Exercise 1: Maximum likelihood estimation for a Gaussian

The Gaussian pdf parametrised by mean μ and standard deviation σ is given by

$$p(x; \theta) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right], \quad \theta = (\mu, \sigma).$$

Assume we have iid data $\mathcal{D} = \{x_1, \ldots, x_n\}$ drawn from this distribution.

(a) Write down the likelihood function $L(\theta)$.

(b) Write down the log-likelihood function $\ell(\theta) = \log L(\theta)$.

(c) Find the maximum likelihood estimates (MLEs) of μ and σ.

Exercise 2: Cancer-asbestos-smoking example: MLE

The cancer-asbestos-smoking example illustrates a causal structure where, the variables are: a (asbestos exposure), s (smoking), and c (cancer). The model assumes that both asbestos exposure and smoking independently contribute to the risk of developing cancer, while they are independent of each other.

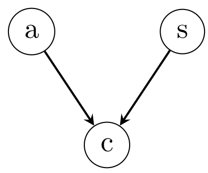

Figure 1: Graphical model for the cancer-asbestos-smoking example.

The distributions of a and s are Bernoulli distributions with success probabilities θ_a and θ_s, respectively:

$$p(a; \theta_a) = \theta_a^a (1-\theta_a)^{1-a}, \quad p(s; \theta_s) = \theta_s^s (1-\theta_s)^{1-s}.$$

The distribution of c given the parents is parametrised as follows:

a	s	$p(c = 1 \mid a, s)$
0	0	θ_c^1
1	0	θ_c^2
0	1	θ_c^3
1	1	θ_c^4

The free parameters of the model are $\boldsymbol{\theta} = (\theta_a, \theta_s, \theta_c^1, \theta_c^2, \theta_c^3, \theta_c^4)$. Assume we observe the following iid data (each row is a data point):

a	s	c
0	1	1
0	0	0
1	0	1
0	0	0
0	1	0

Determine the maximum-likelihood estimates (MLEs) of the parameters $\boldsymbol{\theta}$.

Exercise 3: Cancer-asbestos-smoking example: MAP

Now assume a prior distribution over the parameters $\boldsymbol{\theta}$, given by

$$p(\theta) = \mathcal{U}(0,1), \quad \text{for each } \theta \in \boldsymbol{\theta}.$$

Determine the maximum a posteriori (MAP) estimates of the parameters $\boldsymbol{\theta}$ given the same data as in Exercise 2.

Exercise 4: Cancer-asbestos-smoking example: Posterior Distribution

As above, we assume a prior distribution over the parameters $\boldsymbol{\theta}$, given by

$$p(\theta) = \mathcal{U}(0,1), \quad \text{for each } \theta \in \boldsymbol{\theta}.$$

and assume we observe the same iid data as in Exercise 2. Determine the posterior distribution $p(\boldsymbol{\theta} \mid \mathcal{D})$.

Hint. Beta and Bernoulli distributions are conjugate, meaning that if the prior is a Beta distribution and the likelihood is a Bernoulli distribution, the posterior will also be a Beta distribution. Let $\theta \in [0,1]$ be a parameter with prior:

$$\theta \sim \text{Beta}(\alpha_0, \beta_0)$$

Let $x_1, x_2, \ldots, x_n \sim \text{Bernoulli}(\theta)$ be independent observations. Define:

$$n_1 = \sum_{i=1}^{n} x_i \quad \text{(number of ones)}, \qquad n_0 = n - n_1 \quad \text{(number of zeros)}$$

Then the posterior distribution is:

$$\boxed{\theta \mid x_{1:n} \sim \text{Beta}(\alpha_0 + n_1,\ \beta_0 + n_0)}$$

The posterior mean is:

$$\boxed{\mathbb{E}[\theta \mid x_{1:n}] = \frac{\alpha_0 + n_1}{\alpha_0 + \beta_0 + n}}$$

Special Case: Uniform Prior. If the prior is uniform, i.e., $\theta \sim \text{Beta}(1,1)$, then:

$$\theta \mid x_{1:n} \sim \text{Beta}(1 + n_1,\ 1 + n_0)$$

and the posterior mean becomes:

$$\boxed{\mathbb{E}[\theta \mid x_{1:n}] = \frac{1 + n_1}{2 + n}}$$

Exercise 5: Comparison of MLE and expected value of the posterior

In the previous exercises, we computed the maximum likelihood estimates (MLEs) and the expected values of the posterior distributions for the parameters of the cancer-asbestos-smoking example. Compare the MLEs and the expected values of the posterior distributions for each parameter. What is the conclusion?