Jacobs University Bremen
Keivan Mallahi-Karai

Due: September 27, 2017
Assignment 2

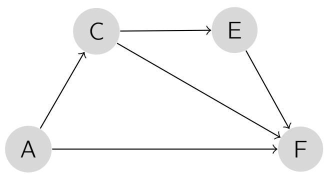
Elements of Probability

Solve only 5 out of the following 6 problems.

(2.1) A network connects computers A and F via intermediate nodes C, E as shown below. For each pair of directly connected nodes, there is a probability $p = 3/4$ that the connection from i to j is up. Assume that the link failures are independent events.

(a) Find the probability that the connection from A to F through at least one of the paths is up.

(b) Suppose that due to some technical work, the connections CE and CF are simultaneously on or off, with the same probability $p = 2/3$. This aside, the other connections are all independent. Under this assumption, compute the probability that the connection from A to F through at least one of the paths is up, and compare the result to part (a).



Solution. Let us denote by C_1, C_2, and C_3 the events that the paths $ACEF$, ACF, and AF are up, respectively. Note that $ACEF$ is up, when all connections AC, CE and EF are up. Since these events are assumed to be independent, we have

$$\mathbb{P}[C_1] = p^3.$$

Similarly, we have

$$\mathbb{P}[C_2] = p^2, \qquad \mathbb{P}[C_3] = p.$$

Let us now consider $C_1 \cap C_2$. Note that $C_1 \cap C_2$ entails that all four paths AC, CE, EF, CF are up, and hence $\mathbb{P}[C_1 \cap C_2] = p^4$. Similarly, we obtain

$$\mathbb{P}[C_1 \cap C_3] = p^4, \qquad \mathbb{P}[C_2 \cap C_3] = p^3, \qquad \mathbb{P}[C_1 \cap C_2 \cap C_3] = p^5.$$

From here, we obtain

$$\mathbb{P}[Connect] = \mathbb{P}[C_1 \cup C_2 \cup C_3] = (p^3 + p^2 + p) - (p^4 + p^4 + p^3) + p^5 = p + p^2 - 2p^4 + p^5.$$

(b) The only difference in part (b) is that since connections CE and CF are both on or off, we have

$$\mathbb{P}[C_1 \cap C_2] = p^3, \qquad \mathbb{P}[C_1 \cap C_3] = p^4, \qquad \mathbb{P}[C_2 \cap C_3] = p^3, \qquad \mathbb{P}[C_1 \cap C_2 \cap C_3] = p^4.$$

From here we obtain

$$\mathbb{P}[Connect] = \mathbb{P}[C_1 \cup C_2 \cup C_3] = (p^3 + p^2 + p) - (p^3 + p^4 + p^3) + p^4 = p + p^2 - p^3.$$

(2.2) Alice and Bob use the following method for determining who pays for lunch. They have three non-standard dice, each with the following numbers on them:

$$(A) : 1, 1, 6, 6, 8, 8$$

$$(B) : 2, 2, 4, 4, 9, 9$$

$$(C) : 3, 3, 5, 5, 7, 7$$

Each side of each one of these dice has probability $1/6$ of coming up. Alice and Bob, each, throws one of the dice, and the one who rolls the smaller number loses and will buy lunch.

(a) If Bob uses die A and Alice uses die B, show that the odds that Alice wins is more than $1/2$.
(b) If Bob uses die B and Alice uses die C, show that the odds that Alice wins is more than $1/2$.
(c) If Bob uses die C and Alice uses die A, show that the odds that Alice wins is still more than $1/2$!

Solution. Suppose that Bob uses die A and Alice uses die B. The set of possible outcomes are given by

$$\Omega = \{(1, 2), (1, 4), (1, 9), (6, 2), (6, 4), (6, 9), (8, 2), (8, 4), (8, 9)\}.$$

Each outcomes has probability $\frac{4}{36} = \frac{1}{9}$. Note that the event that Alice wins is given by

$$A = \{(1, 2), (1, 4), (1, 9), (6, 9), (8, 9)\}$$

Hence

$$\mathbb{P}[A] = \frac{5}{9} > \frac{1}{2}.$$

Parts (b) and (c) can be dealt with in a similar fashion.

(2.3) A fair coin is flipped repeatedly. A *streak* is a consecutive sequence of flips in which the same side comes up. For example, in the sequence

$$THHHTTTTHTTH$$

there is a streak of tails of length 4 starting at the fifth flip, and a streak of length 3 of heads starting at the second flip.

(a) Suppose that the coin is flipped 100 times. What is the expected value of streaks of length 4?
(b) Suppose that the coin is flipped 4 times. What is the probability that it produces a streak of lenth 4?
(c) Suppose that the coin is now flipped 100 times. Show that the probability that there is no streak of length 4 is less than 4 percent.

Hint: For part (a), write the number of streaks as a sum of appropriately defined Bernoulli random variables. For (c), break the sequence of flips into 25 disjoint blocks of size 4. Use part (b) and independence.

Solution. Let S denote the total number of streaks of length 4. We also set S_j to be the Bernoulli random variable which takes value 1 when we have a streak of length 4 starting from the j-th flip. For instance, if the outcomes is given by $THHHTTTTHT..$, then $S_1 = S_2 =$

$S_3 = S_4 = 0$ but $S_5 = 1$ since there is a streak of length 4 starting from the fifth flip. It is clear that

$$S = S_1 + S_2 + \cdots + S_{97}.$$

The reason that the last term is S_{97} is that no streak of length 4 can start after 98th flip. Note that for each $1 \leq j \leq 97$, we have

$$\mathbb{P}[S_j = 1] = \frac{2}{16} = \frac{1}{8},$$

since there are exactly two out of 16 outcomes— $TTTT$ and $SSSS$—which lead to streaks. This implies that

$$\mathbb{P}[S] = \mathbb{P}[S_1 + \cdots + S_{97}] = \mathbb{E}[S_1] + \cdots + \mathbb{E}[S_{97}] = \frac{97}{8} = 12.125$$

(b) As it was indicated above, the probability that a coin, when flipped four times, produces a streak of length 4 equals $1/8$.

(c) Let A_1 denote the event that there is a streak starting from the first flip. Clearly $\mathbb{P}[A_1] = 1/8$. Denote by A_2 the event that there is a streak starting from the fifth flip. Note that $\mathbb{P}[A_2] = 1/8$ and moreover A_1 and A_2 are independent, since they involve disjoint set of flips. More generally, we define A_j, for $1 \leq j \leq 25$ to be the event that there is that flips $4j-3, 4j-2, 4j-1, 4j$ constitute a streak. We know

(a) For each $1 \leq j \leq 20$, we have $\mathbb{P}[A_j] = \frac{1}{8}$, and hence $\mathbb{P}\left[A_j^c\right] = \frac{7}{8}$.

(b) Events $A_1, \ldots, A_{25}$ involves disjoint set of flips and hence are independent.

It follows that that none of these streaks takes place is equal to

$$\mathbb{P}[A_1^c \cap \cdots A_{25}^c] = \mathbb{P}[A_1^c] \cdots \mathbb{P}[A_{20}^c] = \left(\frac{7}{8}\right)^{20} \approx 0.354 < 0.04.$$

This implies that the probability that there are no streaks is less than 0.04.

Remark: Note that the actual probability that there are no streak could be much less than 0.04. In the argument above, we only consider streaks starting from $1, 5, 9, \ldots$ flips. Considering other streaks lead to many dependent events, which makes the problem more complicated.

(2.4) Consider a discrete random variable X that with the PMF given by

$$p_X(x) = \begin{cases} k|x| & \text{if } x = -2, -1, 1, 2, 3 \\ 0 & \text{otherwise} \end{cases}$$

(a) Compute the value of k.
(b) Find $\mathbb{P}[X > 1]$.
(c) Find the PMF of the random variables $Y = X^2$ and $Z = X^3$.

Solution. (a) We have

$$1 = \sum_{x \in \{-2,-1,1,2,3\}} k|x| = 9k.$$

From here it follows that $k = \frac{1}{9}$.

(b) We have

$$\mathbb{P}[X > 1] = \mathbb{P}[X = 2] + \mathbb{P}[X = 3] = \frac{2}{9} + \frac{3}{9} = \frac{5}{9}.$$

(c) Note that $Y = X^2$ takes values $1, 4, 9$, and

$$\mathbb{P}[Y = 1] = \mathbb{P}[X = 1] + \mathbb{P}[X = -1] = \frac{1}{9} + \frac{1}{9} = \frac{2}{9}.$$

Similarly,

$$\mathbb{P}[Y = 4] = \mathbb{P}[X = 2] + \mathbb{P}[X = -2] = \frac{2}{9} + \frac{2}{9} = \frac{4}{9}.$$

And, finally, $\mathbb{P}[Y = 9] = \frac{3}{9}$.

x	1	4	9
$p_Y(x)$	$\frac{2}{9}$	$\frac{4}{9}$	$\frac{3}{9}$

Similarly, $Z = X^3$ takes values $-8, -1, 1, 8, 27$. The probabilities of these values correspond exactly to the probabilities of $-2, -1, 1, 2, 3$. For instance, $\mathbb{P}[Z = 27] = \mathbb{P}[X = 3] = \frac{3}{9}$, etc. We can represent the answer in the following table

x	-8	-1	1	8	27
$p_Z(x)$	2/9	1/9	1/9	2/9	3/9

(2.5) Suppose X is a random variable with a geometric distribution with parameter p.
(a) Compute $\mathbb{P}[X > k]$ for $k = 0, 1, 2, \ldots$.
(b) Show that for all $n, k > 0$, we have

$$\mathbb{P}[X = n + k | X > k] = \mathbb{P}[X = n].$$

Solution. There are two different ways of computing $\mathbb{P}[X > k]$. One can proceed directly: we have

$$\mathbb{P}[X > k] = \sum_{j=k+1}^{\infty} \mathbb{P}[X = j] = \sum_{j=k+1}^{\infty} p(1-p)^{j-1} = p(1-p)^k \sum_{j=0}^{\infty} (1-p)^j = \frac{p(1-p)^k}{p} = (1-p)^k.$$

Alternatively, we can use the fact that the geometric distribution is the distribution of the first success in a sequence of Bernoulli trials (such as flipping a coin) where the probability of success in each round is p. Hence $X > k$ if and only if the first k rounds lead to failure which has probability $(1-p)^k$.

(b)

$$\begin{aligned} \mathbb{P}[X = n + k | X > k] &= \frac{\mathbb{P}[X = n + k \text{ and } X > k]}{\mathbb{P}[X > k]} = \frac{\mathbb{P}[X = n + k]}{\mathbb{P}[X > k]} \\ &= \frac{p(1-p)^{n+k-1}}{(1-p)^k} = p(1-p)^{n-1} = \mathbb{P}[X = n]. \end{aligned} \tag{3}$$

(2.6) For events $A, B \subseteq \Omega$, we write $A \perp B$ if A and B are independent.
(a) Show that if $A \perp B$ then $A \perp B^c$.
(b) Show that if $A \perp B$ then $A^c \perp B^c$.
(c) Is it true that if $A \perp B$ and $C \perp D$ then $A \cup C \perp B \cup D$?

Solution.

$$\mathbb{P}[A \cap B^c] = \mathbb{P}[A - A \cap B] = \mathbb{P}[A] - \mathbb{P}[A \cap B] = \mathbb{P}[A](1 - \mathbb{P}[B]) = \mathbb{P}[A]\mathbb{P}[B^c].$$

(b) follows from a repeated application of (a): $A \perp B$ implies $A \perp B^c$, which is equivalent to $B^c \perp A$, and this in turn implies $B^c \perp A^c$.

(c) is false. Take $\Omega = \{00, 01, 10, 11\}$, where each point in the sample space has probability $1/4$. Take $A = D = \{00, 01\}$, $B = C = \{00, 11\}$. Then $\mathbb{P}[A] = \mathbb{P}[B] = \frac{1}{2}$ and $\mathbb{P}[A \cap B] = \frac{1}{4}$, hence A and B are independent. Since $C = B$ and $D = A$, we also have that C and D are independent. On the other hand

$$A \cup C = B \cup D = \{00, 01, 11\}$$

and it is clear that $A \cup C$ is not independent from $B \cup D$.