HY673 – Tutorial 3

Thomas Marchioro

February 2023

Exercise 1 (Change of Variable)**.** Consider a random variable with normal distribution $X \sim \mathcal{N}(0, \sigma^2)$.

1. Compute the PDF of the r.v. $Y = aX^2 + c$, with $a > 0$. *Hint*: When a function g admits multiple inverses $h_1, \dots, h_k$, the change of variable formula for the PDF is
$$p_Y(y) = \sum_{i=1}^{n} p_X(h_k(y)) \left| \frac{d}{dy} h_k(y) \right| .$$
2. Compute the value of $\mathbb{E}\left[X^3 + aX^2 + bX + c\right]$.

Exercise 2 (Maximum Likelihood Estimation)**.** A random variable X is described by the following parametric piecewise-uniform distribution

$$p_\theta(x) = \begin{cases} 1 - \theta, & 0 \leq x \leq 1 \\ 2\theta, & -\frac{1}{2} \leq x < 0 \end{cases} \tag{1}$$

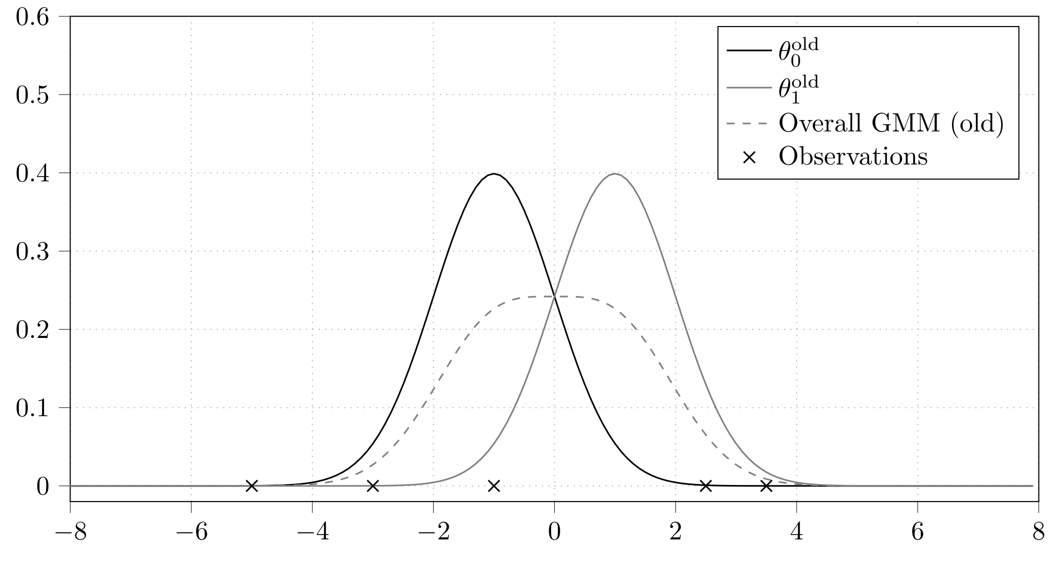
The parameter θ is unknown and needs to be estimated based on the observations $x_1, \dots, x_n$.

1. Let N_1, N_2 be random variables describing the number of observations falling in the regions $[0, 1]$ and $[-1/2, 0)$, respectively, for a sample of n observations. Compute the expectation of both N_1 and N_2.
2. Let n_1, n_2 be the observed values of N_1 and N_2 in the sample $x_1, \dots, x_n$. Derive the maximum likelihood estimator $\hat{\theta}_{\text{MLE}}$ (*Hint*: n_1 and n_2 are all you need to compute the likelihood).
3. Compute the Fisher information $\mathcal{I}(\theta) = -\mathbb{E}_{X \sim p_\theta}\left[\frac{\partial^2}{\partial \theta^2} \log p_\theta(X)\right]$.

Exercise 3 (Toy GMM)**.** Consider a simple 1D Gaussian mixture model (GMM) with two modes θ_0, θ_1. The model is initialized with the following parameters:

- θ_0^{old}: prior π_0^{old}, mean $\mu_0^{\text{old}} = -1$, standard deviation $\sigma_0^{\text{old}} = 1$;
- θ_1^{old}: prior π_1^{old}, mean $\mu_1^{\text{old}} = -1$, standard deviation $\sigma_1^{\text{old}} = 1$.

Suppose the observations are $x = [-5, -3, -1, +2.5, +3.5]$. You can see the initial state and the observed data points in the following figure.



While answering the questions below, you can use Python to perform the necessary calculations. If you do that, you should include your code among the submitted files.

1. **E step**: What are the values of the responsibilities for each observed data point in x? Report them in a table.

Data point	-5	-3	-1	$+2.5$	$+3.5$
Responsibilities of θ_0^{old}					
Responsibilities of θ_1^{old}					

2. **M step**: What are the updated parameters $\pi_0^{\text{new}}, \mu_0^{\text{new}}, \sigma_0^{\text{new}}, \pi_1^{\text{new}}, \mu_1^{\text{new}}, \sigma_1^{\text{new}}$ of the GMM? Provide a plot of its PDF.

3. Compute the overall log-likelihood of the observations x according to the old and new GMM parameters. Did the likelihood increase or decrease after the update? Is this result expected?

4. Sample 1000 points from the new GMM model and make a histogram in the range $[-8, 8]$ with 20 bins. You can use the `plt.hist()` function with the option `density` set to `True`. Compare the histogram with a plot of the PDF for the overall (new) GMM. How do the two graphs compare?

Exercise 4 (Sampling from Multivariate Gaussians)**.** In this exercise you will learn how to sample from any multivariate normal distribution starting from i.i.d. values of a standard normal random variable.

1. Show that if a vector $z \in \mathbb{R}^d$ is a vector with d i.i.d. components $z_i \sim \mathcal{N}(0, 1)$, the linear transformation

$$x = Lz + \mu, \;\; L \in \mathbb{R}^{d \times d}, \mu \in \mathbb{R}^d$$

   follows a multivariate normal distribution $\mathcal{N}(\mu, \Sigma)$ with $\Sigma = LL^\top$. Prove also that any normal random vector can be written using this linear transformation with suitably chosen L and μ.

2. Using the above result, show that the sum of two *dependent* normal random variables $X_1 + X_2$, i.e. with $\text{Cov}(X_1, X_2) \neq 0$ can always be written as the sum of two independent normal random variables $Y_1 + Y_2$.

3. The function `np.random.randn()` in Python can be seen as a generator that produces i.i.d. samples z_i according to a standard normal distribution $\mathcal{N}(0, 1)$. Use this generator to sample $n = 1000$ values from a multivariate distribution $\mathcal{N}(\mu, \Sigma)$, with

$$\mu = \begin{bmatrix} 6 \\ -10 \end{bmatrix}, \;\; \Sigma = \begin{bmatrix} 4 & -3 \\ -3 & 9 \end{bmatrix}$$

   Plot the values on a 2D graph and comment the result.