Chapter 5. Discrete Probability Distributions

• A random variable is a variable whose values are determined by chance.

• The set of values that a random variable can take is called the range (or support) of the random variable.

• There are two major types of random variables: discrete and continuous.

• A discrete random variable has a countable support (either finite or countably infinite).

• A continuous random variable has a support which contains an interval.

• The probability distribution of a discrete random variable specifies the support and the corresponding probabilities. The probability distribution is presented in a table, by a formula, or by a graph.

• If \( x \) is a value in the support \( S \) of discrete random variable \( X \) and its corresponding probability is denoted \( P(x) \), then

\[
0 \leq P(x) \leq 1 \quad \text{and} \quad \sum_{x \in S} P(x) = 1.
\]

Example (of a discrete random variable \( X \) and its probability distribution.)

Let \( X \) denote the number of girls in a 3-child family.

<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) )</td>
<td>( \frac{1}{8} )</td>
<td>( \frac{3}{8} )</td>
<td>( \frac{3}{8} )</td>
<td>( \frac{1}{8} )</td>
</tr>
</tbody>
</table>

• The mean (or expected value) of a discrete random variable \( X \) is a weighted average of the values in the support and the weights are the corresponding probabilities. The mean is denote by the Greek letter \( \mu \) or by the expression \( E(X) \). Specifically,

\[
\mu = E(X) = \sum_{x \in S} x \cdot P(x).
\]

Example. For \( X \) in the previous example, we have

\[
\mu = 0(\frac{1}{8}) + 1(\frac{3}{8}) + 2(\frac{3}{8}) + 3(\frac{1}{8}) = \frac{12}{8} = 1.5
\]

Example. You pay $3 to play the following game. You pick a card from a deck. If the card is an ace, the “house” pays out $10; if the card is a face card the house pays out $5; for any other card, the house pays out $1. What is your expected gain \( X \)?

\[
E(X) = -2 \left( \frac{16}{52} \right) + 2 \left( \frac{12}{52} \right) + 7 \left( \frac{4}{52} \right) = - \frac{20}{52} \approx - .38
\]
The variance of a discrete random variable $X$ is the weighted average of the squared deviations (from the mean) and where the weights are the corresponding probabilities. The variance is denoted by the $\sigma^2$ or $Var(X)$. Specifically

$$
\sigma^2 = \sum_{x \in S} (x - \mu)^2 \cdot P(x) \quad \text{or, algebraically equivalently, } \sigma^2 = \sum_{x \in S} x^2 \cdot P(x) - \mu^2.
$$

- The standard deviation, $\sigma$, of $X$ is the square root of the variance. In other words

$$
\sigma = \sqrt{\sum_{x \in S} x^2 \cdot P(x) - \mu^2}
$$

**Example.** For $X$ in the previous example, since

$$
\sum_{x \in S} x^2 \cdot P(x) = 0^2\left(\frac{1}{8}\right) + 1^2\left(\frac{3}{8}\right) + 2^2\left(\frac{3}{8}\right) + 3^2\left(\frac{1}{8}\right) = \frac{24}{8} = 3,
$$

the variance is given by $\sigma^2 = 3 - (1.5)^2 = .75$, and the standard deviation is given by $\sigma = \sqrt{.75}$.

- A binomial experiment has a fixed number $n$ of independent, identical Bernoulli trials. “Bernoulli trial” here means that each trial results in one of two possible outcomes: “success” or “failure”. “Independent” means that the outcome of any subset of trials does not affect the outcome of any other disjoint subset of trials. “Identical” here means that the probability $p$ of “success” is the same for each trial.

- The total number of successes, say $X$, in a binomial experiment is a discrete random variable called a binomial random variable. We write $X \sim \text{binomial}(n, p)$. The probability distribution is given by the formula

$$
P(x) = \binom{n}{x} p^x (1 - p)^{n-x} \text{ for } x = 0, 1, 2, \ldots, n.
$$

The mean for a binomial random variable is given by

$$
\mu = np
$$

and the variance is given by

$$
\sigma^2 = np(1 - p)
$$

- The TI-83 procedure `binompdf(n, p, x)` calculates binomial probabilities.
- The TI-83 procedure `binomcdf(n, p, x)` calculates cumulative binomial probabilities.

**Example.** Suppose a student fails to study for a test and randomly guesses an answer for all 10 questions on the multiple choice test. If each question has 5 multiple choices, (i) find the probability that the student gets exactly 3 correct, and (ii) find the probability that the student gets 4 or less correct.

(i) TI-83. \[2nd\] \[DIST\] `binompdf(10, 1/5, 3) \quad \text{ENTER} \quad .201326592

(ii) TI-83. \[2nd\] \[DIST\] `binomcdf(10, 1/5, 4) \quad \text{ENTER} \quad .9672065025