Basics of Probability

Let us start with an example.

Example. Roll two dice, what is the probability that the sum is 6 ?

Solution. One can easily see that there are 36 combinations with equal chance, while there are only 5 of them that give us the required sum:

$$(1,5) \quad (2,4) \quad (3,3) \quad (4,2) \quad (5,1)$$

So the probability is $5/36$.

The main ideas we can get from this experiment are:

- The collection of all possible outcomes$\longrightarrow$ ***Sample space/Probability space***, denoted by Ω. In this example, we write

$$\Omega = \Big\{ (1,1),(1,2),\dots,(1,6),(2,1),(2,2),\dots,(2,6),\dots,(3,1),\dots,(6,6) \Big\}.$$

- The collection of some outcomes from the experiment$\longrightarrow$ ***Events***, usually denoted by capital letters. In this example, we write

$$A = \big\{(1,5),(2,4),(3,3),(4,2),(5,1)\big\} = \big\{ \text{ the sum of rolling two dice is 6 } \big\}.$$

  One should notice that events are subsets of the sample space. In particular, the collection of all subsets of Ω is denoted by 2^Ω and is called the ***power set of*** Ω. See problem (1).

- From this example, one can "*guess*" that a "*probability*" is an assignment of a number in $[0,1]$ to an event, which can be informally written as :

$$\mathbb{P} : 2^\Omega \to \big[0,1\big], \quad \textit{e.g.} \quad \mathbb{P}(A) = 5/36.$$

- Something beyond this experiment, which would be given later, is the concept of random variables. For starter, the random variable $X : \Omega \to \mathbb{R}$ is just a function, which can be understood as a gambling profit in this example if formulated correctly. (*Give a try*!)

Problem (1) Prove that for a set A with m elements, its power set 2^A has 2^m elements. (This explains why we use the notation 2^Ω to denote the power set of Ω.)

Now we will begin to state the rigorous definitions. To illustrate mainly the ideas and not focus on the technical details, we only consider the case where the sample space Ω is a nonempty finite set.

Definition 1. Fix an integer $m \geq 1$ and assume that our sample space $\Omega = \{\omega_1,\dots,\omega_m\}$. We call a function $\mathbb{P} : 2^\Omega \to [0,1]$ a probability on Ω if it satisfies the following axioms:

1. $\mathbb{P}(\emptyset) = 0$.

   Some explanation: the notation $\emptyset$ stands for the empty set, which is a set that contains nothing. In probability, $\emptyset$ is an event which for sure does not happen, so its probability must be zero. Does this make sense?

2. Given two events $A \subset B \subset \Omega$, we have $\mathbb{P}(A) \leq \mathbb{P}(B) \leq 1$.

   Some explanation: we have already mentioned in the above example that the event is a subset of our sample space. To be clear, ω_1 is an element of Ω but $\{\omega_1\}$ is a subset of Ω, so $\{\omega_1\}$ is an element of 2^Ω. (Be careful with the brackets { }!)

   $A \subset B$ means that either

   (i) $A = \emptyset$ (as the empty set is a subset of any set.)

   or

   (ii) $A \neq \emptyset$, and any element in A is an element of B. In probability, $A \subset B$ can be understood that if the event A happens, then B must also happen. So that it is reasonable to ask for $\mathbb{P}(B) \geq \mathbb{P}(A)$.

3. $\mathbb{P}(\Omega) = 1$.

   Some explanation: in probability, the whole space Ω is seen as the ***sure event*** that happens with probability 1.

4. $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$ for any two events A, B such that $A \cap B = \emptyset$.

   Some explanation: $A \cap B$ is the intersection of two sets A, B, that is, it is the set of common elements in A and B. Meanwhile, $A \cup B$ is the union of two sets A, B, that is, it is the set that consists of all elements in A ***and*** all elements in B (and nothing else). In probability, $A \cap B$ means that the two events both happen and $A \cup B$ means that the event A or[1] the event B happens. In the case where $A \cap B = \emptyset$ (that is they are disjoint events), the probability that the event A or the event B happens is exactly the probability that A happens plus the probability that B happens. This explains our axiom 4.

Problem (2) Let $\Omega, \mathbb{P}$ be given as in **Definition 1** and $F \subset \Omega$ be such that $\mathbb{P}(F) > 0$. Now we define $Q : 2^\Omega \to [0, 1]$ by setting

$$Q(A) = \frac{\mathbb{P}(F \cap A)}{\mathbb{P}(F)}$$

for each $A \subset \Omega$. Prove that Q is also a probability on Ω.

Hint: You may want to use the de Morgan's law, that is, for any $A, B \subset \Omega$, we have $F \cap (A \cup B) = (F \cap A) \cup (F \cap B)$.

[1] Note A and B can both happen in this situation

Problem (3) Manipulating the set operations is a very basic skill in probability theory. So in this question, we ask students to prove the de Morgan's law: let A, B, C be three sets, then prove the following relations:

- $C \cap (A \cup B) = (C \cap A) \cup (C \cap B)$ • $C \cup (A \cap B) = (C \cup A) \cap (C \cup B)$.

To prove the equality of two sets, it suffices to prove any of them is contained in the other.

Once you have finished the second problem. That is, you prove that Q is a probability on Ω. This probability bears a particular name "the conditional probability on event F" and we write $Q(A) = \mathbb{P}(A|F)$ to specify its relation with the event F. Intuitively speaking, $\mathbb{P}(A|F)$ *in general* is NOT the probability that the event A happens, but it is the probability that the event A happens under the condition that the event F happens. For example, if $F \subset A$, then $\mathbb{P}(A|F) = 1$, which follows from the above definition as well as the intuition that if F happens, then A must happen[2].

What if $\mathbb{P}(A|F) = \mathbb{P}(A)$? What does this equality tell us?

The concept of conditional probability is closely related to another important notion in probability, see the following definition.

Definition 2. Let $\Omega, \mathbb{P}$ be given as in **Definition 1**, we say that two events A and B are independent with respect to the probability $\mathbb{P}$ if any of the following two conditions holds true

(i) $\mathbb{P}(A) = 0$ or $\mathbb{P}(B) = 0$.

(ii) Both A and B happen with strictly positive probability, that is, $\mathbb{P}(A) > 0$ and $\mathbb{P}(B) > 0$. We have $\mathbb{P}(A|B) = \mathbb{P}(A)$ and $\mathbb{P}(B|A) = \mathbb{P}(B)$.

Explanation of (i) and (ii) in the above definition: (i) when the event A has zero probability, then it will not happen for sure, so it is reasonable to say that "*whether the event A happens or not*" does not affect the chance for B to happen; for short, we just say A is independent from B.
(ii) Now suppose $\mathbb{P}(A) > 0$ and $\mathbb{P}(B) > 0$, then the conditional probabilities $\mathbb{P}(\cdot|A)$ and $\mathbb{P}(\cdot|B)$ are well defined. Assume that $\mathbb{P}(B|A) = \mathbb{P}(B)$: this means that conditional on the event A, the probability that B happens is exactly $\mathbb{P}(B)$; in other words, the occurrence of A does not influence that of B, that is, we can say A is independent from B.

Problem (4) (a) Let $\Omega, \mathbb{P}$ be given as in **Definition 2**, show that two events A, B are independent if and only if $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$.

This equivalent is more often used as the definition of independence between two events.

[2]For example: F denotes the event that you will get a basketball tomorrow and A denotes the event that you will get a ball tomorrow, then if F will happen, then for sure A will happen.

Problem (4) (b) Let $\Omega, \mathbb{P}$ be given as in **Definition 2** and two events A, B have strictly positive probabilities. Show that $\mathbb{P}(A|B) = \mathbb{P}(A)$ if and only if $\mathbb{P}(B|A) = \mathbb{P}(B)$.

This equivalent tells us the independence is commutative.

Problem (5) An urn contains eleven balls, six of which are white and five black. We draw three balls, what is the probability that one of them is white and two are black?

Please first formulate the sample space.

Solutions
(1) if $m = 0$, then the empty set is the only element of $2^{\emptyset}$, so the problem is solved; if $m \geq 1$, say, the elements are labelled as $x_1, \ldots, x_m$. When you pick elements from A to form a subset of A, you will decide to pick x_1 or not, pick x_2 or not, ..., pick x_m or not. For each element, you have 2 choices, together you will have $2 \times 2 \times \ldots 2$, product of m times 2, which is equal to 2^m.

(2) We need to check four conditions: $Q(\emptyset) = \frac{\mathbb{P}(F\cap\emptyset)}{\mathbb{P}(F)} = \frac{\mathbb{P}(\emptyset)}{\mathbb{P}(F)} = 0$. For $A \subset B \subset \Omega$, we have $F \cap A \subset F \cap B \subset F$ (why?), so $Q(A) \leq Q(B) \leq 1$. $Q(\Omega) = \frac{\mathbb{P}(F\cap\Omega)}{\mathbb{P}(F)} = \frac{\mathbb{P}(F)}{\mathbb{P}(F)} = 1$, because $F \cap \Omega = F$.
Now assume A, B are two disjoint events, then $F \cap A$ and $F \cap B$ are also disjoint (why?), then by de Morgan's law, we have

$$Q(A \cup B) = \frac{\mathbb{P}(F \cap (A \cap B))}{\mathbb{P}(F)} = \frac{\mathbb{P}\big((F \cap A) \cup (F \cap B)\big)}{\mathbb{P}(F)} = \frac{\mathbb{P}(F \cap A)}{\mathbb{P}(F)} + \frac{\mathbb{P}(F \cap B)}{\mathbb{P}(F)}$$

which is equal to $Q(A) + Q(B)$. This concludes our proof.

(3) The proof is standard. Once I receive students' feedback, I will write the complete solution down. Here I just prove the first equality in the case where all the sets appearing in the formula are not empty:

$$\begin{aligned} x \in C \cap (A \cup B) &\Leftrightarrow x \in C \quad \textbf{and} \quad x \in A \cup B \\ &\Leftrightarrow x \in C \quad \textbf{and} \quad \underline{x \in A \text{ or } x \in B} \\ &\Leftrightarrow \underline{x \in C \text{ and } x \in A} \quad \textbf{or} \quad \underline{x \in C \text{ and } x \in B} \\ &\Leftrightarrow x \in C \cap A \quad \textbf{or} \quad x \in C \cap B \\ &\Leftrightarrow x \in \big(C \cap A\big) \cup \big(C \cap B\big). \end{aligned}$$

(4) When $\mathbb{P}(A), \mathbb{P}(B)$ are strictly positive, then the equivalence follows from the definition; when one of them is zero, say $\mathbb{P}(A) = 0$, then $\mathbb{P}(A)\mathbb{P}(B) = 0$, while $\mathbb{P}(A \cap B) \leq \mathbb{P}(A) = 0$, it must hold that $\mathbb{P}(A \cap B) = 0$. This finishes the proof for (a). The part (b) follows from the following equalities:

$$\mathbb{P}(A|B) = \mathbb{P}(A) \Leftrightarrow \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \mathbb{P}(A) \Leftrightarrow \mathbb{P}(A \cap B) = \mathbb{P}(A) \times \mathbb{P}(B) \Leftrightarrow \frac{\mathbb{P}(B \cap A)}{\mathbb{P}(A)} = \mathbb{P}(B)$$

$\Leftrightarrow \mathbb{P}(B|A) = \mathbb{P}(B)$. The proof of problem (4) is completed now.

(5) We model the sample space as all ***ordered*** draws (hence triplets): more precisely, we label five black balls by b_1, b_2, b_3, b_4, b_5 and six white balls by $w_1, \ldots, w_6$. Then there are $11 \times 10 \times 9 = 990$ possible outcomes of ordered triplets[3] . Now we count through the different possible configurations.

- First ball white, second and third ball black: $6 \times 5 \times 4 = 120$
- First ball black, second ball white, and third ball black: $5 \times 6 \times 4 = 120$
- First and second ball black, third ball white: $5 \times 4 \times 6 = 120$

So the probability we want is equal to

$$\frac{120 + 120 + 120}{990} = \frac{4}{11}.$$

ANY QUESTION? JUST ASK!

[3]Note in this way, the outcome (b_1, b_2, w_1) is different from the outcome (b_2, b_1, w_1).