INSTITUT FÜR THEORETISCHE PHYSIK HEIDELBERG
PROF. DR. LUCA AMENDOLA
COMPUTATIONAL STATISTICS – SUMMER SEMESTER 2022 –

Tutors:

- Group 1: Ziyang Zheng
- Group 2: Nico Lorenz
- Group 3: Ziad Sakr
- Group 4: Adrian Schirra

Please, send (in pdf) or hand in the solution to this exercise sheet to your tutor, following their instructions, and carefully respecting the delivery date shown below. All exercise sheets will be graded. You can solve them individually or in pairs (with another student of the same tutorial group). In the latter case deliver please only one document with the solutions. Write the name of the file in the following format: X.Name1.Surname1 or X.Name1.Surname1-Name2.Surname2.pdf, with X being the number of the corresponding exercise sheet.

---

# Exercise sheet 11

RECEIVED: July.04 - BONUS SHEET (not graded)

## 11.1 Fisher matrix

1) Explain why the marginalized likelihood eq. 3.16 of the lecture notes,

$$L = A \exp\left[-\frac{1}{2}\left(S_2 - \frac{S_1^2}{S_0}\right)\right] . \tag{1}$$

where $S_0 = \sum(1/\sigma_i^2)$, $S_1 = \sum y_i/\sigma_i^2$, $S_2 = \sum y_i^2/\sigma_i^2$, and $y_i = m_i - \mu_i$, is Gaussian in the data $m_i$;

2) Use $y_i$ as data variables, and derive the mean and the covariance matrix of $L$; (*hint*: form the combination $\mathbf{y}^t\mathbf{C}^{-1}\mathbf{y}$ and find the elements of $\mathbf{C}^{-1}$; no need to compute explicitly the matrix $\mathbf{C}$).

3) Assume there are $n$ data and that all the variances $\sigma_i^2$ are equal to $\alpha^2$. What happens in the limit of large $n$?

4) Normalize the likelihood under the same assumptions of point 3

5) If the variances are constant, $\sigma_i^2 = \alpha^2$, what is the MLE for $\alpha$? And what is the MLE under the same assumptions of point 3)?

6) What is the Fisher "matrix" for $\alpha$, assuming all the other parameters are known?

## 11.2 $p$-value

Ten temperatures are measured (in Kelvin), each with a precision of 0.2 K: 10.2, 10.4, 9.8, 10.5, 9.9, 9.8, 10.3, 10.1, 10.3, 9.9.

1. It is suggested that they are all measurements of the same thing, the dispersion being due only to the measurements errors. Is this true? To answer this, define an appropriate hypothesis and statistics, and perform an appropriate hypothesis test to calculate the $p$-value.

2. How does your analysis and result change if it is suggested that they are instead all measurements of the same true value of 10.1 K?

## 11.3 Another $p$-value test

A company measures the length of the rods they produce at two different factories, which are supposed to produce identical rods. In factory A, they find lengths $(5.1, 4.95, 5.15, 5.04, 4.89, 5.05, 5.0, 4.9, 4.96, 5.06, 5.04, 5.12)$ in meters; in factory B, they find $(4.9, 4.96, 4.89, 5.02, 5.03, 5.07, 4.95, 4.98, 5.05, 5.08)$. Assuming a threshold of $\alpha = 0.01$, is the hypothesis that the two factories produce identical rods to be accepted or rejected? (*Hint*: use the $t$-Student test)