Statistics for Beginners

Why do we convert values to standard scores?

Before we answer the question above, let's look at the example as follows:

- **Example 1**

Julie got a mark of 80 from a math class with a mean of 85 and a standard deviation of 5 and her friend Andrea obtained 65 from a chemistry class with a mean of 55 and a standard deviation of 10. Can you tell who got a "better" mark?

At first glance, the mark 80 seems like a "higher/better" mark. But, on second thought, we might realize that it is not true because whether it is a "better" mark or not depends on the other students' marks in the class as well. It might seem weird to say that 65 is a better mark than 80, but it is true because 65 is above the average mark 55 and meanwhile 80 is below the average mark 85. **Therefore, when we compare the two values, we need to consider the means.**

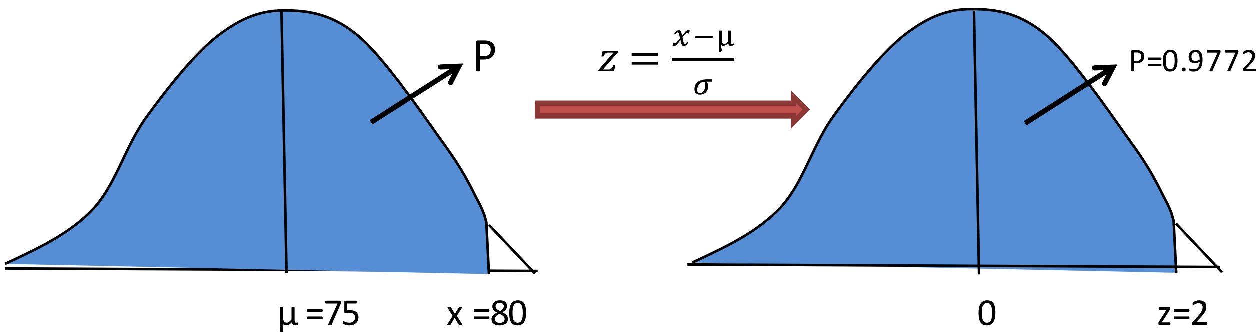
Beside means, we need to take standard deviation into account as well because standard deviation shows how much variation there is from the average (mean, or expected value). A low standard deviation indicates that the data points tend to be very close to the mean, whereas high standard deviation indicates that the data are spread out over a large range of values. Here is another example with same values and means but different standard deviations as follows:

- **Example 2**

Jason got 80 from a history class with a mean of 75 and a standard deviation of 2.5 and his friend Mike obtained 80 from an English class with a mean of 75 and a standard deviation of 5. Can you tell who got a better mark?

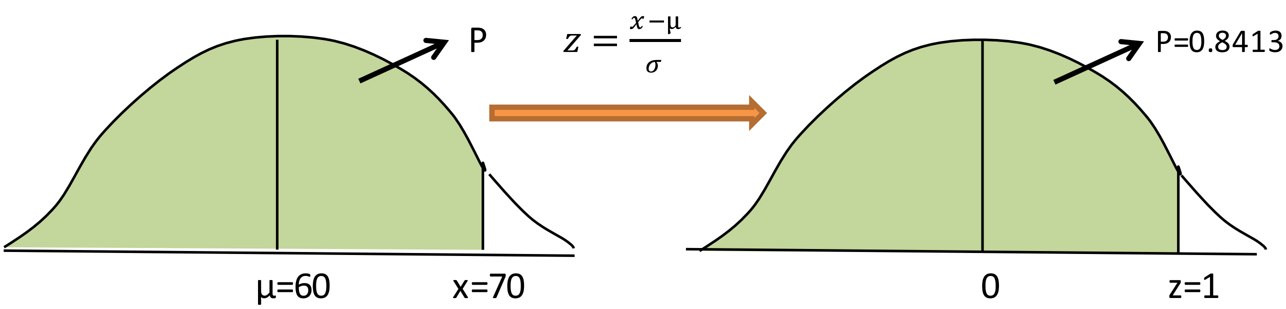
Both marks are 5 points above the average marks but they have different standard deviations. **In order to tell which student has a better mark, we need to convert the two marks to standard scores with same mean $\mu = 0$ and standard deviation $\sigma = 1$ (i.e. under the same standard condition) so that we can tell who got a better mark.** (The standard normal distribution is a normal probability distribution with $\mu = 0$ and $\sigma = 1$, and the total area under its density curve is equal to 1.) We use the formula $z = \frac{x-\mu}{\sigma}$ to convert to values to standard scores (z scores).

- **Jason:** $z = \frac{x - \mu}{\sigma} = \frac{80 - 75}{2.5} = 2$



We use the formula $z = \frac{x - \mu}{\sigma}$ to convert Jason's mark (x=80) to standard score (z=2) which is a positive number showing that this mark is greater than the average mark (z score for the mean μ is 0 and 2 > 0). The greater the positive number it is, the better mark it is. Using Table A-2, we can find the area (probability) below z=2 is 0.9772 which indicates that Jason's mark is greater than 0.9772 (or 97.72%) of the students' marks in the history class. The same procedure works for Mike's mark.

- **Mike:** $z = \frac{x - \mu}{\sigma} = \frac{80 - 75}{5} = 1$



The z score for Mike's mare is 1 which is a positive number and shows that Mike's mark is greater than the mean. Using Table A-2, the area under z=1 is 0.8413 which means that 0.8413 (or 84.13%) and it indicates that Mike's mark is greater than 84.13% of the students' marks in the English class. By comparing Jason and Mike's marks (2 > 1) or the corresponding percentage below the z scores (97.72% > 84.13%), we can see that Jason actually got a "better" mark than Mike did.

In conclusion, to compare two values with different means and standard deviations, we can convert them to standard scores (z scores) by the formula $= \frac{x-\mu}{\sigma}$. Positive z score shows that the value is greater than the mean and negative z score indicates that the value is lower than the mean. The greater the z score is, the better the mark it is or the greater probability below the value is.

Can you figure out the z scores for Julie and Andrea's marks and the probabilities under the two z scores in the example 1? (The answer can be found at "Answer Key for the Example 1".)