Introduction to probability, probability axioms

Saad Mneimneh

1 Introduction and probability axioms

If we make an observation about the world, or carry out an experiment, the outcome will always depend on chance to a varying degree. Think of the weather, the stock market, or a medical experiment. Probability is a mathematical tool to model this dependence on chance.

We start by listing all possible outcomes of the experiment. These possible outcomes form a set S that we call the sample space. Perhaps the most classical experiment is tossing a coin. This has two outcomes (well, assuming the coin does not land on its edge): In this case $S = \{H, T\}$. As an another example, the outcomes of throwing a die form the set $S = \{., .., ..., ..., ..., ...\}$.

Every subset of S is called an event. An example is the event $E = \{.., ..., ...\}$ which represents the event that we throw an even number of dots.

Given a sample space S (which could be infinite), we assign for each outcome s_i is S a "probability" $P(\{s_i\})$ (we then have a probability space) such that:

Probability Axioms

1. $P(s_i) \equiv P(\{s_i\}) \geq 0$ for all $s_i \in S$
2. $P(A \cup B) = P(A) + P(B)$, where $A \cap B = \emptyset$ (exclusive events)
3. $P(S) = 1$

The second axiom suggests that, given an event $E = \{s_1, \dots, s_k\}$, $P(E) = P(s_1) + \dots P(s_k)$, because $E = \{s_1\} \cup \{s_2\} \dots \cup \{s_k\}$ and these are exclusive events. Therefore, by the third axiom $\sum_{si \in S} P(s_i) = 1$. We can also show that $p(\emptyset) = 0$.

The second axiom needs strengthening to handle infinite sample spaces:

Probability Axioms

1. $P(s_i) \equiv P(\{s_i\}) \geq 0$ for all $s_i \in S$
2. $P(A_1 \cup A_2 \cup \dots) = P(A_1) + P(A_2) + \dots$, where $A_i \cap A_j = \emptyset$ (pairwise exclusive events, countably many)
3. $P(S) = 1$

If the probability space is uniform, then $P(s_1) = \dots = P(s_k) = 1/k$ and hence for every event E, $P(E) = |E|/|S|$.

2 Independence and conditioning

We define the probability of A conditioned on B (the probability of A given that B occurred) as:

$$P(A|B) = \frac{P(A,B)}{P(B)}$$

where $P(A,B)$ stands for $P(A \cap B)$, and $P(B) > 0$.

In other words, B acts as our new sample space. Here's an example: Consider the two events when rolling a fair die (uniform probability space):

$$A = \{..,....,......\}$$

$$B = \{...,......\}$$

We can compute the following probabilities: $P(A,B) = 1/6$, $P(A) = 1/2$, and $P(B) = 1/3$. Observe that $P(A,B) = P(A)P(B)$. This is not necessarily true for any two events.

Two events A and B are independent iff $P(A,B) = P(A)P(B)$, i.e. according to our definition above, $P(A) = P(A|B)$ (this also means that $P(B) = P(B|A)$). This definition of independence is motivated by the fact that knowing that B occurred does not change the probability of A: the "size" of $A \cap B$ relative to B is the same as the "size" of A with respect to the entire sample space S.

Note that two events that are independent, may become dependent after conditioning. For instance, while A and B are the events defined above, consider the event $C = \{.,..,...,....,.....\}$.

$P(A) = 1/2$

$P(B) = 1/3$

$P(A,B) = 1/6$ (independent)

$P(C) = 5/6$

$P(A|C) = P(A,C)/P(C) = (1/3)/(5/6) = 2/5$

$P(B|C) = 1/5$

$P(A,B|C) = P(A,B,C)/P(C) = P(\emptyset)/P(C) = 0$

Note that $P(A,B|C) \neq P(A|C)P(B|C)$, so conditioned on C, A and B are not independent anymore.

Finally, consider flipping the coin twice. The sample space $S = \{HH, HT, TH, TT\}$. Each outcome has probability 1/4 because the two flips are independent and, therefore, $P(HH) = P(H)P(H) = 1/4$ for instance. Let's find two events that are dependent. $A = \{HH, HT, TH\}$, and $B = \{HH, TT\}$.

$P(A) = 3/4$

$P(B) = 1/2$

$P(A, B) = 1/4 \neq P(A)P(B)$

$P(B|A) = 1/3 \neq P(B)$

$P(A|B) = 1/2 \neq P(A)$

This example is related to the following (what seems to be a) paradox: The king comes from a family of two children. What is the probability that the king's sibling is a male?