Exam problems

Statistical Analysis of Experimental Data 2022/2023

Please note:

- each problem should be uploaded to Kampus platform as a separate file;
- solutions have to include description and justification of the approach used, as well as discussion of the presented results;
- solutions should be uploaded until Sunday, February 12, 23:55 CET;
- by uploading the solutions to Kampus you declare that they resulted from your own work and that you have not shared nor discussed them with anyone.

Problem 1

We want to measure the rectangular surface of the lecture room with the highest possible precision. The approximate dimensions of the room are: $\ell = 24$ m, $w = 8$ m. We can use electronic lenght meter with $\sigma = 3$ mm precision. It is fully charged, so we can make $N = 100$ lenght measurements. How should we divide this number into width and lenght measurements to obtain highest precision on the measured surface?

Problem 2

Consider production of bottom meson pairs in electron-positron collisions:

$$e^+e^- \to B^0\bar{B}^0.$$

The lifetime of the produced mesons is $\tau = 1.5$ ps. Assuming collision takes place at exactly $t_0 = 0$, both mesons are produced at rest and the two decays are independent:

- calculate the probability density function, $p(t_1)$, for the first decay to be observed, $t_1 = \min(t_B, t_{\bar{B}})$, where t_B and $t_{\bar{B}}$ are the decay times of B^0 and $\bar{B}^0$ meson, respectively;
- calculate the probability density function, $p(t_2)$, for the second decay to be observed, $t_2 = \max(t_B, t_{\bar{B}})$;
- calculate the probability density function, $p(\Delta t)$, for the decay time difference, $\Delta t = t_2 - t_1$;
- confirm analytical result with the Monte Carlo simulation.

Problem 3

Consider measurement of the forward-backward asymmetry in top quark pair-production in e^+e^- colisions:

$$e^+e^- \to t\bar{t}$$

Forward-backward asymmetry is defined based on the polar angle θ of top quark production with respect to the initial electron direction:

$$A_{FB} = \frac{N_F - N_B}{N_F + N_B} = \frac{N(\theta < \frac{\pi}{2}) - N(\theta > \frac{\pi}{2})}{N(\theta < \frac{\pi}{2}) + N(\theta > \frac{\pi}{2})}$$

where the anti-top is expected to go into oposite direction ($\bar{\theta} = \pi - \theta$). However, the sign of the top (and anti-top) quark is correctly recognized only with finite probability p_{tag}. If we recognize top as anti-top and anti-top as top at the same time, the angle θ will not be correctly measured ($\theta \to \pi - \theta$) and the event will be put into wrong category. Also, if only one quark is mis-identified, both quarks in an event will be reconstructed with the same charge and such an event will be rejected.

Data analysis ended in the following reconstruction results:

$$N_F = N(\theta < \frac{\pi}{2}) = 35\,171$$

$$N_B = N(\theta > \frac{\pi}{2}) = 19\,486$$

$$\text{rejected events} \quad N_{rej} = 35\,449$$

Find the values of the forward-backward asymmetry A_{FB} and of the tagging efficiency p_{tag} with corresponding statistical uncertainties.

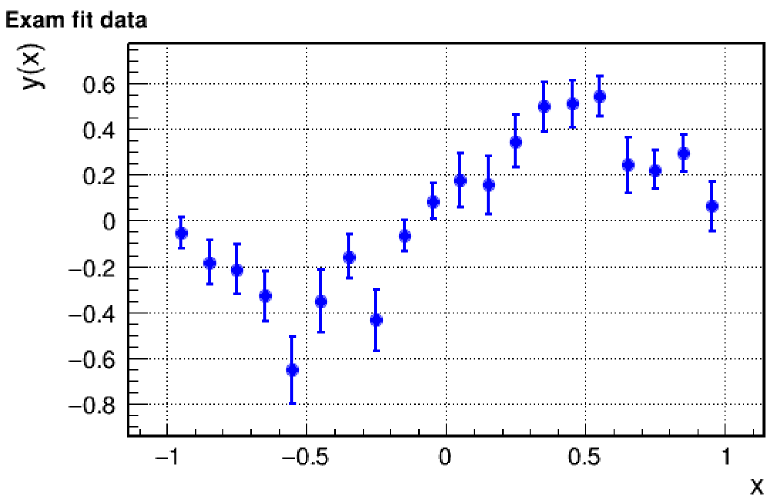
Hint: one of the approaches is to consider χ^2 (or log-likelihood) function with total number of events μ, A_{FB} and p_{tag} as parameters.

Problem 4

Consider the data presented in the plot below (text file with x, y, σ_y values available for download). Fit parameters a, b, A_0 and A_1 of the model:

$$y(x) = A_0 + A_1 \cdot \sin(a\,x + b)$$

Does the result correspond to the global minimum of the likelihood function? Find alternative fit results. Calculate uncertainties and correlation matrix of the parameters.



Problem 5

Event classifier was run on three input samples:

- background Monte Carlo sample with 10 000 events,
- signal Monte Carlo sample with 10 000 events,
- actual data sample with 1 000 events,

and classification results (returned values of decision function) were stored in three files. The goal of the analysis was to find the 95% CL limit on the number of signal events N_{sig} in the actual data sample. Download the three files and assuming the data are consistent with the background only hypothesis (no signal):

- draw ROC curve for the considered classifier,
- using the training samples only, consider the dependence of the expected N_{sig} limit on the cut on the classifier output, find the optimum cut value;
- apply the optimum cut to the actual data to extract the limit on the signal level, compare the observed and expected limit values.