Assignment 2

CSCI2244-Randomness and Computation
Due Wednesday, January 30, 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 and the notes for Lectures 1 and 2. Problems 2 and 3(a,b) require you to use 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. Problems 1(c) and 3(a) require you to use the material on infinite geometric series from the notes for Lecture 1. Problem 3(c) requires you to do some coding and draw plots.

Technically, the durations of the two dice games described in Problem 3 are 'random variables' on the respective sample spaces, but we don't really need this term now.

If you're paying attention you'll notice that a lot of Problems 1 and 2, or simple variations thereof, were solved in class and in the notes, so this is just reinforcement learning. But you'll have to think a bit harder for Problem 3.

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

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

(i) Determine the PMF underlying this model—that is, determine the probability of each of the 8 individual outcomes. (You can continue the analysis above, or reason from a tree diagram, but be sure to explain how you arrived at the result, don't just give the answer.)

(ii) 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

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

for all $i \in S$. 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.

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

(ii) 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.)

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

(iv) Determine (in both models) the probability of the event 'the two cards are in the same suit'. (There are 4 different suits, each with 13 cards.)

3 Dice

This problem concerns 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]

[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

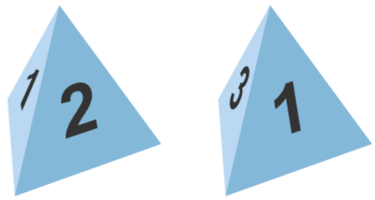

Figure 1: Two views of a tetrahedral die that has landed with its '4' face down.

(a) 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 length i occurs with probability 4^{-i}, so the problem reduces to counting the number of such outcomes for which the game lasts exactly i rounds. Just to get you started—if $i = 2$, there are exactly four such outcomes: $(1, 1), (2, 2), (3, 3), (4, 4)$, so the probability that the game lasts exactly 2 rounds is $4 \times 4^{-2} = 1/4$..)

(b) 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

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.

exactly i rounds. As a reality check, make sure that your probabilities add up to 1.

(c) Simulate the game in (b) 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.)