PD Dr. Torsten Enßlin
Andreas Popp, Philipp Haim,
Vishal Johnson, Fabian Sigler, Hanieh Zandinejad

Exercise sheet 5

Exercise 5 - 1

Assume that for a PDF $P(x)$ only a set of constraints —typically moment constraints— are known, like

$$\begin{aligned} Q\left(x\right) = P\left(x \mid I\right) &= \text{const}, \\ P(x) &> 0, \\ \int P(x)\,\mathrm{d}x &= 1, \\ \int P(x) f_i(x)\,\mathrm{d}x &= \alpha_i. \end{aligned}$$

From the principle of maximum entropy one derives that

$$P(x) = \exp\left(\lambda_0 + \sum_{i=1}^{n} \lambda_i f_i(x)\right)$$

with λ_i such that the constraints are satisfied.

Assume a PDF $P(x|\alpha,\mu)$ with $x \in \mathbb{R}$, $\alpha > 0$, $\mu \in \mathbb{R}$ and $\langle |x-\mu| \rangle_{(x|\alpha,\mu)} = \alpha$. Calculate the PDF of maximum entropy. (3 points)

Exercise 5 - 2

A strictly positive quantity x is bound from above, say by $x \leq 1$. You learn that its natural logarithm is typically the negative number $-l$.

a) How much information did you gain about x (1 point)?

b) What is the least informative value for l you could have gotten (1 point)?

Exercise 5 - 3

You know that a quantity x is in $[1, M]$, where $M \gg b > 1$ is a large positive number and $b > 1$ is a certain base in which digits of numbers are to be expressed. The logarithm $z = \log_b x$ of x in the chosen base b gives you an idea of the order of magnitude of x in this base.

a) Without any additional information, what is the least informative PDF for x? What distribution does this imply for z? (1 point)

b) Now instead assume that z is uniformly distributed. What conclusion can you draw on the least informative distribution of x? (1 point)

c) Assume that the base $b = 10$, use these insights to explain **Benford's law** B, under which a set of positive numbers exhibits the leading digit $d \in \{1, \dots 9\}$ with a probability $P(d|B) = \log_{10}\left(1 + \frac{1}{d}\right)$. Such distributions are observed in many human made, measured, and even mathematical sets of numbers, when these numbers are distributed across many different length scales. Benford's law is used to identify potential tax frauds (3 points).

hint: Find the probability of a generic interval I_k such that $10^k d \leq x < 10^k(d+1)$ for a certain $k \in \mathbb{N}$. Then sum over different ks for fixed d. Finally, consider the limit for large $M \to \infty$.

PD Dr. Torsten Enßlin
Andreas Popp, Philipp Haim,
Vishal Johnson, Fabian Sigler, Hanieh Zandinejad

Exercise 5 - 4

Your knowledge I about a quantity $\mathbf{x} = (x_1, \dots x_n)^\mathrm{T} \in \mathbb{R}^n$ is characterized by being separable, i.e. $\mathcal{P}(\mathbf{x}|I) = \prod_{i=1}^n \mathcal{P}(x_i|I)$ and by being isotropic, i.e. $\mathcal{P}(\mathbf{x}|I) = \mathcal{P}(\mathbf{O}\,\mathbf{x}|I)$ for all orthonormal transformations with $\mathbf{O}^\dagger = \mathbf{O}^{-1}$.

Using the definition of the information Hamiltonian:

$$\mathcal{H}(\mathbf{x}|I) = \sum_{i=1}^n \mathcal{H}(x_i|I) \equiv \sum_{i=1}^n h_i(x_i) \ , \tag{1}$$

we may write

$$\begin{aligned} \prod_{i=1}^n \mathcal{P}(x_i|I) &= \prod_{i=1}^n \mathcal{P}((\mathbf{O}\,\mathbf{x})_i|I) && (2) \\ \sum_{i=1}^n \mathcal{H}(x_i|I) &= \sum_{i=1}^n \mathcal{H}((\mathbf{O}\,\mathbf{x})_i|I) && (3) \\ \sum_{i=1}^n h_i(x_i) &= \sum_{i=1}^n h_i((\mathbf{O}\,\mathbf{x})_i) \ . && (4) \end{aligned}$$

a) Let $\mathbf{x} = r\, e^{(i)}$ be parallel to the i-th unit vector of $\mathbb{R}^n$, $e^{(i)}_j = \delta_{ij}$, and $\mathbf{O}$ such that $\mathbf{O}e^{(i)} = e^{(j)}$ with $i \neq j$. Show that $h_i(r) = h_j(r) + \mathrm{const(i,j)}$. (3 points)

b) Use the result of a) to show that $\mathcal{H}(\mathbf{x}|I) = \sum_i h(x_i) + const.$ with $h(x_i) \equiv h_0(x_i)$. (1 point)

c) Given a general $\mathbf{x} \in \mathbb{R}^n$ with length $r = |\mathbf{x}| = \sqrt{\sum_i x_i^2}$, we can choose $\mathbf{O}$ such that $\mathbf{O}\mathbf{x} = (r, 0, \dots, 0)$. Show that in this case

$$\frac{h'(x_j)}{x_j} = \frac{h'(r)}{r} \quad \forall j \ , \tag{5}$$

where h' denotes the derivative of h. (3 points)

d) Given the result of c), derive the general functional form of h. (2 points)

e) Finally, derive an expression for $\mathcal{P}(\mathbf{x}|I)$. (2 points)

This exercise sheet will be discussed during the exercises.
Group 01, Wednesday 18:00 - 20:00, Theresienstr. 37, A 449,
Group 02, Thursday, 10:00 - 12:00, Theresienstr. 37, A 249,
Group 03, Thursday, 16:00 - 18:00, Theresienstr. 37, A 449

https://wwwmpa.mpa-garching.mpg.de/ ensslin/lectures/lectures.html