Assignment 2

CSCI2244-Randomness and Computation
Due Wednesday, January 31, at 11:30PM

These problems concern discrete probability spaces, and are based on the material in Sections 2.1 and 2.2 of the textbook. The cards problem and the last of the dice problems require you to use some counting techniques to find the sizes of some sets, but it is rather simple counting–little more than the rules for sums and products, and nothing about permutations and combinations.

Technically, the durations of the dice games described in the last two problems are 'random variables' on the respective sample spaces, but we don't really need this term now. The third coin problem and the second dice problem require some essential knowledge about infinite geometric series—this is covered in the appendix. The last dice problem requires some coding and plots.

1 Coins

We toss a coin three times in succession, and model the outcomes by the sample space

$$S = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}.$$

(a). Describe *succinctly in words* the events specified by the subsets of S given in (i,ii) below, and, as sets, the events described in words (iii,iv) below. Determine the probabilities of each of these four events, under the assumption of a uniform PMF.

(i) $E = \{HHH, HHT, HTH, HTT\}$

(ii) $E = \{HHH, TTT\}$

(iii) Exactly two heads.

(iv) At least one tail.

(b). We perform the same experiment as above, and model it with the same sample space, only this time we will change the probability function, because *the coin is magical and remembers what it did on the last toss*. On the first toss, heads and tails are equally likely, but on each subsequent toss, the result is twice as likely to be different from the preceding toss as it is to be the same. For example, the probability of the event 'the first two tosses are heads', *i.e.,* $\{HHT, HHH\}$ is half that of 'the first toss is heads and the second is tails', *i.e.,* $\{HTT, HTH\}$. Since the union of these two disjoint events is 'the first toss is heads', these two probabilities add to $1/2$, and thus

$$P(\{HHT, HHH\}) = 1/6, P(\{HTT, HTH\} = 1/3.$$

(a) Determine the PMF underlying this model—that is, determine the probability of each of the 8 individual outcomes.

(b) Use this to determine the probability of each of the events of parts (i-iv) of the preceding question.

(c). Now consider the experiment where we toss a fair coin repeatedly until we get heads. As we discussed in class, the underlying sample space is

$$S = \{1, 2, 3, \ldots\},$$

giving the duration of the game, and the PMF is

$$P(i) = 2^{-i}.$$

Determine the probabilities of the following events. The answer to each question should be a number, backed up by an explanation and a calculation showing how you got the number.

(i) The game lasts no more than three rounds.

(ii) The game lasts at least 9 rounds.

(iii) The game lasts an odd number of rounds.

2 Cards

We draw two cards in succession from a shuffled deck. There are two versions of this game: In one case, we replace the card we draw in the deck before we draw the second card (sampling with replacement); in the second, we keep the first card out of the deck when we draw the second card (sampling without replacement). In both cases, we model the outcome as an ordered pair of values in the range $\{1, \ldots, 52\}$. The respective sample spaces are

$$S_1 = \{(i, j) : i, j \in \{1, \ldots, 52\}\},$$

for sampling with replacement, and

$$S_2 = \{(i, j) : i, j \in \{1, \ldots, 52\}, i \neq j\},$$

for sampling without replacement.

In both cases, we have no reason for supposing that any one ordered pair is more likely than any other, so we model both versions of the experiment with a uniform PMF. Answer the following questions, providing careful justification in each case.

(a) Determine $|S_1|$ and $|S_2|$.

(b) Determine (in both models) the probability of the event 'both cards are face cards'. (There are 12 face cards in a standard deck, 4 Jacks, 4 Queens, and 4 Kings.)

(c) Determine (in both models) the probability of the event 'the two cards are identical'.

(d) Determine (in both models) the probability of the event 'the two cards have different ranks'. (There are 13 different ranks, each with 4 cards.)

3 Dice

(a) We roll a fair die two times in succession. As usual, we model the set of outcomes by the sample space

$$S = \{1, 2, 3, 4, 5, 6\} \times \{1, 2, 3, 4, 5, 6\},$$

with a uniform probability distribution. For the two parts of the problem below, write each of the two events (described in English) explicitly as a set of outcomes. Then determine if the two events are independent.

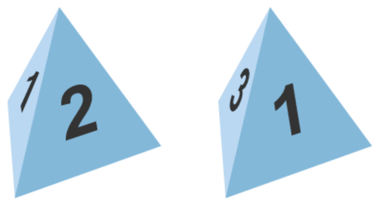

Figure 1: Two views of a tetrahedral die that has landed with its ‘4’ face down.

(i) E_1 : The number on the first die is odd. E_2 : The number on the second die is even.

(ii) E_1 : The number on at least one of the two dice is odd. E_2 : The number on at least one of the two dice is even.

The next two problems concern a fair die with four faces. The problem could just as easily be done with a standard 6-faced die, but using only four faces makes the calculations somewhat easier. [1]

(b) Roll the die repeatedly until the first number that was rolled reappears. This game must last at least two rolls (for example, with an outcome like $(2, 2)$), but there is no upper bound to the number of rounds it might last—for example, you might get an outcome like

$$(4, 3, 3, 1, 2, 3, 1, 1, 2, 4),$$

for a game that lasts 10 rolls. For each $i \geq 2$, find the probability p_i that the game lasts exactly i rolls, and verify that

$$\sum_{i=2}^{\infty} p_i = 1.$$

(HINT: Under the assumption that the die is fair and that the outcomes of individual rolls are independent of one another, each individual outcome of

[1] You don’t need to know this, but the die itself is a regular tetrahedron, with the faces labeled by numbers 1,2,3,4. A tetrahedral die rolled on a table has a ‘down’ face, but no ‘up’ face, so the outcome of a roll is the number that’s *not* showing. The figure shows two views of a die that has been rolled to give an outcome of 4. The tetrahedral dice used by aficionados of Dungeons and Dragons have a more elaborate labeling system to get around this issue.

length i occurs with probability 4^{-i}, so the problem reduces to counting the number of such outcomes.)

(c) Now let's change the game, so that it ends whenever any number that was already rolled reappears. Once again, the game must last at least 2 rounds (*e.g.,* with outcome $(2,2)$) but it cannot last longer than 5 rounds (*e.g.,* $(4,3,1,2,1)$). Solve the problem given in (b) above for this game—that is, find for each $i = 2,\ldots,5$ the probability p_i that the game lasts exactly i rounds. As a reality check, make sure that your probabilities add up to 1.

(d) Simulate the game in (c) for 500 rolls and draw, as a stem plot, the histogram showing the proportion of times that the game lasted for 2,3,4,5 tosses. On the same set of axes, draw the stem plot giving the four probabilities you determined in (c). (To superimpose the two plots, you can, for example, draw one set of stems with x-coordinates 2,3,4,5, and the other with x-coordinates 2.1,3.1,4.1,5.1.)

4 Appendix: Geometric series

Let r be a real number. A *geometric series* with base r is a sum

$$\sum_{j=0}^{n} r^j = 1 + r + r^2 + \cdots + r^n.$$

If $r \neq 1$ then there is a closed-form formula for this sum:

$$\sum_{j=0}^{n} r^j = \frac{1 - r^{n+1}}{1 - r}.$$

If $|r| < 1$ then

$$\lim_{n\to\infty} r^n = 0,$$

so the infinite geometric series converges to the limit

$$\sum_{j=0}^{n} r^j = \frac{1}{1 - r}.$$

Here are a few examples of how to apply this identity. With $r = \frac{1}{2}$ the infinite series is

$$1 + \frac{1}{2} + \frac{1}{4} + \cdots = \frac{1}{1 - \frac{1}{2}} = 2.$$

Now consider the coin-tossing game in which we toss a fair coin repeatedly until it comes up heads. The sample space is $S = \{1, 2, \ldots\}$ and the PMF is given by

$$P(i) = \frac{1}{2^i},$$

so the sum of the values of the PMF is

$$\begin{aligned} \sum_{i \in S} P(i) &= \sum_{i=1}^{\infty} \frac{1}{2^i} \\ &= \frac{1}{2} \cdot \sum_{i=0}^{\infty} \frac{1}{2^i} \\ &= \frac{1}{2} \cdot 2 \\ &= 1, \end{aligned}$$

so this satisfies the requirements for a probability space. More generally, if we play the same game with a possibly biased coin in which the probability of heads is $p > 0$, then the PMF is given by

$$P(i) = q^{i-1} p,$$

where $q = 1 - p$. This is because the game lasts for i turns if and only if the first i tosses are

$$\underbrace{T \cdots T}_{i \text{ times}} H.$$

So now the sum of all values of the PMF is

$$\begin{aligned} \sum_{i=1}^{\infty} q^{i-1} p &= p \sum_{i=0}^{\infty} q^i \\ &= p \cdot \frac{1}{1 - q} \\ &= p \cdot \frac{1}{p} \\ &= 1, \end{aligned}$$

as required.