Lectures 13 - 17
In this chapter, we associate with random variable, a nondecreasing function called distribution function and study its properties.

**Definition 4.1. (Distribution function)** Let $X : \Omega \to \mathbb{R}$ be a random variable on $(\Omega, \mathcal{F}, P)$. The function $F : \mathbb{R} \to \mathbb{R}$ defined by

$$ F(x) = P\{X \leq x\} $$

is called the distribution function of $X$

**Theorem 4.0.13** The distribution function has the following properties

(i) $\lim_{x \to -\infty} F(x) = 0$ and $\lim_{x \to \infty} F(x) = 1$.

(ii) $F$ is non decreasing.

(iii) $F$ is right continuous.

**Proof:**

(i) Let $x_n \downarrow -\infty$, set

$$ A_n = \{X \leq x_n\} $$

Then

$$ A_1 \supseteq A_2 \supseteq \cdots \text{ and } \bigcap_n A_n = \emptyset. $$

Therefore

$$ P(A_n) \to 0 $$

i.e.,

$$ P\{X \leq x_n\} \to 0. $$

Hence

$$ \lim_{x \to -\infty} F(x) = 0. $$

Using similar argument, we can prove

$$ \lim_{x \to \infty} F(x) = 1. $$

(ii) For $x_1 \leq x_2$, $\{X \leq x_1\} \subseteq \{X \leq x_2\}$. Hence $F(x_1) \leq F(x_2)$.

(iii) We have to show that, for each $x \in \mathbb{R}$

$$ \lim_{y \downarrow x} F(y) = F(x). $$

Let $y_n \downarrow x$. Set
\[ A_n = \{ X \leq y_n \}. \]

Then
\[ A_1 \supset A_2 \supset \cdots \]

and
\[ \{ X \leq x \} = \cap_{n=1}^{\infty} A_n. \]

Therefore
\[ F(x) = P\{ X \leq x \} = \lim_{n \to \infty} P(A_n) = \lim_{n \to \infty} F(y_n). \]

**Theorem 4.0.14**  The set discontinuity points of a distribution function \( F \) is countable.

**Proof.** Let \( F \) be the distribution function of a random variable \( X \) and
\[ D = \text{set of all points of discontinuity of } F. \]

Then
\[ D = \{ x \in \mathbb{R} \mid F(x) - F(x-) > 0 \}, \]

where
\[ F(x-) = \lim_{y \uparrow x} F(y) = P\{ X < x \}. \]

Set
\[ D_n = \{ x \in \mathbb{R} \mid F(x) - F(x-) \geq \frac{1}{n} \}. \]

Then
\[ D = \bigcup_{n}^{\infty} D_n. \]

Also note that for \( x \in \mathbb{R} \)
\[ P\{ X = x \} = F(x) - F(x-). \]

If \( x_1, x_2, \ldots, x_m \) be distinct points in \( D_n \). Then
\[ 1 \geq \sum_{k=1}^{m} P\{ X = x_k \} \]
\[ = \sum_{k=1}^{m} F(x_k) - F(x_k-) \]
\[ \geq \frac{m}{n}. \]

Therefore
\[ \#D_n \leq n. \]

Hence \( D \) is countable.

**Theorem 4.0.15** Let \( X \) be a random variable on a probability space \( (\Omega, \mathcal{F}, P) \). For \( B \in \mathcal{B}_{\mathbb{R}} \), define
\[ \mu(B) = P\{X \in B\}. \]

Then \( \mu \) is a probability measure on \((\mathbb{R}, \mathcal{B}_{\mathbb{R}})\).

**Proof.**

\[ \mu(\mathbb{R}) = P\{X \in \mathbb{R}\} = P\Omega = 1. \]

Also

\[ \mu(B) = P\{X \in B\} \geq 0 \]

for all \( B \in \mathcal{B}_{\mathbb{R}} \). Let \( B_1, B_2, \ldots \in \mathcal{B}_{\mathbb{R}} \) be pairwise disjoint, then

\[
\mu(\bigcup_{i=1}^{\infty} B_i) = \mu(\{X \in \bigcup_{i=1}^{\infty} B_i\}) = \mu(\{\bigcup_{i=1}^{\infty} X \in B_i\}) = \sum_{i=1}^{\infty} \mu(B_i) = \sum_{i=1}^{\infty} P\{X \in B_i\}.
\]

[Note that \( B_i \)'s are disjoint imply that \( \{X \in B_i\}'s are disjoint]

This completes the proof.

**Definition 4.2** The probability measure \( \mu \) is called the distribution or law of the random variable \( X \) and we denote it by \( \mathcal{L}(X) \).

Note that if \( F \) is the distribution function of \( X \), then

\[ F(x) = \mathcal{L}(X)(-\infty, x]. \]

**Example 4.0.1** Consider the probability space \((\Omega, \mathcal{F}, P)\) given by

\[ \Omega = \{HH, HT, TH, TT\}, \]

\[ \mathcal{F} = \mathcal{P}(\Omega), \]

\[ P\{HH\} = P\{HT\} = P\{TH\} = P\{TT\} = \frac{1}{4}. \]

Define \( X : \Omega \rightarrow \mathbb{R} \) as follows. For \( \omega \in \Omega \),

\[ X(\omega) = \text{number of heads in } \omega. \]

Then \( X \) takes values from \( \{0, 1, 2\} \). The distribution function of \( X \) is

\[
\mathcal{L}(X)(B) = P\{X \in B\} = \left\{
\begin{array}{ll}
0 & \text{if } \{0, 1, 2\} \cap B = \emptyset \\
\frac{1}{4} & \text{if } \{0, 1, 2\} \cap B = \{0\} \\
\frac{1}{2} & \text{if } \{0, 1, 2\} \cap B = \{1\} \\
\frac{1}{4} & \text{if } \{0, 1, 2\} \cap B = \{2\} \\
1 & \text{if } \{0, 1, 2\} \cap B = \{0, 1\} \\
1 & \text{if } \{0, 1, 2\} \cap B = \{0, 1, 2\}.
\end{array}\right.
\]
The distribution function of the random variable is given by
\[ F(x) = \begin{cases} 
0 & \text{if } x < 0 \\
\frac{1}{4} & \text{if } 0 \leq x < 1 \\
\frac{3}{4} & \text{if } 1 \leq x < 2 \\
1 & \text{if } x \geq 2.
\end{cases} \]

**Example 4.0.2** (Bernoulli distribution) On \((\mathbb{R}, \mathcal{B}_\mathbb{R})\), for \(B \in \mathcal{B}_\mathbb{R}\), define
\[
\mu(B) = \begin{cases} 
0 & \text{if } B \cap \{0, 1\} = \emptyset \\
1 - p & \text{if } B \cap \{0, 1\} = \{0\} \\
p & \text{if } B \cap \{0, 1\} = \{1\} \\
1 & \text{if } B \cap \{0, 1\} = \{0, 1\}.
\end{cases}
\]

Then \(\mu\) is a probability measure on \((\mathbb{R}, \mathcal{B}_\mathbb{R})\) called the Bernoulli distribution.

A random variable with Bernoulli distribution is called a Bernoulli \(\left(p\right)\) random variable.

For example: Let
\[
\Omega = \{H, T\}, \quad \mathcal{F} = \mathcal{P}(\Omega)
\]
and define the probability measure \(P\) on \(\mathcal{F}\) such that
\[
P\{H\} = p, \quad P\{T\} = 1 - p.
\]

Define \(X : \Omega \to \mathbb{R}\) by
\[
X(H) = 1, \quad X(T) = 0.
\]

Then
\[
\mathcal{L}(X) = \mu.
\]

Hence \(X\) is a Bernoulli random variable. The distribution function corresponding to Bernoulli distribution is given by
\[
F(x) = \begin{cases} 
0 & \text{if } x < 0 \\
1 - p & \text{if } 0 \leq x < 1 \\
1 & \text{if } x \geq 1.
\end{cases}
\]

**Example 4.0.3** (Binomial distribution with parameters \((n, p)\)). On \((\mathbb{R}, \mathcal{B}_\mathbb{R})\), the probability measure \(\mu\) given by
\[ \mu(\{k\}) = \binom{n}{k} p^k (1 - p)^{n-k}, \quad k = 0, 1, 2, \ldots, n \]

is said to be a Binomial distribution with parameters \((n, p)\).

It is left as an exercise for the student to write down \(\mu(B)\) for \(B \in \mathcal{B}_\mathbb{R}\).

A random variable with Binomial distribution as its law (distribution) is called a Binomial random variable.

For example: Let

\[ \Omega = \{a_1 a_2 \ldots a_n | a_i \in \{H, T\}, i = 1, 2, \ldots, n\}, \]

\[ \mathcal{F} = \mathcal{P}(\Omega) \]

and probability measure \(P\) on \(\mathcal{F}\) is defined by

\[ P\{a_1 a_2 \ldots a_n\} = \binom{n}{k} p^k (1 - p)^{n-k}, \]

where \(k\) is the number of \(H\) in the string \(a_1 a_2 \ldots a_n\). Define \(X: \Omega \rightarrow \mathbb{R}\) as follows.

\[ X(a_1 a_2 \ldots a_n) = \text{number of } H \text{ in } a_1 a_2 \ldots a_n. \]

Then

\[ \mathcal{L}(X) = \mu. \]

The distribution function corresponding to Binomial distribution \((n, p)\) is given by

\[
F(x) = \begin{cases} 
0 & \text{if } x < 0 \\
\binom{n}{0} (1 - p)^n & \text{if } 0 \leq x < 1 \\
\binom{n}{0} (1 - p)^n + \binom{n}{1} p (1 - p)^{n-1} & \text{if } 1 \leq x < 2 \\
\sum_{i=0}^{k} \binom{n}{i} p^i (1 - p)^{n-i} & \text{if } k \leq x < k + 1, \; k = 2, \ldots, n - 1 \\
1 & \text{if } x \geq n.
\end{cases}
\]

**Example 4.0.29**

(Poisson distribution with parameter \(\lambda\)). On \((\mathbb{R}, \mathcal{B}_\mathbb{R})\) define probability measure

\[ \mu(\{k\}) = \frac{\lambda^k e^{-\lambda}}{k!}, \quad k = 0, 1, 2, \ldots \]

Then \(\mu\) is called Poisson distribution with parameter \(\lambda\).

**Example 4.0.30**
(Normal distribution with parameters $\mu, \sigma$)

On $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ define $\mu$ as follows:

$$\mu(B) = \frac{1}{\sqrt{2\pi}\sigma} \int_B e^{-\frac{(x-\mu)^2}{2\sigma^2}} \, dx.$$  

Then $\mu$ is a probability measure on $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ and is called the normal distribution.

The distribution function $F: \mathbb{R} \to \mathbb{R}$ corresponding to normal distribution with parameters $\mu, \sigma$ is given by

$$F(x) = \mu(-\infty, x] = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^x e^{-\frac{(y-\mu)^2}{2\sigma^2}} \, dy.$$  

**Example 4.0.31**

(Uniform distribution on $(0, 1]$) On $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$, define $\mu$ as follows:

$$\mu(B) = \int_B I_{(0,1]}(x) \, dx,$$

where

$$I_{(0,1]}(x) = \begin{cases} 1 & \text{if } 0 < x \leq 1 \\ 0 & \text{otherwise} \end{cases}.$$  

Then $\mu$ is a probability measure and is called the uniform distribution on $(0, 1]$. The distribution function is given by

$$F(x) = \begin{cases} 0 & \text{if } x \leq 0 \\ x & \text{if } 0 < x \leq 1 \\ 1 & \text{if } x \geq 1. \end{cases}$$

**Example 4.0.32**

(Exponential distribution with parameter $\lambda > 0$) On $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ define $\mu$ as follows:

$$\mu(B) = \int_B \lambda e^{-\lambda x} I_{[0, \infty)}(x) \, dx.$$  

Then $\mu$ is a probability measure and is called the exponential distribution with parameter $\lambda$. The distribution function is given by

$$F(x) = \begin{cases} 0 & \text{if } x \leq 0 \\ 1 - e^{-\lambda x} & \text{if } x > 0. \end{cases}$$
A random variable $X$ with distribution function $F : \mathbb{R} \rightarrow \mathbb{R}$ is said to be a discrete random variable if

$$\sum_{x \in D} (F(x) - F(x^-)) = 1,$$

where $D$ is the set of discontinuities of $F$.

The distributions in Examples 4.0.26, 4.0.27, 4.0.28, 4.0.29 corresponds to discrete random variables.

**Definition 4.5:**
A random variable $X$ with distribution function $F : \mathbb{R} \rightarrow \mathbb{R}$ is said to be continuous if $F$ is continuous.

The distributions given in Examples 4.0.30, 4.0.31, 4.0.32 corresponds to continuous random variable.

**Definition 4.5:**
(Probability mass function)
Let $X$ be a discrete random variable with distribution function $F : \mathbb{R} \rightarrow \mathbb{R}$. Define $f : \mathbb{R} \rightarrow \mathbb{R}$ as follows:

$$f(x) = F(x) - F(x^-)$$

Then $f$ is called the probability mass function (pmf) of $X$.

For example, the pmf of the discrete random variable given in Example 4.0.26 is given by

$$f(x) = \begin{cases} \frac{1}{2} & \text{if } x = 0, 2 \\ \frac{1}{2} & \text{if } x = 1 \\ 0 & \text{otherwise.} \end{cases}$$

It is left as an exercise for the student to write down the pmf of random variables in Examples 4.0.27, 4.0.28, 4.0.29.

The pmf of a continuous random variable is the zero function. Hence the notion of pmf is useless for continuous random variables.

**Definition 4.5:**
(Probability density function)
A continuous random variable $X$ with distribution function $F : \mathbb{R} \rightarrow \mathbb{R}$ is said to have a probability density function (pdf) if there exists a function $f : \mathbb{R} \rightarrow \mathbb{R}$ such that

$$F(x) = \int_{-\infty}^{x} f(y)dy \quad \forall \ x \in \mathbb{R}$$

If $f : \mathbb{R} \rightarrow \mathbb{R}$ exists, then it is called the pdf of $X$.

A continuous random variable with pdf is said to be an **absolutely continuous** random variable.

It is easy to see that if $F$ is differentiable every where and the derivative denoted by $F'$ is a continuous function, then the corresponding random variable $X$ has a pdf and is given by $f = F'$. This is not a necessary condition.
Example 4.0.33:

Define $F: \mathbb{R} \to \mathbb{R}$ as follows.

$$F(x) = \begin{cases} 
  0 & \text{if } x < 0 \\
  x & \text{if } 0 \leq x < \frac{1}{2} \\
  \frac{1}{2} & \text{if } \frac{1}{2} \leq x < 1 \\
  x - \frac{1}{2} & \text{if } 1 \leq x < \frac{3}{2} \\
  1 & \text{if } x \geq \frac{3}{2}.
\end{cases}$$

Student can verify that $F$ corresponds to distribution function of the random variable given by the random experiment of picking a point 'at random' from $[0, \frac{1}{2}] \cup [1, \frac{3}{2}]$.

Then $F$ is a distribution function corresponding to a continuous random variable. But $F$ is not differentiable at $x = \frac{1}{2}, 1$. The function

$$f(x) = \begin{cases} 
  0 & \text{if } x < 0 \\
  1 & \text{if } 0 \leq x < \frac{1}{2} \\
  0 & \text{if } \frac{1}{2} \leq x < 1 \\
  1 & \text{if } 1 \leq x < \frac{3}{2} \\
  0 & \text{if } x \geq \frac{3}{2}.
\end{cases}$$

is the pdf of $F$. 