Geophysical Research Abstracts
Vol. 13, EGU2011-12672, 2011
EGU General Assembly 2011

Multicategorical contingency tables in the precipitation forecast verification

Zoran Pasaric and Josip Juras
University of Zagreb, Faculty of Science, Geophysical department, Zagreb, (pasaric@irb.hr)

Although contingency tables entered the field of forecast verification more then one hundred years ago, they continue to be an area of intensive research. The reason is the ability of contingency table to condense and clearly display the properties of some set of forecasts and corresponding observations. Utilization of contingency tables is particularly useful for the verification of quantitative precipitation forecasts, which is for many reasons difficult to verify. Precipitation itself is governed by various atmospheric processes acting on different spatial and temporal scales, it is strongly affected by local conditions like orography, and typically it is sparsely observed. Furthermore, observation at particular station may not be representative for the respective model grid-box average. By dividing the whole precipitation range into K categories (where the first category could be 'no rain') the forecasting system performance is conveniently represented by the KxK contingency table.

It appears, however, that for continuous monitoring of the forecasting system performance, information carried by the contingency table has to be further condensed and expressed, if possible, by a single number. This is achieved by introducing various numerical scores that are calculated from the table. Unfortunately, the usage of these scores in practice, which started in 1950s, is not yet standardized. This is especially so for the multicategorical (KxK) tables. The Heidke and Peirce scores, well known for 2x2 tables, are readily generalized to KxK tables, but these sores take into account explicitly only the diagonal elements of the table. A new impulse came in 1990s when Gandin and Murphy introduced a new class of scores based on the notion of equitability. Among them there is the Gerrity score which has found somewhat wider application in practice. Nevertheless, recent introduction of a new score (SEEPS) shows that the problem of KxK tables is still open. It could be said that all scores used so far exhibit some undesirable properties. Some scores show the tendency to depend on the frequencies of particular categories, and indirectly on the number of categories. The dependence on bias is also frequently found, implying the possibility that particular score could be improved by hedging.

The key difficulty lies in the well established fact that forecast verification is a multifaceted problem, and each score measures the various facets of forecasts quality in its own way, while no single score itself can encompass all aspects of forecast quality. For this reason, we propose to use scores that focus on single aspect, independently of the others. From many facets of forecast quality we single out three of them with respect to their importance: association, bias and uncertainty. The last two are easily assessed with the probabilities of observing and probabilities of forecasting the precipitation in particular category. The association is, the authors believe, best assessed by the polychoric correlation coefficient, an old measure proposed by Perason. By quantifying the association, bias and uncertainty it is possible to essentially reconstruct the original KxK table. In this sense, the information carried by the table is partitioned between the three facets, while the dimensionality of the problem is essentially reduced from K*K to 2K. Moreover, polychoric correlation coefficient show weak sensitivity on the frequencies of precipitation or forecast categories. Consequently, it should be resistant to hedging and enable comparison of forecasts in different climates. Real life examples will be analyzed from the proposed point of view, and comparison with other scores will be shown.