

# Clustering of loglinear models using likelihood ratio tests p-values to find homogeneous regions regarding drought management<sup>1</sup>

Elsa Moreira and João T. Mexia

*New University of Lisbon, Portugal*

## Abstract

The Alentejo in the southern Portugal is a region that often is subject to severe and extreme droughts. The main concern of the present study was to know if Alentejo should or not be considered a region aiming at drought management. So, monthly time series of Standardized Precipitation Index (SPI) with 67 years correspondent to 40 meteorological stations in the region were selected in order to be analyzed. These series were used to compute contingency tables for the transitions between drought classes, considering 4 levels of drought classes. Then, to this contingency tables appropriate loglinear models were fitted in order to obtain estimates of the drought class transitions probabilities. Finished the model fitting phase, the starting point for a second phase, was in fact the existence of 40 objects, described by the correspondent loglinear models, that we wish to classify and group according to some common characteristics. So, the similarity of the objects reduces to the similarity of their loglinear models. Since the models are members of the same family, the models similarity is defined as the closeness of their vectors of model parameters. The technique used to finding similarity between loglinear model parameters was the Likelihood Ratio Tests (LRT) where the similarity measure is the asymptotic p-value associated with the model linking test. If the p-value is large, the null hypotheses of equality between the vectors of parameters is likely to be true. A p-value similarity matrix was then computed in order to detect possible groups of similar objects. Two model linking tests were performed, one considering that the loglinear model parameters are all of interest and another considering that just some of the parameters are of interest. After several attempts of identify clusters that could delimit regions, none was found. Therefore, according to the present study, the Alentejo could be considered as an homogeneous region aiming at drought management.

**Keywords:** Loglinear models, Drought class transitions, Likelihood ratio test, Model linking, p-value.

---

<sup>1</sup>This work was partially supported by Financiamento Base 2010 ISFL-1-297 from FCT/MCTES/PT.

## References

- Agresti, A. (1990). *Categorical Data Analysis*. J. Wiley & Sons: New York.
- Lehmann, E.L. (1997). *Testing Statistical Hypotheses*. Reprint of the 2nd ed. publ. by Wiley 1986. Springer: New York.
- Moreira, E.E., Paulo, A.A., Pereira, L.S., and Mexia, J.T. (2006). Analysis of SPI drought class transitions using loglinear models. *Journal of Hydrology* 331, 349–359.
- Nelder, J.A. (1974). Loglinear models for contingency tables: a generalization of classical least squares. *Appl Stat.* 23, 323–329.
- Zhu, R. and El-Shaarawi, A.H. (2009). Model clustering and its application to water quality monitoring. *Envirometrics* 20, 190–205.