The exercises this week involve some old material so you can check your learning and understanding.

## Exercise 1 - Maximum Likelihood Estimator

Assume you are given datapoints  $(x_i)_{i=1}^N$  with  $x_i \in \mathbb{R}$  coming from a Exponential distribution. The probability density function of a exponential distribution is given by  $f(x) = \lambda \exp(-\lambda x)$  with  $x \in \mathbb{R}$ . Derive the maximum likelihood estimator of the parameter  $\lambda$ .

## **Exercise 2 - Convolutional Layers**

Consider the following  $4 \times 4 \times 1$  input X and a  $2 \times 2 \times 1$  convolutional kernel K with no bias term

$$X = \begin{pmatrix} 1 & 2 & -1 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 2 \\ 2 & 1 & 0 & -1 \end{pmatrix}, \qquad K = \begin{pmatrix} 1, & 0 \\ 2, & 1 \end{pmatrix}$$

- (a) What is the output of the convolutional layer for the case of stride 1 and no padding?
- (b) What if we have stride 2 and no padding?
- (c) What if we have stride 2 and zero-padding of size 1?

## **Exercise 3 - Computational Parameter Counting**

Use PyTorch to load the vgg11 model and automatically compute its number of parameters. Output the number of parameters for each layer and the total number of parameters in the model.

## **Exercise 4 - Influence Functions**

Let  $\hat{\theta}$  and  $\hat{\theta}(\varepsilon)$  be as defined in class. Show that the first order Taylor expansion of  $\hat{\theta}(\varepsilon)$  around  $\varepsilon = 0$  is given by the equation given in class, i.e. by

$$\hat{\theta}(\epsilon) \approx \hat{\theta} + \epsilon \frac{d\hat{\theta}(\epsilon)}{d\epsilon} \Big|_{\epsilon=0}$$