Probability exercises

Introduction to Machine Learning (CSCI 1950-F), Summer 2011

Solve the following problems. Provide mathematical justification for your answers.

(1) The prosecutor's fallacy

You are on the jury for a murder trial, and have to decide whether the accused person is guilty or innocent. The sole piece of evidence presented against the accused is a DNA match. There is good scientific evidence indicating that the probability of a DNA match for an innocent person is 1/1000 and that the probability of a DNA match for a guilty person is 99/100. A DNA match has been established for the accused person, and the prosecutor claims that this is definitive evidence indicating the guilt of the accused. There are $10,000$ residents in the city, and before seeing the evidence your best judgment is that it is equally likely that anyone of these residents committed the crime. So, summarizing:

$$P(match|guilty) = 99/100,$$

$$P(match|innocent) = 1/1000,$$

$$P(guilty) = 1/10000.$$

What is $P(guilty|match)$? Has the prosecutor presented a convincing argument?

(This type of thing actually happens: look up "Sally Clark" on Wikipedia.)

(2) Mutual independence, pairwise independence, conditional independence

Consider the probability measure space $(\Omega, \mathcal{A}, P)$ with $\Omega = \{1, 2, 3, 4\}$, $\mathcal{A} = 2^{\Omega}$ (i.e. the power set, consisting of all subsets of Ω), and $P(k) = 1/4$ for all $k \in \Omega$. Let $A = \{1, 2\}$, $B = \{1, 3\}$, and $C = \{2, 3\}$.

(a) Are A and B independent?

(b) Are A, B, C mutually independent? Are they pairwise independent?

(c) Are A and B conditionally independent given C?

(3) Bayes' rule for a conditional measure

Assume a probability measure space $(\Omega, \mathcal{A}, P)$. Fix an event $C \in \mathcal{A}$ such that $P(C) > 0$. Show that

$$P(A|B \cap C) = \frac{P(B|A \cap C)P(A|C)}{P(B|C)}.$$

(Note: It is straightforward to prove this identity directly using the definition of conditional probability. On the other hand, this can be thought of as "Bayes' rule when everything is conditioned on C", or more precisely, Bayes' rule for the conditional probability measure defined by $Q(A) = P(A|C)$, since (as you can verify) $Q(A|B) = P(A|B \cap C)$.)

(Challenge problem)

This problem is just for fun. You don't need to turn in a solution — it is optional and will not be graded. This is a good exercise in measure theory.

Assume a probability measure space $(\Omega, \mathcal{A}, P)$. Using the properties of probability measures, prove that for an arbitrary (real-valued) random variable $X : \Omega \to \mathbb{R}$, the function $F : \mathbb{R} \to \mathbb{R}$ defined by $F(x) = P(X \leq x)$ satisfies the following:

(i) F is nondecreasing,

(ii) F is right continuous,

(iii) $\lim_{x\to\infty} F(x) = 1$, and

(iv) $\lim_{x\to-\infty} F(x) = 0$.