



[Course](#) > [Section...](#) > [1.2 Intr...](#) > Norma...

Audit Access Expires Mar 24, 2020

You lose all access to this course, including your progress, on Mar 24, 2020.

Upgrade by Feb 18, 2020 to get unlimited access to the course as long as it exists on the site. [Upgrade now](#)

Normal Distribution

Normal Distribution

the end.

normal
distribution



RAFAEL IRIZARRY: Histogram and density plots

provide excellent
summaries of a
distribution.

But can we summarize
even further?

We often see the average
and the standard
deviation

used as a summary
statistic for a list of



Video

[Download video file](#)

Transcripts

[Download SubRip \(.srt\) file](#)

[Download Text \(.txt\) file](#)

Textbook link

For more information, consult this [textbook section on the normal distribution](#).

Correction

At 3:27 and 3:50, the audio gives incorrect values for the average and standard deviation. The code on screen and the transcript are correct.

Key points

- The normal distribution:
 - Is centered around one value, the *mean*
 - Is symmetric around the mean
 - Is defined completely by its mean (μ) and standard deviation (σ)
 - Always has the same proportion of observations within a given distance of the mean (for example, 95% within 2σ)
- The standard deviation is the average distance between a value and the mean value.
- Calculate the mean using the `mean()` function.
- Calculate the standard deviation using the `sd()` function or manually.
- Standard units describe how many standard deviations a value is away from the mean. The z-score, or number of standard deviations an observation x is away from the mean μ :

$$Z = \frac{x - \mu}{\sigma}$$

- Compute standard units with the `scale()` function.
- **Important:** to calculate the proportion of values that meet a certain condition, use the `mean()` function on a logical vector. Because TRUE is converted to 1 and FALSE is converted to 0, taking the mean of this vector yields the proportion of TRUE.



Equation for the normal distribution

The normal distribution is mathematically defined by the following formula for any mean μ and standard deviation σ :

$$\Pr(a < x < b) = \int_a^b \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

Code

```
# define x as vector of male heights
library(tidyverse)
library(dslabs)
data(heights)
index <- heights$sex=="Male"
x <- heights$height[index]

# calculate the mean and standard deviation manually
average <- sum(x)/length(x)
SD <- sqrt(sum((x - average)^2)/length(x))

# built-in mean and sd functions - note that the audio and printed value
average <- mean(x)
SD <- sd(x)
c(average = average, SD = SD)

# calculate standard units
z <- scale(x)

# calculate proportion of values within 2 SD of mean
mean(abs(z) < 2)
```

Note about the sd function

The built-in R function `sd()` calculates the standard deviation, but it divides by `length(x)-1` instead of `length(x)`. When the length of the list is large, this difference is negligible and you can use the built-in `sd()` function. Otherwise, you should compute σ by hand. For this course series, assume that you should use the `sd()` function unless you are told not to do so.

Learn About Verified Certificates



edX Verified Certificate

Why upgrade?

- Official proof of completion
- Easily shareable certificate
- Proven motivator to complete the course
- Certificate purchases help edX continue to offer free courses

How it works

- Pay the Verified Certificate upgrade fee
- Verify your identity with a webcam and government-issued ID
- Study hard and pass the course
- Share your certificate with friends, employers, and others

edX Learner Stories



My certificate has helped me showcase my knowledge on my resume - I feel like this certificate could really help me land my dream job!

- Christina Fong, edX Learner



I wanted to include a verified certificate on my resume and my profile to illustrate that I am working towards this goal I have and that I have achieved something while I was unemployed.

- Cheryl Troell, edX Learner

Upgrade (\$49)

© All Rights Reserved

