Jan 13

Chapter 4: Probability

Section 4.1: Basic Concepts of Probability

Def **Probability** – a numerical measure of the chance or likelyhood or possibility that a specific event will occur.

Statistics **versus** **Probability**

collect sample data

infer/draw conclusions about population

Know "population" (roll a dice) (deck of cards)

predict likelyhood/chance of getting sample

TERMINOLOGY

1. An **event** is any collection of results (or outcomes) of a procedure/experiment.
2. A **simple event** is an outcome that cannot be further broken down into simpler components.
3. The **sample space** for a procedure consists of all possible simple events.

Ex Sample space of flipping a coin

$S = \{H, T\}$

Set Notation

set = a collecting of things (w/ no order)

↳ notation

set of first 3 #s

$\{1, 2, 3\}$

Ex $\{$Juan, Ryan$\}$

Ex A = set w/ #s between 2 & 10

$= \{2, 3, 4, 5, 6, 7, 8, 9, 10\}$

$A = \{5\}$

NOTATION

Symbol	*Represents*
P	probability
A, B, C	specific events
$P(A)$	probability of event A occurring

Ex: A probability experiment consists of rolling a single six-sided *fair* die.

(a) Determine the sample space.

set of all outcomes

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

(b) What is $P(5)$?

$P(5)$ means "the probability of rolling a 5"

Def $P(A) = \dfrac{\#\text{ outcomes in event } A}{\#\text{ outcomes in sample space}}$

$P(5) = \dfrac{\#A}{\#S} = \dfrac{1}{6}$

$\boxed{P(5) = 1/6 \approx 16.7\%}$

(c) What is $P(even)$?

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

$P(E) = \dfrac{\#E}{\#S} = \dfrac{3}{6} = \boxed{\dfrac{1}{2} = 0.5}$

Important Notes

1. For any event A, $0 \le P(A) \le 1$. (decimals)
2. The probability of an impossible event is zero.
3. The probability of an event that is certain to occur is one.
4. If necessary, round probabilities to three significant digits.

$P(A) = 0 \rightarrow A$ is impossible

$P(B) = 1 \rightarrow B$ will occur

$P(S) = \frac{6}{6} = 1$ for a dice roll.

Def **Unusual Events** An event that has a low probability. *Typically, an event with less than* 5 *% is considered unusual, but it depends on the problem.*

tru dat

impossible events

certain events

$0 \le P(A) \le 1$

unusual events

extremely likely

0 0.05 0.95 1

Rounding Rule: Probabilities round to 3 decimal places OR 3 significant figures (unless told otherwise)

THREE APPROACHES TO PROBABILITY

(1)	(2)	(3)
Relative Frequency Probability: *Based on data.* Formula: $P(A) = \frac{\text{\# frequency of event } A}{\text{\# of total trials}}$	**Classical Probability:** *Based on (hypothetical) equally likely outcomes.* Formula: $P(A) = \frac{\text{\# outcomes in } A}{\text{\# outcomes in } S}$	**Subjective Probability:** *The probability of event A is found by simply guessing or estimating its value based on knowledge of the relevant circumstances.*
Relative Frequency: 200 million people are affected each year by malaria, a disease carried by mosquitos. 600,000 of those people die. What is the probability that someone affected by malaria dies? $P(\text{someone w/ Malaria dies}) = \frac{600{,}000}{200{,}000{,}000} = \boxed{0.003}$	***Classical:*** When two children are born, what's the probability that both are the same gender? $S = \{GG, GB, BG, BB\}$ $A = \{GG, BB\}$ $P(A) = \frac{\text{\# same gender}}{\text{\# of outcomes in } S} = \frac{2}{4} = \boxed{0.5}$	***Subjective:*** What is the probability that the next dollar bill you spend was previously spent by Beyoncé? $\frac{1}{350 \text{ million}}$??? 0.000000001 ?

Ex: A bag has 1 red marble, 1 blue marble, 1 yellow marble, 1 orange marble, and 1 purple marble. The table below shows the results of choosing a marble out of the bag and replacing it each trial. Give answers as decimals and percentages.

(1)

Outcome of the Draw	100 trials	600 trials
Red	33	120
Blue	24	121
Yellow	18	119
Orange	17	122
Purple	8	118

(a) Find the relative frequency probability of drawing a yellow or an orange for the 100 trials.

$P(\text{Yellow or Orange}) = \frac{\#Y \text{ or } \#O}{100} = \frac{18+17}{100} = \frac{35}{100} = \boxed{0.35} = \boxed{35\%}$

(b) Find the relative frequency probability of drawing a marble that is not red for 100 trials.

$P(\text{not Red}) = \frac{\text{not } R}{100} = \frac{100-33}{100} = \frac{67}{100} = \boxed{0.67 = 67\%}$

(c) Find the relative frequency probability of drawing a blue for the 100 trials.

$P(\text{Blue}) = \frac{24}{100} = \boxed{0.24 = 24\%}$

(d) Find the relative frequency probability of drawing a blue for the 600 trials.

$P(\text{Blue}) = \frac{121}{600} = \boxed{0.202 = 20.2\%}$

(e) What is the classical probability of choosing a blue?

$S = \{R, B, Y, O, P\}$
$A = \{B\}$

$P(\text{Blue}) = \frac{\text{\# in } A}{\text{\# in } S} = \frac{1}{5} = \boxed{0.20 = 20\%}$

Def The **complement** (denoted $\bar{A}$) of event A, consists of all outcomes in which event A doesn't occur.

Ex $S = \{1,2,3,4,5,6\}$ $A = \{1,3,5\}$ $\bar{A} = \{2,4,6\}$

THE LAW OF LARGE NUMBERS

As a procedure is repeated again and again, the relative frequency probability of an event tends to approach the ~~actual~~ classical or theoretical probability.

\# trials increases

Ex: Flip coin 10 times, $P(H) = 3/10 = 0.3$
100 times, $P(H) = 55/100 = 0.55$
100,000 times, $P(H) = 50{,}033/100{,}000 = 0.5033$

Classical: $P(H) = 1/2 = 0.5$

Wed 14

~~Section 4.2: Addition Rule and Multiplication Rule~~

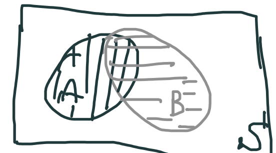
COMPOUND EVENTS

Def A **compound event** is any event combining two or more simple events.

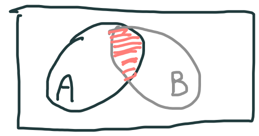
NOTATION: $P(A \text{ or } B)$ denotes the probability that event A occurs or event B occurs (or both.)

Reason we subtract is to undo the double-count!

FORMAL ADDITION RULE	
Symbolic	$\boxed{P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)}$
Meaning	*The probability of event A or event B is the sum of each event's probability of occurring individually, minus the probability of both events occurring simultaneously.*
Note	$P(A \text{ and } B)$ denotes the probability that A and B both occur at the same time. i.e "overlap"



A or B "union"



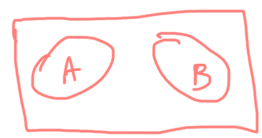
A and B "intersection"

IMPORTANT NOTE

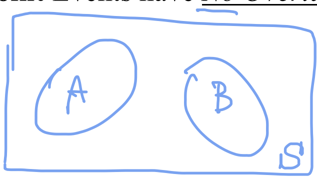
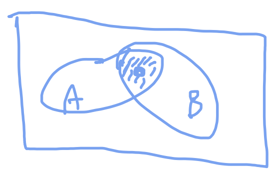
Ex E = roll an even O = roll an odd

Def Two events are **disjoint** (or **mutually exclusive**) if they cannot occur at the same time.

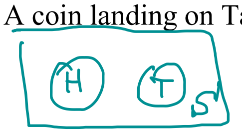
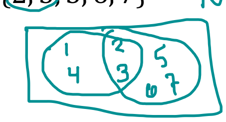
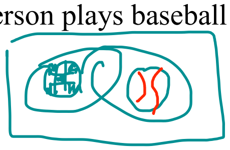
Note: If events A and B are mutually exclusive $\Rightarrow$ $\boxed{P(A \text{ and } B) = 0}$.



Can a compound ***be a*** disjoint ***?***

If NO...	If YES...
Disjoint Events have *No Overlap* 	**Not Disjoint** Events *have Overlap* 

Ex: Determine whether the following are disjoint events.

(a) **A** = A coin landing on Heads **B** = A coin landing on Tails  yes, disjoint! $P(A \text{ and } B) = 0$	(b) **A** = {1, 2, 3, 4} **B** = {2, 3, 5, 6, 7}  No, not disjoint $P(A \text{ and } B) \neq 0$
(c) **A** = person plays soccer **B** = person plays baseball  No, not disjoint	(d) **A** = Roll an even number on a 6-sided fair die. **B** = Roll an odd number on a 6-sided fair die. yes, disjoint.

GTQ

Ex: Let $P(E) = .11$, $P(F) = .78$, $P(G) = .56$, $P(F \text{ and } G) = \cancel{.6}$ 0.4, and events E and F are disjoint. E, F, G

(a) Find $P(F \text{ or } G)$

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

$= 0.78 + 0.56 - 0.4$

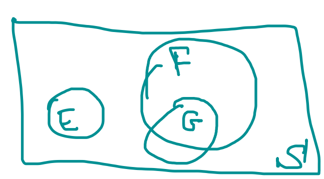
$= \boxed{0.94}$

(b) Find $P(E \text{ or } F)$

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

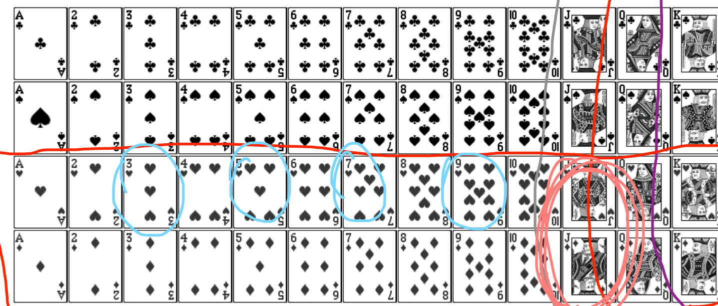
$= 0.11 + 0.78 - 0$

$= \boxed{0.89}$



Ex: Suppose that a single card is selected from a standard 52-card deck, such as shown below.

J Q K



J and R

(a) What is the probability that the card drawn is a king?

$P(K) = \frac{4}{52} \approx \boxed{0.077} = \boxed{7.7\%}$

← unsimplified fraction

(b) What is the probability that the card drawn is a king or a queen?

disjoint

$P(K \text{ or } Q) = P(K) + P(Q) - \underbrace{P(K \text{ and } Q)}$

$= \frac{4}{52} + \frac{4}{52} - 0 = 8/52 = \boxed{0.154} = \boxed{15.4\%}$

(c) What is the probability that the card drawn is a Jack or a red?

$P(J \text{ or } R) = P(J) + P(R) - P(J \text{ and } R)$

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

$= \boxed{0.538} = \boxed{53.87\%}$

(d) What is the probability that the card is an odd number or a heart?

$P(\text{Odd or Heart}) = P(\text{Odd}) + P(\text{Heart}) - P(\text{odd and Heart})$

$= \frac{4 \ast 4}{52} + \frac{13}{52} - \frac{4}{52} = \frac{25}{52} = \boxed{0.481}$

Ex: A study of 1,000 recently deceased people is summarized in the following table.

	Cause of Death			
	Cancer	*Heart Disease*	*Other*	
Smokers	100	180	120	Σ = 400
Non-smokers	100	120	380	Σ = 600
	Σ = 200	Σ = 300	Σ = 500	1000

Find the probability of randomly selecting:

(a) someone who died of cancer.

$P(\text{Cancer}) = \frac{200}{1000} = \boxed{0.2}$

(b) someone who did not die of cancer.

"complement"

$P(\overline{\text{Cancer}}) = \frac{800}{1000} = \boxed{0.8}$

Observation

$P(C) + P(\bar{C}) = 1$

$0.2 + 0.8 = 1$

(c) someone who died of heart disease and cancer.

$P(\text{HD and C}) = \boxed{0}$

(d) someone who died of heart disease or cancer.

$P(\text{HD or C}) = P(HD) + P(C) - \underbrace{P(\text{HD and C})}_{0}$

$= \frac{300}{1000} + \frac{200}{1000} = \frac{500}{1000} = \boxed{0.5}$

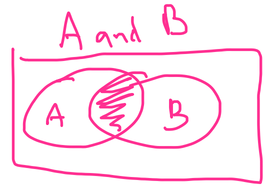
(e) someone who smoked or died of heart disease.

$P(\text{S or HD}) = P(S) + P(HD) - P(\text{S and HD})$

$= \frac{400}{1000} + \frac{300}{1000} - \frac{180}{1000} = \frac{520}{1000} = \boxed{0.52}$

COMPOUND EVENTS

Def A **compound event** is any event combining two or more simple events.



NOTATION: $P(A \text{ and } B)$ denotes the probability that event A occurs in the 1^{st} trial followed by the occurrence of event B in the 2^{nd} trial.

FORMAL MULTIPLICATION RULE	
Symbolic	$\boxed{P(A \text{ and } B) = P(A) * P(B \mid A)}$
Meaning	*The probability of event A followed by event B is found by multiplying the probability of event A by the probability of event B.*
Note	$P(B \mid A)$ denotes the conditional probability of event B occurring after it is assumed that event A has already occurred.

KEY multiply $P(A) * P(B)$

$$\boxed{P(B \mid A) = \frac{P(A \text{ and } B)}{P(A)}}$$

INDEPENDENCE VS. DEPENDENCE

Def Two events are **independent** if the occurrence of one event does not affect the probability of the occurrence of the other event.

Ex Flipping a coin consecutively "each new flip is independent"

Note: *If events A and B are independent* $\Rightarrow \boxed{P(B \mid A) = P(B).}$

If events A and B are independent $\Rightarrow \boxed{P(A \mid B) = P(A).}$

Def If two events are not independent, they are said to be **dependent**.

SUMMARY:

$$P(A \text{ and } B) \begin{cases} = P(A) \cdot P(B \mid A) & \text{(if } A \text{ and } B \text{ are dependent)} \\ = P(A) \cdot P(B) & \text{(if } A \text{ and } B \text{ are independent)} \end{cases}$$

Also works for 3 or more events

EX: **Independent events? Why or why not?**

(a) A = You find a parking spot B = First week of school independent vs dependent? Dependent since first week tends to be busier so it will make finding a parking spot tougher.	(b) C = You pass your class D = Good food at the Piazza independent!

EX: A bag contains an assortment of Jolly Rancher candies. Specifically, there are 5 apple, 8 watermelon, 10 cherry, and 15 grape flavored candies. You get to randomly select three candies without replacement.

(a) Find the probability of picking three grape Jolly Ranchers. (Dependent)

S = 38

$$P(1^{st}\text{ Grape and } 2^{nd}\text{ G and } 3^{rd}\text{ G}) = P(1^{st}\text{ G}) * P(2^{nd}\text{ G} \mid 1^{st}\text{ G}) * P(3^{rd}\text{ G} \mid 1^{st}\,\&\,2^{nd}\text{ G}) = \frac{15}{38} * \frac{14}{37} * \frac{13}{36}$$

$$= \boxed{0.054}$$

(b) Find the probability of not getting any apple Jolly Ranchers.

$$P(\overset{1^{st}}{\bar{A}} \text{ and } \overset{2^{nd}}{\bar{A}} \text{ and } \overset{3^{rd}}{\bar{A}}) = \frac{33}{38} * \frac{32}{37} * \frac{31}{36} = \boxed{0.647}$$

not Apple

$\# \bar{A} = 38 - 5 = 33$

Ex: In the table is the highest level of education information for 50 applicants for a job.

Bachelor's Degree	35
Master's Degree	15

\# S = 50

(a) If two of these fifty applicants names are chosen at random, without replacement, then what is the probability that the 1st selected has a Bachelor's degree and the 2nd has a Master's degree?

$$P(1^{st}\text{ BD and }2^{nd}\text{ MD w/o replacement}) = \frac{35}{50} * \frac{15}{49} = \boxed{0.214}$$

(independent)

(b) What would the probability in (a) be if replacement was allowed?

$$P(1^{st}\text{ BD and }2^{nd}\text{ MD w/ replacement}) = \frac{35}{50} * \frac{15}{50} = \boxed{0.21}$$

The 5% Guideline for Cumbersome Calculations

If a sample size is no more than 5% of the size of the population, treat the selections as being independent.

Ex: A quality control analyst randomly selects 3 different car ignition systems from a manufacturing process that has just produced 200 systems, including 5 that are defective.

$3/200 = 0.015 = 1.5\%$
Can use 5% Guideline
↳ treat as independent

(a) What is the probability that all 3 ignition systems are good?

$$P(1^{st}\text{ Good and }2^{nd}\text{ G and }3^{rd}\text{ G}) = \frac{195}{200} * \frac{194}{199} * \frac{193}{198} = \boxed{0.926}$$

(b) Use the 5% guideline for treating the events as independent, and find the probability that all 3 ignition systems are good.

$$P(1^{st}\text{ G and }2^{nd}\text{ G and }3^{rd}\text{ G}) \underset{\text{indep}}{=} P(G) * P(G) * P(G) = \frac{195}{200} * \frac{195}{200} * \frac{195}{200}$$

$$= \left(\frac{195}{200}\right)^3 = \boxed{0.927}$$

Section 4.3: Complements, Conditional Probability, and ~~Bayes' Theorem~~

COMPLEMENTS: THE PROBABILITY OF "AT LEAST ONE"

at least one day rain = 1R and 2R and 3R .....

Def The **complement** (denoted $\bar{A}$) of event A, consists of all outcomes in which event A doesn't occur.

Note: $\boxed{P(\bar{A}) = 1 - P(A)}$ $\Rightarrow$ $\boxed{P(\text{at least one}) = 1 - P(\text{none})}$

no rain VS at least one day of rain

Ex: Let's say that for the next seven days, the probability of rain is 5%. Assume ~~each~~ the chance of rain each day is independent. What is the probability that it rains ***at least one day*** over the next seven days.

$P(\text{rain}) = 0.05$

$P(\overline{\text{rain}}) = 1 - 0.05 = 0.95$

Ridiculously LONG way…

rain on 1st day only

$P(1R) * P(\overline{2R}) * P(\overline{3R}) \cdot P(\overline{4R}) \cdot P(\overline{5R}) \cdot P(\overline{6R}) \cdot P(\overline{7R})$

$+$ ("or")

rain on 2nd day only

$P(\overline{1R}) \cdot P(2R) \cdot P(\overline{3R}) \cdot P(\overline{4R}) \cdot P(\overline{5R}) \cdot P(\overline{6R}) \cdot P(\overline{7R})$

$+$

+ more where only 1 day of rain

$P(1R) \cdot P(2R) \cdot P(\overline{3R}) \cdot P(\overline{4R}) \cdot P(\overline{5R}) \cdot P(\overline{6R}) \cdot P(\overline{7R})$

+ more ways to have 2 days of rain

+ …… + (last) P(all 7 days rain)

The Sane Way… *What's the* ***complement*** *of at least one day of rain?*

$P(\text{at least one day of rain})$

$= 1 - P(\text{no days of rain})$

$= 1 - P(\text{no rain})$

$= 1 - P(\overline{1R}) \cdot P(\overline{2R}) \cdots P(\overline{7R})$

$= 1 - (0.95)^7 = \boxed{0.302}$

Ex: A bag contains an assortment of Jolly Rancher candies. Specifically, there are 5 apple, 8 watermelon, 10 cherry, and 15 grape flavored candies. You get to randomly select three candies without replacement. Find the probability of getting at least one watermelon.

$\# \text{bag} = 5 + 8 + 10 + 15 = 38$

$\# \bar{W} = 30$

$P(\text{at least one } W) = 1 - P(\text{no } W) = \boxed{0.519}$

$P(\text{no } W) = P(\bar{W} \text{ and } \bar{W} \text{ and } \bar{W}) = \frac{30}{38} \cdot \frac{29}{37} \cdot \frac{28}{36}$

$P(T) = 0.92$

$P(\bar{T}) = 1 - 0.92 = 0.08$

Ex: A satellite defense system has five independent satellites that each have a 0.92 chance of detecting a missile threat.

(a) What's the probability that at least one satellite does detect a missile threat?

T= missile threat one satellite

$P(\text{at least one } T) = 1 - P(\bar{T})$

$= 1 - (0.08)^5 = 0.99996\ldots = \boxed{1.00}$

(why 5? b/c 5 satellites)

(b) What's the probability that at least one satellite does not detect a missile threat?

$P(\text{at least one } \bar{T}) = 1 - P((\bar{\bar{T}}))$

$= 1 - P(T)$

$= 1 - P(\text{all } T)$

$= 1 - (0.92)^5 = \boxed{0.341}$

Q: Are these complements? Why not add up to 1?

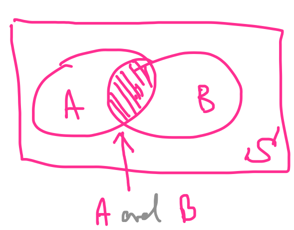
CONDITIONAL PROBABILITY

Def A **conditional probability** of an event is a probability obtained with the additional information that some other event has already occurred.

(See Baye's Rule Videos)

NOTATION: $P(B | A)$ denotes the conditional probability that event B occurs, given that event A has already occurred.

FORMULA: $$\boxed{P(B | A) = \frac{P(A \text{ and } B)}{P(A)}}$$



Reminder Use 3 sig figs!

Ex: Let A = Today is your birthday and B = Your birthday is in this month.

(a) Are events A and B dependent?

yes, dependent b/c B effects A.

(b) $P(A) = \frac{1}{365} = \boxed{0.00274}$

(c) $P(A|B)$

= P(today BD | BD is in Jan) → 0 → depends!! 1/31 ?

depends!

(d) $P(B|A)$ = P(BD is Jan | today is BD) = $\boxed{1}$

b/c if know your BD is Jan 16, then it is Jan!

Ex: The following table gives the mortality data for passengers of the Titanic.

	Men	*Women*	*Children*	
Survived	332	318	56	Σ = 706
Died	1360	104	53	Σ = 1517
	Σ = 1692	Σ = 422	Σ = 109	2223

Find the probability of randomly selecting:

(a) a passenger who died, given that the person was a man.

$P(\text{Died}|\text{Man}) = P(D|M) = \frac{P(D \text{ and } M)}{P(M)} = \frac{1360/\cancel{2223}}{1692/\cancel{2223}} = \frac{1360}{1692} = \boxed{0.804}$

(b) a woman, given that the passenger survived.

$P(W|\text{Survived}) = \frac{P(W \text{ and } S)}{P(\text{Surv})} = \frac{318}{706} = \boxed{0.450}$

use short-cut

(c) a survivor, given that the passenger was a child.

$P(S|C) = \frac{P(S \text{ and } C)}{P(C)} = \frac{56}{109} = \boxed{0.514}$

Ex: The table to the right shows the status of 200 registered college students.

	Part Time	Full Time	Total
Female	80	40	120
Male	60	20	80
Total	140	60	200

(a) What is the probability that a part time student is female?

↳ given part time

$P(\text{Female}_{Fe}|\text{Part time}_{PT}) = \frac{P(Fe \text{ and } PT)}{P(PT)} = \frac{80}{140} = \boxed{0.571}$

(b) What is the probability that a randomly selected student is part time, given that they are a male?

$P(PT|\text{Male}) = \frac{P(PT \text{ and } M)}{P(M)} = \frac{60}{80} = \boxed{0.75}$

(c) What is the probability that a randomly selected female is a full time student?

↳ given that female

$P(FT|Fe) = \frac{P(FT \text{ and } Fe)}{P(F)} = \frac{40}{120} = \boxed{0.333}$