Algorithmic Learning Theory
MSIS 26:711:685
Homework 1

Instructor: Farid Alizadeh

Due Date: Wednesday September 18, 2019, at 11:50PM

last updated on September 21, 2019

Please answer the following questions in **electronic** form and upload and submit your files to Sakai Assignment site, before the due date. Make sure to click on the **submit** button. The file format should be pdf. You may either typeset or hand write your answers. In case of hand-write transform your work into pdf using apps such as **CamScan**. You should have only a single pdf file. For this homework, if you have R or Python scripts, include the text of the script with your pdf submission (no separate file is needed).

1. Let X_1 and X_2 be two discrete random variables, where X_1 can attain values -1, 0, and 1, and X_2 can attain values 2, 4, and 6. The joint probability mass function of these two random variables are given in the table below:

	X_1			
X_2	-1	0	1	$f_{X_2}()$
2	1/10	1/10	1/10	
4	1/8	1/4	1/5	
6	1/10	1/80	1/80	
$f_{X_1}()$				

Answer the following questions: For answers see tha attached python script

1a) Find the marginal density functions $f_{X_1}(t_1)$ and $f_{X_2}(t_2)$. and write them in the appropriate location in the table above.

1b) What are the *expected values* of X_1 and X_2?

1c) What are the *variances* of X_1 and X_2, and the Cov(X_1, X_2)?

1d) Are the random variables X_1 and X_2 independent or not? Justify your answer.

1e) Find the conditional density functions:

- $f_{X_1|X_2}(t_1|X_2 = 2)$

- $f_{X_2|X_1}(t|X_1 = 0)$

1f) What is the *conditional expected value* of X_1 given $X_2 = 2$:

1g) What is the *conditional variance* $\text{Var}_{X_2|X_1}[X_2|X_1 = 0]$?

2. A new mammography test kit is devised to screen for breast cancer. The manufacturer of the kit states that 90% of women who do have breast cancer test positive for the test. The test also has a 9% false-positive rate (that is 9% of women without breast cancer test positive.) About 8% of women who are referred by doctors to take this test (for example, because they saw unusual lumps in their breast) do have breast cancer.

2a) A woman has tested positive. What is the probability that she has breast cancer?

2b) A nice feature of this test is that applying it to the same patient several times, then the outcome of each test is independent of previous results. So if a person who does not have cancer tests positive, the chances of her testing positive on a subsequent test remains 9%. If a patient tests positive for the first time and tests positive again for the second time, what is the probability that she has breast cancer?

2c) A woman has tested negative for breast cancer. What is the probability that she does not have breast cancer?

3. Answer the following questions with proof.

3a) Show that if the pair of events (A, B) are independent, then so are the pairs (not A, B), (A, not B), and (not A, not B).

$$\Pr[\text{not}A\&B] = \Pr[\text{not}A \mid B]\Pr[B] = \big(1-\Pr[A \mid B]\big)\Pr[B] = (1-\Pr[A])\Pr[B] = \Pr[\text{not}A]\Pr[B]$$

Other cases are special case of the above. For A and not$B$ switch A and B. For not A and not B, use the fact that A and not B are independent, then using the above, (replace not B for A, and A for B) we get not A and not B are independent.

3b) Show that the sure event is independent of all other events. Also the impossible event is independent of all other events.

This is a tautology:

$$\Pr[A \cap S] = \Pr[A] = \Pr[A] \times 1 = \Pr[A] \times \Pr[S]$$

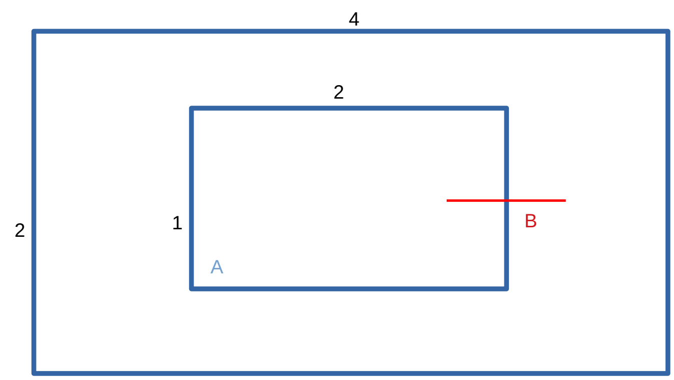
Similarly for the impossible event,

$$\Pr[A \cap \emptyset] = \Pr[\emptyset] = 0 = 0 \times \Pr[A] = \Pr[\emptyset] \times \Pr[A]$$

This result also follows from Question 3a. Note that this case may seem a bit strange. We are saying what is the probability of A given something impossible has happened? However, by definition something impossible cannot happen. So a logical response is if impossible happens (a false statement) then any value can be assigned: In fact, from the definition $\Pr[A \cap \emptyset]/\Pr[\emptyset] = 0/0$ which is undefined. However, if we use the equivalent definition that A and B are independent if and only if $\Pr[A \cap B] = \Pr[A] \times \Pr[B]$, then it makes sense to call $\emptyset$ and A independent. Here is another way of saying it: Since the impossible event never happens, it has no effect on whether any event occurs or not. Similarly for the sure event: Since the sure event always happens, it has no effect on whether another event happens or not.

3c) Consider an infinite sample space, say the set of points (x, y) on the plane where $0 \leq x, y \leq 1$. Is an event of probability zero necessarily independent of all other events? Either prove yes, or give a counterexample.

Consider the following picture:



Let A be the event that a randomly and uniformly chosen point from the 2×4 outer rectangle falls in the 1×2 inner rectangle. Let B be the event that such a randomly chosen point falls on the red line segment. Now by comparing areas, we see that $\Pr[A] = 1/4$ and $\Pr[B] = 0$, but B is not the impossible event. For instance,

$$\Pr[A|B] = 1/2 \neq \Pr[A] = 1/4$$

So, even though B is a probability zero event, it is not independent of A.

3d) Two events, A and B are *mutually exclusive* when both cannot be true at the same time, that is $A \cap B = \emptyset$. When can the A and B be independent?

If A and B are mutually exclusive, then $A \subseteq \text{not}B$. So $\Pr[A\&B] = \Pr[\emptyset] = 0$. So for A and B to be independent we must have $0 = \Pr[A\&B] = \Pr[A] \times \Pr[B]$ This can only happen if at least one of A and B are events of probability zero.

4. A random variable with *geometric distribution* is a discrete variable that can attain values of 1, 2, ... It is used to model the number of trials of a Bernoulli (binary) test until the first success occurs. So, if a Bernoulli trial can have "success" with probability p and "failure" with probability $1-p$, then the number of iid trials before a success happens follows the geometric distribution. The pmf for this distribution is given by

$$f(k) = (1-p)^{k-1}p$$

Because if the first success is seen at the k^{th} trial, it means that we have seen $k-1$ failures in a row and then saw success in the k^{th} trial.

4a) Someone gives you a coin and tells you it may be a fake coin. Furthermore, a fair coin has a 50% probability of heads, while a fake one has only a 20% chance of heads turning up. You toss the coin 3 times in a row, and it turns up tails, and only on the fourth toss, you see the first heads. Based on this observation, would you call this coin fair or fake? Justify your answer.

Since no prior is given, we can assume that prior probabilities of fake and real coins are equal to 1/2. For computations see the attached python script.

4b) Suppose we have two scenarios: the probability of success is p_1, or probability of success is p_2. Given values of $0 \leq p_1 < p_2 \leq 1$, what is the minimum number of times failures should occur before the first success in order for you to select scenario 1 over scenario 2?

To have the likelihood of the event with probability of success p_1 to be larger than the one with probability of success p_2 we must have:

$$(1-p_1)^{k-1}p_1 \geq (1-p_2)^{k-1}p_2$$

$$\left(\frac{1-p_1}{1-p_2}\right)^{k-1} \geq \frac{p_2}{p_1}$$

$$(k-1)\big(\ln(1-p_1) - \ln(1-p_2)\big) \geq \ln(p_2) - \ln(p_1)$$

$$k \geq \frac{\ln(p_2) - \ln(p_1)}{\ln(1-p_1) - \ln(1-p_2)} + 1$$