# Observations to functions

Sanghoon Kim

Department of Political Science
University of Illinois at Urbana-Champaign

# Overview

- Concepts, variables, functions
- Linear functions
- Non-linear functions

# Observations to questions

Sanghoon is interested in studying why some dictators stay in power longer than others. He came up with an answer: since oil-rich countries tend to be dictatorship, there's something with **oil revenue**. The more income the dictator makes from oil, the longer he will stay in office. He brought this idea to his advisor and she said, "OK, it might be interesting. Can you make it into a function so that I can have a better sense of how to operationalize it?" Sanghoon looks away from his advisor to hide his frustration.

# Concepts to variables

Sanghoon is interested in why **some dictators stay in power longer** than others. He came up with an answer: since oil-rich countries tend to be dictatorship, there's something with **oil revenue**. The more income the dictator makes from oil, the longer he will stay in office.

# Concepts to variables

Sanghoon is interested in why **some dictators stay in power longer** than others. He came up with an answer: since oil-rich countries tend to be dictatorship, there's something with **oil revenue**. The more income the dictator makes from oil, the longer he will stay in office.

$$\text{Oil revenue} \rightarrow \text{A dictator's survival}$$

# Concepts to variables

Sanghoon is interested in why **some dictators stay in power longer** than others. He came up with an answer: since oil-rich countries tend to be dictatorship, there's something with **oil revenue**. The more income the dictator makes from oil, the longer he will stay in office.

$$\text{Oil revenue} \rightarrow \text{A dictator's survival}$$

- What are you interested in studying? What is your theory?

# Concepts to variables

Sanghoon is interested in why **some dictators stay in power longer** than others. He came up with an answer: since oil-rich countries tend to be dictatorship, there's something with **oil revenue**. The more income the dictator makes from oil, the longer he will stay in office.

$$\text{Oil revenue} \rightarrow \text{A dictator's survival}$$

- What are you interested in studying? What is your theory?

- We call the item on each side as a *variable*. What is a variable?
- Political scientists are interested in **concepts** such as participation, voting, democracy, international cooperation, war, etc.
- **Concepts** are inventions that we create to understand the world.

# Concepts to variables

Sanghoon is interested in why **some dictators stay in power longer** than others. He came up with an answer: since oil-rich countries tend to be dictatorship, there's something with **oil revenue**. The more income the dictator makes from oil, the longer he will stay in office.

$$\text{Oil revenue} \rightarrow \text{A dictator's survival}$$

- What are you interested in studying? What is your theory?

- We call the item on each side as a *variable*. What is a variable?
- Political scientists are interested in **concepts** such as participation, voting, democracy, international cooperation, war, etc.
- **Concepts** are inventions that we create to understand the world.
- **Variables** are the indicators we develop to measure concepts.

# Variables to functions

- We use *variables* to concepts so that we may use them in formal mathematical expressions.
- *Variables* need to *vary*, i.e. take different values.
  e.g. Which of the following is not a variable?

# Variables to functions

- We use *variables* to concepts so that we may use them in formal mathematical expressions.
- *Variables* need to *vary*, i.e. take different values.
  e.g. Which of the following is not a variable?
  1. Economic growth $\rightarrow$ Civil war onset
  2. Education $\rightarrow$ Partisanship
  3. Nazism $\rightarrow$ World War II

# Variables to functions

- We use *variables* to concepts so that we may use them in formal mathematical expressions.
- *Variables* need to *vary*, i.e. take different values.
  e.g. Which of the following is not a variable?
  1. Economic growth $\rightarrow$ Civil war onset
  2. Education $\rightarrow$ Partisanship
  3. Nazism $\rightarrow$ World War II
- Types of variables
  1. Continuous variable: a variable with infinite number of values. (e.g. economic growth, ethnic diversity (Herfindahl index))
  2. Binary variable: a variable with two possible states (e.g. civil war onset, regime type (DD index))
  3. Nominal variable: containing several categories (e.g. partisanship)
  4. Ordinal variable: a variable used to rank a sample (e.g. education)

# Basics of functions

- Question $\rightarrow$ Concepts $\rightarrow$ Variables $\rightarrow$ Functions

# Basics of functions

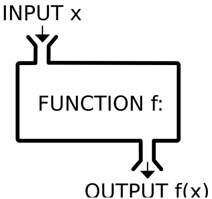
- Question → Concepts → Variables → Functions
- What is a function?
  A mathematical function is a “mapping” (i.e., specific directions), which gives a correspondence from one measure onto exactly one other for that value.
- $f(x) : A \to B$

# Basics of functions

- Question $\rightarrow$ Concepts $\rightarrow$ Variables $\rightarrow$ Functions
- What is a function?
  A mathematical function is a "mapping" (i.e., specific directions), which gives a correspondence from one measure onto exactly one other for that value.
- $f(x) : A \rightarrow B$
- So, a function is **a mapping from one defined space to another**, such as $f : \mathbb{R} \rightarrow \mathbb{R}$ , in which f maps the real numbers to the real numbers (i.e., $f(x) = 2x$), or $f : \mathbb{R} \rightarrow \mathbb{I}$, in which f maps the real numbers to the integers.

# Basics of functions

- Question → Concepts → Variables → Functions
- What is a function?
  A mathematical function is a "mapping" (i.e., specific directions), which gives a correspondence from one measure onto exactly one other for that value.
- $f(x) : A \rightarrow B$
- So, a function is **a mapping from one defined space to another**, such as $f : \mathbb{R} \rightarrow \mathbb{R}$ , in which f maps the real numbers to the real numbers (i.e., $f(x) = 2x$), or $f : \mathbb{R} \rightarrow \mathbb{I}$, in which f maps the real numbers to the integers.



# Variables to functions

*The more income the dictator makes from oil, the longer he can stay in office.*

# Variables to functions

> *The more income the dictator makes from oil, the longer he can stay in office.*

How can we write this into a function?

# Variables to functions

*The more income the dictator makes from oil, the longer he can stay in office.*

How can we write this into a function?

- We know that as one variable increases, the other one does, too.
  $\rightarrow$ A positive relationship

# Variables to functions

*The more income the dictator makes from oil, the longer he can stay in office.*

How can we write this into a function?

- We know that as one variable increases, the other one does, too.
  $\rightarrow$ A positive relationship
- $\text{Dictator's survival} = a \cdot \text{Oil revenue} + b, \quad \text{where } a, b > 0$

# Variables to functions

*The more income the dictator makes from oil, the longer he can stay in office.*

How can we write this into a function?

- We know that as one variable increases, the other one does, too. $\rightarrow$ A positive relationship
- $\textit{Dictator's survival} = a \cdot \textit{Oil revenue} + b, \quad \textit{where } a, b > 0$

  OR
- $y = ax + b$
  *where* $a, b > 0$, *$y$ denotes a dictator's survival and $x$ denotes oil revenue*

# Functions example

"We exploit differences in European mortality rates to estimate the effect of institutions on economic performance. Europeans adopted very different colonization policies in different colonies, with different associated institutions. In places where Europeans faced high mortality rates, they could not settle and were more likely to set up extractive institutions. These institutions persisted to the present." (Acemoglu, Johnson, and Robinson. 2001. "The Colonial Origins of Comparative Development: An Empirical Investigation." *American Economic Review*)

How can we write their argument into mathematical functions?

# Functions example

"We exploit differences in European mortality rates to estimate the effect of institutions on economic performance. Europeans adopted very different colonization policies in different colonies, with different associated institutions. In places where Europeans faced high mortality rates, they could not settle and were more likely to set up extractive institutions. These institutions persisted to the present." (Acemoglu, Johnson, and Robinson. 2001. "The Colonial Origins of Comparative Development: An Empirical Investigation." *American Economic Review*)

How can we write their argument into mathematical functions?

- Settler mortality rates $\rightarrow$ Institutional quality $\rightarrow$ Economic performance
- $InstitutionQuality_i = a + b \cdot SettlerMortality_i$

# Functions example

"We exploit differences in European mortality rates to estimate the effect of institutions on economic performance. Europeans adopted very different colonization policies in different colonies, with different associated institutions. In places where Europeans faced high mortality rates, they could not settle and were more likely to set up extractive institutions. These institutions persisted to the present." (Acemoglu, Johnson, and Robinson. 2001. "The Colonial Origins of Comparative Development: An Empirical Investigation." *American Economic Review*)

How can we write their argument into mathematical functions?

- Settler mortality rates → Institutional quality → Economic performance
- $InstitutionQuality_i = a + b \cdot SettlerMortality_i$
- $C_i = \lambda_c + \beta_c S_i + \mathbf{X}_i' \gamma_c + v_{ci}$

# Linear equations

A **linear equation** is an equation that contains only terms of order $x^1$ and $x^0 = 1$.

$$Survival_i = 5 + 10 \cdot Oil_i$$
$$\Leftrightarrow y = 5 + 10x$$

# Linear equations

A **linear equation** is an equation that contains only terms of order $x^1$ and $x^0 = 1$.

$$Survival_i = 5 + 10 \cdot Oil_i$$
$$\Leftrightarrow y = 5 + 10x$$

- One unit increase in oil revenue leads to a ten-unit increase in dictator's survival.

# Linear equations

A **linear equation** is an equation that contains only terms of order $x^1$ and $x^0 = 1$.

$$Survival_i = 5 + 10 \cdot Oil_i$$

$$\Leftrightarrow y = 5 + 10x$$

- One unit increase in oil revenue leads to a ten-unit increase in dictator's survival.
- With 0.5 units of *Oil*, we expect 10 *years of a dictator's survival* in power.
- With 2 units of *Oil*, we expect 25 *years of a dictator's survival* in power.

# Linear equations

$$y = a + bx$$

- $a$ and $b$ are constants, and $x$ and $y$ are variables.
- The constant $a$ is the **intercept**, where the function crosses the vertical $y$ axis. What does this mean in terms of our analysis? (e.g. oil revenue $\rightarrow$ dictator's survival)

# Linear equations

$$y = a + bx$$

- $a$ and $b$ are constants, and $x$ and $y$ are variables.
- The constant $a$ is the **intercept**, where the function crosses the vertical $y$ axis. What does this mean in terms of our analysis? (e.g. oil revenue $\rightarrow$ dictator's survival)
  $\rightarrow$ A constant can indicate an average value of $y$ *absent* the effects of $x$. Or, the average value that we expect from other factors (economic growth, personality, US ally, etc.)

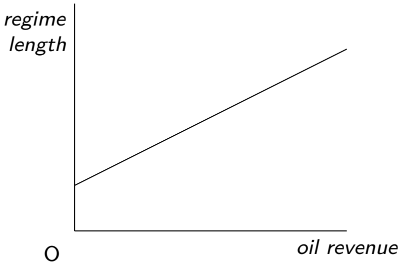
# Linear equations

$$y = a + bx$$

- $a$ and $b$ are constants, and $x$ and $y$ are variables.
- The constant $a$ is the **intercept**, where the function crosses the vertical $y$ axis. What does this mean in terms of our analysis? (e.g. oil revenue $\rightarrow$ dictator's survival)
  $\rightarrow$ A constant can indicate an average value of $y$ *absent* the effects of $x$. Or, the average value that we expect from other factors (economic growth, personality, US ally, etc.)
- The constant $b$ is the slope of a line, or the amount that $y$ changes given a one-unit increase in $x$.

This **linear** aspect of an equation is important, because we are simplifying a complicated relationship into a *single line*.

$$Survival_i = 5 + 10 \cdot Oil_i$$



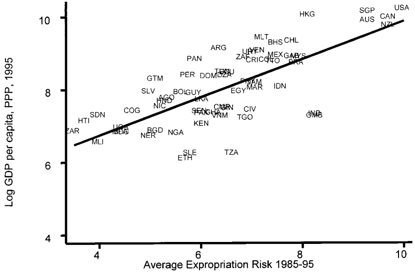
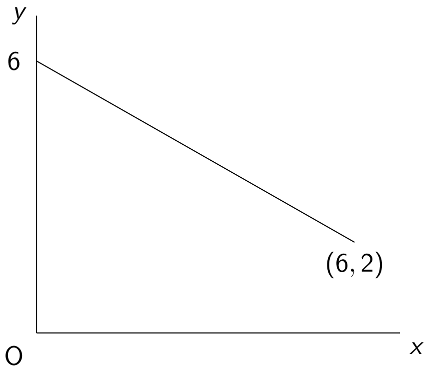



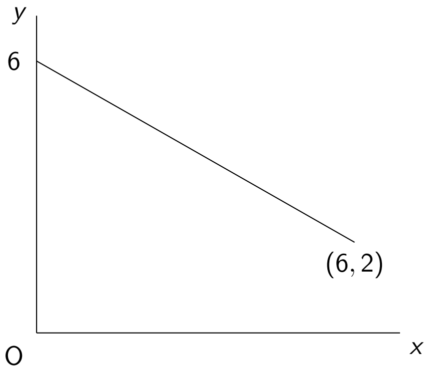

Figure: Acemoglu, Johnson, and Robinson (2001)

# Practice graph




What is the intercept of the graph?:

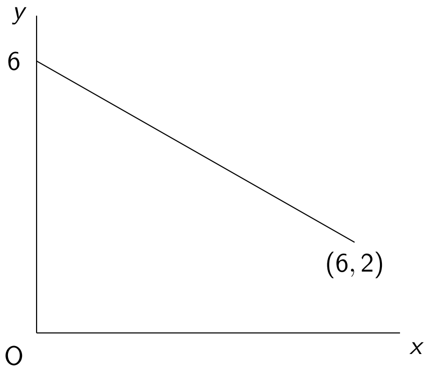
# Practice graph




What is the intercept of the graph?:6

Calculate the slope of the graph:

# Practice graph




What is the intercept of the graph?:6

Calculate the slope of the graph: slope $= \frac{\Delta y}{\Delta x} = \frac{y_1 - y_2}{x_1 - x_2} = \frac{6-2}{0-6} = -\frac{2}{3}$

# Linear models

- Political scientists have widely used linear regression models based on linear equations. (And so will you.)
- They are very simple and intuitive...

# Linear models

- Political scientists have widely used linear regression models based on linear equations. (And so will you.)
- They are very simple and intuitive...

$$Vote_{i,j} = J \times \{\beta_{0,j} + \beta_{1,j} Corruption_{i,j} + \beta_{2,j} Economy_{i,j} + \beta_{3,j} Corruption_{i,j} \times Economy_{i,j}\} + \varepsilon_i, \quad (1)$$

Klasnja and Tucker (2013)

$$tradepolicy_{i,t} = \beta_0 + \beta_1 REGIME_{i,t-1} + \beta_2 IMF_{i,t-1} + \beta_4 OFFICE_{i,t-1} + \beta_5 GDPPC_{i,t-1} + \beta_5 LNPOP_{i,t-1} + \beta_6 ECCRISIS_{i,t-1} + \beta_7 BPCRISIS_{i,t-1} + \beta_8 AVOPEN_{t-1} + u_i + \varepsilon_{i,t}$$

Milner and Kubota (2005)

$$P(t|a) = \alpha + \beta_1 X_1 + \beta_2 X_2 + \varepsilon | a$$

Boix and Stokes (2003)

$$RIGHTS_{i,t} = \alpha_1 RIGHTS_{i,t-1} + \alpha_2 \boldsymbol{Z}_{i,t} + \mu_t + u_{it},$$

Dreher et al. (2012)

# Linear equations

- The linear equation states that the size of the impact of $x$ on $y$ is constant across all values of $x$. What does this mean?

$$Survival_i = 5 + 10 \cdot Oil_i$$

# Linear equations

- The linear equation states that the size of the impact of $x$ on $y$ is constant across all values of $x$. What does this mean?

$$Survival_i = 5 + 10 \cdot Oil_i$$

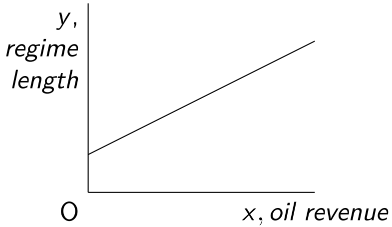
- One unit increase in oil revenue leads to a ten-unit increase in dictator's survival.

# Linear equations

- The linear equation states that the size of the impact of $x$ on $y$ is constant across all values of $x$. What does this mean?

$$Survival_i = 5 + 10 \cdot Oil_i$$

- One unit increase in oil revenue leads to a ten-unit increase in dictator's survival.
- Let's compare Dictator A with oil revenue of 270 billion barrels (Gbbl) and Dictator B with 8 Gbbl. Now suppose a 10 Gbbl increase in oil revenue in both countries. **Do you expect the same ten-unit increase in dictator's survival in these two countries?**

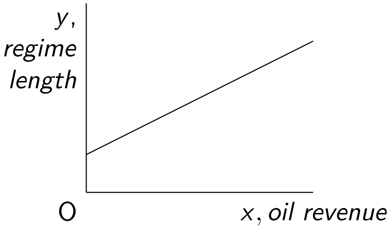


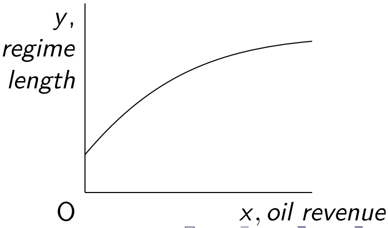
# Linear equations

- The linear equation states that the size of the impact of $x$ on $y$ is constant across all values of $x$. What does this mean?

$$Survival_i = 5 + 10 \cdot Oil_i$$

- One unit increase in oil revenue leads to a ten-unit increase in dictator's survival.
- Let's compare Dictator A with oil revenue of 270 billion barrels (Gbbl) and Dictator B with 8 Gbbl. Now suppose a 10 Gbbl increase in oil revenue in both countries. **Do you expect the same ten-unit increase in dictator's survival in these two countries?**



# Nonlinear fuctions

- Exponents: $x^n = x \times x \times \cdots x$
- Logarithms: $x = a^b \rightarrow \log_a x = b$
- Radicals (Roots): $\sqrt{x}$

# Exponents and exponential function

- **Exponents** are a shorthand for expressing the multiplication of a number by itself: $x^3 = x \times x \times x$
- $x^{-n} = \frac{1}{x^n}$, $x^{\frac{1}{n}} = \sqrt[n]{x}$
- **Multiplication**: $x^m \times x^n = x^{m+n}$
- **Power of a power**: $(x^m)^n = x^{mn}$
- **Division**: $\frac{x^m}{x^n} = x^{m-n}$

# Quadratic functions

- **Quadratic functions** are nonlinear functions that describe a parabola: $y = \alpha + \beta_1 x + \beta_2 x^2$
- E.g. The extent to which governments are transparent (i.e. noncorrupt) varies nonlinearly with the level of political competition. *Corruption is low in a strong autocracy and a full democracy and high in a hybrid regime* (Montinola and Jackman 2002).

# Quadratic functions

- **Quadratic functions** are nonlinear functions that describe a parabola: $y = \alpha + \beta_1 x + \beta_2 x^2$
- E.g. The extent to which governments are transparent (i.e. noncorrupt) varies nonlinearly with the level of political competition. *Corruption is low in a strong autocracy and a full democracy and high in a hybrid regime* (Montinola and Jackman 2002).

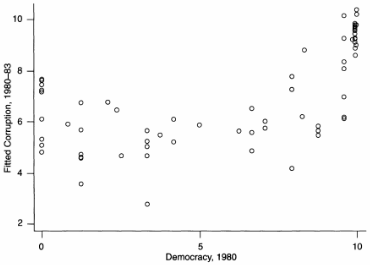
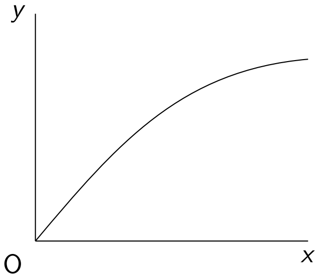



*Fig. 1. Plot of fitted corruption, 1980–83, against democracy, 1980 (fitted values from robust regression estimates in Table 2, column 3)*

# Logarithms

- Logarithms are the inverses of exponents (and vice versa).
- $\ln(1) = 0 \Leftrightarrow x^0 = 1$
- $\ln(x_1 \cdot x_2) = \ln(x_1) + \ln(x_2)$
  $\ln \frac{x_1}{x_2} = \ln(x_1) - \ln(x_2)$
- $\ln(x^b) = b \ln(x)$



# Practice questions

Practice questions! Good luck!

# Inverse function

When a function is defined as $f(x) : A \rightarrow B$, the **inverse** function is $f^{-1}(x) : B \rightarrow A$. For example, if $f(x) = 2x + 3$, then its inverse is $f^{-1}(x) = \frac{x-3}{2}$. Find the inverse of the following functions.

1. $f(x) = 8x - 5$
2. $f(x) = \frac{2}{(x-4)}$
3. $f(x) = \sqrt{(x+3)}$

# Solving equations, finding equilibrium

$$Q_1 = 51 - 3P_1 \qquad Q_1 = 30 - 2P_1$$
$$Q_2 = 6P_2 - 10 \qquad Q_2 = -6 + 5P_2$$

# Solving equations, finding equilibrium

Keynes designed a model of national income. He thought that the amount of money spent in a country (C) depends on the amount of money earned (Y) and a baseline consumption that people spend even when they don't have income. Then, he came up with an equation for national income, which consists of consumption (C), investment (I), and government spending (G). Here, we have sample equations below and our task is to get equilibrium consumption ($C^*$) and income ($Y^*$).

$$Y = C + I + G$$
$$C = 10 + .5Y$$

Find equilibrium $C^*$ and $Y^*$. How many endogenous variables do we have?

We have a more advanced model for national income as follows:

$$Y = C + I + G$$
$$C = 7 + 0.4(Y - T)$$
$$T = 15 + 0.3Y$$

where $T$ and $t$ represents taxation amount and its rates, respectively. How many endogenous variables do we have? What are $Y^*$, $T^*$, and $C^*$?

# Notations

- **Operators**: $+, -, \times, /, x^n, \sqrt[n]{x}, \sum, \prod, !$

$$\sum_{i=k}^{l} x_i = x_k + \cdots + x_l.$$

$$\prod_{i=k}^{l} x_i = x_k \times \cdots x_l.$$

# Notations

## Sequence

A sequence is a function whose domain is the set of positive integers.

We write a sequence as,

$$\{a_n\}_{n=1}^{\infty} = (a_1, a_2, a_3, \ldots, a_N, \ldots)$$

Examples:

- $a_n = \{n\} = (1, 2, 3, 4, \ldots, n, \ldots)$
- $b_n = \{\frac{1}{n^2}\} = (\frac{1}{1^2}, \frac{1}{2^2}, \frac{1}{3^2}, \frac{1}{4^2}, \ldots)$
- $c_n = \{\frac{1}{2^n}\} = (\frac{1}{2^1}, \frac{1}{2^2}, \frac{1}{2^3}, \frac{1}{2^4}, \ldots)$

# Notations

Summations and products: Examples

- Sequence: $a_n = \{n\} = (1, 2, 3, 4, \ldots, n, \ldots)$
- Summation: $\displaystyle\sum_{n=1}^{\infty} a_n = \sum_{n=1}^{\infty} n = 1 + 2 + 3 + 4 + \cdots$
- Product: $\displaystyle\prod_{n=1}^{\infty} a_n = \sum_{n=1}^{\infty} n = 1 \cdot 2 \cdot 3 \cdot 4 \cdots$

# Notations

Summations and products: Examples

- Sequence: $a_n = \{n\} = (1, 2, 3, 4, \ldots, n, \ldots)$
- Summation: $\displaystyle\sum_{n=1}^{\infty} a_n = \sum_{n=1}^{\infty} n = 1 + 2 + 3 + 4 + \cdots$
- Product: $\displaystyle\prod_{n=1}^{\infty} a_n = \sum_{n=1}^{\infty} n = 1 \cdot 2 \cdot 3 \cdot 4 \cdots$

How about...

- $b_n = n^2$
- $\displaystyle\sum_{i=1}^{n} b_i = ?$
- $\displaystyle\prod_{i=1}^{n} b_i = ?$

# Notations

1. When $A = \{A, I, J, K\}, B = \{C, D, J, Q\}$, and $C = \{B, C, E, I\}$, find $A/B \cap C$.
   $A/B \cap C = ?$

2. $C = \{2^1, 2^3, 2^4\},\ c_i \in C.\quad \prod_{i=1}^{n} c_i = ?$

3. When $x \in \{3, 4, 5, 6\}$, and $f(x) = x + 3$,
   $\max_x f(x) = ?$
   $\arg\max_x f(x) = ?$

4. $x \in \mathbb{R}, f(x) = 19 - (x-2)^2$.
   $\max_x f(x) = ?$
   $\arg\max_x f(x) = ?$

The end of the first session. Thank you!