

Percentage multi-level modeling

Michail Belias

25 April 2018

Although normally distributed data are of a high use and importance, sometimes the outcomes do not follow a Normal distribution. A variety of examples such as counts, proportions, categorical and many more are typically used. In this tutorial we will focus on percents. Percents are continuous but $p_i \in [0,1]$, therefore their range is bounded and limited. A result that exceeds the 100% or drops lower than 0% is invalid. Furthermore, due to that reason the analysis conducted is prone to loss of power, misfitting and can not be considered valid. Because the outcome is continuous the right analysis that comes in mind is the linear regression, and if the percentages are near 0.5 is adequately correct, given large samples.

Transforming Percents

An easy solution for modelling percentages is to transform the data. A potential draw-back is that interpretation may become difficult if not impossible. The most common method to handle percentages is to transform them using the **arcsine** function