Chapter Four

Probability

Counting Methods

Probability of a Single Event

Probability of Multiple Events

Expected Value

Combinations

A **combination** is a group of selected items.

The number of possible combinations of r items in a group of n items is ***n* choose *r***, written $\binom{n}{r}$ or $_{n}C_{r}$.

Group of n elements	Chosen group of r elements	Combination Example	Number of possible combinations
7 days of the week	1 day	Thursday	$\binom{7}{1} = 7$
12 months of the year	4 months	June, September, October, December	$\binom{12}{4} = 495$

The **counting principle** states that if two independent events have a and b possible outcomes, respectively, then there are a total of ab possible outcomes for the two events. This can be expanded to abc possible outcomes for three events, etc.

Events	Outcome Example	Number of possible outcomes
1 day of the week and 1 month of the year	Thursday; October	$\binom{7}{1}\binom{12}{1} = 84$
1 day of the week and 4 months of the year	Thursday; June, September, October, December	$\binom{7}{1}\binom{12}{4} = 3465$

Choosing one element at a time

Events	Outcome Example	Number of possible outcomes
Roll 3 six-sided dice	6; 3; 5	$\binom{6}{1}\binom{6}{1}\binom{6}{1} = 216$
Choose 3 favorite colors, in order, out of 8 colors	black; red; orange	$\binom{8}{1}\binom{7}{1}\binom{6}{1} = 336$
Choose a background color, a border color, and a text color, out of 8 colors	red; black; white	$\binom{8}{1}\binom{7}{1}\binom{6}{1} = 336$

For independent events (values can be repeated) like the dice example, the calculation can be simplified to the exponential n^r, in this case 6^3.

For dependent events (values cannot be repeated) like the colors examples, the calculation can be simplified to the permutation ${}_nP_r$, in this case ${}_8P_3$. A **permutation** is a combination in which each item selected is assigned a specific value, such as *first*, *second*, and *third*, or *background*, *border*, and *text*.

Set Notation

A **set** is a combination.

The examples below use the sets $A = \{\text{Saturday, Sunday}\}$ and $B = \{\text{Sunday, Tuesday, Thursday}\}$.

Term	Definition	Notation and Example
Element of *A*	item in A	Saturday $\in A$
Cardinality of *A*	number of elements in A	$\lvert A \rvert = 2$
Intersection of *A* and *B*	set of elements in both A and B	$A \cap B = \{\text{Sunday}\}$
Union of *A* and *B*	set of elements in either A or B	$A \cup B = \{\text{Sunday, Tuesday, Thursday, Saturday}\}$
Complement of *A*	set of elements not in A	$A' =$ the set of weekdays
Universal Set	set of all elements in the given context	$U =$ the set of days of the week
Empty Set	set containing no elements	$\emptyset = \{\ \}$

The probability of an event

The **sample space** of an event is the set of all possible outcomes.

The probability of an event A, written $P(A)$, can be defined as the number of outcomes satsifying A divided by the total number of outcomes in the sample space: $P(A) = \frac{|A|}{|U|}$. For this definition to apply, the outcomes must all be equally likely.

Event A	Outcomes satisfying A	Sample space	Probability of A
Roll higher than 2 on a 6-sided die	3, 4, 5, 6	1, 2, 3, 4, 5, 6	$P(A) = \frac{2}{6}$
Out of three coin flips, two are heads and one is tails.	HHT, HTH, THH	HHH, HHT, HTH, THH, HTT, THT, TTH, TTT	$P(A) = \frac{3}{8}$

Though it is not incorrect to reduce probabilities or convert them to decimals or percents, doing so removes information about the event. Do not do so in this course, except for fractions equaling 0 or 1.

Using counting methods to find probabilities

Many events have too many possible outcomes to list all of them. Instead, the possible outcomes can be counted using the counting principle. The examples below are for rolling three 6-sided dice.

Event *A*	Number of outcomes satisfying *A*	Size of sample space	Probability of *A*
Roll three 6's	1 (6-6-6)	$6 \cdot 6 \cdot 6 = 216$	$P(A) = \frac{1}{216}$
Roll two 6's and a 5	3 (6-6-5, 6-5-6, 5-6-6)	$6 \cdot 6 \cdot 6 = 216$	$P(A) = \frac{3}{216}$

Outcomes can be counted using combinations with the counting principle. This applies in all cases, but can be especially helpful for more complicated probabilities. The first two examples below are the same as above, and the last two are for choosing 5 numbers out of 30 for a lottery ticket.

Event *A*	Number of outcomes satisfying *A*	Size of sample space	Probability of *A*
Roll three 6's	$\binom{3}{3} = 1$	$\binom{6}{1}\binom{6}{1}\binom{6}{1} = 216$	$P(A) = \frac{1}{216}$
Roll two 6's and a 5	$\binom{3}{2}\binom{1}{1} = 3$	$\binom{6}{1}\binom{6}{1}\binom{6}{1} = 216$	$P(A) = \frac{3}{216}$
Choose the winning 5 numbers	$\binom{5}{5} = 1$	$\binom{30}{5} = 142{,}506$	$P(A) = \frac{1}{143506}$
Choose 3 winning and 2 losing numbers	$\binom{5}{3}\binom{2}{2} = 10$	$\binom{30}{5} = 142{,}506$	$P(A) = \frac{10}{143506}$

Given information

Probability is based on the information known, regardless of what has happened. **Conditional** probability takes into account known conditions.

Event	Probability	Given condition	Conditional probability
A card is hearts.	$\frac{13}{52}$	The card is red.	$\frac{13}{26}$
The second card drawn is hearts.	$\frac{13}{52}$	The first card was hearts.	$\frac{12}{51}$
The third card drawn is hearts.	$\frac{13}{52}$	The next 12 cards are all hearts.	$\frac{1}{40}$

Data Snooping

Calculating probabilities without taking into account given information can be a form of **data snooping**, in which the data used to test a hypothesis are the same data that were used to form the hypothesis. Although sometimes useful, this type of circular reasoning runs a high risk of indicating that coincidental events are not coincidental.

Event	Probability	Given condition	Conditional probability
Cassidy rolls a 6, then a 5, then a 4.	$\frac{1}{216}$	It already happened.	100%
Evolution could create eyes.	very low	Eyes exist.	very high

Probabilities of multiple events

For probability problems involving multiple events, the individual probabilities can be multiplied together. Keep in mind that, in some cases, the individual probabilities change based on the events already accounted for, such as fewer cards being left in a deck as more cards are drawn. Such events are called **dependent events**.

Events	Type	Probability	Comment
Roll 3 dice; all land on 6.	independent	$(\frac{1}{6})(\frac{1}{6})(\frac{1}{6}) = \frac{1}{216}$	Each die roll is unaffected by the others.
Draw 3 cards; all are aces.	dependent	$(\frac{4}{52})(\frac{3}{51})(\frac{2}{50}) = \frac{24}{132600}$	Once you know an ace has been drawn, there are only 3 aces possible to draw out of the remaining 51 cards.
Draw 3 cards, putting them back each time; all are aces.	independent	$(\frac{4}{52})(\frac{4}{52})(\frac{4}{52}) = \frac{64}{140608}$	There are always 4 aces available. You could draw the same ace each time.
Draw 3 cards; they are hearts, clubs, king, in that order.	dependent (hearts and clubs), independent (king)	$(\frac{13}{52})(\frac{13}{51})(\frac{4}{52}) = \frac{676}{137904}$	Even though you know you drew a heart and a club, you have no information about whether or not you drew a king. All 52 cards are still equally likely to be a king.

Different arrangements of multiple events

Some events can happen in different orders or ways. For example, rolling a 1 and a 2 on two dice could be (1, 2) or (2, 1). To find probabilities for these, every possible order or way must be included. This can be done by finding the size of the entire sample space and the number of successful outcomes, or by finding the probability of each possibility and adding. In many cases, each possibility has the same probability, so it only has to be calculated once and then multiplied by how many there are.

The examples below refer to pulling three marbles out of a jar of 5 blue, 3 red, and 1 black marble.

Event	By counting possibilities	By multiplying probabilities
All three marbles are the same color.	Choose 3 of the 5 blue or 3 of the 3 red: $\frac{\binom{5}{3}+\binom{3}{3}}{\binom{9}{3}} = \frac{11}{84}$	(blue, blue, blue) or (red, red, red): $\left(\frac{5}{9}\right)\left(\frac{4}{8}\right)\left(\frac{3}{7}\right)+\left(\frac{3}{9}\right)\left(\frac{2}{8}\right)\left(\frac{1}{7}\right)=\frac{66}{504}$
Two marbles are red and one is blue.	Chose 2 of the 3 red, 1 of the 5 blue, and 0 of the 1 black: $\frac{\binom{3}{2}\binom{5}{1}\binom{1}{0}}{\binom{9}{3}} = \frac{15}{84}$	(red, red, blue), but choose which 2 of the 3 are red: $\binom{3}{2}\left(\frac{3}{9}\right)\left(\frac{2}{8}\right)\left(\frac{5}{7}\right)=\frac{90}{504}$
All three marbles are different colors.	Chose 1 of the 3 red, 1 of the 5 blue, and 1 of the 1 black: $\frac{\binom{3}{1}\binom{5}{1}\binom{1}{1}}{\binom{9}{3}} = \frac{15}{84}$	(blue, red, black), but choose which 1 of the 3 is blue and which 1 of the remaining 2 is red: $\binom{3}{1}\binom{2}{1}\left(\frac{5}{9}\right)\left(\frac{3}{8}\right)\left(\frac{1}{7}\right)=\frac{90}{504}$

Conditional probability tables

The probability of A given B is known is written $P(A \mid B)$ and can be calculated $P(A \mid B) = \frac{|A \cap B|}{|B|}$. For example, the probability of a card being hearts, given it is red, is $P(\text{hearts} \mid \text{red}) = \frac{13 \text{ (red hearts)}}{26 \text{ (red cards)}}$.

In some cases with conditional probability it is helpful to make a table. The table below is for a polygraph that identifies 70% of liars and 25% of truth-tellers as liars, assuming that 20% of people tested are liars.

	B: claims is lying	B': claims is not lying	Total
A: liar	$P(A \cap B) =$ $(.20)(.70) = .14$	$P(A \cap B') =$ $(.20)(.30) = .06$	$P(A) =$ $.14 + .06 = .20$
A': not lying	$P(A' \cap B) =$ $(.80)(.25) = .20$	$P(A' \cap B') =$ $(.80)(.75) = .60$	$P(A') =$ $.20 + .60 = .80$
Total	$P(B) =$ $14 + .20 = .34$	$P(B') =$ $.06 + .60 = .66$	1.00

The table makes it easy to calculate probabilities by dividing a probability by the row or column total.

a) A person is lying: $P(A) = \frac{.20}{1.00} = 20\%$

b) A person is lying and is found lying: $P(A \cap B) = \frac{.14}{1.00} = 14\%$

c) A liar is found lying: $P(B \mid A) = \frac{.14}{.20} = 70\%$

d) A person found lying is a liar : $P(A \mid B) = \frac{.14}{.34} = 41\%$

Using the complement

In many cases, the probability of an event's complement is easier to calculate than the probability of the event itself. Since the complement is everything the original event is not, these two probabilities must add up to 100%: $P(A) = 1 - P(A')$. For example, if there is an 80% chance it will rain today, there is a 20% chance it will not rain today.

The examples below refer to rolling four 6-sided dice.

Event A	Complement A'	$P(A')$	$P(A)$
Roll at least one 6.	Roll zero 6's.	$\left(\frac{5}{6}\right)\left(\frac{5}{6}\right)\left(\frac{5}{6}\right)\left(\frac{5}{6}\right) = \frac{625}{1296}$	$1 - \frac{625}{1296} = \frac{671}{1296}$
Roll at least two 6's.	Roll zero or one 6.	$\frac{625}{1296} + \binom{4}{1}\left(\frac{1}{6}\right)\left(\frac{5}{6}\right)\left(\frac{5}{6}\right)\left(\frac{5}{6}\right) = \frac{1125}{1296}$	$1 - \frac{1125}{1296} = \frac{171}{1296}$

Expected Value

The **expected value** of an event is its theoretical average value. It is a weighted average, where the weightings are the probabilities.

In the example below, a player draws two cards and wins $10 if the first card is an ace and $20 if both cards are clubs.

Event	Probability	Value	Product
First card is an ace.	$\frac{4}{52}$	$10	$0.77
Both cards are clubs.	$\left(\frac{13}{52}\right)\left(\frac{12}{51}\right)=\frac{156}{2652}$	$20	$1.18
Total			**$1.95**

In this example, the player wins an average of $1.95 for each play. If there is a per-game price, this must be subtracted from the expected value. For example, if the player must pay $5 each game, the expected value is actually $1.95 – $5.00 = **-$3.05**, that is, the player is expected to lose an average of $3.05 per game.