

Exercise 6: Maximum Likelihood and MAP Estimation
(to be returned on Dec 12th, 8:00)

Prof. Dr. Moritz Diehl, Katrin Baumgärtner, Jakob Harzer, Yizhen Wang, Rashmi Dabir

In this exercise you deepen your knowledge of Maximum Likelihood (ML), Bayesian, and Maximum A Posteriori (MAP) estimation.

1. **Maximum a posteriori (MAP) estimation.** Given is a dataset $Y = \{y_1, \dots, y_N\}$ of independently drawn samples from the same Gaussian distribution $\mathcal{N}(\mu, \sigma^2)$, with mean $\mu \in \mathbb{R}$ and standard deviation $\sigma > 0$. Using the maximum likelihood approach one can show that the maximum likelihood (ML) estimator for the mean, $\hat{\mu}_{\text{ML}}$, equals the sample mean, i.e. $\hat{\mu}_{\text{ML}} = \frac{1}{N} \sum_{n=1}^N y_n$. Now consider the likelihood function:

$$p(Y|\theta) = p(y_1, \dots, y_N|\theta) = \prod_{n=1}^N \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(y_n - \theta)^2}{2\sigma^2}\right)$$

with $\theta = \mu$ and the prior distribution $p(\theta) = \frac{1}{\sqrt{2\pi\sigma_0^2}} \exp\left(-\frac{(\theta - \mu_0)^2}{2\sigma_0^2}\right)$ with $\mu_0 \in \mathbb{R}$, $\sigma_0 \in \mathbb{R}_+$.

PAPER:

- State the optimization that needs to be solved to compute the MAP estimate. (1 point)
 - Now solve the optimization problem to compute the estimate $\hat{\theta}_{\text{MAP}}$. How is the ML estimate $\hat{\mu}_{\text{ML}}$ related to the MAP estimate $\hat{\mu}_{\text{MAP}} = \hat{\theta}_{\text{MAP}}$? (2 points)
 - For a fixed σ_0 , what happens to the relation when $N \rightarrow \infty$? (0.5 point)
 - For a fixed N , what happens to the relation when $\sigma_0 \rightarrow \infty$? (0.5 point)
2. **Maximum likelihood estimation on wind distribution.** Imagine the municipal energy supplier of Freiburg is planning to set up a wind farm on Feldberg. Hence, they want to check whether a wind farm at Feldberg would be profitable or not. In wind energy, one is usually interested in the average power that can be produced at a specific location with a specific wind turbine. A very important quantity for estimating the average power of a wind turbine is the wind speed that varies according to the weather condition and location.

Due to the high complexity, a deterministic wind model would be very difficult to obtain. However, it has been observed that the magnitude of the wind speed throughout a year follows a distribution given by the following conditional probability density function:

$$p(v|\lambda, k) = \begin{cases} \frac{k}{\lambda} \left(\frac{v}{\lambda}\right)^{k-1} e^{-(v/\lambda)^k} & v \geq 0 \\ 0 & v \leq 0. \end{cases} \quad (1)$$

This distribution is known as the Weibull distribution, where $\lambda > 0$ and $k > 0$ are the parameters of the distribution and v is the magnitude of the wind speed at the turbine location. The parameters λ and k depend on the location. The task of this exercise will be to estimate the parameters λ and k using wind speed data taken at Feldberg.

- PAPER: Formulate the negative log-likelihood function $L(v, \lambda, k)$ given N measurements of wind speed v_1, \dots, v_N . Simplify this function as much as possible. (0.5 point)

- (b) PAPER: State the minimization problem to estimate λ and k and simplify again the objective function $L(v, \lambda, k)$. (0.5 point)
Hint: constant terms in the objective function do not alter a minimization problem's solution.
- (c) PAPER: Take a look at the histogram of the probability density function of the provided wind speed data of the German Weather Service taken on Feldberg (the code for plotting the histogram is already provided in the solution template). What do you observe? (1 point)
- (d) CODE: Solve the minimization problem obtained in Task (b) to estimate λ and k .
Hint: To do that you should use function minimize from package scipy.optimize
 Keep in mind that the minimization has the objective function that you derived in Task (b) but also a constraints on the parameters λ and k . Plot the fitted Weibull distribution into your histogram (code for plotting is provided in the solution template). (2 points)

The main motivation on estimating these parameters is the study of the average power that a certain turbine will produce in a certain location. In order to do that, the expected value $\mathbb{E}\{P_{\text{Power}}(v)\}$ of the power profile $P_{\text{Power}}(v)$ of the turbine is calculated using the wind distribution (eq. (1)). As an illustration, a typical power profile of a wind turbine is given in Figure 1.

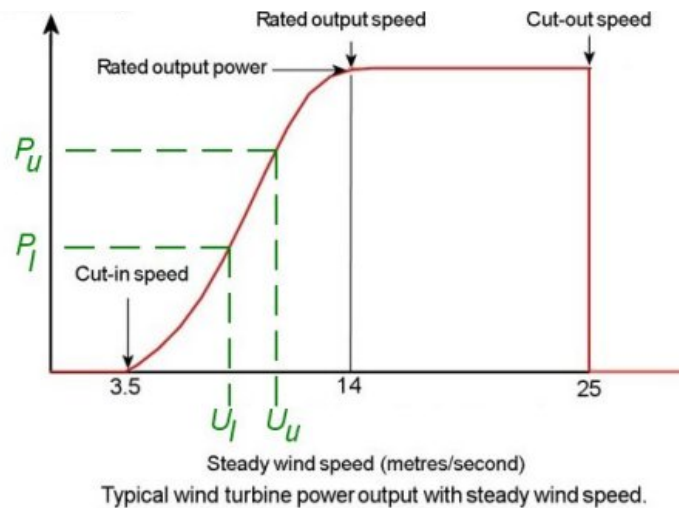


Figure 1: Wind turbine power profile¹

- (e) CODE: Using the data of the power curve of a specific wind turbine given in table (1), compute the expected value of the turbine power on the studied location. Use the trapezoidal rule for integration to compute the expected value of the power. Fill out the provided solution templates.
Hint: First compute the power distribution. (2 points)

Wind speed (m/s)	0	1	2	3	4	5	6	7	8
Output power (kW)	0	0	3	25	82	174	321	532	815
Wind speed (m/s)	9	10	11	12	13	14	15	16	17
Output power (kW)	1180	1580	1900	2200	2480	2700	2850	2950	3020
Wind speed (m/s)	18	19	20	21	22	23	24	25	> 25
Output power (kW)	3020	3020	3020	3020	3020	3020	3020	3020	0

Table 1: Power curve data

This sheet gives in total 10 points

¹Source: http://www.wind-power-program.com/turbine_characteristics.htm