

The normal distribution and probability

[Science](#), [Statistics](#)



The variable of interest that was collected from 5, 000 participants was the “ number of words that a participant can type per minute.” The participants were each made to use a typing software which calculated the rate at which each participant could type per minute after each one was made to type a particular amount of text.

This variable may be considered a ratio level variable because of its attributes. First, differences between different levels of ratio level variables have the same meaning (Rubin, 2009, p. 24). That is, the difference in magnitude between 45 and 50 words per minute is the same as the difference in magnitude between 75 and 80 words per minute. The variable of interest, as with any other ratio variable, has a true zero point. That is, it is possible for a participant to not type any word in one minute but may not be able to type any lower than that value. This true zero point is what separates interval level variables from ratio level variables (Agresti & Finlay, 2009, p. 107).

Ratio level of measurement is considered the highest level of measurement because many descriptive statistics computations and inferential statistics techniques may be performed on such measures (Cohen, 1988, p. 53).

Specifically, it is possible to find the mode (though not applicable on this particular example), median, mean, range, maximum, minimum, and standard deviation for ratio level measures.

This variable is normally distributed because the values that it takes may be scattered on a plot, approximating a symmetric bell-shaped curve (Kazmier, 2004). That is, the graph is more concentrated on the central values and symmetrically decreases on both side of the bell-shaped graph. This is

possible because more people can type the “ average” range of values for words per minute and as these values go to the extremes (whether extreme high or extreme low), fewer people fall into those values. In terms of central tendencies, it is to be expected that the mean, median, and modal value for number of words typed per minute would be most likely the same value (Agresti & Finlay, 2009). Furthermore, since the sample size is quite large, normal distribution of data points almost always follows (Rubin, 2009).

In a normal distribution, the probability value of a particular score is the likelihood of that score occurring in the sample data (which may be carried over to the population). In this particular example, it is the percentage of participants that typed a particular number of words per minute. Moreover, a probability value of . 05 in this example would stand for the likelihood of the scores being in the extreme 5% of the area of the graph under the curve. It may also be interpreted as the scores falling at the extreme values in the normal plot. On a more technical definition, a probability value of . 05 would be taken by those values that are outside two standard deviations from the mean (Kazmier, 2004).

References

- Agresti, A., & Finlay, B. (2009). *Statistical methods for the social sciences* (4th ed.). New Jersey: Prentice Hall.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Routledge.
- Kazmier, L. (2004). *Schaums Outlines: Theory and Problems on Business Statistics* (4th ed.). USA: McGraw-Hill.

Rubin, A. (2009). *Statistics for evidence-based practice and evaluation*. Hoboken, NJ: Cengage Learning.