Lecture 25: Basic probability theory

Nitin Saxena *

IIT Kanpur

1 Probability

All of us encounter situations in our life where we need to take a decision based on the chance/likelihood/probability of some event. We will try to mathematical model these situations and see how this *chance or probability* can be quantified.

1.1 Definitions

Suppose we are interested in calculating/computing the chance of an outcome in a certain experiment. The experiment could be tossing a coin, throw of a die or picking a random number.

The set of all possible outcomes is known as *sample space*, it is a set and is denoted by Ω. We will be studying probability mostly in the context of use in computer science. Hence, our sample sets will be discrete.

Exercise 1. What is the sample space for a coin toss, sequence of coin toss, throw of a die and picking a random number.

For all these experiments, we are interested in a certain outcome or a subset of outcomes from the sample space. A subset of the sample space is known as an *event*. Our task is to model the probability of different events.

Not all subsets of the sample space need to be interesting and it might be a tedious task to define probability on every subset. But for the probability to make sense, if A, B are events, then A^C (*complement* of A) and $A \cup B$ should also be events.

This intuition gives rise to the concept of a *σ-field*. A collection of subsets $\mathcal{F}$ of the sample space Ω is called a σ-field, if,

1. Ω is in $\mathcal{F}$.
2. Complement of a set in $\mathcal{F}$ is in $\mathcal{F}$.
3. (Countable) unions of sets in $\mathcal{F}$ is in $\mathcal{F}$.

Exercise 2. Show that $\mathcal{F}$ is closed under (countable) intersection.

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

With all these definitions, we are ready to define probability function. A function $P : \mathcal{F} \to [0, 1]$ is called a *probability distribution function* (or just probability distribution) if it satisfies,

1. $P(\Omega) = 1$. `(I.e. one of the choices in` Ω `is guaranteed to be 'realized'.)`
2. If A, B are disjoint then $P(A \sqcup B) = P(A) + P(B)$. `(I.e. as we add more choices the 'chance' of one of them being realized grows additively.)`

Exercise 3. Show that the second rule above implies the corresponding property for countable union.

Why does it need countable union?

We can inductively apply the above axiom countably many times, on disjoint sets, and keep track of convergence issues.

* Edited from Rajat Mittal's notes.

Exercise 4. Prove that, in general, $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

Also, if $A \subseteq B$ are events then $P(A) \leq P(B)$.

We've $A \cup B = A \sqcup (B \setminus A)$. The axiom implies $P(A \cup B) = P(A) + P(B \setminus A)$. Also, $(B \setminus A) \sqcup (A \cap B) = B$, implies $P(B \setminus A) = P(B) - P(A \cap B)$.

If $A \subseteq B$ then $B = A \sqcup (B \setminus A)$.

Using the above concepts, we have modeled probability/chance/odds in an experiment. To summarize, say, we perform an experiment and are interested in the probability of various events in the experiment. The set of outcomes of the experiment will be our sample space Ω. Any subset of Ω is an event.

We can define probability of events (that are of interest) by specifying a σ-field. It should be a subset of 2^{Ω} and satisfy the σ-field conditions.

Exercise 5. Why did we define a σ-field? Why not take all the subsets of Ω as the interesting field?

References

1. K. H. Rosen. Discrete Mathematics and Its Applications. *McGraw-Hill*, 1999.
2. N. L. Biggs. Discrete Mathematics. *Oxford University Press*, 2003.
3. D. Stirzaker. Elementary Probability. *Cambridge University Press*, 2003.