Statistical Methods in Physics (14P058)

Prof. Federico Sánchez (federico.sancheznieto@unige.ch)
Dr. Hepeng Yao (hepeng.yao@unige.ch)
Dr. Knut Zoch (knut.zoch@unige.ch)

Final project III – *W* boson polarisation

Submission due on: 24 January 2022

One of the fundamental properties of the weak interaction in particle physics is its parity-violating behaviour: the charged weak interaction, mediated by W bosons, only couples to "left-handed" (components of) particles. In addition, W bosons play a paramount role in the decays of unstable particles, such as heavy quarks or leptons.

The decay of the top quark, the heaviest known elementary particle, is a 2-body decay: $t \to Wb$. Due to the parity violating behaviour of the t–W coupling, the polarisation of W bosons from top-quark decays is not random, but strongly prefers two out of three possible polarisation states. The polarisation can be measured in leptonic decays of W bosons: $W \to \ell\nu_\ell$. The probability to emit the charged lepton (ℓ) under a certain angle θ follows

$$\frac{1}{\sigma}\frac{\mathrm{d}\sigma}{\mathrm{d}\cos\theta^*} = \frac{3}{8}(1-\cos\theta^*)^2 f_L + \frac{3}{4}(1-\cos^2\theta^*) f_0 + \frac{3}{8}(1+\cos\theta^*)^2 f_R. \tag{1}$$

where $\theta^* \in [0,\pi]$ indicates that the angle is measured in the W boson rest frame with respect to the W boson momentum axis in the top-quark rest frame. f_L, f_0, f_R are the *helicity fractions* that describe the fractions of left-handed, longitudinal and right-handed polarisation states of the W boson, respectively. By construction, $f_L + f_0 + f_R \equiv 1$. At lowest order, the fractions are approximately $f_L = 0.3$, $f_0 = 0.7$, $f_R = 0$. Thus, we can drop the last term in Eq. (2) and substitute $f_L = 1 - f_0$. Then, the probability reduces to:

$$\frac{1}{\sigma}\frac{\mathrm{d}\sigma}{\mathrm{d}\cos\theta^*} = \frac{3}{8}(1-\cos\theta^*)^2\,(1-f_0) + \frac{3}{4}(1-\cos^2\theta^*) f_0\,. \tag{2}$$

Task 1: *Basic statistical properties*

Equation (2) fulfils the properties of a probability distribution function. The first task is to assess basic statistical properties of the distribution with the help of Monte Carlo (MC) processes.

a) Demonstrate that Eq. (2) is a probability density function.

b) Build a MC process to generate this probability density function for a large number of data points, $N \gg 1$. Fill your results in a histogram, plot them and verify their correctness by comparing them with the curve of the analytical form in Eq. (2).

c) Using the approximated helicity fractions as listed above, the expectation value of the normalised differential cross-section is

$$\mathbb{E}\left[\cos\theta^*\right] \approx 0.55\,. \tag{3}$$

Compute the sample mean $\langle\cos\theta^*\rangle$ and compare with the expectation value.

d) What does this expectation value mean for the emission of charged leptons? Which values of θ^* are preferred? Are some even forbidden?

e) Compute the variance, skewness and kurtosis of the distribution and comment. You can use an MC-generated dataset to do so.

Task 2: *Convergence*

If the experiment is repeated N_{exp} times, it can be seen that the sample means $\langle\cos\theta^*\rangle$ of the experiments are distributed according to a Gaussian distribution. In the following sub-tasks, we will look at the convergence behaviour of $\langle\cos\theta^*\rangle$. Each of the N_{exp} experiments still contains $N \gg 1$ data points generated with the MC process.

a) Take $N_{exp} = 1$ and show the law of large numbers based on the data points you generated.

b) Take $N_{exp} \gg 1$ and show the validity of the central limit theorem for $\langle\cos\theta^*\rangle$.

c) Compute the variance of $\langle\cos\theta^*\rangle$ and compare it with the value you expect.

Task 3: χ^2 *distribution*

By filling the numbers generated from the MC process into a histogram, we can study the χ^2 properties of the distribution. For now, we fix $N_{exp} = 1$ and $N \gg 1$, according to the probability density function in Eq. (2). Fill the obtained $\cos\theta^*$ values in a histogram.

a) What is the expected p.d.f. for the number of entries per bin?

b) Show that the χ^2 of the obtained entries per bin follows a χ^2 distribution. Instead of calculating the empirical mean value per bin from the experiments, use the nominal value from the p.d.f. formula per bin as the mean value per bin.

c) How does the χ^2 distribution change with the number of bins?

Task 4: *Parameter estimation*

We will now use three different methods to perform parameter estimation and compare the results from the different methods. Let's again start with one single experiment ($N_{exp} = 1$) with $N \gg 1$ data points according to the probability density function in Eq. (2).

N.B. Just a reminder: we assume $f_R = 0$. Thus, we can express the probability as a function of just f_0 as, by construction, $f_L + f_0 + f_R \equiv 1$. So f_0 is our parameter of interest to estimate.

a) Compute the log likelihood $ln\ L(x_1, x_2, ..., x_N \,|\, f_0)$ at a given f_0. Then, use the maximum likelihood method to estimate f_0 and its variance.

b) Compute the goodness of fit at various f_0 and use the least squares method to estimate f_0 and its variance.

c) The parameter f_0 can also be estimated by computing the sample mean $\langle\cos\theta^*\rangle$ of Eq. (2) with the MC integral. Perform this estimation of f_0 and its corresponding variance.

d) Compare the results from the three questions above and discuss them.

Task 5: *Hypothesis testing*

Generate $N_{exp} = 200$ experiments and combine them into pairs, that is, forming a total of 100 pairs of experiments. Now let's assume we obtained these datasets in actual experiments and we would like to – for each pair – test the hypothesis that the two sets of experimental data originate from the same distribution. Set a threshold for the hypothesis testing and perform Kolmogorov tests for all dataset pairs. Then, comment on your outcome with respect to the threshold you set.

Task 6: *Bonus question: Measurement with systematic errors*

In an actual measurement, the obtained data may come with systematic errors. Let's simulate this effect by generating $N = 10^4$ data points which follow Eq. (2). We assume that there is an experimental error in determining the energy values, which follows a Gaussian smearing with standard deviation $\sigma = 0.1$. Now, if we repeat the three estimation techniques in Task 4, how is the outcome influenced by this fluctuation? Which of the estimation techniques is more fragile and which one is more robust? Perform the simulation and discuss.

N.B. This task is not required for your report, but if you hand in a solution, it can be accredited for bonus points.