Z - Scores

*Why is this Important?*

- How do you compare apples and oranges?
- Are you as good a student of French as you are in Physics?
- How many people did better than you on a test?
  - How many did worse?
- Are you abnormal or deviant?
- Dr So and So, are you going to curve the exam?

If you are going to understand a situation, you will need to able compare things. You will need to present the comparisons in a way that facilitates decision-making. You may be asked:

- What percentage of people falls below a given score?
- What is the relative standing of a score in one distribution versus another?
- What score or scores can be used to define an extreme or deviant situation?

You can do all of these with the **Z - Score!**

*Data Types*

For our discussion of the **Z-scores**, we will be using **measurement data** (numerical scores).

*We Will Need:*

- Measures of **central tendency**: The Mean
- Measures of **variability**: Standard Deviation

*Our First Intuitive Example*

You enter your dorm room and find that one of your roommates has a happy face:
But your other roommate is sad:

What's happened! Why are you so happy! But why are You so sad?

They reply:

Hurray, I got a 60 on the Physics test!!!

Wah!, I got a 60 on the French test!!!

They have the same score! Why the different reactions!

As a student, you probably have figured out that the 58 on the French test is bad compared to every one else and that the 58 on the Physics test may be pretty good compared to everyone else.
Our physics major roommate, says "That great old Prof. - she curved the test!".

We need to be more precise - To do that we need to review the Gaussian Distribution! That's how we will compare the French and Physics tests.

**Finding Yourself on the Gaussian Curve**

Remember the **Gaussian curve**. It is the **shape** that quite a lot of measurement data distributions have. That means that the frequency of the scores tend to fall in a very specific shape.

The Gaussian curve is **described by** its **mean** (average) and **standard deviation**.

**Hints:**
- This is the **Gaussian curve** and the line in the middle is the mean. **Mean** - $\Sigma X/N$ or the average.
- Many frequency distributions in Psychology are normal.
- A little box on the graph represents a **score**. In our case it will be an exam score of a particular individual.
- The **red arrow** represents the **standard deviation**. A measure of the width of a measurement data distribution
  $$\sqrt{\frac{\Sigma (X-\text{mean})^2}{N}}$$

**SO WHAT DOES THIS HAVE TO DO WITH THE FRENCH AND PHYSICS TEST!**

**THE BIG IDEA!!**
Because the Physics test was very difficult with a mean of 50, one roommate did better than most (score of 70) in the class as was quite pleased with herself. You can see that most people in the distribution are below her.

It is the opposite for our French-taking roommate. She is scored below most.

**The 2\textsuperscript{nd} Big Idea!!**

We can quantify your place in the distribution. You don’t have to draw a picture every time.

We will do this with the Z-Score.

**The 3\textsuperscript{rd} Big Idea!!**

The Standard Deviation is a Ruler.

Consider the Z - score as the Inch Marks of the Ruler.

A Z-Score Equals the Number of Standard Deviations, the score is from the Mean

$$Z = \frac{(\text{Score} - \text{Mean})}{\text{Standard Deviation}}$$
**THE CONCEPT**

**Rulers**

**Your School Bag**
- 12 Inches Long
- Marked in Inches
- Center is 6 "

**Normal Curve**
- 100% of the Scores
- Marked in Z-Scores
- Center is the Mean

---

**Let’s Look at the Physics Test!**

- We know its mean is 50.
- The Standard Deviation is calculated to be 10 (someone used their calculator).
- Remember the SD is the distance from the Mean from to the Point of Inflection (change in direction of the curve)

![Normal Curve Diagram]

10 Points = 1 SD
2 SDs = 20 points

Mean = 50
Roommate score = 60

---

**THE 4TH BIG IDEA!!**

What’s our Physics Freak Roommate’s Z - score
And Why is She Happy!
1. The roommate’s score is 10 points above the mean or one SD above the mean.
2. Her Z-Score Equals 1 because:

\[ Z = \frac{\text{Score} - \text{Mean}}{\text{SD}} \]

\[ Z = \frac{60 - 50}{10} \]

\[ Z = 1 \]

3. Let’s do it again! The student down the hall runs in to our gang’s room. She yells - I got a 84!! Yahoo!
2. Her Z-Score Equals 3.4 because:

\[ Z = \frac{\text{Score} - \text{Mean}}{\text{SD}} \]

\[ Z = \frac{84 - 50}{10} \]

\[ Z = 3.4 \]

This is a bright young person. Let’s look at the Z-Score ruler on the Gaussian Curve. She gets a very big happy face

She’s done better than almost everyone in the class!

- Oh, Dear - Now we know we our French Taking Room mate is Sad!
Her Z-score in negative and below the mean.

**Nuances About the Z-Score:**
1. The Z-score can be positive or negative.
   - Positive is above the mean.
   - Negative is below the mean.
2. The Mean of the Z scores is Zero.
   - If you got the class average on a test - your Z-score = 0.
   - Because Z would be (Mean-Mean)/0 = 0.
3. The Standard Deviation of the Z distribution = 1. We will discuss this later.

**A REALLY BIG CONCEPT!**

**Z-Scores can be Turned into Percentages**

Here’s the deal - if you know a Z-score - you know what percent or proportion of scores are above and below that score.

- It’s like slicing a pie.

Cut it in quarters and you know that each quarter as 25% of the pie.

- Let’s slice the Gaussian Curve
The distribution holds 100% of the scores or a proportion of 1.0

The gray line cuts it in half. So - half are below and half above.
If you received the average score on the test ($Z = 0.0$), you did better than 50% of the class.

What about cutting the distribution in another place?

- Calculus types have calculated the percents above and below any $Z$-Score.
- We mentioned this is Intro to Stats #1.
- You statistics book has a table with this information in it.

**Some Examples:**

**Our Physics Students:**

1. The score of 60 was 1 Standard Deviation ($Z= 1.0$) above the mean.
   The table would tell me that there are 84.13% of scores below.
2. The score of 84 was 3.4 SDs above the mean ($Z = 3.4$).
   The table tells me that there were 99.97% scores below it.
   *No wonder these kids are happy.*
Our Poor French Student:

The score of 60 was 1 SD below the mean ($Z = -1.0$)

The table would tell me that there were 84.13% scores above it.

You can do this with any score if you have a Gaussian distribution.

More Nuances:

1. There are many variants of this procedure:
   A. Deviation Problem – What’s in Your Tails? - Huh?
      Let's go back to the Physics Test.
      You are the professor and the chair of the department wants to know what scores cut off the top 5% of the students and the lowest 5%.
      The former get an award and the latter get a scolding!

Here’s the Answer

I looked up the $Z$-score that cuts off 5% in the tails. The score is 1.64 as you can see in most Z tables.
Thus, on my $Z$-score ruler, the test scores would have to be plus or minus 1.64 $Z$ units from the mean. One $Z$ unit (our inch, so to speak) equals 10 points.
So I get the picture below and my two cut-offs of 33.6 (low) and 66.4 (high)
2. Cutting of percentages in the tail of a distribution will develop the idea of significance testing:

3. You can find the raw score that corresponds to a percentage by:
   a. Looking up the Z score that cuts off the percent you want
   b. Use the following formula:
      \[ X = \text{Mean} + Z \cdot (\text{SD}) \]
      For example:
      
      ![Normal Distribution Diagram]
      
      I want the score that cuts off 75%.
      \[ X = \text{Mean} + Z \cdot (\text{SD}) \]
      For the French Test =
      \[ X = 70 + 0.67 \cdot (10) = 76.7 \]
      0.67 came from the Z table.

4. The Z-Score is crucial to understanding:
   a. t-test
   b. Correlation Coefficient (Pearson’s r)

5. If the distribution isn’t Gaussian – that’s for more advanced discussions.

**Bottom Line:**

*If you are have a set of measurement scores on different measures using Z-scores you can tell how the scores are placed in their distributions. Then you can compare them. Also, knowing the mean and standard deviation of a distribution that is reasonably Gaussian allows you to determine the percentile rank of that score. So curving a test - means that a professor marks you on a z-score basis - but what if the mean is high? Huh? Think about it.*