8.2 MEAN AND VARIANCE

The next most important property of a random variable, after we know its expected value, is its variance, defined as the mean square deviation from the mean:

\[ \text{Var}(X) = \text{E}[(X - \text{E}(X))^2] \]

If we denote \( \text{E}(X) \) by \( \mu \), the variance \( \text{Var}(X) \) is the expected value of \((X - \mu)^2\). This measures the “spread” of \( X \)'s distribution.

As a simple example of variance computation, let's suppose we have just been offered an offer we can't refuse: Someone has given us two gift certificates for a certain lottery. The lottery organizers sell 100 tickets for each weekly drawing. One of these tickets is selected by a uniformly random process—that is, each ticket is equally likely to be chosen—and the lucky ticket holder wins a hundred million dollars. The other 99 ticket holders win nothing.

We can use our gift in two ways: Either we buy two tickets in the same lottery, or we buy one ticket in each of two lotteries. Which is a better strategy? Let's try to analyze this by letting \( X_1 \) and \( X_2 \) be random variables that represent the amount we win on our first and second ticket. The expected value of \( X_1 \), in millions, is

\[ \text{E}(X_1) = \frac{99}{100} \cdot 0 + \frac{1}{100} \cdot 100 = 1 \]

and the same holds for \( X_2 \). Expected values are additive, so our average total winnings will be

\[ \text{E}(X_1 + X_2) = \text{E}(X_1) + \text{E}(X_2) = 2 \text{ million dollars} \]

regardless of which strategy we adopt.

Still, the two strategies seem different. Let's look beyond expected values and study the exact probability distribution of \( X_1 + X_2 \):

<table>
<thead>
<tr>
<th>Winnings (millions)</th>
<th>( X_1 )</th>
<th>( X_2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( X_1 )</td>
<td>( X_2 )</td>
<td>( X_1 + X_2 )</td>
</tr>
<tr>
<td>same drawing</td>
<td>.9800</td>
<td>.0200</td>
</tr>
<tr>
<td>different drawings</td>
<td>.9801</td>
<td>.0198</td>
</tr>
</tbody>
</table>

If we buy two tickets in the same lottery we have a 98% chance of winning nothing and a 2% chance of winning $100 million. If we buy them in different lotteries we have a 98.01% chance of winning nothing, so this is slightly more likely than before; and we have a 0.01% chance of winning $200 million, also slightly more likely than before; and our chances of winning $100 million are now 1.98%. So the distribution of \( X_1 + X_2 \) in this second situation is slightly