Bayesian inference for the Net Promoter Score

Eliardo G. Costa[1], *Rachel T. Q. Ponte*[2]
[1] Universidade Federal do Rio Grande do Norte, Brazil
[2] Ânima Educação, Brazil

Abbreviated abstract: The NPS is a measure used by several companies as indicator of customer loyalty. Studies that address its statistical properties are still scarce and none of them considered a Bayesian approach neither the sample size determination problem. We provide point and interval estimators for the NPS and discuss the sample size determination. Computational tools in the R language were implemented to use this methodology in practice.

Related publications: (up to 2 references)

– F. F. Reichheld, *Harvard Business Review* 81, 46–55 (2003)
– B. Rocks, *The American Statistician* 70(4), 365-372 (2016)

DEst DEPARTAMENTO DE ESTATÍSTICA
UFRN

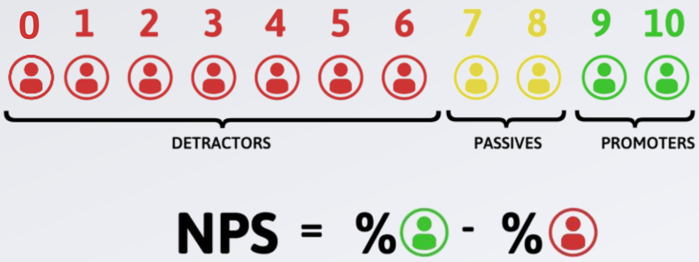
eliardo.costa@ufrn.br

4th Conference on
Statistics and Data Science
Salvador, Brazil (online)
December 1-3, 2022

Introduction

- Reichheld (2003) proposed a statistics called Net Promoter Score (NPS) that may be used by a company as an indicator of customer loyalty;
- Ask the question: “How likely is it that you would recommend [company X] to a friend or colleague?”;
- Reichheld (2003) suggested that the response to the aforementioned questions must be on a 0 to 10 rating scale.

Net Promoter Score



- An estimate of the NPS is computed as the difference between the proportions (or percentages) of “promoters” and “detractors”.
- Rocks (2016) discuss interval estimation techniques for the NPS in a frequentist approach

eliardo.costa@ufrn.br

Methods

- **Model**

- Let $\boldsymbol{\theta} = (\theta_1, \theta_2, \theta_3)$ be the proportions of detractors, passives and promoters in the customer population, respectively. Then, the NPS in the respective population is given by $\Delta = \theta_3 - \theta_1$.
- Let $\boldsymbol{X}_n = (X_1, X_2, X_3)$ be the numbers of customers categorized as detractors, passives and promoters, respectively, in the customer sample of size n.
- Bayesian model: $X_n|\boldsymbol{\theta} \sim \text{Mult}(\boldsymbol{\theta})$; $\boldsymbol{\theta} \sim \text{Dir}(\boldsymbol{\alpha})$. $\longrightarrow$ $\boldsymbol{\theta}|\boldsymbol{x}_n \sim \text{Dir}(\boldsymbol{\alpha} + \boldsymbol{x}_n)$
- With the posterior distribution it is possible to make inference for the NPS drawing values from the Dirichlet;
- We obtained simple expression for the posterior mean and variance of the NPS.

- **Minimum sample size (Average length criterion)**

Minimum n such that : $$\int_{\mathcal{X}} \ell(\boldsymbol{x}_n) g(\boldsymbol{x}_n)\, d\boldsymbol{x}_n \leq \ell_{\max},$$

where $\ell(\boldsymbol{x}_n) = b(\boldsymbol{x}_n) - a(\boldsymbol{x}_n)$ is the HPD interval length of probability $1-\rho$ for Δ, $\mathcal{X}$ is the sample space associated to $\boldsymbol{x}_n$ and $g(\boldsymbol{x}_n)$ is the marginal probability function of the outcomes.

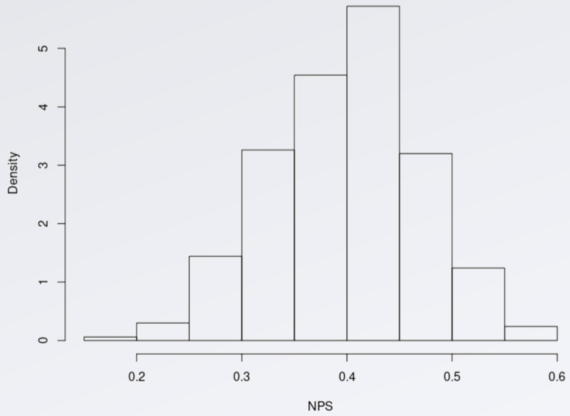
eliardo.costa@ufrn.br

Results and Conclusions

- We developed an R package;
- **Example:** For $\boldsymbol{x}_{100}$ = (14, 31, 55) and $\boldsymbol{\alpha}$ = (1, 1, 1), the Bayesian inference for NPS is:

Mean	SD	HPD (95%)
0.399	0.074	(0.256, 0.542)

Histogram of a NPS posterior sample



ALC minimum sample size

	ρ		
$\ell_{\max}$	0.01	0.05	0.10
0.02	28199	16034	11328
0.04	7153	4085	2805
0.06	3222	1808	1272
0.08	1786	992	721
0.10	1158	655	460
0.12	807	460	315
0.14	591	334	234
0.16	452	249	179
0.18	359	203	142
0.20	291	164	114

eliardo.costa@ufrn.br

4th Conference on Statistics and Data Science
Salvador, Brazil (online)
December 1-3, 2022