Introduction to Statistics

Nels Grevstad

Metropolitan State University of Denver

ngrevsta@msudenver.edu

September 2, 2019

Topics

1. The Empirical Rule
2. Z-Scores

Objectives

Objectives:

- Apply the Empirical Rule to data that follow a bell-shaped distribution.
- Compute and interpret z-scores.
- Use z-scores as measures of the relative standing of individuals from two populations.

The Empirical Rule (3.3)

- The ***Empirical Rule*** (or ***68-95-99.7 Rule***) says that for data that follow a **bell-shaped** distribution:

   1. **68%** of the observations fall within **one standard deviation** of the **mean.**

The Empirical Rule (3.3)

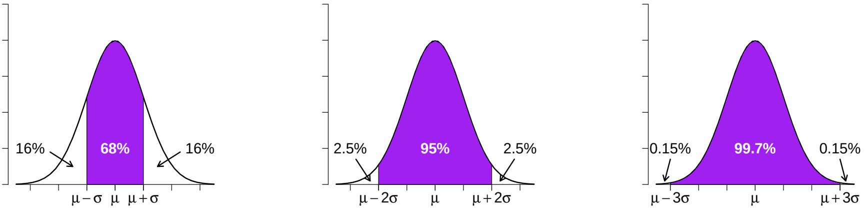
- The ***Empirical Rule*** (or ***68-95-99.7 Rule***) says that for data that follow a **bell-shaped** distribution:

  1. **68%** of the observations fall within **one standard deviation** of the **mean.**
  2. **95%** of the observations fall within **two standard deviations** of the **mean.**

The Empirical Rule (3.3)

- The ***Empirical Rule*** (or ***68-95-99.7 Rule***) says that for data that follow a **bell-shaped** distribution:

  1. **68%** of the observations fall within **one standard deviation** of the **mean.**
  2. **95%** of the observations fall within **two standard deviations** of the **mean.**
  3. **99.7%** of the observations fall within **three standard deviations** of the **mean.**


16%
68%
16%
μ − σ
μ
μ + σ
2.5%
95%
2.5%
μ − 2σ
μ
μ + 2σ
0.15%
99.7%
0.15%
μ − 3σ
μ
μ + 3σ

Exercise

Data on the heights of women aged 20 to 29 show that the **mean** height is **64** inches and the **standard deviation** of the heights is **2.7** inches. A histogram shows that the heights have a **bell-shaped** distribution

Exercise

Data on the heights of women aged 20 to 29 show that the **mean** height is **64** inches and the **standard deviation** of the heights is **2.7** inches. A histogram shows that the heights have a **bell-shaped** distribution

a) What percentage of women are **between 61.3** and **66.7** inches tall?

Exercise

Data on the heights of women aged 20 to 29 show that the **mean** height is **64** inches and the **standard deviation** of the heights is **2.7** inches. A histogram shows that the heights have a **bell-shaped** distribution

a) What percentage of women are **between 61.3** and **66.7** inches tall?

b) What percentage of women are **between 58.6** and **69.4** inches tall?

c) What percentage of women are **shorter** than **66.7** inches?

c) What percentage of women are **shorter** than **66.7** inches?

d) What percentage of women are **taller** than **72.1** inches?

Z-Scores (3.5)

Calculating a Z-Score

- If we know the population mean μ and standard deviation σ of a variable, then for any value x of that variable, the ***z-score*** for that value (also called its ***standardize value***) is:

Z-Score:

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

- If we don't know μ and σ, but we have a sample from the population, the ***z-score*** (***standardize value***) is:

Z-Score (when μ and σ are unknown):

$$z = \frac{x - \bar{x}}{s}$$

Properties and Interpretation of Z-Scores

1. The *z-score* indicates how many **standard deviations above or below the mean** an individual x value lies. These units are called ***standard units***.

Properties and Interpretation of Z-Scores

1. The *z-score* indicates how many **standard deviations above or below the mean** an individual x value lies. These units are called ***standard units***.

2. If an individual's x value is above the mean, their z**-score** will be **positive.** If it's below the mean, their z**-score** will be **negative.** If it's equal to the mean, their z**-score** will be **zero.**

3. The **farther away from zero** a z-score is, the more **"atypical"** the individual is. In particular, for data with a bell-shaped distribution, the Empirical Rule says:

    - A z-score larger than 2.0 or smaller than -2.0 occurs only 5% of the time (i.e. **not very often**).

3. The **farther away from zero** a z-score is, the more **"atypical"** the individual is. In particular, for data with a bell-shaped distribution, the Empirical Rule says:

   - A z-score larger than 2.0 or smaller than -2.0 occurs only 5% of the time (i.e. **not very often**).
   - A z-score larger than 3.0 or smaller than -3.0 occurs only 0.3% of the time (i.e. **very rarely**).

Exercise

The study of women's heights also looked at heights of men between the ages 20 to 29 years. It found that the **mean** height of men is **69.3** inches with a **standard deviation** of **2.8** inches.

a) Suppose a man is **72** inches (six feet) tall. **Calculate** and **interpret** his z**-score**.

Exercise

The study of women's heights also looked at heights of men between the ages 20 to 29 years. It found that the **mean** height of men is **69.3** inches with a **standard deviation** of **2.8** inches.

a) Suppose a man is **72** inches (six feet) tall. **Calculate** and **interpret** his z**-score**.

b) Suppose a man is **84** inches (seven feet) tall. **Calculate** and **interpret** his z**-score**.

The Z-Score as a Measure of Relative Standing

- A z-score indicates where an individual x value stands **relative to the rest of the population.** So we can compare individuals from two ***different*** populations by comparing their z-scores.

Exercise

Recall that the **mean** height for **women** is **64** inches with a **standard deviation** of **2.7** inches, and the **mean** height for **men** is **69.3** inches with a **standard deviation** of **2.8** inches.

Exercise

Recall that the **mean** height for **women** is **64** inches with a **standard deviation** of **2.7** inches, and the **mean** height for **men** is **69.3** inches with a **standard deviation** of **2.8** inches.

If a **woman** is **69** inches tall, and a **man** is **72** inches tall, who's taller *relative to his or her gender*?