## Coming Home to Math

## Questions and Answers

March 15, 2023
Irving P. Herman
A Supplement to Coming Home to Math
dedicated to
Ezra Rhys Herman, who has never answered a math question incorrectly

Please use but do not distribute, copy or modify
Send inquiries to iph1@columbia.edu with subject: Coming Home to Math questions Check for future versions with additional questions and answers

These are questions for the book Coming Home to Math, along with answers. This includes questions addressing the math topics covered in the published version, and those meant to introduce new material, new examples (sometimes extracted and adapted from the news), and updated information. This list will be continually updated.

Most problems are easy, very straightforward and short, and are noted with an *. (When they involve the cultural use of numbers not noted in the book, you can search online about them.) Very slightly more involved questions are denoted with an **. A few problems are more involved or advanced and are noted with an ***.

Updates about this file and further information will be at
http://www.irvingpherman.com/coming-home-to-math/ . More information about Coming Home to Math can also be found at http://www.facebook.com/Coming-Home-to-Math106010094462401/, https://www.amazon.com/Coming-Home-Math-ComfortableNumbers/dp/9811211264 and https://www.worldscientific.com/worldscibooks/10.1142/11540.

## Part I: Our World of Math and Numbers

## Chapter 1. Introduction

* Ch. 1. Teaching about our world of math is constantly changing (a math joke)

Consider the math humor (joke) on The Evolution of Math Teaching:
1960s: A peasant sells a bag of potatoes for $\$ 10$. His costs amount to $4 / 5$ of his selling price. What is his profit?
1970s: A farmer sells a bag of potatoes for $\$ 10$. His costs amount to $4 / 5$ of his selling price, that is, $\$ 8$. What is his profit?
1970s (new math): A farmer exchanges a set P of potatoes with set M of money. The cardinality of the set M is equal to 10 , and each element of M is worth $\$ 1$. Draw ten big dots representing
the elements of M . The set C of production costs is composed of two big dots less than the set M . Represent C as a subset of M and give the answer to the question: What is the cardinality of the set of profits?
1980s: A farmer sells a bag of potatoes for $\$ 10$. His production costs are $\$ 8$, and his profit is $\$ 2$. Underline the word "potatoes" and discuss with your classmates.
1990s: A farmer sells a bag of potatoes for $\$ 10$. His or her production costs are 0.80 of his or her revenue. On your calculator, graph revenue vs. costs. Run the POTATO program to determine the profit. Discuss the result with students in your group. Write a brief essay that analyzes this example in the real world of economics.

Which do you consider the best shown teaching method? How would you update it for more recent decades?
[From: Anon: Adapted from The American Mathematical Monthly, Vol. 101, No. 5, May 1994 (Reprinted by Stan Kelly-Bootle in Unix Review, Oct 94); 10-7-20 website accessed, in 3. Math.education section https://www.math.utah.edu/~cherk/mathjokes.html.]

Answer: The best is a matter of opinion. I like the first method. I won't hazard an opinion on how they could be updated.

## * Ch. 1. Math dislike explored on the classic Andy Griffith TV show, or was it?

On the classic Andy Griffith TV show, Opie Taylor ("Ronnie" Howard) did very poorly on an arithmetic test (on which he received a D , which was much worse than his previous performance on math tests and which was incorrectly portrayed as him failing math) (Season 5, Episode 39, Opie Flunks Arithmetic, April 19, 1965-and often shown in reruns). He said he hated long division.
(a) Do you think that Opie had a genuine dislike of math?

When learning of Opie's performance, Deputy Sheriff Barney Fife (played by Don Knotts) proclaimed to Opie's father Sheriff Andy Taylor (Andy Griffith) in his usual excited manner: "Arithmetic's probably the most important subject there is. How do expect this country to maintain its position in world leadership if our kids are flunkin' arithmetic?" He also said that thought that Opie's future was over unless he shaped up.
(b) Were these reasonable comments to make?
(c) If you were Opie's father, how would have you responded to the situation?

Answer: (a) No. He had some challenges and needed some help. (b) Barney Fife was going overboard as usual, but he was correct that a public with math skills is important for its success and it would be helpful for Opie's if he became more proficient in arithmetic. (c) Opie's father Sheriff Andy Taylor initially ignored Barney, but then realized that Opie's future could be limited by poor math skills. Uncharacteristically, Andy went overboard and made Opie cut out all non-school activities and to study all of the time, but then he found a reasonable middle ground. On his next arithmetic test, Opie did much better and received a $B+$. https://www.metv.com/quiz/how-well-do-you-remember-barney-fifes-last-episode-as-a-mayberry-regular
[https://www.imdb.com/title/tt0798632/]

## * Ch. 1. Division and thinking in the classic Leave it to Beaver TV show

In the classic "Leave it to Beaver" TV show, $5^{\text {th }}$ grader Beaver (Theodore Cleaver) asked his father, Ward: "With business, how often do you have to invert fractions?" (This was because he was having trouble dividing fractions. See Chapter 4.) His wise father Ward, replied: Well almost never, Beaver, but that's not the question. You see, doing problems like these teaches you how to think." (Season 4, Episode 27, Beaver's Report Card, April 1, 1961-and often shown in reruns). Was Ward giving sage advice?

Answer: Yes.

## Chapter 2. We Use Numbers Here, There and Everywhere

* Ch. 2. Getting back to using the number six

What do the expressions "on your six," "I got your six," and "watch your six" mean?
Answer: In military parlance, "on your six" means directly behind you. This comes from the military/pilot designation of position using a clock, with 12:00 in front of you, 3:00 to your right, 6:00 to your back, and 9:00 to your left. It means directly behind you. "I got your six" means I have your back (meaning that I will protect you). To watch your six means to watch your back. [https://www.urbandictionary.com/define.php?term=on\ your\ six
https://yoursix.com/about/
both accessed 1-20-21
https://www.urbandictionary.com/define.php?term=on\ your\ six
https://www.urbandictionary.com/define.php?term=I\ got\ your\ six]

## * Ch. 2. Living better with the number three

What is "The Rule of Threes for Survival"?
Answer: You can survive being $\sim 3$ min with no air, $\sim 3 \mathrm{hr}$ with no shelter (under harsh conditions), $\sim 3$ days with no water, or $\sim 3$ weeks with no food (but with water). These are rough estimates.
[https://en.wikipedia.org/wiki/Rule_of_threes_(survival)]

## * Ch. 2. Expressions with the numbers six and seven

In the 1957 Perry Mason novel by Erle Stanley Gardner The Case of the Screaming Woman, the private detective Paul Drake working with the attorney Perry Mason asks him how the murder trial was going and Mason responded "Sixes and sevens"? (I did know what this meant when I read it.) Was the trial proceeding well or poorly?

Answer: Poorly. The idiom "at sixes and sevens" means in a state of complete disarray and confusion, so the case was proceeding poorly and the best way to proceed was not clear. It is more English than American. This expression first meant careless risk, as in throwing a 6 or 7 in a specific dice game that risked one's money, but then evolved to mean confusion, perhaps because someone would have to be confused to take on such risk. [https://grammarist.com/idiom/at-sixes-and-seven/ ]

## * Ch. 2. Expressions with the numbers seven and six

What would you think the response to a casual question concerning your health and other matters, "the same old seven and six," would mean knowing that the common weekly wage among English workmen sometime in the $19^{\text {th }}$ century was seven shillings and six pence? [https://math.answers.com/Q/Same_old_six and sevenanaser]

Answer: All is usual (as was that wage).

## * Ch. 2. Don't be 86ed

What you hear that someone or something has been or is about to be 86ed, what does it mean?
Answer: It means that someone has been ejected, barred, or refused service (as a customer), or terminated, or something has been rejected, discarded, cancelled or not available. This expression is more common than I had once thought, and I now keep on noticing it in many contexts. It may have been an inspiration for the lead character's code name, Agent 86, in the 1960s comedy show "Get Smart." There are numerous explanations for its origin, so most of them are wrong. One intriguing, though still quite suspect one, relates to the now archaic rotary dial phone in which the digits 1-8 (of 0-9) were at the same location on the rotary dial as sequences of three letters in alphabetical order (all except $Q$ and Z). 8 was linked to TUV and 6 to $M N O$, and so 86 could stand for TO, which could stand for "thrown out." (https://www.stlmag.com/dining/Ask-George-Where-Does-the-Term-86d-Come-From/, retrieved 6-17-2021)

## * Ch. 2. Half of a dozen by any other name

What does the expression " 6 of one and half a dozen of another mean?"
Answer: Since both denote 6, it means that both categories are the essentially equivalent.

## * Ch. 2. The meaning of the numerical name of a TV show

Does the name of the 2014-2017 Canadian police drama TV show, 19-2, indicate a crime statistic, a label, a code for police activity, or another way of saying 17 ?

Answer: It is a label, for car \#2 in the $19^{\text {th }}$ police station.

## * Ch. 2. Does your date of birth follow you your whole life?

Do the numbers that describe your birthday influence your unconscious mind? For those born on February 2 (or $2 / 2$ ) would you think there is a preference for them to live in places with the number 2 referenced in some way, such as in Twin Lakes, Wisconsin, on March 3 (or 3/3) in places with a 3, such as Three Forks, Montana, and on June 6 (or $6 / 6$ ) in places with a 6, such as Six Mile, South Carolina?

Answer: There is a small, but still statistically significant (see Sections 16.1, 16.2.3, and 16.4.1) preference for people to choose to live in places with a number or number reference in its name that relates to their birthday.
[Incognito by David Eagleman (page 63) and sources cited therein]

## * Ch. 2. Fingers do count

The numbers used in different cultures are the same, aside from differences in using commas, periods and so on (Section 4.2). We count numbers the same way, but we use our fingers to count differently in different cultures. If we use three neighboring fingers on one hand, in how many ways could we indicate the number 3 ? How do you do this?

Answer: There are three ways to indicate 3. In the U.S. we use the middle three. In Western Europe they use the thumb and the index and middle fingers. This difference was part of the plot of the movie "Inglorious Basterds." You might want to do a web search to explore the full use of fingers in counting in different cultures. Look online to see the many ways that fingers are used in counting.

## Chapter 3. Numbers are Some of My Favorite Things

* Ch. 3. Flying, movies, and math (a math joke)

On Sesame Street, how did Cookie Monster, playing the role of Alistair Cookie of Monsterpiece Theatre (takeoffs of Alistair Cooke and Masterpiece Theater), introduce the movie "One Flew over the Cuckoo's Nest?"

Answer: By showing a cartoon with the number 1 literally flying over a cuckoo's nest, as in " 1 Flew over the Cuckoo's Nest. https://www.youtube.com/watch?v=dSiVZ524yW4. https://muppet.fandom.com/wiki/Monsterpiece Theater

## * Ch. 3. Number palindromes

(a) Give all the number palindromes for the days in 2018: month-day-year (using 18 for 2018).
(b) Of these, which has the fewest number of different digits? (c) Of these, which has the largest sum of digits?

Answer: (a) 11 of them: 81 18, 810 18, $81118,81218,81318,81418,81518,81618,8$ 1718,81818 , and 81918 . (b) 3 of them have 2 digits ( 1 and 8): 8118,81118 , and 81818. (c) 81818 .

## ** Ch. 3. More number palindromes

(a) Find the 22 palindrome dates in 2021 (all with the year written as 21 ), and find what is the longest stretch of them in a row.
(b) Show that the U.S. Inauguration Day January 20, 2021 is also a 7-digit palindrome date (with the year written as 2021).

Answer: (a) They are: 1-2-21, 1-20-21, 1-21-21, 1-22-21, 1-23-21, 1-24-21, 1-25-21, 1-26-21, 1-27-21, 1-28-21, 1-29-21, 12-1-21, 12-2-21 12-3-21, 12-4-21, 12-5-21, 12-6-21, 12-7-21, 12-8-21, 12-9-21, 12-11-21, and 12-22-21, which includes one four-digit, 19 five-digit, and two six-digit palindrome dates, and 10 in a row-from 1-20-21 to 1-29-21. There are as many as 22 palindrome dates in only two years every century-those ending in 11 and 21. (Why?) (b) 1-202021. This is the first 7-digit palindrome U.S. Inauguration Day in American history, with the next one in 1,000 years on Jan. 20, 3021. There are only 26 seven-digit palindromes dates in the $21^{\text {st }}$ century. Note that 12-2-21 is also the eight-digit palindrome 12-02-2021.
[Inauguration Day falls on a rare palindrome date: That won't happen again for 1,000 years. USA TODAY https://apple.news/ACF4aoIjHSne21LjhwKCnWA]

## * Ch. 3. Math analysis by limericks (math joke) (Also see Chapter 4 and Section 5.4.)

Explain the limerick (a mathematical limerick):
A dozen, a gross, and a score
Plus three times the square root of four
Divided by seven
Plus five times eleven
Is nine squared and not a bit more.
(In limericks, the longer lines 1,2 , and 5 rhyme as do the shorter lines 3 and 4.)

## Answer:

A dozen, a gross, and a score $(12+144+20=176)$
Plus three times the square root of four $(176+3 \times 2=182)$ (because the square root of 4 is 2 )
Divided by seven $(182 / 7=26)$
Plus five times eleven $(26+5 \times 11=81)$
Is nine squared and not a bit more. $(81+0=81)$ (because 81 is the square of 9 )
[https://en.wikipedia.org/wiki/Mathematical_joke and source cited therein Leigh Mercer]

## Part II: The First World of Math: The Eternal Truths of Math (or the Math of What Is, Always Was, and Will Always Be)

## Chapter 4. Linking Numbers: Operations on Numbers

## * Ch. 4. Making subtraction easier by using addition

Show how the subtraction problem 152-79 can be recast as finding the (magnitudes of the) differences between 152 and 100 and then 79 and 100, and then adding them. (Sometimes this approach makes it easier to do arithmetic in your head.)

Answer: The two differences are 52 and 21, which sum to the correct answer 73. The difference between 152 and 100 is obvious. The difference between 79 and 100 can be handled by seeing how much you need to add to 79 to get to 100. This approach can be simpler when there are carryovers (and I use it all of the time).

## ** Ch. 4. Subtraction using what was called the "New Math" - Basics

About a half century ago, renowned satirist (and mathematician) Tom Lehrer wrote a song "New Math" that poked fun at the then new methods of teaching math to school children. It was often called the New Math, and quite a few thought it unduly emphasized approaches that were too advanced and abstract for beginners, with too little emphasis on actual computation. He performed it often, and its lyrics (and preamble), his singing of it, his live performances, and annotated versions of him singing of it are widely available, including online. (a) Without the use of calculators, computers and so on, perform the subtraction example Lehrer (comically) highlighted in his song: 342 - 173. (b) He did this subtraction using the conventional version of "old" math and one (funny) variant of it, along with his version of doing it using the New Math. How do they compare? Which way did you use to solve the problem? Is the New Math method clear to you? Did it help you understand how to do subtraction better?

Answer: (a) 169. (b) All the methods are mathematically equivalent, but you might find the New Math to be too laborious.

## * Ch. 4. Division and upside-down fractions in the classic Leave it to Beaver TV show

 In the classic "Leave it to Beaver" TV show, $5^{\text {th }}$ grader Beaver (Theodore Cleaver) was having trouble in doing a division problem with fractions and his older brother Wally told him to turn the fraction upside down (Season 4, Episode 27, Beaver's Report Card, April 1, 1961-and often shown in reruns). What was Wally talking about?Answer: To divide a fraction by second fraction, you need to invert the second fraction (turn it upside down, meaning to exchange the numerator and denominator) and then multiply them (and so multiply their numerators and then their denominators).

## * Ch. 4. Can you give more than your entire effort?

How should you respond to someone claiming they are giving $110 \%$ effort to a project?
Answer: You should be skeptical. It is nonsense to say you are giving 110\% effort because the maximum is $100 \%$, but not that you are giving to charity this year $110 \%$ of what you gave last year.

## ** Ch. 4. Math thinking: Does it make math sense for parts to sum to more than $\mathbf{1 0 0 \%}$ ? (Chapter 4, Percentages and the whole)

In a video for a course, the instructor said that a 2015 study of 254 autistic individuals showed that $62.6 \%$ of them had had exceptional (brain-related) "talents," $57.5 \%$ of autistic people had exceptional (brain-related) "strengths," and that $7.2 \% \%$ had both traits. Are these numbers consistent with each other?

Answer: No. The fraction of those with any type of such traits cannot exceed 100\%. So, the sum of those with only one of the traits must sum to $100 \%$ or less, after removing the fraction with both (to avoid double counting). The fractions with the first and second traits sum to $120.1 \%$, less $7.2 \%$ gives $112.9 \%$. Something is wrong. (Any round-off errors are too small to account for this large discrepancy.)

The Great Courses, Understanding Disorders of the Brain, Lecture 5, Autism Spectrum Disorders, Part 2, Sandy Neargarder.

## ** Ch. 4. Life's percentages

What is $102 \%$ (a) times, (b) of, (c) more than, and (d) less than 300 ?
Answer: $102 \%$ equals 1.02. (a) times: $1.02 \times 300=306$, (b) of means the same as times: $1.02 \times$ $300=306$, (c) $102 \%$ of 300 is 306 and 306 more than 300 is $300+306=606$, (d) this means $1.02 \times 300=306$ less than 300 , or $300-306=-6$. Apparently, some people are comfortable with using percentages less than $100 \%$, such as $98 \%$, but are not for those larger than $100 \%$, such as 102\%. (see reference)
(https://www.wsj.com/articles/consumers-percentages-marketing-study-11653506251)

## ** Ch. 4. What is the basis of basis points? (Chapter 4, Percent; Section 5.4, Words describing numbers, myriad)

On June 15, 2022, the U.S. Federal Reserve announced the interest rate would increase by 75 basis points. This was reported elsewhere as a $0.75 \%$ rate increase? Why?

Answer: Percent means hundredths. A hundredth of a percent--a ten thousandth--is called a basis point (abbreviated as bp), so 75 basis points and $0.75 \%$ increases are the same. Basis points are commonly used in finance, especially when discussing interest rate changes. When interest rate changes are less than a $1 \%$, the use of basis points enables the use of integers in the description, such as 75 bps, rather than fractional numbers, such as $0.75 \%$. The symbol for a
basis point is similar to that of the percent symbol-except in the lower right portion there are three circles (next to each other), \%oo, instead of one of them for percent, $\%$. One part per ten thousand is (rarel)y called a permyriad, meaning one per 10,000 (or myriad).
(https://www.investopedia.com/ask/answers/what-basis-point-bps/, https://en.wikipedia.org/wiki/Basis_point)

## ** Ch. 4. How square is your acre?

The area of a square is its length squared. A mile is 5,280 feet long. There are 640 acres in a square mile (which, let us say, is a square with length a mile). Let's consider squares that have an area of an acre. What is the length of this square acre?

Answer: The mile square has an area of $5,280 \times 5,280=27,878,400$ square feet. Dividing this into 640 equal sections gives $27,878,400 / 640=43,560$ square feet for each acre. The length of the side of a square acre is the square root of this number, $\sim 208.7$ feet.

## ** Ch. 4. A dunam vs. a square mile

A dunam is a unit of land area that is used in regions of the former Turkish empire, such as in Israel, where it has an area of 900 square meters. A meter is 39.37 inches. (a) What is the length in feet of the side of a square with area of 1 dunam? (b) How many dunams are there in a square mile? (c) Which is larger, an acre or a dunam? (See the previous problem.)

Answer: (a) Because 900 is the square of 30, the length of a dunam square is 30 meters. There are 12 inches in a foot, so this length is $30 \times 39.37=1,181.1$ inches or $(30 \times 39.37) / 12=98.425$ feet. (b) The area of a dunam is $98.425 \times 98.425=9,687.48$ square feet, so there are $(5,280 \times$ $52,80) /(98.425 \times 98.425) \sim 2,878$ dunams in a square mile. (c) The length of the side of a square acre is $\sim 208.7$ feet (previous problem), while that of a dunam square is 98.425 feet, so the acre is larger, and larger by a factor of $(208.7 / 98.425)^{2} \sim 4.496$.

## * Ch. 4. Nonsense math in (TV) space exploration

In the original Star Trek TV series, Captain Kirk tried to locate a missing crew member on the Starship Enterprise by increasing the audio sensitivity of the computer to enable hearing this person's heartbeat. He explained this by saying: "Gentlemen, this computer has an auditory sensor. It can, in effect, hear sounds. By installing a booster, we can increase that capability on the order of one to the fourth power. The computer should bring us every sound occurring on the ship." He succeeded though his instructions were total math nonsense. Why? (Season 1, Episode 15, Court Martial)

Answer: One to the fourth power is one, and so this would mean no increase in sensitivity at all.

## * Ch. 4. The minuses have it, sometimes

What are $-183,596,219$ and $-183,596,220$ ?

Answer: -1 and 1, because a number with a magnitude of 1 (and so, -1 and 1) raised to any power will have magnitude 1, and a negative number raised to an odd power is negative (such as $-1^{1}=-1$ or $-1^{3}=-1 \times-1 \times-1=-1$ ) and one raised to an even power is positive (such as $-1^{2}=-1 \times$ $-1=1$ or $\left.-1^{4}=-1 \times-1 \times-1 \times-1=1\right)$.

## * Ch. 4. This order or that order? (Section 4.1)

(a) What are $2+3+4,(2+3)+4$, and $2+(3+4)$ ?
(b) What are $2 \times 3 \times 4,(2 \times 3) \times 4$, and $2 \times(3 \times 4)$ ?
(c) What are $2+3 \times 4,(2+3) \times 4$, and $2+(3 \times 4)$ ?
(d) What are $2 \times 3+4,(2 \times 3)+4$, and $2 \times(3+4)$ ?

Answer: (a) 9, 9, and 9. The parentheses don't change anything. (b) 24, 24, and 24. The parentheses don't change anything. (c) 14 (the multiplication occurs before the addition), 20 (operations inside the parentheses occur first), and 14 (operations inside the parenthesis occur first; it does not change things here but makes it clearer). (d) 10 (the multiplication occurs before the addition), 10 (operations inside the parentheses occur first; it does not change things here but makes it clearer), and 14 (operations inside the parentheses occur first).

## * Ch. 4. Mind your parentheses? (Section 4.1)

What are (a) $2 \times 3+4 \times 5$, (b) $(2 \times 3)+(4 \times 5)$, (c) $2 \times(3+4) \times 5$, (d) $(2 \times 3+4) \times 5$, and (e) 2 $\times(3+4 \times 5\}$ ?

Answer: By doing what is in a parenthesis first, doing $\times$ before + , and going from left to right: (a) 26, (b) 26, (c) 70, (d) 50, and (e) 46.

## * Ch. 4. Know your pluses and minuses (Section 4.1)

What are (a) $2+3-4+5$, (b) $(2+3)-4+5$, (c) $2+3-(4+5)$, (d) $2+(3-4)+5$, (e) $(2+3-$ 4) +5 , and (f) $2+(3-4+5)$ ?

Answer: By doing what is in a parenthesis first and then going from left to right: (a) 6, (b) 6, (c) -4, (d) 6, (e) 6, and (f) 6. The order rules lead to a difference here only for (c).

## * Ch. 4. Knowing when to multiply and when to divide (Section 4.1)

What are (a) $2 \times 3 \div 4 \times 5$, (b) $(2 \times 3) \div 4 \times 5$, (c) $2 \times 3 \div(4 \times 5)$, (d) $2 \times(3 \div 4) \times 5$, (e) $(2 \times 3 \div$ 4) $\times 5$, and (f) $2 \times(3 \div 4 \times 5)$ ?

Answer: By doing what is in a parenthesis first and then going from left to right: (a) 30/4 = $15 / 2$, (b) $30 / 4=15 / 2$, (c) $6 / 20=3 / 10$, (d) $30 / 4=15 / 2$, (e) $30 / 4=15 / 2$, and (f) $30 / 4=15 / 2$. The order rules lead to a difference here only for (c).
** Ch. 4. The units you use with numbers just about always matter (Section 4.4)
Someone tells you the temperature somewhere is a very cold -40 degrees. You ask if this is in degrees Celsius or Fahrenheit. The person says it doesn't matter. Why?

Answer: -40 degrees corresponds to the same temperature in both scales. For example, starting in Fahrenheit, $\left.T\left({ }^{\circ} C\right)=\left(T\left({ }^{\circ} F\right)-32\right) \times 5 / 9\right)=(-40-32) \times 5 / 9=-72 \times 5 / 9=-72 / 9 \times 5=-8 \times 5=$ -40. Use the equation that gives readings in Fahrenheit from those in Celsius to show this works in reverse as well.
** Ch. 4. Storing vaccines (Section 4.4)
A vaccine needs to be stored at -20 degrees Celsius to remain effective. What does this correspond to in degrees Fahrenheit?

Answer: - 4 degrees Fahrenheit, by using $T\left({ }^{\circ} F\right)=9 / 5 \times T\left({ }^{\circ} C\right)+32=9 / 5 \times(-20)+32=-36+32$ $=-4$.

## * Ch. 4. Much more may not give you much more and much much more may give you much less (Saturation and Relations in Section 4.4)

A certain amount of watering gives you a good yield of crops. Increasing the watering (amount per time interval) from the first amount by $40 \%$ increases the yield by $40 \%$, increasing it from the first amount by $80 \%$ increases it by $60 \%$, increasing it from the first amount by $200 \%$ increases it by $65 \%$, increasing it from the first amount by $400 \%$ increases it by $66 \%$, and then increasing it from the first amount by $600 \%$ decreases it by $90 \%$. What is happening? Does this make sense? Is such a dependence described by a function or only by a relation?

Answer: Increasing watering by up to $400 \%$ show classic linear increases that eventually saturate. Further increases indicate damage to the crop (washing it away, ...) and less yield. This does make sense. Because one level of watering leads to only one yield it is a function

## ** Ch. 4. How daylight savings time functions or Does anybody really know what time it is? (Section 4.4, Linking Numbers by Relations)

(This is a real-life example of when something is or is not a function.) On most days, the relation between the actual time, say as measured as a continually increasing number of minutes, and what the clock says has a well-defined meaning. There is a relation or mapping that uniquely assigns the actual time to the clock time, and so it is a function, and the reverse mapping from clock to actual time is also a function. (A function maps one or more items in one group uniquely to one item in another group. Mapping one item to more than one item is not a function.)
However, in the spring morning that daylight savings time begins, as we spring clocks forward at 2:00 AM to 3:00 AM, and the fall morning when it ends and clocks fall backward from 2:00 AM to 1:00 AM, this mapping from actual to clock time and from clock to actual time may not be functions. Which of these 4 mappings (actual to clock times and vice versa, in both the spring and fall) is a function?

Answer: In the morning of that spring day when daylight savings time begins, both mappings are functions. The clock springs forward at 2:00 to 3:00 and so the clock time proceeds as ... 1:59, 2:00 (which is identical to 3:00), 3:01, ..., and there is a unique mapping in going from the actual time to clock time and vice versa, and so both are functions. In the morning of that fall day when daylight savings time ends, the clock falls backward at 2:00 to 1:00 and so as real time proceeds minute by minute the clock time proceeds as 12:59, 1:00, 1:01, ... 1:59, 2:00 (which is identical to 1:00), 1:01, 1:02, ..., 1:59, 2:00, 2.01, ..., and so on, and so the bold clock times are repeated an hour later as the bold underlined clock times. Therefore, between the times on the clock from 1 and 2 AM, two actual times map to the same clock time. For example, the first 1:30 and then the second one that occurs an hour later both map to 1:30 clock time, and therefore this is a function. However, every clock time between 1 and 2 AM, including 1:30, maps to two actual times (separated by 60 minutes), and therefore this mapping is not a function. (Make a diagram that maps these times to show this.)

## ** Ch. 4. Pi days and perfect squares (Section 4.4 and and Chapter 3)

Show that the 7 digits that comprise the first five perfect squares, when rearranged, are also the first 7 digits in pi after the decimal point. What is the significance of this?

Answer: The first five perfect squares are: 1, 4, 9, 16, and 25, which have the seven digits 1, 4, 9, $1,6,2,5$. Their order can be rearranged to give 1, 4, 1, 5, 9, 2, 6. These are the first seven digits in pi after the decimal point: 3.1415926... This is a coincidence and has no significance (but you might find it interesting).

## ** Ch. 4. Linking differences to the Roman calendar (Section 4.5)

 According to legend, The Roman calendar started with the foundation of the city or Rome by Romulus and Remus on April 21, 753 BC. If 753 BC is considered year 1, what is year is 2020 AD ?Answer: $753+2020-1=2772$, since the year after $1 B C$ was $1 A D$-and there was no year 0 . (The calendar information is from "The Rise of Rome Great Course" Season 1, Episode 2)

## ** Ch. 4. Counting on you to find out how old Hari Seldon was when he died, as noted in the science fiction classic The Foundation? (Differences, Section 4.5)

Isaac Asimov's classic science fiction book, Foundation, which was the first of his three novels in the original The Foundation Trilogy, begins: "HARI SELDON_... born in the 11,988th year of the Galactic Era; died 12,069 . The dates are more commonly given in terms of the current Foundational Era as -79 to the year 1 F.E."." Remembering the need to be careful in counting between B.C. and A.D. years in our usual calendar because there is no year 0 , is this conversion of calendars in The Foundation correct? How old was Hari Seldon when he died?

Answer: No, but for a different reason. The 11,988th year of the Galactic Era corresponds to the year 11,988, because the first year in an era is known as Year 1, and so on. This means that Hari Seldon lived in parts of 82 years $(82=12,069-11,988+1$, so there are 82 numbers
(inclusively) in counting from 11,988 to 12,069)), and he died at the age 81 or 82. Because there seems to be a Year 0 in the Foundational Era calendar, using this calendar Hari Seldon lived in parts of 81 years $(=1-(-79)+1$, so there are 81 numbers (inclusively) in counting from -79 to 1), and he died at age 80 or 81 . If there were no 0 in the Foundational Era calendar, he would have lived in parts of 80 years and died at age 79 or 80 . Something seems to be wrong in converting between the two calendars.

## ** Ch. 4. Linking differences to stacks of money (Section 4.5)

In the TV show "Inspector George Gently" (Season 2, Episode 4, Gently Through the Mill), Chief Inspector Gently asks a murder suspect to remove the 500 pounds (of British money) that the suspect said he had put in his safe, and then asks him for the serial number of the first and last note in the stack of presumably 5-pound notes. The suspect said the serial numbers on the first and last notes were D64741271 and D64741371, which Gently confirmed and then he noted that the notes were sequential. What is mathematically wrong with this?

Answer: It is mathematically inconsistent. If all notes were in the sequence (and so are consecutive), there would be 101 notes or 505 pounds, so something was wrong. (This mistake was not a clue in solving the murder, and was presumably merely an error in the show.) Possibly, there were really 505 and not 500 pounds, one note was missing in the sequence, the assumption that the notes were (entirely or largely) sequential was wrong, or .... If the serial numbers on the first and last notes had been D64741271 and D64741370, all would have been consistent.

## ** Ch. 4. Linking differences to characteristic historical times (Section 4.5)

There are characteristic times in many technical events, but are there such times in history? In the 2019 book Upheaval, the author Jared Diamond noted that there is a common theme within 4 significant pairs of events in German history, in which the latter ones seem to be influenced by the former ones: the failed unification attempt in 1848 and the successful one in 1871, Germany's defeat in WWI in 1918 and it starting WWII in 1939, its defeat in WWI in 1945 and the German student revolt in 1968, and this student revolt in 1968 and German reunification in the 1990s. What is this theme?

Answer: The former and latter ones are respectively separated by 23, 21, 23, and 22 years, and so 21-23 years. This might be just a coincidence or more causal, and due to the amount of time that needed to pass before the former ones help cause the latter ones-and possibly for a new generation to appear.

## ** Ch. 4. Are rankings in different categories linked? (Sections 4.6 and 18.4.1)

If a group is ranked as the first, second, third or so on most in one numerical category and as the first, second, third or so on most in another, what is their ranking in their sum? For example, the total yardage accumulated by a football team is the sum of their yards gained by rushing (running) and that by passing. The teams can be ranked by the number of yards by rushing, by passing, and their sum. In a league of 7 teams, call A to G, team D is ranked \#4 in most yards by
rushing, say with 1,000 yards, and also \#4 in most yards by passing, say also with 1,000 yards. One person says that they would be ranked \#4 in their sum, which would be 2,000 yards. Another says that that is possible, but they could be ranked from \#1 to \#7 in their sum. Who is right? Try to analyze this by using extreme scenarios of different rushing and passing yardage for different terms. (It would be more likely that the league would have an even number of teams so if all play each other on one day, but an odd number helps illustrate the answer more simply.)

Answer: That second person could be right. It would not be shocking if that team would be ranked \#4, or \#3 or \#5, but it need not be so. Consider the following extreme cases.

| Team | Rushing yards/Rank | Passing yards/Rank |  | Total yard |  |  |
| :--- | :---: | :--- | :--- | :--- | :--- | :--- |
| A | 1,030 | 1 | 5 | 7 | 1,035 | 7 |
| $B$ | 1,020 | 2 | 25 | 6 | 1,045 | 6 |
| $C$ | 1,010 | 3 | 45 | 5 | 1,055 | 5 |
| $D$ | 1,000 | 4 | 1,000 | 4 | 2,000 | 1 |
| $E$ | 100 | 5 | 1,030 | 3 | 1,130 | 4 |
| $F$ | 80 | 6 | 1,060 | 2 | 1,140 | 3 |
| $G$ | 60 | 7 | 1,090 | 1 | 1,150 | 2 |

Team D is \#1 overall, in first place here. And now consider

| Team | Rushing yards/Rank | Passing yards/Rank |  | Total yards/Rank |  |
| :--- | :---: | :--- | :--- | :--- | :--- |
| $A$ | 3,200 | 1 | 400 | 7 | 3,600 |
| 4 | 4,700 | 3 |  |  |  |
| $B$ | 3,100 | 2 | 600 | 6 | 3,700 |
| $C$ | 3,000 | 3 | 800 | 5 | 3,800 |
| $D$ | 1,000 | 4 | 1,000 | 4 | 2,000 |
| 7 |  |  |  |  |  |
| $E$ | 950 | 5 | 2,000 | 3 | 2,950 |
|  | 600 | 6 | 2,400 | 2 | 3,300 |
| 5 |  |  |  |  |  |
| $G$ | 850 | 7 | 3,000 | 1 | 3,850 |

Here, Team D is \#7 overall, in last place. These examples need not be so extreme to achieve the same results.

## * Ch. 4. Is the drive for cheaper gas worth it? (Section 4.9)

At the local gas station, gasoline costs $\$ 4.00$ a gallon. It costs $\$ 3.60$ at a station 10 miles away (which is a 20-mile roundtrip). Your car gets 40 miles to the gallon and you need to have 15 more gallons in your tank when you again arrive home. (a) Will you save money if you purchase gasoline at this other station? (b) If you include wear and tear on your car of 20 cents a mile travelled, is it still worth it? (c) If using a gallon of gasoline has a carbon-footprint cost of 9 cents a gallon (https://sustainableclimatesolutions.org/learn-more-about-the-carbon-tax-calculator/) and if you want to include this in your analysis, how does it affect the overall effective cost?

Answer: (a) 15 gallons of gasoline cost $\$ 0.40 \times 15=\$ 6.00$ less at this other station. The trip of 20 miles consumes 0.5 gallons, because your car gets 40 miles to the gallon. So, you will need to buy 15.5 gallons to have 15 more gallons when you return home. So, you will need to buy an
extra half gallon $(0.5 \times \$ 3.60=\$ 1.80)$, and so your savings are $\$ 6.00-\$ 1.80=\$ 4.20$. (b) Wear and tear on your car is (optimistically) estimated to be 20 cents a mile (very, very roughly $\$ 20,000$ per 100,000 miles), so this is a cost of $\$ 0.20 \times 20=\$ 4.00$ for the extra 20 miles, and so you are really saving $\$ 4.20-\$ 4.00=\$ 0.20$. Is this trip worth it to save 20 cents? (c) Moreover, including the carbon footprint cost of $\$ 0.09 \times 0.5=\$ 0.045$ for the extra half gallon, you could argue that buying this less expensive gasoline really saves you and the world $\$ 6.00-\$ 1.80$ $\$ 4.00-\$ 0.045=\$ 0.155$ or only about 15 cents.

## ** Ch. 4. How many spice blends can you make? (Section 4.10)

You want to mix five different spices to make distinctive spice blends. For each, you choose the largest ingredient by weight, and then the next largest (that is perhaps $30 \%$ less by weight, though this is set and not important here), then the third largest, the fourth largest and the fifth, the smallest. How many different blends can you make?

Answer: For the first one you have 5 choices, for the second 4 choices and so on, so you have 5 $\times 4 \times 3 \times 2 \times 1=5!=120$ different blends. (This is the number of permutations of 5 items, for which order counts. Of course, you could also vary the relative out of the spices in each.)

## ** Ch. 4. How many more spice blends can you make with more spices to choose from? (Section 4.10)

You want to mix five different spices to make distinctive spice blends, with the five coming from a set of eight different spices. For each, you choose the largest ingredient by weight, and then the next largest (that is perhaps $30 \%$ less by weight, though this is set and not important here), then the third largest, the fourth largest and the fifth, the smallest. How many different blends can you make?

Answer: For the first one you have 5 choices, for the second 4 choices and so on, so you have 8 $\times 7 \times 6 \times 5 \times 4=8!/ 3!=6,720$ different blends. (These are the number of assemblies of 5 items from 8 items, for which the order of the 5 you choose counts but not that of the 3 that you do not choose. The are many more possibilities than in the previous problem. Of course, you could also vary the relative out of the spices in each.)

## ** Ch. 4. How easy is it find the "combination" of a 10-digit keypad lock and open it? (Permutations, arrangements, and combinations in Section 4.10)

On the TV show "The Blacklist" (Season 3, Episode 3; Eli Matchett (No. 72)), Elizabeth Keen, an FBI agent at the time who was "on the run," and Raymond Reddington, a master criminal and FBI confidential informant, want to enter a building quickly that is locked with a 10-digit keypad lock. To open it, the correct 4 different digits have to be punched and in the correct sequence. (a) Keen sees it and says "Rover keypad. Four-digit pin. Could be thousands of combinations". Is this true? And, exactly how many combinations are there? How long would it take to try, say, half of them if each trial takes 2 sec , and so the probability of getting it right would be $50 \%$ ? (b) To which, Reddington then tosses dirt on the pad and (quite remarkably) sees it stick on 4 numbers, by the way $1,3,4$ and 5 , presumably because those have been touched many times
before and are (unbelievably) sticky, and replies "Only if you don't know the four digits. Now there's only 24 combinations". (should be ... there are ...) Is what he says correct or not, and why or why not?

Answer: They want to know the combination of the keypad, but the numbers they seek are not technically combinations. (a) With the 10 digits $0,1,2, \ldots 9$, there are 10 choices for the first one, nine for the second one, eight for the third and seven for the fourth, and so $10 \times 9 \times 8 \times 7=$ 5,040 ways of doing this. (= 10!/6! because there are 10! permutations of the 10 numbers, but you do not care about the order the six numbers that are not chosen and you care about the order of the four chosen numbers.) So, she is correct that there are thousands of ways of doing this, but these are technically arrangements of 4 (ordered) numbers chosen out of 10, and not combinations-which is used more colloquially and not technically. Punching in half of them $(2,520)$ takes 2 seconds each, or $2,520 \times 2$ seconds $=5,040$ seconds $=84$ minutes, a very long time indeed because more immediate entry is needed. Because half of the possibilities would then have been tested, this would be the median time needed. It could take a shorter time than this or up to twice it. (b) There are now only $4!=4 \times 3 \times 2 \times 1$ ways of doing this (using each number one time, with order important) or 24. Raymond Reddington has the correct number, but these are not technically combinations, but permutations of 4 items. All of them could be tested in only $12 \times 2=24 \mathrm{sec}$.

## ** Ch. 4. How many permutations can you make with a toy train set? (Section 4.10)

A train toy set consists of 9 different cars, each with a magnet in the front that can attach only to magnet at the end of another car. How many permutations can you make using the entire set if all 9 cars are attached in a line (with the back of the first one connected to the front of the second one, and so on)? (Would you want to ask a child to make each one?)

Answer: You can choose the first one in 9 ways. With the 8 remaining cars, you can choose the second one in 8 ways, then 7 for the third, and so on. So, there are $9 \times 8 \times 7 \times 6 \times 5 \times 4 \times 3 \times$ $2 \times 1=362,880$ permutations. (Using the language of factorials, this is 9!.)

## ** Ch. 4. How many combinations can you make with a toy train set? (Section 4.10)

 Repeat the previous problem ("How many permutations can you make with a toy train set") to see how many combinations you could make if 7 of these cars were identical and the other two were different than these, but identical to each other.Answer: If the car were all the same there would be $9 \times 8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1=$ 362,880 permutations. But now 7 are the same, so the permutations of these, $7 \times 6 \times 5 \times 4 \times 3$ $\times 2 \times 1=5,040$, are repeats. Of the remaining 2 identical cars, there are $2 \times 1=2$ repeats. So there would be $362,880 /(5,040 \times 2)=36$ combinations. (Using the language of factorials, this is 9!/(2!7!).)
*** Ch. 4. How many possibilities can you make with a toy train set with symmetrical attachments? (Section 4.10)
Repeat the problem "How many permutations can you make with a toy train set," if the end of any car could be attached to either the (clearly discerned) front or back of another car?

Answer: With one direction per car there are there are $9 \times 8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1=$ 362,880 permutations. But now you can reverse any of the 9 cars and double the possibilities, and so now there are $362,880 \times(2 \times 2 \times 2 \times 2 \times 2 \times 2 \times 2 \times 2 \times 2)=362,880 \times 512=$ 185,794,560 ways of doing this. (Using the language of factorials, this is 9! 29.) However, if the front has not been designated as such, you could just reverse the direction all and repeat a possibility and so then you have to divide by 2, to get 92,897,280 ways of doing this.
** Ch. 4. How many arrangements can you make with a toy train set (Section 4.10) Repeat the problem "How many permutations can you make with a toy train set," to find the number of arrangements you could make using only 4 of these 9 cars.

Answer: You can choose the first one in 9 ways. With the 8 remaining cars, you can choose the second one in 8 ways, and then 7 for the third and 6 for the fourth. So, there are $9 \times 8 \times 7 \times 6=$ 3.024 arrangements. (Using the language of factorials, this is 9!/5!.)
** Ch. 4. How many cars could you have in a toy train set and be able to make all the
permutations during "play time?" (Section 4.10, Estimating Section 8.2)
A train toy set consists of certain different cars, each with a magnet in the front that can attach only to magnet at the end of another car, all in a line. What is the maximum number of cars you could use if it you wanted to make all possible permutations (with each one taking, say, a half a minute, independent of the number of cars) within an allotted time (say an hour)?

Answer: There are 60 minutes in an hour, so you can make 120 assemblies. There are 120 permutations of 5 cars $(5 \times 4 \times 3 \times 2 \times 1)$, so you can do this with 5 cars (or fewer).

## ** Ch. 4. Fingers do count, in all combinations (Arrangements and combinations in Section 4.10)

In the Chapter 2 question "Fingers do count," we asked: "If we use three neighboring fingers, in how many ways could we indicate the number 3 ? How do you do this?" The answer was 3 . We use neighboring fingers to accomplish this because it is easiest to do it in this way, but, all told, how many ways could indicate 3 using one hand. (A such combinations.)

Answer: There are 5 ways to choose the first finger, 4 for the second, and 3 for the third, and so $5 \times 4 \times 3=60$ ways to do this, but since the order of these fingers does not matter, and they can be chosen in $3 \times 2 \times 1=6$ ways, there are $60 / 6=10$ ways of doing this (and so 10 combinations). Equivalently, one can choose the 2 fingers that are not chosen, or $5 \times 4=20$,
dividing by 2 since the order of the two does not matter, and so there are again 20/2 $=10$ ways of doing this. These are also equal to $5!/(3!2!)=10$ ways to do this.

## ** Ch. 4. The combinatorics of cupcakes (Section 4.10)

You want to package 16 circular cupcakes packaged in a $4 \times 4$ square array ( 4 across and down, with the top row clearly being the top of the display). There are 4 chocolate, 4 vanilla, 4 strawberry, and 4 lemon cupcakes. You place any cupcake in the top row on the left, any of the remaining ones next to it, and two more to finish the top row, and then continue into the second row and so on. How many ways can these cupcakes can be packaged this way?

Answer: There would be 16! different arrangements for 16 different types of cupcakes. But since each of the 4 chocolate cupcakes can be chosen in $4!$ ways without changing the display, and it would be the same for the other types, there are $16!/(4!\times 4!\times 4!\times 4!)=63,063,000(63$ million $)$ different cupcake arrangements.

## *** Ch. 4. Bookkeeping in strings of letters (Section 4.10)

(a) How many ways can you form strings of 6 letters? (b) How many ways can you do this if each letter is different? (c) How many ways can you form strings of 6 letters, as three consecutive pairs of the same letters, all pairs are different? (d) How many ways can you form strings of 6 letters, as three consecutive pairs of the same letters, with consecutive pairs being different? (Note the word bookkeeper (and variations of it such as bookkeeping) is the only English word with three consecutive pairs of repeated letters. Subbookkeeping is thought by some to have four consecutive repeated pairs, but it really has a hyphen, sub-bookkeeping. https://jeff560.tripod.com/words4.html .)

Answer: (a) There are 26 ways to choose each of the 6 letters: $26 \times 26 \times 26 \times 26 \times 26 \times 26=$ $(26)^{6}=308,915,776$ (b) There are 26 ways to choose the first letter, 25 for the second, 24 for the third and so on: $26 \times 25 \times 24 \times 23 \times 22 \times 21=165,765,600$. (c) There are 26 ways to choose the first letter, 1 way for the second (the same as the first), 25 for the third (any letter but the first), 1 way for the fourth (the same as the third), 24 for the fifth (any letter but the first and third), and 1 way for the sixth (the same as the fifth) $=26 \times 25 \times 24=15,600$---and only 1 of these is part of word-bookkeeper (d) There are still 26 ways to choose the first letter, still 1 way for the second (the same as the first), still 25 for the third (any letter but the first), still I way for the fourth (the same as the third), but now 25 for the fifth (any letter but the third), and 1 way for the sixth (the same as the fifth) $=26 \times 25 \times 25=16,250$.

## * Ch. 4. How notation factor into more or less strawberry juice (math joke) (Notation and factorials, arrangements and combinations in Section 4.10)

It is important to know if notation is meant to be grammatical or mathematical. For example, if the label on a fruit juice bottle reads: How many strawberries did we squeeze in? 22! ... Does this mean they squeezed 22 factorial ( $\sim 1.1240007 \times 10^{21} \sim 1,120,000,000,000,000,000,000$ ) strawberries and put them in the bottle or is the ! merely for emphasis?)?

Answer: It is very unlikely they could fit 22! crushed strawberries in 1 bottle--or even all the bottles on the planet or in all the bottles in the universe.
[https://www.buzzfeed.com/daves4/funny-math-jokes https://apple.news/AYiITz707QnueM_XwHshNGg]

## ** Ch. 4. Will you be dealt a good hand in poker? (Arrangements and combinations in Section 4.10; betting/gambling, probability in Chapter 15)

If you are dealt a hand with five cards in poker, is it more likely that your hand will be a single pair or better, or that you won't even have a hand as good as one pair. There are $1,302,504$ ways to deal a hand without even one pair.

Answer: There are $52!/(47!5!)=2,598,960$ ways of being dealt 5 cards from 52, in no particular order, So, the probability of being dealt a hand from a fresh deck without even one pair is 1,302,504/2,598,960 $\sim 50.1164 \%$ and so it is very, very slightly less likely you will be dealt one pair or better ( $\sim 49.8836 \%$ ).
[http://www.mathplane.com/yahoo_site_admin/assets/docs/mavericks_solitaire.228103706.pdf]
** Ch. 4. Will you be dealt an even better hand in poker? (Arrangements and combinations in Section 4.10; betting/gambling, probability in Chapter 15)
If you are dealt a hand with five cards in poker and get either a straight ( 5 in a row), a flush (all of the same suit), a full house ( 3 of a kind and a pair), or 4 of a kind, you would have no need to get new cards if this were indeed possible in the type for poker you are playing. So, you would stand "pat" and the hand you have would be called a "pat hand." There are 19,716 possible pat hands. What is the probability of being dealt a pat hand from a fresh deck?

Answer: There are $52!/(47!5!)=2,598,960$ ways of being dealt 5 cards from 52, in no particular order, So, the probability of being dealt a pat hand from a fresh deck is 19,716/2,598,960~7.6\%.

## ** Ch. 4. Can you be dealt an unexpectedly good hand in a modified version of poker? (Arrangements and combinations in Section 4.10; betting/gambling, probability in Chapter 15)

You know that is very unlikely that you will be dealt a pat hand from a fresh deck (above). (a) Whether or not the hand you were dealt from a fresh deck was a pat hand, would you expect, the probability of getting a pat hand in the second draw from the remaining deck to be larger, the same or smaller than from the first drawing? (b) Would you expect it to be unlikely to have a pat hand in either of the first two draws or in both of them, using the same deck? (c) Would you expect it to be unlikely to have a pat hand in each of the first five draws from the same deck (with a total of 25 cards dealt)? (d) Would you expect it to be unlikely that you could re-arrange these 25 cards so they would be 5 pat hands?

Answer: (a) Smaller. (b) Yes and yes. (c) Very unlikely. (d) No, it is very likely for this to occur, because you can make 5 pat hands by re-arranging the order of 25 randomly drawn cards from
one deck $\sim 98.1 \%$ of the time. (Order makes a difference!) This is commonly called Maverick's Solitaire. This is well known from being a "sucker" bet (but for a good cause it turns out!),in the TV show Maverick.

The character Maverick, played by James Garner, in the eponymous TV show (Maverick, Rope of Cards, Season 1, Episode 17) convinces someone that since it is unlikely to be dealt a pat hand from a fresh hand (which is true, see above problems), the probability was miniscule that 5 pat hands would be drawn in succession, from 25 cards) from the initially fresh deck (which is true), and so it would be a safe bet that you could not arrange the 25 drawn cards to be 5 pat hands. He makes his case, for a good cause, but he is able to make 5 pat hands and so win the bet (seemingly against the odds). (Does this make sense to you?)
[Martin J. Chlond, (2012) Five Pat Hands. INFORMS Transactions on Education 12(3):164-165. https://doi.org/10.1287/ ited. 1120.0089 https://pubsonline.informs.org/doi/pdf/10.1287/ited.1120.0089
http://www.solitairelaboratory.com/maverick.html
http://www.mathplane.com/yahoo_site_admin/assets/docs/mavericks_solitaire.228103706.pdf
Scarne J (1961) Scarne's Complete Guide to Gambling (Simon and Schuster, New York).]

## * Ch. 4. How do you pigeonhole the hairs on your head? (Section 4.10, The Pigeonhole Principle)

Do you know if at least two people in New York City (say with population 8,000,000) have the same numbers of hairs on their head, knowing that people have no more than 8,000 hairs on their heads (excluding mustaches and beards)?

Answer: Using the Pigeonhole Principle in Section 4.10, pages 42-3, any place with $8,000+1=$ 8,001 people will have at least two people with the same numbers of hairs on their head. (Of course, this is possible for smaller places as well.)

Mark Kac and Stanislaw M. Ulam, Mathematics and Logic, page 11.
** Ch. 4. How do you pigeonhole your initials? (Section 4.10, The Pigeonhole Principle) How large does a city have to be for at least two people to have the same initials? The initials can have two or three letters of the English alphabet. Ignore "junior," the 3rd and so on.

Answer: Because there are 26 letters, there are $26 \times 26=676$ possible initials with two letters and $26 \times 26 \times 26=17,576$ initials with three letters, and so 18,252 possible initials. Using the Pigeonhole Principle (Section 4.1, pages 42-3), in a town with 18,252+1=18,253 or more people at least two people will have the same initials. (Of course, this is possible for smaller towns as well.)

Mark Kac and Stanislaw M. Ulam, Mathematics and Logic, page 11.

## * Ch. 4. Math thinking about pigeonholing and choosing apples (Section 4.10, The Pigeonhole Principle)

If there are three kinds of apples mixed in a box, how many must you choose to be sure that at least (a) two apples are of one kind? or (b) three apples of one kind?

Answer: (a) If you choose three apples it is possible that they all can be different, so you must choose four to make sure that at two least of them are of one kind. (b) If you choose six apples it is possible that you have two each of one kind, so you must choose seven to make sure that at least three of them are of one kind.

From
The Moscow Puzzles, Boris A. Kordemsky, \#259, pg. 109

## Chapter 5. Words and Numbers: Being Careful

## * Ch. 5. How can you be for this? (Numerology, Chapter 5)

In TV show The Blacklist (Season 5, Episode 12; The Cook (No. 56)) master criminal and FBI confidential informant Raymond Reddington presents a Japanese woman a well-intended and costly gift of four special spiders, spiders that she would very much appreciate, but instead she becomes very angry with him and he is baffled by this. Why?

Answer: Four is traditionally an unlucky number in Japan, meaning that it gives bad luck, because it is sometimes pronounced as the Japanese word for death, and so she became insulted and scared. In fact, sometimes this number is not used for locations in buildings. However, the quite worldly Reddington had no clue about this cultural difference in numerology.

Wikipedia, Japanese superstitions.

## * Ch. 5. Fewer or less than? (Words and number, Section 5.1,)

In TV show The Blacklist (Season 5, Episode 12; The Cook (No. 56)) FBI special agent Donald Ressler notes "Less than $15 \%$ of all arson cases are ever solved." Although it is clear what he means, his statement mathematically (and grammatically) inexact. Why?

Answer: Saying something is less than $15 \%$ is fine, but he is referring to the $15 \%$ of the number of arson cases, and so he is referring to a whole number-and he should have said "fewer than $15 \%$."

## * Ch. 5. What does it mean to measure a cylinder from the inside or the outside? Meaning of words-inside vs. outside diameter (Section 5.1)

Because the wall has a given thickness, the diameter of a hollow metal cylinder is different if it is measured from its outside or its inside. In objects called tubes, the outside diameter is given in specs and in pipes it is the inside diameter. What is the inside diameter of a $1 / 4$ inch metal tube with wall thickness of 0.02 inches?

Answer: Double the wall thickness must be subtracted to obtain the inside diameter, so it is 0.25 $-2 \times 0.02=0.25-0.04=0.21$ inches. This is an example that shows that words have specific numerical meanings.

## * Ch. 5. What does "one day" mean, exactly? (Section 5.1)

Guidance on air travel to the U.S. (on May 25, 2022) includes "Before boarding a flight to the United States, you are required to show a negative COVID-19 test result taken no more than 1 day before travel." If your flight leaves 3 PM on a Friday, when is the earliest you should take this test?

Answer: The wording is ambiguous (to me). One day is 24 hours, so to be careful based on this guidance alone, you could easily think you should be taking the test a full day earlier and no earlier than 3 PM on Thursday, the day before you leave. However, elucidation is provided later in the notice that you can test as earlier as any time on the prior calendar day before you leave, or Thursday here. They could have (and should have) said "no more than 1 day before the day of travel." You need to be careful with words and numbers!
(https://www.cdc.gov/coronavirus/2019-ncov/travelers/international-travel-during-covid19.html, updated May 3, 2022)

## * Ch. 5. When is $\mathbf{1 0 \%}$ really $\mathbf{1 0 0 \%}$ ? (Section 5.1)

What does the common statement that we use only $10 \%$ of our brains suggest and is it correct?
Answer: It suggests that we should be able to expand or mental capabilities by using the rest of our brains, and is just a myth based on unsubstantiated comments made many years ago. This is an example of the possibly deceptive or misleading power of cited numbers.
https://www.wsj.com/articles/is-your-brain-goofing-off-11605868200
Is Your Brain Goofing Off?
By Jo Craven McGinty
Nov. 20, 2020 5:30 am ET

# * Ch. 5. Giving it a second look (math joke) (Section 5.1, Chapter 10) 

Math joke: How many seconds are there in a year?
Answer: Presuming that this is meant to be a joke: 12; January second, February second, March second, ... . (Of course the expected answer presumes the use of seconds as a time unit and not the ordinal number $2^{\text {nd }}$. For the time unit, the answer would be ( 60 seconds/minute) $\times(60$ minutes/hour $) \times(24$ hours/day $) \times(365$ days/year $)=60 \times 60 \times 24 \times 365$ seconds per year or, $31,536,000$ of them; or since an aver a year has closer to $365^{1 / 4}$ days, and so 31,557,600 of them.) This shows the need to be careful with numbers and words. [https://web.sonoma.edu/math/faculty/falbo/jokes.html 10/7/20
Jokes for Mathematics Teachers
8. BIG NUMBERS]

## * Ch. 5. They are just peanuts (The Correct Words, Section 5.1)

A package of mixed nuts has a label noting that is has less than $50 \%$ peanuts (and there is an asterisk that states that the nuts are measured by weight). (a) What does this mean? (b) Does it also mean that if you counted the number of nuts (entire nuts, so peanuts split into two still count as one nut), fewer than half of them would be peanuts? (c) Was is correct for the label to say "less than" instead of "fewer than?"

Answer: (a) It means that peanuts constitute less than $50 \%$ of the total weight of all nuts. (b) No. If the other nuts tend to be heavier than peanuts, on a per nut basis, over half of the nuts would be peanuts. Their labeling is meant to lessen customer complaints. (c) Yes, because the amount in reference is weight and not a number. Fewer is used to refer to non-negative counting numbers, and "less than" to any real fraction or decimal.

## * Ch. 5. When is a pint not a pint? (Section 5.1)

Ben \& Jerry's sells (at least it did in 2021) a box of 3 ice cream bars called PINT slices in a box that says that its totality of its contents has 9 fluid ounces. A pint has 16 fluid ounces. How can you reconcile these numbers?

Answer: You can't reconcile these numbers. Since each of the bars has 3 fluid ounces, neither each one nor three of them constitute a pint.

## * Ch. 5. Are "by" and "to" the same in math? (Section 5.1)

On page 347 in Presidents of War author Michael Beschloss noted in reference to President Woodrow Wilson's interest in the establishment of the League of Nations 'He predicted that the treaty would increase "the probability of peace" by "about ninety-nine percent" '. Explain how the meaning of this sentence changes if "by" were changed to "to." Which do you think would be the correct word here?

Answer: Ninety-nine percent is very nearly 1. As stated, the probability of peace would essentially double. With "to" instead of "by" the probability of peace would be increase, likely many-fold, from a small probability to essentially 1. It is likely that "to" was intended (by Wilson and/or the author). This is an example of why it is important to choose the correct words with numbers.

## ** Ch. 5. Can it be more humid when the humidity is said to be the same? (Section 5.1, What do words mean?)

One day the temperature is $60^{\circ} \mathrm{F}$ and the reported humidity is $60 \%$, while the next day it is $90^{\circ} \mathrm{F}$ and the humidity is $60 \%$. Which day is more humid or are they equally humid?

Answer: This question seems like nonsense, but it is not. Words have specific meaning and one must be careful understanding the numbers that characterize them. The pressure of water vapor in the air (which is proportional to the amount of water vapor in each volume of air) is the absolute humidity. There is a maximum water vapor pressure or maximum absolute humidity that air can have. This maximum amount increases with increasing temperature. (And so, water evaporates faster when it is hotter.) The reported humidity is actually the "relative humidity," which is the fraction: absolute humidity/maximum absolute humidity at that temperature (expressed in per cent). So, the relative humidity is the same on both days, but the absolute humidity is much higher on the second day.

## * Ch. 5. A much, much higher or a much, much, much higher salary? (Section 5.1)

 When it comes to numbers, words matter. A person who makes $\$ 70,000$ a year in a job in the public sector is sure he was told by a private sector recruiter that he would earn four times more in their company. But, when he joined he learned his new salary was "only" \$280,000 and not the $\$ 350,000$ he expected. When confronted with this, the recruiter said, yes, we offered you four times as much. Who is right?Answer: It depends whose memory is accurate. Four times as much as the $\$ 70,000$ salary (or four times the salary) is $\$ 280,000$, while four times more is the original $\$ 70,000$ salary plus $\$ 280,000$, or $\$ 350,000$.

## * Ch. 5. What is it exactly? (Section 5.1)

In Season 2, Episode 39 of the TV sitcom "Leave it to Beaver," the teacher "Miss" Landers assigns the (central character) Beaver's third grade class an assignment to write a composition about an interesting character. In response to a question concerning its required length, she adds it needs to be 100 words long. Does this mean that it needs to be at least 100 words, approximately 100 words lone, or exactly 100 words long?

Answer: Miss Landers should have been clearer. She likely meant that it should be at least 100 words. This is an example that words with numbers have a specific meaning.

## * Ch. 5. What does 6 months really mean (Section 5.1)

In the title and body different accounts, the press presented the results of the same scientific study of how long a vaccine maintained its high level of effectiveness, of say $90 \%$, as either (a) as long as 6 months, (b) 6 months, and (c) 6 months in studies so far. Why do these have different meanings and which do you suspect was actually accurate?

Answer: Their meanings are quite different and illustrate the importance of using numbers and words correctly. (a) means anywhere from 0 to 6 months and no longer, (b) means exactly or essentially 6 months, and (c) means 6 months possibly longer and was this version was actually accurate in one study of how long one of the first COVID-19 vaccine treatments remained effective. One early online account had one version in the title and another in the body, and this account was later corrected.

## * Ch. 5. Be clear about numbers (Section 5.1)

A very sad article in February 2022 reported: "The decomposed body of a 70-year-old Italian woman was found sitting at her table - more than two years after she died, police said. Neighbors said they last saw the elderly woman in September 2019 - and assumed she had moved away at the start of the COVID-19 pandemic. Based on the level of decay, investigators estimated she had been dead for more than two years." Solely from a numerical perspective, what is unclear about this report?
https://nypost.com/2022/02/09/italian-woman-found-sitting-at-table-two-years-after-death/

Answer: Did she actually die at the age of 70 and was found two years later or would she have been 70 when she was found, if she had lived until then? The former version would have made it accurate, but I suspect the latter version was implied.
** Ch. 5. Does it make sense that the average, median and mode gasoline prices in U.S. gas stations greatly differ? (Averaging numbers, Section 5.2)
On Oct. 3, 2022 GasBuddy reported that the average gasoline price in the U.S. was $\$ 3.78$, while the median was $\$ 3.49$ and the mode $\$ 3.29$. Does it make sense that the average was so much higher than the median (middle price) and the mode (most common price)?

Answer: Sure. It depends on the actual distribution. In this case the prices in the very populous California $\sim \$ 6.41$ were much higher than those in the other states and this gave the distribution of the number of gas stations vs. gasoline price a prominent tail at higher prices, making it very asymmetric about the average. (See the bar graph in the cited WSJ article, with data from GasBuddy.) So, California increased the average to a value much larger than that of the other 49 states. However, it did not change the most common price (the mode) found in the other states at all. It makes sense that the mode could be lower than the median and average. The high gas prices in California moved the middle value (median) up some from the distribution without it, but not as much as it increased the average price.

As reported by
https://www.wsj.com/articles/why-the-most-common-gas-price-is-far-from-average11666949402
https://finance.yahoo.com/news/gas-prices-california-2022-highs-national-average192250990.html

## ** Ch. 5. Giving weight to your grade average with weighted averages (Section 5.2)

If during a term in school you took two courses and received an A in one and a C in the other, what would your grade point average (GPA) be for that term? Use 4.0 as the numerical value for an A, 3.0 for a B, and 2.0 for a C, and so on. First find the simple average, and then the average weighted for 4 credits for the former course and 3 credits for the latter. What would the averages be if the grades were reversed? Are these differences significant?

Answer: The simple average is $(4.0+2.0) / 2=3.0$ or a B. The weighted average is $(4 \times 4.0+3$ $\times 2.0) / 7=\sim 3.14$. This is above a $B$ (and approximately halfway between a $B$ and $B+($ which is worth 3.333.... ).) If the grades were reversed, the simple average would be the same and the weighted GPA would be $(4 \times 2.0+3 \times 4.0) / 7=\sim 2.86$ (and approximately halfway between a $B$ and $B$ - (which is worth $2.66666 \ldots$...). These differences are quite significant. Grades on transcripts are always weighted.

## ** Ch. 5. Rolling with rolling averages (Section 5.2)

When new data are received regularly, say daily, averages over several recent periods, such as over the most recent 7 days, are sometimes presented because such "rolling averages" average over fluctuations, so trends are more apparent. You are told the rolling average of the number of newly infected people per day for a disease over the previous 7 days was 21.0. (a) First assume that the number of new infections on each of these 7 days was the same. What is this number? (b) Then, if the next day the number of new infections is 0 , what is the new 7 -day rolling average? (c) Now, let's say the 7-day rolling average is again 21.0, and you know nothing else about the data for each day. Give other possible sets of data with the same rolling average. (d) Again, say the 7-day rolling average is again 21.0, and you know nothing else about the data for each day. Then on the next day, this average has decreased to 0 . Is this result suspicious? Does this mean the data were suspicious?

Answer: (a) The same as the average, 21. (b) The sum of infections would be $6 \times 21.0=126$, which averaged over 7 days is $126 / 7=18.0$. (c) Among the many, many possibilities are: 24, 23, 22, 21, 20, 19, 18, and 19, 26, 21, 14, 23, 28, 16. (d) Such as fast drop-off could seem unlikely and suspicious, but it would be correct if all of the infections in the previous 7-day period had occurred only on the first day-which were therefore included in the first average, but not the next one.

## * Ch. 5. Top halves only (Section 5.2)

You are requested to fill out a recommendation form for someone, which has several evaluation categories. For each you must say whether the candidate is in the top $1 \%, 5 \%, 10 \%, 20 \%$ or
$50 \%$, or that you are unable to judge the candidate in that category. Your math thinking cap sounds an alarm (or it should be sounding an alarm). Why?

Answer: You are not being allowed to say whether or not the person is in the lower $50 \%$, and this is mathematically untenable and illogical, given the other options. A lower-50\% assessment is very different from the "unable to judge" assessment, so that latter option is not appropriate, if your judgement were indeed that the candidate is in the lower 50\%. If that last category were instead "other," the choices given would be logically reasonable, but less meaningful for the evaluation. (I have seen recommendation letter requests with this flaw.)

## * Ch. 5. What does it mean to say you are from millions of years from now? Words and math meanings, big numbers (Section 5.4)

In the Woody Allen movie "Midnight in Paris" the main character Gil Pender travels in time back from 2010 to the 1920s and tells Salvador Dali "I'm from a different time. Another era. The future. Okay, I come from the $2,000^{\text {th }}$ millennium to here." What is mathematically wrong with his statement?

Answer: Gil means to say that he is from the 2000s or the 21st century. A millennium is 1,000 years long, and he is certainly not from 2,000th millennium, which will be roughly 2 million years in the future from real-life-time or Dali-time Gil.
[https://www.imdb.com/title/tt1605783/goofs/?tab=gf\&ref=tt_trv_gf]

## * Ch. 5. Going backward with centuries? (Section 5.4)

A recent book noted: "Psychology worked with the disease model for over 60 years, from about the late 1800 s into the middle part of the 19th century." What is mathematically incorrect with this sentence?

Answer: The authors likely mean "... into the middle part of the 20th century." because they want to indicate a period from the late 1800s to very roughly 1950. This puts the ending date of this range of times in the middle of the $20^{\text {th }}$ century, which extended from January 1, 1901 to December 31, 2000. Furthermore, as written, the final time occurred before the initial one.

## * Ch. 5. The number is Greek to me (Section 5.4)

Why is the first translation of the Hebrew Bible into Greek, supposedly by 70 Jewish scholars, often called the Septuagint?

Answer: In Latin "septuāgintā" means 'seventy'. Knowing that "sept" denotes 7 gets you to most of this answer.

[^0]Answer: In some ways yes. August is currently the eighth month of the year, while October (oct for 8) used to be, when March was the first month of the calendar year.

## * Ch. 5. Is it a million? (Section 5.4.1)

One article or report you read uses M to mean million, while another uses MM. Does this make any sense?

Answer: Yes. Both are meant to signify a million but they originate in different ways and are used in different communities. Usually, $M$ is usually from mega, as noted in the book. In finance and accounting it is not uncommon to denote million by MM, using the Roman numeral for a thousand $M$ and so MM would be a thousand thousands or a million, or mm. Check the context and the standard convention and common usage in an area for all such notation to make sure you are not off by a thousand.

## * Ch. 5. Does it matter if it is a million or a thousand, anyway? (Yes, it does.) (Section 5.4.1)

In presenting a certain number of dollars a financial report notes 100 mm . What does this mean?
Answer: Though M is usually used to be million (from mega, as noted in the book) in much of life (including in my professional life), in finance and accounting it is not uncommon to denote million by MM or mm (which also happens to also to be a symbol for millimeter), with M or $m$ (and not K) meaning a thousand from the Roman numeral for a thousand $M$ - and so MM would mean a thousand thousands or a million. Check the context and the standard convention and common usage in an area for all such notation to make sure you are not off by a thousand.

## Chapter 6. Writing Really Big and Really Small Numbers, and Those In-between

## ** Ch. 6. Faster chips (Section 6.1, Scientific Notation)

November 15, 1971 marked the 50th anniversary of the launch of the Intel 4004 microprocessor. It had 2,300 transistors could perform about 92,000 arithmetic operations a second. The 2021 Apple M1 Max processor has 57 billion transistors that do 10.4 trillion arithmetic operations a second. (a) What has been the increase in the number of transistors per chip and the speed, in scientific notation? (b) How has the speed of the chip per transistor on the chip changed? (The exponential increase of chip speed with time has been characterized by Moore's Law, Sections 13.1.1 and 13.3.)

Answer: (a) The increase in the number of transistors has been by a factor of $57,000,000,000 / 2,300=2.5 \times 10^{7}$ (25 million) and in the computing speed (arithmetic
operations performed per second) this factor has been 10,400,000,000,000/92,000 $=1.1 \times 10^{8}$ (110 million). (b) It has increased by (10,400,000,000,000/57,000,000,000)/(92,000/2,300) = 4.6, which also equals 110 million/ 25 million. (This is arithmetically correct, but the actual speed per individual transistor has increased much faster than this. As noted in the cited article, most of increase in the world's wealth since 1971 can be attributed to introduction of this Intel chip and its subsequent improvement.)

The Chip That Changed the World
By Andy Kessler, Nov. 14, 2021 12:38 pm ET, Wall Street Journal
https://www.wsj.com/articles/the-chip-that-changed-the-world-microprocessor-computing-transistor-breakthrough-intel-11636903999

## ** Ch. 6. Using math to raise or lower your voice (Section 6.2.1)

If you play a musical piece faster, its duration is shorter, such as playing it twice as fast so it plays for $50 \%$ of the originally recorded duration. If you play it slower it will end in a correspondingly longer time, such as playing it half as fast so it plays for twice the time. Moreover, since musical notes corresponding to given number of oscillations per second (Section 6.2.1), when you change the speed of the recording this there is a corresponding change in the time between the oscillations and so the oscillation frequency. So, if you play a musical piece twice as fast, the time between oscillations is half as long and so the oscillation frequency is twice as high, and this corresponds to a higher-pitch note. Similarly, if you play it slower, the notes become lower pitch. (a) How much faster do you need to play music for each note to become 1 octave higher? (b) How much faster do you need to play it so each note becomes one half note higher in pitch?

Answer: (a) One octave higher, the oscillation rate of a note becomes twice as fast, so if you play the original piece twice as fast, each note remains the same note, but one octave higher. (b) The step half notes correspond to increasing the frequency by $2^{1 / 12} \sim 1.0595$, so you need to play it $5.95 \%$ faster for this to occur and, for example, change an $F$ to an $F$ sharp.

Small changes in speed that may not be obvious otherwise, can raise or lower notes in a clearly recognized way. Playing a song with a slight increase in speed relative to the recording speed will slightly increase the pitch and can make the singer seem to be a bit younger. A slight increase in speed relative to the recording speed was used in the released version of the 1963 song Wonderful Summer by Robin Ward (https://en.wikipedia.org/wiki/Robin Ward (singer)) to make it sound better, but it also suggested the singer was younger than she was. A larger increase was purposely used to make the high-pitched "Chipmunk" voices for Alvin and the Chipmunks. (https://en.wikipedia.org/wiki/Alvin_and_the_Chipmunks) In this case, people spoke and sang in normal voices but slower than normal, with the recording speed being at half the usual rate, and words in the sped up (normal rate) version were then at a normal rate, but higher pitched.

## *** Ch. 6. Plotting Zipf's law and the the the the the ... (Log-log plots, Section 6.2.2; Statistics, Chapter 16; Ranking, Chapter 18)

As will be explained in the problems to be presented for Chapter 16 on Statistics, Zipf's law says that how often a word is used in a book, divided by the frequency of the most-used word, approximately equals 1 divided by the rank of that word in the ranked list of decreasing word usage. So, the $\mathrm{n}^{\text {th }}$ ranked word (the word used the $\mathrm{n}^{\text {th }}$ times most frequently) occurs $\sim 1 / \mathrm{n}$ times as often as the most frequent word (the $1^{\text {st }}$ ranked word). If you plot this relative usage frequency, $1 / \mathrm{n}$, vs. the rank, n , on a linear scale, it is 1 for $\mathrm{n}=1$ and it slowly decreases, approaching but never reaching 0 in a nonlinear way for larger $n$. (Here, only integral $n$ make sense because it represents a ranking.) How would this function look if you plotted the $\log$ of the frequency [and so $\log (1 / n)]$ vs. the $\log$ of the rank (and so $\log \mathrm{n}$ )?

Answer: On the log-log scale Zipf's law would look like a straight line, with a negative slope, which is in fact -1 . (This is because $\log (1 / n)=-\log (n)$. You can see this in two (equivalent) ways. $\log (1 / n)=\log (1)-\log (n)=0-\log (n)=-\log (n)$. Also, $1 / n=n^{-1}$ and since $\log \left(n^{x}\right)=x$ $\log (n), \log (1 / n)=\log \left(n^{-1}\right)=-1 \times \log (n)=-\log (n)$. Footnotes for Section 6.2) All of this is true for any log scale, such as base 10 or e.)

## ** Ch. 6. What does plotting with a log bar indicate? (Section 6.2.2)

You see a bar graph in which the bar indicating that the number of antibodies that appear in a person 16 days after vaccination for a disease is twice as long as the bar representing 8 days after vaccination. Does this mean there are twice as many antibodies after 16 days as after 8 days?

Answer: No, it does not necessarily mean that at all! Look at the scales! If the scales are linear, this increase could be very different. If the 8-day results ranged from 0 to 50 and the 167-day one from 0 to 100 this twice as long bar would indicate twice as many antibodies, but if they respectively ranged from 50 to 100 and 100 to 150 (with 50 being the minimum shown in the scale) the increase in antibodies would be only $50 \%$. Such differences can be even more so for a log scale. For example, if the bar spanned from 1 to 10 at 8 days (with 1 being the minimum of the scale shown) and from 1 to 100 at 16 days, the length of the bar doubled but the number of antibodies increased by a factor of 10 . (Note that the $\log 1=0, \log 10=1$, and $\log 100=2$, so the length of the bar would increase by a factor of 2.)

## Chapter 7. Touching All Bases: The Worlds of Logs and Bases

## ** Ch. 7. Saving your life by way of base 2

You are told that you are among a specified number of people who must stay on a circle, and who are numbered from 1 to this number in a given rotation sense, say clockwise. Person 1 is forced to kill $\# 2$, the next surviving person on the circle, $\# 3$, kills the next person in the circle, \#4, and this continues until only 1 person remains. You hope to be that person and are told that you can choose what number you are. A situation similar to this supposedly happened to historian Josephus two millennia ago as he and his troops engaged in mass suicide to avoid capture, and he somehow was at the location as the last surviving person. There is a formal solution to this Josephus Problem (or Permutation) that can be summarized as: Express the total number of people on the circle in base 2. Take the " 1 " in the largest digit place (the leftmost digit
of $1 \ldots$ ) and physically move it to the other side of the number, putting it in the "ones" place, shifting all other number to the left (as in ...1). This is the number you should choose. If there are 41 people in the circle, what place number should you choose?

Answer: $41=1 \times 32+0 \times 16+1 \times 8+0 \times 4+0 \times 2+1 \times 1=1 \times 2^{5}+0 \times 2^{4}+1 \times 2^{3}+0 \times 2^{2}+$ $0 \times 2^{1}+1 \times 2^{0}$, because $2^{0}=1$. So, 41 in base 10 is 101001 in base 2 . Moving the first 1 on the other side gives $010011=10011$, which is $1 \times 2^{4}+0 \times 2^{3}+0 \times 2^{2}+1 \times 2^{1}+1 \times 2^{0}=1 \times 16+0$ $x 8+0 \times 4+1 \times 2+1 \times 1=19$ in base 10 . (For example see https://www.youtube.com/watch? $v=u C s D 3 Z G z M g E a n d ~ h t t p s: / / w w w . g e o g e b r a . o r g / m / E x v v r B b R$. (retrieved 6-22-21))

## ** Ch. 7. Basing holidays on math (math joke) (Base Systems)

Why do some mathematicians confuse Halloween (31 Oct) and Christmas (25 Dec)? (Hint: Think of the month as the base.)

Answer: 31 in base 8 (which is $3 \times 8+1=25$ in base 10) and 25 in base 10 are the same. https://en.wikipedia.org/wiki/Mathematical_joke and refences cited therein

## ** Ch. 7. Subtraction using what was called the "New Math" - In Base 8

As noted above, in "Subtraction using what was called the "New Math" - Basics," Tom Lehrer wrote a song "New Math" that poked fun at the then new methods of teaching math to school children. One of its elements was to teach children to do arithmetic in different base systems, because its proponents expected this would give them increased understanding, but it, in fact, confused many because it lacked everyday context. In Lehrer's rendition of this song (see above), after doing the subtraction example $342-173$, assuming both numbers were in base 10 , he did it accurately, but comically, again now assuming they were in base 8 . What is the answer?

Answer: Expressed in base 8, the difference is 163. We could solve this problem by staying in base 8 and subtracting using the rules for base 10 subtraction modified for base 8, or by converting the numbers to base 10, then subtract them as usual, and then convert the answer to base 8. We do the latter. The numbers 342 and 173 in base 8 are respectively $3 \times 64+4 \times 8+2$ $=226$ and $1 \times 64+7 \times 8+3=123$ in base 10, given that $8^{2}=64$ is in the "hundreds" place and 8 is in the "tens" place. So, the difference is $226-123=103$ in base 10. By subtracting as many units of 64 as possible (keeping the number positive) and then of 8 from the remainder, in base 10 this is seen to be $1 \times 64+4 \times 8+7$. So, in base 8 it is 147 . See how Tom Lehrer does this, remaining in base 8 all of the time.

## ** Ch. 7. When do decimal numbers mean something else base-ically? (Decimals and base systems)

When does 15.2 not mean 15 and 2 tenths and 238.1 not 238 plus 1 tenth?
Answer: The height of horses (at the highest point of the withers, which is highest part of a horse's back, lying at the base of the neck above the shoulders) is often given in units of hands,
with a hand being 4 inches, so a 5-foot tall horse would be 15 hands tall. A horse that is 5 feet 2 inches would be 15 and a half hands tall, but this is written as 15.2 hands, where the number after the "decimal point" stands for quarters (and not tenths) of a hand, so 2 quarters (or a half a hand).[http://www.onlineconversion.com/horse height.htm 5-2-20]

A baseball who pitched $2381 / 3$ innings in a year, completed the equivalent of 238 innings of 3 outs each plus $1 / 3$ of an inning or 1 out. In increasingly used baseball numerical parlance, this is listed as 238.1 innings, where the 1 stands for one third of an inning and not one tenth.

So, the numbers to the left of the decimal point are in base 10 and those to the right of it are respectively in base 4 and 3. This point is not the convention decimal point (in base 10) or the radix point in another base (for which the base is the same for both the integer part to the left of it and to the fraction part to the right of it). (See the book.)

## ** Ch. 7. When do decimal numbers mean something else base-ically, again? (Decimals and base systems)

When can $120.2+45.2=166.1$ ?
Answer: In ordinary notation in base 10, $120.2+45.2=165.4$. Since largest digit is 6 , it could be in base of 7 or higher, but then still $120.2+45.2=165.4$. See the unusual mixed pre- and post- decimal point system in the previous problem. Then this would make sense with base 10 to the left of the point, and in base 3 to the right of it, since (in base 10) 2 thirds +2 thirds is 4 thirds, which is 1 (which is carried over) and one third. (Also see the discussion of radix in the previous problem.)

## * Ch. 7. Elevators and information (Section 7.1)

In a hallway, the wall by an elevator has one light that turns on when it arrives at that floor and it is going up and one light that indicates it is going down. (a) For an elevator passing by a floor, are these two lights like an AND gate or an OR gate? (b) How many bits of information do the lights indicate? (c) If, due to a malfunction, arriving elevators are indicated by the up and down lights both being on, how much information is there? (d) If arriving elevators are instead indicated by the up and down lights both being off, how much information is there?

Answer: (a) An Or Gate. (b) 1 bit. (c) None. (d), Still, none.

## * Ch. 7. More on Greek letter days-going beyond pi day (math joke) (Followup on pi day (Chapter 3), e Section 7.2)

Can there be an e day, like the pi day, March $14^{\text {th }}$ ?
Answer: Using 2.718..., and then 2.7 , it would be the day, February $7^{\text {th }}$, though some would say it would not be possible to do since using 2.718... and then two digits after the decimal point as for pi and so 2.71, it would be the non-existent day, February $71^{s t}$ (and so it would be a joke). [https://www.buzzfeed.com/daves4/funny-math-jokes
https://apple.news/AYiITz707QnueM_XwHshNGg]

# ** Ch. 7. Jenny has a constant as well as a number (math joke) (Transcendental Section 7.2, Chapter 3) <br> We called 8,675,309 Jenny's number, after the song 867-5309/Jenny sung by Tommy Tutone, and examined it in several chapters. Explain why $\left(7^{\mathrm{e}-1 / \mathrm{e}}-9\right) \pi^{2}$ has been defined as Jenny's constant, $J$. 

Answer: It just happens to equal to 867.5309 ... You need to go to 9 places after the decimal point in $e$ and $\pi$ to get this.
[https://en.wikipedia.org/wiki/Mathematical_joke and reference cited therein, and see https://oeis.org/A182369]

## Chapter 8. Numbers Need to be Exact, But It Ain't Necessarily So

* Ch. 8. How high is a mountain, exactly? (Or, what is the pinnacle of success?) (What do numbers and measurements mean?, Section 8.1)
The listed height of a mountain is actually the height of its apex relative to sea level. However, more than one height has been given for the height of the tallest mountain, Mt. Everest, usually a bit over $29,000 \mathrm{ft}$. What are the possible reasons for this?

Answer: Measurement tools improve with time, so the measurement can become more accurate and lead to a different listed height. Another reason is the non-unique definition of what is the top of a snow-capped mountain. Is it the top of the rock, which led China to measure it to be 29,017 ft high, or the top of snow cap, which has led Nepal to measure it to be 29,028 ft high? The actual height also changes due to motion of tectonic plates, earthquakes, ... In December 2020, China and Nepal agreed to call the top to be the top of the snow cap and, with new measurement technology, agreed that Mt. Everest to be 29,032 ft high (rounding off 29,031.69 $f t)$. (And, does the thickness of the snow level change with time?) Also, note that the mountain height is relative to sea level, and not the base of the mountain, and the peak of Mt. Everest is "only" ~~12,000-15,000 ft above the base, depending on location at the base.
[https://www.wsj.com/articles/mount-everest-just-got-taller-11607417239
Mount Everest's Height Just Grew to 29,032 Feet
By Eric Bellman and Krishna Pokharel
Updated Dec. 8, 2020 7:46 am ET
https://www.bbc.com/news/world-asia-55218443
Mt Everest grows by nearly a metre to new height
By Navin Singh Khadka Published 8 December 2020

# https://en.wikipedia.org/wiki/Mount_Everest] 

** Ch. 8. Rescued by significant figures (Section 8.1)

One of the TV ads for the Apple Watch (called Apple Watch Series 7 TV Spot, '911: Bob' https://www.ispot.tv/ad/qeWo/apple-watch-series-7-bobs-911-call) noted an automated call to 9-1-1 was made by the watch after the owner took a fall, was unconscious and missed the phone alert. The phone then gave the dispatcher "his exact coordinate location," with "the emergency location is latitude 47.7 longitude -117.5 with an estimated search radius of forty one meters." Say the presumption is that the latitude is north. The device location accuracy is thought to be 41 m , which is quite good. Still, this message disconcerted me for a math reason. What was it?

Answer: The lack of enough significant figures for it to be useful with a search radius of 41 m . One degree latitude corresponds to $\sim 25,000$ (the earth circumference)/360 degrees or $\sim 69.4$ miles. This is also true for longitude measured at the equator. It is less than this away from the equator, but let's ignore that here. The latitude and longitude are given to one figure after the decimal point, which is 0.1 degrees or $\sim 6.94$ miles, so with the data as given the search radius should be $\sim 7$ miles. This $\sim 6.94$ miles $\times(1609$ meters/mile $) \sim 11,170$ meters, which is much greater than the claimed 41 meters, and the search would take a long time, much longer than one with a 41 meter radius. The problem is that not enough significant figures were given in the verbal and written directions provided in the ad. Presuming that the accuracy is 41 meters, two more significant figures (possibly latitude 47.700 longitude -117.500 or latitude 47.692 longitude -117.519 or ...) would give the location to $\sim 112$ meters and three more (possibly latitude 47.7000 longitude -117.5000 or latitude 47.6924 longitude -117.5186 or ...) to $\sim 11.2$ meters, which is what would be needed for the stated accuracy. It is hoped that such better information is provided in reality.

## * Ch. 8. Precisely, how do dinosaur skeletons age? (math joke) (Section 8.1 precision and Section 8.1.1 rounding off)

"A museum visitor was admiring a Tyrannosaurus fossil, and asked a nearby museum employee how old it was. "That skeleton's sixty-five million and three years, two months and eighteen days old," the employee replied. "How can you be so precise?" she asked. "Well, when I started working here, I asked a scientist the exact same question, and he said it was sixty-five million years old-and that was three years, two months and eighteen days ago." Why is this precision false and silly?

Answer: Rounded off, the answer is still sixty-five million, and added the precision is false and silly (except in the context of it being a joke).
[https://en.wikipedia.org/wiki/Mathematical joke and refence cited therein]

## * Ch. 8. Estimating gallons (Section 8.2)

Estimate the cost of a gallon of a liquid.

Answer: What type? For water it would be nearly zero, for milk and gasoline it is well known, for very costly liquid, well ...
** Ch. 8. Estimating the size of an acre (Estimating Section 8.2)
Refer to the problem: "Ch. 4. How square is your acre?" Estimate the length of a square acre.
Answer: The square of 25 is 625, which is approximately 640. So, a square mile is approximately a square with 25 acres along a side. Rounding off the length of a mile to 5,000 feet, the length of an acre square can be estimated to be 5,000/25 or $\sim 200$ feet, which is close to the correct answer of $\sim 208.7$ feet in the noted problem.
** Ch. 8. Estimating the relative area of a dunam and an acre (Estimating Section 8.2) From "Ch. 4. A dunam vs. a square mile", a square with an area of a dunam (a unit of land area that is used in regions of the former Turkish empire, such as in Israel) has a side of length 98.425 feet, while that for an acre has length 208.7. Estimate how much larger an acre is than a dunam.

Answer: Estimating 208.7 as 200 and 98.425 as 100, an acre is (200/100) ${ }^{2}=4$ times larger. Instead, estimating 208.7 as 210 and 98.425 as 100, an acre is (210/100) $)^{2} 4.4$ times larger, which is closer to the actual answer of $(208.7 / 98.425)^{2} \sim 4.496$.

## * Ch. 8. Estimating daily deaths (Section 8.2)

Estimate the average number of people dying in the U.S. every day.
Answer: With 320 million living to 80, it means 4 million die annually and so roughly 4 million/400 days per year or 10,000 dying daily.

## ** Ch. 8. Are given estimates and statistics all consistent with each other? (Section 8.2)

 In 2021 movie The Map of Tiny Perfect Things, character Margaret tells fellow lead character Mark, both presumably American teenagers, that 150,000 people die every day and then a bit later that 19 million people have a birthday on that day. (We presume this means a birthday anniversary and not the actual day of birth.) Is something strange about these estimates? If so, what?Answer: Presumably these are statistics, but let's make some estimates. 150,000 people dying every day means that 55 million die each year, as is also given in https://ourworldindata.org/births-and-deaths for 2015. If $\sim 1 / 70^{\text {th }}$ die each year (from life expectancies), it means that there are about 4 billion people in the population group. If 19 million people have a birthday on that day, and each day has equal probability as a date of birth (aside from Feb. 29), it would mean there are $\sim 7$ billion people. There are almost 8 billion people on Earth (2021). Given all the assumptions in these estimates and analyses, all is relatively consistent. (Also, both numbers refer to all people, not just to those in the U.S.)

## * Ch. 8. Assessing burn injuries by using nines (Math assessments, estimates, metrics, Sections 8.2 and 16.4.2)

To estimate the percentage of a body's surface area that has burns, physicians use the "Wallace" Rule of Nines for skin to assess treatment options. On average, approximately $9 \%$ of the body skin covers the front and back of the head and neck (of which $1 / 2$ is in the front and $1 / 2$ in back); each arm and hand (of which $1 / 2$ is in the front and $1 / 2$ in back); the chest; the abdomen; upper back; and lower back. Each leg (including foot) is twice that value, or $18 \%$ (of which $1 / 2$ is in the front and $1 / 2$ in back). That makes $99 \%$, and the genital area adds $1 \%$ to make $100 \%$. This is for adults. Generally, how would you expect this rule to be different for children?

Answer: Compared to adults, children tend to relatively have larger heads and shorter legs. In fact, for children, in this rule, the head region fraction is doubled, to $18 \%$, and each leg fraction is decreased by a quarter, to $13.5 \%$.

## ** Ch. 8. How many suitcases can fit in a trunk? (Section 8.2, Estimating; and Modeling)

 It is not trivial to see how to pack a trunk with as many suitcases as possible. Let's make a model to address fitting suitcases in a trunk with a very specific (and limited) model. Say all of your suitcases are 3.0 feet wide, 0.5 feet high and 2.0 feet deep. They are exactly rectangles on each side (as so are rectangular solids (and cubes are a special case with squares on each side)), and you can ignore protruding handles, wheels and so on. Your trunk space is idealized as exactly a rectangular solid space. For our purposes, say you must align the width, height, and depth dimensions of each suitcase respectively along the corresponding width, height, and depth dimensions of the trunk. How many suitcases can fit into this trunk if its dimensions are exactly (a) 6.0 feet plus 1 inch wide, 2.0 feet plus 1 inch high and 4.0 feet plus 1 inch deep, (b) 6.0 feet less 1 inch wide, 2.0 feet less 1 inch high and 4.0 feet less 1 inch deep, and (c) 6.0 feet wide, 2.0 feet high and 4.0 feet deep? Can you do better than this if the suitcases were allowed to be put in differently (or if they had different sizes)?Answer: This question addresses estimating, making a model, packing objects of the same size and shape. (a) You can fit the suitcases $6.0 / 3.0=2$ across, 2.0/0.5 $=4$ high, and $4.0 / 2.0=2$ deep, all with a little extra space because of that extra inch, and so you can fit $2 \times 4 \times 2=16$ suitcases, with little space leftover. (b) Because you have 1 less inch in each dimension you lose a suitcase along each dimension, so you can fit only $1 \times 3 \times 1=3$ suitcases. The loss of that 1 inch has a humongous impact with the given constraints. If you are allowed to place the suitcases differently you can certainly fit many more than 3 (try some packing methods), but never as many as 16. This illustrates how tricky packing can be and how constraints have great impact in modeling and in the real-life situation. (c) As a practical matter you can never fit as many suitcases as in (a) because you need some space between them.

These suitcases have a volume 3.0 feet $\times 0.5$ feet $\times 2.0$ feet $=3.0$ cubic feet. For the trunk volume of 6.0 feet $\times 2.0$ feet $\times 4.0$ feet $=48.0$ cubic feet, you could fit a maximum of a 48.0 cubic feet $/ 3.0$ cubic feet $=48.0 / 3.0=16$ suitcases, if there is a little wiggle room as in (a) and if they are packed in the way described. If the suitcases still had this same volume but various
dimensions each, rarely could you fit as many of them in the trunk. Of course, trunks are not perfect rectangular solids.
(A question on fitting suitcases in car trunks was suggested to the author by Artem Ponomarev.)

* Ch. 8. To interpolate or not to interpolate? That is the question. (Section 8.2.4)

In a given location the mean outdoor temperature on September 10 is $70.4^{\circ} \mathrm{F}$ and that on September 20 is $67.0^{\circ} \mathrm{F}$. Is it proper to interpolate and determine the expected mean temperature on September 15? If so, what would it be?

Answer: Yes. The dates are close enough that a linear interpolation---i.e., assuming a linear change with day-is reasonable. A perfectly linear change with day would give the average, or $68.7^{\circ} \mathrm{F}$. Using the data citied in Section 16.1.1 for Central Park NYC, it is seen not to be exactly a linear change, and it turns out to be $68.3^{\circ} \mathrm{F}$, and so still okay.

## * Ch. 8. More on: "To interpolate or not to interpolate? That is the question." (Section 8.2.4)

In a given location the mean outdoor temperature on January 1 is $33.3^{\circ} \mathrm{F}$ and that on December 1 is $42.5^{\circ} \mathrm{F}$. Is it proper to interpolate and determine the mean temperature on July 1? If so, what would it be?

Answer: (This uses data cited in Section 16.1.1 for Central Park NYC.) It would be if you knew a good fit for many days between these beginning and end-dates, but if you do not and use a linear fit, it would not make sense. There is warming and then cooling between these dates. By the way, for July 1 the mean temperature is $75.5^{\circ} \mathrm{F}$.

## * Ch. 8. To extrapolate or not to extrapolate? That is the question. (Section 8.2.4)

In a given location in New York City the mean outdoor temperatures on April 1 and on August 1 are known. Is it proper to extrapolate outside this range and determine the mean temperature on September 1?

Answer: Because of the change of seasons, this would usually not be reasonable to do this.

## ** Ch. 8. Scaling of motorcycles (Scaling, Section 8.2.5)

You buy a 1:12 toy model of a motorcycle. If it accurately reproduces every facet of the motorcycle-even the materials of each part-how much would it weigh? Assume a typical motorcycle weighs 700 pounds. How does the weight of an actual toy model to compare to this?

Answer: Because volume scales as the cube (third power) of length and weight scales as volume, the model weight would be $700 / 12^{3}=700 / 1,728=0.405$ pounds $\sim 6.5$ ounces. The actual model will be much lighter than this because it is largely made of plastic, which is less dense than metals such as steel (by a factor of $\sim 8$ ).

## *** Ch. 8. How do the challenges in management scale with the size of a department? (Section 8.2.5, Scaling; Section 4.10, Combinations)

World-famous mathematician Stan Ulam once said that the difficulty in running an academic department at a university increases with its size, not as the number of faculty members in it (N) but as its square $\left(\mathrm{N}^{2}\right)$, because this is how the number of pairs of faculty members in that department scales. (That is from his observation as chair that difficulties arose from disputes between two given faculty members and so the more pairs, the more the potential for difficulty. (This has not been my experience.)) (a) Assuming such pairs are random and that a quantitative measure of difficulty were in fact proportional to the number of such pairs, was his math correct?
(b) If there are 60 faculty members, roughly how many pairs are there?

Answer: (a) Yes. There are $N$ ways to choose the first member of a pair and $N-1$ the second, so there would be $N(N-1)$ pairs. However, it does not matter which is the first person of the pair chosen, so only half of these are distinct pairs, and so there are $N(N-1) / 2=N^{2} / 2-N / 2$ of them. (This is the same as the number of combinations of choosing 2 things from $N$ items: $N!(2!\times(N-$ 2)!.).) For $N \gg 1$, the square term dominates and the number of pairs would vary as the square of the department size (aside from that factor of 2 in the denominator). (b) The number of pairs is $(60 \times 59) / 2$, which is roughly $60 \times 30$ or 1,800 pairs.
(The Adventures of a Mathematician, Stanislaw Ulam, pg. 91)

## * Ch. 8. Are supposedly randomly-generated number not random? (Random number generation and usage, Section 8.3)

A 2020 investigation shows that the standard book of random numbers by the Rand Corporation from 1955, may not have been as random and unbiased as expected and thought to be at the time. They were obtained from a physical method: voltage fluctuations in a circuit---which is a good method, and then converted to 0 s and 1 s and then to single digits from 0 to 9 . The numbers were random but their order may not have been, but is this bad?

Answer: Yes. There would be bias in using them unless both the numbers and their order were random. (It is thought that the "computer cards" with data (with all random data on the cards), may be have been dropped and picked up, but not in the original order, which led to some bias in the random number able.
[https://www.wsj.com/articles/rand-million-random-digits-numbers-book-error-11600893049 'A Million Random Digits’ Was a Number-Cruncher’s Bible. Now One Has Exposed Flaws in the Disorder. By Michael M. Phillips Updated Sept. 24, 2020 12:43 pm ET]

## Chapter 9. The Different Types of Numbers Have Not Evolved, But Our Understanding of Them Has

## * Ch. 9. Is your 8 consecutive? (math joke)

Use a calculator (or just copy and paste the following into your browser) to show that an amazing yet seemingly coincidental approximation for the number 8 is $987654321 / 123456789=8.0000000729 \ldots$
[See the very unexpected approximation $\mathrm{e}^{\pi}-\pi \sim 20$ in Chapter 9.
(Fermat'sLibrary@fermatslibrary
https://www.buzzfeed.com/daves4/funny-math-jokes
(https://apple.news/AYiITz707QnueM_XwHshNGg))]

## * Ch. 9. Is your 8 consecutive enough? (math joke) (Chapter 9)

Let's follow up the quite good approximation for 8 given in the previous problem, that $987654321 / 123456789=8.0000000729 \ldots$ This uses count-ups and count-downs between 1 and 9. (a) But, would such a fraction still be a good approximation with shorter numerator and denominator strings of numbers? Say each string stops after 8 numbers (instead of 9), or 7 numbers and so on? (b)What happens if instead the countdowns lists are partly in the numerator and partly in the denominator, such as with $9876 / 12345$ (with all the integers, but 4 of them in the numerator and 5 in the denominator).

Answer: (a) It could still be pretty good. With 4 numbers, 9,876/1,234 = 8.003... (b) 9876/12345 is exactly 0.8, so 9,876/1,234.5 is exactly 8.

## * Ch. 9. Do we positively need negative numbers? Yes! (Chapter 4, Linking numbers; Chapter 9, Different types of numbers)

We grow up with the obvious need for counting, and so with the need for positive integers. So, are all of the other number systems discussed in Chapter 9 needed at all? Yes, we need them for us to use usual mathematical operations, such as addition, subtraction, multiplication and division. When you add or multiply positive integers, you obtain another positive integer. But what happens with, for example, subtraction of only positive integers, and why does this show we need to expand our concept of numbers beyond the counting numbers?

Answer: When you subtract two positive numbers, such as in 5-3 you obtain a positive number, 2, but what happens when you subtract $3-5$ ? You need to expand the number base to negative integers, and then you obtain -2. What happens when you subtract the positive number 3 from itself, 3-3? You get a result that indicates neither a positive nor a negative integer, and you need to introduce 0, to explain the result. The need for other number systems result from similar arguments.

Mark Kac and Stanislaw M. Ulam, Mathematics and Logic, page 28

# * Ch. 9. Is it rational or irrational to contribute to a Pi day fundraiser? (Ch. 3 pi day, Ch. 9 Irrational numbers) 

I was asked by a fundraiser on pi day to contribute $\$ 100$ times pi or $\$ 314$ on pi day, $3 / 14$ or March 14. I said it was irrational. What was irrational, the request?

Answer: No, pi. But, it is rational to use a rational approximation to pi, 3.14, for the request. (But this is really 100 x this rational approximation to this irrational number.)

* Ch. 9. How to "liv" with new professional golf associations? (Ch. 9, Roman numerals) The Professional Golfers' Association of America (PGA) has been the well-established association for professional golfers. It holds tournament with 4 rounds of golf with 18 holes each. In 2022, a new association, the LIV Golf, became more competitive and competitive with the PGA. One feature of it is that its tournaments consist of three rounds of 18 holes. What is the likely origin of its name?

Answer: Three rounds of 18 holes means the tournaments have 54 holes, and LIV is 54 in Roman numerals. $(L=50, I V$ is 4$)$
** Ch. 9. Why don't Roman numeral clocks use the correct Roman numerals? (Chapter 9) After many decades of life, I finally noticed that the Roman numeral representation of 4 in Roman numeral clocks is usually IIII instead of the correct numeral IV, and which became the standard notation a very long time ago. A few clocks do use IV, as does Big Ben in London (https://mol.im $/ \mathrm{a} / 10766885$ ). Why is this different numeral? There are many potential explanations, but no one knows for sure. One good possibility is that IIII balances the symmetry of 4 with 8 (VIII) better than IV does. Another theory is that the number of reasonably-sized molds one would need to cast enough Is, Vs, and Xs for a clock is smaller when IIII is used instead of IV. Why would this be so?

Answer: With IIII, molds are needed to make I, II, III, IIII, V, VI, VII, VIII, IX, X, XI, and XII, so one needs 20 Is , 4 Vs , and 4 Xs. One could make these with one mold with $5 \mathrm{Is}, 1 \mathrm{~V}$, and 1 X , cast four times, with no leftovers. With IV, molds are needed to make I, II, III, IV, V, VI, VII, VIII, IX, X, XI, and XII, so one needs $17 \mathrm{Is}, 5 \mathrm{Vs}$, and 4 Xs. Aside from one big mold that could handle all of these, one would need more than one type to make these with no leftovers (and no partial filling of molds). One possible way of doing this is with one mold producing 4 Is and 1 V cast four times and a second mold with 1 I, 1 V , and 4 Xs cast once.
https://www.mentalfloss.com/article/24578/why-do-some-clocks-use-roman-numeral-iiii https://monochrome-watches.com/why-do-clocks-and-watches-use-roman-numeral-iiii-instead-of-iv/

## Chapter 10. Really, Really Big and Really, Really Small Numbers

## * Ch. 10. The 1234567 password (Section 10.1, Passwords)

You are told to choose a password containing 7 numbers and you choose 1234567 because it meets these requirements, it is easy for you to remember, and it is very unlikely for someone else to guess. You think it is very unlikely because the probability that anyone would guess the first number is $1 / 10$, the second number $1 / 10$, and so on, and since they are independent of each other, the overall probability is the product of seven $1 / 10 \mathrm{~s}$, or one out of ten million. Does this make sense?

Answer: No. It would make sense if a machine with no prior information were to randomly guess all possibilities, but this is not the case. It is well known that 1234567 is a very commonly chosen password, because it is easy to remember. In fact, in a recent survey it was the most common password used by CEOs, so it is clearly a very poor password. 12345, 123456789, 1234, and 12345678 respectively ranked \#3, 4, 6, and 10, while 111111 ranked 9 . What was the second most frequently used password? "password"
(https://mol.im/a/10778135)

## ** Ch. 10. How many big numbers? (Section 10.1, Big Numbers)

How many different 10 -digit numbers can you make using each digit from 0 to 9 only once?
Answer: One might think that the answer is $10 \times 9 \times 8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1=10!=$ 3,628,800 because you have 10 possibilities for the first digit, 9 for the second and so on. However, 0 cannot be the first digit, because if it were it would be a 9-digit number. So, the number is $9 \times 9 \times 8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1=3,265,920$.

From
The Moscow Puzzles, Boris A. Kordemsky \#340A, pg. 157

## Chapter 11. The Whole Truth of Whole Numbers (or The Numbers Racket)

## ** Ch. 11. Smallest number with factors from 1 to 10 (Chapter 11.2. Factoring and Prime Numbers)

What is the smallest integer divisible by every integer from 1 to 10 ?

Answer: 2,250 . The product $1 \times 2 \times 3 \times 4 \times 5 \times 6 \times 7 \times 8 \times 9 \times 10=3,628,800$ is certainly divisible by each, but it is not the smallest number. You need to include only the minimum number of factors of 2, 3 and other prime numbers that multiply to form each number. First keep the prime 2 and 3. 4 has two factors of 2, so to cover the 4 you need a second factor of 2. You need the prime 5. You don't need the 6 because there are already factors of 2 and 3. You need the prime 7. You need a third factor of 2 to cover 8. You need a second factor of 3 for 9. Your existing factors of 2 and 5, cover for 10 . So, this number is $2 \times 2 \times 2 \times 3 \times 3 \times 5 \times 7=2,520$.

It is said that "Scholars discovered 2,250 in hieroglyphs engraved on the stone lid of a tomb in an Egyptian pyramid.", perhaps to honor this.

From
The Moscow Puzzles, Boris A. Kordemsky \#306, pg. 135

## ** Ch. 11. Factoring and primes (Prime numbers, Section 11.2)

2,021 (the year when this problem was written) can be factored into the product of two prime numbers. What are they?

Answer: 43 and 47. The straightforward and laborious way of factoring is to divide the number first by 2, and then by 3, and so on until you get a quotient and no remainder., and then divide that quotient by 2 and so on again and again. In this case you will get a remainder until you reach 43.

## ** Ch. 11. Perfect numbers and access codes (Coding information, Section 7.3; Perfect numbers, Section 11.2)

The series Lewis (aka Inspector Lewis) is a sequel to the famed detective series Inspector Morse. Its pilot centers about Oxford students and faculty highly skilled in math, including a student Danny who is a prime suspect in committing murder. It turns out his access code (of 3 two-digit numbers) to a building involved in the plot was 1247 14. (a) Show how this code relates to the second smallest perfect number, 28. (b) It turns out that Danny used 496 for the combination of his lock and 8128 as computer password. Why?
(Spoiler alert: The murderer in this episode is a fictitious Fields medal winner, who won it for advancing the understanding of Goldbach's conjecture (Section 11.2.2).)

Answer: (a) The factors of 28, which add up to 28, are 1247 14, and so Danny used 124714. (b) They are the third and fourth smallest perfect numbers.

# Part III: The Second World of Math: The Math of Doing 

## Chapter 12. The Math of the Digital World: Modular Arithmetic (or Using Number Leftovers)

## * Ch. 12. Division and remainders in the classic Leave it to Beaver TV show (Remainders

 Section 12.1)In the classic "Leave it to Beaver" TV show, $5^{\text {th }}$ grader Beaver (Theodore Cleaver) was having trouble with division for several reasons including, as he told his older brother Wally, that when you divide whole numbers there can be "a whole bunch of junk left over." Wally replied that there is nothing wrong with that, but Beaver countered "There is when you don't know what to do with it." (Season 4, Episode 27, Beaver's Report Card, April 1, 1961-and often shown in reruns). What were they talking about?

Answer: When you divide one integer by another the remainder may not be zero, and this is the "junk" that concerned Beaver. It most definitely is not junk. The remainder has important meaning and can be expressed as a whole number remainder, as a fraction or in decimal form) and, specifically, as the whole number reminder in modular arithmetic (Section 12.1).

## * Ch. 12. From modular arithmetic to Japanese gangsters (Modular math Section 12.1 and meaning Chapter 3)

The name for Japanese gangsters, $y a k u z a$, is derived from the score one receives with cards having values 8-9-3 in the Japanese card game of oichokabu, which is played with hanafuda "flower cards." (With 8, which is yattsu, shortened to ya; 9, which is ku; and 3, which is san, changed to of za; to give ya-ku-za.) 893 was a bad hand because it indicated "no points" or "useless", which later was interpreted as "useless people" or "gambling people", and so the yakuza. The points you received for this hand is the sum, mod 10. Explain why this score indicated no points would be earned.
https://www.sljfaq.org/afaq/yakuza.html
Answer: $8+9+3=20$, which is $0 \bmod 10$.
[https://www.sljfaq.org/afaq/yakuza.html]

## ** Ch. 12. Modular arithmetic and the combinatorics of flower cards (Modular math Section 12.1, arrangements/combinations Section 4.10))

A variation of this game oichokabu can be played with our usual deck of cards, but with no kings, queens or jacks, and with aces counting as 1s. You choose three cards and add the numbers, mod 10, and that is the number of points you get. (a) How many cards are in the deck, as described? (b) How many ways can you choose 3 cards (with no replacement), if order counts and again if order does not count? To see how frequently you get "no points" or "useless", with
the point you receive being the is the sum of the three cards, mod 10, consider case with 10 cards instead of 40 , as in the next problem.

Answer: (a) $52-12=40$. (b) If order counts: $40 \times 39 \times 38=40!/ 37!=98,760$, and if order does not count: $40 \times 39 \times 38 /(3 \times 2 \times 1)=40!/(37!3!)=9,880$.

## *** Ch. 12. More on: "Modular arithmetic and the combinatorics of flower cards" (Modular math Section 12.1, arrangements/combinations Section 4.10))

(a) Let's simplify the game in the previous problem, by using a deck with only 10 cards, from 110 all of the same suit. How many ways can you choose 3 cards (with no replacement), if order counts and again if order does not count.)? (b) What fraction of the time do you get a 0 (the sum $\bmod 10)$ after choosing three cards if order does not matter. (Consider each of the possibilities, starting with 1 , and then 2 , and then $\ldots$, rather than using the factorial argument you likely used in part a.)

Answer: (a) If order counts: $10 \times 9 \times 8=10!/ 7!=720$ ways, and if order does not count: $10 \times 9$ $\times 8 /(3 \times 2 \times 1)=10!/(7!3!)=120$ ways. (b) These possibilities are, first starting with $1, \ldots:$
1, followed by sequences of two of the remaining number that sum to 9 or 19, so 1-2-7, 1-3-6, 1-4-5, and 1-9-10.
2, followed by combinations of two larger numbers that sum to 8 or 18 (since we already covered combinations with 2 (used once)), so 2-3-5 and 2-8-10.
3, followed by combinations of two larger numbers that sum to 7 or 17 (since we already covered combinations with 2 and 3 (used once), so 3-7-10 and 3-8-9 (because there are none than sum to 7).

4, followed by combinations of two larger numbers that sum to 6 or 16 (since we already covered combinations with 2, 3, and 4 (used once)), so 4-6-10 and 4-7-9 (because there are none than sum to 6).
5, followed by combinations of two larger numbers that sum to 5 or 15 (since we already covered combinations with 2, 3, 4, and 5 (used once)), so 5-6-9 and 5-7-8 (because there are none than sum to 5).
6, followed by combinations of two larger numbers that sum to 4 or 14 (since we already covered combinations with 2, 3, 4, and 5 (used once)), so 5-6-9 and 5-7-8 (because there are none than sum to 4).
7, followed by combinations of two larger numbers that sum to 3 or 13 (since we already covered combinations with 2, 3, 4, and 5 (used once)), so none (because there are none than sum to 3 or 13.).

So, there are 14 such combinations, with order not counting, compared to the total number of ways to draw three cards with order not counting $=10!/(7!3!)=(10 \times 9 \times 8) /(3 \times 2 \times 1)=120$.

There are $14 \times 3!=14 \times 6=84$ ways to draws 3 (distinct)cards whose sum ends with a 0 , with order counting, compared to total numbers of ways you can draw three cards when order matters $=10!/ 3!=10 \times 9 \times 8=720$.

In neither case is the fraction of times the residue 0 exactly $1 / 10$ of the total number (so the final number in the sum is not equally likely to range from 0 to 9). Why?

## Chapter 13. The Math of What Will Be: Growing and Decaying Sequences, Progressions, and Series

* Ch. 13. Arithmetic progression from the number of candles in a box (Section 13.1)

On a certain holiday, on the night of the first day a main candle is lit along with 1 other (of the same type), on the second day the main candle is lit along with 2 others, on the third night it is lit along with 3 others, and so on. A box with all of the candles needed for this holiday contains 44 candles. For how many days does this holiday last?

Answer: 8 days. The straightforward and harder way of solving this is to add up the numbers 2, $3,4,5, \ldots$ until you obtain 44 and you will see there are 8 terms. The easier way is to remember that the sum of the terms in an arithmetic progression is the product of the number of terms in it and their average. For an even number of terms, this equals half of the number of terms times the sum of the first and last terms. 44 is the product of 4 and 11 (and this factor of 11 is revealing), so we could "guess" that there are 4 pairs of terms or 8 terms representing 8 days. Then 11 would be the sum of the first and last terms, so they could be 0 and 11, 1 and 10,2 and 9, 3 and 8, 4 and 7, and so on. But, the first term is 2, so the pair would be expected to be 2 and 9, for which there would be 8 terms: 2, 3, 4, 5, 6, 7, 8, and 9-so there is consistency with the correct answer of 8 days. (This is for the holiday Chanukah.) (Caution: One can be fooled with such an analysis, because there are 45 candles in a few higher-end boxes of candles, because one candle is added in case one is delivered broken.)

## * Ch. 13. Management reporting lines and company sizes- uniform branching (Geometric progressions, Section 13.1)

In a corporation, 5 vice presidents report to the president. 5 highest-level managers report to each of these vice presidents, and this continues with mid-level managers, lowest-level managers, and finally workers. How any people are in this corporation?

Answer: There is 1 president ( $5^{0}$, which is 1 ), and then there are 5 vice presidents ( $5^{1}$ ), $5 \times 5=$ 25 highest-level managers ( $5^{2}$ ), $5 \times 5 \times 5=125$ mid-level managers $\left(5^{3}\right), 5 \times 5 \times 5 \times 5=625$ lowest-level managers $\left(5^{4}\right)$, and $5 \times 5 \times 5 \times 5 \times 5=3,125$ workers $\left(5^{5}\right)$. So, there are $1+5+$ $25+125+625+3,125=3,906$ people in this firm with 6 levels of reporting lines $\left(5^{0}+5^{1}+5^{2}\right.$ $+5^{3}+5^{4}+5^{5}$ ). This is a geometric progression.

## * Ch. 13. Management reporting lines and company sizes - variable branching (Geometric progressions, Section 13.1)

There is no reason there needs to be the same number of people reporting to their respective managers in branching from one rung to the next in the management ladder, making for a geometric progression as in the previous problem. Let's say 6 workers report to lowest-level
managers, 4 of them report to mid-level managers, 5 of them report to highest-level managers, 3 of them report to vice presidents, and 7 of them report to the president. How any people are in this corporation? (Of course, the number of people reporting to the next level need not be the same in the branching across each rung.)

Answer: There is 1 president, and then there are 7 vice presidents, $3 \times 7=21$ highest-level managers, $5 \times 21=5 \times 3 \times 7=105$ mid-level managers, $4 \times 105=4 \times 5 \times 3 \times 7=420$ lowest-level managers, and $6 \times 420=6 \times 4 \times 5 \times 3 \times 7=2,520$ workers. So, there are $1+7+$ $21+105+420+2,520=3,074$ people in this firm with these 6 levels of reporting here.

## * Ch. 13. Geometric progressions in a slowing epidemic (Section 13.1)

There are 1,000 new cases reported at the peak of an epidemic during one week. The number of new cases reported decreases by a factor of 2.0 each successive week. A total of how many new cases are reported, including that first week?

Answer: Each week it becomes the current value/ 2.0 or current value $\times 0.5 \quad 1,000+500+$ $250+125+\ldots=2,000 \quad 1,000 /(1-0.5)=1,000 / 0.5=2,000$. (Ignore fractions, values less than 1, so this is an estimate, based on a geometric progression.)

## ** Ch. 13. Number of possible outcomes in a tournament (Geometric progressions, Section

 13.1; Probability, Chapter 15)In a tournament starting with 64 teams, the teams play and leave the tournament the first time they lose. The 64 teams in 32 games (with teams pre-selected into \#1-\#32 pairs of opponents), the 32 winners then play in 16 games (with winners of pre-selected pairs playing each other, such as the winner of the \#1 pair plays the winner of the \#2 pair, the winner of the \#3 pair plays the winner of the \#4 pair, and so on), and so on. (a) How many games are there in the tournament? (b) How many scenarios are there for winners and losers each time the tournament is played?

Answer: (a) There are 64 teams and each leaves the tournament when it loses, so only 1 team has no losses at the end of the tournament and so there are 63 total games. Also, in the first round there are 32 games, in the second round 16 games and so on, so there are total of $32+16$ $+8+4+2+1=63$ games. This is the sum of a geometric progress with each term decreasing by a factor 2, that starts at 64 and ends at 1. (This is how the NCAA Division I Men's Basketball Tournament was played from 1985 to 2000, with 64 teams with pre-set pairs of teams playing each other, before it was expanded to 65 teams in 2001 and then to 68 in 2011.) (b) There are 2 possible outcomes for each game so there are $2^{63} \sim 9.22 \times 10^{18} \sim 9$ billion billion possibilities. (If the pairs playing each in a given round were not determined by the winners of pre-set pairs in the previous round, there would be even more possibilities.)

## ** Ch. 13. Scaling of expected hurricane damage using a rule of thumb (Scaling Section 8.2.5, geometric progression Section 13.1)

In a Category 1 hurricane the maximum wind speeds are from 74 mph to 95 mph , in a Category 2 storm they are from 96 mph to 110 mph ; for Category 3 from 111 mph to 129 mph ; for Category 4 from 130 mph to 156 mph ; and for Category 5 from 157 mph on up. The damage expected from a hurricane roughly scales by a factor of 4 for every increase in category by 1. (a) Explain why this "rule of thumb" means that the damage increases as a geometric progression (or exponentially) with category number. (b) How much more damage is expected in a 145 mph storm relative to a 85 mph one? (c) How much does wind speed need to increase for the damage to double?
[https://www.wsj.com/articles/storm-isaiass-most-damaging-winds-were-on-its-right11597397402
https://www.weather.gov/jetstream/tc_potential]
Answer: (a) This is how a geometric progression varies, here with the damage multiplied by a factor of 4 for each increase in category number by 1. Long before this was quantified this way, the category number system was devised by empirical levels of observed damage. (b) This is a mid-range Category 4 level storm relative to a mid-range Category 1 one, so the damage is larger by $\sim 4 \times 4 \times 4=4^{3}=64$. (c) The ranges of wind speeds within Categories 1 to 4 are 21, 14, 18, and 26 mph , which average to 20 mph per Category and concomitant increase in damage by a factor 4 , so for an increase in wind speed of 10 mph damage will increase by $\sim 2$.

## ** Ch. 13. The new Moore's law: Huangs' Law (Section 13.1.1)

A follow-up to the so-called Moore's Law that the number of transistors in a chip seems to double every two years or so, is Huang's Law that the performance of silicon chips that power artificial intelligence (AI) more than doubles every two years (due to improvements in hardware, such as the number and speed of transistors, and the quality of software). Between November 2012 and May 2020, the performance of an important class of chips for AI calculations increased by a factor of 317 times. Does this support the Huang's Law assertion?
[Huang's Law Is the New Moore's Law, and Explains Why Nvidia Wants Arm By Christopher Mims Sept. 19, 2020 12:00 am ET
https://www.wsj.com/articles/huangs-law-is-the-new-moores-law-and-explains-why-nvidia-wants-arm-11600488001
Appeared in the September 19, 2020, print edition as 'Moore's Law Is Dead. Long Live Huang's Law. ']
Answer: From November to 2012 to May 2020, would be 7 1/2 years or almost 4 cycles of doubling, so by $\sim 2^{4}$ or 16 times; however, this is slower than what is actually seen. Doubling every year would mean 8 cycles in this time frame or increasing by $2^{8}$ or 256 times, so, perhaps there is a doubling of AI chip performance closer to every year, which is faster than Huang's law would suggest.

## ** Ch. 13. High simple and compound interest rates with loan sharks (Section 13.2)

 Loan sharks charge "excessive" interest", which is sometimes owed on a periodic basis (without paying off the principal). As used on the TV classic The Soprano's, it can be called the "vig" and expressed as a dollar amount or in terms of points, with frequent compounding and increased amounts (and other forms of payment) if not paid on time. Some organization charge $15 \%$ biweekly for pay day loans. What is the effective annual interest rate for this, if you pay the interest only every two week or if you do not pay it back and it compounds every two weeks?Answer: There are 26 two-week periods every year, so if you pay interest only every two weeks over a year (the vig each time), the interest rate per year is $15 \% \times 26=390 \%$. So, you are paying $3.9 \times$ the principal-and also owe the principal. This is simple interest, but still very excessive. If you allowed not to pay it every two weeks and are instead allowed to let it compound, at the end of the year you owe in interest the principal times $(1.15)^{26}-1$, which is an annual, compounded rate of $3,686 \%$ or $36.86 \times$ the principal, and you also have to pay the principal.
[https://www.youtube.com/watch?v=DIW5wYIZtHY https://www.investopedia.com/terms/l/loansharking.asp https://alearningaday.blog/2014/12/03/compound-interest-and-loan-sharks-mba-learnings/amp/]

## ** Ch. 13. How fast can the annual inflation rate fall? (Compound interest, Section 13.2)

 (a) The inflation rate increases by $0.6 \%$ in 12 successive months. What is the annual inflation rate at the end of these 12 months? (In this problem assume these changes add as in simple (and not compound) interest.) (b) Then the monthly rate decreases to $0 \%$ in each of these three next months. What is the annual inflation rate at the end of each of these months? (c) If instead, the monthly rate becomes $-0.6 \%$ in each of these three next months. What is the annual inflation rate at the end of each of these months?Answer: (a) The annual rate is $0.6 \% \times 12=7.2 \%$. (b) At the end of each of these three months, the annual inflation decreases to: $0.6 \% \times 11=6.6 \%, 0.6 \% \times 10=6.0 \%$, and $0.6 \% \times 9=5.4 \%$, respectively. So, if prices do not decrease in these months, the annual rate (for the prior 12 months) can never decrease very rapidly. (c) Now, at the end of each of these three months, the annual inflation changes to: $0.6 \% \times 11+(-0.6 \%)=6.0 \%, 0.6 \% \times 10+(-0.6 \%) \times 2=4.8 \%$, and $0.6 \% \times 9+(-0.6 \%) \times 3=3.6 \%$, respectively.

## ** Ch. 13. Compounding inflation troubles (Compound interest, Section 13.2)

The inflation rate increases by $0.6 \%$ in 12 successive months. What is the annual inflation rate at the end of these 12 months if you correctly include the compounding nature of the changes.

Answer: Prices change by a factor of 1.06 each month so to find the prices at the end of the 12 months one needs to multiply together 12 such factors: (1.006) $)^{12} \sim 1.0744$. So, the increase and actual annual inflation rate is $1.0744-1=0.0744=7.44 \%$. (It was $7.2 \%$ without compounding.)
** Ch. 13. Exponential decay of total neuron length with age (Section 13.4)
(a) The total length of (sheathed with electrical insulation or myelinated) axons in neurons (nerve cells) that hook up to other neurons in the human brain of a 20 -year old male is $176,000 \mathrm{~km}$. How many times would it wrap around the equator? This measurement includes only the brain "wires" that are coated with electrical insulation, called myelin.
(b) Between ages 20-80, the total length of myelinated axons in the brain decreases steadily with age, for a total loss about $45 \%$ for both males and females. The loss of this total length is thought to be one indicator of overall decreased cognitive abilities. Show that this corresponds to a rate of $1 \%$ loss/year and $10 \%$ loss/decade (assuming these rates are independent of age).
[Marked loss of myelinated nerve fibers in the human brain with age
L Marner, JR Nyengaard, Y Tang, B Pakkenberg
Journal of comparative neurology 462 (2), 144-152;
Aging and the human neocortex
B Pakkenberg, D Pelvig, L Marner, MJ Bundgaard, HJG Gundersen, ...
Experimental gerontology 38 (1-2), 95-99;
https://www.psychologytoday.com/us/blog/the-new-brain/201106/brain-wiring
Brain Wiring: After age 20 it's all downhill R. Douglas Fields Ph.D.
The New Brain Posted Jun 21, 2011]
Answer: (a) The Earth circumference at the equator is $40,075 \mathrm{~km}$ (24,901 miles), so the total length is 176,000/40,075-4.39 or about four and a half times, (b) $-\ln (0.55) / 60=0.9964 \% \sim$ $1.0 \%$ and so $\sim 10 \%$ per decade.

## ** Ch. 13. How rounding off can affect estimates of exponential decay of total neuron length (Exponential decay Section 13.4, rounding off Section 8.1.1)

In the previous problem, the loss of total sheathed neuron length could be given was $1 \%$. It turns out it is close to $1.0 \%$, but it would still be $1 \%$ if rounded off from $0.6 \%$ or from $1.4 \%$. Show how the $\%$ decrease in neuron length from ages 20 to 80 would be different for annual rates of decay of $0.6 \%, 1.0 \%$ and $1.4 \%$. Does such rounding-off lead to significant differences?

Answer: For annual decrease of $0.6 \%, 1.0 \%$ and $1.4 \%$, the total decreases would be $30 \%, 45 \%$ and $57 \%$ respectively. Such rounding off leads to very significant differences, and so it is not proper.

## ** Ch. 13. Decreasing risk from in the exponential-growth spreading of infectious diseases (Spreading of disease, Section 13.6)

The understandable fear of infectious diseases increases when they are easily spread and have dire consequences, and when measures to handle the spreading and these consequences are not adequate. In terms of simple math, can substantial and effective social intervention stop the spread? (The answer is yes---if you know the relevant parameters.)

The basic reproduction number $\mathrm{R}_{0}$ (" R nought") is a unitless model parameter that describes how infectious a disease is; the disease spreads when $\mathrm{R}_{0}>1$. It is the product of the average number of
contacts per unit time an infectious contact makes that produce infection and the infectious period. The first term can be modeled as the product of rate of contact between infected and susceptible individuals and the probability of infection for a contact (the transmissibility). Say $\mathrm{R}_{0}$ is estimated from models to be 3.2 for given set of conditions for a disease that is spread airborne between people and without any social intervention, so disease spreading is expected. What is $\mathrm{R}_{0}$ and is spreading expected if: (a) $30 \%$ of all people wear masks (that totally prevent transmission) and the rate of person-to-person contact is the normal rate, (b) $30 \%$ of all people wear such masks and the rate of person-to-person contact decreases by $40 \%$, (c) $60 \%$ of all people wear such masks and the rate of person-to-person contact is the normal rate, and (d) $60 \%$ of all people wear such masks and the rate of person-to-person contact decreases by $75 \%$.

Answer: One or two of the three factors, the rate of contact and the transmissibility, can cause a a decrease in $R_{0}$ here to: (a) $3.2 \times(1.0-0) \times(1.0-0.30)=3.2 \times 0.7=2.24>1.0$, so there is still spreading, leading to an epidemic, (b) $3.2 \times(1.0-0.4) \times(1.0-0.30)=3.2 \times 0.6 \times 0.7=$ $1.344>1.0$, so there is still spreading, leading to an epidemic, (c) $3.2 \times(1.0-0) \times(1.0-0.60)=$ $3.2 \times 0.4=1.28>1.0$, so there is still spreading, leading to an epidemic, and (d) $3.2 \times(1.0-$ $0.75) \times(1.0-0.60)=3.2 \times 0.25 \times 0.4=0.32<1.0$, so there is no spreading, and the number of infected decrease with time.
[https://www.healthline.com/health/r-nought-reproduction-number https://en.wikipedia.org/wiki/Basic reproduction number]

## ** Ch. 13. Herd immunity protection from the exponential growth spreading of infectious diseases (Spreading of disease, Section 13.6)

When a fraction of the populace is no longer susceptible to an infectious disease, often due to vaccination (leading to a new reproduction number called the effective $R$ or $R_{e}$ ), the possibility of spreading of disease decreases and could become below threshold, as it would be when $R_{0}$ is $<1.0$ without intervention. So, if $\mathrm{R}_{0}$ were 5.0 without intervention and it would "effectively" become lower than 1.0 if $<20 \%$ of the susceptible people became immune, as would happen if $>80 \%$ of them were vaccinated. In this case the threshold for such "herd immunity" would be $80 \%$. So, the herd immunity threshold is 1.0 less the reciprocal of $\mathrm{R}_{0}$, often expressed as $\%$. (a) For the childhood disease the measles, the estimated range of $\mathrm{R}_{0}$ is 12-18. What are the herd immunity values for this range? Explain why even if only $10 \%$ of school children are not vaccinated, there can be an outbreak of this disease? (b) For COVID-19 the range of estimated values of $\mathrm{R}_{0}$ was $2.5-4.0$ for the initial strain. Find the range of estimated herd immunity values for this disease.

Answer: (a) For measles: $1.0-1 / 12=91.7 \%$ and $1.0-1 / 18=94.4 \%$. This disease is so contagious that it will not spread only iffewer than $\sim 6-8 \%$ are not vaccinated. (b) For the less but still quite contagious COVID-19: $1.0-1 /(2.5)=60 \%$ and $1.0-1 / 4=75 \%$. [https://en.wikipedia.org/wiki/Herd_immunity]

## ** Ch. 13. Infectious diseases, exponential growth-herd immunity (Spreading of disease Section 13.6)

You are told that the number infected as a function of time is an exponential with an exponent consisting of $\mathrm{R}_{0}$ plus a first number, and then this sum is multiplied by a second number. (a) What is this first number? (Remember that subtraction of a number is the same as addition of the negative of the number.) (b) What can you say about the second number?

Answer: (a) $R_{0}$ plus the first number must be 0 when $R_{0}$ is 1 (and so there is not growth or decay), so then exponent is 0 (and a number raised to 0 is 1 , and so it would mean no change). So, the first number is -1 , and this part of the exponent would be $R_{0}-1$. (b) The second number needs to be positive, because when $R_{0}$ is $>1, R_{0}-1$ is $>0$, and exponent is positive and the term is larger than 1. Also, when $R_{0}$ is $<1, R_{0}-1$ is $<0$, and then the exponent would be negative and the term is smaller than 1 .

## ** Ch. 13. Numerical thinking about the effectiveness of infectious disease vaccination (Spreading of disease, Section 13.6)

A vaccine to protect the public against an infectious disease is tested among a large group of people by giving half of them a placebo and half the vaccine. Then one waits until a certain total specified number have contracted the disease. The number who contracted the disease who had received the placebo is compared to the number who contracted it after having been innoculated the test vaccine. The test is complete when a certain number of people of those in the study, as determined by statistics, say 150 , have contracted the disease. If the number contracting it who had taken the vaccine is equal to the number who contracted it who had received the placebo, the vaccine is $0 \%$ effective (and so it is not effective); if it is $50 \%$ of the latter group, the vaccine is $50 \%$ effective (which could be the minimum level possible for an approved vaccine); if it is $10 \%$, the vaccine has lowered the number of cases by $90 \%$ and it is $90 \%$ effective, and so on. If the probability that an unvaccinated person contracts the disease in a given week is $1 \%$, what is that probability for a vaccinated person if the vaccine is $80 \%$ effective?

Answer: $1 \% \times(1-80 \%)=1 \% \times(20 \%)=0.2 \%$. Of course, if enough people get vaccinated for some "herd immunity" to occur, that contraction probability per week for the unvaccinated will then be below $1 \%$, with corresponding decreases for the vaccinated.

## ** Ch. 13. More on "Numerical thinking about the effectiveness of infectious disease vaccination" (Spreading of disease, Section 13.6)

Explain why if the testing procedure described in the previous problem is rigorously followed, testing an effective vaccine takes longer than a less effective one.

Answer: It will take longer to reach the 150 cases because fewer of the inoculated will get sick.
** Ch. 13. Even more on "Numerical thinking about the effectiveness of infectious disease vaccination" (Spreading of disease, Section 13.6)
If the testing procedure described in the previous two problems is rigorously followed and if testing a perfect vaccine takes 12 weeks, how much longer would it take to fully test a vaccine that turns out to be $80 \%$ effective vaccine than one that is $50 \%$ effective?

Answer: For a 100\% effective vaccine, only those taking the placebo get ill. Whatever the contraction rate for the placebo group, we can say that after 12 weeks there will be 12 weeks of placebo cases. For a $50 \%$ effective vaccine, there are twice as many placebo cases as vaccinated cases. So, after 8 weeks, there will be 8 weeks worth of placebo cases and half as many vaccinated cases, for a total of 12 weeks of placebo cases. For an $80 \%$ effective vaccine, there are five times as many placebo cases as vaccinated cases. So, after 10 weeks, there will be 10 weeks of unvaccinated cases and one fifth as many vaccinated cases, for a total of 12 weeks of unvaccinated cases. So, it would take two weeks longer to test an $80 \%$ effective vaccine than one that is $50 \%$ effective. (This type of problem can be done using algebra. The numbers just work out simply in this problem, so you can do well by just guessing then.)

## ** Ch. 13. Arithmetic vs. geometric means

Is the arithmetic or geometric mean of two positive numbers the larger, or does it depend on what the number actually are? (Try different sets of positive number to test this.)

Answer: The arithmetic mean is always larger-unless the numbers are the same, when the two means would be equal to each other. For example, for 10 and 40 the arithmetic mean is $(10+$ $40) / 2=50 / 2=25$, while the geometric mean is $(10 \times 40)^{1 / 2}=400^{1 / 2}=20$. (This can also be shown by using algebra.)

## ** Ch. 13. Arithmetic vs. geometric vs. harmonic means

The arithmetic mean (or, simple the mean) of $n$ numbers is their sum divided by $n$. The geometric mean is the $n$th root of their product (so the square root for 2 numbers, the cube root for 3 numbers and so on). The harmonic mean is $n$ divided by the sum of their reciprocals (and so equals the reciprocal of the arithmetic mean of the reciprocals). Each weights the numbers-or data-differently, and so conveys different information. What are the mean, geometric, and harmonic means of 10 and 40 ?

Answer: They are respectively $(10+40) / 2=50 / 2=25$; sqrt $(10 \times 40)=\operatorname{sqrt}(10 \times 10 \times 4)=$ sqrt $(10 \times 10 \times 2 \times 2)=10 \times 2=20$; and $2 /(1 / 10+1 / 40)=2 /(4 / 40+1 / 40)=2 /(5 / 40)=2 /(1 / 8)$ $=2 \times 8=16$. Section 13.2.2 describes how arithmetic and geometric means are used to present the rates of return on investments. The later better reflects "compounding. If you calculate the average speed during a round-trip in which there are different speeds (ratios of distances and times) in one direction and the return trip (and so the average is the total distance traveled by the total time), you are calculating the harmonic mean speed of the trip. (Weighted) Harmonic means are used in finance to assess other ratios, such as price-to-earnings ratios.

## Chapter 14. Untangling The Worlds of Probability and Statistics

## ** Ch. 14. Probability vs. statistics

All roads are repaired annually in a town, sequentially from district to district, and first in districts with the largest fraction of roads needing repair, then all those in the district with the next largest fraction, and so on---because it is assumed that this pattern of repair needs is likely to re-occur and that this plan helps the town most. This cycle is set at the beginning of the year on the basis of data of the fraction of roads needing repair each year, averaged over the previous two years. Say there are two districts A and B. In the previous year, 20 of the 100 roads in A needed repair and 40 needed repair in the year before that. Also, in the previous year 50 of the 200 roads in B needed repair and 60 of them needed repair in the year before that. (a) Are these data (i.e., the actual numbers and fractions) probabilities or statistics? (b) Are the repair decisions made on the basis of probabilities or statistics? (c) In the coming year, are the roads in A going to be repaired before those in B or will those in B before those in A? Why?

Answer: (a) The data are statistics. (b) The decisions are made on the basis of projected probabilities (that are derived from the statistics). (c) In A, these fractions are 20/100 $=20 \%$ and $40 / 100=40 \%$, which average to $30 \%$. In B, they are $50 / 200=25 \%$ and $60 / 200=30 \%$, which average to $27.5 \%$, and so the repairs are done first in district $A$.

## * Ch. 14. Probability predictions for sporting events (Probability, Chapter 15)

In the June 16, 2021 NBA (National Basketball Association) playoff game between the Philadelphia 76ers and the Atlanta Hawks, with both teams having won 2 games in that series, the Hawks won "even though" (before the game started) models had a win probability for the 76 ers of $99.7 \%$ and for the Hawks of $0.3 \%$. (a) Is this an example of probability or statistics? (b) What is your evaluation of this model. ("Their Win Probability Was $0.3 \%$. The Hawks Won.: Walls Street Journal, By Ben Cohen, June 17, 2021, 4:00 am ET, https://www.wsj.com/articles/atlanta-hawks-comeback-76ers-nba-playoffs-11623906923)

Answer: (a) This is an example of an assessment of probabilities using models, which may well have included statistics for input. (b) As reported (!), the model prediction is absurd (as so is the model). It says that if they played this game 300 times, the Hawks would be expected to win only one time, and this is ridiculous. This is particularly so, since the Hawks had won 2 of the previous 4 games in the series and had suffered no major injuries to key players. Models give results, but are they valid models, with valid input data?

## Chapter 15. The Math of What Might Be: What are the Odds?-The World of Probability

* Ch. 15. Will you be selected to serve on a jury? (Section 15.1, Basics of Probability) A total of 480 people, including you, are summoned to court for jury duty obligations for a week, and they all comply. During this time, a total of 4 juries are empaneled. Say each jury has 12 members (ignore alternates) selected after voir dire (questioning) of 36 of the jurors summoned that week (who were not already selected to serve on an already chosen jury). (a) If each selection step is random, what is the provability that you will sit on a jury that week? (b) Is each step in fact truly random?

Answer: (a) All that is important is that of those 480 summoned potential jurors $12 \times 4=48$ of them will serve, so this probability is $48 / 480=10 \%$. (b) Each selection step, other than voir dire, is random. Voir dire depends on the questions posed and the personal responses of the potential juror and is not random.

## * Ch. 15. How do you win playing the "numbers?" (Section 15.1.1, Probability of single events)

In the illegal "numbers" games or racket, people would bet on the three-digit number for that day and receive a payoff of typically $\$ 600$ for a $\$ 1$ bet if they had the correct number for that day. (The winning number for that day could be the last three numbers of the total amount bet on a local racetrack that day.) (a) What are the odds of winning? (b) If you regularly "played the numbers," what would your average winnings or losings be?

Answer: (a) The odds of winning are 1 out of 1000, or 999:1 against you. (b) For each \$1,000 you bet, on average you would win once, or $\$ 600$. So, on average, you lose $40 \%$ of what you bet. The attraction was the hope of hitting the jackpot.
https://en.wikipedia.org/wiki/Numbers_game

## * Ch. 15. Finding "lucky" 4-leaf clovers (Section 15.1)

The probability of finding a four-leaf clover among three-leaf ones is $1 / 5,000$. (a) On average, how many clovers would you need to examine to find 6 four-leaf ones? (b) If a lawn has 2.0 clover plants per square inch uniformly distributed on it, how large of an area in square feet would you need to examine to find on average 6 of them?
https://en.wikipedia.org/wiki/Four-leaf_clover (8/12/21)

Answer: (a) To find 1, on average you would need to inspect 5,000 of them, and so $6 \times 5,000$ or 30,000 clover plants need to be examined to find on the average 6. (b) This would be 30,000/2 $=$ 15,000 square inches. There are 12 inches in a foot and so $12 \times 12=144$ square inches in a square foot, so you would need to examine $15,000 / 144=104.1$ square feet, which is a square with length that is the square root of 104.1, or 10.2 feet.
** Ch. 15. Math thinking about probabilities in families with two children (Section 15.1, Basics of Probability; Sec. 15.5 Some Probabilities Depend of Information)
(a) If you have two children and both are not boys, what is the probability that both are girls?
(b) If you have two children and the older one is a boy, what is the probability that both are boys?
(Assume that other things being equal, the probability that a given child is a boy or girl is a half.)
Answer: (a) If you have two children the probability the both are boys is $1 / 4$ (which is $1 / 2 x$ x $1 / 2$ ) (boy/boy, for 1 out of 4 ordered pairs) and that both are girls is also $1 / 4$ (girl/girl, for 1 out of 4). The probability that one is boy and one is a girl is $1 / 2$ (boy/girl and girl/boy, for 2 out of 4). If you know that both are not boys, the probability that both are girls for these remaining three choices is 1 out of $3=1 / 3$. (Also, see the Let's Make a Deal example on page 204 in Sec. 15.5.) (b) if you know the older child is a boy, the younger child could be a boy or a girl, each with probability $1 / 2$, and so the probability that both are boys is $1 / 2$.

From
The Moscow Puzzles, Boris A. Kordemsky \#236, pg. 102

## *** Ch. 15. Why are Friends listed in a particular order? (Section 4.10, Choosing numbers; Sec. 15.1, Basics of Probability)

Friends was a very popular TV sitcom from 1994-2004 that is still very popular in reruns and streaming. It had an ensemble cast of six actors, three males and three females. They contributed equally to the show and were not stars until they appeared in this show. In the opening credits, the names of each of the six actors were presented in six sequential shots showing their character along with their name. I only occasionally watched this show and when I recently watched a rerun I wondered why the three female characters were presented first and then the three male ones. If the three female actors were more well-known when this show started this would be expected, but they were not, and if the order were chosen randomly this order would be quite unlikely I thought. (But why would this be expected to be random?)
(a) In how many ways could the six actors be presented in the credits, if they were chosen randomly? What is the probability that any one of these possibilities would be chosen, if this were indeed random?
(b) In how many ways could the actors be presented in the credits, if they were chosen randomly except for the restriction that the three female actors would presented before the male ones? What is the probability that any one of these restricted random choices would be chosen, if this were random? What is the probability that one of these random sequences in (a) would list all the females before the males, if all were random?
(c) Let's say the actors were presented alphabetically (according to their last names, each of which are different). Would this be considered random? Should probability concept be applied to this? What is the probability that a possibility chosen at random in (a) would be alphabetical? (It turns out the actors were apparently presented alphabetically, presumably to be fair, which I ascertained after watching a second rerun.)
(d) In a related, but quite different question, how many possibilities are there in (a) if the order of the female actors and then the male actors did not matter?

Answer: (a) There are $6!=6 \times 5 \times 4 \times 3 \times 2 \times 1=720$ such possibilities (permutations). If all were random, the odds of any one of them being chosen would be $1 / 720 \sim 0.001389=0.1389 \%$ $\sim 0.14 \%$.
(b) There would be 3! $\times 3!$ (three choices for the first female, then 2 for the second and 1 for the third one, and then three choices for the first male, then 2 for the second and 1 for the third one) $=(3 \times 2 \times 1) \times(3 \times 2 \times 1)=36$ possibilities. If all were random, the odds of any one of them being chosen would be $1 / 36 \sim 0.02777=2.777 \% \sim 2.8 \%$. The probability that one of these random sequences in (a) would list all the females before the males, if all were random, is $36 / 720=0.05=5.0 \%$.
(c) If the list in the credits is alphabetical, as it appears to be, it is not random but fully determined. Probability concepts should not be applied to this. The probability that one of the random possibilities chosen in (a) happens to be alphabetical is $1 / 720 \sim 0.001389=0.1389 \% \sim$ $0.14 \%$. I think that the first three in the alphabetical list of the actors were female was by chance.
(d) This is the number of combinations of six items with three of them being identical and the three others being identical to each other, so $6!/(3!\times 3!)=720 /(6 \times 6)=20$.

## ** Ch. 15. "Zoombombing" (Probability, Section 15.1)

A common pass code to access a video conferencing line has 6 digits chosen at random. (a) How many possibilities are there? (b) What is the probability that you can guess the correct one? (c) If you can type one such possibility every 2 seconds, what is the probability you can type a correct code and gain access at some time during an hour-long meeting? (For the video conferencing line Zoom, gaining such improper access is called "Zoombombing.")

Answer: (a) $10 \times 10 \times 10 \times 10 \times 10 \times 10=10^{6}=1,000,000=1$ million. (b) $1 / 10^{6}=10^{-6}=1$ out of a million. (c) An hour has 60 seconds/hour $\times 60$ minutes/hour $=3,600$ seconds, so you would have 3,600/2 $=1,800$ attempts, each with a probability of success of $1 / 1,000,000$ for a total success probability of $1,800 / 1,000,000=1.8 / 1,000=0.18 \%$. This is also $\sim 1 / 555.6$, so the odds of success are roughly 1 out of 560 . (This assumes that your repetition rate is exactly 2 seconds.)
*** Ch. 15. Probability and ranking, and entering the playoffs with "play-in" games, with equal winning probabilities (Probability, Section 15.1; Ranking, Section 18.4.1)
In the National Basketball Association (NBA) one way that teams are selected to play in the post-season playoffs is to first rank the teams in a conference from 1 to 8 , with descending order of record (winning percentage). In the first-round series, the teams with the better records play the teams with the poorer records (with the team ranked - and so seeded - 1 playing that ranked 8 , and 2 vs. 7,3 vs. 6 , and 4 vs. 5), which purposely gives an advantage to the teams seeded higher, and with most of the games being played at the home of the higher-seed team, which again purposely gives an advantage to the teams seeded higher. This was changed for the playoffs in 2021 , with the teams ranked $7,8,9$, and 10 needing to play each other to earn the right for 2 of them to be seeded 7 and 8 , in a set of 3 "play-in" games (in which these teams try to "play" into the playoffs). The $7^{\text {th }}$-ranked plays at home against the $8^{\text {th }}$-ranked team, and the winner is seeded number 7 in the playoffs. The loser of this game plays at home against the
winner of a game between the $9^{\text {th }}$ - and $10^{\text {th }}$-ranked teams, that is played at the home of the former team, and the winner is seeded number 8. (a) If in each of these 3 games each team has a probability of winning of $50 \%$, independent of their relative records (and prior records against each other and so on), where they play, and so on, what is the probability that each of these teams will make it the first-round of the playoffs? (b) Is this fair?

Answer: (a) Both the $7^{\text {th }}$ - and $8^{\text {th }}$-ranked teams have a $50 \%$ of winning their game and then being seeded $7^{\text {th }}$ in the playoffs. The other team, which had a $50 \%$ chance of losing that game, has a $50 \%$ change of beating the winner of the 9-10 matchup. The probability of the second event is not correlated to that of the first, so that team's probability of entering the playoffs as the $8^{\text {th }}$ seed is $50 \% \times 50 \%$ or $25 \%$. So, there is a $50 \%+25 \%=75 \%$ chance that the $7^{\text {th }}$-ranked will make the playoffs, and the same is true for the $8^{\text {th }}$-ranked team. The $9^{\text {th }}$ - and $10^{\text {th }}$-ranked teams have a $50 \%$ probability of winning their game and then a $50 \%$ chance of winning its game with the 7-8 loser. Because these are not correlated, the probability of the $9^{\text {th }}$-ranked team going to the playoffs is $50 \% \times 50 \%=25 \%$, and the same is true for the $10^{\text {th }}$-ranked team. These 4 probabilities add up to $200 \%(=75 \%+75 \%+25 \%+25 \%)$, which is fine since 2 of these 4 teams will definitely make the playoffs. (b) Fairness is in the eye of the beholder. If the goal is to give the two higher-ranked teams an advantage, this system does it, even when the probability of winning any given game is taken to be 50\%. These probabilities do not reflect the advantage the home team has; this is addressed in the next problem.

## *** Ch. 15. Probability and ranking, and entering the playoffs with "play-in" games, with better winning probabilities when playing at home (Probability, Section 15.1; Ranking, Section 18.4.1)

Repeat the previous problem if the home team, which is the higher-ranked team each time, has a $60 \%$ chance of winning.

Answer: (a) The wining probabilities are now skewed to the home team, which in each case is the team that had been ranked higher. The $7^{\text {th }}$ and $8^{\text {th }}$-ranked team respectively have a $60 \%$ and $40 \%$ of winning their game and then being seeded $7^{\text {th }}$ in the playoffs. The other team, which had either a $40 \%$ (team 7) or $60 \%$ (team 8) chance of losing that first game, has a $60 \%$ change of beating the winner of the 9-10 matchup. The probability of the second event is not correlated to that of the first, so that probability of the $7^{\text {th }}$-ranked entering the playoffs as the $8^{\text {th }}$ seed is $40 \% \times$ $60 \%$ or $24 \%$. So, there is a $60 \%+24 \%=84 \%$ chance that the $7^{\text {th }}$-ranked will make the playoffs. The probability of the $8^{\text {th }}$-ranked team of entering the playoffs as the $8^{\text {th }}$ seed is $60 \% \times 60 \%$ or $36 \%$. So, there is a $40 \%+36 \%=76 \%$ chance that the $8^{\text {th }}$-ranked will make the playoffs. The probability that the $9^{\text {th }}$-ranked team will win its first game is $60 \%$ and then win its second game is $40 \%$, and so the probability of it entering the playoffs is $60 \% \times 40 \%=24 \%$, since the probabilities are not correlated. The probability that the $10^{\text {th }}$-ranked team will win its first game is $40 \%$ and then its second game is $40 \%$, and the probability of it entering the playoffs is $40 \% \times$ $40 \%=16 \%$, again since the probabilities are not correlated. These 4 probabilities add up to $200 \%(=84 \%+76 \%+24 \%+16 \%)$, which is fine since 2 of these 4 teams will definitely make the playoffs. (b) Again, fairness is in the eye of the beholder. If the goal is to give the two higherranked teams an advantage by the 3-game format, this format still does it. Moreover, the chosen probabilities now recognize the impact of the home-court advantage that is given to the higher-
ranked team in each play-in game. The system and analysis now give the $7^{\text {th }}$-ranked team an advantage over the $8^{\text {th }}$-ranked one, and the $9^{\text {th }}$-ranked team an advantage over the $10^{\text {th }}$-ranked one-and also show the $7^{\text {th }}$ - and $8^{\text {th }}$-ranked teams have an even larger advantage in making the playoffs.

## ** Ch. 15. Probability of losing your keys in two ways (Random, but with restrictions, Section 15.1.3)

There is 10.0 per cent probability that you will lose your keys due to one reason and a 5.0 per cent probability of losing them due to a second separate and independent reason. What is the probability you will have your keys?

Answer: There is $90.0 \%$ probability ( $=100.0 \%-10.0 \%$ ) of keeping your keys due to the first reason and a $95.0 \%$ probability ( $=100.0 \%-5.0 \%$ ) of keeping your keys due to the second reason. Because they are independent reasons, the probability of keeping your keys is $90.0 \% \times$ $95.0 \%=0.900 \times 0.950=0.855=85.5 \%$ (and not $100.0 \%-10.0 \%-5.0 \%=85.0 \%$.) This kind of reasoning leads to bigger difference when the probabilities of losing the keys increase.

## ** Ch. 15. Probability of guessing the promotion code (Basics, Section 15.1)

You want to use weekly code for a TV promotion, but you don't have it and want to guess it. You know it is the current date followed by an upper case letter and then a number, both of which you can assume are randomly chosen. What is the probability you would get it correct if you tried 4 times? (You could get locked out if you are wrong after a certain number of trials.)

Answer: The number of possibilities for the letter and number are $26 \times 10=260$, so the probability of getting it correct in your 4 tries is $4 / 260=\sim 0.0154=1.54 \%$. This assumes you have used the correct current date.

## ** Ch. 15. Estimating the probability that the Dow will be the same two consecutive days (Basics, Section 15.1)

(a) Estimate the likelihood that the Dow Jones Average closes at the same (integral) number for any two consecutive business days. Say, the daily changes in this metric uniformly span between $-0.75 \%$ and $+0.75 \%$ and the current value is 30,000 . (b) Now make this estimation for the Dow Jones Average closing at the same value for any two consecutive business days, when it is expressed to the nearest 0.01 .
[https://www.forbes.com/sites/mikepatton/2016/01/29/fast-facts-on-the-dow-jones-stockindex/amp/]

Answer: (a) This means that there are roughly $1.5 \% \times 30,000+1$ or $\sim 451$ (integral) possible changes that are equally likely, including no change, so the probability that it will be the same on any given two consecutive days is $1 / 451=0.22 \%$. (b) This would be $10 \times$ smaller, or $0.0022 \%$.

# ** Ch. 15. Probability determination limited by uncertain binning of information (Basics Section 15.1, Binning Sections 4.8, 15.5, 16.2.2, and 16.3) 

A man is born in 1920 and dies in 1995, but you know nothing about when these events occurred during the calendar year except as given below.
(a) Show that his "age" at the time of his death might be 74 or 75.
(b) What is the probability that his age at death is 74 or that it is 75 ?
(c) If instead you know that he was born in the second half of 1920 and died in the first half of 1995, what are the probabilities he died when he was 74 or 75 ?
(d) If you know he was born in the first half of 1920, what are the probabilities he died when he was 74 or 75 ?
(e) If instead you know that he died in the first quarter of 1995, what are the probabilities he died when he was 74 or 75 ?
(Assume each month has the same 30 days (and so no leap years) and assume that, unless you are told otherwise, birth can occur on any day in the year with equal probability, with the same being true for death. Ignore the possibility that birth and death occurred on the same calendar day.)

Answer: (a) If he died later in the year than when he was born, he would have died while 75 -if before then 74. (b) $1 / 2$ for each since it would be equally probable that he died before or after his birth in the calendar year. (c) He definitely dies before his birthday, so he dies at 74 with a probability 1 and at 75 with a probability 0. (d) $3 / 4$. (Do the math by going through the steps). If he died in the second half, he was 75 when he died. The chance of him dying in the first half is $1 / 2$ and it was $50 \%$ likely then that his birthday date preceded his death day, so him dying in the first half after his birth day has overall probability of $1 / 2 \times 1 / 2=1 / 4$. So, the overall probability of him dying at 75 is $1 / 4+1 / 2=3 / 4$, and at 74 is $1-3 / 4=1 / 4$. (e) During the latter $3 / 4$ of the year he would die while 74. Using the reasoning in (c), the probability of him dying at 74 in the first quarter is $1 / 4 \times 1 / 2=1 / 8$. So, the odds of him dying at 74 is $1 / 8+3 / 4=7 / 8$, and at 75 it is $1 / 8$.

The assumption that death can occur equally likely on any day is subject to season-related reasons. The ignoring births on deaths on the same day, as happened for actress Ingrid Bergman, women's rights activist Betty Friedan, and - tradition has it - William Shakespeare, was a simplification. But, in fact deaths on one's birthday are higher than expected, with someone finding it being $6.7 \%$ higher overall than expected and $25.4 \%$ higher for $20-29$ years old, perhaps due to birthday celebrations and other causes-but these deviations are not certain.
https://en.m.wikipedia.org/wiki/Birthday_effect
https://www.washingtonpost.com/news/wonk/wp/2014/12/29/youre-more-likely-to-die-on-yourbirthday/
You're more likely to die on your birthday Jason Millman
December 29, 2014 at 1:38 PM EST

# ** Ch. 15. Probability of contracting an infectious disease based on risk analysis 

 Probability, Poisson Statistics, Infectious Disease, Risk (Spreading of disease Section 13.6, Basics Section 15.1)During the time an epidemic where a virus is spread by breathing water droplets from someone else's breath, you are told you that the probability of you contracting the virus is $2 \%$ when you are in a particular environment for 1 minute. What is it in 10 minutes and in an hour?

Answer: In 10 minutes it would be $10 \times 2 \%=20 \%$, which is not negligible at all. In an hour, if you assume linearity, it would be $60 \times 2 \%=120 \%>100 \%$, which is nonsense. Though one could do more refined math, it is still easy to say that it is seems to be likely.

## ** Ch. 15. More on "Probability of contracting an infectious disease based on risk analysis Probability, Poisson Statistics, Infectious Disease, Risk" (Spreading of disease Section 13.6, Basics Section 15.1)

During the time an epidemic where a virus is spread by breathing in water droplets from someone else's breath, an infected person who is not wearing a face mask unexpectedly walks very near you. Because of this encounter let's say you now have a $0.5 \%$ probability of contracting the virus. The same person now jogs by you, spending one half as much time near you, but breathing six times faster. What is the probability you will now contract the virus?

Answer: With half as much time of exposure and 6 times the dose, assuming linearity the probability of infection could be estimated to be $0.5 \times 6 \times 0.5 \%=1.5 \%$.

## ** Ch. 15. Even more on "Probability of contracting an infectious disease based on risk analysis Probability, Poisson Statistics, Infectious Disease, Risk" (Spreading of disease Section 13.6, Basics Section 15.1)

At an event you are exposed to an airborne virus that leads to you to have a $5 \%$ probability of contracting an infectious disease. Your partner is with you at the time and has the same probability of contracting it. If because of contact with your partner you have a $100 \%$ of catching the disease from your partner, what is the probability you will become ill?

Answer: There is a 5\% probability you will become ill due to your exposure and since if your partner becomes ill so will you, which is an overall $5 \% \times 100 \%=5 \%$ probability, the probability of you becoming ill is $5 \%+5 \%=10 \%$. Because these probabilities are treated as independent events and are $\ll 1$, you can simply add them to get the estimate.

[^1]Answer: Because these probabilities are not $\ll 100 \%$, you cannot simply add them, as in the previous problem. (If you did, you would get $140 \%$, which is $>100 \%$ and so nonsense. Think of the birthday problem in Section 15.1.3, in which the probabilities of not having the same birthday are evaluated and then multiplied. The probability of you not getting sick is $100 \%$ $70 \%=30 \%$, and that of you not getting sick because of contact with your partner is $30 \%$. So, the probability of you becoming ill is $0.3 \times 0.3=0.09=9 \%$, and the probability of you becoming ill is $100 \%-9 \%=91 \%$.

## ** Ch. 15. Probability, health risk (Ch. 20) (Spreading of disease, Section 13.6; basics, Section 15.1)

At an event you are exposed to an airborne virus that leads to you to have a $70 \%$ probability of contracting an infectious disease. Your partner is with you at the time and has the same probability of contracting it. If because of contact with your partner you have a $50 \%$ of catching the disease from your partner, what is the probability you will become ill?

Answer: Now the probability of you becoming ill from your partner is $70 \% \times 50 \%=35 \%$ and of not becoming ill is $100 \%-35 \%=65 \%$. So, the probability of you becoming ill is $0.3 \times 0.65=$ $0.195=19.5 \%$, and the probability of you becoming ill is $100 \%-19.5 \%=80.5 \% \sim 80 \%$.
** Ch. 15. What do you win when making a winning bet in horse racing? (Section 15.1.1) You can bet on a horse to win, place or show. You win the "win" bet only if your horse comes in $1^{\text {st }}$, the "place" bet if it finishes either $1^{\text {st }}$ or $2^{\text {nd }}$, and the "show" bet if it comes in either $1^{\text {st }}, 2^{\text {nd }}$ or $3^{\text {rd }}$. For example, in a ten-horse race, say Horse 6 wins (comes in $1^{\text {st }}$ ). The announced results could be that it pays out for $\$ 4.80$ for a win bet, $\$ 3.60$ for a place bet, and $\$ 2.40$ for a show bet, each for a $\$ 2$ bet. Horse 2 "places" (comes in $2^{\text {nd }}$ ) and pays out $\$ 9.00$ for a place bet and $\$ 5.40$ for a show bet. Horse 9 "shows" (comes in $3^{\text {rd }}$ ) and pays out $\$ 9.80$ for a show bet. What were the betting odds set for Horse 6 to win the race? (These are the betting odds you see before the race starts, and they depend on how much money was bet on each horse and the cut taken by the track---as we will see below.)

Answer: If $\$ 4.40$ is returned to you for the win, you have netted $\$ 4.80-\$ 2.00=\$ 2.80$ for the $\$ 2$ bet, so the betting odds to win had been set at $\$ 2.80 / \$ 2.00=1.4$ or 1.4 to 1 , which is also expressed as 7 to 5 (because $7 / 5=1.4$ ). This is working "backwards." Working "forwards" 7 to 5 (or 1.40 ) odds return $1.4 \times \$ 2=\$ 2.80$, so you receive $\$ 2.80+\$ 2.00=\$ 4.80$ for the $\$ 2$ winning bet.
(This problem and the following two problems and the one below on the "pari-mutuel" betting system for horse racing are based on information and examples from https://www.liveabout.com/how-to-calculate-betting-odds-and-payoffs-1879161 https://wizardraceandsports.com/win-place-and-show-betting-explained/ https://horseworlddata.com/pmtrcks.html
https://www.thesportsgeek.com/sports-betting/horse-racing/win-place-show/ .)
** Ch. 15. What do you receive when making a win, place or show bet? (Section 15.1.1)
Using the information in the previous problem:
(a) How much do you leave the track with if you bet $\$ 30$ for Horse 6 to (i) win. Repeat for a bet to (ii) place; and for a bet to (iii) show?
(b) How much do you leave the track with if you bet $\$ 30$ for Horse 9 to (i) win. Repeat for a bet to (ii) place; and for a bet to (iii) show?
(c) How much do you leave the track with if you bet $\$ 30$ for Horse 3 to (i) win. Repeat for a bet to (ii) place; and for a bet to (iii) show?

Answer: For each of these $\$ 30$ bets the payoff 15 times that for a $\$ 2$ bet, so (including the initial bet)
(a) (i) $15 \times \$ 4.80=\$ 72.00$, (ii) $15 \times \$ 3.60=\$ 54.00$, (iii) $15 \times \$ 2.40=\$ 36.00$;
(b) (i) and (ii) $15 \times \$ 0=\$ 0$ (you lose), (iii) $15 \times \$ 9.80=\$ 147.00$;
(c) (i), (ii), and (iii) $15 \times \$ 0=\$ 0$ (you lose).

## ** Ch. 15. How does "across the board" horse race betting affect you chances of winning and winning big? (Section 15.1.1)

Using the information in the previous two problems:
(a) How much do you leave the track with if you bet $\$ 30$ in total, with $\$ 10$ separate bets for Horse 6 to win, place, and show? (This is an "across the board" bet.)
(b) How much do you leave the track with if you bet $\$ 30$ in total, with $\$ 10$ separate bets for Horse 9 to win, place, and show?
(c) How much do you leave the track with if you bet $\$ 30$ in total, with $\$ 10$ separate bets for Horse 6 to win, place, and show?
(d) How does across-the-board betting affect you chances of winning and winning big?

Answer: These $\$ 30$ bets are $\$ 10$ to win, place and show, so the payoff is five times that for a $\$ 2$ bet for each of the three, so (including the initial bet)
(a) $(5 \times \$ 4.80=\$ 24.00)+(5 \times \$ 3.60=\$ 18.00)+(5 \times \$ 2.40=\$ 12.00)=\$ 54.00$;
(b) $(5 \times \$ 0=\$ 0)+(5 \times \$ 0=\$ 0)+(5 \times \$ 9.80=\$ 49.00)=\$ 49.00$;
(c) $(5 \times \$ 0=\$ 0)+(5 \times \$ 0=\$ 0)+(5 \times \$ 0=\$ 0)=\$ 0)$.
(d) Your chance of winning something generally increases, but the payouts are generally smaller.
*** Ch. 15. Betting on a horse race that is fixed so you will always win (Section 15.1.1) Follow the math in the upcoming example to show that for the right circumstances you can always win a horse racing bet: if there is a big favorite, not many horses in the field, you bet to win (and not to place or show), and you are absolutely certain that the big favorite will lose (because of some nefarious activity). Suppose the odds to win on the 5 horses in the field are 5/4, $4 / 1,6 / 112 / 1$, and $60 / 1$. Show that you can always win $\$ 60$ with a bet of $\$ 31$ (assuming a minimum bet of $\$ 1$ ) if you do not bet on the $5 / 4$ favorite, bet $\$ 15$ on the horse running with $4 / 1$ odds, bet $\$ 10$ on the horse with $6 / 1$ odds, $\$ 5$ on the $12 / 1$ horse, and $\$ 1$ on the $60 / 1$ horse. What are your net winnings in each case?

Answer: Your total bet is $\$ 15+\$ 10+\$ 5+\$ 1=\$ 31$. A winning bet $\$ 15$ on the horse running with $4 / 1$ odds gets you $4 \times \$ 15=\$ 60$, a winning bet $\$ 10$ on the $6 / 1$ horse gives you $6 \times \$ 10=$ $\$ 60$, a $\$ 5$ winning bet on the $12 / 1$ horse gets you $12 \times \$ 5=\$ 60$, and $a$ winning $\$ 1$ bet on the $60 / 1$ horse gives you $60 \times \$ 1=\$ 60$, so all winning bets provide you winnings of $\$ 60$ for a $\$ 31$ bet. Of course, you do not lose that part of the total $\$ 31$ you bet on the winning horse, so your losing bets are respectively $\$ 16(=\$ 10+\$ 5+\$ 1), \$ 21(=\$ 15+\$ 5+\$ 1), \$ 26(=\$ 15+\$ 10+$ $\$ 1)$, and $\$ 30(=\$ 15+\$ 10+\$ 5)$, and so you respectively would net $\$ 60-\$ 16=\$ 44, \$ 60-\$ 21$ $=\$ 39, \$ 60-\$ 26=\$ 34$, and $\$ 60-\$ 30=\$ 30$. However, you lose if the favorite in fact wins. You lose, if other bettors catch on to this and odds go down before you can bet at these favorable odds. And, you will likely lose if you are involved with this nefarious activity.
[Silent Witness season 8 episode 5. Nowhere Fast. Part 1. 9/19 2004. Nikki Alexander]
** Ch. 15. How are horse racing odds and payouts determined? (Section 15.1.1)
In the dominant "pari-mutuel" betting system for horse racing at a race track, the track first takes a cut out of each betting pool, such as for those bet for a given horse to win, to pay for track operations, payouts to the owners, trainers, and jockeys of the horses that do best (the "purse"), and so on. This is typically roughly $15-20 \%$ depending on the state in the U.S. The odds are set so the remaining amount for that particular betting pool is divided among those with winning bets for that pool, less their own bets. Say $\$ 10,000$ is bet for the horses to win. If the takeout is $18 \%, \$ 10,000-\$ 1,800=\$ 8,200$ is divided among those with tickets for the winning horse, after subtracting the total amount bet on the winning horse to win. If $\$ 1,500$ was bet on the winning horse, $\$ 8,200-\$ 1,500=\$ 6,700$ is divided proportionately among those betting the $\$ 1,500$, so $\$ 6,700 / \$ 1,500 \sim 4.47$ is net for each dollar bet. So, $\$ 4.47$ is returned for each $\$ 1$ bet. These 4.47 to 1 betting odds are rounded off to 4.5 to 1 , or, equivalently, $9 / 2$. (These are not the odds that a given horse will win!) The return on a $\$ 2$ bet, including the original bet is $\$ 4.47 \times 2+\$ 2=$ $\$ 10.94$ and it is announced that that winning horse paid out $\$ 10.94$ for a bet to win. (The algorithm for place and show odds have a few more steps than that for our example of the win pool. Online betting would also be included in these analyses.) What would the odds and payouts be here if the horse that won instead had (a) $\$ 4,000$ or (b) $\$ 400$ bet on it?

Answer: For the example given: $\$ 8,200-\$ 1,500=\$ 6,700$, for $\$ 6,700 / \$ 1,500 \sim 4.47$ to 1 betting odds and a $\$ 4.47 \times 2+\$ 2=\$ 10.94$ payout for a $\$ 2$ bet, including the returned $\$ 2$. For the two new cases:
(a) $\$ 8,200-\$ 4,000=\$ 4,200, \$ 4,200 / \$ 4,000 \sim 1.05$ to 1 betting odds (which is almost 1 to 1 or "even" betting odds) ("favorites" often have odds of 1 (or somewhat smaller or larger than 1) to 1. There is a $\$ 1.05 \times 2+\$ 2=\$ 4.10$ payout, which is relatively small.
(b) $\$ 8,200-\$ 400=\$ 7,800, \$ 7,800 / \$ 400 \sim 19.5$ to 1 (or $\sim 20$ to 1 ) betting odds (which are "long" odds, for a "long shot"). There is a $\$ 19.5 \times 2+\$ 2=\$ 41.00$ payout, which is relatively large.
(See the references given above for "What do you win when making a winning bet in horse racing?".)
** Ch. 15. Do those who bet on the horses usually win on average? (Section 15.1.1) You bet many times on the horses. On average will you come out even or will you win or lose money, and if so, by how much? (Use the information from the previous problem.)

Answer: On average, you will lose the takeout for that track, which is often between $15 \%$ and 20\%.
** Ch. 15. Using the Pareto Principle to set horse racing betting odds (Section 15.1.1) One practical way of establishing odds in a horse race adapts the Pareto Principle (Chapter 2), with $80 \%$ of the winning probability assigned to the top $20 \%$ of the horses, and $20 \%$ to the probability to the remaining $80 \%$ of the field. Let's say there are 3 horses in this top category. If these top three were co-favorites, what would the betting odds on each of them be?

Answer: The estimated probability that any of them wins would be $80 / 3 \%=26.67 \%$. With 3:1 odds the probability of winning are estimated to be $1 /(1+3)=25 \%$, so each would have approximately $3: 1$ odds. $5: 2$ odds would mean a $1 /(1+2.5)=28.6 \%$ winning probability. So, both are approximate (and sufficiently good) answers, and the real answer is in between them. Algebra shows the exact odds are 2.75:1 (or 5.5:2 or 11:4), which is not a standard betting line. https://www.usracing.com/news/horse-betting-101/making-fair-odds-line

## * Ch. 15. Using metrics to control epidemics (Spreading of disease Section 13.6)

During times of infectious disease crisis, public health decisions are often made on the basis of the metric of total daily number of new cases in a location, say a state here called State A, with a rolling daily average over the last specific number of days, say 7 days here, per unit of population, say 100,000 people. What would this metric be if there were a rolling daily average of 80 new cases in this time period in a state of a million people?

Answer: $80 /(1,000,000 / 100,000)=80 /(10.0)=8.0$ (There are 10 units of the 100,000 population.)

## ** Ch. 15. Sharp thresholds of metrics used to control epidemics (Spreading of disease, Section 13.6)

Hard thresholds are often set on the basis of a metric, with total action or consequences for metric values above the threshold value and zero for those below, or vice versa. (This can be expressed in terms of step functions (also called Heaviside functions), which have a value 1 above a threshold value and 0 below it.) But, such abrupt cutoffs often do not make sense. State B will not permit travel into it from another state if that state has infection disease metric (as in the previous problem) that is 10.0 (average new daily cases per 100,000 people) or higher (updated weekly).
(a) Will State B allow travel into it from State C, a state of a million people, which reported 99 new daily cases over the past 7 days?
(b) Will State B allow travel into it from State D, a state of a million people, which reported 101 new daily cases over the past 7 days?

Answer: (a) $9.9<10.0$, so yes. (b) $10.1<10.0$, so no.

## ** Ch. 15. Issues with thresholds due to binning of metrics used to control epidemics (Spreading of disease Section 13.6)

The western half of State E has a million people, and reported 400 new daily cases averaged over the past 7 days. The eastern half of State E also has a million people, and reported 20 new daily cases over the past 7 days. What is infectious disease metric for each half (separately) and for the entire state? Will State B (in the previous Problem) allow travel into it from State E? If the eastern half of State E has a common border with State B, does this seem reasonable and fair? ... and if also State B itself has a metric of 8.0 (average new daily cases per 100,000 people) does this seem even less fair?

Answer: These two halves have a metric of $400 / 10=40$ and $20 / 10=2$, respectively. The whole state has a metric of $420 / 20=21.21>10$, so State $B$ will not allow travel into it from State. Rules based in this coarse binning, especially when the contact region has a low value of the metric (eastern half of State E, 2---with a common border with State B), which may be much lower than value of the metric in the state prohibiting this travel (State B, 8.0).

## ** Ch. 15. Probabilities when rolling one or two four-sided die (Section 15.1)

Some basic probability concepts were introduced in Chapter 15 using a tossed coin, which bas two sides-heads and tails, and the die, which has six sides. For both, the result of tossing it is the top face. Dice are cubes with 6 square sides, each with a different number of dots, $1,2,3,4$, 5 , and 6 dots, and a clear way describing the die value---the number of dots on the top face.
Nothing distinguishes the four faces on the side and it is simpler to assess the top face rather than the bottom face. Imagine 4 -sided die, which are tetrahedra, each face being an equilateral triangle (with three equal sides and angles). If the 4 sides have $1,2,3$, and 4 dots, and you "roll it" the face that is down could (in principle) be used to give the value of this die (since nothing distinguishes any of the three side faces). Some of the questions we asked about the probability of the results of tossing one or more the 6 -sided dice, can be posed for such 4 -sided dice.
(a) Wheat is the probability that when you roll this (unbiased) 4-sided die, you will get a 3 ? (or any other given number, of course)
(b) What values can you get if you roll two of these 4 -sided dice, and what is the probability of getting each?

Answer: (a) $1 / 4=25 \%$ for this and for 1,2 , and 4 too. (b) You can get a total from $2=1+1=2$ (and you can get this in only one way) to $8=4+4$ (one way). Also, you can get a 3 from $1+2$ and $2+1$ (two ways); a 4 from $1+3,2+2$, and $3+1$ (three ways); a 5 from $1+4,2+3,3+$ 2, and $4+1$ (four ways); a 6 from $2+4,3+3$, and $4+2$ (three ways); and a 7 from $3+4$ and $4+3$ (two ways). This gives a total of $1+2+3+4+3+2+1$ (going from the total for 1 to 8 ), which equals $16=4 \times 4$, as expected. So, the probability of rolling a 2 is $1 / 16$, a 3 is $2 / 16=1 / 8$, a 4 is $3 / 16$, a 5 is $4 / 16=1 / 4$, a 6 is $3 / 16$, a 7 is $2 / 16=1 / 8$, and an 8 is $1 / 16$, with $1 / 16=2 / 16+$ $3 / 16+4 / 16+3 / 16+2 / 16+1 / 16=16 / 16$, as expected. (See the results for the usual 6 -sided die in the book.)
** Ch. 15. Probabilities and betting odds when rolling two four-sided die (Section 15.1) The "house" tells you that you win when you roll these two 4 -sided dice if you get any of the four numbers: $2,6,7$ and 8 , but lose only if you roll any of the fewer, three, numbers: 3,4 , and 5. For even money odds (so, if you bet $\$ 1$ and lose you lose your bet and if you win you receive win $\$ 1$ and also get your $\$ 1$ bet returned if you win). Is this in your favor?

Answer: Of course not. The probability of you winning any roll of two such dice 1/16 + 3/16 + $2 / 16+1 / 16=7 / 16$ and of you losing is $2 / 16+3 / 16+4 / 16=9 / 16$, and so the probability of losing a given roll is $9 / 16-7 / 16=2 / 16=1 / 8=12.5 \%$. If you bet $\$ 1$, your net per roll will be $7 / 16 \times \$ 1+9 / 16 \times \$ 0=\$ 0.125$ and $87.5 \%$ is returned per bet. On average, you will lose $\$ 1$ after 8 rolls. This is a bad bet. (See the results for the usual 6 -sided die in the book.)
** Ch. 15. Improving the betting odds when rolling two four-sided die (Section 15.1)
The "house" now tells you they will give even odds for all rolls of these two 4 -sided dice as in the previous problem, except now on Monday you get a $\$ 2$ payout for every $\$ 1$ bet for a roll of 2, and similarly on Tuesday for a roll of 7 . Are these good deals?

Answer: On Monday, for every $\$$ bet on average you will win $\$ 2 \times 1 / 16+\$ 1 \times(3 / 16+2 / 16+$ $1 / 16)=\$ 8 / 16$ and lose $\$ 1 \times(2 / 16+3 / 16+4 / 16)=\$ 9 / 16$, for a net loss of $\$ 1 / 16=\$ 0.0625$ or $6.25 \%$ per bet. It is a better bet, but not great - and on average you will lose $\$ 1$ after 16 rolls get $93.75 \%$ back per bet-which is still worse than the $94.7 \%$ returned in American Roulette. On Tuesdays, you win $\$ 2 \times 2 / 16+\$ 1 \times(1 / 16+3 / 16+1 / 16)=\$ 9 / 16$ and lose $\$ 1 \times(2 / 16+3 / 16+$ $4 / 16)=\$ 9 / 16$, so on average you will not win or lose. The house would not want to do this for long because it will not make money, except to entice new customers---and except for those times when you are on a losing streak and will lose all your money. (See the results for the usual 6sided die in the book.)

## ** Ch. 15. Left with the Beatles (Probability, binomial distribution, Section 15.1.2; combinations, Section 4.10)

When you see pictures of the Beatles performing, two of them (John Lennon and George Harrison are holding their guitars in one direction, while Paul McCartney is holding his in the opposite direction. Why? Because it gives better visual balance? It does, but the real reason is that the former two are right-handed and the latter is left-handed. In fact, two of the four Beatles are left-handed (Paul and Ringo Starr, who for totally unrelated reasons are the two surviving members, as of June 2021). Since $90 \%$ of all people are righties, what is the probability that 2 of 4 randomly chosen people are lefties?

Answer: The probability of being a righty is 0.9 and a being a lefty is 0.1. In choosing two of each the probability is $0.9 \times 0.9 \times 0.1 \times 0.1=0.0081=0.81 \%$. The number of ways you can choose combinations of 2 righties and 2 lefties out of 4 is 6 . You can choose 4 items in $4 \times 3 \times 2$
$\times 1=24$ ways, say of 2 righties and 2 lefties. The order of choosing the righties does not matter, so this is divided by 2, and by another 2 for lefties-so $24 /(2 \times 2)=6$ (or $4!/(2!\times 2!))$. (Also, see Figure 4.6.) Therefore, the probability of choosing two righties and two lefties is $6 \times 0.81 \%$ ~ $4.86 \%$ or 1 in $\sim 20.6$. So, the Beatles were rare not only because of their extraordinary talent but because they had so many left-handed people.
(Using similar reasoning from Section 4.10 and its footnotes, the probability of having 0, 1, 2, 3, and 4 lefties are $65.61 \%, 29.16 \%, 4.86 \%, 0.36 \%$, and $0.01 \%$, respectively, which sum to 1.0 )

## *** Ch. 15. How many "lucky" 4-leaf clovers will you find? (Poisson statistics, Section 15.1.2)

The probability of finding a four-leaf clover among three-leaf ones is $1 / 5,000$. (a) You examine 20,000 clover plants. On average, how many four-leaf ones will you find? (b) Use Poisson statistics to find the probability that you would find either $0,1,2,3,4,5,6,7$, or 8 of them?

Answer: Poisson statistics applies when the probability per event is $\ll 1$, as it is here (1/5,000). As seen in the footnotes of the subsection in the book, the probability for $r$ events occurring if the average number is s is $s^{r} e^{-s} / r!$. That average number here is $20,000 / 5,000=4$. So, the probability of 0 events (of finding a four-leaf clover) occurring is 0.0183 (rounded off), for 1 it is 0.0733 , for 2 it is 0.1465 , for 3 it is 0.1954 , for 4 it is 0.1954 , for 5 it is 0.1563 , for 6 it is 0.1042 , for 7 it is 0.0595 , for 8 it is 0.0298 , ... This monotonically gets smaller for finding more four-leaf clovers, such as 0.0132 for 9 of them-but it is not 0 and so the sum of these probabilities for finding 0 to 8 four-leaf clovers is less than $100 \%$. We see that he probability of finding exactly the average number (4) is about $20 \%$.

## * Ch. 15. Opening locks without key replacement (Probability with restrictions, Section

15.1.3) There are 10 different, but seemingly the same looking, keys in a bag, of which one can unlock a box. You choose one and try it. If it doesn't work you put that key aside and choose another key in the bag and try it, and so one. After how many tries do you have at least a $50 \%$ probability of having unlocked the box?

Answer: You have a 50\% chance of having chosen the correct key with the fifth choice.

## ** Ch. 15. Opening locks with key replacement (Probability with restrictions, Section 15.1.3)

There are 10 different, but seemingly the same looking, keys in a bag, of which one can unlock a box. You choose one and try it. If it doesn't work you put that key back into the bag and choose a key in the bag and try it, and so on. After how many tries do you have at least a $50 \%$ probability of having unlocked the box?

Answer: Now each time you try you have a $10 \%$ chance of opening it. After the first trial there is a $90 \%$ chance the box is not open. After the $6^{\text {th }}$ try there is a $(0.9)^{6}=53.1 \%$ chance it remains unopened and after the $7^{\text {th }}$ try a $(0.9)^{7}=47.8 \%$ one, and so a $52.2 \%$ chance you have opened it.

## ** Ch. 15. When do the outcomes of tennis matches really count? (math joke) (Section

 15.2.1, Standard deviations; Section 16.2.3, Statistically significant)World-famous mathematician Stan Ulam relayed a whimsical comment that physics Nobel Laureate Enrico Fermi would made if he lost a tennis match, say, a match with one set: 6-4 (6 games to 4). "It does not count because the difference is less than the square root of the sum of the games." Why is this a joke?

Answer: Of course, the score counts because it is deterministic, and not statistical. (In tennis sets, the first to win 6 games by no less than two games wins; so 6-4 would win.) If these were statistics, one could say there is a standard deviation of sqrt $10 \sim 3.16$, and 4 and 6 are both within the average (5) plus or minus this amount. But statistics do not apply, and this is meant to be a joke.
(as quoted in The Adventures of a Mathematician, Stanislaw Ulam, pg. 164)

## ** Ch. 15. False positives and negatives in general for one test (Section 15.6)

Medical test A for a given disease has a false negative rate of $19 \%$ and a false positive rate of $11 \%$, while test B has a false negative rate of $7 \%$ and a false positive rate of $2 \%$. Comparing 1,000 people who take test A and the same number taking test B (a) if all taking the test do not have the disease, how many more are falsely told they have the disease in the group taking test A than those taking test B and (b) if all taking the test instead have the disease, how many more are falsely told they do not have the disease in the group taking test A than those taking test B?

Answer: (a) In group $A, 11 \% \times 1,000=110$ are falsely told they are ill, while in group $B, 2 \% \times$ $1,000=20$ are falsely told they are ill, both through false positives, so 90 more (or $(11 \%-2 \%)$ $\times 1,000$ ). (b) In group $A, 19 \% \times 1,000=190$ are falsely told they are not ill, while in group $B$, $7 \% \times 1,000=70$ are falsely told they are not ill, both through false negatives, so 120 more (or $(19 \%-7 \%) \times 1,000)$. Of course, test B is much better.
*** Ch. 15. False positives and negatives in a given population (Section 15.6)
Consider the previous problem with 1,000 people who are given either test A or B , but now with $5 \%$ of the people being ill. For this population base, how many are falsely told they are ill or well, for each test, and how much worse are the results for one test than the other?

Answer: In this population base there are 950 well people and 50 ill people. For those taking test $A, 950 \times 89 \%=845.5$ of the well are correctly told they are well (the true positive rate is $100 \%-11 \%$ (the false positive rate) $=89 \%$ ) and $50 \times 19 \%=9.5$ of the ill people are falsely told they are well (the false negative rate is $19 \%$ ), so $845.5+9.5=855$ are told they are well ( $9.5 / 855=1.11 \%$ of them incorrectly). Also, $50 \times 81 \%=40.5$ of the ill are correctly told they are ill (because the true negative rate is $100 \%-19 \%$ (false negative rate) $=81 \%$ ) and $950 \times$ $11 \%=104.5$ of the well are falsely told they are ill (false positive rate), so $40.5+104.5=145$ are told they are ill (104.5/145 $=72.1 \%$ of them incorrectly). For a perfect test, 950 would have
been correctly told they are well and 50 correctly told they are ill. (Note that 855 (who told they are well) +145 (who are told they are ill) $=1,000$, as expected.)

For those taking test B, $950 \times 98 \%=931$ of the well are correctly told they are well (the true positive rate is $100 \%-2 \%$ (false positive rate) $=98 \%$ ) and $50 \times 7 \%=3.5$ of the ill people are falsely told they are well (the false negative rate is $7 \%$ ), so $931+3.5=934.5$ are told they are well $(3.5 / 934.5=0.37 \%$ of them incorrectly). Also, $50 \times 93 \%=46.5$ of the ill are correctly told they are ill (the true negative rate is $100 \%-7 \%$ (false negative rate) $=93 \%$ ) and $950 \times 2 \%=$ 19 of the well are falsely told they are ill (false positive rate), so $46.5+19=\underline{65.5}$ are told they are ill $(19 / 65.6=29.0 \%$ of them incorrectly). (Note that 934.5 (who told they are well) +65.5 (who are told they are ill) $=1,000$, as expected.) Of course, test $B$ is much better in testing any group with ill and well people. More ill and well people are identified as such, bringing better treatment to the ill and more peace-of-mind to the well.

## ** Ch. 15. A long, but still random walk (Section 15.7.1)

You take a million steps of length 1 foot in random directions. In miles, estimate how far you have moved in total and how far relative to your starting position.

Answer: You have traveled a total of a million feet. A mile is 5,280 feet, which we will round off as 5,000 feet, so you have traveled a total of $\sim 10^{6} /\left(5 \times 10^{3}\right)=200$ miles. After taking one million steps in random directions you have translated on average by the square root of a million, or a thousand, steps, so your final position is on average only roughly 1,000 feet or $\sim 0.2$ miles from your initial position (a factor of 1,000 shorter than your quite exhausting travel of 200 miles).

## ** Ch. 15. Actuary exam probability analysis, of different sets of medical readings that are independent of each other (Probability, Section 15.1; Risk assessment for insurance, Section 20.2)

Actuaries calculate risk for insurers to make sure that their policies generate net positive income (and so no overall losses), on the basis of data. Prospective actuaries take exams that require math thinking, a bit of knowledge of probability, and, in some, relatively simple algebra. This question and the several subsequent ones are samples from those exams (reworded from those provided in The Wall Street Journal article about this, https://www.wsj.com/articles/actuary-credential-test-exam-bad-odds-11640706082). Each requires math and logical thinking in evaluating the various possibilities and applying probability concepts; only the final two sample problems in this series require a bit of algebra

One modified sample question is: In a study in which people are randomly sampled, $14 \%$ have (what is categorized as) high blood pressure, $22 \%$ have low blood pressure, and the rest have normal blood pressure; and $15 \%$ have an irregular heartbeat and the rest have a regular heartbeat. If the level of blood pressure and the regularity of the heartbeat are not correlated (which is not the case in the actual source sample problem, see the next problem), what is the probability that someone has a regular heartbeat and low blood pressure? (https://www.wsj.com/articles/actuary-credential-test-exam-bad-odds-11640706082)

Answer: 0.19. The probability of having normal blood pressure is $1-0.14-0.22=1-0.36=$ 0.64 and of having a regular heartbeat is $1-0.15=0.85$. The probability that someone has a regular heartbeat and low blood pressure is the product of having a regular heartbeat, 0.85, and low blood pressure, 0.22, which is $\underline{\mathbf{0 . 1 9}}$.

## *** Ch. 15. More actuary exam probability analysis, but of different sets of medical readings that are correlated with each other (Probability, Section 15.1; Statistics, Chapter 16; Risk assessment for insurance, Section 20.2)

Repeat the previous problem and find the probability that someone has a regular heartbeat and low blood pressure, but instead of assuming that the blood pressure level and type of heartbeat are uncorrelated, (now, as in the source problem in (https://www.wsj.com/articles/actuary-credential-test-exam-bad-odds-11640706082)) you are told that $1 / 3$ of those with an irregular heart beat have high blood pressure and $1 / 8$ of those with normal blood pressure have an irregular heartbeat.

Answer: 0.20. There are 6 categories describing the 3 blood pressure levels and 2 types of heartbeat regularity, because $3 \times 2=6$. Let's use the supplied information to methodically find the probability fraction in each category. The fraction of people with high, low and normal blood pressure are respectively $0.14,0.22$, and $1.0-0.14-0.22=0.64$. The fraction of people with regular and irregular heartbeats are $1.0-0.15=0.85$ and 0.15 . Because $1 / 8$ of those with normal blood pressure have an irregular heartbeat, $7 / 8$ of them have a regular heartbeat, and so $7 / 8 \times 0.64=\mathbf{0 . 5 6}$ of all of those sampled have normal blood pressure and a regular heartbeat and $1 / 8 \times 0.64=0.08$ of them have normal blood pressure and an irregular heartbeat. Because $1 / 3$ of those with an irregular heartbeat have high blood pressure, $0.15 / 3=\mathbf{0 . 0 5}$ of all sampled have high blood pressure and an irregular hearbeat, and the remainder of those with high blood pressure have a regular heartbeat, or $0.14-0.05=0.09$ of all sampled. Because 0.85 of all sampled have a regular heartbeat, and we just saw that 0.09 of all sampled have a regular heartbeat and high blood pressure and that 0.56 of all sampled have a regular heartbeat and normal blood pressure, we see that $0.85-0.09-0.56=\underline{\mathbf{0 . 2 0}}$ of all those sampled have a regular heartbeat and low blood pressure (which is the answer we seek). Because 0.15 have an irregular heartbeat, and 0.05 have this high and 0.08 have this and normal blood pressure, $0.15-0.05-$ $0.08=\mathbf{0 . 0 2}$ have an irregular heartbeat and low blood pressure (which is also 0.22 (the fraction with low blood pressure) - 0.20 (the fraction with low blood pressure and a regular heartbeat.)).

## ** Ch. 15. More actuary exam probability analysis, binning categories (Probability, Section 15.1; Statistics, Chapter 16; Risk assessment for insurance, Section 20.2; Binning Sections $4.8,15.5,16.2 .2$, and 16.3) <br> Could how the data used for binning the categories affect the analysis in the previous problem?

Answer: Yes. It might easier to set up distinct categories for the data and this may be a good first step, but this might lead to imperfect analysis. Better analysis could occur by using distributions of actual blood pressures. Perhaps having the two distinct categories of heart beat regularity is fine, or perhaps finer binning into different types of heart irregularities would be more helpful.
*** Ch. 15. More actuary exam probability analysis, with independent events and a bit of algebra (Probability, Section 15.1; Risk assessment for insurance, Section 20.2)
This is another sample actuary test problem, but one that requires a bit of algebra. You pick 1 ball from an urn with 4 red and 6 blue balls (balls that are otherwise identical) and 1 from a second urn with 16 red balls and an unknown number of blue balls. The probability that you choose two balls of the same color is 0.44 . What is the number of blue balls in the second urn? (https://www.wsj.com/articles/actuary-credential-test-exam-bad-odds-11640706082)

Answer: 4. There is no correlation between a given ball from one urn and another ball in the second one. The probability that you choose a red ball from the first urn is 4/10, and it is $6 / 10$ for a blue ball. If there are $x$ blue balls in the second urn, then the probability that you choose a red ball from it is $16 /(16+x)$, and it is $x /(16+x)$ for a blue ball. The probability of choosing two red balls is the product $4 / 10 \times 16 /(16+x)$ and that of choosing two blue balls is the product $6 / 10 \times x /(16+x)$. The probability of choosing two balls of the same color is their sum: 4/10 $\times$ $16 /(16+x)+6 / 10 \times x /(16+x)=(64+6 x) /(10)(16+x)$. Setting this equal to 0.44 and solving gives $x=$ 4. $\{$ The solution method is: $(64+6 x) /(10)(16+x)=0.44$; multiplying both sides by 10 and then by $16+x$ give $(64+6 x) /(16+x)=4.4$ and $64+6 x=4.4(16+x)=70.4+4.4 x$; subtracting $4.4 x$ from the left and 64 from right sides gives $1.6 x=6.4$, and then dividing both sides by 1.6, gives $x=4$.\}

## *** Ch. 15. More actuary exam probability analysis, of independent events, requiring some algebra (Probability, Section 15.1; Statistics, Chapter 16; Risk assessment for insurance, Section 20.2)

Another sample actuary exam question concerns the insurance preference of automobile owner. (source https://www.wsj.com/articles/actuary-credential-test-exam-bad-odds-11640706082); this requires a bit of algebra. The actuary learns that (1) an automobile owner is twice as likely to purchase collision coverage as disability coverage, (2) an owner purchasing collision coverage is independent of that owner purchasing disability coverage; and (3) the probability of someone purchasing both types of coverage is 0.15 . What is the probability that an automobile owner purchases neither collision nor disability coverage?

Answer: 0.33. Call the probability of purchasing collision coverage $x$, so the probability of not buying it is $1-x$. Call the probability of purchasing disability coverage $y$, so the probability of not buying it is $1-y$. Because such purchases are independent of each other (point 2), the probability of purchasing both is $x y=0.15$ (point 3). We want to know the probability of purchasing neither, which is $(1-x)(1-y)$ [and this is not equal to $1-x y]$. Point 1 tells us that $x$ $=2 y$, so $x y=(2 y) y=2 y^{2}=0.15$, and therefore $y^{2}=0.075$ and $y$ (which must be positive) $\sim$ 0.274. So, $(1-x)(1-y)=(1-2 y)(1-y) \sim(1-2(0.274))(1-0.274)=(1-0.548)(1-0.274)=$ $0.452 \times 0.726 \sim 0.328$, and therefore the probability of someone not buying either coverage is $\sim$ 0.33 .
** Ch. 15. Who is related to whom, and by how much? (Section 15.6, DNA matching) DNA tests are becoming ever more important in tracing family roots, treating disease, and locating suspected criminals. Overlap in the common DNA, as is now characterized by the unit centimorgan (or cM ), decreases by a factor of two (on average) each time you go up or down the family tree, to a parent or child. This analysis often, but not always, refers only to your autosomal DNA, which is that from your 22 non-sex DNA pairs (and not the $23^{\text {rd }} \mathrm{XX}$ or XY pair), and it depends on the exact details of the DNA analysis. (We choose one method in this question.) Each of us is characterized as having $6,800 \mathrm{cM}$, so you share with each of your parents and children roughly $3,400 \mathrm{cM}$. This is an average and there is a sizable statistical range of what we actually happen to inherit about any such average (for any given method). We share $1 / 2$ of this measured common DNA with full siblings (in one method; it is $3 / 8$ in another common method, which we will not use).

Roughly, how may cMs does a person share on average with each biological (a) grandparent and great-grandparent, and grandchild and great-grandchild, and (b) aunt or uncle, first cousin, and first cousin once removed (the child of a first cousin)?

Answer: You have 6,800 cM, and share 1/2 with each parent and $1 / 2$ with each child, and so you share with each of your parents and children roughly $6,800 c M \times 1 / 2=3,400 c M$ (on average), as noted. (a) Each of your parents shares $1 / 2$ with each of their parents-your grandparents, and each of your children shares $1 / 2$ with each of their children-your grandchildren, and so $1 / 2 \times 1 / 2$ $=1 / 4$ with you, or $6,800 c M \times 1 / 4=1,700 c M$. For each of your great-grandparents and greatgrandchildren there is another factor of $1 / 2$, so the fraction is $1 / 8$ and you share with them 6,800 $c M \times 1 / 8=850 c M$. (b) Each aunt or uncle has $1 / 2$ of their DNA in common with each of their siblings-including your parent here, so $1 / 4$ or $1,700 \mathrm{cM}$. They share $1 / 2$ with each of their children-your first cousins, and so $1 / 8$ or 850 cM is in common with you, and they in turn share 1/2 with their children, your first cousins once remove have in common with you on average 1/16 or $6,800 \mathrm{cM} \times 1 / 16=425 \mathrm{cM}$.

Again, these are averages and the dispersion (spread) in possible values about these averages (due to biology and not to DNA analysis) can lead to uncertainties in identification.
(https://isogg.org/wiki/Autosomal_DNA_statistics)
(https://www.wsj.com/articles/the-obscure-math-exposing-our-genetic-secrets-11653039002)
(https://thegeneticgenealogist.com/)
(https://doi.org/10.1371/journal.pone.0034267)
(International Society of Genetic Genealogy)
(Ancestry.com)

## Chapter 16. The Math of What Was: Statistics-The Good, The Bad, and The Evil

** Ch. 16. Zipf's law and the the the the the ... (Statistics, Chapter 16; Ranking, Chapter 18)

In a book, there is one word that is used more often than any other one, a word that is used the second most, one the third most, and so on. Is there a pattern that describes how the frequency of each word in this sequence varies, aside from that it (of course) decreases? Zipf's law says that in a book or a set of books or other documents, there is an approximate connection between how frequently the same word is used; this is true for all languages. The second most-used word occurs $\sim 1 / 2$ as often as the most frequent word (which is usually the word "the"), the third mostused word occurs $\sim 1 / 3$ as often as the most frequent word, and so on. Therefore, the $\mathrm{n}^{\text {th }}$ most frequent word occurs $\sim 1 / \mathrm{n}$ as often as the most common word. This is not a true or exact law, but it seems to be approximately true, and it is also subject to statistical variations. (See www.youtube.com/watch?v=fCn8zs912OE, https://en.wikipedia.org/wiki/Zipf\'s_law )

In the book we encountered analogous statistics for numbers (Benford's law, Section 16.4.5), the frequency of letters and coding information (Section 7.3, 12.2.5);, and in ranking (Ranking, Chapter 18). We encountered numerical rules of thumb, such as the $80 / 20$ (or Pareto) rule of thumb (Chapter 2).

One illustration of Zipf's law is that in the slightly more than 1 million words of the Brown Corpus of American English text, the word "the" is the most frequently occurring word, occurring 69,971 times and accounting for nearly $7 \%$ of all words; the second-place word "of" occurs 36,411 times and the third-ranked "and" occurs 28,852 times. The most frequent 135 different words account for half of all its words. Do the frequencies of the second and third mostfrequent words seem to follow Zipfs law?
(Fagan, Stephen; Gençay, Ramazan (2010), "An introduction to textual econometrics", in Ullah, Aman; Giles, David E. A. (eds.), Handbook of Empirical Economics and Finance, CRC Press, pp. 133-153, ISBN 9781420070361. P. 139); "Foundations of Statistical Natural Language Processing" Chris Manning and Hinrich Schütze, Foundations of Statistical Natural Language Processing, MIT Press. Cambridge, MA: May 1999.)

Answer: 36,411/69,971~0.520~1/1.92, which is $\sim 1 / 2.28,852 / 69,971 \sim 0.412 \sim 1 / 2.43$, which is not perfectly close to $1 / 3$, but the overall trend for all rankings still follows Zipf's law fairly well.

## ** Ch. 16. Follow up on Zipf's law and the the the the the ... (Statistics, Chapter 16; Ranking, Chapter 18)

(a) What would be true about the product of the frequency of a word in a book and its frequency rank, if Zipf's law were exact, and without fluctuations. (b) Test this for some words in Mark Twain's "Tom Sawyer: the (frequency $=3,332$, rank $=1$ ); and (frequency $=2,972$, rank $=2$ ); a
(frequency $=1,775$, rank $=3$ ); he (frequency $=877$, rank $=10$ ); but (frequency $=410$, rank $=$ 20); one (frequency $=172$, rank $=50$ ); two (frequency $=104$, rank $=100$ ); turned (frequency $=$ 51 , rank $=200$ ); comes (frequency $=16$, rank $=500$ ); begin (frequency $=9$, rank $=900$ ); brushed (frequency $=4$, rank $=2,000$ ).
"Foundations of Statistical Natural Language Processing"
Chris Manning and Hinrich Schütze, Foundations of Statistical Natural Language Processing, MIT Press. Cambridge, MA: May 1999.

Answer: (a) This product would equal a constant, because the the frequency would vary exactly inversely with the rank. (b) These example products are respectively: 3,332; 5,944; 5,235; 8,770; 8,200; 8,600; 10,400; 10,200; 8,000; 8,100; 8,000. (Several less-frequent words, share the same ranking, which is not illustrated here.) These products are typically $\sim 8,000-10,000$, with the largest deviations occurring for the most frequently used words.

## ** Ch. 16. More follow up on Zipf's law and the the the the the ... (Statistics, Chapter 16; Ranking, Chapter 18)

(a) Test Zipf's law, as described in the previous problems, yourself. Pick a random page in a book and count how often each word is used. Rank them in terms of usage. Divide each count by that of the most popular word. Express the reciprocal of that result as a fraction with 1 in the numerator and compare the denominator of that final result with the word frequency rank. (b) If you repeated this for only a single paragraph or for the entire book how would the results differ?

Answer: (a) and (b) The more the data, the better, and it will look more like Zipf's law, with fewer fluctuations and deviations. There should still be some noticeable deviations for a page, many more for a paragraph and still some, but far fewer of them, for an entire book.
** Ch. 16. Is this statistical sanity or insanity? (Getting data for statistics, Section 16.1) A 2022 study claimed that 250,000 people die in the U.S. annually because they have been misdiagnosed in an Emergency Room. This result was obtained from a study of the 503 patients in two Canadian emergency room during a part of 2004, in which the death of one man was attributed to misdiagnosis. The cited 250,000 figure was obtained by multiplying this this death rate of $0.2 \%(=1 / 503)$ by the annual 130 million ER visits annually in the U.S. (It also included statistical range.) Is this statistical analysis reasonable?

Answer: The statistical sample is far too small to merit such a wild extrapolation. This is a very unreasonable use of data.
[A Study Sounds a False Alarm About America's Emergency Rooms
https://www.wsj.com/articles/false-alarm-about-emergency-rooms-ahrq-physicians-er-
misdiagnoses-mortality-rate-us-canada-trust-11672136943]
** Ch. 16. Does the government think you live in an urban or rural area? (Analyzing statistics, Section 16.2; Metrics, 16.4.2; Probability, Section 15.6; Binning, Sections 15.5 and 16.3)

Using the long-time U. S. Census definition of an urban area as one with at least 2,500 people, with the rest being termed rural areas, the fraction of Americans living in rural areas decreased from $19.3 \%$ in 2010 to $18.7 \%$ in 2020, so more people officially lived in urban areas. However, between the 2010 and 2020 censuses the U. S. Census Bureau changed its definition of an urban area from one with at least 2,500 people to one with at least 5,000 . With this change, 4.2 million people who would have had been classified as living in urban areas in 2020 were reclassified as living in rural areas. Estimate the new reported fraction of those living in rural areas in 2020. Explain why letting all know this new policy.

Answer: With roughly 300 million people living in the U.S., this reclassification meant that 4.2 million $/ 300$ million or $1.4 \%$ more people were reclassified as living in rural areas, so $\sim 18.7 \%$ $+1.3 \%=20.0 \%$ were said to be officially in rural areas in 2020. (This estimate leads to the correct answer.) Without knowing this change, you would think that $20.0 \%-19.3 \%=0.7 \%$ more Americans lived in lower population areas in 2020 than in 2010, which is not true.
[And Just Like That, America Becomes More Rural
https://www.wsj.com/articles/and-just-like-that-america-becomes-more-rural-11672963347]
Binning is very important in using and assessing statistics correctly. Sometimes cases when a parameter is lower than a threshold are assigned to a Category $A$ and to Category $B$ when it is above it. Sometimes, there are few cases near the threshold, so the binning is clear. When the changes are gradual and there are always cases near whatever threshold is chosen, this becomes problematic, as in using the level of PSA (prostate serum antigen) to assess the occurrence of prostate cancer (in Section 15.6), it is more problematic. This problem of whether the government calls an area urban or rural is a less serious case of uncertain an uncertain or gradual threshold. It is still one of some consequence as it can be linked to levels of government funding.

## ** Ch. 16. Who shoots free throws well in the NBA? (Section 16.2.1, Covariance and correlation)

Do you expect that taller or shorter NBA (National Basketball Association) players shoot free throws with higher probability? What would this mean for the correlation coefficient of this percentage with player height?

Answer: On average, it is found that shorter players shoot free throws better than taller players, so this correlation coefficient is negative: a greater height, on average, means a smaller percentage of successful free throws. In fact, it is -0.3789 . Much of this is explained by shorter players playing positions in which shooting free throws well is more valued (or playing positions in which poor free throwing shooting is less acceptable). (https://dylan-sivori.github.io/2021-02-05-nba-ft-percent/)

* Ch. 16. Distortions in public reports by selective presentation of data (Section 16.3) A car manufacturer announces that its sale of blue cars increased from one year to the next by $30 \%$ from 1.00 to 1.30 million (or by 300,000 ) and that of red cars by $40 \%$ from 100,000 to 140,000 (or by 40,000 ). In the press report of this, only the $40 \%$ increase in red car sales is noted, perhaps because it has the larger percentage increase in sales and that has become the focus of the news story. Explain why this is a distorted presentation of the data and why the overall story is not presented in proper context.

Answer: This seems to be an attempt to over-represent and over-dramatize the increase in the sales of red cars and to downplay that of blue cars. Playing up only the technically-true, larger percentage increase in red cars-though narrowly true-ignores the much larger absolute increase in the sale of blue cars, as well as the total number of them being sold (and the still hefty percentage increase in the sale of blue cars). Whatever the motivation for this, this selective distortion of data misinforms and misleads the public. All the key representative data should have been presented, in a balanced, albeit brief, way and in context. (When this happens in real life, car statistics are rarely involved.)

## ** Ch. 16. Presenting daily data the best way (Section 16.3)

New infectious diseases cases can be plotted as daily numbers, weekly averages, or as running averages of say 7 days. How are they different and what are their relative advantages?

Answer: Plotting daily rates shows the data as they come in, so it could be considered to be the best type of display, but large daily fluctuations can make them difficult to understand. The fluctuations can be to variations in reliable daily reporting. These can be smoothed out by displaying the average over the most recent $x$ days. If $x$ is too small, the fluctuations could still dominate and if $x$ is too large, and longer than the time after exposure and contracting the disease that symptoms appear---very roughly 7 days for Covid-19---the real trends will not be apparent. A weekly average will show this averaging too but a 7 (or so)-day rolling average displayed daily is often a good compromise.
** Ch. 16. Plotting non-deceptively-with the same units on axes (Section 16.3) Data can be (and in many times have been) presented in plots in misleading and even-more purposely deceptive means. For example, consider the data set of new daily cases of an infectious disease per 100,000 people in two types of regions-one with no added health regulations and one in which such regulations were added on Day 11. (Real data have many more fluctuations that those presented in this simplified data set. This is based on a real data set. See below.)

Areas with no regulations: From
Day 1 to Day 10: constant at 5
Day 11 to 20: linear increase from 5 to 10
Day 21 to 50: constant at 10 .
Areas with regulations: From

Day 1 to Day 10: linear increase from 5 to 10
Day 11 to 20: linear increase from 10 to 20
Day 21 to 34: constant at 20.
Day 35 to 50 : linear decrease from 20 to 15 .
(a) In words, how would you describe each set and how they are different?
(b) Sketch the progression separately for these two areas on the same graph, with cases per day ranging from 0 to 30 (on the vertical $y$ axis, labeled with the is given range) and time from Day 1 to 30 (on the horizontal $x$ axis, with the Day number range shown). What does this way of plotting the data tell you?
(c) Now sketch these data again, EXCEPT now only from Day 35 to 50. Also, now plot the data for the areas with no regulations with the vertical scale ranging from 5 to 15 , BUT plot the data for the areas with regulations with the SAME vertical range BUT with a scale that ranges from 15 to 25 . (You could label the scale for the first one on the right and the second one on the left.) Why is this way of plotting deceptive? What are you being led to think when you see a graph plotted like this?

Answer: (a) On Day 11, the number of cases in areas with no regulations, increases until Day 20, and then it remains steady. On Day 11, the number of cases in areas with regulations, increase even more rapidly until Day 20, and then it slowly decreases, but this number is always larger each day than that in the areas with no regulations. That is all you can conclude from the data. Without more data and information you cannot conclude anything about how the regulations affect the data. (However, by deceptive plotting these data you can fool people.)
(b) It leads you to description in (a).
(c) It is designed to make you think that the number of cases may have decreased with the added regulations, which may be true, but it is exceedingly misleading. Though properly labelled, using different vertical ranges is designed to make you think that there are fewer cases with the regulations than without them; this is never true! By excluding the time period before Day 35, it is making you think that there are no differences in the two types of areas, other than the existence of regulations.
[https://www.wsj.com/articles/kansas-democrats-covid-chart-masks-the-truth11598483406 ?mod=opinion_lead_pos10
Kansas Democrats' Covid Chart Masks the Truth: The state's health secretary fudged the data to make the governor's mask mandate look successful.]

## **Ch. 16. Life expectancies (Section 16.3)

What are the life expectancies for people born on a given day in a population if (a) everyone born that day dies at age 65?, (b) half of them die at age 35 and the other half at age 95?, and (c) $10 \%$ of them die at age 20 and $90 \%$ die at age 70? (Use weighted averaged of lifetimes. (Section 5.2.)

Answer: (a) 65. (b) The average is $(35+95) / 2=130 / 2=65$. (c). The weighted average is $(0.1 \times$ $20+0.9 \times 70)=2+63=65$. So, this case they are all the same, 65 .
** Ch. 16. Are you in the lifespan tail? (Distribution shapes and tails, Section 15.2.2;
Statistics, Actuarial Tables, Section 16.3)
What does it mean to be in the tail of the lifespan distribution?
Answer: In the tails of a distribution, the value, number or fraction of the function decreases rapidly with further increasing or decreasing parameter. In the lifespan distribution, the number or fraction who have reached a given older age decrease with increasing age.
** Ch. 16. Do the actuarial tables of mortality lead you to believe that you will live forever? (Distribution shapes and tails, Section 15.2.2; Statistics, Actuarial Tables, Section 16.3) A man in the U.S. hears that the actuarial tables say that the probability that an average man will live (at least) another 24 hours and wake up the next day is more than $99.9 \%$ for almost every age, and figures that because this is always true, he will live forever since the probability of him dying on any given day is so small. Is he right?

Answer: No. The statistics he has heard are accurate for men only until they are very old, but they really indicate that the average man will not live forever and moreover these are averages and do not tell us the likely fate of any individual.
** Ch. 16. When is it no longer true that it is more than $99.9 \%$ probable that a man will not see another day? (Distribution shapes and tails, Section 15.2.2; Statistics, Actuarial Tables, Section 16.3)
How old does a man need to become before it is no longer true it is more than $99.9 \%$ probable that he will not see another day?

Answer: If it is $99.9 \%$ likely you will see another day, the probability you will die within 24 hours is $100 \%-99.9 \%=0.1 \%=1 / 1,000$. This means your life expectancy is 1,000 days or $\sim 1,000 / 365=2.74$ years. (This is not exactly true.) The life expectancy for men in the U.S., from the U.S. Social Security actuarial tables (https://www.ssa.gov/oact/STATS/table4c6.html) accessed 3/13/23, is 2.86 years for a 95-year old (for whom the probability of living another day is $=99.904 \%>99.9 \%$ ) and 2.69 years for a 96-year old (for whom the probability of living another day is $=99.898 \%<99.9 \%$ ). So, it is true, on average, for a 96 -year old.
** Ch. 16. Why does the expected death age of adult men increase as they age?
(Distribution shapes and tails, Section 15.2.2; Statistics, Actuarial Tables, Section 16.3)
When consulting the life expectancy tables, an adult man notices that for every year he lives his life expectancy decreases by less than a year, and so it is expected that he will live to an older age. Does this make sense?

Answer: Yes, because the cohort (group) being examined is those people who have survived another year. For example, the life expectancy of the "average" 40-year old man in the U.S is 38.74 years, which decreases to 29.88 years at 50 , which is a 8.86 year decrease ( $<10$ years), according to the U.S. Social Security actuarial tables (https://www.ssa.gov/oact/STATS/table4c6.html) accessed 3/13/23). This is also true for women (For a 40-year old woman in the U.S it is 42.76 years, which decreases to 33.50 years at 50 , which is a 9.26 year decrease ( $<10$ years). This trend is not always true for children and teenagers for a variety of reasons, and may not be true for young adults in time of war.


#### Abstract

* Ch. 16. Math assessments, metrics (Section 16.4.2)

The Institute of Electronic and Electrical Engineering (IEEE) is a professional society in which members pay annual dues. If you have been a member long enough and are old enough, you can qualify to become an IEEE lifetime member, who does not need to pay annual dues. (As of 2020) If you are 65 year or older, you become a Life Member if the metric of the sum of your age and your years of IEEE membership equals or exceeds 100. How much does your qualification metric change each year?


Answer: For an active member it increases by 2 every year, 1 for the increase in your age and 1 for the increase in years of membership.

## ** Ch. 16. How does the "Misery Index" metric denote the misery endured by the public due to seemingly poor economic conditions? (Metrics, Section 16.4.2; Weighted sum, Section 5.2)

The "Misery Index" was devised by Arthur Okun (~1971) as a metric of the misery of the U.S. public due to economic issues. As constructed, it is the sum of the unemployment rate and the inflation rate. But, some studies indicate that the public feels worse about an increase in the former than the latter and so a sum weighting the former more is sometimes used, such as multiplying it with a weighting a factor of 2 (by Arthur Oswald, 2001) or a factor of 5 in later studies. Say, the revised metric multiplies the unemployment rate by 3 before adding it to the inflation rate. Using the original Misery Index and then with this revised metric, is the public most miserable with: the unemployment rate, inflation rate = (a) $2 \%, 8 \%$; (b) $4 \%, 4 \%$; or (c) $5 \%$, $2 \%$ ?

Answer: Respectively, for these three cases, the original index is (a) 10\%, (b) 8\%; (c) 7\%, and for the modified metric: (a) $14 \%$, (b) $16 \%$; (c) $17 \%$. With the original index, the public would seem to be most miserable with the first case (a), while with the modified version it would indicate it is most miserable with the last case (c). It is difficult to assess how miserable people are. Though a larger value of (whichever) metric may indicate more misery, it is not linear; if it doubles, the actual misery (however it is measured) might more than double.
[Inflation and Unemployment Both Make You Miserable, but Maybe Not Equally
https://www.wsj.com/articles/inflation-and-unemployment-both-make-you-miserable-but-maybe-not-equally-11668744274]

## * Ch. 16. Gaming the data reported as a metric (Section 16.4.2, Ch. 20 Risk)

The effectiveness of hospital care during stays is sometimes judged by the mortality rate reported 30 days after an event, such as surgery. In a hypothetical example, in given hospital the mortality rate was $2.3 \%$ after 27 days, $2.4 \%$ after 28 days, $2.5 \%$ after 29 days, $2.6 \%$ after 30 days (the reported rate), $3.2 \% \%$ after 31 days, $4.5 \%$ after 32 days, $5.8 \%$ after 33 days, and $5.9 \%$ after 34 days. Is the reported rate suspicious?

Answer: Yes, it is. In seems that special care was given to keep patients alive to 30 days and then nature was allowed to take its course. The true 30-day mortality is closer to $5.8 \%$ than the reported $2.6 \%$. Of course, the severity events differ so the morality rates would differ, and different types of events should be examined separately (an example of binning), and this needs to be understood in examining the statistics. Furthermore, use of this metric causes some hospitals and surgeons to try to avoid serious cases.
*** Ch. 16. The metric for the size of your luggage (Section 16.4.2)
One (non-statistical) metric used for the maximum allowed size of luggage checked in by an airline is the sum of its three dimensions for a rectangular solid bag (the "linear" (total) inches = length + height + width). Does this metric maximize either the volume, area, or length of an object inside it?

Answer: No, no, and no, but it can be optimized for a bag given this metric constraint. A common maximum bag allowed to be checked in by airlines is 62 linear (total) inches, which permits the common bag size of $27^{\prime \prime}$ by $21^{\prime \prime}$ by $14^{\prime \prime}\left(27^{\prime \prime}+21^{\prime \prime}+14^{\prime \prime}=62^{\prime \prime}\right)$, all subject to a common weight limit of 50 pounds. The maximum volume is a bit smaller than the cube of one third of the total linear length of 62, or $(62 / 3)^{3} \sim 8,827$ cubic inches, as can be shown by calculus-which we will not do, and which is larger than the volume of the given bag, 27" $\times 21^{\prime \prime}$ $\times 14^{\prime \prime}=7,938$ cubic inches. The maximum area is a bit smaller than the square of half of the total linear length of 62 , or $(62 / 2)^{2} \sim 961$ square inches, as can be shown by calculus, and which is larger than the largest area of a side of the given bag $27^{\prime \prime} \times 21^{\prime \prime}=567$ square inches. (This ignores that by placing an object across an internal diagonal the area of an object in the bag could be a bit larger.) The maximum length is a bit smaller than the total linear length of 62 inches, as can be shown by calculus, which is longer than the largest length of a side of the given bag, 27". (An item even longer than 27 inches could be placed in the bag well than 27 inches, as along the larger internal diagonals $\left(27^{2}+21^{2}\right)^{1 / 2}=34.2$ inches and $\left(27^{2}+21^{2}+14^{2}\right)^{1 / 2}=37.0$ inches.)

## * Ch. 16. Your normal body temperature may no longer be normal (Chapter 2, Chapter 16 including Section 16.4.2)

98.6 is a very recognizable number because the normal body temperature is $98.6^{\circ} \mathrm{F}$. But is it? For some time I have noticed that my body temperature measured at doctor's visits was always a degree or so below this value (and I did not have a fever each time). Can this make any sense?

Answer: Yes, it could, but ... I always thought the thermometer hadn't reached the final temperature, but that was not the reason. Or, that I had a slightly lower normal than average. (There is expected to be statistical distribution about the average, which has a full width of $\sim 2^{\circ} \mathrm{F}$, and perhaps I was just on the low side of the average. Chapter 16) But, that was not the reason. Perhaps that study in 1869 that provided this widely-used result was flawed or did not adequately correct for the variation in temperature measured at different places in the body-but these are not the reasons. Recent statistical studies show that the average body temperature is now $97.5^{\circ} \mathrm{F}$, and it is definitely lower than those in 1869 - and that the 1869 does not seem to have been flawed. So, my normal is normal! (Of course, 98.6 was never very special tyo those using the Celsius scale. in degrees Celsius the normal average had been $37.0^{\circ} \mathrm{C}$ (and is now recognized to be $36.4^{\circ} \mathrm{C}$ ).)
[https://www.wsj.com/articles/98-6-degrees-fahrenheit-isnt-the-average-any-more11579257001]

## * Ch. 16. Going to war over WAR in baseball (Section 16.4.3, Metrics in sports)

The WAR (Wins Above Replacement) statistic (Section 16.4.3) attempts to summarize a baseball player's value in one number. Should it be used to help determine a player's salary?

Answer: This is a judgement call, and a quite complicated and very controversial one that rears its head in discussions of bargaining agreements between major league players and team owners. Does WAR accurately reflect a player's contribution? Is there a preferred way to calculate it? Does it include all factors, including offensive and defensive statistics, different types of baseball parks (those friendly or not friendly to home runs, batting averages, left or right-hand batters, etc.), and so on? Are there player "intangibles?" Should there be a set salary for a player based on his recent WAR? If so, should it be linear nonlinear?
(https://www.wsj.com/articles/mlb-war-wins-above-replacement-11645548811)

## * Ch. 16. Statistics, correlation, causation (Section 16.4.4)

March 2020 saw a humongous increase in the number of COVID-19 cases in the U.S., a stock market crash, and the first availability of the printed version of Coming Home to Math. Was the similar timing of these events a causal or coincidental correlation?

Answer: The similar timing of at pairs of these may be causal or coincidental, but there is no statistical analysis presented here as would be needed for at least a mathematical determination of a true correlation-whether causal or not. In any case, it is not unlikely that the accelerating spread of COVID helped cause the crash of the stock market. However, it is not likely that the publication of Coming Home to Math was disruptive enough to cause the stock market to crash and it clearly did not affect the COVID pandemic. Though book sales were likely slowed by the pandemic, the depressed sales numbers were likely not enough to hurt the stock market.

## Chapter 17. The Math of Big Data

** Ch. 17. Can more of the people who were vaccinated for an infectious disease be ill than those who are unvaccinated? (Spreading of disease, Section 13.6; Statistics: Presenting Them, Section 16.3; Big Data, Chapter 17)
A vaccine is $90 \%$ effective for a particular infectious disease, so the probability that a vaccinated person contracts this disease is $90 \%$ smaller that for an unvaccinated person. (a) If none of the population of 100,000 is vaccinated, then 1,000 of the residents contract this infectious disease in a given time period. What fraction of the population is ill? (b) If instead they all had been vaccinated, how many would have become ill and what fraction of the entire population would this be? Repeat this if (c) $50 \%$, (d) $90 \%$, or (e) $99 \%$ of the population had been vaccinated, and what fraction of the infected people had been vaccinated? (f) What are the general trends?

Answer: (a) For 0\% vaccinated: 1,000/100,000 $=1.0 \%$ of everyone is ill. (b) For $100 \%$ vaccinated: Only $10 \%$ of the $1,000=100$ become ill, which is $100 / 100,000=0.1 \%$ of the population. (c) For $50 \%$ vaccinated: Of the $50 \%$ who were not vaccinated, which is 50,000 of them, $1 \%$ or 500 of them become ill. Of the $50 \%$ who were vaccinated, or 50,000 of them, $0.1 \%$ or 50 of them become ill. So now $500+50=550$ become ill, which is $550 / 100,000=0.55 \%$ of the population. Of the ill, $500 / 550=90.9 \%$ had not been vaccinated and $50 / 550=9.1 \%$ had been vaccinated. (d) For $90 \%$ vaccinated: Of the $10 \%$ who were not vaccinated, or 10,000 of them, $1 \%$ or 100 of them become ill. Of the $90 \%$ who were vaccinated, or 90,000 of them, $0.1 \%$ or 90 of them become ill. So now $100+90=190$ become ill, which is $190 / 100,000=0.19 \%$ of the population. Of the ill, $100 / 190=52.6 \%$ had not been vaccinated and $90 / 190=47.4 \%$ had been vaccinated. (e) For $99 \%$ vaccinated: Of the $1 \%$ were not vaccinated, or 1,000 of them, $1 \%$ or 10 of them become ill. Of the $99 \%$ who were vaccinated, or 99,000 of them, $0.1 \%$ or 99 of them become ill. So now $10+99=109$ become ill, which is $109 / 100,000=0.11 \%$ of the population. Of the ill, $10 / 109=9.2 \%$ had not been vaccinated and $99 / 109=90.8 \%$ had been vaccinated. (f) As more get vaccinated, fewer become ill and a larger fraction of the ill are those who had been vaccinated----and perhaps even most of the ill had been vaccinated. This is not a contradiction.
--- Added note: The key results of this problem would be the same for any size population group, and it could have been solved by using only the fraction of those who had been vaccinated and however effective the vaccination is---but putting in numbers sometimes makes it easier to visualize. Data need to be presented properly, as seen in the next problem.

## * Ch. 17. Can headlines saying that more vaccinated people are ill than unvaccinated ones be accurate? (Spreading of disease, Section 13.6; Statistics: Presenting Them, Section 16.3; Big Data, Chapter 17)

Using the results of the previous problem, such as part (e), which of the following headlines could be accurate, could be accurate but still be misleading and are not surprising, or could be inaccurate.
(a) Despite massive vaccination more of the ill had been fully vaccinated than not vaccinated
(b) Despite massive vaccination a higher fraction of the ill had been fully vaccinated than not vaccinated
(c) Despite massive vaccination more of the fully vaccinated are ill than of the not vaccinated (d) Despite massive vaccination a higher fraction of the fully vaccinated are ill than of the not vaccinated

Answer: As seen in part (e) of the previous problem, such a headline
(a) Could be accurate but be misleading and not be not surprising, with 99 of the vaccinated becoming ill and "only" 10 of the unvaccinated.
(b) Could be accurate but be misleading and not be not surprising, with $90.8 \%$ of the vaccinated becoming ill and "only" $9.2 \%$ of the unvaccinated.
(c) Could be accurate but be misleading and not be not surprising, with 99 of the vaccinated becoming ill and "only" 10 of the unvaccinated, as in (a) of the problem.
(d) Inaccurate, because $1 \%$ of the unvaccinated are ill and only $0.1 \%$ of the vaccinated.
--- Take home message: Numbers in presented in proper context are not deceptive, but the words that purportedly describe such numbers can be inaccurate and/or misleading. This is of particular concern in presenting statistical data, including "big data."

## Part IV: The Third World of Math: The Math of Making Decisions and Winning

## Chapter 18. The Math of Optimization, Ranking and Voting, and Allocation

## ** Ch. 18. Law of Diminishing Returns - Optimizing whose returns? (Section 18.2, Diminishing returns Section 4.4)

You buy a house and later want to sell it, and considering its cost, improvements, and so on you need to clear $\$ 450,000$. This would be beyond the seller and buyer broker fees that are each $3 \%$ of the selling price. Ignore other closing costs. You realize that if you sell your house for $\$ 500,000$ the broker fees are $\$ 30,000$, so you would clear $\$ 20,000$, which may actually cover some other costs of relocating and buying a new house of $\$ 15,000$, so it would be close to breakeven for you. You convince your (seller) broker to list the house at $\$ 520,000$, so you may net $\$ 40,000$ (aside from these other costs). So, for selling prices of $\$ 500,000$ and $\$ 520,000$ you technically net $\$ 20,000$ and $\$ 40,000$, or, with relocation costs, really $\$ 5,000$ and $\$ 25,000$.

Within one month you get an offer for $\$ 500,000$. (Assume the time needed to close on the house can be ignored.) You think that if the (seller) broker continues to work hard that you could sell for $\$ 520,000$ in three more months (for a total market stay of four months). In any case, you are not in a hurry and want to net more money. However, your broker pushes you to sell after the first offer. Why?
(To analyze this, consider:
(a) What are the total broker fees for selling prices of $\$ 500,000$ and $\$ 520,000$ ?
(b) How much does the (seller) broker earn on average for each per month of hard work for both cases (ignoring other costs)?
(c) How is the optimization math different for each interested party?
(d) Why is this an example of diminishing returns for the broker?)

Answer: This situation begs you to "do the math."
(a) The seller broker fees (3\%) would be $\$ 15,000$ and $\$ 15,600$ respectively.
(b) The seller broker earns on average for each per month $\$ 15,000$ and $\$ 15,600 / 4=\$ 3,900$.
(c) The optimization math tells the seller to wait 4 months for a sale so as to net another
$\$ 20,000$, while seller broker would rather sell in one month and spend all of his/her time in the next 3 months focusing on selling other houses.
(d) The selling broker receives a diminished earning per month with increased time on the market (with less opportunity to sell other houses).

## * Ch. 18. To Wait or not to Wait. This is the question. (Section 18.3.2, Optimization in waiting on line)

That frustrated customers leave lines or queues is called reneging (Section 18.3.2). This affects the statistics of waiting. Would you expect reneging to depend on their position in the line?

Answer: Yes, and it apparently does, particularly for those who are last in line. Studies show that they are 2.5 times as likely to switch lines within 30 seconds and three times as likely to totally leave all lines, than those in front of them. This should be included in the statistical analysis of queues, and suggests that stores should somehow make those in the back of the line happier to improve their business.
(https://www.wsj.com/articles/last-in-line-customers-11651095272)

## ** Ch. 18. How rising and ebbing tides of traffic. (Section 18.3.3, Optimization in traffic; Section 5.2, Distributions and symmetry; Section 15.2, Distributions; Section 15.2.2, The median)

Say there are no cars on the road at 7 AM , and then the number of cars entering the road every minute increases, say linearly with time, until 9 AM , and then it decreases in the same linear manner until it stops at 11 AM . So, the function giving the number of cars entering the road is symmetric about 9 AM , and, for example, it is the same 30 minutes before and after 9 AM , and the same an hour before and after 9 AM. Would you expect the number of cars that are actually on the road at any time to also be symmetric about 9 AM?

Answer: Not necessarily. If the rate of cars entering is low, it could well be symmetric. But if the volume is high enough and/or if traffic accidents occur that slow down traffic, it would likely not be symmetric.
** Ch. 18. Just how much salt is there? (Ranking, Section 18.4.1)
A spice mixture lists five ingredients in decreasing weight with item 4 being salt and item 5 being sea salt. What is the possible range of the total salt content by weight?

Answer: The weight \% of items 4 and 5 could both range between just above $0 \%$ to just below $20 \%$ (which is $100 \% / 5$ ), so the total salt content could be range between just above $0 \%$ to just below $40 \%$. There could be much more salt in the spice blend than you could suspected because there are two "types" of salt and the limited information provided by merely ordering the ingredients by weight.

## *** Ch. 18. In second place after the first round of ranked choice voting, but maybe winning the election, with all providing rankings (Section 18.4.2)

Ranked choice voting, as noted Section 18.4.2 in "Coming Home to Math," seems to becoming more popular, and was used for the first time in the New York City mayoral primary in 2021. How hard is it for the round 1 second-place candidate to win in this system? Say there are 3 candidates, A who received $40 \%$ of the first-ranked votes, B, who received $32 \%$, and C, who received $28 \%$. Because no candidate won at least $50 \%$ of the vote, candidate C is removed from the process and in the next round (the final round in this case because there are three candidates) $A$ and $B$ receive the votes from those who voted for $C$ as first ranked, and $A$ and $B$ as the second choice, respectively. If all those first voting for $C$ did actually give a second-ranked choice (and this is not required), what fraction of C's second choice votes would A and B need to win?

Answer: The winning candidate needs to reach a just a bit over $50 \%$ of all votes in the second round, so $A$ would need $10 \%$ more of the originally-cast votes and $B$ would need $18 \%$. So, $A$ would need a fraction so that $28 \%$ times that fraction is at least $10 \%$, and this is $10 \% / 28 \%$ ~ $35.7 \%$. B would need at least $18 \% / 28 \% \sim 64.3 \%$. (And, of course, these numbers sum to 1.) Another (equivalent) way to approach this problem is to say there are, perhaps, 1,000 votes cast in the first round, with $A, B, C$ and respectively receiving 400, 320, and 280 votes respectively. $A$ needs at least 100 more of C's 280 first-round votes in the second round to reach at least 500 and so $100 / 280 \sim 35.7 \%$ of them and $B$ would need at least 180 or more, so $180 / 280 \sim 64.3 \%$, as before. It likely that $A$ will win, except if $B$ and are very aligned political. (Of course, the result would be the same if we had chosen the number of voters to be other than 1,000.) In general, it has been found that if the second-round candidate needs to "make up" at least $5 \%$ from the firstround vote, he/she is unlikely to win-=but it is still possible. (For example, see https://nypost.com/2021/06/23/nyc-mayoral-primary-eric-adams-lead-should-survive-ranked-choice-experts-say/ (retrieved 6/23/22, https://www.wsj.com/articles/new-york-mayoral-race-what-happens-now-with-vote-count-11624480237 (retrieved 6/24/21))
*** Ch. 18. In second place after the first round of ranked choice voting, but maybe winning the election, but not with all voters providing rankings (Section 18.4.2)
What are these fractions of redistributed votes in the previous problem, if only $60 \%$ of those first voting for C provided a second choice? (Because this is less than $100 \%$, this is an example of "ballot exhaustion," in which not all voters have a say in the final round of ranking.)

Answer: Using the second approach in the previous problem--presuming there are 1,000 voters---now only $80 \%$ of C's 280 first-round votes, $0.6 \times 280=168$ can be redistributed in the secondround. In that second round there will now be a total of $400+320+168=888$ votes and so the candidate now with at least $888 / 2=444$ votes will win. A needs to get $444-400=44$ more votes and B 444-320=124 more votes, which are 44/168~26.2\% of C's voters who made a second ranking for $A$ and 124/168~73.8\% for $B$. This is now a much more challenging task for $B$.
*** Ch. 18. In second place after the first round of ranked choice voting, but maybe winning the election without a "true majority" because of ballot exhaustion (Section 18.4.2) Repeat the previous problem, but now find what fraction of C's second choice ballots need to go to A to provide that candidate with a "true majority" (a majority of all initial voters), and not merely a majority of voters whose ballots had not been exhausted in the final round. (Such a true majority is usually not required).

Answer: A will now need to have been selected as a second choice in at least 100 of the ballots of those who first preferred C and provided a second choice, or $100 / 168=59.5 \%$ of them to get the at least $(400+100=) 500$ votes needed for a true majority. This is a larger fraction than needed for a final round simple majority (of the remaining ballots) if all those first voting for $C$ provided a second choice ( $>35.7 \%$ for $A$, as in two problems up-and this is also a true majority) and if only $60 \%$ provided a second choice ( $>26.2 \%$ for $A$, as in previous problem). (Detractors of ranked-choice voting say that the outcome does not include the say of those voters whose ballots do not count in the final round because of "ballot exhaustion," and this counters claims that with this method one obtains a true majority. (https://www.wsj.com/articles/start-spreading-the-newsranked-choice-voting-is-a-mess-11625178082))

## Chapter 19. The Math of Gaming

Problems to be added.

## Chapter 20. The Math of Risk

## * Ch. 20. Risks in new life insurance policies

Older people are currently (c. 2021) being targeted in TV ads to buy whole life insurance for which no medical exam is required, you cannot be turned down, your premiums will not increase and for which your benefits cannot never be lowered or terminated (as long as you continue to pay the premiums that may increase in age in a clearly stated manner). We estimated how much term life insurance could cost in this chapter-but any such an estimate of life insurance needs to include all factors. What important factor is not being included in this description of such whole life insurance?

Answer: The benefits can never go down, but when do they start in full force or at all? Some, but not all, ads apparently state that payout may be reduced below the full amount in the first few years. Knowing this is important in deciding whether or not buy this insurance and in evaluating whether or not the potential benefits are financially worth the financial risk.

## ** Ch. 20. How risk varies with exposure (Risk/Exposure Section 20.1)

Sketch on one graph the relative risk of contracting a disease vs. dose of exposure, if the measured risk at one higher dose were the same for each graph. Assume an extrapolation to lower doses with either (a) a linear fit, (b) a decreasing slope with increasing dose, or (c) an increasing slope with increasing dose-all with no thresholds, or (d) with a threshold and then a linear increase, or (e) a hormetic fit-which means an initial decrease to negative risk (perhaps suggesting a benefit at low doses) and then increasing to positive risk with a decreasing slope.

Answer: These are seen in Fig. 3 of https://pubmed.ncbi.nlm.nih.gov/14610281/.

## *** Ch. 20. How risk varies with exposure in an unexpected way, with a small sensitive group (Risk/Exposure Section 20.1)

For a given disease, most people ("normal") have a linearly increasing risk of contracting a disease with increasing dose-with no threshold, and that their probability of contracting the disease is very small even for moderate doses. However, a very small fraction of people is much more sensitive ("sensitives"), and their risk increases rapidly with dose and then it saturates to a risk of 1 at moderate doses. (a) Sketch the percentage of all people who would contract the disease vs. dose, including only normals, only sensitives, or both groups. Say that $0.25 \%$ of all are sensitives and that the affected percentage of the total population is the same for normals and sensitives for a dose of 1,500 (in a particular set of units). (b) Explain why the mystery of the shape for the total population is resolved, if it were thought this curve should be linear, when it is understood that a small group is very sensitive to exposure.

Answer: (a) These are seen in Fig. 5 of https://pubmed.ncbi.nlm.nih.gov/14610281/ for the hypothetical case of radiation-induced breast cancer, in which a small fraction of women is very
sensitive to radiation. (b) Most people have a linear dependence, but a small fraction has a much larger risk and one that saturates in the plotted dose range, and so the sum has a decreasing slope with no sign of saturation.

## * Ch. 20. Ranking risks and using such rankings (Ranking, Section 18.4.1; Risk/Exposure Section 20.1)

The risk of contracting the initial strain of COVID-19 from different activities has been assigned values from 1 (as for opening the mail) to 9 (attending a crowded sporting event) for people following recommended safety protocols. You are told that playing tennis is a 2, while working out in a gym is an 8 . Are these mere rankings, denoting less and greater risks, so they only mean that (on the basis of this information) the risk of contracting COVID is very risky for working out in a gym and playing golf has little risk? Or, does it mean the risk when working out in a gym is $4 \times$ larger than that when playing tennis? Or, is it really faster than linear and so the relative risk is much larger than 4 ?
[https://www.fox5ny.com/news/medical-group-releases-chart-ranking-activities-based-on-covid-19-risk]

Answer: As such, these are merely rankings by relative risk levels, but sometimes such rankings and groupings could align with specific quantitative risk assessments that have linear or nonlinear numerical significance (such as rank 8 being $4 x$ as risky as rank 2 or $4^{2} \times$ or $16 \times$ times as risky). This is not clear in this case.
** Ch. 20. Infectious disease transmission, risk (Spreading of disease, Section 13.6) A model of the transmission of an infectious disease that is transmitted in the air says the probability of you being infected increases linearly from 0 to $100 \%$ when the length of time an infected person is anywhere within 6 feet of you increases from 0 to 10 minutes. (For longer times, it stays $100 \%$.) Say, an infected person jogs right next to you and passes you with a speed of 4 mph . (If both of you are walking this is your relative speed.) What is contact time during which you can get infected? (Assume this occurs from when that person is 6 feet in front of you to 6 feet behind you. Also, use the conversion in the book of mph to fps (feet per second).) Estimate the probability that you will become infected from this single encounter? (Ignore any increased risk due to the increased breathing rate of the jogger.)

Answer: $60 \mathrm{mph}=88 \mathrm{fps}$, so $4 \mathrm{mph} \sim 6$ feet per second (Section 11.2.1). So, with 6 feet +6 feet $=12$ feet, the contact time 12 feet $/(6$ feet per second $)=2$ seconds. 10 minutes is 600 seconds so the probability of infection is 2 seconds/(600 seconds) $=1 / 300$ or $\sim 0.3 \%$.

## ** Ch. 20. Weibull model analysis of equipment breaking (Sections 15.2.2 and 20.1.2)

A company makes 3 products. Product A breaks at a per year rate of $1 \%$, but in years $11,12,13$, $\ldots$ this rate increases to $5 \%, 15 \%, 30 \%, \ldots$. Product B breaks at a rate of $3 \%$. Product C breaks at a yearly rate of $20 \%$ in year $1,6 \%$ in year 2 , and then at $2 \%$ for later years. What does this mean and how can this be characterized?

Answer: Product A is subject to breakages that arise after long-term use-with parts wearing out, $C$ to the initial unreliability of brand new parts, and B to neither. The Weibull model is a modified exponential decay distribution that is used to characterize this.

## ** Ch. 20. How to model the probability of falling plates breaking? (Sections 15.2.2 and 20.1.2)

You notice that a plate you frequently use falls to the floor about once a month, and the $19^{\text {th }}$ time it falls it breaks. You hear that others have a similar experience with that type of plate and with a similar floor and falling distance, but some plates break earlier and some later than yours? (a) Which type of statistical distributions should be used to analyze this? (b) Which do you think would be the more relevant variable in this analysis, how the long the plate has been used or the number of times it has fallen?

Answer: (a) The distribution describing breakage from long-term aging is the Weibull distribution with $k>1$, for which the failure rate increases with time. Microcracks develop at during each collision with the floor, and eventually there are so many that it breaks. (b) Probably, the number of falls is most relevant because the falls likely contribute most to aging and eventually breakage, though usage without falls also contributes, so time is not totally irrelevant.

## * Ch. 20. What happens if the retirement income advice changes? (Section 20.3.2, Risk in retirement financial planning)

As a newly retired person, say you receive $\$ 25,000$ per year in Social Security benefits in the first year (with cost of living allowance added in subsequent years) and have \$400,000 invested in a recommended portfolio of stocks and bonds. (a) Say, the prevailing recommendation has been to withdraw $4.0 \%$ of this amount in the first year (with cost of living allowance added in subsequent years), so you would have a strong likelihood of having enough income for 30 years. What would your planned annual income be in that first year? (b) Say, future stock performance has just been judged to be much bleaker than had been thought and the current safe recommendation is now to withdraw only $3.0 \%$ of your portfolio in the first year. Then, what would your annual income be in the first year? (c) Say, that future stock performance has just been judged to be much better than had been thought and the current recommendation is now to withdraw $4.5 \%$ of your portfolio in the first year. Then, what would your annual income be in the first year.

Answer: (a) $\$ 25,000+\$ 400,000 \times 4.0 \%=\$ 25,000+\$ 16,000=\$ 41,000$. (b) $\$ 25,000+$ $\$ 400,000 \times 3.0 \%=\$ 25,000+\$ 12,000=\$ 37,000$. (c) $\$ 25,000+\$ 400,000 \times 4.5 \%=\$ 25,000+$ $\$ 18,000=\$ 43,000$.
[The 4\% Retirement Rule Is in Doubt. Will Your Nest Egg Last?, Anne Tergesen, November 12, 2021, Wall Street Journal]

## ** Ch. 20. IRAs (Retirement planning, Section 20.3.2)

You are told that you need to remove $5.0 \%$ (or more) of your IRA holdings (by a certain date) in year $1,5.2 \%$ of what remains in the account in year $2,5.4 \%$ in year 3 , and so on. Just before your first withdrawal, you have $\$ 100,000$ in the account. (a) If you withdraw only this minimum amount each time, how much do you withdraw in each of the first 3 years and what remains in your IRA after the third withdrawal (assuming the value of your investments change only due to these withdrawals)? (b) Repeat this if the amount in your accounts increase by $2.0 \%$ (due to a sudden increase in IRA investments) just before the withdrawal dates in years 2,3 , and so on.

Answer: (a) In year 1, you withdraw $\$ 100,000 \times 5.0 \%=\$ 5,000$ and then have $\$ 100,000-$ $\$ 5,000=\$ 95,000$ left. In year 2 , you withdraw $\$ 95,000 \times 5.2 \%=\$ 4,940$ and then have $\$ 95,000$ $-\$ 4,940=\$ 90,060$ left. In year 3, you withdraw $\$ 90,060 \times 5.4 \%=\$ 4,863.24$ and then have $\$ 90,060-\$ 4,863.24=\$ 85,196.76$ left. (b) In year 1 , you withdraw $\$ 100,000 \times 5.0 \%=\$ 5,000$ and then have $\$ 100,000-\$ 5,000=\$ 95,000$ left. In year 2 , you have $\$ 95,000 \times 1.02=\$ 96,900$ (Since $1+2 \%=1.02$.) and you withdraw $\$ 96,900 \times 5.2 \%=\$ 5,038.80$ and then have $\$ 96,900-$ $\$ 5,038.80=\$ 91,861.20$ left. In year 3, you have $\$ 91,861.20 \times 1.02=\$ 93,698.42$ and you withdraw $\$ 93,698.42 \times 5.4 \%=\$ 5,059.71$ and then have $\$ 93,698.42-\$ 5,059.71=\$ 88,638.71$ left.

## ** Ch. 20. Disease risk and scaling (Spreading of disease, Section 13.6; Scaling, Section 8.2.5)

You are told that the probability of air-borne infection for a given exposure time is $5 \%$, and that it increases with longer exposures. Your exposure time is four times this time. (a) How would you feel about you risk of infection if you were also told that this probability varies logarithmically, sublinearly, linearly, superlinearly, or exponentially with exposure time? (b) ... or it varies as the square root of the exposure time? (c) ... or it varies as the square of the exposure time (or quadratically with it)?

Answer: (a) For a linear dependence it would be $4 \times 5 \%=20 \%$. For a logarithmic or sublinear dependence, it would be 5-20\%, but you cannot obtain a better answer without more information. Your concern would be greater, but less than if it were linear. For a superlinear or exponential dependence, it would be $>20 \%$, but you cannot obtain a better answer without more information. (b) The square root of 4 is 2, so the probability would be $2 \times 5 \%=10 \%$. (c) The square of 4 is 16, so the probability would be $16 \times 5 \%=80 \%$. Your concern would be greater, and more so than if it were linear. (The exact probability may not be $80 \%$ because of saturation (Section 4.4).)

## * Ch. 20. Disease risk, scaling, and saturation (Spreading of disease, Section 13.6; Scaling Section 8.2.5; Saturation, Section 4.4)

In the previous problem, say the probability of infection increases linearly with time and your exposure time is 40 times this given time. What is the risk of infection?

Answer: For a linear dependence it would seem to be $40 \times 5 \%=200 \%$. This does not make sense because this is larger than $100 \%$, which would mean certainty. This is an example of saturation. In any case, the risk of infection would now be quite high.

## * Ch. 20. Disease risk and inverse scaling (Spreading of disease, Section 13.6; Scaling,

 Section 8.2.5)In the problem before the previous one, say the probability of infection varies inversely with time and your exposure time is 4 times this given time. What would you say concerning the risk of infection?

Answer: Such an inverse dependence with time does not make sense, so the model is wrong.

## * Ch. 20. What does it mean for a percentage of people to test positively during an epidemic? (Spreading of disease, Section 13.6)

At the early stage of an epidemic, 40 out of 1,000 , or $4 \%$, of the people tested for the disease in a community are tested positive for the disease in one week. In a later stage, 80 out of 10,000 , or $0.8 \%$, of the people in that same community test positive in a week-so more people test positive but the rate of positive tests is smaller. What does this mean about the progress of the infection?

Answer: The data may have been taken in a proper manner and the rates of false positives and negatives could have been very small, but little can be concluded without more information. Other than more people testing positive and the rate of positive tests being smaller initially, you conclude nothing from the math as presented. This situation could arise because only those exhibiting symptoms were being tested initially, but even then not all of them were tested, while later all those showing symptoms could have been tested and a network of asymptomatic people could also have been tested.


[^0]:    * Ch. 5. Are August and October the same? Math and words (Section 5.4)

    People born in August and in October both claim to be born in the eighth month of the year. Do both have somewhat reasonable claims?

[^1]:    ** Ch. 15. Probability, health risk (Ch. 20) (Spreading of disease, Section 13.6, Basics, Section 15.1)
    At an event you are exposed to an airborne virus that leads to you to have a $70 \%$ probability of contracting an infectious disease. Your partner is with you at the time and has the same probability of contracting it. If because of contact with your partner you have a $100 \%$ of catching the disease from your partner, what is the probability you will become ill?

