1 Introduction to Random Variables

A random variable is a variable whose value is

1. __________

and 2. __________.

A probability distribution tells us the possible values of a random variable, and the probability of having those values. We use random variables to help us quantify the results of experiments for the purpose of analysis. Random variables are generally denoted using __________, like __________.

Examples of random variables

1. Deal a poker hand. Let $X =$ the number of face cards. $X$ assigns to 3, 7, 10, Q, K the value __________.

2. Roll two six-sided dice. Let $Y =$ the sum of the two numbers rolled. $Y$ assigns to a roll of {2, 4} the value __________.

3. Let $H =$ the difference between the heights of a randomly selected couple (height of husband – height of wife, measured in inches). If my wife and I were the randomly selected couple, then the value of $H$ would be __________.

There are two kinds of random variables: __________ and __________.

1.1 Discrete Random Variables

If $X$ is discrete (has a finite number of possible values), we can describe the distribution of $X$ by making a table that lists each possible value of $X$ and the probability that it occurs.

<table>
<thead>
<tr>
<th>Value of $X$</th>
<th>$x_1$</th>
<th>$x_2$</th>
<th>$x_3$</th>
<th>...</th>
<th>$x_n$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probability</td>
<td>$p_1$</td>
<td>$p_2$</td>
<td>$p_3$</td>
<td>...</td>
<td>$p_n$</td>
</tr>
</tbody>
</table>

Of course, each probability ($p_i$) must be between ______ and ______ (inclusive), and the sum of all the probabilities ($\sum p_i$) must equal __________. You can find the probability of any event by adding the probabilities $p_i$ that correspond to the values $x_i$ that belong to the event.
Example. Roll 2 six-sided dice. Let $X$ = the number of sixes rolled. Find the probability distribution of $X$.

<table>
<thead>
<tr>
<th>Value of $X$</th>
<th>Probability</th>
</tr>
</thead>
</table>

We can also describe this probability distribution graphically, using a **probability histogram**:

What is $P(X \geq 1)$?

What is $P(X$ is even)?

Example. Sometimes we base a model for a random variable on data collected from a sample. For example, an industrial psychologist administered a personality inventory test for passive-aggressive traits to 150 employees. Individuals were rated on a scale from 1 to 5 (passive to aggressive). The results are as follows:

<table>
<thead>
<tr>
<th>Score</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>24</td>
</tr>
<tr>
<td>2</td>
<td>33</td>
</tr>
<tr>
<td>3</td>
<td>42</td>
</tr>
<tr>
<td>4</td>
<td>30</td>
</tr>
<tr>
<td>5</td>
<td>21</td>
</tr>
</tbody>
</table>

Let $T$ be rating of a random employee. Assuming the 150 tested employees are typical of all the employees, construct the probability distribution table for the random variable $T$.

<table>
<thead>
<tr>
<th>Value of $T$</th>
<th>Probability</th>
</tr>
</thead>
</table>

Now compute the following

$P(T > 4) =$

$P(T \leq 2) =$

$P(2 < T < 4) =$

$P(T > 5) =$

$P(T \geq 4) =$

Suppose this test were given to 500 employees. Based on the above probability distribution, about how many would we expect to be extremely aggressive (score of 5)?
2 Continuous Random Variables

A continuous random variable can take all values in an interval of numbers. Therefore the probability distribution of a continuous random variable can’t be described by making a table of values and probabilities. Instead, the probability distribution of a continuous random variable \( X \) is described by a **density curve**. The probability of any event is the area under the density curve and above the values of \( X \) that make up the event.

**Example.** Below is a simple density curve for a random variable \( X \).

\[
P(X > 1) \quad P(0.5 \leq X \leq 1.5) \quad P(X = 1)
\]

2.1 Normal Distributions

The *normal distributions* are a family of continuous probability distributions with distinctive bell-shaped density curves.

\( N(\mu, \sigma) \): normal distribution with mean \( \mu \) and standard deviation \( \sigma \)

If \( X \) is normally distributed, we will write \( X \sim N(\mu, \sigma) \). If \( X \) is approximately normally distributed, we will write \( X \approx N(\mu, \sigma) \).

**Example.** The heights of women 18–24 (according to data that is now several years dated) is approximately normally distributed with a mean of 64.3 inches and a standard deviation of 2.6 inches. If we let \( W \) be the height of a woman, then we can express this as \( W \approx N(64.3, 2.6) \). Let \( M \) be the height of an 18–24 year old man. \( M \approx N(70, 2.8) \).

We have already learned how to work out probabilities using the 68-95-99.7 rule. Some examples of probabilities we could compute this way include
But what do we do if we want to know probabilities like those below?

\[ P(W > 70) \]

\[ P(M < 64.3) \]

\[ P(68 < W < 70) \]

1. For each probability, sketch a density curve and shade the area that represents the probability desired. Be sure to include a scale.

2. Now we want to normalize the scale of our density curves, that is re-express the numbers involved as a number of standard deviations above or below the mean. Normalized values are called Z-scores.

\[ Z = \frac{\text{difference from mean}}{\text{standard deviation}} = \frac{X - \mu_X}{\sigma_X} \]

\( Z \) will be positive if

\( Z \) will be negative if

3. Once the problem has been normalized, we can determine the probability using a table of probabilities for the standard normal distribution, or a computer.

3 Mean, Variance and Standard Deviation of Random Variables

When we want to compute the mean of a list of numbers \( \{x_1, x_2, \ldots, x_n\} \), we simply add the numbers and divide by how many there were

\[ \text{mean} = \frac{x_1 + x_2 + \cdots + x_n}{n} \]

This can be written another way:

\[ \text{mean} = \frac{x_1}{n} + \frac{x_2}{n} + \cdots + \frac{x_n}{n} \]

Or

\[ \text{mean} = x_1 \cdot \frac{1}{n} + x_2 \cdot \frac{1}{n} + \cdots + x_n \cdot \frac{1}{n} \]

Written this last way, we can figure out how to compute the mean\(^1\) of a discrete random variable. Imagine choosing one of our numbers \( \{x_1, x_2, \ldots, x_n\} \) at random, each equally likely. Then \( \frac{1}{n} \) is the probability of choosing a particular number in our list. So we see that we have a sum of values times probabilities:

\[ \text{mean} = \sum x_ip_i \]

We’ll use this same formula to compute the mean of any discrete random variable. That is, to find the mean of \( X \), multiply each possible value by its probability, then add all the products.

---

\(^1\)Note: the mean of a random variable is also called the expected value of the random variable.
Suppose $X$ is a discrete random variable and the distribution of $X$ is given by

<table>
<thead>
<tr>
<th>Value of $X$</th>
<th>$x_1$</th>
<th>$x_2$</th>
<th>$x_3$</th>
<th>...</th>
<th>$x_n$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probability</td>
<td>$p_1$</td>
<td>$p_2$</td>
<td>$p_3$</td>
<td>...</td>
<td>$p_n$</td>
</tr>
</tbody>
</table>

Then to find the mean of $X$, multiply each possible value by its probability, then add all the products:

$$\text{mean} = \mu_X = \sum x_ip_i$$

Examples.
The variance and standard deviation of a discrete random variable is computed similarly:

Suppose $X$ is a discrete random variable and the distribution of $X$ is given by

<table>
<thead>
<tr>
<th>Value of $X$</th>
<th>$x_1$</th>
<th>$x_2$</th>
<th>$x_3$</th>
<th>...</th>
<th>$x_n$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probability</td>
<td>$p_1$</td>
<td>$p_2$</td>
<td>$p_3$</td>
<td>...</td>
<td>$p_n$</td>
</tr>
</tbody>
</table>

Then to find the variance of $X$, multiply the squared difference from the mean for possible value by its probability, then add all the products:

$$\text{variance} = \sigma_X^2 = \sum (x_i - \mu_X)^2 p_i$$

The standard deviation is the square root of the variance: $\sigma_X = \sqrt{\sigma^2}$. 
The means and standard deviations of continuous random variables are computed in a related way, but require more advanced mathematics (integral calculus) to express precisely.

### 3.1 Mean and Standard Deviation of Combinations of Random Variables

There are four rules that can help us compute means and variances of combinations of random variables. A fifth rule is an especially important property of normal distributions. These will be especially useful when we try to understand sampling distributions.

<table>
<thead>
<tr>
<th>Rule</th>
<th>Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>$\mu_{aX+b} = a\mu_X + b$</td>
</tr>
<tr>
<td>2.</td>
<td>$\mu_{X+Y} = \mu_X + \mu_Y$</td>
</tr>
<tr>
<td>3.</td>
<td>$\sigma_{aX+b} = a\sigma_X$</td>
</tr>
<tr>
<td>4. If $X$ and $Y$ are INDEPENDENT,</td>
<td>$\sigma_{X+Y}^2 = \sigma_X^2 + \sigma_Y^2$ (Note: this rule uses variances.)</td>
</tr>
<tr>
<td>5.</td>
<td>If $X$ and $Y$ are normally distributed and independent, then so are $aX + b$ and $X + Y$</td>
</tr>
</tbody>
</table>

**Example.** Do tall men and women tend to marry each other? Or are the heights of spouses independent. If the heights were independent, we could use the rules above to determine the probability that the husband is taller than the wife, $P(M - W > 0)$. 