When analysing joint samples from several random variables, such as a questionnaire with multiple questions, a core challenge is to figure out the dependencies among these variables. Such samples are typically encoded in a multi-dimensional table, where each dimension of the table represents the different outcomes of one random variable. To test certain dependencies, we generate more tables with the same properties as our sample table, and perform a statistical test to see whether our sample table stands out or not. To generate these tables, we randomly transform the sample table using certain steps that maintain the desired properties.

We have analyzed the steps that are necessary when generating these new tables from the sample table. When the number of variables, and therefore the number of dimensions of the table, grows, the number of necessary steps also grows. However, what we have shown is that these steps do not become excessively complex. Instead, we have established that the steps necessary for a larger number of variables can be constructed from those that are necessary for a smaller number of variables. This shows that the complexity of the necessary steps stabilizes.

Random generation algorithms for multi-dimensional tables

Multi-Dimensional Data Model

Questionnaire Data

Measures
Number of times certain combination of answers was given