142.351, 260014: Statistical Methods of Data Analysis

W. Waltenberger[1] (Lecture)
L. Einfalt[1,2] (Tutorials TU Wien and Universität Wien)
F. Reindl[1,2] (Tutorials TU Wien and Universität Wien)

[1] Institut für Hochenergiephysik der Österreichischen Akademie der Wissenschaften, A-1050 Wien, Nikolsdorfer Gasse 18
[2] Technische Universität Wien, Atominstitut, A-1020 Wien, Stadionallee 2

Winter term 2023/2024

Exercise sheet 2

Hand-in by: Friday, 3. November 2023 - 11:00 Uhr
Tutorials: Friday, 3. November 2023 - 14:15 Uhr

Example 2.1

You observe a Poisson process with the mean rate λ. However, an event is only registered with a certain probability p. How is the number of registered/lost events per time unit distributed?

Example 2.2

Warranty A producer of computer monitors would like to offer a 5-year warranty on their products. However, this is only economical if, at maximum, 20% of the monitors fail during that period.

a) How long needs the average faultless running time τ_a to be if one assumes that the running time is exponentially distributed?

b) Which value does τ_b take, assuming a shortened warranty time of 3 years?

c) Now we assume that we have monitors fulfilling the specifications of part b) (min. 90% running after 3 years). How high is the percentage of monitors still running after 5 years?

Example 2.3

You are observing N identical volcanoes which were created at the same time. Let the time to eruption be exponentially distributed with mean τ.

a) What is the mean (expected value / "Erwartungswert") of the distribution of the time to eruption?

b) How is the waiting time until the first eruption of a volcano distributed?

c) How is the waiting time until the eruption of the last volcano distributed?

Example 2.4

You measure an electric current $I = 2.5A$ and voltage $U = 230V$ in a simple circuit containing a resistor of resistance R. The absolute error of the current measurement is $\sigma_I = 0.15A$, this error is to be considered as a pure statistical uncertainty determined by a series of current measurements.

a) Calculate the absolute and relative error of the resistance R using linear error propagation.

b) Does the relative error change, if one uses a different resistor with $R' = \frac{1}{2}R$?

Example 2.5

Compulsory for TU (142.351), optional for University (260014)!
Let X be a random variable that follows a continuous distribution with density function f on the interval $S \in \mathbb{R}$. We define a new random variable $Y = a + bX$ with $a \in \mathbb{R}$ and $b \in \mathbb{R} \setminus \{0\}$ which is thus distributed on the interval $T = \{y = a + bx \,|\, x \in S\}$.

a) What is the density function $g(y)$ for the random variable Y (Hint: transformation theorem for probability densities)?

b) Apply the formula from a) to $f_e(x) = \frac{1}{2\pi} \exp -\frac{1}{2}x^2$.

c) **Optional:** Check your result by 1. drawing random numbers from $f_e(x)$ and then transforming them linearly and 2. drawing random numbers from the distribution you calculated in b).

Example 2.6 (Prog)

Load the dataset „IncPeriodData.csv“ in a Jupyter notebook with Python. This contains the incubation periods of 1084 covid disease cases.

a) Plot a histogram of the data. What do you notice? What is the maximum length of an incubation period that is plausible in our case? Clean the data set accordingly.

b) Calculate the expected value, median, and standard deviation of the data set.

c) Apply a kernel estimator with Gaussian kernel and appropriate width to the data set and plot the resulting density function.

d) Create a sample of size N from the smoothed density function. Reapply a kernel estimator to this sample and plot this new density function as well. Approximately how large must N be for the new density function to resemble the original one?

e) Assume that the time between onset of symptoms in an affected person and contacting the emergency number 1450 is $\Gamma(k = 4, \theta = 1)$-distributed (in Python: scipy.stats.gamma.pdf(x, a=4)). Plot the probability density of the waiting time between infection and contacting the emergency number.