

Bayesian inference in particle physics

Lorenz Gärtner¹, Toni Sculac, Judith Katzy

February 11, 2025

1 Bayesian Higgs

We got our hands on some fresh LHC data. It seems to be invariant mass measurements of the channel ... in the region $m \in [100, 150]$ GeV.

Excitedly, you tell all your friends. One of them tells you that you should analyse the data to obtain a measurement of the signal strength μ . She also tells you, that for this channel the background is expected to have an exponential mass distribution of

$$p(m|m_b, \lambda_b) = \frac{1}{\lambda_b} \exp\left(-\frac{m - m_b}{\lambda_b}\right),$$

where $m_b = 100$ GeV and $\lambda_b = 20$ GeV. The signal has an expected mass distribution given by

$$p(m|m_b, \lambda_b) = \mathcal{N}(m_s, \sigma_s),$$

where $m_s = 125$ GeV and $\sigma_s = 2$ GeV. You expect 10000 background events and 1000 signal events

Hint: Always work with twice negative logarithmic probabilities throughout the exercise.

1. Generate MC events for signal and background and load the data.
2. Construct the likelihood function

$$p(n|\mu, \mu_{bkg}) = \prod_{\text{bins } b} \text{Poisson}(n_b|\nu_b(\mu, \mu_{bkg}))$$

where n_b are your measured data yields per bin b . The expected yields per bin are

$$\nu_b(\mu, \mu_{bkg}) = \mu n_{sig,b} + \mu_{bkg} n_{bkg,b},$$

where $n_{sig,b}$ and $n_{bkg,b}$ are the expected signal and background yields per bin, respectively. Our model parameters are the normalization factors μ and μ_{bkg} . The goal of this analysis will be to infer μ .

3. Think about possible prior choices here. We are relatively confident in the background modeling, but are not so confident in the signal normalization. Think about possible prior choices for μ and μ_{bkg} .

4. Combine (twice negative log) likelihood and prior into a posterior.
5. What is the posterior mode? What is the impact of your priors here?
6. Advanced: Code up your own implementation of the Metropolis-Hastings algorithm.
Hint: Implement it such that it works with twice negative logarithmic probabilities, for numerical stability.
Alternative: Copy my implementation from here.
7. Generate samples from the posterior.
8. Use the samples to obtain estimates for the 1-and 2-dimensional marginal posteriors of μ and μ_{bkg} .
9. What is the 95% upper credible interval on μ ?
10. What are the 68% and 95% highest density intervals of μ and μ_{bkg} ?