Selected Exercises of Machine Learning: A Probabilistic Perspective

Mike Craig

Last Updated September 10, 2017

1 Introduction

This is a collection of certain exercise solutions I did as I was reading through this book. As I am not a mathematician, the "proofs" listed here are rough and probably not rigorous. They only serve as a means to explain the concepts in a way that aids my own understanding.

2 Probability

Exercises

Exercise 2.1. **My neighbor has two children. Assuming that the gender of a child is like a coin flip, it is most likely, a priori, that my neighbor has one boy and one girl, with probability 1/2. The other possibilities, two boys or two girls, have probabilities 1/4 and 1/4.**

a. Suppose I ask him whether he has any boys, and he says yes. What is the probability that one child is a girl?

b. Suppose instead that I happen to see one of his children run by, and it is a boy. What is the probability that the other child is a girl?

a. Let G represent one girl, and B represent one boy. Since the neighbor has two children, we can state the entire sample space:

$$S = \{BB, BG, GB, GG\}$$

When the neighbor answers the question, this changes our beliefs about the other child. Using Bayes' theorem:

$$P(G = 1|B \geq 1) = \frac{P(B \geq 1|G = 1)P(G = 1)}{P(B \geq 1)} \tag{1}$$

$$= \frac{2/2 \times 1/2}{3/4} \tag{2}$$

$$= \frac{2}{3} \tag{3}$$

b. If we instead happen to see one of his children, this is a different way of looking at the problem. In this situation, learning the gender of one child tells us nothing about the gender of the other child. Therefore, the gender of the second child is a coin flip, 1/2.

Exercise 2.2. **Suppose a crime has been committed. Blood is found at the scene for which there is no innocent explanation. It is of a type which is present in 1% of the population.**

a. The prosecutor claims: "There is a 1% chance that the defendant would have the crime blood type if he were innocent. Thus there is a 99% chance that he is guilty". This is known as the prosecutor's fallacy. What is wrong with this argument?

b. The defender claims: "The crime occurred in a city of 800,000 people. The blood type would be found in approximately 8000 people. The evidence has provided a probability of just 1 in 8000 that

the defendant is guilty, and thus has no relevance". This is known as the defender's fallacy. What is wrong with this argument?

a. The defendant sharing the blood type does not mean that the defendant himself has a 99% probability of being guilty, just that he shares the same blood type as the guilty party, just like he shares the same blood type with 1% of the population. In a large enough city, there would be a large number of people fitting this description in a small geographical radius.

b. This statement assumes that the defendent is just as guilty (or just as non-guilty) as anyone else in that group of 8000 people. If there truly is no other evidence to tie this defendent to this crime, then that may be so, but if there were any other evidence (drives a similar car as the criminal, lives in the same area, or frequents the same locations), the probability that the defendant is guilty could be much higher.

Exercise 2.3. **Show that the variance of a sum is $Var[X+Y] = Var[X] + Var[Y] + 2Cov[X,Y]$, where $Cov[X,Y]$ is the covariance between X and Y.**

$$Var[X] + Var[Y] + 2Cov[X,Y] = E[(X-\mu_x)^2] + E[(Y-\mu_y)^2] + 2E[(X-\mu_x)(Y-\mu_y)] \tag{4}$$

$$= E[X^2 - 2X\mu_x + \mu_x^2] + E[Y^2 - 2Y\mu_y + \mu_y^2] + E[2XY - 2X\mu_y - 2Y\mu_x + 2\mu_x\mu_y] \tag{5}$$

$$= E[X^2 - 2X\mu_x + \mu_x^2 + Y^2 - 2Y\mu_y + \mu_y^2 + 2XY - 2X\mu_y - 2Y\mu_x + 2\mu_x\mu_y] \tag{6}$$

$$= E[X^2 + 2XY - 2X(\mu_x + \mu_y) + Y^2 - 2Y(\mu_x + \mu_y) + 2\mu_x\mu_y] \tag{7}$$

Note that $E(X+Y) = E(X) + E(Y) = \mu_x + \mu_y = \mu_{xy}$. Given this,

$$E[X^2 + 2XY - 2X(\mu_x + \mu_y) + Y^2 - 2Y(\mu_x + \mu_y) + 2\mu_x\mu_y] \tag{8}$$

$$= E[X^2 + 2XY - 2X\mu_{xy} + Y^2 - 2Y\mu_{xy} + 2\mu_x\mu_y] \tag{9}$$

$$= E[(X + Y - \mu_{xy})^2] \tag{10}$$

$$= Var[X+Y] \tag{11}$$