Bayesian methods of Machine learning

Assignment 1

1. (2 points) Suppose that we have three coloured boxes r (red), b (blue), and g (green). Box r contains 3 apples, 4 oranges, and 3 limes, box b contains 1 apple, 1 orange, and 0 limes, and box g contains 3 apples, 3 oranges, and 4 limes. If a box is chosen at random with probabilities $p(r) = 0.2, p(b) = 0.2, p(g) = 0.6$, and a piece of fruit is removed from the box (with equal probability of selecting any of the items in the box), then what is the probability of selecting an apple? If we observe that the selected fruit is in fact an orange, what is the probability that it came from the green box?

2. (3 points) Show that for data points from a Gaussian distribution with mean μ and variance σ^2, and $I_{nm} = [n = m]$ it holds that

$$\mathbb{E}[x_n x_m] = \mu^2 + I_{nm}\sigma^2.$$

Use the fact that

$$\mathbb{E}[x^2] = \mu^2 + \sigma^2.$$

Hence derive maximum likelihood estimates for mean and variance of the Gaussian distribution.

3. (3 points) Let $X_1, \dots, X_n \sim \text{Poisson}(\lambda)$. Find the Jeffreys' prior. Find the posterior for this prior. Is this prior conjugate?

4. (3 points) Consider a univariate Gaussian distribution $\mathcal{N}(x|\mu, \tau^{-1})$ with conjugate Gaussian-gamma prior and a data set $\mathbf{X} = \{X_1, \dots, X_n\}$ if i.i.d. observations from this Gaussian distribution. Show that the posterior distribution is indeed a Gaussian-Gamma distribution of the functional form as the prior, and write down expressions for the parameters of this posterior distribution.

We define Gaussian-Gamma as a prior for parameters μ and τ as:

$$p(\mu, \tau) = \mathcal{N}(\mu|\mu_0, (\beta\tau)^{-1})\text{Gam}(\tau|a, b),$$

$a, b, \beta > 0$.