pruebas iid

Eduardo Rubio M

04-09-2024

Generamos la función para calcular la Correlación integral

```r
# Definir la función para calcular la Corelación integral
correlation_integral <- function(data, r, d) {
  N <- length(data)
  count <- 0

  # Iterar sobre todos los pares (i, j) tal que 1 <= i < j <= N-d+1
  for (i in 1:(N-d)) {
    for (j in (i+1):(N-d+1)) {
      # Crear los vectores u_i^d y u_j^d
      u_i_d <- data[i:(i+d-1)]
      u_j_d <- data[j:(j+d-1)]

      # Calcular la distancia euclidiana entre u_i^d y u_j^d
      distance <- sqrt(sum((u_i_d - u_j_d)^2))

      # Verificar si la distancia es menor que r
      if (distance < r) {
        count <- count + 1
      }
    }
  }

  # Calcular C_N(r, d)
  C_N <- (2 / (N^2)) * count
  return(C_N)
}
```

Ahora generamos la función para el X__n(r,d)

```r
# Definir la función X_N(r, d)
X_N <- function(data, r, d) {
  C_N_d <- correlation_integral(data, r, d)
  C_N_d_minus_1 <- correlation_integral(data, r, d-1)
  C_N_d_plus_1 <- correlation_integral(data, r, d+1)

  X_N_value <- (C_N_d^2) / (C_N_d_minus_1 * C_N_d_plus_1)
  return(X_N_value)
}
```

```r
# Caso especial para d = 1
X_N_d_1 <- function(data, r) {
  C_N_1 <- correlation_integral(data, r, 1)
  C_N_2 <- correlation_integral(data, r, 2)

  X_N_value <- (C_N_1^2) / C_N_2
  return(X_N_value)
}
```

Veamos como funciona para múltiples series IID, esta vez con 1000 series de 1000 observaciones.

```r
# Generar múltiples series IID
set.seed(125)
n_series <- 1000  # Número de series
length_series <- 1000  # Longitud de cada serie

# Crear una matriz donde cada fila es una serie IID
data_matrix <- matrix(rnorm(n_series * length_series), nrow = n_series, ncol = length_series)
```

```r
# Parámetros
r <- 0.5  # Umbral para la norma
d_values <- 1:5  # Valores de d

# Inicializar listas para almacenar los resultados
X_N_results <- matrix(0, nrow = n_series, ncol = length(d_values))
colnames(X_N_results) <- paste("d =", d_values)

# Calcular X_{N}(r,d) para cada serie y para cada valor de d
for (i in 1:n_series) {
  series <- data_matrix[i, ]

  # Caso especial para d = 1
  X_N_results[i, 1] <- X_N_d_1(series, r)

  # Calcular X_{N}(r,d) para d = 2, 3, 4, 5
  for (d in 2:length(d_values)) {
    X_N_results[i, d] <- X_N(series, r, d_values[d])
  }
}

# Ver los resultados
head(X_N_results)
```

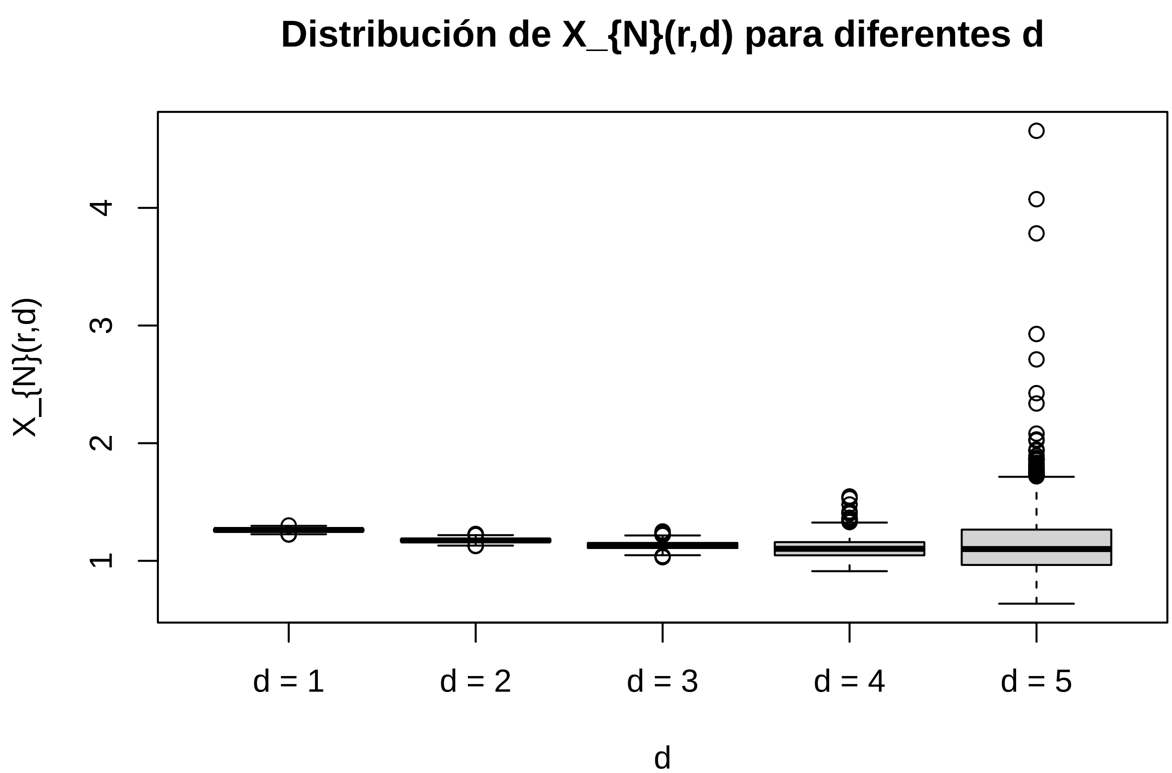
```
##         d = 1    d = 2    d = 3     d = 4     d = 5
## [1,] 1.254635 1.181791 1.057596 1.0182086 0.9569536
## [2,] 1.239122 1.220761 1.249765 1.0935572 1.0406504
## [3,] 1.251379 1.168879 1.132950 0.9641155 0.9905647
## [4,] 1.249430 1.158439 1.090150 1.1040672 0.7674067
## [5,] 1.273359 1.173964 1.109720 0.9662050 1.1260772
## [6,] 1.252905 1.192827 1.135682 1.0229727 0.8511920
```

Veamos como se comporta X_{N}(r,d) para este caso

```r
# Resumen de los resultados
summary(X_N_results)
```

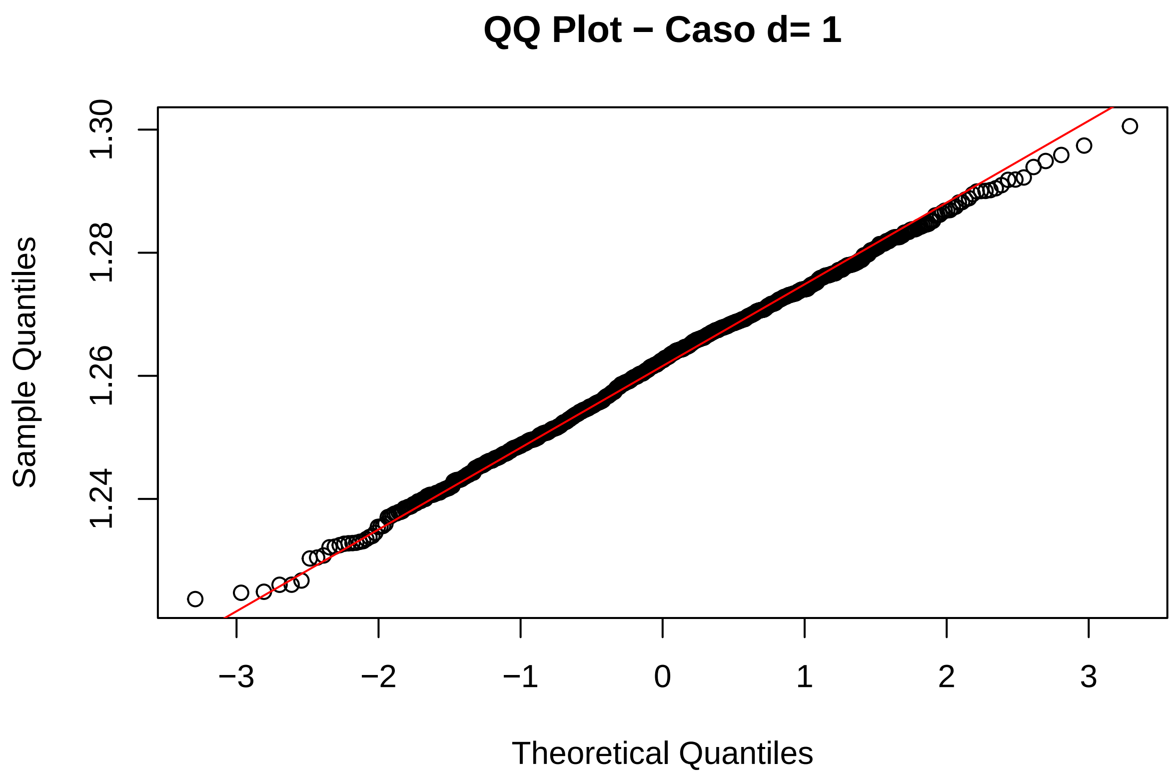
```
##       d = 1           d = 2           d = 3           d = 4
##  Min.   :1.224   Min.   :1.126   Min.   :1.031   Min.   :0.9121
##  1st Qu.:1.253   1st Qu.:1.162   1st Qu.:1.109   1st Qu.:1.0470
##  Median :1.263   Median :1.174   Median :1.130   Median :1.1026
##  Mean   :1.262   Mean   :1.174   Mean   :1.131   Mean   :1.1076
##  3rd Qu.:1.271   3rd Qu.:1.185   3rd Qu.:1.152   3rd Qu.:1.1592
##  Max.   :1.301   Max.   :1.229   Max.   :1.250   Max.   :1.5475
##       d = 5
##  Min.   :0.6360
##  1st Qu.:0.9652
##  Median :1.1003
##  Mean   :1.1514
##  3rd Qu.:1.2654
##  Max.   :4.6550
```

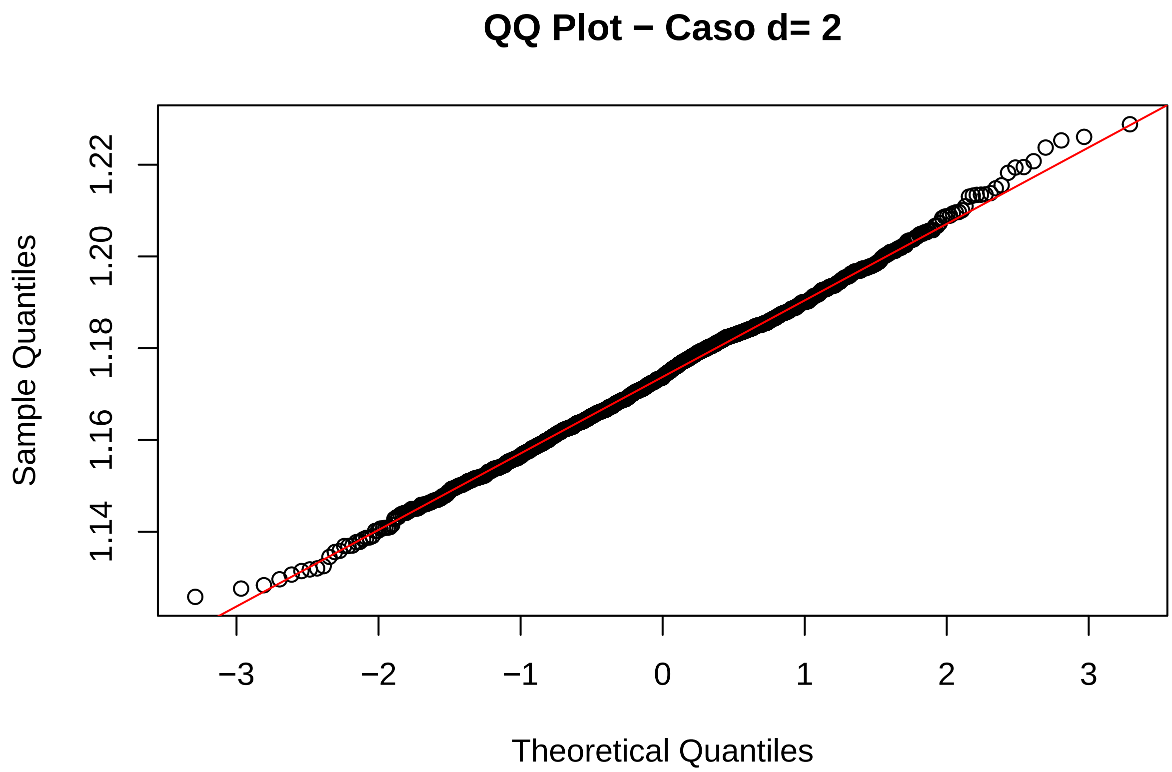
```r
# Graficar los resultados
boxplot(X_N_results, main="Distribución de X_{N}(r,d) para diferentes d", ylab="X_{N}(r,d)", xlab="d")
```

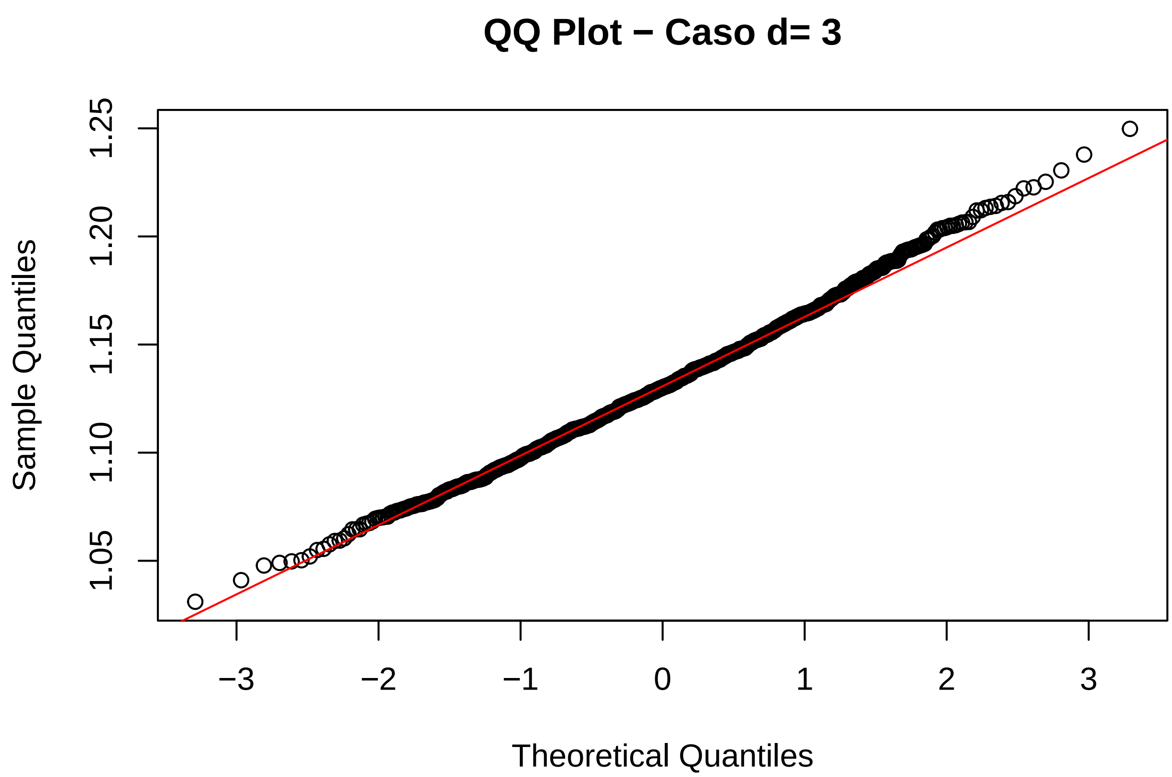


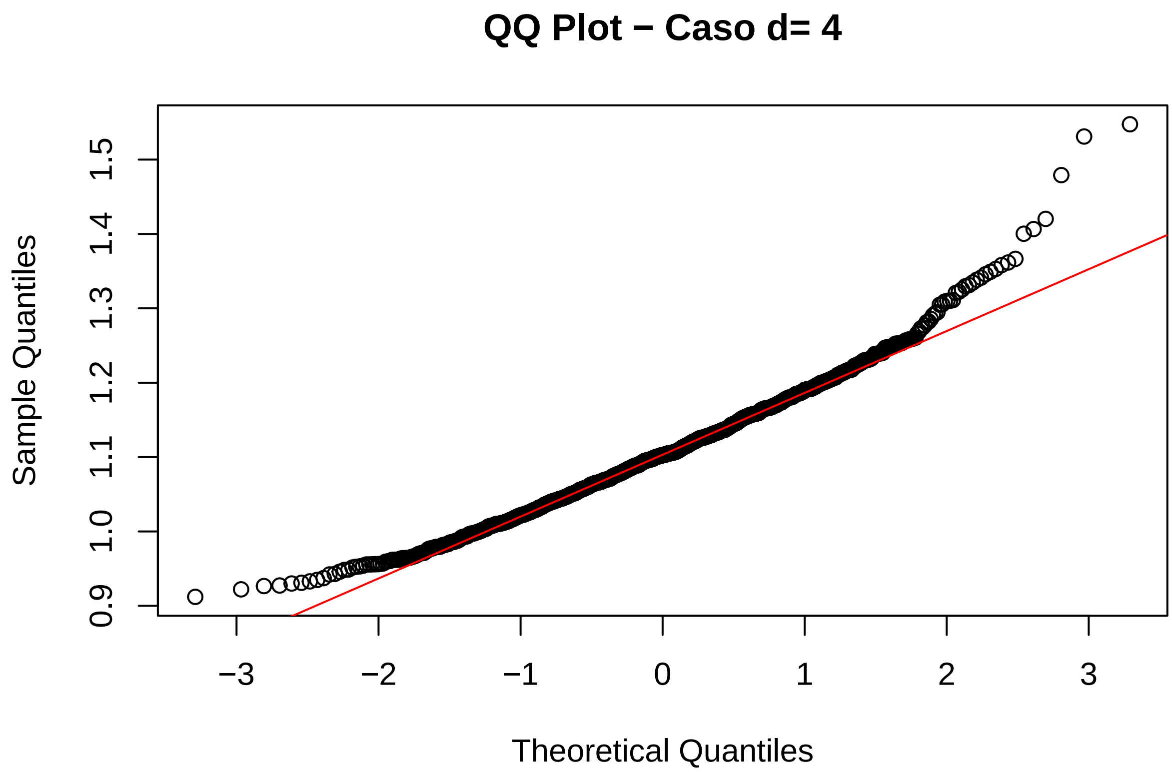

```r
# Crear el QQ plot

for (i in 1:ncol(X_N_results)) {
```

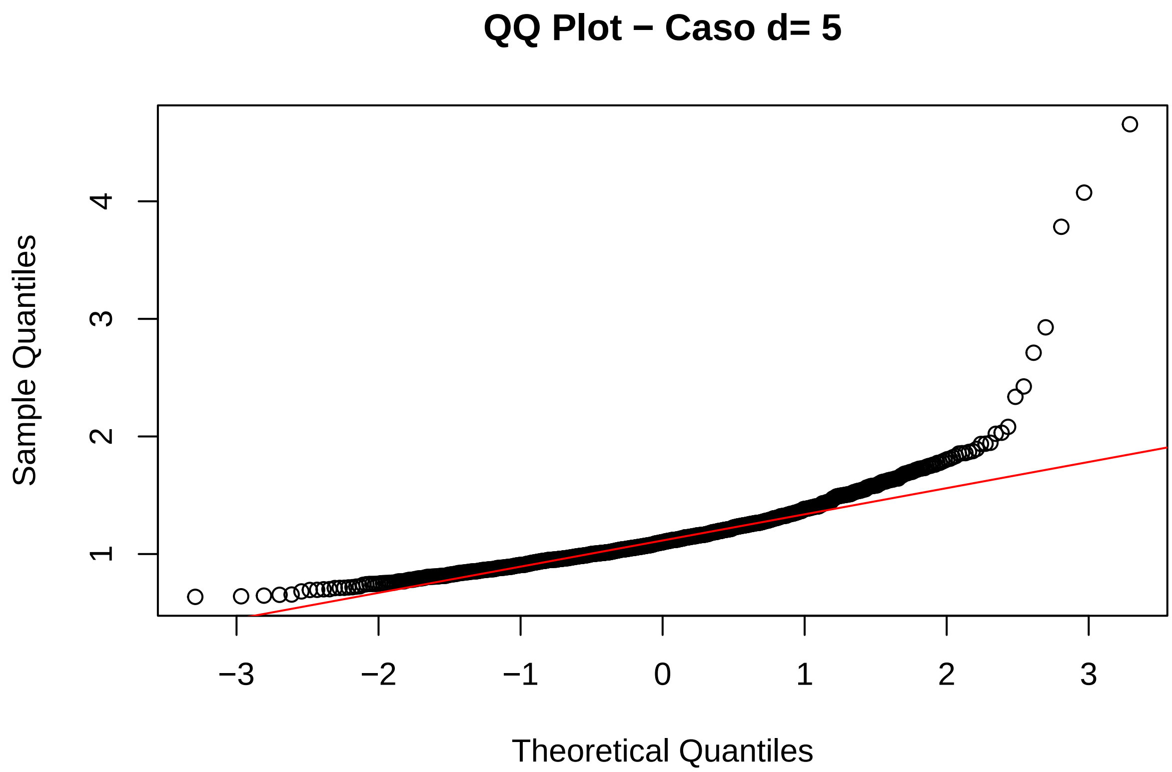
```
  qqnorm(X_N_results[, i], main = paste("QQ Plot - Caso d=", i))
  qqline(X_N_results[, i], col = "red")
}
```




QQ Plot − Caso d= 2
Sample Quantiles
1.14
1.16
1.18
1.20
1.22
−3
−2
−1
0
1
2
3
Theoretical Quantiles


QQ Plot − Caso d= 3
Sample Quantiles
1.05
1.10
1.15
1.20
1.25
−3
−2
−1
0
1
2
3
Theoretical Quantiles


QQ Plot − Caso d= 4
Sample Quantiles
0.9
1.0
1.1
1.2
1.3
1.4
1.5
−3
−2
−1
0
1
2
3
Theoretical Quantiles



Esta vez al aumentar en numero de series se puede observar de manera mas clara la tendencia, sobre todo en los casos d=1,2,3 , mientras que en los casos d=4 y d=5 parece desviarse en los extremos.

Calculamos $\sqrt{N} ln(X_N(r,d))$

```
# Longitud de cada serie
N <- length_series  # Esto es igual a 1000 en nuestro caso

# Calcular sqrt(N)
sqrt_N <- sqrt(N)

# Calcular sqrt(N) * ln(X_{N}(r,d)) para cada serie y cada valor de d
sqrt_N_ln_X_N <- sqrt_N * log(X_N_results)

# Ver los resultados
head(sqrt_N_ln_X_N)
```

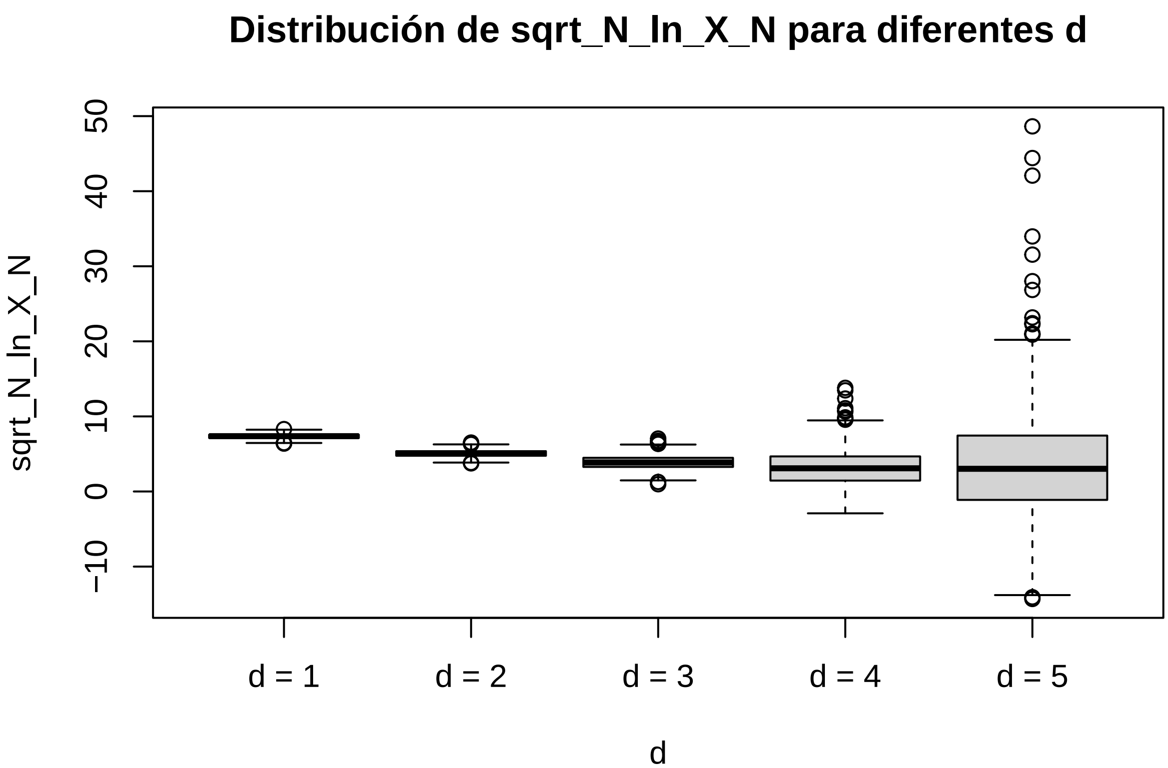
```
##         d = 1    d = 2    d = 3      d = 4      d = 5
## [1,] 7.173452 5.281998 1.770821  0.5706273 -1.3914126
## [2,] 6.780021 6.307930 7.050485  2.8282119  1.2600383
## [3,] 7.091278 4.934576 3.947314 -1.1556274 -0.2997878
## [4,] 7.041995 4.650882 2.729541  3.1306799 -8.3717640
## [5,] 7.641914 5.071842 3.292184 -1.0871663  3.7548925
## [6,] 7.129824 5.575926 4.023463  0.7182414 -5.0949832
```

y vemos como se comporta:

```r
# Resumen de los resultados
summary(sqrt_N_ln_X_N)
```

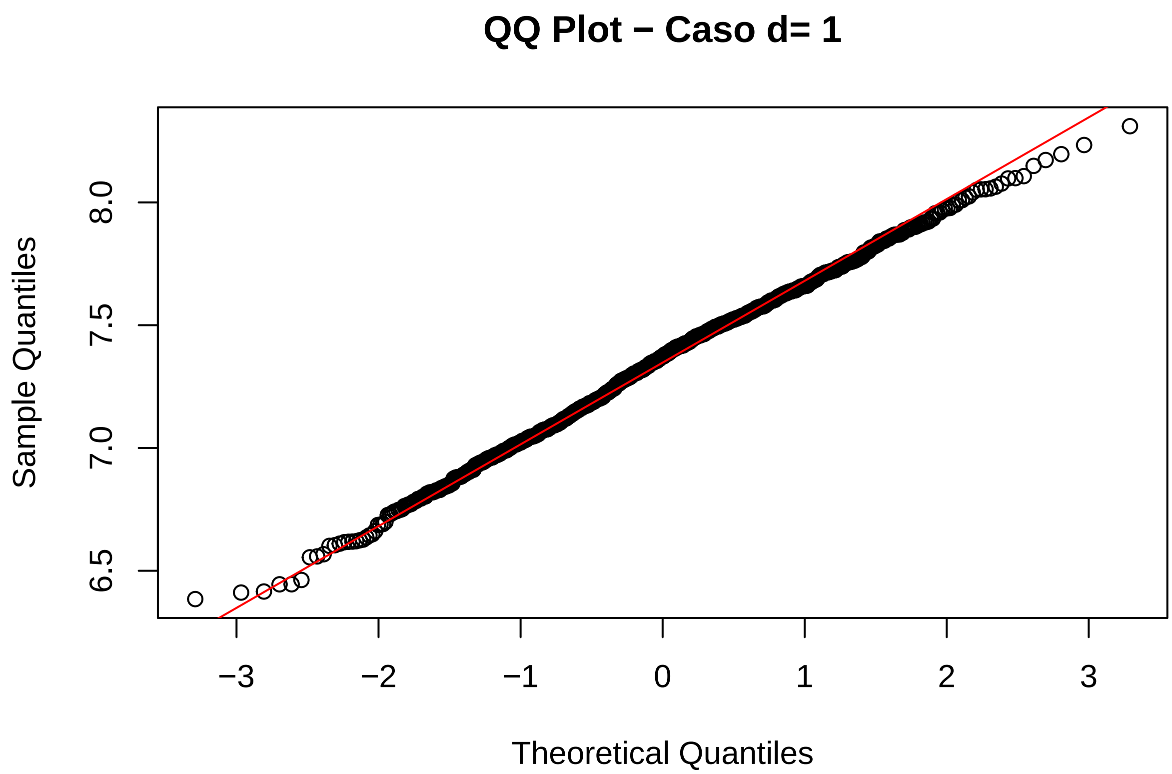
```
##      d = 1           d = 2            d = 3             d = 4
##  Min.   :6.385   Min.   :3.747   Min.   :0.9671   Min.   :-2.911
##  1st Qu.:7.123   1st Qu.:4.761   1st Qu.:3.2744   1st Qu.: 1.453
##  Median :7.372   Median :5.062   Median :3.8683   Median : 3.089
##  Mean   :7.352   Mean   :5.066   Mean   :3.8881   Mean   : 3.137
##  3rd Qu.:7.572   3rd Qu.:5.368   3rd Qu.:4.4852   3rd Qu.: 4.672
##  Max.   :8.310   Max.   :6.516   Max.   :7.0505   Max.   :13.808
##      d = 5
##  Min.   :-14.313
##  1st Qu.: -1.120
##  Median :  3.024
##  Mean   :  3.572
##  3rd Qu.:  7.444
##  Max.   : 48.634
```

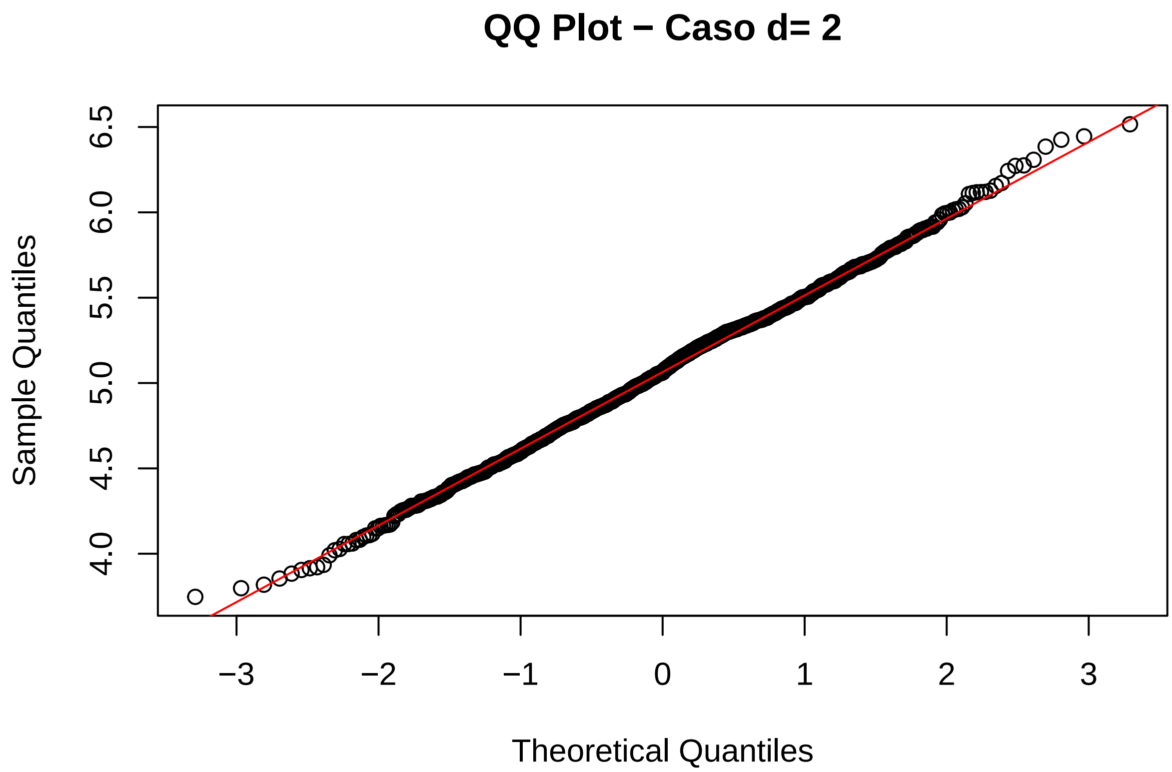
```r
# Graficar los resultados
boxplot(sqrt_N_ln_X_N, main="Distribución de sqrt_N_ln_X_N para diferentes d", ylab="sqrt_N_ln_X_N", xla
```

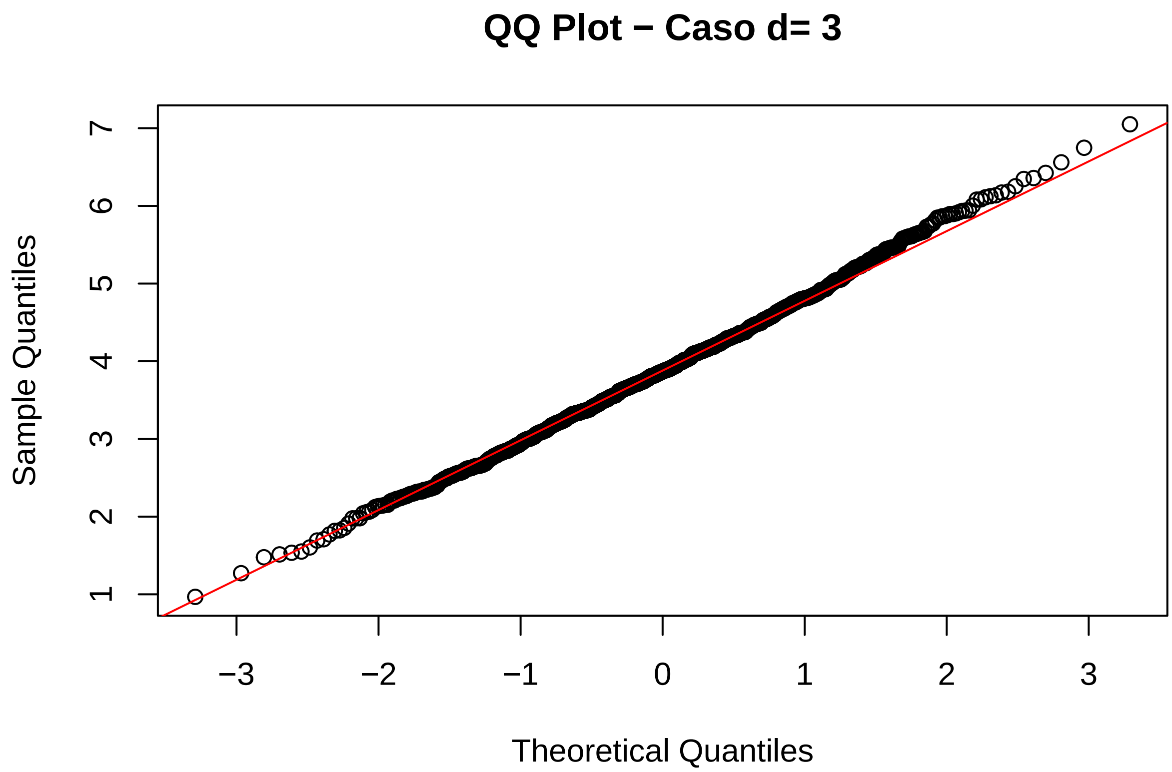


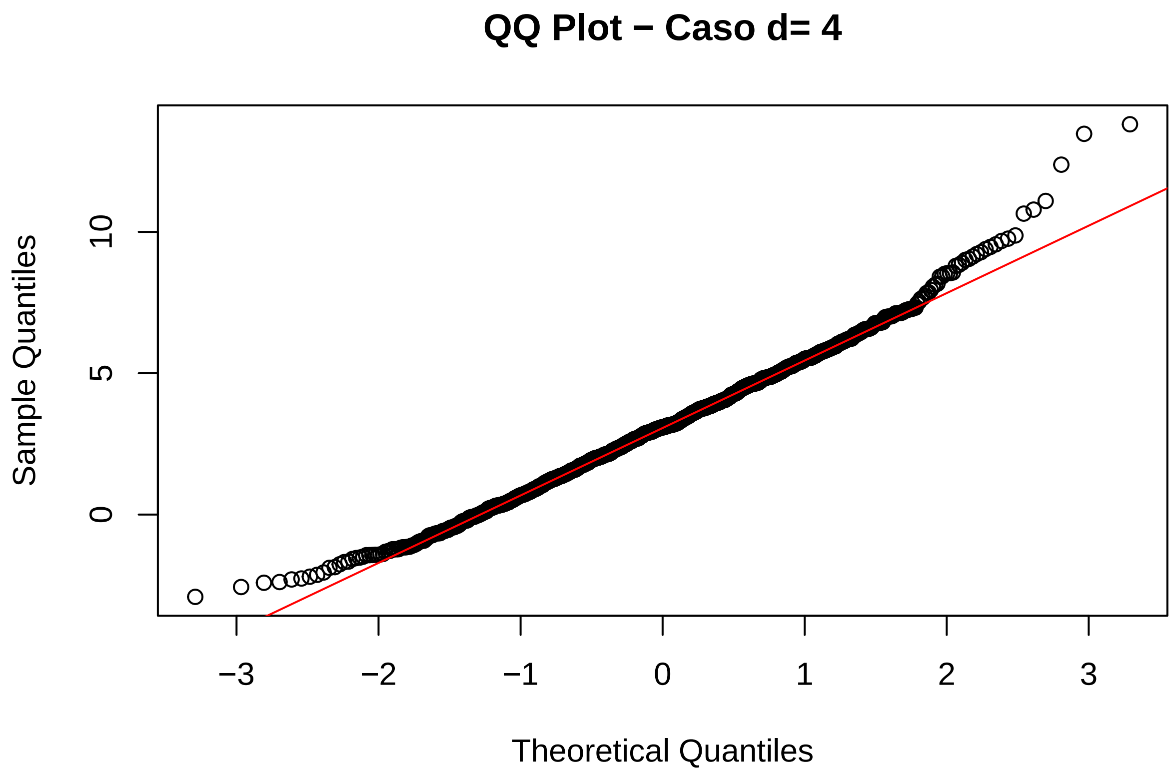

```r
# Crear el QQ plot

for (i in 1:ncol(sqrt_N_ln_X_N)) {
```

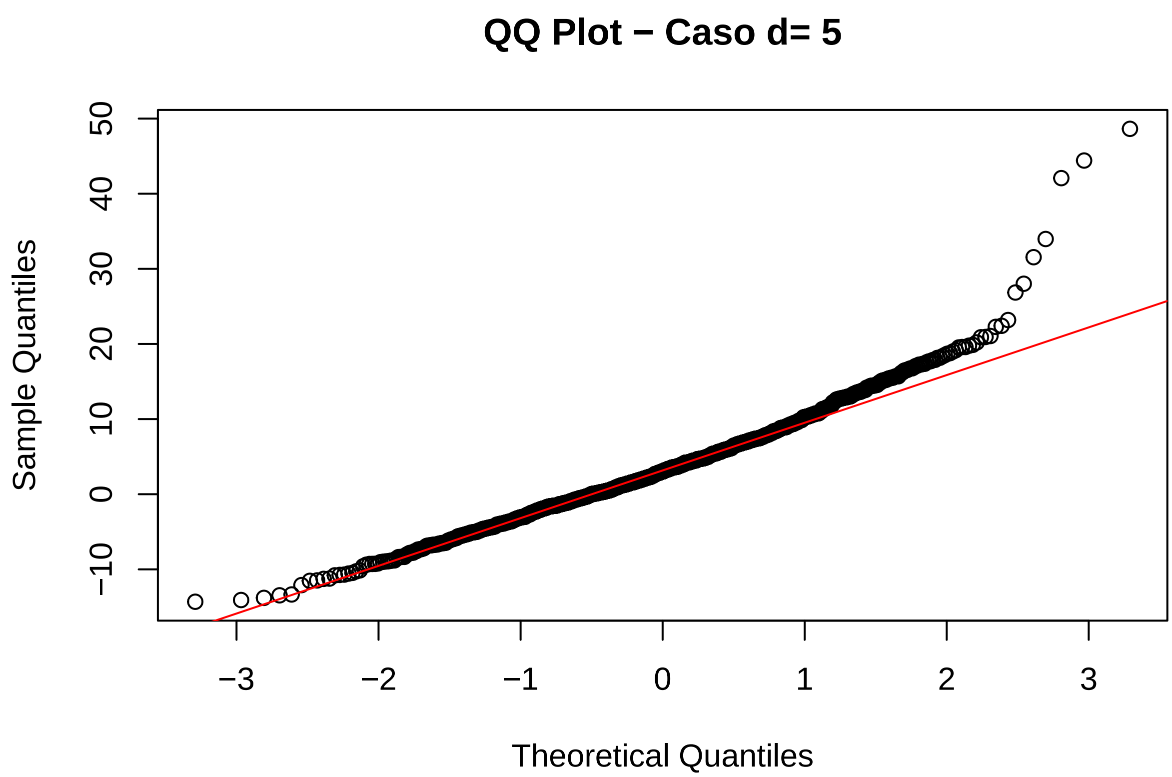
```
  qqnorm(sqrt_N_ln_X_N[, i], main = paste("QQ Plot - Caso d=", i))
  qqline(sqrt_N_ln_X_N[, i], col = "red")
}
```




QQ Plot − Caso d= 2
Sample Quantiles
Theoretical Quantiles
4.0
4.5
5.0
5.5
6.0
6.5
−3
−2
−1
0
1
2
3


QQ Plot − Caso d= 3
Sample Quantiles
1
2
3
4
5
6
7
−3
−2
−1
0
1
2
3
Theoretical Quantiles


QQ Plot − Caso d= 4
Sample Quantiles
0
5
10
−3
−2
−1
0
1
2
3
Theoretical Quantiles



De la misma manera para $\sqrt{N}ln(X_N(r,d))$ se tiene mucho mas clara la tendencia de los datos, esta vez sin tanto desvió como para el caso anterior, mostrando las propiedades de este estadístico.