Evaluating Response Models Performance

Henry Greene is an adjunct professor at the University of Hartford. His research interests include: direct marketing, geodemographic segmentation, database marketing and customer relationship management. He can be reached at HGREENE@IXGUYS.COM .

George Milne is Associate Professor at the University of Massachusetts Amherst. His research interests include: privacy, interactive marketing and consumption. He can be reached at milne@mktg.umass.edu .

ABSTRACT

Traditionally, statistical models are evaluated in terms of goodness of fit. However, it has been argued that superior goodness of fit does not necessarily guarantee superior performance. In the Direct Marketing industry, a variety of descriptive statistics and terminology have been used to evaluate response performance: decile analyses, gains charts, lift charts, whisker plots and banana curves. Less common is the Gini index, traditionally used in economics and other social sciences, which was originally created to measure the disparity of income and wealth among a population. It has also been used to measure other social phenomena such as disparity in educational attainment among groups of people. In a direct marketing context, Gini has been used to indicate disparity of catalog sales among customers. It recently has been used as a general measure to assess response model performance.

In this paper we show mathematically how the Gini index is related to other popular descriptive statistics; namely, the cumulative lift

curve and the cumulative gains chart. In comparison to descriptive statistics that use a set of points, Gini is a single statistic that can be used to investigate the distribution properties of the estimator. Knowledge of the Gini estimate and its distribution properties provides an opportunity for inferential assessments of the Gini statistic. This is not possible with the other commonly used descriptive statistics (lift charts, gains charts, decile analysis tables, whisker plots, etc).

A large Monte Carlo simulation study was conducted under 1,620 different conditions (response rate, file size and Gini index were varied). For each combination, 200 random samples from a large master data file were randomly chosen, the Gini index and the standard deviation of Gini were computed. Upon examining the results of the study, two general observation were: as file size increases, variance in Gini decreases and 2) as response rate increases, variance in Gini decreases. More importantly, when the sample file size is <15,000 (relatively small for a direct marketer), the variability of Gini is excessively large and when the file size is extremely large (N > 100,00 and response rates > .01) the variability of Gini is expected to be very small.

We restricted our assessment of Gini variation to conditions that direct marketers would find most useful. Our original 1,620 test conditions were reduced to 1,136 points. Of the 1,136 test conditions, we selected a sample of 50 conditions for evaluating whether the Gini statistic was normally distributed. In most of the 50 cases, the normality assumption could not be rejected. The set of 1,136 data points were used in a regression to predict the Gini standard deviation as a function of file size, response rate and Gini estimate. The regression formula and the associated statistics are provided. We also tested the formula with samples from a DMEF data file and demonstrated the accuracy of the formula.

In summary, this paper argues that Direct Marketers will find the Gini statistic more useful for assessing model performance when the statistical properties of Gini are known and used in the analysis. When comparing two different models, it is possible to determine if one model's performance falls outside the expected performance range of another model. An example with data from a direct marketing campaign is provided. Also, by utilizing the Gini index and its statistical properties, it is possible to validate a response model with statistical precision.