The Austalian National University
School of Computing
Jo Ciucă & Thang Bui

Semester 2, 2023
Tutorial 8

COMP3670/6670: Introduction to Machine Learning

Question 1 **Bayesian linear regression, one parameter, one training point**

Consider a linear regression problem with a single training point (x, y) where x, y are scalars, and an one-dimensional parameter θ (with no bias/intercept):

$$\text{prior:} \quad p(\theta) = \mathcal{N}(\theta; m_o, v_o) \tag{1}$$

$$\text{likelihood:} \quad p(y|\theta, x) = \mathcal{N}(y; \theta x, v) \tag{2}$$

a) Find the marginal likelihood $p(y|x) = \int \mathrm{d}\theta p(y|\theta, x)p(\theta)$

b) Find the posterior distribution $p(\theta|y, x) = \frac{p(y|\theta,x)p(\theta)}{p(y|x)}$

c) Assuming $x \neq 0$, what happens when v_o is very large?

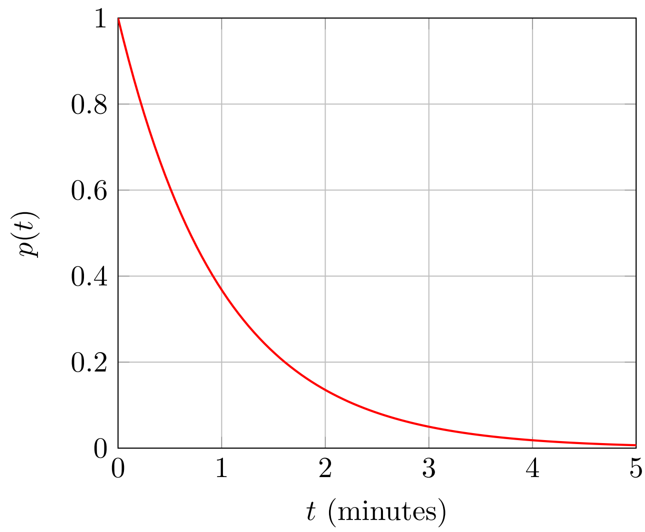
d) What happens when $x = 0$?

Question 2 **Bomb Defusal**

You are a bomb defusal specialist, and you've come across a bomb that has just armed itself. You know from experience with these kinds of bomb, that the time till it explodes is controlled by a random variable T, sampled from the interval $[0, +\infty)$, according to the prior pdf function

$$p(t) = e^{-t} \text{ minutes}$$

which when plotted, looks like this



a) How likely is it for the bomb to take between 1 and 2 minutes to explode?

b) How long on average until the bomb explodes?

c) For any $\phi \in [0, 1)$, how much time t_ϕ would have to pass such that the probability of the bomb having exploded by then is ϕ?

d) Suppose you waited a minute, and the bomb has not yet exploded. What is the posterior distribution $p(t \mid t \geq 1)$ of the bomb's detonation time? Plot $p(t \mid t \geq 1)$ against the prior $p(t) = e^{-t}$.

e) You're a fast defuser, but an even faster runner. The bomb is placing 5 other people in mortal danger. It would take you 15 seconds to move out of the blast radius of the bomb, and 90 seconds to complete a defusal of the bomb. What action maximizes the expected number of lives saved? Attempting a defuse, or running away?

f) You discover a bomb just as it arms itself. This bomb is equipped with a display, reading out the time left till detonation in minutes in seconds (e.g the display says `1: 30` for 90 seconds). Unfortunately, part of the display is damaged, and you can only read the first digit (the minutes digit), which is a 1. What is the posterior $p(t \mid \text{First digit is } 1)$ based on this evidence? Plot this against $p(t)$.

g) Suppose two bombs[1] (with the same distribution $p(t) = e^{-t}$ as before) are armed simultaneously. Let T_1, T_2 denote random variables for the detonation time of each bomb. We define $T_e = \min(T_1, T_2)$, a random variable for the time taken for either bomb to explode, and $T_b = \max(T_1, T_2)$, a random variable for the time taken for both bombs to explode. Find the pdf $p_e(t)$ corresponding to T_e, and $p_b(t)$ corresponding to T_b, satisfying

$$\int_0^t p_e(x)\, dx = P(T_e \leq t) \qquad \int_0^t p_b(x)\, dx = P(T_b \leq t)$$

Plot $p(t) = e^{-t}$ vs. $p_b(t)$ vs. $p_e(t)$. How long, on average, before any bomb explodes? How long, on average, before both bombs explode?

(Hint: Bayes rule is not useful here.)

Question 3 **Definitions of variance**

Recall that given a continuous random variable X defined over a domain $D \subset \mathbb{R}$ with probability distribution function $p(x) : D \to \mathbb{R}$, and a function $f(x) : \mathbb{R} \to \mathbb{R}$, we define [2]

$$\mathbb{E}_X[f(x)] := \int_D f(x)p(x)\, dx$$

The variance $\mathbb{V}[X]$ of a random variable X is defined as

$$\mathbb{V}_X[x] := \mathbb{E}_X\left[(x - \mathbb{E}_X[x])^2\right]$$

It can also be represented in the alternate form

$$\mathbb{V}_X[x] := \mathbb{E}_X\left[x^2\right] - (\mathbb{E}_X[x])^2$$

Prove this!

Question 4 **Substitution of Random variables**

Assume that we have a random variable X on the interval [0,1] characterized by a pdf $p(x) = \frac{3}{2}\sqrt{x}$. Let Y be a random variable on [0,1] such that $Y = X^3$. Compute the pdf of Y

[1] The bombs are placed sufficiently far apart that if one explodes, the other will be undisturbed. The detonation time of each bomb are independent of each other.

[2] Note that $\int_D$ means to integrate over the entire domain D. For example, if $D = [0, 1]$, then $\int_D$ means the same thing as $\int_0^1$.