

## 5.5: Summary of z Scores

Although this section has some math and the z-score formula, it's really to hit the main points of z-scores.

### Summary of z-scores

Suppose Dr. Navarro's friend is putting together a new questionnaire intended to measure “grumpiness”. The survey has 50 questions, which you can answer in a grumpy way or not. Across a big sample (hypothetically, let's imagine a million people or so!) the data are fairly normally distributed, with the mean grumpiness score being 17 out of 50 questions answered in a grumpy way, and the standard deviation is 5. In contrast, when Dr. Navarro takes the questionnaire, she answer 35 out of 50 questions in a grumpy way. So, how grumpy is she? One way to think about would be to say that I have grumpiness of 35/50, so you might say that she's 70% grumpy. But that's a bit weird, when you think about it. If Dr. Navarro's friend had phrased her questions a bit differently, people might have answered them in a different way, so the overall distribution of answers could easily move up or down depending on the precise way in which the questions were asked. So, I'm only 70% grumpy *with respect to this set of survey questions*. Even if it's a very good questionnaire, this isn't very a informative statement.

### Can we standardize?

A simpler way around this is to describe Dr. Navarro's grumpiness by comparing me to other people. Shockingly, out of the friend's sample of 1,000,000 people, only 159 people were as grumpy as Dr. Navarro (which Dr. Navarro believes is realistic), suggesting that she's in the top 0.016% of people for grumpiness. This makes much more sense than trying to interpret the raw data. This idea – that we should describe a person's grumpiness in terms of the overall distribution of the grumpiness of humans – is the idea that standardization attempts to get at. One way to do this is to do exactly what I just did, and describe everything in terms of percentiles. However, the problem with doing this is that “it's lonely at the top”. Suppose that the friend had only collected a sample of 1000 people (still a pretty big sample for the purposes of testing a new questionnaire, I'd like to add), and this time gotten a mean of 16 out of 50 with a standard deviation of 5, let's say. The problem is that almost certainly, not a single person in that sample would be as grumpy as me.

However, all is not lost. A different approach is to convert my grumpiness score into a **standard score**, also referred to as a z-score. The standard score is defined as the number of standard deviations above the mean that my grumpiness score lies. To phrase it in “pseudo-math” the standard score is calculated like this:

$$\text{standard score} = \frac{\text{raw score} - \text{mean}}{\text{standard deviation}}$$

In actual math, the equation for the z-score is:

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

So, going back to the grumpiness data, we can now transform Dr. Navarro's raw grumpiness into a standardized grumpiness score. If the mean is 17 and the standard deviation is 5 then her standardized grumpiness score would be

$$z = \frac{X - \bar{X}}{s} = \frac{35 - 17}{5} = \frac{18}{5} = 3.6$$

To interpret this value, recall that 99.7% of values are expected to lie within 3 standard deviations of the mean. So the fact that my grumpiness corresponds to a z score of 3.6 (3.6 standard deviations above the mean) indicates that I'm very grumpy indeed.

So now that we know that we can compare one person's individual raw score to the whole distribution of scores, what else can the z-score do?

### Compare Across Distributions

In addition to allowing you to interpret a raw score in relation to a larger population (and thereby allowing you to make sense of variables that lie on arbitrary scales), standard scores serve a second useful function. Standard scores can be compared to one another in situations where the raw scores can't. Suppose, for instance, my friend also had another questionnaire that measured extraversion using a 24 items questionnaire. The overall mean for this measure turns out to be 13 with standard deviation 4; and I scored a 2. As you can imagine, it doesn't make a lot of sense to try to compare my raw score of 2 on the extraversion questionnaire to my raw score of 35 on the grumpiness questionnaire. The raw scores for the two variables are about fundamentally different things, so this would be like comparing apples to oranges. BUT, if we calculate the z-scores, we get

$z = \frac{(35 - 17)}{5} = 3.6$  for grumpiness and  $z = \frac{(2 - 13)}{4} = -2.75$  for extraversion. These two numbers *can* be compared to each other. Dr. Navarro is much less extraverted than most people ( $z = -2.75$ ) (the negative sign means that Dr. Navarro is 2.75 standard deviations below the average extraversion of the sample) and much grumpier than most people ( $z = 3.6$ ); but the extent of my unusualness is much more extreme for grumpiness (since 3.6 is a bigger number than 2.75). Because each standardized score is a statement about where an observation falls *relative to its own population*, it is possible to compare standardized scores across completely different variables.

This ability to compare scores across different distributions is the foundation of statistical analyses. It allows us to compare the sample that we have to a probability sample, and then make predictions!

### Proportions & Amounts

But that's not all! Using this z-score formula to turn a raw score into a standardized z-score, we can then use what we know about the proportions of scores in standard normal curves to predict how likely that score was. And if we know the sample size, we can even predict how many people from the sample will score above (or below) that initial raw score. Amazing!

This all is pretty cool, although you probably don't think so yet. Or maybe never will. But the way that the Standard Normal Distribution and these standardized scores (z-scores) function allows much of the rest of statistics to exist.

### Next: Write-Ups

You are in a social science class, not a statistics class, so finding the number is almost never the end of your task. The next section will describe the important components of a good concluding sentence or concluding paragraph.

### Contributors and Attributions

- [Danielle Navarro](#) ([University of New South Wales](#))
- 

[Dr. MO](#) ([Taft College](#))

---

This page titled [5.5: Summary of z Scores](#) is shared under a [CC BY-SA 4.0](#) license and was authored, remixed, and/or curated by [Danielle Navarro](#).