1.4 Random Variable

Motivation example In an opinion poll, we might decide to ask 50 people whether they agree or disagree with a certain issue. If we record a “1” for agree and “0” for disagree, the sample space for this experiment has $2^{50}$ elements. If we define a variable $X =$ number of 1s recorded out of 50, we have captured the essence of the problem. Note that the sample space of $X$ is the set of integers $\{1, 2, \ldots, 50\}$ and is much easier to deal with than the original sample space.

In defining the quantity $X$, we have defined a mapping (a function) from the original sample space to a new sample space, usually a set of real numbers. In general, we have the following definition.

**Definition of Random Variable** A random variable is a function from a sample space $S$ into the real numbers.

**Example 1.4.2 (Random variables)**

In some experiments random variables are implicitly used; some examples are these.

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Random variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Toss two dice</td>
<td>$X =$ sum of the numbers</td>
</tr>
<tr>
<td>Toss a coin 25 times</td>
<td>$X =$ number of heads in 25 tosses</td>
</tr>
<tr>
<td>Apply different amounts of fertilizer to corn plants</td>
<td>$X =$ yield/acre</td>
</tr>
</tbody>
</table>

Suppose we have a sample space

$$S = \{s_1, \ldots, s_n\}$$

with a probability function $P$ and we define a random variable $X$ with range $\mathcal{X} = \{x_1, \ldots, x_m\}$. We can define a probability function $P_X$ on $\mathcal{X}$ in the following way. We will observe $X = x_i$ if and only if the outcome of the random experiment is an $s_j \in S$ such that $X(s_j) = x_i$. Thus,

$$P_X(X = x_i) = P(\{s_j \in S : X(s_j) = x_i\}). \quad (1)$$
Note \( P_X \) is an induced probability function on \( \mathcal{X} \), defined in terms of the original function \( P \). Later, we will simply write \( P_X(X = x_i) = P(X = x_i) \).

**Fact** The induced probability function defined in (1) defines a legitimate probability function in that it satisfies the Kolmogorov Axioms.

**Proof:** \( C \) is finite. Therefore \( \mathcal{B} \) is the set of all subsets of \( \mathcal{X} \). We must verify each of the three properties of the axioms.

1. If \( A \in \mathcal{B} \) then \( P_X(A) = P(\cup_{x_i \in A} \{s_j \in S : X(s_j) = x_i\}) \geq 0 \) since \( P \) is a probability function.

2. \( P_X(\mathcal{X}) = P(\cup_{x_i} \{s_j \in S : X(s_j) = x_i\}) = P(S) = 1. \)

3. If \( A_1, A_2, \ldots \in \mathcal{B} \) and pairwise disjoint then

\[
P_X(\bigcup_{k=1}^{\infty} A_k) = P(\bigcup_{k=1}^{\infty} \{s_j \in S : X(s_j) = x_i\}) = \sum_{k=1}^{\infty} P_X(A_k),
\]

where the second inequality follows from the fact \( P \) is a probability function. \( \square \)

A note on notation: Random variables will always be denoted with uppercase letters and the realized values of the variable will be denoted by the corresponding lowercase letters. Thus, the random variable \( X \) can take the value \( x \).

**Example 1.4.3 (Three coin tosses-II)** Consider again the experiment of tossing a fair coin three times independently. Define the random variable \( X \) to be the number of heads obtained in the three tosses. A complete enumeration of the value of \( X \) for each point in the sample space is

<table>
<thead>
<tr>
<th>s</th>
<th>HHH</th>
<th>HHT</th>
<th>HTH</th>
<th>THH</th>
<th>TTH</th>
<th>THT</th>
<th>HTT</th>
<th>TTT</th>
</tr>
</thead>
<tbody>
<tr>
<td>X(s)</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

The range for the random variable \( X \) is \( \mathcal{X} = \{0, 1, 2, 3\} \). Assuming that all eight points in \( S \) have probability \( \frac{1}{8} \), by simply counting in the above display we see that the induced probability function on \( \mathcal{X} \) is given by
The previous illustrations had both a finite $S$ and finite $\mathcal{X}$, and the definition of $P_X$ was straightforward. Such is also the case if $\mathcal{X}$ is countable. If $\mathcal{X}$ is uncountable, we define the induced probability function, $P_X$, in a manner similar to (1). For any set $A \subseteq \mathcal{X}$,

$$P_X(X \in A) = P\left\{s \in S : X(s) \in A\right\}.$$  \hfill (2)

This does define a legitimate probability function for which the Kolmogorov Axioms can be verified.

### Distribution Functions

**Definition of Distribution** The cumulative distribution function (cdf) of a random variable $X$, denoted by $F_X(x)$, is defined by

$$F_X(x) = P_X(X \leq x), \quad \text{for all } x.$$

**Example 1.5.2 (Tossing three coins)** Consider the experiment of tossing three fair coins, and let $X =$number of heads observed. The cdf of $X$ is

$$F_X(x) = \begin{cases} 0 & \text{if } -\infty < x < 0 \\ \frac{1}{8} & \text{if } 0 \leq x < 1 \\ \frac{1}{2} & \text{if } 1 \leq x < 2 \\ \frac{7}{8} & \text{if } 2 \leq x < 3 \\ 1 & \text{if } 3 \leq x < \infty. \end{cases}$$
Remark:

1. $F_X$ is defined for all values of $x$, not just those in $\mathcal{X} = \{0, 1, 2, 3\}$. Thus, for example,
   \[
   F_X(2.5) = P(X \leq 2.5) = P(X = 0, 1, 2) = \frac{7}{8}.
   \]

2. $F_X$ has jumps at the values of $x_i \in \mathcal{X}$ and the size of the jump at $x_i$ is equal to $P(X = x_i)$.

3. $F_X = 0$ for $x < 0$ since $X$ cannot be negative, and $F_X(x) = 1$ for $x \geq 3$ since $x$ is certain to be less than or equal to such a value.

$F_X$ is right-continuous, namely, the function is continuous when a point is approached from the right. The property of right-continuity is a consequence of the definition of the cdf. In contrast, if we had defined $F_X(x) = P_X(X < x)$, $F_X$ would then be left-continuous.

Theorem 1.5.3

The function $F_X(x)$ is a cdf if and only of the following three conditions hold:

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

b. $F(x)$ is a nondecreasing function of $x$.

c. $F(x)$ is right-continuous; that is, for every number $x_0$, $\lim_{x \downarrow x_0} F(x) = F(x_0)$.

Example 1.5.4 (Tossing for a head) Suppose we do an experiment that consists of tossing a coin until a head appears. Let $p =$ probability of a head on any given toss, and define $X =$ number of tosses required to get a head. Then, for any $x = 1, 2, \ldots,$

\[
P(X = x) = (1 - p)^{x-1}p.
\]

The cdf is

\[
F_X(x) = P(X \leq x) = \sum_{i=1}^{x} P(X = i) = \sum_{i=1}^{x} (1 - p)^{i-1}p = 1 - (1 - p)^x.
\]

It is easy to show that if $0 < p < 1$, then $F_X(x)$ satisfies the conditions of Theorem 1.5.3. First,

\[
\lim_{x \to -\infty} F_X(x) = 0
\]
since \( F_X(x) = 0 \) for all \( x < 0 \), and
\[
\lim_{x \to -\infty} F_X(x) = \lim_{x \to \infty} (1 - (1 - p)^x) = 1,
\]
where \( x \) goes through only integer values when this limit is taken. To verify property (b), we simply note that the sum contains more positive terms as \( x \) increases. Finally, to verify (c), note that, for any \( x \), \( F_X(x + \epsilon) = F_X(x) \) if \( \epsilon > 0 \) is sufficiently small. Hence,
\[
\lim_{\epsilon \downarrow 0} F_X(x + \epsilon) = F_X(x),
\]
so \( F_X(x) \) is right-continuous.

**Example 1.5.5 (Continuous cdf)**

An example of a continuous cdf (logistic distribution) is the function
\[
F_X(x) = \frac{1}{1 + e^{-x}}.
\]
It is easy to verify that
\[
\lim_{x \to -\infty} F_X(x) = 0 \quad \text{and} \quad \lim_{x \to \infty} F_X(x) = 1.
\]
Differentiating \( F_X(x) \) gives
\[
\frac{d}{dx} F_X(x) = \frac{e^{-x}}{(1 + e^{-x})^2} > 0,
\]
showing that \( F_X(x) \) is increasing. \( F_X \) is not only right-continuous, but also continuous.

**Definition of Continuous Random Variable** A random variable \( X \) is continuous if \( F_X(x) \) is a continuous function of \( x \). A random variable \( X \) is discrete if \( F_X(x) \) is a step function of \( x \).

We close this section with a theorem formally stating that \( F_X \) completely determines the probability distribution of a random variable \( X \). This is true if \( P(X \in A) \) is defined only for events \( A \) in \( \mathbb{B}^1 \), the smallest sigma algebra containing all the intervals of real numbers of the form \((a, b), [a, b), (a, b], \) and \([a, b] \). If probabilities are defined for a larger class of events, it is possible for two random variables to have the same distribution function but not the same probability for every event (see Chung 1974, page 27).

**Definition of Identical Random Variables** The random variables \( X \) and \( Y \) are identically distributed if, for every set \( A \in \mathbb{B}^1 \), \( P(X \in A) = P(Y \in A) \).
Note that two random variables that are identically distributed are not necessarily equal. That is, the above definition does not say that \( X = Y \).

**Example 1.5.9 (identically distributed random variables)** Consider the experiment of tossing a fair coin three times. Define the random variables \( X \) and \( Y \) by

\[
X = \text{number of heads observed} \quad \text{and} \quad Y = \text{number of tails observed}.
\]

For each \( k = 0, 1, 2, 3 \), we have \( P(X = k) = P(Y = k) \). So \( X \) and \( Y \) are identically distributed. However, for no sample point do we have \( X(s) = Y(s) \).

**Theorem 1.5.10** The following two statements are equivalent:

a. The random variables \( X \) and \( Y \) are identically distributed.

b. \( F_X(x) = F_Y(x) \) for every \( x \).

**Proof:** To show equivalence we must show that each statement implies the other. We first show that \( (a) \Rightarrow (b) \).

Because \( X \) and \( Y \) are identically distributed, for any set \( A \in \mathbb{B} \), \( P(X \in A) = P(Y \in A) \). In particular, for every \( x \), the set \( (-\infty, x] \) is in \( \mathbb{B} \), and

\[
F_X(x) = P(X \in (-\infty, x]) = P(Y \in (-\infty, x]) = F_Y(x).
\]

The above argument showed that if the \( X \) and \( Y \) probabilities agreed in all sets, then agreed on intervals. To show \( (b) \Rightarrow (a) \), we must prove if the \( X \) and \( Y \) probabilities agree on all intervals, then they agree on all sets. For more details see Chung (1974, section 2.2). \( \square \)