Practice Questions (week 10)

Semester 2, 2019

These questions are about conditional probability, Bayes' formula, law of total probability, generalised Bayes' rule, naive Bayes classifiers.

- Ross, *A first course in Probability* (6th Ed.), Chapter 3.

1. (a) Suppose there is a 40% chance it rains today, and a 25% chance that it rains both today and tomorrow. If it happens to rain later today, what is the chance of rain tomorrow?

   (b) Suppose there is a 10% chance of storm today, and a 30% chance of storm tomorrow if there is a storm today. What is the chance of a storm on both days?

2. Two fair dice are rolled. What is the (conditional) probability of:

   (a) obtaining a sum of at least 10 given at least one 6?

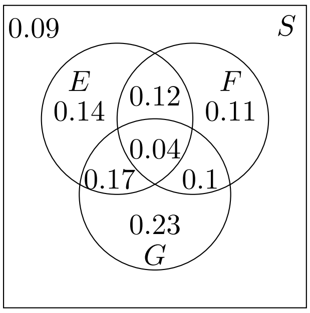
   (b) rolling at least one 6 given the sum is at least 10?

   (c) obtaining a sum of at least 10 given the first dice rolled is a 6?

   (d) obtaining a double digit product given a sum of at most 7?

   (e) obtaining a sum of at most 7 given a double digit product?

3. Consider events A, B, C in a sample space S with the probability of the various intersections as indicated in the following Venn diagram



   Determine each of the following:

   (a) $\Pr(E|F)$

   (b) $\Pr(E \cup F|G^c)$

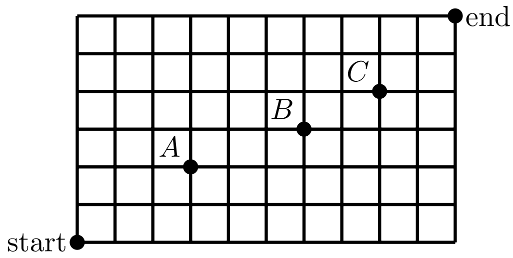
   (c) $\Pr(E \cap G|E \cup F)$

   (d) $\Pr(F^c|G)$

4. A magician shuffles a standard deck of 52 playing cards and you pick one at random. They get to ask one question before trying to guess your card. Which of the following questions should they ask to have the best chance of guessing your card correctly?

- Is it a black card?
- Is it a hearts card?
- Is it a queen?
- Is it the 2 of spades?

5. Consider an ant that must traverse the grid shown below from start to end only moving up or right along each edge. Suppose that every valid path is equally likely. (See also the similar question from last weeks practice questions.)



   (a) What is the probability the ant passes through the point B given it passes through the point A?
   (b) What is the probability the ant passes through the point A given it passes through the point C?
   (c) What is the probability the ant passes through the point C given it passes through the point B?
   (d) What is the probability the ant passes through the point C given it passes through both the points A and B?
   (e) What is the probability the ant passes through the point B given it passes through both the points A and C?

6. A box contains fair and biased coins. Each biased coin has a probability of flipping heads 60% of the time. It is known that 30% of the coins in the box are biased. A coin is randomly selected from the box and flipped.

   (a) What is the probability it is biased if the result is heads?
   (b) Suppose we flip the same coin a second time and it again return heads, what is the probability it is biased.
   (c) How many times in a row would we need to flip heads to be at least 50% certain the chosen coin is a biased one?

7. A movie recommendation website has $70,000$ users that enjoy horror movies and $130,000$ that do not. After the release of a new horror movie all of the users were polled. $55,000$ of the users that enjoy horror movies said they enjoyed the new movie, whereas $90,000$ users that do not enjoy horror also did not enjoy the new movie. Given a randomly chosen user that enjoyed the new movie, what is the probability they enjoy horror movies generally?

8. Let A, B be two events in a probability space S. What can you say about $\Pr(A|B)$ in each of the following cases? (You may assume $\Pr(B) > 0$.)

   (a) $A \subset B$

   (b) $B \subset A$

   (c) $A \subset B^c$

   (d) $B \subset A^c$

   (e) $\Pr(A) = 0.6$, $\Pr(B) = 0.4$ and A, B are independent events

   (f) $\Pr(B|A) = 0.8$, $\Pr(B) = 0.7$ and $\Pr(A) = 0.3$

9. Differential privacy is a technique that can be used to help protect the privacy of individuals participating in surveys (for more on this, see this SIAM News article.)
   Consider a survey asking 'Have you ever smoked weed?'. Participants may hesitate to respond 'yes' truthfully due to fear of potential consequences if they were identified. To obscure the response of the individual, the survey can be designed as follows. If a participant answers 'yes', then record a 'yes', but if a participant answers 'no', then instead randomly record 'yes' or 'no' with $50 : 50$ chance. Let Y be the event of a 'yes' being recorded and let S be the event of the individual having smoked weed before. Suppose $\Pr(S) = 0.2$ and that 90% of participants are anticipated answer truthfully given this survey design (that is $\Pr(Y|S) = 0.9$, i.e. there is still some distrust).

   (a) What is $\Pr(Y)$?

   (b) What is $\Pr(S|Y)$? Explain how this increases the privacy of participants.

   (c) If $\Pr(S)$ is unknown, e.g. let $\Pr(S) = p$, and the survey produces a result of $\Pr(Y) = 0.7$, then what is $\Pr(S)$?

   (d) Why would it not be sensible to instead modify 'yes' responses in the design rather than 'no' responses?

10. You have lost your TV remote. You presume it is equally likely to be in one of three rooms in your house. Let R_1, R_2, R_3 be the event that the remote is in each of the three rooms. You intend to start with a quick look in your bedroom (R_1 say), but it is quite a mess and you are only 60% confident you will find the remote if it is in there. Let F be the event that you find the remote during your quick search of the first room. If your quick search is unsuccessful, what is the probability the remote is in each of the three rooms?

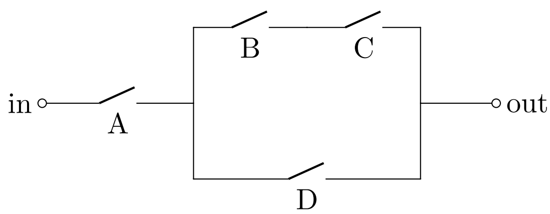
11. Consider rolling two fair die. Let A be the event the first dice roll is even, B be the event the second dice roll is even, and C be the event that their sum is odd.

(a) Show that the three events are pairwise independent (i.e. A, B are independent, A, C are independent and B, C are independent), but not independent (i.e. all together).

(b) Consider now the event D that the multiple of the two dice rolls is odd. What can we say about the independence of A, B, C, D or any pair/triple of these?

12. As hotel guests enter the lobby during the day they greet the receptionist with 'g'day', 'hello' or 'bonjour' if they are from Australia, the USA or France respectively. The receptionist writes each of these down as they occur. The proportion of guests currently staying at the hotel are 30% Australian, 50% American and the remainder are French (and assume each comes and goes equally often). Suppose you pick a random letter from the last greeting written down by the receptionist and find that it is a consonant. What is the probability the last guest to greet the receptionist was an Australian?

13. (a) Consider the following circuit with four switches that operate independently.



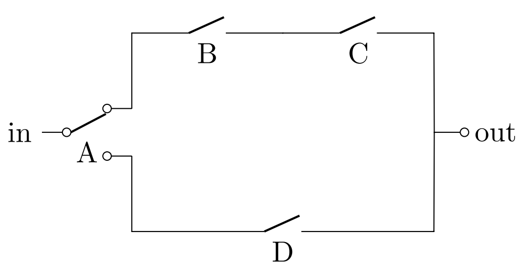
Let A, B, C, D be the events that the respective switches are closed. For convenience let $a = \Pr(A)$, and similar for switches B, C, D.

i. What is the probability that current flows through the circuit (i.e. from 'in' to 'out')?

ii. What is the probability that current flows through the circuit given switch B is closed?

iii. What is the probability that switch D is closed given current flows through the circuit?

(b) Consider the following modification to the previous circuit with the four switches again operating independently.



In this case let A be the event shown where the switch is in the 'up' position. A^c is then the event when the switch is in the 'down' position (i.e. there is no in-between).

i. What is the probability that current flows through the circuit (i.e. from 'in' to 'out')?
ii. What is the probability that current flows through the circuit given switch B is closed?
iii. What is the probability that switch D is closed given current flows through the circuit?

14. Consider 3 biased coins, each with a different bias. Label the coins A, B, C and suppose each flips heads with probability $0.4, 0.5, 0.6$ respectively. Suppose the three coins are placed down in a row in a manner such that any ordering is equally likely. The coins are then flipped. Let L, M, R denote the events that the left, middle and right coin are heads.

    (a) Use a naive Bayes' classifier to determine what the most likely ordering of the coins is given the result $L \cap M^c \cap R$ (i.e. heads, tails and heads going from left to right).
    (b) Explain why this result fits with your intuition.
    (c) Explain why a naive Bayes' classifier is a particularly good approach for this problem.

15. A medical researcher has gathered data on people with and without diabetes. Let D be the event that someone in their database has diabetes. Let W be the event that person hast weight at least 80kg. Let H be the event that person has height at least 160cm. Let A be the event that person has age at least 40 years. Let F be the event that person is female. Suppose that the researcher tabulates some summary statistics from the data which show that

    - $\Pr(W|D) = 0.65$
    - $\Pr(H|D) = 0.35$
    - $\Pr(A|D) = 0.70$
    - $\Pr(F|D) = 0.40$
    - $\Pr(W|D^c) = 0.55$
    - $\Pr(H|D^c) = 0.50$
    - $\Pr(A|D^c) = 0.45$
    - $\Pr(F|D^c) = 0.60$

    They decide to construct a naive Bayes' classifier from this information to decide if their own patients have diabetes. They also use the estimate that 10% of the population has diabetes.

    (a) Given a male patient who weighs 77kg, is 183cm tall, and is 27 years old, are they more likely to have, or not have diabetes based on this classifier?
    (b) Why is a naive Bayes' classifier from this summary data likely to be a poor model/fit?