

Exercise - Hypothesis testing

Recommended reading:

<https://allendowney.blogspot.com/2011/05/there-is-only-one-test.html>

The Higgs notebook (Higgs_Simulation.ipynb) simulates the distribution of the invariant mass of pairs of photons ($m_{\gamma\gamma}$) found in events collected by the ATLAS experiment at the LHC, see accompanying paper, ATLAS collaboration (2012).

The model used for the simulation contains:

- A decreasing exponential background with rate parameter, corresponding to pairs of photons produced independently: $\tau=0.0218\text{GeV}^{-1}$
- A Gaussian distribution corresponding to the decay $H \rightarrow \gamma\gamma$ with H being a Higgs boson of a mass of 126.5 GeV , and with a standard deviation given by the experimental resolution of the measurement of $m_{\gamma\gamma}$ (2GeV).

Let us assume that the normalization set by default in the program ($N_{\text{tot}}=80000$) corresponds to the data collected by the ATLAS detector in one year, and that the performances of the collider and the detector are constant in time: the amount of data to analyze is just proportional to the time used to collect the data.

Let our null hypothesis be “the data follows the exponential distribution with a constant $\tau=0.0218\text{GeV}^{-1}$ ”. Then the p-value of our null hypothesis will quantify the statistical significance of the discovery of a new particle (the Higgs boson).

1. **Build a χ^2 estimator** that tests whether data follows the null hypothesis. Obtain its sampling distribution for many one-year experiments.
2. Let us define the **expected significance** as the expectation value of the p-value of the null hypothesis. What is the expected significance after one year of data taking?
3. How many years of data do we need for the expected significance to be at the level of 5σ that is, $< 2.9 \cdot 10^{-7}$?
4. How many years of data do we need in order to have a 95% probability of the p-value being at the level of that is, $< 2.9 \cdot 10^{-7}$?

Repeat the above numerical experiments using the Kolmogorov-Smirnov test instead of the χ^2 test. Are more or less years of data required?

Discuss in detail the interpretation of the P-values and its distribution as obtained above. **Discuss also in detail** the implications of the results for the design of the experiment and the different efficiency of the two estimators.

Note: the above test is a toy model, while the test used in the paper is much more sophisticated.