Sample space and probability

The following laws of set algebra may be useful:

Commutative law

$$A \cup B = B \cup A$$
$$A \cap B = B \cap A$$

Associative law

$$(A \cup B) \cup C = A \cup (B \cup C)$$
$$(A \cap B) \cap C = A \cap (B \cap C)$$

Distributive law

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$
$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$

De Morgan's law

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

1 Sets

Consider rolling a six-sided die. Let A be the set of outcomes where the roll is an even number. Let B be the set of outcomes where the roll is greater than 3. Calculate the sets on both sides of De Morgan's Laws

$$(A \cup B)^c = A^c \cap B^c, \quad (A \cap B)^c = A^c \cup B^c$$

Solution: We have

$$A = \{2, 4, 6\}, \quad B = \{4, 5, 6\}$$

so $A \cup B = \{2, 4, 5, 6\}$, and

$$(A \cup B)^c = \{1, 3\}$$

On the other hand,

$$A^c \cup B^c = \{1, 3, 5\} \cap \{1, 2, 3\} = \{1, 3\}$$

Similarly, we have $A \cap B = \{4, 6\}$, and

$$(A \cap B)^c = \{1, 2, 3, 5\}$$

On the other hand,

$$A^c \cup B^c = \{1, 3, 5\} \cup \{1, 2, 3\} = \{1, 2, 3, 5\}$$

2 Ghostbusting

Twenty ghostbusters are on their annual camping retreat. Two of them, Abe and Betty, have discovered that another pair, Candace and Dan, are in fact ghosts posing as ghostbusters. Abe and Betty hatch a plan: When all 20 campers are sitting in a circle around the campfire, Abe will fire his proton pack at Candace, and Betty will simultaneously fire her proton pack at Dan, annihilating the ghosts. However, if two proton streams cross (that is, if the paths of their weapons intersect), it means the end of all life on Earth.

If the ghostbusters are arranged randomly around the fire, what are the chances that Abe and Betty will cross streams?

Solution: The chances are $\frac{1}{3}$.

There are 20 ghostbusters, but we only really care about four of them - Abe, Betty, Candace and Dan. The position of the other 16 won't affect the possible crossing of the streams, so let's ignore them.

Fix Abe's spot at the campfire - say he's on the north side. There are then three places his co-ghostbuster Betty can sit - east, west or south. The ghosts, Candace and Dan, will sit in the other two seats. There are $3 \times 2 \times 1 = 6$ possibilities for the seating in the east, west and south seats. In exactly two of these arrangements - the two where Candace occupies the southern seat, forcing Abe to fire across the circle - the ghostbusters will cross their proton streams. Each of these six arrangements is equally likely, so the chances of stream-crossing disaster are $\frac{2}{6} = \frac{1}{3}$.

3 Calculate probabilities in a sample space S

Events A and B are contained within a sample space S. Given that $\Pr(A) = 0.5$, $\Pr(B) = 0.3$ and $\Pr(A \cap B) = 0.1$, find:

a. $\Pr(A \cup B)$

$$\Pr(A \cup B) = \Pr(A) + P(B) - \Pr(A \cap B) = 0.5 + 0.3 - 0.1 = 0.7$$

b. $\Pr(A \cap B^c)$

$$\begin{aligned} \Pr(A) &= \Pr(A \cap B) + \Pr(A \cap B^c) \\ \Pr(A \cap B^c) &= \Pr(A) - \Pr(A \cap B) \\ &= 0.5 - 0.1 = 0.4 \end{aligned}$$

c. $\Pr[(A \cap B^c) \cup (B \cap A^c)]$

$$\begin{aligned}
\Pr[(A \cap B^c) \cup (B \cap A^c)] &= \Pr(A \cap B^c) + \Pr(B \cap A^c) - \Pr(A \cap B^c \cap B \cap A^c) \\
&= \Pr(A \cap B^c) + \Pr(B \cap A^c) - \Pr(\emptyset) \\
&= \Pr(A) - \Pr(A \cap B) + \Pr(B) - \Pr(A \cap B) \\
&= 0.5 - 0.1 + 0.3 - 0.1 = 0.6
\end{aligned}$$

4 Silly campaigns, polling is for political scientists

A political campaign in New Haven, CT decides to conduct an "experiment" to determine the effectiveness of knocking on a door in turning a resident of that house out to vote. The campaign foolishly denies an offer from a team of political scientists to help them design a protocol for this experiment, and instead directs their two teams of volunteers to each select a random group of the 120 total houses in the district and to go knock on as many of those random doors as they can in the week before the election. The campaign manager directs the teams to count the number of doors on which they knock and to record the names of the residents who live in each house, but neglects to ensure that the two teams select a mutually exclusive set of houses, or to set bounds on how many houses each team chooses.

On election day, the Team 1 members return, and proudly report to the campaign manager that they knocked on 70% of the doors in the electoral district. The Team 2 members return shortly after, and report that they knocked on 40% of the doors in the electoral district. In looking over the names the teams recorded, the campaign manager quickly determines that not only was every house in the district contacted, but some houses were contacted by both teams. (This will make drawing inferences about the effectiveness of door knocking ... difficult.)

Use what we have learned about probability to determine how many houses had their doors knocked on by both teams.

Solution: Let A be the event that Team 1 knocked on a given door and B be the event that Team 2 knocked on a given door. We know that $\Pr(A \cup B) = 1$ since every house had their door knocked on at least once. Further, we know that $\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B)$. Plugging in, we get

$$\begin{aligned}
1 &= .7 + .4 - \Pr(A \cap B) \\
-.1 &= -\Pr(A \cap B) \\
.1 &= \Pr(A \cap B)
\end{aligned}$$

Since there are 120 doors in the district, this means that $.1 \times 120 = 12$ doors were knocked on by both teams.

5 Geniuses and chocolate

Out of the students in a class, 60% are geniuses, 70% love chocolate, and 40% fall into both categories. Determine the probability that a randomly selected student is neither a genius nor a chocolate lover.

Solution: Let G and C be the events that the chosen student is a genius and a chocolate lover, respectively. We have $\Pr(G) = 0.6$, $\Pr(C) = 0.7$, and $\Pr(G \cap C) = 0.4$. We are interested in $\Pr(G^c \cap C^c)$, which is obtained with the following calculation:

$$\begin{aligned}
\Pr(G^c \cap C^c) &= 1 - \Pr(G \cup C) \\
&= 1 - (\Pr(G) + \Pr(C) - \Pr(G \cap C)) \\
&= 1 - (0.6 + 0.7 - 0.4) \\
&= 0.1
\end{aligned}$$

6 Rolling the dice

We roll two fair 6-sided dice. Each one of the 36 possible outcomes is assumed to be equally likely.

a. Find the probability that doubles are rolled.

   Each possible outcome has probability 1/36. There are 6 possible outcomes that are doubles, so the probability of doubles is $6/36 = 1/6$.

b. Given that the roll results in a sum of 4 or less, find the conditional probability that doubles are rolled.

   The conditioning event (sum is 4 or less) consists of the 6 outcomes

$$\{(1,1),(1,2),(1,3),(2,1),(2,2),(3,1)\}$$

   2 of which are doubles, so the conditional probability of doubles is $2/6 = 1/3$.

c. Find the probability that at least one die roll is a 6.

   There are 11 possible outcomes with at least one 6, namely $(6,6),(6,i),(i,6)$ for $i = 1,2,\ldots,5$. Thus, the probability that at least one die is a 6 is 11/36.

d. Given that the two dice land on different numbers, find the conditional probability that at least one die roll is a 6.

   There are 30 possible outcomes where the dice land on different numbers. Out of these, there are 10 outcomes in which at least one of the rolls is a 6. Thus, the desired conditional probability is $10/30 = 1/3$.

7 A two-envelope puzzle

The release of two out of three prisoners has been announced, but their identity is kept secret. One of the prisoners considers asking a friendly guard to tell him who is the prisoner other than himself that will be released, but hesitates based on the following rationale: at the prisoner's present state of knowledge, the probability of being released is 2/3, but after he knows the answer, the probability of being released will become 1/2, since there will be two prisoners (including himself) whose fate is unknown and exactly one of the two will be released. What is wrong with this line of reasoning?

Solution: Intuitively, there is something wrong with this rationale. The reason is that it is not based on a correctly specified probabilistic model. In particular, the event where both of the other prisoners are to be released is not properly accounted in the calculation of the posterior probability of release.

To be precise, let A, B, and C be the prisoners, and let A be the one who considers asking the guard. Suppose that all prisoners are a priori equally likely to be released. Suppose also that if B and C are to be released, then the guard chooses B or C with equal probability to reveal to A. Then, there are four possible outcomes:

1. A and B are to be released, and the guard says B (probability 1/3).
2. A and C are to be released, and the guard says C (probability 1/3).
3. B and C are to be released, and the guard says B (probability 1/6).
4. B and C are to be released, and the guard says C (probability 1/6).

Thus,

$$\begin{aligned}\Pr(\text{A is to be released}|\text{guard says B}) &= \frac{\Pr(\text{A is to be released and guard says B})}{\Pr(\text{guard says B})} \\ &= \frac{1/3}{1/3+1/6} = \frac{2}{3}\end{aligned}$$

Similarly,

$$\Pr(\text{A is to be released}|\text{guard says C}) = \frac{2}{3}$$

Thus, regardless of the identity revealed by the guard, the probability that A is released is equal to 2/3, the a priori probability of being released.

8 The birthday problem

Consider n people who are attending a party. We assume that every person has an equal probability of being born on any day during the year, independent of everyone else, and ignore the additional complication presented by leap years (i.e. assume that nobody is born on February 29). What is the probability that each person has a distinct birthday?

Solution: The sample space consists of all possible choices for the birthday of each person. Since there are n persons, and each has 365 choices for their birthday, the sample space has 365^n elements.

Let us now consider those choices of birthdays for which no two persons have the same birthday. Assuming that $n \leq 365$, there are 365 choices for the first person, 364 for the second, etc., for a total of $365 \times 364 \times \cdots (365 - n + 1)$. Thus,

$$\Pr(\text{no two birthdays coincide}) = \frac{365 \times 364 \times \cdots (365 - n + 1)}{365^n}$$