Lecture 28: Basic probability theory

Nitin Saxena *

IIT Kanpur

1 Example– Monty Hall problem

This famous problem is posed in the context of a game show (*Let's make a Deal*, 1960s). There are 3 doors and behind one of them there is a car, while the other two have goats hidden behind them. You are first asked to pick a door, then the game show host (Monty Hall) opens one of the other two doors and reveals a goat. Assuming that you are not interested in a goat, should you switch the door you chose before?

This problem is very famous because of the counter-intuitive nature of the result. We can calculate the exact probability of whether switching helps or not. We will assume the standard assumptions that the car could be behind any door with equal probability. Also if you pick the door with car, Monty will choose the door to be opened uniformly at random (out of the other two).

Suppose the doors are numbered 1,2 and 3. Without loss of generality, we can assume that you pick door 1 and Monty opens door 2. We are interested in the conditional probability that the car is behind door 1, given that Monty opened door 2.

Let D_i be the event that car is behind door i, $P(D_i) = 1/3$. Let B be the event that Monty opens door 2. Then we are interested in $P(D_1 \mid B)$.

$$P(D_1 \mid B) = \frac{P(B \mid D_1) \cdot P(D_1)}{P(B \mid D_1)P(D_1) + P(B \mid D_2)P(D_2) + P(B \mid D_3)P(D_3)} .$$

Note 1. Ideally, all probabilities should be conditioned under the event that you have picked door 1. Since it is common to all events, we have chosen to skip it for brevity.

We know that $P(B \mid D_2) = 0$ and $P(D_1) = P(D_2) = P(D_3) = 1/3$. So,

$$P(D_1 \mid B) = \frac{P(B \mid D_1)}{P(B \mid D_1) + P(B \mid D_3)} .$$

The probability $P(B \mid D_1) = 1/2$ because Monty could have chosen door 2 or door 3. Though $P(B \mid D_3)$ is 1 because Monty's only choice (under the hypotheses) was to open door 2. This tell us that $P(D_1 \mid B) = 1/3$ and hence $P(D_3 \mid B) = 2/3$.

Hence, it is always beneficial for you to switch the doors.

2 Random variable

Many a times in probability theory, we are interested in a *numerical* value associated to the outcomes of an experiment. For example, number of heads in a sequence of tosses, pay out of a lottery, number of casualties after a Hurricane etc. These functions which assign a numerical value to the outcomes of the experiment are called *random variables*.

In this lecture, we will look at random variables, their expectation and properties of expected value.

Given the sample space Ω of an experiment, a random variable is a function $X : \Omega \to \mathbb{R}$. So a random variable assigns a real value $X(\omega)$ to every element ω of the sample space Ω. If the range of X is countable

* Edited from Rajat Mittal's notes.

then X is called a *discrete random variable*. In this course, we will only be interested in discrete random variables.

Given a probability function P on Ω, the natural definition of *probability of the random variable* is,

$$P(X = x) \;:=\; P(\{\omega : X(\omega) = x\}) \;=\; \sum_{\omega : X(\omega) = x} P(\omega)\,.$$

This is called the *probability mass function* of a random variable.

Note 2. If X is a random variable then $g(X)$ is also a random variable, where g is any function from $\mathbb{R}$ to $\mathbb{R}$.

Let us look at some examples of random variables and their probability mass function.

- Suppose the experiment consists of tossing a fair coin 10 times. The sample space is all sequences of length 10 of H, T. Define the random variable to be the number of heads in the sequence. That is, $X(\omega)$ is the number of H in (the length 10 string) ω.

  Exercise 1. Show that the probability of getting a sequence with k heads for a length 10 sequence is,

  $$\binom{10}{k}\left(\frac{1}{2}\right)^k\left(\frac{1}{2}\right)^{10-k}.$$

  Then, the probability mass function of the random variable for k is,

  $$P(X = k) = \binom{10}{k}\left(\frac{1}{2^{10}}\right).$$

  Exercise 2. Generalize the probability mass function if the probability of head is p.

  $\binom{10}{k}p^k(1-p)^{10-k}$.

- For an experiment, we ask the birthday of students in a class one by one. We stop as soon as we find two people with matching birthday. What is the probability mass function for the random variable X which counts the number of students queried?

  Let us find out the probability that we queried k people. Then the first $k-1$ birthdays are distinct and the last one matches at least one of the first $k-1$. First birthday will not match with anyone before, with probability 1. The second one will not match anyone before, with probability 364/365. The third one will not match anyone before, with probability 363/365 and so on. The last (k-th) one will match someone before, with probability $(k-1)/365$.
  Hence,

  $$P(X = k) = \frac{(k-1)}{365^k}\frac{365!}{(366-k)!}.$$

  Exercise 3. For what k does the probability $P(X \leq k)$ exceed 50%? 99.9%?

  23 people. 70 people! Read about Birthday Paradox.

- In a set of 1000 balls, 150 balls have some defect. Say, we choose 50 balls and inspect. Let X be the random variable which denotes the number of defective balls found.

  The probability mass function of X is non-zero in the range $k = 1$ to 50. It is given by,

  $$P(X = k) = \frac{\binom{150}{k}\binom{850}{50-k}}{\binom{1000}{50}}.$$

Like in the case of probability function, two random variables are called *independent* if the product of their probability mass function gives the joint probability mass function, i.e.,

$$P(X = x, Y = y) \;=\; P(X = x)P(Y = y)\,.$$

Exercise 4. Let X be the random variable that assigns 1 if the number on the throw of a dice is even, else it is -1.

Let Y be the random variable that assigns 1 if the number on the throw of a dice is prime, else it is -1.

Show that X and Y are dependent.

$$P(X = 1, Y = 1) = 1/6 \text{ while } P(X = 1) = 1/2 \text{ and } P(Y = 1) = 2/3.$$

References

1. D. Stirzaker. Elementary Probability. *Cambridge University Press*, 2003.