Lab 2: Throwing dice

due date: Friday, Oct 4 in class

Introduction

This week you will complete a short R lab that involves rolling dice and computing sums, and tossing coins. Again and again and again. This is why we use R. Each question is worth 5 points.

You will be graded both on correctness of your answers and on the presentation of your answer. This means you need to *explain* your methodology well.

A list of useful R commands is included as the last page of this handout. You will need to modify the commands to your purposes. A copy of this file is available on the class webpage, if you want to use something like WORD or LaTeX for your write up, or to cut and paste commands.

Problems

1. In this problem, you will explore the sum of the roll of three dice. For sample sizes, try $n = 10, 25, 50, 100$ or other values n of your choice. Suppose you roll the three dice and sum the outcomes repeatedly. If you were to earn a \$1 for the sum, for instance, then a roll of 1,1,1 earns you \$3.

   How much do you expect to earn in an average game? Carefully, explain your answer and methodology below.

   The best answer you could give is $3(3.5) = 10.5$.

   There are many, many ways to arrive at this answer, and I will outline one below. Define the random variables

   $$X: \text{ the roll of a single die,}$$
   $$Y: \text{ the sum of the roll of 3 die.}$$

   The random variable of interest, Y, is a function of X; namely, $Y = X_1 + X_2 + X_3$, where X_i indicates the ith roll. Because each of the three rolls are independent, but identical trials of the same process, we have $E(Y) = E(X_1 + X_2 + X_3) = 3E(X)$. Either directly, or by toying around, it is possible to figure out that $\mathbb{E}(X) = 3.5$. Then by linearity

   $$\mathbb{E}(Y) = \mathbb{E}(3X) = 3 \cdot \mathbb{E}(X) = 3 \cdot (3.5) = 10.5.$$

   Depending on how much you worked with your die roll function, you could have figured this out from being an 'frequentist' too. That is, if you perform the experiment over and over and over again, and averaged the sum of the three rolls, you should get a number quite close to 10.5.

2. You toss a *biased* coin repeatedly, with the probability of heads, $P(H) = p$, for some unknown value of p. Your goal is to give the best estimate of p that you can. To that end, suppose you flip this coin three times and that you earn \$3 for each H that comes up, and lose \$1 for each T that comes up.

There are 3 data sets available for you. They are called *earningsN* where $N = 10, 100, 1000$. Again, your task is to give the best estimate for p you can give from these datasets and explain how you arrived at your answer.

I estimate the value of p to be <u>.238</u> because .....

SOLUTION 1: One of many ways to do this problem would be along the lines of the following. Using the largest dataset, since it includes the most trials of the experiment, count the number of H. We will repeatedly use the fact that each flip of the coin was independent of all the others.

For the counting, notice that an earnings of of \$9 corresponds to 3 H, an earnings of \$5 corresponds to 2 H, and an earnings of 1 corresponds to 1 H (for HTT, THT, or HHT); and an earnings of -\$3 corresponds to 0 H. Using the dataset of size 1000, we tabulate the following:

	Earnings		
	\$ 9	\$ 5	\$ 1
Number	10	131	423

Again, using independence, a good estimate for p would be

$$\frac{10 \cdot (3) + 131 \cdot (2) + 423 \cdot (1)}{3(1000)} = \frac{715}{3000} \approx .238$$

Note that the denominator is $3 \cdot 1000$ since (because of independence) we are counting the total number of coin flips.

SOLUTION 2: First, define the random variable for your earnings to be

$$X: \text{ earnings for three flips of biased coin.}$$

Then X is a discrete random variable with four possible values of x so that $P(X = x) > 0$. These values are $x = -3, 1, 5, 9$ corresponding to rolls of zero H, one H, two H, and three H. The probability function for X is

$$p(x) = \begin{cases} (1-p)^3, & \text{if } x = -3, \\ 3p(1-p)^2, & \text{if } x = 0, \\ 3p^(1-p), & \text{if } x = 1, \\ p^3, & \text{if } x = 9. \end{cases}$$

The expected value of your earning $\mathbb{E}(X)$ is given by

$$\begin{aligned} E(X) &= -3 \cdot (1-p)^3 + 1 \cdot 3p(1-p)^2 + 5 \cdot 3p^2(1-p) + 9 \cdot p^3 \\ &= \quad 12p - 3. \end{aligned}$$

Using the data, we can compute that the (sample) mean earnings is -\$.14. Equating this with the expected value and solving, we find

$$12p - 3 = -.14 \text{ which implies that } p \approx .238,$$

exactly the same answer as above.