CS 107, Probability, Spring 2020
Lecture 19

Michael Poghosyan
mpoghosyan@aua.am

AUA

09 March 2020

Content

- Random Variables
- The Cumulative Distribution Function

$\sim$ Happy March 8[1] $\sim$

[1] To whom it may concern.

Random Variables, prelude

Up to this point when considering Experiments, we have considered

- its Outcomes;
- its Sample Space Ω;
- some Events $A, B \subset \Omega$;
- and their Probabilities $\mathbb{P}(A), \mathbb{P}(B)$.

In many cases, we are interested not exactly in the outcomes, but in some numerical value depending on the outcome.

Random Variables

So we consider functions defined on the Sample Space of an Experiment:

Random Variable

Any function

$$X : \Omega \to \mathbb{R}$$

is called a Random Variable (r.v.) on that Experiment.

Remark: In fact, the definition is not Mathematically rigorous and complete - we need to consider only **measurable** functions (18+ topic)

Remark: From this point on, we will denote by capital letters X, Y, Z Random Variables, and x, y, z will denote their particular values.

Examples:

Example: Our Experiment is rolling a fair die. So the Sample Space is:

$$\Omega = \{1, 2, 3, 4, 5, 6\}.$$

Some examples of r.v. on this Experiment:

$$X(\omega) = \omega, \qquad \forall \omega \in \Omega;$$

$$Y(\omega) = 100 * \omega, \qquad \forall \omega \in \Omega;$$

$$Z(\omega) = \omega^2 - 10.1, \qquad \forall \omega \in \Omega;$$

$$T(\omega) = 100, \qquad \forall \omega \in \Omega;$$

$$U(\omega) =$$

Random Variables, Examples:

Example: We are tossing 3 coins. The Sample Space is[2]

$$\Omega = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}.$$

Now, we can be interested in different r.v.s:

$$X(\omega) = \#\text{Heads in } \omega;$$

$$Y(\omega) = \text{The difference between the numbers of Heads and Tails}$$

$$Z(\omega) =$$

[2] Well, of course, you can model the Experiment with another Sample Space.

Random Variables, Examples:

Example: We are repeating, Independently, a Simple Experiment n times. We have an Event in our Simple Experiment, in our Trial, which we call a *Success*. Then we can consider different R.V.s:

- $X =$ the number of Successes in n trials;
- $Y =$ the number of Failures in n trials;
- $Z =$ the number of Successes minus the number of Failures;
- $T =$ the first time we will have a Success: if we will not have a Success at all, then $T = 0$;
- $U = 100 \cdot X$ etc.

Random Variables, Examples:

We continue by considering some examples of r.v. We consider the following Experiment: we choose a person at random in AUA.

- $W =$ The weight of that person is a r.v.
- $C =$ The eye color of that person is NOT a r.v., since it is not a Number (from $\mathbb{R}$);
- $ED =$ Distance between two ears[3] of that person is a r.v.
- Can you give some other r.v.'s on this experiment?
- Some options: Age, Height, Left Shoe size, Salary, Years of education, Work experience duration, No. of children, No. of missed Probability Classes, No. of Admirers, ...

Remark: One of the great advantages of r.v. is that we do not need to specify the Sample Space! We just forget about the Sample Space, and work with the values of the r.v.

[3]Well, we need to define how to calculate that distance.

Random Variables, Examples:

Some other, real-life Examples of r.v.

- $X =$ the Closing Price of General Electric Stock on the next Wednesday,is a r.v. **Question:** What are the Experiment and Sample Space behind?
- $Y =$ the number of car accidents today in Yerevan is a r.v. **Question:** Again, what are the Experiment and Sample Space behind? Another one is $Y_1 =$ the claim size for a particular Insurance company for today.
- $Z =$ The Mean Grade for our Prob Midterm 2 is a r.v.
- $T =$ The lifetime of some particular car Engine is a r.v.
- $U =$ The no. of customers in a particular shop in a day.
- $V =$ The no. of clicks on a particular webpage in an hour.
- $W =$ The number of air molecules exhaled when reading this sentence is a r.v.

Random Variables

Now, assume $X(\omega)$ is a r.v. defined on Ω. Starting from some point on, we will not specify ω, we will just write X. And usually people are doing this in practice.
Now we want to consider some Events concerning r.v.s, and calculate some Probabilities. Now, assume we have a Probability Measure $\mathbb{P}$ on Ω, defined for all Events.

Can we calculate $\mathbb{P}(X)$, that is, $\mathbb{P}(X(\omega))$?

NO! $X(\omega)$ is not an Event. Say ω is a randomly chosen person in our classroom, and $X(\omega)$ is that person's age (in years). Is it correct to ask to

Calculate the Probability of the Age of the chosen person ?

Of course, no! Instead, we can ask to calculate the Probability that the persons age is 20.1, i.e., $\mathbb{P}(X(\omega) = 20.1)$ or the Probability that the age is less than 25, i.e., $\mathbb{P}(X < 25)$.

Random Variables

So, if X is a r.v., then X **is NOT AN EVENT**, but for any set $A \subset \mathbb{R}$,

$$X(\omega) \in A \qquad \text{or, rather,} \qquad \{\omega : X(\omega) \in A\}$$

is an Event, so we can calculate the Probability $\mathbb{P}(X \in A)$, which is the shorthand of $\mathbb{P}(\{\omega \in \Omega : X(\omega) \in A\})$. Say, if Y is a r.v., then we can calculate $\mathbb{P}(Y < 5)$, $\mathbb{P}(Y = 4)$, $\mathbb{P}(0 \leq Y < 100)$, $\mathbb{P}(|Y| > 2)$, $\mathbb{P}(\tan(Y) < 0.3)$ etc.

Random Variables, General Problem

Assume we have an Experiment with the Sample Space Ω. What we want to know in this Experiment? - We want to know what will be the outcome, the result of that Experiment. Of course, it is not possible to guess that outcome exactly.
Now, assume also that X is a r.v. on that Sample Space. Then we can be interested in what will be the result of X? As above, we cannot predict that with 100%.

Random Variables, General Problem

But what can we get for a r.v. X, if we cannot predict its value? Because we cannot know the exact value X will take after the Experiment will be done, we want to know the Probabilities of different values. In fact, for many r.v. the Probability of assuming a particular value is 0, and we have considered already this kind of examples (not in the r.v. terminology, of course).Say, if X is the height (in m) of a randomly chosen person, a r.v., then the Probability $\mathbb{P}(X = 1.7122123) = 0$. So, rather to consider the Probability of taking a particular value, we consider the Probabilities $\mathbb{P}(X \in A)$, where A is an arbitrary subset of $\mathbb{R}$. In other words, we are studying the **Distribution** of X.

Complete Info about the Random Variable

So we will say that we have the Complete Information about our r.v. X, if we are able to calculate all probabilities

$$\mathbb{P}(X \in A), \qquad \forall A \subset \mathbb{R},$$

and this is, in some way, the General Problem of Probability Theory - to be able to calculate these Probabilities.
It is a remarkable fact that it is enough to be able to calculate all probabilities of type $\mathbb{P}(X \in (-\infty, x])$, for any $x \in \mathbb{R}$. In that case, we will be able to calculate the above Probability for any A.

Cumulative Distribution Function, CDF

So we define:

Cumulative Distribution Function, CDF

Assume X is a r.v. The function

$$F(x) = F_X(x) = \mathbb{P}(X \leq x), \qquad x \in \mathbb{R},$$

is called the Cumulative Distribution Function (CDF) of X.

Remark: X is a r.v. here, and x is a Real Number.
Remark: Having $F(x)$ for any x, we can calculate the Probabilies of the form $\mathbb{P}(X \in A)$.For example,

$$\mathbb{P}(a < X \leq b) = \mathbb{P}(X \leq b) - \mathbb{P}(X \leq a) = F(b) - F(a).$$

Later we will meet a lot this kind of calculations.