<u>ST3009 - Statistical Methods for Computer Science</u>
<u>Week 3 Questions</u>

Brandon Dooley - #16327446

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) Since each roll is an independent event we consider the probability of each event individually for each occurrence. The probability of rolling a specific number on a six sided die is:

$$P(E_1) = \frac{1}{6}$$

As a result, the probability of rolling the sequence 1,1,2,2,3,3 – P(E) is:

$$P(E) = P(E_1) \cap P(E_2) \cap P(E_3) \cap P(E_4) \cap P(E_5) \cap P(E_6)$$

$$P(E) = \left(\frac{1}{6}\right)^6 = 0.00002$$

(b) The probability of exactly four occurrences of 3 can be calculated by calculating all possible outcomes of exactly four 3's and divide this by the total sample space. This can be done as follows:

$$P(E) = \frac{|E|}{|S|}$$

The total number of outcomes containing exactly k occurrences of 3 in n rolls - |E| - can be calculated as follows:

$$|E| = \binom{6}{4} * 5^2 = 375$$

The total number of outcomes over 6 rolls of 6-sided die can be calculated as follows:

$$|S| = 6^6 = 46656$$

Thus, the probability of exactly four occurrences of 3 over 6 rolls can be calculated as:

$$P(E) = \frac{|E|}{|S|} = \frac{375}{46656} = 0.0080$$

(c) The probability of rolling a single one is the probability of a one occurring in one trial and in no others. Similar to above this can be described as:

$$P(exactly\ k\ times\ over\ n\ trials) = \binom{n}{k}\, p(k) * (1-p)^{n-k}$$

$$P(exactly\ 1\ time\ over\ 6\ trials) = \binom{6}{1}\, \frac{1}{6} * (1-\frac{1}{6})^{6-1}\ = 0.4019$$

(d) The probability of one or more 1's occurring is the probability that at least one 1 is rolled. This can be calculated as:

$$1 - P(no\ ones\ are\ rolled)$$

The probability of no ones being rolled is the probability that on each roll any number besides one occurs. This can be calculated as:

$$P(no\ ones\ are\ rolled) = \left(\frac{5}{6}\right)^{6} = 0.3349$$

We can then use this probability to calculate the probability that at least one 1 is rolled:

$$1 - P(no\ ones\ are\ rolled)$$

$$= 1 - 0.3349$$

$$= 0.6651$$

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.*

First we must calculate P(A) and P(B). The probability of A, P(A) is the probability of a 1 occurring on the first roll (we don't know whether this is a 6-sided or 20-sided die. Thus:

$$P(A) = \frac{1}{6} + \frac{1}{20} = 0.2167$$

To calculate P(B) investigate the possible outcomes that can produce a sum of 2. This occurs when both die roll a 1 so there are two possible ways in which this can occur:

$$P(B) = \left(\frac{1}{6} * \frac{1}{20}\right) + \left(\frac{1}{20} * \frac{1}{6}\right) = \ 0.0167$$

Two events are independent if the order in which they occur does not matter. Therefore two events E and F are independent if they satisfy:

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

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

P(A ∩ B) is the probability that the first die comes up a 1 and that the sum of both dies is 2. This can be calculated as follows:

$$P(A \cap B) = P(B|A)P(A)$$

$$P(B|A) = \frac{1}{6} + \frac{1}{20} = 0.2167$$

$$P(A \cap B) = 0.2167 * 0.2167 = 0.0469$$

If the two events are independent the above result for P(A ∩ B) should be equal to P(A) * P(B):

$$0.0469 = 0.2167 * \ 0.0167$$

$$0.0469 \neq 0.0036$$

Therefore the two events A and B are not independent of each other and are in fact dependent.

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) The probability of the hacker choosing a successful password is:

$$P(successful\ password) = \frac{1}{n}$$

However, as the hacker unsuccessfully chooses a password candidate from n she removes this candidate from n this reducing n by one on every attempt.

The probability that her first successful login will be (exactly) on her k-th attempt is the probability that she chooses the wrong option every time until the kth time. Let the random variable X describe the k-th attempt and P(X = k) be the probability that the hacker is successful on her k-th attempt. Therefore:

$$P(X = 1) = \frac{1}{n}$$

$$P(X = 2) = \frac{n-1}{n} * \frac{1}{n-1} = \frac{1}{n}$$

$$P(X = n) = \frac{n-1}{n} * \frac{n-2}{n-1} * \dots * \frac{1}{2} = \frac{1}{n}$$

$$P(X) = \frac{(n-1)!}{n!} = \frac{(n-1)!}{(n-1)! * n} = \frac{1}{n}$$

Therefore the probability of the hacker being successful on her k-th attempt is 1/n if she is deleting the passwords as she progresses.

(b) With n = 6 and k = 3 the value of this probability is as follows:

$$P(success\ on\ 3rd\ attempt) = \frac{5}{6} * \frac{4}{5} * \frac{1}{4} = 0.1667$$

(c) If the hacker was to stop deleting the candidates that she has already tried from the potential options n the probability of her succeeding on her k-th attempt would be:

$$P(success\ on\ kth\ attempt) = \frac{n-1}{n} * \frac{n-1}{n} * \dots * \frac{1}{n}$$

$$P(success\ on\ kth\ attempt) = \left(\frac{n-1}{n}\right)^{k-1} * \frac{1}{n}$$

(d) Now, with n = 6 and k = 3, the probability of the hacker succeeding on her k-th attempt is:

$$P(success\ on\ 3rd\ attempt) = \left(\frac{5}{6}\right)^{2} * \frac{1}{6} = 0.1157$$

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.*

- **P(Human Succeeds Single Test) =** 0.95
- **P(Robot Succeeds Single Test) =** 0.3

(a) Let P(RF) be the probability that a robot is flagged and P(RF') be the probability that a robot is not flagged. The probability that a visitor gets flagged given that they are a robot can be calculated as follows:

$$P(RF) = 1 - P(RF')$$

Given that there are three CAPTCHA tests the probability that a robot is not flagged can be calculated by:

$$P(RF') = P(RF'_{test1}) \cap\ P(RF'_{test2})\ \cap\ P(RF'_{test3})$$

$$P(RF') =\ 0.3 * 0.3 * 0.3$$

$$P(RF') = 0.027$$

Therefore, using this we can calculate the probability that a visitor is flagged that it is actually a robot:

$$P(RF) = 1 - P(RF')$$

$$P(RF) = 1 - 0.027 = 0.973$$

From this we can determine that there is a 97.3% chance of a visitor being flagged if it is a robot.

(b) Let P(HF) be the probability that a human is flagged and P(HF') be the probability that a human is not flagged. The probability that a visitor gets flagged given that they are a human can be calculated as follows:

$$P(HF) = 1 - P(HF')$$

Given that there are three CAPTCHA tests the probability that a robot is not flagged can be calculated by:

$$P(HF') = P(HF'_{test1}) \cap\ P(HF'_{test2})\ \cap\ P(HF'_{test3})$$

$$P(HF') =\ 0.95 * 0.95 * 0.95$$

$$P(HF') = 0.8574$$

Therefore, using this we can calculate the probability that a visitor is flagged that it is actually a human:

$$P(HF) = 1 - P(HF')$$

$$P(HF) = 1 - 0.8574 = 0.1426$$

From this we can determine that there is a 14.26% chance of a visitor being flagged if it is a human.

(c) The probability that a visitor is a robot given that it has been flagged can be calculated using Bayes Rule as follows:

$$P(R|F) = \frac{P(F|R)P(R)}{P(F)}$$

- **P(R|F):** Probability that visitor is a robot given flagged = unknown
- **P(F|R):** Probability that visitor is a flagged given robot = 0.973 (part a)
- **P(R):** Probability of visitor being a robot = 0.1 (given)
- **P(F):** Probability of a visitor being flagged (regardless if robot or human)

We can calculate P(F) as follows:

$$P(F) = P(F|E)P(E) + P(F|E')P(E')$$

$$P(F) = P(FR)P(R) + P(HF)P(H)$$

$$P(F) = (0.973 * 0.1) + (0.1426 * 0.9) = 0.2256$$

Using this value for P(F) we can then solve for P(R|F) using Bayes Rule:

$$P(R|F) = \frac{0.973 * 0.1}{0.2256} = 0.4323$$

Therefore, from this we can determine that the probability of a visitor being a robot given they have been flagged is 0.4323 or there is a chance of 43.23%.