MŰEGYETEM 1782

Budapest University of
Technology and Economics

Applied Mathematics Master of Sciences Thesis

The use of Fleishman distribution in the empirical investigation of statistical tests

Author:
Tamás Ferenci

External supervisor:
Prof. Jenő Reiczigel DSc
Internal supervisor:
Dr. Márton Balázs PhD

2012

Abstract

Statistical tests are among the most important tools of modern statistics. Given a population which was only observed partly, we can not decide whether a statement H_0 is true if it pertains to the whole population. However, statistical test can give a probabilistic answer on whether H_0 stands or not, based only on the observed part of the population (called sample). Statistical tests typically impose more or less presumptions about the population; they are guaranteed to produce valid (probabilistic) answers only if these presumptions are met. Practically, the two most important property of a statistical test is robustness and power. Robustness measures how valid the test remains if its presumptions are not met, while power measures how well the test can detect that H_0 is not true, if it in fact does not stand. Practically, the most problematic task is to investigate robustness for tests which have a distribution presumption on the population (called parametric tests), and to investigate power for tests that do not have such assumption (called non-parametric tests). The common point in these (which presents their difficulty) is that both require to investigate the operation of a test on a population with non-specified distribution. (In robustness testing for the parametric case, there is a specified distribution in the presumptions, but we want to investigate the effect of deviating from it; in power testing at the non-parametric case, we stick to the presumptions of the test, but they do not specify a distribution.) For practical reasons, these non-specified distributions are usually characterized by their first four moments, so the first problem is to find a distribution family that can be parametrized to have arbitrary skewness and kurtosis. (Setting arbitrary mean and variance is trivial by linear scaling.) Common distributions do not have this property; in this thesis, we introduce – in detail – a well-known distribution family, the Fleishman distribution, that does have. In most of the cases, the analytical investigation with Fleishman distribution (or any other similar distribution) is infeasible or impossible. Thus, robustness and power are typically investigated empirically, using Monte Carlo simulation: many sample from the "non-specified" distribution (i.e. a distribution with pre-specified skewness and kurtosis) are generated, the test is applied to them, and results are recorded to determine robustness or power. With enough replications, this sufficiently converges to the true value of the investigated parameter. The method can be than re-iterated for different levels of skewness and kurtosis. This, however, has enormous computation requirement, which is infeasible even on modern personal computers. To alleviate this issue, we developed a program according to the novel principle of General Purpose GPU-computing, which harnesses the GPUs of modern video cards, as they can offer a performance that was previously only achievable with supercomputers for certain problems. More specifically: for problems that are computationally intensive (not I/O-intensive) and exhibit high data parallelism – just like Monte Carlo simulation for investigating statistical tests. We implemented a program, a simulational environment that is able to perform more than 60 million hypothesis testings per second (one-sample t-test, $n = 10$) even on a low-end video card. The environment is modular and highly flexible, so that it facilitates further extension. The program is introduced in detail in this thesis. Using this program, we systematically investigated many commonly used statistical test for robustness (parametric tests) and power (non-parametric tests). The results are presented and briefly discussed in this thesis. As a conclusion, we also briefly point out one limitation of the application of Fleishman distribution for this purpose to present a balanced discussion.