COMP 2804 — Solutions Assignment 3

Question 1: Write your name and student number.

Solution:

• Name: Johan Cruijff

• Student number: 14

Question 2: When Tri¹ is a big boy, he wants to have four children. Assuming that the genders of these children are uniformly random, which of the following three events has the highest probability?

1. All four kids are of the same gender.

2. Exactly three kids are of the same gender.

3. Two kids are boys and two kids are girls.

As always, justify your answer.

Solution: The sample space is the set

$$S = \{(g_1, g_2, g_3, g_4) : \text{ each } g_i \in \{B, G\}\}.$$

Observe that (g_1, g_2, g_3, g_4) is an ordered sequence, where g_1 is the gender of the oldest child, g_2 is the gender of the second oldest child, g_3 is the gender of the third oldest child, and g_4 is the gender of the youngest child.

The sample space has size $2^4 = 16$, and each outcome has a probability of 1/16.

1. All four kids are of the same gender.

We denote this event by A. As a subset of the sample space, we can write $A = \{(B, B, B, B), (G, G, G, G)\}$. Since we have a uniform probability,

$$\Pr(A) = \frac{|A|}{|S|} = \frac{2}{16} = \frac{1}{8}.$$

2. Exactly three kids are of the same gender.

We denote this event by B. What is the size of B: There are 4 ways to have one girl and three boys, and there are 4 ways to have one boy and three girls. Thus, |B| = 4 + 4 = 8. This gives

$$\Pr(B) = \frac{|B|}{|S|} = \frac{8}{16} = \frac{1}{2}.$$

¹your friendly TA

3. Two kids are boys and two kids are girls.

We denote this event by C. What is the size of C: There are $\binom{4}{2} = 6$ ways to have two girls and two boys. Thus, |C| = 6. This gives

$$\Pr(C) = \frac{|C|}{|S|} = \frac{6}{16} = \frac{3}{8}.$$

Sanity check: Since A, B, and C cover all possibilities, these three probabilities must add up to 1.

From these three answers, we see that

$$Pr(B) > Pr(C) > Pr(A)$$
.

Question 3: In this exercise, we assume that, when a child is born, its gender and day of birth are uniformly random and independent of other children. Thus, for each $G \in \{\text{boy, girl}\}$ and each

$$D \in \{\text{Sun, Mon, Tue, Wednes, Thurs, Fri, Satur}\},\$$

the probability that a child has gender G and is born on a Dday is equal to 1/14.

Anil Maheshwari² has two children. You are given that at least one of Anil's kids is a boy who was born on a Sunday. Determine the probability that Anil has two boys.

Solution: We will use the set

 $\mathcal{D} = \{\text{Sunday, Monday, Tueday, Wednesday, Thursday, Friday, Saturday}\}.$

For any kid, there are two sources of randomness: gender and day of birth. Based on this, we will represent a kid using these two parameters.

Initially, we are told that Anil has two kids. Thus, the sample space is the set

$$S = \{(g_1, d_1, g_2, d_2) : g_1, g_2 \in \{B, G\} \text{ and } d_1, d_2 \in \mathcal{D}\},\$$

where g_1, d_1 represents the oldest kid and g_2, d_2 represents the youngest kid. Observe that we have a uniform probability.

Consider the following two events:

$$X =$$
 "both kids are boys"
= $\{(B, d_1, B, d_2) : d_1, d_2 \in \mathcal{D}\}$

and

Y = "at least one kid is a boy who was born on a Sunday".

²the guy in the office next to my office

We have to determine the conditional probability

$$\Pr(X \mid Y) = \frac{\Pr(X \cap Y)}{\Pr(Y)}.$$

We are going to write Y as the union of two events, depending on whether the oldest or the youngest kid is a boy who was born on a Sunday:

$$Y_1 = \{(B, \text{Sunday}, g_2, d_2) : g_2 \in \{B, G\}, d_2 \in \mathcal{D}\}$$

and

$$Y_2 = \{(g_1, d_1, B, \text{Sunday}) : g_1 \in \{B, G\}, d_1 \in \mathcal{D}\}.$$

Since $Y = Y_1 \cup Y_2$ and we have a uniform probability,

$$\Pr(Y) = \Pr(Y_1 \cup Y_2)$$

$$= \frac{|Y_1 \cup Y_2|}{|S|}$$

$$= \frac{|Y_1| + |Y_2| - |Y_1 \cap Y_2|}{|S|}$$

$$= \frac{2 \cdot 7 + 2 \cdot 7 - 1}{|S|}$$

$$= \frac{27}{|S|}.$$

Using a Venn diagram, you will see that

$$X \cap Y = X \cap (Y_1 \cup Y_2) = (X \cap Y_1) \cup (X \cap Y_2)$$
.

This gives

$$|X \cap Y| = |(X \cap Y_1) \cup (X \cap Y_2)|$$

$$= |X \cap Y_1| + |X \cap Y_2| - |X \cap Y_1 \cap Y_2|$$

$$= 7 + 7 - 1$$

$$= 13.$$

This gives

$$\Pr(X \cap Y) = \frac{|X \cap Y|}{|S|} = \frac{13}{|S|}.$$

We conclude that

$$Pr(X \mid Y) = \frac{Pr(X \cap Y)}{Pr(Y)}$$

$$= \frac{13/|S|}{27/|S|}$$

$$= \frac{13}{27}$$

$$\approx 0.48.$$

Question 4: You are given a fair red die and a fair blue die. Each of these two dice has the letter a on one face, the letter b on two faces, and the letter c on three faces. You roll both dice uniformly at random and independently of each other. Define the events

A = "at least one of the two rolls results in the letter b"

and

B = "both rolls result in the same letter".

• Determine Pr(A), Pr(B), and $Pr(A \mid B)$.

Solution: The sample space is the set

$$S = \{(x, y) : x, y \in \{a, b, c\}\},\$$

where x represents the red die and y represents the blue die. To determine Pr(A), we consider the complement of A:

$$\overline{A} = \text{``}x \neq b \text{ and } y \neq b\text{''}$$

= \{(a, a), (a, c), (c, a), (c, c)\}.

This gives

$$\Pr(\overline{A}) = \Pr((a, a) \text{ or } (a, c) \text{ or } (c, a) \text{ or } (c, c))$$
$$= \Pr(a, a) + \Pr(a, c) + \Pr(c, a) + \Pr(c, c).$$

Since the results of both rolls are independent, we get

$$\Pr(\overline{A}) = \Pr(a) \cdot \Pr(a) + \Pr(a) \cdot \Pr(c) + \Pr(c) \cdot \Pr(a) + \Pr(c) \cdot \Pr(c)$$

$$= 1/6 \cdot 1/6 + 1/6 \cdot 3/6 + 3/6 \cdot 1/6 + 3/6 \cdot 3/6$$

$$= 4/9.$$

This implies that

$$Pr(A) = 1 - Pr(\overline{A}) = 1 - 4/9 = 5/9.$$

Using a similar reasoning, we get

$$Pr(B) = Pr((a, a) \text{ or } (b, b) \text{ or } (c, c))$$

$$= Pr(a, a) + Pr(b, b) + Pr(c, c)$$

$$= Pr(a) \cdot Pr(a) + Pr(b) \cdot Pr(b) + Pr(c) \cdot Pr(c)$$

$$= 1/6 \cdot 1/6 + 2/6 \cdot 2/6 + 3/6 \cdot 3/6$$

$$= 7/18$$

and

$$Pr(A \cap B) = Pr(b, b)$$

$$= Pr(b) \cdot Pr(b)$$

$$= 2/6 \cdot 2/6$$

$$= 1/9.$$

We conclude that

$$Pr(A \mid B) = \frac{Pr(A \cap B)}{Pr(B)}$$
$$= \frac{1/9}{7/18}$$
$$= 2/7.$$

Question 5: In a standard deck of 52 cards, each card has a *suit* and a *rank*. There are four suits (spades \spadesuit , hearts \heartsuit , clubs \clubsuit , and diamonds \diamondsuit), and 13 ranks (Ace, 2, 3, 4, 5, 6, 7, 8, 9, 10, Jack, Queen, and King).

A hand of cards is a subset consisting of five cards. A hand of cards is called a *straight*, if the ranks of these five cards are consecutive and the cards are not all of the same suit.

An Ace and a 2 are considered to be consecutive, whereas a King and an Ace are also considered to be consecutive. For example, each of the three hands below is a straight:

$$8 \spadesuit, 9 \heartsuit, 10 \diamondsuit, J \spadesuit, Q \clubsuit$$
$$A \diamondsuit, 2 \heartsuit, 3 \spadesuit, 4 \spadesuit, 5 \clubsuit$$
$$10 \diamondsuit, J \heartsuit, Q \spadesuit, K \spadesuit, A \clubsuit$$

• Assume you get a uniformly random hand of cards. Determine the probability that this hand is a straight.

Solution: The probability is given by

$$Pr(straight) = \frac{number of straights}{number of hands}.$$

The number of hands is equal to $\binom{52}{5}$.

We determine the number of straights: Consider the sequence

$$A, 2, 3, 4, 5, 6, 7, 8, 9, 10, J, Q, K, A.$$

1. The first card of a straight can be any of the cards

Thus, there are $10 \cdot 4 = 40$ choices for the first card.

2. For i = 2, 3, 4, 5, once we have chosen the first i - 1 cards, there are 4 choices for the i-th card.

This gives $4 \cdot 4 \cdot 4 \cdot 4 = 4^4$ choices for the cards 2—5. We have to subtract 1, because we have included the case when cards 2—5 all have the same suit as the first card. Thus, given our choice for the first card, the number of ways to complete the straight is equal to

$$4^4 - 1$$
.

3. By the Product Rule, the total number of straights is equal to

$$40(4^4-1)$$
.

We conclude that

$$Pr(straight) = \frac{40(4^4 - 1)}{\binom{52}{5}}.$$

Using Wolfram Alpha, this gives

$$\Pr(\text{straight}) = \frac{5}{1274} \approx 0.0039.$$

Question 6: In this exercise, we consider a standard deck of 52 cards.

• We choose, uniformly at random, one card from the deck. Define the events

A = "the rank of the chosen card is Ace",

B = "the suit of the chosen card is diamonds".

Are the events A and B independent? As always, justify your answer.

• Assume we remove the Queen of hearts from the deck. We choose, uniformly at random, one card from the remaining 51 cards. Define the events

C = "the rank of the chosen card is Ace",

D = "the suit of the chosen card is diamonds".

Are the events C and D independent? Again, justify your answer.

Solution: We start with the first part of the question: Since there are four Aces out of 52 cards, we have

$$Pr(A) = 4/52 = 1/13.$$

Since there are 13 diamonds out of 52 cards, we have

$$Pr(B) = 13/52 = 1/4.$$

Since there is only one way for $A \cap B$ to happen, we have

$$\Pr(A \cap B) = 1/52.$$

Since

$$Pr(A \cap B) = Pr(A) \cdot Pr(B),$$

the events A and B are independent.

Next, we consider the second part of the question. Since there are four Aces out of 51 cards, we have

$$Pr(C) = 4/51.$$

Since there are 13 diamonds out of 51 cards, we have

$$Pr(D) = 13/51.$$

Since there is only one way for $C \cap D$ to happen, we have

$$\Pr(C \cap D) = 1/51.$$

Since

$$\Pr(C \cap D) \neq \Pr(C) \cdot \Pr(D)$$
,

the events C and D are not independent.

Question 7: Let $n \geq 2$ be an integer. Assume we have n balls and 10 boxes. We throw the balls independently and uniformly at random in the boxes. Thus, for each k and i with $1 \leq k \leq n$ and $1 \leq i \leq 10$,

Pr(the k-th ball falls in the i-th box) = 1/10.

Define the event

 $A_n =$ "there is a box that contains at least two balls"

and let $p_n = \Pr(A_n)$.

- Determine the smallest value of n for which $p_n \geq 1/2$.
- Determine the smallest value of n for which $p_n \ge 2/3$.

Solution: In class, we have seen the birthday problem: If there are d days in one year and we have n people with $2 \le n \le d$, then

Pr(at least two people have the same birthday) =
$$1 - \frac{d!}{(d-n)!d^n}$$
.

This question is the birthday problem with d = 10. Using Wolfram Alpha, we get

$$p_2 = 1/10 = 0.1,$$

 $p_3 = 7/25 = 0.28,$
 $p_4 = 62/125 = 0.496,$
 $p_5 = 436/625 = 0.6976.$

Thus, for both parts of this question, n = 5.

Question 8: Nick³ is taking the course SPID 2804 (The Effect of Spiderman on the Banana Industry). The final exam for this course consists of one true/false question. To answer this question, Nick uses the following approach:

- 1. If Nick knows that the answer to the question is "true", he answers "true".
- 2. If Nick knows that the answer is "false", he answers "false".
- 3. If Nick does not know the answer, he flips a fair coin.
 - (a) If the coin comes up heads, he answers "true".
 - (b) If the coin comes up tails, he answers "false".

You are given that Nick knows the answer to the question with probability 0.8. Define the event

A = "Nick gives the correct answer to the question".

• Determine Pr(A).

Hint: Use the event B = "Nick knows the answer". What are the conditional probabilities $Pr(A \mid B)$ and $Pr(A \mid \overline{B})$?

Solution: Using a Venn diagram, you will see that

$$A = (A \cap B) \cup (A \cap \overline{B}).$$

This implies that

$$\Pr(A) = \Pr\left((A \cap B) \cup (A \cap \overline{B}) \right).$$

Since the two events $A \cap B$ and $A \cap \overline{B}$ are disjoint, we get

$$Pr(A) = Pr(A \cap B) + Pr(A \cap \overline{B})$$
$$= Pr(A \mid B) \cdot Pr(B) + Pr(A \mid \overline{B}) \cdot Pr(\overline{B}).$$

It remains to determine the four terms on the right-hand side.

³your friendly TA

- 1. We are given that Pr(B) = 8/10 and $Pr(\overline{B}) = 2/10$.
- 2. What is $Pr(A \mid B)$: If the event B occurs, then Nick knows the answer. It follows from the algorithm that in this case, Nick gives the correct answer, i.e., the event A occurs. Thus,

$$Pr(A \mid B) = 1.$$

3. What is $Pr(A \mid \overline{B})$: If the event \overline{B} occurs, then Nick does not know the answer. It follows from the algorithm that in this case, Nick gives a uniformly random answer, i.e., the event A occurs with probability 1/2. Thus,

$$\Pr\left(A \mid \overline{B}\right) = 1/2.$$

We conclude that

$$Pr(A) = Pr(A \mid B) \cdot Pr(B) + Pr(A \mid \overline{B}) \cdot Pr(\overline{B})$$
$$= 1 \cdot 8/10 + 1/2 \cdot 2/10$$
$$= 9/10.$$

Question 9: You are asked to design a random bit generator. You find a coin in your pocket, but, unfortunately, you are not sure if it is a fair coin. After some thought, you come up with the following algorithm GENERATEBIT(n), which takes as input an integer $n \ge 1$:

```
Algorithm GenerateBit(n):

// all coin flips made are mutually independent
flip the coin n times;
k = the number of heads in the sequence of n coin flips;
if k is odd
then return 0
else return 1
endif
```

In this exercise, you will show that, when $n \to \infty$, algorithm GenerateBit(n) returns a uniformly random bit.

Let p be the real number with 0 , such that, if the coin is flipped once, it comes up heads with probability <math>p and tails with probability 1 - p. (Note that algorithm GenerateBit does not need to know the value of p.) For any integer $n \ge 1$, define the two events

$$A_n =$$
 "algorithm GENERATEBIT(n) returns 0"

and

 B_n = "the *n*-th coin flip made by algorithm GENERATEBIT(n) results in heads",

and define

$$P_n = \Pr\left(A_n\right)$$

and

$$Q_n = P_n - 1/2.$$

- Determine P_1 and Q_1 .
- For any integer $n \geq 2$, prove that

$$P_n = p + (1 - 2p) \cdot P_{n-1}.$$

Hint: Express the event A_n in terms of the events A_{n-1} and B_n .

• For any integer $n \geq 2$, prove that

$$Q_n = (1 - 2p) \cdot Q_{n-1}.$$

• For any integer $n \geq 1$, prove that

$$Q_n = (1 - 2p)^{n-1} \cdot (p - 1/2).$$

• Prove that

$$\lim_{n \to \infty} Q_n = 0$$

and

$$\lim_{n \to \infty} P_n = 1/2.$$

Solution: We start with determining P_1 and Q_1 : For n = 1,

 A_1 = "algorithm GenerateBit(1) returns 0"

= "when running algorithm GenerateBit(1), k = 1"

= "the single coin flip results in heads".

It follows that

$$P_1 = \Pr\left(A_1\right) = p$$

and

$$Q_1 = P_1 - 1/2 = p - 1/2.$$

Let $n \geq 2$. We derive the recurrence for P_n : The event A_n occurs if and only if the number of heads in n coin flips is odd. That is, one of the following two cases occurs:

1. In the first n-1 coin flips, the number of heads is odd, and the n-th coin flip results in tails. This is the event $A_{n-1} \wedge \overline{B}_n$.

2. In the first n-1 coin flips, the number of heads is even, and the n-th coin flip results in heads. This is the event $\overline{A}_{n-1} \wedge B_n$.

This shows that

$$A_n = \left(A_{n-1} \wedge \overline{B}_n \right) \vee \left(\overline{A}_{n-1} \wedge B_n \right).$$

Since the two events $A_{n-1} \wedge \overline{B}_n$ and $\overline{A}_{n-1} \wedge B_n$ are disjoint, it follows that

$$P_{n} = \Pr(A_{n})$$

$$= \Pr((A_{n-1} \wedge \overline{B}_{n}) \vee (\overline{A}_{n-1} \wedge B_{n}))$$

$$= \Pr(A_{n-1} \wedge \overline{B}_{n}) + \Pr(\overline{A}_{n-1} \wedge B_{n}).$$

Since the coin flips are independent, we get

$$P_{n} = \operatorname{Pr}(A_{n-1}) \cdot \operatorname{Pr}(\overline{B}_{n}) + \operatorname{Pr}(\overline{A}_{n-1}) \cdot \operatorname{Pr}(B_{n})$$

$$= P_{n-1} \cdot (1-p) + (1-P_{n-1}) \cdot p$$

$$= p + (1-2p) \cdot P_{n-1}.$$

Note that we have obtained a recurrence for the numbers P_n . Our goal is to solve this recurrence. As we will see, by expressing the P_n 's in terms of the Q_n 's, we obtain a simpler recurrence.

How to obtain the recurrence for Q_n : We know that

$$P_n = p + (1 - 2p) \cdot P_{n-1}.$$

This is the same as

$$(P_n - 1/2) + 1/2 = p + (1 - 2p) \cdot (P_{n-1} - 1/2) + (1 - 2p) \cdot 1/2.$$

This is the same as

$$Q_n + 1/2 = p + (1 - 2p) \cdot Q_{n-1} + 1/2 - p,$$

which is the same as

$$Q_n = (1 - 2p) \cdot Q_{n-1}.$$

How to solve the recurrence for Q_n : We know that $Q_1 = p - 1/2$. We obtain Q_n by taking the previous term Q_{n-1} and multiplying it by 1 - 2p. Therefore, we get, for each $n \ge 1$,

$$Q_n = (1 - 2p)^{n-1} \cdot (p - 1/2).$$

(If you want, you can prove this by induction. I do not think this is necessary, because it should be clear to you.)

Since 0 , we have <math>-1 < 1 - 2p < 1. Thus,

$$\lim_{n \to \infty} (1 - 2p)^{n-1} = 0.$$

It follows that

$$\lim_{n \to \infty} Q_n = 0$$

and

$$\lim_{n \to \infty} P_n = \lim_{n \to \infty} (Q_n + 1/2) = 0 + 1/2 = 1/2.$$

Question 10: Let p be a real number with $0 . You are given two coins <math>C_1$ and C_2 . The coin C_1 is fair, i.e., if you flip this coin, it comes up heads with probability 1/2 and tails with probability 1/2. If you flip the coin C_2 , it comes up heads with probability p and tails with probability p and tails with probability p and tails with probability p and the coin p and p are independent of each other. Define the events

A = "the first coin flip results in heads", B = "the second coin flip results in heads".

• Determine Pr(A).

Hint: Express Pr(A) in terms of conditional probabilities, depending on which coin is chosen.

- Assume that p = 1/4. Are the events A and B independent? As always, justify your answer.
- Determine all values of p for which the events A and B are independent. Again, justify your answer.

Solution: Following the hint, we will use the event

$$D =$$
 "we choose coin C_1 ".

Using a Venn diagram, you will see that

$$A = (A \cap D) \cup \left(A \cap \overline{D}\right).$$

This implies that

$$\Pr(A) = \Pr((A \cap D) \cup (A \cap \overline{D})).$$

Since the two events $A \cap D$ and $A \cap \overline{D}$ are disjoint, we have

$$Pr(A) = Pr(A \cap D) + Pr(A \cap \overline{D})$$

=
$$Pr(A \mid D) \cdot Pr(D) + Pr(A \mid \overline{D}) \cdot Pr(\overline{D}).$$

It remains to determine the four terms on the right-hand side.

1. Since we pick one of the coins C_1 and C_2 uniformly at random, Pr(D) = 1/2 and $Pr(\overline{D}) = 1/2$.

- 2. What is $Pr(A \mid D)$: If the event D occurs, then we choose coin C_1 . Since C_1 is a fair coin, we have $Pr(A \mid D) = 1/2$.
- 3. What is $\Pr(A \mid \overline{D})$: If the event \overline{D} occurs, then we choose coin C_2 . Thus, $\Pr(A \mid \overline{D}) = p$.

We conclude that

$$Pr(A) = 1/2 \cdot 1/2 + p \cdot 1/2 = 1/4 + p/2.$$

By exactly the same reasoning, we see that

$$Pr(B) = 1/4 + p/2.$$

Next, we determine $Pr(A \cap B)$: Since

$$A \cap B = (A \cap B \cap D) \cup (A \cap B \cap \overline{D})$$
,

and the right-hand side is a disjoint union of two events, we have

$$Pr(A \cap B) = Pr ((A \cap B \cap D) \cup (A \cap B \cap \overline{D}))$$

$$= Pr (A \cap B \cap D) + Pr (A \cap B \cap \overline{D})$$

$$= Pr (A \cap B \mid D) \cdot Pr(D) + Pr (A \cap B \mid \overline{D}) \cdot Pr (\overline{D})$$

$$= 1/2 \cdot 1/2 \cdot 1/2 + p \cdot p \cdot 1/2$$

$$= 1/8 + p^2/2.$$

Assume that p = 1/4. To determine if the events A and B are independent, we verify if

$$\Pr(A \cap B) \stackrel{?}{=} \Pr(A) \cdot \Pr(B).$$

Since p = 1/4, this becomes

$$\frac{1}{8} + \frac{1/4^2}{2} \stackrel{?}{=} \left(\frac{1}{4} + \frac{1/4}{2}\right) \cdot \left(\frac{1}{4} + \frac{1/4}{2}\right).$$

The left-hand side is equal to 5/32, whereas the right-hand side is equal to 9/64. Since these two values are not equal, A and B are not independent.

For the final part, we have to find all values of p for which

$$\frac{1}{8} + \frac{p^2}{2} = \left(\frac{1}{4} + \frac{p}{2}\right)^2.$$

If we multiply both sides by 16, we get

$$2 + 8p^2 = (1 + 2p)^2.$$

After expanding and simplifying, we get

$$4p^2 - 4p + 1 = 0,$$

which is equivalent to

$$p^2 - p + 1/4 = 0,$$

which is equivalent to

$$(p - 1/2)^2 = 0,$$

which is equivalent to

$$p = 1/2$$
.

Conclusion: The events A and B are independent if and only if p=1/2.