

**Statistics - Week 3 Questions**  
Owen Burke, 15316452

February 18, 2019

## Question 1

Say we roll a fair 6-sided die six times. Using the fact that each roll is an independent random event, what is the probability that we roll:

### a) The sequence 1,1,2,2,3,3 ?

Firstly, let's define the events.

Let's say event  $R_1$  is the event that we roll the desired number on the first dice (a 1 in this case).

The same for  $R_2$  (also a 1),  $R_3$  (a 2 for this dice) and so on.

Given that we know the events are independent we can say

$$P(R_1 \cap R_2 \cap R_3 \dots R_6) = P(R_1)P(R_2)P(R_3)\dots P(R_6).$$

from the rule

$$P(E \cap F) = P(E)P(F), \text{ for some } E \text{ and } F.$$

Let's call event  $(R_1 \cap R_2 \cap R_3 \dots R_6)$ , event  $E$  where  $E$  is the event that we roll the required sequence.

From the rule above, we can get this probability by simply multiplying the probabilities of the individual events whose intersections make up event  $E$ .

We already know that the probability of rolling a certain number on a six sided dice is  $\frac{1}{6}$ .

$$\text{So } P(E) = \frac{1}{6} * \frac{1}{6} * \frac{1}{6} * \frac{1}{6} * \frac{1}{6} * \frac{1}{6}$$

Therefore the result is

$$P(E) = \frac{1}{46656}$$

### b) A three exactly 4 times?

The logic behind this answer is to find how many ways the three's can appear in a certain position among the six rolls.

Then multiply that result by how many positions they can hold among the rolls.

So, for example, let's begin by saying the 4 three's appear in the first, second, third and fourth roll. We know that they can appear in this layout in  $1*1*1*1*5*5 = 25$  different ways as the last two rolls must be anything but three's.

We then find how many layouts the 4 three's can occupy. This is simply given by  $\binom{6}{4} = 15$ .

So our result is  $15*25 = 375$ .

Let's call this event, that we roll exactly 4 three's,  $E$ .

So now we know  $\#E = 375$ .

So to get  $P(E)$  we use the formula

$$P(E) = \frac{\#E}{\#S}, \text{ where } S \text{ is the sample space.}$$

We know the sample space size is  $6*6*6*6*6*6 = 46656$  as that's the total number of ways you can roll a dice six times. (We also know it from the previous question).

So, now  $P(E)$  is

$$P(E) = \frac{\#E}{\#S} = \frac{375}{46656}$$

### c) A single 1

This is a similar approach to part b).

Let's say the 1 is the first roll. We know the number of ways it can appear in the first dice is  $1*5*5*5*5*5$  (the others are 5 as they must be anything but a roll of 1).

This  $1*5*5*5*5*5 = 3125$ .

We then multiply this by how many ways a single 1 can appear among 6 rolls, which is given by  $\binom{6}{1} = 6$ .

This results in  $3125*6 = 18750$ .

Call event E the event that we roll a single 1.

The probability of E is given by

$$P(E) = \frac{\#E}{\#S}, \text{ where } S \text{ is the sample space.}$$

giving us, as we know the sample space from the previous part,

$$P(E) = \frac{18750}{46656}$$

### d) One or more 1s

Again, this is a similar process to the previous parts.

However, we must start with one 1, find the result and then add that to the result of two 1's, and three 1's and so on.

So, firstly, for a single 1, we know it's given by

$$\text{One 1} = (1*5*5*5*5*5) * \binom{6}{1} = 5^5 * \binom{6}{1}.$$

And by the same logic for two 1's it is

$$\text{Two 1's} = (1*1*5*5*5*5) * \binom{6}{2} = 5^4 * \binom{6}{2}.$$

This is repeat as so

$$\text{Three 1's} = (1*1*1*5*5*5) * \binom{6}{3} = 5^3 * \binom{6}{3}.$$

$$\text{Four 1's} = (1*1*1*1*5*5) * \binom{6}{4} = 5^2 * \binom{6}{4}.$$

$$\text{Five 1's} = (1*1*1*1*1*5) * \binom{6}{5} = 5^1 * \binom{6}{5}.$$

Six 1's =  $(1*1*1*1*1*1) * \binom{6}{6} = 1 * \binom{6}{6}$ .

We then simply add these results like so

$$(5^5 * \binom{6}{1}) + (5^4 * \binom{6}{2}) + (5^3 * \binom{6}{3}) + (5^2 * \binom{6}{4}) + (5^1 * \binom{6}{5}) + (1 * \binom{6}{6})$$

giving us, if we call event E the event that we roll at least one 1

$$\#E = 31031$$

So, as before

$$P(E) = \frac{\#E}{\#S} = \frac{31031}{46656}$$

## Question 2

Suppose one 6-sided and one 20-sided die are rolled. Let A be the event that the first die comes up 1 and B that the sum of the dice is 2. Are these events independent? Explain using the formal definition of independence.

The formal definition of independence states two events (E and F) are independent if

$$P(E \cap F) = P(E)P(F)$$

So, we need to find  $P(A)$ ,  $P(B)$  and  $P(A \cap B)$  and then determine if the formula above holds true.

We know the probability of event A is  $\frac{1}{6}$  as that is the probability of getting any one number on a six sided dice.

Event B is the event that the two dice sum to 2, and there is clearly only one way this can happen. If each dice lands on a 1.

So  $B = (1,1)$ .

$P(B) = \frac{\#B}{\#S}$  where S is the sample size (in this case, the number of outcomes of rolling both a 6-sided and 20-sided dice).

This is given by  $6*20 = 120 = \#S$ .

So  $P(B) = \frac{1}{120}$ .

$P(A \cap B)$  is the probability that the first die comes up 1 and that they sum to 2. This is clearly the same as  $P(B)$ , as there's no other way the first dice can land on a 1 and the dice sum to 2 other than  $(A \cap B) = (1,1)$ .

So, does the formula hold true?

No, it does not as

$$\left(\frac{1}{120}\right) \neq \left(\frac{1}{120}\right) * \left(\frac{1}{6}\right) \text{ from the definition of independence stated above.}$$

**So the events are not independent.**

### Question 3

Say a hacker has a list of  $n$  distinct password candidates, only one of which will successfully log her into a secure system.

**a) If she tries passwords from the list uniformly at random, deleting those passwords that do not work, what is the probability that her first successful login will be (exactly) on her  $k$ -th try?**

To start with the hacker has  $n$  passwords on her list.

So the probability of her picking the correct one at random will be  $\frac{1}{n}$  for her first go.

However for her subsequent tries, we must take into account the odds of the previous go failing, plus her  $k$ -th try being successful.

This gives the series  $\frac{n-1}{n} * \frac{n-2}{n-1} * \frac{n-3}{n-2} \dots \frac{1}{n-k+1}$  for some  $n$  and  $k$ .

The denominator (as well as the numerator) is being reduced as the list of passwords is made smaller after an incorrect guess.

This series gives the odds of all previous tests failing and then the odds of the  $k$ -th one being successful.

This can also be given :  $\frac{1}{(n-k)+1} * \prod_{i=0}^{i < (k-1)} \frac{n-(i+1)}{n-i}$

**b) When  $n=6$  and  $k=3$ , what is the value of this probability ?**

We simply insert these values into the series from part a).

This is given as

$$\left(\frac{1}{(6-3)+1}\right) * \left(\frac{6-(0+1)}{6-0}\right) * \left(\frac{6-(1+1)}{6-1}\right) = \frac{1}{6}$$

**c) Now say the hacker tries passwords from the list at random, but does not delete previously tried passwords from the list. She stops after her first successful login attempt. What is the probability that her first successful login will be (exactly) on her  $k$ -th try?**

This scenario is similar to the previous one above, but somewhat different. Given that she does not delete previous passwords, we will not obtain a series.

So we must get her odds of guessing incorrectly  $k$  times from a constant list of  $n$  items plus one correct guess.

Her correct guess is simply  $\frac{1}{n}$  as you would expect when selecting randomly from a list of  $n$  items.

Her incorrect guesses are given as  $\left(\frac{n-1}{n}\right)^{k-1}$ .

This says for all her previous  $(k-1)$  tries, the odds of her guessing incorrectly

are all the wrong passwords  $(n-1)$  over the whole list  $(n)$ .  
So the resulting formula is

$$\left(\frac{n-1}{n}\right)^{k-1} * \left(\frac{1}{n}\right)$$

**d) When  $n= 6$  and  $k= 3$ , what is the value of this probability ?**

We simply insert these values into the formula from part c).  
This is given as

$$\left(\frac{6-1}{6}\right)^{3-1} * \left(\frac{1}{6}\right) = \frac{25}{216}$$

## Question 4

A website wants to detect if a visitor is a robot. They decide to deploy three CAPTCHA tests that are hard for robots and if the visitor fails in one of the tests, they are flagged as a possible robot. The probability that a human succeeds at a single test is 0.95, while a robot only succeeds with probability 0.3. Assume all tests are independent.

**a) If a visitor is actually a robot, what is the probability they get flagged?**

We know that the probability that a robot succeeds in a single test is 0.3.  
So the probability that a robot succeeds all three tests =  $0.3 * 0.3 * 0.3$ , as all the tests are independent.  
It also applies, that the probability of a robot being flagged is 1 - (the probability they pass all tests).  
This is simply

$$1-(0.3)^3 = 0.973$$

**b) If a visitor is human, what is the probability they get flagged?**

We know that the probability that a human succeeds in a single test is 0.95.  
So the probability that a human succeeds all three tests =  $0.95 * 0.95 * 0.95$ , as all the tests are independent.  
It also applies, that the probability of a human being flagged is 1 - (the probability they pass all tests).  
This is simply

$$1-(0.95)^3 = 0.142625$$

c) The fraction of visitors on the site that are robots is 1/10. Suppose a visitor gets flagged. What is the probability that visitor is a robot? Hint: use Bayes Rule.

Bayes rule states

$$P(E | F) = \frac{P(F|E)P(E)}{P(F)}, \text{ for some E and F.}$$

Given that we know the visitor gets flagged in this scenario, we want to know the probability it is a robot.

This is the same as  $P(R|F)$ , where R is the event that the visitor is a robot and F is the event that the visitor is flagged.

From Bayes rule this is given by  $\frac{P(F|R)P(R)}{P(F)}$ .

We know  $P(F | R)$  is 0.973 from part a) (This is the probability of being flagged given it is a robot).

We also know  $P(R)$  from the question ("The fraction of visitors on the site that are robots is 1/10") is 0.1.

So we must find  $P(F)$ .

$P(F)$  is the same as 1 - (the probability of not being flagged at all).

The probability of not being flagged / passing all tests is

((the probability of a human passing) \* (the probability it is a human)) + ((the probability of a robot passing) \* (the probability it is a robot)) which is  $((0.95)^3 * 0.9) + ((0.3)^3 * 0.1)$  as we know the probability it is a human is 1 - (the probability it is a robot) = 0.9.

$$\text{This gives } P(F) = 1 - 0.7743375 = 0.2256625$$

By applying this to Bayes rule, we get

$$P(R | F) = \frac{(0.973)P(0.1)}{0.2256625} = \mathbf{0.431174874}$$