Lesson 5

Random Variables and Their CDFs

It is not always easy to work with Sample Spaces like $\{H,T\}^n$, especially when we want to quantify ~~the~~ properties of such sample spaces. Therefore, we work with a Random Variable.

Loosely speaking, a random variable is a <u>function</u> from the <u>sample space</u> to <u>real numbers</u>.

Def. Assume that $(\Omega, F, \mathbb{P})$ is a probability space. The function $X: \Omega \to \mathbb{R}$ is a random variable if $\forall c \in \mathbb{R}$, $X^{\leftarrow}((-\infty, c])$ is F-measurable, i.e. it is an event, In other words

$$X^{\leftarrow}((-\infty, c]) = \{ \qquad \}$$

is an event for all real numbers $c \in \mathbb{R}$.

Example: Consider $(\Omega, \mathcal{F}, \mathbb{P})$

where $\Omega = \{\alpha, \beta, \gamma\}$

$$\mathcal{F} = \{\emptyset, \{\alpha, \beta\}, \{\gamma\}, \Omega\}$$

and $\mathbb{P}(\emptyset) = 0$, $\mathbb{P}(\Omega) = 1$

$\mathbb{P}(\{\alpha, \beta\}) = 1/2$, $\mathbb{P}(\{\gamma\}) = 1/2$

Consider $g : \Omega \to \mathbb{R}$

$$g(\omega) = \begin{cases} 1 & \omega \in \{\alpha, \beta\} \\ 2 & \omega \in \{\gamma\} \end{cases}$$

$\mathbb{P}(g(\omega) = 1) = \mathbb{P}(\{\omega | \qquad \})$

$= \mathbb{P}(\qquad) =$

$\mathbb{P}(g(\omega) = 2) = \mathbb{P}(\{\omega | \qquad \})$

$= \mathbb{P}(\qquad) =$

$$\mathbb{P}(g(\omega) \leq c) = \begin{cases} & c < 1 \\ & 1 \leq c < 2 \\ & c \geq 2 \end{cases}$$

For all values of c, $\mathbb{P}()$ is well-defined/ not defined

Next, consider $h : \Omega \to \mathbb{R}$

$$h(\omega) = \begin{cases} 0 & \omega = \alpha \\ 1 & \omega = \beta \\ 2 & \omega = \gamma \end{cases}$$

$$\mathbb{P}(h(\omega) = 0) = \mathbb{P}(\{\omega \mid \quad \})$$

$$= \mathbb{P}(\{ \quad \}) =$$

Issue: The σ-field on which $\mathbb{P}$ is defined is not rich/large enough.

Convention: Upper-case letters are used for r.v.'s: X, Y, Z, and lower case letters are used for numerical values that they can take.

It is an appropriate practice to show the dependence of a random variable on $\omega \in \Omega$:

$X(\omega), Y(\omega), Z(\omega)$

Example: Indicator Function of A Measurable Set.

Let $(\Omega, \mathcal{F})$ be a measurable space and $A \in \mathcal{F}$.

The indicator function of A:

$$I_A : \Omega \to \{0,1\}$$

$$I_A(\omega) = \begin{cases} 1 & \omega \in A \\ 0 & \omega \notin A \end{cases}$$

is a random variable:

$$\overleftarrow{X}((-\infty, 2]) =$$

$\overleftarrow{X}((-\infty,-2]) =$

$\overleftarrow{X}((-\infty,\tfrac{1}{2}]) =$

In general

$$\overleftarrow{X}((-\infty, c]) = \begin{cases} & c < 0 \\ & 0 \leq c < 1 \\ & c \geq 1 \end{cases}$$

Note: It is common to show the event

$\overleftarrow{X}((-\infty, 2])$ as

$\{X \leq 2\}$

Example: If $(\Omega, \mathcal{F})$ is a measurable space equipped with probability measure $\mathbb{P}$, and $A \notin \mathcal{F}$, I_A, the indicator function of A, is NOT a random variable:

$$I_A : \Omega \to \{0,1\}$$

$$I_A(\omega) = \begin{cases} 1 & \omega \in A \\ 0 & \omega \notin A \end{cases}$$

$$\{X \leq 1/2\} = X^{\leftarrow}((-\infty, 1/2])$$

$$= \{\omega \mid X(\omega) \leq 1/2\} =$$

When $A \in \mathcal{F}$, I_A is called a "Bernoulli Random Variable"

Bernoulli Random Variable represents an experiment that has two types of outcome: success and failure

If A occurs (outcome belongs to A), success occured

If A^c occurs (outcome doesn't belong to A), failure occured.

$\mathbb{P}(\text{success}) = \mathbb{P}(X=1)$

$= \mathbb{P}(I_A = 1) = \mathbb{P}(\{\omega \mid I_A(\omega)=1\}$

$= \mathbb{P}(\quad)$

$\mathbb{P}(\text{failure}) = \mathbb{P}(X=0)$

$= \mathbb{P}(I_A = 0) = \mathbb{P}(\{\omega \mid I_A(\omega)=0\})$

$= \mathbb{P}(\quad)$

Example: A function of a r.v.

Assume that X is a r.v. on $(\Omega, \mathcal{F})$. Also assume that $Y : \Omega \to \mathbb{R}$

$$Y(\omega) = X^3(\omega)$$

Is Y a random variable?

In order to show that Y is a random variable, we must show that

$$\overleftarrow{Y}((-\infty, c]) =$$

is an event for all $c \in \mathbb{R}$.

Remark: The term random variable is a misnomer, because a random variable is <u>neither random, nor is it a variable</u>. A random variable, as we observed, is

just a function from the sample space Ω to real numbers $\mathbb{R}$.

The term "random" actually refers to the randomness in observing a particular member (or particular members) of a sample space Ω, as the outcome of an experiment. The random variable maps the outcome (ω) to a "fixed" real number, $X(\omega)$.

One can imagine different members of the sample space are selected randomly, and are mapped into different real numbers, through the random variable. However, the random variable does not assign real numbers

to the members of the sample space "randomly"

Remark: It is also VERY IMPORTANT to note that probability is assigned to subsets of the sample space (events)

so: $\mathbb{P}: \mathcal{F} \longrightarrow [0,1]$, but

a random variable assigns a

real number to each element

of the sample space, therefore

$X: \Omega \longrightarrow \mathbb{R}$.

Borel Sets and The Probability Law of A Random Variable

Def (σ-field generated by a set)

Let C be any collection of

subsets of Ω, i.e. $C \subseteq 2^{\Omega}$, which

is <u>not</u> necessarily a σ-field.

Define $\sigma(C)$ to be <u>the smallest</u>

<u>σ-field</u> that contains C, which

means that if D is a

σ-field and $\mathcal{C} \subseteq \mathcal{D}$, then

$\sigma(\mathcal{C}) \subseteq \mathcal{D}$.

Example: We previously saw that the smallest σ-field that contains some $A \subset \Omega$ is:

$\mathcal{F} = \{\emptyset, A, A^c, \Omega\}$.

$\mathcal{F}$ is the σ-field generated by $\mathcal{C} = \{A\} \subseteq 2^{\Omega}$.

It is obvious that any other σ-field that contains $\mathcal{C} = \{A\}$

is a superset of $\sigma(C) = \{\emptyset, A, A^c, \Omega\}$. For example, if $B \neq A$, the σ-field that has both B, A as its members, is a superset of $\sigma(C) = \{A, A^c, \Omega, \emptyset\}$.

Exercise: Try to enumerate the members of the σ-field that has both A and B as its members.

Borel σ-field:

Def: The Borel σ-field of subsets of $\mathbb{R}$, denoted as $\mathcal{B}(\mathbb{R})$, or simply $\mathcal{B}$, is a σ-field <u>generated</u> by all of the intervals

$(-\infty, c]$, $c \in \mathbb{R}$.

Question: Are intervals in the form $(c, +\infty)$, $c \in \mathbb{R}$ and (c, d), $c, d \in \mathbb{R}$ members of $\mathcal{B}$?

It can be proved that all intervals, their countable unions and intersections, and any set that is a result of countable applications of $\cup$, $\cap$, and $\setminus$ to intervals are in $\mathcal{B}(\mathbb{R})$.

Remark: It can be proved that any singleton $\{a\}$ is a Borel Set, $\forall a \in \mathbb{R}$, e.g $\{1\}$, $\{2.3\}$, $\{\pi\}$ are Borel Sets.

- Any finite/countably infinite subset of $\mathbb{R}$ is a Borel Set.

- Remark: Borel Sets should not be confused with the Borel σ-field itself.

Borel Sets are the members of $\mathcal{B}(\mathbb{R})$, the Borel σ-field.

- Proposition: The pre-image $\overleftarrow{X}(B)$ of any Borel Set $B \in \mathcal{B}(\mathbb{R})$ is an event, if X is a r.v. on $(\Omega, \mathcal{F}, \mathbb{P})$

(Sketch of Proof): Any Borel Set is a countable union of intervals of the form $(-\infty, c]$ or their complements $(c, +\infty)$, so,

$$B = \bigcup_n A_n$$

where A_n is either $(-\infty, c_n]$ or $(-\infty, d_n]^c$.

$\overleftarrow{X}(B) =$

The Probability Law of A Random Variable

Since $\forall B \in \mathcal{B}(\mathbb{R})$, $\overleftarrow{X}(B)$ is an event (is $\mathcal{F}$-measurable), we can define the probability of $\overleftarrow{X}(B)$:

$$\mathbb{P}(\overleftarrow{X}(B)) = \mathbb{P}(\{\omega | \qquad \})$$

$$= \mathbb{P}_X(B)$$

Precisely, $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathbb{P}_X)$ is a probability space induced/created by X, where

$$\mathbb{P}_X(B) = \mathbb{P}(\qquad)$$

Example: Assume that

$\Omega = \{1, 2, 3, 4, 5, 6\}$ and $\mathbb{P}(A) = \frac{|A|}{6}$

$\forall \omega \in \Omega$, $X_1(\omega) = 2$ and $X_2(\omega) = \omega^2$.

$X_3(\omega) = \sqrt{\omega}$

Calculate $\mathbb{P}_{X_1}(B)$ and $\mathbb{P}_{X_2}(B)$

$\mathbb{P}_{X_3}(B)$

when:

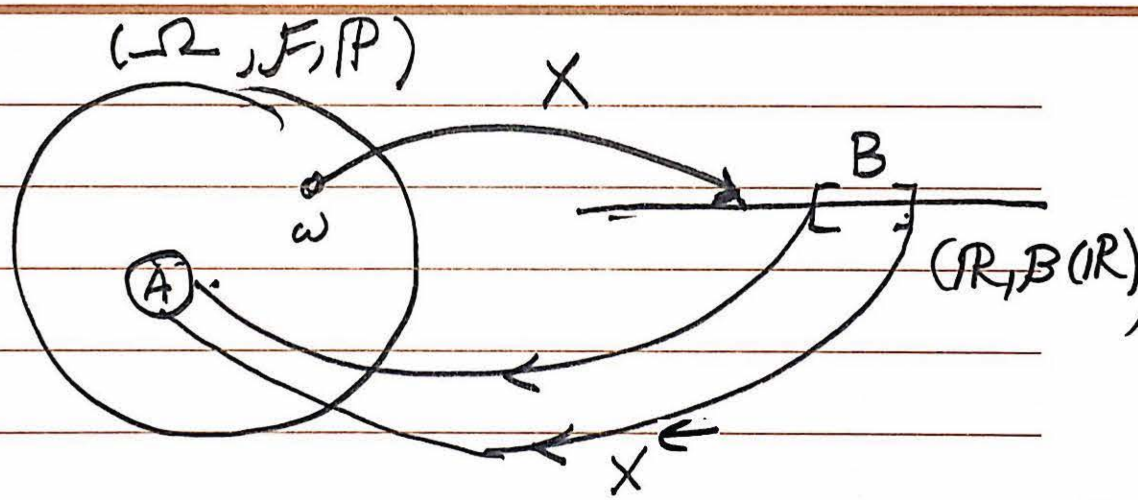
a) $B = (-\infty, -10]$

(b) $B = [1.5, 2.5]$

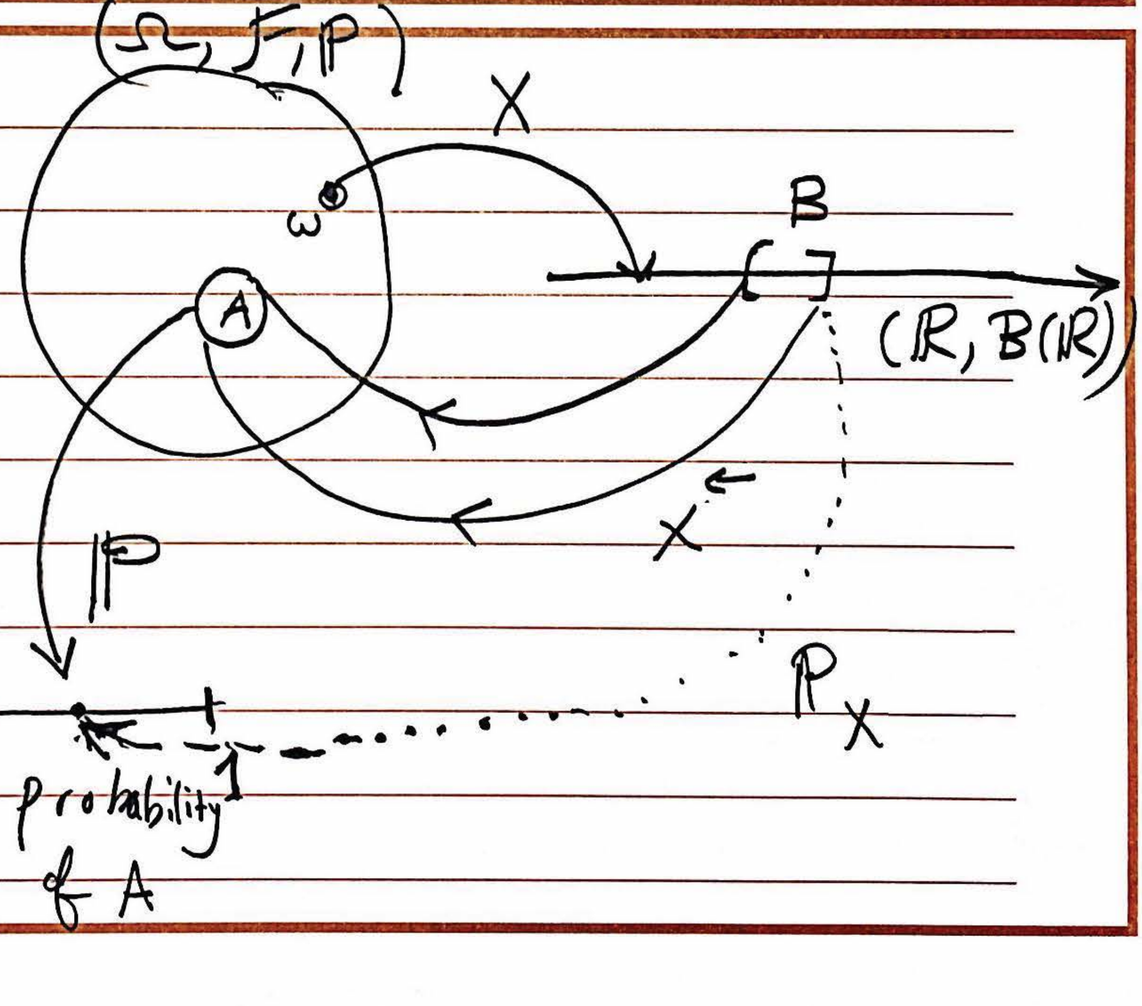
(c) $B = \mathbb{N}$

Remark: Borel sets shed light on the usefulness of the concept of a random variable. A random variable helps creating a probability space on $\mathbb{R}$,

whose σ-field is $\mathcal{B}(\mathbb{R})$, and its probability measure is $\mathbb{P}_X(\cdot)$. Instead of dealing with general sample spaces and σ-fields, by using a random variable, we can deal with $\mathbb{R}$ and $\mathcal{B}(\mathbb{R})$, and compute probabilities of "numerical" measurable sets such as intervals.



In this picture, X is a random variable that maps every element ω in the sample space to a real number, in a way that the pre-image of every Borel Set $B \in \mathcal{B}(\mathbb{R})$ is an event $A \in \mathcal{F}$.



In this picture, $\overleftarrow{X}$ maps any Borel subset of $\mathbb{R}$ into an event A, and the probability measure $\mathbb{P}$ maps the event into a number in $[0,1]$.

Therefore, the probability law of X, $\mathbb{P}_X$, can be viewed as composition of $\mathbb{P}$ and $\overleftarrow{X}$;

$$\mathbb{P}_X(B) = \mathbb{P}(\overleftarrow{X}(B)) = \mathbb{P} \circ \overleftarrow{X}(B)$$

Theorem (Construction of r.v.'s)

(a) If $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space, and $A \in \mathcal{F}$ (is an event), the corresponding indicator function of A, $I_A(\omega)$ is a random variable, where

$$I_A(\omega) = \begin{cases} 0 & \omega \notin A \\ 1 & \omega \in A \end{cases}$$

(b) If $A_1, A_2, \ldots, A_n \in \mathcal{F}$ and $\alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbb{R}$, then:

$$X(\omega) = \sum_{i=1}^{n} \alpha_i I_{A_i}(\omega)$$

is a random variable.

(c) Any continuous real function is a random variable. So, if $\Omega = \mathbb{R}$, $X: \mathbb{R} \to \mathbb{R}$ is a random variable if ~~X~~ is a continuous function, i.e.

$$\lim_{\omega \to \omega_0} X(\omega) = X(\omega_0) \quad \forall \omega_0 \in \mathbb{R}$$

(d) Let $f: \mathbb{R} \to \mathbb{R}$ be a continuous function (i.e. $\lim_{x \to x_0} f(x) = f(x_0) \; \forall x_0 \in \mathbb{R}$) then $Y = f(X)$ is also a random variable

Remark: (d) doesn't mean that if f is not continuous, $f(X)$ is NECESSARILY not a random variable. $f(X)$ CAN still be a random variable if f is not continuous

(e) (Functions of Multiple Random Variables) Assume that $X_1, \ldots, X_n$ are random variables on $(\Omega, \mathcal{F}, P)$ and assume that $f: \mathbb{R}^n \to \mathbb{R}$ is a continuous function.

Then, $Y = f(X_1, X_2, \ldots, X_n)$ is also a random variable.

Example: $y = f(x_1, x_2) = x_1 + x_2$ and $y = g(x_1, x_2) = x_1 x_2$ are continuous functions, so $X_1 + X_2$ and

$X_1 X_2$ are random variables

if X_1 and X_2 are random

variables.

The Cumulative Distribution Function (CDF)

Def: The cdf (cumulative distribution function) of a r.v. X, $F_X : \mathbb{R} \to [0,1]$ is defined as:

$$F_X(x) = \mathbb{P}(\underbrace{X \leq x}_{\text{event}})$$

$$\{X \leq x\} =$$

Ques: Explain why any random variable has a well-defined CDF, i.e. $\forall X$, $F_X(x)$ exists for all $x \in \mathbb{R}$.

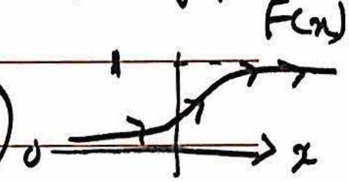
Lemma : A cdf of X, a random variable defined on $(\Omega, \mathcal{F}, \mathbb{P})$, has the following properties:

(a) $\lim_{x \to -\infty} F_X(x) =$

$$\lim_{x \to +\infty} F_X(x) =$$

(b) $x < y \Rightarrow F_X(x) \leq F_X(y)$

(It is monotonic).



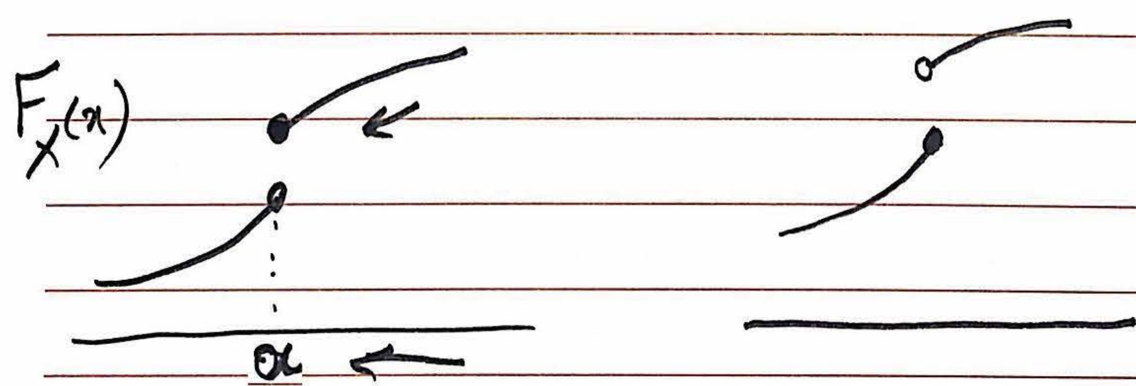
(c) F is right-continuous,

i.e.

$$\lim_{x \to a^+} F_X(x) = F_X(a)$$

or

$F(x+h) \to F(x)$ when $h \to 0^+$



Proof:

(a) Define $B_n = \{\omega \in \Omega \mid X(\omega) \leq -n\}$

To prove $\lim_{x \to \infty} F_X(x)$, we similary need to define an appropriate set of B_n's. It is left as an exercise.

(b)

(C) Again, another appropriate set of events has to be defined. Then, one should use continuity of probability.

Example(Constant Random Variable)

Let $c \in \mathbb{R}$, $X : \Omega \to \mathbb{R}$

$X(\omega) = c \quad \forall \omega \in \mathbb{R}$

What is the cdf of X?

Lemma (Properties of cdf).

Let X be a r.v. on $(\Omega, \mathcal{F}, \mathbb{P})$

and F_X be the cdf of X.

Then:

(a) $\mathbb{P}(X > x) = \mathbb{P}(X \in \underbrace{(x, +\infty)}_{\text{Borel set}})$

$=$

(b) $\mathbb{P}(x < X \leq y) = \mathbb{P}(X \in \underbrace{(x, y]}_{\text{Borel set}})$

$=$

c) $\mathbb{P}(X = x) = \mathbb{P}(X \in \underbrace{\{x\}}_{\text{Borel Set}})$

$=$

Example (Constant Random Variable)

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

$$X(\omega) = c$$

What is $\mathbb{P}(\{X = c\})$?

Exercise: Assume that X is a r.v. on $(\Omega, \mathcal{F}, \mathbb{P})$ and F_X is its CDF.

Is $G(x) = [1 - F(x)] \log(1 - F(x)) + F(x)$ a cdf of a random variable?

Exercise: Are the following statements equivalent

(1) $F_X(a)$ is continuous for all $a \in \mathbb{R}$

(2) $\mathbb{P}\{X=a\}=0 \quad \forall a \in \mathbb{R}$

Exercise: Let X be a r.v. with continuous cdf F_X. Find the cdf of the following random variables:

(a) X^2

(b) $\sqrt{X}$

(c) $F_X(X)$

(d) $G^{-1}(X)$, where G is a continuous and strictly increasing function

Exercise:

Let X have the CDF

$$F_X(x) = \begin{cases} 0 & x < 0 \\ \frac{1}{2}x & 0 \le x \le 2 \\ 1 & x > 2 \end{cases}$$

and let $Y = X^2$. Find

(a) $\mathbb{P}(1 \le X < 2)$

(b) $\mathbb{P}(2X + Y \leq 3)$

Independence of Random Variables

Recall that the events $A, B \in \mathcal{F}$ ~~are~~ are said to be independent iff:

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\,\mathbb{P}(B)$$

Based on the concept of independence for events we define independence for r.v.'s:

Def. Assume X_1, X_2 are r.v.'s on $(\Omega, \mathcal{F}, \mathbb{P})$. They are said to be independent, if for all Borel

sets $B_1, B_2 \in \mathcal{B}(\mathbb{R})$:

$\mathbb{P}(\underbrace{X_1 \in B}, \underbrace{X_2 \in B})$

$= \cdot|$

Question: How can we extend the definition of independence a set of r.v.'s $\{X_i \mid i \in I\}$, where I is an index set?

Def(Joint cdf):

Assume that X_1 and X_2 are r.v.'s on $(\Omega, \mathcal{F}, \mathbb{P})$. The joint cdf of X_1 and X_2 is defined as $F_{X_1, X_2}(x_1, x_2) =$

More generally, if $X_1, X_2, \ldots, X_n$ are random variables on $(\Omega, \mathcal{F}, \mathbb{P})$, the joint cdf of $X_1, X_2, \ldots, X_n$ is denoted as $F_{X_1, X_2, \ldots, X_n}$ and is defined as:

$$F_{X_1, X_2, \ldots, X_n}(x_1, x_2, \ldots, x_n)$$

$$= \mathbb{P}$$

Lemma: X_1, X_2 are independent if and only if $\forall x_1, x_2 \in \mathbb{R}$

$$F_{X_1,X_2}(x_1,x_2) =$$

More generally, $X_1, X_2, \ldots, X_n$ are independent if and only if

$\forall x_1, x_2, \ldots x_n \in \mathbb{R}$

$$F_{X_1,X_2,\ldots,X_n}(x_1,x_2,\ldots,x_n) =$$

Remark: The previous lemma states that instead of checking the following equation for all possible Borel sets A, B,

$$\mathbb{P}(X_1 \in A, X_2 \in B) = \mathbb{P}(X_1 \in A)\,\mathbb{P}(X_2 \in B)$$

one can check it only for A's and B's in the following form $A = (-\infty, x_1]$, $B = (-\infty, x_2]$

$$\mathbb{P}(X_1 \in A, X_2 \in B) = \mathbb{P}(X_1 \in (-\infty, x_1], X_2 \in (-\infty, x_2])$$