Lecture 6

A. Random Variables (5.1, 5.2)

- **Examples.**

  **Example 1: Tossing a fair coin 3 times.**
  1. Let X stand for the number of Heads in the 3 tosses.
  2. Let Y stand for the number of Tails in the 3 tosses.
  3. Let Z stand for the difference in the number of Heads and the number of Tails in the 3 tosses.

  **Example 2: Counting Cars.**
  At an intersection, an observer counts the number of cars passing by until a new Ford is spotted. Let X be that number.

  **Example 3: Waiting for Bus.**
  Suppose you are waiting for a bus at a bus stop at 9:00 am, and the bus can arrive at the bus stop at any time between 9:am and 10:am. Let X be the waiting time.

  **Example 4: Measuring Heights.**
  Randomly select an MSU student, measure his/her height. Let Y be the height of the student.

- **Definition (Random Variable):** A random variable associates a numerical value with each outcome of an experiment.
  - It is a variable because it will typically have different values
  - It is random because the value it takes depends on the outcome of the experiment.

- **Two Types of Random Variables.**
  1. **Discrete Random Variable:** A random variable that can take a finite number of values or infinitely many values that can be arranged in a
sequence is called a discrete random variable.

Example: Random variables in Example 1, 2.

2. **Continuous Random Variable:** A random variable that can take any numerical value within an interval is called a continuous random variable.

Example: Random variables in Example 3, 4.

**B. Probability Distribution of Discrete Random Variables (5.3)**

- Example 1: Let X stand for the number of Heads in 3 tosses. The probability distribution of X is

<table>
<thead>
<tr>
<th>X</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>P(X=x)</td>
<td>1/8</td>
<td>3/8</td>
<td>3/8</td>
<td>1/8</td>
</tr>
</tbody>
</table>

- Example 2: You roll a die. If it comes up a 6, you win $100. If not, you get to roll it again. If you get a 6 the second time, you win $50. If not, you lose $100. Let X be the amount of money you win. The probability distribution of X is

<table>
<thead>
<tr>
<th>X</th>
<th>100</th>
<th>50</th>
<th>-100</th>
</tr>
</thead>
<tbody>
<tr>
<td>P(X=x)</td>
<td>1/6</td>
<td>(5/6)*(1/6)</td>
<td>(5/6)*(5/6)</td>
</tr>
</tbody>
</table>

- Suppose X is a discrete random variable which take values $x_1, x_2, \ldots, x_k$. The probability distribution of a discrete random variable X is described as the function

  \[ f(x_i) = P(X=x_i) \]

or a table

<table>
<thead>
<tr>
<th>X</th>
<th>$x_1$</th>
<th>$x_2$</th>
<th>\ldots</th>
<th>$x_k$</th>
</tr>
</thead>
<tbody>
<tr>
<td>P(X=x)</td>
<td>$f(x_1)$</td>
<td>$f(x_2)$</td>
<td>\ldots</td>
<td>$f(x_k)$</td>
</tr>
</tbody>
</table>

which gives the probability for each value and satisfies

1. $f(x_i) \geq 0$ for each value $x_i$ of X.
2. $\sum_{i=1}^{n} f(x_i) = 1.$

**C. Expectation**
• **Why this concept?**

Suppose a die is tossed 20 times and the following data are obtained.

\[
4 \ 3 \ 4 \ 2 \ 5 \ 1 \ 6 \ 6 \ 5 \ 2 \\
2 \ 6 \ 5 \ 4 \ 6 \ 2 \ 1 \ 6 \ 2 \ 4
\]

The mean of these observations, called the sample mean, is calculated as

\[
\overline{X} = \frac{4+3+4+2+5+1+6+6+5+2+2+6+5+4+6+2+1+6+2+4}{20} = \frac{76}{20} = 3.8
\]

Alternatively, we can calculate the mean as

\[
\overline{X} = \frac{1(2)+2(5)+3(1)+4(4)+5(3)+6(5)}{20}
\]

\[
= 1\left(\frac{2}{20}\right) + 2\left(\frac{5}{20}\right) + 3\left(\frac{1}{20}\right) + 4\left(\frac{4}{20}\right) + 5\left(\frac{3}{20}\right) + 6\left(\frac{5}{20}\right) = 3.8
\]

The second calculation illustrates the formula:

\[
\text{Sample Mean } \overline{X} = \sum (\text{value} \times \text{relative frequency})
\]

Suppose a large number of tosses are made. Then the relative frequency approaches the probability. The mean can be calculated as

\[
1\left(\frac{1}{6}\right) + 2\left(\frac{1}{6}\right) + 3\left(\frac{1}{6}\right) + 4\left(\frac{1}{6}\right) + 5\left(\frac{1}{6}\right) + 6\left(\frac{1}{6}\right) = 3.6
\]

• **Expectation (Mean):** The expectation or mean of a discrete random variable is defined as

\[
E(X) = \mu = \sum (\text{value} \times \text{probability}) = \sum x_i f(x_i)
\]

Here the sum extends over all the distinct values \(x_i\) of \(X\).

• **Example**

1. A commuter must pass 5 traffic lights on her way to work, and will have to stop at each one that is red. She estimates the probability model for the number of red lights she hits, as shown below.
<table>
<thead>
<tr>
<th>X = # of red</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>P(X=x)</td>
<td>0.05</td>
<td>0.25</td>
<td>0.35</td>
<td>0.15</td>
<td>0.15</td>
<td>0.05</td>
</tr>
</tbody>
</table>

How many red lights should she expect to hit each day?

2. A guy toes to casino. Now there is a game. You draw a card from 10 cards, 7 of which are black and 3 of them are red. If the outcome is a red one, you win $10; if the outcome is a black one, you lose $10. Let X be the amount of money you’ll win. Would you play this game? Why?

D. Variance and Standard Deviations of Discrete Random Variables

The variance of a random variable tells us something about the spread of the possible values of the random variable.

1. Variance of a discrete random variable X is defined as

\[
\text{Var}(X) = \sigma^2 = \sum (x_i - \mu)^2 f(x_i)
\]

2. Standard deviation of a discrete random variable X is defined as

\[
\text{Sd}(X) = \sigma = \sqrt{\text{Var}(X)}
\]

Example: Calculate the variance and the standard deviation of the random variable X in the traffic light example.

Solution:

<table>
<thead>
<tr>
<th>X</th>
<th>f(x)=P(X=x)</th>
<th>xf(x)</th>
<th>(x- \mu)</th>
<th>(x- \mu)^2</th>
<th>(x- \mu)^2f(x)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.05</td>
<td>0</td>
<td>-2.25</td>
<td>5.0625</td>
<td>0.25313</td>
</tr>
<tr>
<td>1</td>
<td>0.25</td>
<td>0.25</td>
<td>-1.25</td>
<td>1.5625</td>
<td>0.39063</td>
</tr>
<tr>
<td>2</td>
<td>0.35</td>
<td>0.70</td>
<td>-0.25</td>
<td>0.0625</td>
<td>0.02188</td>
</tr>
<tr>
<td>3</td>
<td>0.15</td>
<td>0.45</td>
<td>0.75</td>
<td>0.5625</td>
<td>0.08438</td>
</tr>
<tr>
<td>4</td>
<td>0.15</td>
<td>0.60</td>
<td>1.75</td>
<td>3.0625</td>
<td>0.45938</td>
</tr>
<tr>
<td>5</td>
<td>0.05</td>
<td>0.25</td>
<td>2.75</td>
<td>7.5625</td>
<td>0.37813</td>
</tr>
<tr>
<td>Total</td>
<td>1</td>
<td>\mu = 2.25</td>
<td></td>
<td>\sigma^2 = 1.5875</td>
<td></td>
</tr>
</tbody>
</table>

So, Var(X) = \sigma^2 = 1.5875, Sd(X) = \sigma = \sqrt{1.5875} = 1.25996.

Remark: Alternative formula for calculating variance

\[
\text{Var}(X) = \sigma^2 = \sum x_i^2 f(x_i) - \mu^2
\]