

Command	Purpose	Example
plot(x, y)	Basic scatter or line plot.	plot(x, y, type = "p")
hist(x)	Histogram of data.	hist(x, breaks = 30, col = "lightblue")
boxplot(x)	Boxplot to show spread and outliers.	boxplot(x)
barplot(h)	Bar chart for categorical data.	barplot(table(gender))
lines(x, y)	Add a line to an existing plot.	lines(xs, dnorm(xs), col = "red")
points(x, y)	Add points to an existing plot.	points(x, y, col = "blue")
curve(expr)	Plot a mathematical function.	curve(dnorm(x), from=-3, to=3)
abline()	Add a straight line (horizontal, vertical, or regression).	abline(h=0, col="red")
legend()	Add a legend to a plot.	legend("topright", legend=c("Data","Model"), col=c("black","red"), lty=1)
density(x) + lines()	Kernel density estimate of a distribution.	lines(density(x), col="darkgreen")
ecdf(x)	Empirical cumulative distribution function (ECDF).	plot(ecdf(x), main="Empirical CDF", col="blue")

# Simulation of a Random Variable with Density $f(x) = \frac{4}{x^5}$ , $x > 1$

## 1. Theoretical Background

The given density function:

$$f(x) = \frac{4}{x^5}, \quad x > 1$$

is a **Pareto distribution** with parameters  $x_m = 1$  and  $\alpha = 4$ .

The cumulative distribution function (CDF) is:

$$F(x) = 1 - \frac{1}{x^4}, \quad x > 1$$

By inversion of the CDF:

$$F^{-1}(u) = (1 - u)^{-1/4}, \quad u \in (0, 1)$$

Thus, to generate random variables with this density, we can use the transformation:

$$X = (1 - U)^{-1/4}, \quad \text{where } U \sim \text{Uniform}(0, 1)$$

## 2. Simulation in R

The following R code simulates  $n = 10000$  random values from the given distribution, and compares the empirical results with the theoretical density and CDF.

```
set.seed(123)
n <- 10000
u <- runif(n)
# Inverse transform method
x <- (1 - u)^(1/4)
# Theoretical density and CDF
f_theor <- function(x) ifelse(x > 1, 4/x^5, 0)
F_theor <- function(x) ifelse(x > 1, 1 - 1/x^4, 0)

# 1. Histogram and theoretical density
hist(x, breaks = 50, freq = FALSE, col = "lightblue", main = "Simulated Data vs Theoretical Density", xlab = "x", xlim = c(1, 5))
curve(f_theor, from = 1, to = 5, add = TRUE, col = "red", lwd = 2)
legend("topright", legend = c("Histogram", "Theoretical Density"), col = c("lightblue", "red"), lty = c(1, 1), lwd = c(10, 2), bty = "n")

# 2. Empirical and theoretical CDF
plot(ecdf(x), main = "Empirical vs Theoretical CDF", col = "blue", lwd = 2, ylab = "F(x)", xlab = "x", xlim = c(1, 5))
curve(F_theor, from = 1, to = 5, add = TRUE, col = "red", lwd = 2, lty = 2)
legend("bottomright", legend = c("Empirical CDF", "Theoretical CDF"), col = c("blue", "red"), lty = c(1, 2), lwd = c(10, 2), bty = "n")
```

## 3. Comparison Results

- The histogram closely follows the theoretical density curve  $f(x) = \frac{4}{x^5}$ .
- The empirical cumulative distribution function aligns well with the theoretical CDF  $F(x) = 1 - \frac{1}{x^4}$ .

The following figures illustrate the comparison results:

## 4. Theoretical Moments

For a Pareto distribution with  $\alpha = 4$ ,  $x_m = 1$ :

$$E[X] = \frac{\alpha x_m}{\alpha - 1} = \frac{4}{3}, \quad \text{Var}[X] = \frac{\alpha x_m^2}{(\alpha - 1)^2(\alpha - 2)} = \frac{4}{18} \approx 0.222$$

Empirical estimates from the simulated data should be close to these theoretical values.



Figure 1: Histogram and theoretical density.



Figure 2: Empirical and theoretical cumulative distribution functions.