Karlsruher Institut für Technologie

Due Date: Computerpraktikum am 02.05. und 03.05.2024

Verantwortlich für Vorlesung bzw. Computerpraktikum:
Dr. P. Goldenzweig, Dr. J. Kieseler, Dr. S. Stefkova
Tutoren: Dr. G. De Pietro, Dr. R. Quishpe, M. Mormile, J. Eppelt, A. Brusamolino, Dr. X. Zuo

Computerpraktikum zur Vorlesung Moderne Methoden der Datenanalyse - Blatt 1

Introduction — Getting started

- The ILIAS-web page of this course

  https://ilias.studium.kit.edu/ilias.php?baseClass=ilrepositorygui&ref_id=2071903

  will be used to provide additional information, material and links for this course. It will be gradually updated during the course. You might want to set a bookmark to this page in your browser.

- To work on the exercises, a Jupyter Hub server is provided which can be accessed from any device with a browser via the URL

  https://jupytermachine.etp.kit.edu.

  Use your KIT account credentials to log in. Your KIT account must be registered for access of the Physik-Pool. If this is not the case, please register your account under

  https://comp.physik.kit.edu/Account/.

  On this page you can also find a link for the prolongation of an existing account.

- Once logged on to the Jupyter Hub, you can spawn a server instance to work on. Choose the **Datenanalyse** image and spawn your server. From this server, you will be able to access your home directory of the Physik-Pool in the File Browser on the left side of your Jupyter Lab window.

- The exercises will be provided as Python 3 Jupyter Notebooks (e.g. `Exercise1.ipynb`). These contain the instructions, additional hints and templates you can use to solve the exercises. They will be provided to you by means of the git repository

  https://gitlab.etp.kit.edu/Lehre/dataanalysisexercises_forstudents.git

  which will be updated each week to contain the latest exercise sheet, the notebook template, as well as all additionally required information.

- You need to clone this git repository in Jupyter Lab by pressing the **Git Clone** button in the upper right corner of your File Browser or via the **Git** tab of the menu on the left-hand side. Enter the Clone URI

  https://gitlab.etp.kit.edu/Lehre/dataanalysisexercises_forstudents.git

in the pop-up window and press **Clone**. The directory for the repository will be created in the directory you are currently in and you should now find it under the name `dataanalysisexercises_forstudents` in your file browser.

- To get the newest exercises, you need to update the `dataanalysisexercises_forstudents` repository. To do so, go to the `Git` menu tab while in the `DataAnalysisExercises_ForStudents` directory. Then press **Pull last changes** button on the top right of the Git Interface menu. Alternatively, you can also navigate to `Git -> Pull from Remote` in the top bar.
- It is assumed that you followed the course *Rechnernutzung in der Physik* and are familiar with the basics of Linux, Python and ROOT. Additionally, some hints will be provided in the exercise notebooks.
- Feel free to ask questions and discuss problems and solutions with the tutors and other students.

Exercises 1: Programming and Standard Error Propagation

To complete the exercises, implement the remaining code in the exercise notebook. Generally there are two ways to solve the exercises. First, you can use dedicated Python libraries (numpy, matplotlib,...). Second, you can import and use ROOT in Python. Hints for both approaches are given in the notebook.

- **Exercise 1.1:**

  Write a code snippet (function, class, etc.) that creates N Gaussian distributed random numbers with mean = 0 and sigma = 1 and plots these numbers as a histogram. N should be an argument of the code snippet. To obtain Gaussian distributed random numbers you can use `random.normal()` from numpy or `gRandom->Gaus()` from ROOT.

- **Exercise 1.2:**

  Extend the code from Exercise 1.1 so that the histogram data is written to a file.

- **Exercise 1.3:**

  Load the histogram from the file you created in Exercise 1.2 again and plot it.

- **Exercise 1.4:**

  Fit a Gaussian function to the histogram you created in the previous exercises.

- **Exercise 1.5:**

  Make the plot nicer and save it as vector graphic file (e.g., eps or pdf). Use filled blue boxes with error bars for the histogram and a red line with thickness 3 for the fitted function. Label the axes "x" and "Entries". Display mean, standard deviation, fit probability and fitted parameters with uncertainties in a statistics box, but nothing else.

- **Exercise 1.6:** **obligatory**

  Use your knowledge from the previous exercises to fill a histogram with a function

  $$f(x_1, x_2) = x_1/x_2,$$

  where x_1 and x_2 are two Gaussian distributed random numbers with mean = 2, sigma = 1.5 and mean = 3, sigma = 2.2, respectively. Assuming standard error propagation without correlations:

  $$\sigma_f^2 = \sum_i \left(\frac{\partial f}{\partial x_i}\right)^2 \sigma_i^2 \tag{1}$$

  calculate the propagated uncertainty for this function (using the mean values for x_1 and x_2). How does the result compare with the properties of the created histogram?