Lecture 29: Basic probability theory

Nitin Saxena [⋆]

IIT Kanpur

1 Expectation

We introduced random variables because of our interest in numerical values associated with the set of outcomes. Many a times we are interested in the *average*, mean or expectation of this numerical value. Taking an example, in a sequence of 10 tosses, you get Rs.1 every time it turns out to be head. What is your expected earning?

The expectation is easy to define if all outcomes are equally likely. In that case, if $\Omega = \{\omega_1, \cdots, \omega_n\}$, we would expect to get the average $\left(\frac{X(\omega_1)+X(\omega_2)+\cdots+X(\omega_n)}{n}\right)$.

Taking this idea further, the expected value of a random variable X is defined as,

$$E[X] := \sum_{x \in \mathbb{R}} P(X = x) \cdot x \,.$$

Is the above sum convergent? Since random variables are discrete in our case, most of the time range of X will be much smaller. By range of X, we mean values attained by X with a non-zero probability. Suppose the range of X is R, then

$$E[X] := \sum_{x \in R} P(X = x) \cdot x \,.$$

It is a common misinterpretation, probably because of the name, when it is stated that X will attain value $E[X]$ with "high" probability. That is clearly not the case. It is easy to construct examples where $E[X]$ might not even be in R.

Exercise 1. Construct a random variable such that $E[X]$ is not in R.

For instance, X may take integer values but the expectation becomes fractional.

The correct interpretation is: if we independently repeat the experiment multiple times, then with high probability the average value will be *close to* the expectation. This intuition will be formalized in the later sections.

Exercise 2. In an experiment, you get Rs.100 every time an odd number shows up on a dice. You loose Rs.100 every time an even number shows up. What is your expected earning?

Consider some more examples of expectation.

– Your friend is ready to give you Rs.100 if on a throw of a dice, a prime > 2 turns up. How much should you give him otherwise?

  Exercise 3. What is the random variable in this case?

  You cannot be certain to win this game if you bet any positive amount. The bet will be profitable to you if the "expected profit" for you is greater than zero (at least when the experiment is repeated multiple times).
  If you bet Rs.x, then the expected earning should be greater than zero, i.e., $1/3 \times 100 + 2/3 \times (-x) \geq 0$. So you can agree to pay any amount less than 50. This example suggests that in a *fair* bet, the expected profit should be zero.

[⋆] Edited from Rajat Mittal's notes.

Exercise 4. Suppose the expected value of a random variable X is zero. What is the expected value of $-X$?

- You toss a coin till you get head. What is the expected number of tosses?
  The random variable is the number of tosses.

  Exercise 5. Show that $P(X = k) = (1/2)^k$.

  The expected number of tosses is,
  $$E[X] = \sum_{k \geq 1} k \, (1/2)^k \, .$$
  Let $S = \sum_k k \, (1/2)^k$, then $(1/2) \, S = \sum_k k \, (1/2)^{k+1}$. Subtracting and then further manipulation tells us that $S = 2$.

- Your friend asks you to bet on the rise/fall of stock market. Both of you put Rs.100 in a pot and guess. If the guesses are the same then both get Rs.100, otherwise the one with the correct guess gets all the money.
  Assuming that stock market rises/falls with equal probability, this bet is fair. Now your friend wants to include her brother's share also. So she will put 200 Rs. and also the guess for her and her brother. Should you take the bet?
  Suppose your friend puts opposite guesses for her and her brother all the time. If your guess is correct then you will get Rs.$300/2 = 150$, while if you guess incorrectly then you get Rs.0. The expected value of your pot earning is Rs.75, which is less than the amount you put in the pot. So it is not advisable to bet.

One point of caution is that expectation need not be defined all the time. Let X be a random variable such that $P(X = k) = \frac{6}{\pi^2 k^2}$. It can be shown that $\sum_{k \geq 1} P(X = k) = 1$, but $\sum_k P(X = k) \cdot k$ is not convergent.

If X is a random variable then so is $g(X)$, where g is any function from $\mathbb{R}$ to $\mathbb{R}$. Then,
$$E[g(X)] = \sum_{x \in \mathbb{R}} g(x) \cdot P(X = x) \, .$$

Note 1. We need to assume that $\sum_x |g(x)| \cdot P(X = x)$ converges (i.e., *absolute convergence*).

1.1 Linearity of expectation

One of the most important property of expectation is that it is linear. This means that given two random variables X and Y,
$$E[X + Y] = E[X] + E[Y] \, .$$

Note 2. Say $X : \Omega \to \mathbb{R}$ and $Y : \Omega \to \mathbb{R}$, then the random variable $X + Y : \Omega \to \mathbb{R}$ is defined as $(X + Y)(\omega) := X(\omega) + Y(\omega), \ \forall \omega \in \Omega$.

This property is known as *linearity of expectation.*

Note 3. Nothing is assumed about the relationship between X and Y. They might be dependent or independent random variables.

Proof. The expectation $E[X+Y]$ is given using the joint probability mass function $P(X=x, Y=y)$.

$$\begin{aligned}
E[X+Y] &= \sum_{x,y}(x+y)P(X=x,Y=y) \\
&= \sum_{x,y} xP(X=x,Y=y) + yP(X=x,Y=y) \\
&= \sum_{x} x \sum_{y} P(X=x,Y=y) + \sum_{y} y \sum_{x} P(X=x,Y=y) \\
&= \sum_{x} xP(X=x) + \sum_{y} yP(Y=y) \\
&= E[X] + E[Y]\,.
\end{aligned}$$

□

Exercise 6. Extend the linearity of expectation to more than two random variables using induction.

References

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