Chapter Five: Standard Deviation and the Normal Distribution

We're all in a higher educational institution here, so chances are good that you have either taken the ACT and SAT. Suppose we wanted to compare ACT and SAT scores, but these scores are on different scales. How might we compare them?

Recall, that for nice-looking distributions we can use the average or mean as a measure of center and the standard deviation as the measure of spread. Using the mean and the standard deviation we can ______________________________ the scores to compare them.

Z-Scores

A ______________________________ is essentially a unitless measure of how extreme an observation is compared to its ______________________________.

Let:

- y represent the quantitative variable of interest
- $\bar{y}$ represent the mean of the variable of y.
- σ represent the standard deviation of the quantitative variable y.

To calculate a z-score we use the following equation:

Properties

-
-

  Distributions of Z have:

-
-
-

Interpretation

z is the ____ an observation is away from the mean.

-
-
-

Example *Suppose the average score on the ACT English Section is 21 with a standard deviation of 4 and the average score of the SAT Verbal section is 520 with a standard deviation of 100. Suppose we want to compare two students, Ann scores a 27 on the ACT English exam, whereas Denise scores a 770 on the SAT Verbal exam. Who scores better on their standardized test? Denise or Ann?*

Example *Suppose the average on the ACT Math exam is 21 points with a standard deviation of 4.1 points. Similarly, suppose the average on the SAT Math exam is 510 with a standard deviation of 100. let's compare two students, Jim who scored a 15 on the ACT Math section and Joe who scored a 340 on the SAT Math section. Who has the better score?*

What we are doing by calling these z-scores is known as ______________________________ which allows us to compare observations between different quantitative variables. If we knew specifically the distributional information of the quantitative variables we are comparing, we could compare using percentiles.

The Normal Distribution

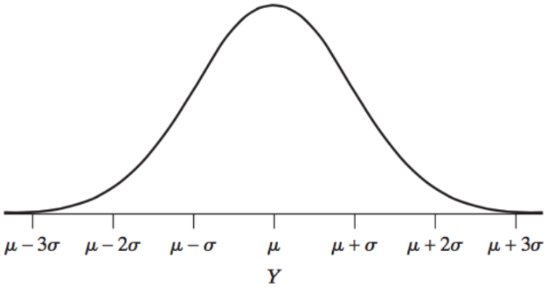
In statistics we use what's known as ___________________________________ to represent processes/phenomenon in our world. We can then use these models to analyze the data we observe.

Models are chosen based on their:

-
-
-

One of the most important models in statistics, and the only one you'll learn in this course is the ___________________________________.

- Shape
    -
    -
- Notation
    -
    -
    -



Some implications of these characteristics:

-
-
-

When we are talking about the Normal Distribution it's important to note that the μ and σ ____.
They are considered parameters rather than summary statistics.

Many times, data from the real world doesn't exactly follow the Normal model. So how can we tell if a dataset is distributed Normally?

-
-

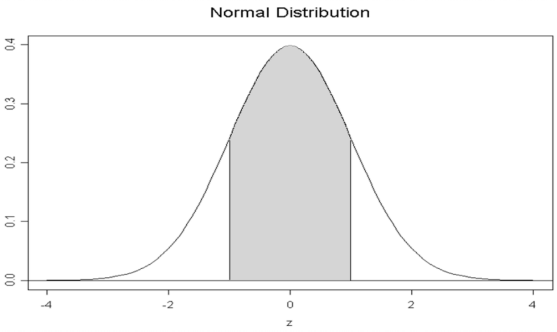
68-95-99.7 Rule

The following holds for all Normal Distributions:

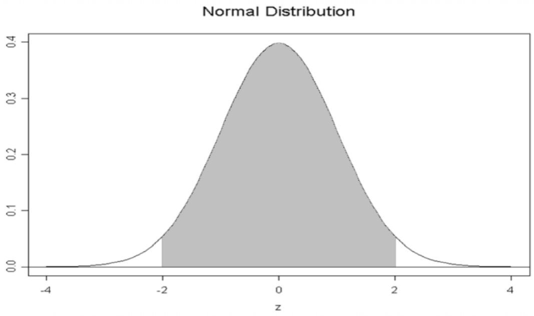
-
-

-

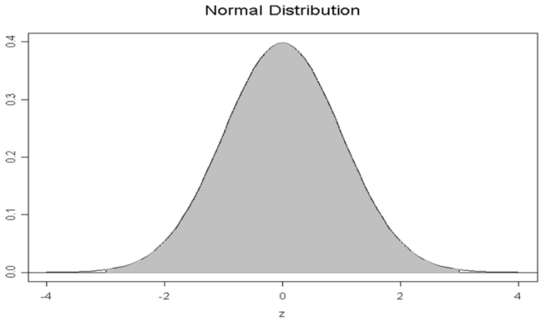
We note that if $\mu = 0$ amd $\sigma = 1$, then ________________________________ are between -1 and 1.



If $\mu = 0$ amd $\sigma = 1$, then ________________________________ are between -2 and 2.



if $\mu = 0$ amd $\sigma = 1$, then ________________________________ are between -3 and 3.



Example *Recall that the heights of men are approximately distributed N(70,3). Using the 68-95-99.7 rule, answer the following questions.*

a) *The middle 68% of men have heights between which two values?*

b) *What percentage of men have heights greater than 73 inches?*

c) *What percentage of men have heights between 64 and 73 inches?*