Why more is better: The power of multiple proofs

Steven J. Miller, Williams College
Steven.J.Miller@williams.edu
http://web.williams.edu/Mathematics/sjmiller/public_html/

Hampshire College
Prime Time Talk, July 31, 2014

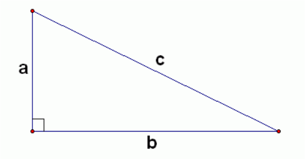
Goals of the Talk

- Often multiple proofs: Say a proof rather than the proof.
- Different proofs highlight different aspects.
- Too often rote algebra – explore!
- General: How to find / check proofs: special cases, 'smell' test.
- Specific: Pythagorean Theorem, Dimensional Analysis, Sabermetrics.

My math riddles page: http://mathriddles.williams.edu/.

Pythagorean Theorem

Geometry Gem: Pythagorean Theorem



Theorem (Pythagorean Theorem)

Right triangle with sides a, b and hypotenuse c, then $a^2 + b^2 = c^2$.

Most students know the statement, but the proof?

Why are proofs important? Can help see big picture.

Geometric Proofs of Pythagoras

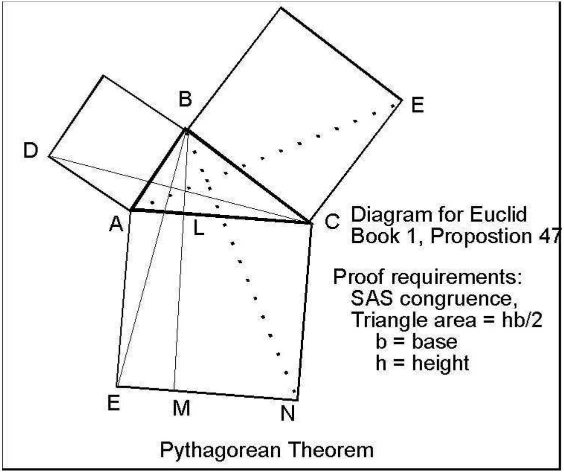


Figure: Euclid's Proposition 47, Book I. Why these auxiliary lines? Why are there equalities?

Geometric Proofs of Pythagoras

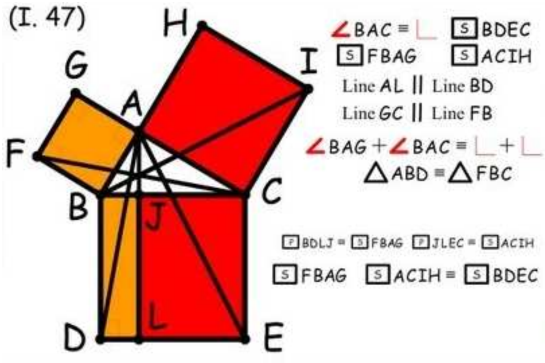


Figure: Euclid's Proposition 47, Book I. Why these auxiliary lines? Why are there equalities?

Geometric Proofs of Pythagoras

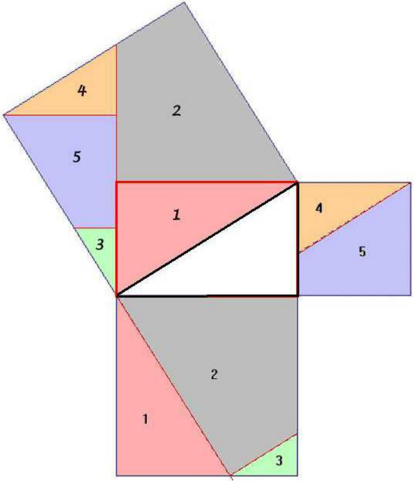


Figure: A nice matching proof, but how to find these slicings!

Geometric Proofs of Pythagoras

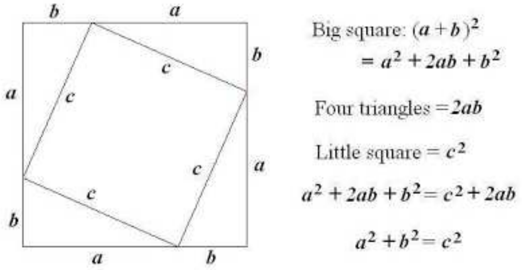


Figure: Four triangles proof: I

Geometric Proofs of Pythagoras

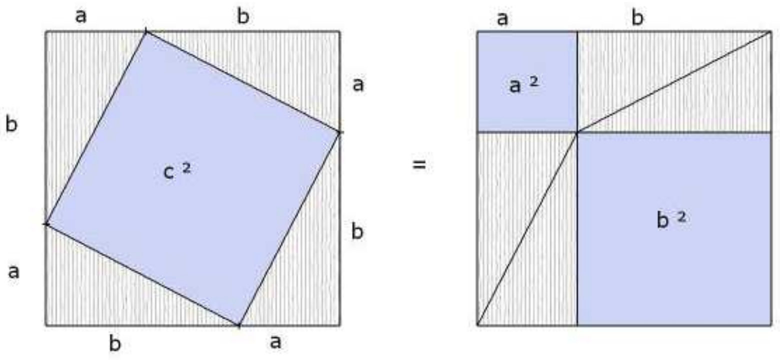


Figure: Four triangles proof: II

Geometric Proofs of Pythagoras

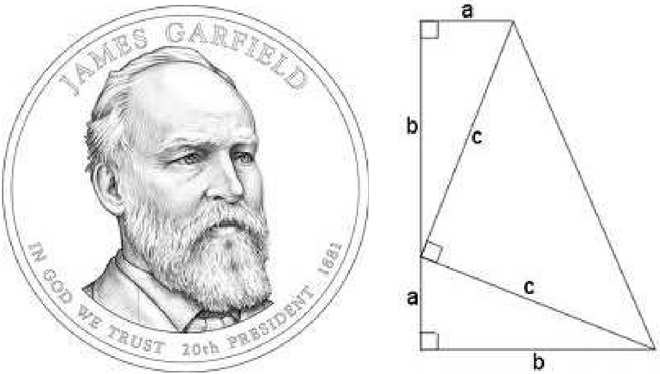


Figure: President James Garfield's (Williams 1856) Proof.

Geometric Proofs of Pythagoras

Lots of different proofs.

Difficulty: how to find these combinations?

At the end of the day, do you know *why* it's true?

Feeling Equations

Sabermetrics

Sabermetrics is the art of applying mathematics and statistics to baseball.

Danger: not all students like sports (Red Sox aren't making life easier!).

Lessons: not just for baseball; try to find the right statistics that others miss, competitive advantage (business, politics).

Estimating Winning Percentages

Assume team A wins p percent of their games, and team B wins q percent of their games. Which formula do you think does a good job of predicting the probability that team A beats team B? Why?

$$\frac{p+pq}{p+q+2pq}, \quad \frac{p+pq}{p+q-2pq}$$

$$\frac{p-pq}{p+q+2pq}, \quad \frac{p-pq}{p+q-2pq}$$

Estimating Winning Percentages

$$\frac{p+pq}{p+q+2pq}, \quad \frac{p+pq}{p+q-2pq}, \quad \frac{p-pq}{p+q+2pq}, \quad \frac{p-pq}{p+q-2pq}$$

How can we test these candidates?

Can you think of answers for special choices of p and q?

Estimating Winning Percentages

$$\frac{p+pq}{p+q+2pq}, \quad \frac{p+pq}{p+q-2pq}, \quad \frac{p-pq}{p+q+2pq}, \quad \frac{p-pq}{p+q-2pq}$$

Homework: explore the following:

◇ $p = 1$, $q < 1$ (do not want the battle of the undefeated).

◇ $p = 0$, $q > 0$ (do not want the Toilet Bowl).

◇ $p = q$.

◇ $p > q$ (can do $q < 1/2$ and $q > 1/2$).

◇ Anything else where you ‘know’ the answer?

Estimating Winning Percentages

$$\frac{p+pq}{p+q+2pq}, \quad \frac{p+pq}{p+q-2pq}, \quad \frac{p-pq}{p+q+2pq}, \quad \frac{p-pq}{p+q-2pq}$$

Homework: explore the following:

◇ $p = 1$, $q < 1$ (do not want the battle of the undefeated).

◇ $p = 0$, $q > 0$ (do not want the Toilet Bowl).

◇ $p = q$.

◇ $p > q$ (can do $q < 1/2$ and $q > 1/2$).

◇ Anything else where you 'know' the answer?

Estimating Winning Percentages

$$\frac{p - pq}{p + q - 2pq} = \frac{p(1-q)}{p(1-q) + (1-p)q}$$

Homework: explore the following:

$\diamond$ $p = 1$, $q < 1$ (do not want the battle of the undefeated).

$\diamond$ $p = 0$, $q > 0$ (do not want the Toilet Bowl).

$\diamond$ $p = q$.

$\diamond$ $p > q$ (can do $q < 1/2$ and $q > 1/2$).

$\diamond$ Anything else where you ‘know’ the answer?

Estimating Winning Percentages: ‘Proof’

Figure: First see how A does, then B.

Estimating Winning Percentages: 'Proof'

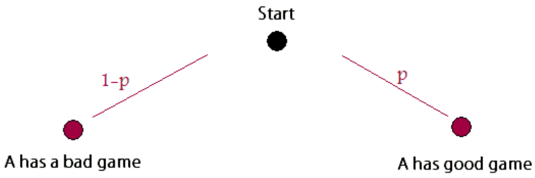


Figure: Two possibilities: A has a good day, or A doesn't.

Estimating Winning Percentages: ‘Proof’

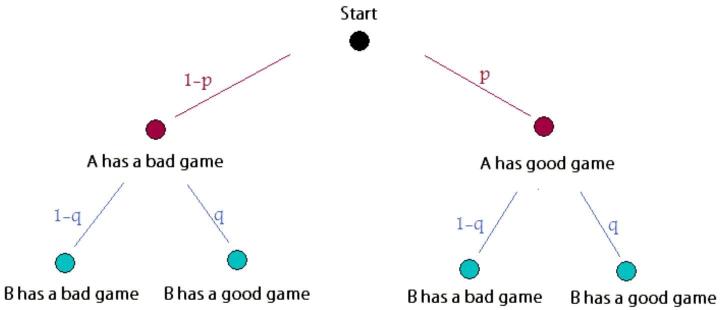


Figure: B has a good day, or doesn’t.

Estimating Winning Percentages: ‘Proof’

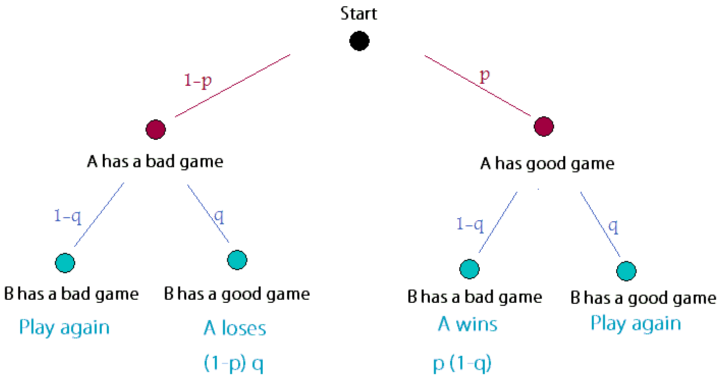
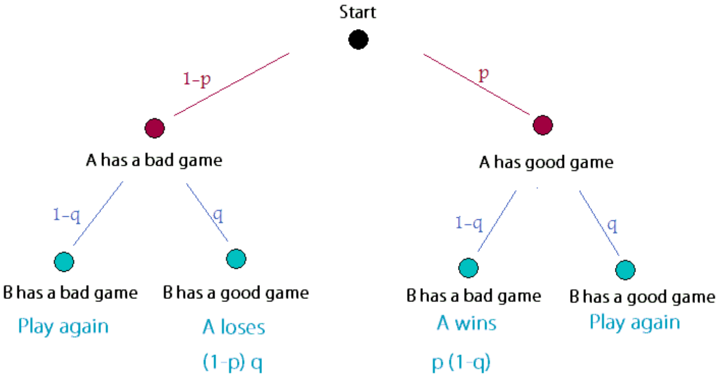


Figure: Two paths terminate, two start again.

Estimating Winning Percentages: ‘Proof’



$$\text{Probability } A \text{ wins is } \frac{p(1-q)}{p(1-q)+(1-p)q} = \frac{p-pq}{p+q-2pq}$$

Figure: Probability A beats B.

Lessons

Special cases can give clues.

Algebra can suggests answers.

Better formula: Bill James' Pythagorean Won-Loss formula.

Numerical Observation: Pythagorean Won-Loss Formula

Parameters

- RS_{obs}: average number of runs scored per game;
- RA_{obs}: average number of runs allowed per game;
- γ: some parameter, constant for a sport.

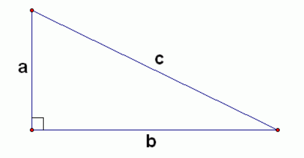
James' Won-Loss Formula (NUMERICAL Observation)

$$\text{Won} - \text{Loss Percentage} \; = \; \frac{\text{RS}_{\text{obs}}{}^{\gamma}}{\text{RS}_{\text{obs}}{}^{\gamma} + \text{RA}_{\text{obs}}{}^{\gamma}}$$

γ originally taken as 2, numerical studies show best γ is about 1.82. Used by ESPN, MLB.
See http://arxiv.org/abs/math/0509698 for a 'derivation'.

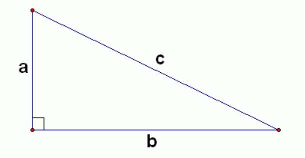
Dimensional Analysis

Possible Pythagorean Theorems....



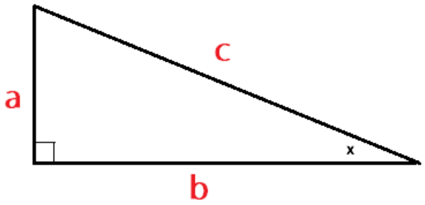
- ⋄ $c^2 = a^3 + b^3$.
- ⋄ $c^2 = a^2 + 2b^2$.
- ⋄ $c^2 = a^2 - b^2$.
- ⋄ $c^2 = a^2 + ab + b^2$.
- ⋄ $c^2 = a^2 + 110ab + b^2$.

Possible Pythagorean Theorems....



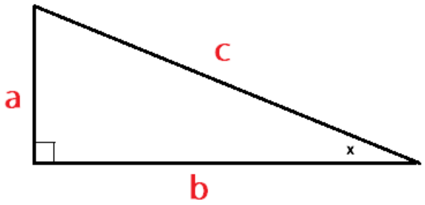
- $c^2 = a^3 + b^3$. No: wrong dimensions.
- $c^2 = a^2 + 2b^2$. No: asymmetric in a, b.
- $c^2 = a^2 - b^2$. No: can be negative.
- $c^2 = a^2 + ab + b^2$. Maybe: passes all tests.
- $c^2 = a^2 + 110ab + b^2$. No: violates $a + b > c$.

Dimensional Analysis Proof of the Pythagorean Theorem



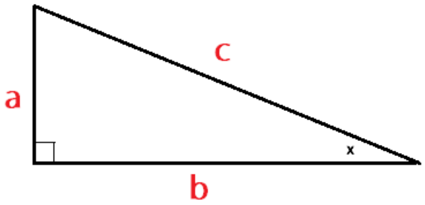
◇ Area is a function of hypotenuse c and angle x.

Dimensional Analysis Proof of the Pythagorean Theorem



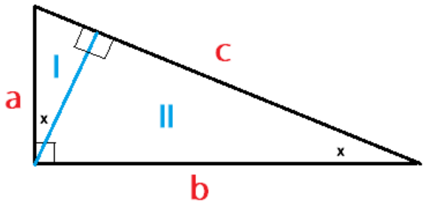
- Area is a function of hypotenuse c and angle x.
- $\mathrm{Area}(c, x) = f(x)c^2$ for some function f (similar triangles).

Dimensional Analysis Proof of the Pythagorean Theorem



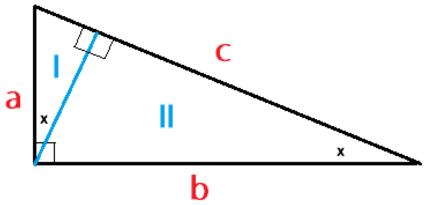
- Area is a function of hypotenuse c and angle x.
- $\text{Area}(c, x) \; = \; f(x)c^2$ for some function f (similar triangles).
- Must draw an auxiliary line, but where? Need right angles!

Dimensional Analysis Proof of the Pythagorean Theorem



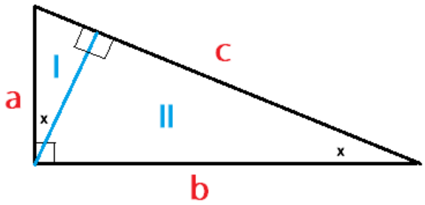
- Area is a function of hypotenuse c and angle x.
- $\text{Area}(c, x) \ = \ f(x)c^2$ for some function f (CPCTC).
- Must draw an auxiliary line, but where? Need right angles!

Dimensional Analysis Proof of the Pythagorean Theorem



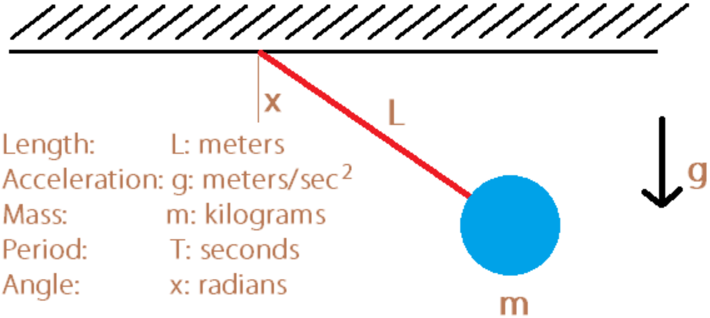
- Area is a function of hypotenuse c and angle x.
- $\mathrm{Area}(c, x) = f(x)c^2$ for some function f (CPCTC).
- Must draw an auxiliary line, but where? Need right angles!
- $f(x)a^2 + f(x)b^2 = f(x)c^2$

Dimensional Analysis Proof of the Pythagorean Theorem

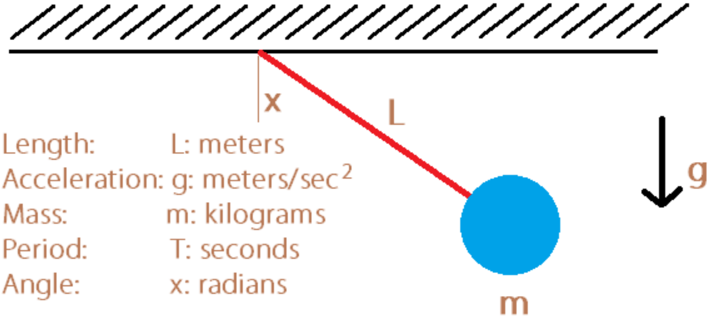


- Area is a function of hypotenuse c and angle x.
- $\text{Area}(c, x) = f(x)c^2$ for some function f (CPCTC).
- Must draw an auxiliary line, but where? Need right angles!
- $f(x)a^2 + f(x)b^2 = f(x)c^2 \Rightarrow a^2 + b^2 = c^2$.

Dimensional Analysis and the Pendulum

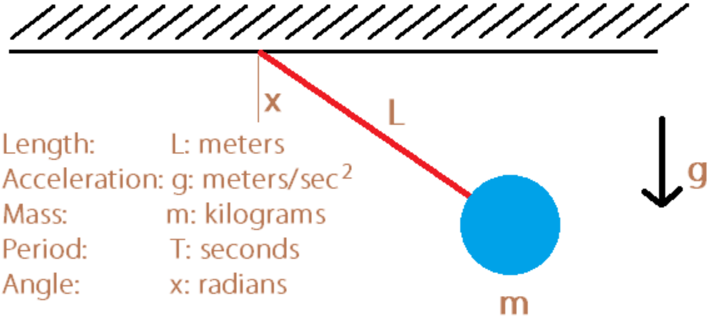


Dimensional Analysis and the Pendulum



Period: Need combination of quantities to get seconds.

Dimensional Analysis and the Pendulum



Period: Need combination of quantities to get seconds.

$$T = f(x)\sqrt{L/g}.$$

Dimensional Analysis Examples

Consider $\int x^{17} e^{ax} dx$.

What are the features of the solution?

Dimensional Analysis Examples

Consider $\int x^{17} e^{ax} dx$.

What are the features of the solution?

$$\frac{e^{ax}}{a^{18}} \left(\sum_{k=0}^{17} c_k a^k x^k \right).$$

Other Gems

Sums of Integers

$$S_n \; := \; 1 + 2 + \cdots + n \; = \; \frac{n(n+1)}{2} \; \approx \; \frac{1}{2}n^2.$$

Sums of Integers

$$S_n := 1 + 2 + \cdots + n = \frac{n(n+1)}{2} \approx \frac{1}{2}n^2.$$

Proof 1: Induction.

Proof 2: Grouping:

$2S_n = (1 + n) + (2 + (n-1)) + \cdots + (n + 1).$

Sums of Integers

$$S_n \;:=\; 1 + 2 + \cdots + n \;=\; \frac{n(n+1)}{2} \;\approx\; \frac{1}{2}n^2.$$

Proof 1: Induction.
Proof 2: Grouping:
$2S_n \;=\; (1 + n) + (2 + (n - 1)) + \cdots + (n + 1)$.

Instead of determining sum useful to get sense of size.

Sums of Integers

$$S_n := 1 + 2 + \cdots + n = \frac{n(n+1)}{2} \approx \frac{1}{2}n^2.$$

Proof 1: Induction.
Proof 2: Grouping:
$2S_n = (1 + n) + (2 + (n - 1)) + \cdots + (n + 1).$

Instead of determining sum useful to get sense of size.

Have $\frac{n}{2}\frac{n}{2} \le S_n \le n^2$; thus S_n is between $n^2/4$ and n^2, have the correct order of magnitude of n.

Sums of Integers

$$S_n := 1 + 2 + \cdots + n = \frac{n(n+1)}{2} \approx \frac{1}{2}n^2.$$

Proof 1: Induction.
Proof 2: Grouping:
$2S_n = (1 + n) + (2 + (n-1)) + \cdots + (n + 1)$.

Instead of determining sum useful to get sense of size.

Have $\frac{n}{2}\frac{n}{2} \leq S_n \leq n^2$; thus S_n is between $n^2/4$ and n^2, have the correct order of magnitude of n.

Can improve: divide and conquer again: lather, rinse, repeat....

$$\frac{n}{4}\frac{n}{4} + \frac{n}{4}\frac{2n}{4} + \frac{n}{4}\frac{3n}{4} \leq S_n, \quad \text{so} \quad \frac{6}{16}n^2 \leq S_n.$$

Stirling's Formula

Stirling's Formula

We have

$$n! \approx n^n e^{-n}\sqrt{2\pi n}\left(1+\frac{1}{12n}+\cdots\right).$$

Can prove / get close by Integral Test, Euler-Maclaurin Formula.

Stirling's Formula ($n! \approx n^n e^{-n}\sqrt{2\pi n}$: Approximations

To illustrate ideas not worrying about rounding issues.

$[1, 2, \dots, \frac{n}{2}]\ [\frac{n}{2} + 1, \frac{n}{2} + 2, \dots, n]$.

$1^{n/2}(n/2)^{n/2} \le n! \le (n/2)^{n/2} n^{n/2}$, or $(n/2)^{n/2} \le n! \le n^n \sqrt{2}^{-n}$.

Have $\sqrt{2} \approx 1.414$ vs $e \approx 2.718$.

Stirling's Formula ($n! \approx n^n e^{-n}\sqrt{2\pi n}$: Approximations

To illustrate ideas not worrying about rounding issues.

$[1, 2, \dots, \frac{n}{2}]\ [\frac{n}{2}+1, \frac{n}{2}+2, \dots, n]$.

$1^{n/2}(n/2)^{n/2} \le n! \le (n/2)^{n/2} n^{n/2}$, or $(n/2)^{n/2} \le n! \le n^n \sqrt{2}^{-n}$.

Have $\sqrt{2} \approx 1.414$ vs $e \approx 2.718$.
$[1, 2, \dots, \frac{n}{4}]\ [\frac{n}{4}+1, \dots, \frac{n}{2}][\frac{n}{2}+1, \frac{n}{2}+2, \dots, \frac{3n}{4}][\frac{3n}{4}+1, \dots, n]$.

Upper bound now
$n! \le (n/4)^{n/4}(n/2)^{n/4}(3n/4)^{n/4}n^{n/4} = n^n(32/3)^{-n/4} = n^n(\sqrt[4]{32/3})^{-n}$.

Have $\sqrt[4]{32/3} \approx 1.8072$ vs $e \approx 2.718$.

Stirling's Formula ($n! \approx n^n e^{-n}\sqrt{2\pi n}$: Approximations

Use $xy \leq \left(\frac{x+y}{2}\right)^2$.

$[1, 2, \dots, \frac{n}{2}]\ [\frac{n}{2}+1, \frac{n}{2}+2, \dots, n]$.

$n! \leq (n/4)^{2(n/4)}(3n/4)^{2(n/4)}$, or $n! \leq n^n(16/3)^{-n/2}$.

Have $(4^2/3)^{1/2} \approx 2.3094$ vs $e \approx 2.718$ (much better than 1.414).

Stirling's Formula ($n! \approx n^n e^{-n}\sqrt{2\pi n}$: Approximations

Use $xy \leq \left(\frac{x+y}{2}\right)^2$.

$[1, 2, \dots, \frac{n}{2}]\ [\frac{n}{2}+1, \frac{n}{2}+2, \dots, n]$.

$n! \leq (n/4)^{2(n/4)}(3n/4)^{2(n/4)}$, or $n! \leq n^n(16/3)^{-n/2}$.

Have $(4^2/3)^{1/2} \approx 2.3094$ vs $e \approx 2.718$ (much better than 1.414).
$[1, 2, \dots, \frac{n}{4}]\ [\frac{n}{4}+1, \dots, \frac{n}{2}][\frac{n}{2}+1, \frac{n}{2}+2, \dots, \frac{3n}{4}]\ [\frac{3n}{4}+1, \dots, n]$.

Upper bound now
$n! \leq (n/8)^{2(n/8)}(3n/8)^{2(n/8)}(5n/8)^{2(n/8)}(7n/8)^{2(n/8)} = n^n(8^4/7!!)^{-n/4}$.

Have $(8^4/7!!)^{1/4} \approx 2.49915$ vs $e \approx 2.718$ (much better than 1.8072).

Stirling's Formula ($n! \approx n^n e^{-n}\sqrt{2\pi n}$: Approximations

Use $xy \le \left(\frac{x+y}{2}\right)^2$.

$[1, 2, \dots, \frac{n}{2}]\ [\frac{n}{2}+1, \frac{n}{2}+2, \dots, n]$.

$n! \le (n/4)^{2(n/4)}(3n/4)^{2(n/4)}$, or $n! \le n^n(16/3)^{-n/2}$.

Have $(4^2/3)^{1/2} \approx 2.3094$ vs $e \approx 2.718$ (much better than 1.414).
$[1, 2, \dots, \frac{n}{4}]\ [\frac{n}{4}+1, \dots, \frac{n}{2}][\frac{n}{2}+1, \frac{n}{2}+2, \dots, \frac{3n}{4}]\ [\frac{3n}{4}+1, \dots, n]$.

Upper bound now
$n! \le (n/8)^{2(n/8)}(3n/8)^{2(n/8)}(5n/8)^{2(n/8)}(7n/8)^{2(n/8)} = n^n(8^4/7!!)^{-n/4}$.

Have $(8^4/7!!)^{1/4} \approx 2.49915$ vs $e \approx 2.718$ (much better than 1.8072).

Next $(16^8/15!!)^{1/8} \approx 2.60473$, then $(32^{16}/31!!)^{1/16} \approx 2.66047$.

Homework: Note $e \approx 2.71828182845905$.

Can derive

$$n! \ \leq \ n^n \left((2^k)^{2^{k-1}} / (2^k - 1)!! \right)^{-n/2^{k-1}}.$$

k	$\left((2^k)^{2^{k-1}} / (2^k - 1)!! \right)^{1/2^{k-1}}$
2	2.30940107675850
3	2.49915194953620
4	2.60472929511376
8	2.71093864109117
12	2.71782189208667
16	2.71825307857336
20	2.71828003157610
24	2.71828171615380

Useful fact:

$$(2m-1)!! \ = \ \frac{(2m)!}{2^m m!}.$$

Geometric Irrationality Proofs: http://arxiv.org/abs/0909.4913

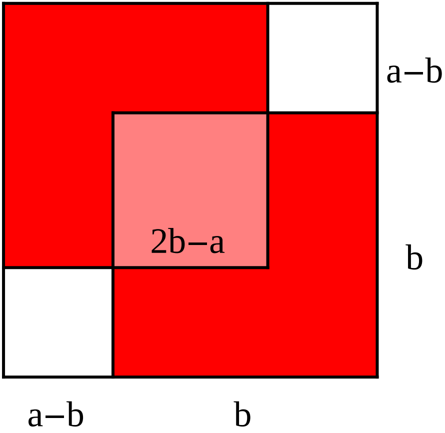


Figure: Geometric proof of the irrationality of $\sqrt{2}$.

Geometric Irrationality Proofs: http://arxiv.org/abs/0909.4913

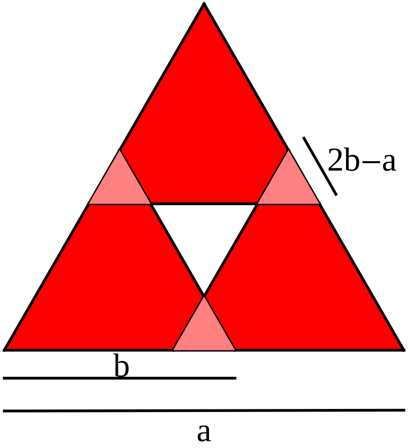


Figure: Geometric proof of the irrationality of $\sqrt{3}$

Geometric Irrationality Proofs:
http://arxiv.org/abs/0909.4913

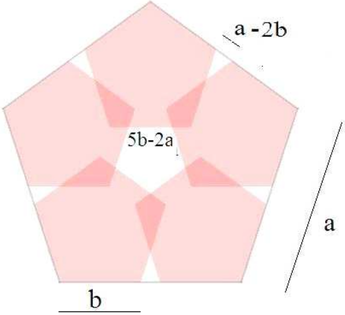


Figure: Geometric proof of the irrationality of $\sqrt{5}$.

Geometric Irrationality Proofs:
http://arxiv.org/abs/0909.4913

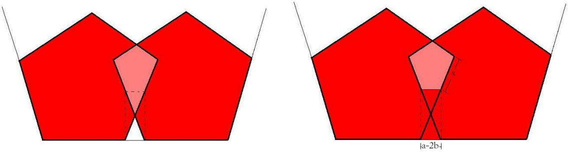


Figure: Geometric proof of the irrationality of $\sqrt{5}$: the kites, triangles and the small pentagons.

Geometric Irrationality Proofs:
http://arxiv.org/abs/0909.4913

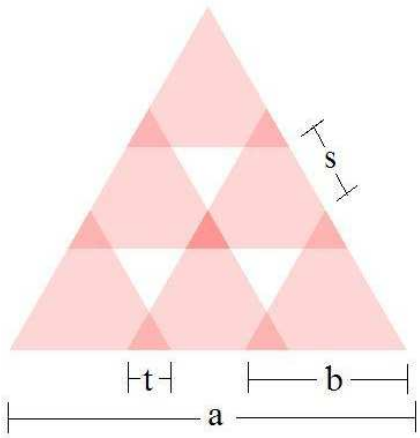


Figure: Geometric proof of the irrationality of $\sqrt{6}$.

Preliminaries: The Cookie Problem

The Cookie Problem

The number of ways of dividing C identical cookies among P distinct people is $\binom{C+P-1}{P-1}$.

Preliminaries: The Cookie Problem

The Cookie Problem

The number of ways of dividing C identical cookies among P distinct people is $\binom{C+P-1}{P-1}$.

Proof: Consider $C + P - 1$ cookies in a line.
Cookie Monster eats $P - 1$ cookies: $\binom{C+P-1}{P-1}$ ways to do.
Divides the cookies into P sets.

Preliminaries: The Cookie Problem

The Cookie Problem

The number of ways of dividing C identical cookies among P distinct people is $\binom{C+P-1}{P-1}$.

Proof: Consider $C + P - 1$ cookies in a line.

Cookie Monster eats $P - 1$ cookies: $\binom{C+P-1}{P-1}$ ways to do.

Divides the cookies into P sets.

Example: 8 cookies and 5 people ($C = 8$, $P = 5$):

Preliminaries: The Cookie Problem

The Cookie Problem

The number of ways of dividing C identical cookies among P distinct people is $\binom{C+P-1}{P-1}$.

Proof: Consider $C+P-1$ cookies in a line.
Cookie Monster eats $P-1$ cookies: $\binom{C+P-1}{P-1}$ ways to do.
Divides the cookies into P sets.
Example: 8 cookies and 5 people ($C=8$, $P=5$):

Preliminaries: The Cookie Problem

The Cookie Problem

The number of ways of dividing C identical cookies among P distinct people is $\binom{C+P-1}{P-1}$.

Proof: Consider $C + P - 1$ cookies in a line.
Cookie Monster eats $P - 1$ cookies: $\binom{C+P-1}{P-1}$ ways to do.
Divides the cookies into P sets.
Example: 8 cookies and 5 people ($C = 8$, $P = 5$):

Preliminaries: The Cookie Problem

The Cookie Problem

The number of ways of dividing C identical cookies among P distinct people is $\binom{C+P-1}{P-1}$. Solved $x_1 + \cdots + x_P = C$, $x_i \geq 0$.

Proof: Consider $C + P - 1$ cookies in a line.
Cookie Monster eats $P - 1$ cookies: $\binom{C+P-1}{P-1}$ ways to do.
Divides the cookies into P sets.
Example: 8 cookies and 5 people ($C = 8$, $P = 5$):

Conclusion

Conclusion

- ◇ Math is not complete – explore and conjecture!
- ◇ Different proofs highlight different aspects.
- ◇ Get a sense of what to try / what might work.