

1. Introduction

While many variables have distributions that are approximately normal, there are just as many situations when the distribution is not well represented by the normal curve. Recall that the median is an alternative way to describe CENTER and so there are alternative ways to describe SPREAD.

2. Minimum, Maximum, Range, Percentiles and the IQR (Chapter 5.4)

A. The lowest and highest values of a variable are the minimum and maximum.

B. The range is maximum-minimum.

C. Definition: a number y is the n th PERCENTILE for the data if $n\%$ of the data are less than or equal to y .

D. The QUARTILES are the 25th, 50th, and 75th percentiles and are denoted Q_1 , Q_2 , and Q_3 . (Another name for Q_2 is the median.)

E. The INTER-QUARTILE RANGE, or IQR, is defined as $IQR = Q_3 - Q_1$; the IQR measures the spread of the middle 50% of the data. It is used in some cases as a replacement for the SD.

THOSE ARE NICE ROBUST MEASURES (e.g. relatively resistant to extreme observations) and they are good for getting an idea of what the distribution looks like (e.g. center, spread) especially when the normal approximation does not accurately represent a variable (e.g. income, education, housing prices).

3. Percentiles and the Normal Curve (Chapter 5.5)

Standard (Z) scores vs. percentiles: If all one does with standard scores is convert them to percentiles, then why have both?

Percentiles and standard scores have slightly different information in them. In other words, the move from standard scores to their normal curve percentile equivalents is not a smooth one-to-one transformation. Percentiles shows how a score relates to other scores in a distribution, it is the score value below which the proportion of the cases in the distribution falls. A standard score (Z) can be used to answer a few more questions:

- *What percentage of people fall below a given score?*
- *What is the relative standing of a score in one distribution versus another distribution?*
- *What score or scores can be used to define an extreme or deviant situation?*

Very large differences between standard scores far away from the mean (in either direction) correspond to small differences in percentiles; likewise, very small differences in standard scores near the mean correspond to large differences in percentiles.

4. Changing Scales (Chapter 5.6)

Definition: changing the original values in a list of numbers by adding, subtracting, multiplying or dividing each value by a constant. All a Z score really does is change the scale of an existing set of values by subtracting the mean from a value and then dividing by the standard deviation. What effect does this have? It affects both "center" and "spread".

*The mean will be changed (added to, subtracted from, multiplied or divided by the constants we used)
The standard deviation will be changed (multiplied by or divided by the constants we used)*

Essentially, this is how the standard normal curve in Table A-105 works. Every set of numbers converted to Standard Units will have a mean of 0 and a SD=1.

5. Last words on Using the Normal Curve

A. **IF** the data are normally distributed, then raw scores can be converted into standard units to find percentages; also, percentages can be converted into standard units and then converted into raw scores.

B. If the data are NOT normally distributed, then using the normal curve will give the wrong answer! DO NOT ASSUME THAT DATA ARE NORMAL.

