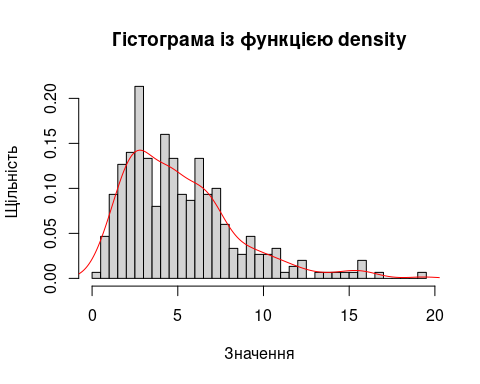
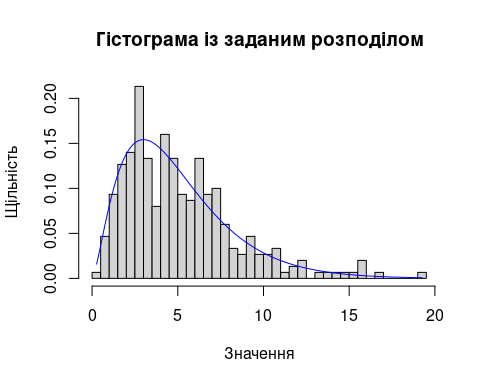
Лабораторна робота №2

Відкриємо наші дані і подивимось на гістограму і щільність розподілу Подивимось на використання функції density з гістограмою і просто функцію щільності для розподілу ксі квадрат із ступенем свободи 5.

data <- read.table("distr8.txt", header = TRUE)  
  
X <- sort(data$X)  
  
hist(X, breaks = 50, freq = FALSE, main = "Гістограма із функцією density", xlab = "Значення", ylab = "Щільність")  
lines(density(X), col = "red")



hist(X, breaks = 50, freq = FALSE, main = "Гістограма із заданим розподілом", xlab = "Значення", ylab = "Щільність")  
x <- seq(min(X), max(X), by = 0.1)  
y <- dchisq(x, 5)  
lines(x, y, col = "blue")



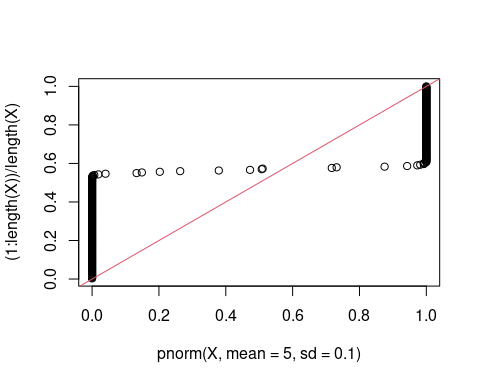
Додатково перевіримо наші дані використовуючи QQ і PP діаграми

Подивимось на QQ і PP діаграми із такими розподілами :

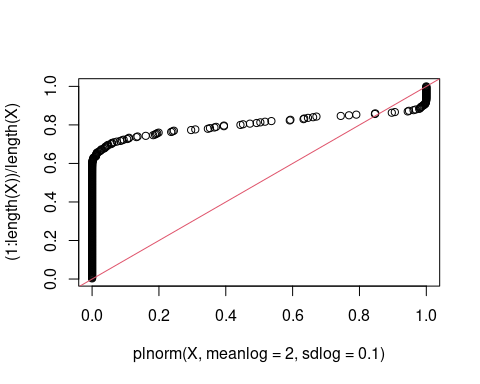
(а) нормальний; (б) логнормальний; (в) експоненційний; (г) χ²;

Подивимось на діаграми і оцінимо їх якість. Використаємо для цього різні розподіли із різними параметрами і побачимо, що ксі квадрат розподіл із ступенем свободи 5 найкраще описує наші дані

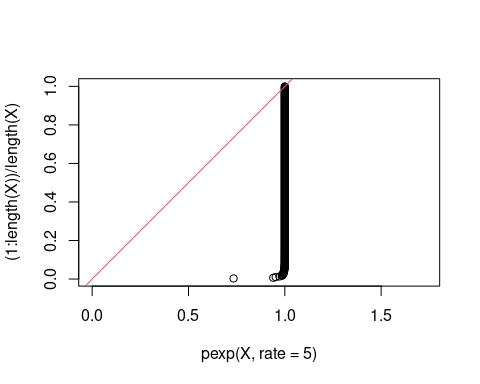
plot(pnorm(X, mean = 5, sd = 0.1), (1:length(X))/length(X))  
abline(0,1,col=2)



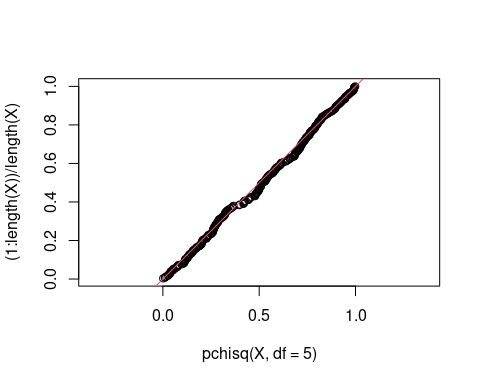
plot(plnorm(X, meanlog = 2, sdlog = 0.1), (1:length(X))/length(X))  
abline(0,1,col=2)



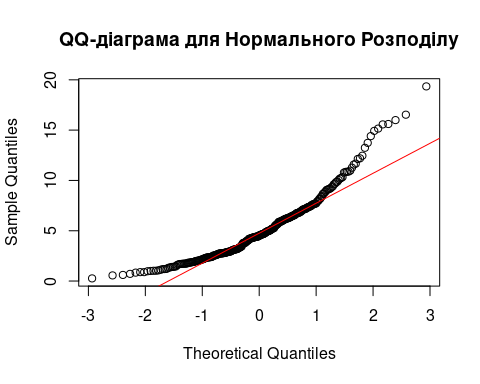
plot(pexp(X, rate=5), (1:length(X))/length(X), asp=1)  
abline(0,1,col=2)



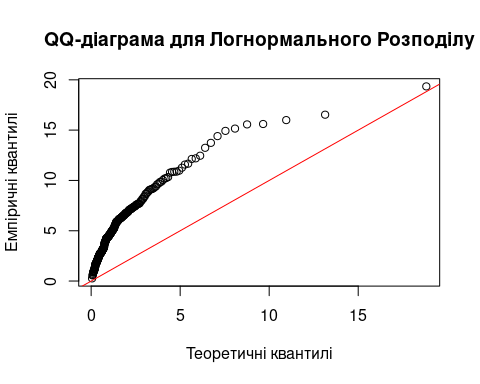
plot(pchisq(X, df = 5), (1:length(X))/length(X), asp=1)  
abline(0,1,col=2)



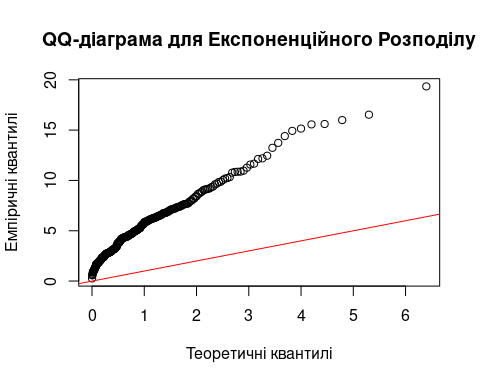
qqnorm(X, main = "QQ-діаграма для Нормального Розподілу")  
qqline(X, col = "red")



qqplot(qlnorm(ppoints(X)), X, main = "QQ-діаграма для Логнормального Розподілу", xlab = "Теоретичні квантилі", ylab = "Емпіричні квантилі")  
abline(a = 0, b = 1, col = "red")



qqplot(qexp(ppoints(X)), X, main = "QQ-діаграма для Експоненційного Розподілу", xlab = "Теоретичні квантилі", ylab = "Емпіричні квантилі")  
abline(a = 0, b = 1, col = "red")

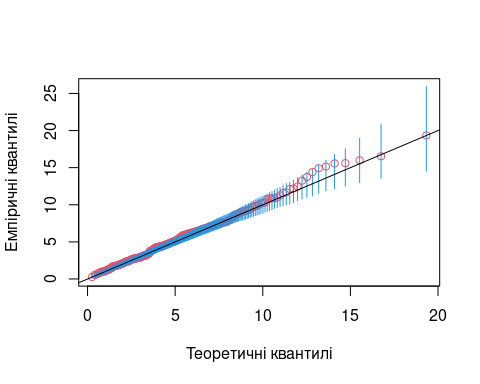


qqplot(qchisq(ppoints(X), df = 5), X, main = "QQ-діаграма для χ² Розподілу", xlab = "Теоретичні квантилі", ylab = "Емпіричні квантилі")  
abline(a = 0, b = 1, col = "red")



Як бачимо, то дійсно розподіл ксі квадрат із ступенем свободи 5 найкраще описує наші дані. Тепер подивимось і проаналізуємо графік QQ діаграми із прогнознимим інтервалами. Переробимо і використаємо функцію QQplot із підручника “Комп’ютерна статистика” Р.Майборода. Маємо:

QQplot <- function(x, df, K=1000, alpha=0.05) {  
 n <- length(x)  
 chiSquareQ <- qchisq(ppoints(n), df = df)  
 sx <- sort(x)  
 W <- matrix(rchisq(K\*n, df = df), nrow = n, ncol = K)  
 W <- apply(W, 2, sort)  
 tops <- apply(W, 1, quantile, probs = 1 - alpha/2)  
 bots <- apply(W, 1, quantile, probs = alpha/2)  
 plot(c(chiSquareQ, chiSquareQ, chiSquareQ), c(tops, bots, sx), type = "n", xlab = "Теоретичні квантилі", ylab = "Емпіричні квантилі")  
 points(chiSquareQ, sx, col = 2)  
 segments(chiSquareQ, bots, chiSquareQ, tops, col = 4)  
 abline(0, 1, col = 1)  
}  
  
QQplot(X, 5)



Отже, дійсно наш розподіл збігається із розподілом ксі квадрат із ступенем свободи 5