

# Data Analysis - 2020

Exercise sheet no 6:

4. December 2020

The maximum likelihood method

---

## Exercise 1: Maximum Likelihood Estimator (6 Points)

Consider a random variable  $x$ , distributed according to the probability density function

$$P(x|\alpha) = \frac{1}{2}(1 + \alpha x) \quad \text{for } -1 \leq x \leq 1.$$

$n = 20$  measurements of this random variable gave the sample found in `MLE.txt` on the webpage.

- Calculate the negative-log-likelihood function  $\text{NLL}(\alpha) = -\ln L(\alpha)$  for these data as a function of  $\alpha$ . Plot it for  $0 \leq \alpha \leq 1$ .
- Determine the maximum likelihood estimator  $\hat{\alpha}$  using your plot of the negative-log-likelihood.

## Exercise 2: Maximum Likelihood and least squares (14 Points)

In the file `exponential_data.txt` there are 1000 decay time measurements between 0 and  $5 \mu\text{s}$  of muons, which have a mean lifetime of  $\tau = 2 \mu\text{s}$ . The PDF used to generate this data is:

$$P(t) = \frac{1}{\tau \cdot (1 - e^{-5/\tau})} \cdot e^{-t/\tau}$$

where  $\tau$  is the mean decay time in  $\mu\text{s}$  and  $t$  is the measured decay time.

- Calculate twice the negative-log-likelihood ( $2 \cdot \text{NLL}(\tau)$ ) as a function of the lifetime  $\tau$ . Shift the values of the  $2 \cdot \text{NLL}(\tau)$  such that the minimum is zero and plot it for  $1.8 < \tau < 2.2 \mu\text{s}$ .
- Bin the time measurements into a histogram of 40 equally sized bins, from 0 to  $5 \mu\text{s}$ . Use the binned maximum likelihood method to calculate twice the negative-log-likelihood ( $2 \cdot \text{NLL}(\tau)$ ) for these 40 bins as a function of  $\tau$ . Shift the values of the  $2 \cdot \text{NLL}$  such that the minimum is zero and overlay the plot with the unbinned calculation in part a).
- Now calculate the  $\chi^2$  function as a function of  $\tau$  and overlay it on the NLL plot with the other two calculations in parts a) and b). Shift the  $\chi^2$  also such that the minimum is at zero. Compare the results.
- Bin the dataset into a wider histogram of 2 bins and calculate the binned  $2 \cdot \text{NLL}$  and the  $\chi^2$  function again. How is the agreement now between them? How do they agree with the unbinned calculation in part a)?

**Deadline for submission: Friday, 4. December 2020 14:00**

**Form: Please submit your solutions to `da@physik.uzh.ch`. The solutions should**

**Data Analysis (PHY231) HS20 - Exercise sheet no 6: The maximum likelihood method**

---

be submitted as a single python script with answers to specific questions in the comments.