Predicting Repeat Purchases When Customers' Heterogeneities Follow Any Distributions

ABSTRACT

Customers of online business are highly heterogenous. They are geographically scattered and have their own shopping frequencies. Also, since everyone's tolerance towards service quality varies, some customers may dropout and switch to other vendors. To predict customers' repeat purchases, modelers need to effectively account for heterogeneities in these behaviors, specifically, in both purchasing rate (number of purchases in a time range) and dropout rate (probability of becoming inactive).

The classical Beta-Geometric/Negative Binomial Distribution (BG/NBD) model employs Gamma distributions to capture the heterogeneity on shopping frequencies and Betta distributions on dropout rate. By doing so, the BG/NBD model becomes mathematically tractable and has achieved great successes in prediction. However, since both Gamma and Betta distributions are simplifications of the realities, when the underlying distributions become complicated, deriving the close form solution is no longer possible. We can not extend the BG/NBD model without losing its tractability.

In this paper, we aim to model the heterogeneities of customers' purchase pattern and dropout rate using two arbitrary distributions. The extension retains the elegance of BG/NBD framework while at the same time extending the model to any diverse customer base. The generalization is made possible through using the Gaussian quadrature. Our empirical and simulation results show that the proposed method outperforms the tradition models in predictive performance.

Keywords: NBD, BGNBD, Gaussian Quadrature, Count Model, Repeat Purchase