

## CHAPTER FOUR

### The Normal Distribution

#### Introduction

Stating an individual score or value of a data distribution along without stating the central tendency (e.g. the mean) and/or variability (e.g. the standard deviation) often does not tell us anything about the significance or magnitude of that score. For example, if Tony's intelligence quotient (*IQ*) score is 70 without know the distribution statistics for *IQ*, we cannot tell whether Tony's is a smarter than most or not. This score is more meaningful if we know that from a sample of 100,000 individuals, the mean *IQ* was 100 and the standard deviation (*SD*) was 15. Knowing this additional information and comparing Tony's score to these tell us that Tony's *IQ* is below average and specifically that Tony's score is exactly two standard deviations below average ( $100 - 2(15) = 70$ ).

The **standard score** (or *z-score*) is simply a way of telling how far a score is from the mean in *standard deviation units*. Knowing the z-score or standard score for a particular data point, *X* not only tells us how far that data is from the mean but also what percent of the distribution or data set is below or above that point, *X*.

The following example illustrates why it is useful to transform individual scores into some standardized scores that take in consideration the central tendency and variability of the distribution. Let us suppose that Tony scored 75 ( $X = 75$ ) on an algebra

exam. How well Tony scored relative the rest of the class depends on the data distribution of all the scores in the class. Let us assumed that the class average was 70 ( $M = 70$ ). If in one case the class's standard deviation was 2 ( $SD = 2$ ) and in another case the class's standard deviation was 10 ( $SD = 10$ ), then the position of Tony's score relative to the rest of the class is different for these two distribution, even though they have the same mean. Figure 4.1.1a shows that when the standard deviation is 2, a lot more students scored less than Tony on the algebra exam and Tony's score is one of the highest in the class; however, as Figure 4.1.1b shows, when the  $SD = 10$ , fewer students scored less than Tony when  $SD = 2$ .

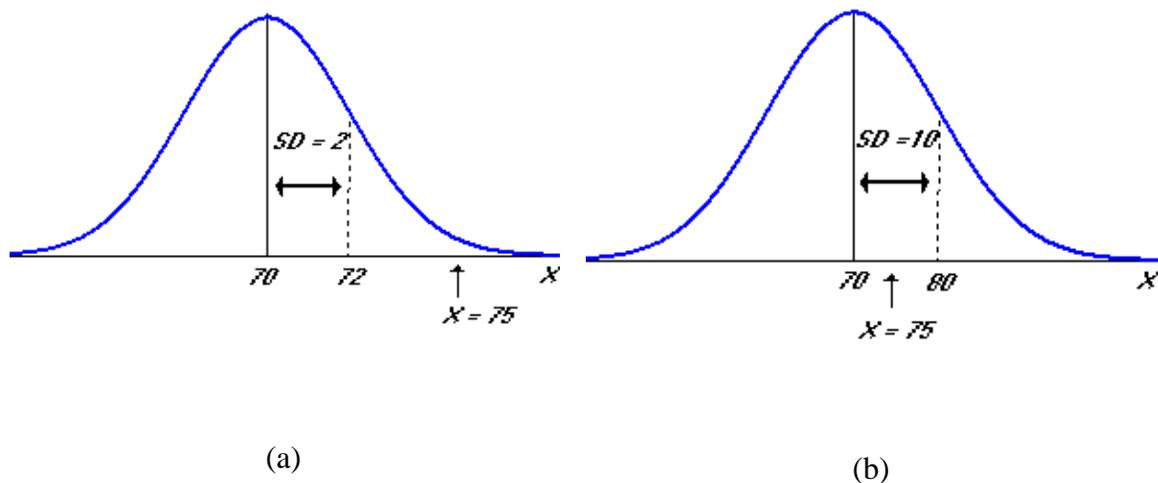


Figure 4.1.1 Illustration of Two Distributions with Different Standard Deviations.

The graph or curve of a normal distribution is symmetrical or looks like a bell-shaped curve. The area under the normal distribution curve and the standard scores of individual scores of a sample are related and will be discussed in the next section on the normal curve. We use the standard normal table to determine the percentiles of scores.

## 4.1 The Standard Score

It is a bit awkward when discussing a score,  $X$ , to say that it is “2 standard deviations above the mean” or “1.5 standard deviations below the mean.” The  $z$ -score was developed to state this fact; this, in effect, says the same thing but in more meaningful ways. Figure 4.1.2 shows a sample distribution with various positions of the  $z$ -scores values. A symmetric distribution or normal distribution with a mean of zero ( $M = 0$ ) and standard deviation of one ( $SD = 1$ ) is called a **standard normal distribution**.

The  **$z$ -score** ( $z$ ) is a measure of the number of standard deviation a score is from the mean. It is the distance a given score is from the mean divided by the standard deviation.

A **standard normal distribution** is a normal distribution with a mean of zero and a standard deviation of one.

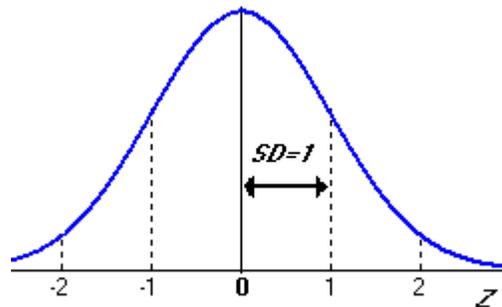


Figure 4.1.2 The Standard Normal Distribution: Distribution of  $z$ -scores

Note that, when the  $z$ -score is positive, it is located above the mean, and when negative it is below the mean. A  $z$ -score of zero (0) tells us that 50% of the data is above or below the mean. Observed that the  $z$ -score consists of two parts: a sign (+ or -) and a

magnitude. Both parts are needed to describe where an individual score is located within a distribution. The formula for converting any score,  $X$  into its corresponding  $z$ -score is:

$$z = \frac{X - M}{S}$$

where  $z$  is the  $z$ -score

$X$  is the observed score or data point being examined

$M$  is the mean of the distribution of scores

$S$  is the standard deviation

Example 4.1.1 From the ODE data table, we know that for the *pass9th* variable that  $M=65.86$  and  $S=13.61$ . If we would like to know how the Arlington schools did relative to the rest of the schools for the *pass9th* variable in ODE table we would compute Arlington's  $z$ -score or standard score.

Arlington's observed score,  $X$  was 84, so its  $z$ -score is:

$$z = \frac{X - M}{S} = \frac{84 - 65.86}{13.61} = 1.33$$

The positive 1.33 tells us that Arlington school scored above the mean of all the schools or their mean score is **1.33** standard deviations above the

If we would like to know how the Lima schools did relative to the rest of the schools for the *pass9th* variable in the ODE table we would compute Lima's  $z$ -score or standard score. Lima's observed score,  $X$  was 40, so its  $z$ -score is:

$$z = \frac{X - M}{S} = \frac{40 - 65.86}{13.61} = -1.90$$

The negative 1.90 tells us that Lima schools scored below the mean of all the schools or their mean score is **-1.90** standard deviations below the mean.

## The Standardized Distribution

If we convert every score,  $X$  of a distribution into its corresponding  $z$ -score, we get a distribution that is very useful. This new distribution is called a *standardized distribution* and when this transformed distribution is symmetrically distributed we call it the *standard normal distribution*. Figure 4.1.3 shows two distributions: (a) is the original *IQ* distribution scores with mean of 100 and standard deviation of 15, and (b) is the transformed  $z$ -score distribution with mean of zero and standard deviation of one.

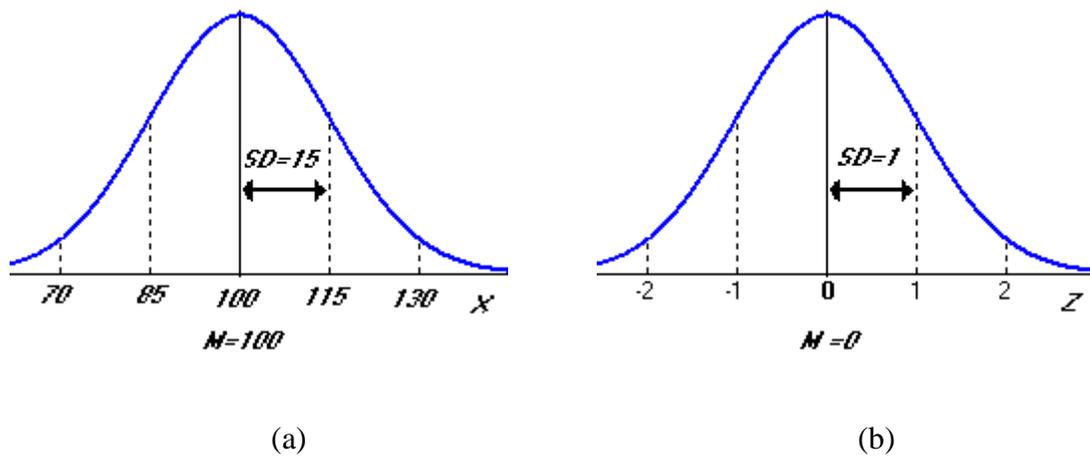


Figure 4.1.3 IQ and Its Transformed  $z$ -score Distribution

This transformed distribution of  $z$ -scores will have the following properties:

1. The shape of the  $z$ -score distribution will be the same as the original distribution. If the original distribution is skewed, the  $z$ -score distribution will also be skewed.
2. The mean of the  $z$ -score distribution will always be zero. In Figure 4.1.3a, the original *IQ* mean was 100. To transform any  $X$  score to its

corresponding  $z$ -score we first subtract it from the mean ( $X - 100$ ) and then divide by the standard deviation  $(X - 100)/15$ , so for  $X = 100$ , we get a  $z$ -score of 0.

3. The standard deviation of the  $z$ -score will always be one. If the mean  $IQ$  is 100 and its standard deviation is 15, then an  $IQ$  of 115 is transformed into  $z = + 1.00$ . Similarly, an  $IQ$  of 85 is transformed into  $z = - 1.00$ .

### SPSS $z$ -score Procedure

SPSS has a simple procedure for computing the  $z$ -scores for every score of a variable. To compute the  $z$ -scores for any variable of a dataset, go to the **Analyze** menu, then select the **Descriptives** option and select variable to be transformed to  $z$ -score values and select the “*Save standardized values as variables*” option. SPSS then generates a variable with the standard scores for each value. Figure 4.1.4 shows the SPSS procedure menu options for transforming the scores of any variable to their corresponding  $z$ -scores.

Table 4.1.1 below is an example of doing this with a table of values:

Table 4.1.1 *Tabulated  $z$ -score Example*

$X$	$z\text{-score} = (X - M)/SD$
12	-0.42696
13	-0.07763
14	0.271704
15	0.621037
10	-1.12563
12	-0.42696
16	0.970371
18	1.669038
9	-1.47496
$M = 13.22, SD = 2.86$	

CT\_Data.sav - SPSS Data Editor

File Edit View Data Transform Analyze Graphs Utilities Add-ons Window

1 : ZVAR00001 -1.48197728879088

	VAR00000 1	ZVAR00001	var	v
1	82.00	-1.48198		
2	83.00	-1.08144		
3	84.00	-.68091		
4	85.00	-.28037		
5	85.00	-.28037		
6	86.00	.12016		
7	86.00	.12016		
8	87.00	.52069		
9	89.00	1.32176		
10	90.00	1.72230		

Descriptives

Variable(s):  
# VAR00001

# Zscore(VAR00001) [ZV]

Save standardized values as variables

OK  
Paste  
Reset  
Cancel  
Help  
Options...

Figure 4.1.4 SPSS procedure: Standard scores

## Stats4U z-score Procedure

Stats4U has a simple procedure for computing the z-scores for every score of a variable. To compute the z-scores for any variable of a dataset, go to the **Analyze** menu, then select the **Descriptive Statistics** option and select variable to be transformed to z-score values and select the “z-score to *Grid*” option. Stats4U then generates a variable with the standard scores for each value. Figure 4.1.5 shows the Stats4U procedure menu options for transforming the scores of any variable to their corresponding z-scores.

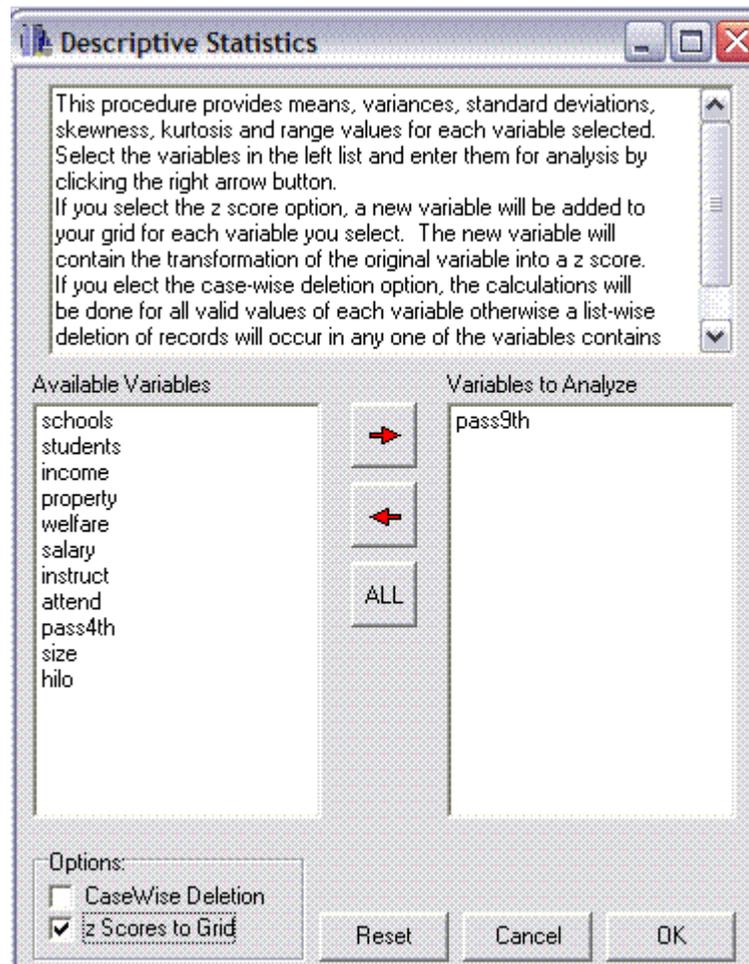


Figure 4.1.5 Stats4U Procedure: Standard Scores

Figure 4.1.6 shows the Stata4U outputs for the z-score transformation of the pass9th variable of the ODE dataset. A z-score computed example is shown for score 59.

DISTRIBUTION PARAMETER ESTIMATES (Output 1)

```
pass9th (N = 94)  Sum =      6191.000
Mean =      65.862  Variance =    185.217  Std.Dev. =    13.609
Std.Error of Mean =      1.404
Range =      72.000  Minimum =    28.000  Maximum =    100.000
Skewness =     -0.258  Std. Error of Skew =      0.249
Kurtosis =      0.447  Std. Error Kurtosis =      0.493
```

The z-scores for 1<sup>st</sup> 17 data points

Example: for pass9th = 59,  $z = -0.50$  [(59 - 65.862)/13.609]

pass9th	pass9thz (z-scores)
85	1.40625
73	0.52451
68	0.15712
65	-0.06332
62	-0.28375
40	-1.90027
72	0.45103
68	0.15712
63	-0.21027
<b>59</b>	<b>-0.50419</b>
56	-0.72462
77	0.81842
74	0.59799
74	0.59799
66	0.01016
37	-2.12071
100	2.50843

Figure 4.1.6 Stata4U Standard Scores Outputs

## Other Standard Scores

Using the z-score values you can calculate other standard scores using the formula: Standard score =  $sz + \mu$ . Table 4.1.2 shows the corresponding  $s$  and  $\mu$  values of each standard score system.

Table 4.1.2 *Standard Score Systems*

Standard Score System	S	$\mu$
z-score	0	1
T score	10	50
IQ	15	100
GATB: General Aptitude Test Battery	100	20
CEEB: College Entrance Examination Board	500	100

To compute the standard score for any of the above system, first compute the z-score by the formula

$$z = \frac{X - M}{S}$$

then use the computed z-score value above and the corresponding mean and standard deviations for that system in the formula =  $sz + \mu$ .

Examples: The T score from a z-score of -1.48 = 35.2 since  $10(-1.48) + 50 = 35.2$ .

The IQ standard score would be 77.8, since  $IQ = sz + \mu = 15(-1.48) + 100 = 77.8$ .