2 Topics: Counting, axiomatic definition of probability, conditional probability

2.1 Prerequisites: Lecture 4

Exercise 2- 1: (Suggested for personal/peer tutorial)

Explain, without direct calculation that, for $k, N \in \mathbb{N}, k \leq N$,

$$\sum_{n=k}^{N} \binom{n}{k} = \binom{N+1}{k+1}.$$

Use a proof where you only comment on sampling from sets of an appropriate cardinality. You might want to write out the sum on the left hand side as

$$\sum_{n=k}^{N} \binom{n}{k} = \binom{k}{k} + \binom{k+1}{k} + \cdots + \binom{N}{k}.$$

Solution: The right hand side $\binom{N+1}{k+1}$ gives us the number of possibilities of selecting $k+1$ elements from the set with labels $\{1, \ldots, N+1\}$. Hence, the maximal element selected will take values in $\{k+1, \ldots, N+1\}$.

- How many possibilities are there such that $k+1$ is in fact the maximal element in the selection? – 1 (when we select $\{1, 2, \ldots, k+1\}$), which corresponds to $\binom{k}{k}$.
- How many possibilities are there such that $k+2$ is in fact the maximal element in the selection? – Here we pick k out of the set $A = \{1, \ldots, k+1\}$, with $\mathrm{card}(A) = k+1$ and we select $k+2$. So there are $\binom{k+1}{k}$ possibilities of having $k+2$ as the maximal element in the selection.
- More generally, for any $n \in \{k, \ldots, N\}$, there are $\binom{n}{k}$ possibilities of selecting $k+1$ elements such that $n+1$ is the maximal selected element.
- Now we only need to sum over all possible $n \in \{k, \ldots, N\}$ to conclude that $\sum_{n=k}^{N} \binom{n}{k} = \binom{N+1}{k+1}$.

2.2 Prerequisites: Lecture 5

Exercise 2- 2: Given two events $E, F \subseteq \Omega$, prove that the probability of *one and only one* of them occurring is

$$\mathrm{P}(E) + \mathrm{P}(F) - 2\mathrm{P}(E \cap F).$$

Solution: Let G="exactly one occurs". Then $G = (E \cap F^c) \cup (E^c \cap F)$, where $(E \cap F^c) \cap (E^c \cap F) = \emptyset$, and by Axiom (iii)

$$\mathrm{P}(G) = \mathrm{P}(E \cap F^c) + \mathrm{P}(E^c \cap F).$$

Also, using that $\Omega = F \cup F^c$ and $\Omega = E \cup E^c$

$$E = E \cap \Omega = E \cap (F \cup F^c) = (E \cap F) \cup (E \cap F^c), \quad \text{where } (E \cap F) \cap (E \cap F^c) = \emptyset,$$

hence, by axiom (iii) of the definition of a probability measure,

$$\mathrm{P}(E) = \mathrm{P}(E \cap F) + \mathrm{P}(E \cap F^c) \Longrightarrow \mathrm{P}(E \cap F^c) = \mathrm{P}(E) - \mathrm{P}(E \cap F).$$

Similarly,

$$F = F \cap \Omega = F \cap (E \cup E^c) = (E \cap F) \cup (E^c \cap F), \quad \text{where } (E \cap F) \cap (E^c \cap F) = \emptyset,$$

hence, by axiom (iii) of the definition of a probability measure,

$$\mathrm{P}(F) = \mathrm{P}(E \cap F) + \mathrm{P}(E^c \cap F) \Longrightarrow \mathrm{P}(E^c \cap F) = \mathrm{P}(F) - \mathrm{P}(E \cap F).$$

Altogether, we can deduce that

$$\mathrm{P}(G) = \mathrm{P}(E \cap F^c) + \mathrm{P}(E^c \cap F) = \mathrm{P}(E) + \mathrm{P}(F) - 2\mathrm{P}(E \cap F).$$

Exercise 2- 3: Consider the following statements, which are claimed to be true for events A_1, A_2 in a sample space Ω:

$$\begin{array}{lll}
(a) & \mathrm{P}(A_1) = 0 & \Longrightarrow \mathrm{P}(A_1 \cup A_2) = 0 \\
(b) & \mathrm{P}(A_1) = \mathrm{P}(A_2^c) & \Longrightarrow A_1^c = A_2 \\
(c) & A_1 \subseteq A_2 \text{ and } \mathrm{P}(A_1) = \mathrm{P}(A_2^c) & \Longrightarrow \mathrm{P}(A_1) \leq 1/2 \\
(d) & \mathrm{P}(A_1^c) = x_1,\ \mathrm{P}(A_2^c) = x_2 & \Longrightarrow \mathrm{P}(A_1 \cup A_2) \geq 1 - x_1 - x_2
\end{array}$$

In each case, either prove that the statement is true for all Ω, A_1, A_2, or provide a specific counter-example to show that there exists Ω, A_1, A_2 for which it is false

Solution:

(a) FALSE : (e.g. $A_1 = \emptyset$, $A_2 = \Omega$).

(b) FALSE : (e.g. coin toss, let $A_1 = A_2 = \{H\}$).

(c) TRUE : $A_1 \subseteq A_2 \Longrightarrow \mathrm{P}(A_1) \leq \mathrm{P}(A_2) = 1 - \mathrm{P}(A_2^c) = 1 - \mathrm{P}(A_1) \Longrightarrow 2\mathrm{P}(A_1) \leq 1$.

(d) TRUE : $\mathrm{P}(A_1 \cup A_2) = \mathrm{P}(A_1) + \mathrm{P}(A_2) - \mathrm{P}(A_1 \cap A_2) \geq (1 - x_1) + (1 - x_2) - 1 = 1 - x_1 - x_2$.

Exercise 2- 4: Show the so-called *Boole's inequality*: For any events $A_1, \ldots, A_n$ with $n \in \mathbb{N}$, we have

$$\mathrm{P}\left(A_1 \cup A_2 \cup \cdots \cup A_n\right) \leq \mathrm{P}(A_1) + \cdots + \mathrm{P}(A_n).$$

Solution: We prove this by induction. The inequality holds for $n = 2$ as $\mathrm{P}(A \cup B) = \mathrm{P}(A) + \mathrm{P}(B) - \mathrm{P}(A \cap B) \leq \mathrm{P}(A) + \mathrm{P}(B)$. So assume true for n; then

$$\mathrm{P}(A_1 \cup ... \cup A_n \cup A_{n+1}) \leq \mathrm{P}(A_1 \cup ... \cup A_n) + \mathrm{P}(A_{n+1}) \leq \sum_{i=1}^{n+1} \mathrm{P}(A_i)$$

and hence true for $n + 1$.

Exercise 2- 5: Show the so-called *inclusion-exclusion principle*: For any events $A_1, \ldots, A_n$ with $n \in \mathbb{N}$, we have

$$\begin{aligned}
&\mathrm{P}\left(A_1 \cup A_2 \cup \cdots \cup A_n\right) \\
&= \sum_{1 \leq i \leq n} \mathrm{P}\left(A_i\right) - \sum_{1 \leq i_1 < i_2 \leq n} \mathrm{P}\left(A_{i_1} \cap A_{i_2}\right) + \sum_{1 \leq i_1 < i_2 < i_3 \leq n} \mathrm{P}\left(A_{i_1} \cap A_{i_2} \cap A_{i_3}\right) - \cdots \\
&\quad + (-1)^n \sum_{1 \leq i_1 < i_2 < \cdots < i_{n-1} \leq n} \mathrm{P}\left(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_{n-1}}\right) \\
&\quad + (-1)^{n+1} \mathrm{P}\left(A_1 \cap A_2 \cap \cdots \cap A_n\right).
\end{aligned}$$

Solution: This can be proved by induction. The case when $n = 2$ is covered by Theorem 4.2.3. Assume now that the formula holds for n and we want to show that it holds for $n + 1$. Then we use first the associative law for sets, then Theorem 4.2.3 and finally the distributive law for sets to conclude that

$$\begin{aligned}
&\mathrm{P}\left(A_1 \cup A_2 \cup \cdots \cup A_n \cup A_{n+1}\right) \\
&= \mathrm{P}\left(\left(A_1 \cup A_2 \cup \cdots \cup A_n\right) \cup A_{n+1}\right) \\
&= \mathrm{P}\left(A_1 \cup A_2 \cup \cdots \cup A_n\right) + \mathrm{P}\left(A_{n+1}\right) - \mathrm{P}\left(\left(A_1 \cup A_2 \cup \cdots \cup A_n\right) \cap A_{n+1}\right) \\
&= \mathrm{P}\left(A_1 \cup A_2 \cup \cdots \cup A_n\right) + \mathrm{P}\left(A_{n+1}\right) - \mathrm{P}\left(\left(A_1 \cap A_{n+1}\right) \cup \left(A_2 \cap A_{n+1}\right) \cup \cdots \cup \left(A_n \cap A_{n+1}\right)\right)
\end{aligned}$$

We note that the first and the last term contains unions of n sets to which we can apply our induction hypothesis. In particular, we have

$$\begin{aligned}
&\mathrm{P}\left(A_1 \cup A_2 \cup \cdots \cup A_n\right) && (2.1) \\
&= \sum_{1 \leq i \leq n} \mathrm{P}\left(A_i\right) && (2.2) \\
&- \sum_{1 \leq i_1 < i_2 \leq n} \mathrm{P}\left(A_{i_1} \cap A_{i_2}\right) && (2.3) \\
&+ \sum_{1 \leq i_1 < i_2 < i_3 \leq n} \mathrm{P}\left(A_{i_1} \cap A_{i_2} \cap A_{i_3}\right) - \cdots && (2.4) \\
&+ (-1)^n \sum_{1 \leq i_1 < i_2 < \cdots < i_{n-1} \leq n} \mathrm{P}\left(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_{n-1}}\right) && (2.5) \\
&+ (-1)^{n+1} \mathrm{P}\left(A_1 \cap A_2 \cap \cdots \cap A_n\right) && (2.6)
\end{aligned}$$

and

$$\begin{aligned}
&- \mathrm{P}\left(\left(A_1 \cap A_{n+1}\right) \cup \left(A_2 \cap A_{n+1}\right) \cup \cdots \cup \left(A_n \cap A_{n+1}\right)\right) \\
&= - \sum_{1 \leq i \leq n} \mathrm{P}\left(A_i \cap A_{n+1}\right) && (2.7)
\end{aligned}$$

$$+ \sum_{1\le i_1<i_2\le n} \mathrm{P}\left(A_{i_1}\cap A_{i_2}\cap A_{n+1}\right) \tag{2.8}$$
$$- \sum_{1\le i_1<i_2<i_3\le n} \mathrm{P}\left(A_{i_1}\cap A_{i_2}\cap A_{i_3}\cap A_{n+1}\right)+\cdots \tag{2.9}$$
$$-(-1)^n \sum_{1\le i_1<i_2<\cdots<i_{n-1}\le n} \mathrm{P}\left(A_{i_1}\cap A_{i_2}\cap\cdots\cap A_{i_{n-1}}\cap A_{n+1}\right) \tag{2.10}$$
$$-(-1)^{n+1}\mathrm{P}\left(A_1\cap A_2\cap\cdots\cap A_n\cap A_{n+1}\right). \tag{2.11}$$

Now we need to add all terms together: $\mathrm{P}(A_{n+1})$ added to (2.2) gives $\sum_{1\le i\le n+1}\mathrm{P}\left(A_i\right)$, the sum of (2.3) and (2.7) give all probabilities of intersections of two sets, the sum of (2.4) and (2.8) give all probabilities of intersections of three sets ect., and the sum of (2.6) and (2.10) give all probabilities of intersections of n sets and (2.11) is the final term including an intersection of $n+1$ sets. So overall, we have

$$\begin{aligned}
&\mathrm{P}\left(A_1\cup A_2\cup\cdots\cup A_n\cup A_{n+1}\right)\\
&= \sum_{1\le i\le n+1}\mathrm{P}\left(A_i\right) - \sum_{1\le i_1<i_2\le n+1}\mathrm{P}\left(A_{i_1}\cap A_{i_2}\right) + \sum_{1\le i_1<i_2<i_3\le n+1}\mathrm{P}\left(A_{i_1}\cap A_{i_2}\cap A_{i_3}\right)-\cdots\\
&\quad +(-1)^{n+1}\sum_{1\le i_1<i_2<\cdots<i_n\le n+1}\mathrm{P}\left(A_{i_1}\cap A_{i_2}\cap\cdots\cap A_{i_n}\right)\\
&\quad +(-1)^{n+2}\mathrm{P}\left(A_1\cap A_2\cap\cdots\cap A_{n+1}\right),
\end{aligned}$$

which means that the formula indeed holds for $n+1$.

2.3 Prerequisites: Lecture 6

Exercise 2- 6: Consider a standard 52-card deck which has been shuffled well. You pick two cards at random, one at a time without replacement. We denote by A the event that the first card is a spade and by B the event that the second card is black. Find $\mathrm{P}(A|B)$ and $\mathrm{P}(B|A)$.

Solution: We use the naive definition of probability and the multiplication rule: We have

$$\mathrm{P}(A)=\frac{1}{4},$$

since all four suits are equally likely. Further, we have

$$\mathrm{P}(A\cap B)=\frac{13\cdot 25}{52\cdot 51}=\frac{25}{204},$$

since we have 13 possibilities of choosing a spade and then 25 remaining black cards for the second draw. Also,

$$\mathrm{P}(B)=\frac{26\cdot 51}{52\cdot 51}=\frac{1}{2},$$

since there are 26 black cards to choose from for the second draw and the first card can be any other card. [We do not need the chronological order in the multiplication rule.] Then

$$\mathrm{P}(A|B)=\frac{\mathrm{P}(A\cap B)}{\mathrm{P}(B)}=\frac{25/204}{1/2}=\frac{25}{102}\approx 0.245,$$

and

$$\mathrm{P}(B|A) = \frac{\mathrm{P}(A \cap B)}{\mathrm{P}(A)} = \frac{25/204}{1/4} = \frac{25}{51} \approx 0.490.$$