# Assignment 4: Probabilistic Modeling

The Swiss AI Lab IDSIA, USI, SUPSI

24th of November, 2023

**Due date: 8th of December 2023**

**Submission instructions** Please submit your answers in LaTeX (e.g. `http://overleaf.com`) as a single PDF file. Name the 'file name.lastname.pdf' and upload it to the iCorsi website before the deadline. Late submissions are not accepted. Keep your answers brief. Note that there are a total of *33 points in this assignment.*

**How to get help** We encourage you to use the tutorials to ask questions or to discuss exercises with other students. However, do not look at any report written by others or share your report with others. You can also ask questions on the iCorsi discussion forum or contact anand@idsia.ch.

**Problem** Consider a data set of binary (black and white) images. Each image is arranged into a vector of pixels by concatenating the columns of pixels in the image. The data set has $N$ images $\{\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(N)}\}$ and each image has $D$ pixels, where $D$ is (number of rows $\times$ number of columns) in the image. For example, image $\mathbf{x}^{(n)}$ is a vector $(x_1^{(n)}, \ldots, x_D^{(n)})$ where $x_d^{(n)} \in \{0, 1\}$ for all $n \in \{1, \ldots, N\}$ and $d \in \{1, \ldots, D\}$.

**Question 1**. Explain why a multivariate Gaussian would not be an appropriate model for this data set of images. Give at least three reasons. (5 points)

**Question 2**. Now assume that the images were modeled as independently and identically distributed samples from a $D$-dimensional multivariate Bernoulli distribution with parameter vector $\mathbf{p} = (p_1, \ldots, p_D)$, which has the form

$$P(\mathbf{x} \mid \mathbf{p}) = \prod_{d=1}^{D} p_d^{x_d} (1 - p_d)^{(1-x_d)} \tag{1}$$

where both $\mathbf{x} \in \{0, 1\}^D$ and $\mathbf{p} \in [0, 1]^D$ are $D$-dimensional vectors. Here, the values 0 and 1 taken by $\mathbf{x}$ are for black and white pixels respectively.

What is the likelihood function of $\mathbf{p}$ given the data set of $N$ images $\{\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(N)}\}$? (3 points)

**Question 3**. Derive the log-likelihood. Include all intermediate steps and simplify the final result. (3 points)

**Question 4**. What is the equation for the maximum likelihood (ML) estimate of $\mathbf{p}$? You can assume the critical point to be the maximum, no second derivatives are required. Include all intermediate steps and simplify the final result. (10 points)

**Question 5**. Assuming independent Beta priors on the parameters $p_d$

$$P\left(p_d\right)=\frac{1}{B(\alpha, \beta)} p_d^{\alpha-1}\left(1-p_d\right)^{\beta-1}$$

and $P(\mathbf{p})=\prod_d P\left(p_d\right)$, what is the maximum a posteriori (MAP) estimate of $\mathbf{p}$? $B(\alpha, \beta)=\frac{\Gamma(\alpha) \Gamma(\beta)}{\Gamma(\alpha+\beta)}$ is the normalization constant. Include all intermediate steps and simplify the final result. Hint: Maximize the log posterior with respect to $\mathbf{p}$. Like before, you can assume the critical point to be the maximum. (12 points)