A TEST OF HYPOTHESES FOR FUZZY MATCHED PAIR DATA

Ronei Marcos de Moraes e Sydney Gomes da Silva
Departamento de Estatística / CCEN / UFPB

In several application fields the need exists of testing statistical hypotheses formulated starting from qualitative data. Applications of tests of hypotheses on data obtained by subjective answers, to request a type of statistical approach generally given by the nonparametric statistics. The Signs Test is known and diffused in the statistical way for the comparison of small matched samples, admitting observations that obey the nominal scale at least. However, the qualitative data can be analised with a new mathematical approach done by the fuzzy sets theory [L. A. Zadeh, Fuzzy Sets. *Information and Control*, 8, 1965, 338-353]. The fuzzy sets are a generalization of the concept of the function characteristic of the Theory of the Classic sets, introducing one function more generic denominated pertinence function. In this paper, we suggest a new methodology to resolve this problem, using fuzzy numbers [Dubois, D. and Prade, H.; *Possibility Theory. An Approach to Computerized Processing of Uncertainty.* New York: Plenum Press, 1988]. Paired data obtained of this waycan be treated by a classical Test of Sign [Lehmann, E. L. and D'Abrera, H. J. M; *Nonparametrics: Statistical Methods Based on Ranks.* San Francisco: Holden-Day, 1975]. In this test, if the null hipothesis is true, we would expect about half of the differences to be negative and half to be positive. Let be fuzzy numbers $\widetilde{A}_i$ and $\widetilde{B}_i$, $(i = 1, ..., N)$, and the difference between them given by:

$$\widetilde{d}_i = (\widetilde{A}_i - \widetilde{B}_i)(z) = \sup_{z=x-y} \min \left\{ \widetilde{A}_i(x) - \widetilde{B}_i(y) \right\}, i = 1, ..., N.$$

where N is the size of the sample and the fuzzy values resultants can have negative, positive signs or they be approximately null. To test the

hypothesis of difference among the two populations the is equal to test the median of the difference between the samples is or not significantly different from zero. Thus, everybody the resulting values of those differences they are compared with a fuzzy number C of linguistic meaning "approximately zero", and the values of the tolerances are given by the own researcher in agreement with the convenience of the specific study. It is necessary to obtain a relationship of order of each fuzzy value $\widetilde{d}_i$with that private fuzzy number. The tolerances of the fuzzy number C can be null and then C is the real number zero.

Being known how many of those fuzzy values then is larger, smaller or approximately the same C, it is obtained the proportions of values larger, or smaller than C, denoted, respectively for $p_+$and p_-, given for: $p_+ = \frac{n_+}{N}$ and $p- = \frac{n_-}{N}$; where n_+ is the number of larger fuzzy values than C; n_- is the number of smaller fuzzy values than C; and N is the size of the sample.

As the values $n_+$and n_-they are not fuzzy values and the statistics of the test is built about those proportions, it can be made the same suppositions of the Test of the Sign. In the case of small samples ($N < 25$), the probability associated with the occurrence of a number peculiar of positive or negative values, it can be determined by the reference to a distribution binomial with parameters $p_+ = p- = 0.5$ and with N observations.

Thus, we propose a test of hypotheses where the uncertainty is included in the analysis through the concept of fuzzy number to generalize the Test of the Sign. When the data are given by crisp values, this new methodology return to classical test.

Endereço para correspondência:Ronei Marcos de Moraes, Departamento de Estatística, Centro de CCEN, UFPB, Campus Universitário I, João Pessoa - PB. CEP 58.059-970. e-mail: ronei@de.ufpb.br