8.1 DEFINITIONS

We have been using x-notation in a more general sense here than defined in Chapter 2: The sums in (8.1) and (8.4) occur over all elements \( w \) of an arbitrary set, not over integers only. However, this new development is not really alarming; we can agree to use special notation under a \( \sum \) whenever nonintegers are intended, so there will be no confusion with our ordinary conventions. The other definitions in Chapter 2 are still valid; in particular, the definition of infinite sums in that chapter gives the appropriate interpretation to our sums when the set \( \Omega \) is infinite. Each probability is nonnegative, and the sum of all probabilities is bounded, so the probability of event A in (8.4) is well defined for all subsets \( A \subseteq \Omega \).

A random variable is a function defined on the elementary events \( w \) of a probability space. For example, if \( \Omega = \{1, 2, 3, \ldots, 6\} \), we can define \( S(w) \) to be the sum of the spots on the dice roll \( w \), so that \( S(\boxed{1 \ m}) = 6 + 3 = 9 \). The probability that the spots total seven is the probability of the event \( S(w) = 7 \), namely

\[
\Pr(\boxed{1 \ m}) + \Pr(\boxed{2 \ m}) + \Pr(\boxed{3 \ m}) + \Pr(\boxed{4 \ m}) + \Pr(\boxed{5 \ m}) + \Pr(\boxed{6 \ m})
\]

With fair dice (\( \Pr = \Pr_{\text{fair}} \)), this happens with probability \( \frac{1}{6} \); with loaded dice (\( \Pr = \Pr_{\text{loaded}} \)), it happens with probability \( \frac{1}{16} + \frac{1}{64} + \frac{1}{64} + \frac{1}{64} + \frac{1}{64} + \frac{1}{16} = \frac{3}{16} \), the same as we observed for doubles.

It’s customary to drop the ‘\( w \)’ when we talk about random variables, because there’s usually only one probability space involved when we’re working on any particular problem. Thus we say simply ‘\( S = 7 \)’ for the event that a 7 was rolled, and ‘\( S = 4 \)’ for the event \( \{ \boxed{1 \ m}, \boxed{2 \ m}, \boxed{3 \ m} \} \).

A random variable can be characterized by the probability distribution of its values. Thus, for example, \( S \) takes on eleven possible values \( \{2, 3, \ldots, 12\} \), and we can tabulate the probability that \( S = s \) for each \( s \) in this set:

<table>
<thead>
<tr>
<th>( s )</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \Pr_{\text{fair}}[S = s] )</td>
<td>( \frac{1}{36} )</td>
<td>( \frac{2}{36} )</td>
<td>( \frac{3}{36} )</td>
<td>( \frac{4}{36} )</td>
<td>( \frac{5}{36} )</td>
<td>( \frac{6}{36} )</td>
<td>( \frac{7}{36} )</td>
<td>( \frac{8}{36} )</td>
<td>( \frac{9}{36} )</td>
<td>( \frac{10}{36} )</td>
<td>( \frac{11}{36} )</td>
</tr>
<tr>
<td>( \Pr_{\text{loaded}}[S = s] )</td>
<td>( \frac{4}{64} )</td>
<td>( \frac{4}{64} )</td>
<td>( \frac{5}{64} )</td>
<td>( \frac{6}{64} )</td>
<td>( \frac{7}{64} )</td>
<td>( \frac{8}{64} )</td>
<td>( \frac{9}{64} )</td>
<td>( \frac{10}{64} )</td>
<td>( \frac{11}{64} )</td>
<td>( \frac{12}{64} )</td>
<td>( \frac{13}{64} )</td>
</tr>
</tbody>
</table>

If we’re working on a problem that involves only the random variable \( S \) and no other properties of dice, we can compute the answer from these probabilities alone, without regard to the details of the set \( \Omega = \{1, 2, 3, \ldots, 6\} \). In fact, we could define the probability space to be the smaller set \( \Omega = \{2, 3, \ldots, 12\} \), with whatever probability distribution \( \Pr(s) \) is desired. Then ‘\( S = 4 \)’ would be an elementary event. Thus we can often ignore the underlying probability space \( \Omega \) and work directly with random variables and their distributions.

If two random variables \( X \) and \( Y \) are defined over the same probability space \( \Omega \), we can characterize their behavior without knowing everything