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 7

RECEIVED: June.07 - DELIVERY: June.13

Exercise 7.1: Exponential distribution (6pt)

The lifetime t of a radioactive nucleus is distributed as the exponential $\exp(-t/\lambda)$ with $\lambda > 0$ (to be normalized). In an experiment, we measure the lifetime of many such nuclei, obtaining n values $t_1, t_2, ...t_n$. The measures are independent of each others.

a) Find mean and variance of t (1pt)
b) Find the maximum likelihood estimator $\hat{\lambda}$ of λ. (2pt)
c) Is $\hat{\lambda}$ an unbiased estimator? (1pt)
d) Find the variance of $\hat{\lambda}$. (2pt)

Exercise 7.2: Shannon entropy (7pt)

The Shannon entropy of a PDF is a functional defined as

$$S = -\int p(x) \ln p(x) \mathrm{d}x. \tag{1}$$

This quantity characterizes the average surprise from the outcome of a random variable x. The maximum value of S signifies that we have the least information on the outcome of x. This happens when the probability is evenly distributed so what comes next is not known. On the contrary, if the result is certain, i.e. $p = 0$ or $p = 1$, we see that $S = 0$. That is, if a value has strong probability over the others, we expect it next. The optimal combination, i.e. the one that gives the highest value for S, would then be when all $p(x)$ are equal.

a) Enforce the normalization of $p(x)$ with a Lagrange-multiplier λ and compute $\delta S = 0$ (i.e., vary S with respect to $p(x)$) to find the distribution that maximizes S. (A Lagrange multiplier is a term added to an action functional as $S + \lambda \cdot C$, where C is a constraint that we want to put to zero. Variation with respect to λ gives $C = 0$ thereby enforcing the constraint. In the present case, for instance, $C = \int dx p(x) - 1$.) (2pt)

b) Enforce again the normalization with λ as in a) and now keep also the variance σ^2 fixed with another Lagrange-multiplier μ. Compute $\delta S = 0$ to find the distribution that maximizes S. (2pt)

c) Show that the Shannon entropy is additive, i.e., for statistically independent x, y

$$S_{x,y} = S_x + S_y. \tag{2}$$

(1pt)

d) The relative Shannon entropy or Kullback-Leibler divergence is defined as

$$\Delta S = \int p(x) \ln \frac{p(x)}{q(x)} \mathrm{d}x. \tag{3}$$

Compute ΔS for the case that p and q are Gaussian distributions with different mean and variance. (2pt)