1. Fill in the blanks: we saw in class that, if you roll a *single* fair, six-sided die, and let X be the number that comes up, then $E[X] =$ ___3.5___. So, now, suppose you roll *two* fair, six-sided dice. Let X_1 be number that comes up on the first die, X_2 the number that comes up on the second die, and X the sum of these two numbers. Then, $X = X_1 + X_2$ so, by the ___sum___ rule for expected value,

$$E[X] = E[X_1] + \underline{\quad E[X_2] \quad} = \underline{\quad 3.5 \quad} + 3.5 = \underline{\quad 7 \quad}.$$

Now let's consider the sum of two dice from a perspective that blends an empirical point of view with a theoretical point of view: instead of actually rolling two dice repeatedly, we'll think about what we'd expect to happen if we did this.

2. For this problem you want to recall, as we did in class, that if X is (again) the sum showing on two fair, six-sided dice, then

$$P(X=2) = \frac{1}{36}, \quad P(X=3) = \frac{2}{36}, \quad P(X=4) = \frac{3}{36}, \quad P(X=5) = \frac{4}{36},$$
$$P(X=6) = \frac{5}{36}, \quad P(X=7) = \frac{6}{36}, P(X=8) = \frac{5}{36}, \quad P(X=9) = \frac{4}{36},$$
$$P(X=10) = \frac{3}{36}, \quad P(X=11) = \frac{2}{36}, \quad P(X=12) = \frac{1}{36}.$$

Given these probabilities, imagine that you roll your two fair dice 3600 times. Fill in, in the table below, the number of times out of these 3600 that you would expect the sum to equal the given number (assuming things played out exactly according to their probabilities). For example, in the blank space to the right of the "4" in the "Sum on dice" column, write in how many times you would expect to see a sum of 4, if you rolled the dice 3600 times.

Sum on dice	Expected number of times	Sum on dice	Expected number of times
2	100	8	500
3	200	9	400
4	300	10	300
5	400	11	200
6	500	12	100
7	600		

3. What is the mean (average), call it $\overline{x}$, of the 3600 observations (data points) represented by the above data? Recall that, to find the mean of a set of numerical data points, you add them all up, and divide by however many data points you have. Hint: group your data, rather than trying to add up 3600 separate numbers. That is: if the number 5 occurs 800 times (for example), then instead of adding up 800 5's as part of your calculation of the mean, just add in $5 \cdot 800$.

Answer:

$$\overline{x} = \frac{\begin{pmatrix} 2\cdot 100 + 3\cdot 200 + 4\cdot 300 + 5\cdot 400 + 6\cdot 500 + 7\cdot 600 \\ +8\cdot 500 + 9\cdot 400 + 10\cdot 300 + 11\cdot 200 + 12\cdot 100 \end{pmatrix}}{3600} = 7.$$

4. How does your mean $\overline{x}$ compare with your expected value $\mu = E[X]$ from problem 1?

They're the same.

5. Fill in the blanks: Suppose the result of some experiment is a random variable X, and let x denote a generic value of X. (In the above thought experiment of rolling two dice, x can equal any one of the numbers __2__ through __12__.)

Now suppose this experiment repeated n times. (For example, in the above thought experiment, $n =$ __3600__.) Suppose that f_x denotes the number of times X actually takes the value x, in these n trials. (For example, in the above thought experiment, $f_2 = 100, f_3 =$ __200__, $\ldots, f_7 =$ __300__, $\ldots f_{12} =$ __100__.)

Then, by grouping together data points that have the same value x, we find that the mean, or average, of the n data values of X is

$$\overline{x} = \sum_{\substack{\text{values} \\ x \text{ of } X}} \frac{x \cdot \underline{\quad f_x \quad}}{n}. \qquad (*)$$

However note that, since f_x is the *number* of times X takes the value x, out of n trials, we see that f_x/n is just the *proportion* of the time, out of the n trials, that X takes the value __x__. If the experiment plays out according to probabilities, we would expect this proportion to equal the *probability* of X equalling x; that is, we would expect f_x/n to equal $P(X =$ __x__$)$.

But under these assumptions, $(*)$ gives us

$$\overline{x} = \sum_{\substack{\text{values} \\ x \text{ of } X}} x \cdot P(X = \underline{\quad x \quad}).$$

And the right hand side is exactly the formula for the __expected__ __value__ of X!

Conclusion: if an experiment, repeated many times, plays out according to its probabilities, then the average, or __mean__, of the data obtained is just the __expected__ __value__ of that data.

6. To finish, roll your two dice and put a tally mark in the appropriate space below. Repeat this many times. Then compute the average sum that comes up, for your many trials. Hint: as in problem 3 above, group your data.

Sum on dice	Tally	Sum on dice	Tally
2	21	8	78
3	35	9	56
4	40	10	53
5	68	11	28
6	87	12	21
7	92		

Answer: $\overline{x} = 7.6635$