

## Lesson 6

### Z-Scores

#### Outline

Linear Transformation

-effect of addition/subtraction

-effect of multiplication/division

Z-Transformation

Z-Distribution

-properties

Using Z-scores to compare values

#### **Linear Transformation**

Anytime we change a distribution by using a constant we perform a linear transformation. For example if I measure the heights of everyone registered in this course, but then found the tape measure I was using was missing the first two inches (so it started at inch two instead of zero), what would I have to do to find the true heights of everyone? If you think about it you will see that I must subtract two inches from each measurement to get the true heights (because the start position was too high). This example of a linear transformation is one in which we simply shift the numbers up on the same scale. Notice that even though all the numbers move, the relationship between values is not affected.

<u>X</u>	<u>X + 2</u>
5'5"	5'7"
5'7"	5'9"
5'8"	5'10"
5'10"	6'0"

You will need to know how linear transformations affect the mean and standard deviation of a distribution as well. How does adding or subtracting a constant affect the mean and standard deviation? How does multiplying and dividing a constant affect the mean and standard deviation?

When adding or subtracting a constant from a distribution, the mean will change by the same amount as the constant. The standard deviation will remain unchanged. This fact is true because, again, we are just shifting the distribution up or down the scale. We do not affect the distance between values. In the following example, we add a constant and see the changes to the mean and standard deviation.

X	X + 5
1	6
2	7
3	8
4	9
5	10
$\mu = 3$	$\mu = 8$
$\sigma = 1.41$	$\sigma = 1.41$

The effect is a little different when we multiply or divide by a constant. For these transformations the mean will change by the same amount as the constant, but this time the standard deviation will change too. That is because when we multiply numbers together, for example, we change the distance between values rather than just shifting them up or down the scale. In the following example, we multiply a constant and see the changes to the mean and standard deviation.

X	X * 5
1	5
2	10
3	15
4	20
5	25
$\mu = 3$	$\mu = 8$
$\sigma = 1.41$	$\sigma = 7.91$

### Z-Transformation

The z-transformation is a linear transformation, just like those we have discussed. Transforming a raw score to a Z-score will yield a number that expresses exactly how many deviations from the mean a score lays. Here is the formula for transforming a raw score in a population to a Z-score:

$$z = \frac{X - \mu}{\sigma}$$

Notice that the distance a score lies from the mean is now relative to how much scores deviate in general from the mean in the population. Regardless of what the raw score values are in the population, when we use the Z-transformation we obtain a **standard** measure of the distance of the score from the mean. Anytime  $Z=1$ , the raw score that produced the Z is exactly one standard deviation from the mean for that population. Anytime  $Z=1.5$ , the raw score that produced the Z is exactly 1.5 standard deviations from the mean for that population.

Think for a minute about what it means to know how many standard deviations from a mean a score lays. Consider our simple distribution example.

X
1
2
3
4
5
$\mu = 3$
$\sigma = 1.41$

What z-score will we expect the value “3” to have in this example? That is, how many standard deviations from the mean is “3”? The answer is that it is at the mean, so it is zero standard deviations from the mean and we will get a z-score of zero for the original value of three. Now consider the value “1” in the distribution. What z-score will we expect to get for this score? Will it be less than one standard deviation or more than one standard deviation away from the mean? You can estimate the z-score by counting from the mean. One standard deviation is 1.41 units. Counting down from the mean the value “2” is one unit from the mean. That’s a little less than one standard deviation. We have to go down 1.41 units from the mean before we reach one standard deviation. So, when we get to “1” on the scale, we are two units from the mean and a little more than one standard deviation below the mean.

Let’s transform the simple distribution into a distribution of z-scores by plugging each value into the z-formula:

X	Z-Transformation	Z
1	$z = \frac{1-3}{1.41} =$	-1.42
2	$z = \frac{X-\mu}{\sigma} =$	-.71
3	$z = \frac{X-\mu}{\sigma} =$	0
4	$z = \frac{X-\mu}{\sigma} =$	.71
5	$z = \frac{X-\mu}{\sigma} =$	1.42
$\mu = 3$		$\mu = 0$
$\sigma = 1.41$		$\sigma = 1$

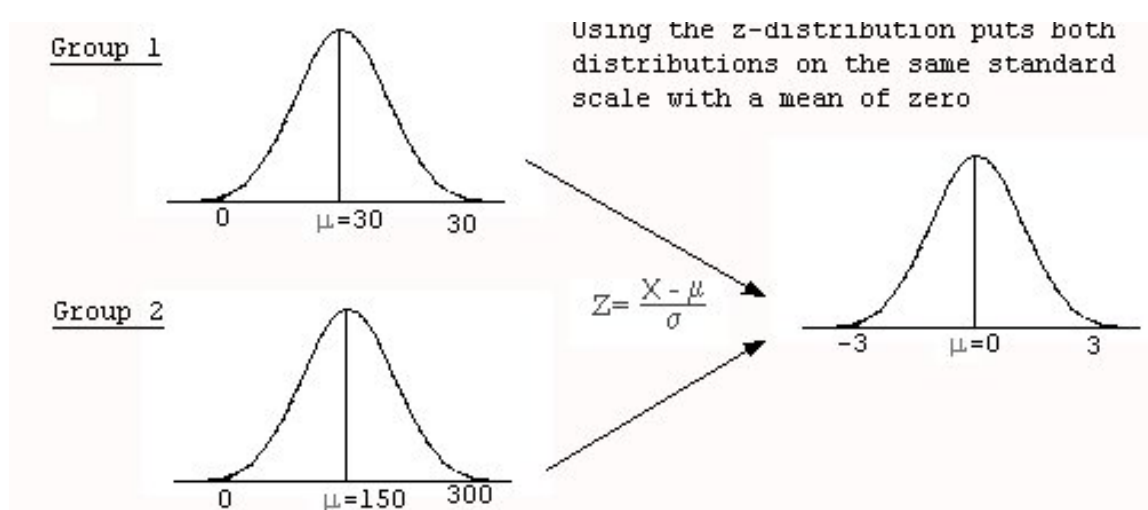
The value of “1” is 1.42 standard deviations below the mean. The value “2” is .71 standard deviations below the mean, and so on. Notice that the mean is at zero, so any scores below the mean in the original distribution will always have a negative z-score and any score above the mean will have a positive z-score.

Properties of the z-distribution. Also notice in the above example that we had to compute the mean and standard deviation of the simple x-distribution in order to compute the z-score. We can compute the mean and standard deviation of the resulting z-distribution as well. The mean of the z-distribution will always be zero, and the standard deviation will always be one. These facts make sense because the mean is always zero standard deviations away from the mean, and units on the z-distribution are themselves standard deviations.

### Using Z-scores to Compare Values

Since z-scores reflect how far a score is from the mean they are a good way to standardize scores. We can take any distribution and express all the values as z-scores (distances from the mean). So, no matter scale we originally used to measure the variable, it will be expressed in a standard form. This standard form can be used to convert different scales to the same scale so that direct comparison of values from the two different distributions can be directly compared.

For example, say I measure stress in the average college sophomore on a scale between 0 and 30 and find the mean is 15. Another researcher measures stress with the same population, but uses a scale that ranges from 0 to 300 with a mean of 150. Who has more stress, a person with a stress score of 15 from the first distribution or a person with a stress score of 150 from the second group? The value of 150 is a much larger number than 15, but both are at the mean of their own distribution. Thus, they will both have the same z-score of zero. Both values indicate an average amount of stress.



Consider another example. Let’s say that Joan got an  $x = 88$  in a class that had a mean score of 72 with a standard deviation of 10 ( $\mu = 72, \sigma = 10$ ).

In a different class let's say Bob got a  $x = 92$ . The mean for Bob's class, however, was 87 with a standard deviation of 5 ( $\mu = 87, \sigma = 5$ ).

Who had the higher grade relative to their class?

If you think about it for a second you will know that Joan's score of 88 is much higher relative to the average of 72 compared to Bob's score of 92 to the average of 87. We can easily compare the values, however, if we simply compute the z-score for each.

Joan

$$Z = \frac{88 - 72}{10} = \frac{16}{10} = 1.6$$

Joan has the higher score because she is 1.6 standard deviations above the mean, and Bob's score is only 1 standard deviation above the mean.

Bob

$$Z = \frac{92 - 87}{5} = \frac{5}{5} = 1$$