CSE 404: Introduction to Machine Learning (Fall 2019)
Homework #3

1. (10 points) Suppose that we have three colored boxes r (red), b (blue), and g (green). Box r contains 2 apples, 4 oranges, and 4 limes, box b contains 2 apple, 3 orange, and 0 limes, and box g contains 4 apples, 3 oranges, and 3 limes. If a box is chosen at random with probabilities $p(r) = 0.1$, $p(b) = 0.3$, $p(g) = 0.6$, 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 blue box?

   Solution: Let us denote apples, oranges and limes by a, o and l respectively.

   The marginal probability of selecting an apple is given by:

$$\begin{aligned} p(a) &= p(a|r)p(r) + p(a|b)p(b) + p(a|g)p(g) \\ &= 0.2 \times 0.1 + 0.4 \times 0.3 + 0.4 \times 0.6 = 0.38 \end{aligned}$$

   The conditional probability $p(g|o)$ is given by:

$$\begin{aligned} p(b|o) &= \frac{p(o|b)p(b)}{p(o)} = \frac{p(o|b)p(b)}{p(o|r)p(r) + p(o|b)p(b) + p(o|g)p(g)} \\ &= \frac{0.6 \times 0.3}{0.4 \times 0.1 + 0.6 \times 0.3 + 0.3 \times 0.6} = 0.45 \end{aligned}$$

2. (10 points) We are given a set of data points $x_1, x_2, \ldots, x_n$ that are i.i.d. drawn from the density function:

$$f(x|\sigma) = \frac{1}{2\sigma} \exp\left(-\frac{|x|}{\sigma}\right), -\infty < x < \infty, \sigma > 0$$

   Find the maximum likelihood estimate of σ.

   Solution: The log-likelihood function of $x_1, x_2, \ldots, x_n$ is

$$\begin{aligned} \ell(\sigma) &= \ln p(x_1, x_2, \ldots, x_n|\sigma) \\ &= \ln \prod_{i=1}^{n} \frac{1}{2\sigma} \exp\left(-\frac{|x_i|}{\sigma}\right) \\ &= \sum_{i=1}^{n} \left(-\ln 2 - \ln \sigma - \frac{|x_i|}{\sigma}\right) \\ &= -n \ln 2 - n \ln \sigma - \frac{\sum_{i=1}^{n} |x_i|}{\sigma}. \end{aligned}$$

   Taking the derivative:

$$\frac{d}{d\sigma}\ell(\sigma) = -\frac{n}{\sigma} + \frac{\sum_{i=1}^{n} |x_i|}{\sigma^2}$$

   Set $\frac{d}{d\sigma}\ell(\sigma) = 0$, we get that $\sigma_{ML} = \frac{1}{n}\sum_{i=1}^{n} |x_i|$.