Statistical Modelling 2023-2024
(Prof. C. Ley, S. Loizidou)
Practical 1

Question 1 (Calculating basic quantities of probability distributions)

Let X be a continuous random variable with probability density function (pdf) given by

$$f(x) = \begin{cases} c\left(\frac{1}{3}x + \frac{1}{2}\right), & 1 < x < 4, \\ 0, & \textit{otherwise}. \end{cases}$$

a) Calculate the value of c.

b) Compute the cumulative distribution function (cdf) of X, $F(x)$.

c) Compute $\mathbb{E}[X]$ and $Var(X)$.

Question 2 (The Pareto distribution)

An interesting distribution to model heavy-tailed data on $\mathbb{R}^+$ is the Pareto distribution with density

$$f_{\alpha,\theta}(x) = kx^{-\alpha} I_{[x>\theta]};$$

with parameters $\theta > 0$ and $\alpha > 1$.

a) Calculate the normalizing constant k.

b) Give the cumulative distribution function F_X.

c) Calculate the moment of order $s (s \in \mathbb{N}_0)$. What assumptions need to be fulfilled for its existence? Deduce from here $\mathbb{E}[X]$ and $Var[X]$.

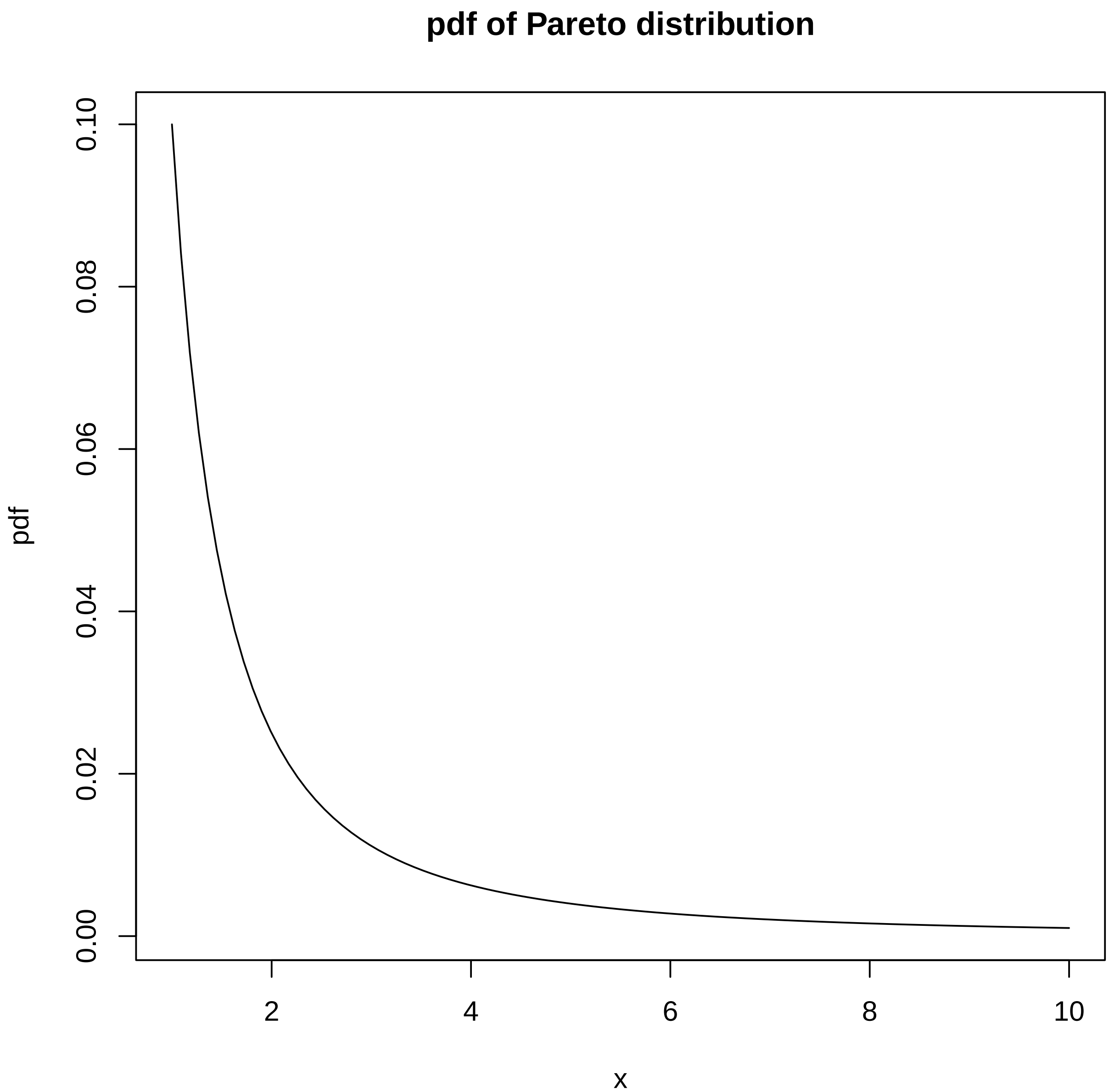
d) What is the distribution of the transformed random variable $(\alpha - 1)\ln(X/\theta)$.

e) The pdf of the Pareto distribution can be plotted in R as follows. What are the limitations of this distribution? You can copy the following code in R and change the parameters *alpha* and *theta* to see how the pdf changes.

```r
# We define a function that estimates the pdf of the Pareto distribution for
# user-specified values for alpha and theta. We could also have used the function
# dpareto() from the package EnvStats. Note that the parameters for dpareto() are
# defined differently compared to our function.

pdf_pareto <- function(data, alpha, theta){
  pdf <- (alpha - 1) * data^(-alpha) / (theta^(1 - alpha)) * as.numeric(data > theta)
  return(pdf)
}

curve(pdf_pareto(data = x, alpha = 2, theta = 1/10), from = 1, to = 10,
main = 'pdf of Pareto distribution', ylab = 'pdf')
```



Question 3 (The Cauchy distribution)

Let X and Y be independent standard normal random variables. Then the random variable $C = X/Y$ follows a Cauchy distribution.

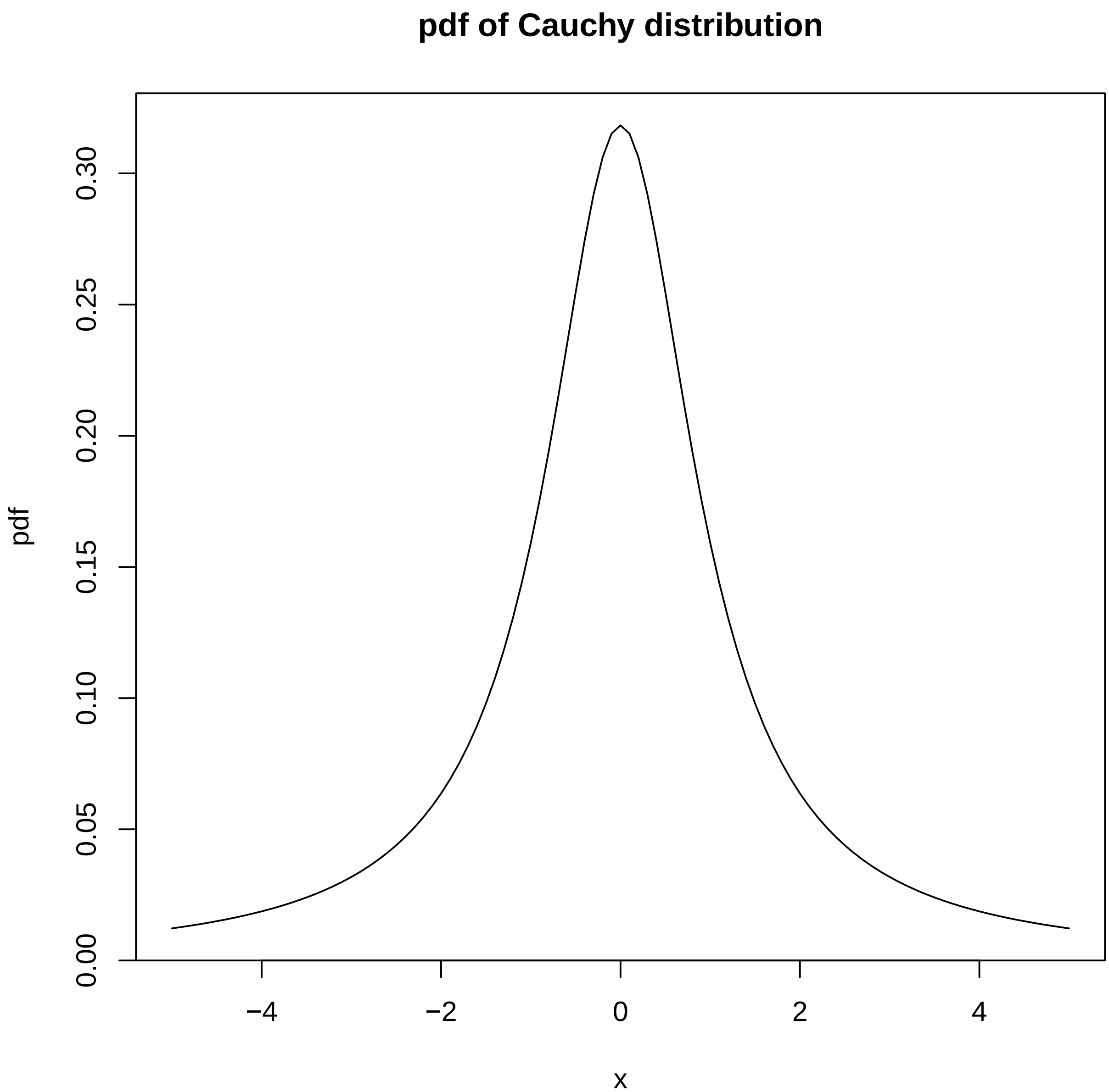
a) Calculate the density $f_C(x)$ of C.

b) What is the distribution of $1/C$?

c) Calculate $\int_{-R}^{R} x f_C(x) dx$ and $\int_{-2R}^{R} x f_C(x) dx$. What can you conclude from this?

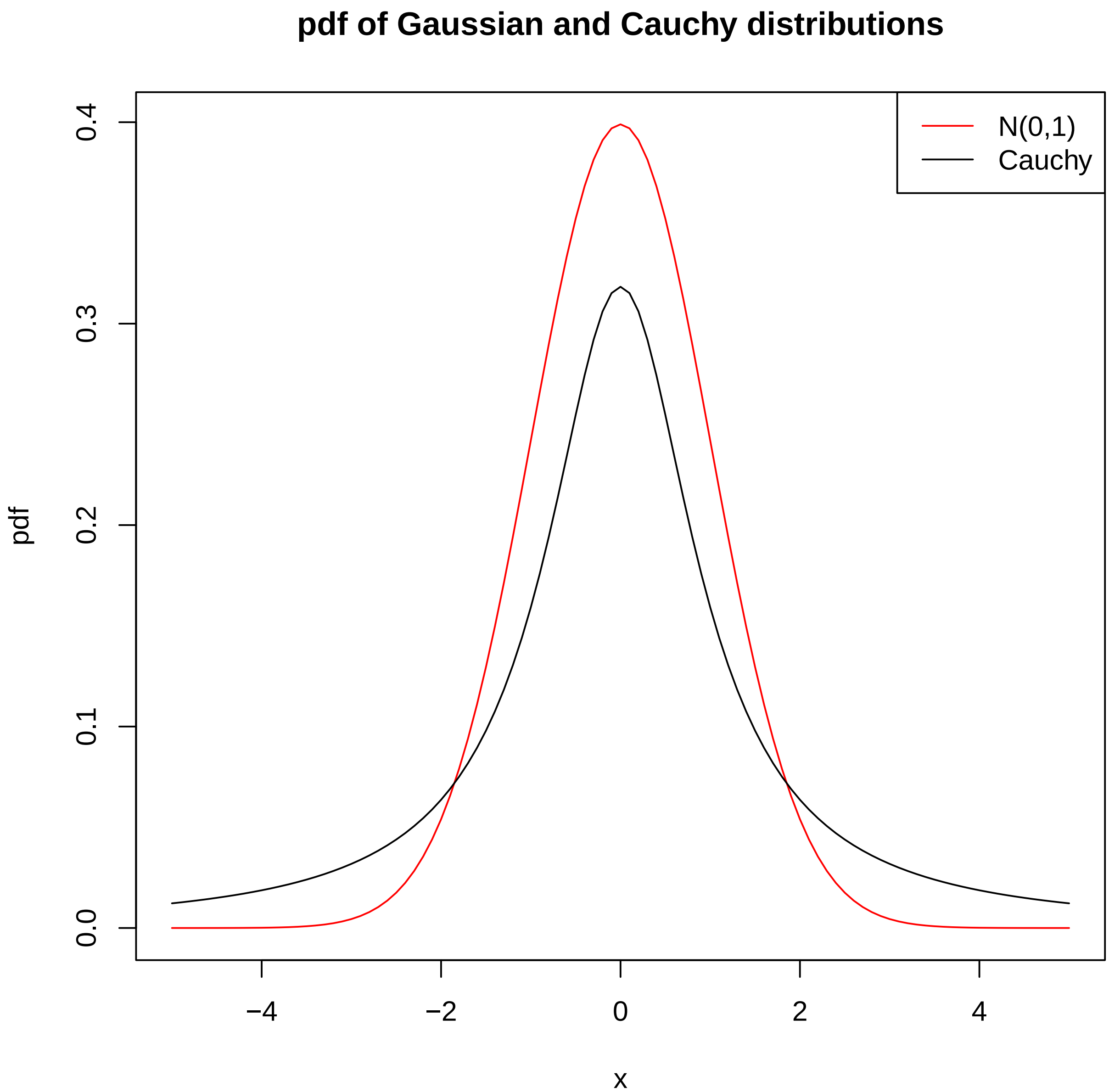
d) The pdf of the Cauchy distribution can be plotted in R as follows. What are the limitations of this distribution?

```r
# Here we use the built-in function in R for the estimation of the pdf.
curve(dcauchy(x), from = -5, to = 5, main = "pdf of Cauchy distribution",
ylab = 'pdf')
```



The Cauchy distribution has heavier tails than the Gaussian.

```r
curve(dnorm(x), from = -5, to = 5, main = "pdf of Gaussian and Cauchy distributions",
ylab = 'pdf', col = 'red')
curve(dcauchy(x), from = -5, to = 5, ylab = 'pdf', add = T)
legend("topright", c('N(0,1)', "Cauchy"), col = c("red", "black"), lty = 1)
```


pdf of Gaussian and Cauchy distributions
N(0,1)
Cauchy
pdf
0.0
0.1
0.2
0.3
0.4
−4
−2
0
2
4
x