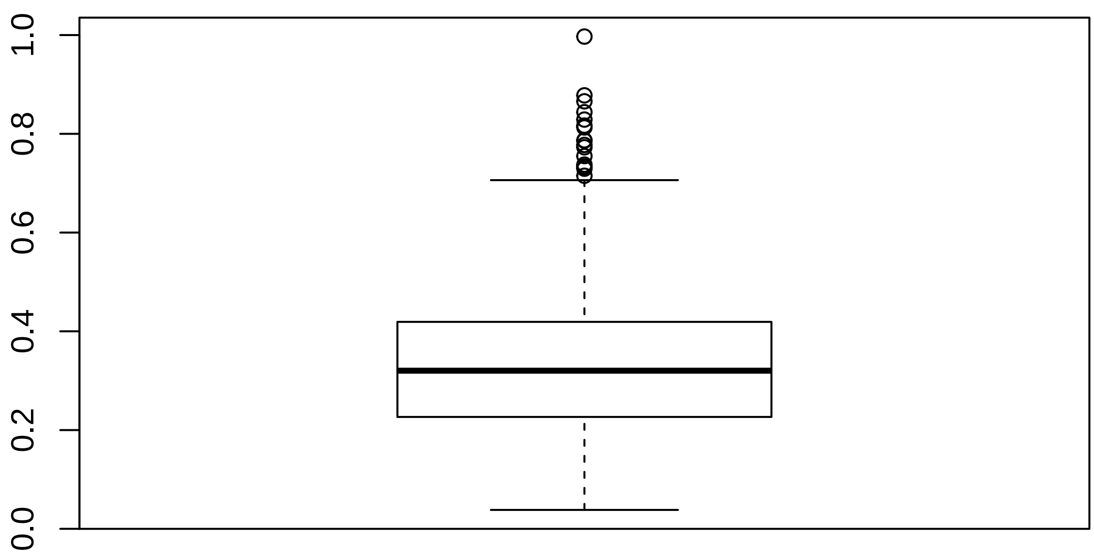
Histograma

```
hist(mediasB)
```



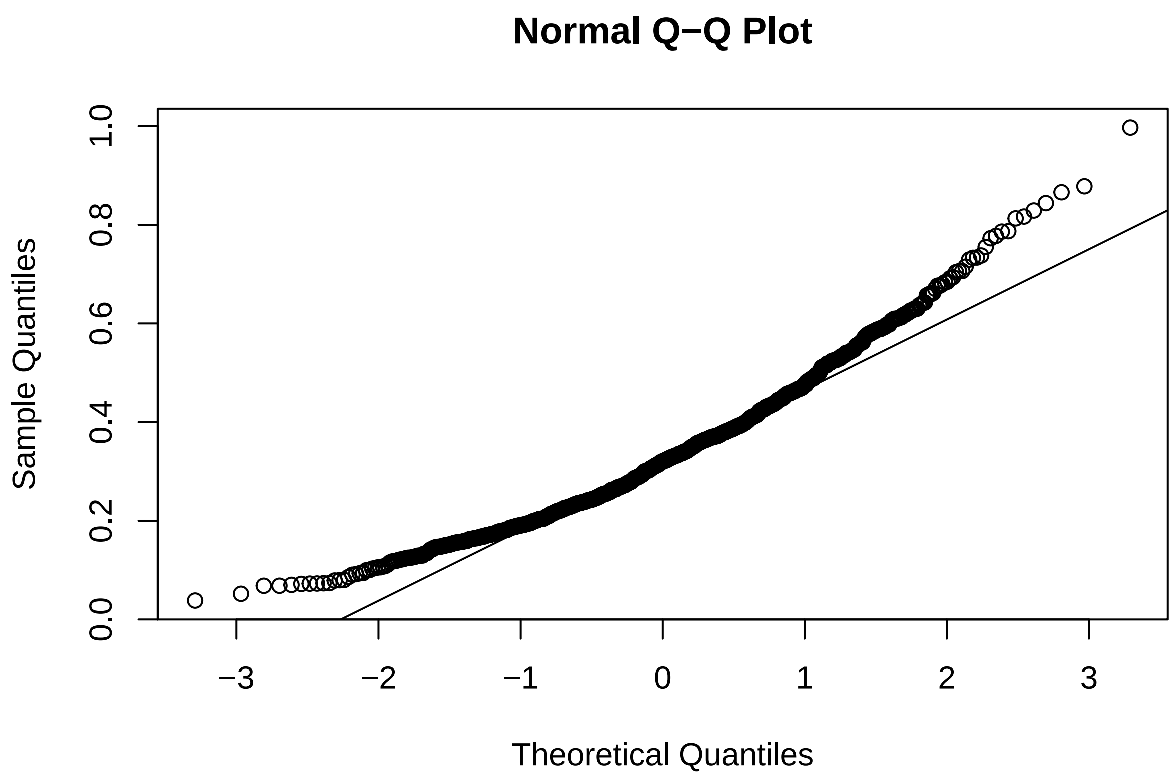
Boxplot

```r
boxplot(mediasB)
```



Q-Q Plot

```r
qqnorm(mediasB)
qqline(mediasB)#La cola del plot
```

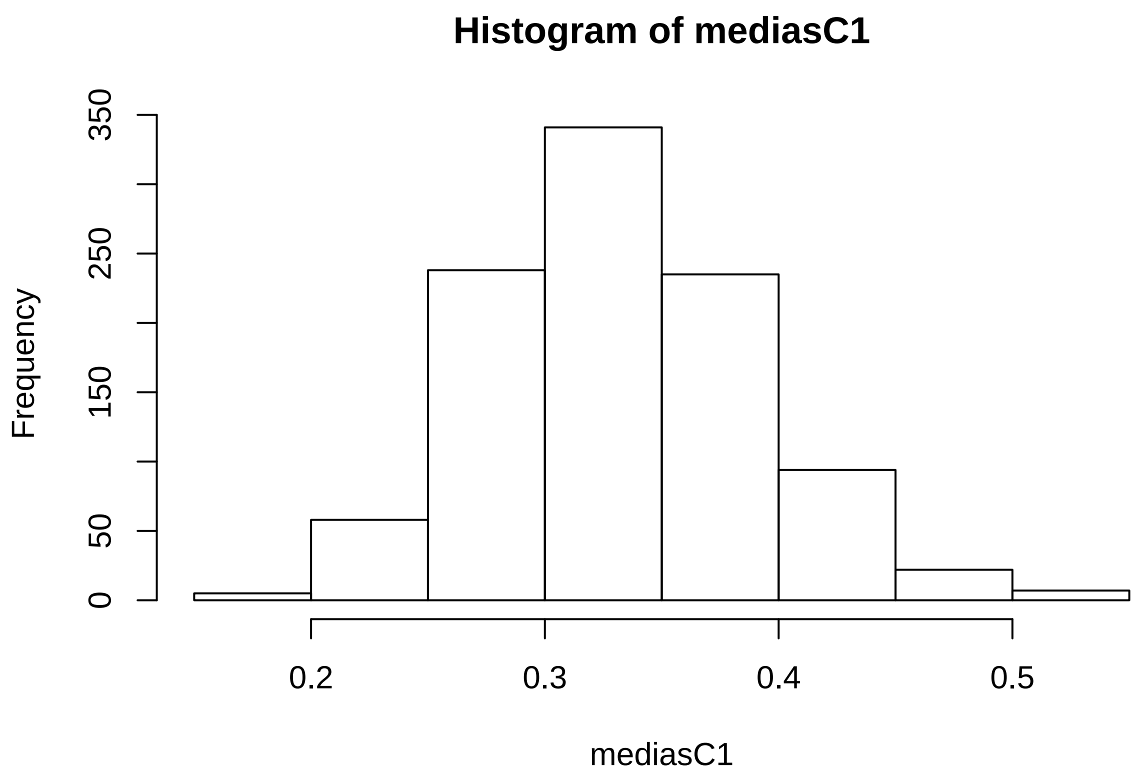


Se observa que el histograma empieza a tender a una distribucion normal El boxplot hace notar mas las colas pesadas y el Q-Q plot sigue manteniedo la correcta simetria, lo cual tiene a una normal.

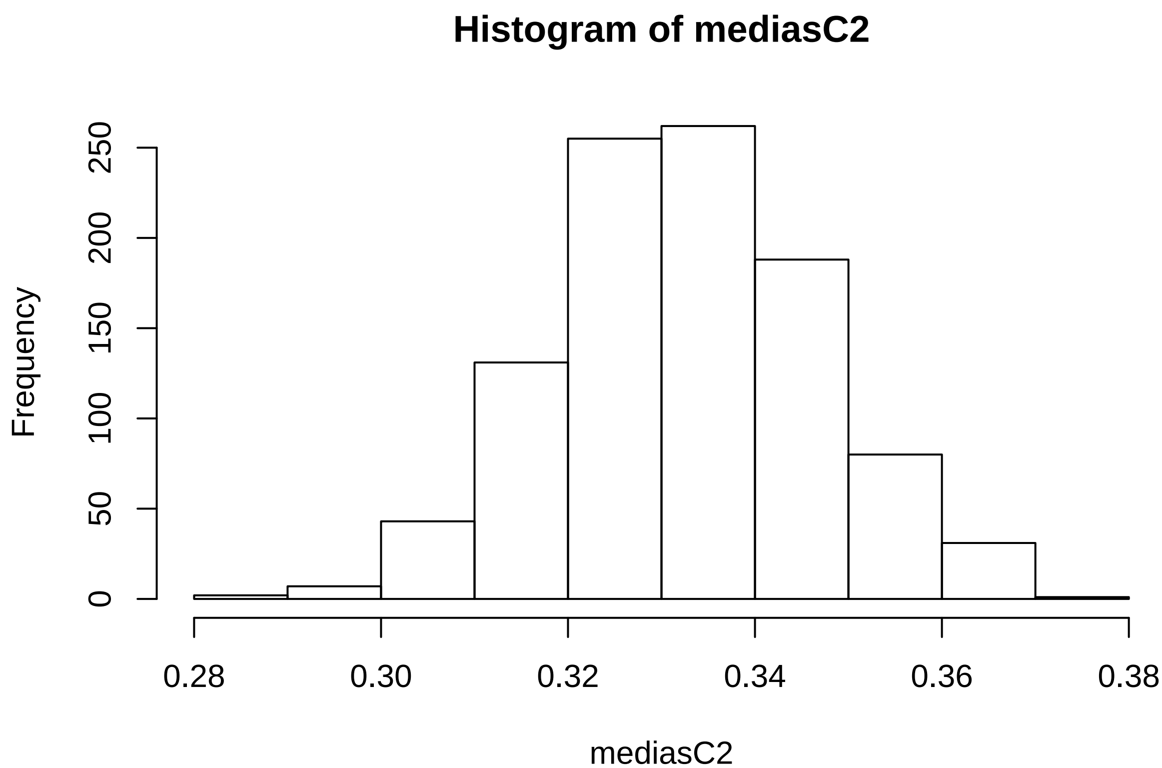
c)

Histograma

```
hist(mediasC1)
```

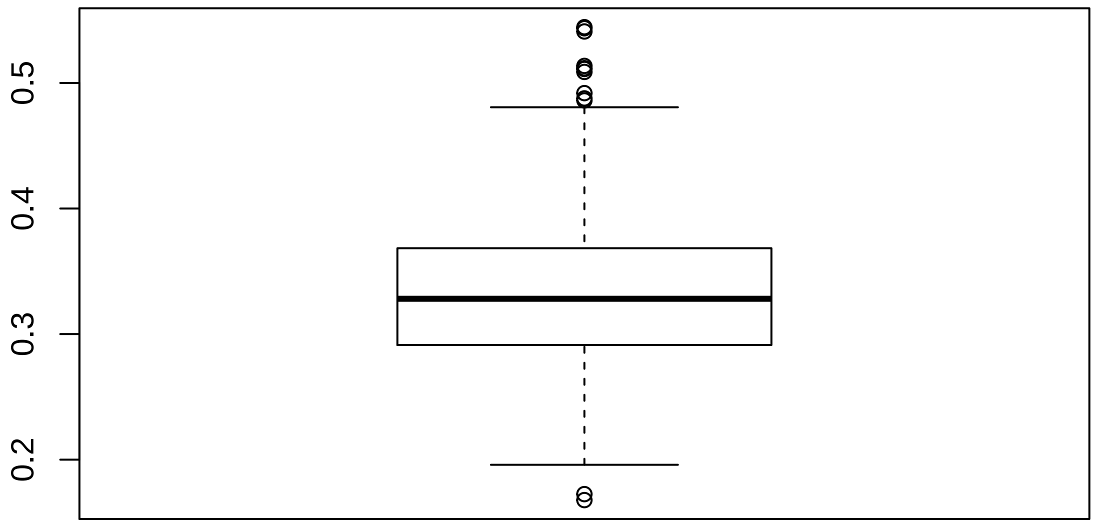



```r
hist(mediasC2)
```

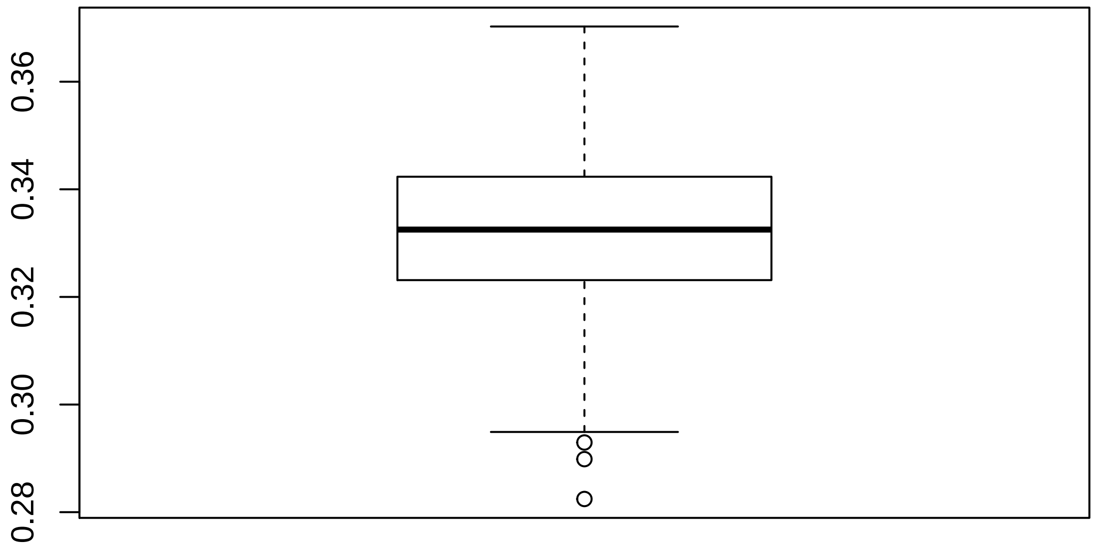


Boxplot

```r
boxplot(mediasC1)
```

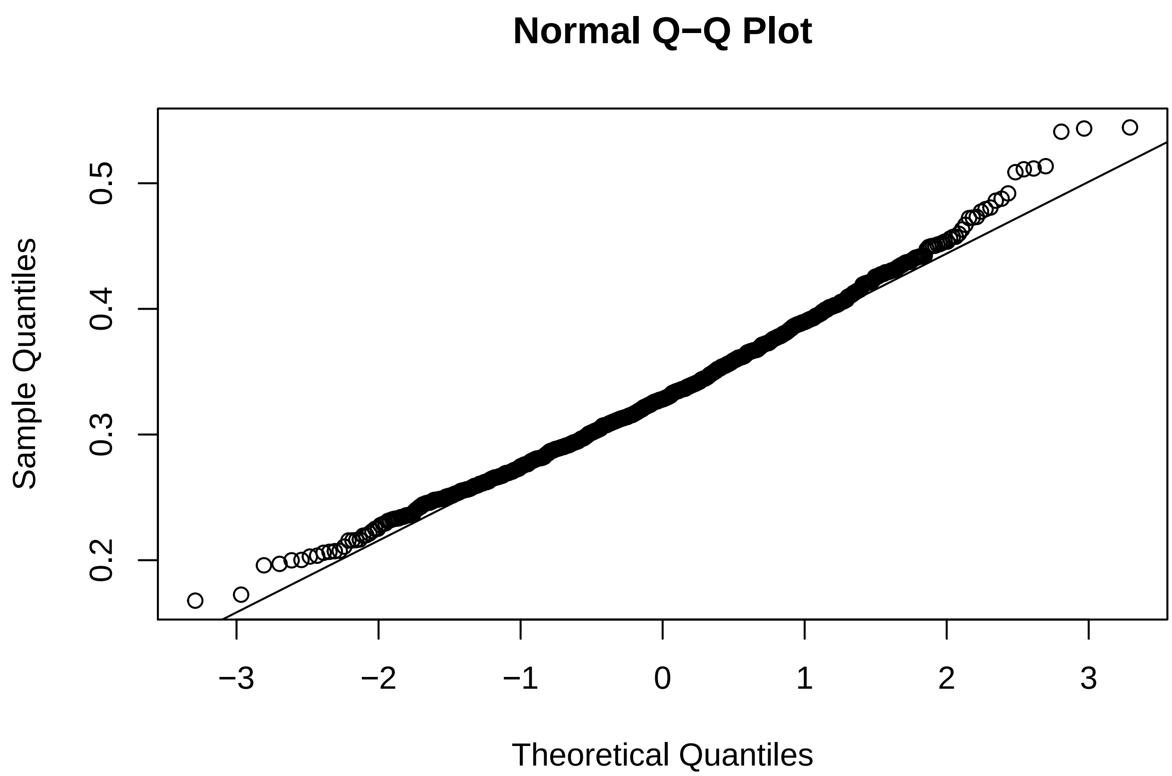



```r
boxplot(mediasC2)
```

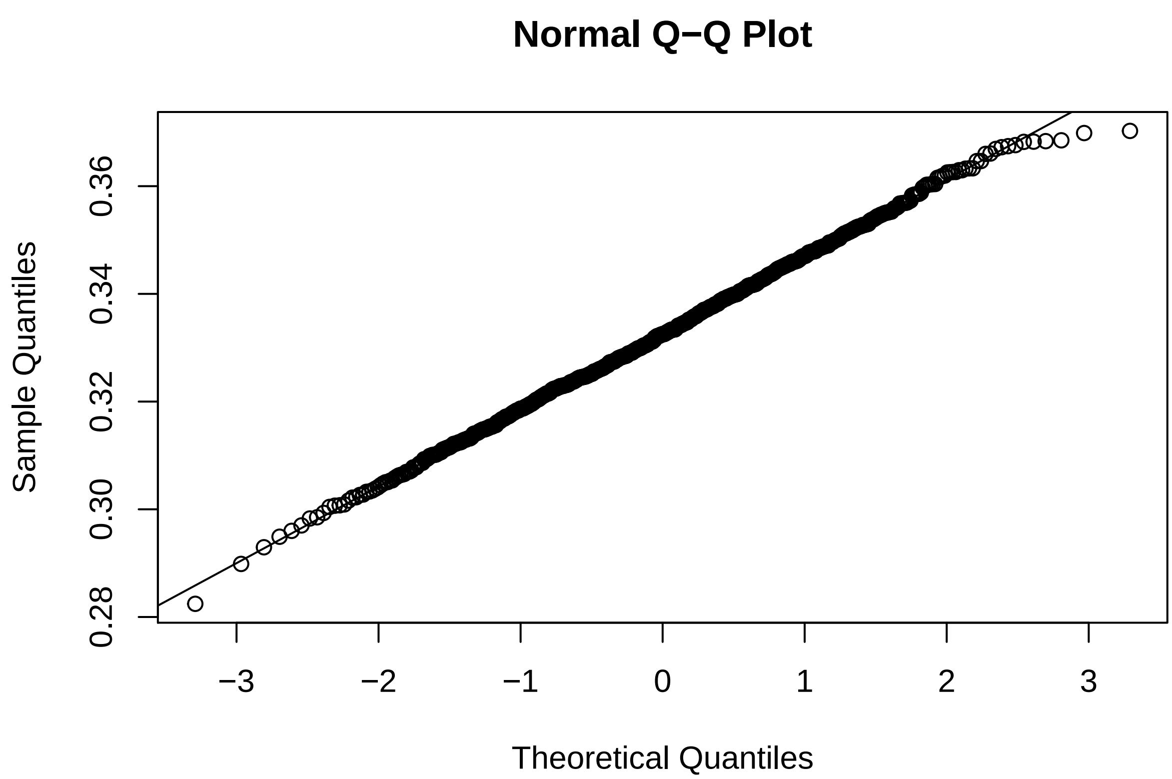


Q-Q Plot

```r
qqnorm(mediasC1)
qqline(mediasC1)#La cola del plot
```

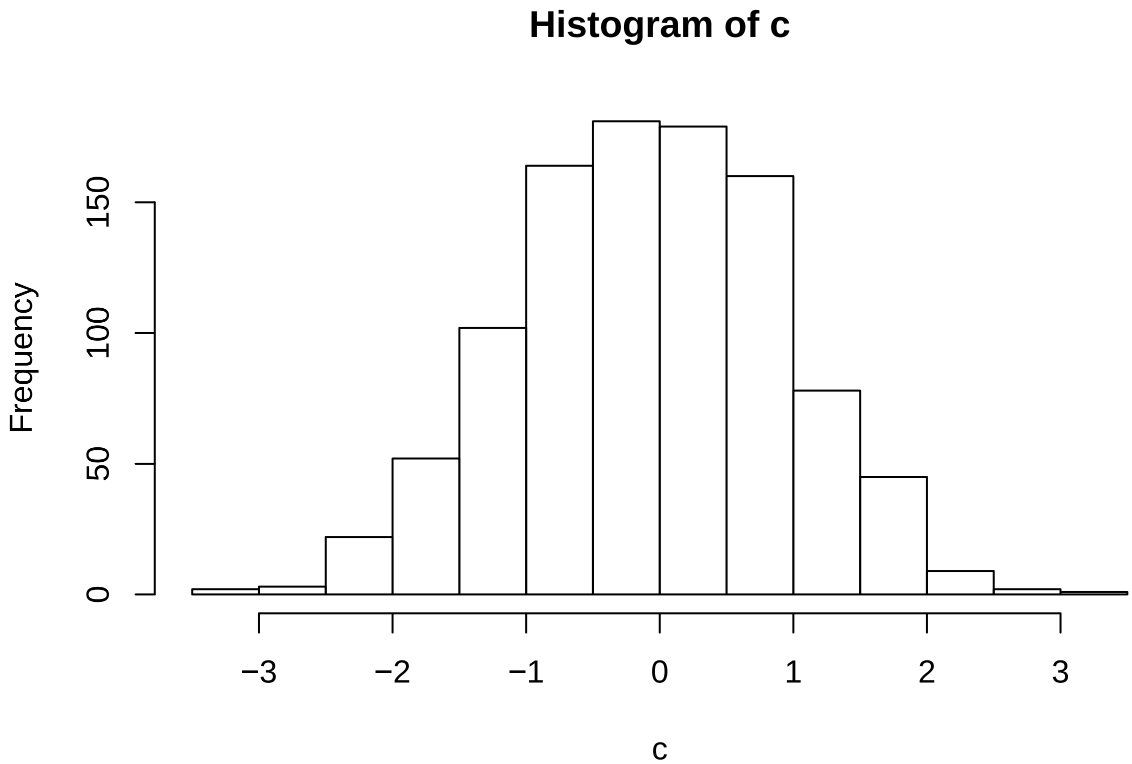



```r
qqnorm(mediasC2)
qqline(mediasC2)
```

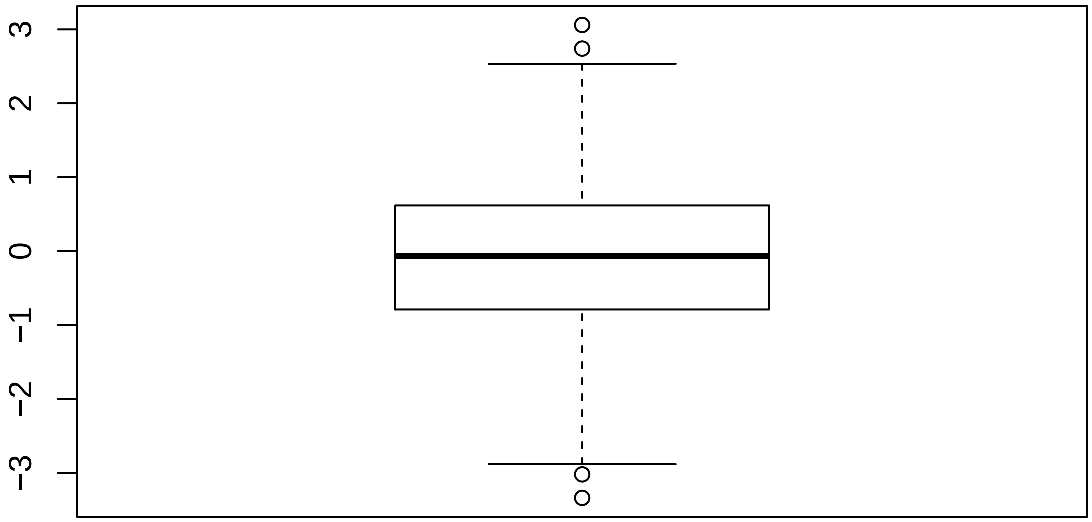


##d)

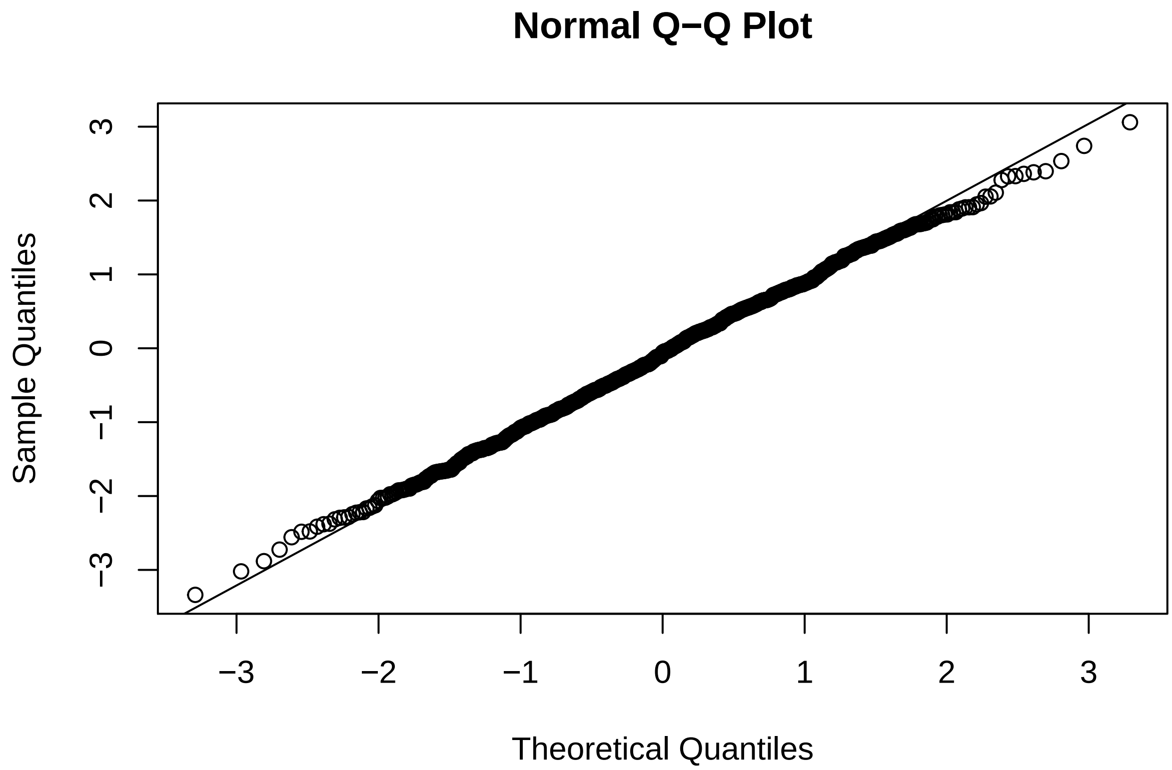
```r
hist(c)
```




```r
boxplot(c)
```




```r
qqnorm(c)
qqline(c)
```

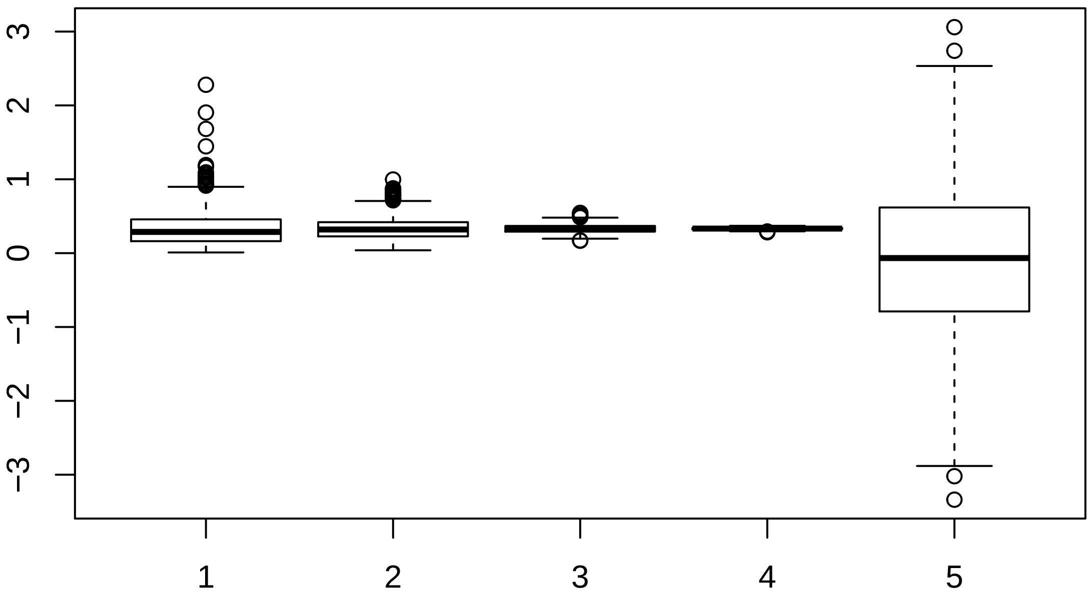


Acá tenemos los gráficos de una Normal(0,1), y notamos que se cumple la Ley de los Grandes Números al ver que a medida que aumentamos el n se van pareciendo más a este gráfico. Si aumentaramos todavía más el n, los gráficos serían más parecidos aún al de la Normal. Se nota con muchas mas fuerza en el histograma la distribucion normal. Lo mismo con el boxplot, el cual era el unico hasta el momento que no parecia tender a la normal. Ahora con una gran seguridad podemos confirmar que tiene a una normal con muy pocos outliers. Y el Q-Qplot se aferra con mucha mas fuerza a una distribucion normal.

e)

Boxplot

```
boxplot(mediasA,mediasB,mediasC1,mediasC2,c)
```



Obsevamos que con mayor muestra se puede verificar su tendencia a una distribucion normal. Dado que en el punto a, presuponia que poseia colas pesadas, tendiendo a la cola superior, con bastantes ouliers, siendo que cada vez que aumentabamos las muestras estos outleirs disminuian y las colas pesadas tendian a desaparecer y tender cada vez mas a la normal.

Ejercicio 3

a)

Estamos suponiendo que se refiere a la esperanza y varianza muestral.

```
#MediaX1
mediaX1
```

```
## [1] 0.3347528
```

```
#varX1
varX1
```

```
## [1] 0.05707561
```

```
#MediaX2
mediaX2
```

```
## [1] 0.335202
```

```
#varX2
varX2
```

```
## [1] 0.02173452
```

```
#MediaX3
mediaX3
```

```
## [1] 0.3321062
```

```
#varX3
varX3
```

```
## [1] 0.003370017
```

```r
#MediaX4
mediaX4
```

```
## [1] 0.3326577
```

```r
#varX4
varX4
```

```
## [1] 0.000201832
```

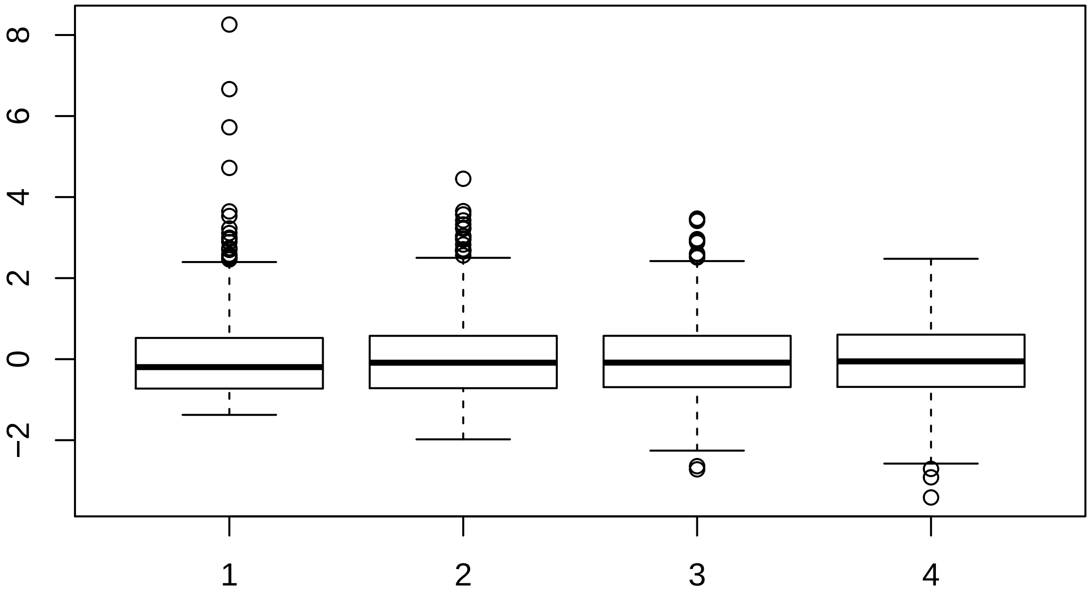
b)

No nos quedó muy claro que era lo que había que hacer en este punto. Entendimos que Xn es el promedio muestral aplicado sobre las medias

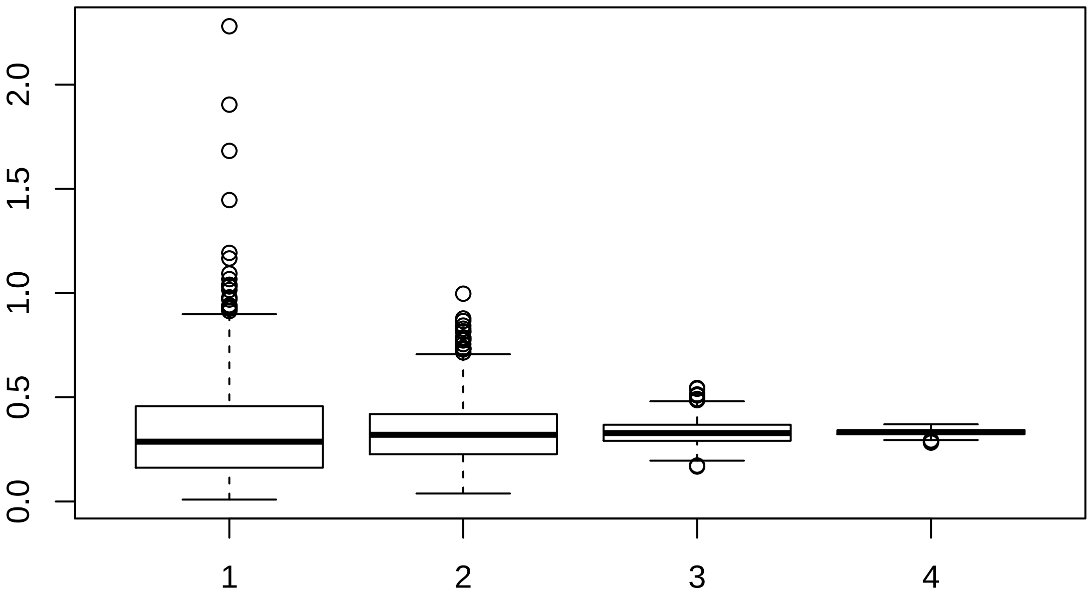
```r
#Transformación: (1/3 es la esperanza de la exp(3) )
transformacionesA <- seq(1000)
for (i in 1:1000){
transformacionesA[i] <- (mediasA[i] - (1/3))/(sqrt((1/9)/2))
}
#transformacionA

transformacionesB <- seq(1000)
for (i in 1:1000){
transformacionesB[i] <- (mediasB[i] - (1/3))/(sqrt((1/9)/5))
}

#transformacionB <- (mediaX2 - (1/3))/(sqrt((1/9)/5))
#transformacionB

transformacionesC1 <- seq(1000)
for (i in 1:1000){
transformacionesC1[i] <- (mediasC1[i] - (1/3))/(sqrt((1/9)/30))
}

#transformacionC1 <- (mediaX3 - (1/3))/(sqrt((1/9)/30))
#transformacionC1

transformacionesC2 <- seq(1000)
for (i in 1:1000){
transformacionesC2[i] <- (mediasC2[i] - (1/3))/(sqrt((1/9)/500))
}

#transformacionC2 <- (mediaX3 - (1/3))/(sqrt((1/9)/1000))
#transformacionC2
```

Boxplot de las transformaciones, abajo los de las medias.

```r
boxplot(transformacionesA, transformacionesB, transformacionesC1, transformacionesC2)
```

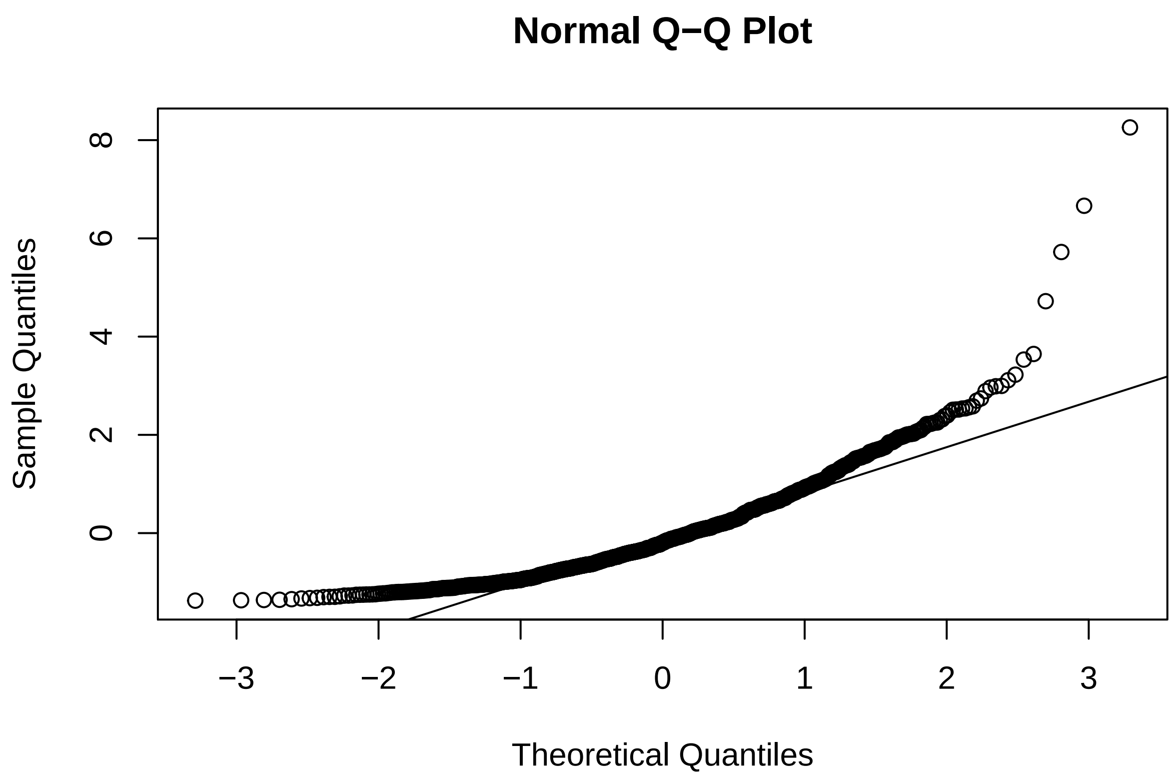



```r
boxplot(mediasA, mediasB, mediasC1, mediasC2)
```

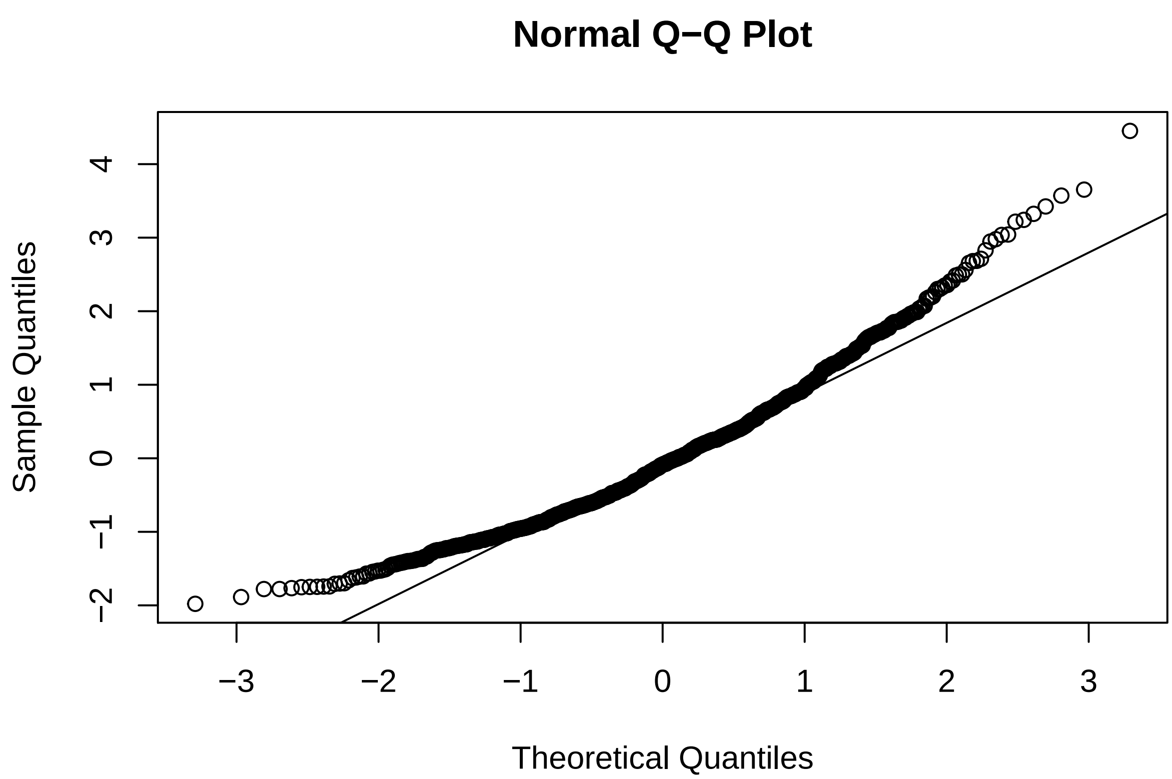


Q-Q Plots. De las medias y después de las transformaciones.

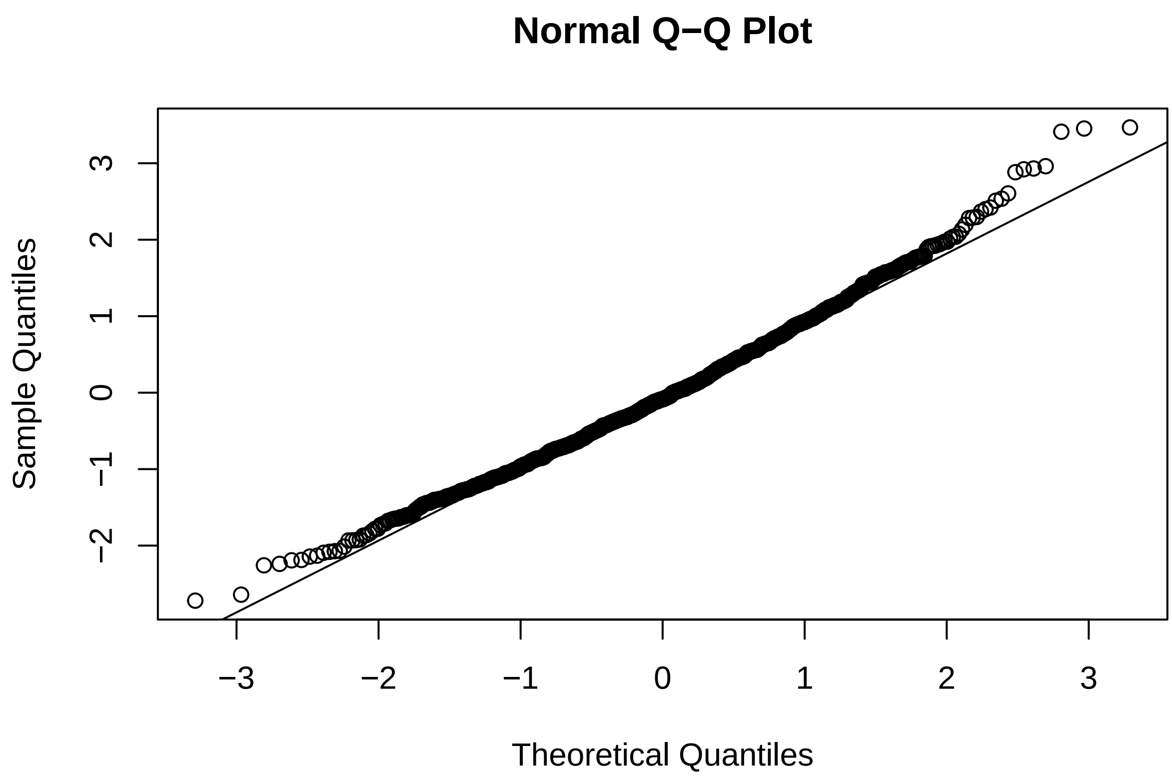
```r
qqnorm(transformacionesA)
qqline(transformacionesA)
```



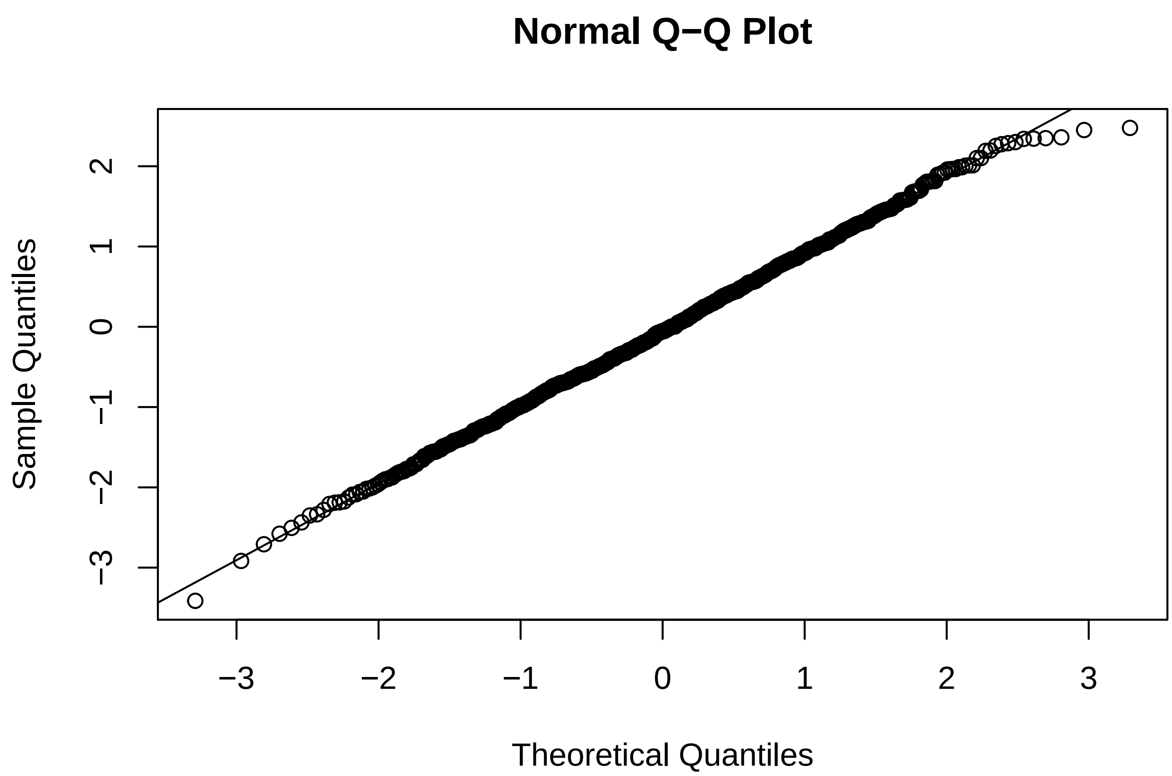

```r
qqnorm(transformacionesB)
qqline(transformacionesB)
```



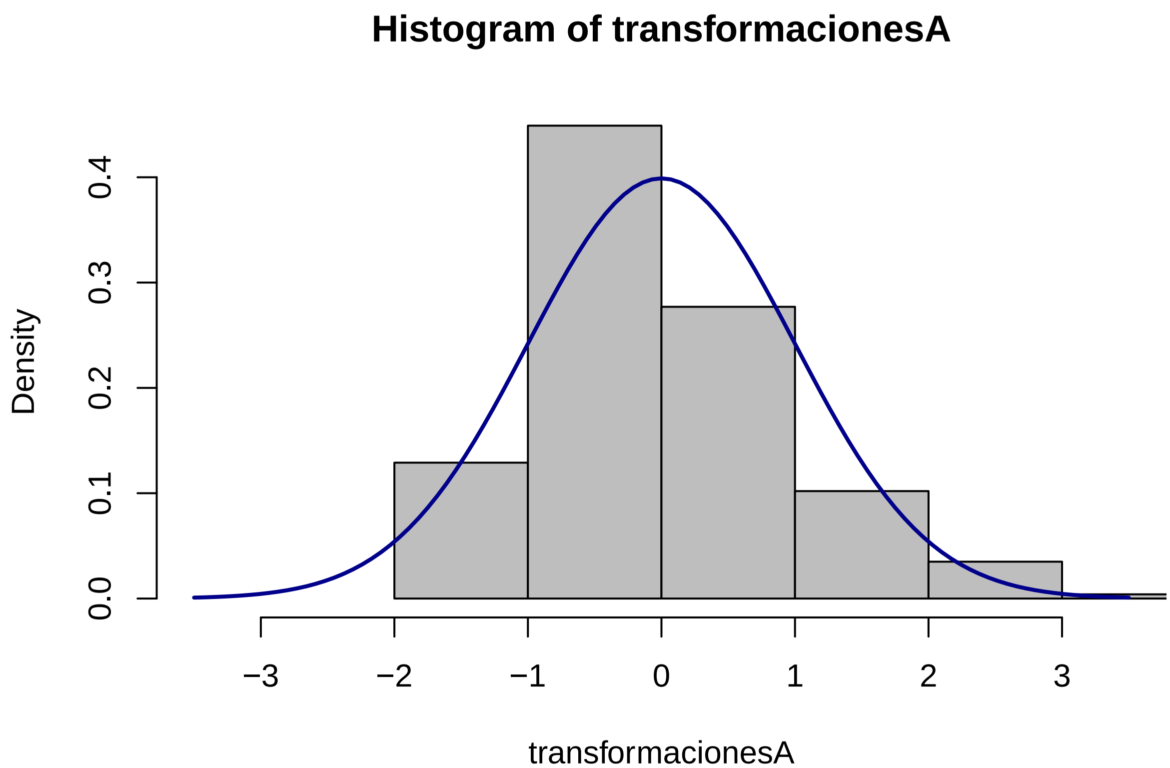

```r
qqnorm(transformacionesC1)
qqline(transformacionesC1)
```



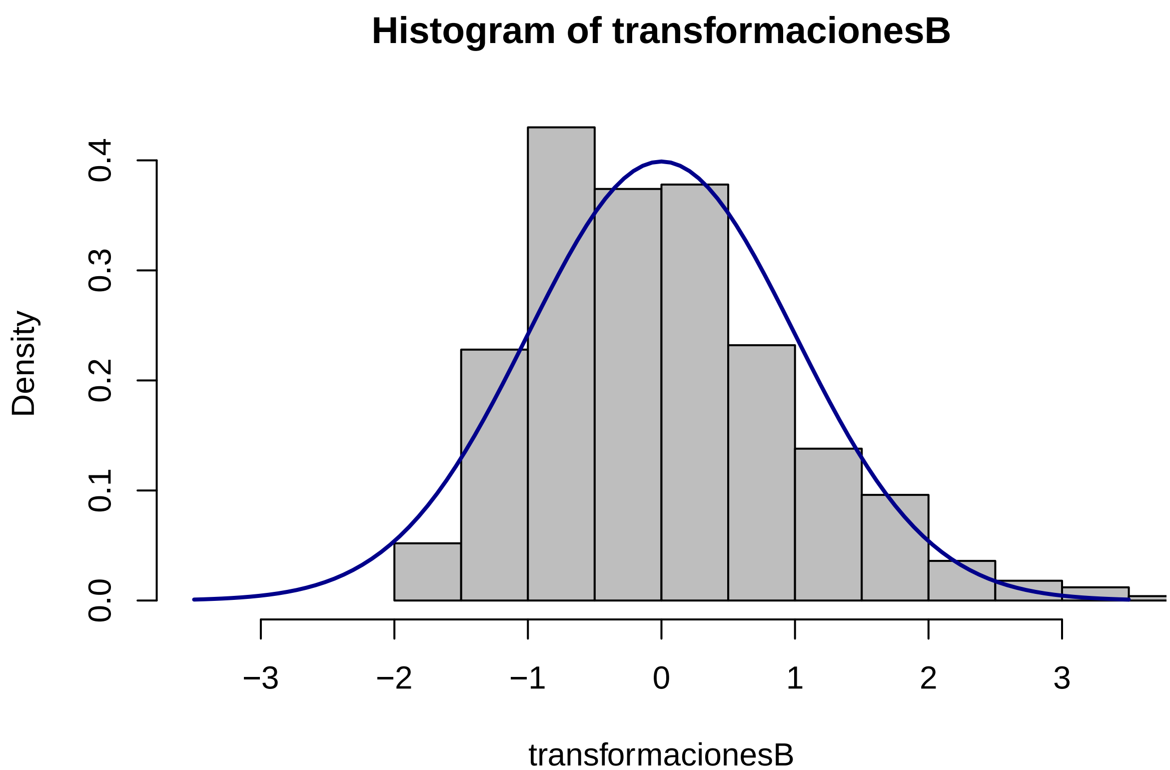

```r
qqnorm(transformacionesC2)
qqline(transformacionesC2)
```



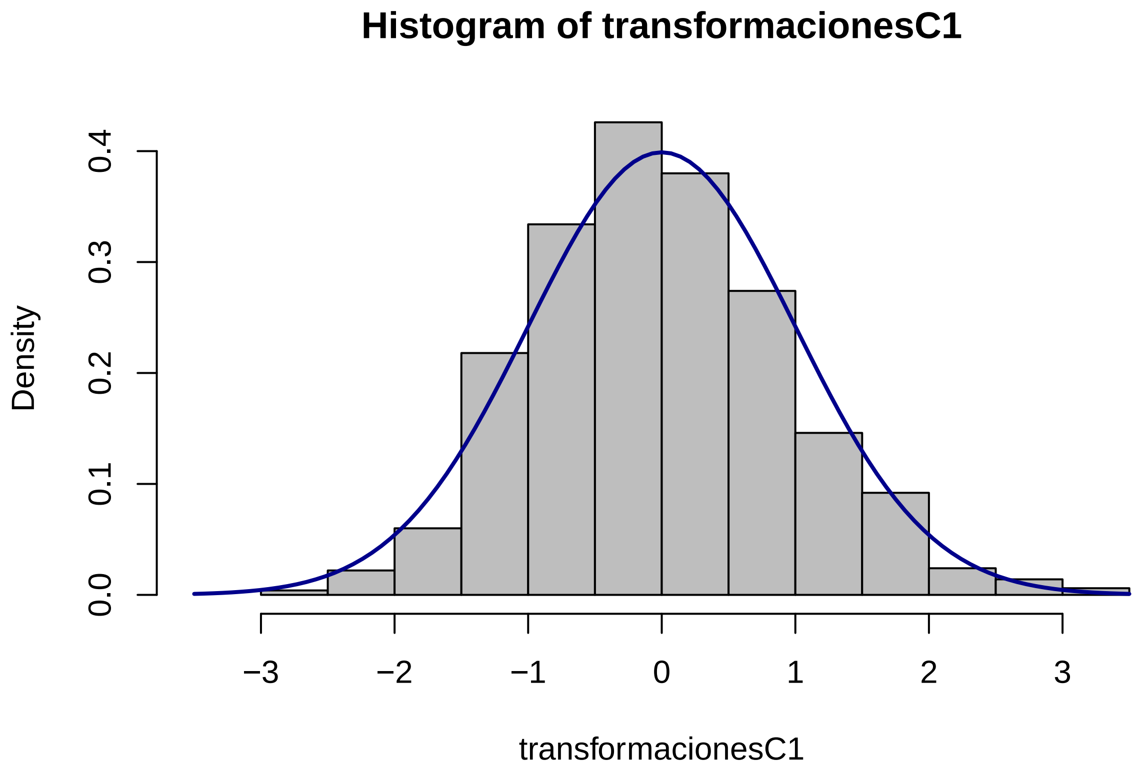
c)

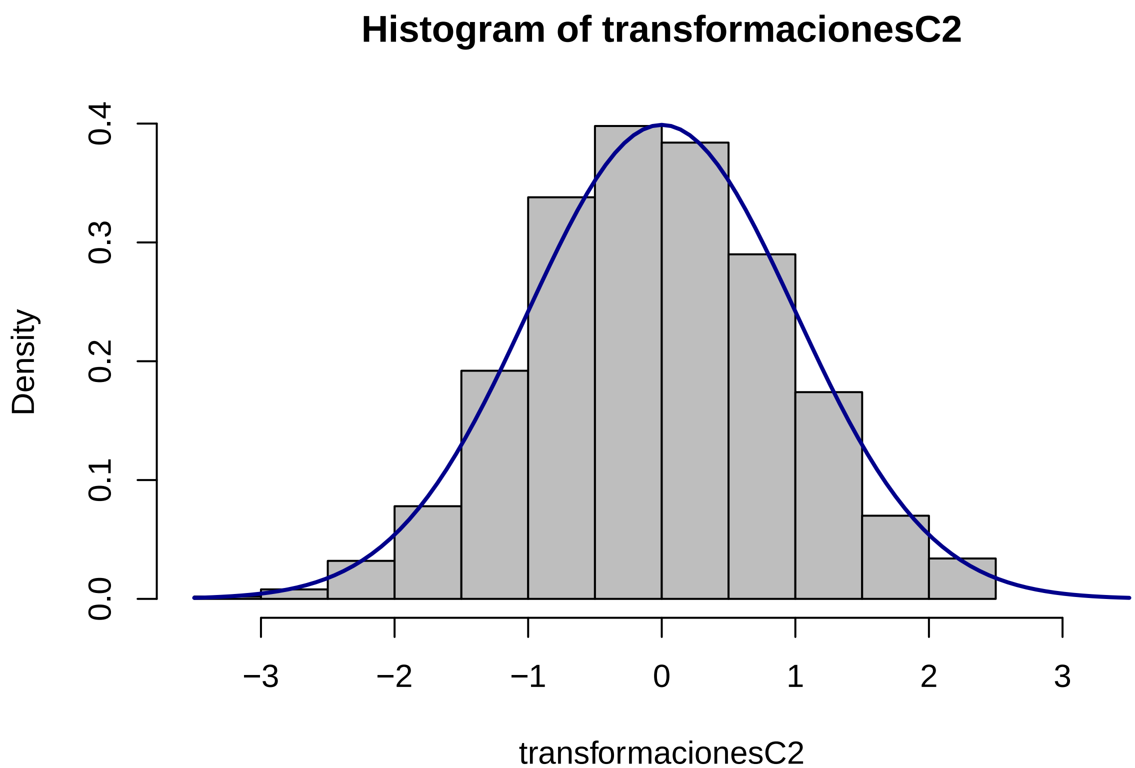


Histogram of transformacionesA
Density
0.0
0.1
0.2
0.3
0.4
−3
−2
−1
0
1
2
3
transformacionesA



Histogram of transformacionesB
Density
0.0
0.1
0.2
0.3
0.4
−3
−2
−1
0
1
2
3
transformacionesB



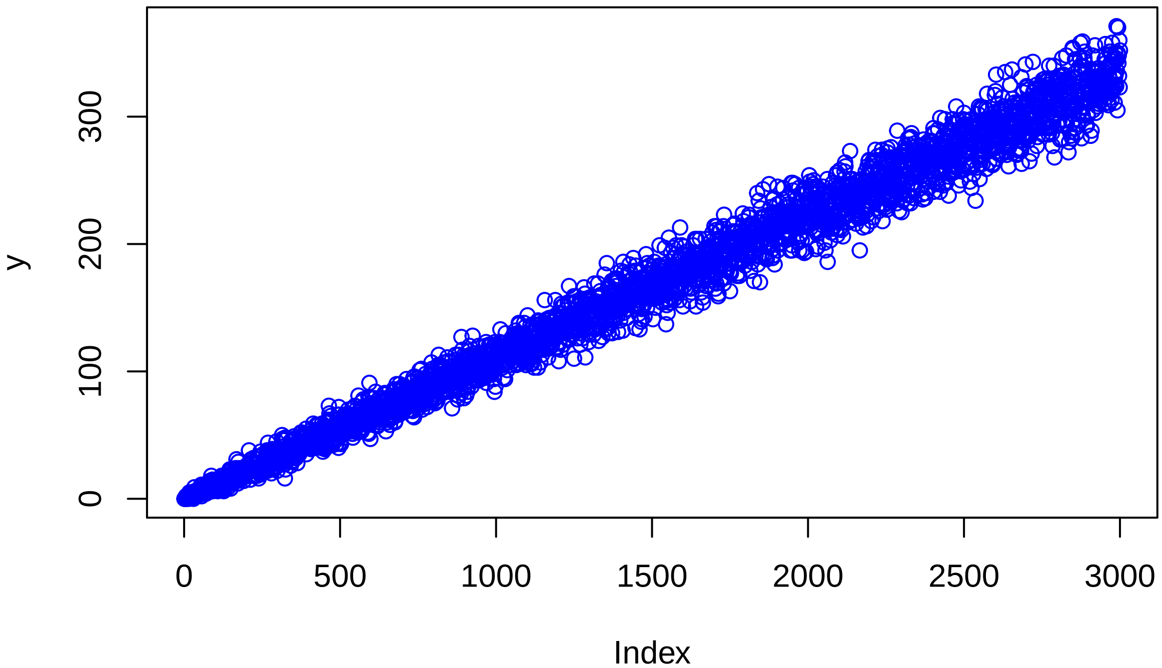


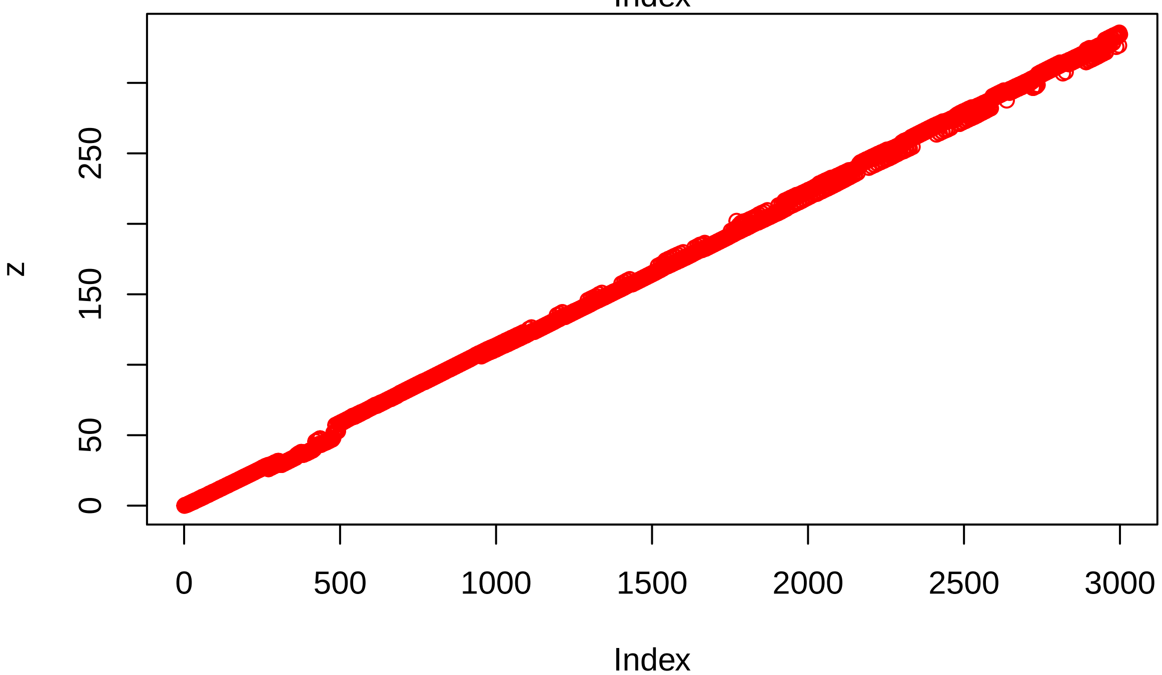
d)

Con respecto a los histogramas,`mediasA` `mediasB`, `mediasC1` y `mediasC2`, se puede observar como los graficos empiezan a represtan una "campana" al estilo de una normal, lo cual nos informa, que nuestra muestra fue lo suficientemente buena, para poder estimar dicha distribucion aleatoria.

Ejercicio 4

Punto 1



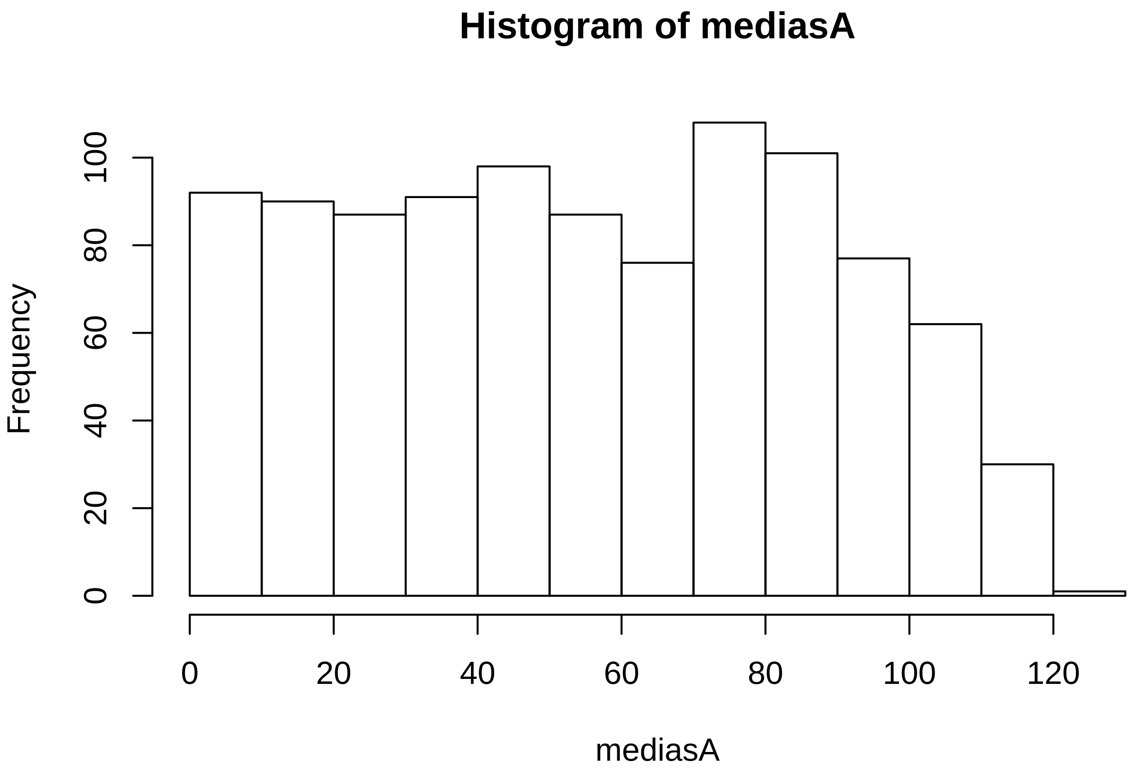


Se puede verificar mediante la Ley de Grandes Numeros, la media real y la media estimada, son casi exactas, con un posible error de (+0,05 o -0,05). En el plot de promedio, se observa como por la LGN, la misma converge a la media cuando su tamaño de muestras tiende a infinito. Se puede observar el comportamiento asintotico del promedio muestral.

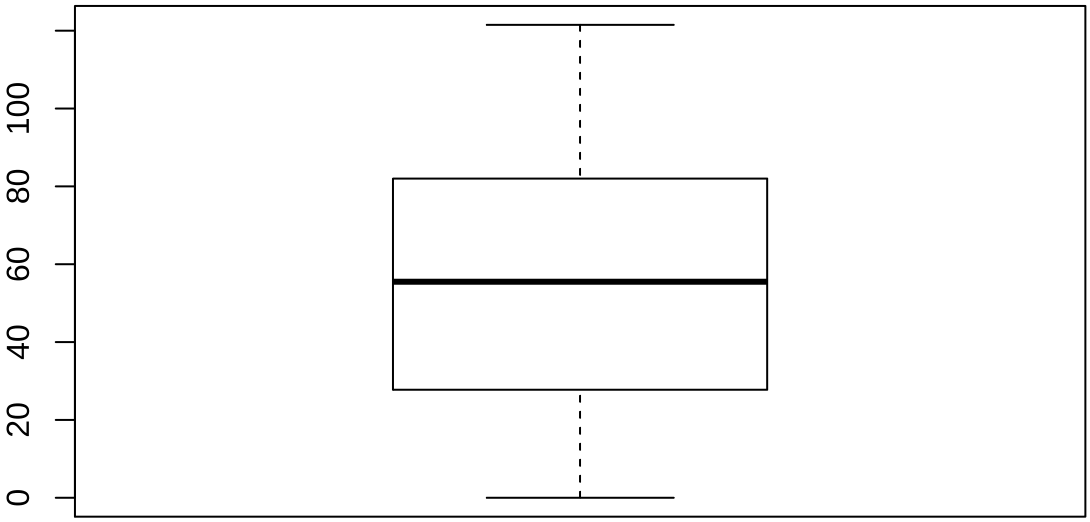
Punto 2

a)

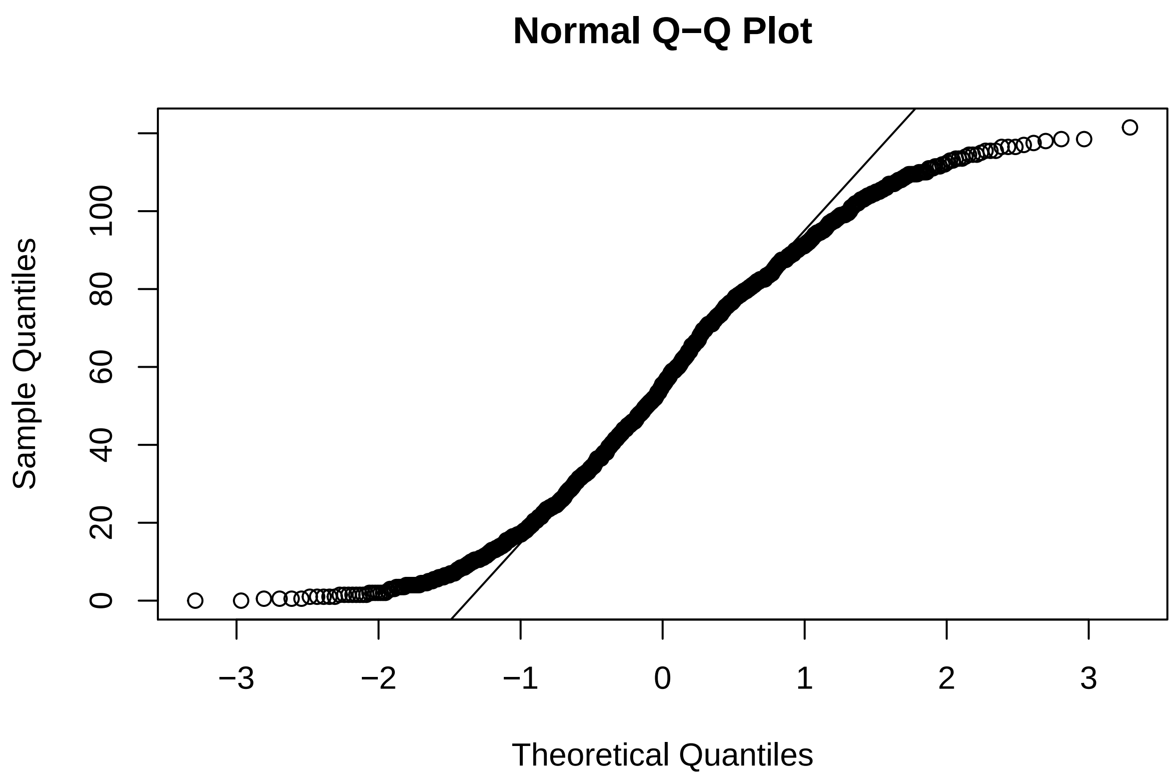
```r
#Histograma
hist(mediasA)
```




```r
#Boxplot
boxplot(mediasA)
```



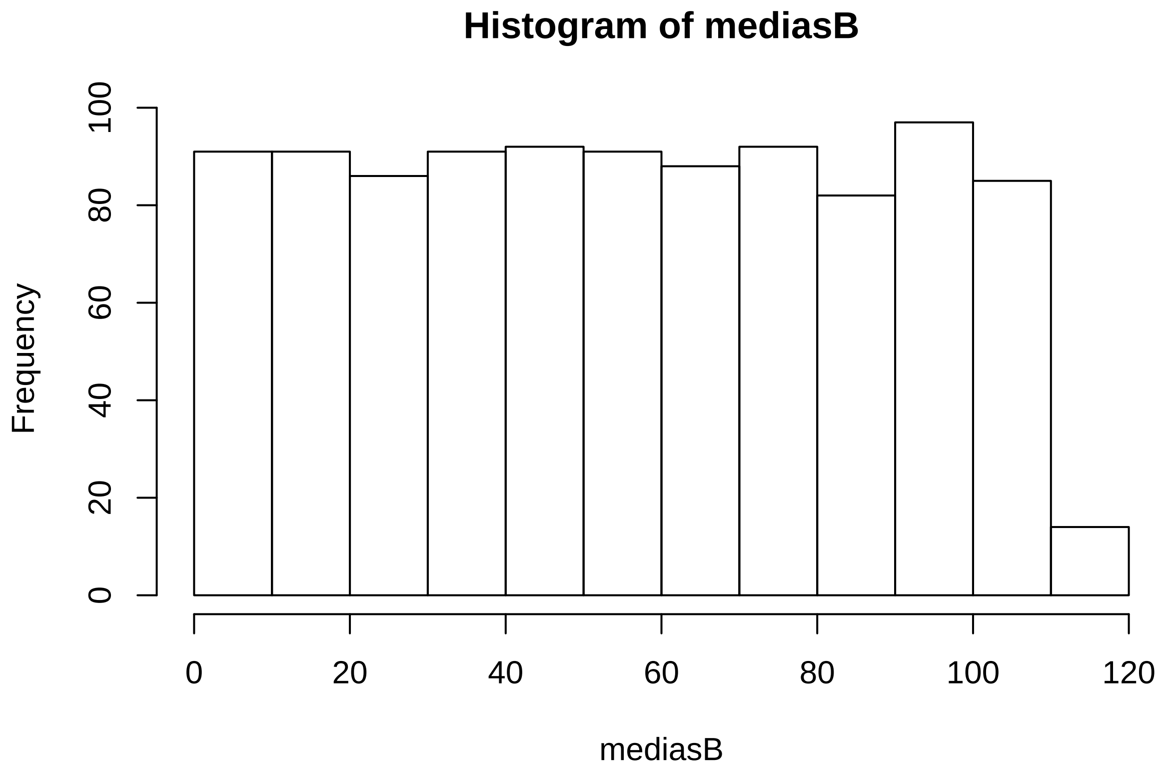

```r
#Q-Q Plot
qqnorm(mediasA)
qqline(mediasA)#La cola del plot
```



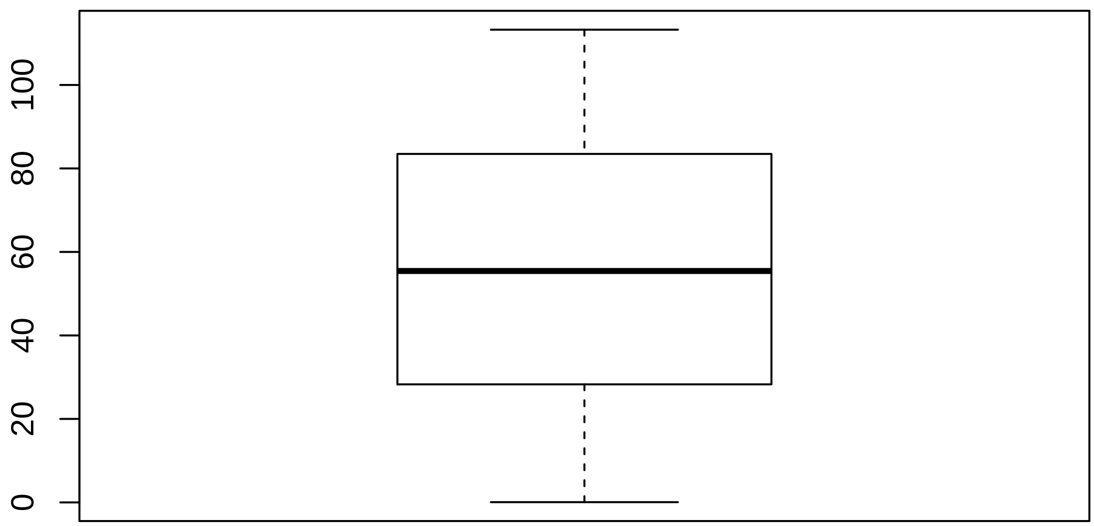
Del histograma se obveserva simetria de las colas livianas. En el Q-Q plot se observa una simetria de las colas livianas. Se observa que el boxplot tiende a una normal.

b)

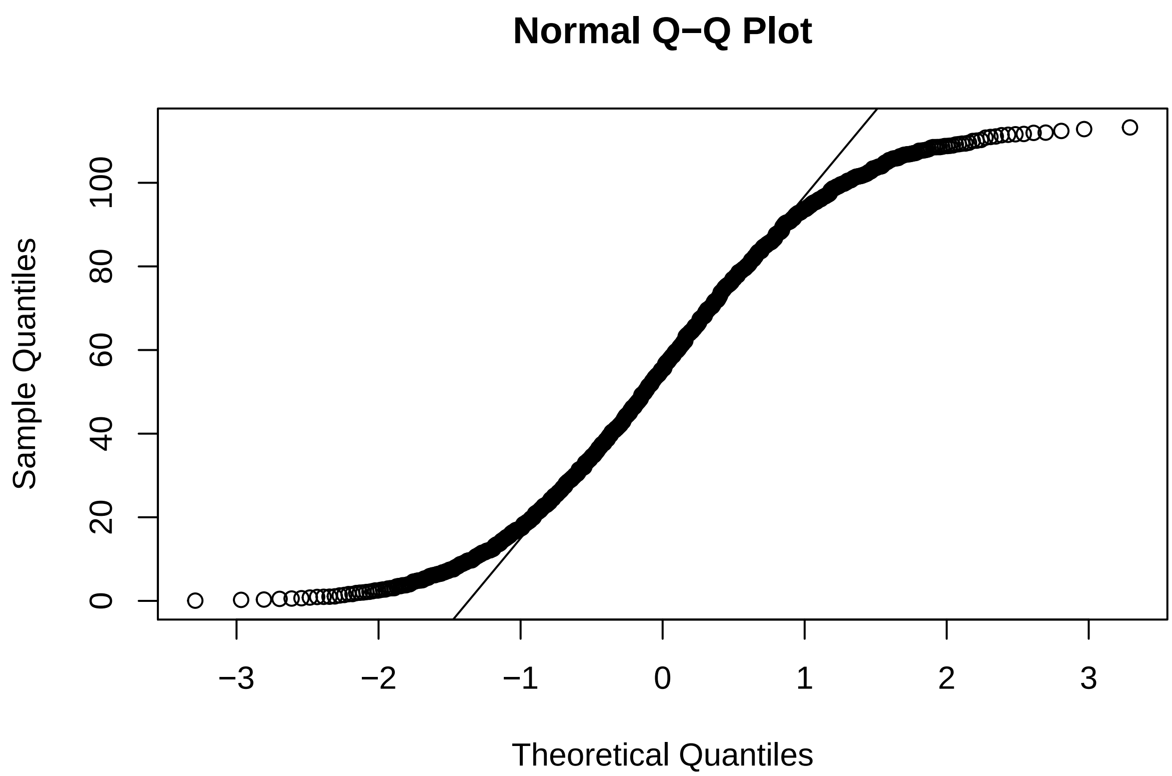
```
#Histograma
hist(mediasB)
```




```
#Boxplot
boxplot(mediasB)
```



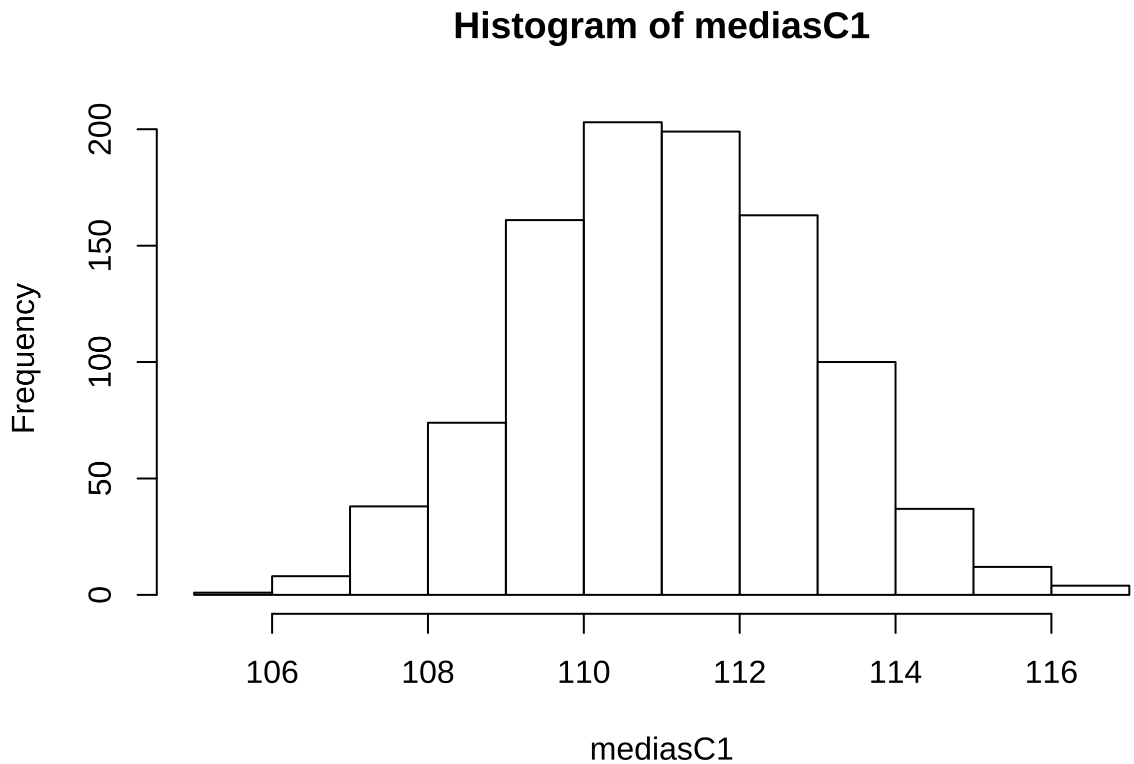

```
#Q-Q Plot
qqnorm(mediasB)
qqline(mediasB)#La cola del plot
```



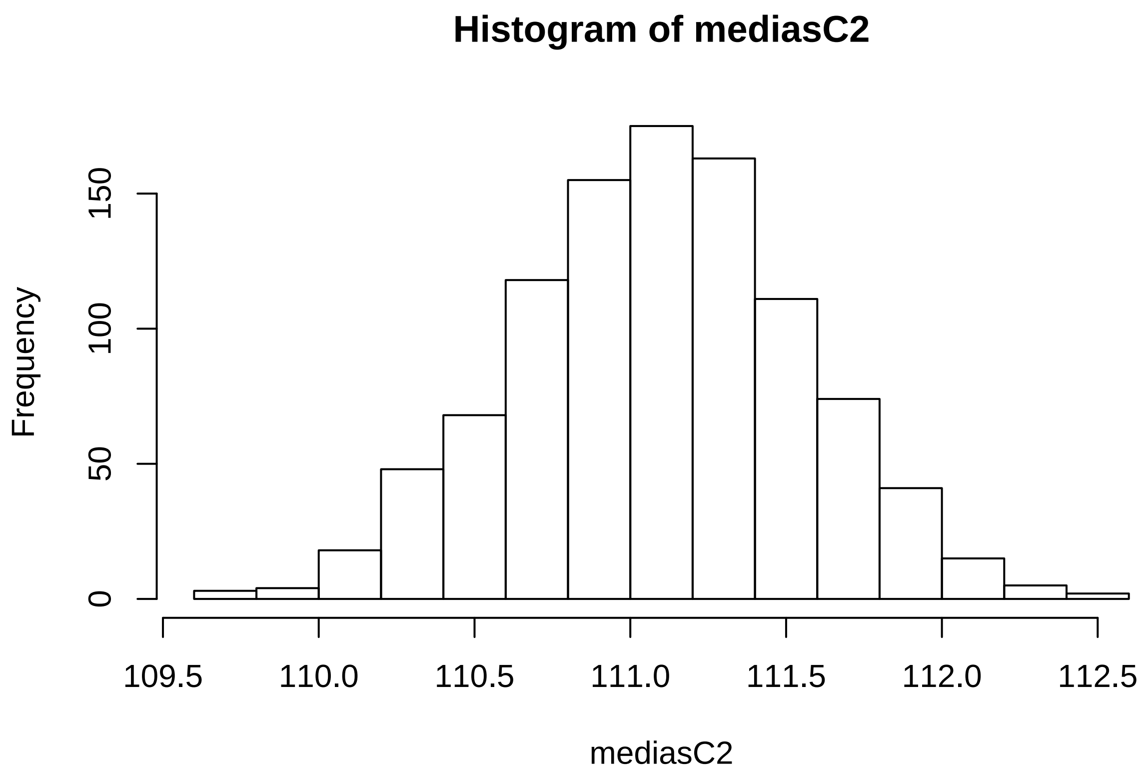
Se observa que el histograma simetria de las colas livianas es mas fuerte. El boxplot es una "perfecta" normal y el Q-Q plot sigue manteniedo la simetria de las colas livianas.

c)

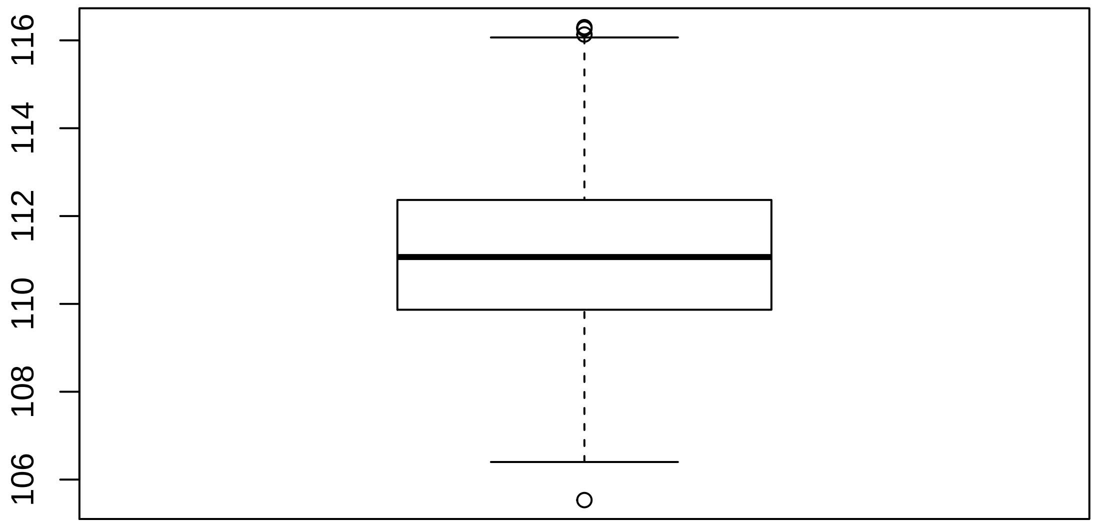
```
#Histograma
hist(mediasC1)
```



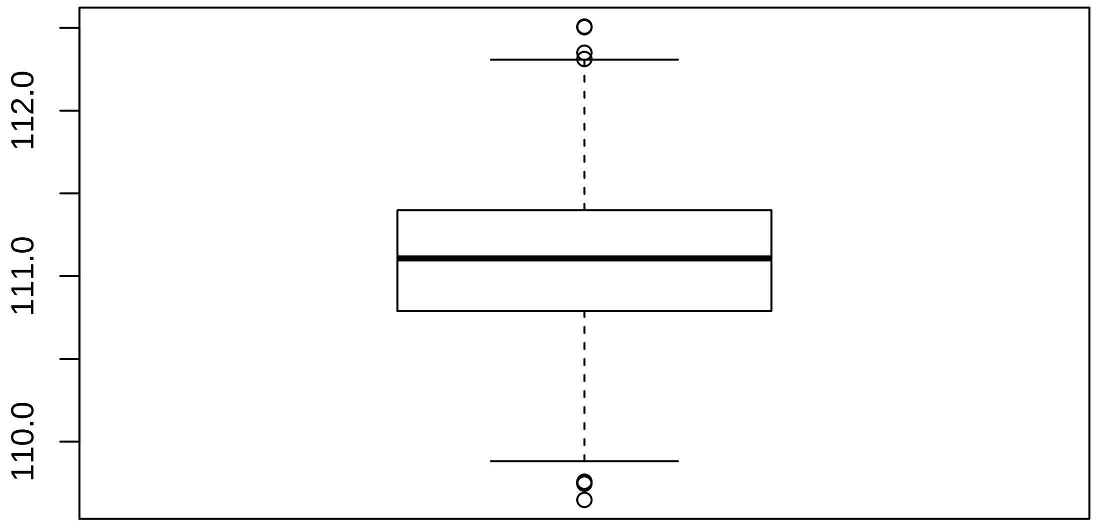

```
hist(mediasC2)
```



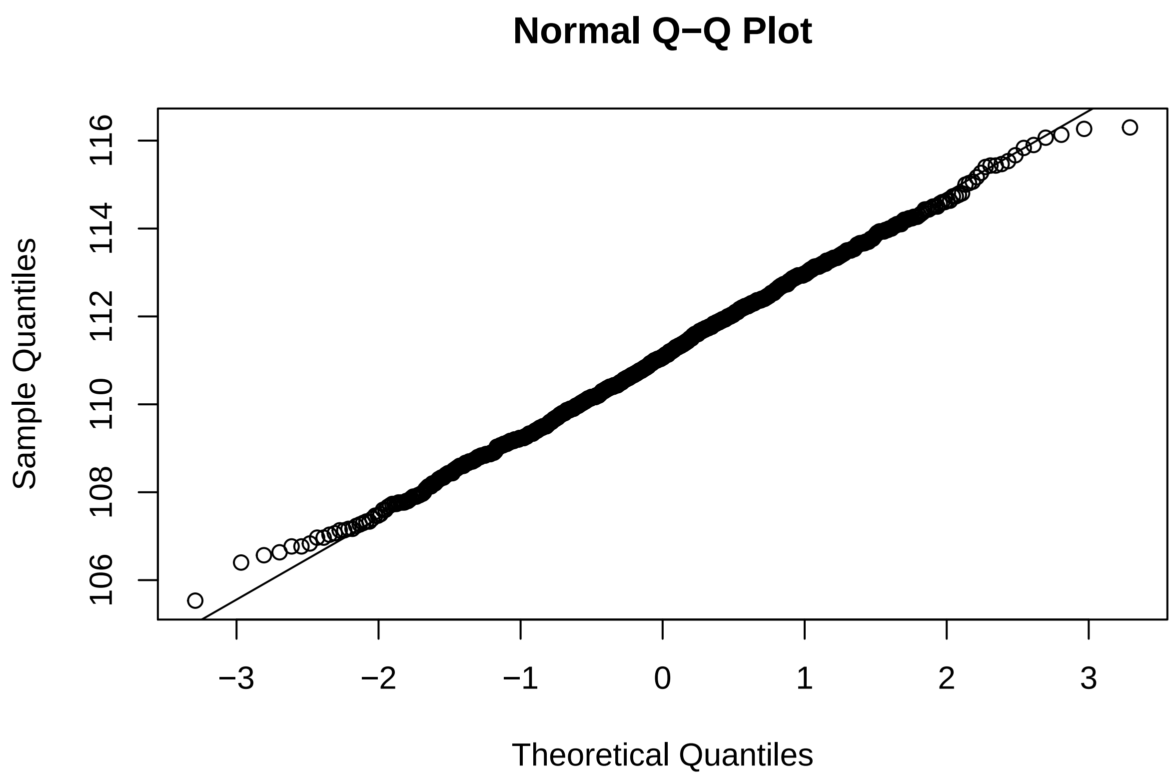

```
#Boxplot
boxplot(mediasC1)
```



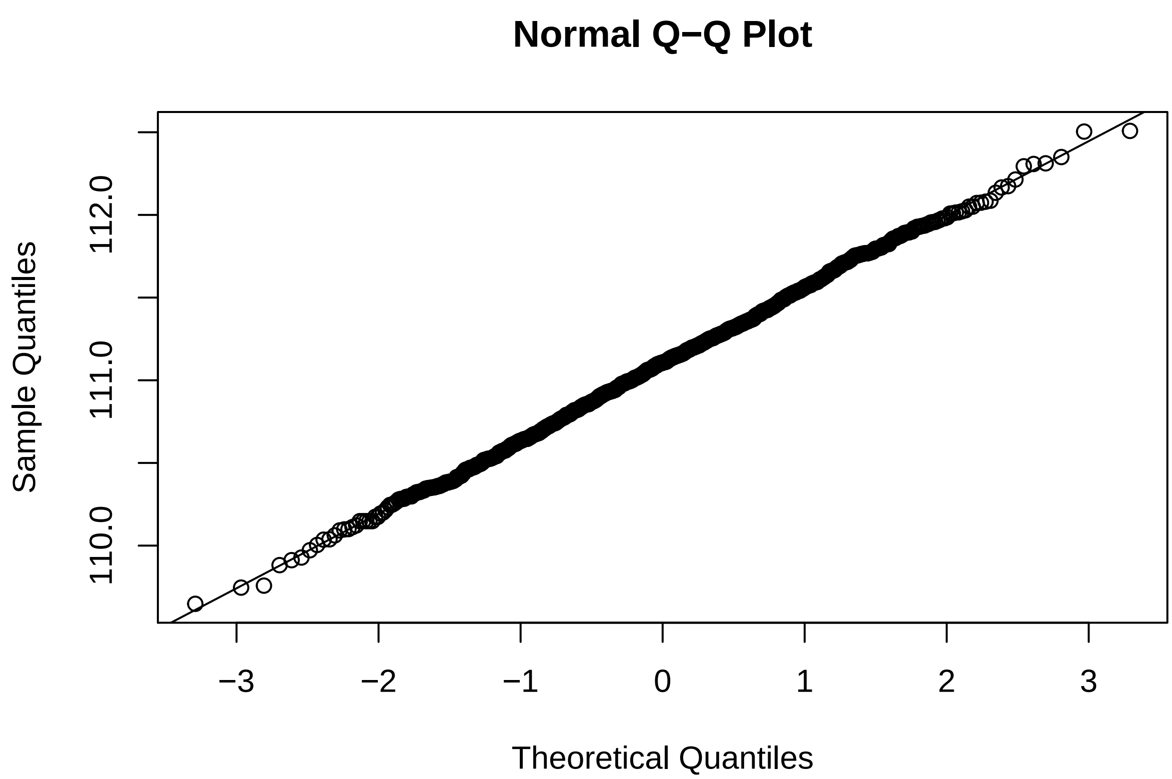

```r
boxplot(mediasC2)
```

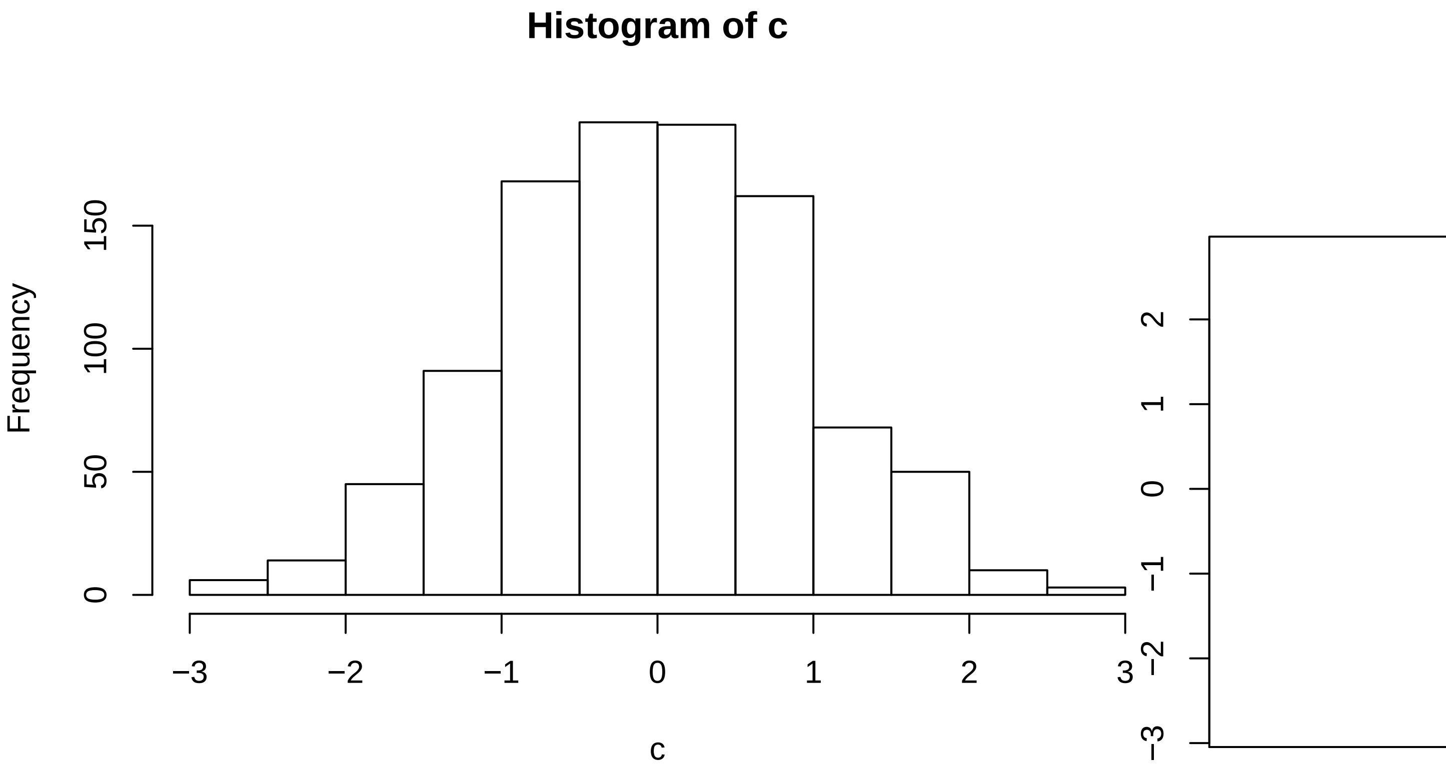


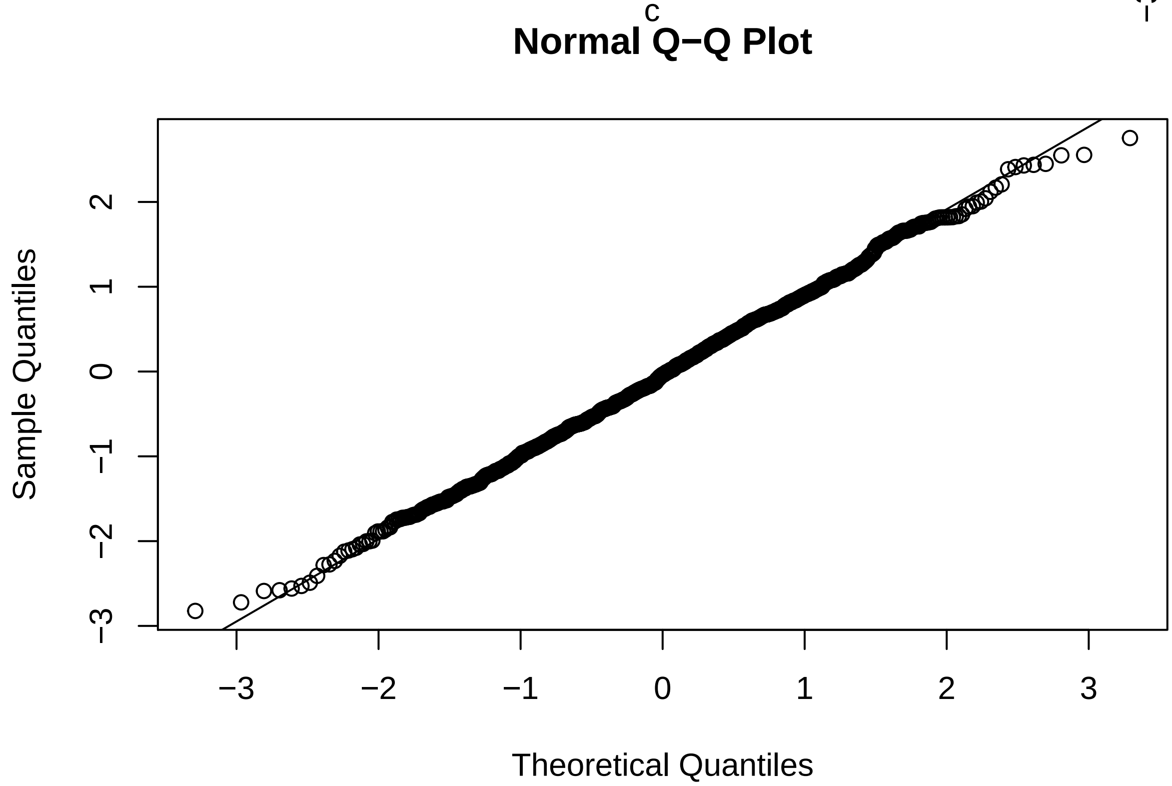

```r
#Q-Q Plot
qqnorm(mediasC1)
qqline(mediasC1)#La cola del plot
```




```r
qqnorm(mediasC2)
qqline(mediasC2)
```



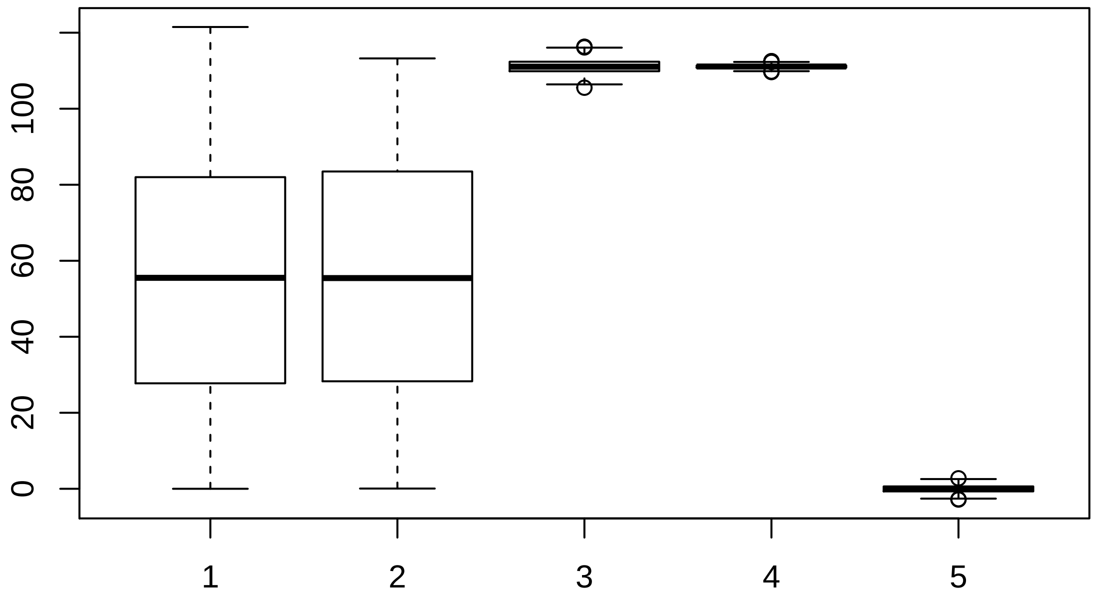




Se nota con muchas mas fuerza en el histograma la distribucion normal. Lo mismo con el boxplot, el cual era el unico hasta el momento que no parecia tender a la normal. Ahora con una gran seguridad podemos confirmar que tiene a una normal con muy pocos outliers. Y el Q-Qplot se aferra con mucha mas fuerza a una distribucion normal.

e)

```r
boxplot(mediasA,mediasB,mediasC1,mediasC2,c)
```



Obsevamos que con una menor muestra se puede verificar su tendencia a una distribucion normal. Si aumentamos el n, el boxplot empieza a tener colas pesadas y deja de tender a una normal.

Punto 3

a)

```r
#mediaX1
mediaX1
```

```
## [1] 55.5015
```

```r
#varX1
varX1
```

```
## [1] 1045.549
```

```r
#mediaX2
mediaX2
```

```
## [1] 55.60766
```

```r
#varX2
varX2
```

```
## [1] 1032.131
```

```r
#mediaX3
mediaX3
```

```
## [1] 111.1069
```

```r
#varX3
varX3
```

```
## [1] 3.302835
```

```r
#mediaX4
mediaX4
```

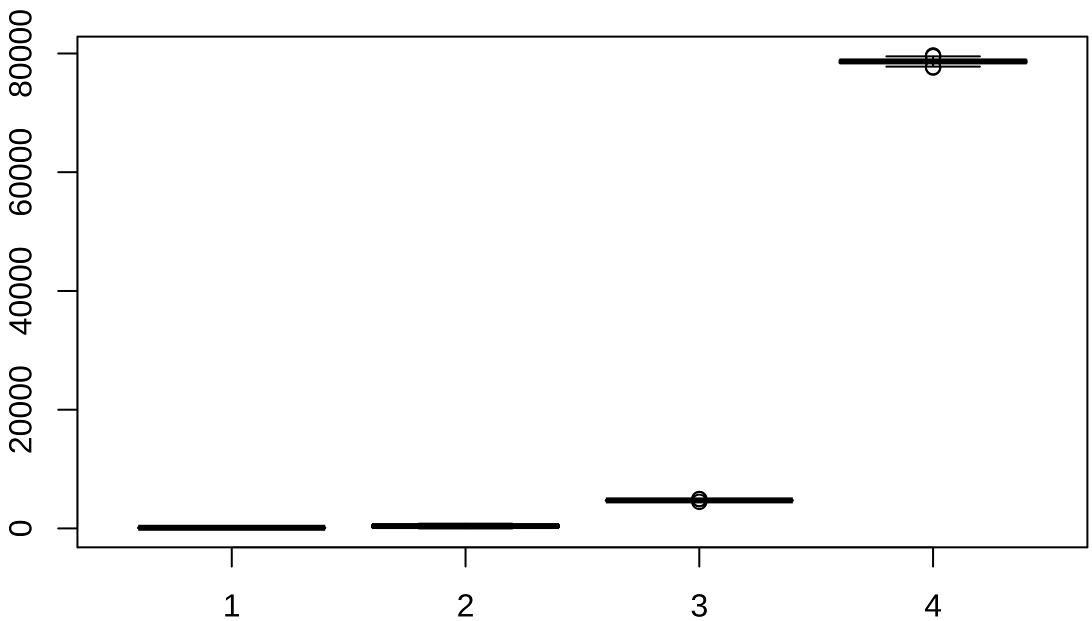
```
## [1] 111.0976
```

```r
#varX4
varX4
```

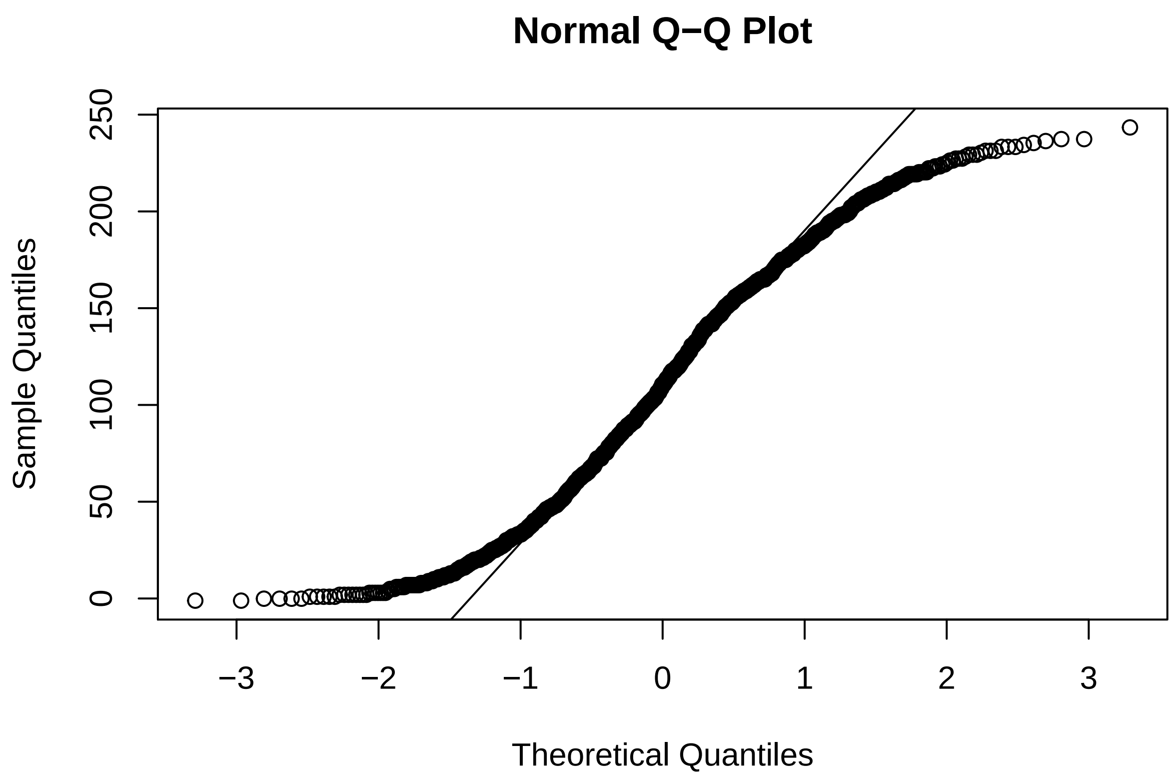
```
## [1] 0.2102568
```

b)

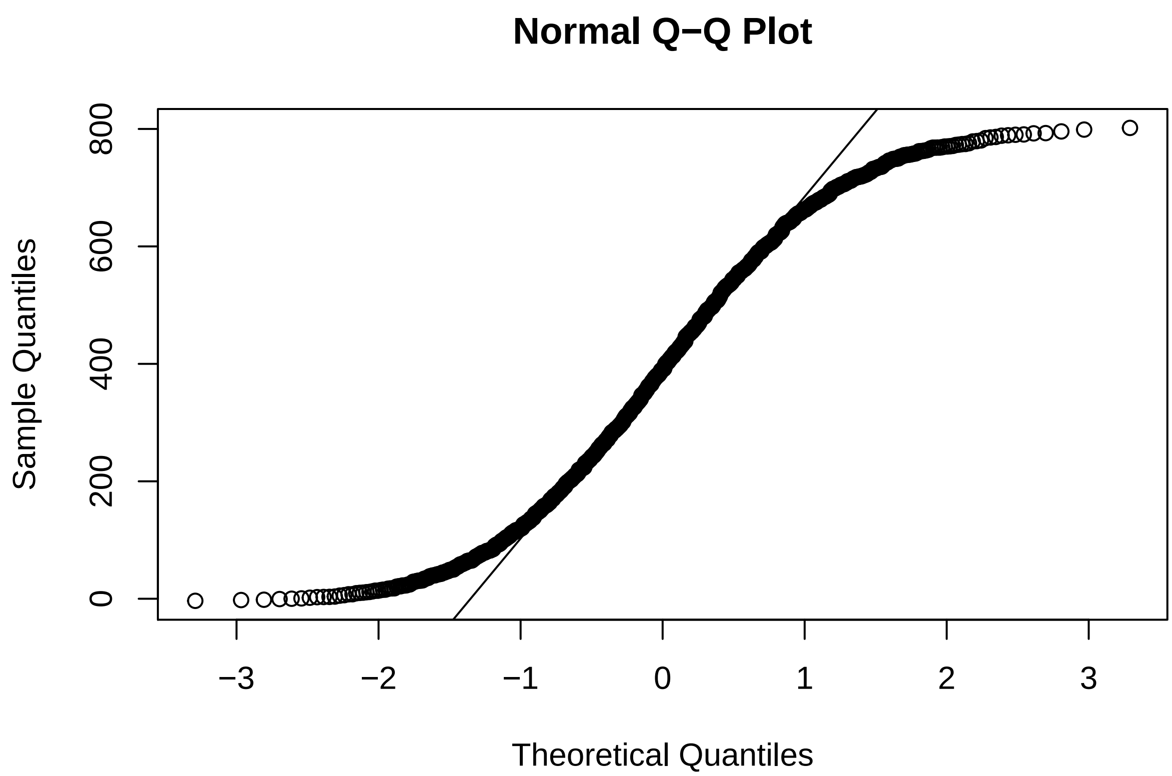
```r
boxplot(transformacionesA, transformacionesB, transformacionesC1, transformacionesC2)
```



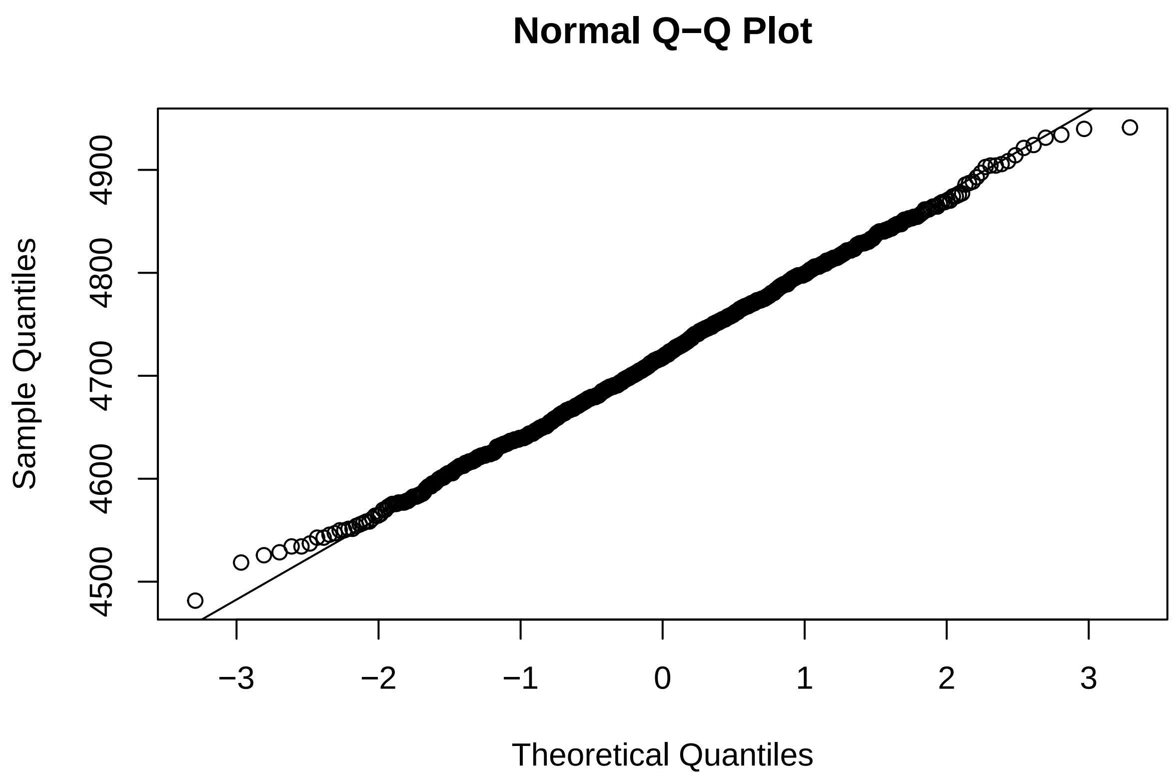

```r
qqnorm(transformacionesA)
qqline(transformacionesA)
```



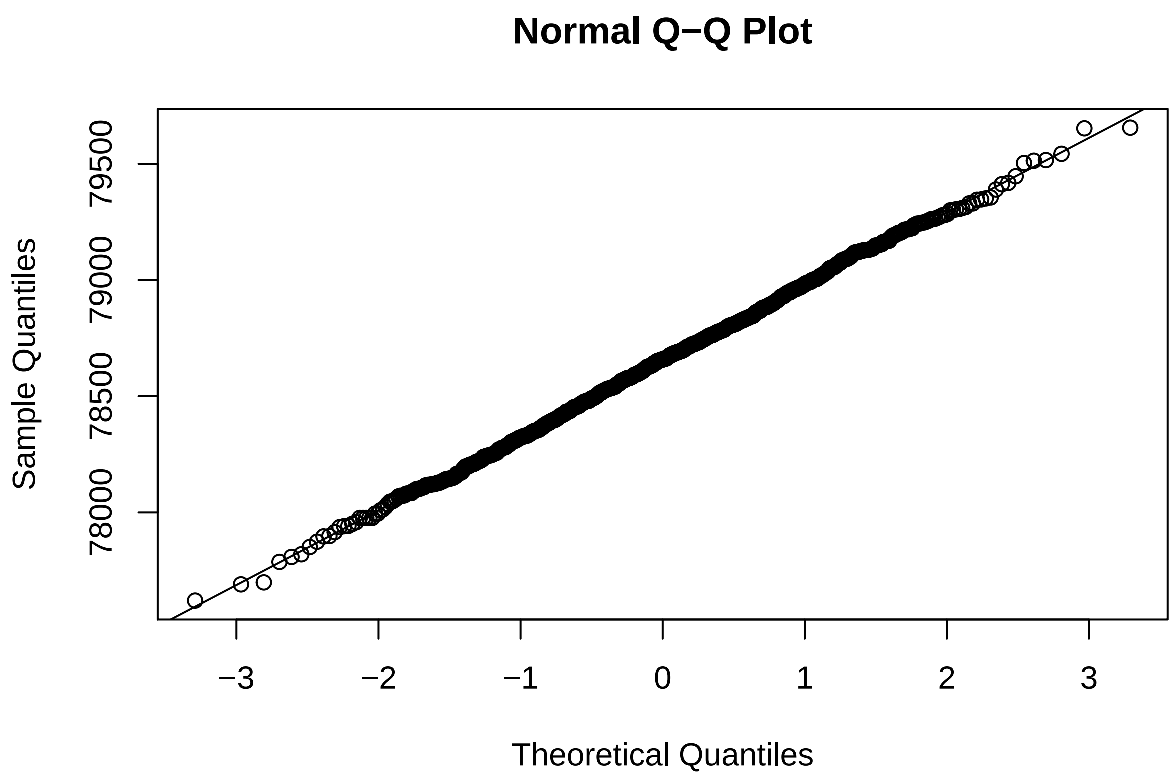

```r
qqnorm(transformacionesB)
qqline(transformacionesB)
```



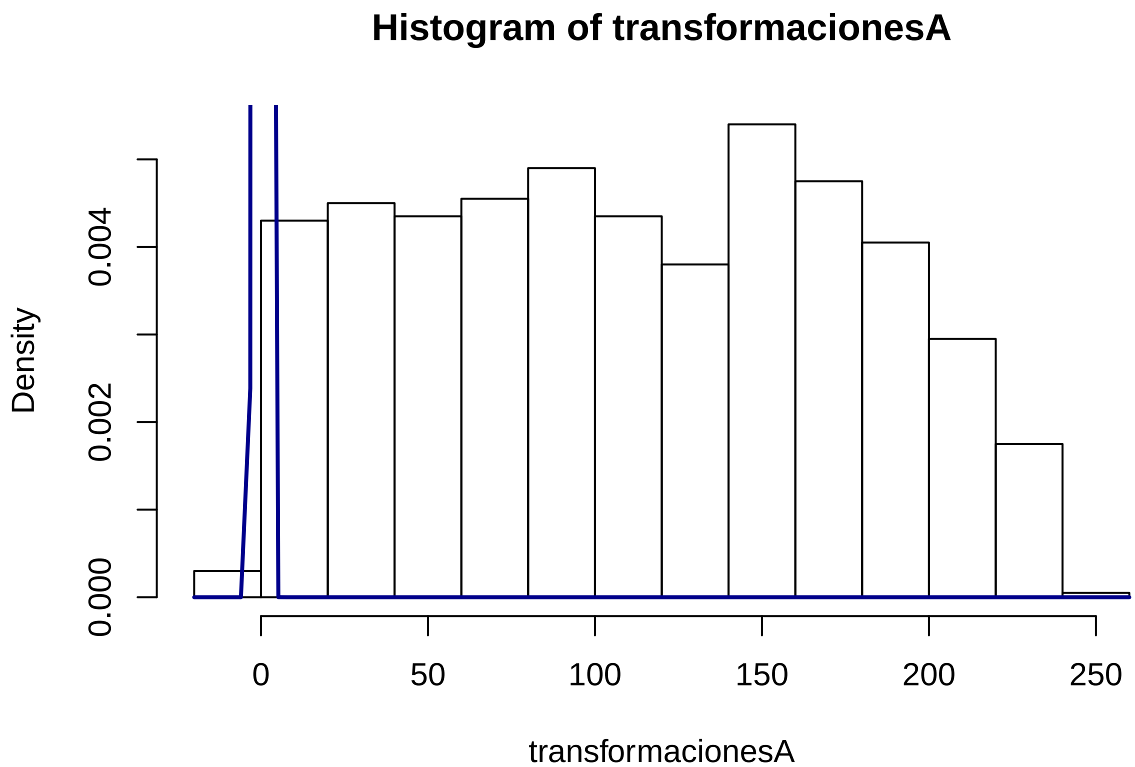

```r
qqnorm(transformacionesC1)
qqline(transformacionesC1)
```

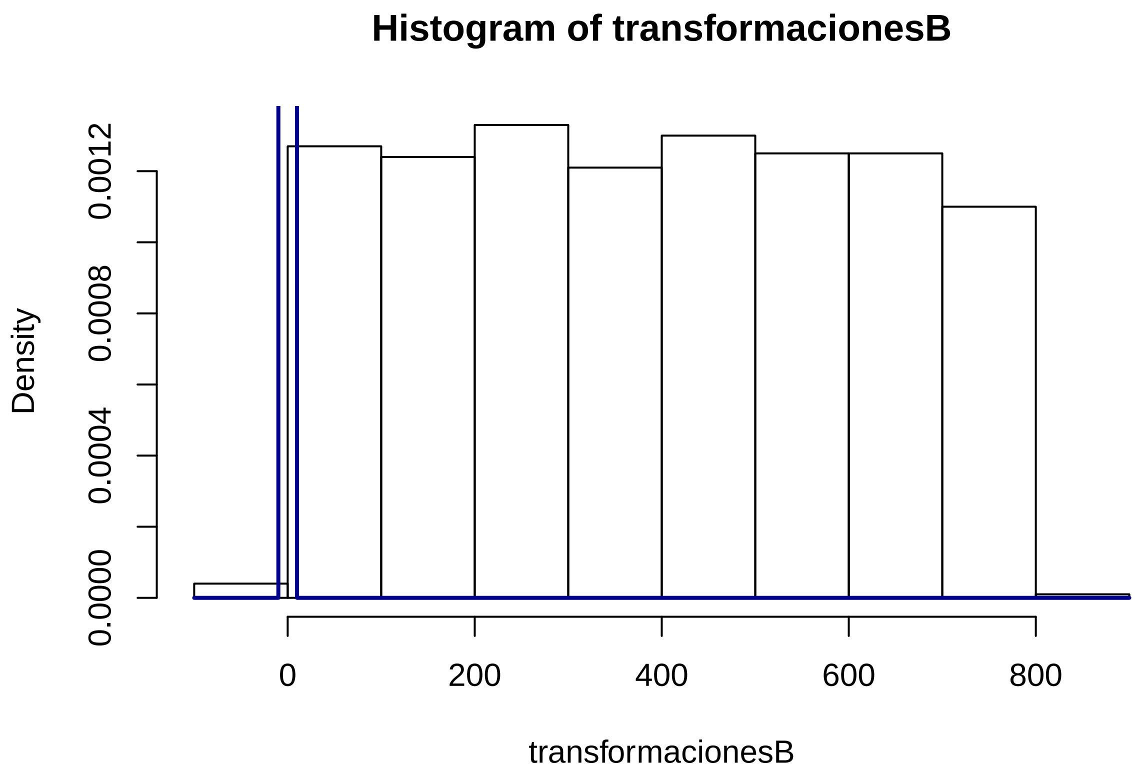


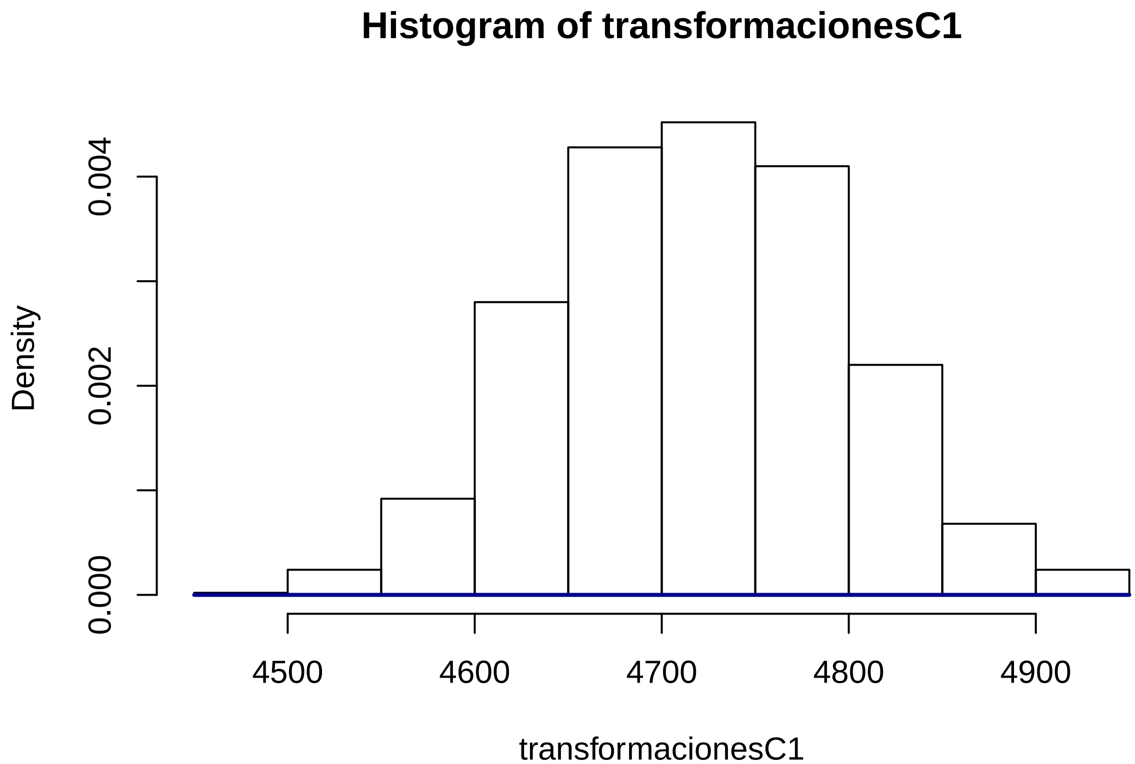

```
qqnorm(transformacionesC2)
qqline(transformacionesC2)
```

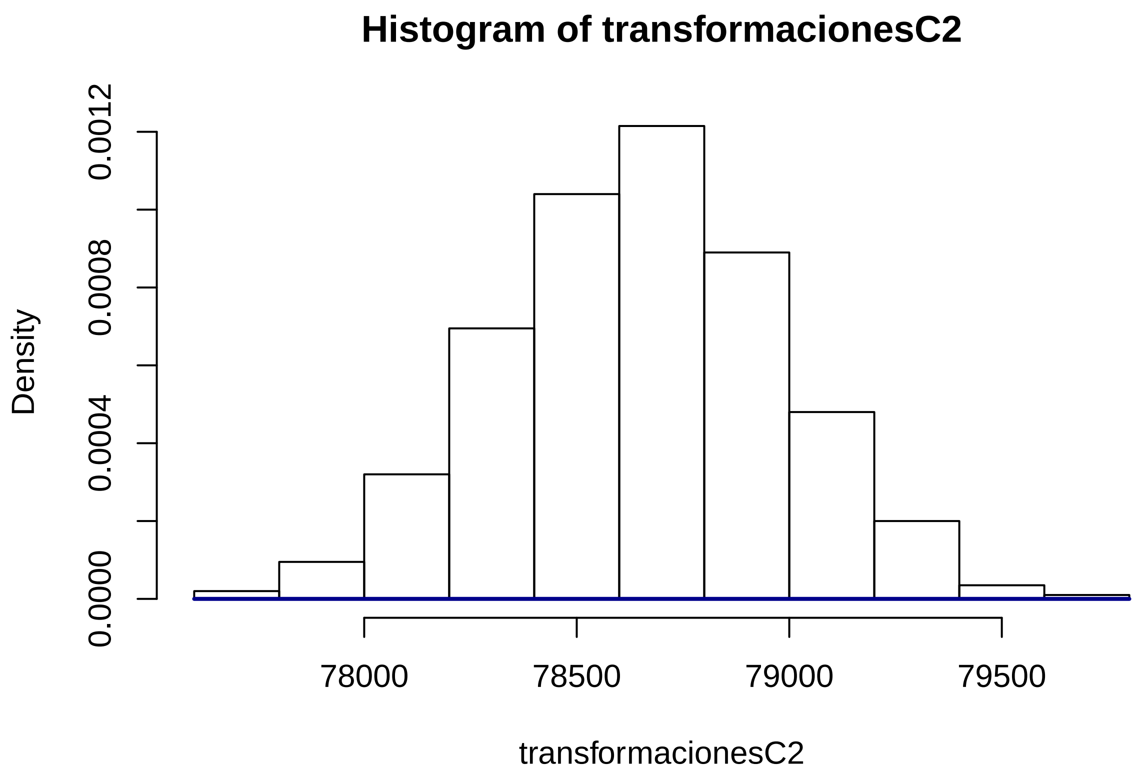


c)


Histogram of transformacionesA
Density
0.000
0.002
0.004
0
50
100
150
200
250
transformacionesA


Histogram of transformacionesB
Density
0.0000
0.0004
0.0008
0.0012
0
200
400
600
800
transformacionesB





Se observa en el histograma `mediasA` y en `mediasB`, poseen una asimetria a la derecha, dado a que la muestra parece no ser lo suficientemente grande, lo cual no nos da informacion respecto a si tiende a una normal. Con respecto a los histogramas, `mediasC1` y `mediasC2`, se puede observar como los grafico represtan una "campana" al estilo de una normal, lo cual nos informa, que nuestra muestra fue lo suficientemente buena, para poder estimar dicha distribucion aleatoria. Siendo la mas precisa, el histograma de `mediasC1`