Assignment I - CompStat2023

Name of the Team

Team Member A, Team Member B, Team Member C

Contents

1 Montecarlo Simulation II

1.1 Point a)

To prove that the random variable $Y = U^{-\frac{1}{\gamma}}$, where U is Uniform in $(0, 1)$, follows a Pareto distribution, we need to show

Let's begin by finding the cumulative distribution function (CDF) of Y. We have:\

$$F_Y(y) = P(Y \leq y) = P\left(U^{-1/\gamma} \leq y\right)$$

Taking the reciprocal of both sides, we get:\

$$P\left(\frac{1}{U^{1/\gamma}} \leq \frac{1}{y}\right)$$

Since U follows a standard uniform distribution on the interval $(0, 1)$, the above expression is equivalent to:\

$$P\left(U^{1/\gamma} \geq y^{-1}\right) = P\left(U \geq y^{-\gamma}\right)$$

Since U is uniformly distributed, the probability $P(U \geq y^{-\gamma})$ is equal to $1 - P(U < y^{-\gamma})$. Therefore, the CDF of Y can be written as:\

$$F_Y(y) = 1 - P(U < y^{-\gamma})$$

The probability $P(U < y^{-\gamma})$ represents the CDF of a standard uniform distribution evaluated at $y^{-\gamma}$. Since the CDF of a standard uniform distribution is $F_U(u) = u$ for $0 < u < 1$, we have:\

$$F_Y(y) = 1 - y^{-\gamma}$$

Now, let's differentiate the CDF to obtain the PDF of Y:\

$$f_Y(y) = \frac{d}{dy}F_Y(y) = \gamma y^{-\gamma-1}$$

Comparing this PDF with the PDF of a Pareto distribution, we can see that they are identical:\

$$f_Y(y) = \gamma y^{-\gamma-1} = \gamma x^{-(\gamma+1)}$$

Therefore, we have shown that the random variable $Y = U^{-\frac{1}{\gamma}}$ follows a Pareto distribution with parameter γ.\

1.2 point b)

To find the distribution of $Y = \log(X)$, where X follows a Pareto distribution with parameter γ, we need to determine the probability density function (PDF) of Y.\

Let's start by expressing the relationship between X and Y using the transformation function:\

$$Y = \log(X)$$

To find the PDF of Y, we can use the cumulative distribution function (CDF) method. Let's begin by finding the CDF of Y.\

We have:\

$$F_Y(y) = P(Y \leq y) = P(\log(X) \leq y)$$

Taking the exponential of both sides, we get:\

$$P(X \leq e^y)$$

Since X follows a Pareto distribution, its CDF is given by:\

$$F_X(x) = 1 - \left(\frac{1}{x}\right)^\gamma \text{ for } x \geq 1$$

Using the CDF of X, we can express the CDF of Y as follows:\

$$F_Y(y) = P(X \leq e^y) = 1 - \left(\frac{1}{e^y}\right)^\gamma = 1 - e^{-\gamma y}$$

To find the PDF of Y, we differentiate the CDF with respect to y:\

$$f_Y(y) = \frac{d}{dy}F_Y(y) = \gamma e^{-\gamma y}$$

Therefore, the PDF of Y is given by:\

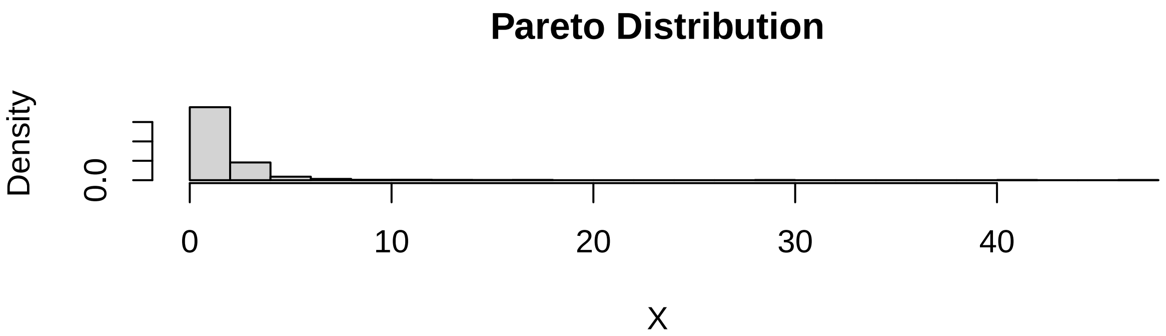
$$f_Y(y) = \gamma e^{-\gamma y}$$

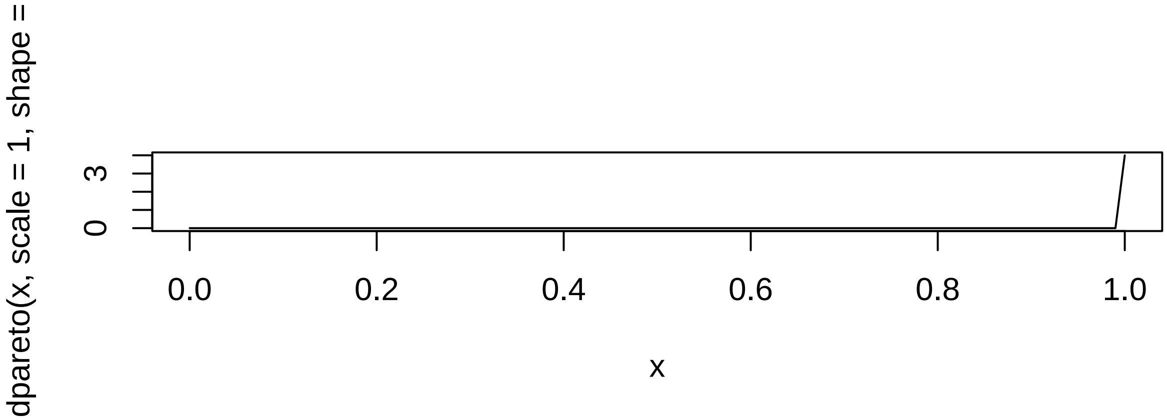
In conclusion, when $Y = \log(X)$ and X follows a Pareto distribution with parameter γ, the resulting distribution of Y is an exponential distribution with parameter γ.\

1.3 point c)

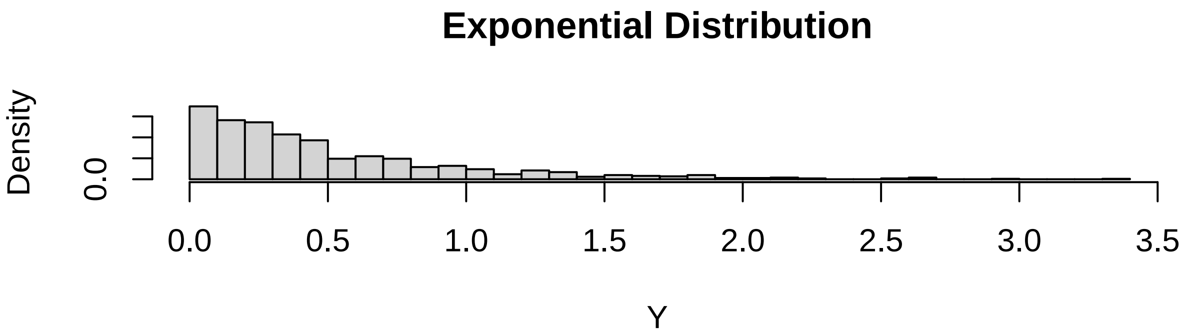
```
set.seed(123)

# Function to generate samples from a Pareto distribution
pareto_sampler <- function(n, gamma) {
  u <- runif(n)  # Generate n samples from a uniform distribution
  x <- (1/u)^(1/gamma)  # Transform the uniform samples to follow a Pareto distribution
  return(x)
}

# Function to generate samples from the logarithmic distribution
logarithmic_sampler <- function(n, gamma) {
  x <- pareto_sampler(n, gamma)  # Generate n samples from a Pareto distribution
  y <- log(x)  # Transform the Pareto samples to follow a logarithmic distribution
  return(y)
}
```

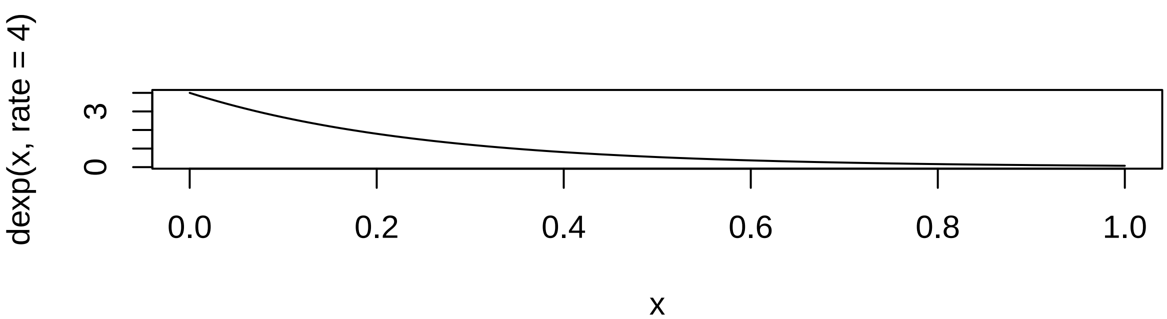
```r
# Parameters
n <- 1000  # Number of samples
gamma <- 2  # Parameter of the Pareto distribution

# Generate samples from the Pareto and logarithmic distributions
x_samples <- pareto_sampler(n, gamma)
y_samples <- logarithmic_sampler(n, gamma)

# Plot histograms
par(mfrow = c(2, 1))
hist(x_samples, breaks = 30, main = "Pareto Distribution", xlab = "X", freq = FALSE, xlim = c(0, max(x_s
curve(dpareto(x, scale=1, shape = 4))
```



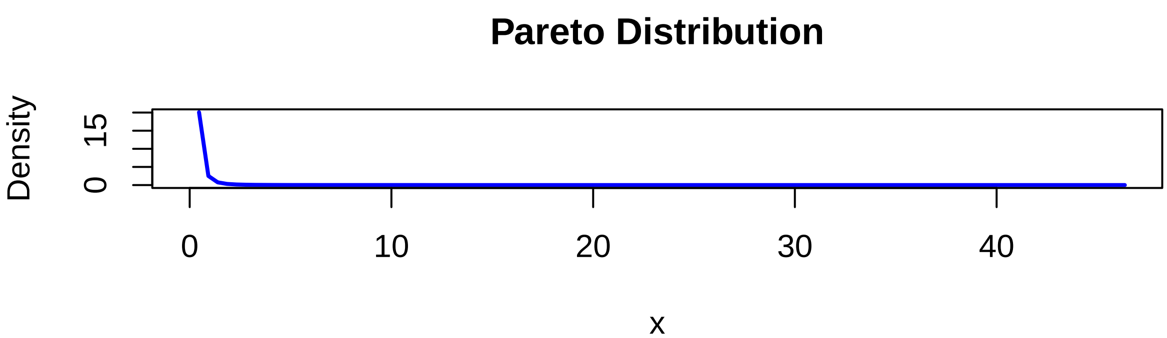


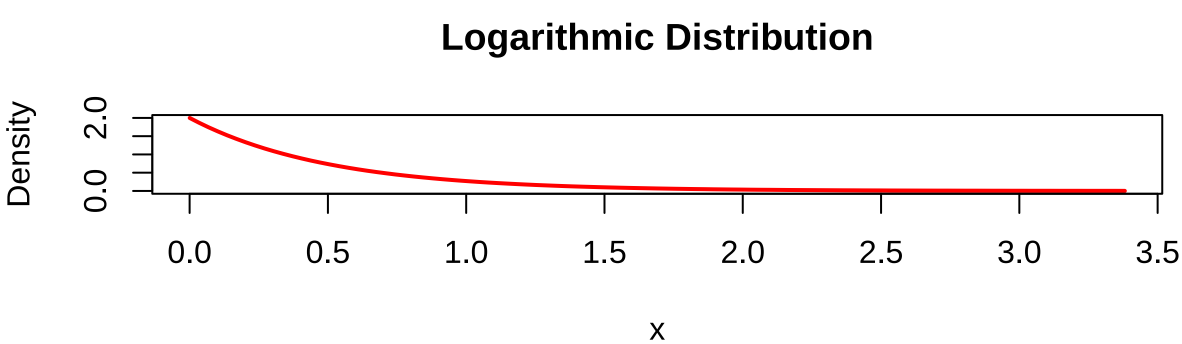

```r
hist(y_samples, breaks = 30, main = "Exponential Distribution", xlab = "Y", freq = FALSE, xlim = c(min(y
curve(dexp(x, rate=4))
```






```
# Plot density plots
curve((gamma * x^-(gamma + 1)), from = 0, to = max(x_samples), col = "blue", lwd = 2, ylab = "Density",
curve(dexp(x, gamma), from = min(y_samples), to = max(y_samples), col = "red", lwd = 2, ylab = "Density
```





1.4 point d)

```r
# Analytical approach
gamma <- 2  # Pareto parameter
p_analytical <- 1 - (1/5)^gamma

# Monte Carlo estimate
num_samples <- 1000000  # Number of samples
samples <- rpareto(num_samples,scale=1, shape=gamma)
p_mc <- mean(num_samples < 5)

# Output results
p_analytical
```

```
## [1] 0.96
```

```r
p_mc
```

```
## [1] 0
```