Section 8.1 Distributions of Random Variables

Random Variable  A random variable is a rule that assigns a number to each outcome of a chance experiment. There are three types of random variables:

1. **Finite Discrete:** The random variable has a finite number, \( n \), of values it can take on, and the random variable can assume any countable collection of values, like \( \{0, 1/2, 1, 3/2, 2, \ldots, n\} \). For this class, discrete mostly means the random variable takes on whole number values, like \( \{0, 1, 2, \ldots, n\} \).

2. **Infinite Discrete:** The random variable has an infinite number of values it can take on. Again, in this class, infinite discrete mostly means the random variable assumes whole number values, like \( \{0, 1, 2, 3, \ldots\} \).

3. **Continuous** The random variable has an infinite number of values it can take on, and the random variable can assume any value in a continuous interval, like \( 0 \leq X \leq 1 \).

1. Consider the following.

\[ X = \text{The number of times a die is thrown until a 2 appears} \]

Give the range of values that the random variable \( X \) may assume.

\[ X = \{1, 2, 3, \ldots\} \]

Classify the random variable.

**Infinite Discrete**

2. Consider the following.

\[ X = \text{The number of hours a child watches television on a given day} \]

Give the range of values that the random variable \( X \) may assume.

\[ X = \{x \in \mathbb{R} | 0 \leq x \leq 24\} \text{ set builder} \]

Classify the random variable.

**Continuous**
3. Cards are selected one at a time without replacement from a well-shuffled deck of 52 cards until an ace is drawn. Let \( X \) denote the random variable that gives the number of cards drawn. What values may \( X \) assume?

\[ X = \{1, 2, 3, \ldots, 52, 53\} \]

4. Determine the possible values of the given random variable and indicate as your answer whether the random variable is finite discrete, infinite discrete, or continuous.

A marble is drawn at random and then replaced from a box of 7 red and 6 green marbles. Let the random variable \( X \) be the number of draws until a red marble is picked.

What are the possible values of \( X \)?

\[ X = \{1, 2, 3, \ldots, 7\} \]

Classify \( X \).

\[ \text{Infinite Discrete} \]

**Probability Distribution for a Random Variable \( X \)**

If \( X = \{x_1, x_2, \cdots, x_n\} \) is a random variable with the given set of values, then the probability distribution for the random variable is a table where the entries in the first row are all the possible values \( X \) can assume (\( x_1, x_2, \cdots, x_n \)) and the entries in the second row are all their corresponding probabilities (\( P(X = x_1), P(X = x_2), \ldots, P(X = x_n) \)).

<table>
<thead>
<tr>
<th>( x )</th>
<th>( x_1 )</th>
<th>( x_2 )</th>
<th>( \cdots )</th>
<th>( x_n )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( P(X = x) )</td>
<td>( P(x_1) )</td>
<td>( P(x_2) )</td>
<td>( \cdots )</td>
<td>( P(x_n) )</td>
</tr>
</tbody>
</table>
5. The probability distribution of the random variable $X$ is shown in the accompanying table.

<table>
<thead>
<tr>
<th>$x$</th>
<th>-10</th>
<th>-5</th>
<th>0</th>
<th>5</th>
<th>10</th>
<th>15</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P(X = x)$</td>
<td>.20</td>
<td>.10</td>
<td>.25</td>
<td>.15</td>
<td>.05</td>
<td>.15</td>
<td>.10</td>
</tr>
</tbody>
</table>

Find the following.

(a) $P(X = -10) = .20$

(b) $P(X \geq 5) = .15 + .05 + .15 + .10 = .45$

(c) $P(-5 \leq X \leq 5) = .10 + .25 + .15 = .50$

(d) $P(X \leq 20) = 1$

(e) $P(X < 5) = .20 + .10 + .25 = .55$

(f) $P(X = 2) = 0$

6. A survey was conducted by the Public Housing Authority in a certain community among 1000 families to determine the distribution of families by size. The results are given below.

<table>
<thead>
<tr>
<th>Family Size</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency of Occurrence</td>
<td>300</td>
<td>209</td>
<td>207</td>
<td>80</td>
<td>69</td>
<td>12</td>
<td>123</td>
</tr>
</tbody>
</table>

Find the probability distribution of the random variable $X$, where $X$ denotes the number of persons in a randomly chosen family. (Give answers as fractions.)
7. Two cards are drawn from a well-shuffled deck of 52 playing cards. Let $X$ denote the number of aces drawn. Find the probability distribution of the random variable $X$. (Round answer to three decimal places.)

<table>
<thead>
<tr>
<th>Family Size</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P(X = x)$</td>
<td>$\frac{300}{1000} = 0.30$</td>
<td>$\frac{100}{1000} = 0.10$</td>
<td>$\frac{20}{1000} = 0.02$</td>
<td>$\frac{10}{1000} = 0.01$</td>
<td>$\frac{5}{1000} = 0.005$</td>
<td>$\frac{2}{1000} = 0.002$</td>
<td>$\frac{1}{1000} = 0.001$</td>
</tr>
</tbody>
</table>

$P(X = 2) = \frac{4 \binom{4}{2} \binom{48}{2}}{\binom{52}{2}} = 0.005$

$P(X = 0) = \frac{\binom{48}{2}}{\binom{52}{2}} = 0.851$

$P(X = 1) = \frac{4 \binom{1}{1} \binom{48}{1}}{\binom{52}{2}} = 0.145$

8. Let $X$ denote the random variable that gives the sum of the faces that fall uppermost when two fair dice are rolled. Find $P(X = 7)$. (Round answer to two decimal places.)

$P(\text{Sum} = 7) = \frac{6}{36} = 0.17$
9. A box has 5 yellow, 7 gray, and 3 black marbles. Three marbles are drawn at the same time (i.e. without replacement) from the box. Let $X$ be the number of gray marbles drawn. Find the following. (Round answers to three decimal places.)

(a) $P(X = 2) = \frac{\binom{7}{2}\binom{15}{1}}{\binom{28}{3}} = .367$

(b) $P(X \leq 2) = 1 - P(\text{don’t win})$

$= 1 - P(3 \text{ gray})$

$= 1 - \frac{\binom{7}{3}}{\binom{15}{3}} = .923$

**Histograms** A histogram is a graphical representation of a probability distribution of a random variable $X$. The horizontal axis represents all the possible values the random variable $X$ may assume, while the vertical axis represents their corresponding probabilities.

10. An examination consisting of ten true-or-false questions was taken by a class of 100 students. The probability distribution of the random variable $X$, where $X$ denotes the number of questions answered correctly by a randomly chosen student, is represented by the accompanying histogram. The rectangle with base centered on the number 8 is missing. What should be the height of this rectangle?

\[ P(X = 8) = 1 - .05 - .10 - .25 - .20 - .05 \]

\[ = 1 - .65 = .35 \]
Section 8.2 Expected Value

Average or Mean The average or mean of the \( n \) numbers

\[ x_1, x_2, \ldots, x_n \]

is \( \bar{x} \) (read “x bar”), where

\[ \bar{x} = \frac{x_1 + x_2 + \ldots + x_n}{n} \]

Expected Value of a Random Variable \( X \) Let \( X \) denote a random variable that assumes the values \( x_1, x_2, \ldots, x_n \) with associated probabilities \( p_1, p_2, \ldots, p_n \), respectively. The expected value of \( X \), denoted by \( E(X) \), is given by

\[ E(X) = x_1 p_1 + x_2 p_2 + \ldots + x_n p_n \]

Median and Mode

The median of a group of numbers arranged in increasing or decreasing order is (a) the middle number if there is an odd number of entries or (b) the mean of the two middle numbers if there is an even number of entries.

Note: The mean associated with a probability distribution, is the number \( m \) such that \( P(X \leq m) \geq \frac{1}{2} \) and \( P(X \geq m) \geq \frac{1}{2} \).

The mode of a group of numbers is the number in the group that occurs most frequently (or the number with the highest probability).

1. In an examination given to a class of 20 students, the following test scores were obtained.

40 45 50 50 55 60 70 75 75 80
85 85 85 85 85 90 95 95 95 100

Find the mean (or average) score, the mode, and the median score.

\[
\text{Mean} = \frac{\text{Add up } \#s}{20} = \frac{1495}{20} = 74.75
\]

\[
\text{Median} = \frac{80 + 85}{2} = 82.5
\]

Even \# = average of two middle \#s

\[
\text{Mode} = 85
\]
2. The frequency distribution of the hourly wage rates (in dollars) among blue-collar workers in a certain factory is given in the following table. Find the mean (or average) wage rate, the mode, and the median wage rate of these workers. (If necessary, round answers to two decimal places.)

<table>
<thead>
<tr>
<th>Wage Rate</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>14.40</td>
<td>60</td>
</tr>
<tr>
<td>14.50</td>
<td>90</td>
</tr>
<tr>
<td>14.60</td>
<td>75</td>
</tr>
<tr>
<td>14.70</td>
<td>120</td>
</tr>
<tr>
<td>14.80</td>
<td>60</td>
</tr>
<tr>
<td>14.90</td>
<td>45</td>
</tr>
</tbody>
</table>

Mean $\bar{X} = 14.64$

Median = 14.65

Mode = 14.70

3. A panel of 76 economists was asked to predict the average unemployment rate for the upcoming year. The results of the survey follow.

<table>
<thead>
<tr>
<th>Unemployment Rate, %</th>
<th>Economists</th>
</tr>
</thead>
<tbody>
<tr>
<td>10.7</td>
<td>3</td>
</tr>
<tr>
<td>11.1</td>
<td>6</td>
</tr>
<tr>
<td>11.5</td>
<td>8</td>
</tr>
<tr>
<td>11.9</td>
<td>24</td>
</tr>
<tr>
<td>12.3</td>
<td>18</td>
</tr>
<tr>
<td>12.7</td>
<td>12</td>
</tr>
<tr>
<td>13.1</td>
<td>5</td>
</tr>
</tbody>
</table>

On the basis of this survey, what does the panel expect the average unemployment rate to be next year? (Round answer to two decimal places.)

Expected Value $= \bar{X} = 12.05$
4. On the basis of past experience, the manager of the VideoRama Store has compiled the following table, which gives the probabilities that a customer who enters the VideoRama Store will buy 0, 1, 2, 3, or 4 DVDs. How many DVDs can a customer entering this store be expected to buy? (Enter answer to two decimal places.)

<table>
<thead>
<tr>
<th>DVDs</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probability</td>
<td>0.41</td>
<td>0.35</td>
<td>0.16</td>
<td>0.06</td>
<td>0.02</td>
</tr>
</tbody>
</table>

Expected Value \( E(X) = X = 0.93 \)

5. A man purchased a $22,000, 1-year term-life insurance policy for $400. Assuming that the probability that he will live for another year is 0.988, find the company’s expected net gain.

\[
\begin{array}{c|c|c|c|c|c}
X & 400 & 400 - 22000 \\
\hline
P(X) & 0.998 & 1 - 0.998 \\
\hline
\end{array}
\]

\[
E(X) = (400 \times 0.998) + (400 - 22000 \times (1 - 0.998)) = 136
\]

Expected net gain is $136

6. A man wishes to purchase a life insurance policy that will pay the beneficiary $30,000 in the event that the man’s death occurs during the next year. Using life insurance tables, he determines that the probability that he will live another year is 0.96. What is the minimum amount that he can expect to pay for his premium? Hint: The minimum premium occurs when the insurance company’s expected profit is zero.

\[
\begin{array}{c|c|c|c|c|c}
X & x & x - 30000 \\
\hline
P(X) & 0.96 & 1 - 0.96 \\
\hline
\end{array}
\]

\[
E(X) = 0 \Rightarrow (x)(0.96) + (x - 30000)(1 - 0.96) = 0
\]

\[
0.04x + 0.04x - (0.04)(30000) = 0
\]

\[
x - 1200 = 0 \Rightarrow x = 1200
\]

Minimum Payment = $1200
7. A game consists of you rolling one fair 5-sided die and observing the number that appears uppermost. If you roll a 1 or a 5, then you win $6. If you roll any other number, then you do not win any money. Find your expected net winnings for this game if it costs $1 to play.

\[
\begin{array}{c|c|c|c|c|c|c}
X & 6 & 1 & 5 & 2 & 3 & P(X) \\
\hline
6 \cdot \frac{1}{5} & 1 \cdot \frac{1}{5} & 1 \cdot \frac{2}{5} & \frac{2}{5} & \frac{3}{5} & \frac{1}{5} \\
\hline
\end{array}
\]

\[E(X) = (6 \cdot \frac{2}{5}) + (-1 \cdot \frac{3}{5}) = \$1.4\]

8. A lottery has a grand prize of $100,000, three runner-up prizes of $10,000 each, nine third-place prizes of $2500 each, and twenty-two consolation prizes of $1000 each. If 300,000 tickets are sold for $1 each and the probability of any one ticket winning is the same as that of any other ticket winning, find the expected net winnings on a $1 ticket. (Round your answer to two decimal places.)

\[
\begin{array}{|c|c|c|c|c|c|}
\hline
X & 100000 & 10000 & 2500 & 1000 & 0 \rightarrow L1 \\
P(X) & \frac{1}{300000} & \frac{3}{300000} & \frac{9}{300000} & \frac{22}{300000} & \frac{35}{300000} \rightarrow L2 \\
\hline
\end{array}
\]

\[E(X) = \bar{X} = -0.42\]

\[\text{Expect to lose } \$0.42\]

**Odds In Favor Of and Odds Against** If the probability of an event \(E\) is \(P(E) = \frac{a}{b}\), where this probability is a reduced fraction, then

1. The odds in favor of \(E\) occurring are

\[a \text{ to } (b - a)\]

2. The odds against \(E\) occurring are

\[(b - a) \text{ to } a\]

**Probability of an Event (Given the Odds)** If the odds in favor of an event \(E\) occurring are \(a\) to \(b\), then the probability of \(E\) occurring is

\[P(E) = \frac{a}{a + b}\]
9. Carmen, a computer sales representative, believes that the odds are 8 to 3 that she will clinch the sale of a minicomputer to a certain company. What is the (subjective) probability that Carmen will make the sale? (Round answer to three decimal places.)

Odds to probability:
\[ P(\text{sale}) = \frac{8}{8+3} = \frac{8}{11} \approx 0.727 \]

10. A salesman has a 28% chance of a sale at his first stop, a 40% chance of a sale at his second stop and a 10% chance of a sale at both stops. Find the odds that

\( P(1\text{st}) = .28 \), \( P(2\text{nd}) = .40 \)
\( P(1\text{st} \cap 2\text{nd}) = .10 \)

(a) he will make a sale at both stops.
\[ P(\text{both}) = .10 = \frac{1}{10} \]
Odds in favor \( 1 \text{ to } \frac{10-1}{1+9} = 9 \text{ to } 41 \)

(b) he will make a sale only at the first stop.
\[ P(\text{only 1st}) = .18 = \frac{9}{50} \]
Odds in favor \( 9 \text{ to } (50-9) = 9 \text{ to } 41 \)

(c) he will make a sale at the second stop.
\[ P(2\text{nd}) = .40 = \frac{4}{10} = \frac{2}{5} \]
Odds in favor \( 2 \text{ to } (5-2) = 2 \text{ to } 3 \)
Section 8.3 Variance and Standard Deviation

Variance of a Random Variable $X$
Suppose a random variable has the probability distribution

<table>
<thead>
<tr>
<th>$x$</th>
<th>$x_1$</th>
<th>$x_2$</th>
<th>...</th>
<th>$x_n$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P(X = x)$</td>
<td>$p_1$</td>
<td>$p_2$</td>
<td>...</td>
<td>$p_n$</td>
</tr>
</tbody>
</table>

and expected value (or mean)

$$E(X) = \mu \text{ (pronounced “mu”) }$$

Then the variance of the random variable $X$ is

$$Var(X) = p_1(x_1 - \mu)^2 + p_2(x_2 - \mu)^2 + \ldots + p_n(x_n - \mu)^2$$

The variance is one measure of the spread of a probability distribution about its mean.

Standard Deviation of a Random Variable $X$
The standard deviation of a random variable $X$ denoted $\sigma$ (pronounced “sigma”), is defined by

$$\sigma = \sqrt{Var(X)}$$

Calculator Steps

STAT [EDIT] ENTER. Enter the random variable values into $L_1$ and the probabilities/frequencies into $L_2$.

STAT [CALC] 1. You should see 1-Var Stats on the screen. Then click 2ND, 1, 2ND, 2, ENTER.

$\sigma x$ is the standard deviation and if you square that value, $(\sigma x)^2$, you will get the variance. Also, as mentioned in last class, $\bar{x}$ is the mean or expected value.

1. The probability distribution of a random variable $X$ is given.

<table>
<thead>
<tr>
<th>$x$</th>
<th>410</th>
<th>490</th>
<th>530</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P(X = x)$</td>
<td>.2</td>
<td>.5</td>
<td>.3</td>
</tr>
</tbody>
</table>

Find the mean, standard deviation, and variance of $X$. (Round answers to two decimal places.)
2. Find the variance of the probability distribution for the histogram shown. (Round answer to two decimal places.)

![Histogram](image)

\[
\text{Var}(X) = (6\bar{x})^2
\]
\[
= (2.39)^2
\]
\[
= 5.71
\]

3. The following table gives the math SAT scores for males and females in the past five years. Find the mean, standard deviation, and variance for each group. (Round answers to two decimal places.)

<table>
<thead>
<tr>
<th></th>
<th>Males</th>
<th>Females</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>504</td>
<td>454</td>
</tr>
<tr>
<td></td>
<td>505</td>
<td>452</td>
</tr>
<tr>
<td></td>
<td>503</td>
<td>450</td>
</tr>
<tr>
<td></td>
<td>496</td>
<td>451</td>
</tr>
<tr>
<td></td>
<td>500</td>
<td>447</td>
</tr>
</tbody>
</table>

\[
\bar{x}_\text{Males} = 501.6
\]
\[
\text{Var}(x) = (3.26)^2
\]
\[
= 10.63
\]

\[
\bar{x}_\text{Females} = 450.8
\]
\[
\text{Var}(x) = (2.32)^2
\]
\[
= 5.38
\]

Which group is varying more?

Highest Variance

\[\text{Males}^{\square}\]
4. The distribution of the number of chocolate chips ($x$) in a cookie is shown in the following table.

<table>
<thead>
<tr>
<th>$x$</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P(X = x)$</td>
<td>.01</td>
<td>.04</td>
<td>.05</td>
<td>.11</td>
<td>.12</td>
<td>.24</td>
<td>.22</td>
<td>.17</td>
<td>.04</td>
</tr>
</tbody>
</table>

Find the mean and the variance of the number of chocolate chips in a cookie. (Round answers to two decimal places.)

STAT-CALC $\rightarrow$ Mean $= \bar{x} = 4.98$

$\text{Var}(X) = (\bar{x}^2) = (1.78)^2 = 3.17$