

## SOLUTION FOR HOMEWORK ASSIGNMENT NO. 10

Nils Hoyer, Maurice Morgenthaler

### Exercise 10.1

In this exercise we are asked to determine the equally tailed intervals for the 68.0 %, 95.0 % and 99.7 % confidence levels for the position  $y = 150 \mu\text{m}$  with a Gaussian uncertainty of  $\sigma = 10 \mu\text{m}$  and a sum of two Gaussian uncertainties where  $\sigma_1 = 10 \mu\text{m}$  with a total contribution of 90 % and  $\sigma_2 = 200 \mu\text{m}$  and a total contribution of 90 %.

To do this we will calculate

$$\alpha = \int_{x_{\alpha,1}}^{x_{\alpha,2}} dx \cdot f_{\mu}(x) \quad (1)$$

where  $\alpha$  is the confidence level and  $x_{\alpha,1/2}$  the range in  $x$  corresponding to this level.

In this exercise we use two Gaussian distributions which are symmetric with respect to their mean  $\mu$ . This is why we calculate

$$\frac{\alpha}{2} = \int_{\mu}^{x_{\alpha}} dx \cdot f_{\mu}(x). \quad (2)$$

a) We are given a single Gaussian with  $\sigma = 10 \mu\text{m}$  and  $\mu = 150 \mu\text{m}$ .

$$\begin{aligned} \alpha &= 2 \cdot \int_{\mu}^{x_{\alpha}} dx \cdot \frac{1}{\sqrt{2\pi}\sigma} \exp \left\{ -\frac{(x-\mu)^2}{2\sigma^2} \right\} \\ &= 2 \frac{1}{\sqrt{2\pi}\sigma} \cdot \sqrt{\frac{\pi}{2}} \sigma \cdot \text{Erf} \left( \frac{x_{\alpha} - \mu}{\sqrt{2}\sigma} \right) \\ &= \text{Erf} \left( \frac{x_{\alpha} - \mu}{\sqrt{2}\sigma} \right) \\ &\rightarrow x_{\alpha} = \mu + \sqrt{2}\sigma \cdot \text{Erf}^{-1}(\alpha) \end{aligned} \quad (3)$$

Here  $\text{Erf}^{-1}$  denotes the inverse error function.

For  $\alpha = 0.680$ :  $x_{\alpha} \in [140.055, 159.945]$ .

For  $\alpha = 0.950$ :  $x_{\alpha} \in [130.400, 169.599]$ .

For  $\alpha = 0.997$ :  $x_{\alpha} \in [120.322, 179.677]$ .

b) As we are given a sum of two Gaussians we can reuse the calculation from a). The normed distribution is

$$f_{\mu}(x) = \frac{1}{\sqrt{2\pi} [a_1\sigma_1 + a_2\sigma_2]} \cdot \left( a_1 \cdot \exp \left\{ -\frac{(x-\mu)^2}{2\sigma_1^2} \right\} + a_2 \cdot \exp \left\{ -\frac{(x-\mu)^2}{2\sigma_2^2} \right\} \right) \quad (4)$$

where  $a_{1/2}$  indicates the strength of the contribution of that function. Following the calculation given in a) we eventually get to

$$\alpha = a_1 \cdot \text{Erf} \left( \frac{x_{\alpha} - \mu}{\sqrt{2}\sigma_1} \right) + a_2 \cdot \text{Erf} \left( \frac{x_{\alpha} - \mu}{\sqrt{2}\sigma_2} \right). \quad (5)$$

For  $\alpha = 0.680$ :  $x_\alpha \in [138.485, 161.515]$ .  
For  $\alpha = 0.950$ :  $x_\alpha \in [15.103, 284.897]$ .  
For  $\alpha = 0.997$ :  $x_\alpha \in [-284.018^1, 584.018]$ .

## Exercise 10.2

We were given a note on how the XENON collaboration uses a likelihood base method to set a limit on the Dark Matter - Standard Model Nuclei interaction. The following questions had to be answered.

- i. Explain in a few words how the Profile Likelihood Ratio statistical approach works. The test statistic  $q_\sigma$  is constructed via a profile likelihood  $\lambda(\sigma)$ , as written below. By definition  $q_\sigma \geq 0$ . It equals zero if  $\hat{\sigma} = \sigma$  indicating a compatibility with a signal hypothesis.
- ii. Why is this method advantageous compared to a “hard cut in the space parameter of interest” method? How much better is the limit with the likelihood method?  
Methods which use the “hard cut” firstly have the drawback that they are only able to give an upper limit without being able to claim a detection directly. Second, they don’t take systematic uncertainties into account. Both disadvantages are not present in the presented method. ....How much better is the limit, tho?
- iii. Why is it important to model correctly the background and how is this done in XENON100? Modelling the background precisely is of great importance as the cross-section of WIMPs is set to be lower than  $10^{-43}$  in specific energy ranges given in figure 6 of the paper. As a result of using a wrongfully modelled background one risks of claiming the detection of WIMPs despite it only being recoils of  $\beta$ -radiation.
- iv. Which ones are the nuisance parameters and how are they included in the full likelihood? There are five nuisance parameters which were introduced in the model.
  - expected numbers of background events  $N_b$
  - the probabilities  $\epsilon_s = \{\epsilon_s^j\}$  and  $\epsilon_b = \{\epsilon_b^j\}$
  - relative scintillation efficiency  $\mathcal{L}_{eff}$
  - escape velocity  $v_{esc}$

The likelihood  $\mathcal{L}$  has the form

$$\mathcal{L} = \mathcal{L}_1\{\sigma, N_b, \epsilon_s, \epsilon_b, \mathcal{L}_{eff}, v_{esc}; m_\chi\} \times \mathcal{L}_2\{\epsilon_s\} \times \mathcal{L}_3\{\epsilon_b\} \times \mathcal{L}_4\{\mathcal{L}_{eff}\} \times \mathcal{L}_5\{v_{esc}\} \quad (6)$$

Therefore the nuisance parameters are included via the main likelihood  $\mathcal{L}_1$  as well as the subsidiary likelihoods  $\mathcal{L}_2$  to  $\mathcal{L}_5$ .

- v. Which are the test statistics for the exclusion and the discovery case?  
In both cases the profile likelihood  $\lambda(\sigma)$  is used.

$$\lambda(\sigma) = \frac{\max_{\sigma \text{ fixed}} \mathcal{L}\{\sigma; \mathcal{L}_{eff}, v_{esc}, N_b, \epsilon_s, \epsilon_b\}}{\max \mathcal{L}\{\sigma; \mathcal{L}_{eff}, v_{esc}, N_b, \epsilon_s, \epsilon_b\}} \quad (7)$$

For the exclusion case the test statistic  $q_\sigma$  is defined as

$$q_\sigma = \begin{cases} -2 \ln \lambda(\sigma) & \hat{\sigma} < \sigma \\ 0 & \hat{\sigma} > \sigma \end{cases} \quad (8)$$

As  $0 \leq \lambda(\sigma) \leq 1$  the test statistics  $q_\sigma$  will be  $q_\sigma \geq 0$  in which a smaller value indicates a better compatibility between data and signal hypothesis.

---

<sup>1</sup>In this case a negative value is fine as we are talking about positions of particles.

For the discovery case the test statistic  $q_0$  is defined as

$$q_0 = \begin{cases} -2 \ln \lambda(\sigma = 0) & \hat{\sigma} > \sigma \\ 0 & \hat{\sigma} < \sigma \end{cases} \quad (9)$$

In this method one tries to reject the background-only hypothesis.

- vi. What is the difference between the exclusion sensitivity and the profiled limit? What is the observed limit in this analysis?

The exclusion sensitivity is the expected limit for given conditions concerning the experiment. It can be estimated before using actual data as the median of the test statistic distribution  $f(q_\sigma|H_0)$  with background-only data simulated according to a Poisson distribution. The profiled limit on the other hand is calculated with the profile likelihood ratio method. The limit of this analysis is  $\sigma^{up} < 2.4 \times 10^{-44} \text{cm}^2$  for WIMP's with a mass of about  $m_\chi = 50 \text{GeV}/c^2$