

**Question 1**

Recall from Section 8.3.8 in Prof. Veraart's notes that the p.d.f. of the standard Cauchy distribution is

$$f_X(x) = \frac{1}{\pi(1+x^2)}, \quad \text{with support } x \in \mathbb{R}.$$

- (a) Show that  $f_X$  is a probability density function (p.d.f.) and plot  $y = f_X(x)$ .
- (b) Compute the first (raw) moment of  $X$ ,  $\mu = \mu'_1 = E(X)$ .
- (c) Compute the  $k$ th central moment of  $X$ ,  $\mu_k = E((X - \mu)^k)$  for  $k \in \{2, 3, \dots\}$ .
- (d) Compute the second raw moment of  $X$ ,  $\mu'_2 = E(X^2)$ .

**Question 2**

Suppose that the random variable  $X$  is known to only take non-zero values in the bounded range  $[a, b]$ , i.e. the support of  $X$  is  $[a, b]$ .

- (a) Derive the expression  $(X - a)(X - b) = (X - \frac{a+b}{2})^2 - \frac{(b-a)^2}{4}$ .
- (b) Use the expression from (a) to prove that  $\text{Var}(X) \leq \frac{(b-a)^2}{4}$ .
- (c) Conclude that if  $X \sim \text{Bern}(p)$ , for some  $p \in [0, 1]$ , then  $\text{Var}(X) \leq \frac{1}{4}$ .
- (d) Is the bound  $\text{Var}(X) \leq \frac{(b-a)^2}{4}$  tight? In other words, is there a distribution  $F_X$  with support  $[a, b]$  for which  $\text{Var}(X) = \frac{(b-a)^2}{4}$ ?

**Question 3 (using R)**

- (a) Use R to generate 10 observations from a normal distribution with mean 3 and variance 2. Save the values in a vector  $\mathbf{x}$ .
- (b) Use the built-in R commands to compute the sample mean, variance and standard deviation of  $\mathbf{x}$ .
- (c) Write your own R functions to compute the sample mean and sample variance of  $\mathbf{x}$ .

**Question 4 (using R)**

Suppose there is a file named `data1.txt` which contains the following data:

```
x,y  
2,3  
4,6  
6,9  
8,12
```

(Either download the file from Blackboard, or copy-paste the data into a file and name it `data1.txt`. If you copy-paste, be sure to include an additional blank line after the line with '8, 12'.)

- (a) Use the function `read.table` to read the data from `data1.txt` into a data frame object named `df`.  
**Hint:** because the data is comma-separated, use `sep=','` in the `read.table` command.
- (b) Extract a vector named `x`, containing values (2, 4, 6, 8) from the data frame `df`. Similarly, extract a vector named `y`, containing values (3, 6, 9, 12) from the data frame `df`.
- (c) Create a vector named `z` which is the mean of the two vectors `x` and `y`, i.e. `z` contains four values, the first of which is  $(2 + 3)/2 = 2.5$ .
- (d) Add the vector `z` to the data frame `df` so that `df` contains three columns, `x`, `y` and `z`.
- (e) Write the data frame `df` to a file named `data2.txt`, so that this file contains:

```
x,y,z  
2,3,2.5  
4,6,5  
6,9,7.5  
8,12,10
```