Lecture 3 Worksheet

Chrysafis Vogiatzis

Every worksheet will work as follows.

1. You will be entered into a Zoom breakout session with other students in the class.
2. Read through the worksheet, discussing any questions with the other participants in your breakout session.
   - You can call me using the "Ask for help" button.
   - Keep in mind that I will be going through all rooms during the session so it might take me a while to get to you.
3. Answer each question (preferably in the order provided) to the best of your knowledge.
4. While collaboration between students in a breakout session is highly encouraged and expected, each student has to submit their own version.
5. You will have 24 hours (see Compass) to submit your work.

Worksheet 1: Game of dies

Problem 1

Consider a game where you roll two fair dice (where each number from 1 to 6 has an equal probability of appearing). What is the probability that the sum of the numbers on the two dice is 7?

Answer to Problem 1.

Problem 2

What is the probability the sum of the numbers on the two dice is 7 assuming that the first dice rolled on a 3?

Answer to Problem 2.

Problem 3

Based on your answers on part (a) and (b), what can you claim about the independence of the events "the sum of the two dice is 7" and "the first dice is a 3"? Is that true for all pairs of events "the sum of the two dice is 7" and "the first dice is a i" where $i = 1, 2, 3, 4, 5, 6$?

Answer to Problem 3.

Problem 4

Prove or disprove[1] the following statement.

- When throwing two dies, the two events "the sum of the two dies is $j = 2, 3, \ldots 12$" and "the first die is a $i = 1, 2, \ldots, 6$" for $i < j$ are independent events.

[1] To disprove a statement, you may simply find an example where the statement is **not** true.

Answer to Problem 4.

Worksheet 2: Quality control revisited

Problem 5

A manufacturing facility is making 2 different products. Every product can be classified as defective (D) or non-defective (ND). In addition to that, some products appear to have cosmetic damage (C) or not (NC). The company has collected data for both products over the last 400 items for each. [2]

[2] You may treat these numbers as "probabilities": for example a product 1 is defective and has cosmetic damage with probability $5/400$, whereas a product 2 that is known to be defective has cosmetic damage with probability $2/20$.

Product 1:

	Cosm. dam.		
Def.	Yes (C)	No (NC)	Total
Yes (D)	5	23	28
No (ND)	24	348	372
Total	29	371	400

Product 2:

	Cosm. dam.		
Def.	C	NC	Total
D	2	18	20
ND	38	342	380
Total	40	360	400

You pick up an item from the recent production of Product 1. If you see it has cosmetic damage, does this alter your perception that the product is defective?

Answer to Problem 5.

Problem 6

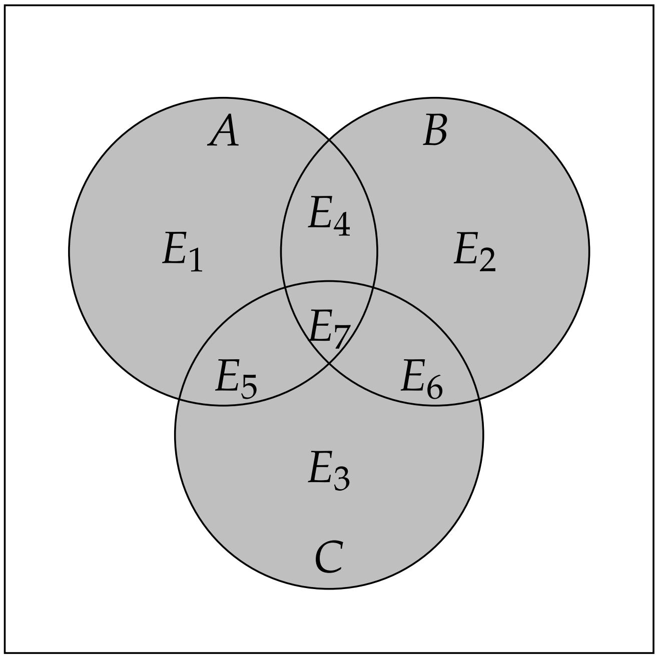
Using the data provided from the previous problem, what can you deduce about Product 2? Does knowing that it has cosmetic damage alter your perception that the product is defective?

Answer to Problem 6.

Worksheet 3: Deriving the probability of the union of more than 2 events

Problem 7

Assume that in the following picture, S is the whole rectangle and A, B, and C are some events.



You will notice that we have already marked **7 mutually exclusive** events for you. Define them (in mathematical terms, using unions, intersections, complements, relative complements, etc.). [3]

[3] Hint: E_7 is easy to define based on sets A, B, and C... It looks like the intersection of all 3...

Answer to Problem 7.

E_1:

E_2:

E_3:

E_4:

E_5:

E_6:

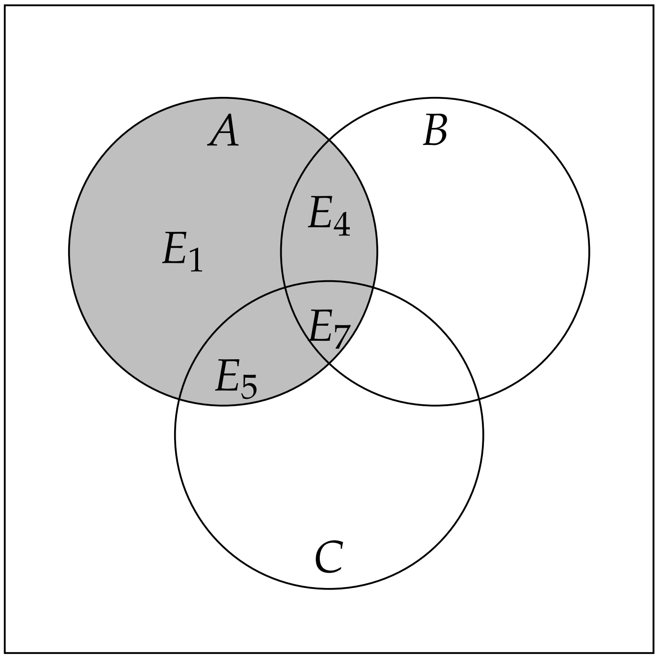
E_7:

From the fact that the previous sets are mutually exclusive, we may deduce that

$$P(A \cup B \cup C) = P(E_1) + P(E_2) + \ldots + P(E_7) = \sum_{i=1}^{7} P(E_i). \tag{1}$$

Problem 8

Let us focus on event A and the events E_1, E_4, E_5, E_7 that belong to it.



Consider the probability of the set E_7: how can you write $P(E_7)$ as a function of sets A, B, C? Then, consider the probability of the first set ($P(E_1)$). How can you write this one as a function of $P(A)$, $P(E_4)$, $P(E_5)$, and $P(E_7)$? Finally, do the same for the probabilities of events E_4, E_5.

Answer to Problem 8.

$P(E_7) =$

$P(E_4) =$

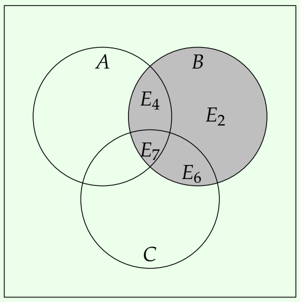
$P(E_5) =$

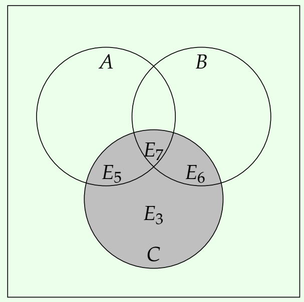
$P(E_1) =$

Problem 9

Without redoing all of the calculations, also complete based on your answers on Problem 8, the probabilities according to set B and C:

Answer to Problem 9.





$P(E_7) =$	$P(E_7) =$
$P(E_4) =$	$P(E_5) =$
$P(E_6) =$	$P(E_6) =$
$P(E_2) =$	$P(E_3) =$

Problem 10

Combine all of the above in (1) to derive what $P(A \cup B \cup C)$ is equal to as a function of A, B, C, $A \cap B$, $A \cap C$, $B \cap C$, $A \cap B \cap C$. [4]

Answer to Problem 10.

[4] If you get stuck, take a look at the lecture notes and the general case for $m > 3$ events.

Worksheet 4: The birthday problem

Our class has 90 students. If our class had 365 students (assume for now with me that February 29th does *not* exist and every year has 365 days), then we would be guaranteed that at least two of you share the same birthday.

The question though becomes: in a class of 90, what is the probability that two of you have the same birthday?

Problem 11

Let's start simple. If there are only two of you, what is the probability that you share the same birthday?

Answer to Problem 11.

Problem 12

With your answer in Problem 11 in mind, add a third person in the mix. What is the probability that the third person has the same birthday with either the first or the second person? [5]

Answer to Problem 12.

[5] Hint: consider the event that you do not have the same birthday as E and then calculate $P(\overline{E})$. Also: how many possible triplets of birthday dates can you create such that no two birthdays are the same? And how many possible triplets of birthday dates can you create in total? Probability can de calculated as the first number over the second on, if only we knew how to count the number of events...

Problem 13

Based on your reasoning in Problems 11 and 12, you must have reached a probability of

$$\frac{364}{365} \cdot \frac{363}{365} \cdot \frac{362}{365} \cdot \ldots \cdot \frac{365-n+1}{365}$$

for the event that no two people share a birthday in a group of n people. Using the fact that $P(E) = 1 - P(\overline{E})$, we can deduce that the probability that two people share a birthday is:

$$P(\textit{share birthday}) = 1 - \frac{364}{365} \cdot \frac{363}{365} \cdot \frac{362}{365} \cdot \ldots \cdot \frac{365-n+1}{365} \tag{2}$$

What does expression (2) evaluate to in a class of 90 students? What does it evaluate to in a class of 25 students?

Answer to Problem 13.

Next time a person in any class of a significant size shares the same birthday with you, remember that this is not the biggest coincidence in the world, but a rather common observance.

Say we had run this for a many values of n, starting from $n = 1$ (one person alone has a 0% chance of sharing the birthday with someone else) to $n = 2$ (two people sharing a birthday with a $1/365$ chance) and so on, until $n = 150$ people. We would have then obtained a figure like the one in Figure 1. What do you observe?

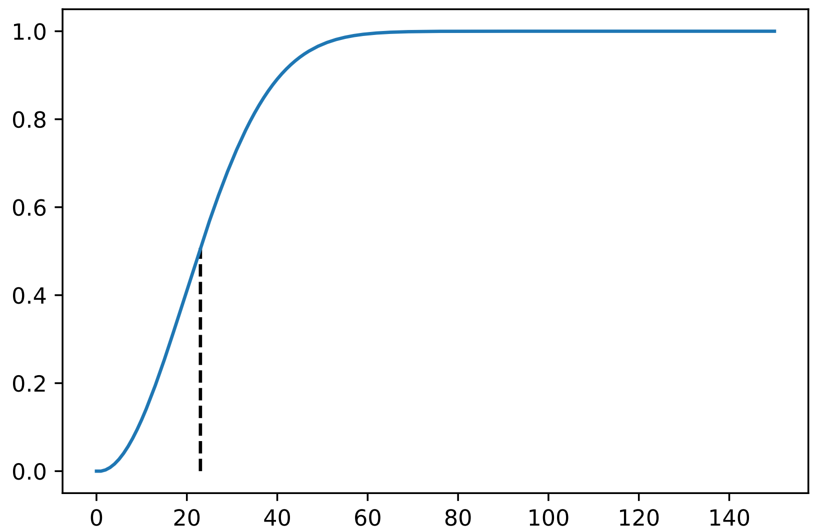


Figure 1: The birthday problem probabilities, visualized. It is at 23 people that this is roughly equal to 50%!