

Know your Limits: Minimal Requirements for Pareto/NBD Distributed Data Sets

Lydia Simon

University of Duisburg-Essen

Jost Adler

University of Duisburg-Essen

Cite as:

Simon Lydia, Adler Jost (2019), Know your Limits: Minimal Requirements for Pareto/NBD Distributed Data Sets. *Proceedings of the European Marketing Academy*, 48th, (8372)

Paper presented at the 48th Annual EMAC Conference, Hamburg, May 24-27, 2019.



Know your Limits: Minimal Requirements for Pareto/NBD Distributed Data Sets

Abstract

The Pareto/NBD model is one of the best-known models in customer base analysis. When being applied to empirical data sets, the true distribution and parameters of the data are obviously not known. Here arises a lack in research on how many customers or which duration of calibration are necessary to achieve reliable estimation results. Furthermore, it was never examined how well the underlying parameters can be recovered at all. Therefore, we perform a broad simulation study on Pareto/NBD distributed data sets to find minimal requirements for its usage. The results show that a calibration period between 18 months and five years with at least 5,000 customers is recommended. Furthermore, the heterogeneity parameters are overrated which leads to an overestimation in the expected values of the purchase and dropout rate.

Keywords: *Pareto NBD model; customer base analysis; synthetic data*

Track: Methods, Modelling & Marketing Analytics