1 EXPLORING DATA

1.07 Z-scores

What you see here is the so-called “tattoo density” of football players, expressed in the percentage of the body covered with tattoos. The dot plots and the standard deviations show that there is much more variability in the distribution of Team 2 than in the distribution of Team 1.

Sometimes, researchers ask the question if a specific observation is common or exceptional. To answer that question, they express a score in terms of the number of standard deviations it is removed from the mean. This number is what we call a ***z*-score**. In this video I’ll explain how you can compute *z*-scores and I’ll tell you why they can be useful.

Let’s first take a look at the distribution of Team 1. The mean is 15 and the standard deviation is 2.5. To compute *z*-scores we use *this* formula. It is not very complicated. It tells you to compute for the value you’re interested in the difference between that value and the mean, and to divide the outcome by the standard deviation. Let’s see what that means for a tattoo density of 10.8 percent. The z-score of that value is 10.8 minus 15 divided by 2.5. That equals -1.68. So, the z-score is -1.68. You can do that for all the values in your distribution. If you do that here, *these* are the results.

Notice that you end up with negative z-scores and positive z-scores. Negative z-scores represent values below the mean. And positive z-scores represent values above the mean. Because the mean is the balance point of your distribution, the negative and positive z-scores cancel each other out. In other words, if you add up all z-scores you will get a value of 0.

Okay, that’s nice, but how do you know if a certain z-score is low or high? Well, that depends on your distribution and on context. A good rule is that IF the histogram of your variable is bell-shaped, 68 percent of the observations fall between z-scores of minus 1 and 1; 95 percent between z-scores of minus 2 and 2; and 99 percent between z-scores of minus 3 and 3.

This means that for *this* type of distribution, a z-score of more than 3 or less than minus 3 can be conceived of as rather exceptional. However, if a distribution is strongly skewed to the right, as in *this* graph, large positive z-scores are more common, because there are more extreme values on the right side of the distribution. Similarly, if a distribution is strongly skewed to the left, large negative z-scores are more common, because there are more extreme values on the left side of the distribution. A rule that applies to any distribution, regardless shape, is that 75 percent of the data must lie within a z-score of plus or minus 2 and 89 percent within a z-score of plus or minus 3.

So, in itself a z-score gives you, to a certain extent, information about how extreme an observation is. Z-scores are even more useful if you want to compare different distributions. Let’s, for example, look at the question whether a body weight of 19.3 is common or not. Well, in Team 1 it is not *that* common. The z-score is 19.3 minus 15 divided by 2.5. That equals 1.72. In Team 2 the value of 19.3 corresponds with a z-score of 19.3 minus 15 divided by 8 equals 0.54. This indicates that in Team 2 the value of 19.3 is much more common than in Team 1. In Team 2 it is only 0.54 standard deviations removed from the mean. In Team 1 it is 1.72 standard deviations removed from the mean.

If we recode original scores into z-scores, we say that we **standardize** a variable. Standardization means that we replace the scores measured in the original metric by scores expressed in standard deviations from the mean. The advantage is that we can see at a glance whether a specific score is relatively common or exceptional.

So, is a football player covering about one-fifth of his body with tattoos exceptional or not? Well, that depends on his football team – or whichever other group you want to compare him with!