If you need this document in another format then please contact course organiser Ian.Stark@ed.ac.uk

Homework 4 is currently running: ask your tutor if you have any questions about it; you can also discuss the work with other students in your tutorial.

Task A

This is a task to explore the difference between 'mutually exclusive' and 'independent' events.

Suppose you roll two fair six-sided dice and look for the following three events.

- Event A: First die lands on 1
- Event B: Second die shows larger number than first die
- Event C: Both dice show same number

Which of these events are independent and which are mutually exclusive?

Task B

A random variable X has the following probability distribution.

x_i	1	3	4	6
p_i	0.25	0.3	0.3	0.15

Compute the cumulative distribution function (CDF) and from this calculate the probability $P(2 \leq X \leq 5)$.

Plot the probability mass function (PMF) and the cumulative distribution function.

Task C

This is a task to explore how the technical term "expected" in "expected value" relates to an everyday understanding of the word.

An exam is made up of 4 multiple-choice questions. Each question has 4 answers, of which exactly 1 is correct. A total of 256 students sit this exam. If each student answers each question randomly then what is the expected number of exam scripts with no correct answers? With one correct answer? With 2, with 3, with 4?

Consider the expected number of exam scripts with all four answers correct. What is the probability that the actual number of all-correct scripts is exactly the "expected" number? What is the probability that it is lower than that? Higher?

Task D

Alice and Bob play a game by taking turns throwing a fair six-sided die. Alice throws first and the winner is the first one to throw a 5 or 6. What are the probability of the events $A = \{\text{Alice wins}\}$ and $B = \{\text{Bob wins}\}$?

(You may use the fact that $\frac{1}{1-x^2} = 1 + x^2 + x^4 + x^6 + x^8 + \cdots$.)

Solution Notes

Task A

One way to work out the probability of the various events is to set out a 6×6 table of all 36 outcomes, each of probability $1/36$, and then identify the regions corresponding to the three events.

From this we discover that for events A and B we have $P(B|A) = 5/6$ and $P(B) = 15/36$. These are not equal so A and B are **dependent**. In addition $P(A \cap B) = 5/36$ is non-zero so they are not mutually exclusive.

For events A and C we calculated $P(C|A) = 1/6$ and $P(C) = 6/36 = 1/6$. These are identical so A and C are **independent**. Again $P(A \cap C) = 1/36$ is non-zero so they are not mutually exclusive.

For events B and C we have $P(B \cap C) = 0$ and they are **mutually exclusive**. As a result we also have $P(C|B) = 0$ which is different to $P(C) = 1/6$ meaning that B and C are **dependent**.

Summary: A and C are independent; B and C are mutually exclusive.

For any pair of events where both have non-zero probability, if they are mutually exclusive then they cannot be independent. As a result such a pair will either be mutually exclusive (B and C), independent (A and C), or neither (A and B), but never both.

Task B

Here is the CDF of X.

$$F_X(x) = \begin{cases} 0 & \text{if } x < 1 \\ 0.25 & \text{if } 1 \leq x < 3 \\ 0.55 & \text{if } 3 \leq x < 4 \\ 0.85 & \text{if } 4 \leq x < 6 \\ 1 & \text{if } 6 \leq x \end{cases}$$

Figure 1 shows plots of the PMF and CDF of X. We can use the CDF to calculate $P(2 \leq X \leq 5)$ as follows:

$$P(2 \leq X \leq 5) \;=\; P((X = 3) \cup (X = 4)) \;=\; F(5) - F(2) \;=\; 0.85 - 0.25 \;=\; 0.6\ .$$

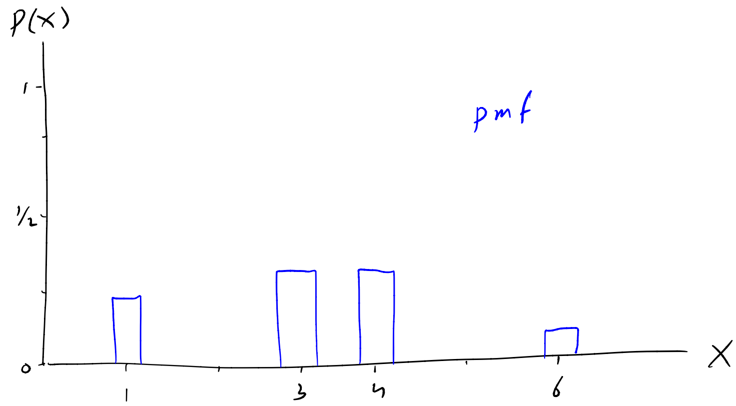
Task C

For each student the guessing of correct answers is a binomial experiment composed of $n = 4$ trials (4 questions). Since the student is randomly guessing each answer, the 4 trials are independent and the probability of success (i.e. guessing the correct answer) in each trial is $1/4$.

Let random variable X be the number of correct answers guessed by a single student. This quantity is distributed according to a binomial distribution $X \sim \text{Bin}(n = 4, p = 1/4)$. Therefore the probability that $X = x$ for $x = 0, 1, 2, 3, 4$ is

$$P(X = x) \;=\; \binom{4}{x} (1/4)^x (3/4)^{4-x}\ .$$

Let random variable Y_x be the number of exam scripts with exactly x correct answers. Then these are also binomially distributed $Y_x \sim \text{Bin}(n = 256, p = P(X = x))$ for $x = 0, 1, \ldots, 4$. We



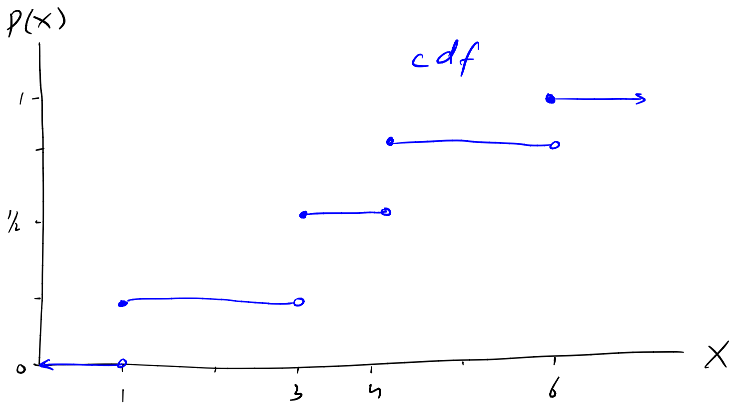


Figure 1: Task B: Probability mass function (PMF) and cumulative distribution function (CDF) of random variable X

deduce that their expected value $\mathrm{E}(Y_x) = 256 \cdot P(X = x)$ and they are distributed over the values of x as follows:

$$\mathrm{E}(Y_x) \;=\; 256 \cdot P(X = x) \;=\; \begin{cases} 81 & x = 0 \\ 108 & x = 1 \\ 54 & x = 2 \\ 12 & x = 3 \\ 1 & x = 4 \end{cases}$$

Consider the scripts where every answer is correct. Among 256 students we "expect" one of these as $\mathrm{E}(Y_4) = 1$. But the actual number is a binomially distributed random variable, $Y_4 \sim \mathrm{Bin}(256, 1/256)$.

$$\begin{aligned} P(Y_4 = 1) &= \binom{256}{1} \cdot \left(\frac{1}{256}\right)^1 \cdot \left(\frac{255}{256}\right)^{255} &&= 0.37 \quad (2\,\text{s.f.}) \\ P(Y_4 < 1) &= P(Y_4 = 0) \;=\; \binom{256}{0} \cdot \left(\frac{255}{256}\right)^{256} &&= 0.37 \quad (2\,\text{s.f.}) \\ P(Y_4 > 1) &= 1 - P(Y_4 = 1) - P(Y_4 = 0) &&= 0.26 \quad (2\,\text{s.f.}) \end{aligned}$$

So the chance that no student gets all questions right is 37%, that two or more get all questions right is 26%, and the chance of the "expected" number of exactly one all-correct student script is 37%. It is more likely than not random variable Y_4 will fail to take its expected value. The meaning of "expected" in work on probability does not always align with informal understanding of the word, which can seem surprising.

Task D

Assume that Alice and Bob go on throwing for ever, and define events A_k and B_k for $k = 1, 2, \ldots$ as below.

$$\begin{aligned} A_k &= \{\text{on the } k\text{th throw Alice gets a 5 or 6}\} \\ B_k &= \{\text{on the } k\text{th throw Bob gets a 5 or 6}\} \end{aligned}$$

The probability that on any throw a player gets a 5 or 6 is $p = 2/6 = 1/3$. Therefore $P(A_k) = P(B_k) = p = 1/3$, and $P(A_k^c) = P(B_k^c) = q = 1 - p = 2/3$. Taking event A as "Alice wins" we have the following equivalence:

$$A \;=\; A_1 \cup (A_1^c \cap B_1^c \cap A_2) \cup (A_1^c \cap B_1^c \cap A_2^c \cap B_2^c \cap A_3) \cup \cdots .$$

Since the events A_k and B_k are independent and the events A_1 and $A_1^c \cap B_1^c \cap A_2$ and $A_1^c \cap B_1^c \cap A_2^c \cap B_2^c \cap A_3$ are pairwise mutually exclusive, we can multiply and add their probabilities:

$$\begin{aligned} P(A) &= p + qqp + qqqqp + \ldots \;=\; p(1 + q^2 + q^4 + \ldots) \\ &= \frac{p}{1 - q^2} \;=\; 0.6 \,. \end{aligned}$$

With event B as "Bob wins" we can calculate the probability of that in the same way.

$$\begin{aligned} B &= (A_1^c \cap B_1) \cup (A_1^c \cap B_1^c \cap A_2^c \cap B_2) \cup \ldots \\ P(B) &= qp + qqqp + qqqqqp + \cdots = pq(1 + q^2 + q^4 + \ldots) \\ &= \frac{pq}{1 - q^2} \;=\; 0.4 \,. \end{aligned}$$

We can also approach this using recursion: note that if Alice fails to throw a 5 or 6 on her first go, and Bob does the same, then we are back to the starting situation. Alice wins either if she succeeds on her first throw or if both Alice and Bob fail and then Alice goes on to win. Those are two mutually exclusive possibilities, which gives us the following equation.

$$P(A) = \frac{1}{3} + \frac{2}{3} \cdot \frac{2}{3} \cdot P(A)$$

We can solve this algebraically.

$$\begin{aligned} P(A) &= \frac{1}{3} + \frac{4}{9} \cdot P(A) \\ 9P(A) &= 3 + 4P(A) \\ 5P(A) &= 3 \\ P(A) &= 0.6 \end{aligned}$$

A similar equation gives Bob's probability of winning.

$$P(B) = \frac{2}{3} \cdot \frac{1}{3} + \frac{2}{3} \cdot \frac{2}{3} \cdot P(B)$$

From this $P(B) = 0.4$ again follows algebraically.