International Journal of Innovative Research and Advanced Studies (IJIRAS)
Volume 4 Issue 7, July 2017

ISSN: 2394-4404

Study On Unaffected Distribution Of The Linear Transformations

Dr. Sajjad Zahir

Associate Professor, Department of Statistics,
D.N.College, Masaurhi, Patna

Abstract: ***This paper reveals that a linear transformation is a transformation of the form X' = a + bX. If a measurement system approximated an interval scale before the linear transformation, it will approximate it to the same degree after the linear transformation. Other properties of the distribution are similarly unaffected. For example, if a distribution was positively skewed before the transformation, it will be positively skewed after.***

Keywords: linear transformation, positively skewed, multiplicative component, etc.

I. INTRODUCTION

The symbols in the transformation equation, $X'_i = a + bX_i$, have the following meaning. The raw score is denoted by X_i, the score after the transformation is denoted by X'_i, read X prime or X transformed. The "*b*" is the *multiplicative component* of the linear transformation, sometimes called the *slope*, and the "*a*" is the *additive component*, sometimes referred to as the *intercept*. The "a" and "b" of the transformation are set to real values to specify a transformation.

The transformation is performed by first multiplying every score value by the multiplicative component "b" and then adding the additive component "a" to it. For example, the following set of data is linearly transformed with the transformation $X'_i = 20 + 3*X_i$, where a = 20 and b = 3.

Linear Transformation - a=20, b=3

X	X' = a + bX
12	56
15	65
15	65
20	80
22	86

The score value of 12, for example, is transformed first by multiplication by 3 to get 36 and then this product is added to 20 to get the result of 56.

II. PURPOSE OF THE STUDY

The main aim of this paper is to analyse that the effect of the linear transformation on the mean and standard deviation of the scores is of considerable interest. For that reason, both, the additive and multiplicative components, of the transformation will be examined separately for their relative effects.

III. ANALYSIS OF THIS STUDY

If the multiplicative component is set equal to one, the linear transformation becomes X' = a + X, so that the effect of the additive component may be examined. With this transformation, a constant is added to every score. An example additive transformation is shown below:

Linear Transformation - a=20, b=1

	X		X' = a + bX
	12		32
	15		35
	15		35
	20		40
	22		42
$\bar{X}$=	16.8	$\bar{X}'$=	36.8
s_X=	4.09	$s_{X'}$=	4.09

The transformed mean, $\bar{X}'$, is equal to the original mean, $\bar{X}$, plus the transformation constant, in this case a=20. The standard deviation does not change. It is as if the distribution was lifted up and placed back down to the right or left, depending upon whether the additive component was positive or negative. The effect of the additive component is graphically presented below.

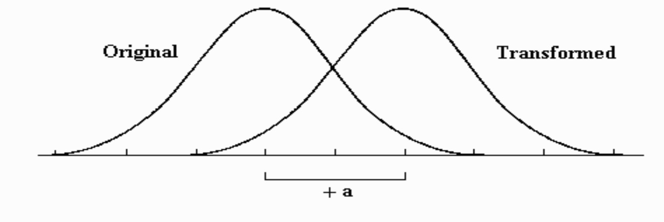

Figure 1

The effect of the multiplicative component "b" may be examined separately if the additive component is set equal to zero. The transformation equation becomes X' = bX, which is the type of transformation done when the scale is changed, for example from feet to inches. In that case, the value of b would be equal to 12 because there are 12 inches to the foot. Similarly, transformations to and from the metric system, i.e. pounds to kilograms, and back again are multiplicative transformations.

An example multiplicative transformation is presented below, where b=3:

Linear Transformation - a=0, b=3

	X		X' = a + bX
	12		36
	15		45
	15		45
	20		60
	22		66
$\bar{X}$=	16.8	$\bar{X}'$=	50.4
s_X=	4.09	$s_{X'}$=	12.26

Note that both the mean and the standard deviation of the transformed scores are three times their original value, which is precisely the amount of the multiplicative component. The multiplicative component, then, effects both the mean and standard deviation by its size, as illustrated below:

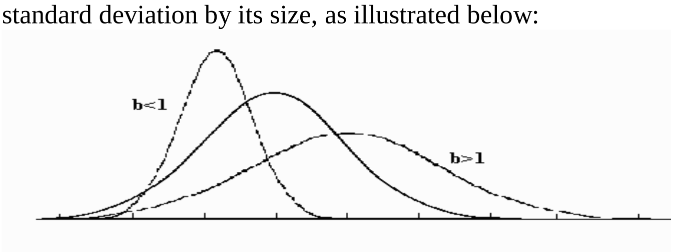

Figure 2

STANDARD SCORES OR Z-SCORES

Another possible transformation is so important and widely used that it deserves an entire section to itself. It is the standard score or z-score transformation. The standard score transformation is a linear transformation such that the transformed mean and standard deviation are 0 and 1 respectively. The selection of these values was somewhat arbitrary, but not without some reason.

Transformation to z-scores could be accomplished using the procedure described in the earlier section to convert any distribution to a distribution with a given mean and standard deviation, in this case 0 and 1. This is demonstrated below with the example data.

$$b = \frac{s_{X'}}{s_X} = \frac{1}{4.09} = 0.244$$

$$a = \bar{X}' - b\,\bar{X} = 0 - (0.244 * 16.8) = -4.11$$

Linear Transformation - a=-4.11, b=0.244

	X		X' = a + bX
	12		-1.18
	15		-0.45
	15		-0.45
	20		0.77
	22		1.26
$\bar{X}$=	16.8	$\bar{X}'$=	-0.01
s_X=	4.09	$s_{X'}$=	0.997

Note that the transformed mean and standard deviation are within rounding error of the desired figures.

Using a little algebra, the computational formulas to convert raw scores to z-scores may be simplified. When converting to standard scores ($\bar{X}'$=0 and $s_{X'}$=1.0), the value of a can be found by the following:

$$b = \frac{s_{X'}}{s_X} = \frac{1}{s_X}$$

$$a = \bar{X}' - b\,\bar{X} = 0 - \frac{1}{s_X} * \bar{X} = \frac{-\bar{X}}{s_X}$$

The value for b can then be found by substituting these values into the linear transformation equation:

$$X' = a + b\,X$$

$$X' = \left(-\frac{\bar{X}}{s_X}\right) + \left(\frac{1}{s_X} * X\right)$$

$$X' = \frac{X}{s_X} - \frac{\bar{X}}{s_X}$$

$$X' = \frac{X - \bar{X}}{s_X}$$

The last result is a computationally simpler version of the standard score transformation. All this algebra was done to demonstrate that the standard score or z-score transformation was indeed a type of linear transformation. If a student is unable to follow the mathematics underlying the algebraic transformation, he or she will just have to "Believe!" In any case, the formula for converting to z-scores is:

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

ISSN: 2394-4404

(Note that the "z" has replaced the "X'")

Application of this computational formula to the example data yields:

Z-score Transformation

	X		Z
	12		-1.17
	15		-0.44
	15		-0.44
	20		0.78
	22		1.27
$\bar{X}$ =	16.8	$\bar{X}'$ =	0.0
s_X=	4.09	$s_{X'}$=	.997

Note that the two procedures produce almost identical results, except that the computational formula is slightly more accurate. Because of the increased accuracy and ease of computation, it is the method of choice in this case.

IV. CONCLUSION AND SUMMARY

This paper concludes that transformations are performed to interpret and compare raw scores. Of the two types of transformations described in this text, percentile ranks are preferred to interpret scores to the lay public, because they are more easily understood. Because of the unfortunate property of destroying the interval property of the scale, the statistician uses percentile rank transformations with reluctance. Linear transformations are preferred because the interval property of the measurement system is not disturbed.

Using a linear transformation, a distribution with a given mean and standard deviation may be transformed into another distribution with a different mean and standard deviation. Several standards for the mean and standard deviation were discussed, but standard scores or z-scores are generally the preferred transformation. The z-score transformation is a linear transformation with a transformed mean of 0 and standard deviation of 1.0. Computational procedures were provided for this transformation.

REFERENCES

[1] Anselin, Luc (1988). Spatial Econometrics: Methods and Models. Dorddrecht: Kluwer Academic Publishers.

[2] Barry, Ronald, and R. Kelley Pace. (1999). "A Monte Carlo Estimator of the Log Determinant of Large Sparse Matrices." Linear Algebra and its Applications, 289, 41–54.

[3] Bavaud, Francois (1998). "Models for Spatial Weights: A Systematic Look." Geographical Analysis, 50, 155–71.

[4] Besag, Julian (1975). "Statistical Analysis of Non-lattice Data," The Statistician, 24, 179–95.