Summer School on Mathematical Philosophy for Female Students

Introduction to Probability Theory, Algebra, and Set Theory

Catrin Campbell-Moore and Sebastian Lutz*

July 28, 2014

Question 1. Draw Venn diagrams for the following sets and write them in a simpler way:

- $A \cap (B \cup A)$
- $A \setminus (B \cap (A \cup C))$

Answer. I'll omit the Venn diagrams, but $A \cap (B \cup A) = A$ and $A \setminus (B \cap (A \cup C)) = A \setminus B$.

Question 2. Let $\Omega = \{\omega_1, \omega_2, \omega_3\}$ and $\mathcal{F} = \mathcal{P}(\Omega)$. Let $P(\{\omega_1\}) = 0.1$, $P(\{\omega_2\}) = 0.3$ and $P(\{\omega_3\}) = 0.6$. Find the probabilities for all the other events in $\mathcal{F}$.

Answer. Since $\{a\}$, $\{b\}$, and $\{c\}$ are the atoms of the Boolean algebra, their probabilities give all the other events' probabilities simply by summation.

$$
\begin{aligned}
P(\varnothing) &= 0\\
P(\{a,b\}) &= P(\{a\}) + P(\{b\}) = 0.1 + 0.3 = 0.4\\
P(\{a,c\}) &= P(\{a\}) + P(\{c\}) = 0.7\\
P(\{b,c\}) &= P(\{b\}) + P(\{c\}) = 0.9\\
P(\{a,b,c\}) &= P(\Omega) = 1
\end{aligned}
$$

Question 3. Two fair, independent, four-sided dice, one red and one green, are rolled. Let the event A be "The sum of the faces showing is an even number." Define an appropriate Ω, list the states in A and say what $P(A)$ is.

Answer. Let $\ulcorner\langle i, j\rangle\urcorner$ stand for $\ulcorner$The red die shows i and the green die shows j.$\urcorner$. Then set of states is $\Omega = \{\langle i, j\rangle : 1 \leqslant i, j \leqslant 4\}$ and the atoms of the algebra of events are the singleton sets of Ω. Since the dice fair and their results independent, each atom has the same probability, namely $\frac{1}{4\cdot 4} = \frac{1}{16}$. The event described by A is $\{\langle i, j\rangle : i + j \in \{2, 4, 6, 8\}\}$. The probability of this event is given by the sum of the probabilities of the atoms whose union is the event. I

*Email `catrin@ccampbell-moore.com` with any mistakes or comments

bet there are combinatorical formulas to find out how many states there are in the event, but I don't know them. So I will write out all the states:

$$
\begin{aligned}
\{\langle i,j\rangle : i+j \in \{2,4,6,8\}\} = \{&\langle 1,1\rangle \\
&\langle 1,3\rangle, \langle 2,2\rangle, \langle 3,1\rangle \\
&\langle 2,4\rangle, \langle 3,3\rangle, \langle 4,2\rangle \\
&\langle 4,4\rangle\}
\end{aligned}
$$

There are thus 16 states, the singleton set of which make up an atom, each of which has probability $\frac{1}{16}$. The probability of the event is thus $\frac{8}{16} = \frac{1}{2}$.

Question 4. Consider the following two events:

- A: It rains in Munich tomorrow
- B: The burglars who stole 300,000 litres of beer a few weeks ago will be caught

Assume $P(A) = 0.4$, $P(B) = 0.6$ and that A and B are probabilistically independent according to P.

Calculate:

1. $P(A^c)$
2. $P(A \cap B)$
3. $P(A \cup B)$

Answer.

$$
\begin{aligned}
P(A^c) &= 1 - P(A) = 0.6 \\
P(A \cap B) &= P(A) \cdot P(B) = 0.24 \\
P(A \cup B) &= 1 - P((A \cup B)^c = 1 - P(A^c \cap B^c) = 1 - P(A^c) \cdot P(B^c) \\
&= 1 - (1 - P(A)) \cdot (1 - P(B)) = 1 - 0.6 \cdot 0.4 = 0.24
\end{aligned}
$$

The last calculation uses the fact that if two events are independent, so are their complements: If $P(A|B) = P(B)$, then $P(A \cap B) = P(A) \cdot P(B)$ and hence $P(A^c|B^c) = \frac{P(A^c \cap B^c)}{P(B^c)} = \frac{P(A^c) \cdot P(B^c)}{P(B^c)} = P(A^c)$.

Question 5. Consider $\Omega = \{\omega_1, \omega_2\}$ and $\{\omega_1\}, \{\omega_2\} \in \mathcal{F}$. Let $P(\{\omega_1\}) = 0.2$ and $P(\{\omega_2\}) = 0.8$. Calculate the expected values of the following variables.

	ω_1 $P(\{\omega_1\}) = 0.2$	ω_2 $P(\{\omega_2\}) = 0.8$
X	5	5
Y	6	8
Z	1	10

Answer.

$$E[X] = \sum_{r \in \mathsf{range}(X)} r \cdot P\{X = r\} \tag{1}$$
$$= \sum_{r \in \{5\}} r \cdot P\{X = r\} \tag{2}$$
$$= 5 \cdot P\{\omega_1, \omega_2\} \tag{3}$$
$$= 5 \cdot 1 \tag{4}$$
$$= 5 \tag{5}$$

$$E[Y] = \sum_{r \in \mathsf{range}(Y)} r \cdot P\{X = r\} \tag{6}$$
$$= \sum_{r \in \{6,8\}} r \cdot P\{X = r\} \tag{7}$$
$$= 6 \cdot P\{\omega_1\} + 8 \cdot P\{\omega_2\} \tag{8}$$
$$= 6 \cdot 0.2 + 8 \cdot 0.8 \tag{9}$$
$$= 7.6 \tag{10}$$

$$E[Z] = \sum_{r \in \mathsf{range}(Z)} r \cdot P\{X = r\} \tag{11}$$
$$= \sum_{r \in \{1,10\}} r \cdot P\{X = r\} \tag{12}$$
$$= 1 \cdot P\{\omega_1\} + 10 \cdot P\{\omega_2\} \tag{13}$$
$$= 1 \cdot 0.2 + 10 \cdot 0.8 \tag{14}$$
$$= 8.2 \tag{15}$$

Question 6. For the roll of one four-sided and one six-sided die which are fair and independent,

1. what is the probability of the total score being 3?
2. what is the expectation value of the total score?
3. what is the expectation value of the maximum score?

Answer. Similar to the answer to question 3, the set of states can be written as $\Omega = \{\langle i, j\rangle : 1 \leqslant i \leqslant 4 \text{ and } 1 \leqslant j \leqslant 6\}$. Again, each singleton set of a state is an event, and its probability is $\frac{1}{4 \cdot 6} = \frac{1}{24}$.

The probability of the total score being 3 is the probability of the event $\{\langle i, j\rangle : i + j = 3\} = \{\langle 1, 2\rangle, \langle 2, 1\rangle\}$, and thus is $P(\{\langle 1, 2\rangle, \langle 2, 1\rangle\}) = P(\{\langle 1, 2\rangle\}) + P(\{\langle 2, 1\rangle\}) = \frac{1}{12}$.

To determine the expected value of the total score, the previous calculation has to be done for all the possible total scores:

$$\begin{aligned}
\text{Score 2:} \quad & P(\{\langle 1,1\rangle\}) = \frac{1}{16} \\
\text{Score 3:} \quad & P(\{\langle 1,2\rangle, \langle 2,1\rangle\}) = \frac{2}{16} \\
\text{Score 4:} \quad & P(\{\langle 1,3\rangle, \langle 2,2\rangle, \langle 3,1\rangle\}) = \frac{3}{16} \\
\text{Score 5:} \quad & P(\{\langle 1,4\rangle, \langle 2,3\rangle, \langle 3,2\rangle, \langle 4,1\rangle\}) = \frac{4}{16} \\
\text{Score 6:} \quad & P(\{\langle 1,5\rangle, \langle 2,4\rangle, \langle 3,3\rangle, \langle 4,2\rangle, \langle 5,1\rangle\}) = \frac{5}{16} \\
\text{Score 7:} \quad & P(\{\langle 1,6\rangle, \langle 2,5\rangle, \langle 3,4\rangle, \langle 4,3\rangle\}) = \frac{4}{16} \\
\text{Score 8:} \quad & P(\{\langle 2,6\rangle, \langle 3,5\rangle, \langle 4,4\rangle\}) = \frac{3}{16} \\
\text{Score 9:} \quad & P(\{\langle 3,6\rangle, \langle 4,5\rangle\}) = \frac{2}{16} \\
\text{Score 10:} \quad & P(\{\langle 4,6\rangle\}) = \frac{1}{16}
\end{aligned}$$

The expectation value of the total score is then

$$\sum_{s \in S} s \cdot P(\{\langle i,j\rangle : i+j=s\}) =$$
$$2 \cdot \frac{1}{16} + 3 \cdot \frac{2}{16} + 4 \cdot \frac{3}{16} + 5 \cdot \frac{4}{16} + 6 \cdot \frac{5}{16} + 7 \cdot \frac{4}{16} + 8 \cdot \frac{3}{16} + 9 \cdot \frac{2}{16} + 10 \cdot \frac{1}{16}$$
$$= \frac{2+6+12+20+30+28+24+18+10}{16} = \frac{75}{8}, \quad (16)$$

where S is the set of total scores.

The expectation value of the maximum score (where one just looks at the die showing the highest value) is determined by the following probabilities:

$$\begin{aligned}
\text{Score 1:} \quad & P(\{\langle 1,1\rangle\}) = \frac{1}{16} \\
\text{Score 2:} \quad & P(\{\langle 1,2\rangle, \langle 2,2\rangle, \langle 2,1\rangle\}) = \frac{3}{16} \\
\text{Score 3:} \quad & P(\{\langle 1,3\rangle, \langle 2,3\rangle, \langle 3,3\rangle, \langle 3,2\rangle, \langle 3,1\rangle\}) = \frac{5}{16} \\
\text{Score 4:} \quad & P(\{\langle 1,4\rangle, \langle 2,4\rangle, \langle 3,4\rangle, \langle 4,4\rangle\langle 4,3\rangle, \langle 4,2\rangle, \langle 4,1\rangle\}) = \frac{7}{16} \\
\text{Score 5:} \quad & P(\{\langle 1,5\rangle, \langle 2,5\rangle, \langle 3,5\rangle, \langle 4,5\rangle\}) = \frac{4}{16} \\
\text{Score 6:} \quad & P(\{\langle 1,6\rangle, \langle 2,6\rangle, \langle 3,6\rangle, \langle 4,6\rangle\}) = \frac{4}{16}
\end{aligned}$$

The expectation value of the maximum score is then

$$\sum_{s\in S} s\cdot P(\{\langle i,j\rangle : \max\{i,j\}=s\}) =$$

$$1\cdot\frac{1}{16}+2\cdot\frac{3}{16}+3\cdot\frac{5}{16}+4\cdot\frac{7}{16}+5\cdot\frac{4}{16}+6\cdot\frac{7}{16}=\frac{112}{16}=7\ , \quad (17)$$

Where S this time is the set of all maximum scores.

Question 7. Assume that for the role of a die, $X(\{i \text{ on top}\}) = 2i$ and $Y(\{i \text{ on top}\}) = i^2$. Assume X is applied to the outcomes of a roll of an eight-sided die and Y is applied to the outcomes of a roll of a six-sided die. What is the probability of $X + Y = 6$?

Answer. Similar to the answer to the previous question, the set of states can be written as $\Omega = \{\langle i,j\rangle : 1 \leqslant i \leqslant 8 \text{ and } 1 \leqslant j \leqslant 6\}$. The probability we are looking for is $P(\{X+Y=6\}) = P(\{\langle i,j\rangle : X(\langle i,j\rangle)+Y(\langle i,j\rangle)=6\}) = P(\{\langle i,j\rangle : 1\leqslant i\leqslant 8 \text{ and } 1\leqslant j\leqslant 6 \text{ and } 2i+j^2=6\}) = P(\{\langle 1,2\rangle\}) = \frac{1}{8\cdot 6} = \frac{1}{48}$.

Question 8. Suppose there is a medical diagnostic test for a disease. The sensitivity of the test is .95. This means that if a person has the disease, the probability that the test gives a positive response is .95. The specificity of the test is .90. This means that if a person does not have the disease, the probability that the test gives a negative response is .90, or that the false positive rate of the test is .10. In the population, 1% of the people have the disease. What is the probability that a person tested has the disease, given the results of the test is positive? Let D be the event that the person has the disease and T be the event that the test gives a positive result.

Answer. We have the following information:

- $P(T|D) = 0.95$
- $P(T^c|D^c) = 0.9$
- $P(D) = 0.01$

And we are asked to work out $P(D|T)$.

To do this we will use the (extended) Bayes theorem:

$$P(B_m|A) = \frac{P(A|B_m)\cdot P(B_m)}{\sum_{i=1}^{k} P(A|B_i)\cdot P(B_i)}$$

$$P(D|T) = \frac{P(T|D)\cdot P(D)}{P(T|D)\cdot P(D) + P(T|D^c)\cdot P(D^c)}$$

From the given information we can also calculate:

- $P(D^c) = 1 - 0.01 = 0.99$
- $P(T|D^c) = 1 - P(T^c|D^c) = 1 - 0.9 = 0.1$

The first follows from the rule: $P(A^c) = 1 - P(A)$. The second was actually given in the question, but it can be shown by proving that in general $P(A|B) + P(A^c|B) = 1$. This can be done as follows: $(A \cap B)$ and $(A^c \cap B)$ are disjoint so $P((A \cap B) \cup (A^c \cap B) = P(A \cap B) + P(A^c \cap B)$, and $(A \cap B) \cup (A^c \cap B) = B$, so

$$
\begin{aligned}
P(A|B) + P(A^c|B) &= \frac{P(A \cap B)}{P(B)} + \frac{P(A^c \cap B)}{P(B)} && (18) \\
&= \frac{P(A \cap B) + P(A^c \cap B)}{P(B)} && (19) \\
&= \frac{P((A \cap B) \cup (A^c \cap B))}{P(B)} && (20) \\
&= \frac{P(B)}{P(B)} && (21) \\
&= 1 && (22) \\
& && (23)
\end{aligned}
$$

We can finally put these together:

$$
\begin{aligned}
P(D|T) &= \frac{P(T|D) \cdot P(D)}{P(T|D) \cdot P(D) + P(T|D^c) \cdot P(D^c)} && (24) \\
&= \frac{0.95 \cdot 0.01}{0.95 \cdot 0.01 + 0.1 \cdot 0.99} && (25) \\
&\approx 0.08756 \qquad < 10\% && (26)
\end{aligned}
$$

Question 9 (Monty Hall)**.** Suppose you are on a gameshow. There are three doors: D_A, D_B and D_C. Behind one of the doors there is an expensive sports car, behind the other two there are goats, but you don't know which. (You want the sports car!). Monty Hall, the gameshow host, asks you to pick one of the doors but not to open it yet. He will then pick one of the other two doors which has a goat behind it and show you the goat. He then gives you the opportunity to change your mind. Should you switch? Justify your answer using probability theory. Note: If you are standing in front of the door with the car, Monty will open one of the remaining doors at random.

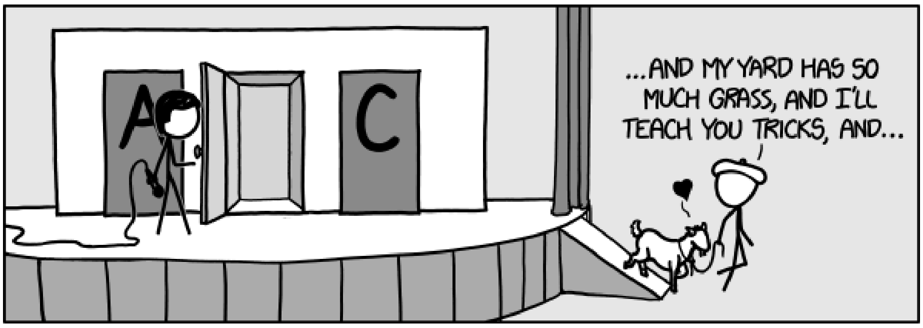

Figure 1: xkcd

Answer. At the beginning of the game the car could be behind any of the doors D_A, D_B, D_C with equal probabilities, that is, $P(C_A) = P(C_B) = P(C_C) = \frac{1}{3}$, where $\ulcorner C_X \urcorner$ stands for $\ulcorner$The car is behind door $C_X\urcorner$. The letters labelling the doors do not matter, so just assume you are standing in front of door D_A and Monty has opened door D_B. Since behind D_B there is a goat, you just need to know the probability of the car being behind door D_C given that Monty has opened door D_B (which I write as 'O_B'). Thus you need to know $P(C_C|O_B)$. You only know the conditional probabilities in the other direction, though: You know that Monty does not open the door behind you, nor the door with the car. And you know that if you are standing in front of the car, he will open one of the other doors randomly. You also know that the car is behind one of the doors, that is, C_A, C_B, and C_C are mutually exclusive and exhaustive, so one can use the law of total probability. All of this can be plugged into Bayes' theorem:

$$P(C_C|O_B) = \frac{P(O_B|C_C)P(C_C)}{P(O_B|C_A)P(C_A) + P(O_B|C_B)P(C_B) + P(O_B|C_C)P(C_C)}$$
$$= \frac{1 \cdot \frac{1}{3}}{\frac{1}{2} \cdot \frac{1}{3} + 0 \cdot \frac{1}{3} + 1 \cdot \frac{1}{3}} = \frac{2}{3} \quad (27)$$

You can play the game yourself and read a non-formal explanation of the result at the New York Times' website: `http://www.nytimes.com/2008/04/08/science/08monty.html`.

Question 10. Show that a partition of Ω induces a partition on every element of $\mathcal{F}$. I.e. if $B_1, \ldots, B_n$ partitions Ω then for every $A \in \mathcal{F}$, $(A \cap B_1), \ldots, (A \cap B_n)$ partitions A.

Answer. For every $i \neq j$, $(A \cap B_i) \cap (A \cap B_j) = A \cap (B_i \cap B_j) = A \cap \varnothing = \varnothing$. Further it has to be shown that $(A \cap B_1) \cup (A \cap B_2) \cup \cdots \cup (A \cap B_n) = A$. This is easier to show by working backwards and doing some monkey-business up front: $A = A \cap \Omega = A \cap (B_1 \cup B_2 \cup \cdots \cup B_n) = (A \cap B_1) \cup \big(A \cap (B_2 \cup B_2 \cup \cdots \cup B_n)\big) = (A \cap B_1) \cup (A \cap B_2) \cup \big(A \cap (B_3 \cup B_4 \cup \cdots \cup B_n)\big) = \cdots = (A \cap B_2) \cup (A \cap B_2) \cup \cdots \cup (A \cap B_n)$.

Question 11. Show that the following follow from the formalism for events:

- If $A, B \in \mathcal{F}$ then $A \cap B \in \mathcal{F}$.
- If $A, B \in \mathcal{F}$ then $A \backslash B \in \mathcal{F}$.

Answer. If $A, B \in \mathcal{F}$, then $A^c \in \mathcal{F}$ and $B^c \in \mathcal{F}$. Therefore $A^c \cup B^c \in \mathcal{F}$ as well, and also $(A^c \cup B^c)^c \in \mathcal{F}$. Since $(A^c \cup B^c)^c = A \cap B$, it therefore holds that $A \cap B \in \mathcal{F}$.

If $A, B \in \mathcal{F}$ then $B^c \in \mathcal{F}$ and hence, by the previous proof, $A \cap B^c \in \mathcal{F}$. Since $A \cap B^c = A \backslash B$, $A \backslash B \in \mathcal{F}$.

Question 12. Prove that the following follow from the axioms of probability:

1. $P(\varnothing) = 0$
2. $P(A \cup B) + P(A \cap B) = P(A) + P(B)$
3. If $C_1, \ldots, C_k$ is a partition of A then $P(A) = P(C_1) + \ldots + P(C_k)$[1]

[1] I have added this as an explicit part since I will use it as a lemma to prove the law of total probability, and also for question 13 so I want to be able to refer back to it.

4. The law of total probability.

Answer. 1)$P(\varnothing) = 0$: Observe that $\varnothing \cap \Omega = \varnothing$. Therefore by finite additivity, $P(\varnothing \cup \Omega) = P(\varnothing) + P(\Omega)$. We can see that $\varnothing \cup \Omega = \Omega$ and therefore $P(\Omega) = P(\varnothing) + P(\Omega)$. This directly shows that $P(\varnothing) = 0$.

2)$P(A \cup B) + P(A \cap B) = P(A) + P(B)$: $A \cap B$ and $A \cap B^c$ are disjoint and their union is A, so by the law of finite additivity

$$P(A) = P(A \cap B) + P(A \cap B^c)$$

Also B and $(A \cap B^c)$ are disjoint and their union is $A \cup B$, so by the law of finite additivity

$$P(A \cup B) = P(B) + P(A \cap B^c)$$

Putting these together we get:

$$P(A \cup B) = P(B) + P(A) - P(A \cap B)$$

Which leads directly to the desired result.

3)Work by induction on k to show that if $C_1, \ldots, C_k$ is a partition of A then $P(A) = P(C_1) + \ldots + P(C_k)$.

Base case: $k = 1$ is trivial since the only partition is $C_1 = A$.

Induction step: suppose that for any k-sized partition, $C'_1, \ldots, C'_k$ of A is such that $P(A) = P(C'_1) + \ldots + P(C'_k)$. Consider the partition $C_1, \ldots, C_{k+1}$. Observe that therefore $C_1, \ldots, C_{k-1}, (C_k \cup C_{k+1})$ is a k-sized partition of A. Therefore by the induction hypothesis we have that $P(A) = P(C_1) + \ldots + P(C_{k-1}) + P(C_k \cup C_{k+1}))$. C_k and C_{k+1} are disjoint so by the axiom of finite additivity we have that $P(C_k \cup C_{k+1}) = P(C_k) + P(C_{k+1})$. Plugging this into the previous equation this suffices to show that $P(A) = P(C_1) + \ldots + P(C_k) + P(C_{k+1})$.

4) The law of total probability: Suppose that $B_1, \ldots, B_k$ is a partition of Ω. We need to show that

$$P(A) = \sum_{i=1}^{k} P(A|B_i) \cdot P(B_i)$$

By question 10, $(A \cap B_1), \ldots, (A \cap B_n)$ partitions A. Therefore by part 3,

$$P(A \cap B_1) + P(A \cap B_2) + \ldots + P(A \cap B_k) = P(A)$$

. Then observe that if $P(B_i) > 0$ then $P(A \cap B) = P(A|B_i) \cdot P(B_i)$ by rearranging the ratio formula. If $P(B_i) = 0$ then by monotonicity also $P(A \cap B) = 0$, so we can without loss of generality assume that each B_i has $P(B_i) > 0$ (otherwise delete the B_i). By substituting this into the equation we directly get the desired result:

$$P(A) = \sum_{i=1}^{k} P(A \cap B_i) \tag{28}$$

$$= \sum_{i=1}^{k} P(A|B_i) \cdot P(B_i) \tag{29}$$

Question 13. [2]Suppose that $\mathcal{F}$ is finite. Show that P satisfies the axioms of probability if and only if there is a $P^- : \{A|A \text{ is an atom of } \mathcal{F}\} \to \mathbb{R}$ such that

- $P^-(A) \geqslant 0$ for all atoms A
- $\sum_{A \text{ is an atom of } \mathcal{F}} P^-(A) = 1$

with

$$P(B) = \sum_{A \text{ is an atom and } A \subseteq B} P^-(A)$$

Hint: first show that if $B \in \mathcal{F}$ then

$$P(B) = \sum_{A \textit{ is an atom with } A \subseteq B} P(A)$$

Answer. To give the answer to this question we first show that for any $B \in \mathcal{F}$, the collection of atoms which are a subset of B are a partition of B.

Firstly, we show that the atoms are pairwise disjoint: Consider atoms $A_1 \neq A_2$. Since these are $\in \mathcal{F}$, also $A_1 \cap A_2 \in \mathcal{F}$. $A_1 \cap A_2 \subseteq A_1$, so since A_1 is an atom, $A_1 \cap A_2 = \varnothing$ or $A_1 \cap A_2 = A_1$. Since A_2 is also an atom, $A_1 \not\subset A_2$, and therefore $A_1 \cap A_2 \neq A_1$. This allows us to conclude that they must be disjoint.

Let $B \in \mathcal{F}$, we show that

$$\bigcup_{A \text{ is an atom with } A \subseteq B} = B$$

Let $\omega \in B$. Then

$$\bigcap_{E \in \mathcal{F} \text{ with } \omega \in E} E$$

is an atom of $\mathcal{F}$ containing ω. This shows that $\bigcup_{A \text{ is an atom with } A \subseteq B} \supseteq B$. For each atom $A \subseteq B$, $A \cap B = A$ since otherwise A would not be an atom. Then

$$B = B \cap \left(\bigcup_{A \text{ is an atom with } A \subseteq B} A \right) \tag{30}$$

$$= \bigcup_{A \text{ is an atom with } A \subseteq B} (A \cap B) \tag{31}$$

$$= \bigcup_{A \text{ is an atom with } A \subseteq B} A \tag{32}$$

as required.

We also observe as a corollary of this, also using item 3 of question 12, for any event $B \in \mathcal{F}$,

$$P(B) = \sum_{A \text{ is an atom with } A \subseteq B} P(A) \tag{33}$$

$\Longrightarrow$:

Suppose that P satisfies the axioms. Define P^- by for A an atom of $\mathcal{F}$, $P^-(A) = P(A)$. P^- satisfies

[2]There was a part of this question which I'd forgotten to write up on the original problem sheet

- $P^-(A) \geqslant 0$ for all atoms A
- $\sum_{A \text{ is an atom of } \mathcal{F}} P^-(A) = 1$

the first by the axiom of positivity, the second by eq. (33)

We finally need to show that for any $B \in \mathcal{F}$,

$$P(B) = \sum_{A \text{ is an atom and } A \subseteq B} P^-(A)$$

but this follows directly from eq. (33) and the choice of P^-.

$\Longleftarrow$: Let $P^-(A)$ be defined on atoms with

- $P^-(A) \geqslant 0$ for all atoms A
- $\sum_{A \text{ is an atom of } \mathcal{F}} P^-(A) = 1$

Define

$$P(B) = \sum_{A \text{ is an atom and } A \subseteq B} P^-(A)$$

. We need to show that P satisfies the axioms of probability.

Positivity is easy to see.

Normalisation:

$$P(\Omega) = \sum_{A \text{ is an atom and } A \subseteq \Omega} P^-(A) = \sum_{A \text{ is an atom of } \mathcal{F}} P^-(A) = 1$$

Finite additivity: Let $B_1 \cap B_2 = \varnothing$. Then observe that there is no atom A with $A \subseteq B_1$ and $A \subseteq B_2$. Therefore

$$P(B_1 \cup B_2) = \sum_{A \text{ is an atom with } A \subseteq (B_1 \cup B_2)} P(A) \tag{34}$$

$$= \sum_{A \text{ is an atom with } A \subseteq B_1} P(A) + \sum_{A \text{ is an atom with } A \subseteq B_2} P(A) \tag{35}$$

$$= P(B_1) + P(B_2) \tag{36}$$

Question 14. Prove that if A and B are independent then from knowing $P(A)$ and $P(B)$ one can find the probabilities of all the events in the Boolean algebra generated by A and B.

Answer. Let $P(A) = a$ and $P(B) = b$. Observe that the atoms of the algebra are: $A \cap B, A \cap B^c, A^c \cap B$ and $A^c \cap B^c$. We show that we can write the probabilities of each of these in terms of the probabilities of A and B..

Since A and B are independent, we know that $P(A \cap B) = P(A) \cdot P(B) = a \cdot b$.

$P(A \cap B^c) = P(A) - P(A \cap B) = a - (a \cdot b)$ (using the axiom of finite additivity)

$P(A^c \cap B) = P(B) - P(A \cap B) = b - (a \cdot b)$ similarly

$P(A^c \cap B^c) = 1 - P(A \cup B) = 1 - (P(A) + P(B) - P(A \cap B)) = 1 - a - b + a \cdot b$.

By eq. (33) we see that this suffices to find the probabilities of the other members of the Boolean algebra by taking for any $E \in \mathcal{F}$

$$P(E) = \sum_{A \text{ is an atom with } C \subseteq E} P(C)$$

Question 15. Show that if A is probabilistically independent of B then B is probabilistically independent of A. I.e. if $P(A|B) = P(A)$ then $P(B|A) = P(B)$.

Answer. Suppose $P(A|B) = P(A)$. Then by Bayes theorem

$$P(B|A) = \frac{P(A|B) \cdot P(B)}{P(A)} \tag{37}$$

$$= \frac{P(A) \cdot P(B)}{P(A)} \tag{38}$$

$$= P(B) \tag{39}$$

as required.

Question 16. Random variable X is an *indicator function* for $A \in \mathcal{F}$ if and only if

$$X(\omega) = \begin{cases} 1 & \text{if } \omega \in A \\ 0 & \text{if } \omega \notin A. \end{cases}$$

1. Show that the expected value of the indicator function of A is the probability of A.
2. Express the indicator functions for $A \cap B$, $A \cup B$ and A^c through the indicator functions for A and B.

Answer. Let X be the indicator function for A. Then

$$\begin{aligned} E[X] &= 0 \cdot P(\{X = 0\}) + 1 \cdot P(\{X = 1\}) \\ &= P(\{X = 1\}) = P(\{\omega : X(\omega) = 1\}) = P(\{\omega : \omega \in A\}) \\ &= P(A) \ . \end{aligned}$$

This has a nice effect: It shows that the laws of large numbers can give us information about the probabilities of the events by way of the expectation values of the events' indicator functions.

Now let X be the indicator function for A and Y be the indicator function for B. The indicator function for $A \cap B$ is 1 for all and only for $\omega \in A \cap B$, otherwise its 0. Thus we need a function made up from X and Y that behaves like that. Playing around with algebraic combinations of X and Y might lead you to $X \cdot Y$, and this function does exactly what is needed: $X \cdot Y(\omega)$ is 1 for all and only for ω that are both in A and in B, since otherwise $X(\omega) = 0$ and thus $X(\omega) \cdot Y(\omega) = 0 \cdot Y(\omega) = 0$, or $Y(\omega) = 0$ and thus $X(\omega) \cdot Y(\omega) = X(\omega) \cdot 0 = 0$.

The indicator function for A^c should "flip" X, the indicator function of A: When $X(\omega) = 1$, the indicator function of A^c should have the value 0, and vice versa. One function that does this is $(1 - X)^2$.

The indicator function for $A \cup B$ must be 1 if $\omega \in A$ or $\omega \in B$. $X + Y$ almost does the trick: It is not 0 if and only if $\omega \in A \cup B$. But whenever ω is both in A and in B ($\omega \in A \cap B$), $X + Y(\omega) = 2$, which is not what we want. But we already have an indicator function for $A \cap B$, and so we can express the indicator function for $A \cup B$ as $X + Y - X \cdot Y$. Note that this expression parallels the theorem about independent probabilities $P(A \cup B) = P(A) + P(B) - P(A) \cdot P(B)$.

Question 17. [3] Let Ω be infinite. Show that there can be no probability function defined on $\mathcal{F} = \mathcal{P}(\Omega)$ where for each $\omega, \omega' \in \Omega$, $P(\{\omega\}) = P(\{\omega'\}) > 0$.

Answer. Since $\{\omega_1, \ldots, \omega_m\} \cap \{\omega_m + 1\} = \varnothing$ for any two $\omega_1, \ldots, \omega_{m+1} \in \Omega$ and $\omega_i \neq \omega_j$ when $1 \leqslant i \neq j \leqslant m+1$, $P(\{\omega_1, \ldots, \omega_{m+1}\}) = P(\{\omega_1, \ldots, \omega_m\}) + P(\{\omega_{m+1}\})$ for any $n \in \mathbb{N}^{>0}$. Therefore $P(\{\omega_1, \ldots, \omega_n\}) = \sum_{i=1}^{n} P(\{\omega_i\})$ for all n. Choose $n > \frac{1}{P(\omega_1)}$. Then $P(\{\omega_1, \ldots, \omega_n\}) = \sum_{i=1}^{n} P(\{\omega_i\}) = n \cdot P(\omega_1) > \frac{1}{P(\omega_1)} \cdot P(\omega_1) = 1$, which is incompatible with the axioms of probability.

Question 18 (The Birthday Problem)**.** There are n people in the room. Assume that peoples birthdays are equally likely to be on any day of the year. (And that for two different people, where their birthdays lie are independent). Ignore leap years.

What is the probability that at least one of them has the same birthday as you? How large does this need to be for the probability to be more than 0.5?

What is the probability that two people have their birthday on the same day? How large does this need to be for the probability to be more than 0.5?

Answer. The probability that at least one person out of n people has my birthday is $\frac{n}{365}$. For this to be > 0.5 n must be $\geqslant 183$.

Let A_k be the event that the people $p_1, \ldots, p_k$ all have different birthdays.

$P(A_1) = 1$

$P(A_{k+1}) = P(A_k) \cdot \left(1 - \frac{k}{365}\right)$

Therefore

$$P(A_n) = \prod_{0 \leqslant i < n} \left(1 - \frac{i}{365}\right) \tag{40}$$

$$= \prod_{0 \leqslant i < n} \left(\frac{365 - i}{365}\right) \tag{41}$$

$$= \frac{365!}{(365 - n)! \cdot 365^n} \tag{42}$$

By putting this into a calculator we find that when $n > 23$ this value is > 0.5.

[3]Note, this question has changed since the original problem sheet since the original question was just wrong.