Chapter 5
Probability Concepts

What is Probability?

Consider the frequency table used in Chapter 2:

Selling Price (Class)	Tally	Frequency (f)	Cumulative Frequency	Relative Frequency
$5,000 up to $10,000	|	1	1	0.0625
$10,000 up to $15,000	|||	3	4	0.1875
$15,000 up to $20,000	~~||||~~ |	6	10	0.3750
$20,000 up to $25,000	||||	4	14	0.2500
$25,000 up to $30,000	||	2	16	0.1250
Totals		16		1.0000

If we selected one sale from the above 16 sales, what is the *likelihood* that the sale was for an amount between $15,000 and $20,000?

The *relative frequency* of outcomes in that range represents the *likelihood* or **probability** that any **randomly selected** sale will fall within the limits of that class.

We can say that the 16 sales recorded on the above day represent an **experiment** conducted to obtain observations.

The actual sales dollars recorded for each sales range represent the **events** we measure by the experiment.

The **sample space** for our experiment represents all the possible simple events (sale prices) that are covered within the experiment.

We will denote the probability of an event occurring in our experiment by **P** and the event for which the probability is measured by **A, B, C, etc.**
Hence, the probability of event A occurring in our experiment would be denoted by **P(A),** the probability of event B would be **P(B)**, etc.

In accordance with the relative frequency approach, the probability of the occurrence of each event for this sample of 16 days can be described as

P(A) = (occurrences of event A) / (total events observed in the experiment)

That is, probability is a ratio of the number of outcomes of interest divided by the total number of possible outcomes.

From data that has been observed, probability can be written as the ratio of two numbers (note that this equation is the same as that for the relative frequency of an event):

$$P(x) = f_x / \Sigma f$$

What is the probability that a random sale occurs in each of the classes shown above?

What is the total of all the probabilities?

What is the probability of observing a sale between $5,000 and $30,000?

What is the probability of observing a sale of less than $5,000?

We note that it is **absolutely certain** that any sale we observe from this sample will be between $5,000 and $30,000. The probability of the sale falling within this range would be the sum of all possible probabilities in the experiment, or **1.**

We also note that it is absolutely **impossible** for a sale to be less than $5,000. Therefore, the probability of this outcome occurring must be **0.**

Therefore, for any outcomes x, the following is true:

$$\text{Impossible} \le P(x) \le \text{Certain}$$

$$0 \le P(x) \le 1$$

and

$$\Sigma P(x) = 1.0$$

We can also observe that, the more sales we observe, the more our relative frequency will represent the actual probability of the sales events. That is, the larger the number of events measured, the more accurate the results. This is the **Law of Large Numbers.**

Examples: Find the probability that a person rolling two dice will roll a 4 on any given roll.

The **complement** of event **A** is all events in which event **A** does not occur. We denote the complement by **~A.** Therefore,

$$P(A) = 1 - P(\sim A)$$

and

$$P(A) + P(\sim A) = 1$$

What is the probability that the dice roller above will NOT roll a 4?

Homework: pp. 127-128, Exercises 1, 3, 5, 8, 9

Compound Events

When two or more events are combined, the result is a **compound event** and the probability of their occurrence is a **joint probability.**

Consider the following events measured in an experiment:

	Aspirin	Tylenol	Control	Total
Upset Stomach	20	5	20	45
No upset	10	30	15	55
	30	35	35	100

What is the probability that a participant took aspirin?

What is the probability that a participant took Tylenol?

What is the probability that a participant took a placebo?

What is the probability that a participant took aspirin, Tylenol, or a placebo?

What is the probability that a person either took Aspirin, took Tylenol, or both?

Then, **P(A or B) = P(A) + P(B)** **[Special Rule of Addition]**

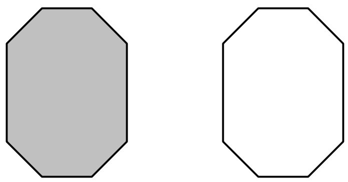
This is true because the events are **mutually exclusive**. That is, a participant either took aspirin or Tylenol, but not both. Taking one EXCLUDED taking the other.

What is the probability that a person either took Tylenol, got an upset stomach, or both? Note that some of the participants taking Tylenol as well as some NOT taking Tylenol got upset stomachs. Since both these events can happen to the same individual, taking Tylenol and getting an upset stomach are **NOT mutually exclusive,** and we can't count those more than once. Then:

P(A or B) = P(A) + P(B) – P(A and B)   [General Rule of Addition]

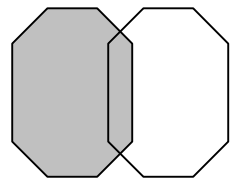
Note that for mutually exclusive events **P(A and B) = 0.**

This can be illustrated with the following Venn Diagrams



Event A   Event B

Mutually Exclusive Events



Event A   Event B

Overlapping Events

Example: Among 500 seats at a New York Theater, 60 are aisle seats and 100 are orchestra level seats (including 15 aisle seats). Find the probability that a randomly selected theater-goer has an aisle seat OR a seat in the orchestra section.

Example: The driving records of 650 New York drivers that had received traffic tickets were examined to determine if an accident with a DUI was involved. The following data was collected:

	Accident	No Accident	Total
DUI	100	150	
Not DUI	50	350	
Total			

a. Find the probability that a randomly selected driver had an accident?

b. Find the probability that a randomly selected driver had an accident or DUI?

c. Find the probability that a randomly selected driver was not DUI.

d. Find the probability that a randomly selected driver was not DUI or did not have an accident.

Homework: pp. 133-134, Exercises 11, 13, 15, 16, 17, 18, 19, 21

Looking again at our Aspirin/Tylenol example:

	Aspirin	Tylenol	Control	Total
Upset Stomach	20	5	20	45
No upset	10	30	15	55
	30	35	35	100

What is the probability that a participant took Tylenol AND had no upset stomach?

What is the probability that a participant took Aspirin AND had no upset stomach?

What is the probability that someone had NO stomach upset **given** that they took Tylenol?

What is the probability that someone had NO stomach upset **given** that they took Aspirin?

Note that the probability of both event A and event B happening is

P(A and B) = P(A) x P(B|A) **[General Rule of Multiplication]**

which can be written **P(B|A) = P(A and B) / P(A)**

where **P(B|A)** is the probability that event B happens given that event A has **already happened** (the probability of B given A). **P(B|A) is called a conditional probability.**

If event B is **independent** of event A, then **P(B|A) = P(B),** and the formula becomes

P(A and B) = P(A) x P(B) [Special Rule of Multiplication]

Event B is said to be independent of event A if the occurrence of A has no effect on the probability of B occurring. Thus, if P(A and B) = P(A) x P(B), that is, if P(B|A) = P(B), then the events are independent.

Two **events** that are **not independent** are said to be **dependent**.

Example: A test to assess job applicant cognitive ability has three questions – a 3 answer multiple choice question, a 2 answer multiple choice question, and a True/False question. What is the probability that a job applicant can guess at all three answers and get a perfect score on the test?

What is the complement to the above probability (i.e., what is the probability that a guesser would not get all three answers correct)?

Example: You have a box that contains 4 ping-pong balls. One is painted red and the other three are painted green. Without looking into the box, you are to draw ping-pong balls from the box.

What is the probability that you will pull the red ball from the box on the first try?

R

G P(Red) = 1/4

G

G

What is the probability that you would pull two green balls out first, assuming that you set aside each ball after drawing it from the box? (Sampling without replacement). Would the two events be dependent or independent?

Note: P(B|B) is not equal to P(B), which implies dependence

What is the probability that you would pull two green balls out first if you placed the first ball you had drawn back in the box before drawing the second ball? (Sampling with replacement). Would the two events be dependent or independent?

For **more than two independent events**, the probability of any sequence of events is simply the product of the probabilities corresponding to the individual events:

P(A and B and C and D) = P(A) x P(B) x P(C) x P(D)

For more than **two dependent events**, the probability formula becomes

P(A and B and C and D) =
P(A) x P(B|A) x P(C|A and B) x P(D|A and B and C)

Homework: p. 141, Exercises 23, 24, 25, 27, 28, 32

Counting Rules

Remember, the probability of any outcome is defined as a ratio, with the numerator being the number of ways to get that outcome and the denominator being the total number of possible outcomes.

$$P(x) = f_x / \Sigma f$$

It is apparent from this formula that, in order to calculate a probability, it is necessary to know both the numerator (number of possible ways to get outcome "x") AND the denominator (total number of possible outcomes). When we have a large number of possible outcomes and dependencies in a problem, it is frequently very difficult to determine one or both of these numbers and therefore impossible to determine probabilities. However, some tools are available to help determine the number of outcomes in such situations.

Multiplication Formula (Fundamental Counting Rule)

For a sequence of 2 events in which the first event can occur **m** ways and the second event can occur **n** ways, the events together can occur a total of **(m x n)** ways.

Suppose event 1 can occur 4 (m) ways and event 2 can occur 3 (n) ways. Then the tree diagram of possible outcomes looks like this:

Therefore, m and n can occur together a total of 12 possible ways. The 12 can be counted from the tree diagram, or found by multiplying m x n.

Example: South Carolina has auto license plates that have 3 numbers followed by 3 letters. How may unique license plates can be issued before it is necessary to come up with a new numbering scheme for the plates?

Example: The United States phone system uses a three-digit area code, three-digit exchange, and four-digit number to direct calls through the system. For example, (803) 555-2171. How many possible phone numbers are available for assignment in the United States?

Factorial Rule

The factorial rule applies when you need to know the number of orders or ways in which a set of items can be arranged. It is stated as follows:

n different items can be arranged in order **n!** ways.

Suppose you have three items, denoted by 1, 2, and 3 (n = 3). The first selection is from three possible items – 1, 2, or 3. After making this first selection, only two are left, so the second selection has only 2 available items. After having made 2 selections, the third selection has to be the remaining available item.
The tree diagram of these selections looks as follows:

The tree diagram leads directly to the result that 3 items can be arranged 3! ways:

Since "n" can be any integer, we can generalize as follows:

$n! = n \times (n-1) \times (n-2) \times ... \times 1$

Note: By definition, $0! = 1$

Example: A child has a set of 5 blocks each of which has painted on it a letter of her name, April. How many ways can the child arrange the blocks?

If April is 8 months old and is playing with the blocks, what is the probability that she will arrange the blocks to spell her name?

Permutation and Combination Rules
(for sampling without replacement)

Frequently, we need to select groups of items from larger groups of items. For example, a computer operator may be able to batch 4 jobs at a time. If each night he has 20 jobs to run, how many different batch sequences can he have? For situations like this we use either the permutation rule or the combination rule, depending on whether the order of selection is important. In other words, if he selects jobs 2, 5, 6, 19 to run, would that be considered the same batch as 5, 2, 19, 6 or would they be considered distinct batches (i.e., outcomes)?

If order is to be considered in our outcomes, then the number of sequences can be calculated from the formula below. When we say that order is important, that means that if A happens and then B happens we have a different outcome than if B happens and then A happens. The order of selection matters. Selection 1,2,3 and 2,1,3 are different outcomes.

The letter **P** stands for number of **permutations** (arrangements, sequences, or orders), **n** the number of available items, and **r** the number of items to be selected. Then the number of permutations (**P**) or arrangements of “n” things selected “r” at a time is:

$$_nP_r = (n!) / (n-r)!$$

Example: Calculate the number of possible outcomes if we select 3 items from 5 available items, and the same three items selected in different orders are considered different outcomes.

Example: A child, April, has a set of alphabet blocks, each of the 26 blocks containing a single letter of the alphabet. She randomly selects 5 blocks at a time in an attempt to spell her name. How many different orders of blocks could she select?

What is the probability that April, randomly selecting blocks, will spell her name?

Note that the factorial rule is just a special case of the permutations rule. When April had only five blocks to chose from and selected 5 at a time, there were 120 different way in which she could select (arrange) the blocks. If we set that up as a permutation formula, (order is important here), then we have

$${}_5P_5 = 5! / (5-5)! = 5! / 0! = 5! / 1 = 5! = 120$$

In other words, the permutation rule becomes the factorial rule when $n = r$.

Example: A paper route consists of delivering newspapers to 10 homes. The delivery person gets bored driving the same route every morning, so decides to drive a different route every day until he has exhausted all possible routes. How many days will it take him to achieve his goal?

Combinations are the same as permutations except that the order in which they are selected does not matter. That is 1, 2, 3 is considered to be the same outcome as 2,1,3, is the same as 3,1,2, etc. Therefore, all things being equal, there are fewer possible combinations than permutations.
Let **C** denote the number of combinations, then

$$\mathbf{{}_nC_r = (n!) / [(n - r)!\ r!]}$$

Example: Calculate the number of possible outcomes if we select 3 items from 5 available items, and selection in any order is the same outcome (order of selection doesn't matter).

Example: An artist wants to paint a picture using only 3 colors. He has available 15 colors. How many possible combinations of colors can he select?

Order here is unimportant because three colors selected in any order are the same three colors – combinations.

The artist randomly selects 3 colors to use in his painting. What is the probability that the painting will be umber, white, and purple?

Example: The South Carolina Powerball lottery requires the player to match 5 numbers out of a possible 55 numbers and in addition match the single powerball number selected from 42 possible numbers. Does the order of selection matter?

What is the probability that an individual buying a single Powerball ticket will win the grand prize?

Homework: pp. 146-147, Exercises 33, 34, 35, 36, 39, 40
pp. 148-152, Exercises 42, 43, 44, 48, 49, 51, 53, 59, 61, 67, 70, 73