Modeling Continuous Variables

Lecture 19
Sections 6.1 - 6.3.1

Robb T. Koether

Hampden-Sydney College

Tue, Feb 21, 2012

Outline

Outline

Models

Definition (Mathematical model)

A mathematical model is a mathematical abstraction and, therefore, a simplification, of a real situation, one that retains the essential features.

- Real situations are usually much too complicated to deal with in all their details.
- For example, when we model a coin toss, we assume that the probability of heads is exactly 50%.

Example

- The “bell curve” is a model (an abstraction) of many populations.
- Real population distributions have all sorts of bumps and twists and irregularities.
- The bell curve is smooth and perfectly symmetric.
- It fits many populations very well.
- It fits no population perfectly.
- In statistics, the bell curve is called the normal curve, or normal distribution.

Models

- Our models will be models of distributions, presented either as histograms or as continuous distributions.

Outline

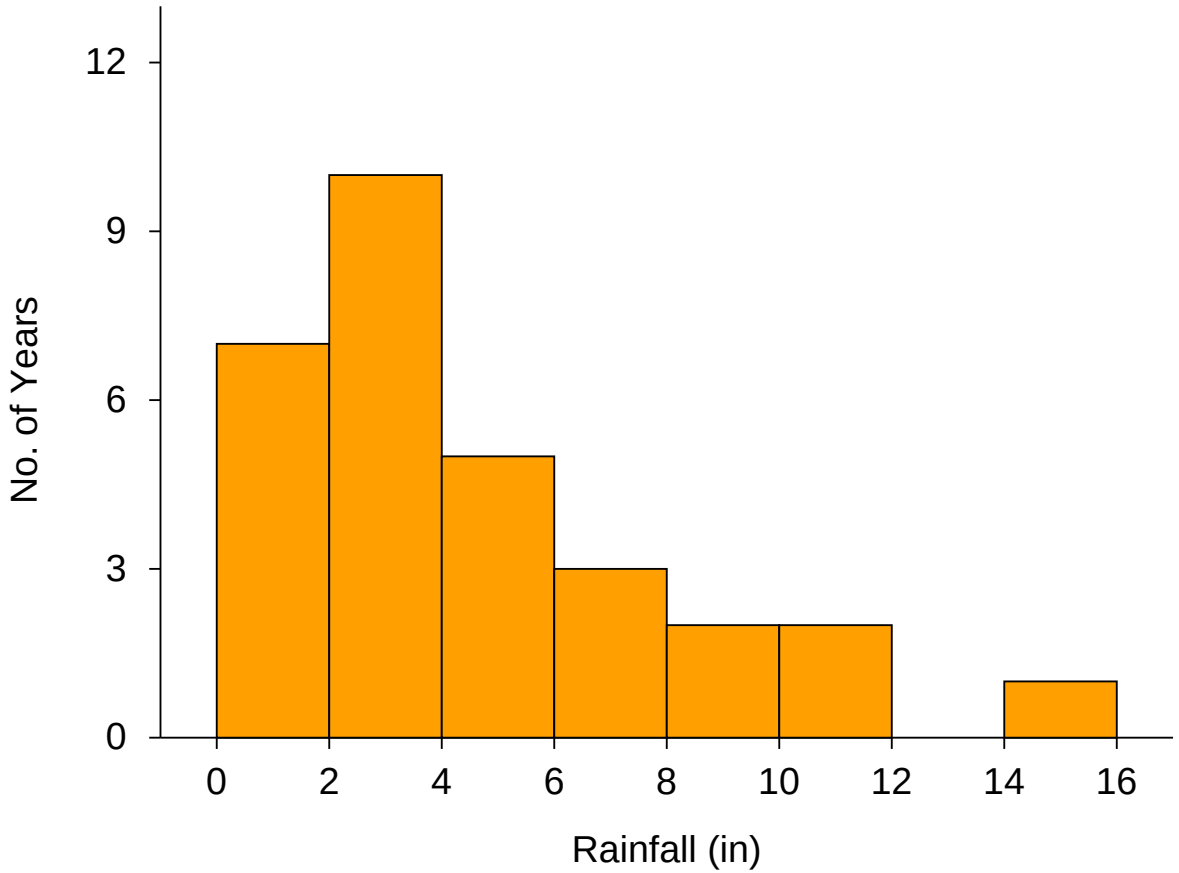
Outline

Histograms and Area

- In a histogram, frequency is represented by area.
- Consider the following distribution of rainfall data.

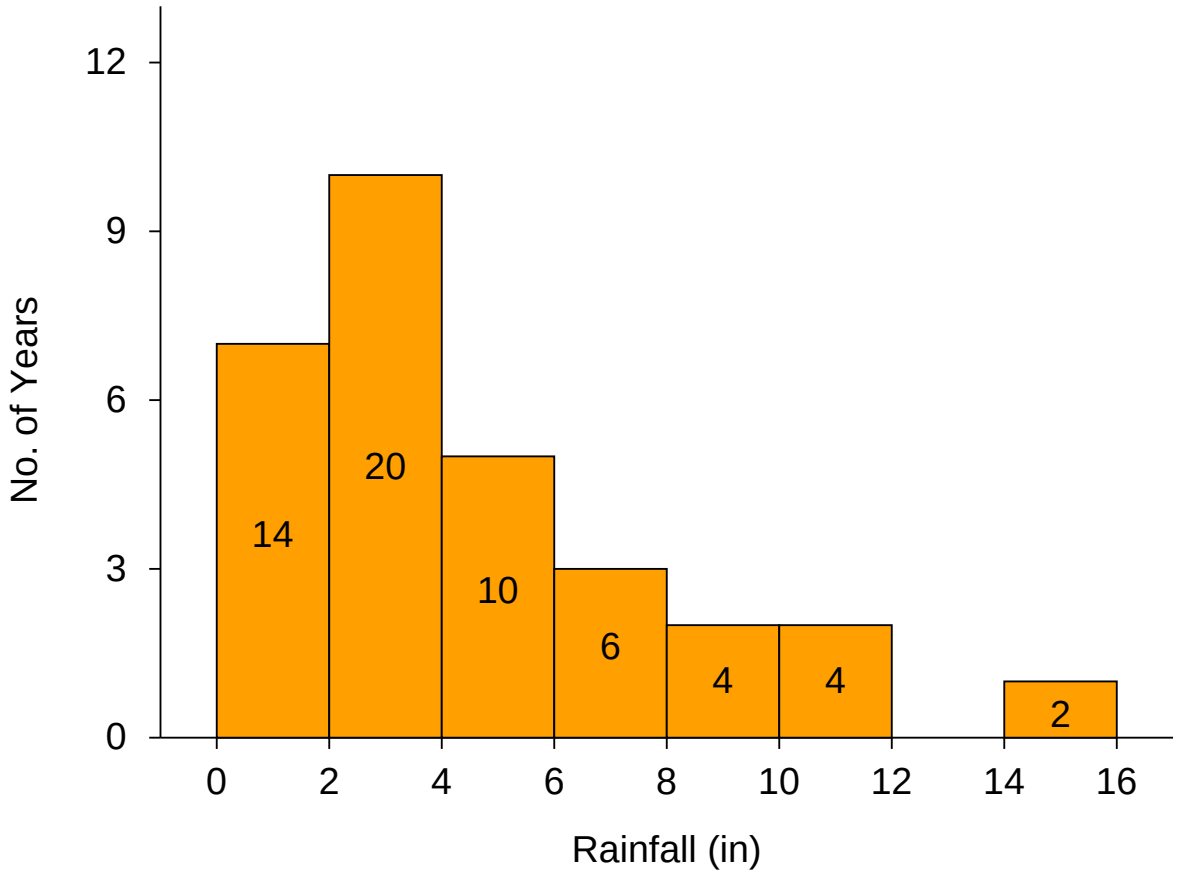
Rainfall (in)	Frequency
$[0, 2)$	7
$[2, 4)$	10
$[4, 6)$	5
$[6, 8)$	3
$[8, 10)$	2
$[10, 12)$	2
$[12, 14)$	0
$[14, 16)$	1

Histograms and Area



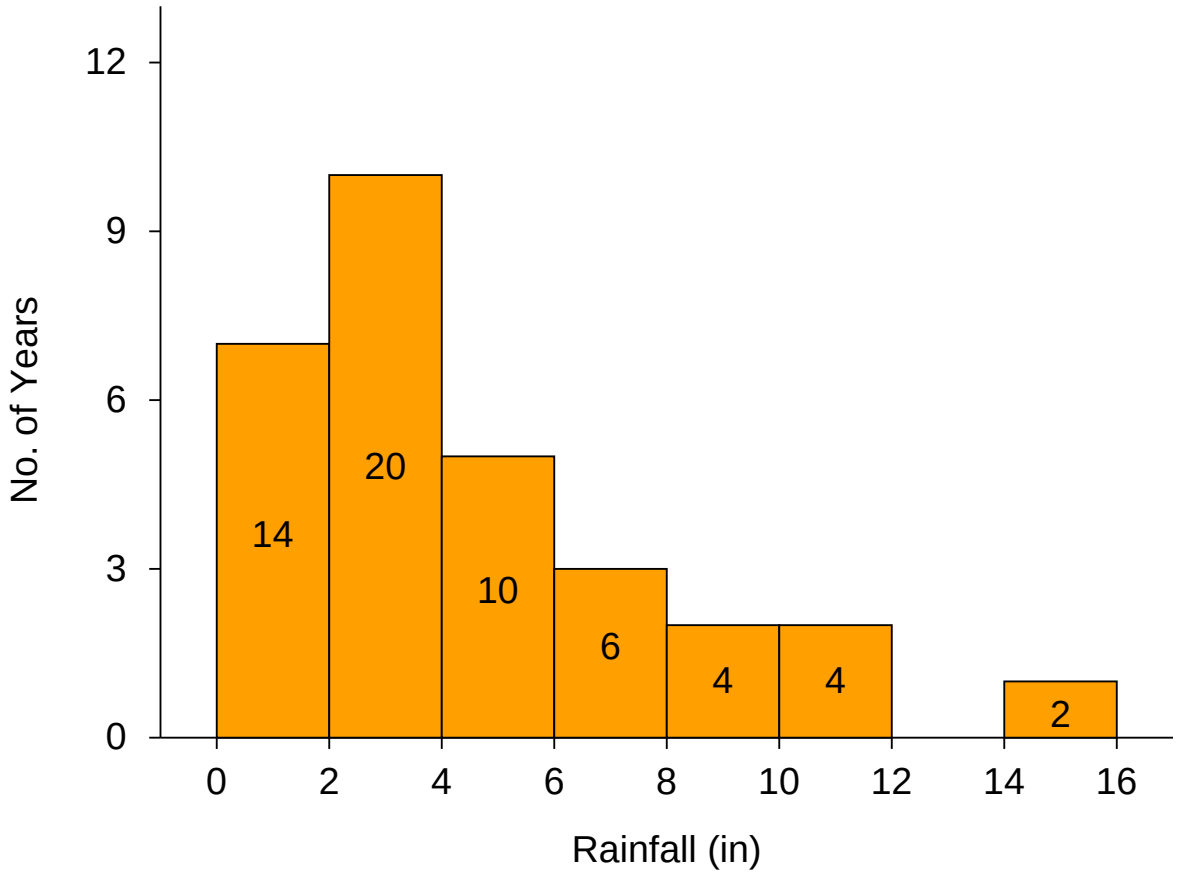
What is the total area of this histogram?

Histograms and Area



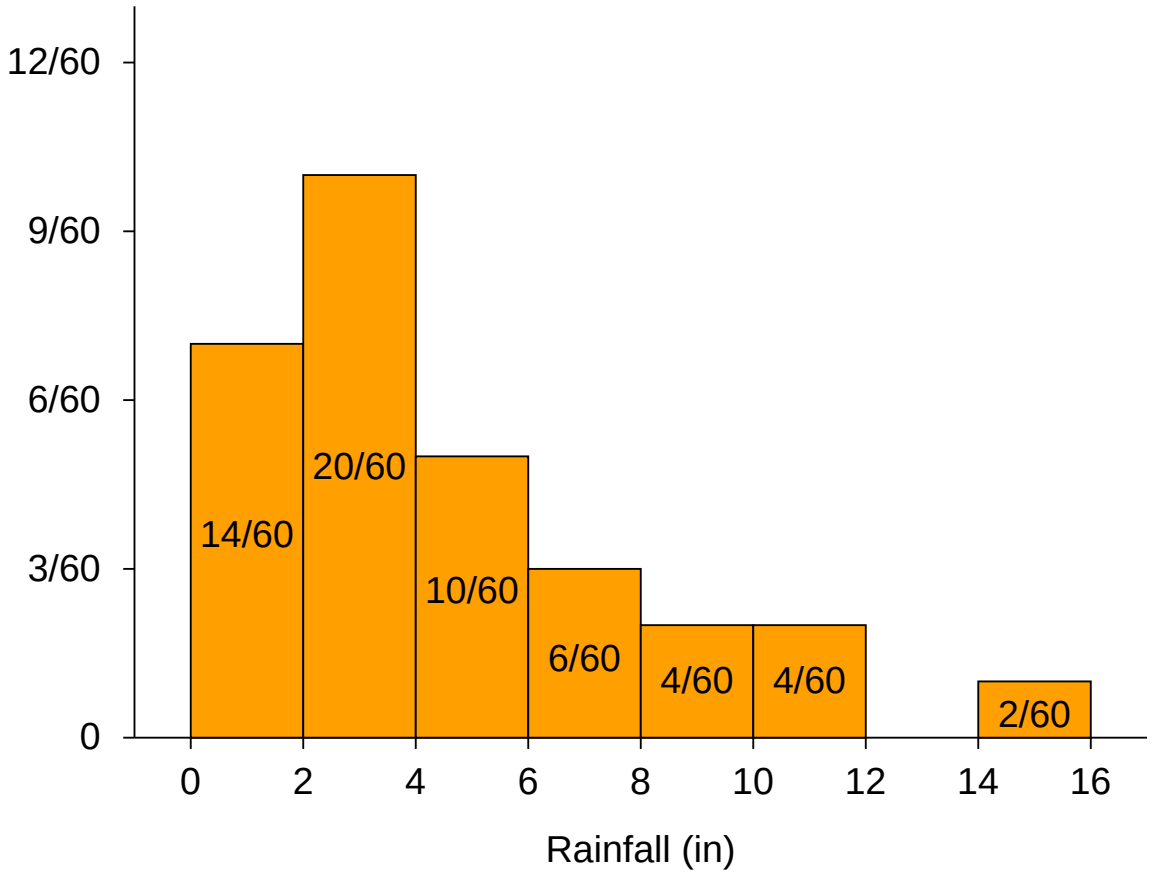
$$\text{Total area} = 14 + 20 + 10 + 6 + 4 + 4 + 2 = 60$$

Histograms and Area



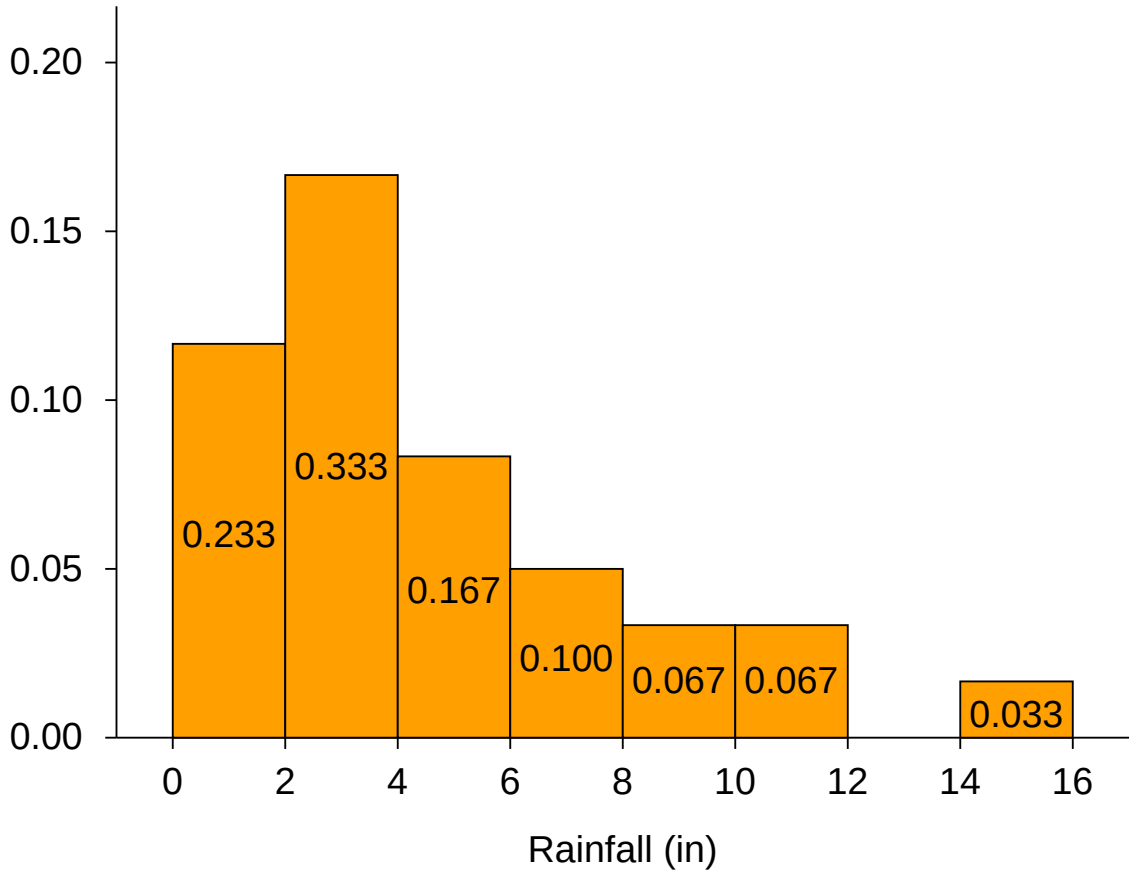
Notice that this is the class width (2) times the total frequency (30)

Histograms and Area



To make the total area = 1, shrink the vertical scale by a factor of 60

Histograms and Area

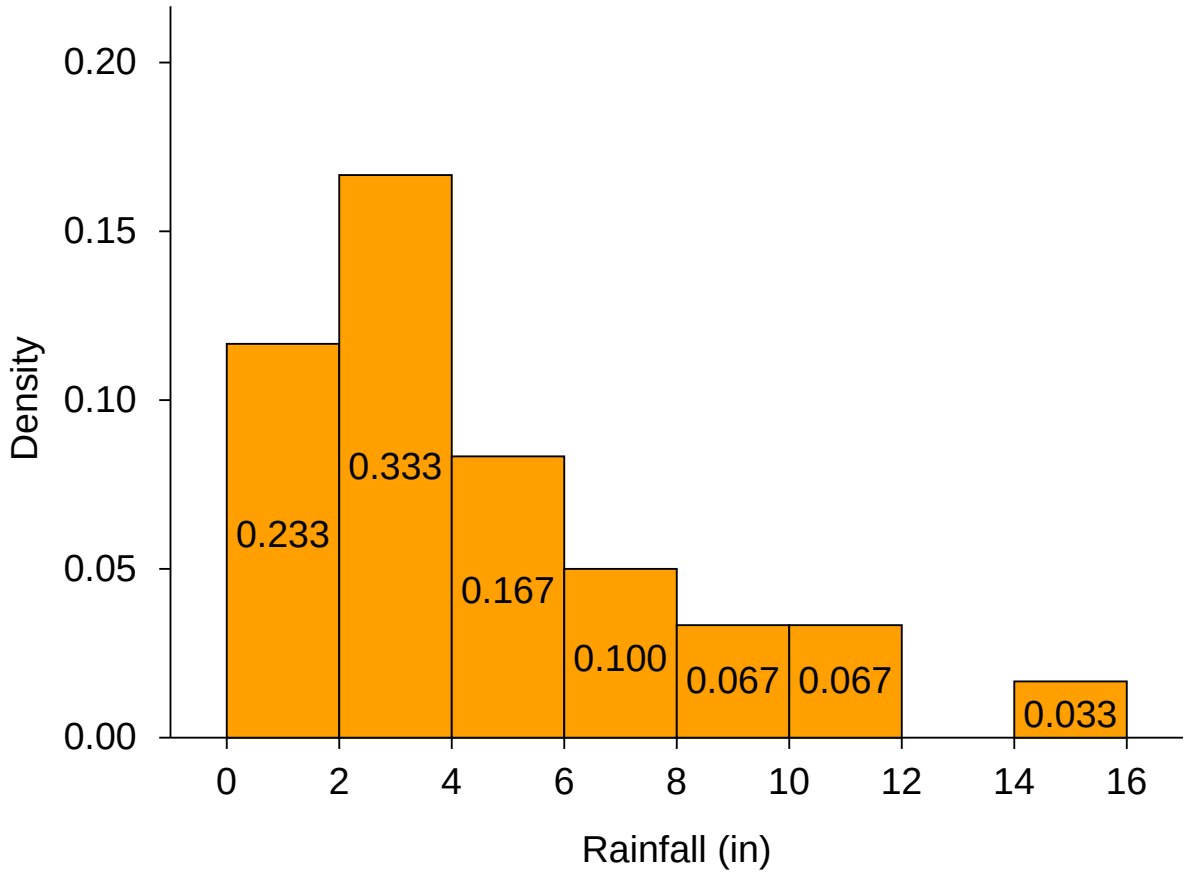


What is the probability a September rainfall with be between 4 and 8 inches?

Histograms and Probability

- We will rescale the vertical axis so that the total area equals 1, representing 100%.
- This new scale is called the density.

Histograms and Area



The probability density function (pdf) for September rainfall

Outline

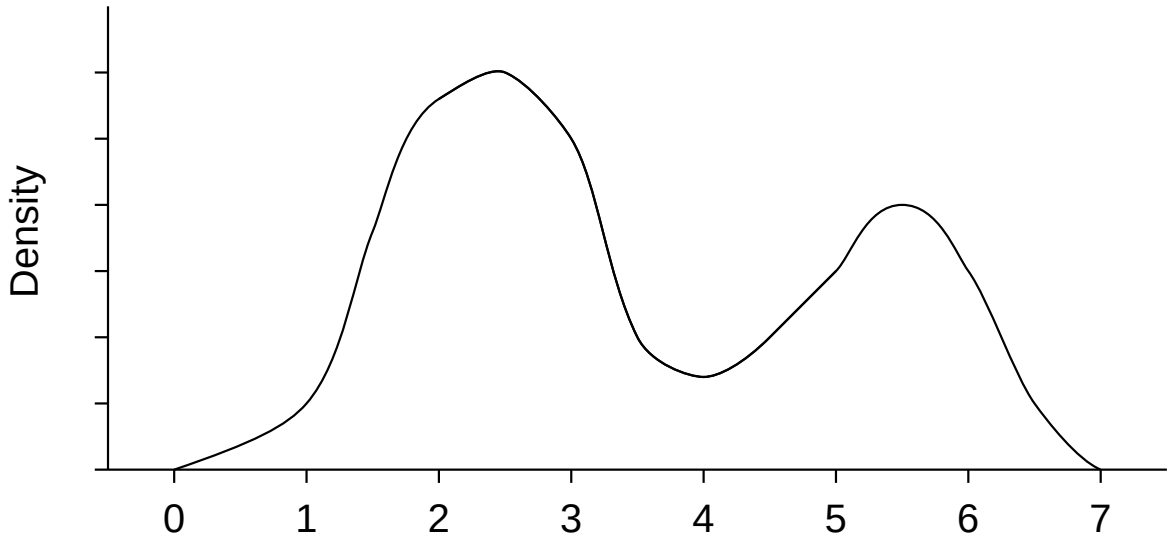
Density Functions

AREA = PROPORTION = PROBABILITY

- This is the fundamental property that connects the graph of a continuous model to the population that it represents.
- The area under the graph between two points on the x-axis represents the proportion of the population that lies between those two points.

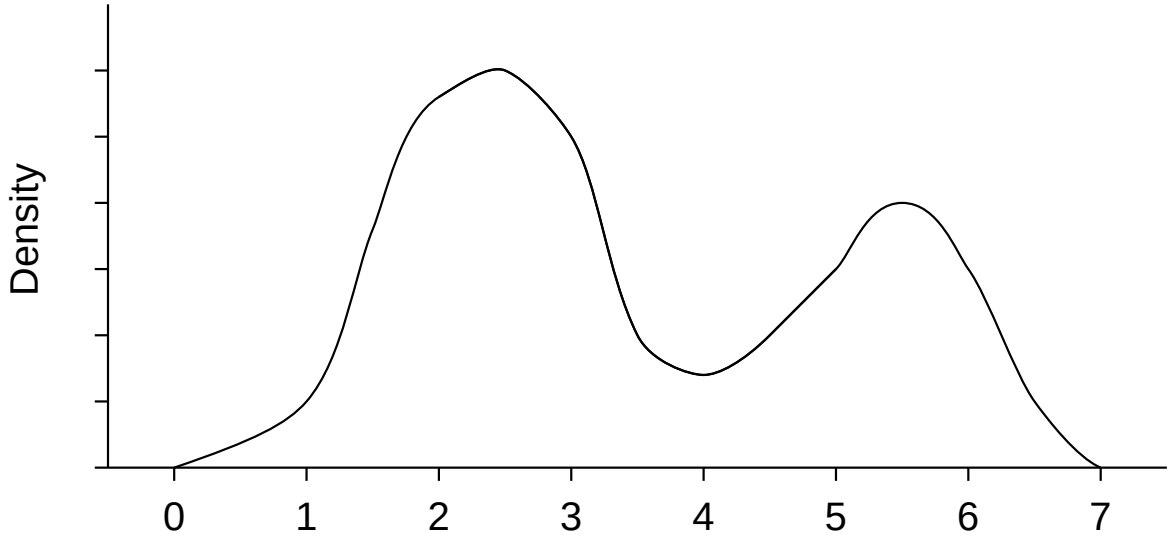
Density Functions

- Consider an arbitrary continuous distribution.



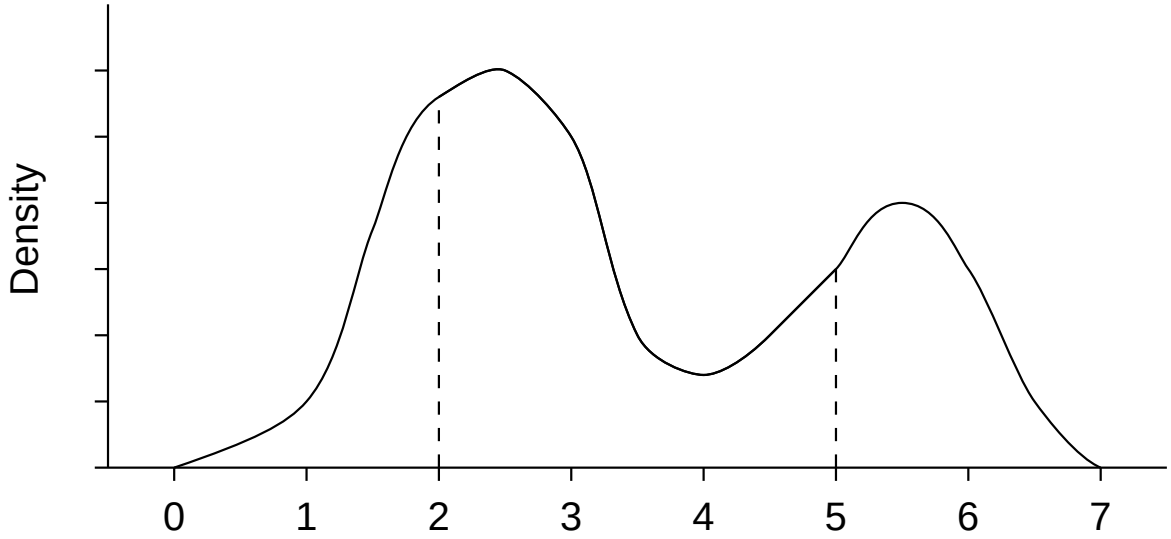
Density Functions

- The area under the curve between 2 and 5 is the proportion of the values of x that lie between 2 and 5.



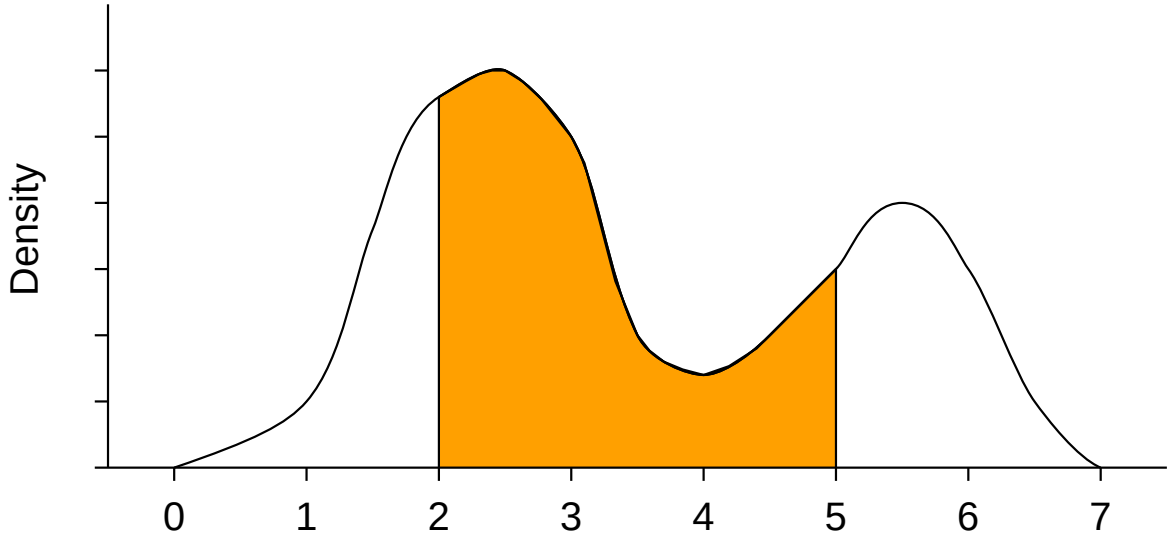
Density Functions

- The area under the curve between 2 and 5 is the proportion of the values of x that lie between 2 and 5.



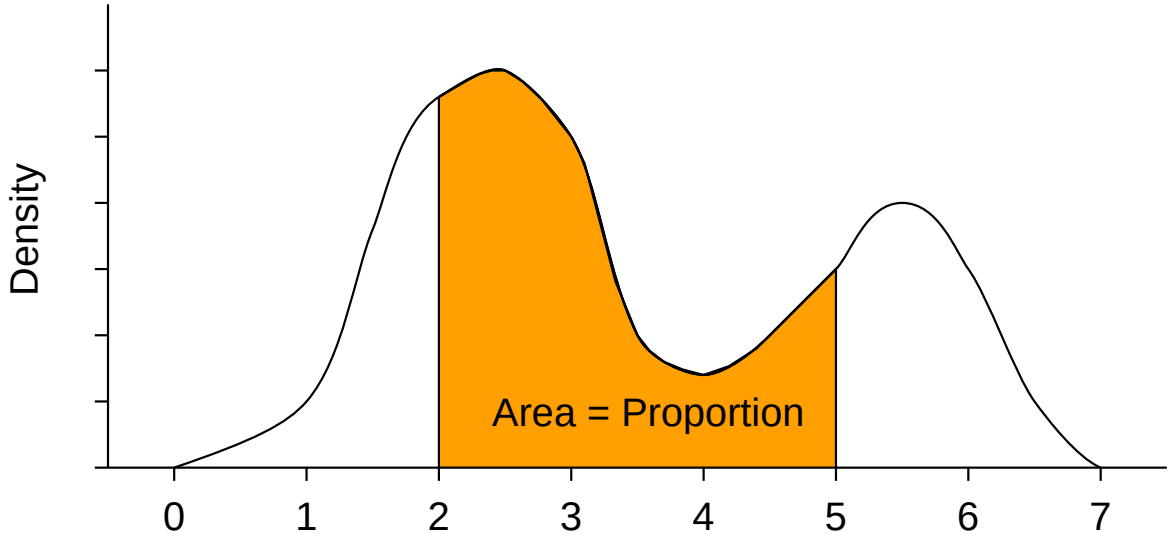
Density Functions

- The area under the curve between 2 and 5 is the proportion of the values of x that lie between 2 and 5.



Density Functions

- The area under the curve between 2 and 5 is the proportion of the values of x that lie between 2 and 5.



Outline

Outline

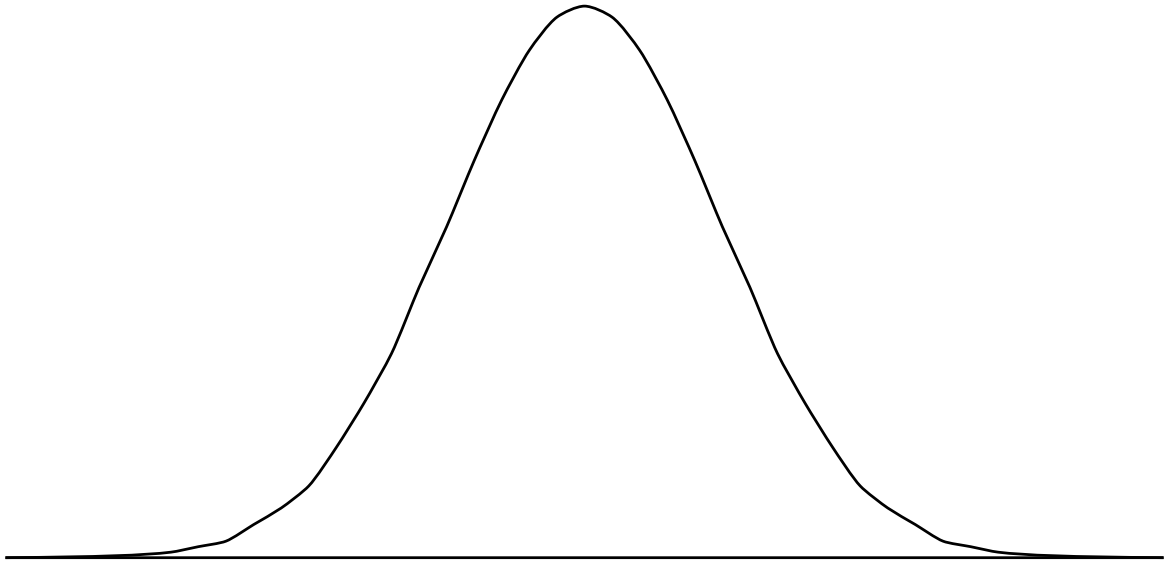
The Normal Distribution

Definition (Normal distribution)

The normal distribution is the statistician's name for the bell curve.

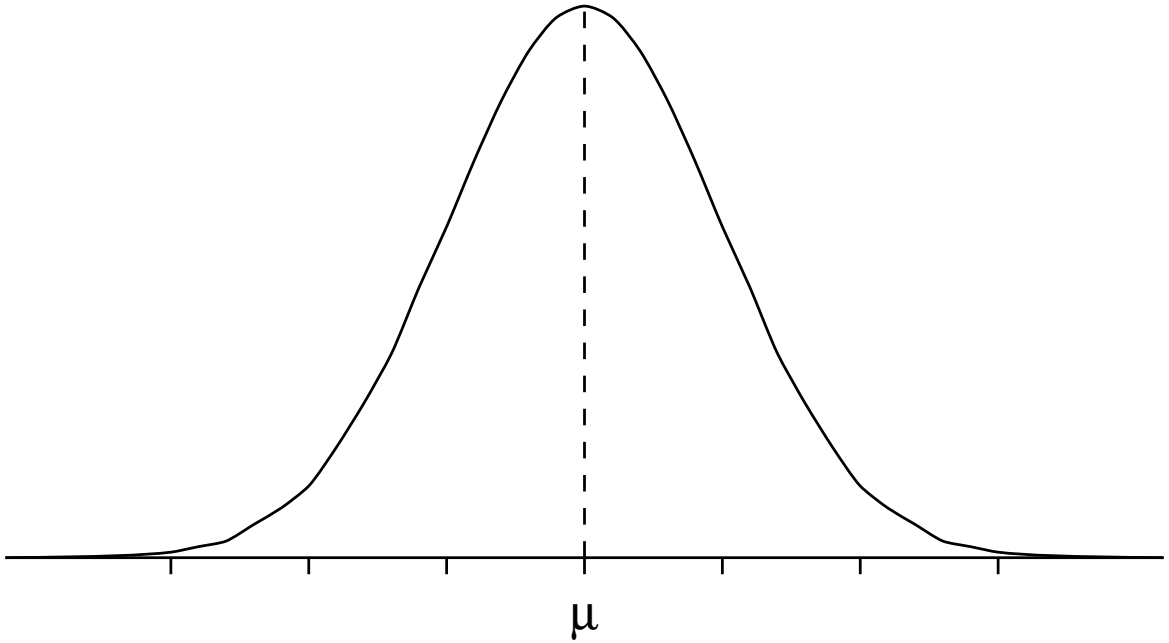
- It is a density function in the shape of a "bell."
  - Symmetric.
  - Unimodal.
  - Extends over the entire real line (no endpoints).
  - "Main part" lies within $\pm 3\sigma$ of the mean.

The Normal Distribution



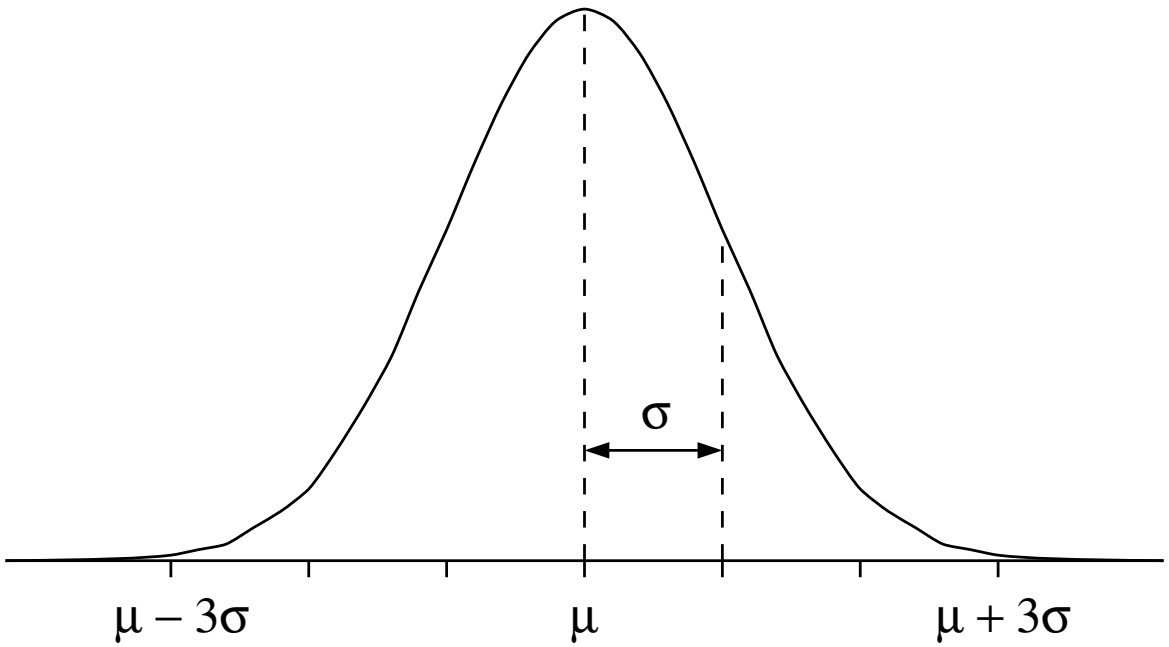
The curve has a bell shape, with infinitely long tails in both directions

The Normal Distribution



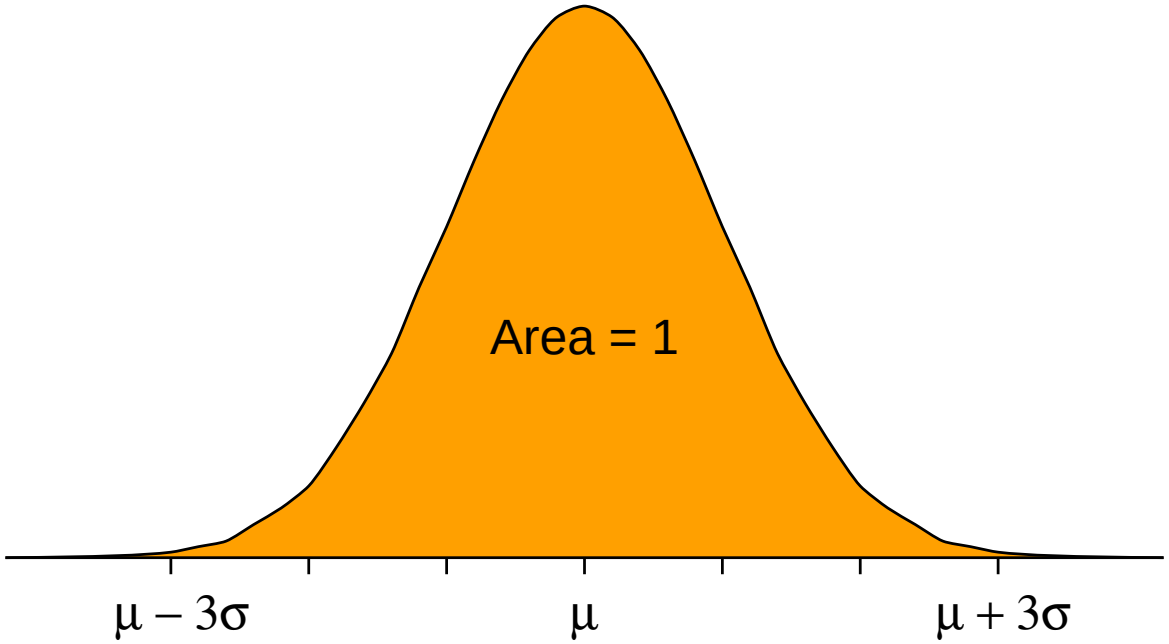
The mean μ is located in the center, at the peak

The Normal Distribution



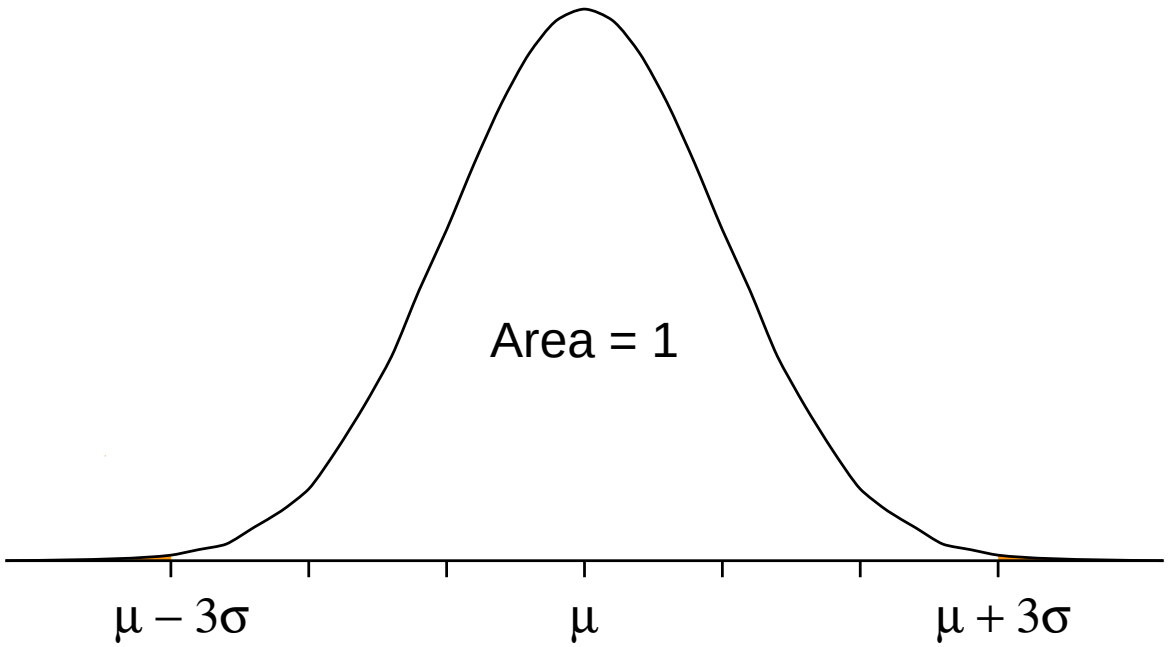
The width of the “main” part of the curve is 6 standard deviations wide

The Normal Distribution



The area under the entire curve is 1

The Normal Distribution



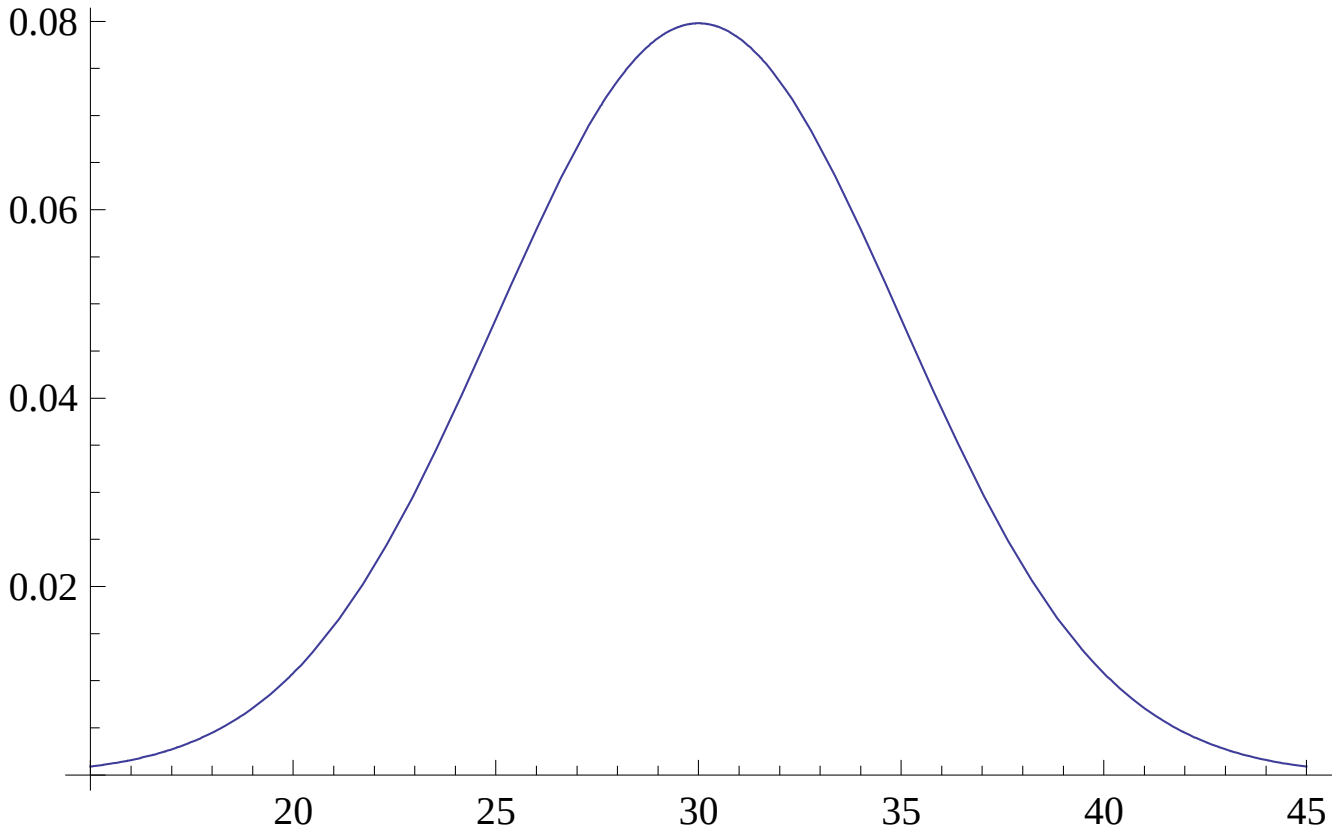
The area outside of $\mu \pm 3\sigma$ is approx. 0.0027

The Normal Distribution

- The normal distribution with mean μ and standard deviation σ is denoted $N(\mu, \sigma)$.
- For example, if X is a variable whose distribution is normal with mean 30 and standard deviation 5, then we say that "X is $N(30, 5)$."

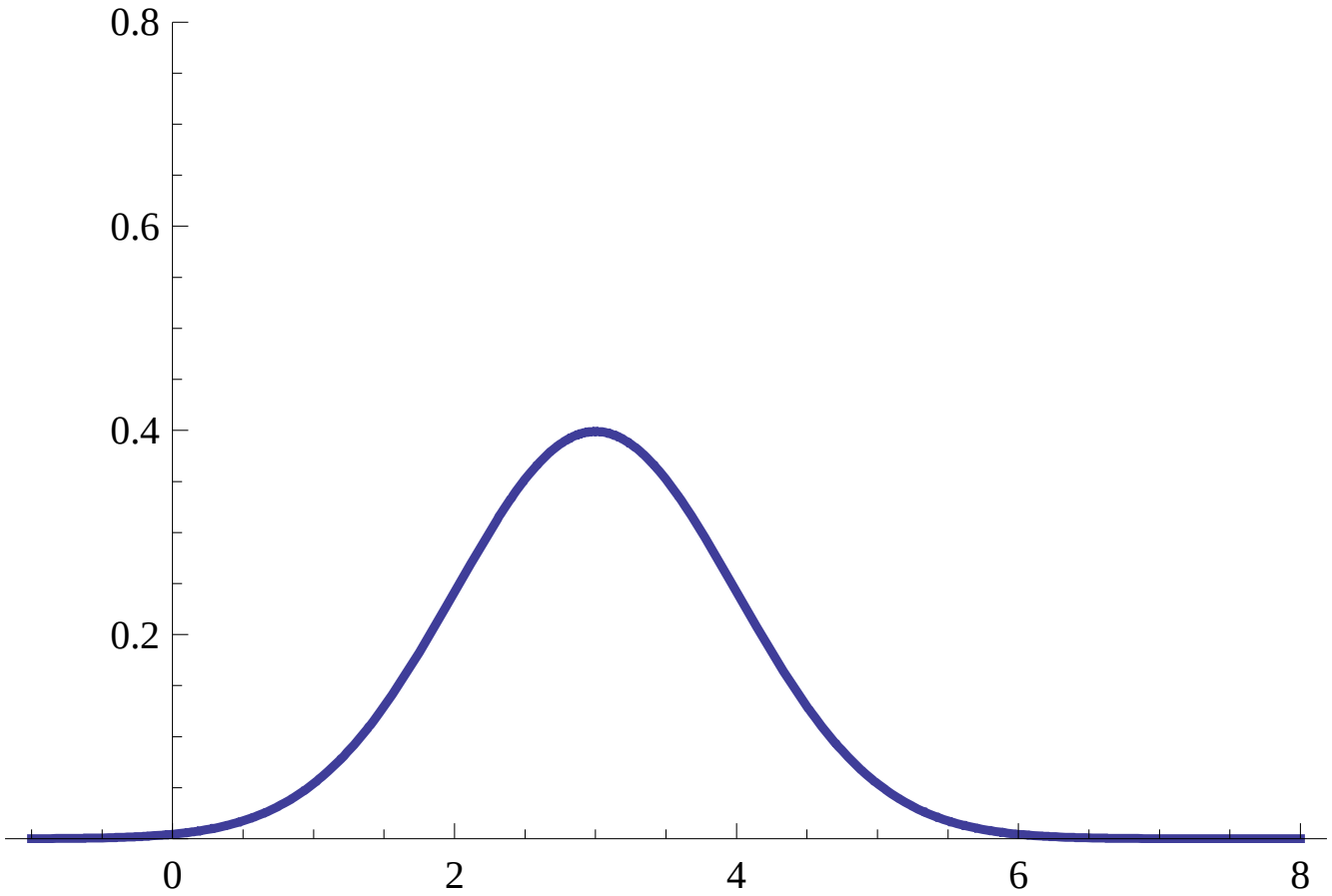
Outline

The Normal Distribution



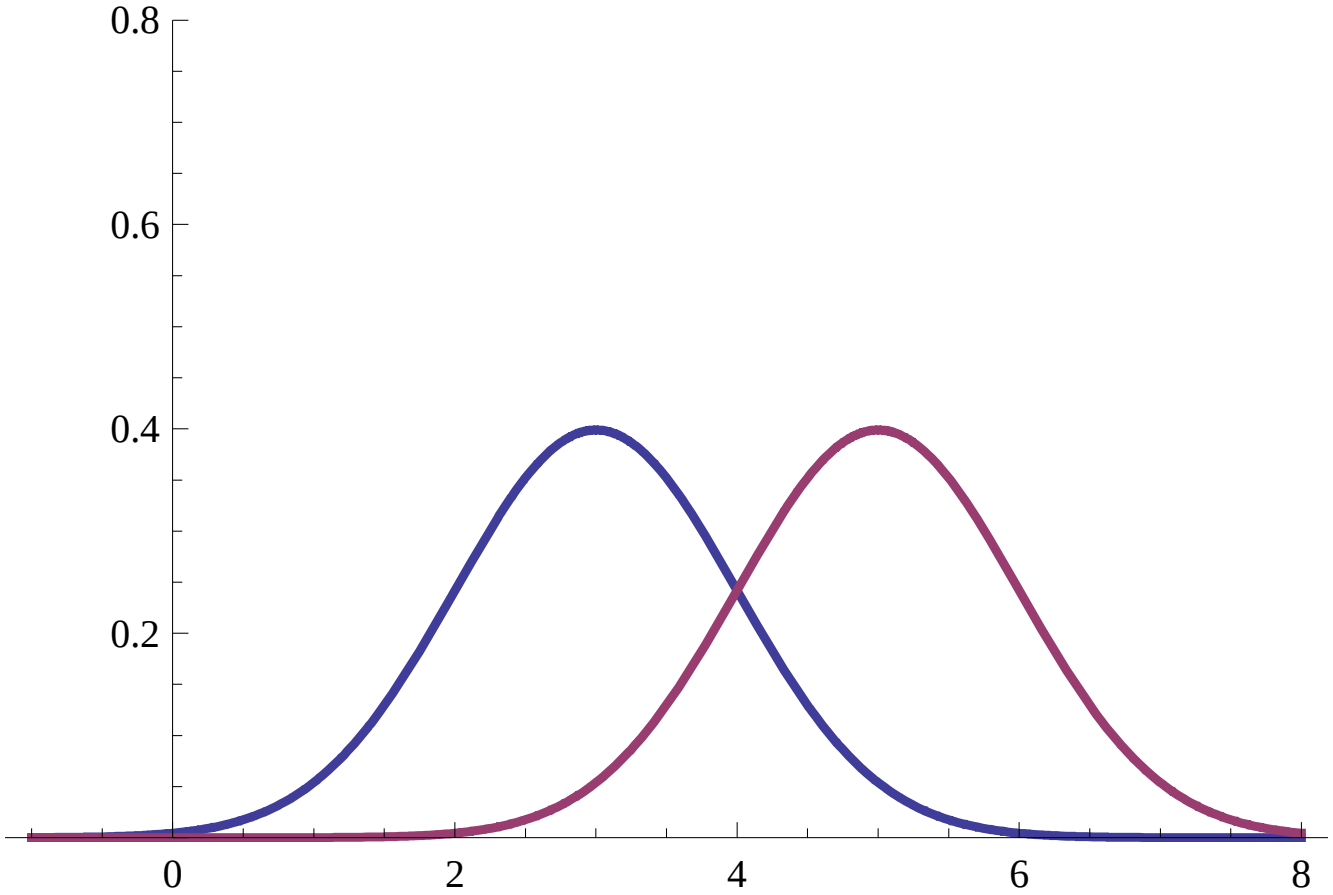
This is the graph of $N(30, 5)$

Some Normal Distributions



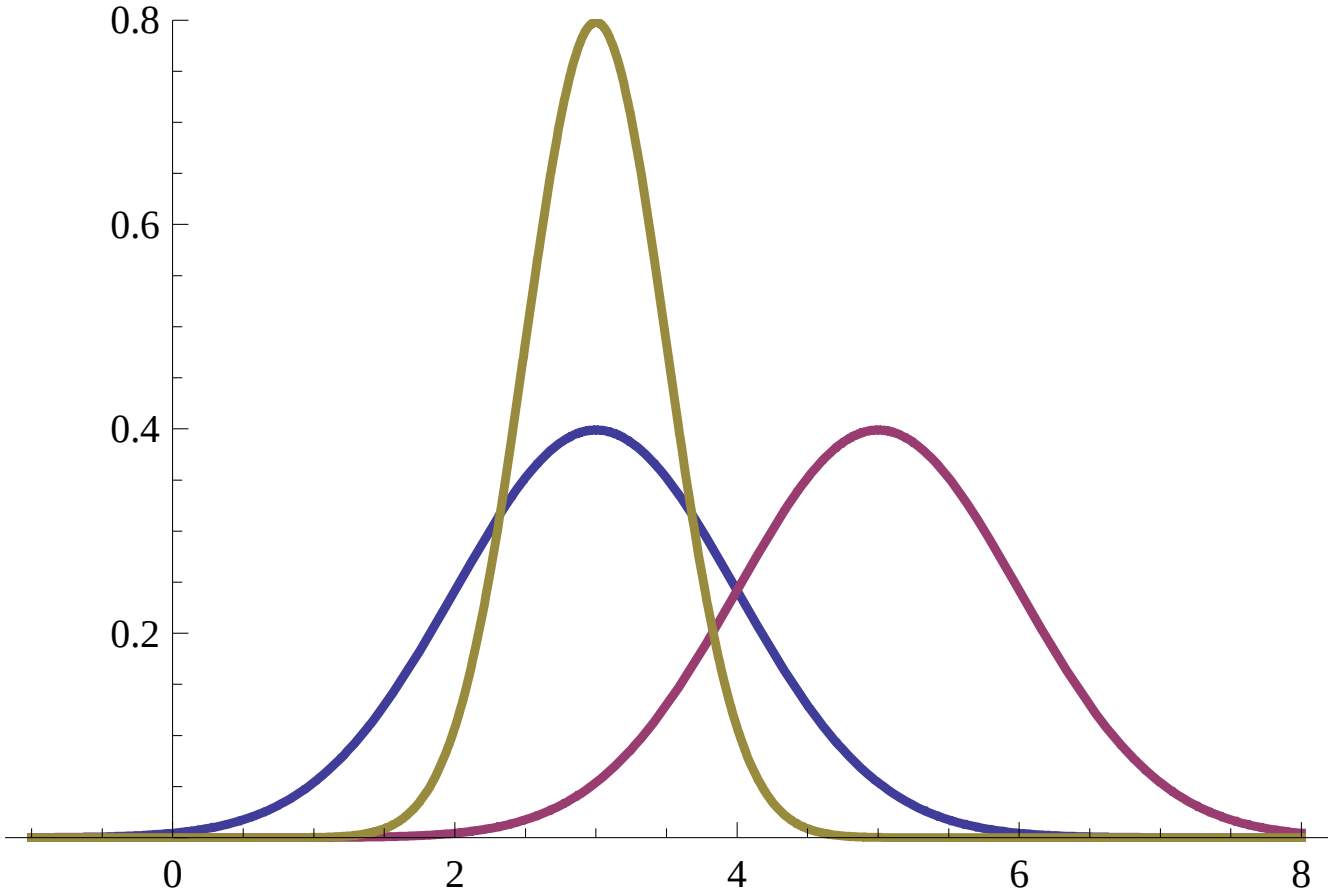
Some other examples: $N(3, 1)$

Some Normal Distributions



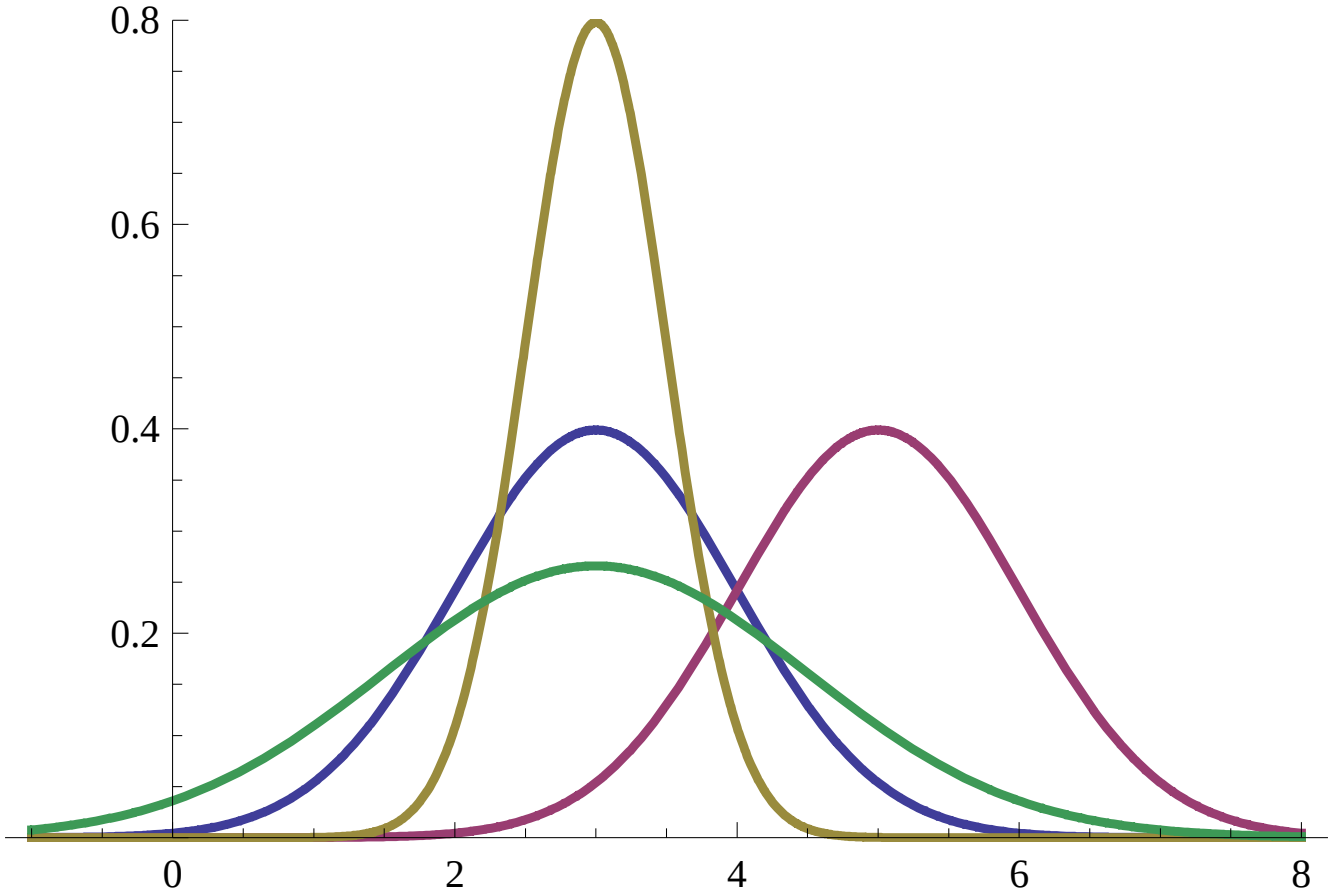
Some other examples: $N(5, 1)$

Some Normal Distributions



Some other examples: $N\left(2, \frac{1}{2}\right)$

Some Normal Distributions



Some other examples: $N\left(3\frac{1}{2}, 1\frac{1}{2}\right)$

Outline

Assignment

Homework

- Read Section 6.1 - 6.3.1, pages 357 - 362.
- Let's Do It! 6.1.
- Exercises 1, 2, 3, 32, page 376.