PROBABILITY THEORY FOR DATA SCIENCE

Prof. Ishapathik Das

Department of Mathematics and Statistics

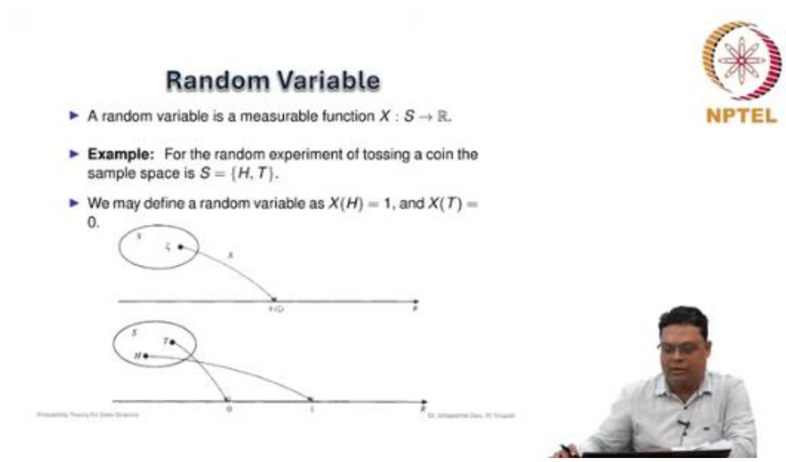
Indian Institute of Technology Tirupati

Week - 02

Lecture - 08

Random Variable

Now we will learn about the concept of a random variable. A random variable is a measurable function from the sample space, S, to the set of real numbers, $\mathbb{R}$. To understand what a measurable function is, we need to know about the sigma field. So, let us recall the sigma field. A sigma field is defined on a non-empty set.



So, suppose S is a non-empty set. Let's consider this set S, and we take a collection of subsets of S, denoted by C. Now, let's look at an example. Remember the random experiment—let's say we're considering a random experiment like tossing a coin. So, if you toss a coin, what will be the outcome? What are the possible outcomes?

We don't know in advance whether it will be heads or tails, but we know the possible outcomes: heads and tails of this random experiment. So, S is the sample space containing all possible outcomes, heads and tails. To define a σ-field, no random experiment is actually required; we're just giving an example here. Any non-empty set can be used for S; for example, S could be R. Here, we are considering S as the set {head, tail}, and C is some collection of subsets of S.

So, we can consider this set of subsets. Suppose C contains only two subsets: the null set ∅, which is also a subset of S, and S itself, which is also a subset of S. Now, let's take another example—suppose C_2 contains ∅, H, and S. Here, we've included H as another subset of S. Now, a σ-field is defined as a collection, C, of subsets of S that satisfies certain properties. So, C is a σ-field if the following properties are satisfied:

1. S must belong to C: Since S is a subset of itself, if C is to be a σ-field, then S has to be in C.

2. If any set A belongs to C (meaning $A \subseteq S$), then A^c (the complement of A) also has to be in C.

3. If $A_1, A_2, ..., A_n$ is a countably infinite collection of subsets of S, and each $A_i \in C$ (for i = 1, 2, ...), then the union of A_i from i = 1 to infinity must also be in C.

These are the three conditions. If these properties are satisfied by this collection of subsets of S, then C is called a σ-field.

So, this is a σ-field. For example, if you consider this collection of subsets, since it is a collection of subsets of S, you can see that S belongs to C, and only S and ∅ are in C. That's why we only need to check these elements to verify the properties. Now, if S belongs to C, then S^c—what is S^c? S^c is just the null set ∅—so that also belongs to C.

We see that ∅ is in C, and $\emptyset^c$, which is S, is also in C. So, this property is satisfied by C. Now, for the condition involving $A_1, A_2, ...$ as an infinite collection. Here, we only have two sets, so if you take their union, $S \cup \emptyset$, that also belongs to C. That's why even though it's for an infinite collection, any finite collection would also belong to C because we can include ∅ and meet this condition.

If you consider an infinite collection like $A_1 = S$, then ∅, ∅, ∅, and so on, and take the union, it would be $S \cup \emptyset$, union ∅, etc., which is just S, and S belongs to C. So, this collection satisfies all the properties and is, therefore, a σ-field.

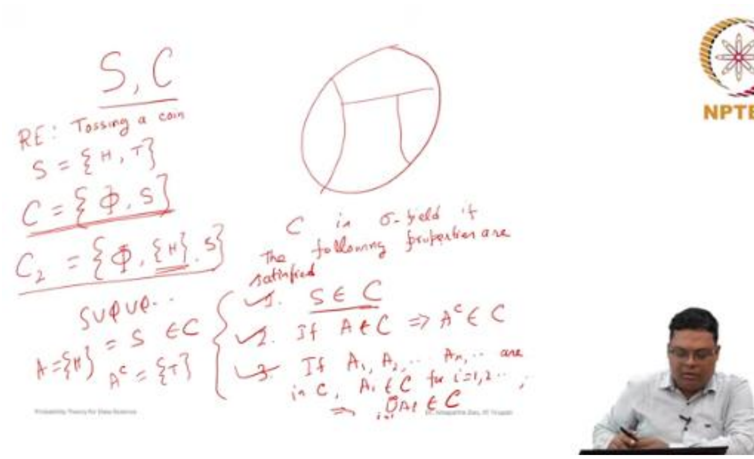
Now, if you consider this collection C_2, you can see that S belongs to C_2, so that condition is satisfied. Now we have to check the second condition, which says that if an element A belongs to C, then A^c must also belong to C. For S, the complement of S is ∅, and ∅ belongs to C_2, so no problem there.

For ∅, ∅ belongs to C_2, and $\emptyset^c$, which is S, also belongs to C_2. So, this is satisfied in C_2. Now, if we consider the set A as H, which is in this collection C_2, then the complement of H—since S contains both H and T—is the set T. But we did not include T in this set. So, that's why C_2 does not satisfy this property, meaning it is not a σ-field because all three properties need to be satisfied for a σ-field.

Now we have to check the second condition. The second condition says that if an element A belongs to C, then it must imply that A^c also belongs to C. Now, for S, the complement of S is ∅, which belongs to C, so no problem there. For ∅, ∅ belongs to C, and $∅^c$, which is just S, also belongs to C. So, this is actually condition C_2 we're considering.

We have to think about C_2. Since both S and ∅ satisfy this, they belong to C_2. Now, if you consider the set A equal to H, then H belongs to this collection C_2. But if you look at the complement of H, the complement of H is actually the set T, since S contains both H and T. So, if you consider the complement of H, it's the set T, which we didn't include in this set. That's why this property of C_2 is not satisfied, so this is not a σ-field, because for a σ-field, these three properties have to be satisfied.

So, this is not a σ-field. Now, we will discuss what a measurable function is. Let's consider another σ-field. Suppose you are doing a random experiment by rolling a die. What will be the sample space?



We will denote the sample space as S_2, which consists of {1, 2, 3, 4, 5, 6}. I just want to mention that C_2 is not a σ-field, as we found. The problem is that whenever we consider A = S, its complement does not belong to T. Now, we want to make C_2 a σ-field, so what corrections do we have to make? Keeping all the elements in this set, we want to add more subsets of S so that it will be a σ-field.

Let's consider adding ∅, S, and also T. We want to see if it is a σ-field. In this case, we can see that the properties will be satisfied. We are adding more elements to ensure it becomes a σ-field. For the element H, H's complement is T, which is also in C. If we add T, then T's complement, which is H, is also in C. Now, if you take any finite union or consider distinct elements, the property will be satisfied. If you consider unions like H ∪ T, that results in S, and S ∪ ∅ is still S. So, all possible unions will be in C, which means C_2 will be a σ-field because the third property is also satisfied. This set should satisfy all conditions because if you consider all possible subsets of S, since it is a finite set, it will contain 2^n number of elements. For n = 2, this results in 4 subsets. So, C_2 contains 1, 2, 3, and 4 subsets. Therefore, it contains all the subsets, which is known as the power set.

Whenever you consider all subsets, if C contains all of them, let S be a non-empty set. Let Ω denote the set containing all possible subsets of S. So, the set Ω is known as the power set. The power set includes all subsets of S, and it is a σ-field because it satisfies all the properties. $S \in C$, $A \in C$, and $A^c \in C$ since we have included all subsets. So, C_2 is a power set.

So, let us consider another σ-field. Suppose you are doing a random experiment, like rolling a die. In that case, what will be the sample space? Let's denote the sample space as S_2, which is just {1, 2, 3, 4, 5, 6}. Now, another thing I just want to mention is that we found C_2 is not a σ-field. The problem is that whenever we consider A = S, its complement does not belong to T.

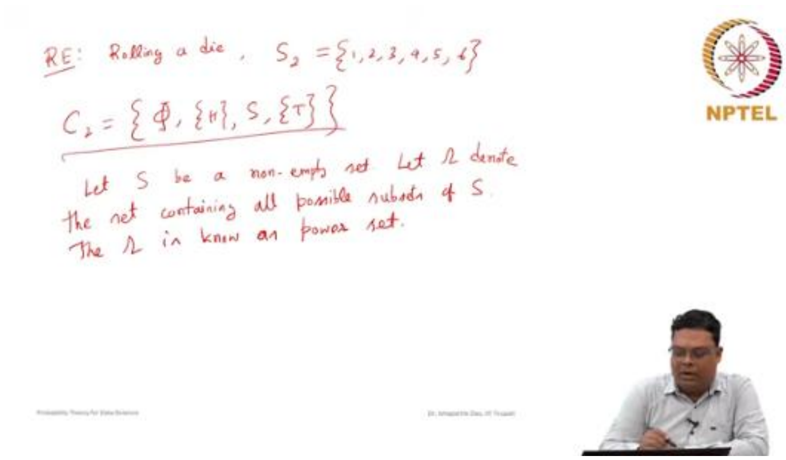
So now, suppose we want to make C_2 into a σ-field; what correction would we have to make? To do this, we'd keep all the elements in this set and just add more subsets of S so that it forms a σ-field. How can we do that? Let's take C_2 with ∅, S, H, and now add T as well. So, let's see if this forms a σ-field or not.

So then, you see, in that case, this property will be satisfied. Basically, here we are adding more elements—we want to add T as well—so that it becomes a σ-field. Then, you can see that for the element H, its complement, which is T, is also in C. Now, if I add T, then T's complement, which is H, is also in C. If you take any union, finite or infinite, or distinct elements, this condition will still be satisfied.

So for finite collections, this property will also be satisfied in C. Now, if you consider all kinds of unions, like $H \cup T$, which is just S, and $S \cup \emptyset$, which is still S, then all possible unions will be in C. That's why C_2 will also be a σ-field, as this third property is also satisfied. So, this set should satisfy the requirements because if you consider all possible subsets of S, then, since S is finite, it will contain 2^n elements. So, for 2^2, we get 4, meaning C_2 also contains 4 subsets.

This way, it includes all the subsets, which is why it's known as a power set. So, it's known as a power set whenever you consider all subsets. Let S be a non-empty set, and let Ω denote the set containing all possible subsets of S. So, Ω is known as the power set. A power set means it contains all subsets of S.

Obviously, the power set is a σ-field because it satisfies all these properties. So, if $S \in C$, and $A \in C$, then $A^c \in C$, since all the subsets are included. So, C_2 is a power set; C_2 is a power set. Now, let us next discuss this example: a random experiment of rolling a die. This is a subset, a non-empty set S_2.



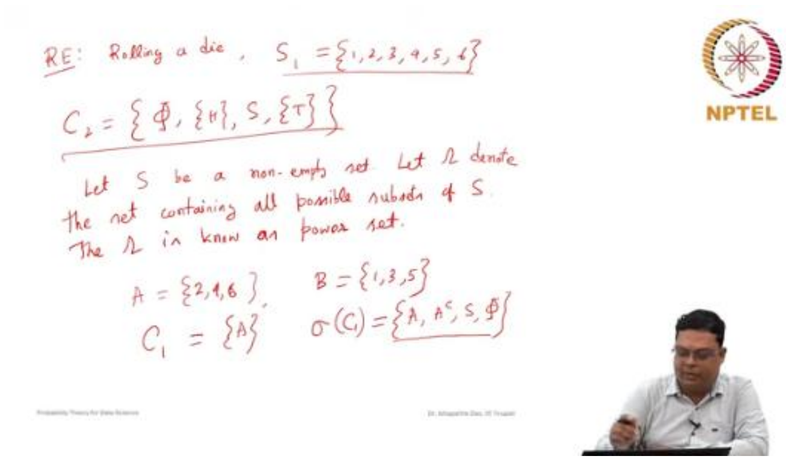
S_2 is the sample space for rolling a die, so it contains {1, 2, 3, 4, 5, 6}—6 elements. If you consider the power set, then it will contain 2^6 elements in the power set. Now, let us consider not the power set, but some non-trivial σ-field. Let us consider one set A = {1, 2, 3}, and another set B = {2, 4, 6}. B is the subset {1, 3, 5}.

So now we want to find a σ-field generated by, let's say, a collection of subsets. Suppose C_2 is here; we already denoted C_2. Let us consider this as S_1, and let's call it C_1. C_1 is just the set containing A as the only subset. But note that it is not a σ-field because S is not here, ∅ is not here, and A^c is not here.

So we want to make this a σ-field. We want to add more sets. So, the σ-field generated by C_1 will include A, A^c (which is B), S, and also the null set. This is the σ-field generated by C_1. You can add more sets as well, so that's a σ-field.

But maybe we consider it as a smaller σ-field generated by C because if you remove any of the subsets here, any of the elements here, then it will not be a σ-field. So, it is called the smallest σ-field generated by C_1. So, this is a σ-field. Basically, A^c is nothing but B. So, this set A, it is a σ-field, okay.

Now, we want to define—so there are two sets, sample spaces, here. So, S_1 is one set. Let us consider S_1 and C_1. C_1 is denoted as σ, so we consider this. Let us denote this as F_1.



F_1 is the σ-field generated by C_1. F_1 is a non-empty subset, and this is the σ-field. S is nothing but the head and tail we talked about, and C, which we have already discussed, is nothing but ∅ and S. This is another set, or we can consider C_2. This is also a σ-field, so no problem.

Let us consider C_2 as S and C. Now, this is also a σ-field. We want to define it. This notation is called $\mathcal{F}$; it is the standard notation. Here, this is also $\mathcal{F}$, not a small f.

Now, we want to define a function between S_1 and S, some function. f: $S_1 \rightarrow S$, and we want to define that function. Suppose S_1 contains {1, 2, 3, 4, 5, 6}, and S is the set {H, T}; there are two elements. Suppose we want to map all even numbers to heads, or let us consider a constant function. If we consider a constant function, let f(x) = H for any element $x \in S_1$. So, this is a function.

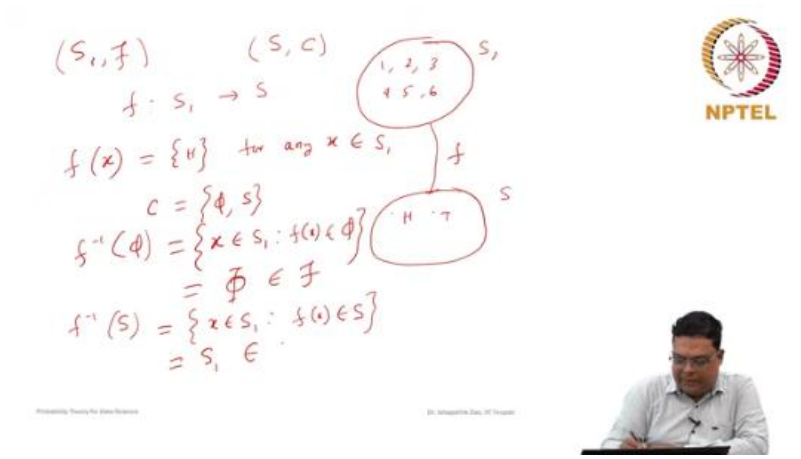
For any $x \in S_1$, this is the function I am defining, so basically all points go to H. This is a constant function. Now, we want to check what C contains. C contains just two elements, so you can remember that C is ∅ and S. Now, I just want to find what is $f^{-1}(\emptyset)$.

The $f^{-1}(\emptyset)$ is defined by all $x \in S_1$ such that $f(x) \in \emptyset$. Now, you can see that it does not contain any elements satisfying this because f(x) = H for any $x \in S_1$. So, ∅ means it is the null set. None of the elements will satisfy this. But f(x) = H for any x, so this will again be a null set.

That is why ∅ belongs to $\mathcal{F}$; you can see that. Now, if you consider another set in C, let's take $f^{-1}(C)$. This is defined as all $x \in S_1$ such that $f(x) \in C$. So, what is C? We know that C is a σ-field.

We want to take this set S, and we want to take another element of C. That is why I just realized notationally it is wrong; it is actually S. All elements belong to S. So, note that if you consider x such that $f(x) \in C$, now what is f(x)? f(x) = H, and S contains {H, T}.

So, for any x you consider in S_1, then f(x) has to belong to S. So, all x satisfy this condition. So, all x means here it is nothing but all of S_1, and this also belongs to the σ-field $\mathcal{F}$. You can see that for this function f, if you consider any element in the σ-field and take the inverse of that, it actually belongs to this σ-field. This kind of function is known as a measurable function.



So, what is a measurable function? Let's denote S_1, $\mathcal{F}_1$, and S_2, $\mathcal{F}_2$ as two σ-fields, and a function f: $S_1 \to S_2$ is said to be a measurable function if for any element $B \in \mathcal{F}_2$, the notation should be $\mathcal{F}$. Otherwise, you will conclude with the usual functional notation. For any element $B \in \mathcal{F}_2$, $f^{-1}(B)$ must belong to $\mathcal{F}_1$. This means that if you take the inverse of this, $f^{-1}(B)$ is all $x \in S_1$ such that $f(x) \in B$.

This is called $f^{-1}(B)$. If this belongs to $\mathcal{F}_1$ for any $B \in \mathcal{F}_2$, because it is S_1 to S_2, then it is known as a measurable function.

Now, we will discuss a special kind of σ-field. So, because these concepts are a bit abstract, like head and tail when throwing a die, the outcomes are {1, 2, 3, 4, 5, 6}, which are real numbers. However, most phenomena in real life are described using some abstract types that may not be real numbers.

If we can observe something as a real number or relate it to a real number, it will be very helpful for our understanding. In the real number system, we have learned many concepts, such as algebra from our school and college education, including algebraic properties, multiplication, addition, and subtraction. We also learn about limits, continuity, sequences, and series, all of which are developed within the real number system.

If we take any abstract set, we need to develop those concepts again. While this can be done, we must verify it or develop all these concepts and results with respect to what we want to use.

Since the sample space is often abstract, like head and tail for events such as earthquakes, we need to find a way to relate these abstract concepts to real numbers in order to understand them and perform computations. Therefore, we want to learn about some special σ-fields, which are special collections of subsets of the real numbers. So, if you consider this set of real numbers S, we are assuming it is a non-empty set.