Data Science: Supervision 1 (Additional Exercises)

The following questions are intended for further practice.

Maximum Likelihood Estimate (MLE)

Exercise 1 : A very inexperienced archer shoots an arrow n times at a disc of (unknown) radius θ. The disc is hit every time, but at completely random places. Let $r_1, \ldots, r_n$ be the distances of the various hits to the center of the disck.

1. Show that given θ the pdf is $f_\theta(r) = \frac{2r}{\theta^2}$.
2. Determine the Maximum Likelihood estimate for θ.

Exercise 2 : Suppose that $x_1, \ldots, x_n$ is a dataset, which is a realisation of a random sample from a Rayleigh distribution, which is the continuous distribution with pdf $f_\theta(x) = \frac{x}{\theta^2} \exp\{-\frac{1}{2}\frac{x^2}{\theta^2}\}$ for $x \geq 0$. Determine the maximum likelihood estimator for θ.

Exercise 3 : Suppose that $x_1, \ldots, x_n$ is a dataset, which is a realisation of a random sample from a distribution with pdf

$$f_\theta(x) = \frac{\theta}{(x+1)^{\theta+1}} \text{ for } x > 0$$

Determine the MLE for θ.

Exercise 4 : Suppose that $x_1, \ldots, x_n$ is a dataset, which is a realisation of a random sample from a distribution with pdf

$$f_\theta(x) = \begin{cases} e^{\theta - x} & \text{for } \ x > \theta \\ 0 & \text{for } \ x \leq \theta \end{cases}$$

1. Determine the MLE for θ.
2. Is there anything weird with this distribution?

Exercise 5 : (+) Suppose that $x_1, \ldots, x_n$ is a dataset, which is a realisation of a random sample from a distribution with pdf

$$f_\theta(x) = \frac{1}{2} e^{-|x-\theta|} \text{ for } -\infty < x < \infty$$

Determine the maximum likelihood estimator for θ.

Exercise 6 : Suppose that $x_1, \ldots, x_n$ is a dataset, which is a realisation of a random sample from a distribution with pdf

$$f_{\mu,\lambda}(x) = \left(\frac{\lambda}{2\pi x^3}\right)^{1/2} \exp\left\{-\lambda(x-\mu)^2/(2\mu^2 x)\right\} \text{ for } x > 0$$

Determine the maximum likelihood estimator for μ and λ.

Exercise 7 : Suppose that $x_1, \ldots, x_n$ is a dataset, which is a realisation of a random sample from a binomial distribution with parameters (k, p), where p is known and k is unknown. (For example, this could correspond to flipping a coin with a known probability, but not knowing how many times the coin was flipped).

1. Write an expression $L(k|\mathbf{x}, p)$ for the likelihood.
2. Show that $L(k|\mathbf{x}, p) = 0$ for $k < \max_i x_i$.
3. Explain why if the MLE is k, then

$$\frac{L(k, |\mathbf{x}, p)}{L(k-1, |\mathbf{x}, p)} \geq 1 \text{ and } \frac{L(k+1, |\mathbf{x}, p)}{L(k, |\mathbf{x}, p)} < 1$$

4. (+++) Show that there is a unique k that satisfies this. [Do not spend too much time on this]

Exercise 8 : (+) The unigram model in natural language processing models the probability of a sentence as s as $\mathbb{P}(s) = p_{s_1} \cdot p_{s_2} \cdot \ldots \cdot p_{s_n}$ where $s_1, \ldots, s_n$ are the n words of the sentence. Given M sentences $s^1, \ldots s^M$, show that the MLE for the parameters p_w are $\frac{c_w}{W}$, where c_w is the number of times w occurs in all sentences and W is the total number of words in all sentences.

Linear independence

Exercise 9 : Are the following vectors independent?

1. $v_1 = (1, 2)$ and $v_2 = (-5, 3)$
2. $v_1 = (1, 2)$ and $v_2 = (-4, -8)$
3. $v_1 = (2, -1, 1)$, $v_2 = (3, -4, 2)$ and $v_3 = (5, -10, -8)$

Exercise 10 : Show that the if a vector v belongs to the span of vectors v_1, v_2, v_3 then $\{v, v_1, v_2, v_3\}$ are not linearly independent.

Exercise 11 : Given $n + 1$ vectors in $\mathbb{R}^n$, can they be independent? (Do not give a proof for this).

Exercise 12 : What is the minimum number k of vectors in $\mathbb{R}^n$ that can be linearly dependent?

Exercise 13 : Give n vectors in $\mathbb{R}^n$ which are linearly independent.

Variable transformations

Exercise 14 : Explain why the softmax transformation solves Exercise 1.6. Find a different transformation that also solves this problem.

Exercise 15 : (Optional +) Find a linear time algorithm to find the shortest distance between a pair of points in 2D, given that the distance between (x_1, y_1) and (x_2, y_2) is given by $|x_1 - x_2| + |y_1 - y_2|$.