



Cramér-von Mises statistic for continuous distributions: A Monte-Carlo study for calculating its associated probability

Lorentz Jäntschi

Department of Physics and Chemistry,
Technical University of Cluj-Napoca, Cluj, Romania
lorentz.jantschi@chem.utcluj.ro

Literature is abundant on the uses of the Cramér-von Mises test in biological and earth sciences (see for instance [1]), for comparing observed distributions of phenomena against theoretical models and for testing if certain data, like gene expression levels and physiological measurements follow a theoretical distribution as prerequisite in applying parametric statistical methods.

The Cramér-von Mises criterion [2, 3] is used for judging the goodness of fit of a cumulative distribution function. Well-known alternatives to Cramér-von Mises include Watson U^2 , Kolmogorov-Smirnov, Kuiper V , and Anderson-Darling. All listed here have something important in common: belongs to the category of order statistics, fundamental tools in non-parametric statistics and inference. A key element related to order statistics of random samples from a continuous distribution, is the fact that the cumulative distribution function reduces the analysis to the case of order statistics of the uniform distribution. In a nutshell, having a very good generator for uniform distributed samples, significantly increases the chances of success for any Monte-Carlo experiment with the order statistics.

A general procedure, described in [4], applicable for any order statistic was particularized for the Cramér-von Mises criterion and a large amount of data was generated. Furthermore, regression analysis was deployed in order to obtain the dependence of the probability as function of the the value of the statistic and of the sample size. The analysis includes a variational analysis, so for the obtained regression coefficients also confidence limits were provided.

Keywords: Monte-Carlo simulation, Cramér-von Mises criterion, goodness-of-fit tests

MSC2020: 62G30, 62P10, 65C05

References

- [1] O. Mazdiyasni, A. AghaKouchak, Substantial increase in concurrent droughts and heat-waves in the United States, *Proceedings of the National Academy of Sciences*, 112:11484–11489, 2015.
- [2] H. Cramér, On the composition of elementary errors, *Scandinavian Actuarial Journal*, First paper: Mathematical deductions, 1928:13–74, Second paper: Statistical applications, 1928:141–180, 1928.
- [3] R. Von Mises, *Statistik und wahrheit*, Julius Springer 20, 1928.
- [4] L. Jäntschi, Detecting Extreme Values with Order Statistics in Samples from Continuous Distributions, *Mathematics*, 8:216, 2020.