

Trabajo Práctico Probabilidad y Estadística (c)

Luis Greco - Nicolas Hertzulis - Ruslan Sanmartin Sobol

20 de noviembre de 2017

Contents

Ejercicio 1	2
Ejercicio 2	4
a)	4
b)	6
c)	8
d)	11
e)	13
Ejercicio 3	14
a)	14
b)	14
c)	17
d)	19
Ejercicio 4	20
Punto 1	20
Punto 2	21
a)	21
b)	23
c)	25
d)	28
e)	29
Punto 3	29
a)	29
b)	30
c)	33
d)	35



Figure 1: Tendencia Probabilística Revolucionaria

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

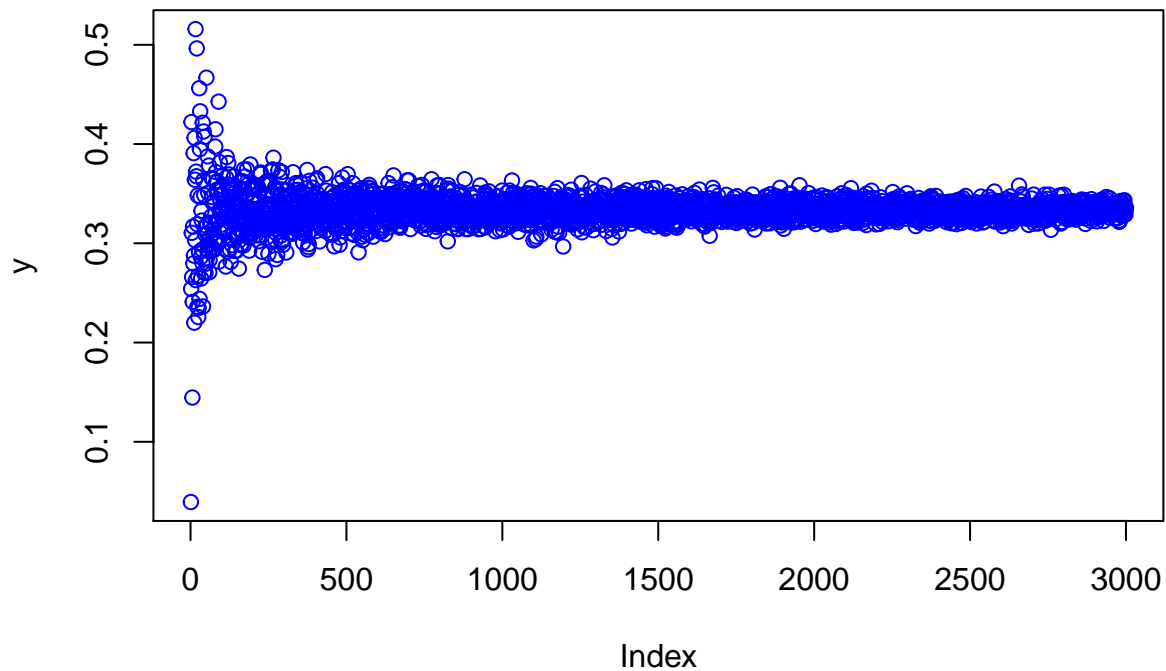
```
set.seed(1109)
```

Ejercicio 1

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

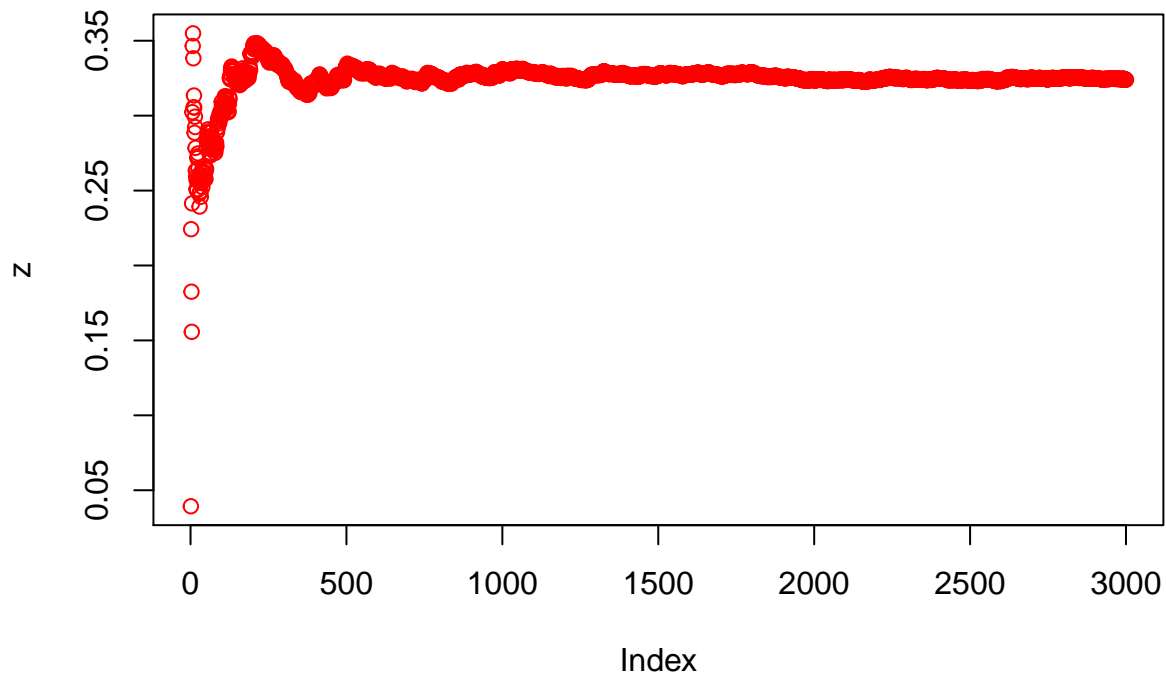
Ahora veamos los gráficos. Este primer gráfico es el que hicimos con el set.seed() “global”.

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



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

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



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 subíndice indicado, pero luego las simulaciones siguientes irán cambiando en función de la anterior, o sea que al fijar `set.seed()` localmente, en cada generación de números no necesariamente tienen que dar iguales.

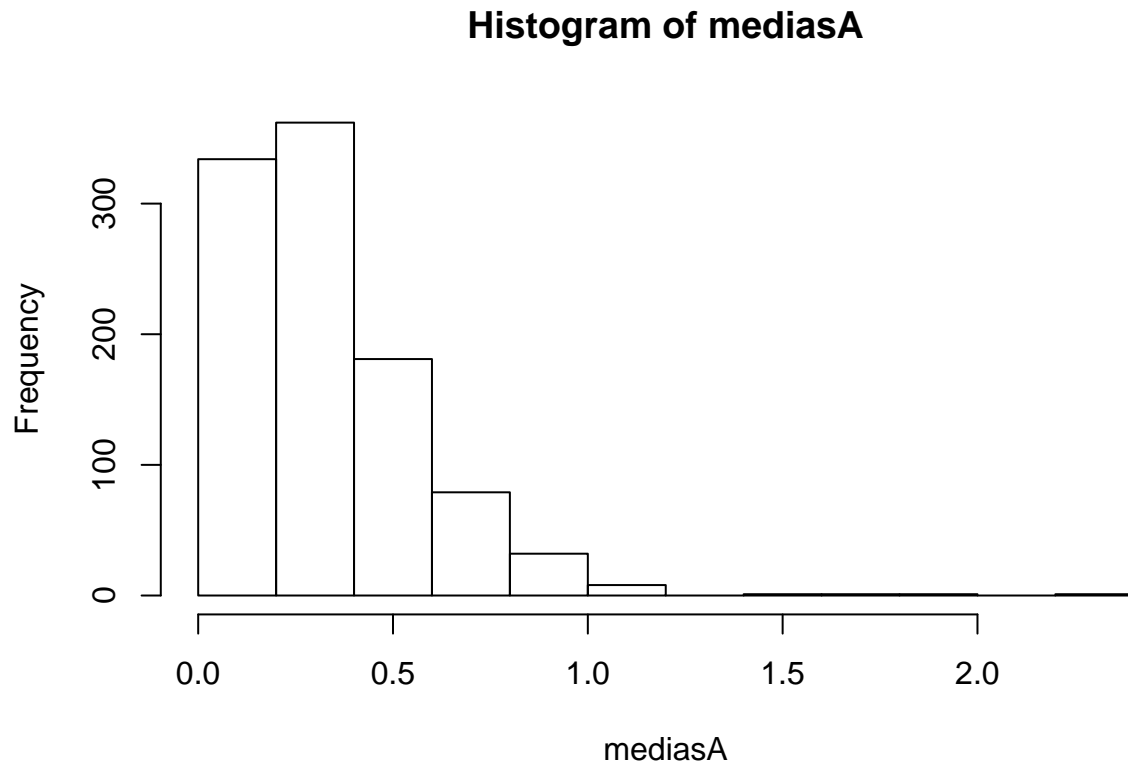
También pudimos observar el cumplimiento de la Ley de los Grandes Números, al ver que a medida que aumenta el tamaño de la muestra, la media de la misma tiende al valor de la esperanza teórica.

Por lo que explicamos anteriormente sobre el uso de `set.seed()`, en el segundo gráfico se nota una tendencia muchísimo más marcada.

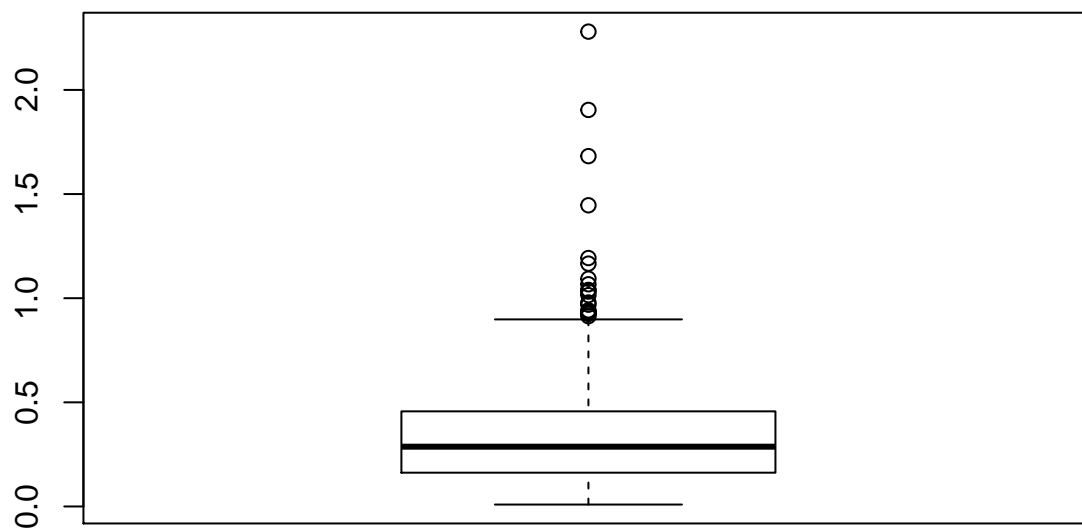
Ejercicio 2

a)

```
hist(mediasA)
```

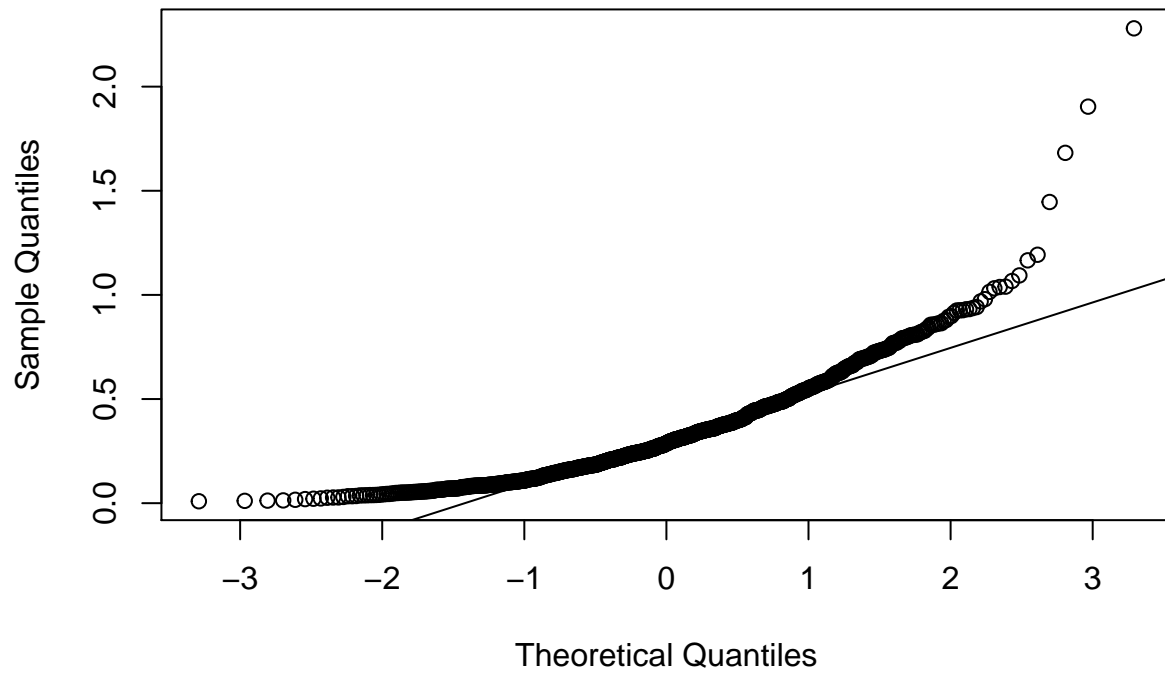


```
boxplot(mediasA)
```



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

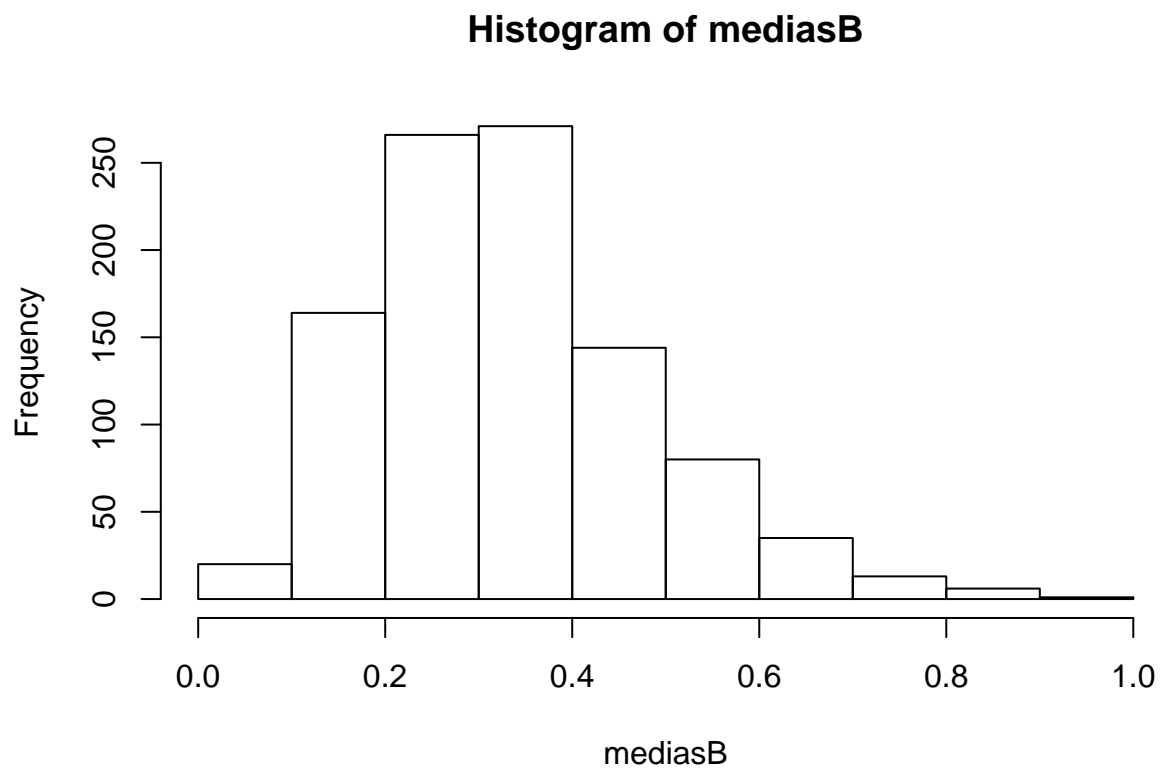
Normal Q-Q Plot



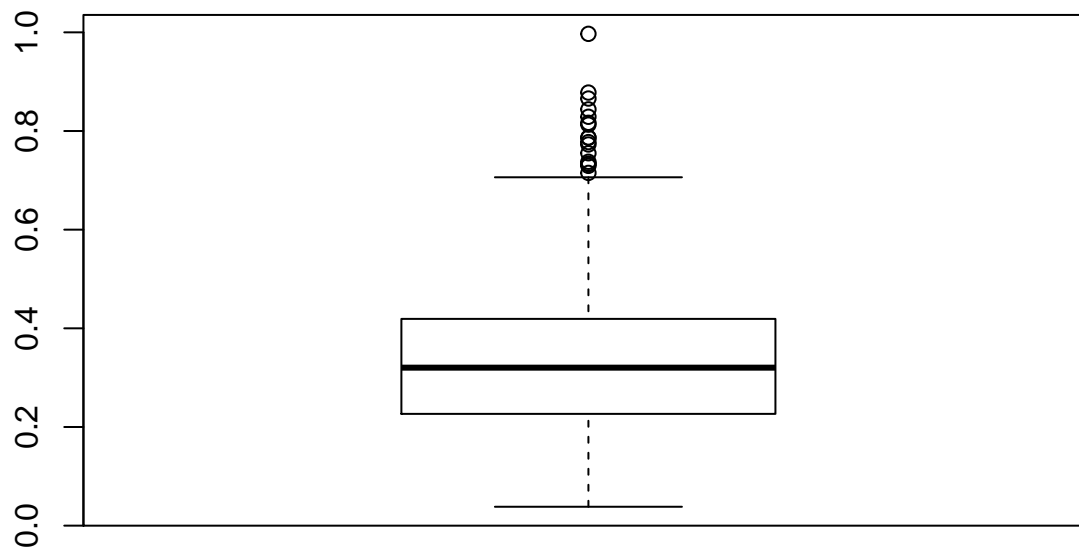
Observamos que tanto el histograma, como en el boxplot y q-q plot, los gráficos son asimétricos, lo que es consistente con una distribución exponencial como esta.

b)

```
hist(mediasB)
```



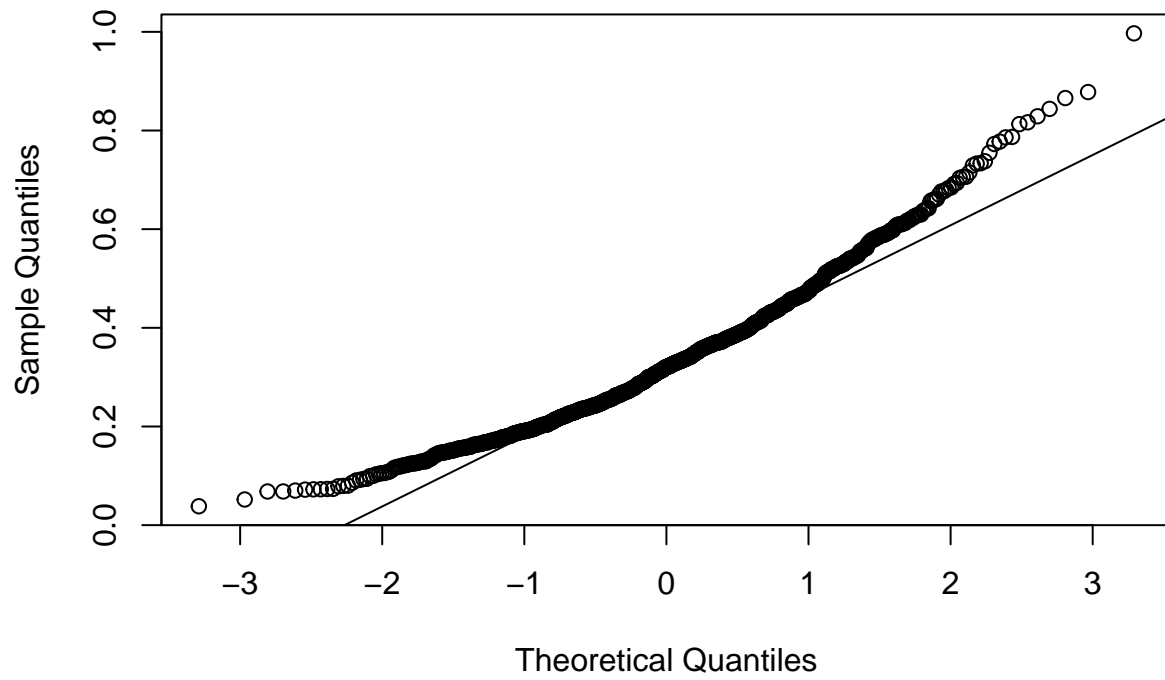
```
boxplot(mediasB)
```



```
qqnorm(mediasB)
```

```
qqline(mediasB)
```

Normal Q-Q Plot

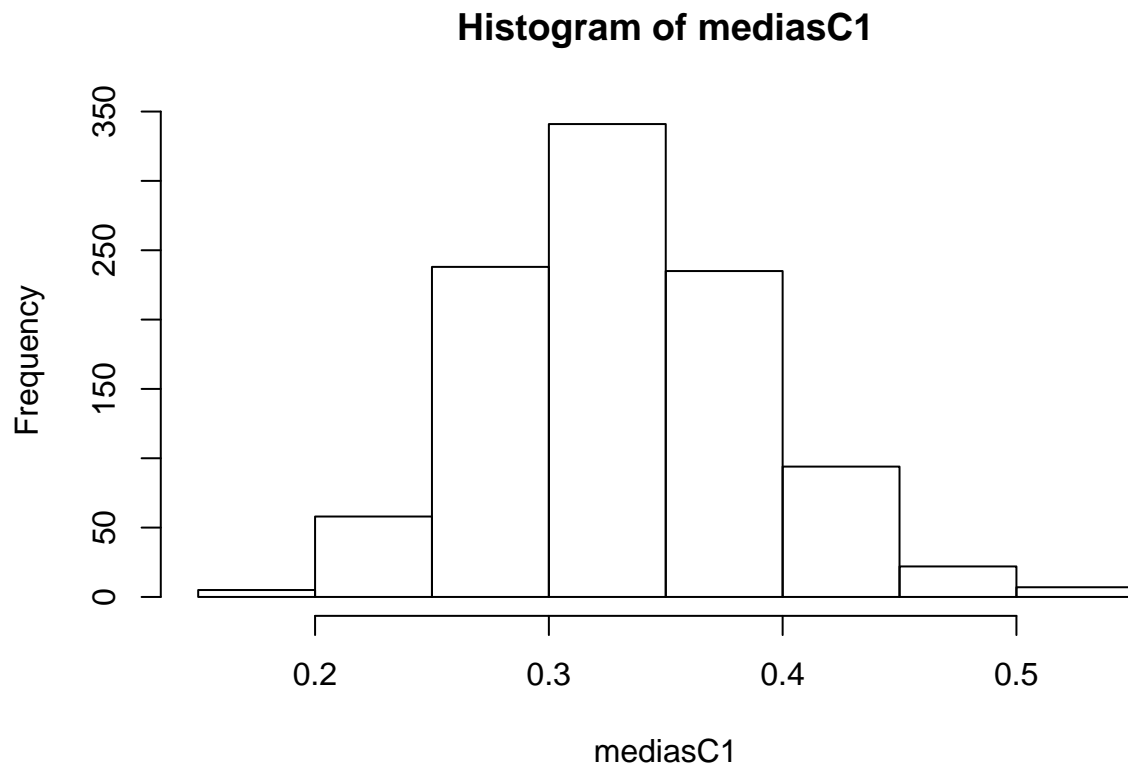


Se observa que el histograma empieza a tender a una distribución normal. En el boxplot pudimos notar colas más pesadas, y por último el Q-Q plot sigue manteniendo una cierta simetría, lo que hace que se vaya pareciendo a una distribución Normal.

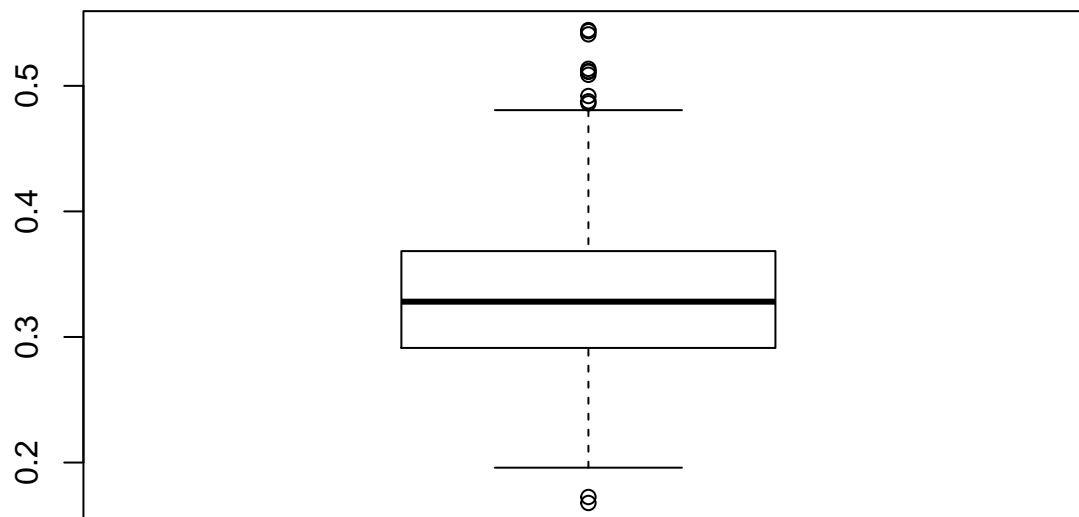
c)

Con n = 30

```
hist(mediasC1)
```



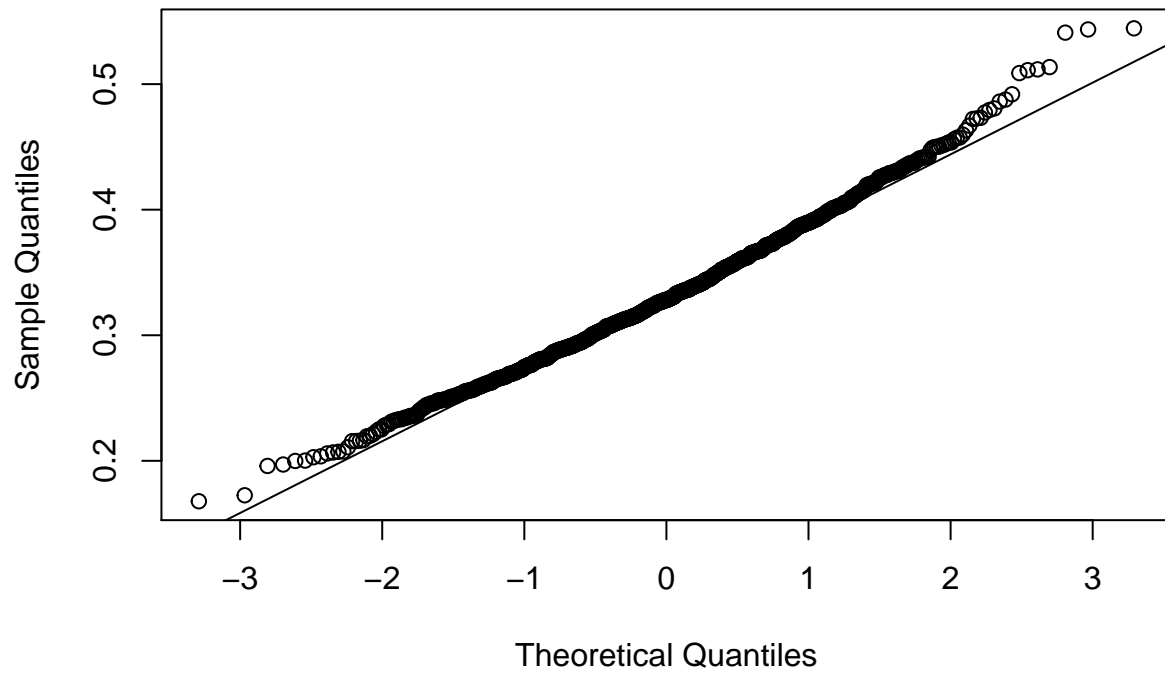
```
boxplot(mediasC1)
```



```
qqnorm(mediasC1)
```

```
qqline(mediasC1)
```

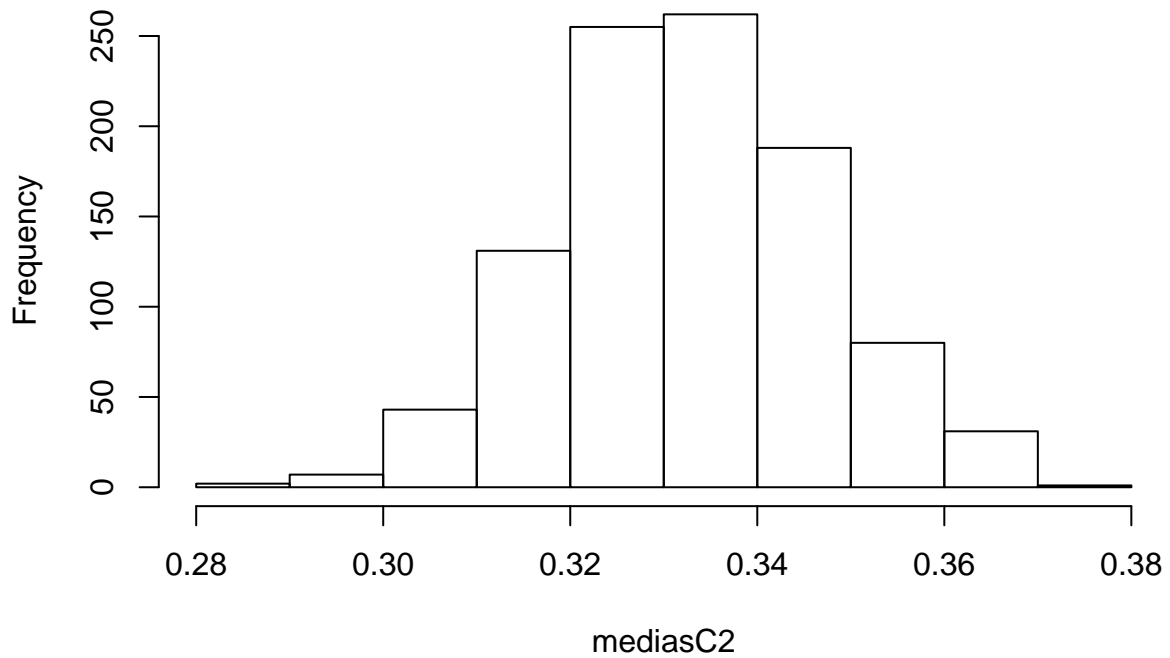

Normal Q-Q Plot



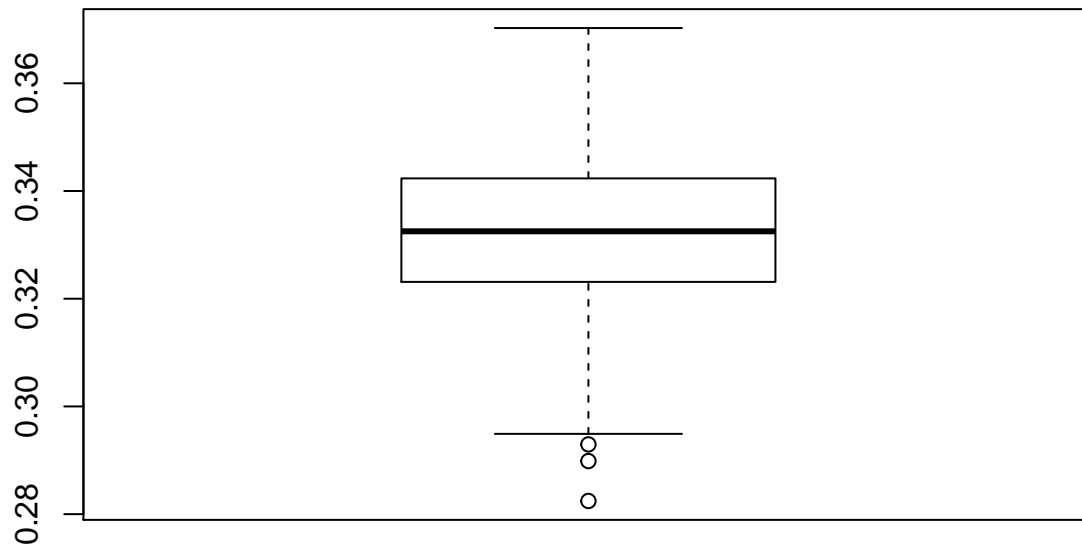
Con n = 500

```
hist(mediasC2)
```

Histogram of mediasC2

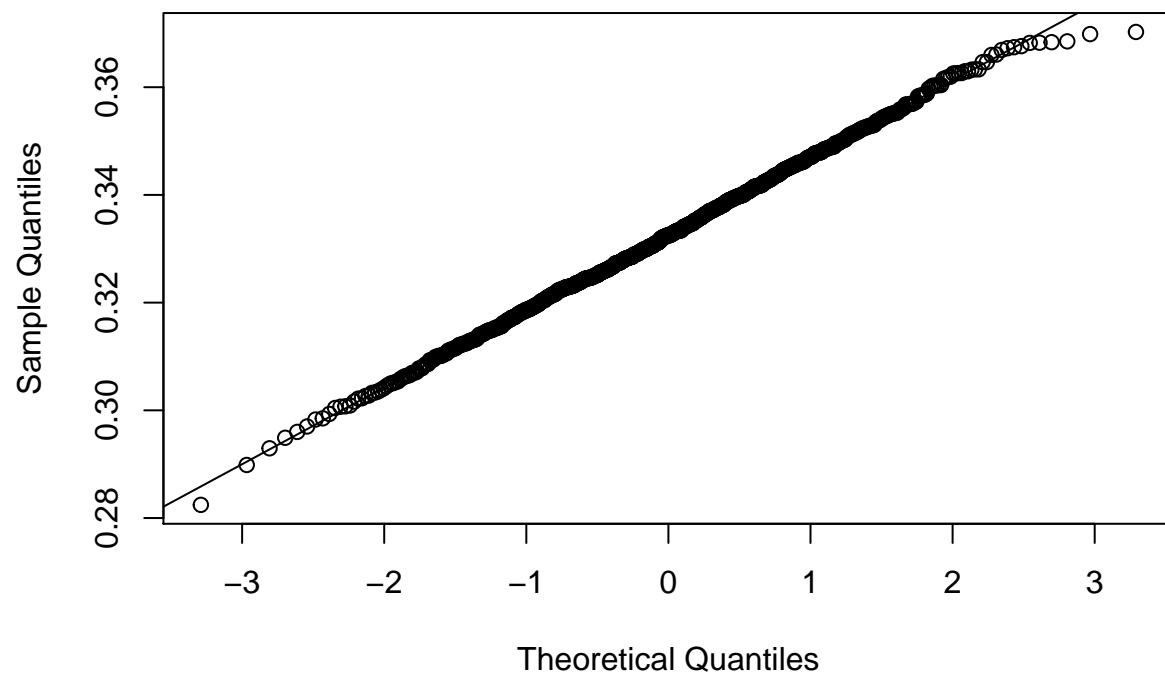


```
boxplot(mediasC2)
```



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

Normal Q-Q Plot

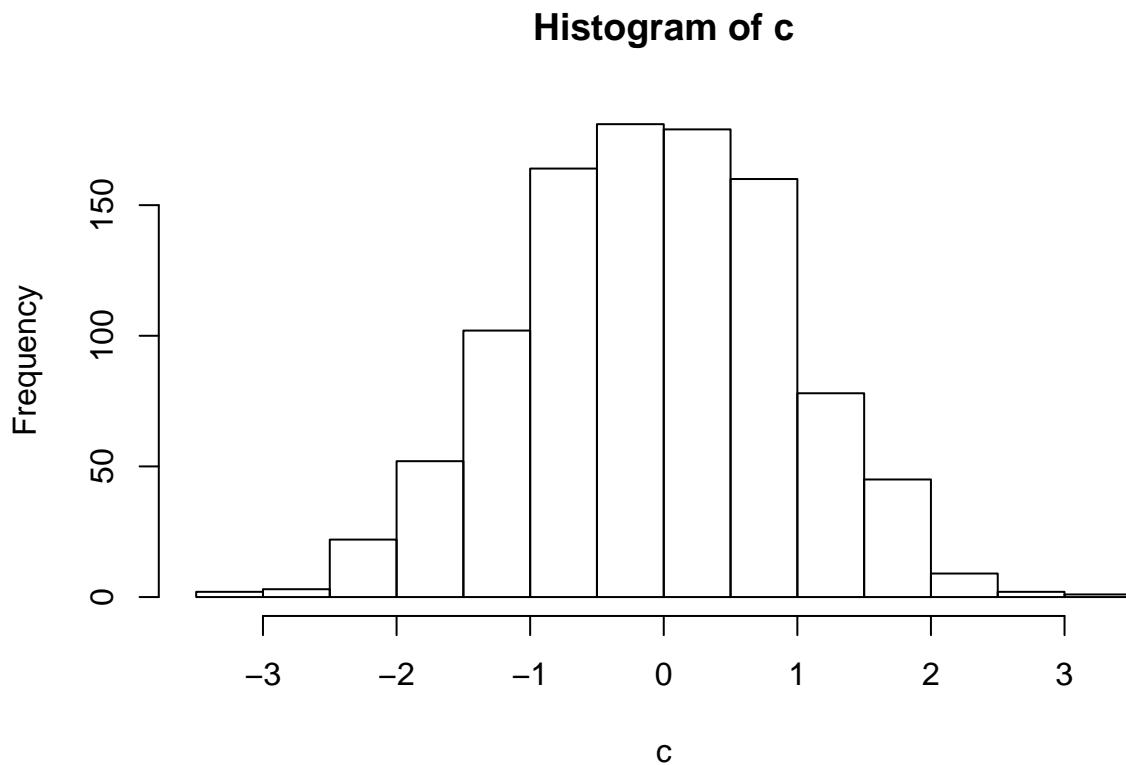


d)

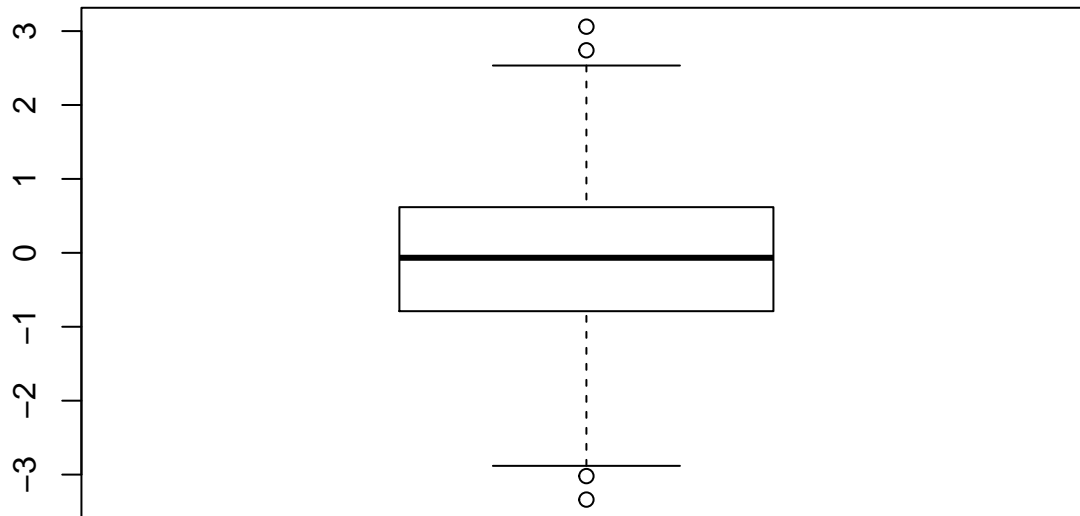
Acá nos pareció oportuno adjuntar el gráfico de una distribución normal standard a modo de comparación con los gráficos anteriores, y pudimos notar el cumplimiento de la Ley de los Grandes Números, al ver que a medida que aumentamos el “n” se van pareciendo más a una distribución normal.

Si aumentáramos más el “n”, los gráficos serían aún más parecidos a los de una normal. Algo similar pudimos observar en los boxplots, en los primeros tres gráficos no se notaba tanto una forma normal, pero con un “n” más grande también lo empezamos a notar y con $n = 500$ ya podemos decir con seguridad que el boxplot se parece a una normal.

```
c <- rnorm(1000, 0, 1)
hist(c)
```

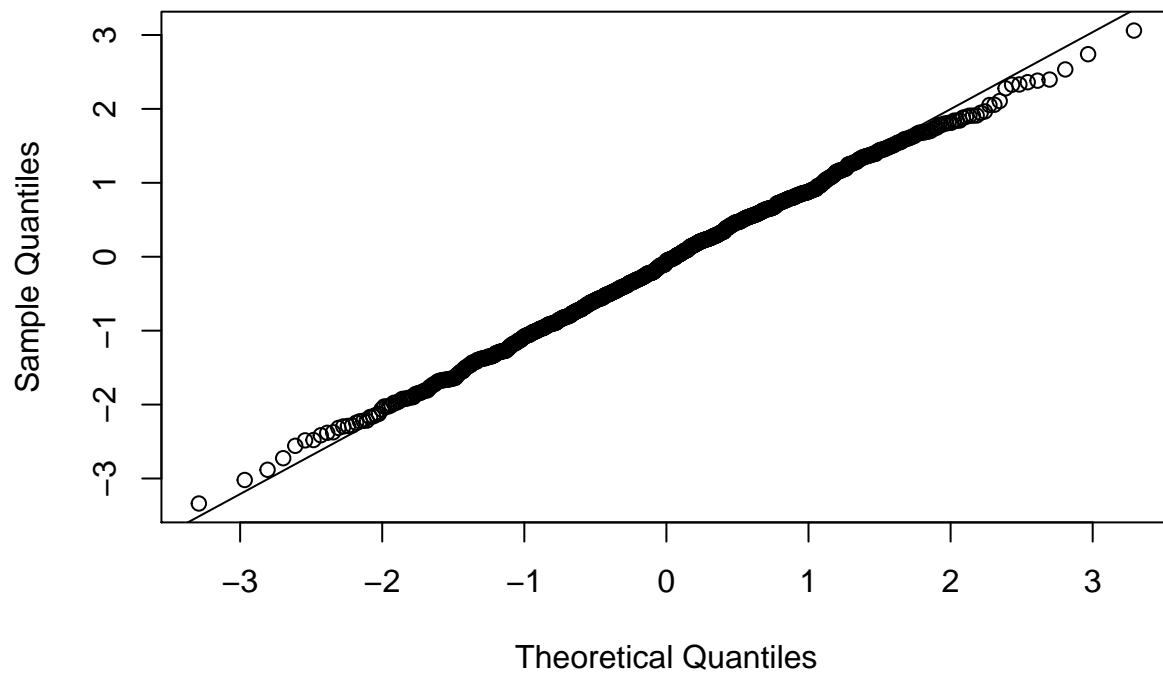


```
boxplot(c)
```



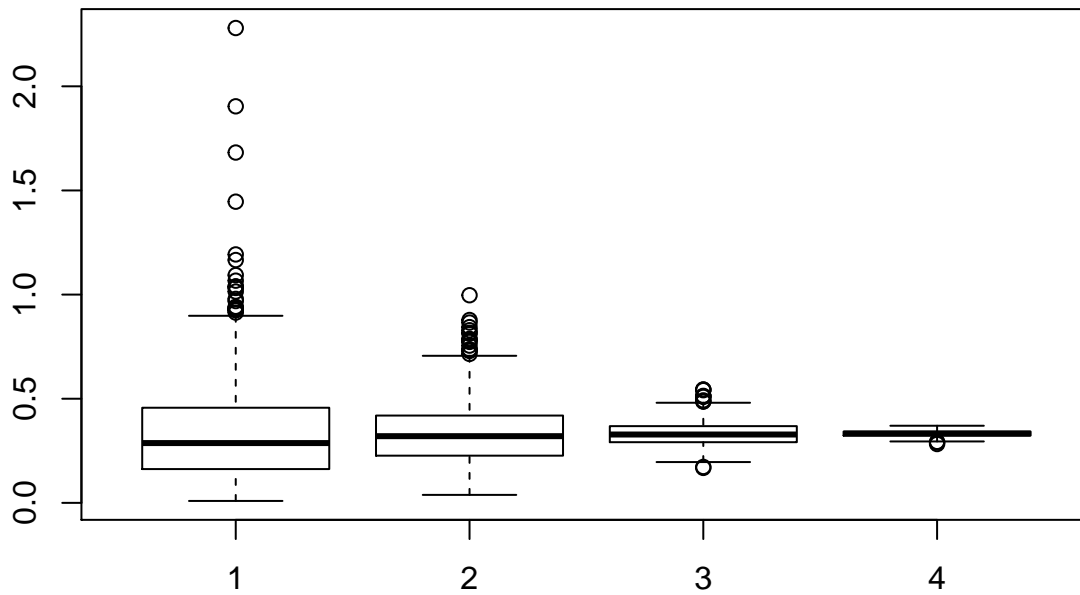
```
qqnorm(c)
qqline(c)
```

Normal Q-Q Plot



e)

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



Obseavamo que a mayor tamaño de la muestra, mayor es su tendencia a una distribución normal.

En el punto a) poseía colas pesadas, tendiendo a la cola superior y con un número mayor de outliers. Luego cada vez que aumentábamos el tamaño de nuestras muestras estos outliers disminuían y se iban haciendo cada vez más simétricos los boxplots.

En el gráfico no se pueden apreciar con mucho detalle los boxplots correspondientes a los “n” más grandes por la escala, sin embargo ya pudimos ver anteriormente como su forma se correspondía con la de una distribución normal.

Ejercicio 3

a)

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

```
mediaX1
```

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

```
varX1
```

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

```
mediaX2
```

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

```
varX2
```

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

```
mediaX3
```

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

```
varX3
```

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

```
mediaX4
```

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

```
varX4
```

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

b)

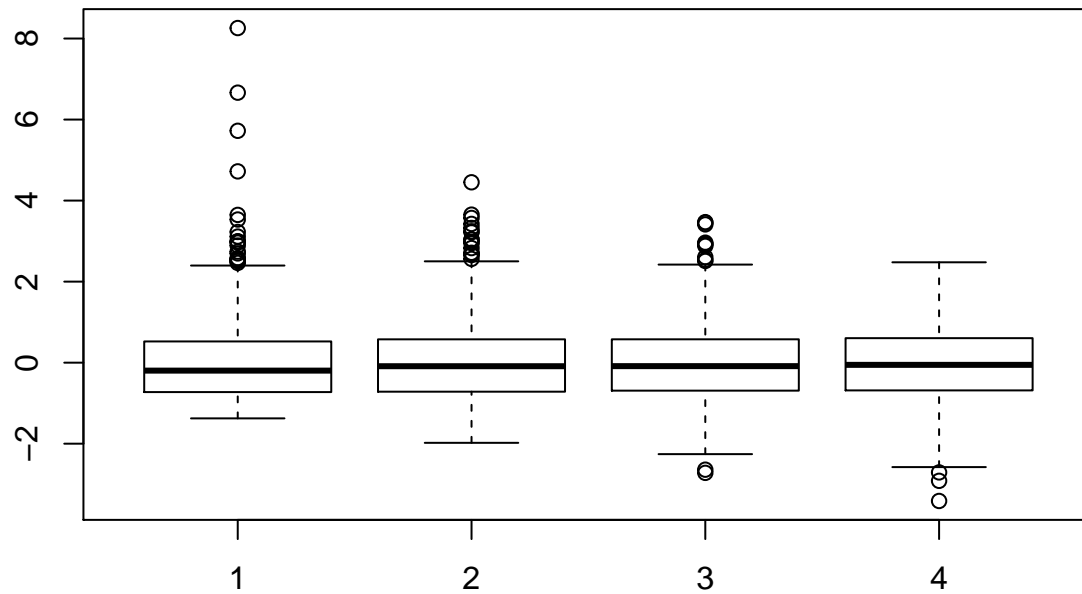
transformacionesA -> n = 2

transformacionesB —> n = 5

transformacionesC1 —> n = 30

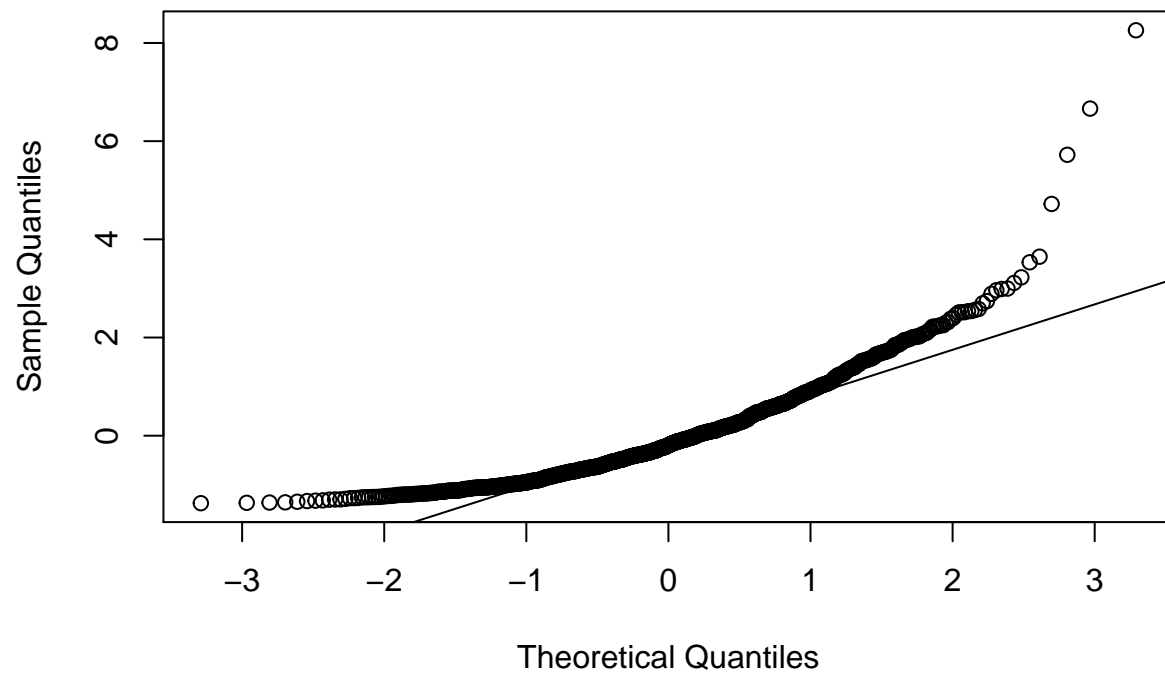
transformacionesC2 —> n = 500

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



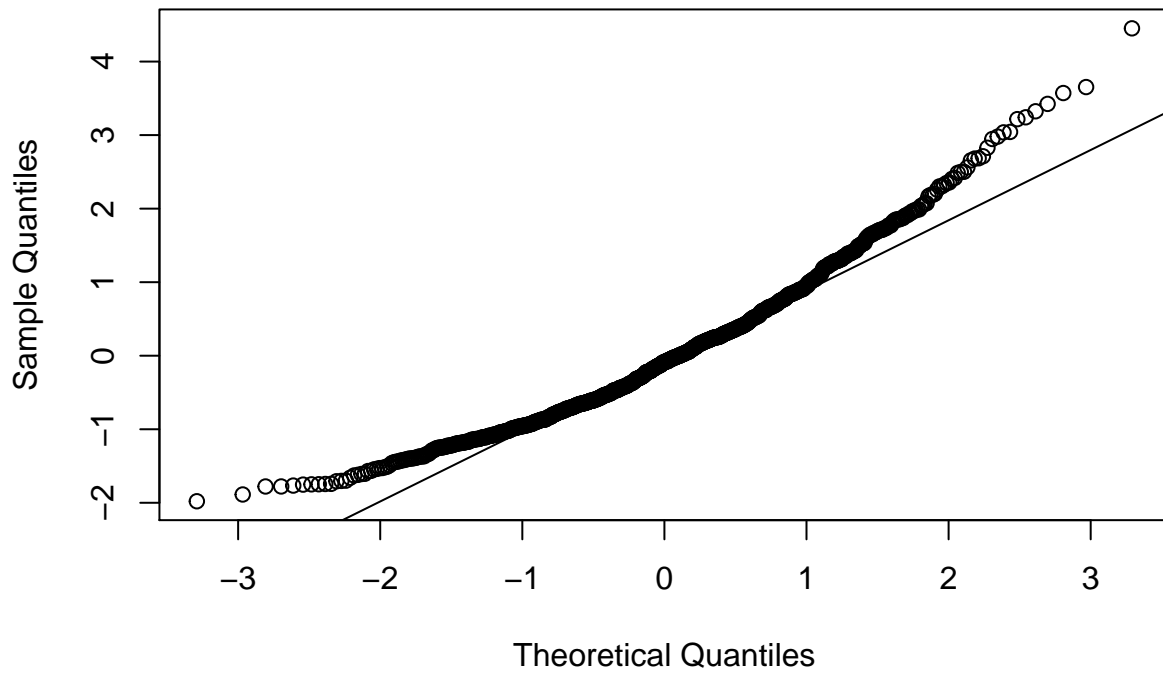
```
qqnorm(transformacionesA)
qqline(transformacionesA)
```

Normal Q-Q Plot



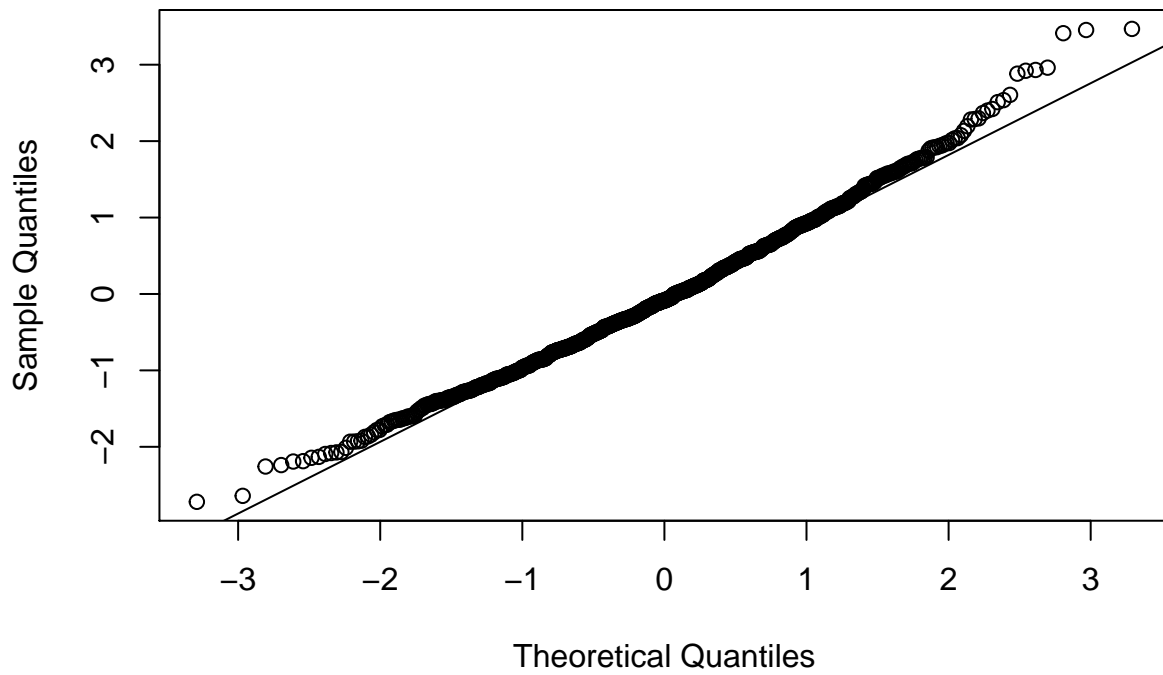
```
qqnorm(transformacionesB)
qqline(transformacionesB)
```

Normal Q-Q Plot



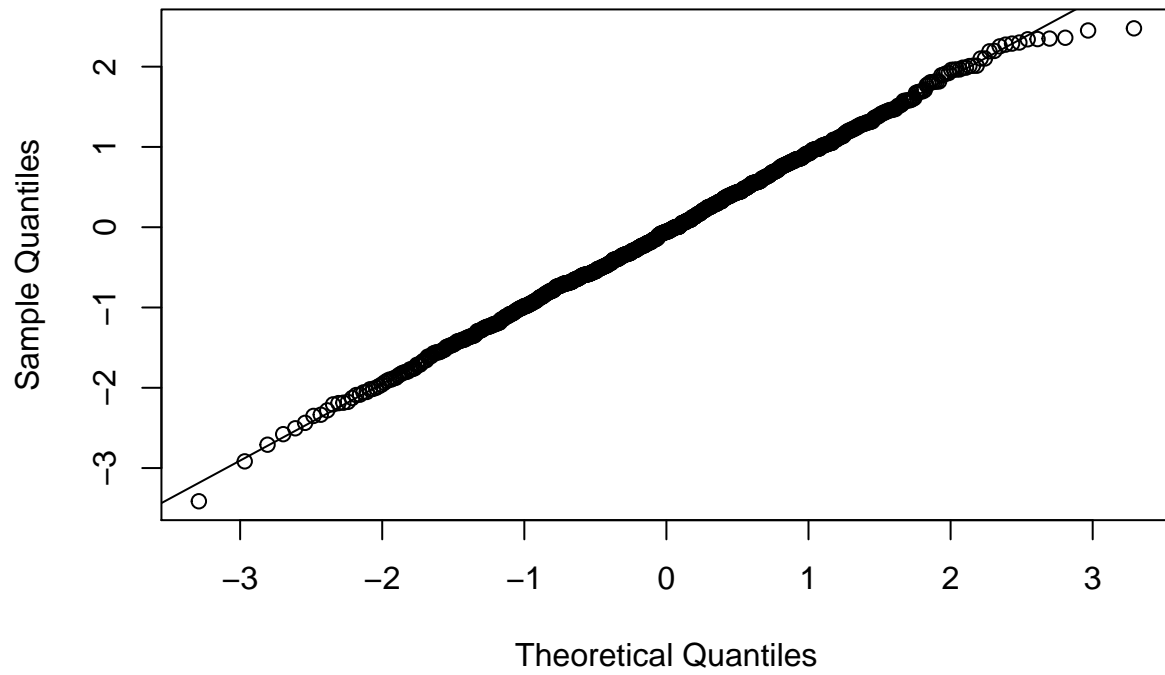
```
qqnorm(transformacionesC1)  
qqline(transformacionesC1)
```

Normal Q-Q Plot



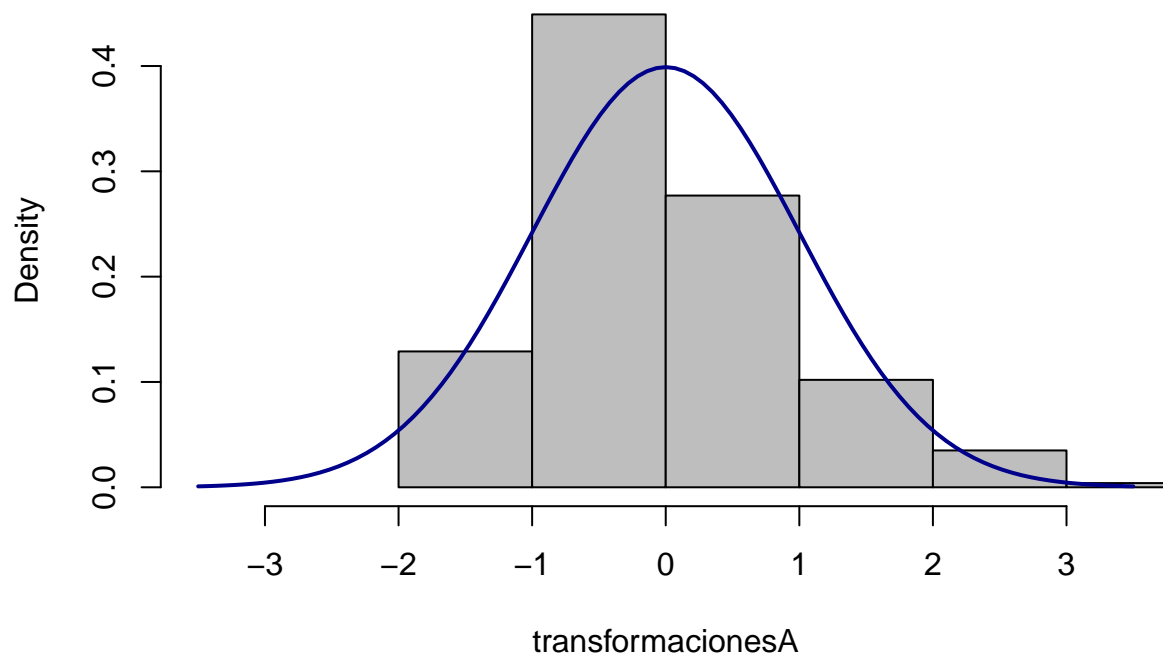
```
qqnorm(transformacionesC2)  
qqline(transformacionesC2)
```


Normal Q-Q Plot

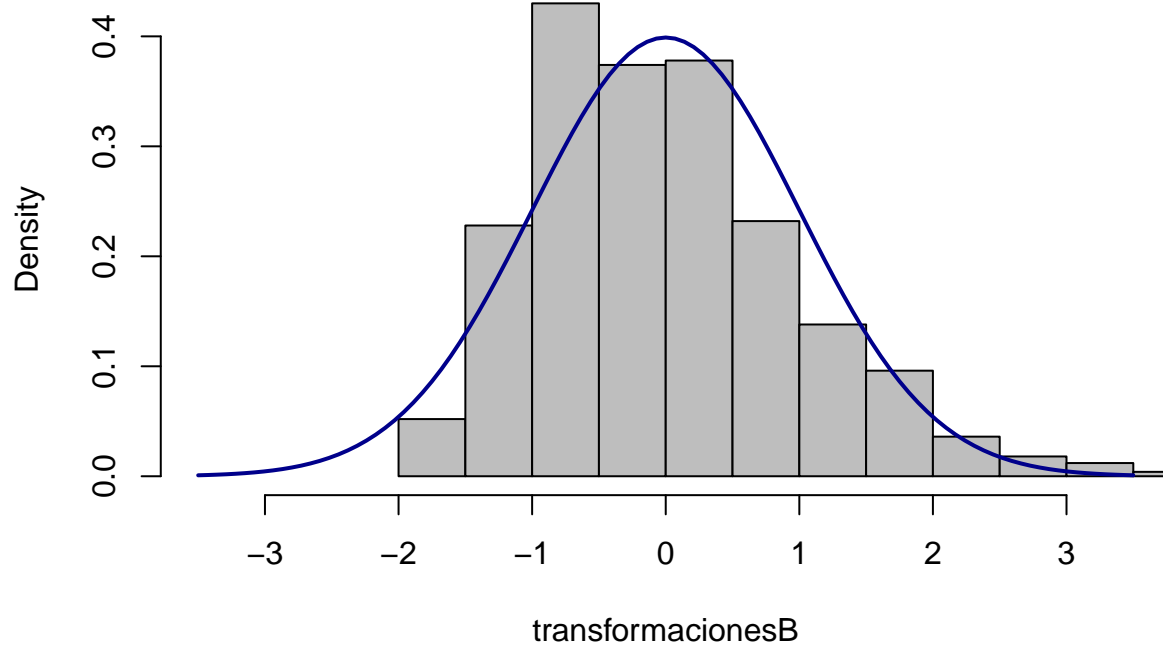


c)

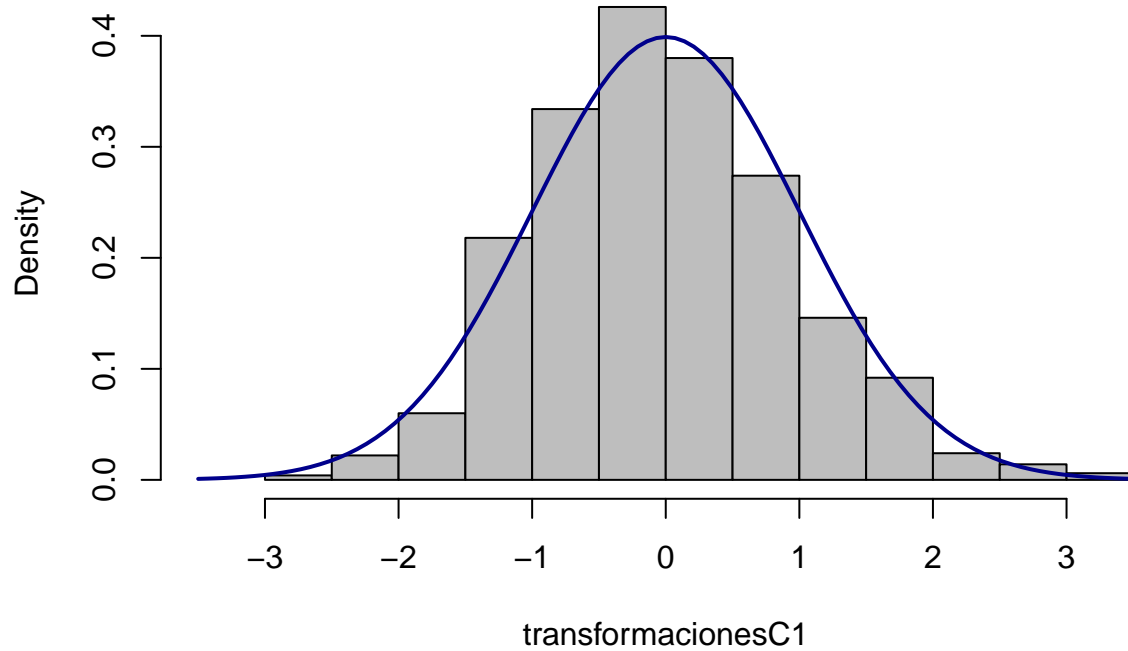
Histogram of transformacionesA



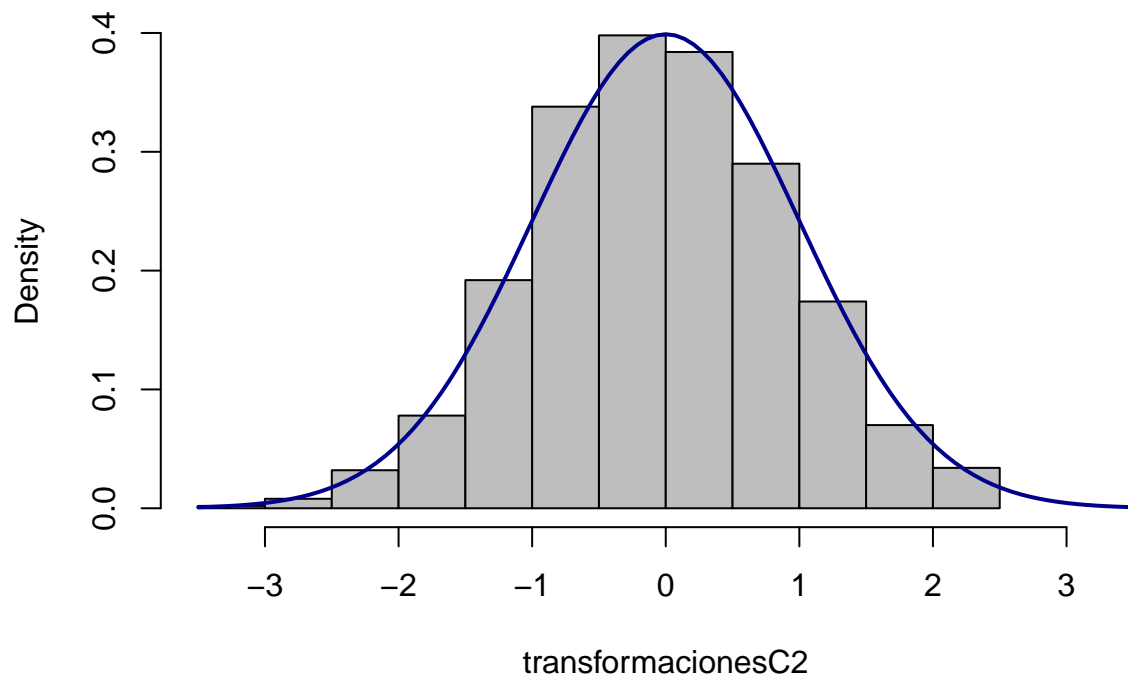
Histogram of transformacionesB



Histogram of transformacionesC1



Histogram of transformacionesC2

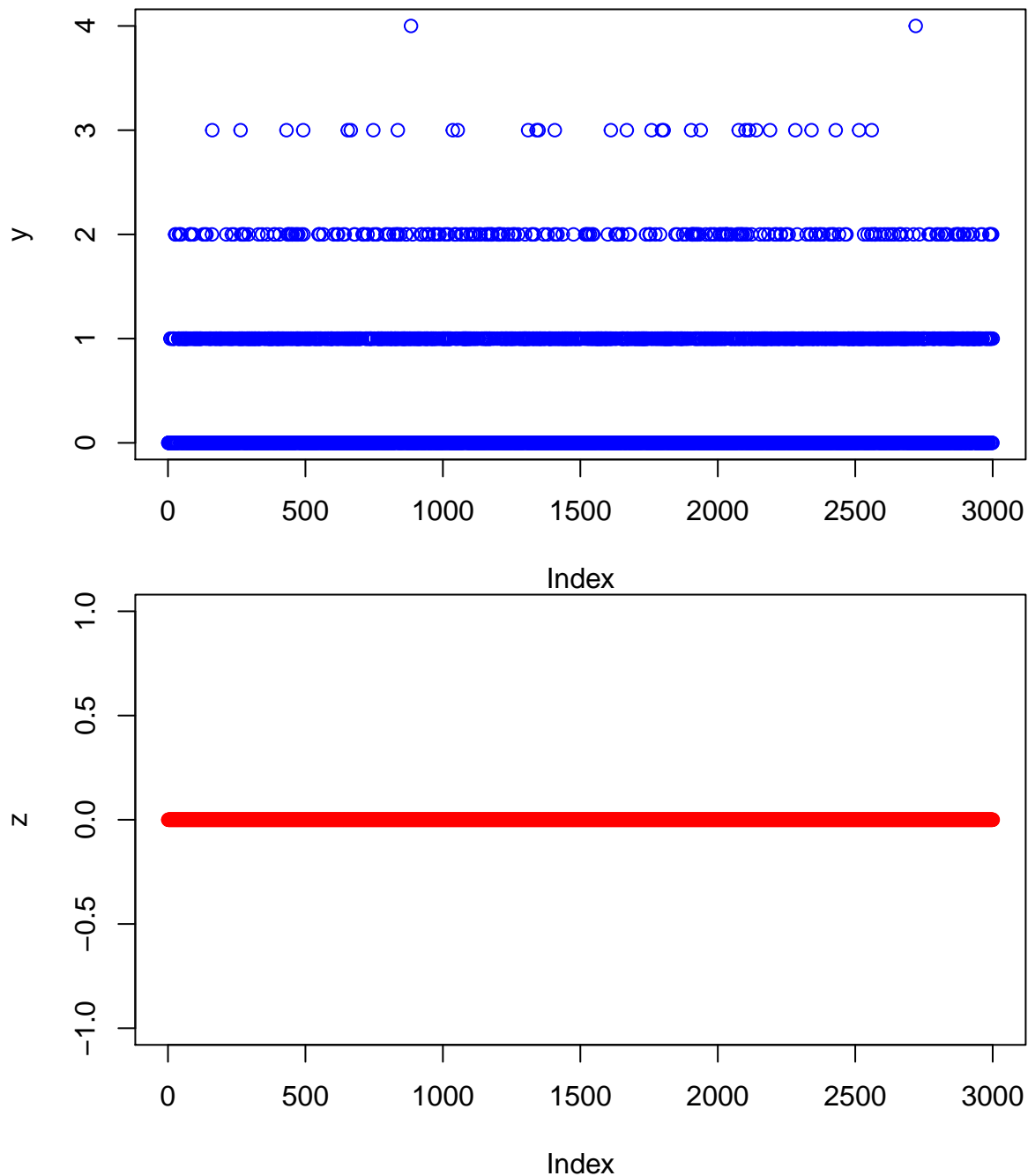


d)

En los diferentes histogramas vemos aún con más claridad lo que explicamos en los incisos anteriores sobre como a medida que aumenta el “n”, la tendencia de la transformación hacia una normal standard se hace más y más evidente.

Ejercicio 4

Punto 1



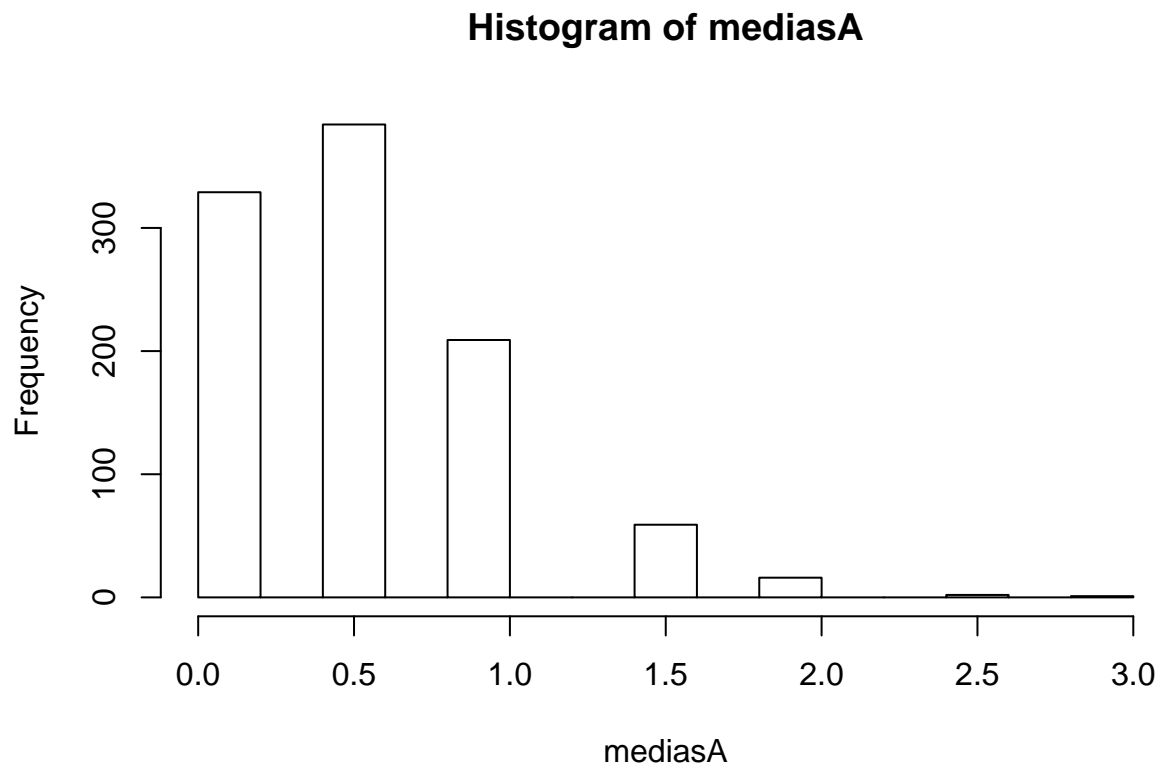
Para este punto vale la misma explicación que para la exponencial, con la diferencia que en este caso se trata de una variable discreta. Debido a esto la diferencia en el gráfico es mucho más brusca, dado que en este caso tenemos un espacio de probabilidad finito y bastante acotado.

El primer gráfico se corresponde con una binomial de los parametros utilizados, y como al fijar `set.seed()` dentro de las funciones se repite el mismo experimento se explica que en ese caso el gráfico correspondiente a esa función tenga esa forma de recta.

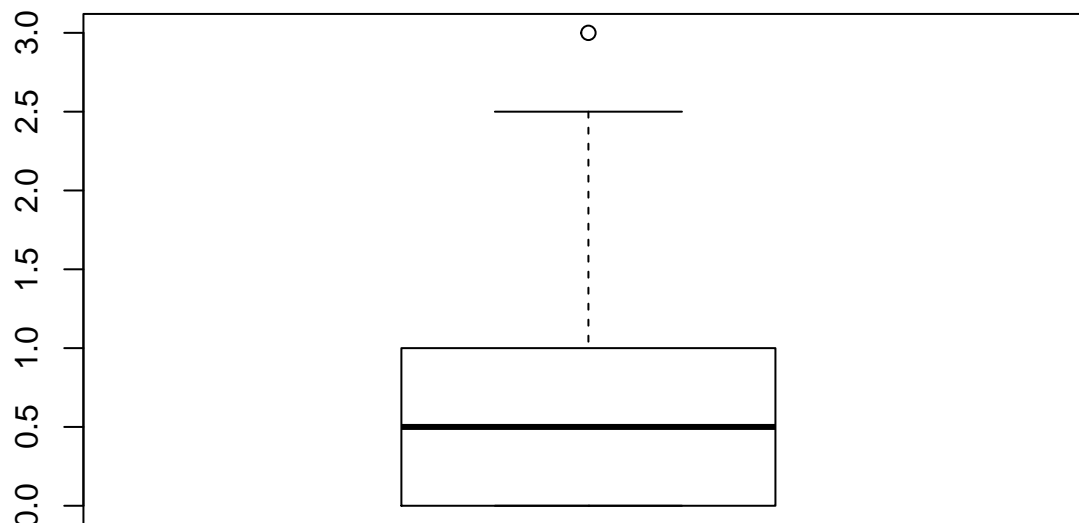
Punto 2

a)

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

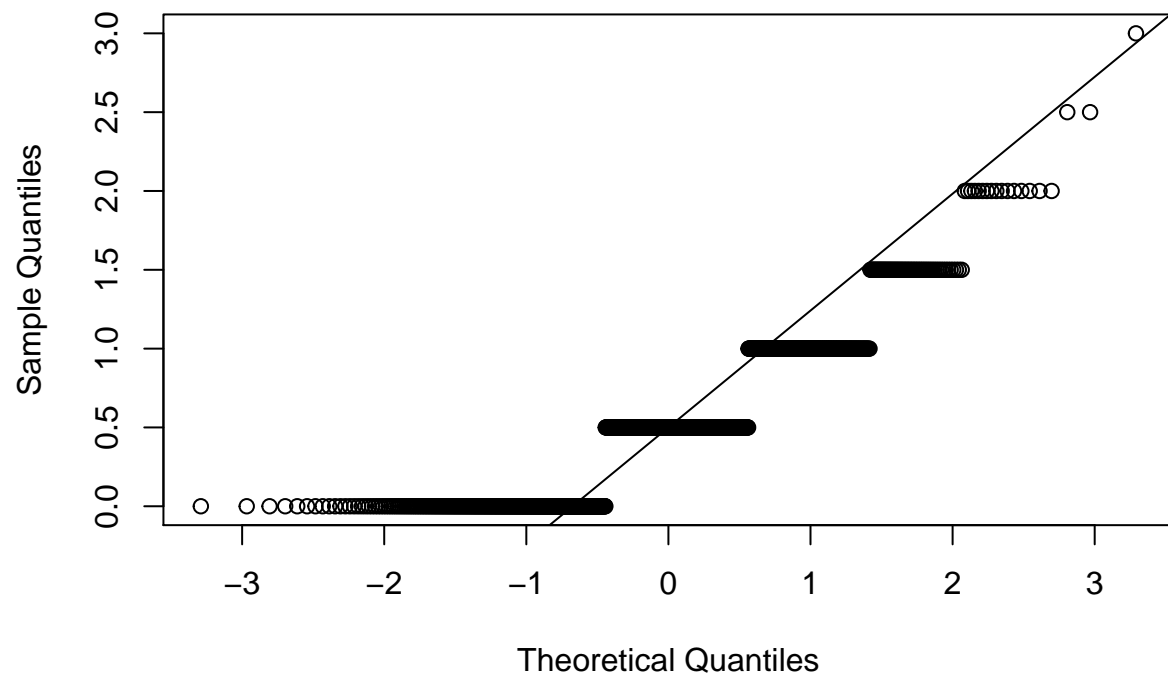


```
#Boxplot  
boxplot(mediasA)
```



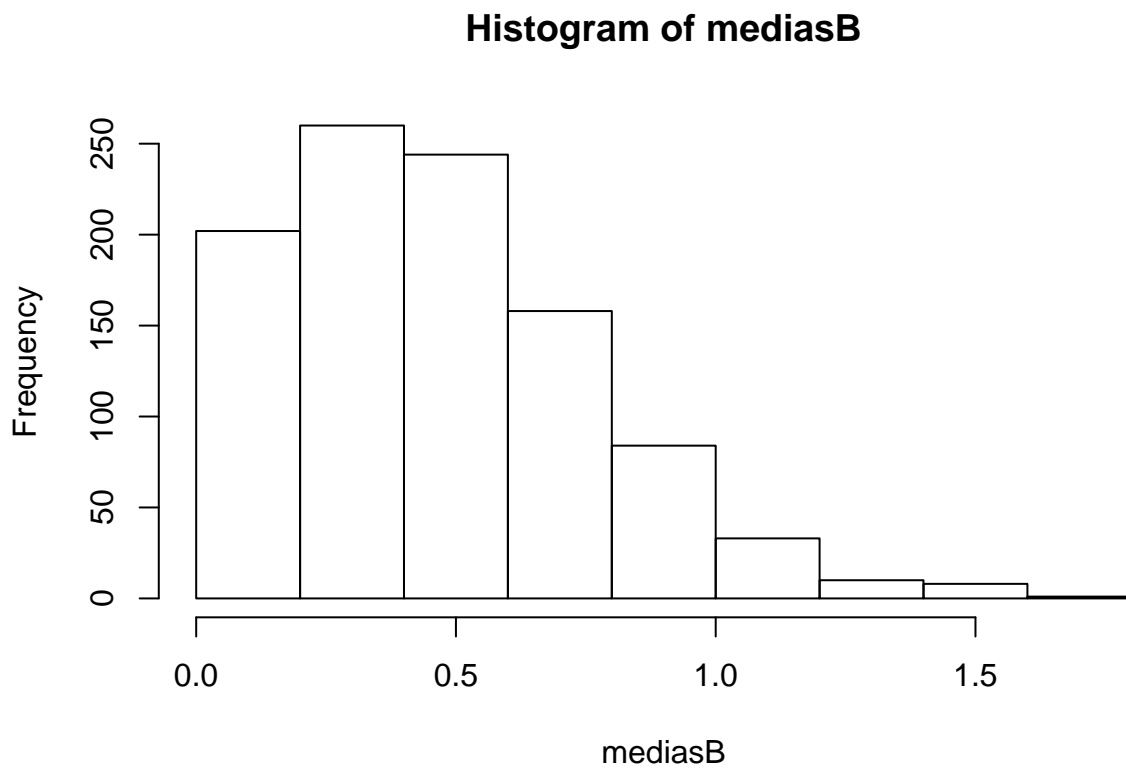
```
#Q-Q Plot  
qqnorm(mediasA)  
qqline(mediasA) #La cola del plot
```

Normal Q-Q Plot

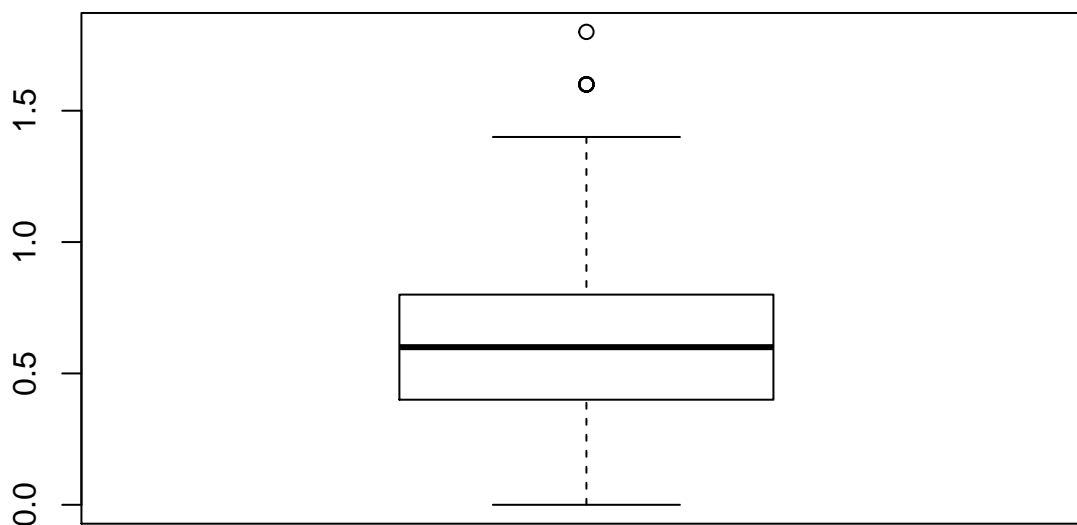


b)

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

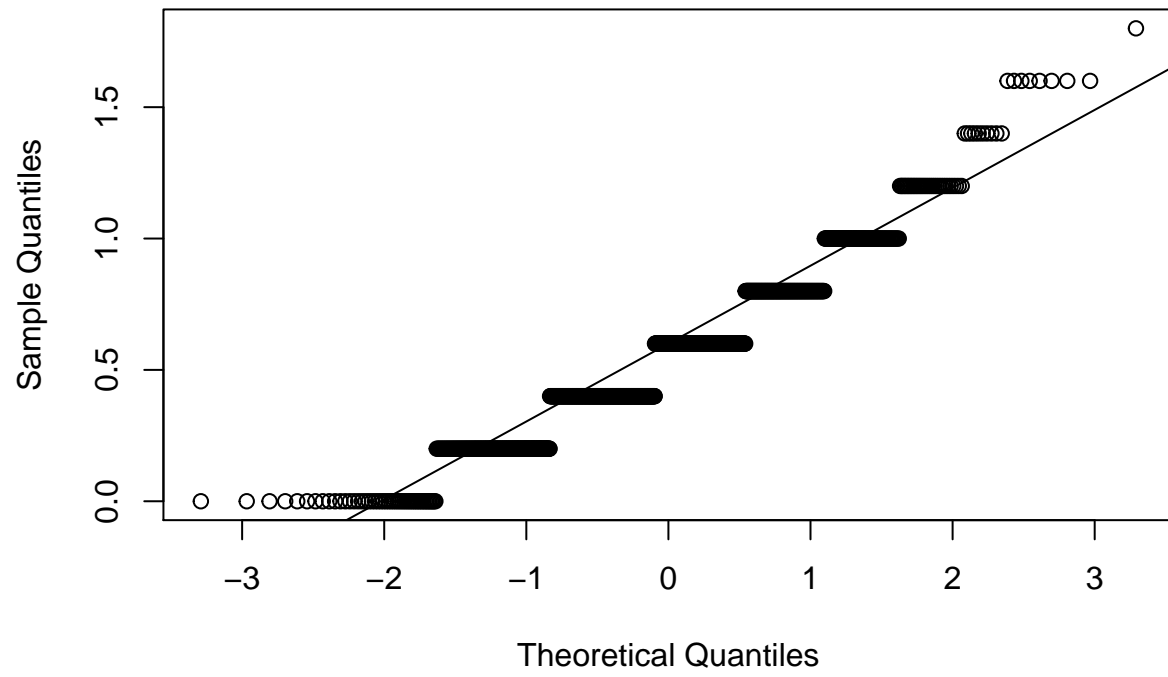


```
#Boxplot  
boxplot(mediasB)
```



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

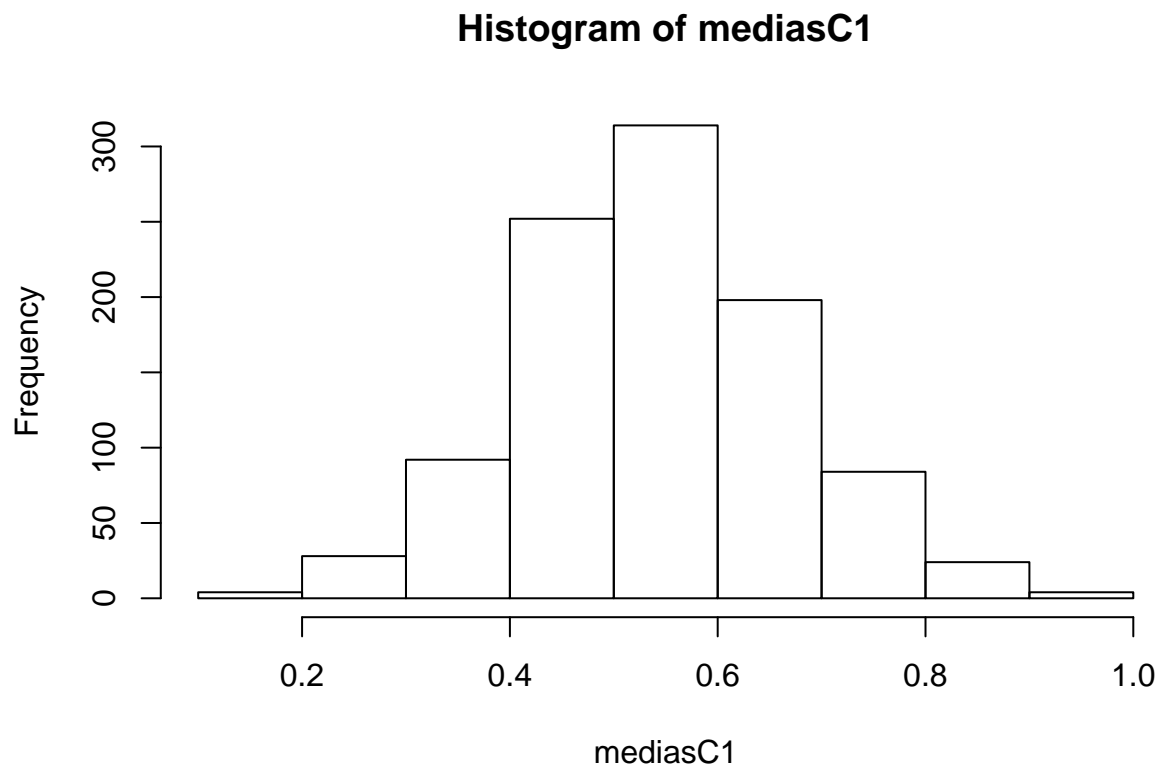
Normal Q-Q Plot



c)

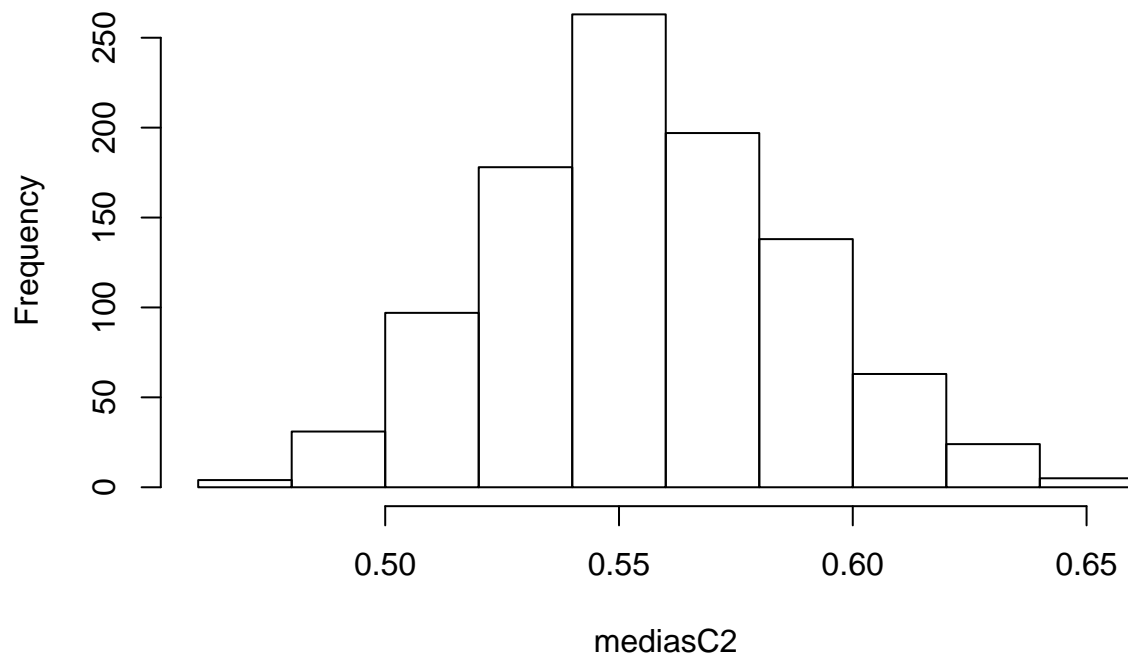
```
#Histograma
```

```
hist(mediasC1)
```



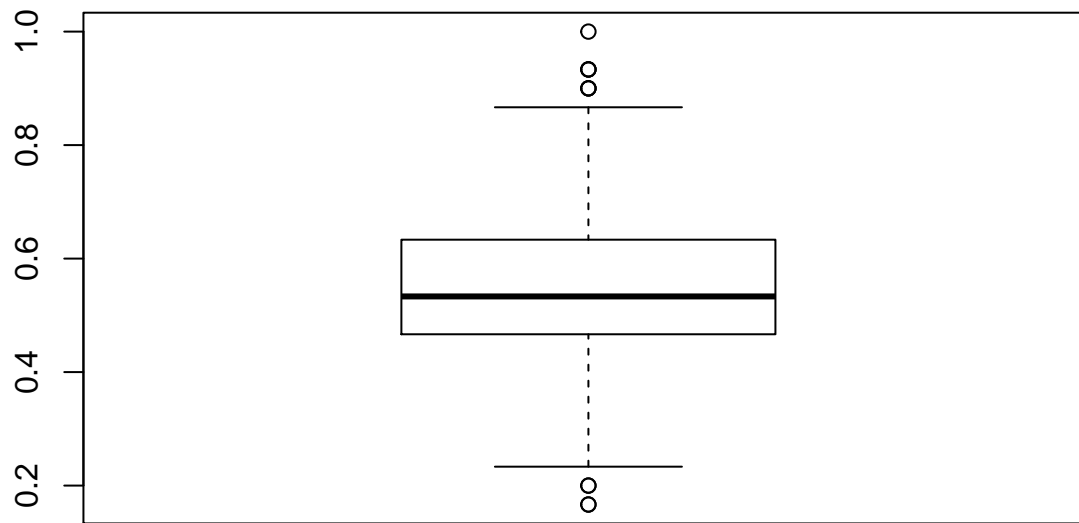
```
hist(mediasC2)
```

Histogram of mediasC2

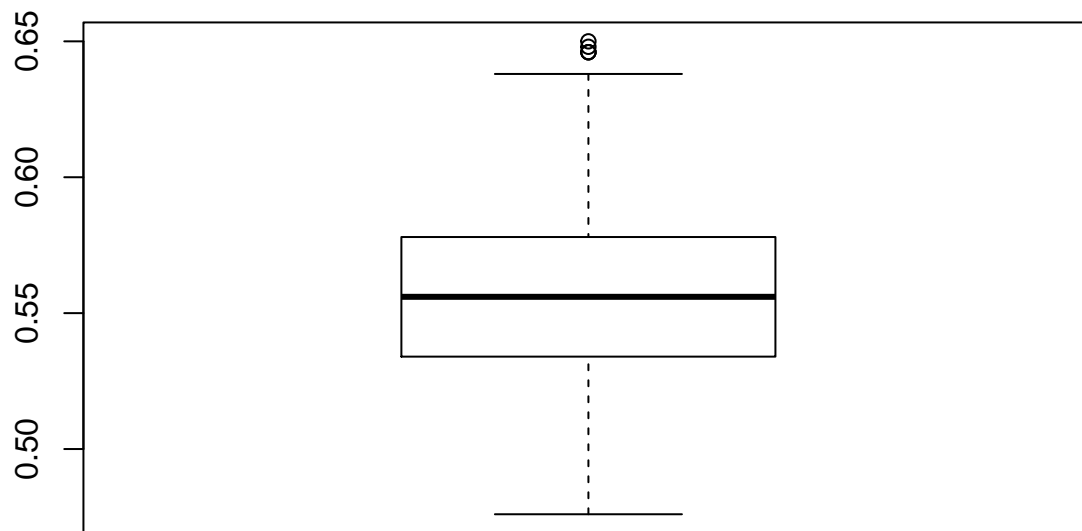


#Boxplot

```
boxplot(mediasC1)
```



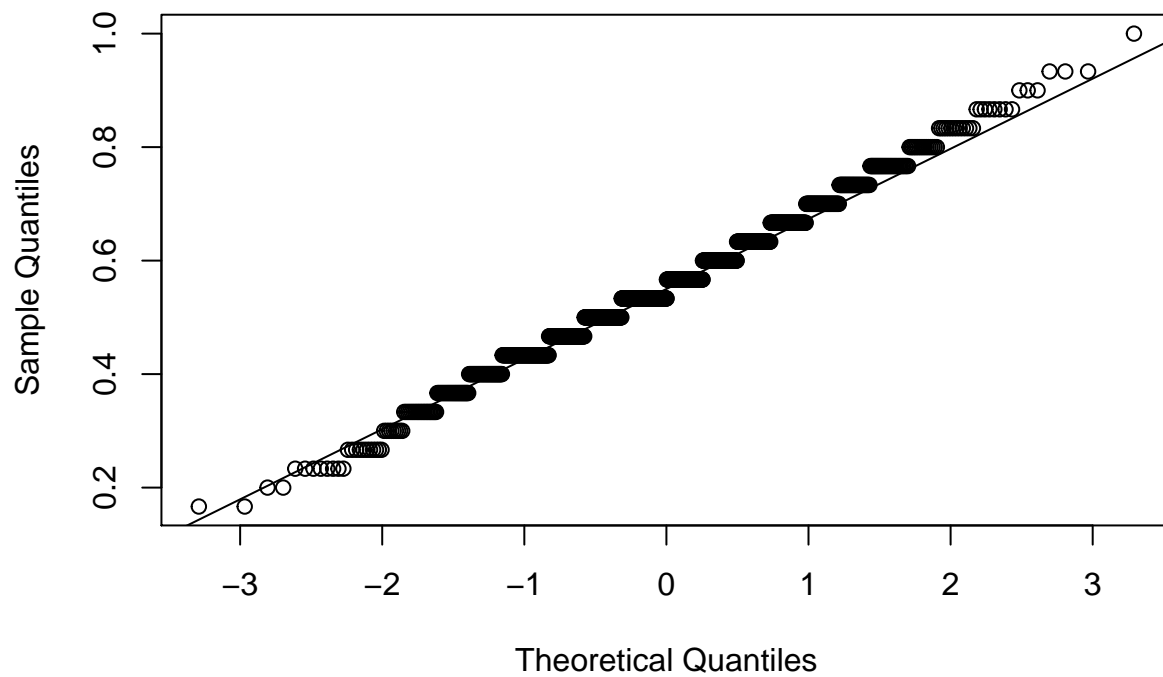
```
boxplot(mediasC2)
```



```
#Q-Q Plot
qqnorm(mediasC1)

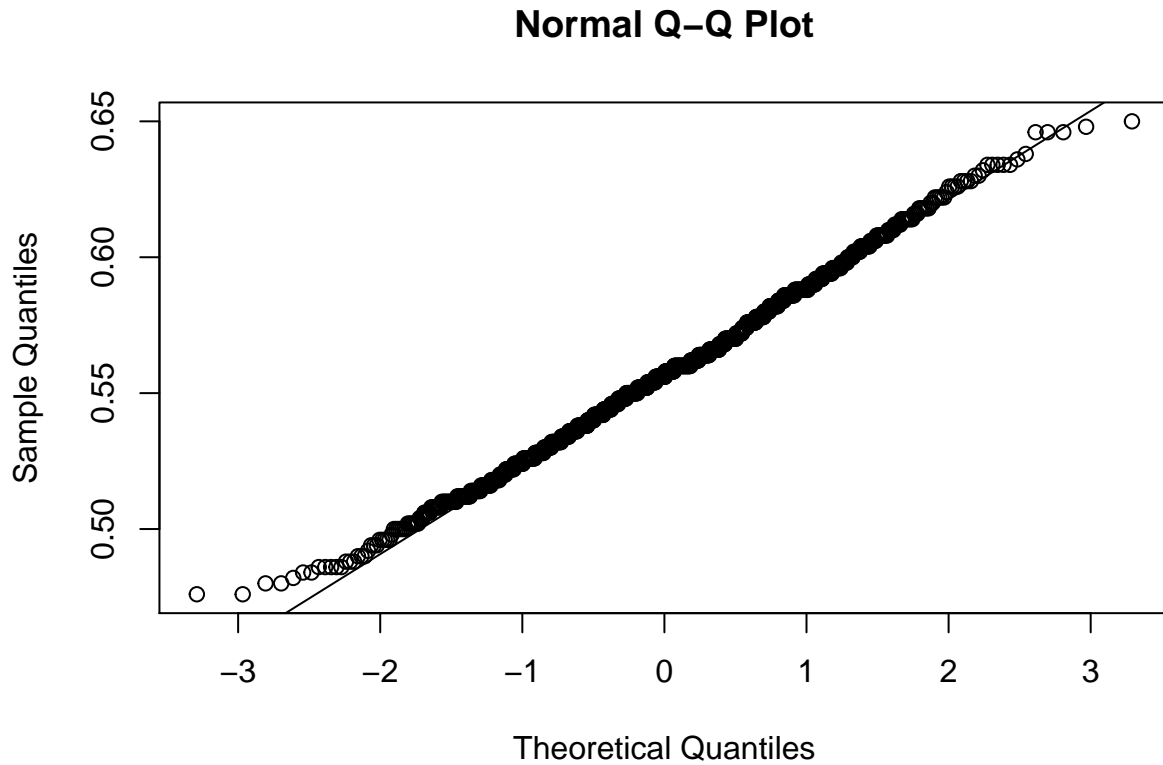
qqline(mediasC1) #La cola del plot
```

Normal Q-Q Plot



```
qqnorm(mediasC2)

qqline(mediasC2)
```



d)

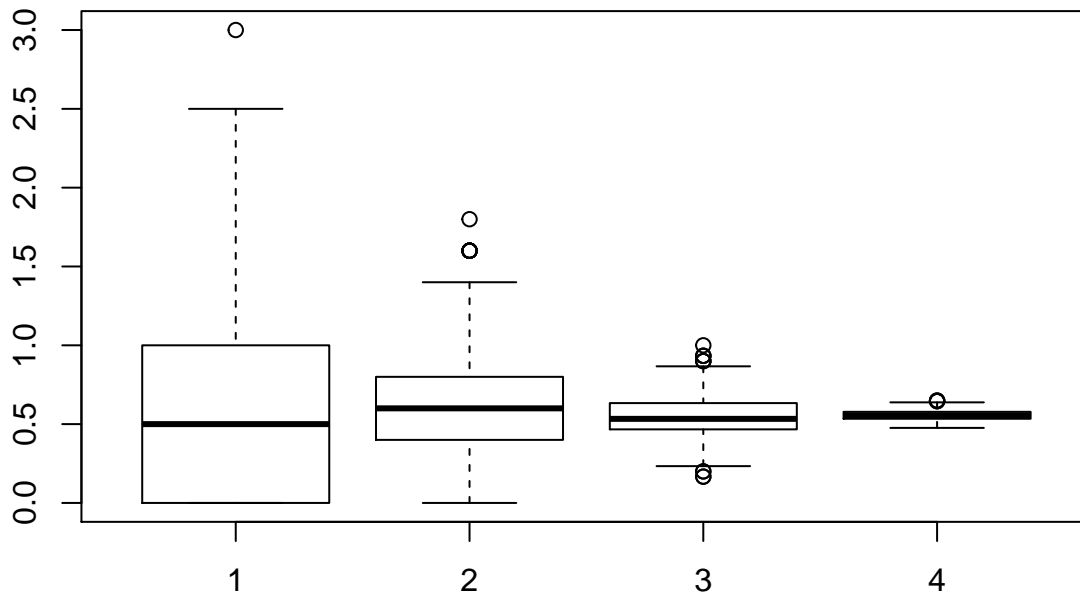
En el primer punto vemos gráficos que se corresponden con una binomial. Al ser una distribución discreta se explica que el Q-Q Plot tenga esa forma escalonada.

Con $n = 5$ el histograma va adquiriendo una ligera forma de normal, el cambio más notable lo vemos en el boxplot que tiene pocos outliers y brazos relativamente simétricos.

Ahora en los casos de $n = 30$ podemos ver como todos los gráficos van pareciéndose más al de una normal, y más aún cuando $n = 500$, los gráficos inexorablemente tienen la forma de una distribución normal, la mayor diferencia antes la notabamos en el Q-Q Plot, pero con una muestra de este tamaño ya no queda ninguna duda.

e)

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



Vemos que pasa algo similar a lo que observamos cuando hicimos este mismo experimento con una exponencial, tal vez la mayor diferencia sea que con $n = 5$ acá ya podemos notar el patrón de una distribución normal.

Punto 3

a)

```
#mediaX1  
mediaX1
```

```
## [1] 0.5295
```

```
#varX1  
varX1
```

```
## [1] 0.2431229
```

```
#mediaX2  
mediaX2
```

```
## [1] 0.5592
```

```
#varX2  
varX2
```

```
## [1] 0.09883419
```

```
#mediaX3  
mediaX3
```

```
## [1] 0.5565667
```

```
#varX3  
varX3
```

```
## [1] 0.01701389
```

```
#mediaX4
```

```
mediaX4
```

```
## [1] 0.557012
```

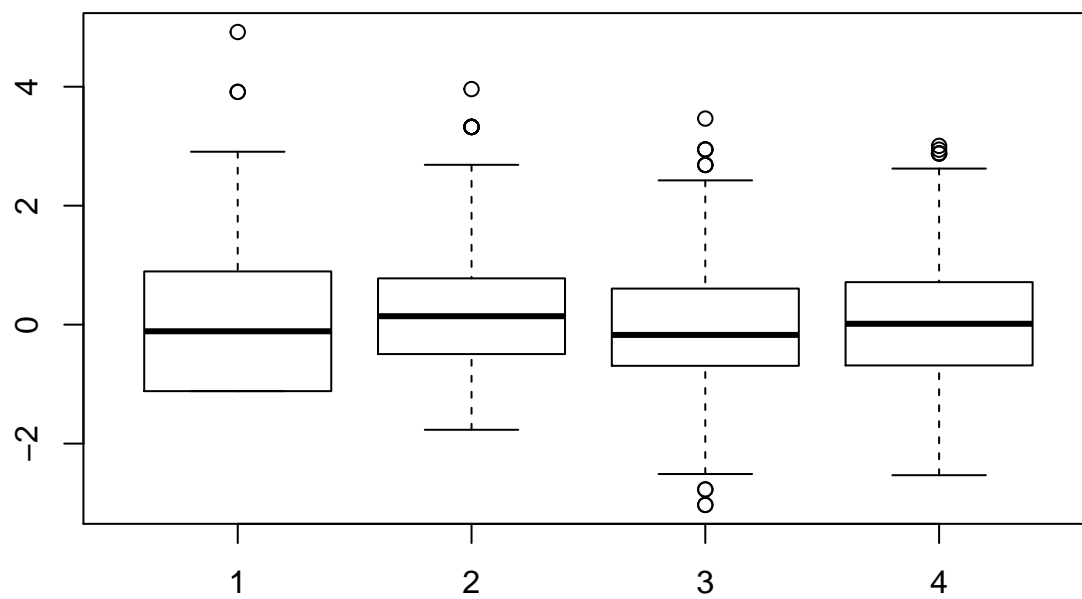
```
#varX4
```

```
varX4
```

```
## [1] 0.001017105
```

b)

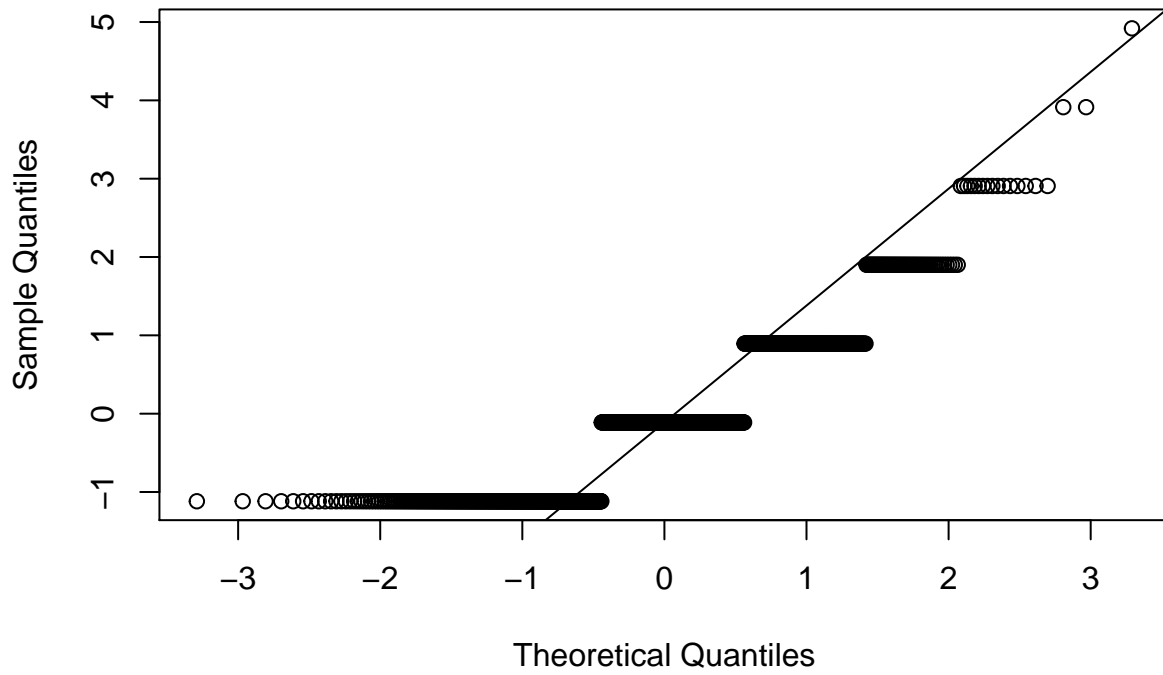
```
boxplot(transformacionesA, transformacionesB, transformacionesC1, transformacionesC2)
```



```
qqnorm(transformacionesA)
```

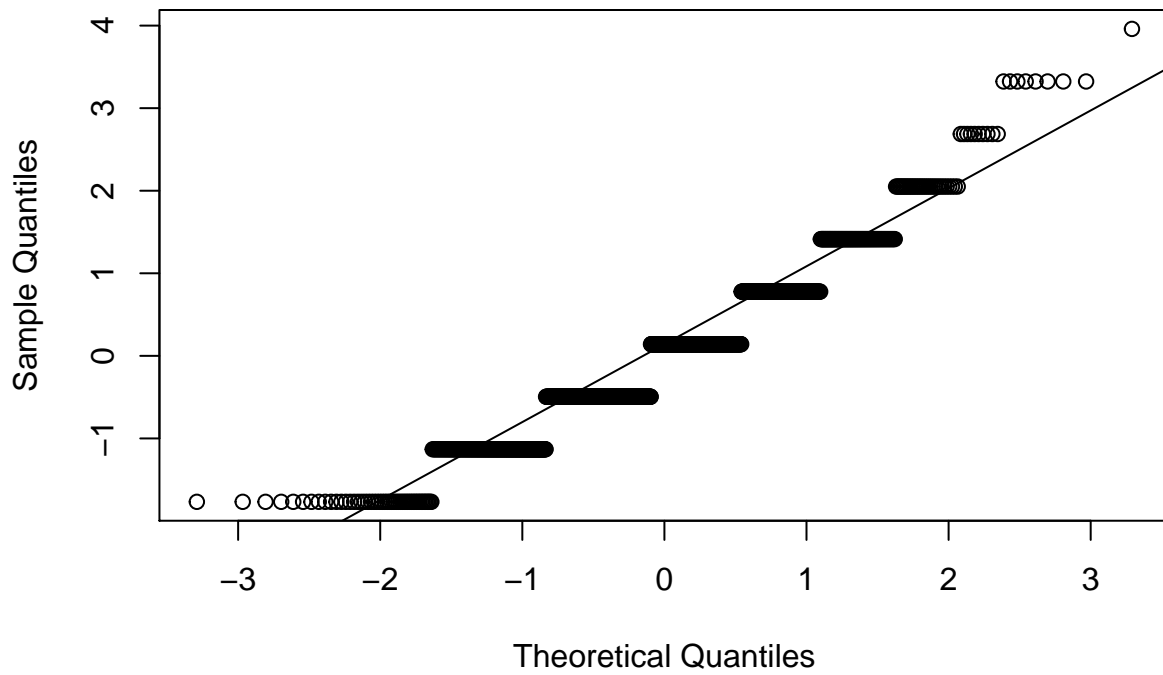
```
qqline(transformacionesA)
```

Normal Q-Q Plot



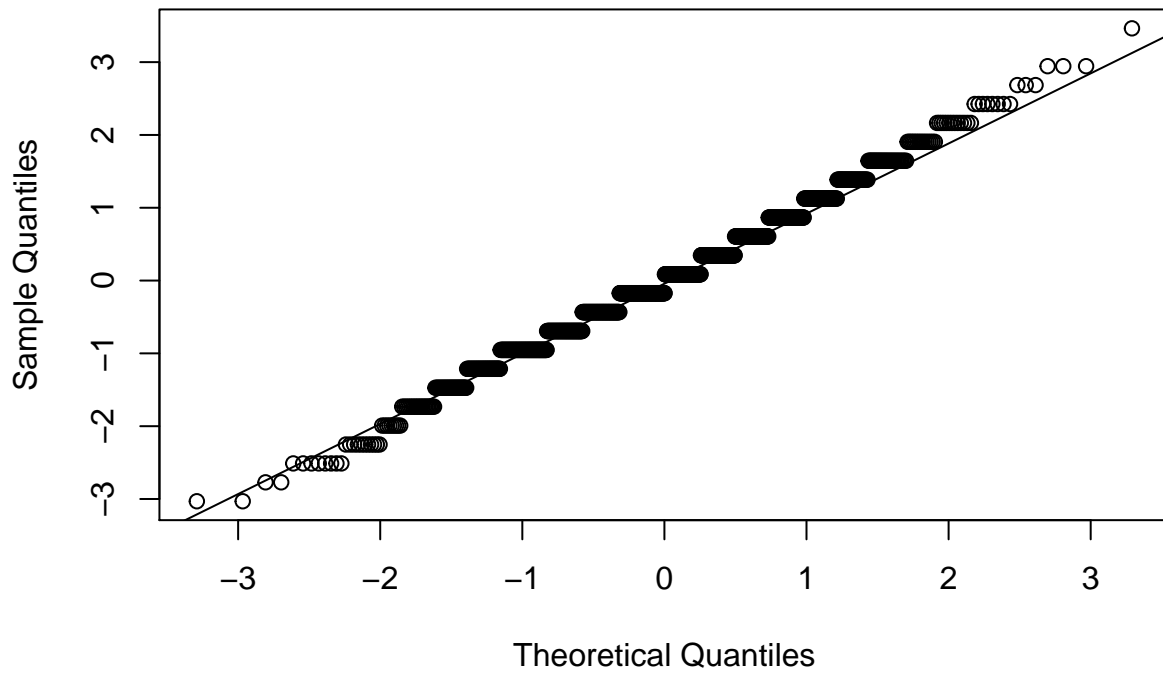
```
qqnorm(transformacionesB)
qqline(transformacionesB)
```

Normal Q-Q Plot



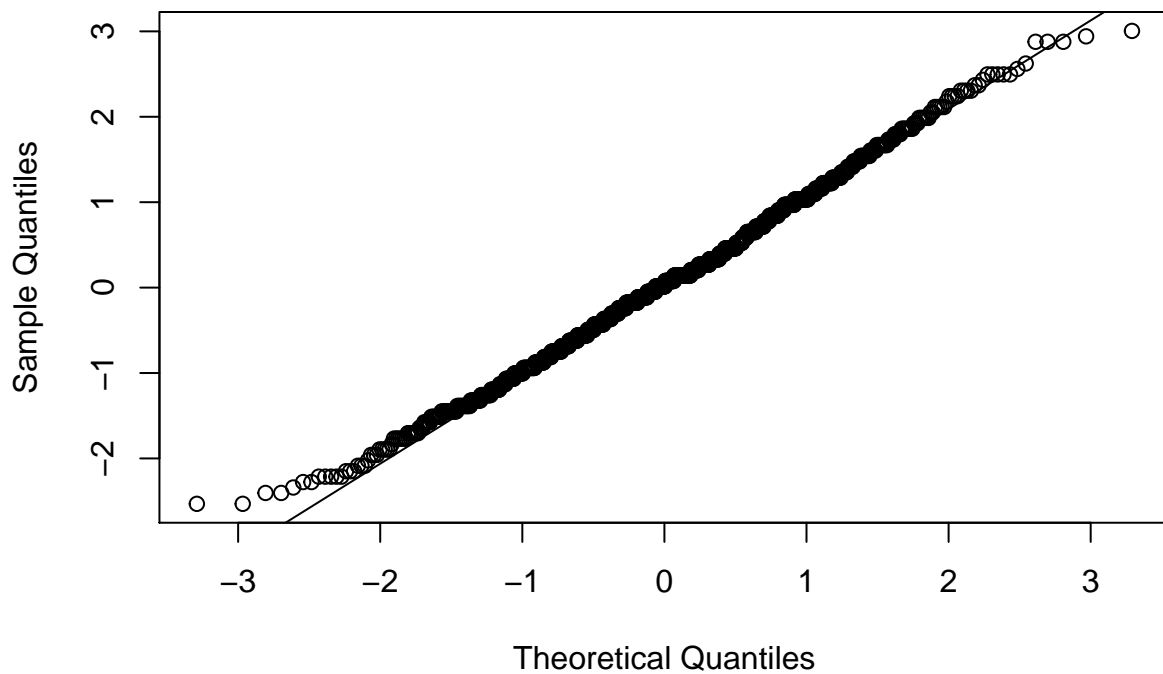
```
qqnorm(transformacionesC1)
qqline(transformacionesC1)
```

Normal Q-Q Plot



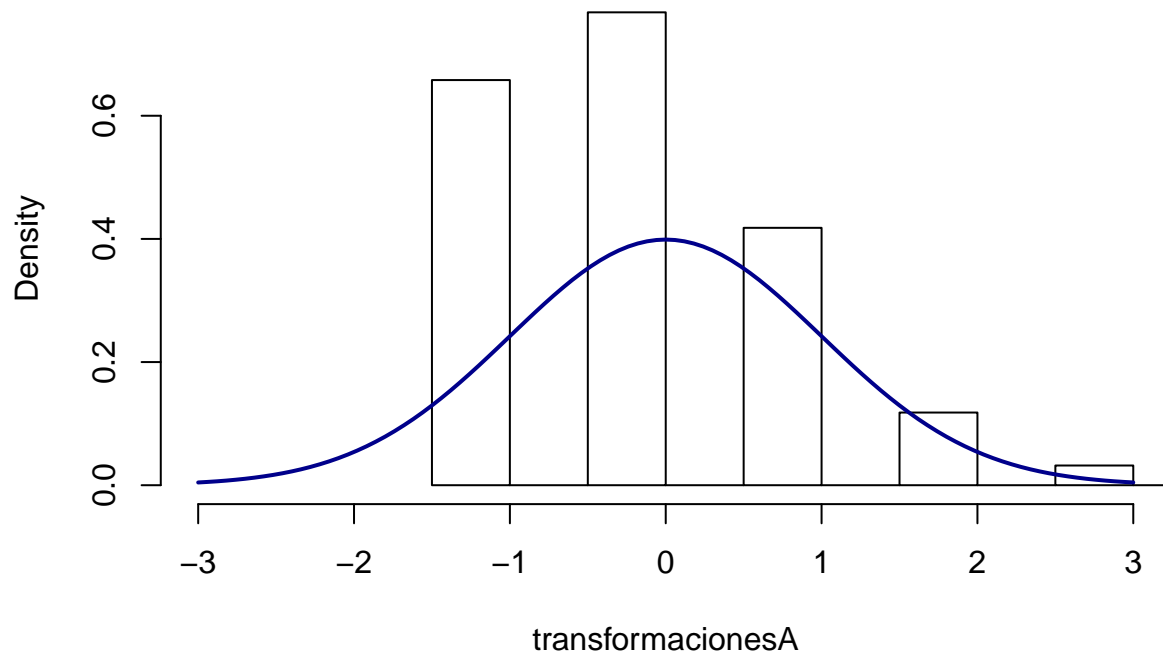
```
qqnorm(transformacionesC2)  
qqline(transformacionesC2)
```

Normal Q-Q Plot

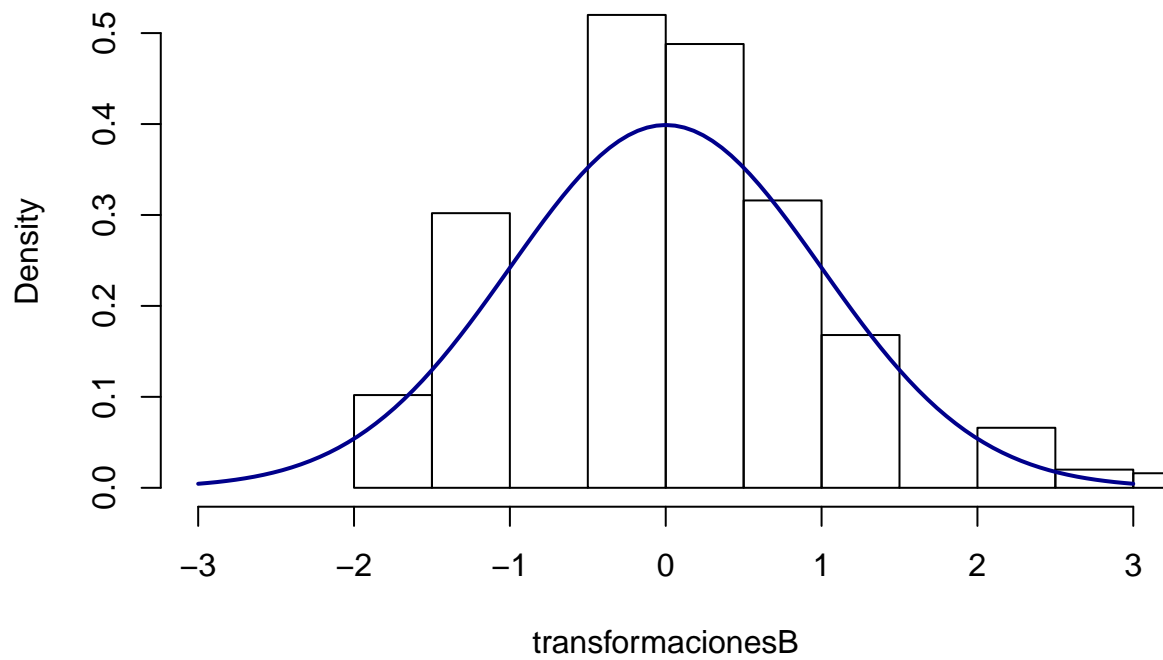


c)

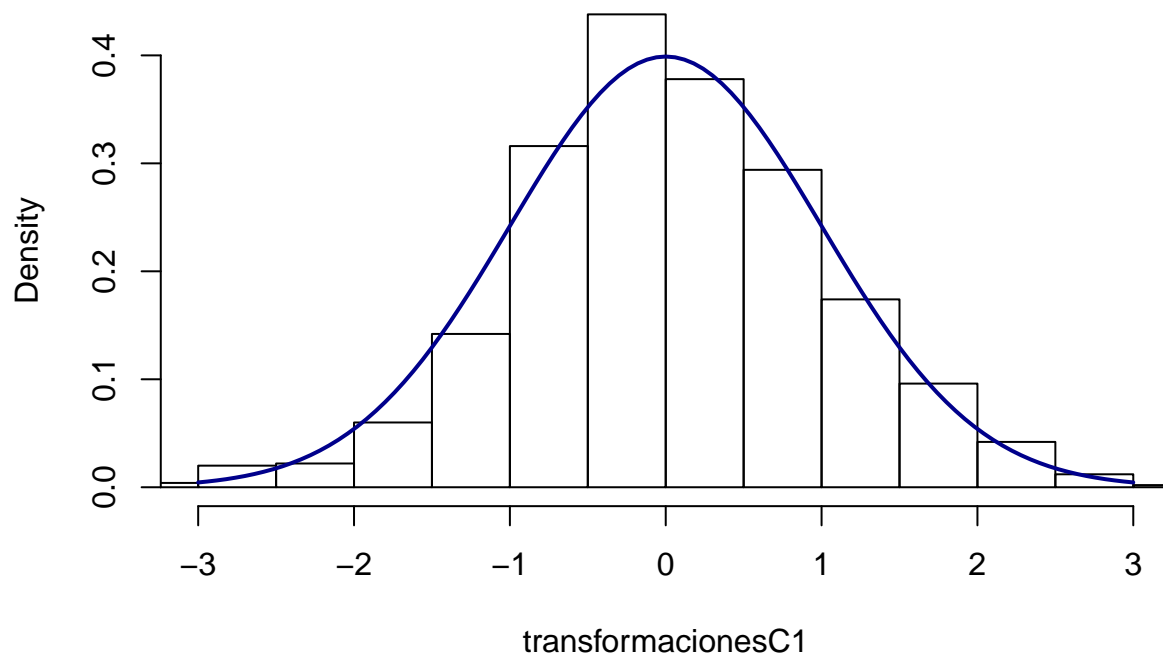
Histogram of transformacionesA



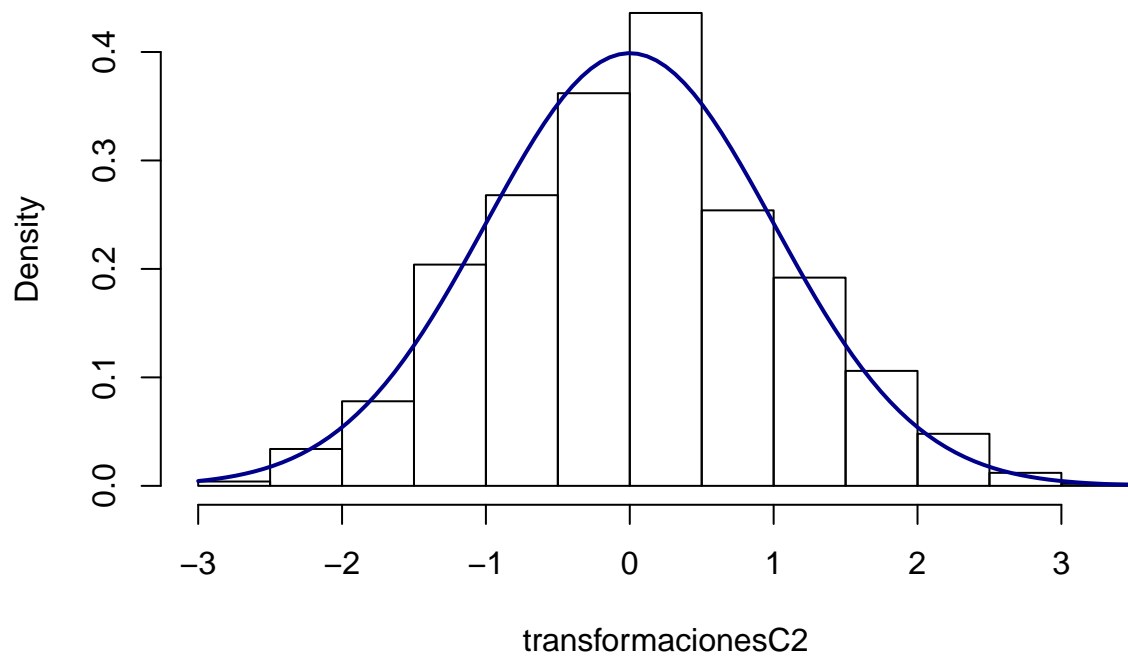
Histogram of transformacionesB



Histogram of transformacionesC1



Histogram of transformacionesC2



d)

Observamos lo mismo que al llevar a cabo este ejercicio con la exponencial. Anteriormente vimos como la binomial tendía a una normal por la Ley de los Grandes Números, y ahora vimos como gracias a las propiedades de la distribución normal pudimos aplicar esta transformación y obtener algo que tiende aproximadamente a una distribución normal standard.