Probability and Statistics

Topic 7 - The Addition Rule and Complements

Aamir Alaud Din, PhD

October 30, 2023

TABLE OF CONTENTS

TABLE OF CONTENTS

OBJECTIVES

After learning this topic and studying, you should be able to:

1. Use the addition rule for disjoint events
2. Use the general addition rule
3. Compute the probability of an event using complement rule

TABLE OF CONTENTS

THE WHY SECTION

- The probability of an event can be computed by 1) empirical method or 2) classical method.
- If two events occur simultaneously and the outcomes of the events are mutually exclusive, then, how to compute the probability of these mutually exclusive events.
- The logic suggests to add the probabilities of such events because the events are mutually exclusive.
- What if the events are not mutually exclusive?
- Our target to study this topic is to learn how to compute the probabilities of 1) mutually exclusive and 2) mutually non-exclusive events.

TABLE OF CONTENTS

ADDITION RULE FOR DISJOINT EVENTS

- Before presenting more rules for computing probabilities, we must discuss disjoint events.

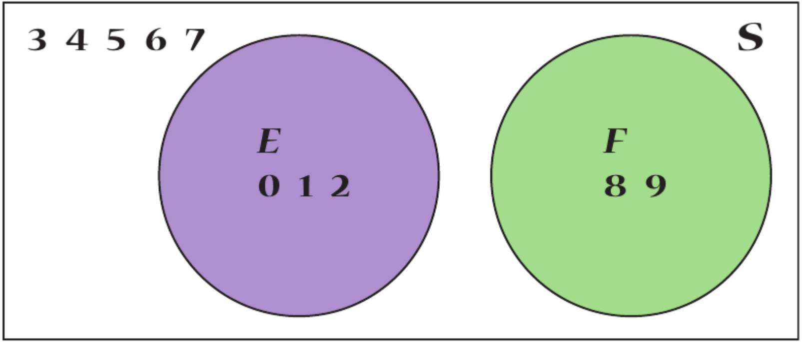
Disjoint or Mutually Exclusive Events

Two events are disjoint if they have no outcomes in common. Another name for disjoint events is mutually exclusive events.

- We can use Venn diagrams to represent events as circles enclosed in a rectangle.
- The rectangle represents the sample space, and each circle represents an event.
- For example, suppose we randomly select chips from a bag.
- Each chip is labeled 0, 1, 2, 3, 4, 5, 6, 7, 8, 9.

ADDITION RULE FOR DISJOINT EVENTS

- Let E represent the event "choose a number less than or equal to 2," and let F represent the event "choose a number greater than or equal to 8."
- Because E and F have no outcomes in common, they are disjoint.



ADDITION RULE FOR DISJOINT EVENTS

- Notice that the outcomes in event E are inside circle E, and the outcomes in event F are inside circle F.
- All outcomes in the sample space that are not in E or F are outside the circles, but inside the rectangle.
- From this diagram, we know that

$$P(E) = \frac{N(E)}{N(S)} = \frac{3}{10} = 0.3 \text{ and } P(F) = \frac{N(F)}{N(S)} = \frac{2}{10} = 0.2$$

- In addition,

$$P(E \text{ or } F) = \frac{N(E \text{ or } F)}{N(S)} = \frac{5}{10} = 0.5$$

ADDITION RULE FOR DISJOINT EVENTS

- In other words

$$P(E \text{ or } F) = P(E) + P(F) = 0.3 + 0.2 = 0.5$$

- This result occurs because of the Addition Rule for Disjoint Events.

Addition Rule for Disjoint Events

If E and F are disjoint (or mutually exclusive) events, then

$$P(E \text{ or } F) = P(E) + P(F)$$

- The Addition Rule for Disjoint Events can be extended to more than two disjoint events.

ADDITION RULE FOR DISJOINT EVENTS

- In general, if E, F, G, $\ldots$ each have no outcomes in common (they are pairwise disjoint), then

$$P(E \text{ or } F \text{ or } G \text{ or } \ldots) = P(E) + P(F) + P(G) + \ldots$$

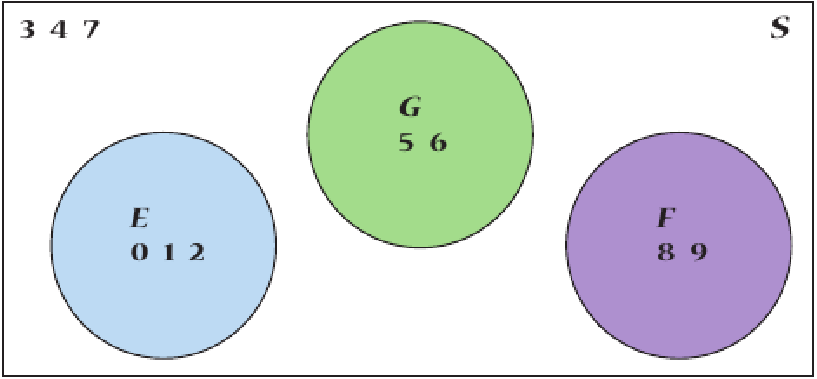
- Let event G represent "the number is a 5 or 6."
- The Venn diagram in Figure below illustrates the Addition Rule for more than two disjoint events using the chip example.
- Notice that no pair of events has any outcomes in common.
- So, from the Venn diagram, we can see that

$$P(E) = \frac{N(E)}{N(S)} = \frac{3}{10} = 0.3,\ P(F) = \frac{N(F)}{N(S)} = \frac{2}{10} = 0.2, \text{ and } P(G) = \frac{N(G)}{N(S)} = \frac{2}{10} = 0.2.$$

ADDITION RULE FOR DISJOINT EVENTS

- In addition

$$P(E \text{ or } F \text{ or } G) = P(E) + P(F) + P(G) = 0.3 + 0.2 + 0.2 = 0.7$$



ADDITION RULE FOR DISJOINT EVENTS

EXAMPLE 1

Our number system consists of the digits 0, 1, 2, 3, 4, 5, 6, 7, 8, and 9. Because we do not write numbers such as 12 as 012, the first significant digit in any number must be 1, 2, 3, 4, 5, 6, 7, 8, or 9. We may think that each digit appears with equal frequency so that each digit has a 1/9 probability of being the first significant digit, but this is not true. In 1881, Simon Newcomb discovered that digits do not occur with equal frequency. The physicist Frank Benford discovered the same result in 1938. After studying lots and lots of data, he assigned probabilities of occurrence for each of the first digits, as shown in Table 5. The probability model is now known as Benford's Law and plays a major role in identifying fraudulent data on tax returns and accounting books.

Table 5

Digit	1	2	3	4	5	6	7	8	9
Probability	0.301	0.176	0.125	0.097	0.079	0.067	0.058	0.051	0.046

ADDITION RULE FOR DISJOINT EVENTS

EXAMPLE 1 (CONTINUED)

1. Verify that Benford's Law is a probability model.
2. Use Benford's Law to determine the probability that a randomly selected first digit is 1 or 2.
3. Use Benford's Law to determine the probability that a randomly selected first digit is at least 6.

EXAMPLE 2

Suppose that a single card is selected from a standard 52-card deck, such as the one shown in Figure below.

1. Compute the probability of the event $E =$ "drawing a king."
2. Compute the probability of the event $E =$ "drawing a king" or $F =$ "drawing a queen" or $G =$ "drawing a jack."

ADDITION RULE FOR DISJOINT EVENTS

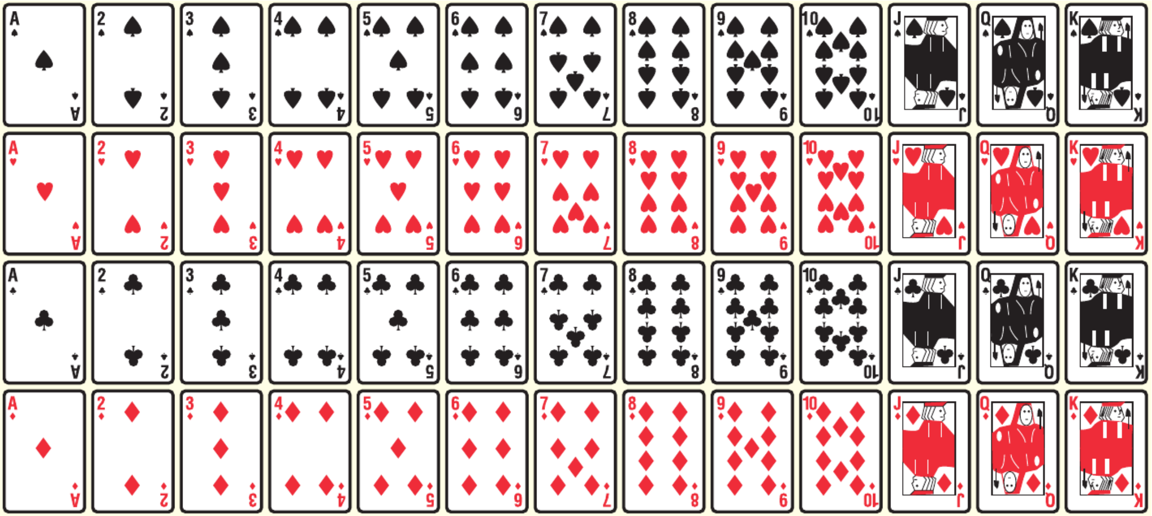
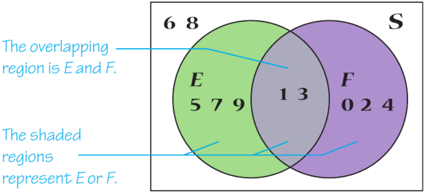


TABLE OF CONTENTS

GENERAL ADDITION RULE

- What happens when you need to compute the probability of two events that are not disjoint?
- Suppose we are randomly selecting chips from a bag.
- Each chip is labeled 0, 1, 2, 3, 4, 5, 6, 7, 8, or 9.
- Let E represent the event "choose an odd number," and let F represent the event "choose a number less than or equal to 4."
- Because $E = 1, 3, 5, 7, 9$ and $F = 0, 1, 2, 3, 4$ have the outcomes 1 and 3 in common, the events are not disjoint.
- Figure below shows a Venn diagram of the events.

GENERAL ADDITION RULE



- We can compute $P(E or F)$ directly by counting because each outcome is equally likely.
- There are 8 outcomes in E or F and 10 outcomes in the sample space, so

$$P(E \text{ or } F) = \frac{N(E \text{ or } F)}{N(S)} = \frac{8}{10} = 0.8$$

GENERAL ADDITION RULE

- Notice that using the Addition Rule for Disjoint Events to find $P(E \text{ or } F)$ would be *incorrect:*

$$P(E \text{ or } F) \neq P(E) + P(F) = \frac{5}{10} + \frac{5}{10} = 1$$

- This implies that the chips labeled 6 and 8 will never be selected, which contradicts our assumption that all the outcomes are equally likely.
- Our result is incorrect because we counted the outcomes 1 and 3 twice: once for event E and once for event F.
- To avoid this double counting, we have to subtract the probability corresponding to the overlapping region, E and F.
- That is, we have to subtract $P(E \text{ and } F) = \frac{2}{10}$ from the result and obtain the solution.

GENERAL ADDITION RULE

$$
\begin{aligned}
P(E \text{ or } F) &= P(E) + P(F) - P(E \text{ and } F) \\
&= \frac{5}{10} + \frac{5}{10} - \frac{2}{10} \\
&= \frac{8}{10} = 0.8
\end{aligned}
$$

- It agrees with the result we found by counting.
- The following rule generalizes these results.

The General Addition Rule

For any two events E and F,

$$P(E \text{ or } F) = P(E) + P(F) - P(E \text{ and } F)$$

GENERAL ADDITION RULE

EXAMPLE 3

Suppose a single card is selected from a standard 52-card deck. Compute the probability of the event $E =$ "drawing a king" or $F =$ "drawing a diamond.

- Consider the data shown in Table 6, which represent the marital status of males and females 15 years old or older in the United States in 2013..
- This table is called a contingency table or two-way table, because it relates two categories of data.
- The row variable is marital status, because each row in the table describes the marital status of an individual.

GENERAL ADDITION RULE

Table 6

	Males (in millions)	Females (in millions)
Never married	41.6	36.9
Married	64.4	63.1
Widowed	3.1	11.2
Divorced	11.0	14.4
Separated	2.4	3.2

GENERAL ADDITION RULE

- The column variable is gender.
- Each box inside the table is called a cell.
- For example, the cell corresponding to married individuals who are male is in the second row, first column. Each cell contains the frequency of the category: There were 64.4 million married males in the United States in 2013.
- Put another way, in the United States in 2013, there were 64.4 million individuals who were male and married.
- Contingency tables are very helpful in statistical studies specially in surveys as they quickly give a check on quantities.

EXAMPLE 4

Using the data in Table 6,

1. Determine the probability that a randomly selected U.S. resident 15 years old or older is male.
2. Determine the probability that a randomly selected U.S. resident 15 years old or older is widowed.
3. Determine the probability that a randomly selected U.S. resident 15 years old or older is widowed or divorced.
4. Determine the probability that a randomly selected U.S. resident 15 years old or older is male or widowed.

TABLE OF CONTENTS

PROBABILITY OF AN EVENT USING COMPLEMENT RULE

- Suppose that the probability of an event E is known and we would like to determine the probability that E does not occur.
- This can easily be accomplished using the idea of complements.

Complement of an Event

Let S denote the sample space of a probability experiment and let E denote an event. The complement of E, denoted E^c, is all outcomes in the sample space S that are not outcomes in the event E.

- Because E and E^c are mutually exclusive,

$$P(E \text{ or } E^c) = P(E) + P(E^c) = P(S) = 1$$

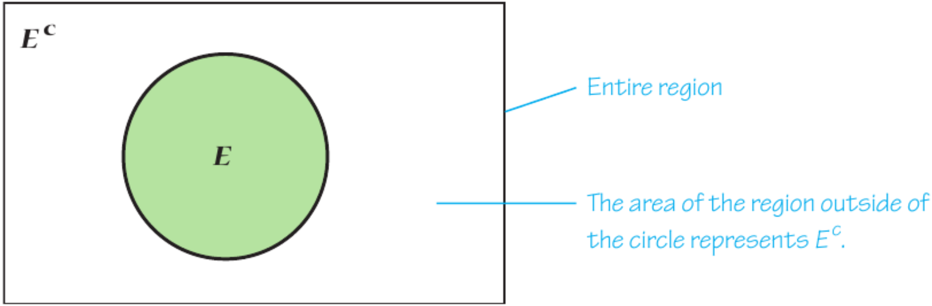
PROBABILITY OF AN EVENT USING COMPLEMENT RULE

Complement Rule

If E represents any event and E^c represents the complement of E^c, then

$$P(E^c) = 1 - P(E)$$

- Figure below illustrates the Complement Rule using a Venn diagram.



PROBABILITY OF AN EVENT USING COMPLEMENT RULE

EXAMPLE 5

According to the National Gambling Impact Study Commission, 52% of Americans have played state lotteries. What is the probability that a randomly selected American has not played a state lottery?

EXAMPLE 5

The data in Table 7 represent the income distribution of households in the United States in 2013. Compute the probability that a randomly selected household earned the following incomes in 2013:

1. \$200,000 or more
2. Less than \$200,000
3. At least \$10,000

PROBABILITY OF AN EVENT USING COMPLEMENT RULE

Table 7

Annual Income	Number (in thousands)	Annual Income	Number (in thousands)
Less than $10,000	8940	$50,000 to $74,999	21,659
$10,000 to $14,999	6693	$75,000 to $99,999	14,687
$15,000 to $24,999	13,898	$100,000 to $149,999	15,266
$25,000 to $34,999	12,756	$150,000 to $199,999	6463
$35,000 to $49,999	16,678	$200,000 or more	5913

TABLE OF CONTENTS

SUMMARY

- For disjoint events E and F, the probability of two events is obtained by adding the individual probabilities of the two events *i.e.*,

$$P(E \text{ or } F) = P(E) + P(F)$$

- For non-disjoint events E and F, the probability is computed by the formula

$$P(E \text{ or } F) = P(E) + P(F) - P(E \text{ and } F)$$

- Contingency tables are two-way tables and are very helpful in a lot of statistical studies, specially surveys.
- The probability of the complement of an event E^c given the probability of event E is computed by the relation

$$P(E^c) = 1 - P(E)$$


Thank You!