

4 Exercises/Discussion

4.1 A toy example

Suppose our instrument has collected $N = 1000$ observations which appear to suggest that an excess of events at 125 GeV is present. For simplicity, we assume that the background was completely resolved and the goal is to understand what is the exact location of these signal events. The mass distribution is assumed to be a Gaussian with mean μ and variance σ^2 (both unknown). We want to test the hypotheses

$$H_0 : \mu = 125 \quad \text{versus} \quad H_1 : \mu \neq 125$$

For this scenario, specify the following.

1. The log-likelihood function. (profile likelihood)

$$f(y; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(y-\mu)^2}$$

$$\ell(\mu; \sigma^2, \underline{y}) = \log \prod_{i=1}^n f(y_i; \mu, \sigma^2) = \sum_{i=1}^n \log \left[\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(y_i-\mu)^2} \right]$$

$$= -\frac{n}{2} \log 2\pi \hat{\sigma}_{\mu}^2 - \frac{1}{2\hat{\sigma}_{\mu}^2} \sum_{i=1}^n (y_i - \mu)^2$$

2. A formula for $\Lambda(125)$

Hint: $X \sim N(\mu, \sigma^2)$

$$\hat{\mu} = \bar{X} = \frac{1}{n} \sum_{i=1}^n x_i$$

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \hat{\mu})^2$$

$$\Lambda(125) = -2 \left[\ell(125; \hat{\sigma}_{125}^2, \underline{y}) - \ell(\hat{\mu}; \hat{\sigma}_{\hat{\mu}}^2, \underline{y}) \right]$$

$$= n \log(2\pi \hat{\sigma}_{125}^2) + \frac{1}{\hat{\sigma}_{125}^2} \sum_{i=1}^n (y_i - 125)^2 - n \log(2\pi \hat{\sigma}_{\hat{\mu}}^2)$$

$$- \frac{1}{\hat{\sigma}_{\hat{\mu}}^2} \sum_{i=1}^n (y_i - \hat{\mu})^2$$

$$= n \log \frac{\hat{\sigma}_{125}^2}{\hat{\sigma}_{\hat{\mu}}^2} + \frac{\sum_{i=1}^n (y_i - 125)^2}{\hat{\sigma}_{125}^2} - \frac{\sum_{i=1}^n (y_i - \hat{\mu})^2}{\hat{\sigma}_{\hat{\mu}}^2}$$

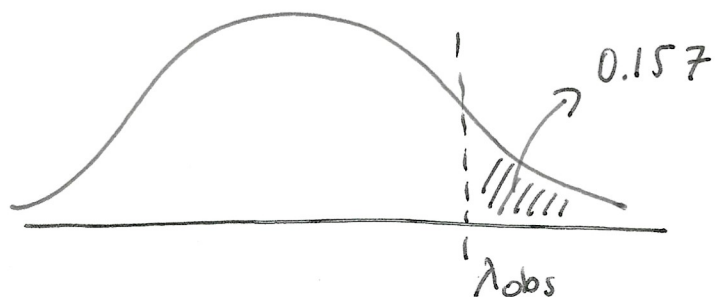
$$= n \log \frac{\hat{\sigma}_{125}^2}{\hat{\sigma}_{\hat{\mu}}^2}$$

3. The asymptotic distribution of Λ , when H_0 is true.

$$\lambda(125) \sim \chi^2_1$$

4. Suppose $\Lambda_{obs} = 2$, specify a formula for the p-value.

$$P(\chi^2_1 > 2) = 0.157$$

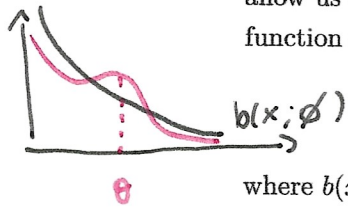


5. The p-value is 0.157, can we claim that $\mu = 125\text{GeV}$ is likely to be the correct location of the signal?

we
 $\alpha = 0.05$
fail to reject H_0
(p-value $> \alpha$)

4.2 Discussion: Identifying situations of non-regularity

Suppose X is the number of events detected by our instrument (e.g., photon emissions over a given range of energy) and let's assume that the resolution of the latter is sufficiently high to allow us to treat the data as a continuous stream of data (no binning). The probability density function of X specifies as follows:



$$f(x; \eta, \phi, \theta) = (1 - \eta)b(x, \phi) + \eta s(x, \theta) \quad (4)$$

$$\eta \in [0, 1]$$

where $b(x, \phi)$ is the distribution of the background and ϕ is a parameter characterizing it. For instance, if b is a power-law, ϕ is its slope. Whereas, $s(x, \theta)$ is the distribution of the signal and θ the parameter characterizing it. For instance, if s is a Gaussian with variance 1, θ is the mean. Finally, η is the intensity of the signal, i.e., the proportion of events that we expect to come from the signal and it is assumed to be non-negative. Finally, you are given a sample of $N = 10000$ observations which you may use to assess the validity of the model in (4).

A few questions for you...

1. How would you specify your hypotheses in order to assess if a signal is present or not?

$$H_0: \eta = 0$$

$$H_1: \eta > 0$$

2. Suppose θ is known to be 125. Which among the necessary conditions needed for the χ^2 -approximation of the LRT fail (see Section 3.4)? Why?

under H_0
condition C1 fails because $\eta = 0$ which is on the boundary of the parameter space

Chernoff (1954) $\lambda_{LRT} \not\approx \chi^2_1$ $\lambda_{LRT} \approx \frac{1}{2} \chi^2_1 + \frac{1}{2} \delta(0)$

$$p\text{-value} = \frac{P(\chi^2_1 > \lambda_{obs})}{2}$$

3. Suppose θ is unknown. Which among the necessary conditions needed for the χ^2 -approximation of the LRT fail (see Section 3.4)? Why?

- Condition C_1 fails (For the same reason we have seen in Q2)
- Condition C_2 fails because whenever $M=0$ for any value of σ we get the same model \Rightarrow we have non-identifiability

LOOK-ELSEWHERE EFFECT Gross-Vitells 2011

4. Suppose you decided to test $H_0 : b(x, \phi)$ versus $H_1 : s(x, \theta)$. Which among the necessary conditions needed for the χ^2 -approximation of the LRT fail?

C_3 fails \rightarrow we are testing non-nested models

You can find a self-contained review on solutions for the problems above in

Algeri et al., 2020. *Searching for new phenomena with profile likelihood ratio tests*. Nature Reviews Physics. <https://www.nature.com/articles/s42254-020-0169-5>