# Bayesian statistics - Exam

Claudia Nuñez
*Facultat de Ciències, Universitat Autònoma de Barcelona, Til·lers, 08193 Bellaterra, Barcelona*

**Abstract:** In this work we use Bayesian inference and a Metropolis-Hastings algorithm to estimate two free parameters ($A_{FB}$ and $\beta$) in the angular distribution of $b\bar{b}$ quark pairs, produced in electron-positron beam collisions. To this end, we use an experimental dataset containing $10^4$ values of the cosine of the polar angle with respect to the beam-line. We obtain the values $A_{FB} = 0.102 \pm 0.010$ and $\beta = 1.053 \pm 0.067$, and check that the probability density function for the simulated data follows a Gaussian distribution, for both estimators. Finally, we check that the value obtained for $\beta$ is compatible with the prediction of the Standard Model of Particle Physics, $\beta = 1$.

## I. INTRODUCTION

In this work we use Bayesian inference to estimate the value of the two free parameters in the angular distribution of $b\bar{b}$ quark pairs, produced in electron-positron beam collisions. We use an experimental dataset ($D$), which contains $10^4$ values of the cosine of the polar angle with respect to the beam-line ($x$). The values of $x$ range from $-a$ to $a$, where $a = 0.95$ is the angular acceptance of the detector.

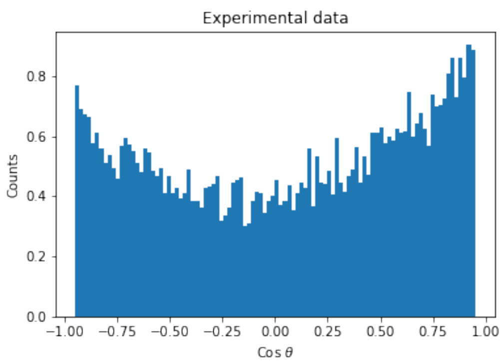


FIG. 1: Histogram showing the values of $x$ in the dataset.

The expected probability density function (PDF) for $x$ is

$$f(x) = \left(1 + \frac{8}{3}A_{FB}x + \beta x^2\right)\frac{1}{2a(1+\beta a^2/3)}, \tag{1}$$

with $A_{FB}$ (forward-backward asymmetry) and $\beta$ the free parameters we want to estimate.

The estimators of both parameters are computed using Bayesian inference plus a Metropolis-Hastings algorithm, with $10^6$ Monte Carlo steps and a *burn-in* period of $10^4$ steps. The values obtained for both estimators (sample mean $\pm$ sample standard deviation) are

$$A_{FB} = 0.102 \pm 0.010 \qquad \beta = 1.053 \pm 0.067, \tag{2}$$

and present a correlation coefficient

$$r = 0.224. \tag{3}$$

Inserting these parameters in the function (1), we obtain the orange curve showed in FIG.2, that fits the data in $D$.

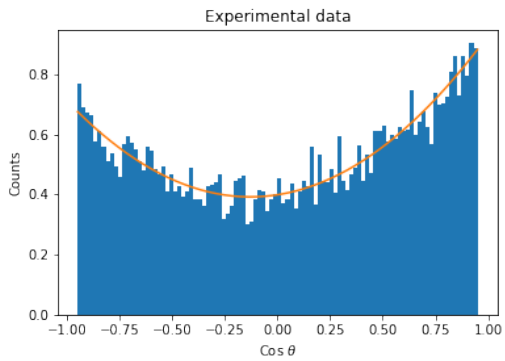


FIG. 2: Histogram of the data and curve for (1) applying the values of the estimators for $A_{FB}$ and $\beta$, obtained with the MCMC method (orange).

Comparing the value obtained for $\beta$ with the prediction given by the Standard Model of Particle Physics (SM), $\beta = 1$, one finds that is compatible with this value.

## II. METHODOLOGY

Bayesian inference is a method where Bayes theorem is used to update an initial (prior) PDF with some new information to obtain a new PDF (posterior). These method can be used for parameter estimation. Given a dataset $D$ and its PDF $f(x;\vec{\theta})$, with some free parameters we want to determine, $\vec{\theta} = (\eta_1, \eta_2, ...)$, we can define a prior PDF for these parameters $\pi(\vec{\theta})$ and build the logarithm of the posterior PDF as

$$\log(p(\vec{\theta}|D)) = \log\left(\frac{(L(D|\vec{\theta}))\pi(\vec{\theta})}{N}\right), \tag{4}$$

where $L(D|\vec{\theta})$ is the likelihood of the data given the parameters,

$$\log(L(D|\vec{\theta})) = \sum_{i=1}^{n} \log(f(x_i; \vec{\theta})), \tag{5}$$

and $N$ a normalization integral.

Since we don't need the whole PDF, only an estimation of the parameters, we sample $\vec{\theta}$ from the logarithm of the posterior PDF and use these samples to compute the statistics of $\vec{\theta}$, that in our case is $\vec{\theta} = (A_{FB}, \beta)$.

One way to perform the sampling is using the Metropolis-Hastings algorithm, a Markov-chain Monte Carlo (MCMC) technique, with which we don't need to compute $N$. Then, choosing the prior PDF to be uniform, we can take the logarithm of the likelihood as the logarithm of the posterior.

The MCMC are numerical methods that use random numbers to explore the phase space with a stochastic process (Markov chain). Our goal is to find the values of $\vec{\theta}$ that maximize the likelihood. That is, the maximum in the posterior PDF. In the Metropolis-Hastings algorithm we proceed as follows:

1. Start at an arbitrary value of the parameters $\vec{\theta}_0$.

2. Using this point, select a trial point $\vec{\theta}^*$, throwing a random number following a proposal distribution. In our case, the distributions selected for both parameters are normal distributions, centered in $\vec{\theta}_0$ with standard deviation $\sigma = 0.01$. We tried different values of $\sigma$ and decided to use these ones. The choice of $\sigma$ is relevant, cause if it is very small you can end up in a local maximum, and don't have enough range when you perform the change to escape its neighbourhood. On the other hand, if the value of $\sigma$ is very large, you will not explore the phase space with enough resolution to find the extreme.

3. Compute the values of the logarithm of the posterior for both points. If the proposal point gives a higher or equal value of the likelihood, we accept the change. If not, we throw a uniform random number between 0 and 1 and accept the change if the ratio between the likelihood of the proposed point over the likelihood of the old one is greater than this number. This last step is also a way to try to avoid getting stuck in a local maximum.

4. Finally, if the change is accepted $\vec{\theta}_0 = \vec{\theta}^*$. If not, it remains unchanged. We repeat the procedure from point 2. We call each run through all the steps a Monte Carlo step.

To obtain good estimators of the parameters, we need the values given by the algorithm to be statistical independent. The Markov chain presents strong correlations with the initial state in the first Monte Carlo steps of the algorithm. We mitigate the effects of this correlations applying a *burn-in* period. That is, removing the first Monte Carlo steps of the algorithm.

## III. RESULTS

We apply the Metropolis-Hastings algorithm to find the parameters $A_{FB}$ and $\beta$ in equation (1), using the dataset plotted in FIG.1 and sampling from the logarithm of the PDF in equation (4), with the considerations explained in the previous section. Additionally, we impose the prior conditions $-3/8 < A_{FB} < 3/8$ and $\beta > 0$. When the values of the parameters doesn't fulfill these conditions, $p(\theta|D) = 0$. That is, $\log(p(\theta|D)) = -\infty$. We run the algorithm for $10^6$ Monte Carlo steps and apply a *burn-in* period of $10^4$ steps.

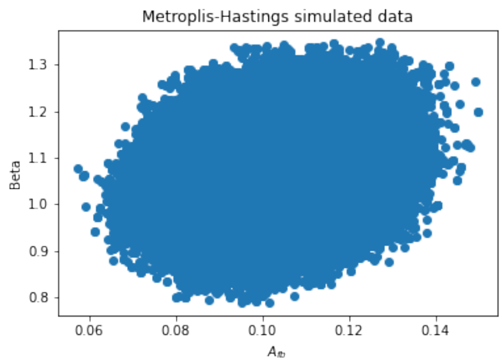


FIG. 3: 2D scattering plot of the Metropolis-Hastings simulated data for $A_{FB}$ and $\beta$.

We obtain the estimators (sample mean $\pm$ sample standard deviation)

$$A_{FB} = 0.102 \pm 0.010 \qquad \beta = 1.053 \pm 0.067, \tag{6}$$

that present a correlation coefficient

$$r = 0.224. \tag{7}$$

Plotting the one-dimensional histograms for both variables, one can see that they both follow a Gaussian distribution.

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/2\sigma^2}, \tag{8}$$

where $\mu$ is the mean and $\sigma$ the standard deviation (FIG.4 and FIG.5).

Using the two one-dimensional histograms for $A_{FB}$ and $\beta$, we compute the values of the median and the mode for both parameters with the two-sided 68% confidence level intervals around them. That is, the intervals where the true value of the parameters (median and mode) will be located the 68% of the times. The values obtained for the estimator $A_{FB}$ are

$$\text{median}(A_{FB}) = 0.102 \quad \text{CL}(68\%) = (0.091, 0.112),$$

$$\text{mode}(A_{FB}) = 0.101 \qquad \text{CL}(68\%) = (0.090, 0.111),$$

and the ones for $\beta$,

$$\text{median}(\beta) = 1.054 \qquad \text{CL}(68\%) = (0.979, 1.118),$$

$$\text{mode}(\beta) = 1.054 \qquad \text{CL}(68\%) = (0.979, 1.118).$$

Finally, comparing the estimated value $\beta = 1.053 \pm 0.067$ with the prediction of the Standard Model of Physics (SM), $\beta = 1$, we see that they are compatible. It is less than $1\sigma$ away of this value. If we also look at the values of the median and the mode, the prediction of the SM is inside the 68% confidence level interval in both cases.

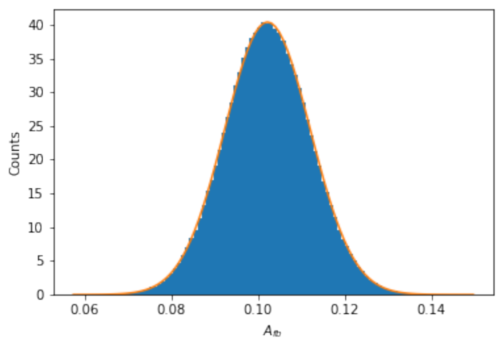


FIG. 4: Histogram of the Metropolis-Hastings simulated data for $A_{FB}$ and corresponding Gaussian distribution (orange).

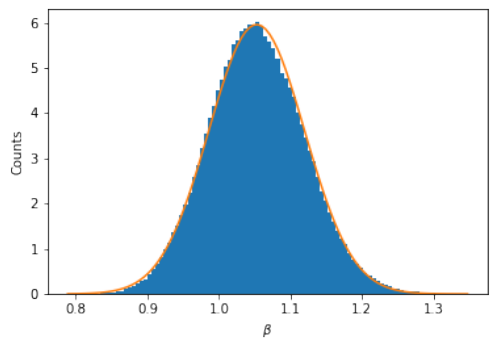


FIG. 5: Histogram of the Metropolis-Hastings simulated data for $\beta$ and corresponding Gaussian distribution (orange).