Computer exercise: to be done before October 16

This work can be done in pairs (ce travail peut être effectué en binôme). You have to upload on eCampus a notebook *NAME1-NAME2.ipynb* before october 16. Your code must be carefully commented (in code cells) and explained (in text cells), the results and figures must be displayed when I print the notebook .

The file "*rainfall.txt*" contains the dataframe of precipitation measurements (*rainfall*) observed during storms in the state of Illinois between 1960 and 1964 [1].

1. Draw an histogram of the data; comment on the shape of the distribution; what is the average rainfall per storm? the median rainfall per storm? explain why these measures differ.

We propose to fit these data with a gamma law parameterized by a shape parameter $a > 0$ (`scipy.stats.gamma`):

$$PDF(x) = \frac{x^{a-1}e^{-x}}{\Gamma(a)} \quad \text{for } x \geq 0.$$

2. Find the method of moments estimate $\widehat{a}$ of a (explain how you get the formula). Plot the estimated density on the same figure as the histogram.

   Draw a quantile-quantile plot of the sample with the estimated gamma distribution. What do you think of the resulting fit?

3. Estimate the sampling distribution and the standard error of $\widehat{a}$ by using the bootstrap. Explain (and comment) your code. Draw an histogram of the sampling distribution.

   Compute a 95% bootstrap confidence interval for a.

4. Write a function `logL` which takes as argument the parameter a and returns the value of the negative log-likelihood at point a.

   Draw the graph of the negative log-likeliood.

5. Find the maximum likelihood estimate (MLE) of a (use an optimization algorithm such as `minimize`).

6. (*bonus*) Find an approximate value of the standard error of the MLE of a. Compute a 95% confidence interval for a based on the MLE.

[1] John Rice (Mathematical Statistics and Data Analysis, 1995) Chapter 10 Problem 42