Math Camp: Lesson 1

Basics, Notation, Pre-Calculus

UW–Madison Political Science

August 19, 2019

Welcome

First…

Numbers

Math: not just numbers

Math is a general framework for manipulating various concepts, one of which is "numbers"

Types of numbers:

- Integers (whole numbers, including negative): $\mathbb{Z}$
- Real numbers (the continuous number line): $\mathbb{R}$
- Positive and negative real numbers: $\mathbb{R}^+$ and $\mathbb{R}^-$
- Real numbers in n dimensions: $\mathbb{R}^n$
- Complex numbers: $\mathbb{C}$
  - $1 + 2i$
  - where $i = \sqrt{-1}$

Special numbers

0 and 1:

- Useful for identities in addition and multiplication
  - $4 + 0 = 4$
  - $1 * x = x$
- Useful for turning "off" and "on" pieces of an equation
  - We have an equation: $4 + z(5x) + (1 - z)3x$
  - What happens if $z = 0$?
  - If $z = 1$?

Special numbers

π: Approximately 3.1415926 ...

- Ratio of a circle's circumference to its diameter
- Useful for angles, circles, polar coordinates (which you may use!)

e: Euler's number

- Approximately 2.718281828 ...
- Derivative of e^x is equal to e^x
- Very important for probability and statistics

i: The imaginary number, $\sqrt{-1}$. For advanced algebra

∞: Infinity. Not really a number but a boundless quantity

Numbers and variables

We can manipulate numbers with simple operations

- addition: $1 + 2$
- subtraction: $3 - 4$
- multiplication: 5×6
- division: $7 \div 8$

We can generalize these statements by using variables

- $1 + x$
- $3 - y$
- $5 \times a$
- $7 \div m$

Variable: a symbol to represent an entity that could take different values

With me so far?

Equations

Statements about equality and inequality

Equations

Equality: left-hand and right-hand sides (LHS and RHS) are substitutable.

If $a = 7x$, then $8 * a$ and $8 * 7x$ are equivalent statements

- Usually used to find new ways to express a known idea or to derive implications from known ideas

Inequalities show whether LHS or RHS is greater

- Usually we use them to find the conditions under which a statement holds

Solving equalities and inequalities

Isolate an unknown quantity to express it in terms of known quantities

Solve for x:

$$-7x + 3 = 12$$

Solve for x:

$$-7x + 3 > 12$$

Remember to flip an inequality if you multiply by a negative number!

Data

Data

The information we record about what we study

- Cases: The units being studied (rows)
- Variables: Characteristics that describe units (columns)
- Values: Specific realization of a variable (cells)

Establishment	Location	Coffee	Vibe	Notable Flaw
Aldo's	Campus	7	good	no plugs
Ancora	Capitol	8	great	hours
Barriques	Capitol	6	good	bathroom key
Colectivo	State St.	7	fair	spotty wifi
Fair Trade	State St.	8	good	expensive, tables
Michelangelo's	Capitol	5	meh	bathroom key
Steep & Brew	Bascom Hill	0	fair	bad coffee (espresso OK)

Classifying data

Quantitative vs. Qualitative analysis

- Broad, sometimes contentious, arguably artificial divides in the study of politics
- Quantitative: larger n, statistical description and inference
- Qualitative: smaller n, rich description, non-statistical inference

Quantitative vs. Qualitative variables

- Quantitative: countable, numeric (age, number of toes)
- Qualitative: not countable but descriptive (gender, party preference)
- Unlike qualitative research, qualitative variables can still be organized in a data table

I advise not getting too hung up these classification systems. Use them only until you can lose them

Discrete vs. Continuous

Discrete

- Variables take specific values from a finite set of possible values
- Could be categories, could be discrete numbers
- Layer of the atmosphere, number of parties, country of origin

Dichotomous

- Special type of discrete variable, two possible values
- 0 or 1, yes or no, war or peace, win or lose, voted or not

Continuous

- Variables take values from a continuous number line
- Could be a bounded number line
- GDP, vote share, percentage of turnout, unemployment rate

Related: the "levels of measurement"

Nominal / Categorical

- Unordered categories
- e.g. party affiliation, gender, country of origin, occupation

Ordinal

- Ordered or ranked outcomes
- Could be categories, but numbers are possible (e.g. rankings)
- No fixed "distance" between levels
- e.g. highest level of educational attainment, ranking of countries by the level of corruption, issue prioritization

Related: the "levels of measurement"

Interval

- Ordered values with fixed distance between levels
- But no true zero point
- Issue scales, day of the year, Likert scales (debatable)

Ratio:

- Ordered, fixed intervals, and true zero
- Vote percentage, turnout, minutes in line to vote

Why do these categorizations matter?

Different types of variables lend themselves to different types of analysis

- Types of statistics you calculate
- Possibilities for data visualization
- Statistical models (you will learn if you take 812, 813, MLE, etc.)
    - Discrete/Nominal: logistic regression (logit), multinomial logit, χ^2 (chi-square)
    - Ordinal: ordinal logit, also χ^2
    - Interval: least-squares regression
    - Ratio: least-squares, count/rate models, duration/survival models

Everyone on board?

Sets

(useful for "speaking math" about data)

Sets

Collections of objects or entities

Sets contain "elements" or "members"

- if o is in set A, then we write $o \in A$
- o "is an element of" A

Sets could contain individual numbers, but they could contain other sorts of entities

- vectors, matrices
- functions, probability distributions

We need only some set notation to help us work with data

What's in a set?

A set could contain individual elements, written as

$$A = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$$

A set could contain intervals

- $B = [1, 10]$ (square brackets imply that endpoints are included)
- Is it the case that $A = B$?
- $C = (0, 11)$ (parentheses indicate that endpoints are not included)
- Does $B = C$?

Set notation

$\cup$: the union of two sets

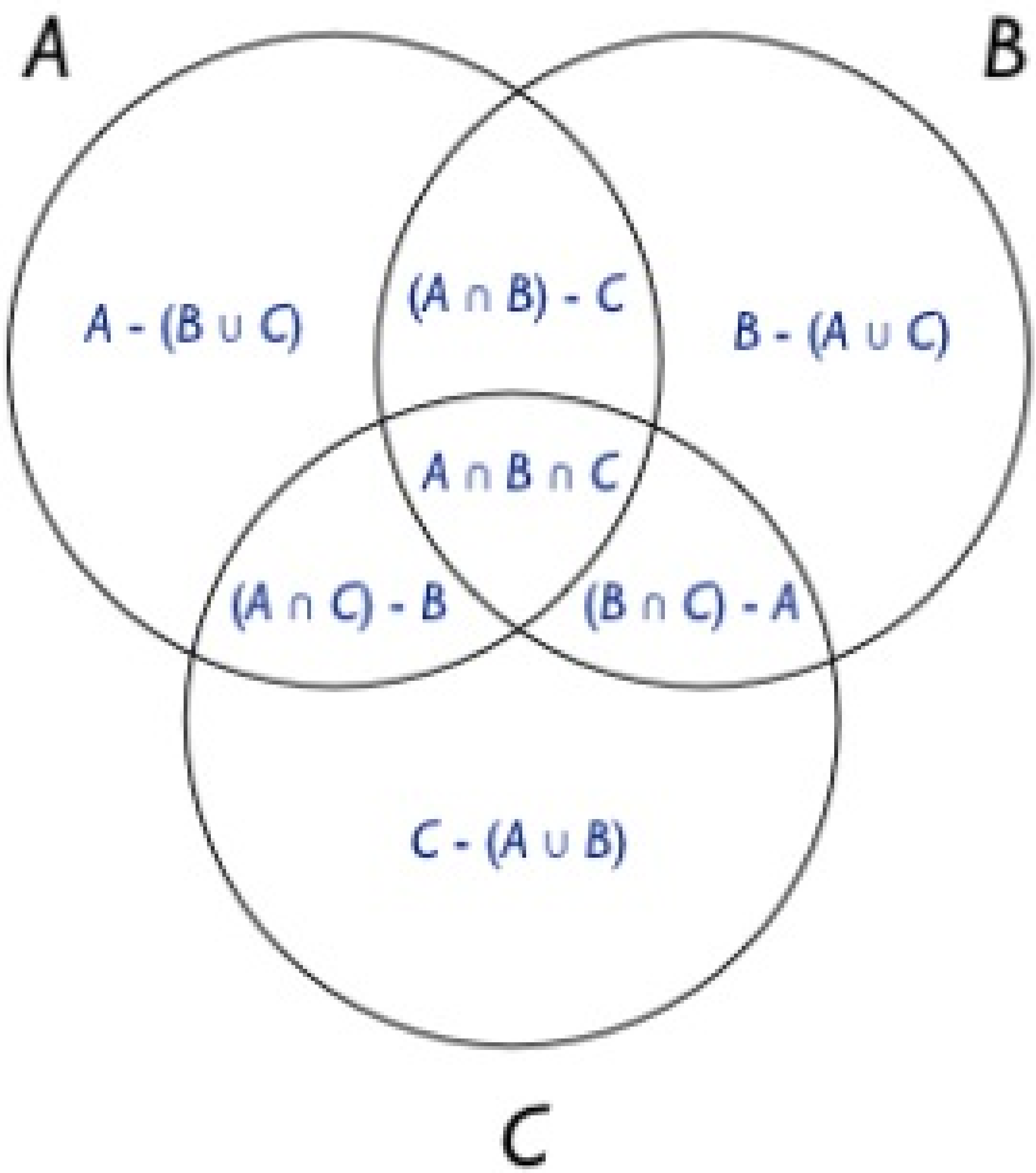
- elements that are members of either set
- if $A = \{1, 2\}$ and $B = \{2, 3\}$, then $A \cup B = \{1, 2, 3\}$

$\cap$: the intersection of two sets

- elements that are members of both sets
- if $A = \{1, 2\}$ and $B = \{2, 3\}$, then $A \cap B = \{2\}$

(these concepts are helpful for probabilities)

$\emptyset$: the empty set (null set)



https://bosker.wordpress.com/2013/07/10/venn-diagram-partitioning/

Symbols and Set Notation

A handful of symbols are commonly used when we represent data mathematically.

Symbol	Meaning
$>$	greater than
$\geq$	greater than or equal to
$<$	less than
$\leq$	less than or equal to
$\approx$	approximately equal to $(x \approx y)$
$\equiv$	equivalent to (for establishing identities)
$\propto$	proportional to $(4x \propto x)$

Symbols and Set Notation

Symbol	Meaning
$\in$	is an element of a set ($x_i \in \mathbf{X}$)
$\neg$	not
$/$	not (if through a symbol: $a_i \notin \mathbf{B}$)
$\mid$	given that ($A \mid B$)
$\rightarrow$	implies ($A \rightarrow B$)
$\leftrightarrow$	if and only if ($[x = y] \leftrightarrow [y = x]$)

Subsets

if $A = \{1, 2\}$ and $B = \{1, 2, 3\}$:

- $A \subseteq B$ means "A is a subset of B"
- Can A be a subset of A?
- using $\subseteq$, yes it can

"Proper subset"

- $A \subset B$
- A isn't a proper subset of A
- proper subsets can't be equivalent to their superset

Indexing

Variables can be sets, and they can take different values for different individuals in a dataset. It is convenient to index individual observations using a subscript (typically i).

Student	Math Courses in College
1	3
2	0
3	1
4	4

If x represents the number of math courses, x_i refers to the ith observation in x

- $x_1 = 3$
- $x_2 = 0$
- $x_3 = ?$
- $x_4 = ?$

Set practice

Consider the following sets

$A = \{10, 20, 30\}$

$B = [1, 10]$

$C = (0, 40)$

- What is $A \cap B$?
- Is $B \subseteq C$?
- Is $A_2 \in B$?
- Which elements in A are subsets of B? Subsets of C?
- Is $(A \cup C) \subset C$?

Set notation in actual research

Almost verbatim from a paper about congressional votes ("roll call votes")*

- The data consist of n legislators voting on m different roll calls [bills].
- Each roll call $j = 1, \ldots, m$ presents legislators $i = 1, \ldots, n$ with a choice between a 'Yea' position ζ_j and a 'Nay' position Ψ_j, locations in $\mathbb{R}$
- Let $y_{ij} = 1$ if legislator i votes Yea on the jth roll call and $y_{ij} = 0$ otherwise.

Legislator (i)	Bill (j)	Vote (y)
1	1	0
1	2	1
$\vdots$	$\vdots$	$\vdots$
n	m	y_{nm}

* - Clinton, Jackman, and Rivers. "The Statistical Analysis of Roll Call Data." APSR 2004

Functions

Functions

Operations or rules of assignment that map an input to a unique (that is, exactly one) output

- inputs: also called arguments
- outputs: also called values
- mapping: also called the definition
- $\text{value} = f(\text{argument})$
- $f = [\text{insert definition}]$

Analogies include:

- algorithms, machines, black box, routinized process

Functions

We can generically refer to functions using $f(x)$ or $g(x)$, just as we might use variables to stand in for values.

- e.g. let $f(p) = \log\left(\frac{p}{1-p}\right)$
- $f(0.5) = ?$

Other symbols also are fine, e.g.

- $\Phi\left(\cdot\right)$
- $\Gamma(\cdot)$
- $B(\cdot)$
- $\Lambda(\cdot)$

Operators

Operators are components of a function that tell what to do with the inputs to produce the outputs

Several common operators serve as the building blocks of mathematical functions

- $+$
- $-$
- $\times$
- $/$

Order of operations:

- Operations within parentheses
- Exponents
- Multiplication and division (left to right)
- Addition and subtraction (left to right)

Function examples

x	y	z	$f(x,y) = x - y$	$g(z) = 2z - 1$	$h(x,y,z) = \frac{x+y}{z}$
5	0	5			
2	5	8			
0	3	9			
3	2	0			
8	4	2			
1	2	4			

Function examples

x	y	z	$f(x,y) = x - y$	$g(z) = 2z - 1$	$h(x,y,z) = \frac{x+y}{z}$
5	0	5	5	9	1
2	5	8	-3	15	.875
0	3	9	-3	17	.333
3	2	0	1	-1	undefined
8	4	2	4	3	6
1	2	4	-1	7	.75

Nested functions

Given that all functions do is map an input to an output, we can nest functions

- Imagine we perform one function $f(\cdot)$
- Output from $f(\cdot)$ serves as the input for $g(\cdot)$
- $g(f(\cdot))$

Functions are evaluated "inside out" following the order of operations

Suppose $x = 3$, $f(x) = 2x + 5$, $g(x) = x - 7$

$$f(g(x)) = ?$$

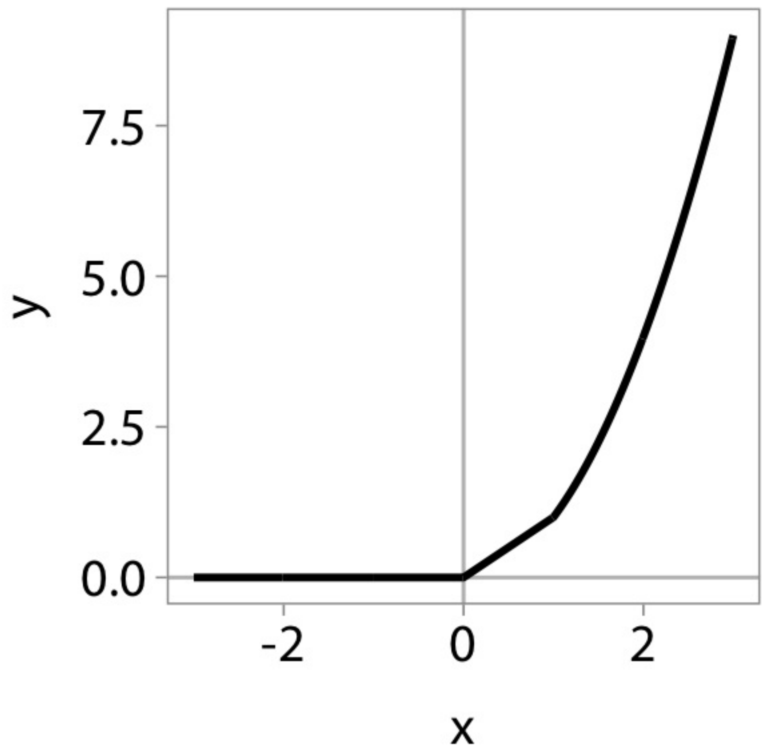
$$g(f(x)) = ?$$

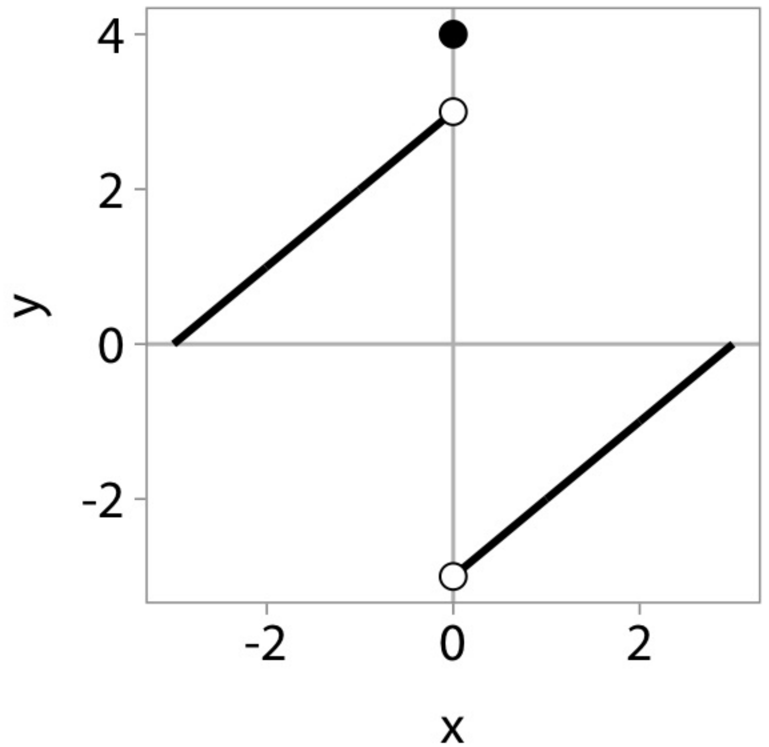
Piecewise-defined functions

A piecewise-defined function has different definitions for different regions of the domain (domain = function inputs)

$$f(x) = \begin{cases} 0 & \text{if } x < 0 \\ x & \text{if } x \in [0, 1] \\ x^2 & \text{if } x > 1 \end{cases}$$

$$g(x) = \begin{cases} x + 3 & \text{if } x \in (-\infty, 0) \\ 4 & \text{if } x = 0 \\ x - 3 & \text{if } x \in (0, \infty) \end{cases}$$





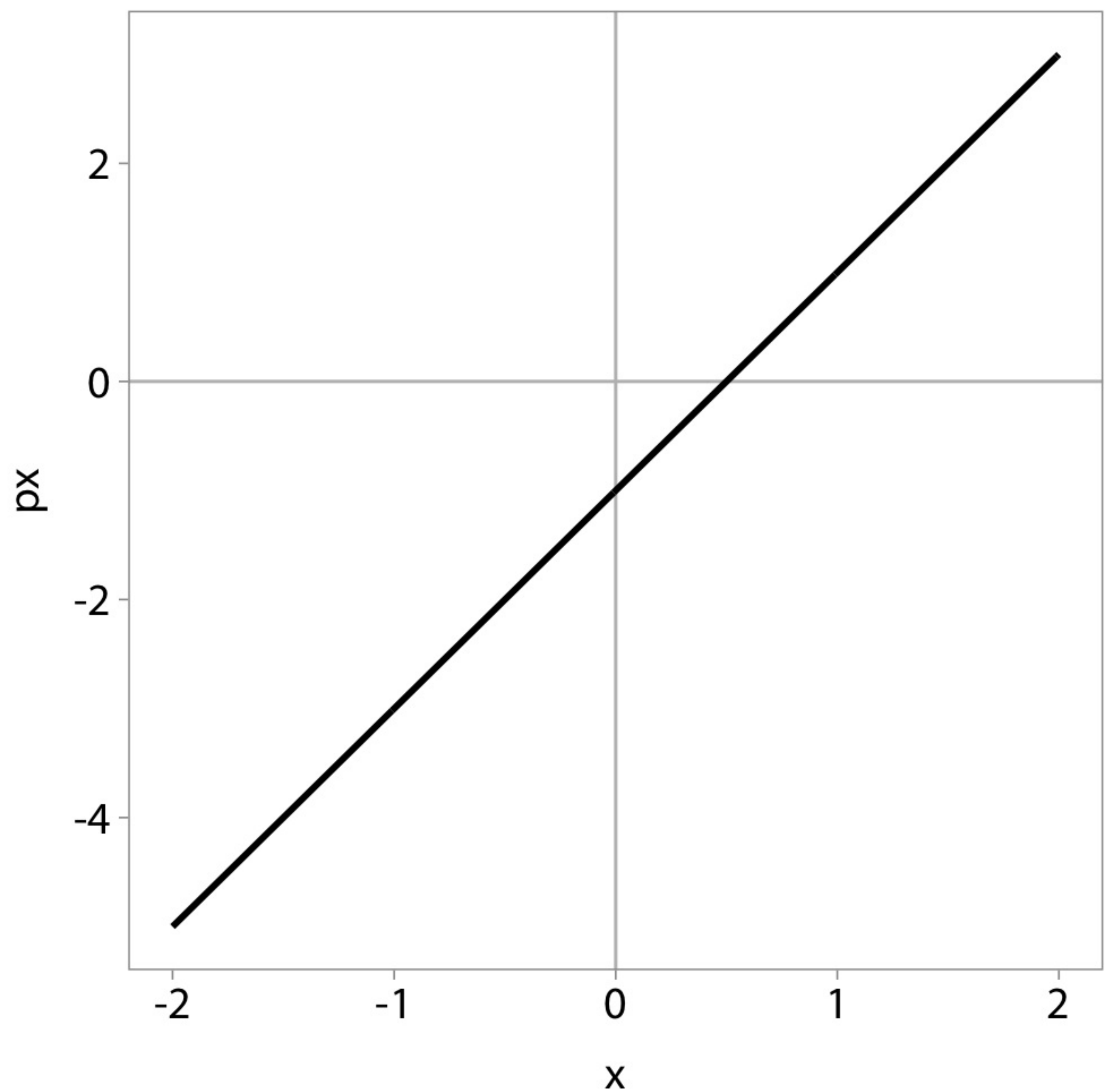
Function practice

Practice

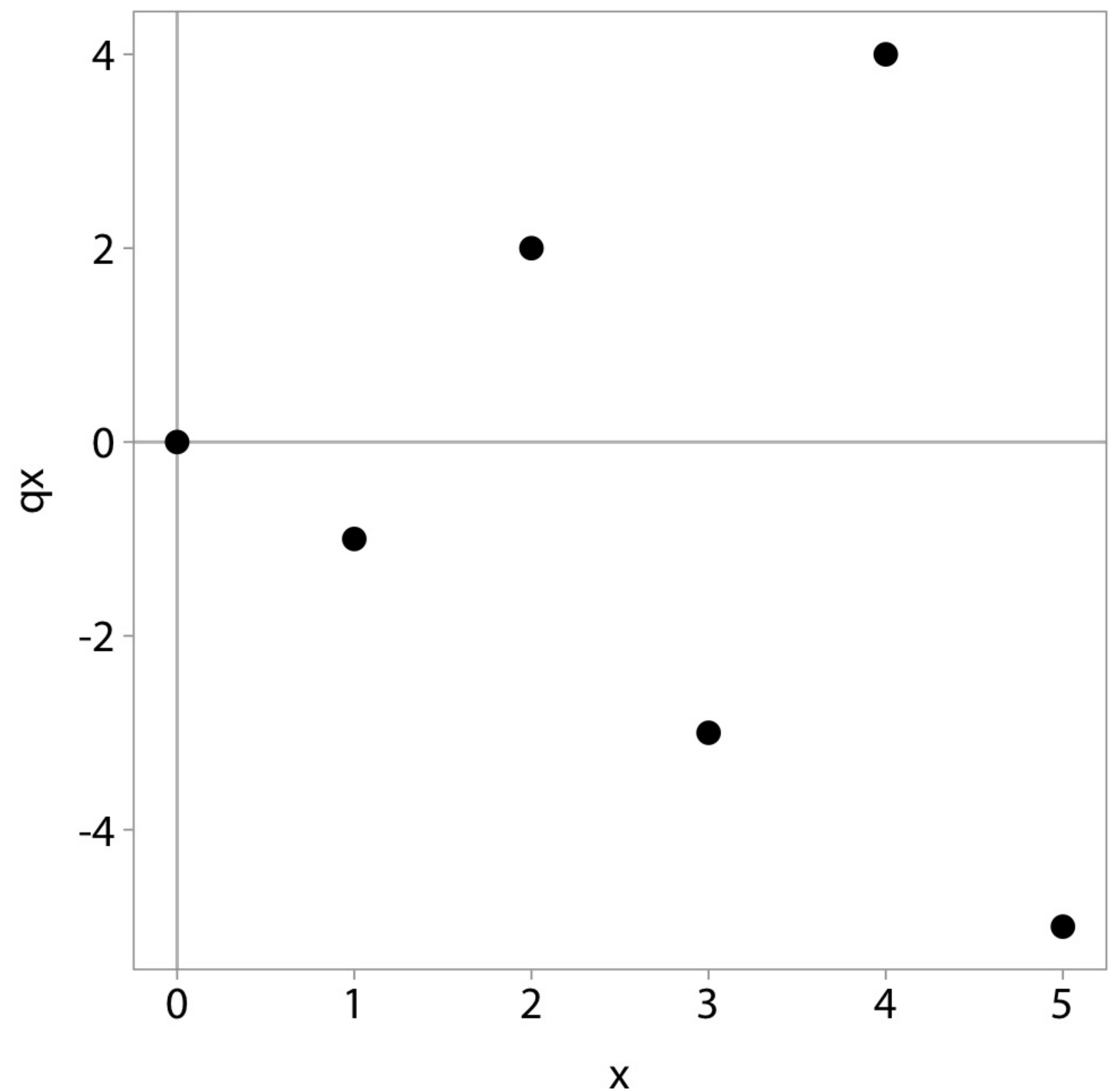
Sketch graphs of the following functions:

- $p(x) = 2x - 1$, on the interval $[-2, 2]$
- $q(x) = x * -1^x$, for integers $\{0, 1, 2, 3, 4, 5\}$
- $r(x) = 2x^2 - 3x + 4$, on the interval $(0, 4)$

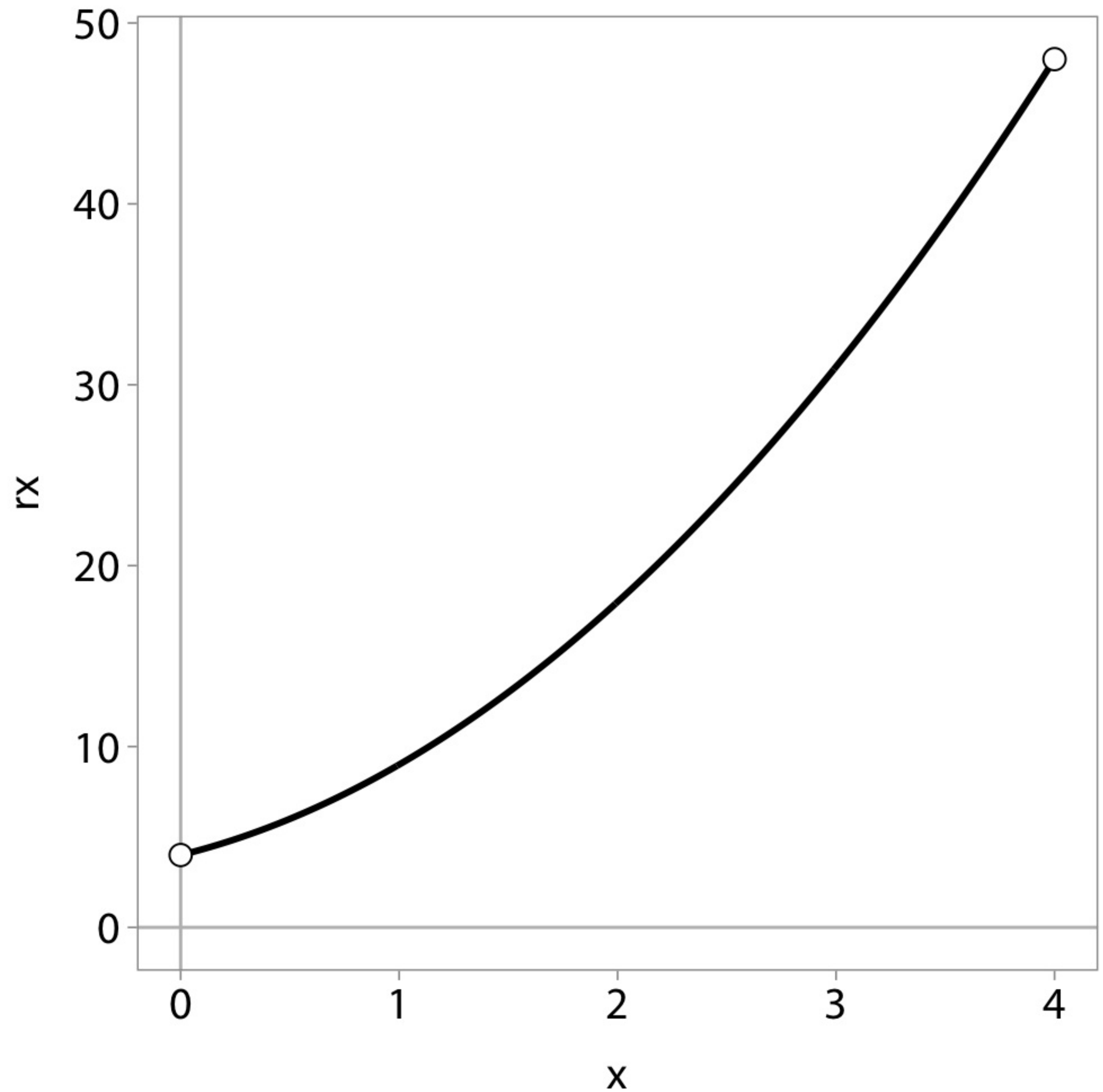
$p(x) = 2x - 1$, on the interval $[-2, 2]$



$q(x) = x * -1^x$, for integers $\{0, 1, 2, 3, 4, 5\}$



$r(x) = 2x^2 - 3x + 4$, on the interval $(0, 4)$



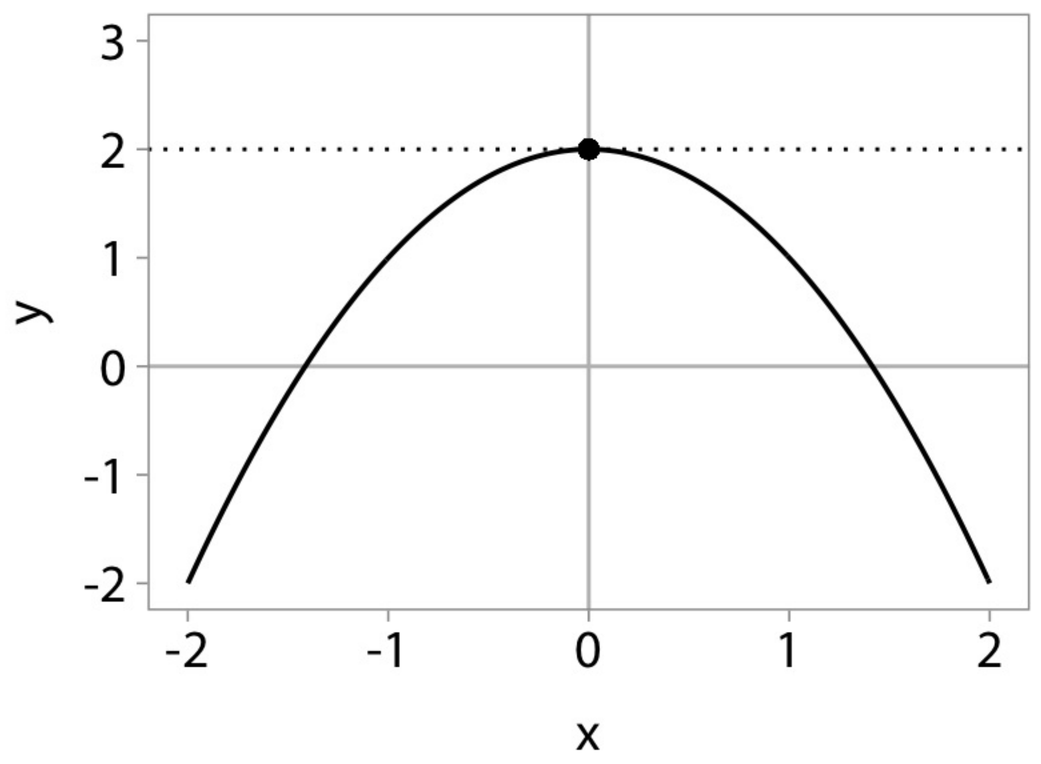
Important functions, routines, and properties

Limits

Limits will help us formally define concepts in this and other lectures.

A limit describes a function's behavior at a given input: $\lim_{x \to 0} \left(2 - x^2\right) = 2$

...or as the input value changes: $\lim_{x \to \infty} \left(2 - x^2\right) = -\infty$

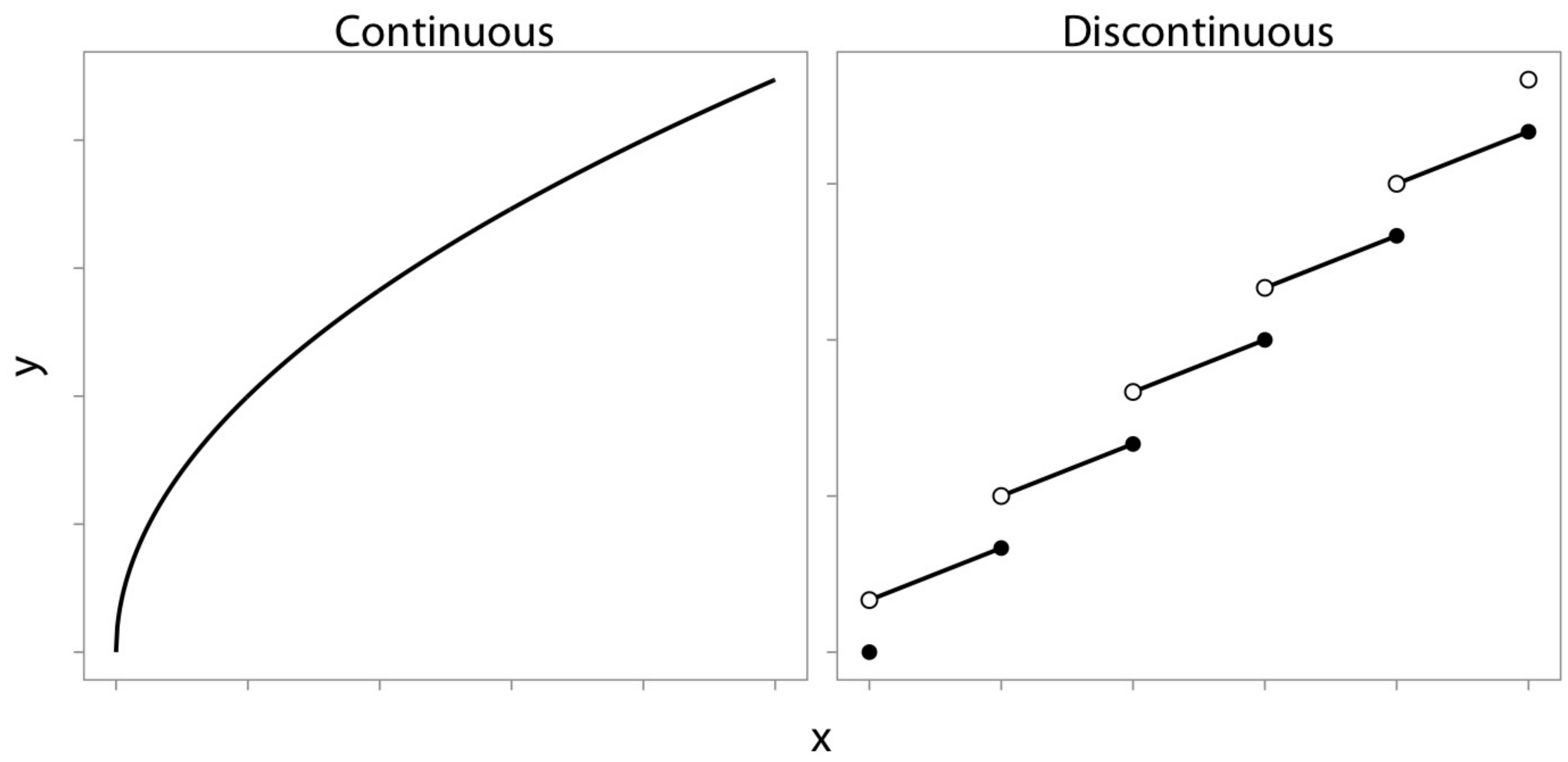


Continuity

A function is continuous if it has no gaps or jumps.

$$\lim_{x \to a} f(x) = f(a)$$

Meaning... small changes in input produce small changes in output



Continuity and discontinuity in political science research

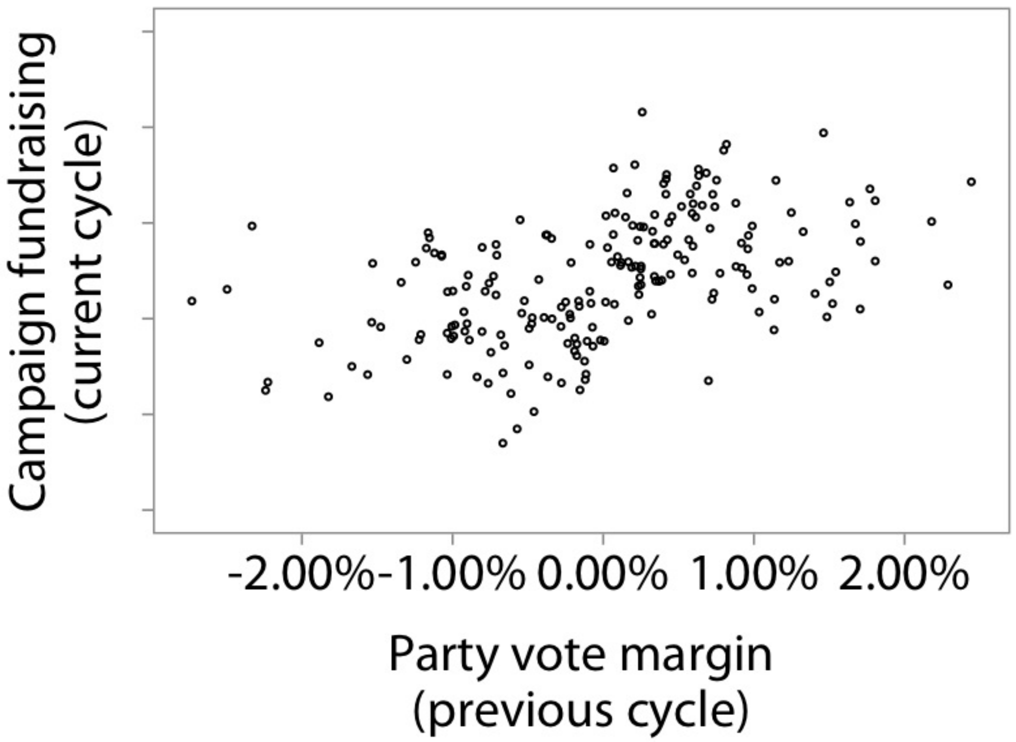
Why we care about continuity

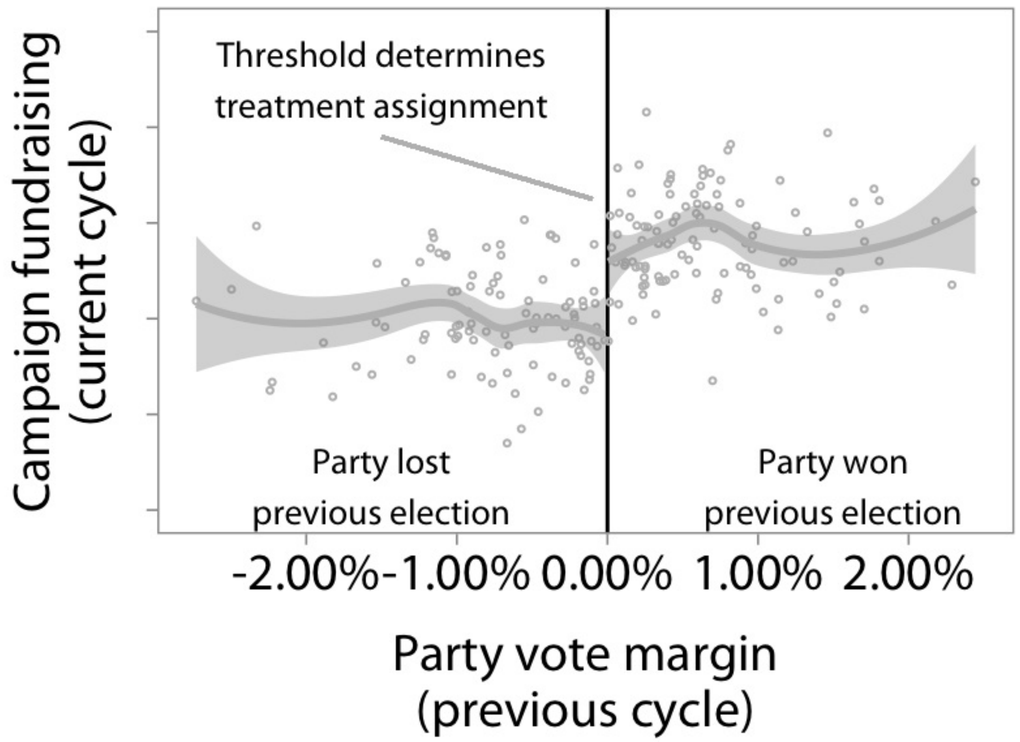
- Differentiation (derivatives), which is good for...
- Utility functions (game theory)
- How some variable affects another variable (statistical models)
- More on Wednesday

Continuity and Discontinuity

Why we care about discontinuity: regression discontinuity

- Causal inference method: units are "treated" if above/below a threshold
- Popular example: vote share to wins and losses, wins and losses affect future fundraising

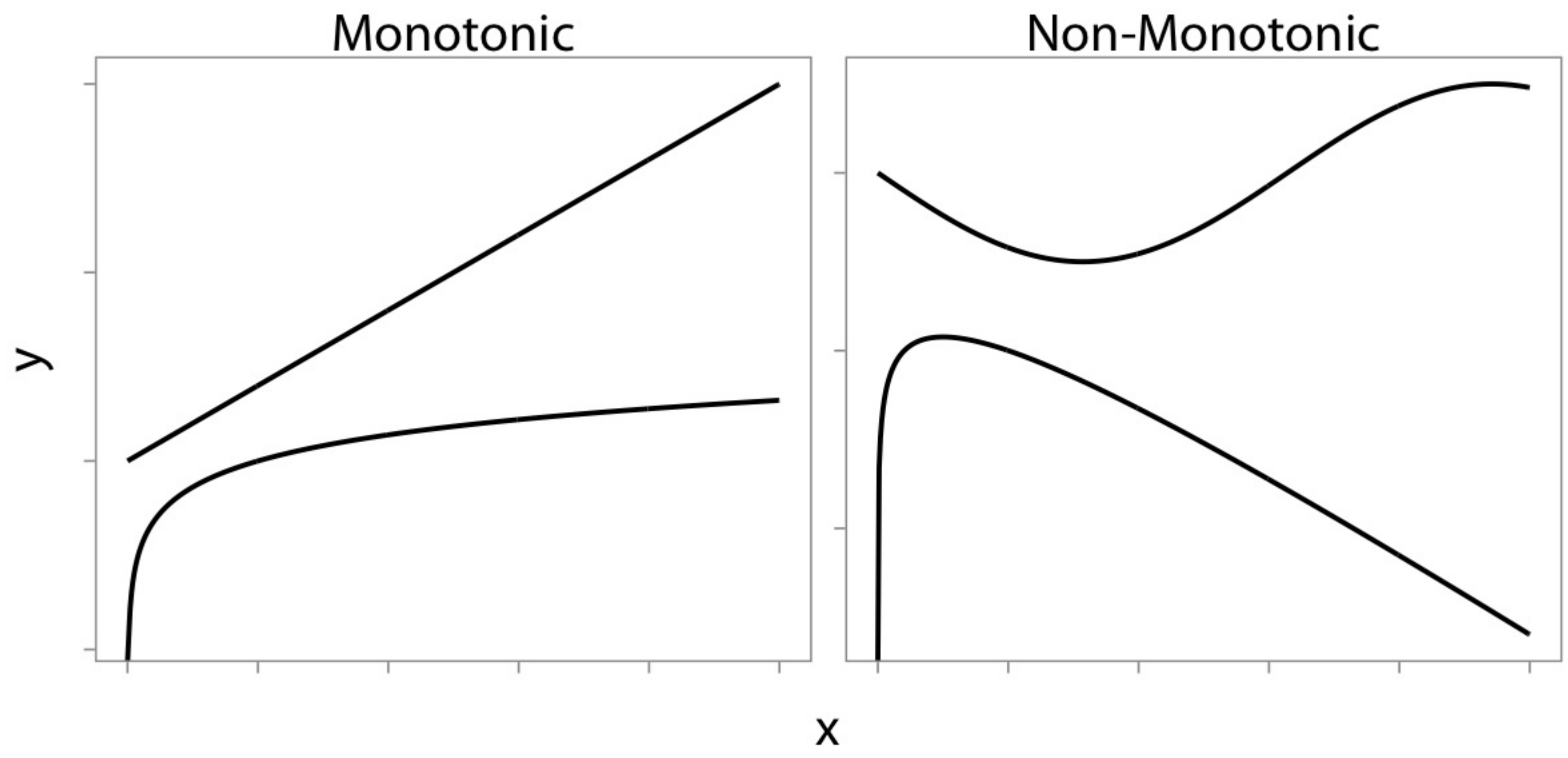




Monotonicity

A function is monotonic if it always increases (monotonically increasing) or always decreases (monotonically decreasing)

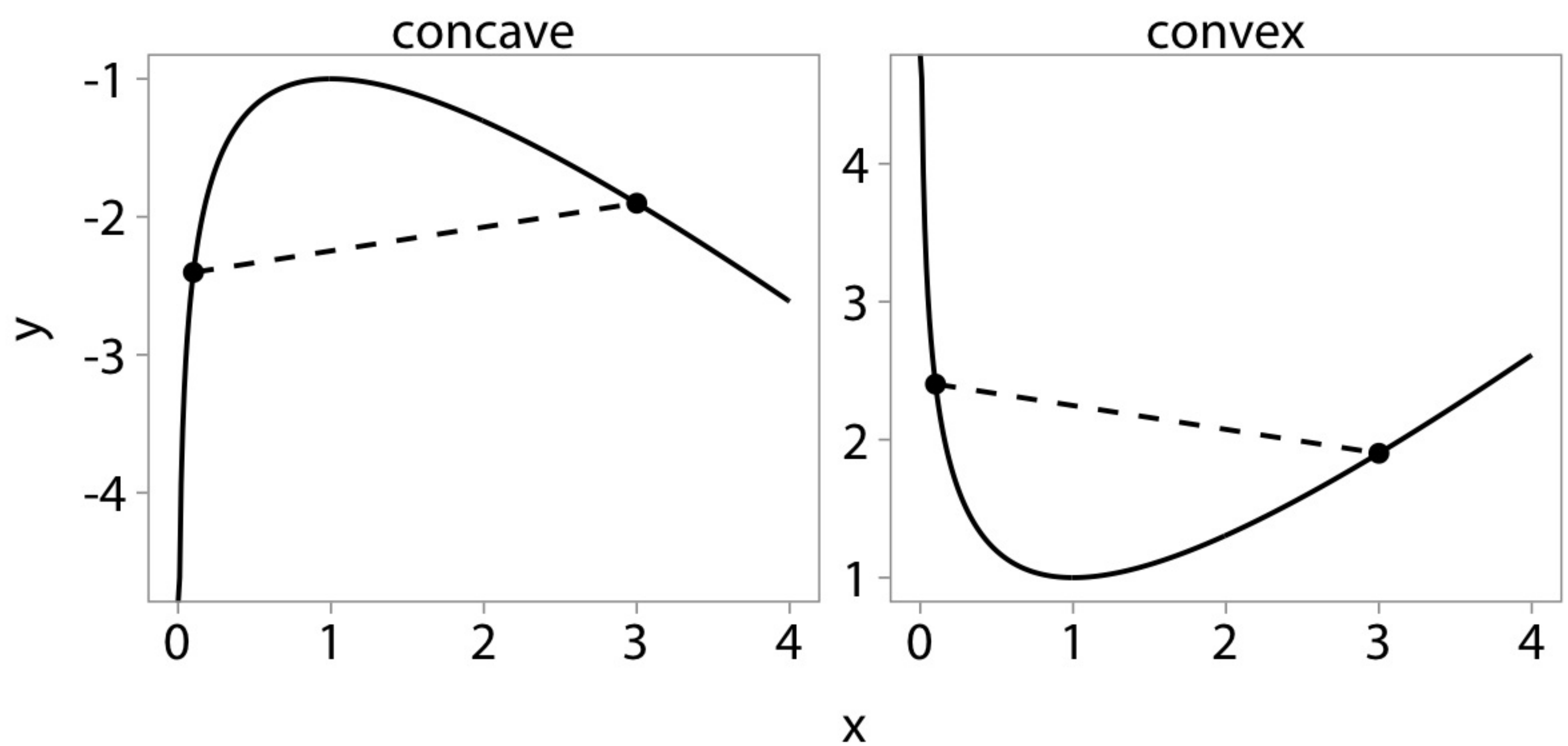
- increasing: for any $x_1 > x_2$, $f(x_1) > f(x_2)$
- decreasing: for any $x_1 < x_2$, $f(x_1) < f(x_2)$



Concavity and Convexity

Imagine you draw a line between two points along a function. A function (or segment of a function) is concave if this line is below the function, and convex if the line is above the function.

- concave: $\frac{f(x_1)+f(x_2)}{2} < f\left(\frac{x_1+x_2}{2}\right)$
- convex: $\frac{f(x_1)+f(x_2)}{2} > f\left(\frac{x_1+x_2}{2}\right)$

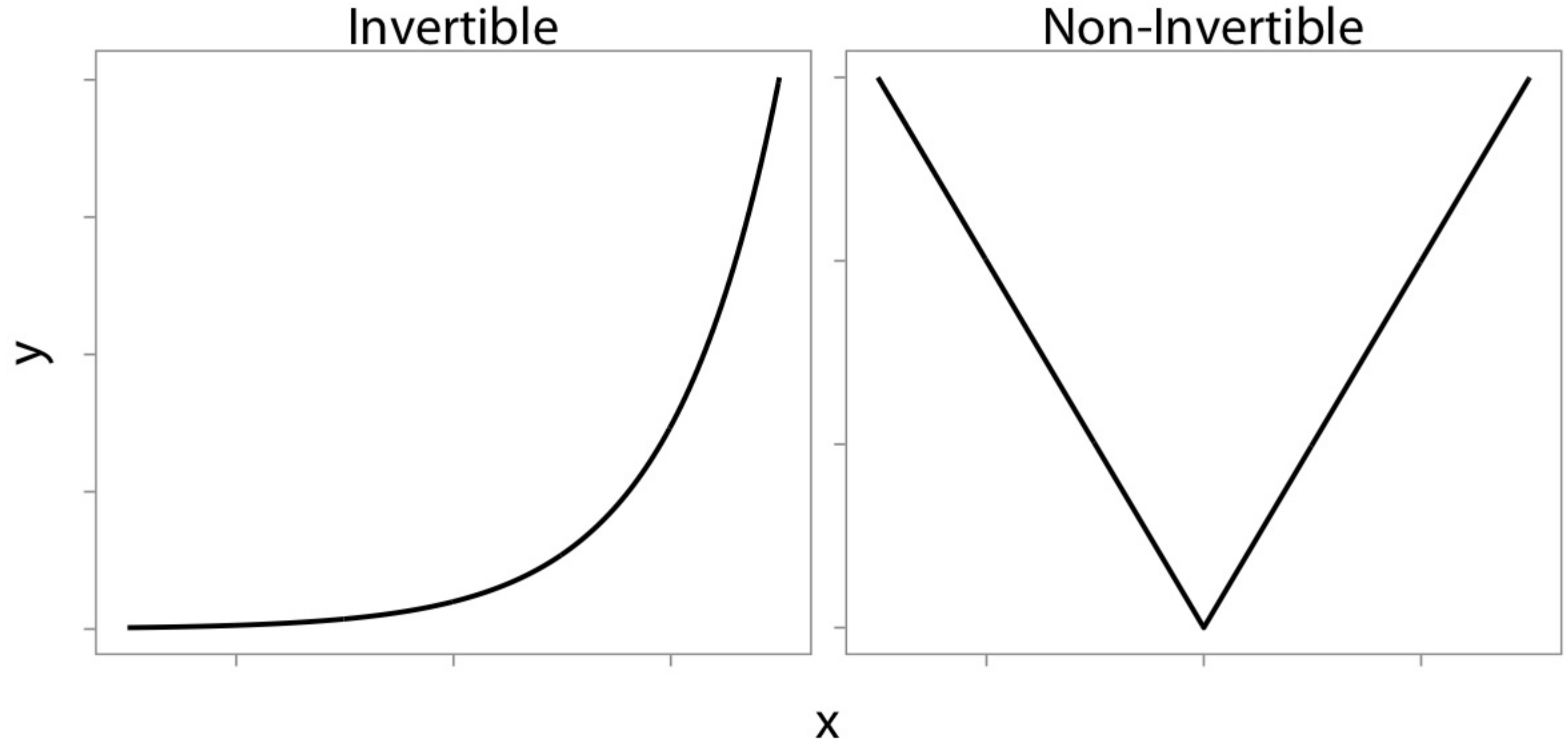


Invertible functions

A function maps an input to an output. A function is invertible if there exists a reverse function that maps the output back to the input.

Formally: if $y = f(x)$, then $f^{-1}(y) = x$

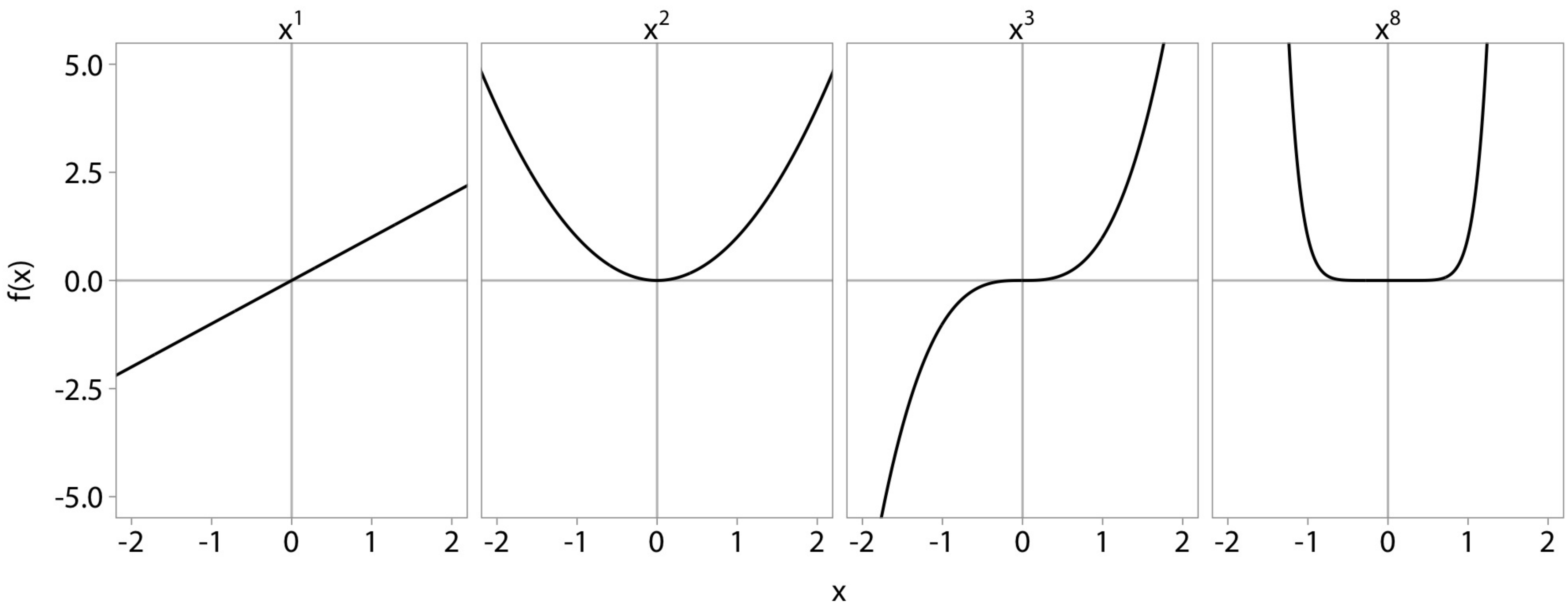
Also: $f^{-1}(f(x)) = x$



Exponents

The exponent operator multiplies a number by itself the number of times indicated in the exponent

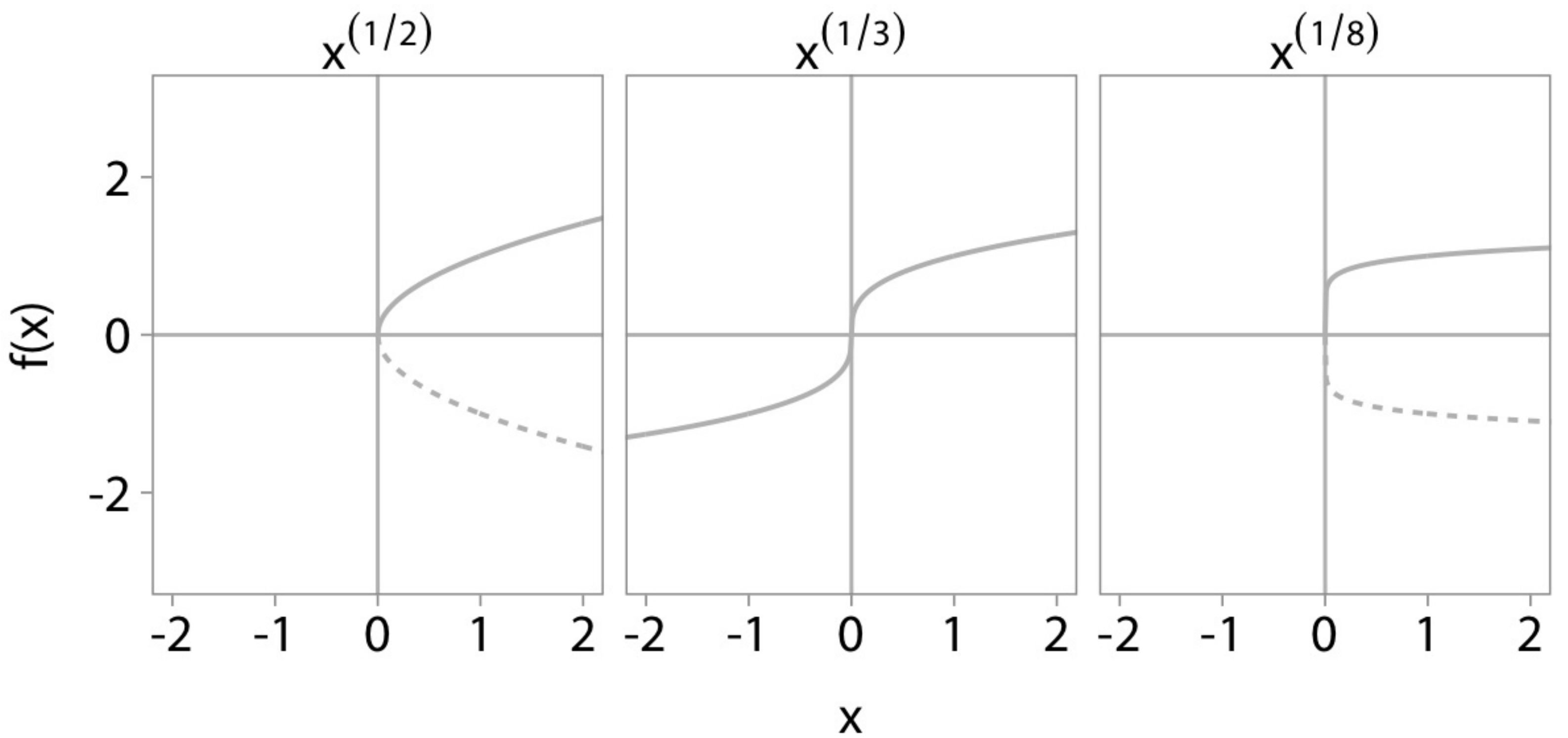
$$x^n = x * x * \ldots * x \quad (n \text{ times})$$



Roots

Root operators return the number that, when multiplied by itself the number of times indicated in the root, is equal to the input. When no number is given, that indicates the square root.

$$x = \sqrt[n]{x} * \sqrt[n]{x} \ldots \sqrt[n]{x} \quad (n \text{ times})$$



Exponents and Roots

All roots can be expressed as exponents (with the same properties)

$$\sqrt[n]{x} \equiv x^{\frac{1}{n}}$$

Important properties for exponents and roots

Zeroth power	$x^0 = 1$
Negative powers	$x^{-n} = \frac{1}{x^n}$
Inversion using exponents	$x^{-1} = \frac{1}{x}$
Distribution of powers (multiplication)	$(x * y)^n = x^n * y^n$
Distribution of powers (division)	$(\frac{x}{y})^n = \frac{x^n}{y^n}$
Product of powers	$x^n * x^m = x^{n+m}$
Nested powers	$(x^n)^m = x^{n*m}$

Two important continuous, monotonic, invertible functions: Exponentials and Logarithms

Exponentials and Logarithms

The logarithm (log) of some value y (with base b) is...

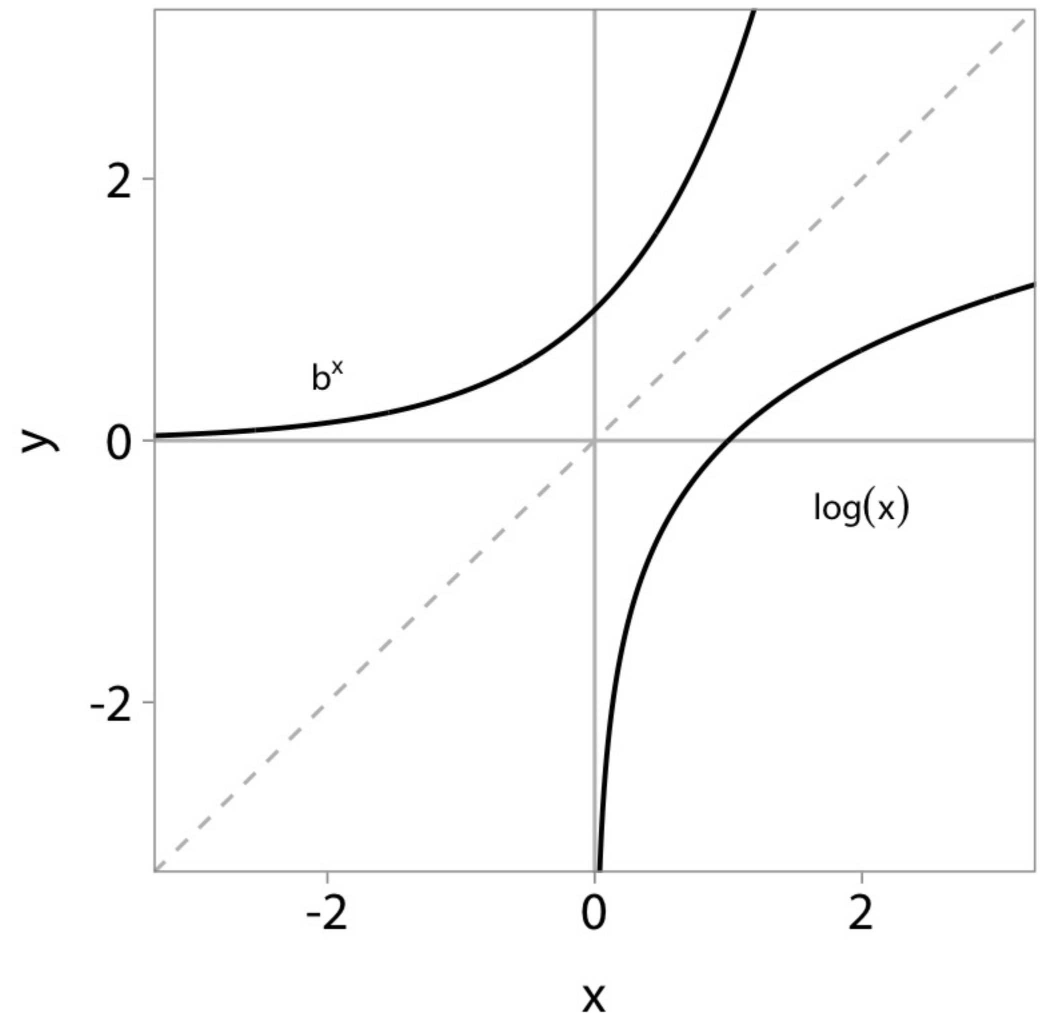
...the power to which that base would need to be raised to equal y

$$\text{If } b^x = y \text{, then } \log_b(y) = x$$

Exponentials and logarithms are inverse functions. Logs "undo" exponentials, and exponentials "undo" logs.

Exponentials and Logarithms

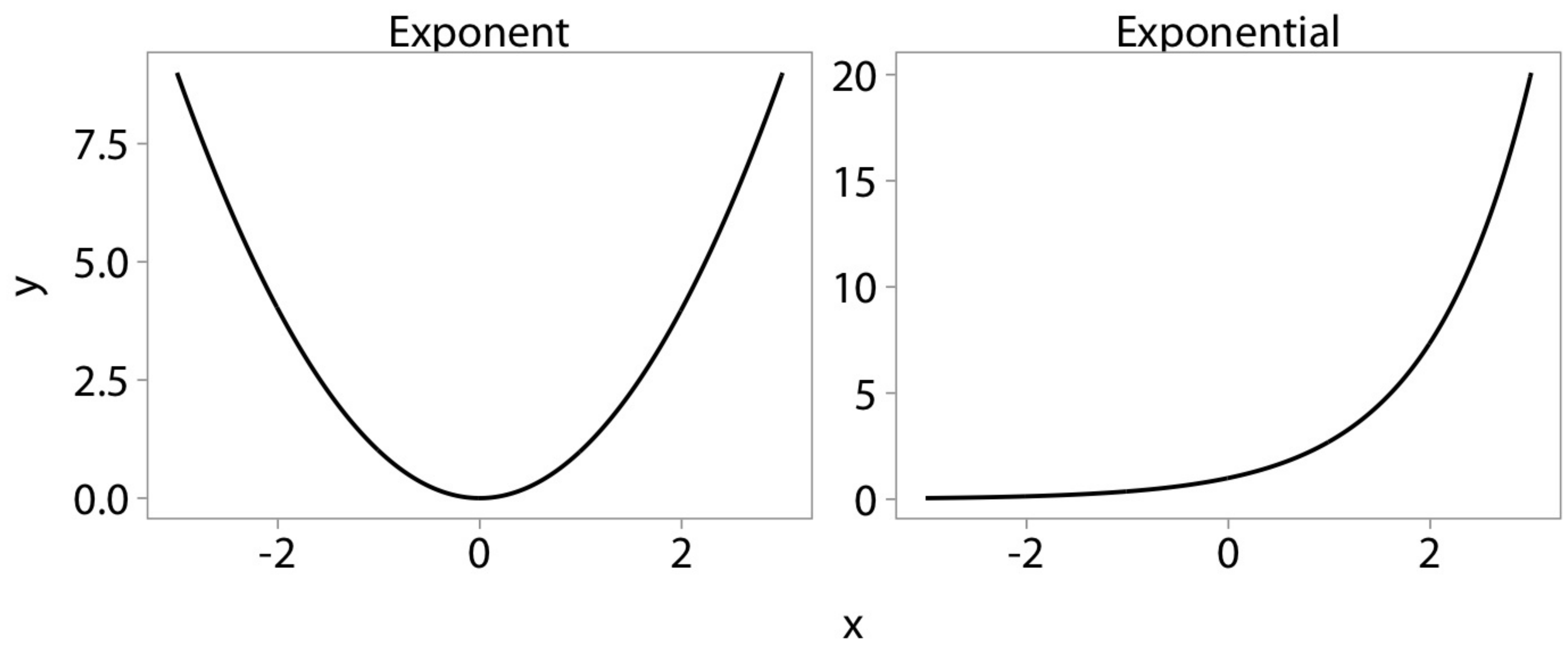
We can see this because exponentials and logs are reflections of each other over $y = x$ (one way to identify inverse functions)



Exponents $\neq$ Exponentials

Exponent: x^2 (x is the base)

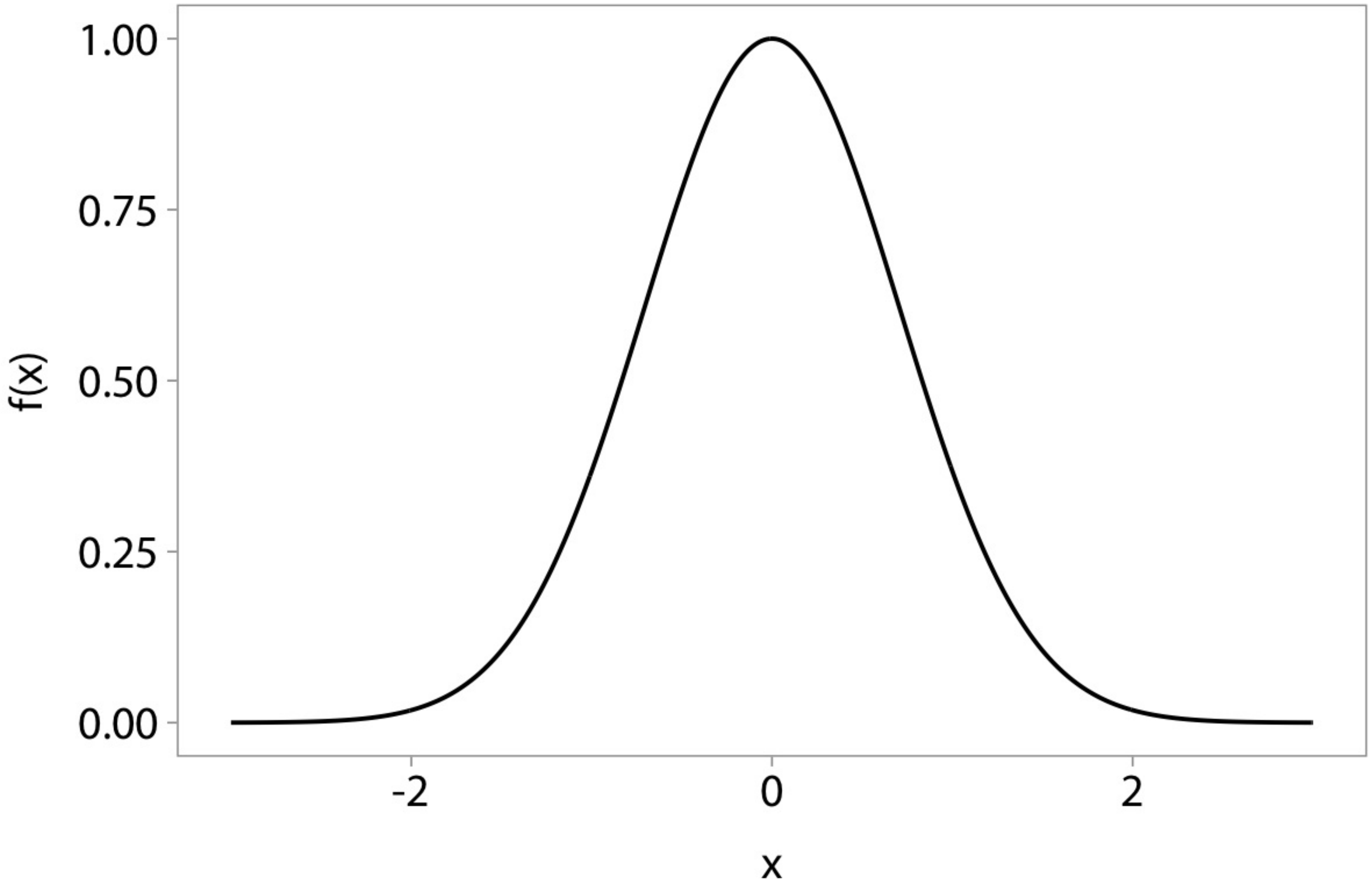
Exponentials: 2^x (x is the power)



Better yet...

What happens when you exponentiate a parabola?

$$f(x) = e^{-x^2}$$



Rules for logarithms

The following apply to all logs, regardless of base

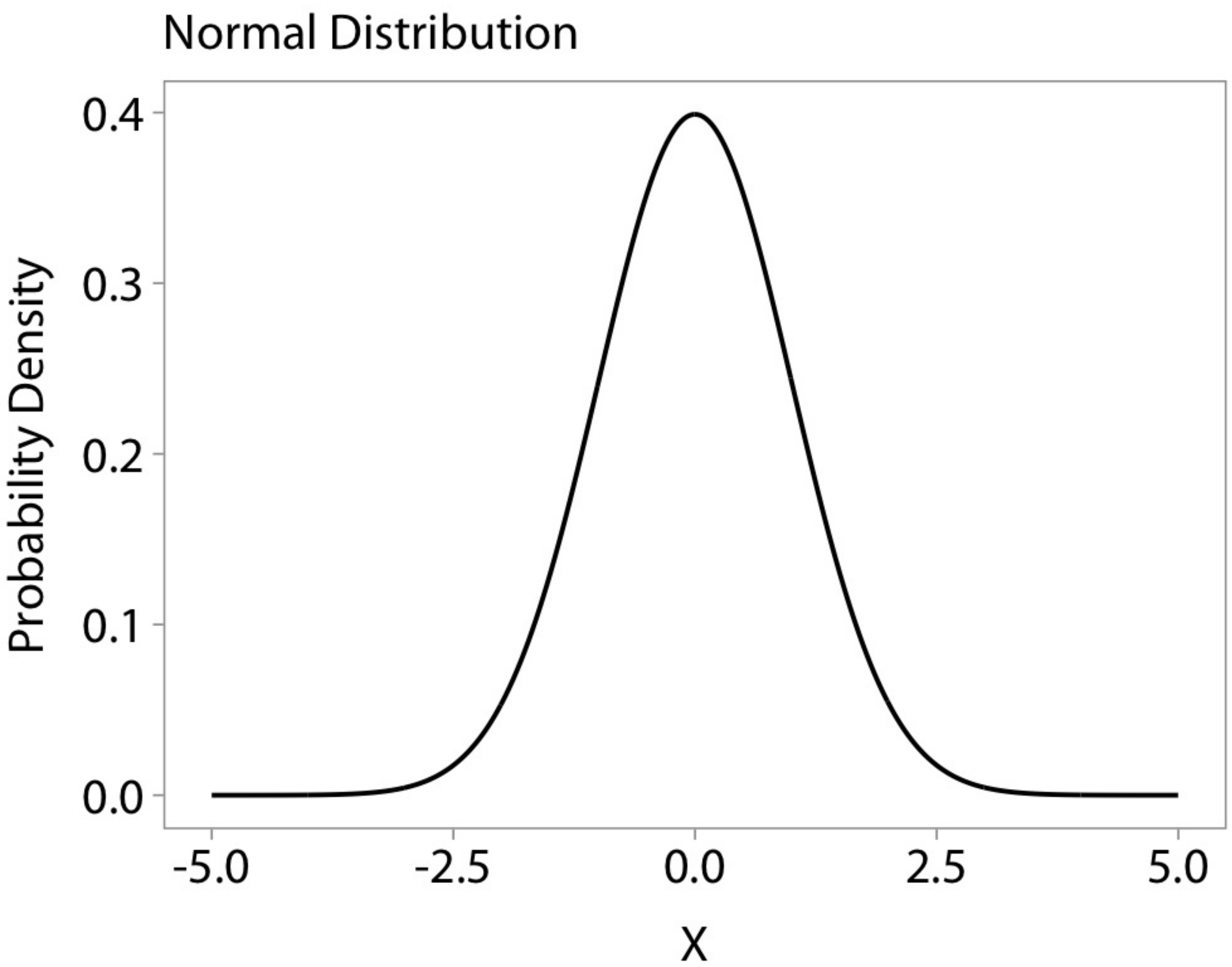
Log 1	$\log(1) = 0$
Log 0	undefined, approaches $-\infty$
Multiplication	$\log(x * y) = \log(x) + \log(y)$
Division	$\log(\frac{x}{y}) = \log(x) - \log(y)$
Exponentiation	$\log(x^a) = a * \log(x)$
Basis	$\log_b(b^x) = x$

(potentially) helpful video for understanding logs here

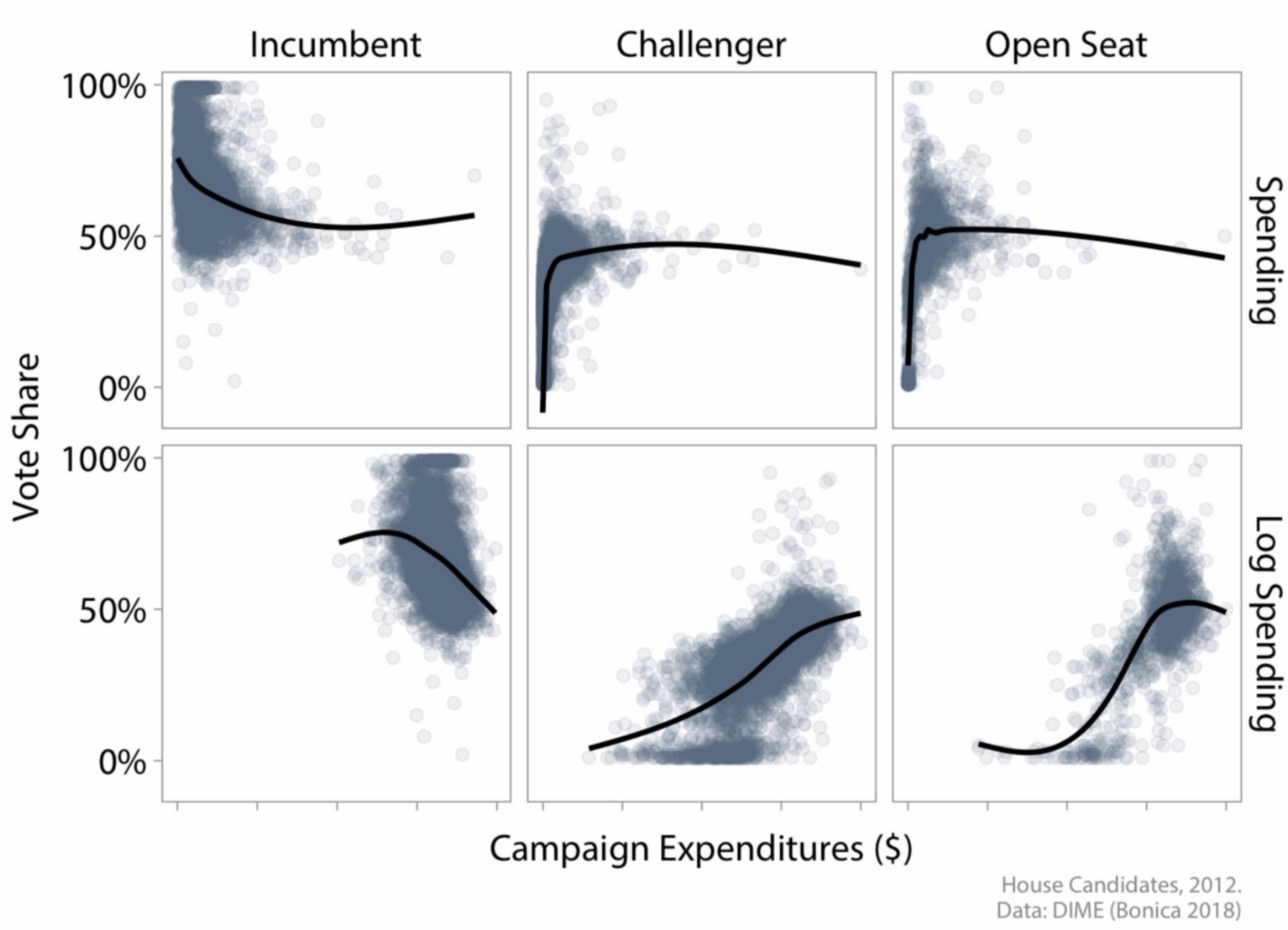
You WILL use logs and exponents

They are important for probability distributions

$$f(x \mid \mu, \sigma) = \frac{1}{2\pi\sigma^2} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$



Common transformations



They are helpful for analytic manipulation of equations

e.g. maximum likelihood estimation

We flip a coin $n = 5$ times and get $y = 4$ heads. What's the probability of a heads (π)?

Probability mass function for a binomial outcome (n independent trials, y successes, success probability π): $\Pr\left(y = k \mid \pi\right) = \binom{n}{k}\pi^k(1-\pi)^{n-k}$

Plug in our data: $\Pr\left(y = 4 \mid \pi\right) = \binom{5}{4}\pi^4(1-\pi)^{5-4}$

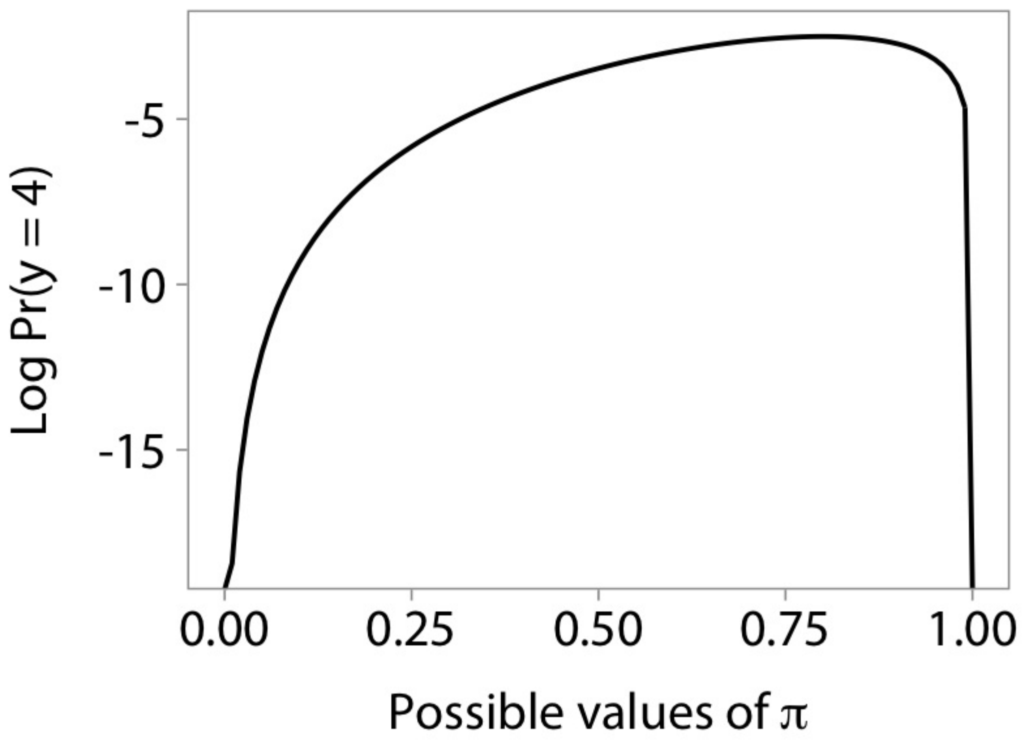
Log probability: $\log(\Pr(y = 4 \mid \pi)) = \log\binom{5}{4} + 4\log(\pi) + (5-4)\log(1-\pi)$

This function defines the (log) probability of our data, as a function of an unknown π.

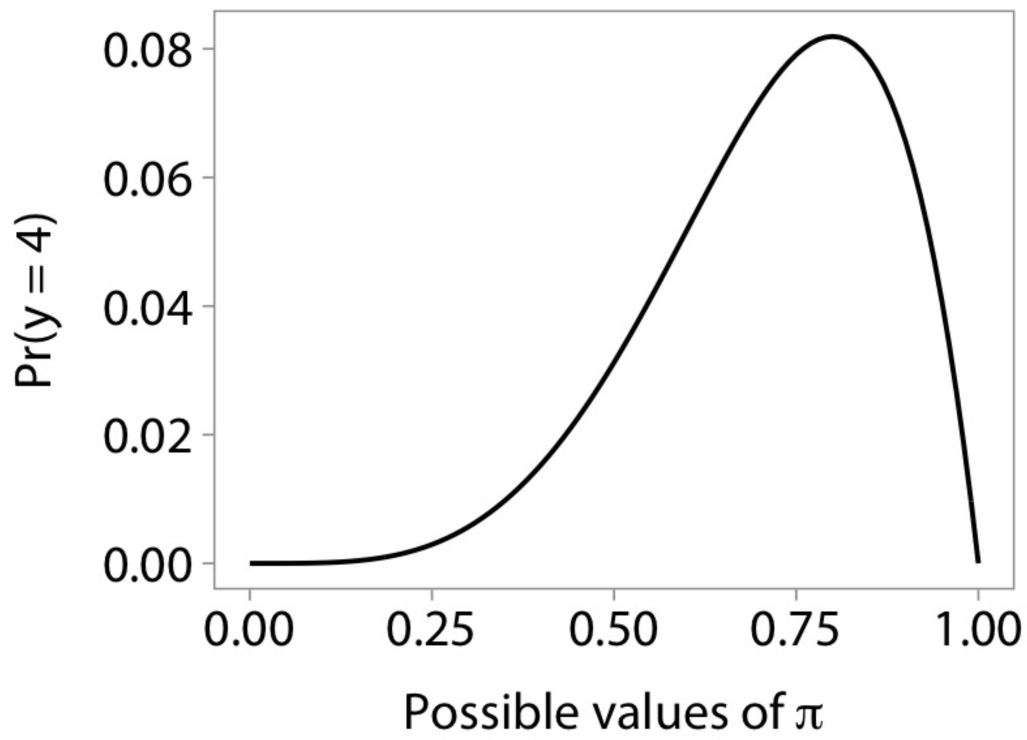
If we maximize this function with respect to π, we find the π value that gives us the greatest probability. That is, the most likely value of π that could give us these data.

This is how maximum likelihood works

Maximizing the log likelihood



On the unlogged scale



Point being: we use logs in MLE

(Note what the log transformation does to the y-axis)

Logs get easier with experience, which you will have

Base e

Although many early examples with logs use some arbitrary base (like base 10, the "common log"), most applications use base e ($\log_e$, the "natural log", $\ln$)

- e = Euler's number, approximately 2.718281828 ...

- $e^1 = e$
- $\ln(e) = 1$

- $e^0 = 1$
- $\ln(1) = 0$

Why base e?

It turns out that e is very important for functions that have constant and continuous growth rates (e.g. compounding interest)

Similar properties in probability

If you roll an n-sided die n times, the probability of getting side n exactly zero times approaches $\frac{1}{e}$ (as n increases)

We will emphasize the value of base e when we talk about probability (and the concept of odds)

Some practice with logs and exponents

Logs and Exponents

If $f(x) = \log_{10}(x)$, what is $f(1000)$?

- $\log_{10}(1000) = 3$, because $10^3 = 1000$

If $g(x) = \ln\left(\frac{x}{e}\right)$, what is $g(4)$?

- $$\begin{aligned} \ln\left(\frac{4}{e}\right) &= \ln(4) - \ln(e) \\ &\approx 1.386 - 1 \\ &\approx 0.386 \end{aligned}$$

If $h(x) = \log_2\left(x^5\right)$, what is $h(4)$?

- $$\begin{aligned} \log_2\left(4^5\right) &= 5\log_2(4) \\ &= 5 * 2 \\ &= 10 \end{aligned}$$

Solve:

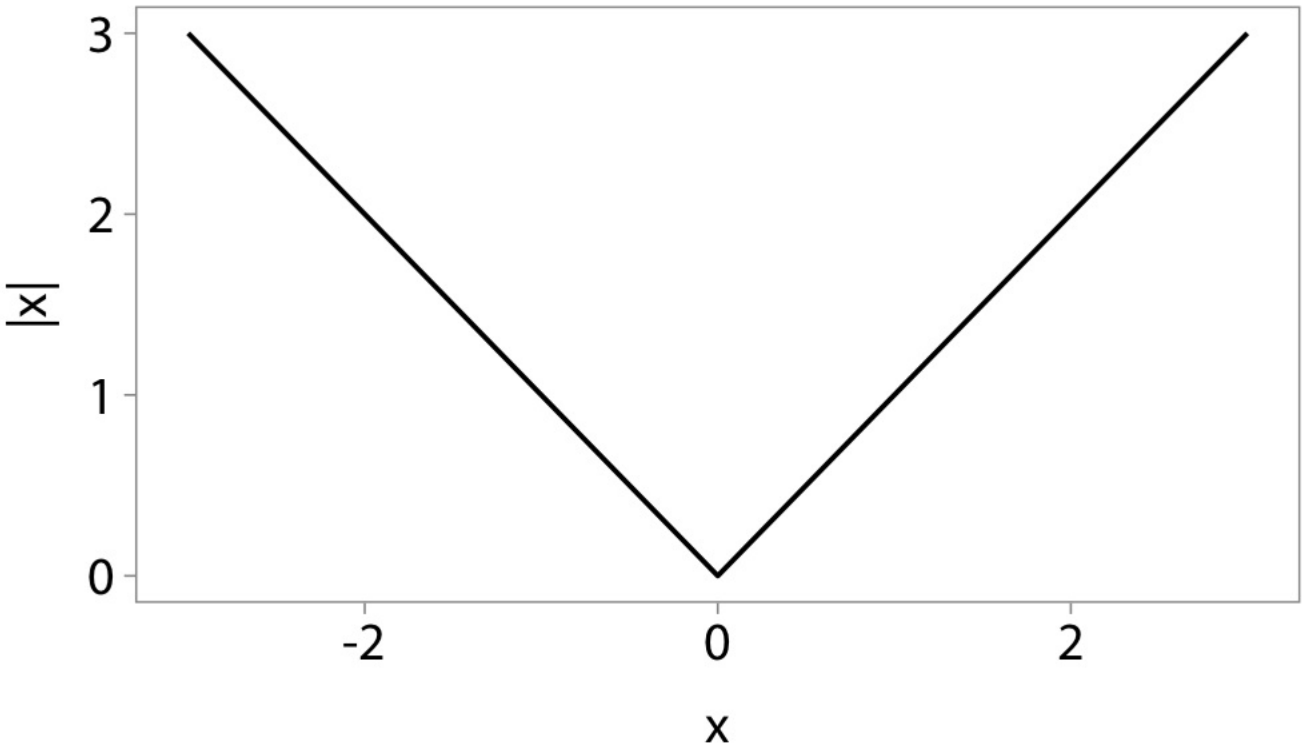
- $\log_2\left(4^3\right)$
- $\ln\left(\frac{x}{y} * q^4 * e\right)$

Now for something slightly easier

Absolute value

The absolute value operator returns the positive representation of a number

$$|x| = \begin{cases} x & \text{if } x \text{ is positive} \\ -x & \text{if } x \text{ is negative} \end{cases}$$



Multiplying polynomials

Polynomial: an expression with variables and coefficients, using only addition, subtraction, multiplication, and non-negative integer exponents

$$x^3 + 5x^2 + 4x^2 + 7$$

Which are variables, which are coefficients?

Multiply them by distributing: every element of each polynomial must be multiplied by every element of the other polynomial, then group terms by the powers of the variables.

Assume a, b, and c are coefficients, what is $ax\left(bx^2 + c\right)$?

$$ax\left(bx^2 + c\right) = \left(ax \cdot bx^2\right) + (ax \cdot c)$$

$$ax\left(bx^2 + c\right) = abx^3 + acx$$

Multiplying polynomials: FOIL

First, Outside, Inside, Last: for polynomials that each have two terms

$$(ax + b) \cdot (cx + d)$$

How do we FOIL?

- First: $ax \cdot cx$
- Outside: $ax \cdot d$
- Inside: $b \cdot cx$
- Last: $b \cdot d$

What do we get?

$$acx^2 + adx + bcx + bd$$

Longer polynomials

They work the same way, just keep track of all the terms. (FOIL is the same as distributing)

$$(2x^4 + 5x^3) \cdot (8x^2 + x + 3) = ?$$

$$= (2x^4 \cdot 8x^2) + (2x^4 \cdot x) + (2x^4 \cdot 3) + (5x^3 \cdot 8x^2) + (5x^3 \cdot x) + (5x^3 \cdot 3)$$

$$= 16x^6 + 2x^5 + 6x^4 + 40x^5 + 5x^4 + 15x^3$$

$$= 16x^6 + 42x^5 + 11x^4 + 15x^3$$

Polynomial practice

Find the products:

- $(x^2 + 3) \cdot (x - 2)$
- $(3p + 4q) \cdot (p - 2q)$

Factorials

The factorial operator (denoted with an exclamation mark $!$) returns the product of an integer with all lesser integers.

$$x! = x \cdot (x-1) \cdot (x-2) \cdot \ldots \cdot 2 \cdot 1$$

An example: $10! = 10 * 9 * 8 * 7 * \ldots * 3 * 2 * 1 = 3628800$

Special factorials:

$$1! = 1$$

$$0! = 1$$

These come in handy when we do probabilities (combinations and permutations)

Let's call it a day

Homework is online

https://github.com/shirikov/math-camp-2019