Exercises

- The probability density function of the univariate Gaussian distribution is given by:

$$p(x \mid \mu, \sigma^2) = \frac{1}{(2\pi\sigma^2)^{1/2}} e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}}$$

where x is the scalar random variable and μ and σ^2 the mean and variance, respectively.

1) Prove that $\int_{-\infty}^{+\infty} \frac{1}{(2\pi\sigma^2)^{1/2}} e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}} dx = 1$

2) Given a set of samples, x_i, i = 1…N, find the values of μ and σ^2 for which the following function of μ and σ^2 (known as log-likelihood) has a maximum:

$$\sum_{i=1}^{N} \log p(x_i \mid \mu, \sigma^2)$$

3) Let us now assume that each sample, x_i, comes accompanied by a weight, w_i. Find the values of μ and σ^2 for which the following function of μ and σ^2 (known as weighted log-likelihood) has a maximum:

$$\sum_{i=1}^{N} \mathrm{w_i} \log p(x_i \mid \mu, \sigma^2)$$

4) Let us now assume that each sample, x_i, has a normal unitary weight, but it is generated from the following pdf:

$$p(x \mid \mu, \sigma^2, u) = \frac{1}{\left(2\pi \frac{\sigma^2}{u}\right)^{1/2}} e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2/u}}$$

For each sample, x_i, there is an accompanying value, u_i. Find the values of μ and σ^2 for which the following function of μ and σ^2 has a maximum:

$$\sum_{i=1}^{N} \log p(x_i \mid \mu, \sigma^2, u_i)$$

- The probability distribution of a discrete random variable, x, with K possible outcomes is defined by K probability values, $0 \leq p_k \leq 1$, k = 1…K, subject to constraint $\Sigma_k\, p_k = 1$ (the interval where the p_k can take value is called the *simplex*). Such K probability values can be regarded as the parameters, $\theta = [p_1, \ldots, p_K]$, of the probability distribution which can be written as $p(x|\theta)$.

  We now want to estimate θ with maximum likelihood from a training set, X, with N samples. We assume that the training set has N_1 samples with outcome 1, N_2 samples with outcome 2, … N_K samples with outcome K, with $\Sigma_k\, N_k = N$. Prove that the maximum-likelihood estimate for θ is:

$$\theta_{ML} = \left[p_1 = \frac{N_1}{N}, \cdots, p_K = \frac{N_K}{N}\right]$$

  that is, nothing else than the fraction of samples of each outcome.