Day 1: Basics

Ikuma Ogura

Ph.D. student, Department of Government, Georgetown University

August 19, 2019

Why Do I Need to Study Mathematics?

- Statistical analysis/Formal modeling
  - Standard tools for political science research
  - Applied field of mathematics
- We study mathematics to...
  - read textbooks
  - read articles using statsitics/formal modeling

Real Stats

Specifically, the expression for the sum of squared residuals for any given estimates of $\hat{\beta}_0$ and $\hat{\beta}_1$ is

$$\sum_{i=1}^{N} \hat{\epsilon}_i^2 = \sum_{i=1}^{N} (Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i)^2$$

Game Theory for Applied Economists

where $y_i(e_i) = e_i + \varepsilon_i$. The first-order condition for (2.2.4) is

$$(w_H - w_L)\frac{\partial \text{Prob}\{y_i(e_i) > y_j(e_j^*)\}}{\partial e_i} = g'(e_i). \qquad (2.2.5)$$

That is, worker i chooses e_i such that the marginal disutility of extra effort, $g'(e_i)$, equals the marginal gain from extra effort, which is the product of the wage gain from winning the tournament, $w_H - w_L$, and the marginal increase in the probability of winning.

By Bayes' rule,[12]

$$\begin{aligned}
\text{Prob}\{y_i(e_i) > y_j(e_j^*)\} &= \text{Prob}\{\varepsilon_i > e_j^* + \varepsilon_j - e_i\} \\
&= \int_{\varepsilon_j} \text{Prob}\{\varepsilon_i > e_j^* + \varepsilon_j - e_i \mid \varepsilon_j\} f(\varepsilon_j)\, d\varepsilon_j \\
&= \int_{\varepsilon_j} [1 - F(e_j^* - e_i + \varepsilon_j)] f(\varepsilon_j)\, d\varepsilon_j,
\end{aligned}$$

Elements of Statistical Learning

orm. The least squares estimate of $a^T\beta$ is

$$\hat{\theta} = a^T\hat{\beta} = a^T(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y}. \qquad (3.17)$$

Considering $\mathbf{X}$ to be fixed, this is a linear function $\mathbf{c}_0^T\mathbf{y}$ of the response ector $\mathbf{y}$. If we assume that the linear model is correct, $a^T\hat{\beta}$ is unbiased ince

$$\begin{aligned} \mathrm{E}(a^T\hat{\beta}) &= \mathrm{E}(a^T(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y}) \\ &= a^T(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{X}\beta \\ &= a^T\beta. \qquad (3.18) \end{aligned}$$

No Worries!

Specifically, the expression for the sum of squared residuals for any given estimates of $\hat{\beta}_0$ and $\hat{\beta}_1$ is

$$\sum_{i=1}^{N} \hat{\epsilon}_i^2 = \sum_{i=1}^{N} (Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i)^2$$

= Today

No Worries!

where $y_i(e_i) = e_i + \varepsilon_i$. The first-order condition for (2.2.4) is

$$(w_H - w_L)\frac{\partial \text{Prob}\{y_i(e_i) > y_j(e_j^*)\}}{\partial e_i} = g'(e_i). \qquad (2.2.5)$$

Tuesday + Wednesday

That is, worker i chooses e_i such that the marginal disutility of extra effort, $g'(e_i)$, equals the marginal gain from extra effort, which is the product of the wage gain from winning the tournament, $w_H - w_L$, and the marginal increase in the probability of winning.

By Bayes' rule,[12]

Friday

$$\begin{aligned}
\text{Prob}\{y_i(e_i) > y_j(e_j^*)\} &= \text{Prob}\{\varepsilon_i > e_j^* + \varepsilon_j - e_i\} \\
&= \int_{\varepsilon_j} \text{Prob}\{\varepsilon_i > e_j^* + \varepsilon_j - e_i \mid \varepsilon_j\} f(\varepsilon_j)\, d\varepsilon_j \\
&= \int_{\varepsilon_j} [1 - F(e_j^* - e_i + \varepsilon_j)] f(\varepsilon_j)\, d\varepsilon_j,
\end{aligned}$$

Wednesday

No Worries!

form. The least squares estimate of $a^T\beta$ is

$$\hat{\theta} = a^T\hat{\beta} = a^T\underline{(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T}\mathbf{y}. \quad \text{Thursday} \tag{3.17}$$

Considering $\mathbf{X}$ to be fixed, this is a linear function $\mathbf{c}_0^T\mathbf{y}$ of the response vector $\mathbf{y}$. If we assume that the linear model is correct, $a^T\hat{\beta}$ is unbiased since

$$\begin{aligned} \mathrm{E}(a^T\hat{\beta}) &= \mathrm{E}(a^T(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y}) \\ &= a^T(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{X}\beta \\ &= a^T\beta. \end{aligned} \tag{3.18}$$

Note

- The mathematics you need depends on what kinds of research you want to do.
- This lecture will cover basic mathematics that you often encounter.

Do I Need to Memorize Formulas?

- NO!
- Rather focus on understanding the meanings of the formulas and when/why/how you should use them.

Today

- Basics
  - Algebra review
  - Set
  - Variable
  - Functions
    - Graphing equations/inequalities
    - Exponential and log functions
  - Summation & product operators

- (Today's lecture is a bit all over the place...)

Algebra Review: Order of Calculation

- Perform the following calculations.
    1. $3 + 5 \times 2 - 17 =$
    2. $5 \times (7 - 4) + (4 - 2) =$
    3. $(8 - 3) - (9 - 13) \times 4 =$

Algebra Review: Fraction

- Perform the following calculations.
  1. $\frac{2}{3} - \frac{1}{5} =$
  2. $\frac{4}{5} \times \frac{2}{3} =$
  3. $\frac{2}{5} \div \frac{3}{8} =$
  4. $\frac{\frac{2}{7}}{\frac{3}{4}} =$

Algebra Review: Power, Root, Factorial, & Absolute value

- Perform the following calculations.
  1. $2^3 \times 4^2 =$
  2. $(2^3)^2 =$
  3. $3^{-1} =$
  4. $\sqrt[3]{8} =$
  5. $5! =$
  6. $|-3| =$

Set

- **Set**: a collection of distinct objects
- **Element**: an object consisting a set

Set: Notation

- We write elements of a set within $\{\}$.
  - $A = \{1, 2, 4, 6, 7\}$
- When we can write elements of a set using a general form, we use the notation $\{\text{general form}|\text{definition}\}$
  - e.g., $B = \left\{x^2|x \text{ is an integer}, 1 \leq x \leq 4\right\}$
- If x is an element of the set A, we write $x \in A$.
- If x is not an element of the set A, we write $x \notin A$.

Set: Subset

- If every element of A is also in B, A is called the **subset** of B, and denoted as $A \subseteq B$.

- Example: Let $A = \{1, 2, 4, 6, 7\}$ and $B = \{1, 2, 7\}$. Then, $B \subseteq A$.

Set Operation

- **Intersection**: the intersection of A and B, denoted as $A \cap B$, is the set of common elements to both sets.
- **Union**: the union of A and B, denoted as $A \cup B$, is the set of all elements contained in either set.
- **Complement**: the complement of A, denoted as A^c, is the set that contains elements that are not contained in A.
- **Difference**: the difference of sets A and B, denoted as A/B, is the set of elements in A but not in B.
  - A/B is the same as $A \cap B^c$.
  - A/B is generally different from B/A.

Set: Number System

- **Natural numbers** ($\mathbb{N}$): set of 0 and positive integers ($\{0, 1, 2, \cdots\}$)

- **Integers** ($\mathbb{Z}$): set of all integers ($\{\cdots, -2, -1, 0, 1, 2, \cdots\}$)

- **Rational numbers** ($\mathbb{Q}$): set of all numbers which can be represented as the fraction of integers

- **Real numbers** ($\mathbb{R}$): set of all numbers whose squares are larger than or equal to 0.
  - Real numbers which cannot be represented as the fraction of integers (e.g., $\sqrt{2}$) are called the **irrational numbers**

- Using the set notation, $\mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R}$.

Set: Additional Notations/Terms

- **Empty set**: a set with no elements, denoted as $\emptyset$
- If sets A and B have no elements in common (i.e., $A \cap B = \emptyset$), we say A and B are **disjoint**
- We denote the union and the intersection of a sequence of sets $(A_1, A_2, \cdots, A_n)$ as follows.

$$\bigcap_{i=1}^{n} A_i = A_1 \cap A_2 \cap \cdots \cap A_n$$

$$\bigcup_{i=1}^{n} A_i = A_1 \cup A_2 \cup \cdots \cup A_n$$

Set: Open/Closed Intervals

- **Interval**: a set of real numbers such that any numbers between the two points in the set are also included in the set.

- **Open interval** is an interval that does not include the end points as elements. e.g., $A = \{a < i < b | i \in \mathbb{R}\}$
  - We denote open intervals using parentheses ().
  - e.g., $A = (a, b)$

- **Closed interval** is an interval that include the end points as elements. e.g., $A = \{a \leq i \leq b | i \in \mathbb{R}\}$
  - We denote closed intervals using square brackets [].
  - e.g., $A = [a, b]$

Set: Exercises

1. List all the elements of the set $A = \{x^3 | x \in \mathbb{Z}, -2 \leq x \leq 2\}$

2. Let $A = \{-3, 1, 4, 6, 8, 13\}$ and $B = \{-5, -3, 1, 8, 11, 13\}$. Then what is
   2.1 $A \cap B$
   2.2 $A \cup B$
   2.3 A/B
   2.4 B/A

Variable/Constant

- **Variable**: a symbol which represent an arbitrary number
  - value of a variable is not fully specified, so it can change/vary
- **Constant**: a quantity whose value does not change

Basic Number Properties

- **Commutative Property**: the order of addition/multiplication does not affect the outcome

$$\begin{aligned} a + b &= b + a \\ a \times b &= b \times a \end{aligned}$$

- **Associative Property**: the order of addition/multiplication does not matter as long as the sequence of operation is not changed

$$\begin{aligned} (a + b) + c &= a + (b + c) \\ (a \times b) \times c &= a \times (b \times c) \end{aligned}$$

- **Distributive Property**: distribution of multiplication over addition/subtraction

$$a(b + c) = ab + ac$$

Expansion

- Remove parentheses () from the product of polynomials.

- Monomial/Polynomial
  - **Monomial**: product of a constant and a variable raised to some value, which looks like ax^k
    - ⋆ The constant (a) is called the **coefficient**, and the number of exponent (k) is called the **degree** of the monomial
  - **Polynomial**: sum of monomials
    - ⋆ e.g., $2x + 3$, $x^3 - 4x + 5$
    - ⋆ The degree of a polynomial is the highest degree of its monomials

- How to perform expansion?
  - Repeatedly apply the distribution rule!

Expansion (cont.)

- Question: Multiply out $(2x+3)(x^2-8x+5)$
  - Answer:

$$\begin{aligned} & (2x+3)(x^2-8x+5) \\ = \;& 2x(x^2-8x+5)+3(x^2-8x+5) \\ = \;& 2x^3-16x^2+10x+3x^2-24x+15 \\ = \;& 2x^3-13x^2-14x+15 \end{aligned}$$

Expansion (cont.)

- Question: Multiply out $(x+2)^3$
  - Answer:

$$\begin{aligned} & (x+2)^3 \\ = \; & (x+2)(x+2)^2 \\ = \; & (x+2)\left\{x(x+2)+2(x+2)\right\} \\ = \; & (x+2)(x^2+4x+4) \\ = \; & x(x^2+4x+4)+2(x^2+4x+4) \\ = \; & x^3+6x^2+12x+8 \end{aligned}$$

Factoring

- **Factoring**: writing a polynomial as a product of polynomials of lower degrees.
- Factoring is the reverse of expanding parentheses.
- (Factoring is a bit harder than expansion, and you need to get used to it...)
- Tips
  - If you find common factors, try grouping the terms containing them
    - e.g., $x^3 - 5x^2 = x^2 \times x + x^2 \times (-5) = x^2(x-5)$
  - Try applying the rules on the next page

Factoring/Expansion Rules

- Some of the rules you often encounter:
    1. $(x+a)(x+b) = x^2 + (a+b)x + ab$
    2. $(ax+b)(cx+d) = acx^2 + (ad+bc)x + bd$
    3. $(x+a)^2 = x^2 + 2ax + a^2$
    4. $(x-a)^2 = x^2 - 2ax + a^2$
    5. $(x+a)(x^2 - ax + a^2) = x^3 + a^3$
    6. $(x-a)(x^2 + ax + a^2) = x^3 - a^3$
    7. $(x+y+z)^2 = x^2 + y^2 + z^2 + 2xy + 2yz + 2zx$

- (Again, you don't neet to memorize them...)

Factoring/Expansion: Exercises

1. Multiply out:
   1.1 $(x+6)(x-7) =$
   1.2 $(x^2+3)(x^2+x+7) =$
2. Factor:
   2.1 $x^2 - 2x - 3 =$
   2.2 $6x^2 - x - 1 =$
   2.3 $x^3 - 8 =$
   2.4 $x^3 - 3x^2 - 10x =$

Solving Equations/Inequalities

- This is when we want to find values of a variable satisfying an equation/range of values of a variable satisfying an inequality
- Tips
  - Rearrange terms so that the variable of interest is isolated
  - Make sure to perform the same operations on both sides of equality/inequality
    - Be careful when multiplying by negative numbers in solving inequalities!
  - Check answer

Solving Equations/Inequalities (cont.)

- Question: Temperatures in Celsius (say x) is converted to those in Fahrenheit (y) using the following equation.

$$y = \frac{9}{5}x + 32$$

Then,

1. Find the temperature when both are equal.
2. When do temperatures measured in Fahrenheit are higher than those in Celsius?

Solving Quadratics

- After you rearrange the terms in the form $(\mathrm{i})ax^2 + bx + c = 0/(\mathrm{ii})ax^2 + bx + c \leq 0/(\mathrm{iii})ax^2 + bx + c \geq 0...$

- If you can easily factor as $a(x + \alpha)(x + \beta)$ $(\alpha > \beta, a > 0)$, then
  - (i): $(x + \alpha) = 0$ and/or $(x + \beta) = 0 \Rightarrow x = -\alpha, -\beta$
  - (ii): $(x + \alpha) \geq 0$ and $(x + \beta) \leq 0 \Rightarrow -\alpha \leq x \leq -\beta$
  - (iii): $(x + \alpha) \leq 0$ or $(x + \beta) \geq 0 \Rightarrow x \leq -\alpha, x \geq -\beta$
  - How should we do when $a < 0$?

- When we cannot easily factor... $\rightarrow$ **Complete the square**
  - Transform the quadratic into the form $a(x \pm \alpha)^2 = \beta$ and solve for x by taking the square root of those terms.

Solving Quadratics: Complete the Square

1. Move the constant to the RHS, and divide the equation by the coefficient on the squared term
   - $x^2 + \frac{b}{a}x = -\frac{c}{a}$
2. Divide the coefficient on x by 2, square the value, and add to both sides
   - $x^2 + 2 \cdot \frac{b}{2a}x + \frac{b^2}{4a^2} = -\frac{c}{a} + \frac{b^2}{4a^2}$
3. Factor the LHS into the form $(x \pm \alpha)^2$ and simplify the RHS
   - $(x + \frac{b}{2a})^2 = \frac{b^2-4ac}{4a^2}$
4. Take the square root of both sides
   - $x + \frac{b}{2a} = \pm\sqrt{\frac{b^2-4ac}{4a^2}} = \pm\frac{\sqrt{b^2-4ac}}{2a}$
5. Solve for x
   - $x = \frac{-b\pm\sqrt{b^2-4ac}}{2a}$

Solving Quadratics: Exercises

- Solve the following equations/inequalities for x
  1. $2x^2 - 4x - 3 = 0$
  2. $x^2 + x - 6 \leq 0$
  3. $x^3 = x^2 + 12x$

Function

- **Function** is a relation between sets which associate every element of the first set to exactly one element of the second set.
  - describes how the values of the LHS variable (outputs) changes with values of RHS variables (inputs)
  - input variables are often called the *inpdependent* variables
  - output variable is often called the *dependent* variable

- Domain, image, range
  - **Domain**: set over which a function is defined
  - **Image**: output value of the function
  - **Range**: set of the images of all elements of the domain

Function: Notation

- Let f be a function which relates every element $x \in X$ to exactly one element $y \in Y$. Then we denote the relationship as

$$y = f(x)$$

or

$$f : X \to Y$$

  - Other letters often used to denote functions: g, h, u, v...

Function: Examples

- $f(x) = x + 4$ $(x \in \mathbb{R})$

- $f(x) = x^2 + 2$ $(x \in \mathbb{R})$
  - What is the range of this function?

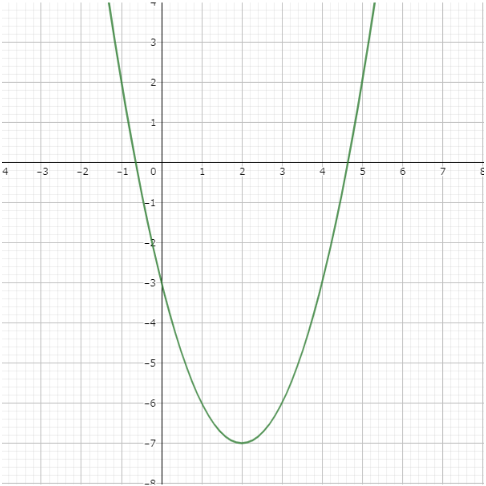
- $f(x_1, x_2) = 2x_1 - x_2 + 7$ $(x_1, x_2 \in \mathbb{R})$
  - f maps/associates every element $(x_1, x_2) \in X$ to exactly one element of $y \in Y$.

- $x^2 + y^2 = 1$ is not a function!
  - e.g., when $x = 0$, $y = \pm 1$

Graphing Functions

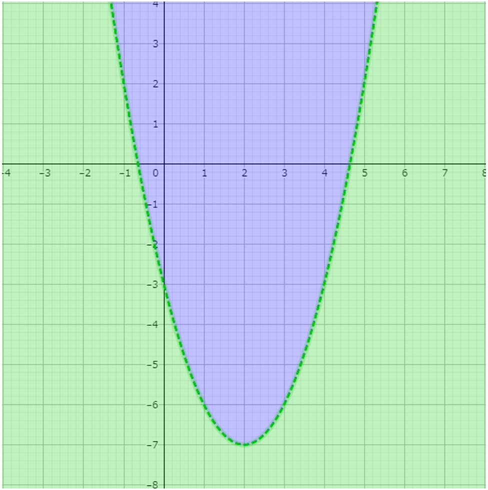
- Why do you want to graph a function?
  - Is it increasing/decreasing?
  - How fast is it increasing/decreasing?
  - ...
- Intersection of algebra and geometry

Graphing Functions (cont.)



- Graph of $y = f(x)$ is the collection of points satifying the relationship.
- Compute the points $(x_0, f(x_0))$ for multiple x_0 and connect them!

Graphing Functions (cont.)



- Areas above the curve (areas in blue): set of points satisfying the relationships $y > f(x)$
- Areas below the curve (areas in green): set of points satisfying the relationships $y < f(x)$

Graphing Functions (cont.)

- Linear function (polynomial of degree 1): $y = a + bx$
  - Must pass the y-intercept $(0, a)$.
- Quadratic function
  - If we can factor as $a(x+b)(x+c) \rightarrow$ must pass the x-intercepts $(-b, 0)$ and $(-c, 0)$
  - Complete the square and get $a(x+b)^2 + c \rightarrow$ must pass the point $(-b, c)$
    - ⋆ If $a > 0$: $(-b, c)$ is the global minimum
    - ⋆ If $a < 0$: $(-b, c)$ is the global maximum

Graphing Functions: Exercises

- Graph the following relationships.
  1. $y = x^2 - 2x + 5$
  2. $y = 2x^2 + 2x - 12$
  3. $y \geq \frac{3}{2}x - 4$

Functions Often Used in Social Science

- Exponential function
- Logarithmic function
- (Trigonometric function)

Exponential Function

- An exponential function of x is some constant (say a) raised to x: $f(x) = a^x$
  - $f(x) = 4^x$
  - $f(x) = (-2)^x$

- Euler's constant e
  - $e = \lim_{x\to\infty}(1 + \frac{1}{h})^h = 2.71828...$
  - $e^x = \exp(x)$

Exponential Function: Basic Rules

1. $a^{x_1} \times a^{x_2} = a^{x_1+x_2}$ $(x_1, x_2 \in \mathbb{R})$
2. $(a^{x_1})^{x_2} = a^{x_1 x_2}$
3. $a^0 = 1$
4. $a^{-x} = \frac{1}{a^x}$
5. $a^{\frac{1}{x}} = \sqrt[x]{a}$

Logarithmic Funtion

- Logarithm is the inverse of an exponential function. Logarithm of x to base a $(a \neq 1, a > 0)$ is the number such that a power of it equals to x. Formally,

$$y = \log_a x \Leftrightarrow x = a^y$$

- Log of x to base e is called the **natural log**, and often denoted as

$$\log_e x = \log x = \ln(x)$$

- Question: what is the domain of a log function?

Logarithmic Funtion: Basic Rules

1. $\log_a x^b = b \log_a x$
2. $\log_a \frac{1}{x} = \log_a x^{-1} = -\log_a x$
3. $\log_a xy = \log_a x + \log_a y$
4. $\log_a \frac{x}{y} = \log_a (x \times y^{-1}) = \log_a x - \log_a y$
5. $\log_x x = 1$
6. $\log_a 1 = \log_a a^0 = 0$
7. (Change of base) $\log_b x = \frac{\log_a x}{\log_a b}$

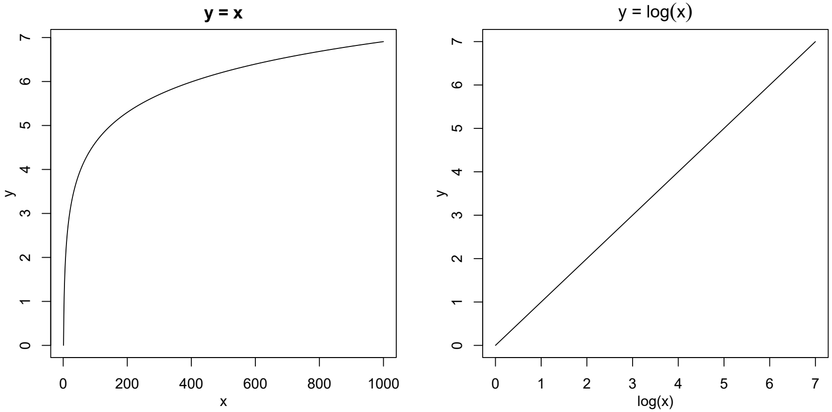
Exponential and Logarithmic Functions: Exercises

1. Simplify the following expressions.
   1.1 $2^5 \times 4^{\frac{3}{2}} =$
   1.2 $2\log x - \log(3x) =$
   1.3 $a^{\log_a x} =$
2. Derive the change of base formula. (Hint: by definition, x is represented as...)

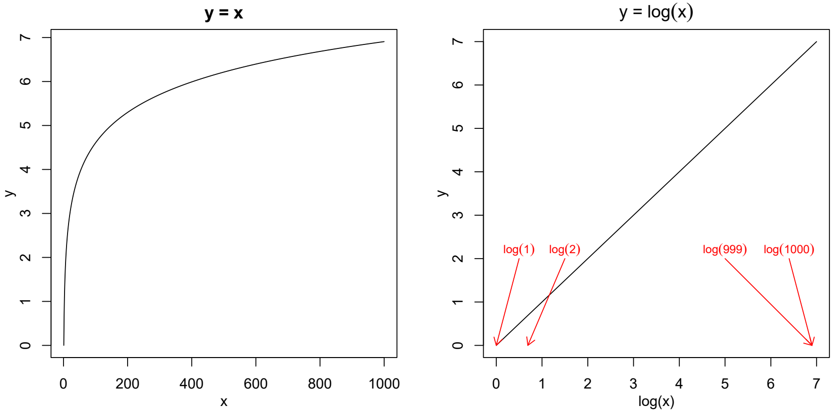
Why Do People Use Log?

- Log function is (one of) the most frequently used function(s) in political science.
- But why?
- Log function stretches out the scale when the value of the input is small while compresses the scale when it's large.
  - Dealing with skewness of the data
  - Express nonlinear relationship
  - ...

Why Do People Use Log? (cont.)



Why Do People Use Log? (cont.)



Summation and Product Operators

- Assume a sequence of variables/values $x_m, x_{m+1}, \cdots, x_n \ (m < n)$

- We denote their sum as follows.

$$\sum_{i=m}^{n} x_i = x_m + x_{m+1} + \cdots + x_n$$

- We denote their product as follows.

$$\prod_{i=m}^{n} x_i = x_m \times x_{m+1} \times \cdots \times x_n$$

Summation and Product Operators: Examples

- $\sum_{i=1}^{10} i = 1 + 2 + \cdots + 10 = 55.$
- $\prod_{i=1}^{7} i = 7! = 840.$
- $\sum_{i=1}^{6} x_i = x_1 + x_2 + x_3 + x_4 + x_5 + x_6.$
- $\prod_{j=2}^{5} x_j = x_2 \times x_3 \times x_4 \times x_5.$

Summation and Product Operators: Exercises

1. Perform the following calculation.
   1.1 $\sum_{i=3}^{7} i^2$
   1.2 $\sum_{i=1}^{n} a$

2. Show that
   - $\sum_{i=1}^{n}(x_i + y_i) = \sum_{i}^{n} x_i + \sum_{i}^{n} y_i$
   - $\sum_{i=1}^{n}(ax_i + b) = a\sum_{i=1}^{n} x_i + nb$

Tomorrow

- Problem set 1 $\rightarrow$ review in the morning

- Tomorrow
  - Limits
  - Derivative
  - Unconstrained optimization
  - Moore & Siegel, Chapters 4-6 & 8