Performances of Shannon's Entropy Statistic in Assessment of Distribution of Data

Lorentz JÄNTSCHI[1,2,*] and Sorana D. BOLBOACĂ[3]

[1] *Technical University of Cluj-Napoca, Department of Physics and Chemistry, Muncii Blvd., No. 103-105, 400641 Cluj-Napoca, Romania; lorentz.jantschi@gmail.com*
[2] *Babeş-Bolyai University, Institute for Doctoral Studies, Kogălniceanu Street, No. 1, 400084 Cluj-Napoca, Romania*
[3] *Iuliu Haţieganu University of Medicine and Pharmacy, Department of Medical Informatics and Biostatistics, Louis Pasteur Street, No. 6, 400349 Cluj-Napoca, Romania; sbolboaca@umfcluj.ro*

Abstract. Statistical analysis starts with the assessment of the distribution of experimental data. Different statistics are used to test the null hypothesis (H_0) stated as *Data follow a certain/specified distribution*. In this paper, a new test based on Shannon's entropy (called Shannon's entropy statistic, H1) is introduced as goodness-of-fit test. The performance of the Shannon's entropy statistic was tested on simulated and/or experimental data with uniform and respectively four continuous distributions (as *error function*, *generalized extreme value*, *lognormal*, and *normal*). The experimental data used in the assessment were properties or activities of active chemical compounds. Five known goodness-of-fit tests namely Anderson-Darling, Kolmogorov-Smirnov, Cramér-von Mises, Kuiper V, and Watson U^2 were used to assess the performances of H1.

Keywords: Shannon's entropy; statistic; continuous distribution; tests of goodness-of-fit.

---

* Corresponding author. *E-mail address*: lorentz.jantschi@gmail.com; *Tel.*: +40-264-401-775 (Lorentz Jäntschi)