Verantwortlich für Vorlesung bzw. Computerpraktikum:
Dr. P. Goldenzweig, Dr. R. Wolf, F. Metzner
Tutoren: M. Bauer, P. Ecker, M. Horzela, Dr. S. Stefkova, Dr. N. Trevisani, T. Voigtländer

Computerpraktikum zur Vorlesung Moderne Methoden der Datenanalyse - Blatt 3

Exercise 3: Maximum Likelihood and χ^2 Methods

Fitting parametrized functions to measured data is daily business in research. By this, models can be tested against experimental data. Moreover, parameters of the models and their uncertainties can be determined. The physicist often refers to this process as "*fitting*" — in general it is called "*parameter estimation*".

- **Exercise 3.1: Decay** **obligatory**

  Generate uniformly distributed random numbers. Then apply the transformation method to generate random numbers following an exponential distribution $\exp(-x/\tau)$ for $x > 0$. These values can be interpreted as measurements of decay times t (e.g., of radioactive particles) corresponding to a lifetime τ, which have the following distribution:

$$f(t, \tau) = \frac{1}{\tau} \cdot \exp\left(-\frac{t}{\tau}\right)$$

  a) Show analytically that the maximum likelihood estimator for τ is the mean $\hat{\tau}$ of the sample ($\hat{\tau}$ = mean of all measured decay times t_i).

  b) Generate 1000 samples with τ=1, each with $N = 10$ values of t. Evaluate the mean $\hat{\tau}$ for each sample and create a histogram of the resulting means. Compare the mean of $\hat{\tau}$ with the true value τ=1.

  c) Assume that the probability density function (p.d.f.) has been parametrized in terms of $\lambda = 1/\tau$, which means:

$$f(t, \lambda) = \lambda \cdot \exp(-\lambda \cdot t)$$

  Create a histogram of the estimations $\hat{\lambda}$. Compare the mean value of $\hat{\lambda}$ with the true value λ=1, and determine the bias numerically for $N = 5, 10, 100$. Calculate the bias also for the experiments made in the exercise part **b)** and compare the results of the two approaches **b)** and **c)**.

d) Compare the results of the maximum likelihood method and the χ^2 method: Make three different histograms with 1000 bins from 0 to 10 containing N generated decay times t (try $N = 10, 1000, 100000$). Fit the function $f(t,\tau)$ to each histogram using the χ^2 method and the binned likelihood method. Compare the fitted parameters and the χ^2 values of both methods and discuss the results.

- **Exercise 3.2: MINUIT** **voluntary**

  The goal of this exercise is to make you familiar with the minimizer package `MINUIT` which was developed at CERN in the 70s in `FORTRAN`. This well-tested toolbox provides different minimization algorithms, the most famous one being `MIGRAD`. The package is particularly liked by physicists due to it's sophisticated methods for the parameter uncertainty estimation.

  For the purpose of this exercise, it is suggested to use the Python frontend to `MINUIT`, which is available in the form of the package `iminuit`.

  Take the function $f(t,\tau)$ and the generated data set from the previous Exercise 3.1, and perform an unbinned log likelihood fit for $N = 10, 1000, 100000$ entries.

  Plot a histogram from 0 to 10 with the N entries and the fitted function normalized to the number of entries. Display the value of the negative logarithmic likelihood as a function of the fit parameter τ from 0.5 to 5. How is this plot related to the uncertainty of the fitted parameter?

  The `iminuit` Python package provides predefined cost function classes which also cover the unbinned case we are interested in. However, to learn how to define your own cost function, you should use the `scipy`-like interface which allows you to provide your own cost function in form of the argument `fun`. You can use either of the two approaches and take a look at the output that MINUIT provides after the cost function is minimized.

Appendix: Formulas

χ^2-function:

$$\chi^2(a) = \sum_{i=1}^{m} \frac{(y_i - \lambda(x_i;a))^2}{\sigma_i^2}$$

with m the number of bins or points, (x_i, y_i) the measurement points or, respectively, x_i the position of the i-th bin and y_i the value of the entry, $\lambda(x_i;a)$ the fit function with parameter a, and σ_i the Gaussian uncertainty (standard deviation) of y_i (in case of a binned distribution this corresponds to the Poisson uncertainty).

General unbinned likelihood:

$$L(a) = \prod_{i=1}^{n} f(x_i;a)$$

with n the size of the sample, x_i the measurement values, and $f(x_i; a)$ a probability density function with parameter a.

Binned likelihood for the case of a Poisson distribution:

$$L(a) = \prod_{i=1}^{m} p(n_i; \nu)$$

with m the number of bins, discrete values n_i (number of events in the i-th bin), and $p(n_i; \nu)$ the Poisson function with parameter ν, which is:

$$p(n_i; \nu) = \frac{\nu^{n_i}}{n_i!} \exp(-\nu)$$
