Defining Probability

August 1, 2019

Probability

When we talked about the small sample malaria experiment, what we really wanted to know was, if the independence model is correct, what is the *probability* that we'd see a difference as large as 64.3%?

- Probability forms the foundation of statistics.
- You already know about a lot of these ideas!
  - You may not have thought about them much, but you deal with probability automatically all the time.
- We are going to formalize these concepts.

Example: Rolling a Die

If you play any kind of dice-based tabletop games, you are probably familiar with weighing your options before making your next roll. This the kind of probability concept we want to formalize!

Suppose we have a six-sided die (d6). If we roll our d6 one time, what are the chances that we roll a 1?

Example: Rolling a Die

- We assume that our d6 is a fair die, so it's not weighted toward any number in particular.
- This means that all 6 numbers are equally likely.
- Therefore there is a 1-out-of-6 chance that we roll that `1`.
- When talking about probability, we write 1-out-of-6 as a fraction or decimal: $1/6 = 0.167$.
- We might also say that we have a 16.7% chance of rolling a `1`.

Example 2: Rolling a Die

Suppose we need to roll at least a 4 to succeed in some game move. What are the chances that we succeed?

Example 2: Rolling a Die

Suppose we need to roll at least a `4` to succeed in some game move. What are the chances that we succeed?

- To succeed, we can roll a `4`, `5`, or `6`.
- Our d6 has 6 sides and there are 3 numbers that result in success.
- Thus there is a 3-out-of-6 chance that we succeed, or $3/6 = 1/2 = 0.5$, a 50% chance.

Example 3: Rolling a Die

What if we are interested in rolling a `1`, `2`, `3`, `4`, `5`, or `6`?

Example 3: Rolling a Die

What if we are interested in rolling a `1`, `2`, `3`, `4`, `5`, or `6`?

- This is all of the possible sides.
- We have to roll at least one of those numbers (we cannot fail to roll a `1`, `2`, `3`, `4`, `5`, *or* `6`).
- There is a 6-out-of-6 chance that we roll one of these numbers, or $6/6 = 1$ a 100% chance.

Example 4: Rolling a Die

What if we are happy as long as we do **not** roll a 1?

Example 4: Rolling a Die

What if we are happy as long as we do **not** roll a 1?

- The chances of rolling a 1, 2, 3, 4, 5, or 6 are 100%.
- The chances of rolling a 1 are 16.7%.
- So the chances of rolling a 2, 3, 4, 5 , or 6 (but not a 1) are $100\% - 16.7\% = 83.3\%$

Example 4: Rolling a Die

What if we are happy as long as we do **not** roll a 1?

- Alternately, we can calculate this directly: not rolling a 1 means rolling a 2, 3, 4, 5, or 6.
- The chances of rolling a 2, 3, 4, 5, or 6 are 5-out-of-6, or $5/6 = 0.833$, 83.3%.

Example 5: Rolling a Die

What if we have 2d6? What is the chance that we roll two `1`s?

Example 5: Rolling a Die

What if we have 2d6? What is the chance that we roll two `1`s?

- We know that there is a $1/6$ chance that the first die is a `1`.
- Then, *of those 1/6 times*, there is a $1/6$ chance that the second die is a `1`.
- Then the chance that both dice roll a `1` is $(1/6) \times (1/6) = 1/36$ or 2.78%.

Example 5: Rolling a Die

We can also picture this in a table:

		first die					
		1	2	3	4	5	6
	1	X					
	2						
	3						
second die	4						
	5						
	6						

There are 36 possible combinations (6 sides on the first die $\times$ 6 sides on the second die) and only one of them results in two ones: $1/36$.

Probability

Whenever we mentioned the chance of something happening, we were also talking about the **probability** of something happening.

- We use probability to describe and understand **random processes** and their **outcomes**.
- In the previous examples, the random process is *rolling a die* and the outcome is *the number rolled*.

Probability

The **probability** of an outcome is the proportion of times the outcome would occur if we were able to observe the random process an infinite number of times.

Probability

- Probability is defined as a proportion and it *always takes values between 0 and 1.*
  - If you ever calculate a probability and get a number outside of 0 and 1, recalculate!
- As a percentage, it takes values between 0% and 100%.
- A probability of 0 (0%) means the outcome is impossible.
- A probability of 1 (100%) means that the outcome has to happen (all other outcomes are impossible).

Law of Large Numbers

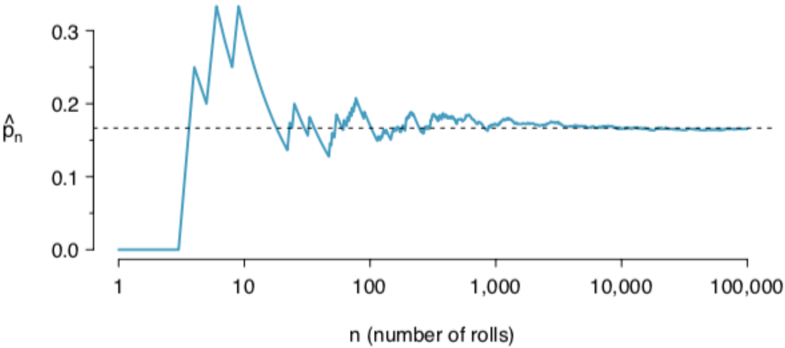
We can illustrate probability by thinking about rolling a d6 and estimating the probability that we roll a `1`.

- We estimate this probability by counting up the number of times we roll a `1` and dividing by the number of times we rolled the d6.
- Each time we roll, we recalculate and our estimate will change a little bit.
- We denote this estimate $\hat{p}_n$, where n is the number of rolls.
- We denote the true probability of rolling a `1` as $p = 1/6$.

Law of Large Numbers

- As the number of rolls, n, increases, $\hat{p}_n$ will get closer and closer to the true value of 1/6, or 16.7%.
- We say that $\hat{p}_n$ *converges* to the true probability.
- The tendency for $\hat{p}_n$ to converge to the true value as n gets large is called the **Law of Large Numbers**.
  - This is another case of more data = better information!

Law of Large Numbers



With real-world data, we usually don't get a chance to see what happens when n gets really big... but with simulations, we can see the Law of Large Numbers in action.

Probability Notation

We have some shorthand notation for talking about probabilities.

- We denote "the probability of rolling a 1" as P(rolling a 1).
- As we get more comfortable with our notation, (assuming it's clear that we're talking about rolling a die) we may shorten this further to P(1).
- So we can write

$$P(\text{rolling a } \texttt{1}) = P(1) = 1/6.$$

Random Processes

Can you think of any other random processes we might want to examine? What are the possible outcomes?

Random Processes

Here are a few random processes:

- Flipping a coin
- Wait time (in minutes) at the DMV
- How many hours of sleep you get each night

Some of these aren't completely random (the DMV is probably less crowded on, say, Tuesday mornings), but we may still want to model them based on random processes.

Disjoint Outcomes

Two outcomes are **disjoint** or **mutually exclusive** if they cannot both happen.

- If we roll our d6 only one time, we cannot roll a 1 and a 2.
  - On any single roll, the outcomes "rolling a 1" and "rolling a 2" are disjoint.
- If one of a set of disjoint outcomes happens, it is impossible that any of the others can also happen.

Disjoint outcomes

It's easy to calculate probabilities for disjoint outcomes.

- $P(\text{rolling a } \mathtt{1} \textbf{ and } \text{rolling a } \mathtt{2}) = P(\mathtt{1} \text{ and } \mathtt{2}) = 0$
  - We can roll either a 1 or a 2, but not both (on the same roll).
- $P(\mathtt{1} \textbf{ or } \mathtt{2}) = P(\mathtt{1}) + P(\mathtt{2}) = 1/6 + 1/6 = 1/3$
  - If we want to roll a 1 or 2, we have a 2-out-of-6 or $2/6 = 1/3$ chance.

Addition Rule for Disjoint outcomes

We can formalize this relationship with the **addition rule for disjoint outcomes**. Suppose A_1 and A_2 are two disjoint outcomes. Then

$$P(A_1 \text{ or } A_2) = P(A_1) + P(A_2).$$

This can be extended to many disjoint outcomes $A_1, \ldots, A_k$ where the probability that at least one of these outcomes will occur is

$$P(A_1) + P(A_2) + \cdots + P(A_k).$$

Example

Recall our contingency table for `homeownership` and `apptype`:

		`homeownership`			
		Rent	Mortgage	Own	Total
`apptype`	Individual	3496	3839	1170	8505
	Joint	362	950	183	1495
	Total	3858	4789	1353	10000

1. Are the outcomes Rent, Mortgage, and Own disjoint? Are Rent and Individual disjoint?
2. What is the probability that someone applied for a joint loan? That someone is a renter and applied for an individual loan?
3. Compute the probability that someone has a mortgage or owns their home.

Example

Are the outcomes Rent, Mortgage, and Own disjoint? Are Rent and Individual disjoint?

		homeownership			
		Rent	Mortgage	Own	Total
apptype	Individual	3496	3839	1170	8505
	Joint	362	950	183	1495
	Total	3858	4789	1353	10000

Rent, Mortgage, and Own are disjoint outcomes. Someone either rents *or* has a mortgage *or* owns their home outright.

Rent and Individual are *not* disjoint outcomes. It is possible to be a renter and apply individually for a loan.

What is the probability that someone applied for a joint loan? That someone is a renter and applied for an individual loan?

		homeownership			
		Rent	Mortgage	Own	Total
`apptype`	Individual	3496	3839	1170	8505
	Joint	362	950	183	1495
	Total	3858	4789	1353	10000

$P(\text{joint loan}) = 1495/10000 = 0.1495$ or 14.95%.

$P(\text{rent and individual loan}) = 3496/10000 = 0.3496$ or 34.96%.

Example

Compute the probability that someone has a mortgage or owns their home.

		homeownership			
		Rent	Mortgage	Own	Total
apptype	Individual	3496	3839	1170	8505
	Joint	362	950	183	1495
	Total	3858	4789	1353	10000

We decided that these are disjoint, so we use the addition rule:

$$
\begin{aligned}
P(\text{mortgage or own}) &= P(\text{mortgage}) + P(\text{own}) \\
&= (4789/10000) + (1353/10000) \\
&= 6142/10000 \text{ or } 61.42\%
\end{aligned}
$$

Sets and Events

It is common to work with **sets** of outcomes instead of individual outcomes. We call these sets **events**.

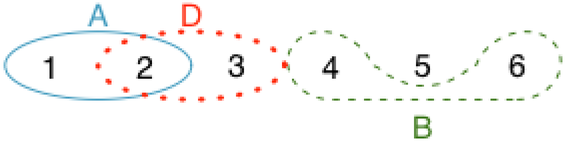
- Let A be the event that rolling a d6 results in a `1` or a `2`.
- Let B be the event that rolling a d6 results in a `4` or a `6`.
- We write these out as $A = \{1, 2\}$ and $B = \{4, 6\}$.

Since events A and B have no elements (outcomes in a set) in common, they are disjoint.

Sets and Events

Keep $A = \{1, 2\}$ and $B = \{4, 6\}$, and let D be the event that rolling a die results in a 2 or a 3 ($D = \{2, 3\}$).

Sometimes it's helpful to draw a picture when thinking about sets and probability:



Now we can see that A and B are disjoint; D and B are disjoint; but A and D are *not* disjoint.

Addition Rule for Sets?

The addition rule applies to sets in the same way that it applies to outcomes.

Keep $A = \{1, 2\}$ and $B = \{4, 6\}$. For our die, $P(A) = 1/3$ and $P(B) = 1/3$, so

$$P(A \text{ or } B) = P(A) + P(B) = (1/3) + (1/3) = 2/3$$

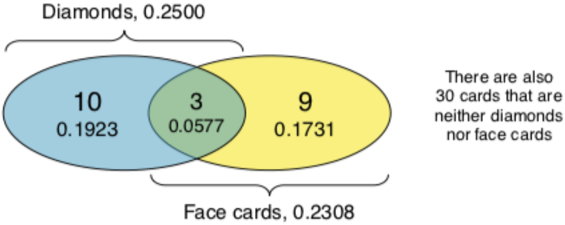
Probability for Non-Disjoint Events

We will use a standard 52 card deck to discuss disjoint events. If you are unfamiliar with the 52 card deck, it looks something like this:

2♣	3♣	4♣	5♣	6♣	7♣	8♣	9♣	10♣	J♣	Q♣	K♣	A♣
2♢	3♢	4♢	5♢	6♢	7♢	8♢	9♢	10♢	J♢	Q♢	K♢	A♢
2♡	3♡	4♡	5♡	6♡	7♡	8♡	9♡	10♡	J♡	Q♡	K♡	A♡
2♠	3♠	4♠	5♠	6♠	7♠	8♠	9♠	10♠	J♠	Q♠	K♠	A♠

- 2 colors (red and black)
- 4 suits (clubs ♣, spades ♠, diamonds ♢, and hearts ♡)
- In each suit, there are 13 cards labeled 2, 3, ..., 10, J (jack), Q (queen), K (king), A (ace).
- The cards J, Q, and K are called the "face cards".

A few slides ago, I suggested that drawing a picture might be helpful. A **Venn Diagram** is a good way to visualize the relationship between events.

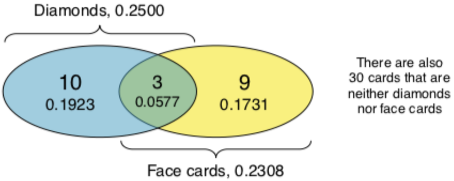


This Venn Diagram shows the events `Diamonds` and `Face Cards` as ovals. There are 3 face cards in the diamond suit, so the ovals overlap.

Probability for Non-Disjoint Events

What if we want to know the probability that a randomly selected card is a `diamond` or a `face card`?

Probability for Non-Disjoint Events

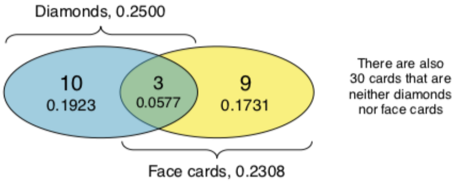


- We start by adding up the probabilities

$$P(\diamondsuit) + P(\texttt{face card}) = 13/52 + 12/52$$

- But this double counts the 3 cards in the overlap!

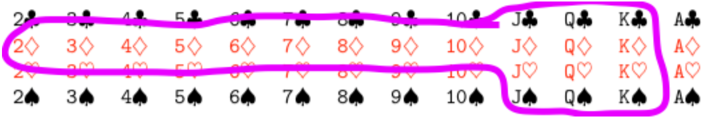
Probability for Non-Disjoint Events



- We need to correct for this double count:

$$\begin{aligned} P(\diamondsuit \text{ or } \texttt{face}) &= P(\diamondsuit) + P(\texttt{face}) - P(\diamondsuit \text{ and } \texttt{face}) \\ &= 13/52 + 12/52 - 3/52 \\ &= 22/52 \end{aligned}$$

We can also confirm that this works by checking our deck.



All of the cards that are a `diamond` or a `face card` are circled in purple. We can count that there are 22 of them. $22/52 = 0.42$ or 42%.

General Addition Rule

For any two events A and B, the probability that at least one of them will occur is

$$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$$

where $P(A \text{ and } B)$ is the probability that both events occur.

Note: In statistics, whenever we say "or" we mean "and/or". If we say that "A or B occurs", that means A, B, or both A and B occur.

General Addition Rule

By definition, for disjoint events $P(A \text{ and } B) = 0$ (they can never occur simultaneously), so the general addition rule will work for both disjoint and non-disjoint events.

Example

In the loans data set describing 10000 loans, 1495 loans were from joint applications, 4789 applicants had a mortgage, and 950 had both of these characteristics. Create a Venn diagram for this setup.

Example

Using the Venn diagram, find the probability a randomly selected loan is from a joint application where the couple had a mortgage. What is the probability that the loan had either of these attributes (joint or mortgage)?

Example

Using the Venn diagram, find the probability a randomly selected loan is from a joint application where the couple had a mortgage.

$$P(\text{joint and mortgage}) = 950/10000 = 0.095$$

Example

What is the probability that the loan had either of these attributes (joint or mortgage)?

$$
\begin{aligned}
P(\text{joint or mortgage}) &= P(\text{joint}) + P(\text{mortgage}) - P(\text{joint and mortgage}) \\
&= (1495/10000) + (4789/10000) - (950/10000) \\
&= 0.5334
\end{aligned}
$$

Probability Distributions

A **probability distribution** shows all possible (disjoint) outcomes and their corresponding probabilities.

Dice sum	2	3	4	5	6	7	8	9	10	11	12
Probability	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

This is the probability distribution for the sum of two six-sided dice.

Rules for Probability Distributions

A probability distribution is a list of the possible outcomes with corresponding probabilities that satisfies three rules:

1. The outcomes listed must be disjoint.
2. Each probability must be between 0 and 1.
3. The probabilities must sum to 1.

We can use these rules to check whether something is a valid probability distribution.

Example: Rules for Probability Distributions

Let's start by checking our sums for the two six-sided dice.

Dice sum	2	3	4	5	6	7	8	9	10	11	12
Probability	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

1. The outcomes are disjoint (the two dice can't simultaneously sum to 3 *and* 4).
2. Each probability is between 0 and 1 (the minimum is 1/36 and the maximum is 6/36).
3. The probabilities sum to 1

$$\frac{1}{36}+\frac{2}{36}+\frac{3}{36}+\frac{4}{36}+\frac{5}{36}+\frac{6}{36}+\frac{5}{36}+\frac{4}{36}+\frac{3}{36}+\frac{2}{36}+\frac{1}{36}=\frac{36}{36}$$

Example: Rules for Probability Distributions

The table below suggests 3 different distributions for household income in the US. Which one is a valid probability distribution? Why?

Income Range	\$0-25k	\$25k-50k	\$50k-100k	\$100k+
(a)	0.18	0.39	0.33	0.16
(b)	0.38	-0.27	0.52	0.37
(c)	0.28	0.27	0.29	0.16

Example: Rules for Probability Distributions

Income Range	\$0-25k	\$25k-50k	\$50k-100k	\$100k+
(a)	0.18	0.39	0.33	0.16
(b)	0.38	-0.27	0.52	0.37
(c)	0.28	0.27	0.29	0.16

Let's check (a):

1. The outcomes listed must be disjoint. (TRUE)
2. Each probability must be between 0 and 1. (TRUE)
3. **The probabilities must sum to 1.**
    - $0.18 + 0.39 + 0.33 + 0.16 = 1.16$

Example: Rules for Probability Distributions

Income Range	\$0-25k	\$25k-50k	\$50k-100k	\$100k+
(a)	0.18	0.39	0.33	0.16
(b)	0.38	-0.27	0.52	0.37
(c)	0.28	0.27	0.29	0.16

Checking (b),

1. The outcomes listed must be disjoint. (TRUE)
2. **Each probability must be between 0 and 1.**
   - $P(\$25\text{k-}50\text{k}) = -0.27$
3. The probabilities must sum to 1. (TRUE)

Example: Rules for Probability Distributions

Income Range	\$0-25k	\$25k-50k	\$50k-100k	\$100k+
(a)	0.18	0.39	0.33	0.16
(b)	0.38	-0.27	0.52	0.37
(c)	0.28	0.27	0.29	0.16

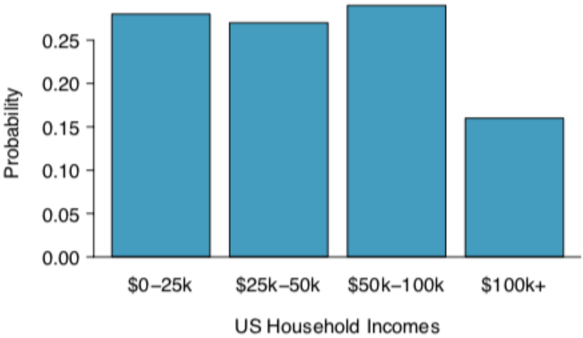
And checking (c),

1. The outcomes listed must be disjoint. (TRUE)
2. Each probability must be between 0 and 1. (TRUE)
3. The probabilities must sum to 1. (TRUE)

So (c) is our valid probability distribution.

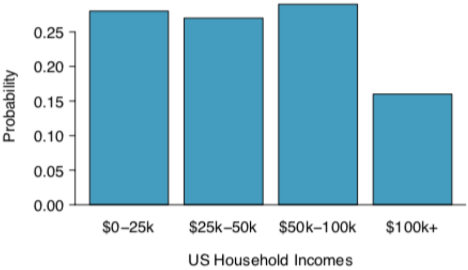
Visualizing Probability Distributions

We can visualize this probability distribution using a bar plot.



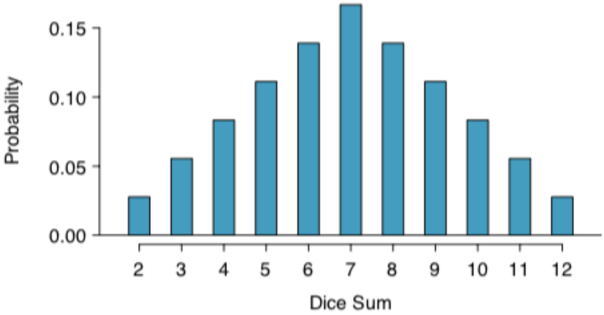
This is very similar to what we did when we created bar plots for proportion-based summary tables.

Visualizing Probability Distributions



In these bar plots, the heights of the bars represent the probabilities of each event.

Visualizing Probability Distributions



This plot shows the probability distribution for the sum of the two six-sided dice.

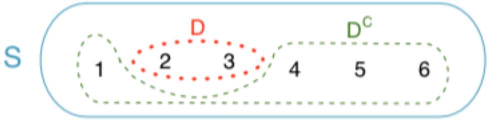
Sample Space

- Rolling our six-sided die results in some event in the set $S = \{1, 2, 3, 4, 5, 6\}$.
- We call this set our **sample space** (S).
- The sample space is defined as the set of all possible outcomes.

Complement of an Event

Let $D = \{2, 3\}$ be the event that a single roll of our d6 is a 2 or a 3.

- The **complement** of D is the set of events in the sample space that are *not* in D.
- We denote the complement by D^c.
- Then $D^c = \{1, 4, 5, 6\}$



Example: Complement of an Event

Let $D = \{2, 3\}$ be the event that a single roll of our d6 is a 2 or a 3.

Find $P(D \text{ or } D^c)$.

Example: Complement of an Event

Let $D = \{2, 3\}$ be the event that a single roll of our d6 is a 2 or a 3.

Find $P(D \text{ or } D^c)$.

- First, note that an event and it's complement are always disjoint!
- So

$$\begin{aligned} P(D \text{ or } D^c) &= P(D) + P(D^c) \\ &= (1/3) + (2/3) \\ &= 1 \end{aligned}$$

Example: Complement of an Event

Think back to possible rolls for our six-sided die. Let $A = \{1, 2\}$, the event that we roll a `1` or a `2`, and $B = \{4, 6\}$, the event of a `4` or a `6`.

1. What do A^c and B^c represent?
2. Compute $P(A^c)$ and $P(B^c)$.
3. Compute $P(A) + P(A^c)$ and $P(B) + P(B^c)$.

Properties of the Complement

- Every possible outcome not in A is in A^c, so (A or A^c) encompasses the entire sample space.
- So $P(A \text{ or } A^c)$ is the same as $P(S)$
- $P(S) = 1$, always!
  - S is all possible outcomes and there is a 100% chance that we observe at least one of the possible outcomes.

Properties of the Complement

So we can write

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

and

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

Using this relationship with the complement can help us deal with more complex probability problems down the line.

Example

Dice sum	2	3	4	5	6	7	8	9	10	11	12
Probability	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Use the probability distribution to compute the following probabilities for rolling two six-sided dice:

1. The sum of the dice is not 6.
2. The sum is at least 4. That is, determine the probability of the event $B = \{4, 5, ..., 12\}$.
3. The sum is no more than 10. That is, determine the probability of the event $D = \{2, 3, ..., 10\}$.

Example: Find $P(\text{sum not } 6)$.

We could add all of the probabilities for the sums that are not 6... or we could use the complement!

- Let $A = \{\text{not } 6\}$ be the event that the sum is not 6.
- Then $A^c = \{6\}$, the event that the sum is 6.
- Recall

$$\begin{aligned} P(A) &= 1 - P(A^c) \\ &= 1 - P(6) \\ &= 1 - \frac{5}{36} \\ &= 31/36 \end{aligned}$$

Example: Find P(sum at least 4).

Now, we want to know if the sum is *at least* 4. This means that we want to know if the sum is greater than *or equal to* 4.

- Let $B = \{4, 5, \ldots, 12\}$ be the event that the sum is at least 4.
- Then $B^c = \{2, 3\}$, the event that the sum is less than 4.

$$\begin{aligned} P(B) &= 1 - P(B^c) \\ &= 1 - P(\{2, 3\}) \\ &= 1 - [P(3) + P(2)] \\ &= 1 - \left[\frac{2}{36} + \frac{1}{36}\right] \\ &= 1 - (3/36) \\ &= 11/12 \end{aligned}$$

Example: Find P(sum no more than 10).

Now, we want to know if the sum is *no more than* 10. This means that we want to know if the sum is less than *or equal to* 10.

- Let $D = \{2, 3, \ldots, 10\}$ be the event that the sum is no more than 10.
- Then $D^c = \{11, 12\}$, the event that the sum is greater than 10.

$$\begin{aligned} P(D) &= 1 - P(D^c) \\ &= 1 - P(\{11, 12\}) \\ &= 1 - [P(11) + P(12)] \\ &= 1 - \left[\frac{2}{36} + \frac{1}{36}\right] \\ &= 1 - (3/36) \\ &= 11/12 \end{aligned}$$