Day 2, Session 1: Graphs

Brian D. Williamson

EPI/BIOST Bootcamp 2016

26 September 2016

Graphs

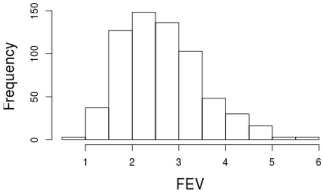
- Why do we use graphs?
  - Describe relationships in the data
  - Visualize functions
- Graphs are very useful in exploratory analyses, or for description

Example data analysis: FEV (from Scott Emerson, MD PhD)

Based on numerous studies, we believe that smoking tends to impair lung function. Much of the data to support this claim arises from studies of long-term adult smokers. A natural question is: can deleterious effects of smoking be detected in children who smoke?

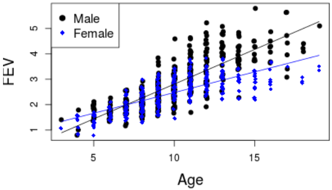
- Data on 654 children seen during routine check-up at pediatric clinic
- Outcome: forced expiratory volume (FEV)
  - measures how much air you can blow out of your lungs in a short period of time
  - higher FEV typically associated with better respiratory function
- Predictor of interest: smoking status
- Other variables: sex, age, height

Common types of graphs in data analysis



(a) Histogram

(b) Boxplot



(c) Scatterplot

What do graphs tell us?

- Histograms: summaries of one-dimensional distributions
  - Counts or frequencies of each occurrence
- Boxplots: summaries of two-dimensional distributions
  - measures of center (typically median)
  - measures of spread (typically inter-quartile range)
- Scatterplots: summaries of two-dimensional distributions
  - Can visualize the whole data
  - Trends in two or more dimensions by using different colors/shapes for strata

Linear trends

- A common way to describe data (think linear regression!)

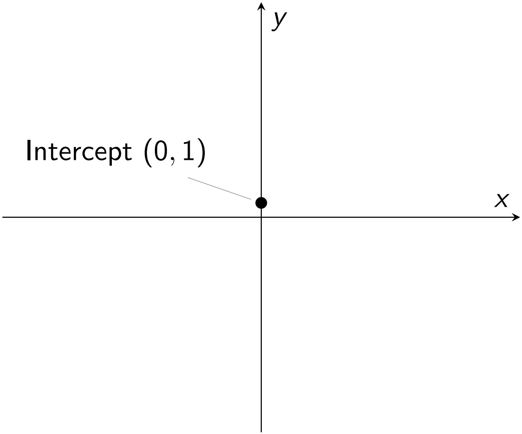
- Lines are easy to compute
  - Only need a point and a slope

  - Two common forms of linear equations

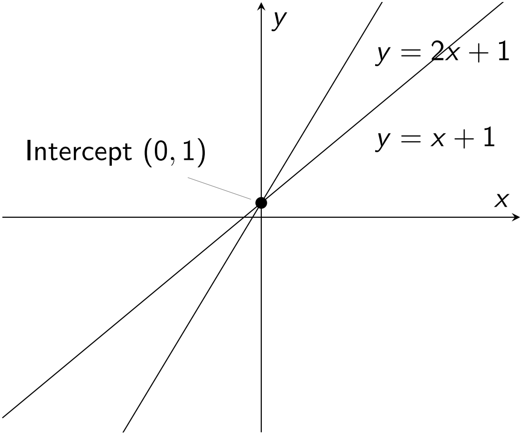
Slope-intercept form

- $y = mx + b$
- Slope: m
  - Rate of change, i.e. how does y change with each one unit change in x?
  - Example: speed, the distance traveled with each unit change in time
- Intercept: b
  - The point where the line crosses the $y-$axis

Slope-intercept form: determining a line



Slope-intercept form: determining a line



Point-slope form

- $y - y_1 = m(x - x_1)$
- Point: (x_1, y_1)
  - A point on the line (can be any point! Even the intercept!)
- Slope: m
  - Same as in slope-intercept form!
- Example: $y - 1 = 2(x - 0)$ is the same as $y = 2x + 1$ in slope-intercept form!

Exercise: slopes and intercepts

1. What is the slope of the line $y = 2x - 3$?
2. What is the $y-$intercept of the line $y = 2x - 3$?
3. What is the slope of the line $y + 1 = 2(x - 1)$?
4. What point did we use to create the line $y + 1 = 2(x - 1)$?

Solution: slopes and intercepts

1. The equation is in slope-intercept form, so the slope is 2
2. The equation is in slope-intercept form, so the intercept is -3
3. The equation is in point-slope form, so the slope is 2
4. The equation is in point-slope form, so the point is $(1, -1)$

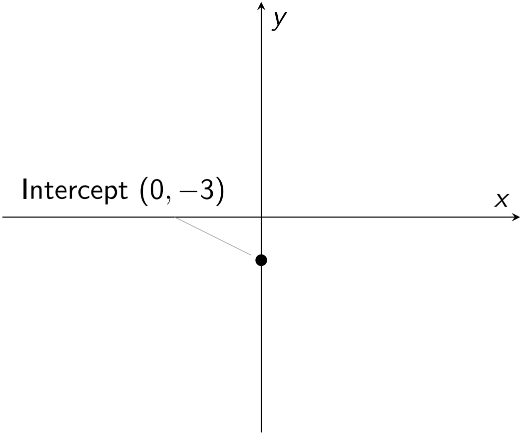
Creating a graph using an equation

- Steps:
  1. Draw axes
  2. Place a point at the $y-$intercept (slope-intercept form) or at the starting point (point-slope form)
  3. Increase x by one unit, increase y by m units, place a new point
  4. Draw a line between the old point and the new point!

Example: creating a graph using an equation

- Equation $y = 2x - 3$
- Slope: 2, Intercept: -3

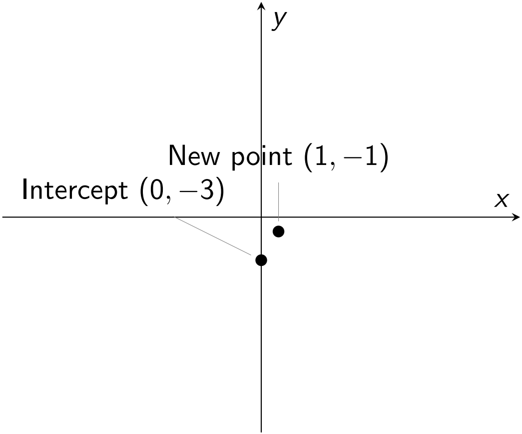
1. Draw a point at $(0, -3)$



Example: creating a graph using an equation

- Equation $y = 2x - 3$
- Slope: 2, Intercept: -3

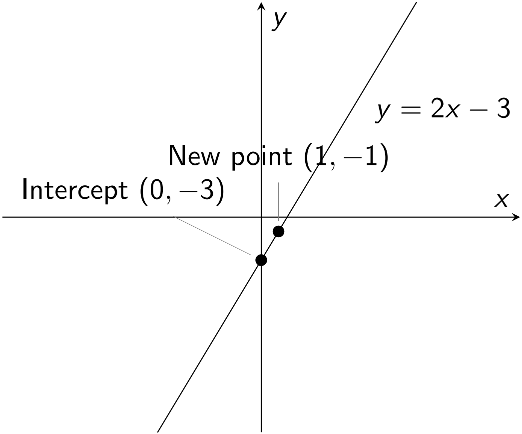
2. Increase x to 1, increase y by 2. New point at $(1, -1)$



Example: creating a graph using an equation

- Equation $y = 2x - 3$
- Slope: 2, Intercept: -3

3. Draw a line!

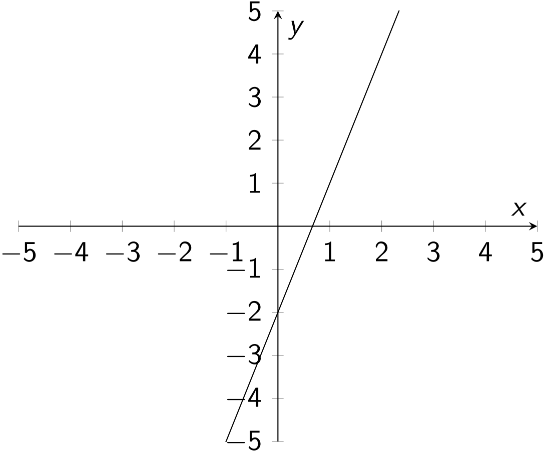


Reading an equation from a graph

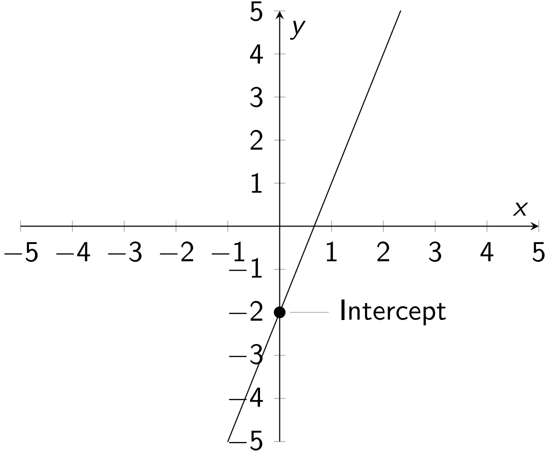
- Two options:
  1. Slope-intercept form
     1.1 Find the $y-$intercept
     1.2 Find the slope: how much does y change with each 1 unit difference in x?

  2. Point-slope form
     2.1 Choose any point on the line
     2.2 Find the slope: how much does y change with each 1 unit difference in x?

Example: reading an equation from a graph

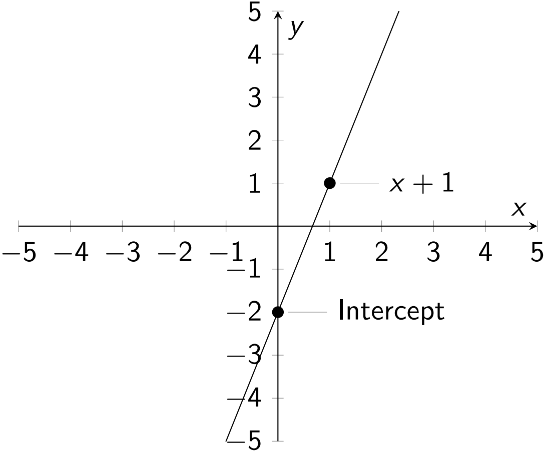


Example: reading an equation from a graph



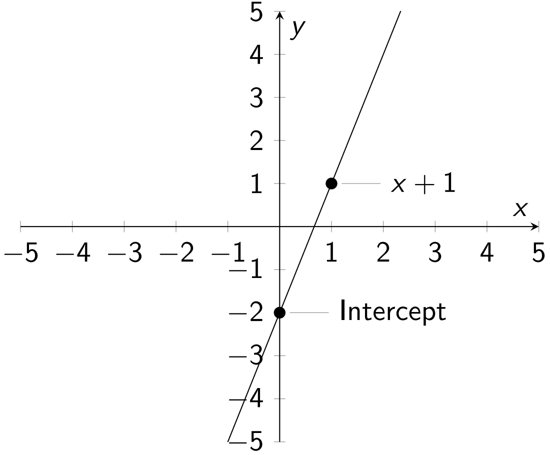
Example: reading an equation from a graph

1. Find the intercept. Here it is -2



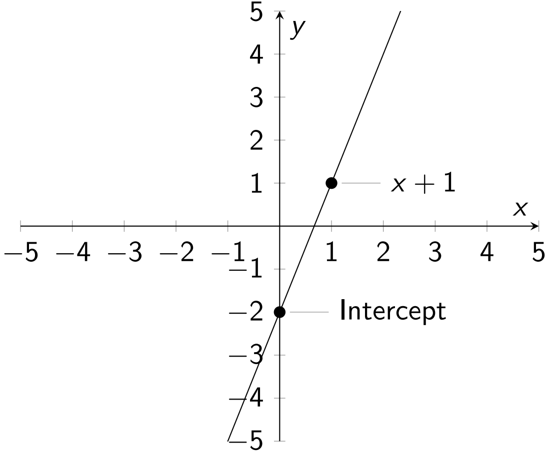
Example: reading an equation from a graph

2. Increase x by one to find the slope. The y value at $x = 1$ is 1, so y changed by 3



Example: reading an equation from a graph

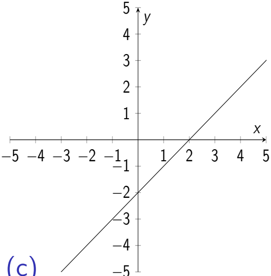
3. Slope-intercept: $y = 3x - 2$, point-slope: $y + 2 = 3(x - 0)$



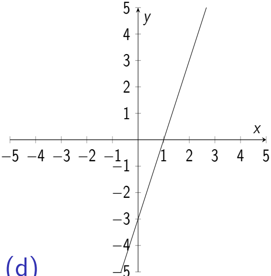
Exercise: matching graphs to equations

1. Which is the graph of $y = 2x - 3$?

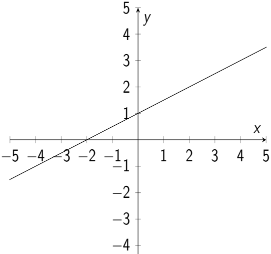
(a)



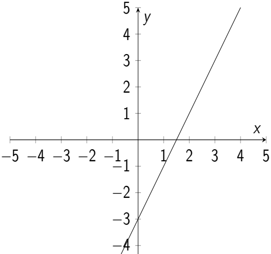
(b)



(c)



(d)



Solution: matching graphs to equations

(a) Intercept: -2, slope: 1

(b) Intercept: -3, slope: 3

(c) Intercept: 1, slope: 1

(d) Intercept: -3, slope: 2 ✓

Quadratics

- Sometimes we believe a trend is higher order than linear
- Higher order terms allow more flexibility in modeling
- Quadratics are the natural next step from linear terms, and are shaped like parabolas

Defining a quadratic

- Standard form: $y = ax^2 + bx + c$

- a determines the direction of the tails and the degree of curvature
  - $a > 0$ means the curve faces up (convex)
  - $a < 0$ means the curve faces down (concave)
  - large $|a| > 1$ means steep slope
  - $0 < |a| < 1$ means shallow slope

- b and a together determine the $x-$coordinate of the vertex: $x = -\frac{b}{2a}$

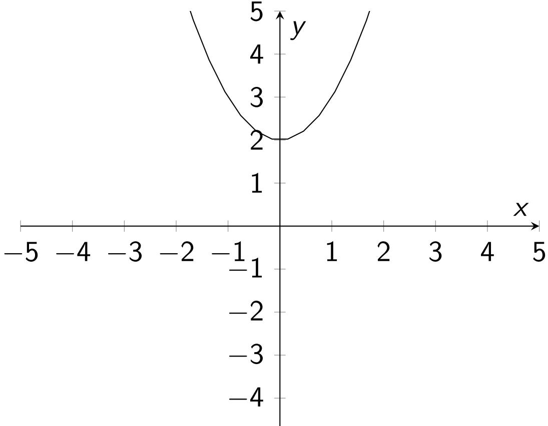
- c controls the height

Example: quadratics

- $y = x^2 + 2$

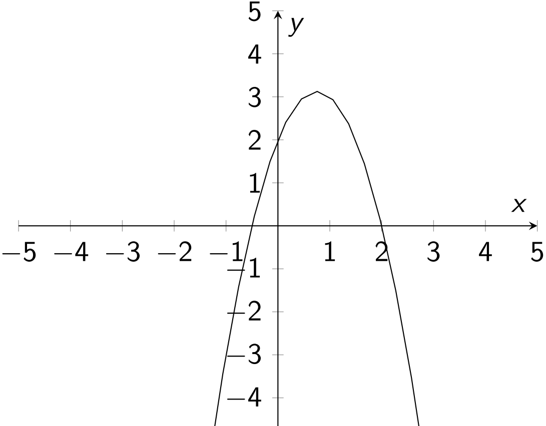
- $a = 1,\ b = 0,\ c = 2$

- Not too steep, convex, and vertex is at $(0, 2)$



Example: quadratics

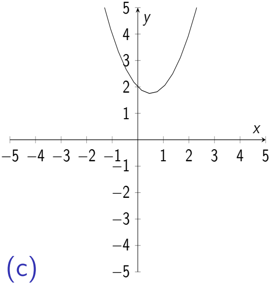
- $y = -2x^2 + 3x + 2$
- $a = -2$, $b = 3$, $c = 2$
- Steeper than before, concave, and vertex is at $(3/4, 50/16)$



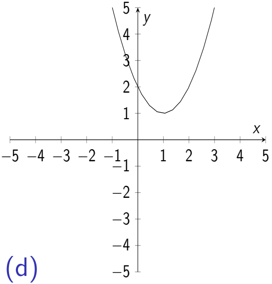
Exercise: quadratics

1. Which is a plausible plot of $x^2 - x + 2$?

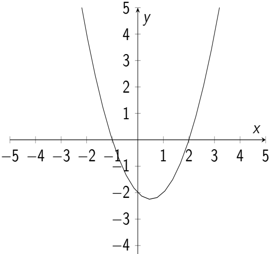
(a)



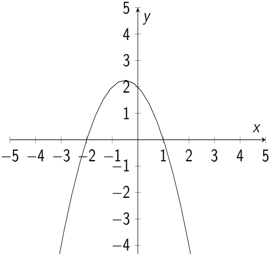
(b)



(c)



(d)



Solution: quadratics

1. $a = 1$, $b = -1$, $c = 2$
   - We are looking for a plot where the vertex has $y-$coordinate near 2
   - This rules out (b)
   - Now we want a plot with $y = 2$ when $x = 1$, ruling out (b) [and (d)]
   - Of the two remaining, (d) is concave, so it has $a < 0$
   - (a) is the solution!

Transforming graphs

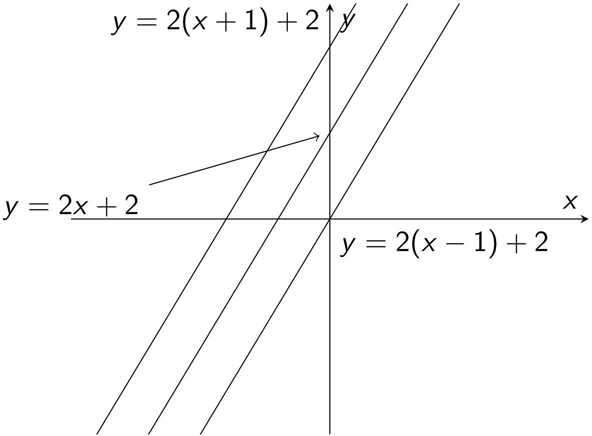
- Once we have one graph, how do we get another?

- Two main types of transformations: shifting and stretching

Shifting graphs

- Sometimes we want to change the interpretation of the intercept
- Once we know the properties of a graph, shifting doesn't change much!
- Shift left: add to x
- Shift right: subtract from x
- Shift up: add to intercept
- Shift down: subtract from intercept
- Why? Adding to x: smaller x's now have the same y. Subtracting from x: larger x's now have the same y.

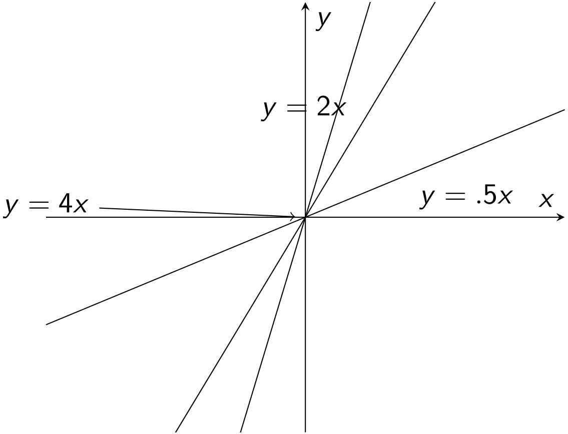
Example: shifting graphs



Stretching graphs

- Make the graph steeper, or shrink it: make $|m|$ larger in a linear equation, and make $|a|$ larger in a quadratic equation
- Make the graph shallower, or stretch it: make $|m|$ smaller in a linear equation, and make $|a|$ smaller in a quadratic equation

Example: stretching graphs



Exercise: transforming graphs

1. How do we shift the graph of $y = 2x + 3$ one unit right?

2. How do we transform the graph of $y = 2x + 3$ to have a shallower slope?

3. How do we transform the graph of $y = 2x + 3$ to have a slope of 1 and a $y-$intercept of 4?

Solution: shifting graphs

1. Subtract 1 from x! New equation: $y = 2(x - 1) + 3$
2. Multiply by a number less than 1; for example, take $x/2$. This gives new equation $y = 2(x/2) + 3$, or $y = x + 3$
3. To get a slope of 1, divide x by 2. To make the $y-$intercept 4, shift left by adding $1/2$ to x. New equation: $y = 2 * (x/2 + 1/2) + 3$, or $y = x + 4$

Summary

- Graphs are useful tools to describe relationships in data or visualize functions

- Histograms, boxplots, and scatterplots are common and useful types of graphs

- Linear trends can be described in:
  - Slope-intercept form — $y = mx + b$
  - Point-slope form — $y - y_1 = m(x - x_1)$

- Reading an equation from a graph involves finding the intercept and calculating the slope

- Graphs can be transformed by adding/subtracting from x, or multiplying/dividing x