

Statistics in Education: A mission statement¹

Vaclav Brezina, Lancaster University

Numbers are an essential part of our everyday life. We count the coins in our pocket, the minutes before the next bus arrives or the sunny days in a rainy year. Numbers and quantitative information are also very important for students and educators. Statistical literacy – the ability to produce and interpret quantitative information – belongs to the basic set of academic skills that, despite its importance, has so far received only limited attention.

Many students (and academics) are afraid of statistics – think about what your first reaction is to the equation in Figure 1 below.

$$SD = \sqrt{\frac{\sum_{i=1}^N (X - \bar{X})^2}{N-1}}$$

Figure 1: The equation of standard deviation (mathematical form)

This is because statistics is often misconstrued as the art of solving extremely complicated equations or a mysterious magic with numbers. Statistics, however, is first and foremost about *understanding* and making sense of numbers and quantitative information. For this, we need to learn the basic principles of collecting, organising and interpreting quantitative information. Critical thinking is thus much more important for statistics than the number crunching ability. After all, computers are very good at processing numbers and solving equations and we can happily leave this task to them. For example, many even complex statistical tasks can be achieved by using tools such as the [Lancaster Stats Tool online](#), where the researcher can merely copy-paste their data (in an appropriate format) and press one button to receive the answer.

Humans, on the other hand, outperform computers in the interpretation skills. This is because we have the knowledge of the context in which numbers appear and we can therefore evaluate the relative importance of different quantitative results. We as teachers, linguists, sociologists, scientists etc. can provide the underlying meaning to numbers and equations and relate them to our experience and the knowledge of the field. For example, the equation in Figure 1 can be simplified as follows:

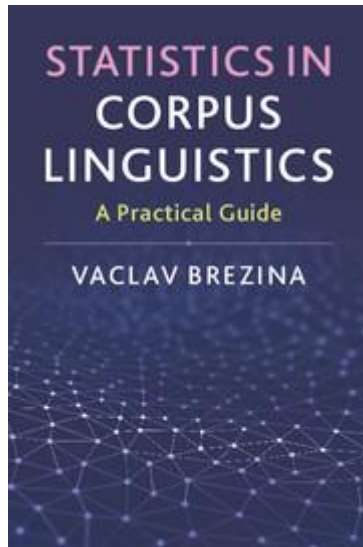
$$SD = \sqrt{\frac{\text{Sum of squared distances from the mean}}{\text{Number of cases}-1}}$$

= amount of variation from the central tendency in a group

Figure 2: The equation of standard deviation (conceptual)

¹ The work on the statistics project was supported by ESRC grants no. EP/P001559/1 and ES/K002155/1. This blog post was originally written as a follow-up on a seminar given at the University of Sterling in 2016; thanks to Vander Viana for organising the event.

When we relate this to what we know about the world, we can see that the question we are asking in Figure 2 is how much variation there is in our data, a question about variability, difference in tendencies and preferences and overall diversity. This is something that we can relate to in our everyday experience: *Will I ever find a twenty-pound note in my pocket? Is the wait for the bus longer in the evening? Is the number of sunny days different every year?* When talking about statistics in education, I consider the following point crucial: as with any subject matter, we need to strive to find ways of highlighting the link between statistics and our daily experience.



Brezina, V. (2018). [*Statistics in Corpus Linguistics: A Practical Guide*](#). Cambridge: Cambridge University Press.

Do you use language corpora in your research or study, but find that you struggle with statistics? This practical introduction will equip you to understand the key principles of statistical thinking and apply these concepts to your own research, without the need for prior statistical knowledge. The book gives step-by-step guidance through the process of statistical analysis and provides multiple examples of how statistical techniques can be used to analyse and visualise linguistic data. It also includes a useful selection of discussion questions and exercises which you can use to check your understanding.

The book comes with a Companion website, which provides additional materials (answers to exercises, datasets, advanced materials, teaching slides etc.) and [Lancaster Stats Tools online](#), a free click-and-analyse statistical tool for easy calculation of the statistical measures discussed in the book.