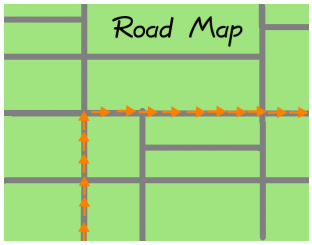
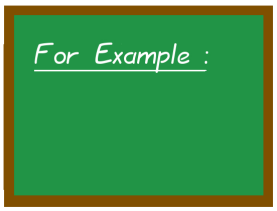
Module 4.2: The Permutation and Factorial Principles



In the previous module, we learned the exponent principle and the multiplication principle. In contrast to those broad principles, which are concepts used in solving a combinatorics problem, there are shortcut formulas that are useful for solving the most common cases of problems in combinatorics and probability. In this module, we're going to learn two of those shortcuts: factorials and permutations.



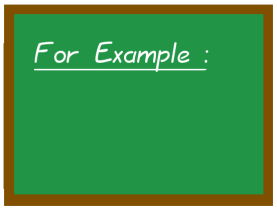
4-2-1

Imagine that you're planning a major event, like the annual shareholders meeting of a corporation. Suppose there are five speakers for the event. How many different ways are there to order the speakers?

There are five choices for the first speaker, then four choices (any but the first) for the second speaker, three choices (any but the first or second) for the next speaker, followed by two choices (any but the first three) and finally, for the last speaker, there is only 1 choice. This results in a final answer of

$$5 \times 4 \times 3 \times 2 \times 1 = 120$$

possible orderings. As you can see, we've used the multiplication principle.

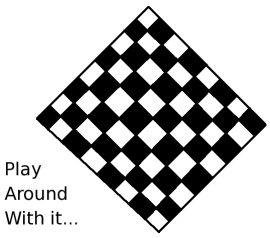


4-2-2

Let's consider a variation of the previous box. What if there were 6 speakers?

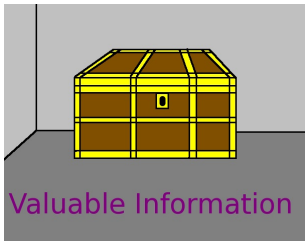
Then there are six choices for the first speaker, five for the second speaker, four for the third speaker, three for the fourth speaker, two for the fifth speaker, and one for the sixth speaker.

$$6 \times 5 \times 4 \times 3 \times 2 \times 1 = 720$$



4-2-3

- What if there were only four speakers? [Answer: $4 \times 3 \times 2 \times 1 = 24$.]
- What if there were seven speakers? [Answer: $7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1 = 5040$.]



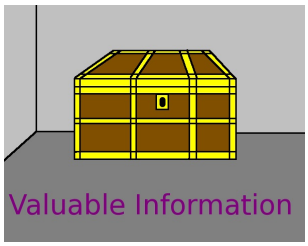
The long multiplication

$$n \times (n-1) \times (n-2) \times \cdots \times 3 \times 2 \times 1$$

occurs very often. Therefore, it is nice to have an abbreviation for it. This product can be abbreviated "$n!$" in mathematical writing.

This is spoken as "n factorial" in formal situations or "n bang" during informal chatter. The numerical value of $n!$ can get extremely large even for relatively small n. The *factorial* is a mathematical operation which is not very rare but also not very common.

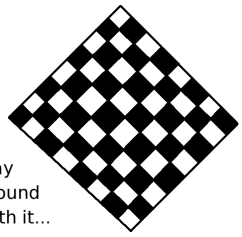
COPYRIGHT NOTICE: This is a work in-progress by Prof. Gregory V. Bard, which is intended to be eventually released under the Creative Commons License (specifically agreement # 3 "attribution and non-commercial.") Until such time as the document is completed, however, the author reserves all rights, to ensure that imperfect copies are not widely circulated.



In the previous few boxes, we examined the number of ways to order four, five, six, or seven speakers at a corporate meeting. I could have easily asked about ordering cities on a tour, or ordering gadgets in a product demo. The number of orderings will be the same whether you are ordering speakers, distinct cities, or different gadgets for a demo.

In general, we've learned that there are $5! = 120$ orderings for 5 items, there are $6! = 720$ orderings for six items, there $4! = 24$ orderings for four items, and also there are $7! = 5040$ orderings for seven items. (I sincerely hope that the speakers at the annual meeting are not insulted by the fact that I have just now referred to them as "things.")

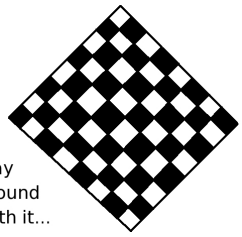
In general, there are $n!$ distinct ways to order n items. This is the *factorial principle*.



4-2-4

There is a certain style of elections in situations with many candidates, called "Instant Run-off Voting." In an IRV election, each voter is presented with a list of candidates. They must rank each candidate, indicating who is their first choice, second choice, third choice, fourth choice, and so forth. For example, in an election with 8 candidates, voters would have to place the numbers 1, 2, 3, 4, 5, 6, 7, and 8 on the ballot, with one number next to each candidate's name.

- How many possible ways are there to order the 8 candidates on this ballot? [Answer: $8! = 40,320$.]
- If there were only 6 candidates, how many ways are there to order the candidates? [Answer: $6! = 720$.]

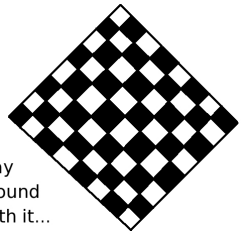


4-2-5

Suppose a student at that charity auction has obtained part-time work, staffing the coat check. Unfortunately, this student took advantage of the "open bar" and got drunk, and didn't keep track of which coat belongs to which guest. He's going to return the coats randomly to each guest instead. Suppose there are nine guests. What is the probability that each person gets the correct coat?

Hint: Out of all the possible orderings in which the student could hand the coats to the guests waiting in line, there is only one ordering which is correct.

- How many orderings are there of nine coats? [Answer: 362,880.]
- Since only one of those orderings is the correct ordering, what is the probability that each person gets the correct coat? [Answer: 1/362,880.]



4-2-6

Repeat the above problem for a spring event, where only 5 guests needed to use the coat check. (No one else had a coat or jacket.)

- How many ways are there to assign 5 coats to 5 people? [Answer: 120.]
- What are the chances that returning the coats randomly will be totally correct? [Answer: 1/120.]

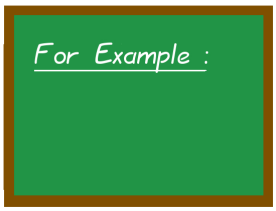
While this is vastly better than before, surely it is a fairly low probability, being less than 1%.

COPYRIGHT NOTICE: This is a work in-progress by Prof. Gregory V. Bard, which is intended to be eventually released under the Creative Commons License (specifically agreement # 3 "attribution and non-commercial.") Until such time as the document is completed, however, the author reserves all rights, to ensure that imperfect copies are not widely circulated.



The problem of the previous box is another example of how rapidly the output of a combinatorial problem can change when the inputs change only slightly.

We saw examples of rapid change in combinatorics previously, on Page 291 and on Page 413.



4-2-7

Now I'm going to tell you about a simple card game which I thought up while a graduate student. We take the cards 2♡, 3♡, 4♡, 5♡, 6♡, 7♡, and 8♡, from an ordinary deck of cards. The dealer will shuffle them, then the player will shuffle them some more, and the dealer will shuffle them again. The cards will be carefully flipped over one at a time. If the cards happen to end up in order (either decreasing or increasing) then the player wins $ 2000. Otherwise, the player has to pay the dealer $ 1.

I think that a lot of people would be willing to play this game. Is this a good game for the player to be playing? Let's make that question precise by calculating the probability that the player will win the game.

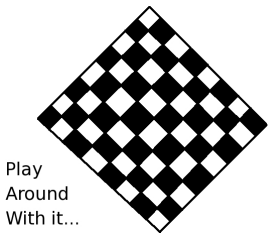
I think that when many students first see this game, they might be *very* attracted to playing it. As you can tell by the tone of my writing in this box, the game is not a good one for the player, though it is a great game for the dealer. First, let's ask ourselves, how many ways are there to order these cards?

There are seven cards. Therefore, there are $7! = 5040$ ways of ordering the cards. Of all those 5040 orderings, only 2 orderings are winners: perfectly increasing and perfectly decreasing. Therefore, the probability of wining is

$$\frac{2}{5040} = \frac{1}{2520} = 0.000396825\cdots$$



We will analyze the game from the previous box, in more detail, during the module "Expected Value and Insurance," on Page **??**. For the time being, we can state that since the chance of winning is 1 in 2520, the payment of $ 2000 is not enough. If the payment were $ 2521 or more, then the game would favor the player. If the payment is $ 2519 or less, then the game would favor the dealer. Since $ 2000 is significantly less than $ 2520, there is a strong bias toward the dealer in this game.



4-2-8

Let's modify the game in the previous two boxes, so that it has fewer cards. What is the probability of a win...

- ...using six cards? [Answer: 1/360 or $0.00277\overline{7}$.]
- ...using five cards? [Answer: 1/60 or $0.016\overline{6}$.]

COPYRIGHT NOTICE: This is a work in-progress by Prof. Gregory V. Bard, which is intended to be eventually released under the Creative Commons License (specifically agreement # 3 "attribution and non-commercial.") Until such time as the document is completed, however, the author reserves all rights, to ensure that imperfect copies are not widely circulated.



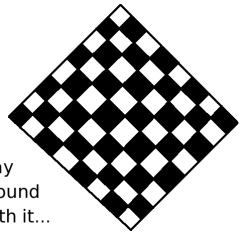
The previous problem is an example of how, in combinatorics, a small change in the inputs can lead to a huge change in the outputs. Let's examine how changing the number of cards affects the probability of a win.

- Five Cards: $1.6666\cdots\%$.
- Six Cards: $0.27777\cdots\%$.
- Seven Cards: $0.0396825\cdots\%$.

Surely, this is counter-intuitive. The probabilities changed a great deal each time we added just one card. Behavioral experiments have shown that humans are not very good at numerically estimating probabilities without first performing a calculation. As humans, we simply are unable to guess probabilities well. Quite frankly, this is the reason that casinos are extremely profitable.

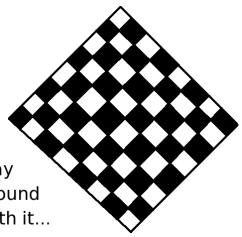
Therefore we should either calculate probabilities precisely, or at least engage in an approximate calculation.

Since the factorial is a mathematical tool that you might or might not have seen before, I think it is prudent to take a few moments to practice with it.



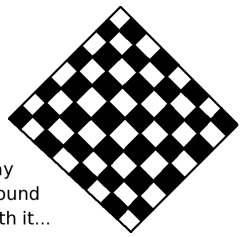
4-2-9

- What is $(2 \cdot 4)!$ equal to? [Answer: 40,320.]
- What is $2(4!)$ equal to? [Answer: 48.]
- Is the statement $(2 \cdot 4)! = 2(4!)$ true? [Answer: Nope.]
- In general, can we assume that $(2x)! = 2(x!)$ is true? [Answer: Certainly not.]



4-2-10

- What is $(3 + 4)!$ equal to? [Answer: 5040.]
- What is $3! + 4!$ equal to? [Answer: 30.]
- Is the statement $(3 + 4)! = 3! + 4!$ true? [Answer: Nope.]
- In general, can we assume that $(x + y)! = x! + y!$ is true? [Answer: Certainly not.]



4-2-11

- What is $(3 \cdot 2)!$ equal to? [Answer: 720.]
- What is $(3!)(2!)$ equal to? [Answer: 12.]
- Is the statement $(3 \cdot 2)! = (3!)(2!)$ true? [Answer: Nope.]
- In general, can we assume $(xy)! = (x!)(y!)$ is true? [Answer: Certainly not.]

COPYRIGHT NOTICE: This is a work in-progress by Prof. Gregory V. Bard, which is intended to be eventually released under the Creative Commons License (specifically agreement # 3 "attribution and non-commercial.") Until such time as the document is completed, however, the author reserves all rights, to ensure that imperfect copies are not widely circulated.



If you're wondering what the three previous boxes are about, they deal with misconceptions that I've seen while grading student final exams. The three common misconceptions dealing with the factorial are given below:

WRONG!	$\rightarrow$	$(xy)!$	$=$	$x(y!)$	$\leftarrow$	WRONG!
WRONG!	$\rightarrow$	$(x+y)!$	$=$	$x!+y!$	$\leftarrow$	WRONG!
WRONG!	$\rightarrow$	$(xy)!$	$=$	$(x!)(y!)$	$\leftarrow$	WRONG!



There's a cool mathematical property

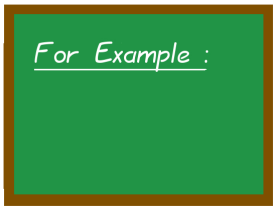
$$n^n > n! > 2^n$$

for any integer $n > 3$. Maybe you'll be able to see why this is true if you take a few minutes to think about it. If you get stuck, try calculating n^n and $n!$ as well as 2^n for some moderate sized n, like 10, on your calculator.

Alternatively, suppose you were trying to explain to a younger sibling what 10^{10}, 10!, and 2^{10} actually mean. Imagine you were explaining it using a pencil, and not a calculator. In so doing, you might reveal the mathematical pattern that results in an explanation of the above inequalities.

If you cannot think of an explanation, don't panic. This is a very hard challenge—but one worth thinking about. The answer is given at the end of the module, on Page 451.

In the next few boxes, we're going to explore how the topic of combinatorics can come up in entrepreneurship, by looking at a business launching a new website.



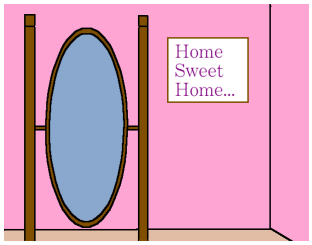
4-2-12

Here is a model for an anonymous backup and file-sharing service. You can upload your files anonymously to a server, and you get a secret code back. Anytime that you wish to retrieve your files (such as if your laptop gets stolen), then you can just enter the secret code and download them. In this way, you could also share your files with others, by giving them your secret code. Suppose that the secret code will be a sequence of four letters without repeats. How many codes are there? (Assume the standard English alphabet is used.)

We know there are 26 letters of the alphabet, and we are forbidden to have any repeats. There are 26 choices for the first letter, and then 25 choices for the second letter (since the first letter is forbidden). Next, there are 24 choices for the third letter (since two letters are now forbidden), and finally there are 23 choices for the fourth letter (since three letters are now unavailable). Using the multiplication principle, we get a final answer of

$$26 \times 25 \times 24 \times 23 = 358,800$$

possible secret codes.

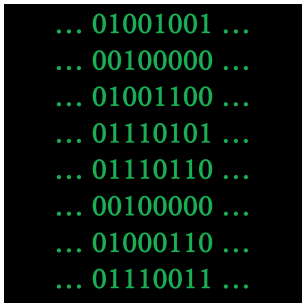


A Pause for Reflection...

As a future business person, do you see a problem with the circumstance of the previous box? Surely, we must take care to make sure that every new user gets a new code, instead of someone else's code.

Numerous websites have a million users or more. Some of the most popular websites in the world have hundreds of millions of users. We cannot be sure that this new business will reach that level of popularity. However, a new business should not ruin its chances for success by prematurely capping the maximum number of possible users at only 358,800.

COPYRIGHT NOTICE: This is a work in-progress by Prof. Gregory V. Bard, which is intended to be eventually released under the Creative Commons License (specifically agreement # 3 "attribution and non-commercial.") Until such time as the document is completed, however, the author reserves all rights, to ensure that imperfect copies are not widely circulated.

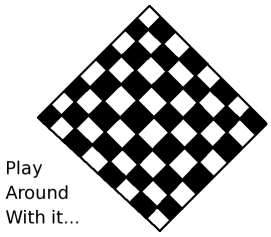


Of course, if the secret code is too long, it might be intimidating to users. Also, if the codes are too long, users might make the error of writing the code down, instead of memorizing it. Writing a code down opens up the possibility of some unauthorized person getting access to the files, by finding the note where the code was written.

A new website should be easy to use, and that's an argument for making sure that the secret codes aren't too long. This sort of balance between usability and absolute security is one of the hallmarks of the subject of computer security.

While it is a young subject, computer security is a very important subject. Luckily, it is also relatively easy. I would strongly encourage any business student who has any interest in e-commerce or other online activities to learn something about computer security.

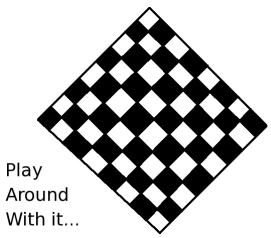
For a start, I highly recommend the textbook *Elementary Information Security*, 2nd edition, by Richard E. Smith, published by Jones & Bartlett Learning in 2015.



4-2-13

We return now to the previous example about codewords. Recall, our codewords were sequences of letters from the English alphabet. However, no letter can be repeated. How many possible codes are there using:

- Five letters? [Answer: 7,893,600.]
- Six letters? [Answer: 165,765,600.]
- Seven letters? [Answer: 3,315,312,000.]
- Eight letters? [Answer: 62,990,928,000.]



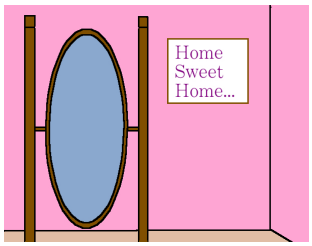
4-2-14

Continuing with the previous box, for reasons of usability, let's suppose that some research is done on letters that are frequently confused with each other. Suppose a Human-Computer Interaction Lab (HCIL) is hired to look at this matter of confused letters. They conclude that `Q` and `O`, and `I` and `J` are the most confused pairs. Accordingly, the website decides to ban these four letters from the randomly assigned user codes. All the other letters are okay. How many possible codes are there now, using:

- Five letters? [Answer: 3,160,080.]
- Six letters? [Answer: 53,721,360.]
- Seven letters? [Answer: 859,541,760.]
- Eight letters? [Answer: 12,893,126,400.]
- Nine letters? [Answer: 180,503,769,600.]

The question of balancing security and usability is an important one. No one wants to be the victim of identity theft, but over-zealous security systems can be extremely difficult to use. There is even an annual conference on this topic, called "SOUPS" which stands for "Symposium On Usable Privacy and Security."

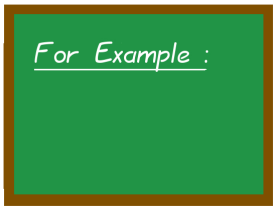
COPYRIGHT NOTICE: This is a work in-progress by Prof. Gregory V. Bard, which is intended to be eventually released under the Creative Commons License (specifically agreement # 3 "attribution and non-commercial.") Until such time as the document is completed, however, the author reserves all rights, to ensure that imperfect copies are not widely circulated.



A Pause for Reflection...

At the time of the writing of this book, the world's population was slightly over seven billion people. Therefore, we can see that 8 letters is more than sufficient for the file-storing business being discussed in the previous few boxes. On the one hand, some power-users might have more than one account, or even ten accounts. On the other hand, this is balanced out by the fact that it is unlikely that every living human being would be a customer of this specific business. Using 9 letters might be okay, but eventually the codewords become so long that they become hard to remember.

The possibility of 12.8 billion users is in contrast with perhaps using six letters, where only 53.7 million users could be supported. For example, on January 29th, 2014, Facebook announced that they have 1.23 billion users who visit at least once a month. While it is unlikely that our anonymous backup and file-sharing business will ever grow that large, it seems a pity to hamstring the business before it begins.



4-2-15

This notion of multiplying a sequence of integers, each smaller than the previous by exactly one, comes up a lot in combinatorics. Therefore, to simplify writing, there is a name and an abbreviation for it. This notation is easier to understand by looking at specific cases first.

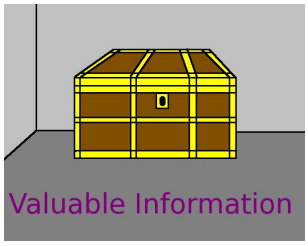
Looking at the final products from some of the last few boxes, we can define:

$$P_{26,4} = \underbrace{26 \times 25 \times 24 \times 23}_{4 \text{ numbers}}$$

$$P_{26,8} = \underbrace{26 \times 25 \times 24 \times 23 \times 22 \times 21 \times 20 \times 19}_{8 \text{ numbers}}$$

$$P_{22,9} = \underbrace{22 \times 21 \times 20 \times 19 \times 18 \times 17 \times 16 \times 15 \times 14}_{9 \text{ numbers}}$$

By the way, some books will write $P_{22,9}$ as P_9^{22} or as ${}_{22}P_9$, a point we will clarify on Page 433 of this module, and on Page 458 of the next module.

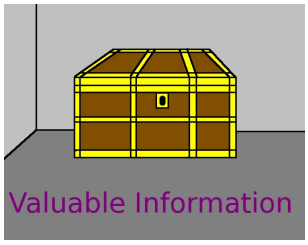


In the general case,

$$P_{n,x} = \underbrace{(n)(n-1)(n-2)(n-3)\cdots(n-x+3)(n-x+2)(n-x+1)}_{x \text{ numbers}}$$

Remember that you start with n, and then write the next few integers in decreasing order, until you have written x integers. By the way, almost all modern calculators have a button for this operation, so you do not have to multiply the numbers together manually.

This operator is called the *permutations* operator.



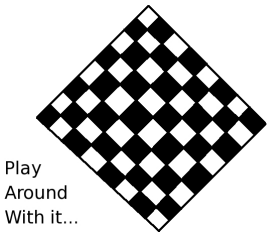
When one must make x selections, from a set of n options, and order matters, but repeats are not allowed, then the number of possibilities is $P_{n,x}$. This is the *permutation principle*.

This is the second member of a four-entry chart. This chart is phenomenally useful in solving combinatorial problems. We saw this chart for the first time on Page 407, in the previous module, and we'll see it again on Page 456, in the next module.

	Repeats OK	No Repeats
Order Matters	Exponent Princ.	Permutation Princ.
Order Doesn't Matter	???	???

We will add to this chart on Page 476, in the module "Which Combinatorial Formula Should I Use?" later in this chapter.

COPYRIGHT NOTICE: This is a work in-progress by Prof. Gregory V. Bard, which is intended to be eventually released under the Creative Commons License (specifically agreement # 3 "attribution and non-commercial.") Until such time as the document is completed, however, the author reserves all rights, to ensure that imperfect copies are not widely circulated.



4-2-16

Returning again to the example of the file retrieval and backup service, let's consider how many secret codes there are if the codes are 10 letters long. However, we'll write our answers using the "permutations" notation.

- Using all 26 letters: [Answer: $P_{26,10}$.]
- Using only 22 letters, because some letters can be confused: [Answer: $P_{22,10}$.]



Let's look at the first of the two answers from the previous box.

As you can see, $P_{26,10}$ is a much more concise notation and "name" for the number which is the answer. That number is normally called

$$19,275,223,968,000$$

which is a long and unwieldy name for a number. Likewise, it is easier to say $P_{22,10}$ than to say

$$2,346,549,004,800$$

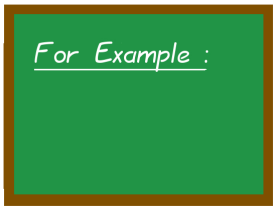
As my former colleague Prof. Brian Knaeble once pointed out to me, one can conclude that the symbol $P_{26,10}$ is a new, short and compact name for the number 19,275,223,968,000. Not only is this representation more convenient, it also transmits useful information, telling us how the number was obtained. In contrast, the name "19,275,223,968,000" is just a sequence of digits, and is therefore less informative.



Note that some books will write ${}_{26}P_{10}$ in place of $P_{26,10}$. However, this notation can be very confusing in the middle of a large formula, and therefore I will not use that notation in this book.

Another unwise system of notation is to use P_{10}^{26}, but this looks like some sort of exponent (the 26th power), or maybe a chemical element. Accordingly, I will not use that notation in this book, either.

We will discuss this matter further on Page 458, of the next module.



4-2-17

Suppose a realty office has six realtors, and it is a slow season. They decide to assign "walk-in" customers to realtors by rolling an ordinary 6-sided die. The problem with this is that it is possible some realtor might get multiple new customers while someone else gets zero. Suppose that four new customers arrive this week. What is the chance that one realtor, by coincidence, gets two or more customers, by luck of the dice?

This problem seems heavy but it can be approached by cutting into small pieces. First, four customers will result in four dice rolls. How many outcomes are there? Well, surely order matters. That's because if Customer #1 goes to Alice, and Customer #2 goes to Bob, that's not the same thing as Customer #1 goes to Bob, and Customer #2 goes to Alice. Since we are talking about dice, repeats are entirely possible. The four dice could easily come out 1-3-1-5, having repeated the one. Since order matters and repeats are allowed we know that we are using the exponent principle. Therefore, there are $6^4 = 1296$ possible assignments of new customers to realtors. We will continue in the next box.

COPYRIGHT NOTICE: This is a work in-progress by Prof. Gregory V. Bard, which is intended to be eventually released under the Creative Commons License (specifically agreement # 3 "attribution and non-commercial.") Until such time as the document is completed, however, the author reserves all rights, to ensure that imperfect copies are not widely circulated.

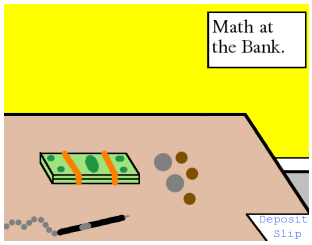
Continuing with the previous box, we should ask how many of those assignments are free of repeats. If we want to compute with repeats disallowed, then we move from the exponent principle to the permutations principle. Then there are $P_{6,4} = 360$ possible outcomes free of repeats.

Third, dividing these two numbers, we get the probability that no two customers get the same realtor. That turns out to be

$$\frac{P_{6,4}}{6^4} = \frac{360}{1296} = \frac{5}{18} = 0.27\overline{7} \approx 27.77\%$$

and therefore the probability that one realtor gets two (or more) customers is given by

$$1 - \frac{P_{6,4}}{6^4} = 1 - \frac{5}{18} = \frac{13}{18} = 0.72\overline{2} \approx 72.22\%$$



The previous box is actually rather striking. Since the customers are assigned to realtors "entirely at random," a naïve manager might think that the system would be perfectly fair. However, the term "fair" is notoriously hard to define. In one sense, the system is fair (if the dice aren't loaded) in that no one realtor has an advantage over another realtor. There can be no favoritism or cronyism.

On the other hand, the system is surprisingly unfair in that it is possible for one realtor to have two customers, while other realtors are sitting idle. Surely, this is not good management.

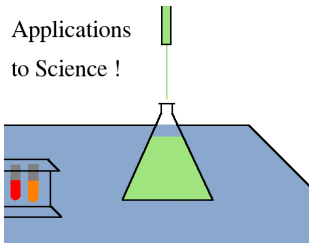
For this reason, managers should not make such assignments at random.

Our previous example shows excellent use of the complement principle. We had a probability of 5/18 of something happening (all customers getting a distinct realtor), therefore the probability of that *not happening* is

$$1 - \frac{5}{18} = \frac{13}{18}$$

Similarly, there were 1296 possible assignments in general. In particular, 360 of them had each customer going to a distinct realtor. Therefore, $1296 - 360 = 936$ assignments would result in some realtor getting two or more customers.

What do we get if we divide 936 by 1296? We get $0.72\overline{2} = 13/18$.



In biomedical studies, it often is useful to assign subjects a secret code. This is particularly useful if privacy-sensitive information, such as venereal diseases, illegal drug use, or HIV status is involved. The secret codes are much safer than using real names, because no matter how good the passwords and computer-security design might be for the main database, it is always possible that private data can be leaked.

The secret codes should be generated randomly (or pseudorandomly) by a computer, dice, or a smart phone app. The reason for this is that if subjects were instead permitted to chose their own code, then they might choose codes like "1111," "45678," or perhaps their birthday. Such codes are easier to guess than random codes.

Now let's look at how permutations and codes can interact.

COPYRIGHT NOTICE: This is a work in-progress by Prof. Gregory V. Bard, which is intended to be eventually released under the Creative Commons License (specifically agreement # 3 "attribution and non-commercial.") Until such time as the document is completed, however, the author reserves all rights, to ensure that imperfect copies are not widely circulated.

For Example :

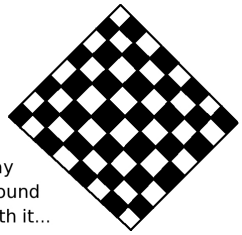
4-2-18

With the previous box in mind, let's consider two students, Charlie and Dave, who are participating in a medical study, and therefore receiving a secret code. The code will be a collection of five letters, without repeats. How many codes are there? If each subject in the study chooses uniformly at random (for example, with the aid of an app on the smart phone), what is the probability that Charlie and Dave accidentally get the same code?

We know there are 26 letters of the alphabet, and we are forbidden to have any repeats. Therefore, $P_{26,5}$ will be the answer, so we just have to calculate that:

$$P_{26,5} = \frac{26!}{21!} = 26 \times 25 \times 24 \times 23 \times 22 = 7,893,600$$

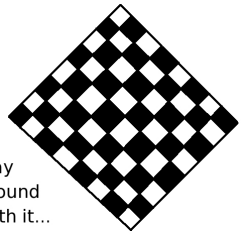
Because the choices are being made uniformly at random, the "equally likely assumption" is justified. With that in mind, the probability of Charlie and Dave getting the same code by accident is 1/7,893,600.



4-2-19

In the situation of the above box, how would the probability change...

- ...if it were a 3-letter code? [Answer: 1/15,600.]
- ...if it were a 4-letter code? [Answer: 1/358,800.]



4-2-20

Continuing with the previous box, suppose each subject were required to select a 5-letter code. However, to avoid confusing `M` and `N`, as well as `E` and `F`, not to mention `O` and `Q`, those six letters are forbidden. How many codes are there now? [Answer: 1,860,480.]

For Example :

4-2-21

Now suppose a subject of the study, Ed, has not been paying attention. He doesn't know about the forbidding of those six letters: `M, N, E, F, O, Q`, and uses an older version of the smartphone app which uses the entire English alphabet. What is the probability that his secret code doesn't have any of those six letters, by coincidence?

Well, there are 1,860,480 codes without those six letters, and 7,893,600 codes overall, as we calculated in the last three boxes.

The probability then is

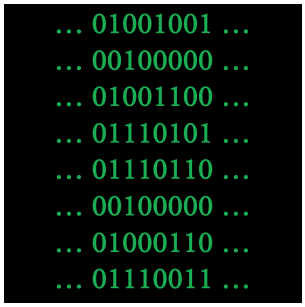
$$\frac{1,860,480}{7,893,600} = 0.235694\cdots \approx 23.56\%$$

We can conclude that Ed will receive, by coincidence, a secret code without those six letters with probability 23.56%. This means that with probability

$$1 - 0.235694 = 0.764305\cdots \approx 76.43\%$$

his secret code will have one of the six forbidden letters. Once again, we have used the complement principle.

COPYRIGHT NOTICE: This is a work in-progress by Prof. Gregory V. Bard, which is intended to be eventually released under the Creative Commons License (specifically agreement # 3 "attribution and non-commercial.") Until such time as the document is completed, however, the author reserves all rights, to ensure that imperfect copies are not widely circulated.

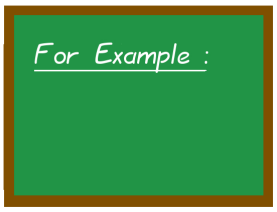


You're probably wondering why, over the last four boxes, I made such an effort to tell you that the passwords were chosen uniformly at random, by an "app" on a smart phone.

In the most serious computer-security situations, users cannot choose their own passwords. Instead, they are given passwords. However, home users and low-to-mid-level employees at most businesses are permitted to choose their own passwords. On the one hand, home users have no or few restrictions. On the other hand, at work or for banking-related websites, often there are several restrictions, forcing the employee/customer to choose a reasonably good password. Nonetheless, in all of these cases, the person gets to choose.

In contrast, for the IT-staff and high-ranking financiers, it is forbidden for them to choose their passwords. These must be randomly generated. The reason for the distinction is that if a particular employee is sloppy, and chooses a moderately-easy-to-guess password, then probably only their own files will be leaked. However, if someone on the IT-staff ends up with their password being compromised, then *all of the secrets of the entire company* might be leaked, which is obviously much worse. Likewise, high-ranking financiers can write checks on behalf of the company, and if their password ends up hacked, huge withdrawals of cash from the company's accounts might occur—or long term competitive plans might be leaked.

Because randomly generated passwords are much harder to guess than user-chosen passwords, the passwords for IT-staff and high-ranking financiers are almost always generated randomly.



4-2-22

Let's say that the Office of Residential Life at Fordham University is concerned about the quality of laundry services in the dorms. Therefore, they randomly select three students from each dorm building, and ask them to take a survey in return for a free gift of some sort. When it comes to O'Hare Hall (which has only 5 floors), the survey manager is surprised that two of the three randomly chosen students are actually from the same floor. Is this a coincidence? Or is the random-number-generating software broken? To measure if he is right to be surprised or not, we should calculate the probability of the students being from different floors, in general. Namely, what is the probability that three students, chosen at random from this building, are all from different floors?

We are told that the computer has chosen the students at random, and we can safely assume, since each floor has an identical floor plan, that the various floors have equal capacity and therefore are equally likely. To select floors for three students, without restriction, is possible $5^3 = 125$ ways. To select floors for three students, without a repeat, is possible $P_{5,3}$ ways. Therefore, the probability of the students being on different floors comes to

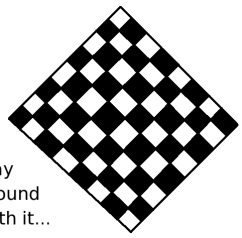
$$\frac{P_{5,3}}{5^3} = \frac{5 \times 4 \times 3}{5 \times 5 \times 5} = \frac{60}{125} = \frac{12}{25} = 0.48$$

In conclusion, the probability of the students being on different floors is only 48%. Therefore, the survey manager should not be surprised at all. There is a 52% chance that the selectees will have a floor in common.



Did you notice that we used the exponent principle and the permutation principle in the same problem? This is how we will typically use the four-entry chart that you saw four boxes ago. In that chart, we indicated that if order matters, and repeats are not allowed, then we should be using the permutation principle. If order matters, and repeats are okay, then we should be using the exponent principle. In the next module, we will fill in more of that table.

COPYRIGHT NOTICE: This is a work in-progress by Prof. Gregory V. Bard, which is intended to be eventually released under the Creative Commons License (specifically agreement # 3 "attribution and non-commercial.") Until such time as the document is completed, however, the author reserves all rights, to ensure that imperfect copies are not widely circulated.

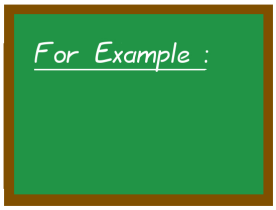


4-2-23

Let's re-examine the problem of the previous box. If it were instead Walsh Hall, which has 13 floors, how would your answers change?

- How many ways is it possible to select three floors with repeats permitted? [Answer: $13^3 = 2197$ ways.]
- How many ways is it possible to select three floors with no repeats? [Answer: $P_{13,3} = 1716$ ways.]
- What is the probability that the students are from different floors? [Answer: $1716/2197 \approx 0.781065\cdots \approx 78.10\%$.]

Even in this case, the survey manager should not be surprised. The probability is $78.1065\cdots\%$ that we get all three floors being different, which means $21.8934\cdots\%$ that we get a repeated floor.



4-2-24

Just to round out the analysis above, we might want to distinguish between the case of two repeated floors, with a third different floor, and the case of three identical floors. In Walsh Hall, there are 13 ways to have all the students from the same floor, because it is a 13-floor dorm. Therefore, the probability of all three students being from the same floor is given by the following:

$$\frac{13}{13^3} = \frac{1}{13^2} = \frac{1}{169} = 0.00591715\cdots \approx 0.59\%$$

Of course, we can just subtract this, and find out the probability that two floors are repeated, and the third is different. We get

$$21.8934\cdots\% - 0.591715\cdots\% = 21.3017\cdots\%$$

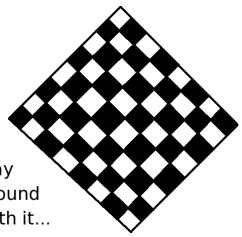
and therefore we can write a probability distribution:

- All three floors distinct: $78.1065\cdots\%$
- Two repeated floors, one different: $21.3017\cdots\%$
- All three floors the same: $0.591715\cdots\%$

Last but not least, we should add those, and make sure that they add to 100%. We get

$$78.1065\% + 21.3017\% + 0.591715\% = 99.999915\%$$

which is accurate to six decimal places! The rest is clearly just rounding error.



4-2-25

The previous problem is a bit hard. Therefore, don't be alarmed if you find this a bit difficult. Let's repeat the above process for O'Hare hall.

- What is the probability that the students are all from the same floor? [Answer: $0.04 = 4\%$ exactly.]
- What is the probability that two students are from the same floor, but the third is distinct? [Answer: $52\% - 4\% = 48\%$ exactly.]

Note, we can check that these add to 100%, as we did in the previous box. We obtain

$$48\% + 48\% + 4\% = 100\%$$

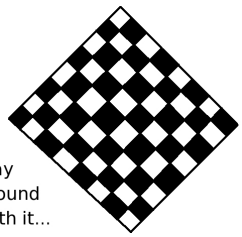
which is awesome.

COPYRIGHT NOTICE: This is a work in-progress by Prof. Gregory V. Bard, which is intended to be eventually released under the Creative Commons License (specifically agreement # 3 "attribution and non-commercial.") Until such time as the document is completed, however, the author reserves all rights, to ensure that imperfect copies are not widely circulated.

The problem of the last several boxes was a bit extensive, so it might be nice to summarize the data in a chart. To make the chart easier to read, I'm going to round down to the next basis point.

Situation	Probability O'Hare	Probability Walsh
All from the same floor	4%	0.59%
Two the same, one different	48%	21.30%
All three from distinct floors	48%	78.10%

As you can see, the survey manager has good cause to be worried that the random number generator is broken if all three students are from the same floor in Walsh Hall, because the probability of that is under 1%. In O'Hare hall, there really isn't cause to be *worried*, because the probability is 4%, but it might be a good idea to double check by generating another threesome of students as a test case, even if those three additional students will not, in fact, be surveyed. If it happens a second time then he should be worried. In fact, we'll learn in a later module that the probability of such a repeat would be $0.04^2 = 0.0016$.

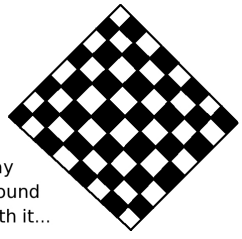


4-2-26

Now let's return to the train problem that we first saw on Page 254, in the module "A Formal Introduction to Probability." Let's suppose that 3 trains are arriving simultaneously. The station still has 20 tracks, and the computer is broken again. The drivers must choose a random track. If the drivers all choose different tracks then they will be safe. However, if even two drivers pick the same track, then there will be a crash.

- Hint: Since Train 101 arriving on Track 5 and Train 302 arriving on Track 6 is a very different thing than Train 302 arriving on Track 5 and Train 101 arriving on Track 6, we can conclude that order matters. However, this is not obvious, which is why I'm giving you a hint.
- In how many ways can 3 distinct tracks be chosen from the 20 tracks? [Answer: 6840.]
- In how many ways can 3 tracks be chosen, in general, from the 20 tracks? [Answer: 8000.]
- What is the probability that the chosen tracks are all different? [Answer: $0.855 = 85.5\%$.]
- What is the probability of a collision? [Answer: $0.145 = 14.5\%$.]

As you can see, we're getting a lot of use out of the complement principle.

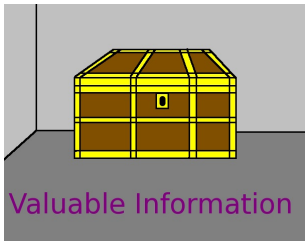


4-2-27

Now let's repeat the problem of the previous box. However, imagine Grand Central Station, and the first trains of the morning are arriving into the empty train station. There are 6 arrivals, and they have to choose from 67 tracks. In the event of a computer failure, forcing them to choose randomly, what is the probability of two (or more) trains choosing the same track?

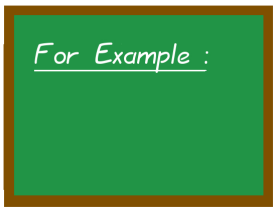
[Answer: $0.205680\cdots \approx 20.56\%$.]

COPYRIGHT NOTICE: This is a work in-progress by Prof. Gregory V. Bard, which is intended to be eventually released under the Creative Commons License (specifically agreement # 3 "attribution and non-commercial.") Until such time as the document is completed, however, the author reserves all rights, to ensure that imperfect copies are not widely circulated.



Since permutations and factorials are appearing in the same module, you might wonder if there is a relationship between the two concepts. Actually, there is a formula which relates these two ideas. The formula is

$$P_{n,x} = \frac{n!}{(n-x)!}$$



The proof of the formula of the previous box can be found in the next box. Before we examine that proof, here are some quick examples to verify that the formula works.

$$\begin{array}{lllllllllll}
P_{7,3} & = & (7)(6)(5) & = & 210 & \text{meanwhile} & \frac{7!}{(7-3)!} & = & \frac{7!}{4!} & = & \frac{5040}{24} & = & 210 \\
P_{7,4} & = & (7)(6)(5)(4) & = & 840 & \text{meanwhile} & \frac{7!}{(7-4)!} & = & \frac{7!}{3!} & = & \frac{5040}{6} & = & 840 \\
P_{7,5} & = & (7)(6)(5)(4)(3) & = & 2520 & \text{meanwhile} & \frac{7!}{(7-5)!} & = & \frac{7!}{2!} & = & \frac{5040}{2} & = & 2520
\end{array}$$

4-2-28



It is important to realize that you cannot cancel the factorial symbols. Some students imagine that

$$\frac{7!}{3!} = \frac{7\cancel{!}}{3\cancel{!}} = \frac{7}{3} \qquad \leftarrow \text{ WRONG!}$$

but this is not true.

If you like, you can check the following computation with your calculator.

$$\frac{7!}{3!} = \frac{5040}{6} = 840 \text{ as compared to } \frac{7}{3} = 2.3\overline{3}$$



This box, and the two following, are only for those students who are curious about how one might prove the formula relating the permutations principle and the factorial principle.

We're going to start with the formula $n!/(n-c)!$, and compute with it a bit, and show that this equals $P_{n,x}$.

$$\begin{aligned}
\frac{n!}{(n-x)!} &= \frac{n(n-1)(n-2)\cdots(3)(2)1}{(n-x)(n-x-1)(n-x-2)\cdots(3)(2)1} \\
&= \frac{[n(n-1)(n-2)\cdots(n-x+2)(n-x+1)]\,[(n-x)(n-x-1)(n-x-2)\cdots(3)(2)1]}{(n-x)(n-x-1)(n-x-2)\cdots(3)(2)1} \\
&= \frac{[n(n-1)(n-2)\cdots(n-x+2)(n-x+1)]\,[\cancel{(n-x)}\cancel{(n-x-1)}\cancel{(n-x-2)}\cdots\cancel{(3)}\cancel{(2)}\cancel{(1)}]}{\cancel{(n-x)}\cancel{(n-x-1)}\cancel{(n-x-2)}\cdots\cancel{(3)}\cancel{(2)}\cancel{(1)}} \\
&= n(n-1)(n-2)\cdots(n-x+2)(n-x+1) \\
&= P_{n,x}
\end{aligned}$$

As you can see, the right-hand bracketed product is completely cancelled out by the denominator, leaving only our desired terms in the numerator, and a mere "one" in the denominator.

The proof in this box is a bit symbol-heavy, and some readers might find it confusing. Therefore, let's look at a specific example in the next box, with numbers instead of variables.

COPYRIGHT NOTICE: This is a work in-progress by Prof. Gregory V. Bard, which is intended to be eventually released under the Creative Commons License (specifically agreement # 3 "attribution and non-commercial.") Until such time as the document is completed, however, the author reserves all rights, to ensure that imperfect copies are not widely circulated.



Because the proof of the previous box is a bit messy, let's take a moment to consider the special case of $P_{11,6}$. The goal is to show that the two methods produce the same answer. The first method is computing 11!/6! and the second method is starting at 11, counting downward until you have 6 numbers written down, and then multiplying those 6 numbers together. Let's see what happens:

$$\begin{aligned}\frac{11!}{5!} &= \frac{(11)(10)(9)(8)(7)(6)(5)(4)(3)(2)(1)}{(5)(4)(3)(2)(1)} \\ &= \frac{(11)(10)(9)(8)(7)(6)\cancel{(5)(4)(3)(2)(1)}}{\cancel{(5)(4)(3)(2)(1)}} \\ &= \frac{(11)(10)(9)(8)(7)(6)}{1} \\ &= \underbrace{(11)(10)(9)(8)(7)(6)}_{6 \text{ numbers}} = P_{11,6}\end{aligned}$$

As you can see, the 5! in the denominator of the original fraction destroys the last 5 entries of the 11!. This leaves six terms remaining, and they are the six terms that we'd actually want when computing $P_{11,6}$. That's why we get the same answer.



Personally, I view the multiplication principle as leading naturally into the permutation formula. I would tend to calculate $P_{11,5}$ using the multiplications

$$P_{11,5} = 11 \times 10 \times 9 \times 8 \times 7 = 55,440$$

and not with

$$P_{11,5} = \frac{11!}{(11-5)!} = \frac{11!}{6!} = \frac{39,916,800}{720} = 55,440$$

which just seems to be more work.

Breaking the calculation into two factorials is (to me) a numerical "coincidence," even if it can be a very useful aid to prove theorems about permutation problems during an advanced course in pure mathematics.

In all fairness, I have to point out that nearly all students simply use the "permutations" button on their calculator to compute something like $P_{11,5}$.



There is a hazard that I should point out. It is dangerous to overuse the factorial operator. Consider the number of ways to draw 3 cards from a deck of 52 cards, when order matters. We have 52 choices for the first card, 51 choices for the second card, and 50 choices for the third card. Therefore, there are

$$P_{52,3} = 52 \times 51 \times 50 = 132,600$$

possibilities. This is an easy calculation for us at this point.

However, if you were unwisely to depend on the factorial operator, you'd be calculating $P_{52,3} = 52! \div 49!$ and those are huge numbers. This is usually a problem. As it turns out

$$52! \approx 8.06581751 \cdots 10^{67}$$

which I'm sure we can all agree is a huge number. There are two possibilities here for how your calculator will react. It might object with an "overflow" related error; alternatively, it might try to represent that number approximately.

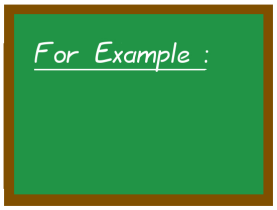
COPYRIGHT NOTICE: This is a work in-progress by Prof. Gregory V. Bard, which is intended to be eventually released under the Creative Commons License (specifically agreement # 3 "attribution and non-commercial.") Until such time as the document is completed, however, the author reserves all rights, to ensure that imperfect copies are not widely circulated.

The reason that your calculator cannot represent 52! exactly is because

$$52! = 80,658,175,170,943,878,571,660,636,856,403,766,975,289,505,440,883,277,824,000,000,000,000$$

is the exact integer, and that has 57 digits. Your hand-calculator probably can only display 8, 10, or 12 significant digits.

You've probably heard about the next problem before, because it is very famous. It is called "the birthday paradox." Most every textbook about probability has this problem—so this textbook has to have it too.



4-2-29

Suppose there are 40 people in a room. They are asked about their birthdays. Specifically, they are asked to state the month and the day, but not the year. What is the chance that two (or more) people in the room share a birthday?

At first glance, a pair of people with the same birthday seems unlikely. There are 365 days of the year, or 366 if you include leap years, and we only have 40 people in the room. Calculation will reveal a surprising answer. First, let's compute the number of ways that we can have 40 birthdays, without repeats. Here, order matters. That's because saying that Bob has a birthday on Dec 4th, while Alice has a birthday on Jan 8th, is different from saying that Alice has a birthday on Dec 4th, while Bob has a birthday on Jan 8th. Since order matters, and repeats are not allowed, we know that we should use the permutations principle. Let's ignore leap years, so that there are $P_{365,40}$ ways for birthdays to be assigned without repeats.

Second, let's compute how many ways that we can have 40 birthdays, but with repeats being allowed. Order matters as before, but because repeats are allowed we must use the exponent principle. Therefore, there are 365^{40} ways for birthdays to be assigned, allowing for possible repeats.

These numbers are so large that they probably overflow your calculator. We'll talk now about how to handle that.



The numbers that we need from the previous box, $P_{365,40}$ and 365^{40} are so huge that we cannot compute them on most calculators. Either the calculator will return an error, or it will return a decimal approximation. However, the computer algebra system Sage is very good with large numbers. Using Sage, I found the values

$$P_{365,40} = 337,455,476,433,028,093,130,395,325,716,779,597,960,373,443,552,267,062,349,764,611,493,458,502,539,556,028,031,935,971,328,000,000,000$$

as well as

$$365^{40} = 3,102,519,917,525,622,220,066,362,032,333,600,168,597,425,643,900,453,198,929,195,853,371,717,515,992,713,742,889,463,901,519,775,390,625$$

These numbers are so large, that they do not really have a meaning. Also, they will cause problems for both calculators and ordinary mathematical software (like MS-Excel) because of phenomena called "overflow" and "underflow." While I am tempted to discuss overflow and underflow with you, that conversation would be long and take us too far astray.

In any case, there is a better way to look at this problem.

COPYRIGHT NOTICE: This is a work in-progress by Prof. Gregory V. Bard, which is intended to be eventually released under the Creative Commons License (specifically agreement # 3 "attribution and non-commercial.") Until such time as the document is completed, however, the author reserves all rights, to ensure that imperfect copies are not widely circulated.



An alternative approach is to look at the computation of the previous two boxes as an unfinished long multiplication. We have

$$P_{365,40} = \underbrace{(365)(364)(363)(362)(361)\cdots(328)(327)(326)}_{40 \text{ numbers}}$$

in contrast with

$$365^{40} = \underbrace{(365)(365)(365)(365)(365)\cdots(365)(365)(365)}_{40 \text{ numbers}}$$

Now if we divide them, we get

$$\frac{P_{365,40}}{365^{40}} = \frac{(365)(364)(363)(362)(361)\cdots(328)(327)(326)}{(365)(365)(365)(365)(365)\cdots(365)(365)(365)}$$

which can be thought of as

$$\frac{P_{365,40}}{365^{40}} = \frac{365}{365}\cdot\frac{364}{365}\cdot\frac{363}{365}\cdot\frac{362}{365}\cdot\frac{361}{365}\cdot\cdots\cdot\frac{328}{365}\cdot\frac{327}{365}\cdot\frac{326}{365}$$

We will continue in the next box.



Continuing with the previous box, that last equation, with a collection of fractions being multiplied, has an advantage. Namely, all of those fractions are between 0 and 1. Therefore, there is no possibility of the calculator having an overflow in that case. If we multiply those fractions out (or if we ask a computer to do it for us) then we will have an accurate answer.

Similarly, to do this on a hand calculator, you should type

$$365 \div 365 \times 364 \div 365 \times 363 \div 365 \times 362 \div 365 \times 361 \div 365 \times \cdots$$

which definitely will not cause an overflow or underflow. The final answer is given in the next box.

Returning to our previous example, using Sage I obtained the final answer:

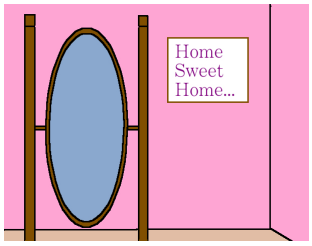
$$0.108768190182051\cdots \approx 10.87\%$$

This means that there is a $0.108768\cdots$ probability that all 40 people have a different birthday. Another way to say that is there is a

$$1 - 0.108768\cdots = 0.891231\cdots \approx 89.12\%$$

probability that there is at least one pair of people that shares a birthday.

In summary, it is a really good bet that a pair of people share a birthday, even though there are only 40 people in the room, and there are 365 days in the year.



A Pause for Reflection...

As we often do, we made the "equally likely assumption" in the previous example. Is this fair and accurate? Do we have any reason to believe that someone's birthday being one day of the year is as likely as another day of the year? Are all the days of the year equally likely?

Take a moment to think about this, and then I'll share my thoughts with you in the next few boxes.

COPYRIGHT NOTICE: This is a work in-progress by Prof. Gregory V. Bard, which is intended to be eventually released under the Creative Commons License (specifically agreement # 3 "attribution and non-commercial.") Until such time as the document is completed, however, the author reserves all rights, to ensure that imperfect copies are not widely circulated.

As it comes to pass, when people study the distribution of birthdays, several phenomenon are discovered. Birthdays are not uniformly distributed.

- Some months are longer than others. For this reason, we expect January (with 31 days) more often than we do February (with 28 or 29 days, depending on leap years). However, our previous box worked day-by-day, not month-by-month, so we have steered clear of this problem.
- Many children (including myself) were born via a Caesarian Section. Since doctors do not like to perform surgery on the weekends, this tends to slightly favor the work week, and slightly disfavor the weekends, for the date of birth—even though Caesarian Sections are not all that common. However, year after year, the day of the week will change for a particular date. For example, February 14th, was a Monday, a Tuesday, a Thursday, a Friday, and a Saturday in 2011, 2012, 2013, 2014, and 2015. Therefore, this effect vanishes when looking at a large number of years.
- However, if the "room" of the previous example were for a convention of Scorpios, then we'd have a problem. Scorpios are born between late October and late November. Likewise, a convention for any other sign of the Zodiac will cause a problem. Moreover, if it were a convention of twins, then we would have a problem. On the other hand, these are extremely fringe situations and therefore we should not lose sleep over this.

I'd like to reflect on two other thoughts in the next box.

Continuing with the previous box, we are reflecting on aspects of our assumption that birthdays are uniformly distributed.

- It turns out that statisticians have studied the distribution of birthdates by month. It turns out that July, August, September, and to a much lesser extent, October, have significantly more births than expected, even after adjusting for the fact that some months have 31 days and others have 30 days. At first, this might be puzzling.
- Further thought will reveal the cause. If we subtract nine months, we will see that the July through October birth months imply conception occurring during the months of October through January. There is a cluster of holidays during the late November through January period, and it is customary for people to take vacation time during these months. Naturally, this will affect the rate of conception, as it is much easier to conceive a child during vacation or when at home, than while at the office.
- The previous bullet explains the increase in births for late August through October, but not for July or early August. If we subtract nine months from July, we get October, and early August results in early November. While October and early November do not have a concentration of holidays, we should remember that some babies are born prematurely, before the completion of nine months. That possibility of premature birth, along with the phenomenon of the previous bullet, might explain an increase in births during July and early August.



You can read more about the distribution of birthdays in the article "An Analysis of the Distribution of Birthdays in a Calendar Year" by the actuary Roy Murphy.
`http://www.panix.com/~murphy/bday.html`

In summary, it is really good to think about whether or not the "equally likely" assumption holds. Practicing this thought process is useful, especially if you're ever going to use probability or statistics in a research internship or career.

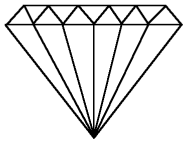
In any case, after careful reflection, we can conclude that assuming birthdays are evenly distributed is not an entirely justified assumption. It might be okay for a rough approximation, but it does not qualify as a sound mathematical practice.

COPYRIGHT NOTICE: This is a work in-progress by Prof. Gregory V. Bard, which is intended to be eventually released under the Creative Commons License (specifically agreement # 3 "attribution and non-commercial.") Until such time as the document is completed, however, the author reserves all rights, to ensure that imperfect copies are not widely circulated.

You're probably wondering why I'm including this problem on the birthdays, especially because the numbers are so difficult to calculate without a computer. Clearly, the problems are not well-suited for examinations (though you should ask your instructor about that).

There are two reasons that I've included the birthday problem. The minor reason is that it is a famous problem, and therefore it would be bad to leave it out. The major reason is that the same phenomenon comes up for digital signatures, in cryptography, the science of codes.

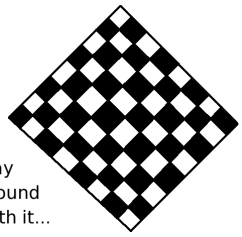
Hard but Valuable!



The rest of this module contains some advanced material that not all instructors will want to discuss. In particular,

- The relationship of the birthday paradox, to cryptography.
- A cool story about permutations, the mafia, and Sicilian local elections.
- A discussion on the value of 0!, zero factorial.
- A trick for solving cool equations involving factorials.

You might want to check with your instructor, to see if these topics are relevant or not.



Play Around With it...

4-2-30

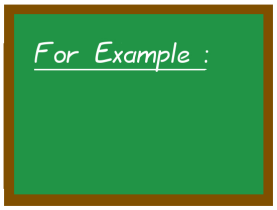
One of the key ingredients of a digital signature is called a "digest." It is a code that represents the document being signed. Let's suppose that the digital signature system used by a corporation will output a 6-character digest, where each character is either a capital letter or a digit. For example, valid digests can look like `U8RL4M` or `7OLM5M`, and so forth. Just to be clear, order matters and repeats are allowed.

- How many possible digests are there? [Answer: 2,176,782,336.]
- If the designers were to add the option for lower-case letters, in addition to capital letters and numerals, then how many possible digests are there? [Answer: 56,800,235,584.]



Let's look at the two possible schemes in the previous box, for generating digests for digital signatures. It would take a very long time to explain, but if two documents ever have the same digest, then there are all sorts of cryptographic attacks that hackers can exploit. Let's not go into detail about how hackers can take advantage of that, but let's summarize by saying that hackers can generate forgeries whenever they have two messages that end up with the same digest.

COPYRIGHT NOTICE: This is a work in-progress by Prof. Gregory V. Bard, which is intended to be eventually released under the Creative Commons License (specifically agreement # 3 "attribution and non-commercial.") Until such time as the document is completed, however, the author reserves all rights, to ensure that imperfect copies are not widely circulated.



4-2-31

Continuing with the previous checkerboard box, let's compute the probability that at least one pair of documents will have the same digest Perhaps the CIO (Chief Information Officer) says that they anticipate needing to sign 20,000 documents over the lifetime of the system. Since there are 20,000 documents which must be assigned a digest from among 2,176,782,336 possible digests, then that's analogous to 40 people who must be assigned a birthday from among 365 birthdays.

We would have to compute

$$\frac{P_{2,176,782,336;\ 20,000}}{2,176,782,336^{20,000}} = 0.912219\cdots \approx 91.22\%$$

Alternatively, if we allow the lowercase letters in addition to the capital letters and numerals, then we have 56,800,235,584 possible digests. We would have to calculate

$$\frac{P_{56,800,235,584;\ 20,000}}{56,800,235,584^{20,000}} = 0.996486\cdots \approx 99.64\%$$

Your calculator probably cannot handle these numbers. Let's not worry about that now, but let's try to understand the consequences of these numbers. We'll continue our analysis in the next box.



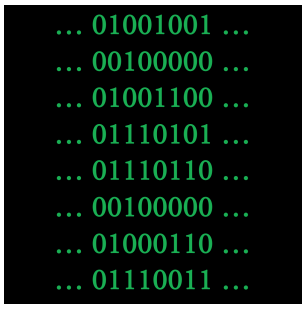
Continuing with the previous box, the cryptographic attack in the case of capital-letters and numerals will be possible with probability

$$1 - 0.912219\cdots = 0.0877804\cdots \approx 8.77\%$$

whereas if we include the lower-case letters, the cryptographic attack will be possible only with probability

$$1 - 0.996486\cdots = 0.00351357\cdots \approx 0.35\%$$

which is an extraordinarily different probability. As you can see, we've gotten a lot of mileage out of the complement principle.



What ideas are we supposed to take away from the above discussion of digital signatures? First, that the numbers involved in cryptography tend to be very large. It is not uncommon for me to encounter 600-digit numbers while teaching *Math-380: Cryptography* at the University of Wisconsin—Stout.

Second, and far more important, we should realize that every detail matters. Even very tiny and seemingly insignificant details about a digital signature system, such as whether or not to permit lower-case letters in the digest, can have a major impact on the probability of a successful cryptographic attack on the system. In turn, that can have catastrophic business consequences for the company involved.

In actual cases used in real e-commerce situations, all of the numbers involved would be much larger, and the percentages would be much smaller. However, those numbers would be so large (hundreds of digits), and those probabilities so small, that typing them on a printed page would be very awkward.

Typically, digital signatures are a bit string of length 128, 160, 192, or 256 bits. (Aren't you glad that we talked about bit strings on Page 407?)

COPYRIGHT NOTICE: This is a work in-progress by Prof. Gregory V. Bard, which is intended to be eventually released under the Creative Commons License (specifically agreement # 3 "attribution and non-commercial.") Until such time as the document is completed, however, the author reserves all rights, to ensure that imperfect copies are not widely circulated.



We return now to the example about Charlie, Dave, and Ed (see Page 435), who were participating in a medical study. If we use the 5-letter codes, where six particular letters were prohibited, then we have 1,860,480 possible codes. Suppose there are 300 people in the study. The chance that each subject gets a distinct code is given by

$$\frac{P_{1,860,480;\ 300}}{1,860,480^{300}} \approx 0.976180\cdots$$

and the chance that two subjects get the same code is given by

$$1 - 0.976180\cdots = 0.0238196\cdots \approx 2.38\%$$

which is a bit too high, but not terrible.

We will contrast this with using a 26-letter alphabet, in the next box.



Continuing with the previous box, if we alternatively use all 26 letters of the English alphabet, then there are 7,893,600 possible codes. The chance that each subject gets a distinct code is given by

$$\frac{P_{7,893,600;\ 300}}{7,893,600^{300}} \approx 0.994334\cdots$$

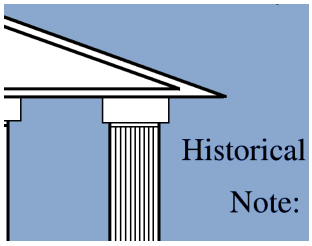
and the chance that two subjects get the same code is given by

$$1 - 0.994334\cdots = 0.00566578\cdots \approx 0.56\%$$

which is probably much more acceptable.

In the next few boxes, we're going to explore how a very transparent system of voting was circumvented in Victorian-era Sicily, by use of the factorial principle.

The story told in the next two boxes is one of my favorite mathematical stories of all time. It was told to me when I worked for the National Security Agency (NSA), and the person who told me was one of the nation's foremost experts on organized crime in general, and the Sicilian Mafia in particular. He had many successes in damaging the mafia. So many successes, in fact, that he had to spend a large number of years in the witness protection program.



In late 19th-century Sicily, the very local branch of government was a town or village council. Often, there would be about two to three thousand voters in a typical small town or large village. A typical size for the council would be perhaps seven positions, and ordinarily there would be about twenty candidates. Each voter would write the names of up to seven candidates on a blank ballot sheet. Whichever seven candidates had received the largest number of votes would constitute the new town council.

The Italian central government in Rome had been concerned about Mafia interference in elections, and had established tight controls. Only registered living voters could vote. Ballot boxes were closely watched, so ballot-stuffing was infeasible. Furthermore, all of the ballots (without the voter's name on them) would be publicly posted so that anyone could inspect them. This way, anyone could inspect the ballots to verify that the outcome was correct.

Does this sound like a good plan? To the extremely naïve eye it might seem so, but not to someone who understands combinatorics!

COPYRIGHT NOTICE: This is a work in-progress by Prof. Gregory V. Bard, which is intended to be eventually released under the Creative Commons License (specifically agreement # 3 "attribution and non-commercial.") Until such time as the document is completed, however, the author reserves all rights, to ensure that imperfect copies are not widely circulated.



Continuing the story in the previous box, the Mafia would send a Mafioso to each household, and he would tell the householder whom to vote for—namely, seven candidates who were either actual Mafiosi or oath-bound to the Mafia.

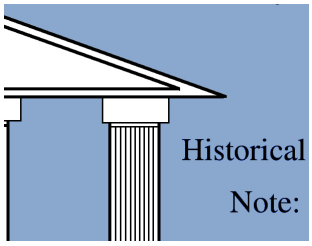
However, the visiting Mafioso would also tell them precisely and exactly in which order they must write the names in. With seven names, there are $7! = 5040$ orderings possible, via the factorial principle. Usually, these villages did not have 5000 voters in them—and therefore great care would be taken to ensure that each household be instructed to use a distinct ordering. Since the number of voters was smaller than 5040, this was always possible.



Still continuing the story from the previous box, the Mafiosi would be careful to inspect the ballots where they had been posted in the town hall. If one of the assigned orderings was not found among the posted ballots, then the Mafia would know that the household which was assigned that ordering had been disobedient.

Naturally, most households would be obedient out of fear, yet some would disobey. Since each household was issued a different ordering, it would be easy to identify precisely which households disobeyed. Of course, that disobedient voter would be murdered, and often their children and other family members would be murdered also.

As a result, voters tended to be highly obedient, and Mafia control over the local governments was maintained, much to Rome's frustration. In turn, the town's public prosecutor, chief of police, and other important functionaries would be selected by the town council, ensuring that those positions were held by Mafiosi (or alternatively, men who had become oath-bound to the Mafia).

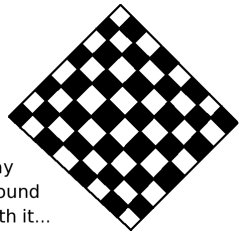


When I had first heard the above story, I was skeptical. How would mere criminals know enough mathematics to be able to use the factorial principle? However, I realized that running casinos, lotteries, and other forms of gambling, such as horse races, are mainstays of the Sicilian Mafia.

As you will learn throughout this chapter, gambling and probability are highly intertwined, and therefore anyone running a casino would have access to someone who was educated in probability.

If you'd like to read more about the NSA's accomplishments in destroying the power of the Sicilian Mafia, I recommend the book *Body of Secrets: Anatomy of the Ultra-Secret National Security Agency: from the Cold War through the Dawn of a New Century*, by James Bamford, published by Doubleday in 2001, and reprinted by Anchor Books in 2002. It is a general history of the NSA from 1952 until roughly the year 1999.

Now we're going to talk about why $0! = 1$, over the next few boxes.

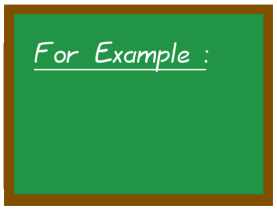


4-2-32

Imagine you're the manager of a small Cable TV station (or a digital streaming website, if you prefer), and you're trying to decide what to do with the 1 AM to 4 AM time slots—which are an hour long. You decide to rerun three hour-long old shows. Suppose there are five possible TV-shows that your cable channel can choose to rerun. Of course, the order matters, as few people are watching at 4 AM, but slightly more are watching at 2 AM.

- How many three-show schedules can be built out of these five shows? [Answer: 60.]
- If you also free-up the 4 AM to 5 AM time slot, and must now compute a schedule for 1 AM to 5 AM, then how many four-show schedules are there? [Answer: 120.]

COPYRIGHT NOTICE: This is a work in-progress by Prof. Gregory V. Bard, which is intended to be eventually released under the Creative Commons License (specifically agreement # 3 "attribution and non-commercial.") Until such time as the document is completed, however, the author reserves all rights, to ensure that imperfect copies are not widely circulated.



4-2-33

Now suppose, continuing from the previous box, that you get the midnight to 1 AM time slot also. Scheduling from midnight to 5 AM, you now have five time slots, and five shows. So clearly you have to run each of the shows, it is just a question of the ordering. If you do this with the permutation formula, you get

$$P_{5,5} = \frac{5!}{0!} = \frac{120}{???}$$

but what is 0! going to be?

Imagine that you ask a friend for help. Your friend then reminds you that one can calculate the number of ways to order 5 objects using the factorial principle directly, avoiding permutations entirely. That calculation would be $5! = 120$. Therefore, it is clear that 0! must equal 1, in order to make both answers come out correctly.



For most readers, the previous box is a sufficient argument to cause one to believe that $0! = 1$. However, here is another argument, based on the idea that

$$n! = n \times (n-1)!$$

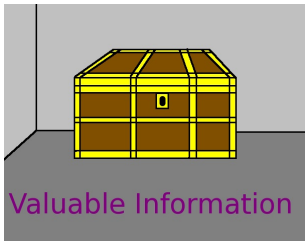
which we should first prove:

$$\begin{aligned} n! &= (n)(n-1)(n-2)(n-3)(n-4)\cdots(3)(2)(1) \\ n! &= (n) \times \underbrace{(n-1)(n-2)(n-3)(n-4)\cdots(3)(2)(1)} \\ n! &= n \times (n-1)! \end{aligned}$$

Using that, let us observe the pattern

$$\begin{array}{rcrcrcl} 7! & = & 5040 & = & (7)(720) & = & 7 \times 6! \\ 6! & = & 720 & = & (6)(120) & = & 6 \times 5! \\ 5! & = & 120 & = & (5)(24) & = & 5 \times 4! \\ 4! & = & 24 & = & (4)(6) & = & 4 \times 3! \\ 3! & = & 6 & = & (3)(2) & = & 3 \times 2! \\ 2! & = & 2 & = & (2)(1) & = & 2 \times 1! \\ 1! & = & 1 & = & (1)(0!) & = & 1 \times 0! \end{array}$$

which works so well until the last line. The only way to make the last line work as well as all the others is to define $0! = 1$.



The factorial of zero equals one.

$$0! = 1$$

COPYRIGHT NOTICE: This is a work in-progress by Prof. Gregory V. Bard, which is intended to be eventually released under the Creative Commons License (specifically agreement # 3 "attribution and non-commercial.") Until such time as the document is completed, however, the author reserves all rights, to ensure that imperfect copies are not widely circulated.



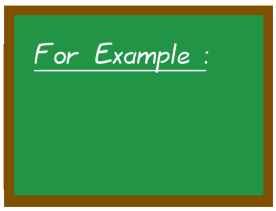
What we really did over the last three boxes, is expand our definition for the factorial, which made sense for the positive integers, $\{1, 2, 3, 4, \ldots\}$. We expanded that definition, to include zero, so that it will make sense for what pure mathematicians sometimes call "the natural numbers," but what most newer textbooks call "the non-negative integers." Of course, I'm talking about $\{0, 1, 2, 3, \ldots\}$. (We discussed why the phrase "the natural numbers" is being phased out back on Page 163 of the module "Intermediate Set Theory and Irrationality.")

However, such an expansion of the definition is only good if we keep the important properties. Those properties include

$$n! = n \cdot (n-1)!$$

as well as a desire that $P_{5,5} = 120$, in the television scheduling example. In general, we want $n! = P_{n,n}$, and this requires us to accept that $0! = 1$.

Now that the issue of 0! is resolved, I'd like to show you how to solve some bizarre equations involving factorials.



4-2-34

Let's try to solve the following equation:

$$(n+2)! = 56n!$$

The trick here is to realize that

$$\frac{(n+2)!}{n!} = P_{n+2,2} = (n+2)(n+1)$$

or alternatively you might think of it this way:

$$\begin{aligned} \frac{(n+2)!}{n!} &= \frac{(n+2)(n+1)(n)(n-1)(n-2)\cdots}{(n)(n-1)(n-2)\cdots} \\ &= \frac{(n+2)(n+1)\cancel{(n)}\cancel{(n-1)}\cancel{(n-2)}\cdots}{\cancel{(n)}\cancel{(n-1)}\cancel{(n-2)}\cdots} \\ &= (n+2)(n+1) \end{aligned}$$

With this in mind, we can solve the equation, in the next box.

Continuing with the previous box,

$$\begin{aligned} (n+2)! &= 56n! \\ \frac{(n+2)!}{n!} &= 56 \\ P_{n+2,2} &= 56 \\ (n+2)(n+1) &= 56 \\ (n+2)(n+1) &= (8)(7) \\ n &= 6 \end{aligned}$$

COPYRIGHT NOTICE: This is a work in-progress by Prof. Gregory V. Bard, which is intended to be eventually released under the Creative Commons License (specifically agreement # 3 "attribution and non-commercial.") Until such time as the document is completed, however, the author reserves all rights, to ensure that imperfect copies are not widely circulated.

For Example :

4-2-35

To practice some more with this new technique, let's try to solve the following equation:

$$(n+1)! = 1320(n-2)!$$

The trick here is to realize that

$$\frac{(n+1)!}{(n-2)!} = \frac{(n+1)!}{(n+1-3)!} = P_{n+1,3} = (n+1)(n)(n-1)$$

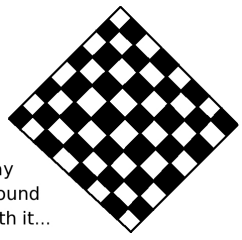
or alternatively you might think of it this way:

$$\begin{aligned} \frac{(n+1)!}{(n-2)!} &= \frac{(n+1)(n)(n-1)(n-2)(n-3)\cdots}{(n-2)(n-3)(n-4)(n-5)\cdots} \\ &= \frac{(n+1)(n)(n-1)\cancel{(n-2)(n-3)}\cdots}{\cancel{(n-2)(n-3)(n-4)(n-5)}\cdots} \\ &= (n+1)(n)(n-1) \end{aligned}$$

With this in mind, we can solve the equation, in the next box.

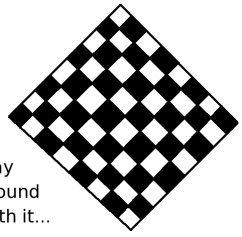
Continuing with the previous box,

$$\begin{aligned} (n+1)! &= 1320(n-2)! \\ \frac{(n+1)!}{(n-2)!} &= 1320 \\ P_{n+1,3} &= 1320 \\ (n+1)(n)(n-1) &= 1320 \\ (n+1)(n)(n-1) &= (12)(11)(10) \\ n &= 11 \end{aligned}$$



4-2-36

- First, as a warm up, what is
$$\frac{(n+4)!}{(n+2)!}$$
in the notation of $P_{n,x}$? [Answer: it is $P_{n+4,2}$.]
- Second, using that, solve the equation $(n+4)! = 870(n+2)!$ for n. [Answer: $n = 26$.]



4-2-37

- First, as a warm up, what is
$$\frac{(n+6)!}{(n+4)!}$$
in the notation of $P_{n,x}$? [Answer: it is $P_{n+6,2}$.]
- Second, using that, solve the equation $(n+6)! = 182(n+4)!$ for n. [Answer: $n = 8$.]

COPYRIGHT NOTICE: This is a work in-progress by Prof. Gregory V. Bard, which is intended to be eventually released under the Creative Commons License (specifically agreement # 3 "attribution and non-commercial.") Until such time as the document is completed, however, the author reserves all rights, to ensure that imperfect copies are not widely circulated.



Here is the explanation for the very hard reflection box that was given earlier on Page 430. You were asked to explain why it is true that

$$n^n > n! > 2^n$$

for all $n > 3$. This is an extremely hard question so please don't be disappointed if you didn't figure it out.

We will explain this by way of using $n = 10$. The three items being compared are

$$\begin{array}{rcccccccccccccccccc} 10^{10} & = & 10 & \times & 10 & \times & 10 & \times & 10 & \times & 10 & \times & 10 & \times & 10 & \times & 10 & \times & 10 \\ 10! & = & 10 & \times & 9 & \times & 8 & \times & 7 & \times & 6 & \times & 5 & \times & 4 & \times & 3 & \times & 2 \\ 2^{10} & = & 2 & \times & 2 & \times & 2 & \times & 2 & \times & 2 & \times & 2 & \times & 2 & \times & 2 & \times & 2 \end{array}$$

Let's compare 10^{10} versus $10!$. Surely in each case, the term for 10^{10} is $\geq$ the term for $10!$. Moreover, only once are the terms equal. Therefore, we can be sure that $10^{10} > 10!$. However, we used no special mathematical property of 10 to show this. The same argument can work for 9 or 11 as well. It turns out that for all integers $n > 2$, we have $n^n > n!$. This is a reasonably logical argument, by the way, but it does not constitute a mathematical proof.

Next, let's compare $10!$ versus 2^{10}. Again in each case, the term for $10!$ is $\geq$ the term for 2^{10}. Moreover, only once are the terms equal. Therefore, we can be sure that $10! > 2^{10}$. As before, we used no special mathematical property of 10 to show this. The same argument can work for 9 or 11 as well. It turns out that for all integers $n > 3$, we have $n! > 2^n$.



Okay, so the previous box has convinced us that

$$n^n > n! > 2^n$$

for all integers $n > 3$, but why should this matter?

The reason that this property of $n^n > n! > 2^n$ is thought of as "cool" is because of the following idea. For any $n > 1$, we have

$$n < n^2 < n^3 < n^4 < n^5 < n^6 < n^7 < n^8 < n^9 < \cdots$$

from which arises the idea of polynomials "of the first degree," "of the second degree," "of the third degree," and so on. The degree of the polynomial is indicating an "order of growth." A higher degree polynomial grows faster in value, as n increases, than a lower degree polynomial.



In fact, this property holds for more complicated polynomials. Consider

$$f(x) = 123,456,789x^2 \qquad \text{and} \qquad g(x) = 0.01x^3$$

Surely, for small x, such as 1, 2, 3, ..., 10, we can see without consulting a calculator that $f(x)$ is bigger. However, because $g(x)$ is degree three, and $f(x)$ is degree two, the theory of "orders of growth" guarantees that after some particular x-value, $g(x)$ will be bigger than $f(x)$.

In our case, it turns out that

$$\text{for all } x > 12,345,678,900 \text{ it is the case that } g(x) > f(x)$$

but what matters is not how I found the value 12,345,678,900, but rather that $g(x)$ eventually wins for all x that are "sufficiently large."

COPYRIGHT NOTICE: This is a work in-progress by Prof. Gregory V. Bard, which is intended to be eventually released under the Creative Commons License (specifically agreement # 3 "attribution and non-commercial.") Until such time as the document is completed, however, the author reserves all rights, to ensure that imperfect copies are not widely circulated.



Because this is a mathematics book and not a theology book, I urge you not to take the above statement on faith. Make up a few numbers, all greater than 12,345,678,900, and see that it is true for each and every one of them that $g(x)$ is greater than $f(x)$.

Let's try 12,400,000,000 and see what happens!

$$\begin{aligned} f(1,240,000,000) &= 123,456,789(12,400,000,000)^2 = 1.89827\cdots \times 10^{28} \\ g(1,240,000,000) &= 0.01(12,400,000,000)^3 = 1.90662\cdots \times 10^{28} \end{aligned}$$

As you can see, $g(x)$ is larger, as promised. Okay, now pick some numbers for yourself, and see that it works!



However, one can prove that 2^x grows faster than any of those. It is possible to prove that 2^x grows faster than x to the millionth power, to the billionth power, or to the trillionth power.

With that in mind, 2^x can be thought of as somewhat similar to "infinite degree." Really, no one should ever say something like "infinite degree" because infinity is not a real number. More precisely, 2^n grows faster than any polynomial, regardless of its degree.



Extraordinary claims require extraordinary evidence. Due to "overflow" it is hard to check

$$2^x > x^{1,000,000} \text{ for all } x \geq 24,549,171$$

on a calculator. Remember, the way that I got the value 24,549,171 does not matter. Instead, we will take the logarithm of both sides, and check

$$\ln 2^x > \ln x^{1,000,000} \text{ for all } x \geq 24,549,171$$

or more simply

$$x \ln 2 > 1,000,000 \ln x \text{ for all } x \geq 24,549,171$$

Let's check with $x = 24,549,172$ and see what happens! We'll have to jump to nine digits of precision for this one, instead of our usual six digits.

$$\begin{aligned} (24,549,172)(\ln 2) &= (24,549,172)(0.693147180\cdots) = 17,016,189.3\cdots \\ (1,000,000)(\ln 24,549,172) &= (1,000,000)(17.0161886\cdots) = 17,016,188.6\cdots \end{aligned}$$

As you can see, the inequality holds.



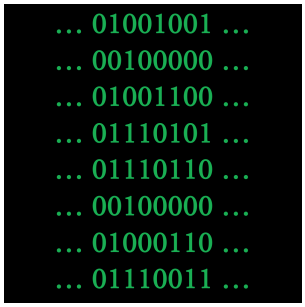
There is now an entire hierarchy of exponential functions

$$2^x < 3^x < 4^x < 5^x < 6^x < 7^x < 8^x < \cdots$$

all of them growing faster than any polynomial, including millionth, billionth, and trillionth degree polynomials.

As if that were not enough, $n!$ grows faster than 2^n, 3^n, 4^n, or any of those. Yet even faster and more rapidly growing than $n!$ is the function n^n, which reaches very high values as early as $n = 20$.

COPYRIGHT NOTICE: This is a work in-progress by Prof. Gregory V. Bard, which is intended to be eventually released under the Creative Commons License (specifically agreement # 3 "attribution and non-commercial.") Until such time as the document is completed, however, the author reserves all rights, to ensure that imperfect copies are not widely circulated.



Amazingly enough, it turns out that these ideas are immensely important in "complexity theory," a branch of computer science that helps computer programmers figure out how long computational tasks will take to run. Using tools such as "big Θ-notation" and "big-Oh notation," which are based on these concepts, they can design fast computer programs, capable of handling large data sets.

You have now completed the module. Thank you very much for sticking with me through this very long reading!

COPYRIGHT NOTICE: This is a work in-progress by Prof. Gregory V. Bard, which is intended to be eventually released under the Creative Commons License (specifically agreement # 3 "attribution and non-commercial.") Until such time as the document is completed, however, the author reserves all rights, to ensure that imperfect copies are not widely circulated.