

Student Name: Raymond Shum
Due Date: 1-25-2022
CST 383 - Intro to Data Science
Dr. Glenn Bruns

Lab: sample spaces and events

1. What's the probability that a roll of one die will give a 5? (use naive probability).
 - $1/6$
2. What's the probability that a roll of two dice will give a sum of 5?
 - $P(\{1,4\}) + P(\{4,1\}) + P(\{2,3\}) + P(\{3,2\}) = 4/36 = 1/9$
3. A wooden cube with painted faces is sawed up into 1000 little cubes, all the same size. Then one little cube is chosen randomly. What's the probability the little cube is painted on exactly two sides?
 - Cube has dimensions: $10 \times 10 \times 10 = 1000$
 - 12 edges per cube
 - 10 sub-cubes per edge (2 with 3 painted faces & 8 with 2 painted faces)
 - # cubes with 2 painted faces = $12 \text{ edges} \times 8 \text{ cubes/edge} = 96 \text{ cubes}$
 - $96/1000 = .096$
4. What are the two components of a probability space?
 - Sample Space: S
 - Probability Function: P
5. I have a probability space. The sample space S has three outcomes: a , b , and c . What is the total number of events in this sample space?
 - $\{a, b, c\}, \{a,b\}, \{b,c\}, \{a,c\}, \{a\}, \{b\}, \{c\}$
 - 7
6. Let P be the probability function of the probability space of the previous problem. If $P(\{a,b\}) = 0.6$ then what is $P(\{c\})$?
 - $1 - 0.6 = 0.4$
7. Suppose we draw two cards from a shuffled deck of 4 cards marked 1-4.
 - o what is the sample space if the first card is put back in the deck after it is drawn?
 $S = \text{a pair of cards from 1-4, minus one of the possible numbers}$
 - o what is the sample space if the first card is not put back?
 $S = \text{a pair of cards from 1-4 minus 2 of the possible numbers}$

8. In Python, simulate 10,000 rolls of a die, and calculate the fraction of those rolls that are 5. Use NumPy function `numpy.random.randint()`.

```
NumPy.random.randint(low=1,high=7,size=10000)
```

9. Repeat the previous problem, but this time simulate 10,000 rolls of two dice. In rolling two dice we look at the sum of the dice. Now try again, but this time calculate the fraction of the rolls that add up to 3 (instead of 5). Comment on the values you got for 3 and 5.

```
NumPy.random.randint(low=1,high=7,size=10000) +  
NumPy.random.randint(low=1,high=7,size=10000)
```

Hints:

1. There are six outcomes, equally likely.
2. There are 36 outcomes, equally likely. In how many of the outcomes do the values on the two dice sum to 5?
3. There are 1000 outcomes, equally likely. In how many of the outcomes is a cube painted on exactly two sides? Be careful about the corners, where three sides are painted. The correct way to calculate the answer is to consider the number of non-corner pieces on each edge. There are 12 edges, and 8 non-corner pieces on each edge.
4. A probability space consists of a sample space S and a probability function P .
5. An event is a subset of S . How many subsets of $\{a,b,c\}$ are there?
6. To solve this problem, use the 3 things we know about probability functions. For example, $P(\{\}) = 0$.
7. In the first case, a sample is a pair of numbers (a,b) where each of a and b can range from 1 to 4. It's similar in the second case, but the value of b cannot be the same as a .
8. You can simulate 1000 dice rolls like this:

```
np.random.choice(6, 10000) + 1
```

With this function, the first argument can be a list or a number. In this case, we use a number, and get 10,000 random numbers from 0 to 5. We add 1 to get numbers from 1 to 6.

9. First, separately generate the rolls for each of two dice. Next, add the rolls together. Remember that you can do vectorized arithmetic in NumPy.