

# Hypothesis testing

How many samples do we need to verify threshold effects?

November 21, 2024

## 1 Assumptions

We examine the scattering of (almost) massless fermions,  $\bar{f}f$ . There can be three distinct kinds of events happening; No scattering ( $\bar{f}f \rightarrow \bar{f}f$ ), observable scattering ( $\bar{f}f \rightarrow \Phi \rightarrow \bar{f}f$ ), and decaying into the hidden or unknown sector ( $\bar{f}f \rightarrow \Phi \rightarrow \bar{g}g$ ). To simplify the argument, we combine the first and third events as background (B), and the second one is considered as a signal (S). Then the chance of getting signal is proportional to total cross section of the process  $\bar{f}f \rightarrow \Phi \rightarrow \bar{f}f$ . To further simplify the analysis, we will impose some assumptions:

1. A process in interest is dominant in  $\bar{f}f$  scattering. That is, the contribution of other scattering processes are negligible.
2. The detector classifies the background and the signal perfectly, which means that it does not confuse between no scattering events and observable scattering events.
3. There is no systematic error or background noise. The only exception is the uncertainty of momentum measurement, which is dealt with binned counting or Gaussian smearing.

In likelihood-ratio testing, we fix the mass of the hidden sector particle to prevent non-identifiability problem in Wilks' test[1]. However, in practice we need to implement statistical corrections to prevent look-elsewhere effect. The other way around is using Pearson's  $\chi^2$  test. It is convenient if we only need to reject the null hypothesis, but the drawback is that we cannot really tell much about the alternative models.

It is impossible to predict the sample count needed without knowing the specific experimental procedure. In this report, I assumed that the invariant masses of the incoming particles are uniformly distributed along  $M_{\text{peak}} - 2\Gamma < m_{\text{in}} < M_{\text{peak}} + 2\Gamma$ , where  $M_{\text{peak}}$  is the peak of resonance (real-on-shell mass) and  $\Gamma$  is the total decay width of  $\Phi$ .

## 2 Likelihood-Ratio Test

The null hypothesis and the alternative hypothesis are

$$H_0 : P_0(\mu; A, M, \Gamma) = \frac{As}{(\mu^2 - M^2)^2 + \Gamma^2},$$
$$H_1 : P_1(\mu; A, M, \Gamma, C \neq 0) = \frac{As}{|\mu^2 - M^2 + i\Gamma + C\Pi(\mu, m)|^2}.$$

We can see that the alternative hypothesis has one extra degree of freedom in fitting parameters ( $C$ ), since we have fixed  $m$ .

### 2.1 Unbinned Analysis without Momentum Uncertainties

If there are no uncertainties in momentum measurement, then we may perform the unbinned analysis. Suppose we got  $N$  samples  $x_1, \dots, x_N$ . For each sample  $x_i$ ,  $\mu_i = \mu(x_i)$  is its invariant mass and  $S(x_i)$  is its

measurement result, where  $S(x_i) = 1$  if the detector counted it as a signal and  $S(x_i) = 0$  for a background. Then the likelihood functions are given by

$$L_0(\theta) = \prod_{i=1}^N \left[ S(x_i) P_0(\mu_i; \theta) + (1 - S(x_i))(1 - P_0(\mu_i; \theta)) \right],$$

$$L_1(\theta, C) = \prod_{i=1}^N \left[ S(x_i) P_1(\mu_i; \theta, C) + (1 - S(x_i))(1 - P_1(\mu_i; \theta, C)) \right].$$

It is convenient to use log-likelihood functions instead:

$$\ln L_0(\theta) = \sum_{i=1}^N \ln \left[ S(x_i) P_0(\mu_i; \theta) + (1 - S(x_i))(1 - P_0(\mu_i; \theta)) \right],$$

$$\ln L_1(\theta, C) = \sum_{i=1}^N \ln \left[ S(x_i) P_1(\mu_i; \theta, C) + (1 - S(x_i))(1 - P_1(\mu_i; \theta, C)) \right].$$

The best fit is given by maximising likelihoods.

Suppose that the alternative hypothesis is true:

$$P(\mu) = P_1(\mu; \theta_t, C_t), \quad C \neq 0.$$

Then the best fit parameters for  $\theta$  and  $C$  in the alternative hypothesis should be sufficiently close to the true values for big enough  $N$ . Hence, as  $N \rightarrow \infty$ ,

$$\begin{aligned} \max_{\theta, C} \left[ \ln L_1(\theta, C) \right] &\simeq \sum_{i=1}^N \ln \left[ S(x_i) P(\mu_i) + (1 - S(x_i))(1 - P(\mu_i)) \right] \\ &\simeq \sum_{i=1}^N \left[ P(\mu_i) \ln P(\mu_i) + (1 - P(\mu_i)) \ln(1 - P(\mu_i)) \right] \\ &\simeq \frac{N}{4\Gamma_t} \int_{M_t - 2\Gamma_t}^{M_t - 2\Gamma_t} d\mu \left[ P(\mu) \ln P(\mu) + (1 - P(\mu)) \ln(1 - P(\mu)) \right]. \end{aligned}$$

The final quantity can be evaluated once we fix the true parameters. Similarly, for null hypothesis,

$$\ln L_0(\theta) \simeq \frac{N}{4\Gamma_t} \int_{M_t - 2\Gamma_t}^{M_t - 2\Gamma_t} d\mu \left[ P(\mu) \ln P_0(\mu; \theta) + (1 - P(\mu)) \ln(1 - P_0(\mu; \theta)) \right],$$

so we can maximise  $\ln L_0(\theta)$  with numerical computation.

One we have calculated  $\mathcal{L}_0 = \max_{\theta} [\ln L_0(\theta)]$  and  $\mathcal{L}_1 = \max_{\theta, C} [\ln L_1(\theta, C)]$ , the log-likelihood-ratio

$$t := 2 \ln \mathcal{L}_1 - 2 \ln \mathcal{L}_0$$

is said to be following the  $\chi^2(1)$  distribution. However, since the physically meaningful value of the extra parameter  $C$  lies on the nonnegative half plane, our test statistics  $t$  actually follows the  $\frac{1}{2}\chi^2(1)$  distribution[2]. Then, if we want  $5\sigma$  significance level, we should have  $t > 23.66$ . Alternatively, since  $t \propto N$ , we can estimate the minimum sample number  $N$  to get  $5\sigma$  significance to reject the null hypothesis for given  $P(\mu)$ .

## References

- [1] Sara Algeri, Jelle Aalbers, Knut Dundas Morå, and Jan Conrad. Searching for new phenomena with profile likelihood ratio tests. *Nature reviews physics*, 2(5):245–252, 2020.
- [2] Glen Cowan, Kyle Cranmer, Eilam Gross, and Ofer Vitells. Asymptotic formulae for likelihood-based tests of new physics. *The European Physical Journal C*, 71:1–19, 2011.