PROBABILITY THEORY FOR DATA SCIENCE

Prof. Ishapathik Das

Department of Mathematics and Statistics

Indian Institute of Technology Tirupati

Week - 07

Lecture - 32

Properties of the Joint Cumulative Distribution Function of a Bivariate Random Variable

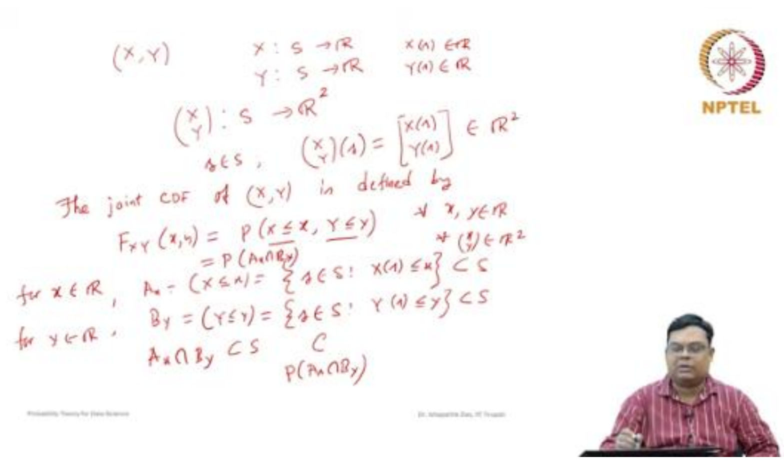
We will start the discussion on the properties of the cumulative distribution function. We have already defined the cumulative distribution function and the joint cumulative distribution function of a bivariate random variable. This is for two random variables, X and Y. Both are random variables from a sample space. X is a function, and it is a measurable function from S to R. Similarly, Y is also a measurable function from S to R.

We define (X, Y) as a measurable function from S to R, defined for any $s \in S$. (X, Y) of s is nothing but F(x, y), so (X, Y) of s = X(s) and Y(s). Both are real numbers because $X(s) \in R$, and $Y(s) \in R$. Therefore, this is in R^2. So, (X, Y) is a measurable function from S to R^2. This is a bivariate random variable, as we have already discussed.

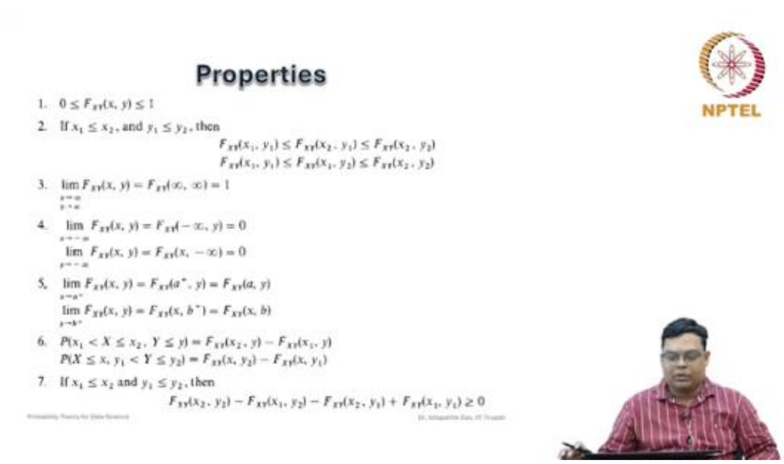
We then define the joint probability distribution function, also known as the joint CDF (cumulative distribution function). The joint CDF of the random variables (X, Y) is defined by $F_{X,Y}(x, y)$, which is the probability that $X \leq x$ and $Y \leq y$, for all $(x, y) \in R^2$.

Now, we have already defined those concepts. We denote A_x as the event $X \leq x$, which is the set of all $s \in S$ such that $X(s) \leq x$. Similarly, for a given x, where $x \in R$, A_x is defined. For $y \in R$, B_y is the event $Y \leq y$, which is the set of all $s \in S$ such that $Y(s) \leq y$. This is a subset of S. Similarly, B_y is a subset of S. Therefore, $A_x \cap B_y$ is also a subset of S.

Additionally, since X and Y are measurable functions, A_x is in the sigma field, and B_y is also in the sigma field, which is the class of events we are considering. Since $A_x \cap B_y$ is part of the sigma field, it should be inside C. Therefore, it will be a subset of S, and we can then discuss the probability of $A_x \cap B_y$, which is the probability in this case. So, this is nothing but the probability of $A_x \cap B_y$.



This probability has been discussed. Now, we will discuss some of the properties of the cumulative distribution function. This cumulative distribution function is defined, so there is a probability. So, it is very clear that since $F_{X,Y}(x, y)$ represents a probability, it must belong to the range [0, 1]. The next property is that if $x_1 \leq x_2$ and $y_1 \leq y_2$, then let us first discuss the case when you take two points on the real line: $x_1 \leq x_2$, and $y_1 \leq y_2$.



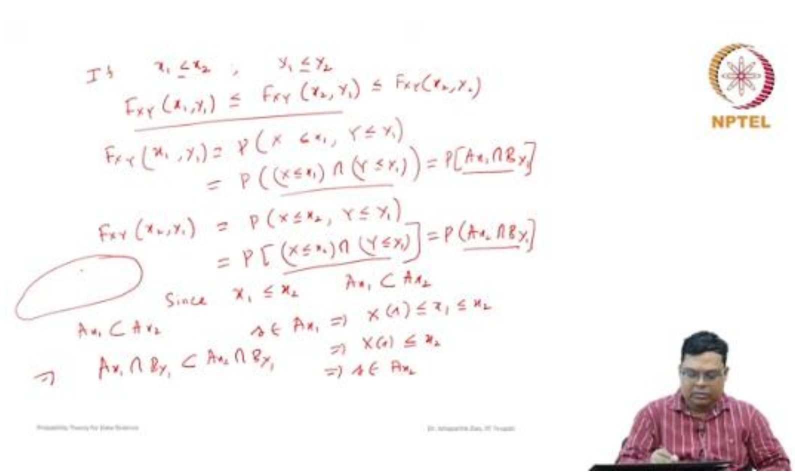
Then, if $y_1 \leq y_2$, $F_{X,Y}(x_1, y_1) \leq F_{X,Y}(x_2, y_1)$, which in turn is $\leq F_{X,Y}(x_2, y_2)$. In these inequalities, you can see that we are fixing y_1 as one coordinate while the other coordinate may increase. So, let us prove this. Similarly, we can prove that for x_1 and y_1, F_{X,Y}(x_1, y_1) is nothing but the probability that $X \leq x_1$ and $Y \leq y_1$. This is the probability by definition.

So now, this is nothing but the event $X \leq x_1$ and $Y \leq y_1$. Now, what is F_{X,Y}(x_2, y_1)? This is nothing but the probability that $X \leq x_2$ and $Y \leq y_1$. This is the probability of the event $A_x \leq x_2$ and $Y \leq y_1$. So, now what is the relation here?

In both cases, you can see that, for example, if we write this as the probability of $A_{x1} \cap B_{\gamma 1}$, it is nothing but the probability of $A_{x2} \cap B_{\gamma 1}$. Now, what is the relation with this? Now, since $x_1 < x_2$, we have already shown that $A_{x1} \subseteq A_{x2}$. If they are equal, then these two sets will be equal. However, if $x_1 < x_2$, then any element that belongs to A_{x1} means it satisfies the condition $X \leq x_1$. Sorry, A_{x1} means that if an element, say s, belongs to A_{x1}, then X(s) $\leq x_1$, by the definition of A_{x1}.

Now, since this is $\leq x_2$, it implies that $X(s) \leq x_2$, which means s belongs to A_{x2}. This is the definition of A_{x2}, where all s satisfy $X(s) \leq x_2$. Hence, $A_{x1} \subseteq A_{x2}$. Now, if you take the intersection of the same sets, since we are taking A_{x1} as a subset of A_{x2}, it implies that if you take the intersection of the same sets on both sides, $A_{x2} \cap B_{\gamma 1}$ will be equal to $B_{\gamma 1}$. This is because $A_{x1} \subseteq A_{x2}$, so when you take the intersection on both sides, you get the result.

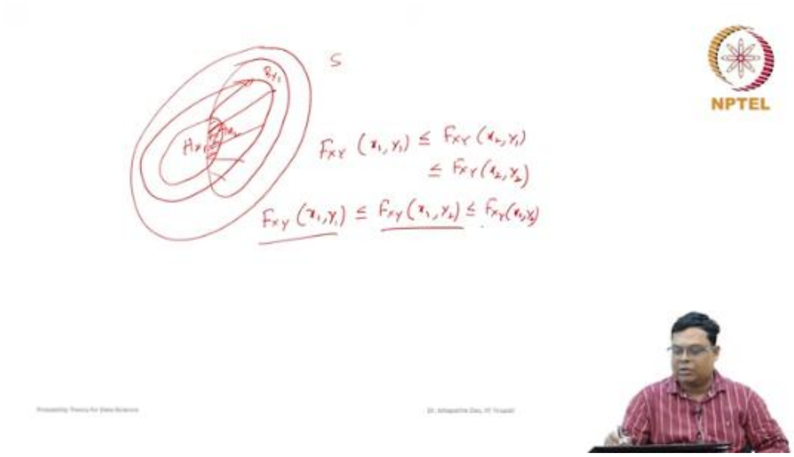
Basically, if you write A_{x1} and A_{x2}, A_{x2} will be the superset. This means A_{x1} is a subset, and A_{x2} is the larger set. Now, if you take $B_{\gamma 1}$, one minute... Suppose this is A_{x2}, and let us draw a larger graph. One minute.



This is the sample space S. Suppose this is A_{x2}, and inside it is A_{x1}. This is A_{x2}, and then $B_{\gamma1}$ is supposed to be here. Now, the subset of $A_{x1} \cap B_{\gamma1}$ is nothing but $B_{\gamma1}$. So, $B_{\gamma1} \cap A_{x1}$ is this part, and $B_{\gamma1} \cap A_{x2}$ will be the whole part.

This means that $A_{x1} \cap B_{\gamma1}$ will be a subset of $A_{x2} \cap B_{\gamma1}$. So, this implies that $P(A_{x1} \cap B_{\gamma1}) \leq P(A_{x2} \cap B_{\gamma1})$. Therefore, we can conclude that $F_{X,Y}(x_1, y_1) \leq F_{X,Y}(x_2, y_1)$. Similarly, you can show that if you keep x_2 fixed and then increase the value of the second coordinate, it will be less than or equal to. Similarly, another relationship will be the same.

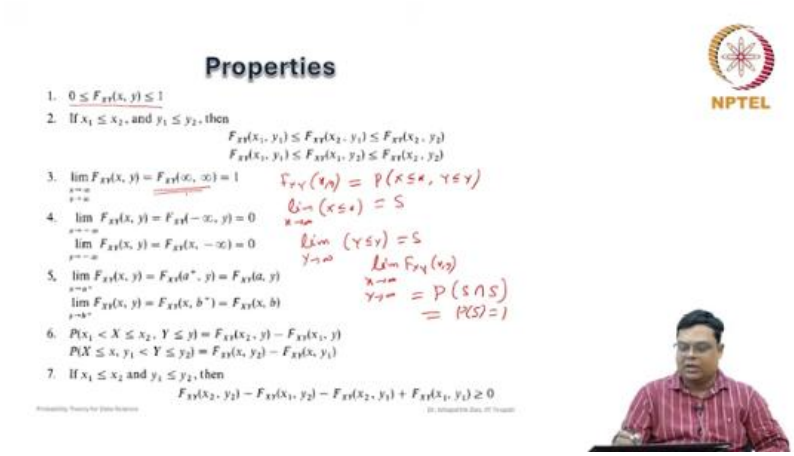
For (x_1, y_1), we first fix y_1 and then $F_{X,Y}(x_1, y_2) \leq F_{X,Y}(x_2, y_2)$. So, this is the relationship. So, both of these proofs can be shown very similarly. Now, these are some of the basic properties, and this is another property: if you take $x \to \infty$ and $y \to \infty$, $F_{X,Y}(x, y)$ is nothing but $F_{X,Y}(\infty, \infty)$. We're not directly putting this value; it's just the limiting value that we want to find. We won't go into the rigorous proof; we will just try to find some intuition.



So, what is $F_{x\gamma}(x, y)$? This is nothing but the probability that $X \leq x$, and $Y \leq y$. Now, if you take the limit as $x \to \infty$, what will happen to this set? As $x \to \infty$, this small x becomes equivalent to saying $x \leq \infty$. So, it will satisfy all the conditions.

Essentially, X is a random variable, which is a measurable function from S to $\mathbb{R}$. So, as $x \to \infty$, it will satisfy this relationship, and this is essentially the whole set S. Similarly, if you take $y \to \infty$, $Y \leq y$ will also represent the whole set S. Hence, as $x \to \infty$ and $y \to \infty$, $F_{XY}(x, y) = P(S \cap S)$, which is P(S). This is equal to 1. So, this is the result we arrive at.

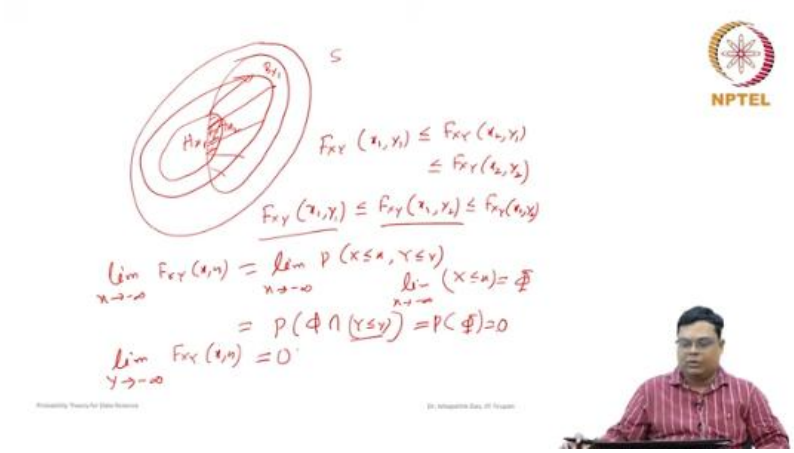
Now, in the case where you take the limit of x as it goes to $-\infty$, suppose you consider the limit of one of the variables going to $-\infty$. What will be the value of F(x, y)? This is nothing but the limit as $x \to -\infty$ of $P(X \leq x \text{ and } Y \leq y)$. So now, we are trying to understand what the value should be. The proof will be different, but here we are just trying to understand the concept.



When $x \to -\infty$, what will this event be? As $x \to -\infty$, what will happen? Since X is a function from S to $\mathbb{R}$, we assume that all values are included. So, here, $X \leq x$ represents all $s \in S$ such that $X(s) \leq x$. Now, if you let x tend to $-\infty$, none of the values in S will satisfy this relationship.

Therefore, this will be the empty set. The actual proof will be different, but this is just to help understand the concept. So, now if this is the empty set, then the probability will be nothing but the probability of the empty set $\cap$ $Y \leq y$. Whatever the set $Y \leq y$ is, when intersected with the empty set, it will again be the empty set. Therefore, this is equal to 0.

Similarly, you can show that if you take the limit as $y \to \infty$, one of the coordinate components of this vector will tend to ∞. If you take $y \to -\infty$, $F_{XY}(x, y)$, one of the components, will tend to 0. So, these are some of the properties to just understand. Now, this will be a right-continuous function. We are not going to prove it, just to understand it.



Basically, for univariate cases, we have also checked that this function will be right-continuous. So, right-continuous means that if you take the limit, for example, $F_{x\gamma}$, and suppose you take the limit as $h \rightarrow 0^+$, this is nothing but the positive limit of $F_{x\gamma}$ as a + h, y. Now, this value may or may not be equal to $F_{x\gamma}(a, y)$. If it is equal, then this limit is called right-continuous. So, the function is right-continuous.

Similarly, for the other coordinate, we can also show that $F_{x\gamma}(x, b + h)$ holds. So, for some value b, this is actually the limit. Similarly, we can write that as the limit, where $h \rightarrow 0^+$, $F_{x\gamma}(x, b + h)$, which is nothing but defined as $F_{x\gamma}(x, b^+)$, which is equal to $F_{x\gamma}(x, b)$. This is the definition of right continuity. Both coordinates are right-continuous, so both components of $F_{x\gamma}$ are right-continuous functions.

Now, what about the other properties? As mentioned, it is right-continuous, meaning $b^+ = F_{x\gamma}(b)$. These are usual properties that we have already seen in univariate cases. We have already discussed these properties in the context of univariate random variables.

Now, we will discuss some additional properties. This property is also equivalent; it is just an extension of the properties for univariate random variables. Now, the second property says that if you consider how we will find the probability that $Y \leq y$, but $X \in [x_1, x_2]$, it is simply $F_{x\gamma}(x_2, y)$ - $F_{x\gamma}(x_1, y)$. Let us find out how this can be proved. It states that, suppose we consider $x_1 < x_2$. If $x_1 = x_2$, it becomes a trivial case. So, if $x_1 = x_2$, there is nothing to prove.

However, if $x_1 < x_2$, we consider $F_{x\gamma}$ for $X \leq x_2$, $X > x_1$. Oh, sorry, we want to find the probability where $x_1 < x_2$, $Y \in \mathbb{R}$, and $x_1, x_2 \in \mathbb{R}$ with $x_1 < x_2$. The probability that $X \leq x_2$,

$X > x_1$, and $Y \leq y$—what will this probability be? This probability will be $F_{XY}(x_2, y) - F_{XY}(x_1, y)$. In the univariate case, it would simply be $F_X(x_2) - F_X(x_1)$.

However, since Y is also involved here, by fixing Y, we can calculate this probability. So, how can we prove this? Now, how can we represent this event? The probability in question is essentially the probability that $X \in [x_1, x_2]$ and $Y \leq y$. What is this probability?

This probability is simply the intersection of two events: $X \in [x_1, x_2]$ and $Y \leq y$. Now, let us first determine what this event represents. So, $x_1 < X \leq x_2$ is the set of all $s \in S$ satisfying $X(s) \leq x_2$ and $X(s) > x_1$. This is the set. Now, this set can be represented as $s \in S$ such that this condition holds.

So, can we represent this as the set of s where $X(s) \leq x_2$ - the set of s where $X(s) \leq x_1$? Basically, what is this? Now we are talking about the range between x_1 and x_2. Now we want to find the probability that this random variable is taking a value between x_1 and x_2. So, we are saying, what is the probability or the event that the random variable is taking all values in the interval $(-\infty, x_2)$ - the values in $(-\infty, x_1)$?

This is the interval we are finding. So, this is simply $X \leq x_2 - X \leq x_1$. We can write this event in this way. Now, this is the same as taking the intersection of both sets, which will result in the same outcome. So, hence we can say that if $x_1 < X \leq x_2$, and you take the intersection with any set Y, this can be represented as $X \leq x_2 \cap Y \leq y$.

This set, formed by subtracting this from that, is the same as taking the intersection in both cases. So, this should be equal to the result of subtracting this. Similarly, it will be found that $X \leq x_1 \cap Y \leq y$. So, that's actually how we can show it. So, using the distributive property, you can also show this.

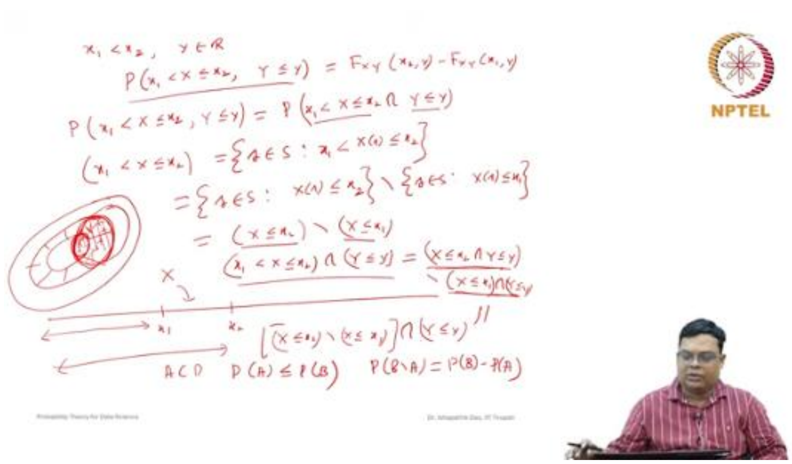
This can be represented as $(X \leq x_2 - X \leq x_1) \cap (Y \leq y)$. Then, using the distributive property, it becomes $(X \leq x_2 \cap Y \leq y) - (X \leq x_1 \cap Y \leq y)$. So, in any of these ways, you can understand this. Or, if you use a Venn diagram, you can also find it. Suppose this is $X \leq x_2$.

Actually, one minute, there is something else important to note. $X \leq x_2$ is a subset of $X \leq x_1$. So, that minus part is nothing but this. Now, if you take the intersection with any other $Y \leq y$, this is what you will get. So, this intersection part is nothing but this part only.

So, now how will we find that this is in this set, intersected with this set? So, this minus this set is the intersection, and then this set intersects with this set. After that, we will get the whole set, which is also included. So, this minus this gives us this set. That is actually explained here. One way you can see this is true is that this part will be a subset of this part. Sorry, this part will be a subset of this set. So, basically, this whole set will be a subset of this set. Then, we have theory 1.1, which we have already proved: whenever one set is a subset of another, for example, $A \subseteq B$, we know that $P(A) \leq P(B)$. We will apply that here.

Sorry, and not only that, the probability of B - A in Theorem 1 is nothing but P(B) - P(A). We will apply that here. So, now what do we have? Let's write down this probability. We started from here.

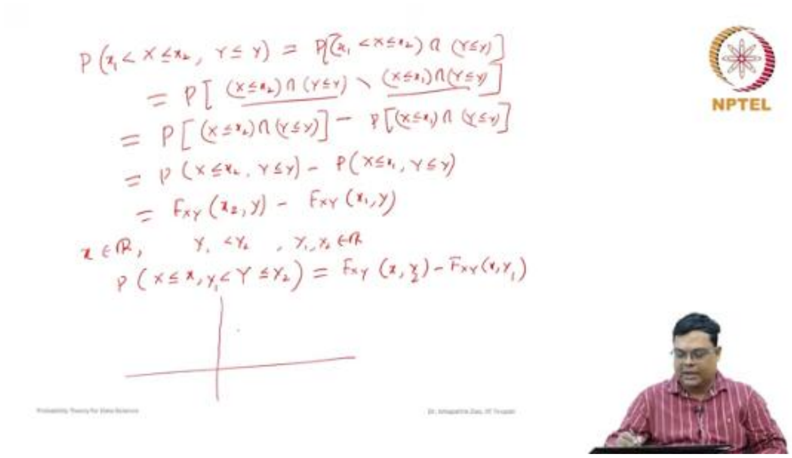
So, what will we get? The probability that $X \in [x_1, x_2]$ and $Y \leq y$. This is nothing but $P(X \in [x_1, x_2] \cap Y \leq y) = P(X \leq x_2 \cap Y \leq y) - P(X \leq x_1 \cap Y \leq y)$, which we have already shown. Just now we have shown it.



Since this set is a subset of this set, by Theorem 1.1, we can prove that the probability of $X \leq x_2$ is the probability of this minus the probability of $X \leq x_1 \cap Y \leq y$, which is the probability of B - the probability of A. So, this is nothing but $P(X \leq x_2 \cap Y \leq y) - P(X \leq x_1 \cap Y \leq y)$. By the definition of the joint cumulative distribution function, this is $F_{x\gamma}(x_2, y) - F_{x\gamma}(x_1, y)$. So, we have proved this. This is one of the properties we wanted to prove, and this is the proof.

So, now, these things we have already proved. Similarly, you can prove this. What does this mean? It means that these are the other coordinates, actually. Let x be any real number, and let y_1 and y_2 be two real numbers such that $y_1 < y_2$.

Then, the probability that $X \leq x$, and $Y \in [y_1, y_2]$, is given by $F_{x\gamma}(x, y_2) - F_{x\gamma}(x, y_1)$. So, this will be $F_{x\gamma}(x, y_2) - F_{x\gamma}(x, y_1)$. So, this is very similar. You can prove that as well. You can represent it graphically to show how it can be done.



Let's see that on the next page. What we are saying is that for $x \in \mathbb{R}$, and for y_1 and $y_2 \in \mathbb{R}$ such that $y_1 < y_2$, then we can show that the probability that $X \leq x$ and $Y \in (y_1, y_2)$ can be represented as $F_{xy}(x, y_2)$ - $F_{xy}(x, y_1)$. So, that is what we wrote just before. Yeah.