Assessing Models in Non-Location Scale Family: its application in Biostatistics

Claudia Castro-Kuriss[1], *Victor Leiva*[2]

[1]ccastro@itba.edu.ar, Department of Mathematics, Instituto Tecnológico de Buenos Aires, Argentina.

[2]victorleivasanchez@gmail.com, Facultad de Ingeniería y Ciencias, Universidad Adolfo Ibañez, Chile.

A common problem in statistics consists of describe data by a specific parametric model. Goodness-of-fit (GOF) techniques have been widely studied for doing so, in particular for distributions belonging to the location-scale family. However, one could be interested in making the assessing for distributions that do not belong to this family. In particular, several truncated and skewed distributions arise in biostatistics. For example, because of time or money decisions, some epidemiological studies end before all the patients enrolled with a disease die and the number of the survivors remains unknown. Also environmental problems that arise from unusual events like amounts of precipitations due to storms, concentration of minerals under volcano activity, concentrations of micro-particles in polluted meltwater, measurement of environment contamination or radiation due to leaks of radioactive material can be described with skewed distributions denoting those abnormal situations. We review the available GOF tests and graphical tools based on these tests with uncensored and censored data for distributions in the non-location-scale (NLS) family. Anderson-Darling, Kolmogorov-Smirnov and Michael GOF statistics are considered. We specify the proposed results for NLS life distributions and apply them to real-world data sets from different areas of science to illustrate their potential, with emphasis in some Birnbaum-Saunders and inverse Gaussian distributions. More precisely, one example consists of survival times of mice inoculated with a carcinogen while other deals with airborne measurements of radiation recorded after a reactor accident. We find for these data sets the NLS distributions that best describe them. In addition, we illustrate the fit with probability plots that indicate a very good specification of the postulated hypothetical distribution.

Keywords: GOF tests; Non-Location-Scale Family.

AMS: Primary 65C20, Secondary 62F03