Module 2: Probability

Lecture 3: Probability

Overview

Table 1: Supplementary Reading

Topics	**Relevant Ch**
Terminology, definitions, rules and axioms, LLN, Venn diagrams	OIS: 3.1.1–3.1.4
Compliment of an event, independence	OIS: 3.1.6, 3.1.7

RbE = R by Example

OIS = OpenIntro Statistics

Probability

Probability

Probability theory is the branch of mathematics concerned with probability, the analysis of random phenomena.

> *"Misunderstanding of probability may be the greatest of all impediments to scientific literacy."*

-Stephen Jay Gould

(Evolutionary Biologist and Science Historian)

Background, Motivation

- Probability is one of the most interesting and misunderstood areas of mathematics.
- An understanding of probability is fundamental in order to understand statistics (amongst other things).
- We will be assuming no previous knowledge of probability.
- Fair warning: many of you will find this section the trickiest part of the course (while other's may find it the easiest!).

Terminology

- A *random process* is a situation in which we know what outcomes *could* happen, but we don't know which particular outcome *will* happen.
- Probability is a way to quantify the uncertainty surrounding the outcome of an experiment.
- An experiment is a process by which we observe something random. Some examples:
  - tossing a coin, rolling a die, weighing a loaf of bread, obtaining blood types from a group of individuals

  An outcome is the result of an experiment. For example:
  - experiment = rolling a die,
  - outcome= number of dots facing up when it lands.

Probability

- One possible interpretation of probability defines it as a *proportion*.
- That is, the *probability* of an outcome is the proportion of times the outcome would occur if we observed the random process an infinite number of times
- Consequently, probability must always take on a value between 0 and 1 (inclusively). Alternatively, it may also be displayed as a percentage between 0% and 100%.

Law of large numbers

Theorem 1 (Law of Large Numbers, or LLN)

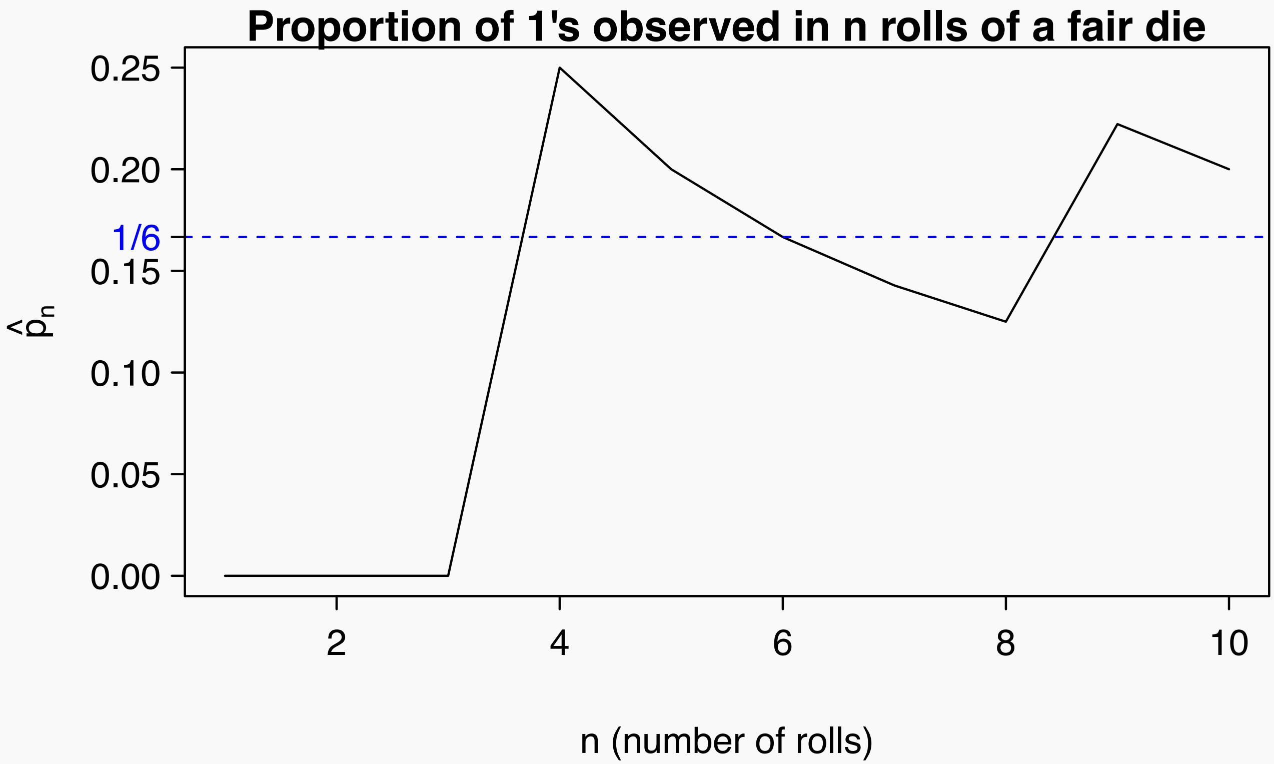
The Law of large numbers (LLN) states that as more observations are collected, the proportion of occurrences with a particular outcome, $\hat{p}_n$, converges to the probability of that outcome, p.

- Consider the example of rolling a die in order to determine p, the probability of rolling a 1.
- With the help of R, I will roll the virtual die n times and compute the proportion of 1s I observe in those n tosses, i.e.

$$\hat{p}_n = \frac{\text{number of 1s observed in } n \text{ rolls}}{n}$$

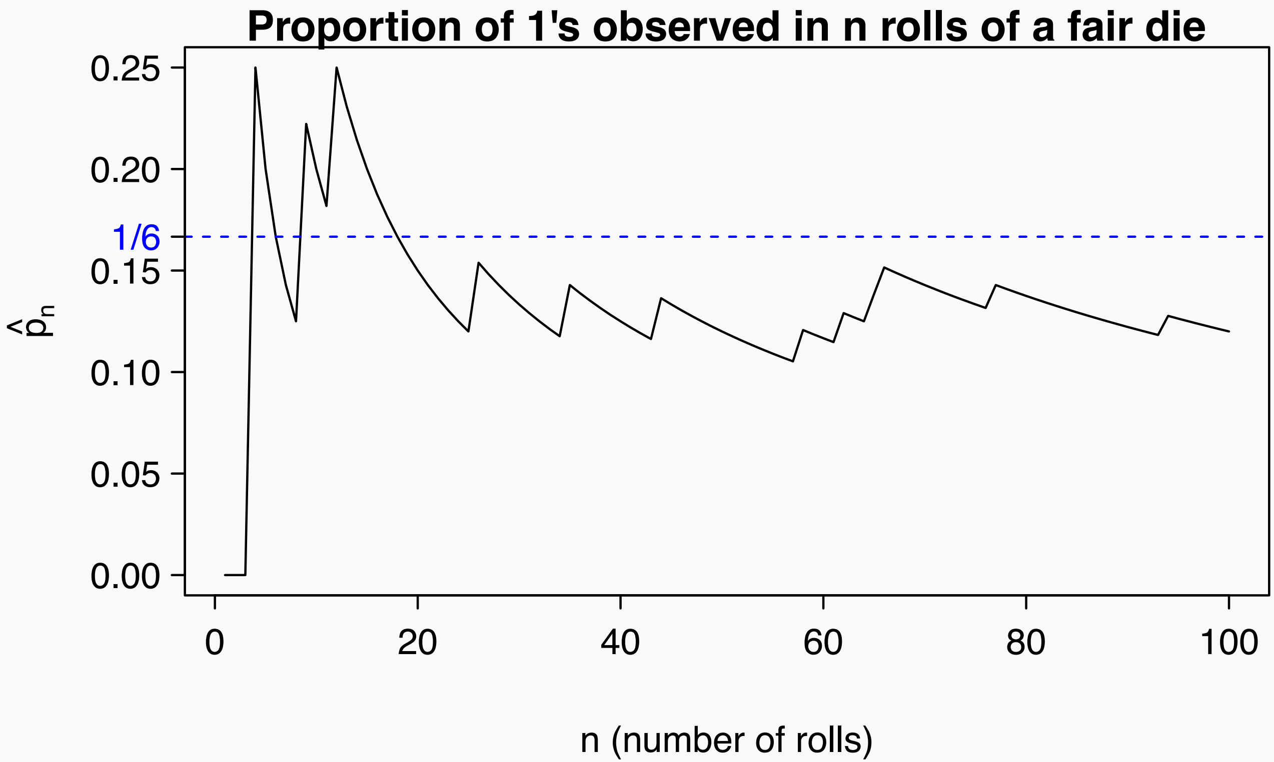
Law of Large numbers

For the first 10 rolls: 6 3 3 1 3 2 5 3 1 5



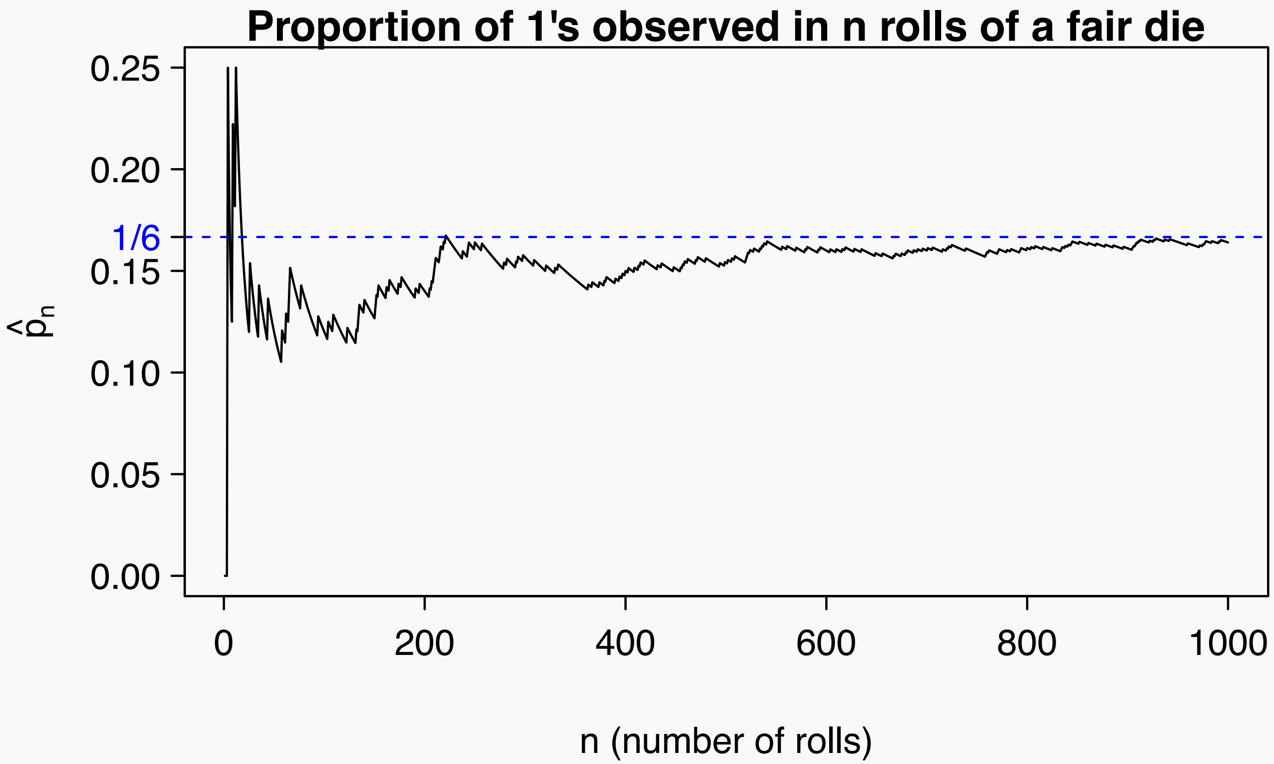
Law of Large numbers

For 100 rolls...



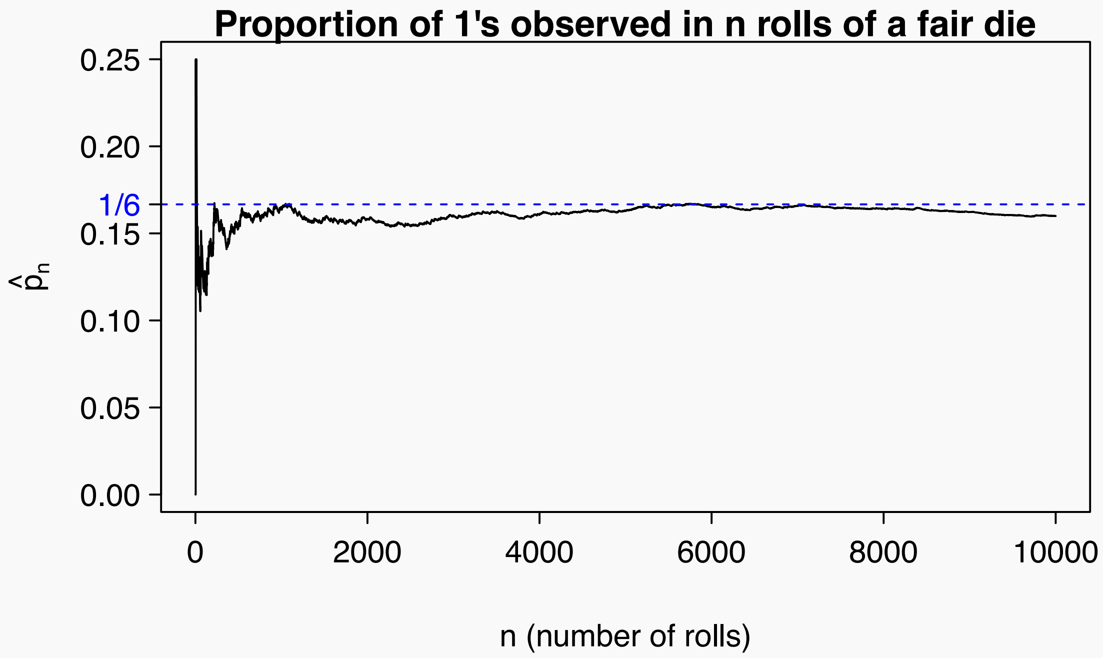
Law of Large numbers

For 1000 rolls...



Law of Large numbers

For 10000 rolls...



The proportion tends to get closer to the probability $p = 1/6$ as the number of rolls increases.

Law of large numbers (cont.)

Clicker 1

*When tossing a **fair** coin, if heads comes up on each of the first 10 tosses, what do you think the chance is that another head will come up on the next toss? 0.5, less than 0.5, or more than 0.5?*

$$\underline{H}\ \underline{H}\ \underline{H}\ \underline{H}\ \underline{H}\ \underline{H}\ \underline{H}\ \underline{H}\ \underline{H}\ \underline{H}\ \underline{?}$$

(a) < 0.5

(b) 0.5

(c) > 0.5

(d) not enough information to say

Law of large numbers (cont.)

Clicker 1

*When tossing a **fair** coin, if heads comes up on each of the first 10 tosses, what do you think the chance is that another head will come up on the next toss? 0.5, less than 0.5, or more than 0.5?*

$$\underline{H}\ \underline{H}\ \underline{H}\ \underline{H}\ \underline{H}\ \underline{H}\ \underline{H}\ \underline{H}\ \underline{H}\ \underline{H}\ \underline{?}$$

(a) < 0.5

(b) 0.5

(c) > 0.5

(d) not enough information to say

Law of large numbers (cont.)

- The probability is still 0.5, or there is still a 50% chance that another head will come up on the next toss.

$$P(H \text{ on } 11^{th} \text{ toss}) = P(T \text{ on } 11^{th} \text{ toss}) = 0.5$$

Law of large numbers (cont.)

- The probability is still 0.5, or there is still a 50% chance that another head will come up on the next toss.

$$P(H \text{ on } 11^{th} \text{ toss}) = P(T \text{ on } 11^{th} \text{ toss}) = 0.5$$

- The coin is not “due” for a tail.

Law of large numbers (cont.)

- The probability is still 0.5, or there is still a 50% chance that another head will come up on the next toss.

$$P(H \text{ on } 11^{th} \text{ toss}) = P(T \text{ on } 11^{th} \text{ toss}) = 0.5$$

- The coin is not "due" for a tail.
- The common misunderstanding of the LLN is that random processes are supposed to compensate for whatever happened in the past; this is just not true and is also called *gambler's fallacy* (or *law of averages*).

Sample Spaces and Events

We call the collection/set of all possible outcomes of an experiment the sample space, often denoted by S (or capital omega Ω).

- eg. the sample space for the roll of a die experiment is $S = \{1, 2, 3, 4, 5, 6\}$.

The sample space can contain a finite or infinite number of possible outcomes.

- For example, consider the experiment or repeatedly tossing a coin until first tail shows up. The possible outcomes are:
  - A finite number of heads followed by a tail
  - An infinite number of heads

$$S = \{H, HT, HHT, HHHT, \dots\}$$

Sample Spaces and Events

- The sample space is comprised of subsets called events.
- An event is any subset of the sample space, often denoted by the first letters of the alphabet, eg. $A, B, C, \ldots$
  - For the roll of a die, the set odd numbers $E = \{1, 3, 5\}$.
- A simple event is one that consists of exactly one outcome.
- A compound event is one that consists of more than one outcome.
- In general, exactly one simple event will occur, but many compound events can occur simultaneously.

Simple Events Example

Example 1

Consider an experiment in which each of three vehicles taking a particular highway exit either turn left (L) or right (R) at the end of the exit ramp. The eight possible outcomes that comprise the sample space are:

$$S = \{LLL, RLL, LRL, LLR, LRR, RLR, RRL, RRR\}$$

There are eight simple events:

- $E_1 = LLL$
- $E_2 = RLL,$
- $E_3 = LRL,$
- $E_4 = LLR$
- $E_2 = RRL,$
- $E_2 = RLR,$
- $E_2 = LRR,$
- $E_8 = RRR$

Compound events example

We could also define the following compound events:

- $A = \{RLL, LRL, LLR\}$ = the event that exactly one of the three vehicles turns right

Compound events example

We could also define the following compound events:

- $A = \{RLL, LRL, LLR\}$ = the event that exactly one of the three vehicles turns right
- $B = \{LLL, RLL, LRL, LLR\}$ = the event that at most one of the vehicles turns right

Compound events example

We could also define the following compound events:

- $A = \{RLL, LRL, LLR\}$ = the event that exactly one of the three vehicles turns right
- $B = \{LLL, RLL, LRL, LLR\}$ = the event that at most one of the vehicles turns right
- $C = \{LLL, RRR\}$ = the event that all three vehicles turn in the same direction

Compound events example

We could also define the following compound events:

- $A = \{RLL, LRL, LLR\}$ = the event that exactly one of the three vehicles turns right
- $B = \{LLL, RLL, LRL, LLR\}$ = the event that at most one of the vehicles turns right
- $C = \{LLL, RRR\}$ = the event that all three vehicles turn in the same direction

Suppose that all three vehicles turn left. In that case:

- the simple event $E_1 = LLL$ has occurred

Compound events example

We could also define the following compound events:

- $A = \{RLL, LRL, LLR\}$ = the event that exactly one of the three vehicles turns right
- $B = \{LLL, RLL, LRL, LLR\}$ = the event that at most one of the vehicles turns right
- $C = \{LLL, RRR\}$ = the event that all three vehicles turn in the same direction

Suppose that all three vehicles turn left. In that case:

- the simple event $E_1 = LLL$ has occurred AND
- the compound events B and C have occurred (but not A).

Rules of Probability

We denote the probability of event E occurring by $P(E)$. $P(E)$

For an experiment with possible outcomes $E_1, E_2, \ldots, E_n$, the $P(E)$
probability $P(E_i)$ must obey the following rules: $Pr(E)$

1. probabilities must be non-negative real numbers
2. the probability of the entire sample space must be equal to 1
3. If two events are disjoint[1] (i.e. they cannot happen at the same time), the probability that either of the events happens is the sum of the probabilities that each happens

These rules are often stated as Axioms ...

[1] also known as mutually exclusive

Kolmogorov axioms (AKA the Axioms of Probability)

Let S be the sample space comprised of events $\{E_1, E_2, \ldots, E_n\}$.

Kolmogorov axioms (AKA the Axioms of Probability)

Let S be the sample space comprised of events $\{E_1, E_2, \ldots, E_n\}$.

Axiom 1 (Non-negative probability)

$P(E_i) \geq 0$ *for all* E_i

Kolmogorov axioms (AKA the Axioms of Probability)

Let S be the sample space comprised of events $\{E_1, E_2, \ldots, E_n\}$.

Axiom 1 (Non-negative probability)

$P(E_i) \geq 0$ *for all* E_i

Axiom 2

$P(S) = P(E_1) + P(E_2) + \ldots + P(E_n) = 1$

Kolmogorov axioms (AKA the Axioms of Probability)

Let S be the sample space comprised of events $\{E_1, E_2, \ldots, E_n\}$.

Axiom 1 (Non-negative probability)

$P(E_i) \geq 0$ *for all* E_i

Axiom 2

$P(S) = P(E_1) + P(E_2) + \ldots + P(E_n) = 1$

Axiom 3

If $E_i \cup E_j$ *are mutually exclusive for all* $i \neq j$,

$$P(\cup_{i=1}^{n} E_i) = P(E_1) + P(E_2) + \ldots P(E_n)$$

Kolmogorov axioms (AKA the Axioms of Probability)

Let S be the sample space comprised of events $\{E_1, E_2, \ldots, E_n\}$.

Axiom 1 (Non-negative probability)

$P(E_i) \geq 0$ *for all* E_i

Axiom 2

$P(S) = P(E_1) + P(E_2) + \ldots + P(E_n) = 1$

Axiom 3

If $E_i \cup E_j$ *are mutually exclusive for all* $i \neq j$*,*

$$P(\cup_{i=1}^{n} E_i) = P(E_1) + P(E_2) + \ldots P(E_n)$$

More details on Axiom 3 to come!

Fundamental Probability

- The classical probability concept was first stimulated by the need of gamblers. What is my chance to win?
- For example, if a fair die is rolled (the experiment), what is the probability of it landing with 5 dots facing up?
- Intuitively, there are 6 equally possible outcomes and there is only 1 number 5, so the probability would be 1 out of 6 or 0.167.
- That is, if all possible outcomes of an experiment are equally, the probability of each outcomes is

$$\frac{1}{\text{total number of all possibilities}}$$

Probability for Equally Likely Outcomes

- More generally, if all possible outcomes of an experiment are equally likely and E is comprised of r of these outcomes, then
$$P(E) = \frac{r}{n}$$
where n is total number of possible outcomes.
- in set notation we often denote this using the absolute symbols, i.e. $n = |S|$.
- n is sometimes referred to as the size (also called cardinality) of the sample space
- The cardinality of a set is a measure of the number of elements in the set.
    - eg $A = \{1, 3, 5\}$ contains 3 elements, and therefore has a cardinality of 3, or $|A| = 3$, or A

Fundamental Probability

Theorem 2 (Equally likely events)

Suppose an experiment has n equally possible outcomes, if event E comprise r of these outcomes, then

$$P(E) = \frac{r}{n} = \frac{|E|}{|S|}$$

Consider the experiment of rolling a fair die. If we define event E as rolling an even number, then the $P(E)$ can be found:

$E = \{2, 4, 6\}$ $|E| = 3$

$S = \{1, 2, 3, 4, 5, 6\}$ $|S| = 6$

$$P(E) = \frac{|E|}{|S|} = \frac{r}{n} = \frac{3}{6} = \frac{1}{2}$$

$$= 0.5$$

Fundamental Probability

Theorem 2 (Equally likely events)

Suppose an experiment has n equally possible outcomes, if event E comprise r of these outcomes, then

$$P(E) = \frac{r}{n} = \frac{|E|}{|S|}$$

Consider the experiment of rolling a fair die. If we define event E as rolling an even number, then the $P(E)$ can be found:

$$S = \{1, 2, 3, 4, 5, 6\} \quad |S| = 6 = n$$

Fundamental Probability

Theorem 2 (Equally likely events)

Suppose an experiment has n equally possible outcomes, if event E comprise r of these outcomes, then

$$P(E) = \frac{r}{n} = \frac{|E|}{|S|}$$

Consider the experiment of rolling a fair die. If we define event E as rolling an even number, then the $P(E)$ can be found:

$$S = \{1, 2, 3, 4, 5, 6\} \quad |S| = 6 = n$$
$$E = \{2, 4, 6\} \quad |E| = 3 = r$$

Theorem 2 (Equally likely events)

Suppose an experiment has n equally possible outcomes, if event E comprise r of these outcomes, then

$$P(E) = \frac{r}{n} = \frac{|E|}{|S|}$$

Consider the experiment of rolling a fair die. If we define event E as rolling an even number, then the $P(E)$ can be found:

$$S = \{1, 2, 3, 4, 5, 6\} \quad |S| = 6 = n$$

$$E = \{2, 4, 6\} \quad |E| = 3 = r$$

$$P(E) = \frac{3}{6} = 0.5$$

Standard deck of cards

The following example is based on a *standard deck of cards*. A standard deck of cards is made up of the following:

- Four suites: ♡(hearts), ♣ (clubs), ♠ (spades), and ♢(diamonds).
- Each suite has thirteen cards: ace (A), 2, 3, 4, 5, 6, 7, 8, 9, 10, jack (J), queen (Q) and king (K).
- Thus the entire deck has 52 cards total.
- The 12 cards represented by the jacks, queens, and kings are called *face cards*.
- Hearts (♡) and diamonds (♢) are typically coloured red, while clubs (♣) and spade (♠) are typically coloured black.
- Unless otherwise specified, you may assume the deck is well shuffled

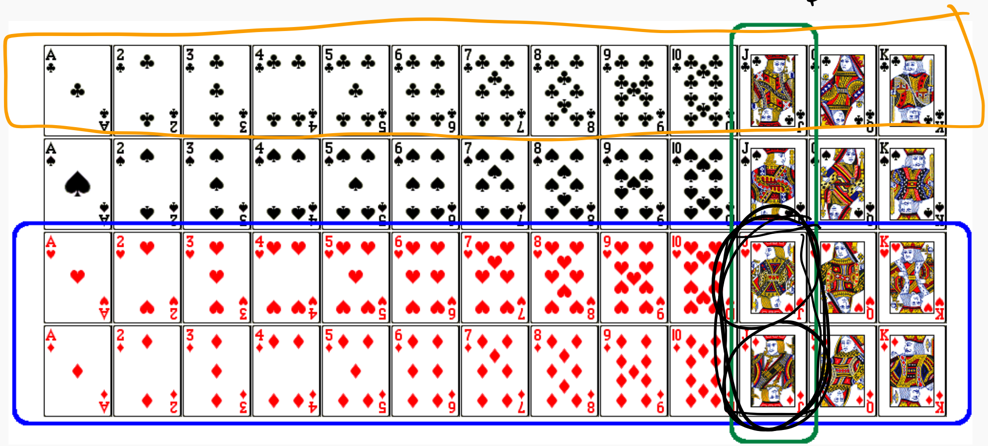
Example 2

What is the probability of drawing a jack or a red card from a well shuffled full deck?

Ex 3
$P(\text{red or club})$

B

C



A

$P(\text{red}) + P(\text{club}) - 0$

$= \frac{26}{52} + \frac{13}{52}$

$= \frac{3}{4}$

We *could* write down all the elements in this event (being sure not to count the same outcome twice) and use Theorem 2 ...

Figure from *http://www.milefoot.com/math/discrete/counting/cardfreq.htm*.

Union of non-disjoint events

Let A be the event of selecting black card or a jack.

$$A = \{A\heartsuit, 2\heartsuit, 3\heartsuit, 4\heartsuit, 5\heartsuit, 6\heartsuit, 7\heartsuit, 8\heartsuit, 9\heartsuit, 10\heartsuit, J\heartsuit, Q\heartsuit, K\heartsuit,$$
$$A\diamondsuit, 2\diamondsuit, 3\diamondsuit, 4\diamondsuit, 5\diamondsuit, 6\diamondsuit, 7\diamondsuit, 8\diamondsuit, 9\diamondsuit, 10\diamondsuit, J\diamondsuit, Q\diamondsuit, K\diamondsuit, J\clubsuit, J\spadesuit\} \quad (1)$$

Note: being sure not to count the $J\diamondsuit, J\heartsuit$ twice!

There are $n = 52$ possible outcomes and there are $r = 28$ that satisfy event A, so

round(28/52, 4)

$$P(A) = \frac{r}{n} = \frac{28}{52} = \boxed{0.5384}615 \approx 0.5385$$

... but this takes a considerable amount of time and care

Addition Rule

Instead we might make use of the general addition rule:

Rule 1 (General addition rule)

$$\frac{26}{52} + \frac{4}{52} - \frac{2}{52} = \frac{28}{52}$$

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

- When A and B, are *mutually exclusive*, the probability that A or B will occur is the sum of the probability of each event.
- That is to say, for *disjoint events* $P(A \text{ and } B) = 0$, and the above formula simplifies to $P(A \text{ or } B) = P(A) + P(B)$.

Disjoint and non-disjoint outcomes

Disjoint (AKA mutually exclusive) outcomes: Cannot happen at the same time.

- The outcome of a single coin toss cannot be a head and a tail.
- A student both cannot fail and pass a class.
- A single card drawn from a deck cannot be an ace and a queen.

Disjoint and non-disjoint outcomes

Disjoint (AKA mutually exclusive) outcomes: Cannot happen at the same time.

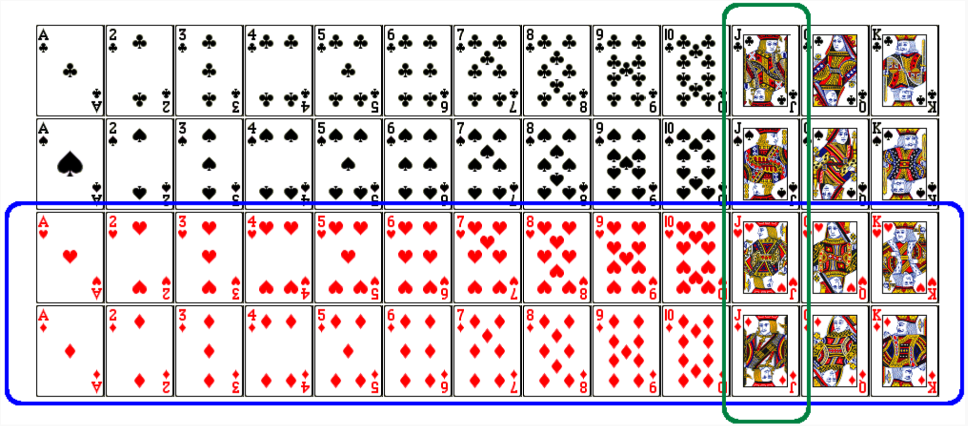
- The outcome of a single coin toss cannot be a head and a tail.
- A student both cannot fail and pass a class.
- A single card drawn from a deck cannot be an ace and a queen.

Non-disjoint outcomes: Can happen at the same time.

- A student can get an A in Stats and A in Econ in the same semester.

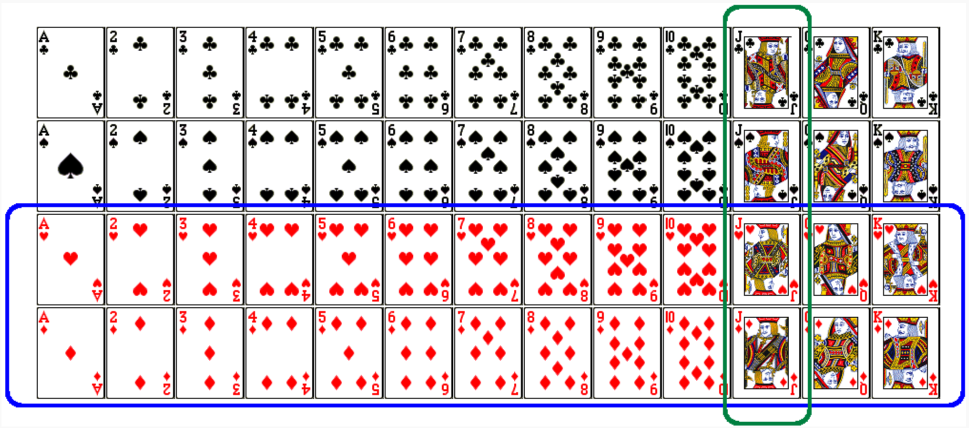
Union of non-disjoint events

Let's return to Example 2 and use Rule 1 to determine the probability of drawing a jack or a red card from a well shuffled deck



Union of non-disjoint events

Let's return to Example 2 and use Rule 1 to determine the probability of drawing a jack or a red card from a well shuffled deck



$$P(jack\ or\ red) = P(jack) + P(red) - P(jack\ and\ red)$$

$$= \frac{4}{52} + \frac{26}{52} - \frac{2}{52} = \frac{28}{52}$$

Figure from *http://www.milefoot.com/math/discrete/counting/cardfreq.htm*.

Clicker 2

What is the probability that a randomly sampled student is double vaccinated or they agree with their parents' political views? or both

	A	B	C
	Share Parents' Politics		
Double vaccinated	No	Yes	Total
1 No	11	40	51
2 Yes	36	78	114
3 Total	47	118	$\boxed{165}$

(a) $\frac{40+36-78}{165}$

(b) $\frac{114+118-78}{165}$

(c) $\frac{78}{165}$

(d) $\frac{78}{188}$

(e) $\frac{11}{47}$

A = that the student shares parents' political view

B = the student is double vaccinated

THERE IS OVERLAP !! (A and B are NOT mutually ex.)

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

$= \frac{118}{165} + \frac{114}{165} - \frac{78}{165} =$

Clicker 2

What is the probability that a randomly sampled student is double vaccinated <u>or</u> they agree with their parents' political views?

	Share Parents' Politics		
Double vaccinated	No	Yes	Total
No	11	40	51
Yes	36	78	114
Total	47	118	165

(a) $\frac{40+36-78}{165}$

(b) $\frac{114+118-78}{165}$

(c) $\frac{78}{165}$

(d) $\frac{78}{188}$

(e) $\frac{11}{47}$

Multiple Events: Mutually Exclusive

- What would happen in more complicated situations? Eg, what if we were rolling 3 die? Would we have to write out all of the possible outcomes?

- What if we rolled a die and chose a card — imagine writing out all of those outcomes ($6 \times 52 = 312$ in all)!

- Thankfully, we do not need to write out all of these outcomes — there are rules we can use... (more on this next lecture)

- ... but first some more definitions and concepts.

Venn diagrams

- A *Venn diagram* uses overlapping circles to show the logical relationship between two or more sets. Let's demonstrate its primary use in statistics with an example...

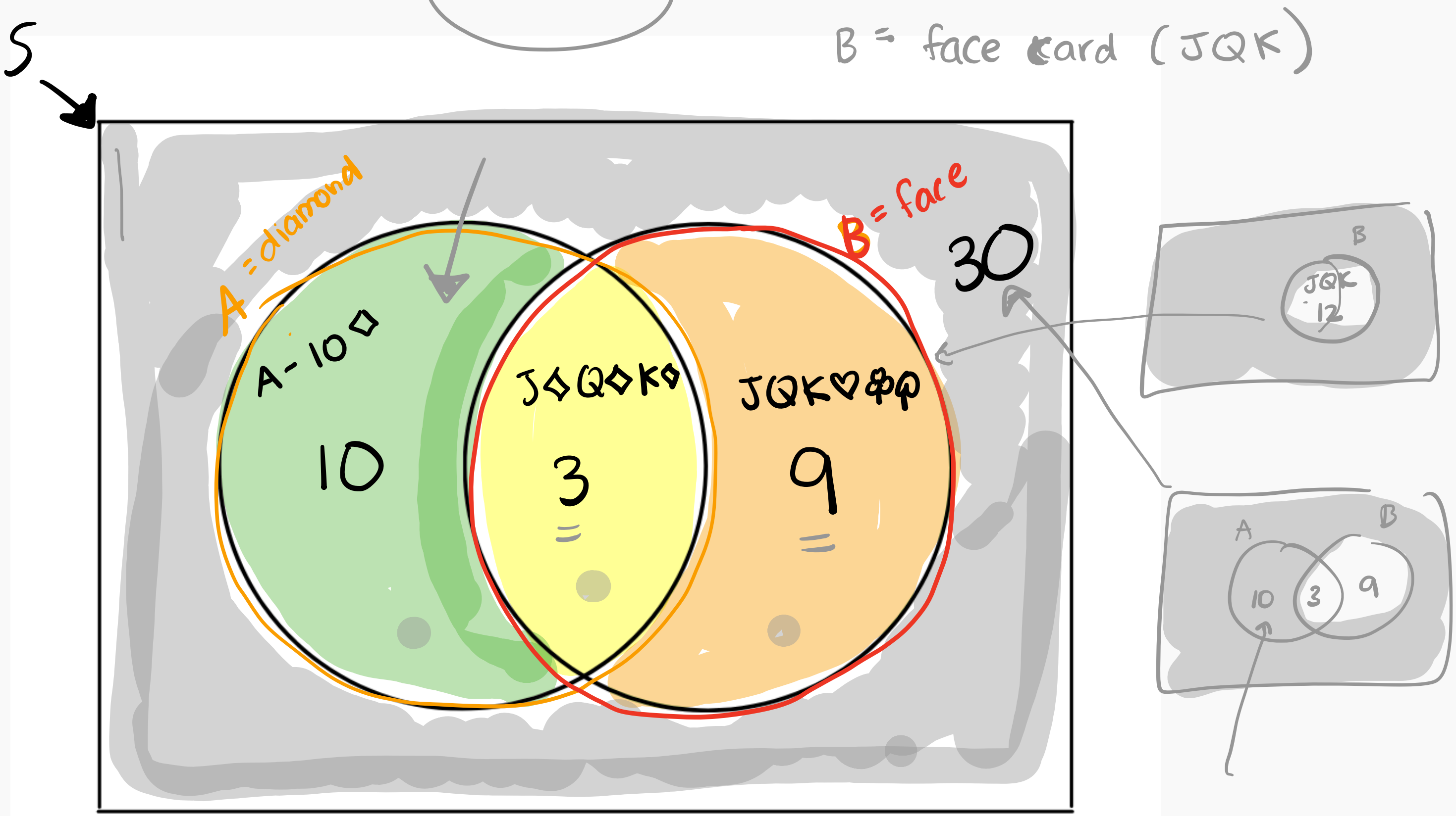
Example 3

Consider selecting a random card from a standard deck of cards.
Let A be the event that the card is a diamond.
Let B be the event that the card is a face card (ie. J, Q, K)

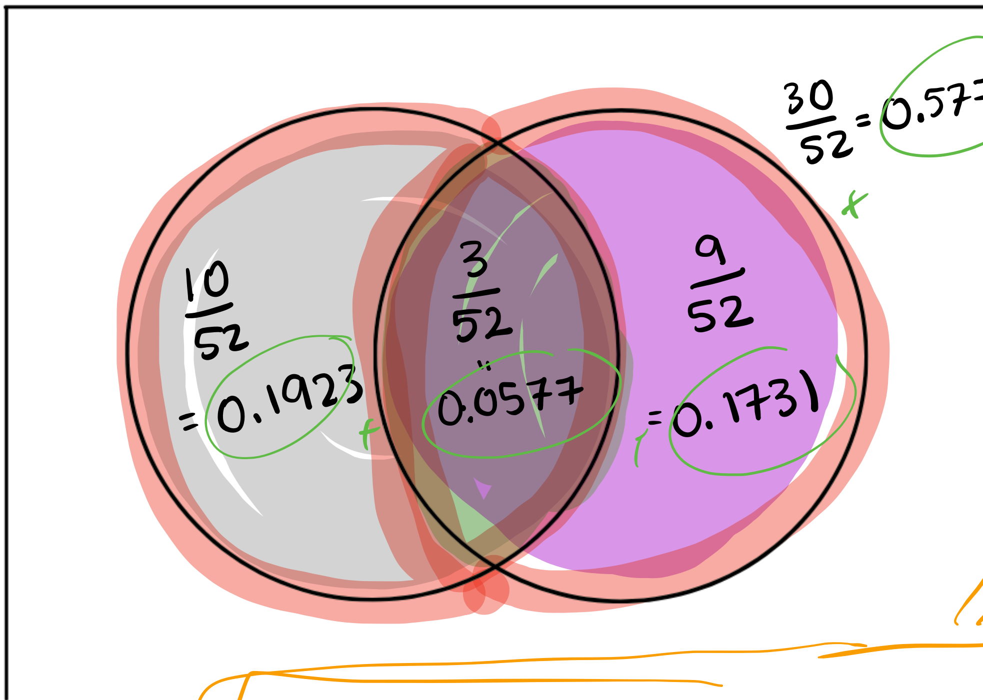
- We can use *Venn diagrams* to visualize the probability of these events.

Venn diagram's were popularized by John Venn in the 1880s.

Venn Diagrams using Counts



Venn Diagrams using Probabilities



$= 1.0001$

$P(A \cup B) = P(A) + P(B) - P(A \cap B)$

$= \frac{10}{52} + \frac{3}{52} + \cancel{\frac{3}{52}} + \frac{9}{52}$

$- \cancel{\frac{3}{52}}$

$P(A \cup B)$

$= \frac{10}{52} + \frac{3}{52} + \frac{9}{52}$

Venn Diagrams

- More commonly in statistics you will see Venn Diagrams begin used with probabilities rather than with counts.
- If we add up all the counts in the first Venn Diagram they should add to 52 (total number of cards in a deck)
- If we add up all the probabilities in the first Venn Diagram they should add to 1 *
- Since Venn diagram's represents the entire sample space for two events, A and B, the probabilities should always add to 1 (or with counts, the total should sum to $|S|$) to satisfy Axiom 3

---

*Note: * if you sum up the rounded probabilities it actually works out to be 1.001 due to rounding error!*

Set Theory

Often events are combinations of two or more events formed by taking unions, intersections, and complements.

These terms are borrowed from set theory but it might help you interchange these terms with the following:

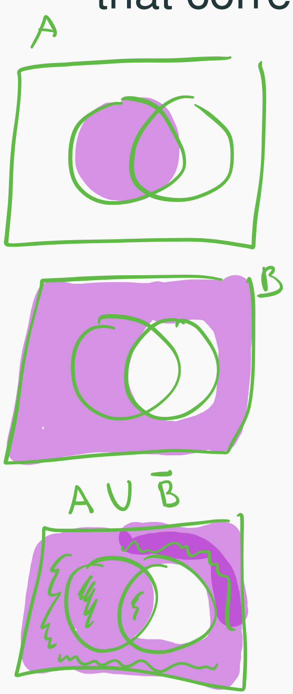
- *union* with *or*, $\cup$
- *intersection* with *and* $\cap$
- *compliment* with *not* $\bar{A}$ A^c

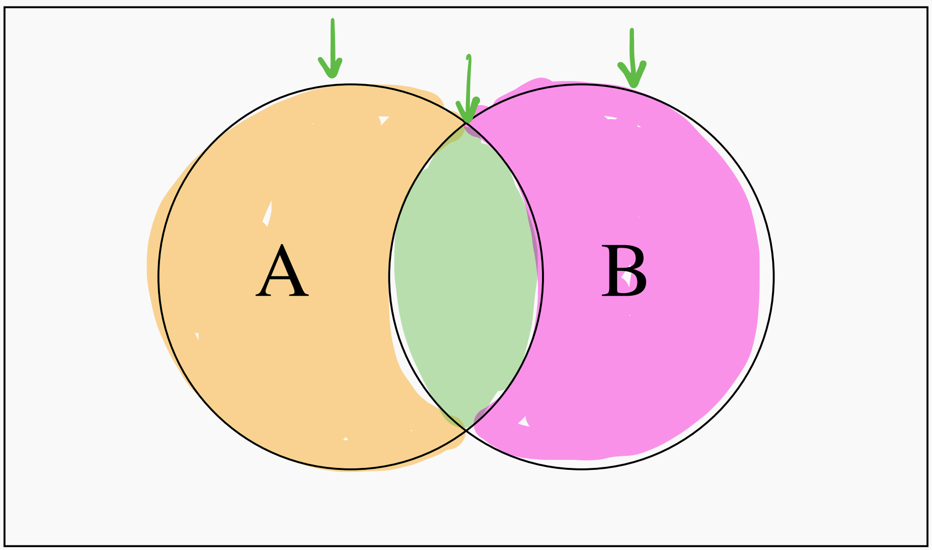
Tip 1 ("or" is inclusive)

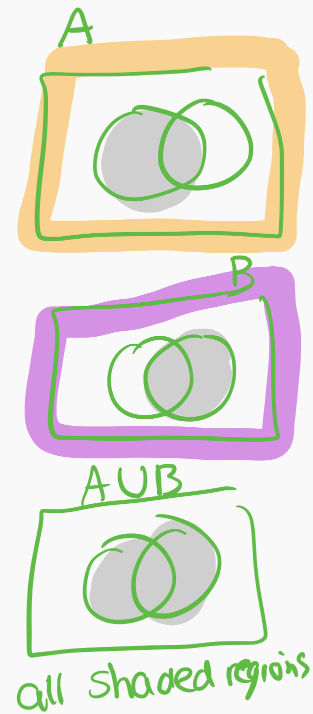
When we use "or" in statistics, we mean "and/or" unless we explicitly state otherwise. Thus, "A or B" occurring means, either A, B, or both A and B occurred. $A \cup B$

Unions

The *union of A and B*, denoted $\boxed{A \cup B}$ and read “A or B”, is the set of all elements from A and/or B. Thus, $A \cup B$ means A, B, or both A and B occur. Shade in the appropriate area in the Venn diagram that corresponds to $A \cup B$:







Unions

Let's return to Example 3

Clicker

Example 3 (cont'd)

Consider selecting a random card from a standard deck of cards.

Let A be the event that the card is a diamond.

Let B be the event that the card is a face card (ie. J, Q, K)

In words, the $P(A \cup B)$ is the probability that the random card selected is

a) a diamond and face card (J$\diamondsuit$, Q$\diamondsuit$, K$\diamondsuit$)
b) a diamond but not a face card (A, 2, . . . , 9, 10 of $\diamondsuit$)
c) either a diamond or a face card or both
d) either a diamond or a face card but not both
e) none of the above

Unions

Let's return to Example 3

Example 3 (cont'd)

Consider selecting a random card from a standard deck of cards.

Let A be the event that the card is a diamond.

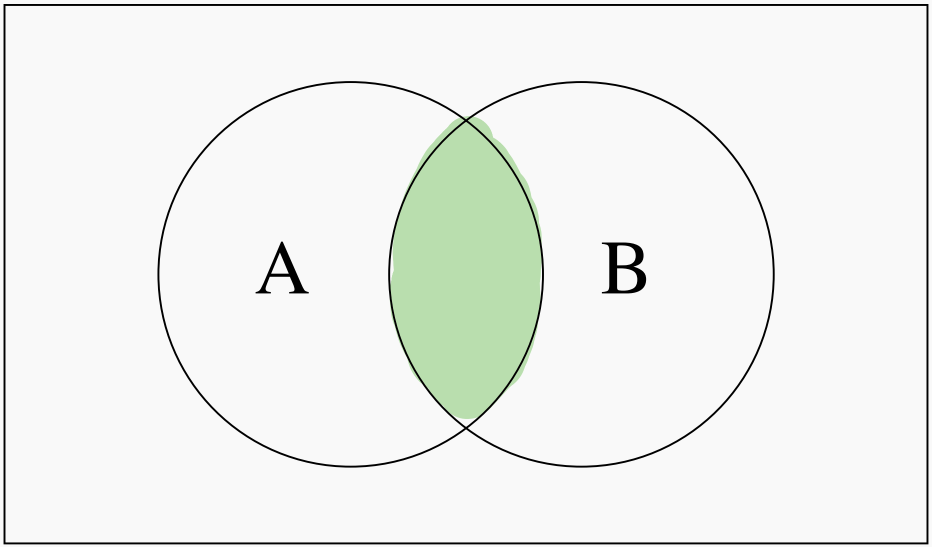
Let B be the event that the card is a face card (ie. J, Q, K)

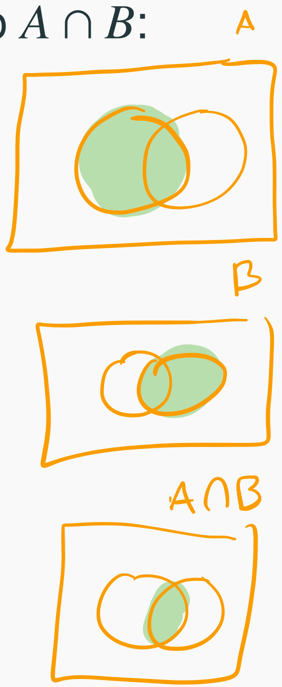
In words, the $P(A \cup B)$ is the probability that the random card selected is

a) a diamond and face card (J$\diamondsuit$, Q$\diamondsuit$, K$\diamondsuit$)
b) a diamond but not a face card (A, 2, . . . , 9, 10 of $\diamondsuit$)
c) either a diamond or a face card or both
d) either a diamond or a face card but not both
e) none of the above

Intersections

The *intersection of A and B* denoted by $A \cap B$ and read "A and B', is the set of all elements that A and B have in common. Shade in the appropriate area in the Venn diagram that corresponds to $A \cap B$:





Intersections

Let's return to Example 3

Example 3 (cont'd)

Consider selecting a random card from a standard deck of cards.

Let A be the event that the card is a diamond.

Let B be the event that the card is a face card (ie. J, Q, K)

In words, the $P(A \cap B)$ is the probability that the random card selected is

a) a diamond and face card (J$\diamondsuit$, Q$\diamondsuit$, K$\diamondsuit$)
b) a diamond but not a face card (A, 2, . . . , 9, 10 of $\diamondsuit$)
c) either a diamond or a face card or both
d) either a diamond or a face card but not both
e) none of the above

Intersections

Let's return to Example 3

Example 3 (cont'd)

Consider selecting a random card from a standard deck of cards.

Let A be the event that the card is a diamond.

Let B be the event that the card is a face card (ie. J, Q, K)

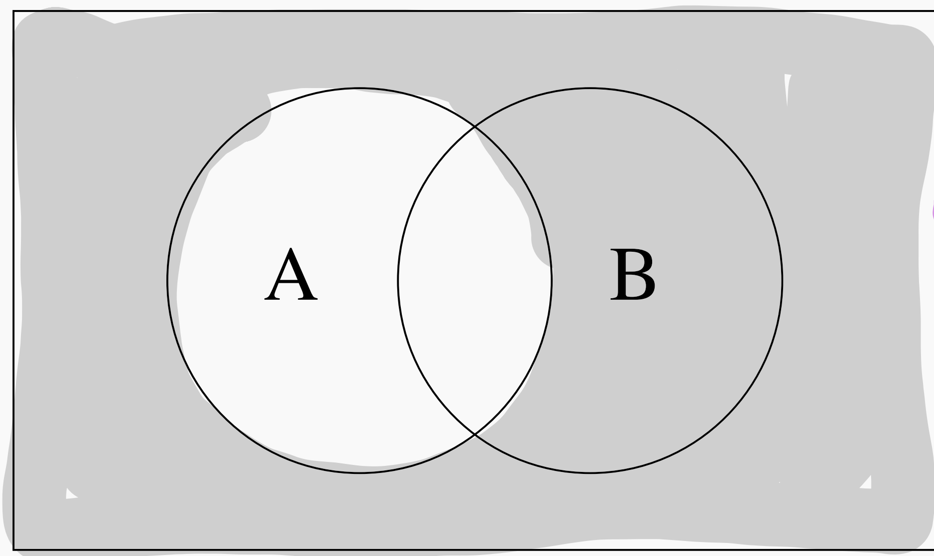
In words, the $P(A \cap B)$ is the probability that the random card selected is

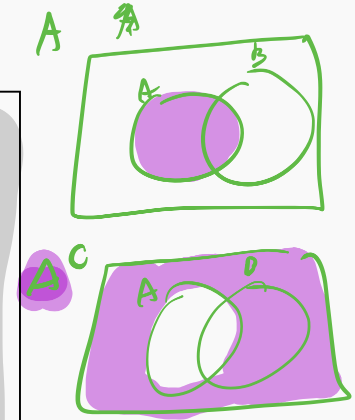
a) a diamond and face card (J$\diamondsuit$, Q$\diamondsuit$, K$\diamondsuit$)
b) a diamond but not a face card (A, 2, . . . , 9, 10 of $\diamondsuit$)
c) either a diamond or a face card or both
d) either a diamond or a face card but not both
e) none of the above

Compliments

The *complement of A*, denoted $\overline{A}$ *or* A^c and read “A complement” or “not A”, is all elements in S that are not in A. Naturally, the $P(A) + P(A^c)$ should be 1. Shade in the appropriate area in the Venn diagram that corresponds to A^c:

A^c



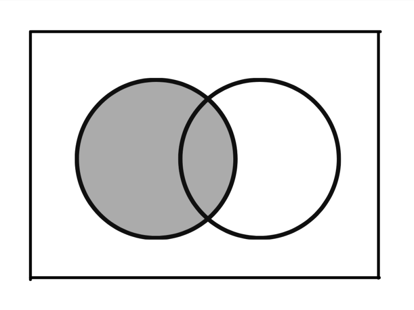
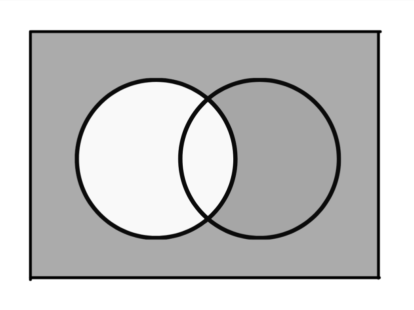
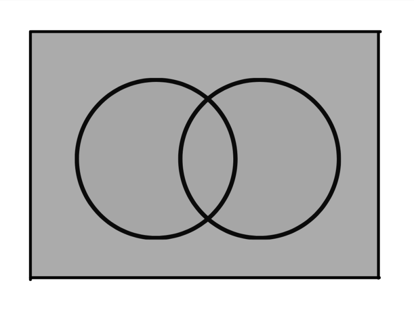


Compliment Rule

By the rule of total probability:

$$P(A) + P(\overline{A}) = 1$$

$P(S) = 1$


+

=


Compliments

Let's return to Example 3

Example 3 (cont'd)

Consider selecting a random card from a standard deck of cards.

Let A be the event that the card is a diamond.

Let B be the event that the card is a face card (ie. J, Q, K)

In words, the $P(\overline{B})$ is the probability that the random card selected is

probability B "not" ~~of~~ Prob B "compliment"

a) either a spade (♠) or a clubs (♣)
b) either a spade (♠) or a clubs (♣) or a heart (♡)
c) either a A, 2, . . . , 9, 10 of diamonds (♢)
d) none of the above

Compliments

Let's return to Example 3

Example 3 (cont'd)

Consider selecting a random card from a standard deck of cards.

Let A be the event that the card is a diamond.

Let B be the event that the card is a face card (ie. J, Q, K)

In words, the $P(\overline{B})$ is the probability that the random card selected is

a) either a spade (♠) or a clubs (♣)
b) either a spade (♠) or a clubs (♣) or a heart (♡)
c) either a A, 2, . . . , 9, 10 of diamonds (♢)
d) none of the above

either a A, 2, . . . , 9, 10 (i.e. non-face card) of *any* suit (♢, ♡, ♠,♣)

Venn Diagrams for more than two events

Example 4 (Venn Diagrams for Multiple Events)

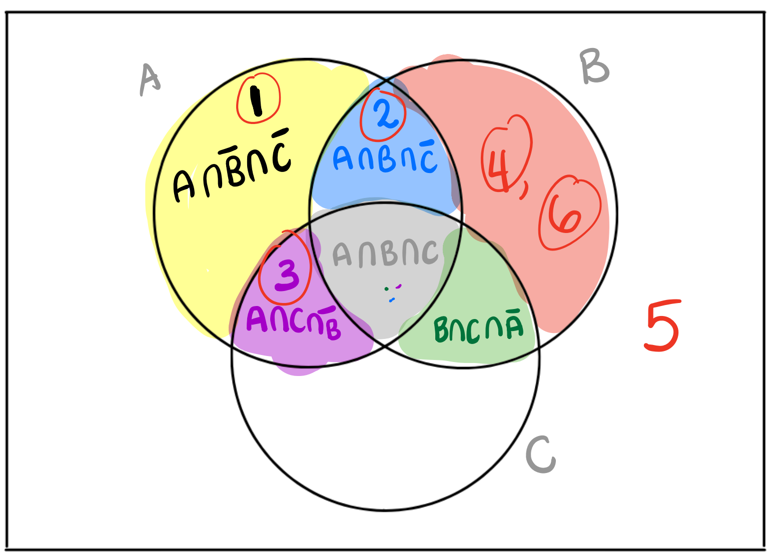
Consider rolling a fair die. Draw the Venn Diagram where:

A is the event of rolling a one, two, or three: $A = \{1, 2, 3\}$

B is the event of rolling an even number: $B = \{2, 4, 6\}$

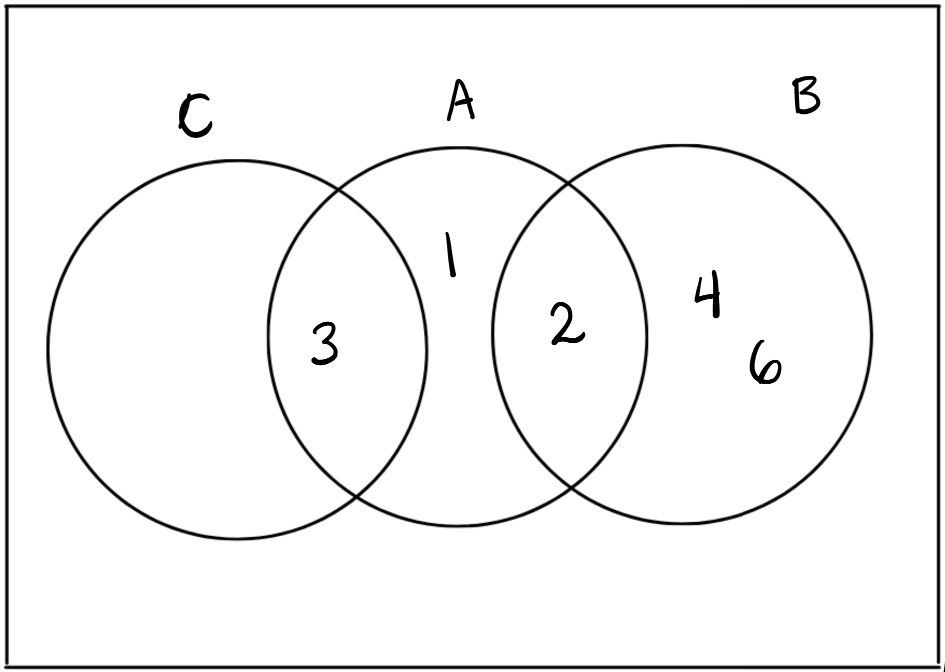
C is the event of rolling a 3: $C = \{3\}$.

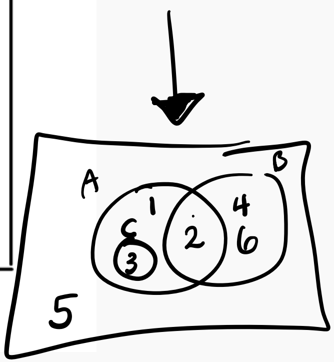
$S = \{\cancel{1}, \cancel{2}, \cancel{3}, \cancel{4}, 5, \cancel{6}\}$



Venn Diagrams for more than two events

$$A \cap B \cap C = \{\} = \emptyset$$





Mutually Exclusive/Disjoint Events

- Sometimes two events will have no outcomes in common.
- Two events are said to be *disjoint* or *mutually exclusive* if they have no elements in common *i.e.* the events cannot occur at the same time.
- In that case, we use $\emptyset$ to denote the null event/empty set (i.e the event consisting of no outcomes whatsoever).
- For example, B = $\{2, 4, 6\}$ and C = $\{3\}$ have no elements in common/we cannot roll a number that is both even and a 3. Therefore B and C are mutually exclusive and $B \cap C = \emptyset$.

Mutually Exclusive/Disjoint Events

Definition 1 (mutually exclusive)

When $A \cap B = \emptyset$, A and B are said to be mutually exclusive or disjoint events.

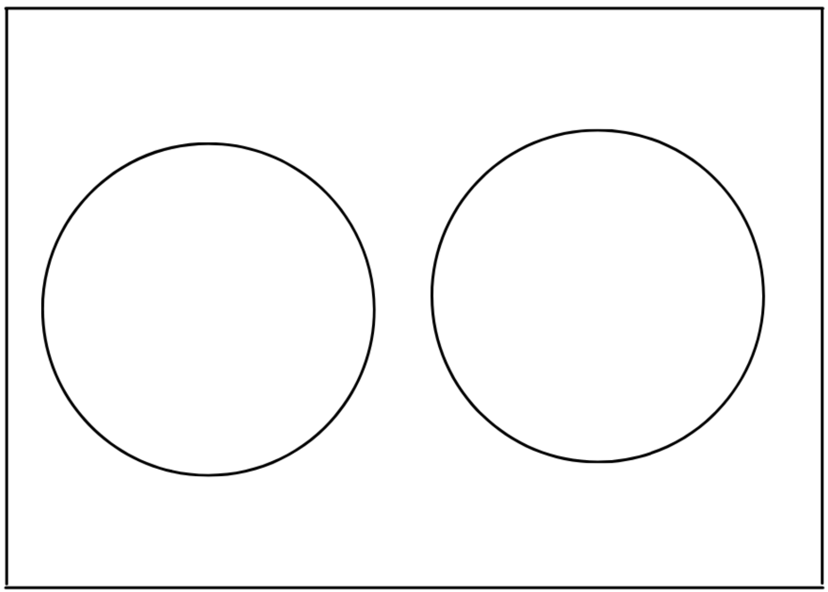
How might mutually exclusive events look on a Venn Diagram?

Mutually Exclusive/Disjoint Events

Definition 1 (mutually exclusive)

When $A \cap B = \emptyset$, A and B are said to be mutually exclusive or disjoint events.

How might mutually exclusive events look on a Venn Diagram?



Summary of Concepts

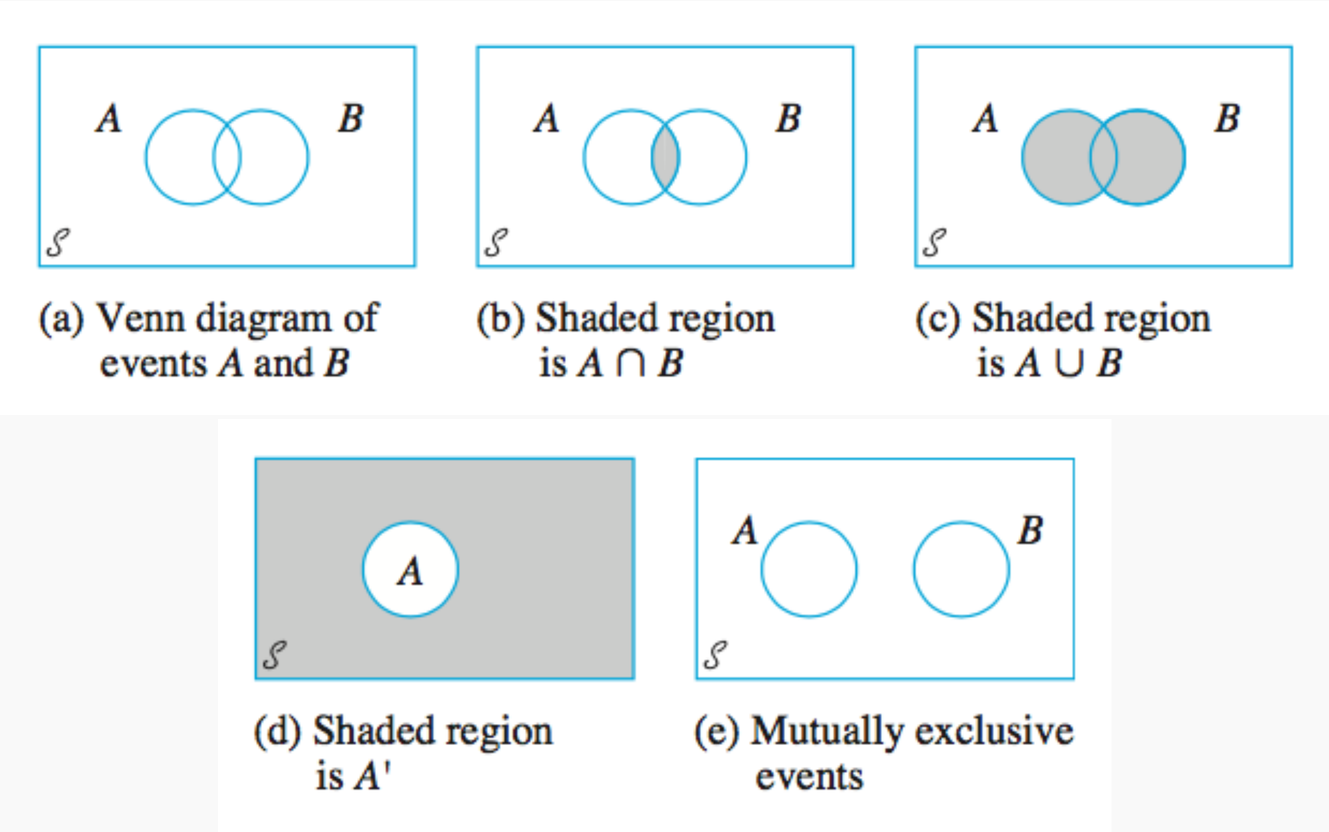


Figure 1: Figure 2.1 Venn diagrams from Devore

Subset

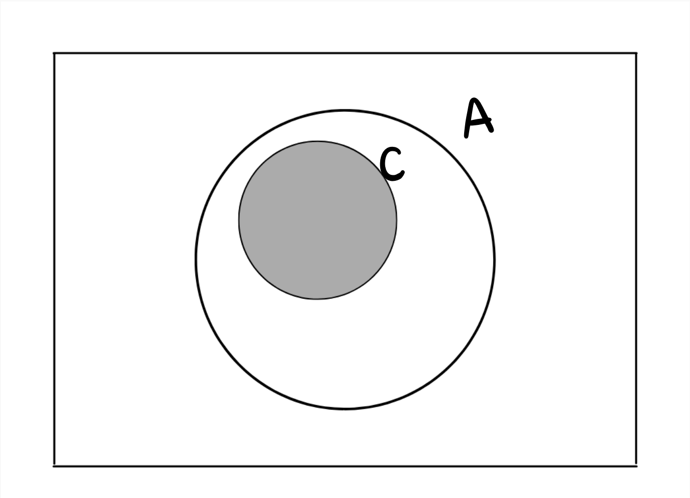
- C is *contained in A*, denoted $C \subseteq A$, read "C subset A", if all elements in C are elements in A, ie $A \subseteq S$ for all events A.
- Thus *any* event (ie any set of outcomes of an experiment) is a subset of the sample space (and any subset of the sample space is an event) S

How might this look like on a Venn Diagram?See this video for how does independence look likes.

Subset

- C is *contained in A*, denoted $C \subseteq A$, read "C subset A", if all elements in C are elements in A, ie $A \subseteq S$ for all events A.
- Thus *any* event (ie any set of outcomes of an experiment) is a subset of the sample space (and any subset of the sample space is an event)

How might this look like on a Venn Diagram?See this video for how does independence look likes.



Clicker 3

What does the following shaded region in the Venn diagram correspond to?

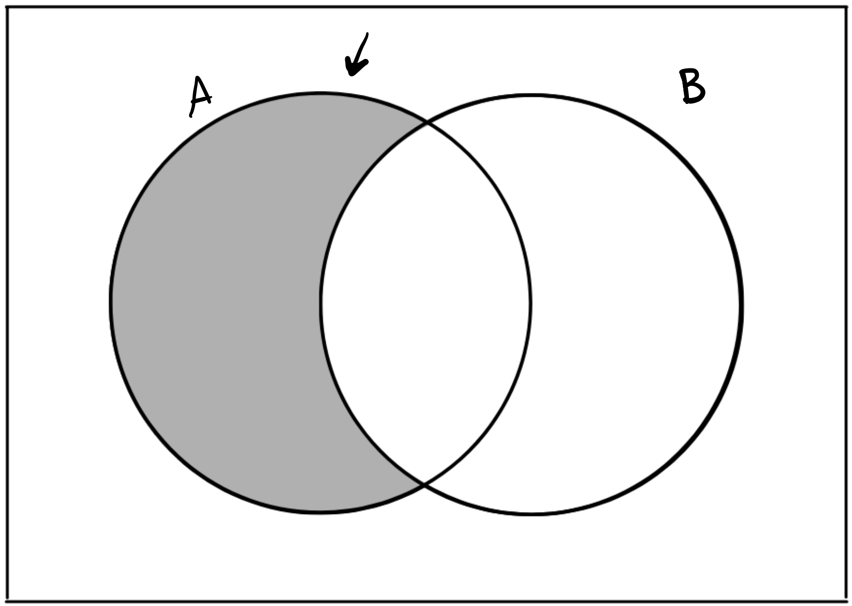
a) $P(A \cap B)$

b) $P(A \cup B)$

c) $P(A \cap \overline{B})$

d) $P(\overline{A} \cup B)$

e) none of the above



Clicker 3

What does the following shaded region in the Venn diagram correspond to?

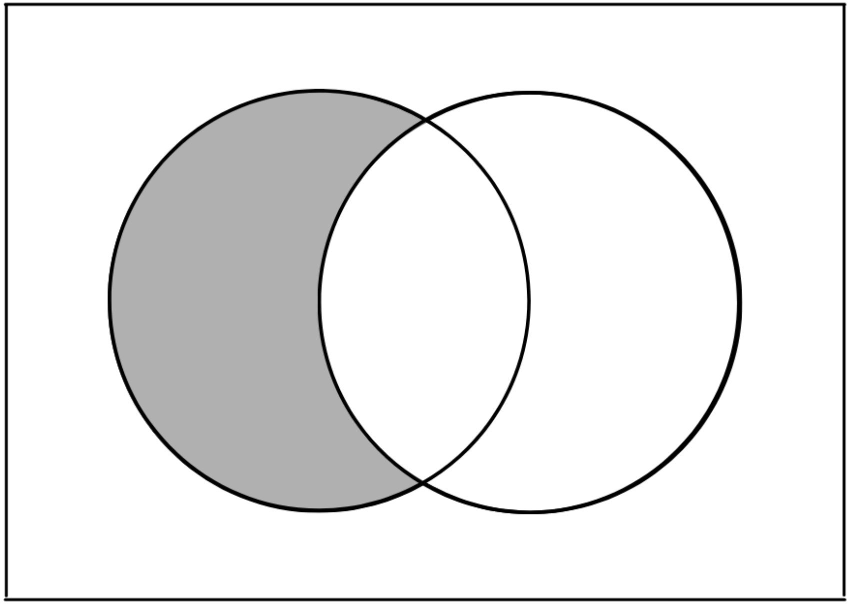
a) $P(A \cap B)$

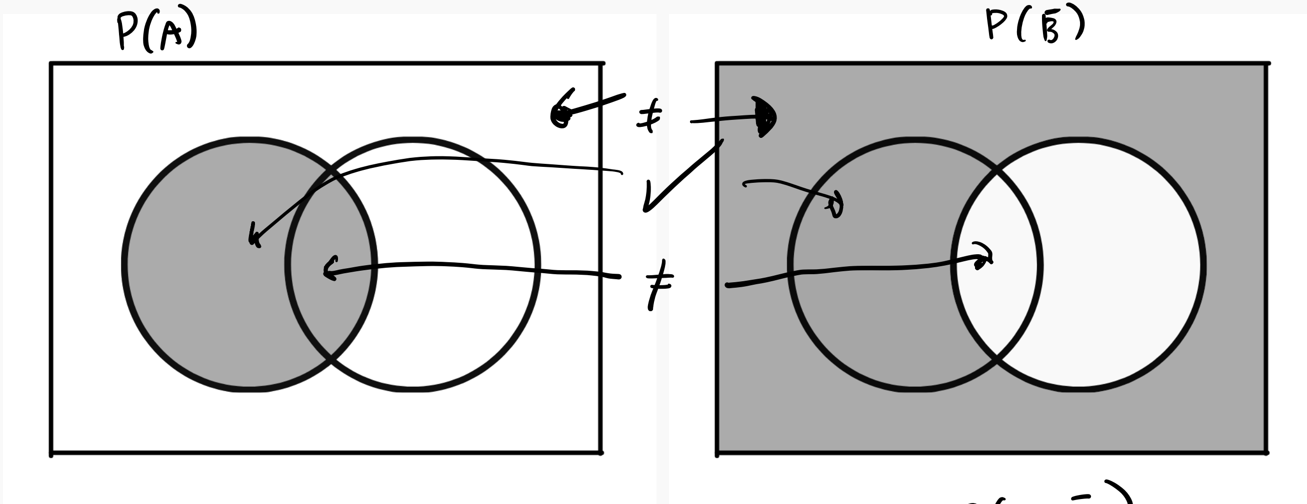
b) $P(A \cup B)$

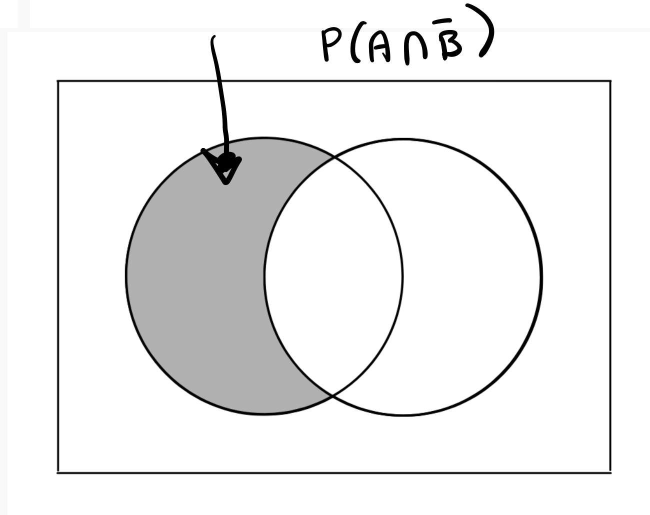
c) $P(A \cap \overline{B})$

d) $P(\overline{A} \cup B)$

e) none of the above







Select *common* shaded region

Clicker 4

What does the following shaded region in the Venn diagram correspond to?

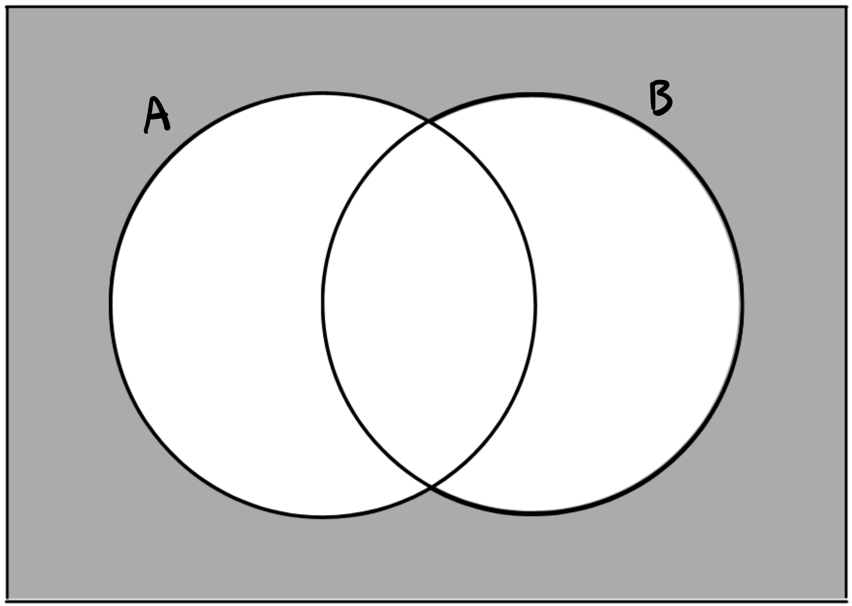
a) $P(\overline{A \cap B})$

b) $P(\overline{A \cup B})$

c) $P(\overline{A} \cup \overline{B})$

d) $P(\overline{A} \cup B)$

e) none of the above



Clicker 4

What does the following shaded region in the Venn diagram correspond to?

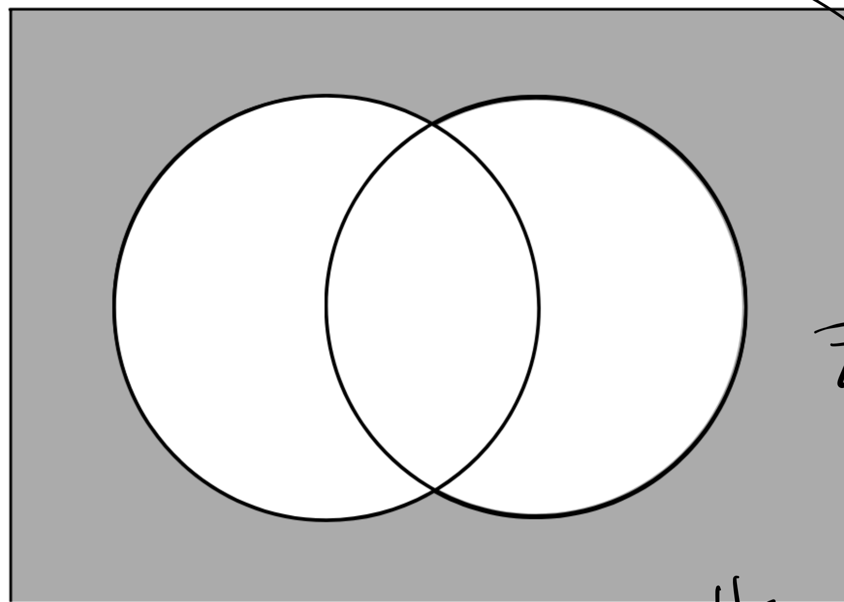
a) $P(\overline{A \cap B})$

b) $P(\overline{A \cup B})$ $= P(\bar{A} \cap \bar{B})$

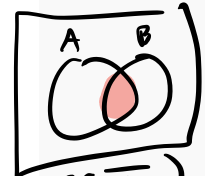
c) $P(\bar{A} \cup \bar{B})$

d) $P(\bar{A} \cup B)$

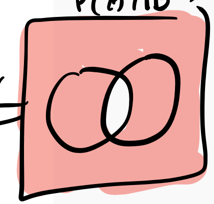
e) none of the above



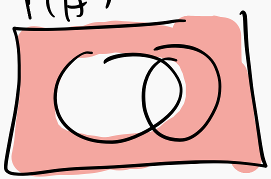
$P(A \cap B)$



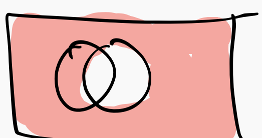
$P(\overline{A \cap B})$ $\neq$



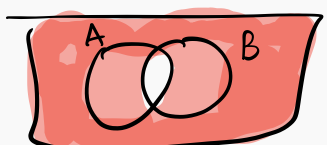
$P(\bar{A})$

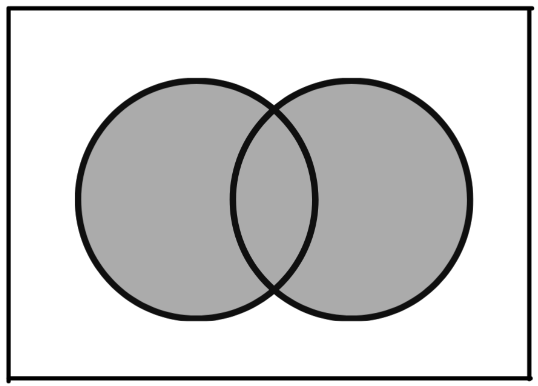


$P(\bar{B})$

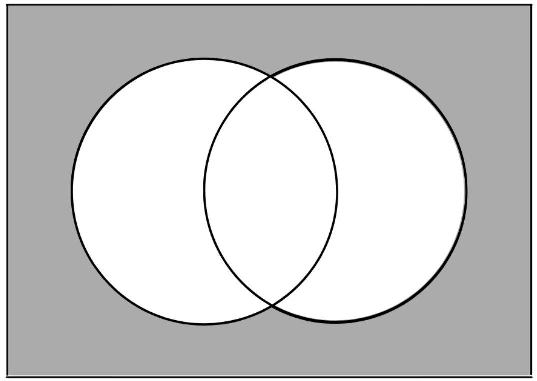


$P(\bar{A} \cup \bar{B})$ $\neq$





unshaded region = compliment



Clicker 5

What does the following shaded region in the Venn diagram correspond to?

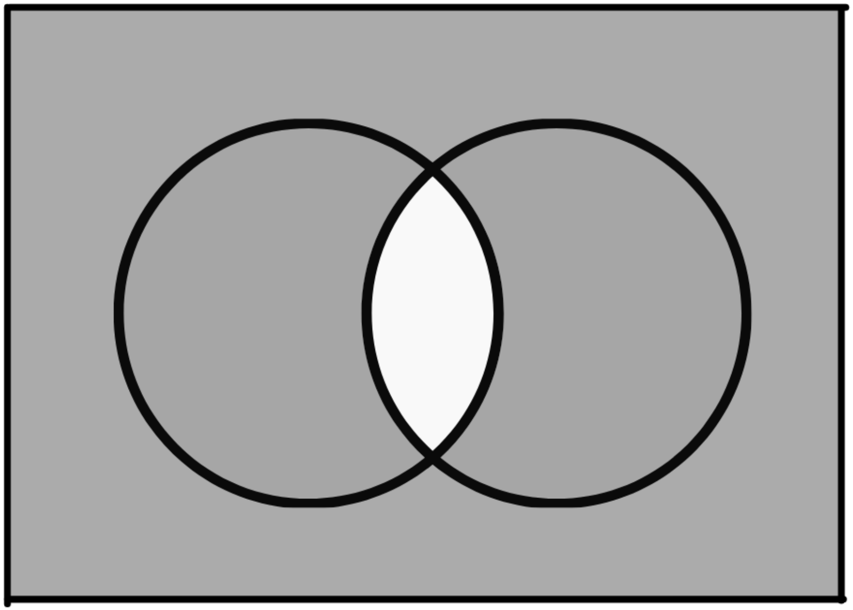
a) $P(\overline{A \cap B})$

b) $P(\overline{A \cup B})$

c) $P(\overline{A} \cup \overline{B})$

d) a) & c)

e) none of the above



Clicker 5

What does the following shaded region in the Venn diagram correspond to?

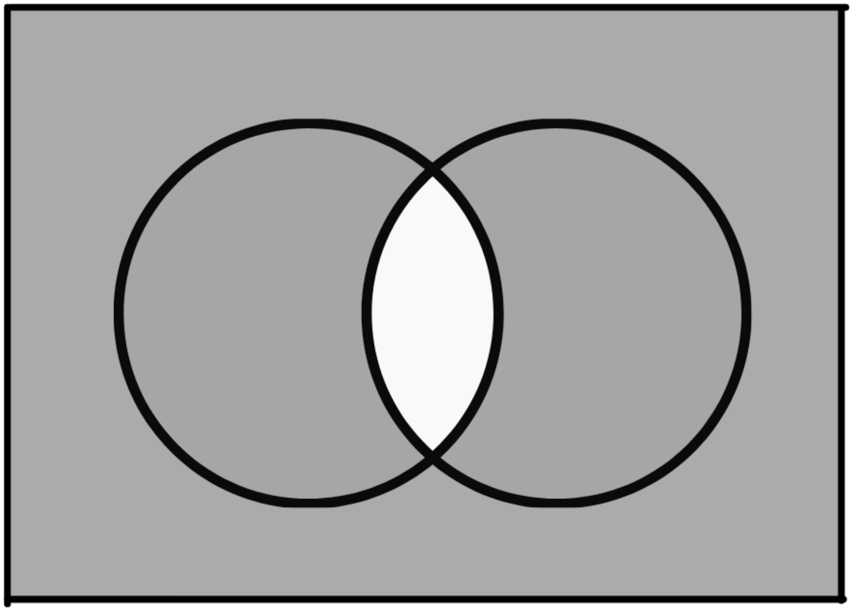
a) $P(\overline{A \cap B})$

b) $P(\overline{A \cup B})$

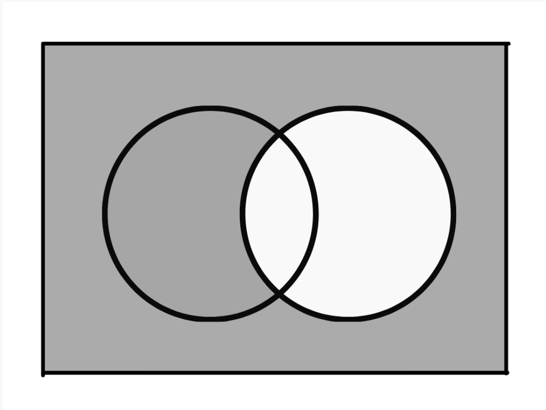
c) $P(\overline{A} \cup \overline{B})$

d) a) & c)

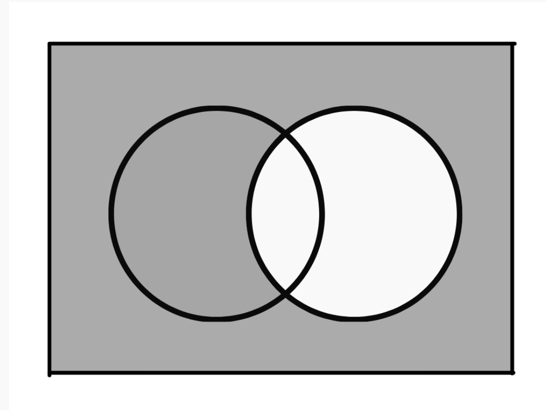
e) none of the above



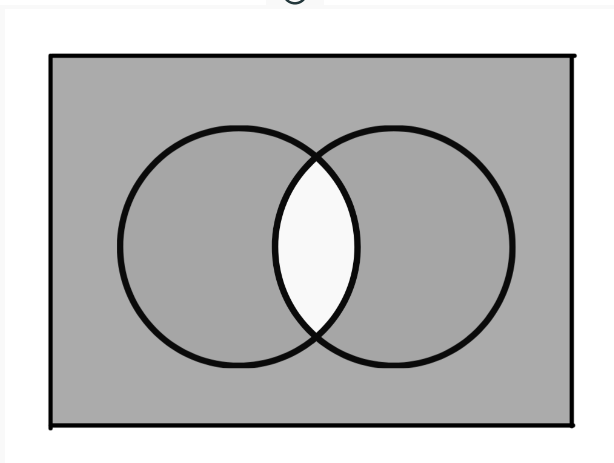
Union example



$\cup$

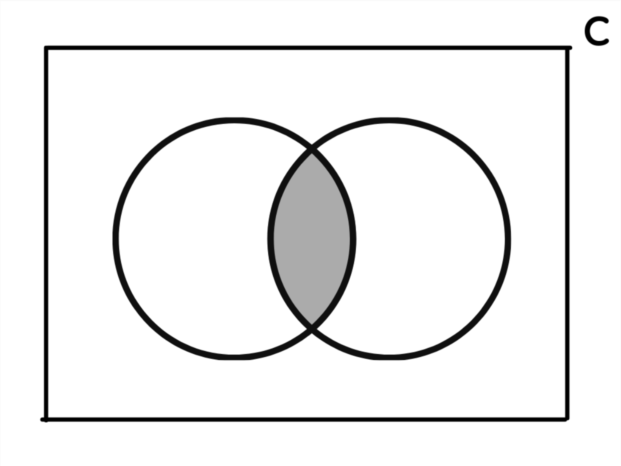


$=$

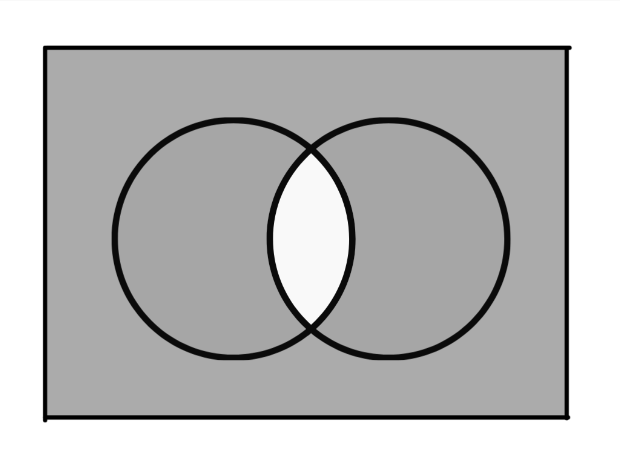


Collect *all* shaded region for $\overline{A} \cup \overline{B} = \overline{A \cap B}$

Compliments example



=



$$\overline{A \cap B}$$

OR Rule

While we have seen this rule already, let's restate it using the notation we just learnt.

Theorem 3 (Addition Rule AKA the "OR" Rule)

*For any two events E and F, the probability of E **or** F occurring is given by*

$$P(E \cup F) = P(E) + P(F) - P(E \cap F)$$

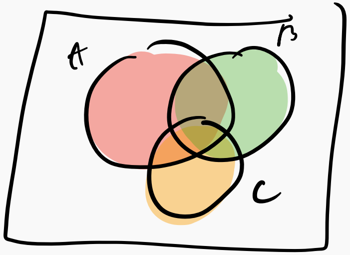
If E and F M.E.

$$P(E \cup F) = P(E) + P(F) (-0)$$

Addition (OR) Rule

Note that if E and F are mutually exclusive then $P(E \cap F) = 0$ and we have the special case formula previously discussed:

$$P(E \cup F) = P(E) + P(F)$$



More generally ...

Theorem 4 (Addition rule (OR rule) for disjoint outcomes)

If there are many disjoint outcomes $A_1, \ldots, A_k$, then the probability that one of these outcomes will occur is simply the sum of their individual probabilities:

(Ax. 3.)

$$P(A_1 \cup A_2 \cup \cdots A_k) = P(A_1) + P(A_2) + \cdot + P(A_k)$$

Addition (OR) Rule

Clicker 6

A coin is flipped 3 times, what is the probability that it lands heads at least twice?

A) 1. 3/8

B) 2. 1/8

C) 3. 1/2

D) 4. 3

E) 5. none of the above

Addition (OR) Rule

Clicker 6

THH, HTH, HHT

A coin is flipped 3 times, what is the probability that it lands heads at least twice? HHT, HHH

M.E.

A = the coin lands heads (exactly) twice

B = the coin lands heads (exactly) three times

$S = \{HHH, HHT, HTH, THH, TTT, TTH, THT, HTT\}$

1. 3/8
2. 1/8
3. 1/2
4. 3
5. none of the above

$$P(\text{at least } 2) = P(A \cup B)$$
$$= P(A) + P(B)$$
$$= \boxed{\frac{3}{8}} + \boxed{\frac{1}{8}}$$
$$= \frac{4}{8} = \frac{1}{2}$$

S = {HHH, TTT, HHT, TTH}

Conditional Probability

Conditional Probability

Conditional probability, aims at finding the probability of an event, given that a certain event has already occurred.

Definition 2

The conditional probability of A given B is

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)},$$

provided $P(B) > 0$.

$$P(B) \cdot P(A|B) = P(A \cap B)$$

Example 5

Suppose a fair die is rolled. (i) What is the probability that the number of dots landing face up is odd? (ii) What is the probability that the die lands odd given that the number is greater than 2?

A = die lands odd

$\boxed{B = \text{die lands} > 2}$

$A = \{1, 3, 5\}$

$B = \{3, 4, 5, 6\}$

$A \cap B = \{3, 5\}$

(i) $P(A) = \frac{3}{6} = \frac{1}{2} = 0.5$

(ii) $P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{2/\cancel{6}}{4/\cancel{6}} = \frac{1}{2} = 0.5$

Example 6

Two cards are selected from a well shuffled pack. If the first card drawn was a jack, what is the probability second card drawn is also a jack?

J_1 = first card is Jack

J_2 = second card is Jack

A♡2♡
A♡3♡
⋮
K♧Q♧

$S = \{ JA, J2, JJ, JQ \ldots \ldots J \}$

$$P(J_2 \mid J_1) = \frac{3}{51}$$

Multiplication (AND) Rule

Definition 3 (Multiplication Rule AKA "AND Rule")

If E and F are two events then

$$P(E \cap F) = P(E)P(F \mid E)$$

$$P(F \cap E) = P(F)P(E \mid F)$$

where $F \mid E$ means the occurrence of an event F given that an event E has already occurred and $E \mid F$ means the occurrence of an event E given that an event F has already occurred.

Example 7

without replacement.

Two cards are selected from a well shuffled pack. What is the probability that they are both jacks?

Let J_1 be the event that the first card is a jack and

Let J_2 be the event that the second card is a jack, then

$$P(J_1 \cap J_2) = P(J_1) \cdot P(J_2 | J_1)$$

$$= \frac{4}{52} \cdot \frac{3}{51} = \frac{1}{221}$$

J♠ J♡

J♣ J◇

⋮

$S = \{AA, A2, JQ. \ldots$

Example 8

There are 3 red and 4 green balls in a bag. Two balls are selected consecutively, at random without replacement. What is the probability that first ball is red and the second green?

Generalizations to more than 2 events

More generally ...

Theorem 5 (Generalization of Theorem 3)

If sample space S is comprised of k events: $A_1, A_2, \ldots, A_k$ such that $P(A_i \cap A_j) \neq 0$ for all $i \neq j$, then

$$P(\cap_{i=1}^{k} A_k) = P(A_1) \cdot P(A_2|A_1) \cdots P(A_k|A \cap_{i=1}^{k-1} A_k).$$

For instance, if $k = 3$:

If A, B, and C are any three events in a sample space S such that $P(A \cap B) \neq 0$, then

$$P(A \cap B \cap C) = P(A) \cdot P(B|A) \cdot P(C|A \cap B).$$

Example 9

There are 3 red, 4 green and 7 blue balls in a bag. Three balls are selected consecutively, without replacement. at random. What is the probability that the first is a red, the second green and the third blue?

Let R_1 be the event that the first ball is red

Let G_2 be the event that the second ball is green

Let B_3 be the event that the third ball is blue

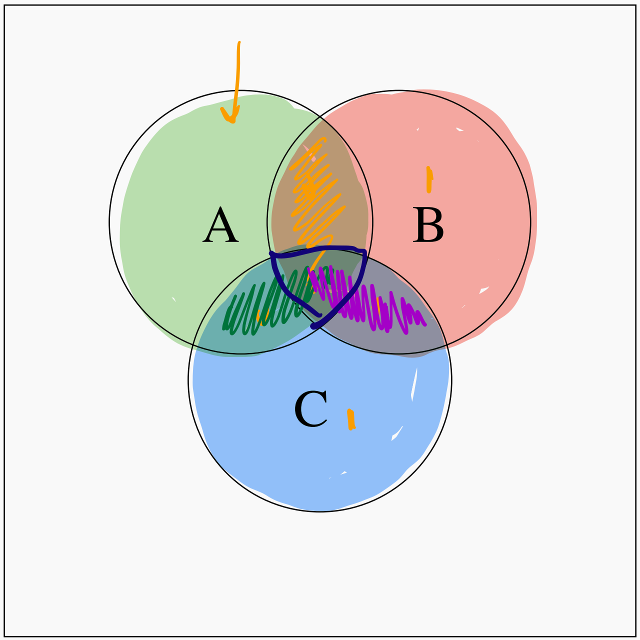
$$P(R_1 \cap G_2 \cap B_3) = P(R_1) \cdot P(G_2 | R_1) \cdot P(B_3 | R_1 \cap G_2)$$

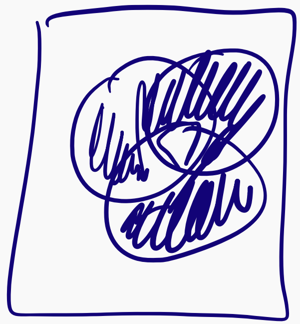
$$= \frac{3}{14} \cdot \frac{4}{13} \cdot \frac{7}{12} = \frac{1}{26}$$

Theorem 6 (Generalization of Theorem 3)

If A, B and C are three events in a sample space S, then

$$
\begin{aligned}
\underline{P(A \cup B \cup C)} \quad = \quad & P(A) + P(B) + P(C) - \cancel{P(A \cap B)} \\
& -\cancel{P(A \cap C)} - \cancel{P(B \cap C)} \boxed{+ P(A \cap B \cap C).}
\end{aligned}
$$





Statistical Independence

Independence

Two processes are *independent* if knowing the outcome of one provides no useful information about the outcome of the other.

Independence

Two processes are *independent* if knowing the outcome of one provides no useful information about the outcome of the other.

- Knowing that the coin landed on a head on the first toss does not provide any useful information for determining what the coin will land on in the second toss. → Outcomes of two tosses of a coin are independent.

Independence

Two processes are *independent* if knowing the outcome of one provides no useful information about the outcome of the other.

- Knowing that the coin landed on a head on the first toss does not provide any useful information for determining what the coin will land on in the second toss. → Outcomes of two tosses of a coin are independent.
- Knowing that the first card drawn from a deck is an ace does provide useful information for determining the probability of drawing an ace in the second draw. → Outcomes of two draws from a deck of cards (without replacement) are dependent.

(with replacement → independent)

Independence

Definition 4 (Independence)

If events E and F are independent:

$$P(F \mid E) = P(F) \quad \text{and} \quad P(E \mid F) = P(E)$$

- In essence: two events are independent if the outcome of one has no effect over the outcome of the other.
- Note that this is very different from mutually exclusive - many students mix the two definitions up!
- In fact, if two events E and F are mutually exclusive, then the following is true:
  - $P(F \mid E) = 0$
  - $P(E \mid F) = 0$

Some Notes

- There is an important special case of the AND rule — the case where the events are independent. In fact, statistical independence is often defined in terms of the AND rule.

Definition 5 (independence)

Two events E and F are said to be independent if

$$P(E \cap F) = P(E)P(F)$$

with non-ind events:

$$P(E \cap F) = P(E) \cdot \underbrace{P(F \mid E)}_{P(F)}$$

Example 10 (Ex:1)

A card is selected from a well shuffled pack and a die is rolled. What's the probability of obtaining a red card & an even number?

Let R be the event that the card is red

Let E be the event that the die shows an even number

$$P(R \cap E) = P(R) \cdot P(E) = \left(\frac{1}{2}\right) \cdot \left(\frac{1}{2}\right) = \frac{1}{4}$$

$$= P(R) \cdot \underline{\underline{P(E|R)}} = \left(\frac{1}{2}\right) \cdot \left(\frac{1}{2}\right) = \frac{1}{4}$$

Example 11

A card is selected from a well shuffled pack and a coin is tossed. What is the probability of obtaining a queen and a tail?

$= P(Q) \cdot P(T \mid Q)$

$P(Q \cap T) = P(Q) \cdot P(T)$

$= \frac{4}{52} \cdot \frac{1}{2} = \frac{1}{26}$

Let Q be the event that the card is a queen

Let T be the event that the coin shows a tail

The Complement Rule

Theorem 7 (Complement Rule)

For any event E, $P(E) = 1 - P(\overline{E})$.

Note that basic probability rules discussed in the previously, also hold conditionally, i.e.

$$P(A \mid B) = 1 - P(\overline{A} \mid B)$$

$$P(A \cup B \mid C) = P(A \mid C) + P(B \mid C) - P(A \cap B \mid C)$$

Example 12

A coin is flipped 20 times, what is the probability that it lands heads at least once? $P(E)$ $1 - P(\bar{E})$

exactly.

- Let H_i be the event that i coins land head facing up

$\to$ H T T T T T
$\to$ T H T T . . . .

Let E be the event that the coin lands heads at least once

$P(H_1) = P(HTT\cdots T) + P(THTT\cdots) +$

TTT . . . . TT

$(\frac{1}{2})^{20}$ P_1 $P_{20} = 1$

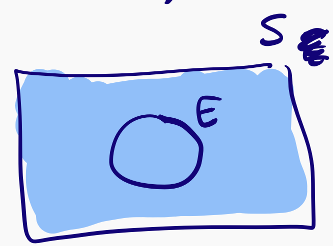
$S = \underbrace{\{ H_0, H_1, H_2 \cdots , H_{20} \}}$

$E = \{ \quad , H_1, H_2, \ldots H_{20} \}$

$\overline{E} = E^c = \{ H_0 \}$

Multiplication Rule w/ ind events.

$P(\overline{E}) = P(H_0) = P(T_1 \cap T_2 \cap \cdots \cap T_{20}) = \left(\frac{1}{2}\right)\cdot\left(\frac{1}{2}\right)\cdots\left(\frac{1}{2}\right) = \left(\frac{1}{2}\right)^{20}$



compliment Rule

$P(E) = 1 - P(\overline{E}) = \boxed{1 - \left(\frac{1}{2}\right)^{20}} = 0.99999046$

Some Food for thought

In Ex:12, what would happen if we had to work out the probability of it landing on heads 11 times?

- We need to study counting and some further probability to work this out.
- Come to next lecture to find out how!

Exercises to try at home

Exercise 1 Two fair dice are rolled.

1. What is the probability that both numbers are odd? Hint: write out all of the possible outcomes; there are 36.
2. What is the probability that both numbers are the same?

Exercise 2 Two cards are chosen at random from a standard deck. What is the probability that (i) both are red? (ii) None are red?

Exercise 3 A die is rolled 10 times. What is the probability that:

1. The number 6 does not show at all?
2. No odd numbers show?
3. An odd number shows at least once?

Exercises to try at home

Exercise 4 Lotto rules: Choose 6 numbers from 1–49, each number is only selected once (in other words, the combination $\{1,30,32,32,5,8\}$ is impossible because the number 32 cannot arise twice). Order of selection does not matter.

1. What is the probability of winning the Lotto?
2. What is the probability of matching at least one number in the Lotto?

Exercise 5 Suppose that a class consists of eleven females (including the professor) and four males. If two students are chosen at random, what is the probability that:

1. Both are female?
2. Both are male?

Exercise 6 A card is chosen from a well shuffled pack. What is the probability that it is:

1. A queen or a heart?
2. A 3 or a black card?
3. An odd numbered card or a diamond?

Exercise 7 A bag contains nine balls numbered 1 – 9; balls 1 – 3 are red, balls 4 – 6 are blue and balls 7 – 9 are green. A ball is selected at random. What is the probability that:

1. It is blue?
2. It has an odd number on it?
3. It is blue or it has an odd number on it?