MODULE 2

RANDOM VARIABLE AND ITS DISTRIBUTION

LECTURES 7 -11

Topics

2.1 RANDOM VARIABLE

2.2 INDUCED PROBABILITY MEASURE

2.3 DISTRIBUTION FUNCTION AND ITS PROPERTIES

2.4 TYPES OF RANDOM VARIABLES: DISCRETE, CONTINUOUS AND ABSOLUTELY CONTINUOUS

MODULE 2

RANDOM VARIABLE AND ITS DISTRIBUTION

LECTURE 7

Topics

2.1 RANDOM VARIABLE
2.2 INDUCED PROBABILITY MEASURE

2.1 RANDOM VARIABLE

Let $(\Omega, \mathcal{F}, P)$ be a probability space. On many occasions we may not be directly interested in the whole sample space Ω. Rather we may be interested in some numerical characteristic of the sample space Ω, as the following example illustrates.

Example 1.1

Let three distinguishable dice be labeled as A, B and C. Consider the random experiment of rolling these three dice. Then the sample space is $\Omega = \{(i, j, k): i, j, k \in \{1, 2, \dots, 6\}\}$; here an outcome $(i, j, k) \in \Omega$ indicates that the dice A, B, and C show, respectively, i, j and k number of dots on their upper faces. Suppose that our primary interest is on the study of random phenomenon of sum of number of dots on the upper faces of three dice. Here we are primarily interested in the study of the function $X: \Omega \to \mathbb{R}$, defined by

$$X\big((i, j, k)\big) = i + j + k, \ \ (i, j, k) \in \Omega. \blacksquare$$

Moreover, generally, the sample space Ω is quite abstract and thus may be tedious to deal with. In such situations it may be convenient to study the probability space $(\Omega, \mathcal{F}, P)$ through the study of a real-valued function defined on Ω.

Example 1.2

Consider the random experiment of tossing a fair coin twice. Here the sample space $\Omega = \{\mathrm{HH}, \mathrm{HT}, \mathrm{TH}, \mathrm{TT}\}$, where H and T stand for head and tail respectively and in an outcome (e.g., HT) the first letter (e.g., H in HT) indicates the result of the first toss and the second letter (e.g., T in HT) indicates the result of the second toss. Since we are more comfortable in dealing with real numbers it may be helpful to identify various outcomes in Ω with different real numbers (e.g., identify $\mathrm{HH}, \mathrm{HT}, \mathrm{TH}$ and TT with $1, 2, 3$ and 4 respectively). This amounts to defining a function $X: \Omega \to \mathbb{R}$ on the sample space (e.g., $X: \Omega \to \mathbb{R}$, defined as $X(\mathrm{HH}) = 1,\ X(\mathrm{HT}) = 2,\ X(\mathrm{TH}) = 3,$ and $X(\mathrm{TT}) = 4$). $\blacksquare$

The above discussion suggests the desirability of study of real valued functions $X: \Omega \to \mathbb{R}$ defined on the sample space Ω.

Consider a function $X: \Omega \to \mathbb{R}$ defined on the sample space Ω. Since the outcomes (in Ω) of the random experiment cannot be predicted in advance the values assumed by the function X are also unpredictable. It may be of interest to compute the probabilities of various events concerning the values assumed by function X. Specifically, it may be of interest to compute the probability that the random experiment results in a value of X in a given set $B \subseteq \mathbb{R}$. This amounts to assigning probabilities,

$$P_X(B) \stackrel{\text{def}}{=} P(\{\omega \in \Omega: X(\omega) \in B\}), \ \ B \subseteq \mathbb{R},$$

to various subsets of $\mathbb{R}$. Note that, for $B \subseteq \mathbb{R}$, $P_X(B) = P(\{\omega \in \Omega:\ X(\omega) \in B\})$ is properly defined only if $\{\omega \in \Omega:\ X(\omega) \in B\} \in \mathcal{F}$ .This puts restrictions on kind of functions X and/or kind of sets $B \subseteq \mathbb{R}$ we should be considering. An approach to deal with this issue is to appropriately choose an event space (a sigma-field) $\mathcal{B}$ of subsets of $\mathbb{R}$ and then put restriction(s) on the function X so that $P_X(B) = P(\{\omega \in \Omega : X(\omega) \in B\})$ is properly defined for each $B \in \mathcal{B}$, i.e., $\{\omega \in \Omega:\ X(\omega) \in B\} \in \mathcal{F},\ \forall B \in \mathcal{B}$.

Let $\mathcal{P}(\mathbb{R})$ and $\mathcal{P}(\Omega)$ denote the power sets of $\mathbb{R}$ and Ω, respectively. Define $X^{-1}: \mathcal{P}(\mathbb{R}) \to \mathcal{P}(\Omega)$ by

$$X^{-1}(B) = \{\omega \in \Omega:\ X(\omega) \in B\},\ B \in \mathcal{P}(\mathbb{R}).$$

The following proposition, which follows directly from the definition of X^{-1}, will be useful for further discussion (see Problem 2).

Lemma 1.1

Let $A, B \in \mathcal{P}(\mathbb{R})$ and let $A_\alpha \in \mathcal{P}(\mathbb{R}),\ \alpha \in \Lambda$, where $\Lambda \subseteq \mathbb{R}$ is an arbitrary index set. Then

(i) $X^{-1}(A - B) = X^{-1}(A) - X^{-1}(B)$. In particular $X^{-1}(B^C) = \left(X^{-1}(B)\right)^c$;

(ii) $X^{-1}(\bigcup_{\alpha\in\Lambda} A_\alpha) = \bigcup_{\alpha\in\Lambda} X^{-1}(A_\alpha)$ and $X^{-1}(\bigcap_{\alpha\in\Lambda} A_\alpha) = \bigcap_{\alpha\in\Lambda} X^{-1}(A_\alpha)$;

(iii) $A \cap B = \phi \Rightarrow X^{-1}(A) \cap X^{-1}(B) = \phi$. ■

Let $\mathcal{J}$ denote the class of all open intervals in $\mathbb{R}$, i.e., $\mathcal{J} = \{(a, b): -\infty \le a < b \le \infty\}$. In the real line $\mathbb{R}$ an appropriate event space is the Borel sigma-field $\mathcal{B}_1 = \sigma(\mathcal{J})$, the smallest sigma-field containing $\mathcal{J}$. Now, for $P_X(B) = P(\{\omega \in \mathbb{R}: X(\omega) \in B\})$ to be properly defined for every Borel set $B \in \mathcal{B}_1$, we must have

$$X^{-1}(B) = \{\omega \in \Omega: X(\omega) \in B\} \in \mathcal{F},\ \forall B \in \mathcal{B}_1.$$

This leads to the introduction of the following definition.

Definition 1.1

Let $(\Omega, \mathcal{F}, P)$ be a probability space and let $X: \Omega \to \mathbb{R}$ be a given function. We say that X is a *random variable* (r.v.) if $X^{-1}(B) \in \mathcal{F}, \forall B \in \mathcal{B}_1$. ■

Note that if $\mathcal{F} = \mathcal{P}(\Omega)$ then any function $X: \Omega \to \mathbb{R}$ is a random variable. The following theorem provides an easy to verify condition for checking whether or not a given function $X: \Omega \to \mathbb{R}$ is a random variable.

Theorem 1.1

Let $(\Omega, \mathcal{F}, P)$ be a probability space and let $X: \Omega \to \mathbb{R}$ be a given function. Then X is a random variable if, and only if, $X^{-1}((-\infty, a]) = \{\omega \in \Omega: X(\omega) \le a\} \in \mathcal{F},\ \forall a \in \mathbb{R}$.

Proof. First suppose that X is a random variable. Then $X^{-1}(B) \in \mathcal{F}, \forall B \in \mathcal{B}_1$ and, in particular

$$X^{-1}((c,d)) \in \mathcal{F}, \text{ whenever } -\infty \le c < d \le \infty \text{ (since } \mathcal{J} \subseteq \mathcal{B}_1).$$

Fix $a \in \mathbb{R}$.Then

$$(-\infty, a) = \cup_{n=1}^{\infty}\left(-n,\ a-\frac{1}{n}\right) \text{ and } \{a\} = \cap_{n=1}^{\infty}\left(a-\frac{1}{n}, a+\frac{1}{n}\right).$$

Therefore

$$\begin{aligned}(-\infty, a] &= (-\infty, a) \cup \{a\}\\ &= \left(\bigcup_{n=1}^{\infty}\left(-n,\ a-\frac{1}{n}\right)\right) \cup \left(\bigcap_{n=1}^{\infty}\left(a-\frac{1}{n}, a+\frac{1}{n}\right)\right).\end{aligned}$$

Now using Lemma 1.1 (ii), it follows that

$$X^{-1}((-\infty,a]) = \underbrace{\underbrace{\left(\bigcup_{n=1}^{\infty}\underbrace{X^{-1}\left(-n, a-\frac{1}{n}\right)}_{\in\mathcal{F},\forall n\ge 1}\right)}_{\in\mathcal{F}} \cup \underbrace{\left(\bigcap_{n=1}^{\infty}\underbrace{X^{-1}\left(a-\frac{1}{n}, a+\frac{1}{n}\right)}_{\in\mathcal{F},\forall n\ge 1}\right)}_{\in\mathcal{F}}}_{\in\mathcal{F}}$$

i.e., $X^{-1}([-\infty, a]) \in \mathcal{F}$.

Conversely suppose that$X^{-1}\big((-\infty, a]\big) \in \mathcal{F},\ \forall a \in \mathbb{R}$. Then, for $-\infty \le c < d \le \infty$,

$$(-\infty, d) = \cup_{n=1}^{\infty}\left(-\infty,\ d-\frac{1}{n}\right),$$

and

$$\begin{aligned}X^{-1}\big((c,d)\big) &= X^{-1}((-\infty,d)) - ((-\infty,c])\\ &= X^{-1}((-\infty,d)) - X^{-1}((-\infty,c]) \quad \text{(using Lemma 1.1 (i))}\\ &= X^{-1}\left(\bigcup_{n=1}^{\infty}(-\infty, d-\frac{1}{n}]\right) - X^{-1}((-\infty,c])\\ &= \underbrace{\underbrace{\bigcup_{n=1}^{\infty}\underbrace{X^{-1}((-\infty, d-\frac{1}{n}])}_{\in\mathcal{F},\forall n\ge 1}}_{\in\mathcal{F}} - \underbrace{X^{-1}((-\infty,c])}_{\in\mathcal{F}}}_{\in\mathcal{F}}\end{aligned}$$

$$\Rightarrow X^{-1}(I) \in \mathcal{F}, \;\; \forall I \in \mathcal{J}. \qquad (1.1)$$

Define,

$$\mathcal{D} = \{A \subseteq \mathbb{R}: \; X^{-1}(A) \in \mathcal{F}\}.$$

Using Lemma 1.1 it is easy to verify that $\mathcal{D}$ is a sigma-field of subsets of $\mathbb{R}$. Thus $\mathcal{D} = \sigma(\mathcal{D})$. Using (1.1) we have $\mathcal{J} \subseteq \mathcal{D} = \sigma(\mathcal{D})$, i.e., $\mathcal{J} \subseteq \sigma(\mathcal{D})$. This implies that $\sigma(\mathcal{J}) \subseteq \sigma(\mathcal{D}) = \mathcal{D}$, i.e., $\mathcal{B}_1 \subseteq \mathcal{D}$. Consequently $X^{-1}(B) \in \mathcal{F}, \; \forall B \in \mathcal{B}_1$, i.e., X is a random variable. ■

The following theorem follows on using the arguments similar to the ones used in proving Theorem 1.1.

Theorem 1.2

Let $(\Omega, \mathcal{F}, P)$ be a probability space and let $X: \Omega \to \mathbb{R}$ be a given function. Then X is a random variable if, an only if, one of the following equivalent conditions is satisfied.

(i) $X^{-1}((-\infty, a)) \in \mathcal{F}, \quad \forall a \in \mathbb{R}$;
(ii) $X^{-1}((a, \infty)) \in \mathcal{F}, \quad \forall a \in \mathbb{R}$;
(iii) $X^{-1}([a, \infty)) \in \mathcal{F}, \quad \forall a \in \mathbb{R}$;
(iv) $X^{-1}((a, b]) \in \mathcal{F}$, whenever $-\infty \leq a < b < \infty$;
(v) $X^{-1}([a, b)) \in \mathcal{F}$, whenever $-\infty < a < b \leq \infty$;
(vi) $X^{-1}((a, b)) \in \mathcal{F}$, whenever $-\infty \leq a < b \leq \infty$. ■

2.2 INDUCED PROBABILITY MEASURE

Let $(\Omega, \mathcal{F}, P)$ be a probability space and let $X: \Omega \to \mathbb{R}$ be a random variable. Define the set function $P_X: \mathcal{B}_1 \to \mathbb{R}$, by

$$P_X(B) = P\big(X^{-1}(B)\big) = P(\{\omega \in \mathbb{R}: \;\; X(\omega) \in B\}), \;\; B \in \mathcal{B}_1,$$

where $\mathcal{B}_1$ denotes the Borel sigma-field.

Since X is a r.v., $X^{-1}(B) \in \mathcal{F}, \; \forall B \in \mathcal{B}_1$ and, therefore, P_X is well defined.

Theorem 2.1

$(\mathbb{R}, \mathcal{B}_1, P_X)$ is a probability space.

Proof. Clearly,

$$P_X(B) = P\big(X^{-1}(B)\big) \geq 0, \qquad \forall B \in \mathcal{B}_1.$$

Let $B_1, B_2, \cdots$ be a countable collection of mutually exclusive events in $\mathcal{B}_1$. Then $X^{-1}(B_1), X^{-1}(B_2), ...$ is a countable collection of mutually exclusive events in $\mathcal{F}$ (Lemma 1.1 (iii)). Therefore

$$P_X\left(\bigcup_{i=1}^{\infty} B_i\right) = P\left(X^{-1}\left(\bigcup_{i=1}^{\infty} B_i\right)\right)$$

$$= P\left(\bigcup_{i=1}^{\infty} X^{-1}(B_i)\right) \qquad \text{(using Lemma 1.1 (ii))}$$

$$= \sum_{i=1}^{\infty} P\left(X^{-1}(B_i)\right)$$

$$= \sum_{i=1}^{\infty} P_X(B_i),$$

i.e., P_X is countable additive.

We also have

$$P_X(\mathbb{R}) = P\left(X^{-1}(\mathbb{R})\right) = P(\Omega) = 1.$$

It follows that P_X is a probability measure on $\mathcal{B}_1$, i.e.,$(\mathbb{R}, \mathcal{B}_1, P_X)$ is a probability space. ■

Definition 2.1

Let $(\Omega, \mathcal{F}, P)$ be a probability space and let $X: \Omega \to \mathbb{R}$ be a r.v.. Let $P_X: \mathcal{B}_1 \to \mathbb{R}$ be defined by $P_X(B) = P\left(X^{-1}(B)\right), B \in \mathcal{B}_1$. The probability space $(\mathbb{R}, \mathcal{B}_1, P_X)$ is called *the probability space induced by* X and P_X is called the probability *measure induced by* X. ■

Our primary interest now is in the induced probability space $(\mathbb{R}, \mathcal{B}_1, P_X)$ rather than the original probability space $(\Omega, \mathcal{F}, P)$.

Example 2.1

(i) Suppose that a fair coin is independently flipped thrice. With usual interpretations of the outcomes HHH, HHT, ..., the sample space is

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

Since Ω is finite we shall take $\mathcal{F} = \mathcal{P}(\Omega)$. The relevant probability measure $P: \mathcal{F} \to \mathbb{R}$ is given by

$$P(A) = \frac{|A|}{8}, \quad A \in \mathcal{F},$$

where $|A|$ denotes the number of elements in A. Suppose that we are primarily interested in the number of times a head is observed in three flips, i.e., suppose that our primary interest is on the function $X: \Omega \to \mathbb{R}$ defined by

$$X(\omega) = \begin{cases} 0, & \text{if } \omega = \mathrm{TTT} \\ 1, & \text{if } \omega \in \{\mathrm{HTT}, \mathrm{THT}, \mathrm{TTH}\} \\ 2, & \text{if } \omega \in \{\mathrm{HHT}, \mathrm{HTH}, \mathrm{THH}\} \\ 3, & \text{if } \omega = \mathrm{HHH} \end{cases}.$$

Since $\mathcal{F} = \mathcal{P}(\Omega)$, any function $Y: \Omega \to \mathbb{R}$ is a random variable. In particular the function $X: \Omega \to \mathbb{R}$ defined above is a random variable. The probability space induced by r.v. X is $(\mathbb{R}, \mathcal{B}_1, P_X)$, where $P_X(\{0\}) = P_X(\{3\}) = \frac{1}{8}, P_X(\{1\}) = P_X(\{2\}) = \frac{3}{8}$, and

$$P_X(\mathrm{B}) = \sum_{i \in \{0,1,2,3\} \cap B} P_X(\{i\}), \quad B \in \mathcal{B}_1.$$

(ii) Consider the probability space $(\mathbb{R}, \mathcal{B}_1, P)$, where

$$P(A) = \int_0^\infty e^{-t} \mathrm{I}_{\mathrm{A}}(t)\, dt$$

$$= \int_{-\infty}^\infty e^{-t} I_{A \cap [0,\infty)}(t)\, dt,$$

and, for $B \subseteq \mathbb{R}$, $I_B(\cdot)$ denotes the indicator function of B $(\text{i.e.}, I_B(t) = 1, \text{if } t \in B, = 0, \text{if } t \notin B)$. It is easy to verify that P is a probability measure on $\mathcal{B}_1$.

Define $X: \mathbb{R} \to \mathbb{R}$ by

$$X(\omega) = \begin{cases} \sqrt{\omega}, & \text{if } \omega > 0 \\ 0, & \text{if } \omega \leq 0 \end{cases}.$$

We have

$$X^{-1}((-\infty, \mathrm{a}]) = \begin{cases} \phi, & \text{if } a < 0 \\ (-\infty, \mathrm{a}^2], & \text{if } a \geq 0 \end{cases}$$

$$\in \mathcal{B}_1, \qquad \forall \mathrm{a} \in \mathbb{R}.$$

Thus X is a random variable. The probability space induced by X is $(\mathbb{R}, \mathcal{B}_1, P_X)$, where, for $B \in \mathcal{B}_1$

$$\begin{aligned}
P_X(B) &= P(\{\omega \in \mathbb{R}:\ \ X(\omega) \in B\}) \\
&= P\left(\left\{\omega \in \mathbb{R}:\ \omega > 0, \sqrt{\omega} \in B\right\}\right) + P(\{\omega \in \mathbb{R}: \omega \leq 0,\ \ 0 \in B\}) \\
&= \int_0^\infty e^{-t} I_B\left(\sqrt{t}\right) dt + 0 \\
&= 2\int_0^\infty z e^{-z^2} I_B(z) dz. \blacksquare
\end{aligned}$$