3.4: Union and Intersection

Learning Objectives

- Find the probability of the intersection of events
- Find the probability of the union of events

Mutually Exclusive Events

When two events cannot happen at the same time, they are called **mutually exclusive** or **disjoint events.**

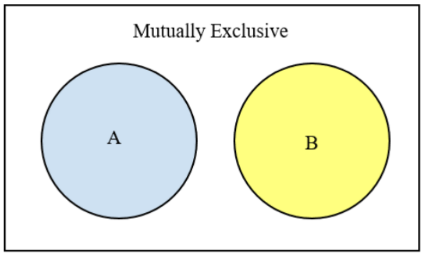


Figure $3.4.1$

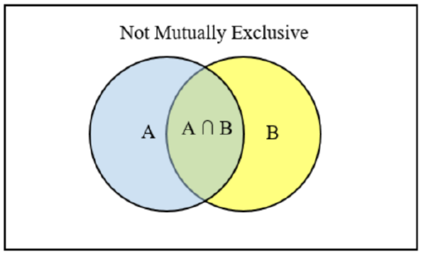


Figure $3.4.2$

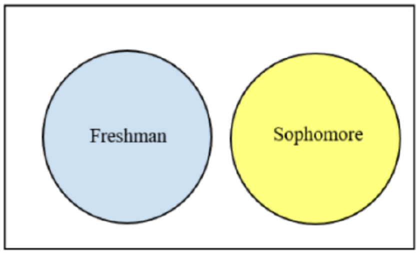


Figure $3.4.3$

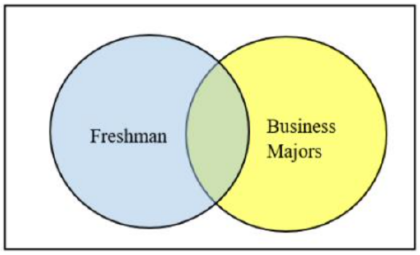


Figure $3.4.4$

For example, a student cannot be a freshman and a sophomore at the same time, see Figure $3.4.3$. These are mutually exclusive events. A student could be freshman and a business major at the same time so the event "freshman" and the event "business majors" are not mutually exclusive, as shown in Figure $3.4.4$.

Intersection

When we are finding the probability of both *A* **and** *B* happening at the same time, we denote this as $P(A \cap B)$. This overlap is called the intersection.

Definition: **Intersection**

When two events *A* and *B* occur at the same time, it is called the **intersection** of *A* and *B* and is denoted as $(A \cap B)$. Think of the symbol ∩ as an A in "and."

Formula

If two events are mutually exclusive, then $A \cap B$ = { }, the empty set (also denoted as ∅) and $P(A \cap B)$ = 0.

Union

When either event *A*, event *B*, or both occur, then it is called the **union** of *A* or *B*, which is denoted as $A \cup B$. When finding the probability of *A* or *B*, it is denoted as $P(A \cup B)$. When we write "or" in statistics, we mean "and/or" unless we explicitly state otherwise. Thus, "*A* or *B* occurs" means *A*, *B*, or both *A* and *B* occur.

Figure $3.4.5$ is a Venn diagram for the union of sets *A* or *B*, which is both sets put together.

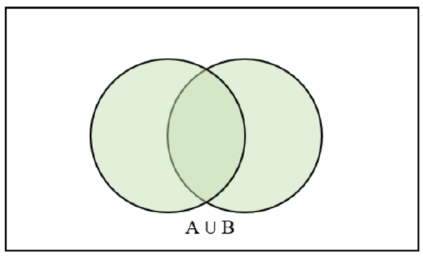


Figure $3.4.5$

Formula: **Addition Rule**

$$\textbf{Addition Rule: } P(A \cup B) = P(A) + P(B) - P(A \cap B) \quad .$$

If two events are mutually exclusive, then the probability of them occurring at the same time is $P(A \cap B) = 0$. So if *A* and *B* are mutually exclusive, then the $P(A \cup B) = P(A) + P(B)$ as shown in Figure $3.4.1$. It is best to write out the rule with the intersection so that you do not forget to subtract any overlapping intersection.

Example $3.4.1$

The family college data set contains a sample of 792 cases with two variables: teen and parents. The teen variable is either college or not, where the college label means the teen went to college immediately after high school. The parents variable takes the value degree if at least one parent of the teenager completed a college degree. Make a Venn Diagram for the data.

Example from OpenIntroStatistics.

Teen/Parents	Degree	No Degree	Total
College	231	214	445
No College	49	298	347
Total	280	512	792

Solution

Find the relative frequencies by dividing each cell by the grand total, 792.

Teen/Parents	Degree	No Degree	Total
College	0.29	0.27	0.56
No College	0.06	0.38	0.44
Total	0.35	0.65	1

Use shapes to show the overlaps between the groups, as show in the completed Venn diagram in Figure $3.4.6$. Note that you do not need to use circles to represent the sets.

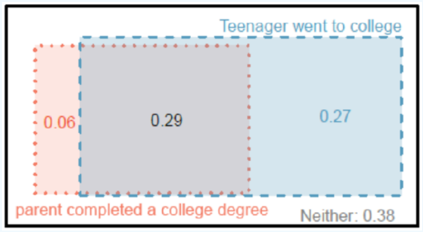


Figure $3.4.6$

✔ Example $3.4.2$

A random sample of 500 people was taken from the 2020 United States Census. Their marital status and race were recorded in the following contingency table using the census labels. A person is randomly chosen from the census data. Find the following.

Marital Status/Race	American Indian	Black	Asian	White	2 Major Races	Total
Divorced	0	6	1	30	1	38
Married	1	25	23	156	4	209
Single	2	33	21	155	11	222
Widowed	0	7	2	22	0	31
Total	3	71	47	363	16	500

a. P (Single ∩ American Indian)
b. P(Single ∪ American Indian)
c. Probability that the person is Asian or Married.
d. P(Single ∩ Married)
e. P(Single ∪ Married)

Solution

a. The intersection for a contingency table is found by simply finding where the row and column meet. There are 2 Single American Indians, therefore the P(Single ∩ American Indian) = P(Single and American Indian) = $\frac{2}{500} = \frac{1}{250}$ = 0.004.
b. There are 222 Single people, there are 3 American Indians, but we do not want to count the 2 Single American Indians twice, therefore the P(Single ∪ American Indian) = P(Single or American Indian) = $\frac{222}{500} + \frac{3}{500} - \frac{2}{500} = \frac{223}{500}$ = 0.446.
c. The union for a contingency table is found by either using the Addition Rule or adding up all the numbers in the corresponding row and column. There are 47 Asian people, there are 209 Married people, but we do not want to count the 23 Married Asian people twice. Thus, P(Asian ∪ Married) = P(Asian or Married) = P(Asian) + P(Married) - P(Asian ∩ Married) = $\frac{47}{500} + \frac{209}{500} - \frac{23}{500} = \frac{233}{500}$ = 0.466.

d. The events Single and Married are mutually exclusive so P(Single ∩ Married) = 0. Alternatively, there is no place in the table where the Single row and Married row meet.
e. The events Single and Married are mutually exclusive so the P(Single ∪ Married) = P(Single) + P(Married) – P(Single ∩ Married) = $\frac{222}{500}+\frac{209}{500}-0=\frac{431}{500}$ = 0.862.

Example 3.4.3

Use a random experiment consisting of rolling two dice and adding the numbers on the faces.

a. Find the probability of rolling a sum of 8.
b. Find the probability of rolling a sum of 8 or a sum of 5.
c. Find the probability of rolling a sum of 8 or a double (each die has the same number).

Solution

a. There are 36 possible outcomes for rolling the two dice as shown in the following sum table. There are 5 pairs where the sum of the two dice is an 8, the pairs (2,6), (3,5), (4,4), (5,3) and (6,2). Note that the events (2,6) and (6,2) are different outcomes since the numbers come from different dice.

1st die/2nd die	**1**	**2**	**3**	**4**	**5**	**6**
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

Thus, P(8) =$\frac{5}{36}$ = 0.1389.

b. Highlight all the places where a sum of 5 or a sum of 8 occurs. There are 9 pairs where the sum of the two dice is a 5 or an 8. Note that rolling a sum of 5 is mutually exclusive from rolling a sum of 8 so the probability is zero for the intersection of the two events.

1st die/2nd die	**1**	**2**	**3**	**4**	**5**	**6**
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

Thus, the P(5 ∪ 8) = P(5) + P(8) – P(5 ∩ 8) = $\frac{4}{36}+\frac{5}{36}-0=\frac{9}{36}=\frac{1}{4}$ = 0.25.

c. The events rolling a sum of 8 and rolling doubles are not mutually exclusive since the pair of fours (4, 4) falls into both events. An easy way is to highlight all the places a sum 8 or doubles occur, count the highlighted values, and divide by the total. Thus, P(8 or Doubles) = $\frac{10}{36}=\frac{5}{18}$ = 0.2778.

1st die/2nd die	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

When using the Addition Rule, we subtract this overlap out one time to account for the pair of 4s being counted twice, once in both events. P(8 ∪ Doubles) = P(8) + P(Doubles) – P(8 ∩ Doubles) = $\frac{5}{36}+\frac{6}{36}-\frac{1}{36}=\frac{10}{36}=\frac{5}{18}$ = 0.2778.

Example 3.4.4

Randomly pick a card from a standard deck. A standard deck of cards, not including jokers consists of 4 suits called clubs = ♣, spades = ♠, hearts = ♥, and diamonds = ♦. The clubs and spades are the black cards. The hearts and diamonds are the red cards. Each suit has 13 cards. The numbered cards shown in Figure 3.4.7, are Ace = 1 or A, 2, 3, 4, 5, 6, 7, 8, 9, 10. The face cards are the Jack = J, Queen = Q, and King = K.

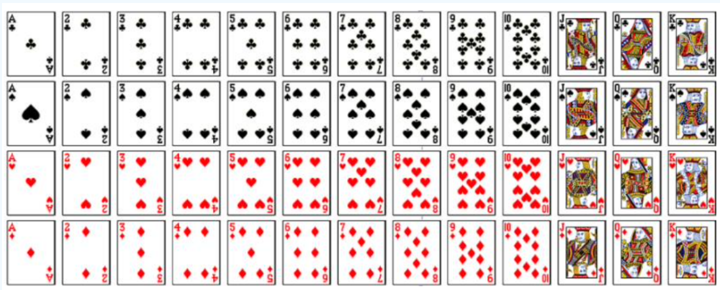
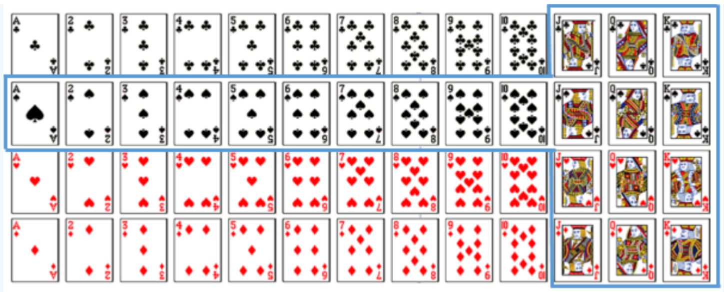


Figure 3.4.7

a. Find the probability of selecting a card that shows a club.
b. Find the probability of selecting a heart or a spade card.
c. Find the probability of selecting a spade or a face card.

Solution

a. There are 52 cards in a standard deck and 13 clubs. Thus, P(♣) = $\frac{13}{52}=\frac{1}{4}$ = 0.25.
b. There are 52 cards in a standard deck and 13 cards of each suit. Hearts and spades are mutually exclusive since there are no hearts that are also spades. Thus, P(♥ ∪ ♠) = P(♥) + P(♠) - P(♥ ∩ ♠) = $\frac{13}{52}+\frac{13}{52}-0=\frac{26}{52}=\frac{1}{2}$ = 0.5.
c. There are 13 spades and 12 face cards. However, there are 3 cards that are both spades and face cards. P(♠ ∪ FC) = P(♠) + P(FC) – P(♠ ∩ FC) = $\frac{13}{52}+\frac{12}{52}-\frac{3}{52}=\frac{22}{52}=\frac{11}{26}$ = 0.4231.



Since the sample space is small, you could just count how many spades and face cards there are.

Words Are Important!

When working with probability, words such as “more than” or “less than” can drastically change the answer. Figure 3.4.8 shows some of the common phrases you may run into while reading a problem. It will be essential later in the course that you can correctly match these phrases with their correct symbol.

=	≤	≥
is the same as	is less than or equal to	is greater than or equal to
is equal to	is at most	is at least
is exactly the same as	is not greater than	is not less than
has not changed from	within	

≠	>	<
is not	more than	less than
is not equal to	greater than	below
is different from	above	lower than
has changed from	higher than	shorter than
is not the same as	longer than	smaller than
	bigger than	decreased
	increased	reduced

Figure 3.4.8

Example 3.4.5

Use a random experiment consisting of rolling two dice and adding the numbers on the faces.

a. Find the probability of rolling a sum of less than 5.
b. Find the probability of rolling a sum of 5 or less.

Solution

a. Let X be the sum of the faces when rolling two dice. A sum less than 5 would not include 5. For notation, we use $P(X < 5)$, which is read as the “probability that X is less than five.” Highlight all the sums that are less than 5.

1st die/2nd die	1	2	3	4	5	6
1	2	3	4	5	6	7

2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

Thus, P(X < 5) = $\frac{6}{36} = \frac{1}{6}$ = 0.1667.

b. Let X be the sum of the faces when rolling two dice. A sum of 5 or less includes 5. For notation, we can use P(X ≤ 5), which is read as the "probability that X is less than or equal to five." Highlight all the sums that are 5 or less.

1st die/2nd die	**1**	**2**	**3**	**4**	**5**	**6**
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

Thus, P(X ≤ 5) = $\frac{10}{36} = \frac{5}{18}$ = 0.2778.