Introductory Statistics: A Problem-Solving Approach

by Stephen Kokoska

Chapter 4

Probability

Experiments

An **experiment** is an activity in which there are at least two possible outcomes, and the result of the activity cannot be predicted with absolute certainty.

Examples:

- Roll a six-sided die and record the number that land face up.
- Using a radar gun, record the speed of pitch at a Red Sox baseball game.
- Count the number of patients who arrive at the emergency room of a city hospital during a 24-hour period.
- Select two Ninja Blenders and inspect each for flaws in materials and workmanship.

Outcomes

Because we don’t know for sure what will happen when we conduct an experiment, we need to consider all possible outcomes. This sounds easy (just think about all the things that can happen), but it can be tricky. Sometimes it involves a lot of counting, but often outcomes can be visualized using a tree diagram.

Example: Drone Pilot

Suppose two drone pilots flying devices that weigh more than 0.55 lb are selected at random and their drones are checked for registration. How many possible outcomes are there, and what are they?

Step 1: If the drone is registered, denote the observation by R (registered). If the drone is not registered, use U (unregistered).

Step 2: Each outcome is a pair of observations, one on each drone. There are four possible outcomes: RR, RU, UR, UU.

The first letter indicates the observation on the first drone, and the second letter indicates the observation on the second drone.

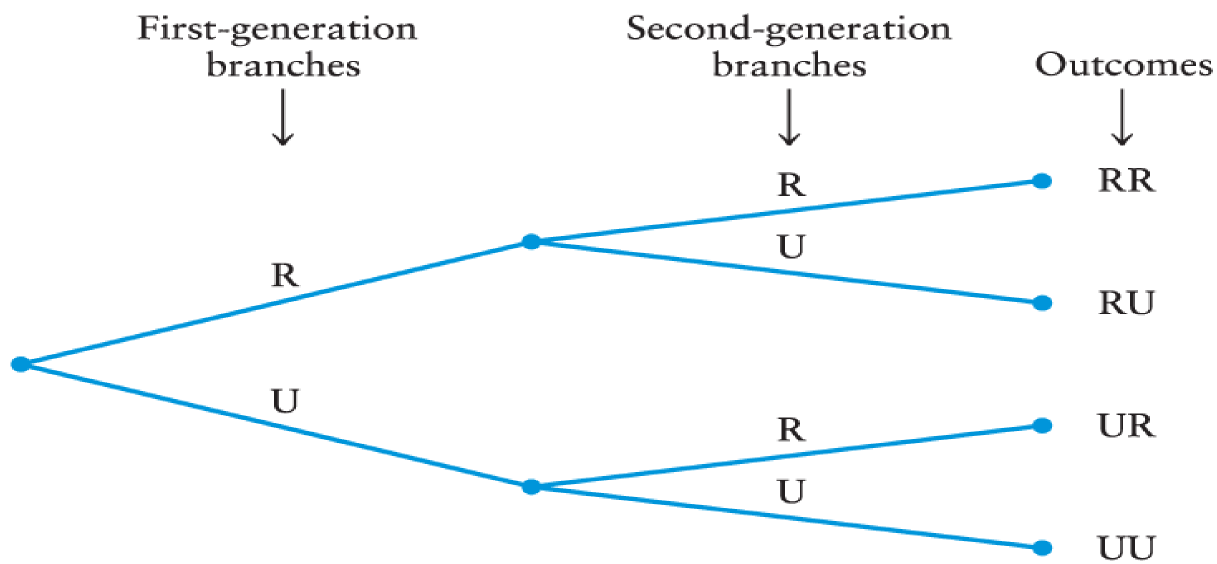
RU is a different outcome from UR.

RU means the first drone was registered and the second drone was not.

UR means the first drone was unregistered and the second drone was registered.

Example: Drone Pilot

All of the outcomes from the experiment in this example can be determined by constructing a **tree diagram**, a visual road map of possible outcomes.



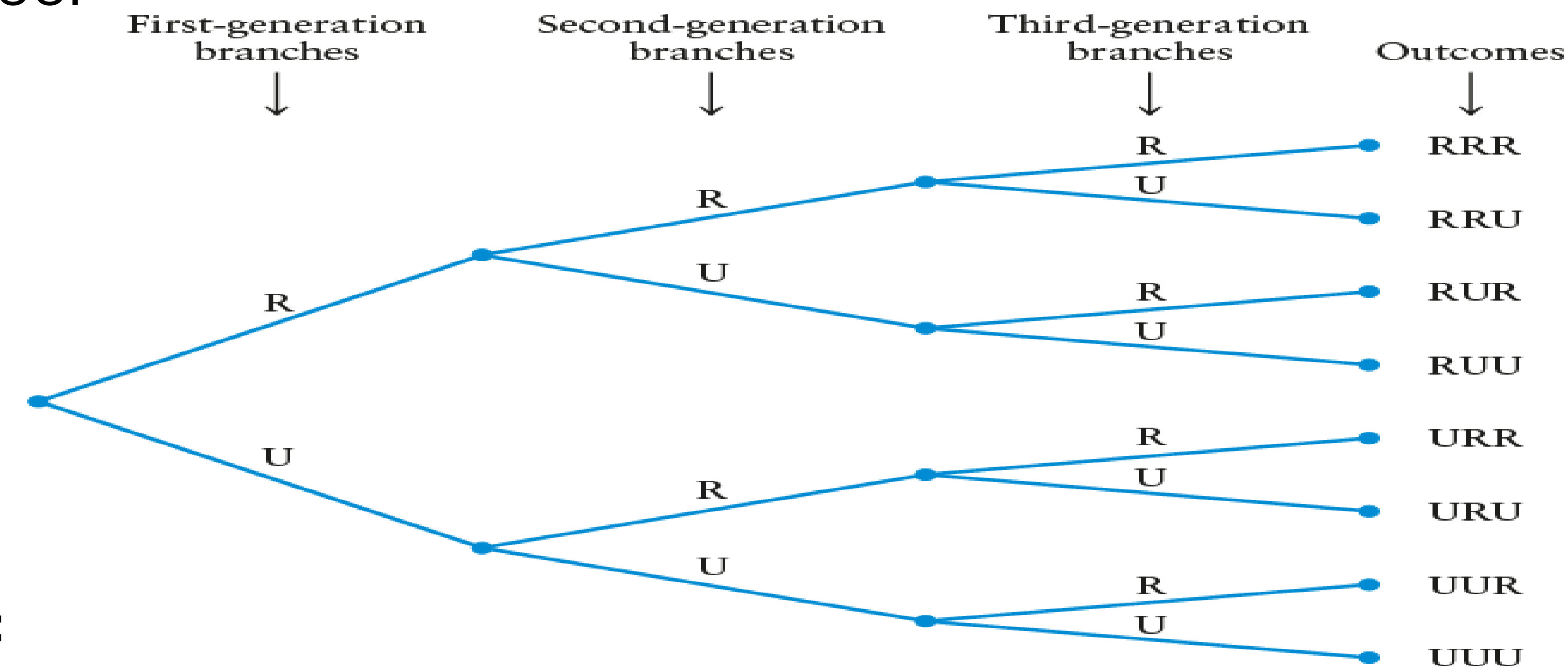
Kokoska, *Introductory Statistics*, 3e © 2020 W. H. Freeman and Company

- The first-generation branches indicate the possible choices associated with the first drone, and the second-generation branches represent the choices for the second drone.
- A path from left to right represents a possible experimental outcome.

Example: Drone Pilot

Extend the previous example. How many outcomes are there if we select three drones and record their registration status?

Now there are eight possible outcomes: RRR, RRU, RUR, RUU, URR, URU, UUR, UUU.



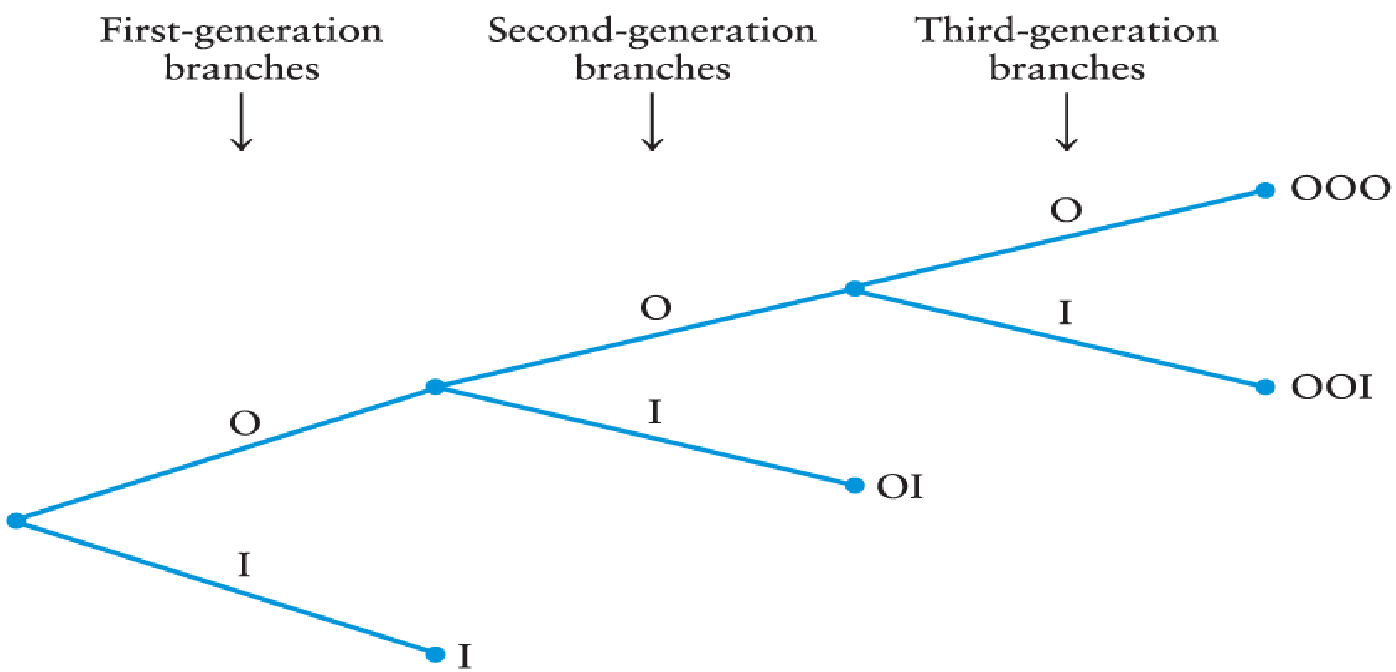
Notes:

1. Tree diagrams are a fine technique for finding all the possible outcomes for an experiment. However, they can get very large, very fast.
2. A tree diagram does not have to be symmetric. The branches and paths depend on the experiment. Example on the next slide

EXAMPLE: Breakfast of Champions

A consumer is searching for a box of his favorite breakfast cereal. He will check all three grocery stores in town if necessary but will stop if the cereal is found. The experiment consists of recording the cereal search result. How many possible outcomes are there, and what are they?

- If the cereal is in stock, use the letter I; if it is out of stock, use O.
- The tree diagram shows four possible paths from left to right.
- This tree diagram is not symmetric.



Kokoska, *Introductory Statistics*, 3e © 2020 W. H. Freeman and Company

The Sample Space and Event

The **sample space** associated with an experiment is a listing of all possible outcomes using set notation. It is the collection of all outcomes written mathematically, in set notation using braces { }, and is denoted by *S.*

Examples:
Drone pilot experiment: S = {RR, RU, UR, UU}.
Extended drone pilot experiment: S = {RRR, RRU, RUR RUU, URR, URU, UUR, UUU}.
Cereal experiment: S = {I, OI, OOI, OOO}.

- An **event** is any collection (or set) of outcomes from an experiment (i.e., a subset of the sample space).
- A simple event is an event consisting of exactly one outcome.
- An event may be given in standard set notation, or it may be defined in words.
- Notation:
  - ✓ Events are denoted with capital letters, A, B, C, ….
  - ✓ Simple events are often denoted by E_1, E_2, E_3, ….
- An event may be empty. An event containing no outcomes is denoted by { } or ∅

More Related Definition

Complement of an Event
Let *A* be an event associated with a sample space *S*.
The event ***A* complement**, denoted A', consists of all outcomes in the sample space *S* that are not in *A*.

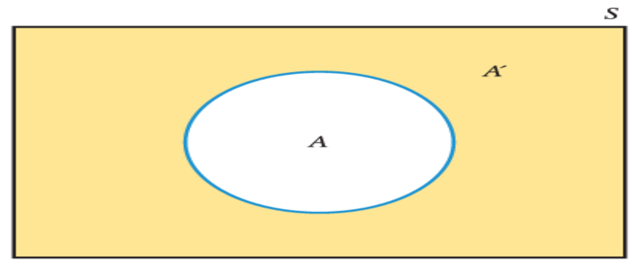
Union of Two Events
Let A and B denote two events associated with a sample space S.
The event A union B, denoted $A \cup B$, consists of all outcomes that are in A or B or both.

Intersection of Two Events
Let A and B denote two events associated with a sample space S.
The event A intersects B, denoted $A \cap B$, consists of all outcomes that are in both A and B.

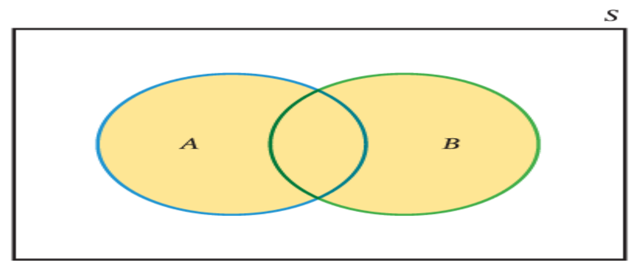
Mutually Exclusive Events
Let A and B denote two events associated with a sample space S.
If A and B have no elements in common, they are disjoint or mutually exclusive, written $A \cap B = \{\ \}$.

Venn Diagrams

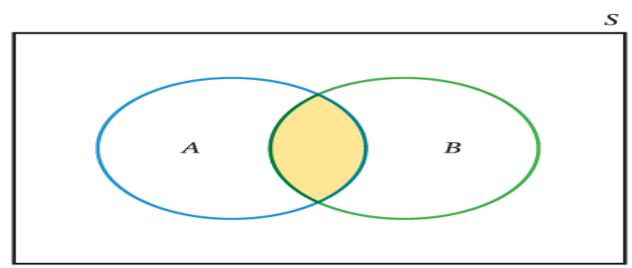
1. Venn diagrams are used to visualize a sample space, events, and combinations of events.
2. Draw a rectangle to represent the sample space.
3. Figures (often circles) are drawn inside the rectangle to represent events.
4. Plane regions represent events.



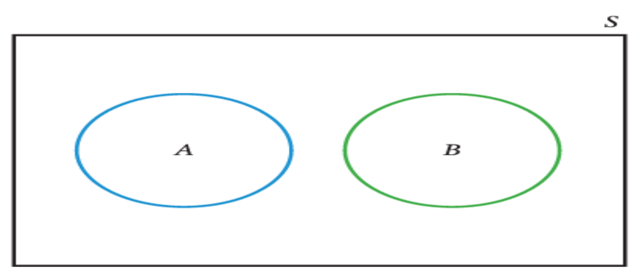
A complement: *A*′



A union *B*: $A \cup B$



A intersection *B*: $A \cap B$



A and *B* are disjoint: $A \cap B = \{\ \}$

Introduction to Probability

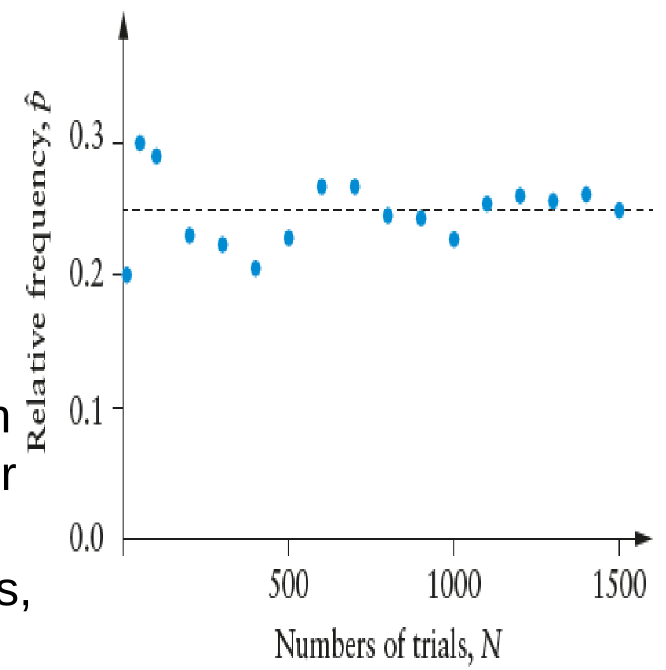
- Given an experiment, some events are more likely to occur than others.
- For an event *A*, we need to assign a number that conveys the *likelihood of occurrence*.
- The likelihood that *A* will occur is simply the **probability** of the event *A*.
- The probability of an event A, denoted P(A), is a number between 0 and 1 (including those endpoints) that measures, or conveys, the likelihood that A will occur.
  - ✓ If the probability of an event is close to 1, then the event is likely to occur
  - ✓ If the probability of an event is close to 0, then the event is not likely to occur
- The **relative frequency of occurrence** of an event is the number of times the event occurs divided by the total number of times the experiment is conducted.
- If an experiment is conducted N times and an event occurs n times, then the probability of the event is approximately n/N (the relative frequency of occurrence). **The probability of an event A**, P(A), is the proportion of times that the event A will occur in the long run.

EXAMPLE: Pick a Card, Any Card

A regular 52-card deck contains 13 clubs, 13 diamonds, 13 hearts, and 13 spades. Suppose an experiment consists of selecting one card from the deck and recording the suit. What is the probability of selecting a club?

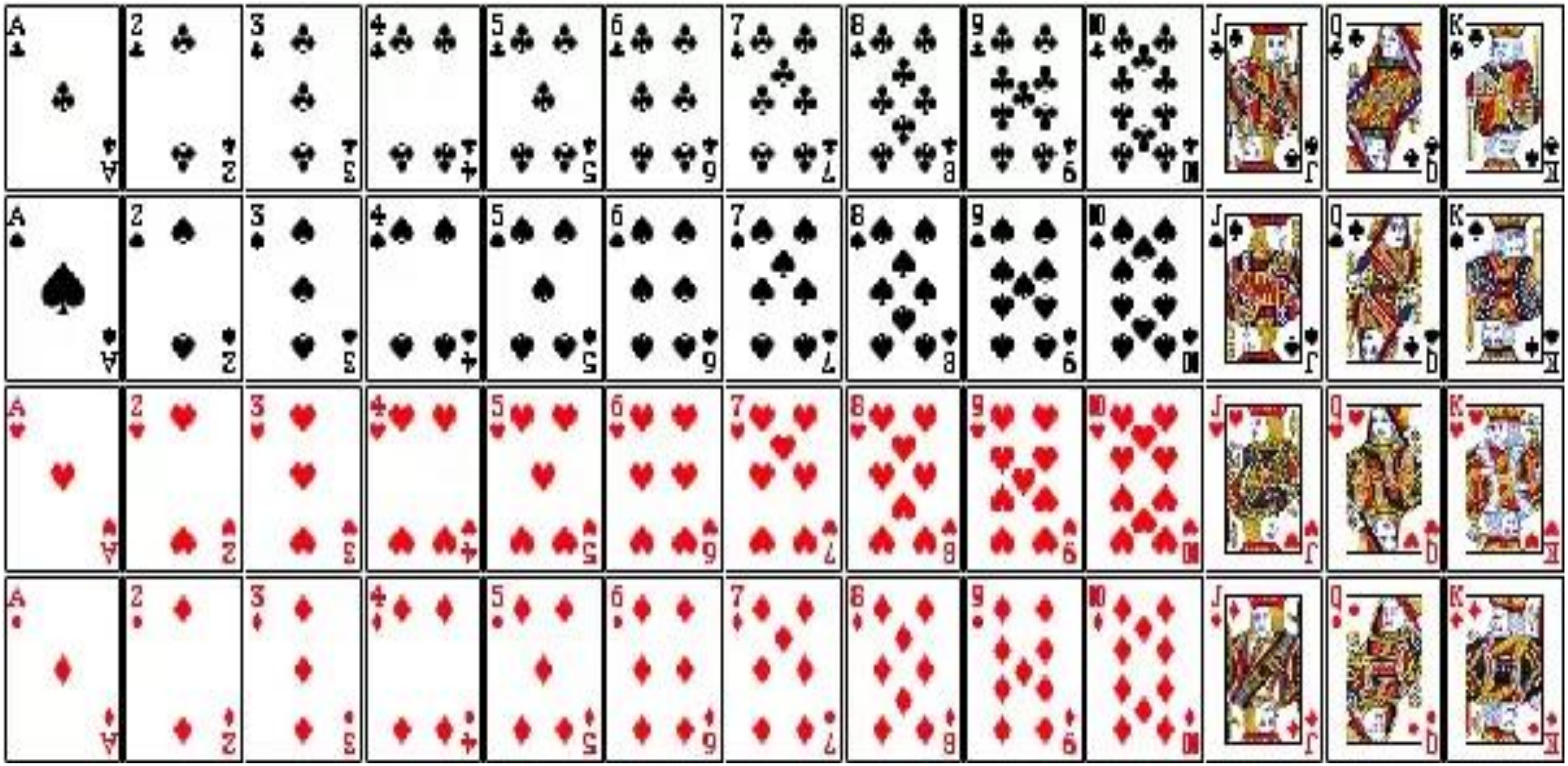
To estimate the probability of C, P(C) it seems reasonable to conduct the experiment several times and see how often a club is selected. If C occurs often (we get a club a lot of the time), then the likelihood (probability) should be high. If C occurs rarely, then the probability should be close to 0.
To estimate the likelihood of selecting a club, we use the relative frequency of occurrence of a club:

$$\text{Relative frequency} = \frac{\text{number of times a club is selected}}{\text{total number of selections}}$$



- Suppose that after 10 tries, a club was selected only twice. The relative frequency is 2/10 = 0.2. This is an estimate of P(C).
- Suppose we try the experiment a few more times. With more observations, we should be able to make a better guess at P(C).
- The graph shows a remarkable pattern. As N increases, the points come noticeably closer to the dashed line.

clubs (♣), diamonds (♦), hearts (♥) and spades (♠)



The probability of an event A, P(A)

- The probability of an event A, P(A), is the limiting relative frequency, the proportion of time that the event A will occur in the long run. This is a basic and sensible definition, a rule for assigning probability to an event.
- Given an event, all we need to do is finding the limiting relative frequency. Although this definition makes sense, and 52-card Example and its Figure support and illustrate our intuition, it presents a real practical problem.
- We cannot conduct experiments over and over, compute relative frequencies, and only then estimate the true probability.
  - ✓ How will we ever know the true limiting relative frequency?
  - ✓ How large should N be?
  - ✓ When are we close enough?
  - ✓ Will we ever hit the limiting relative frequency exactly?
- The definition is nice, but it seems to offer little hope of ever finding the true probability of an event.
- Fortunately, there is another way to determine the exact probability in some cases. Let’s discuss two examples in the next slide.

EXAMPLES

1. Suppose an experiment consists of tossing a fair coin and recording the side that lands face up. The event H is the coin landing with heads face up. Find P(H).

   - ➢ There are only two possible outcomes on each flip of the coin, and both are equally likely to occur.
   - ➢ In the long run, we expect heads to occur half of the time. Therefore, P(H) = 1/2.
   - ➢ Without flipping the coin thousands of times, making estimates, or guessing at the limiting relative frequency, we are certain the probability is 1/2.

2. An experiment consists of tossing a fair six-sided die and recording the number that lands face up. Consider the event E={1}, rolling a one. Find P(E).

- ➢ There are six possible outcomes for each roll of the die, and all are equally likely to occur.
- ➢ In the long run, we expect a 1 to occur one-sixth of the time. Therefore, P(E)=1/6.

These two examples suggest that it is indeed possible to find the limiting relative frequency! They are special cases, however, because in each experiment, all of the outcomes are equally likely.

Properties of Probability

1. For any event A, $0 \leq P(A) \leq 1$.
2. For any event A, P(A) is the sum of the probabilities of all of the outcomes in A.
3. The sum of the probabilities of all possible outcomes in a sample space is 1: $P(S) = 1$.
4. The probability of the empty set, or empty event, is 0: $P(\{\ \}) = P(\varnothing) = 0$. This event contains no outcomes and cannot happen.

Finding Probabilities in an Equally Likely Outcome Experiment

In an equally likely outcome experiment, the probability of any event A is the number of outcomes in A divided by the total number of outcomes in the sample space S.

$$\mathrm{P}(A) = \frac{\mathrm{N}(A)}{\mathrm{N}(S)}$$

Probability Rules

Complement Rule

For any event A, $P(A) = 1 - P(A')$

Addition Rule

1. For any two events A and B,

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

2. For any two **disjoint** (mutually exclusive) events A and B,

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

Conditional Probability Rule

$$\boxed{P(E_1 \mid E_2) = \frac{P(E_1 \text{ and } E_2)}{P(E_2)}}$$

where $P(E_2) > 0$

"given that"

e.g., Purchasing something given that the price has been lowered

Probability Rules

Multiplication Rule

$$P(E_1 \mid E_2) = \frac{P(E_1 \text{ and } E_2)}{P(E_2)}$$

$$P(E_1 \text{ and } E_2) = P(E_1)\, P(E_2 | E_1)$$

If E_1 and E_2 are independent, then $P(E_2|E_1) = P(E_2)$, and the multiplication rule simplifies to

$$P(E_1 \text{ and } E_2) = P(E_1) P(E_2)$$

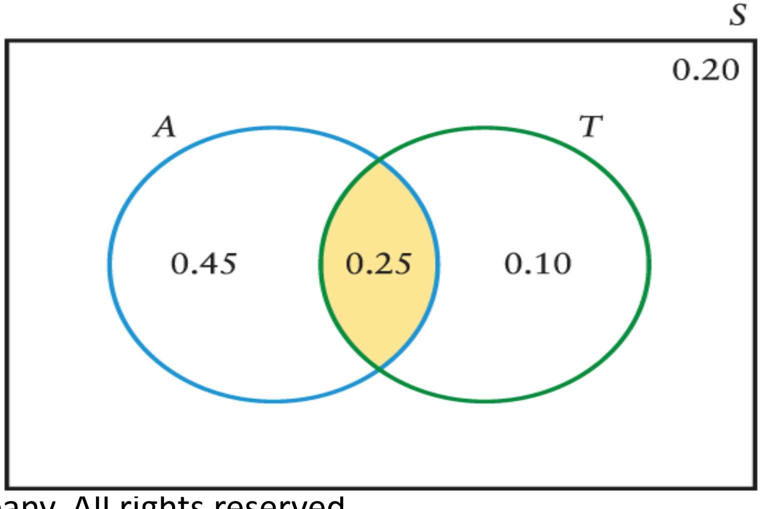
- Conditional probability considers how knowing extra information may change a probability assignment.
- If we know the events A and B are independent, then
  $P(A \mid B) = P(A)$ and $P(B \mid A) = P(B)$
  Similarly, if either one of these equations is true, then the other is also true, and the events A and B are independent.

Example: Super Bowl Ads

The 2018 Super Bowl ads included Morgan Freeman lip syncing and Eli Manning and O'Dell Beckham Jr. dancing. According to *Variety*, two of the best commercials were for Amazon's Alexa voice service and Tide laundry detergent. A survey indicated that 70% of all people watching the Super Bowl saw the ad for Alexa, 35% saw the ad for Tide, and 25% saw both. Suppose a person who watched the Super Bowl is selected at random.

a. Draw a Venn diagram to illustrate the events in this problem.
b. What is the probability that the person saw at least one of these two ads?
c. What is the probability that the person saw neither ad?
d. What is the probability that the person saw just the ad for Alexa?
e. What is the probability that the person saw just one of these two ads?

Let A = the person saw the Alexa ad; P(A) = 0.70
Let T = the person saw the Tide ad; P(T) = 0.35



Example: Super Bowl Ads

b. What is the probability that the person saw at least one of these two ads?
c. What is the probability that the person saw neither ad?
d. What is the probability that the person saw just the ad for Alexa?
e. What is the probability that the person saw just one of these two ads?

b. The probability of seeing at least one ad means seeing the ad for Alexa, or the ad for Tide, or both. That's a union of two events.

$$P(A \cup T) = P(A) + P(T) - P(A \cap T)$$
$$= .70 + .35 - .25 = .80$$

c. Saw neither means did not see either the Alexa ad or the Tide ad. Because the event $A \cup T$ means saw the Alexa ad or the Tide ad, neither suggests the complement of $A \cup T$.

$$P(A \cup T)^{/} = 1 - P(A \cup T)$$
$$= 1 - .80 = .20$$

d. Saw just the ad for Alexa means saw the Alexa ad but not both ads.

$$P(\text{just Alexa}) = P(A) - P(A \cap T)$$
$$= .70 - .25 = .45$$

e. Saw just the ad for Alexa means saw the Alexa ad but not both ads

$$P(\text{Exactly one}) = P(A \cup T) - P(A \cap T)$$
$$= .80 - .25 = .55$$

More Examples:

Example1: If you draw one card randomly from the 52 deck cards, find the probability of drawing black or red card.

Let A = black card & B = red card

$P(AUB) = P(A) + P(B) = 26/52 + 26/52 =$

Example2: If you roll one die and the event A denotes an even number and the event B denotes a number divisible by 3. Find P(AUB)

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

$A = \{2,4,6\}$

$B = \{3,6\}$

$P(AUB) = P(A) + P(B) - P(AB)$

$= 3/6 +$

Example3: A bag of candy contains 4 different flavors: lemon, strawberry, orange, and blueberry. What is the probability of selecting a lemon OR an orange piece of candy?

Flavor	Count
Lemon	8
Strawberry	12
Orange	5
Blueberry	10
TOTAL	35

E_1 = lemon candy $\quad P(E_1) = (8/35) = 0.23$

E_2 = orange candy $\quad P(E_2) = (5/35) = 0.14$

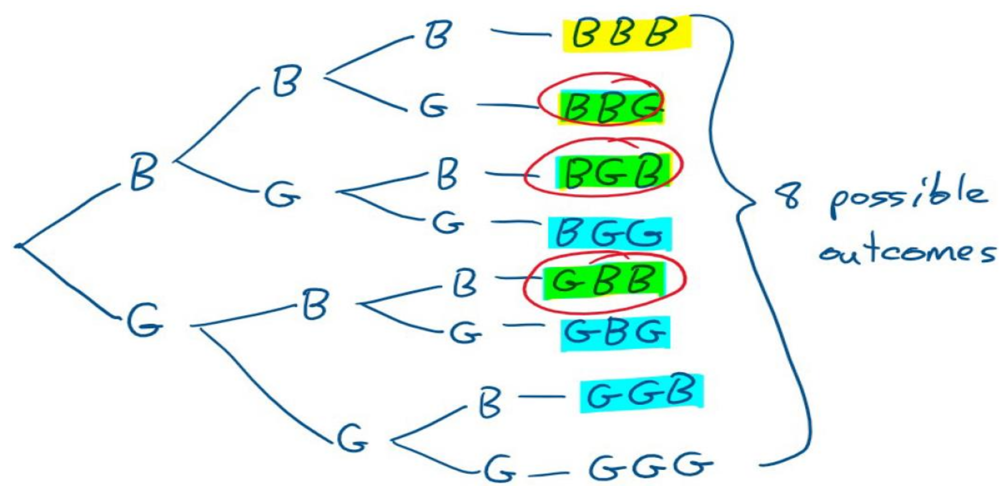
$$P(E_1 \cup E_2) = P(E_1) + P(E_2)$$
$$= 0.23 + 0.14 = 0.37$$

Examples on Multiplication Rule

Example1: If you draw one card randomly from the 52 deck cards and an event A denotes to number one (an ace) card and event B denotes to a red card. Is it A and B independent events (Find P(AB))?

$$P(E_1 \cap E_2) = P(E_1)P(E_2)$$

Example2: If you choose a family randomly included 3 children. If you know that: E_1=represents a family with at most one girl
E_2=represents a family with both gender (boys & girls)
Are A and B independent events?
S = {BBB, …………………, GGG}
E_1 = {BBG, BGB, GBB, }
E_2 = { }
E_1 E_2 = { }



$$P(E_1 \cap E_2) = P(E_1)P(E_2)$$

Example 3: The same last example but with 2 children. HW

Examples on Multiplication Rule

Example 4: A study of a telemarketing company's data showed that of 1000 calls placed. The experimental probability of a call receiver staying on the line for at least a minute was 16%. The conditional probability of a call resulting in a sale, given that a receiver stayed on the line for at least a minute, was 10%.

a) What is the probability that a customer stays on the line for at least a minute and will result in a sale

Let A = a call receiver staying on the line for at least a minute

B = a call results in a sale

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

$$= (0.16)(0.1) = 0.016 \text{ or approximately } 1.6\%$$

b) How many sales were made?

No. of sales made = 0.016 (1000) = 16 sales

Example: Who Needs a Car?

According to a recent study, the percentage of households with vehicles in Boston, Massachusetts, is 66.2%; in Cambridge, Massachusetts, 63.2%; and in Worcester, Massachusetts, 80.7%. Suppose a household from each city is randomly selected.

a) Find the probability that all three households have vehicles.

Let B, C, and W represent a household from the respective city having a vehicle and assume these three events are independent.

We need $P(B \cap C \cap W) = P(B) \cdot P(C) \cdot P(W)$

$$= (0.662)(0.632)(0.807) = 0.3376$$

b) Find the probability that none of the three households has a vehicle.

$$P(B' \cap C' \cap W') = P(B')\,P(C')\,P(W')$$

$$= [1 - P(B)] \cdot [1 - P(C)] \cdot [1 - P(W)]$$

$$= (1 - 0.662)(1 - 0.632)(1 - 0.807) = 0.0240$$

c) Find the probability that exactly one of the three households has a vehicle

$$P(B \cap C' \cap W') + P(B' \cap C \cap W') + P(B' \cap C' \cap W)$$

$$= P(B) \cdot P(C') \cdot P(W') + P(B') \cdot P(C) \cdot P(W') + P(B') \cdot P(C') \cdot P(W)$$

$$= (0.662)(0.368)(0.193)+(0.338)(0.632)(0.193)+(0.338)(0.368)(0.807) = 0.1886$$

Permutations and Combinations (Counting Techniques)

For any positive whole number n, the symbol $n!$ (read "***n* factorial**") is defined as $n! = n(n-1)(n-2)\cdots(3)(2)(1)$

In addition, $0! = 1$ (0 factorial is 1)

Example: $7! = (7)(6)(5)(4)(3)(2)(1) = 5040$

Permutations

Given a collection of n different items, an ordered arrangement, or subset, of these items is called a permutation. The number of permutations of n items, taken r at a time, is given by

$$_nP_r = \frac{n!}{(n-r)!}$$

A distinguishing characteristic of a permutation is that **order matters.**

Combinations

Given a collection of n different items, an unordered arrangement, or subset, of these items is called a combination. The number of combinations of n items, taken r at a time, is given by

$$_nC_r = \binom{n}{r} = \frac{n!}{r!(n-r)!}$$

$$= \frac{_nP_r}{r!}$$

Permutations and Combinations

Example1: Intelligence Division

A fan of *Chicago P.D.* has recorded *nine* episodes from the most recent season of this show. However, he has time to watch only four episodes. Suppose he selects *four* shows at random.

a) How many different ordered arrangements of episodes are possible?

$$_{9}P_{4} = 9! / (9 - 4)! = 3024$$

b) If the season finale is recorded, what is the probability he will select and watch this episode last?

Let A = the last recording selected is the season finale.

There are four positions to fill, but the last slot is fixed (with the season finale). The first three positions can be filled by any of the remaining eight recordings, so

$$\mathrm{P}(A) = ({}_{8}P_{3}) / ({}_{9}P_{4}) = 336 / 3024 = 0.1111$$

Example2: Jury Duty

How many different ways are there to select a jury of 12 people from a pool of 20 candidates?

$${}_{20}C_{12} = 20! / [12!\,(20 - 12)!] = 125{,}970$$

Example3: An Apple a Day

A produce bin at Giant Supermarket contains 18 apples; 12 are McIntosh and 6 are Cortland. A shopper randomly selects four apples from the bin. What is the probability that

a. all of the apples selected are McIntosh?

b. the shopper selects three McIntosh apples?

The order in which the apples are selected does not matter.

Let A = select all McIntosh apples.

$$P(A) = \frac{N(A)}{N(S)}$$

$$= \frac{{}_{12}C_4}{{}_{18}C_4} = \frac{495}{3060} = 0.1618$$

Let B = select three McIntosh apples (and, therefore, 1 Cortland apple)

$$P(B) = \frac{N(B)}{N(S)}$$

$$= \frac{{}_{12}C_3 \times {}_{6}C_1}{{}_{18}C_4} = \frac{1320}{3060} = 0.4314$$

Conditional Probability

- Probability questions so far have all been examples of *unconditional* probability. No special conditions are imposed, nor is any extra information given.
- Sometimes two events are related so that the probability of one event *depends* on whether the other has occurred.
- In this case, knowing something extra may affect the probability assignment.
- **Definition**: Suppose A and B are events with $P(B) > 0$. The conditional probability of the event A given the event B has occurred, $P(A \mid B)$, is

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

$$P(A \mid B) \neq P(B \mid A)$$

Example1: It's a Digital World

The 2018 Global Digital suite of reports suggests that more than 4 billion people around the world use the Internet. The number of people using social media, such as Facebook, Twitter, and Instagram, has also increased dramatically. The following **joint probability table** lists the probabilities of people using Facebook corresponding to age (in years) and gender.

		Age				
		13–24 (A)	25–44 (B)	45–64 (C)	65+ (D)	
Gender	Male (M)	0.21	0.26	0.07	0.02	0.56
	Female (F)	0.15	0.19	0.08	0.02	0.44
		0.36	0.45	0.15	0.04	1.00

a. Find the probability that a randomly selected Internet user is female and 45–64 years old.

$P(F \cap C) = 0.08$

b. Suppose the person is male. What is the probability he is 13–24 years old?

$P(A \mid M) = P(A \cap M) / P(M) = 0.21 / 0.56 = 0.375$

c. If the person is 25–44 years old, what is the probability she is female?

$P(F \mid B) = P(F \cap B) / P(B) = 0.19 / 0.45 = 0.4222$

More Examples:

Example2: A recent survey conducted by a major financial publication yielded the following results:

	Percentage of retirement investments in the stock market					
Age of investor	<5%	5-10	10-30	30-50	50 or more	Total
<30 years	70	240	270	80	55	715
30-50	90	300	630	1120	1420	3560
50-65	110	305	780	530	480	2205
65 or older	200	170	370	260	65	1065
Total	470	1015	2050	1990	2020	7545

What is the probability that someone 65 or older will have 50% or more of his/her retirement funds invested in the stock market?

Let E_1 = 65 year or older

E_2 = 50% or more

$$P(E_2 | E_1) = P(E_1 \cap E_2) / P(E_1) = .06$$

More Examples:

Example3: Of the cars on a used car lot, 70% have air conditioning (AC) 40% have a CD player (CD), and 20% of the cars have both.

a. What is the probability that a car has a CD player, **given that** it has AC?
i.e., we want to find P(CD | AC)

$$P(CD|AC) = \frac{P(CD\text{ and }AC)}{P(AC)}$$

$$= \frac{0.2}{0.7} = 0.2857$$

	CD	No CD	Total
AC	**0.2**	**0.5**	**0.7**
No AC	**0.2**	**0.1**	**0.3**
Total	**0.4**	**0.6**	**1.0**

b. What is the probability that a car has no CD player, given that it has AC? P(no CD|AC) HW

c. What is the probability that a car has no CD player, given that it has no AC? P(no CD | no AC) HW