arXiv:2303.00835v1 [stat.ME] 1 Mar 2023

净推荐值的贝叶斯推断

Eliardo G. Costa*[1] 和 Rachel Tarini Q. Ponte[2]

[1]北里奥格兰德联邦大学统计系，纳塔尔/北里奥格兰德州，巴西[2]^Anima Educa˜cao，纳塔尔/北里奥格兰德州，巴西

2023年3月3日

摘要

净推荐值是一种简单的指标，多家公司用它来衡量客户忠诚度。探讨该指标统计性质的研究仍然很少，而且没有研究考虑过样本量确定问题。我们采用贝叶斯方法，为净推荐值提供点估计量和区间估计量，并讨论样本量的确定。我们还开发了计算工具，以便在实践中应用这一方法。文中还给出了一个使用金融服务数据的示例。

关键词：客户忠诚度、多项分布、狄利克雷分布、样本量、平均长度准则。

1 引言

Reichheld（2003年）提出了一种名为净推荐值（NPS）的统计指标，公司可将其用作客户忠诚度的指标。作者对一组

*eliardo.costa@ufrn.br

特定行业的客户进行了问卷调查，并结合每位客户的购买历史，确定哪些问题与重复购买或推荐的统计相关性最强。在大多数行业中，其中一个问题的表现更好：“您向朋友或同事推荐[公司 X] 的可能性有多大？”Reichheld（2003年）建议，这个问题的回答应采用 0 到 10 分的评分量表。回答 9 或 10 分的客户被视为“推荐者”，回答 7 或 8 分的客户被视为“被动者”，回答 0 到 6 分的客户则被视为“贬损者”。其理念是，公司 X 的“推荐者”越多，增长就越快。净推荐值的估算方法是“推荐者”与“贬损者”的比例（或百分比）之差。Keiningham 等人（2008年）探讨了净推荐值是衡量公司增长能力最可靠的单一指标，且优于客户满意度指标等说法。Rocks（2016年）简要总结了对净推荐值的一些批评，详见其中的参考文献。

在统计建模的背景下，Rocks（2016年）采用频率学派方法，通过 Wald 区间和 Score 方法，重点研究净推荐值的区间估计。此外，作者进行了模拟研究，以评估所提出区间估计的覆盖概率，并得出结论：调整 Wald 方法的变体和迭代 Score 方法表现更佳。Markoulidakis 等人（2021年）通过机器学习算法，将客户体验视为一个净推荐值分类问题。对于相关工作，我们还可以引用 Eskildsen 与 Kristensen（2011年）以及 Kristensen 与 Eskildsen（2014年）。研究该指标统计性质的文献仍然十分有限，据我们所知，其中没有研究考虑样本量确定问题。在此背景下，我们提出一个贝叶斯模型，以便对净推荐值进行推断，并建立一种样本量确定方法。关于贝叶斯方法在营销中的实用性，请参见 Rossi 与 Allenby（2003年）的论述。

在第2节中，我们介绍贝叶斯模型，以及用于获得净推荐值点估计和区间估计的方法。第3节讨论并实现了最小样本量的确定问题。在第4节中，我们通过金融服务领域的数据给出一个示例。

我们将在第5.2节中作一些总结。

2 贝叶斯模型

设 $\boldsymbol{\theta} = (\theta_1, \theta_2, \theta_3)$，其中 θ_1，θ_2 和 θ_3 分别是客户总体中贬损者、被动者和推荐者的比例。那么，该总体的净推荐值为 $\Delta = \theta_3 - \theta_1$，即感兴趣参数。在包含 n 位客户的样本中，我们根据客户对上述问题的回答，统计各类别中的客户人数。设 $\boldsymbol{X}_n = (X_1, X_2, X_3)$，其中 X_1，X_2 和 X_3 分别是样本中被归类为贬损者、被动者和推荐者的客户人数。

给定 $\boldsymbol{\theta}$，我们假设计数 $\boldsymbol{X}_n$ 服从多项分布，并记作 $\boldsymbol{X}_n|\boldsymbol{\theta} \sim \text{Mult}(n, \boldsymbol{\theta})$。相应的概率分布为

$$\mathbb{P}\left[X_1 = x_1, X_2 = x_2, X_3 = x_3\right] = \frac{n!}{x_1!x_2!x_3!}\theta_1^{x_1}\theta_2^{x_2}\theta_3^{x_3},$$

其中 $x_1, x_2, x_3 = 0, 1, \ldots, n$，满足 $x_1 + x_2 + x_3 = n$，且 $\theta_1 + \theta_2 + \theta_3 = 1$.

$\boldsymbol{\theta}$ 的自然（共轭）先验分布选择是狄利克雷分布，记作 $\boldsymbol{\theta} \sim \text{Dir}(\boldsymbol{\alpha})$，其相应的概率密度函数为

$$\pi(\boldsymbol{\theta}) = \frac{\Gamma(\alpha_1 + \alpha_2 + \alpha_3)}{\Gamma(\alpha_1)\Gamma(\alpha_2)\Gamma(\alpha_3)}\theta_1^{\alpha_1 - 1}\theta_2^{\alpha_2 - 1}\theta_3^{\alpha_3 - 1},$$

其中 $\theta_1 + \theta_2 + \theta_3 = 1$，$\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \alpha_3)$ 是由正超参数组成的向量，$\Gamma(\cdot)$ 是伽马函数。该模型可按以下层次结构表示

$$\boldsymbol{X}_n|\boldsymbol{\theta} \sim \text{Mult}(\boldsymbol{\theta}); \quad \boldsymbol{\theta} \sim \text{Dir}(\boldsymbol{\alpha}). \tag{1}$$

在此设定下，给定对 $\boldsymbol{X}_n$的一次观测 $\boldsymbol{x}_n$，我们有 $\boldsymbol{\theta}$ 的后验分布为参数为 $\boldsymbol{\alpha} + \boldsymbol{x}_n$的狄利克雷分布，即，$\boldsymbol{\theta}|\boldsymbol{x}_n \sim \text{Dir}(\boldsymbol{\alpha} + \boldsymbol{x}_n)$（Turkman 等，2019）。此外，贝叶斯更新变得十分简单，因为后验分布的当前参数

可用作下一次抽样中先验分布的超参数 $\boldsymbol{X}_n$。鉴于存在一种从狄利克雷分布生成随机值的方法，这为我们提供了一种简单的方法，可以从 Δ 的后验分布中抽取值，从而近似获得后验汇总，例如均值、中位数、方差、分位数等，并对净推荐值进行推断。

获取大小为 N 的 Δ 后验分布样本的算法如下所述。

1. 设置 $\boldsymbol{\alpha}$、 $\boldsymbol{x}_n$ 和 N 的值（例如， $N = 1000$）。

2. 从参数为 $\boldsymbol{\alpha} + \boldsymbol{x_n}$ 的狄利克雷分布中抽取一个 $\boldsymbol{\theta} = (\theta_1, \theta_2, \theta_3)$ 值。

3. 计算 $\Delta = \theta_3 - \theta_1$ 并保留该值。

4. 重复步骤 2-3 N 次。

众所周知，狄利克雷分布的边际分布是贝塔分布。令 $\boldsymbol{\alpha}^* = \boldsymbol{\alpha} + \boldsymbol{x}_n = (\alpha_1^*, \alpha_2^*, \alpha_3^*)^\top$。则可得

$$\theta_1|\boldsymbol{x}_n \sim \text{Beta}(\alpha_1^*, \alpha_2^* + \alpha_3^*) \quad \text{and} \quad \theta_3|\boldsymbol{x}_n \sim \text{Beta}(\alpha_3^*, \alpha_1^* + \alpha_2^*),$$

由此可得，净推荐值后验分布的均值为

$$\mathbb{E}\left[\Delta\middle|\boldsymbol{x}_n\right] = \mathbb{E}\left[\theta_3 - \theta_1\middle|\boldsymbol{x}_n\right] = \frac{\alpha_3^* - \alpha_1^*}{\alpha_0^*}, \tag{2}$$

其中 $\alpha_0^* = \alpha_1^* + \alpha_2^* + \alpha_3^*$。该均值可用作净推荐值的点估计量。相应的方差为

$$\begin{aligned} \text{Var}\left[\Delta\middle|\boldsymbol{x}_n\right] &= \text{Var}\left[\theta_3\middle|\boldsymbol{x}_n\right] + \text{Var}\left[\theta_3\middle|\boldsymbol{x}_n\right] - 2\text{Cov}(\theta_3, \theta_1|\boldsymbol{x}_n) \\ &= \frac{\alpha_1^*\alpha_2^* + \alpha_2^*\alpha_3^* + 4\alpha_1^*\alpha_3^*}{(\alpha_0^*)^2(\alpha_0^* + 1)}. \end{aligned} \tag{3}$$

我们可以构造一个基于 (2) 和 (3) 的可信区间，即,$\mathbb{E}\left[\Delta\middle|\boldsymbol{x}_n\right] \pm \gamma\sqrt{\text{Var}\left[\Delta\middle|\boldsymbol{x}_n\right]}$，其中 γ 是一个固定常数。我们开发了一个

Excel 电子表格，用于根据 (2) 和 (3) 计算该可信区间和一个点估计。更多详情请参阅补充材料。

另一种可信区间是最高后验密度（HPD）区间。在这种情况下，我们采用蒙特卡罗方法来近似求得 HPD 区间。换言之，我们从 Δ 的后验分布中抽取样本；由于后验分布是狄利克雷分布，因此很容易做到这一点。更多详情请参见 Turkman 等人（2019年，第 47–48 页）。

3 最小样本量

为确定以预先指定的精度估计 Δ 所需的最小样本量，我们考虑一种基于可信区间平均长度的准则。后验可信区间反映了净推荐值的大小，这有助于公司了解何时进行差距分析并制定业务行动计划，以提高净推荐值，即提高净推荐值，直到公司的推荐者多于贬损者（$\Delta > 0$）。

设 $a(\boldsymbol{x}_n)$ 和 $b(\boldsymbol{x}_n)$ 分别为 Δ 的最高后验密度区间的下限和上限。这里的思路是设定最小贝叶斯覆盖概率$1-\rho$，并通过要求最高后验密度区间的长度 $\ell(\boldsymbol{x}_n) = b(\boldsymbol{x}_n) - a(\boldsymbol{x}_n)$ 满足以下条件来确定最小样本量

$$\int_{\mathcal{X}} \ell(\boldsymbol{x}_n) g(\boldsymbol{x}_n)\, d\boldsymbol{x}_n \leq \ell_{\max}, \tag{4}$$

其中，$\ell_{\max}$ 是最高后验密度区间的最大允许长度，$\mathcal{X}$ 是与 $\boldsymbol{x}_n$ 相关的样本空间，而 $g(\boldsymbol{x}_n)$是结果的边际概率函数。这称为平均长度准则（ALC）。有关该准则的更多详情，请参见 Costa 等人（2021年）及其中的参考文献。

由于实际上无法通过解析方法求得 Δ的最高后验密度区间的下界和上界，我们采用蒙特卡罗方法（Chen 与 Shao，1999）求得相应的上下界以及相应的积分。

满足此标准的最小样本量算法概述如下

概述如下。

1. 设置 $\ell_{\max}$、 $\boldsymbol{\alpha}$、 ρ 的值，并取 $n = 1$。 2. 从 $\boldsymbol{x}_n$中抽取一个样本，样本量为 L（例如， $L = 1000$）；要抽取 $\boldsymbol{x}_n$，先从参数为 $\boldsymbol{\alpha}$ 的狄利克雷分布中抽取一个 $\boldsymbol{\theta}$ 值，再根据该值从参数为 $\boldsymbol{\theta}$的多项分布中抽取 $\boldsymbol{x}_n$。 3. 对抽取的每个 $\boldsymbol{x}_n$，求概率为 $1-\rho$的最高后验密度区间，并计算相应的区间长度：对于步骤 2 中抽取的每个值，按照 Chen 与 Shao（1999）所述，求概率为$1-\rho$的最高后验密度区间的下限和上限。然后，计算每个抽取值对应的上限与下限之差，得到区间长度。 4. 计算 L 个最高后验密度区间长度的平均值。 5. 如果该平均值小于或等于 $\ell_{\max}$，则停止。此步骤得到的值 n 即为所需值。否则，设置 $n = n + 1$ 并返回步骤 2。

我们开发了一个 R 软件包（R 核心团队，2022年），其中提供了一个函数，可通过上一节所述的蒙特卡洛模拟获得点估计和区间估计。此外，该软件包还提供了一个函数，可通过 ALC 的最高后验密度区间计算估计净推荐值所需的最小样本量（见补充材料）。

表1–4展示了在图1所示的 $\boldsymbol{\theta}$ 先验分布以及 $\ell_{\max}$ 和 ρ的若干取值下，针对所有情形，使用通过 ALC 计算的最高后验密度来估计净推荐值所需的最小样本量。对于其他情形，可以使用 R 软件包。 $\alpha_1 = \alpha_2 = \alpha_3 = 1$和 $\alpha_1 = \alpha_2 = \alpha_3 = 5$ 对应的情形中，净推荐值的先验期望值（Δ）为零，但变异性不同。$\alpha_1 = 2$、 $\alpha_2 = 5$ 和 $\alpha_3 = 8$ 对应的情形中，净推荐值的先验期望值（Δ）为正；而在 $\alpha_1 = 8$、 $\alpha_2 = 5$ 和 $\alpha_3 = 2$的情形中，净推荐值的先验期望值（Δ）为负，但相应的方差相等。

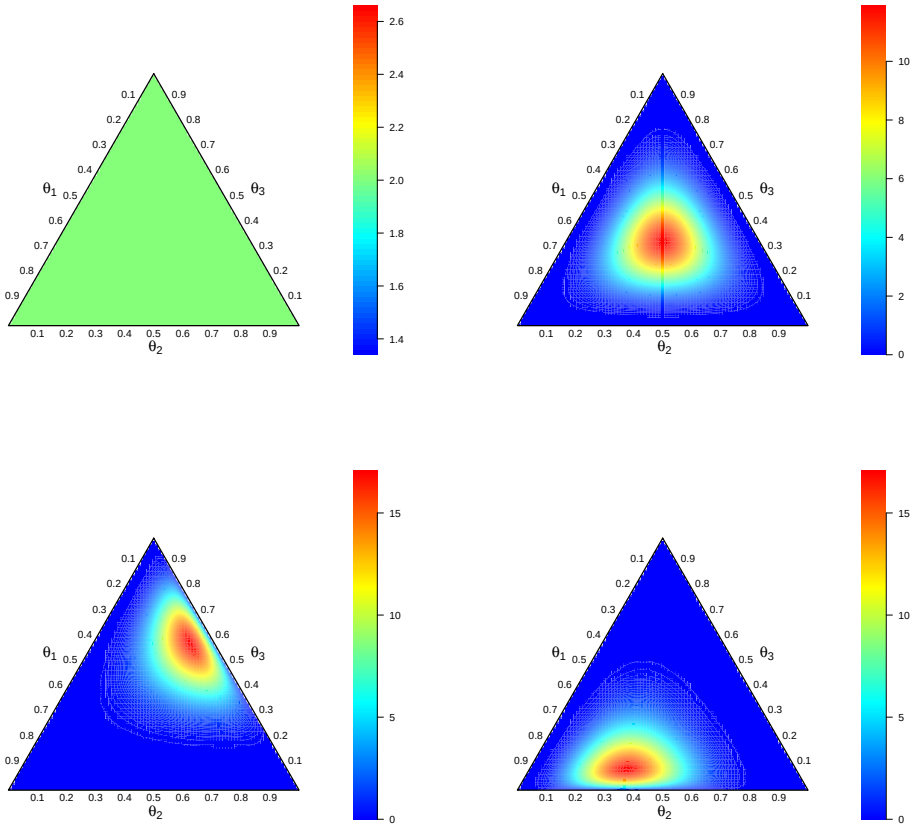

图 1：狄利克雷分布的三元图：顶行分别为 $\boldsymbol{\alpha} = (1, 1, 1)$ 和$\boldsymbol{\alpha} = (5, 5, 5)$；底行分别为 $\boldsymbol{\alpha} = (2, 5, 8)$ 和 $\boldsymbol{\alpha} = (8, 5, 2)$。

当 ρ （$\ell_{\max}$）固定时，随着 $\ell_{\max}$ （ρ）增大，最小样本量会减少，这是符合预期的（表1–4）。在所有 α_i均相等且逐渐增大的情况下，无论 $\ell_{\max}$ 和 ρ 取何值，最小样本量似乎都会增加（表1–2）。当 $\alpha_1 = 2$、$\alpha_2 = 5$ 和 $\alpha_3 = 8$ 取值时，所需的最小样本量与 $\alpha_1 = 8$ $\alpha_2 = 5$、$\alpha_3 = 2$ 取值且 $\ell_{\max}$ 和 ρ 相同时大致相等（表3–4）。这很合理，因为这两种情形在期望值上互为“补充”，但方差相同。

对于所采用的模型参数，计算最小样本量所需的运行时间从 47 秒到 4.69 小时不等，具体取决于

表 1：通过最高后验密度区间，基于 ALC 估算净推荐值的最小样本量，其中 $\alpha_1 = \alpha_2 = \alpha_3 = 1$ 为 $\boldsymbol{\theta}$的先验分布。

	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

设置。 $\ell_{\max}$ 和/或 ρ 的值越小，运行时间越长。所使用计算机的配置如下：操作系统为 Linux Ubuntu 20.04，内存为 7.7 GB，处理器为 AMD PRO A8-8600B。

4 示例

在尚未获得样本的情况下，我们可以通过设置 α_1、 α_2、α_3、 $\ell_{\max}$ 和 ρ，利用 ALC 确定获得净推荐值最高后验密度区间所需的样本量。例如，若考虑 $\alpha_1 = \alpha_2 = \alpha_3 = 1$、 $\ell_{\max} = 0.10$和 $\rho = 0.05$，则最小样本量为 655（表 1）。即，为了获得最大长度为 0.10 且相应概率为 0.95 的最高后验密度区间，我们应向 655 人询问净推荐值问题，然后将他们归类为贬损者、被动者和推荐者，以分别获得 X_1、 X_2 和 X_3的观测值。

鉴于从公司获取真实的净推荐值数据集十分困难，因为此类信息非常敏感，我们考虑使用一个关于2021年三个市场的金融服务的假设数据集（见补充材料），以模拟这些方法在实际情境中的应用。

为说明该方法和贝叶斯更新，我们考虑

表 2：针对先验分布 $\boldsymbol{\theta}$，使用 $\alpha_1 = \alpha_2 = \alpha_3 = 5$ 通过最高后验密度估计净推荐值的 ALC 最小样本量。

	ρ		
$\ell_{\max}$	0.01	0.05	0.10
0.02	37737	21516	15157
0.04	9521	5396	3810
0.06	4243	2386	1681
0.08	2391	1330	944
0.10	1515	860	599
0.12	1046	590	409
0.14	764	433	298
0.16	585	326	224
0.18	464	254	175
0.20	371	204	139

表 3：针对先验分布 $\boldsymbol{\theta}$，使用 $\alpha_1 = 2$、$\alpha_2 = 5$ 和 $\alpha_3 = 8$ 通过最高后验密度估计净推荐值的 ALC 最小样本量。

	ρ		
$\ell_{\max}$	0.01	0.05	0.10
0.02	28494	16105	11346
0.04	7116	4018	2816
0.06	3198	1790	1258
0.08	1779	1004	698
0.10	1137	636	442
0.12	788	442	307
0.14	577	322	225
0.16	433	243	166
0.18	342	189	130
0.20	275	150	100

表4：使用 ALC，根据 $\boldsymbol{\theta}$的先验分布，通过最高后验密度及 $\alpha_1 = 8$、 $\alpha_2 = 5$ 和 $\alpha_3 = 2$ 估计净推荐值所需的最小样本量。

	ρ		
$\ell_{\max}$	0.01	0.05	0.10
0.02	28360	16040	11380
0.04	7160	4028	2861
0.06	3190	1794	1263
0.08	1795	1003	702
0.10	1151	640	443
0.12	781	440	308
0.14	579	322	219
0.16	438	240	164
0.18	345	186	129
0.20	278	152	100

墨西哥市场第一季度和第二季度的数据。对于第一季度，我们没有先验知识，因此设定 $\alpha_1 = \alpha_2 = \alpha_3 = 1$。本季度贬损者、被动者和推荐者的数量分别为 136、82 和 188，这意味着后验狄利克雷分布的向量参数为 $\boldsymbol{\alpha}^* = (137,83,189)^\top$，对应于 $\boldsymbol{\theta}$。从该后验分布中抽取样本并计算其汇总统计量后，我们得到净推荐值的点估计为 0.127，最高后验密度 95%区间为 $[0.038, 0.206]$。对于第二季度，我们可以将第一季度的后验参数用作本季度的先验参数，即 $\alpha_1 = 137$、 $\alpha_2 = 83$ 和 $\alpha_3 = 189$。第二季度贬损者、被动者和推荐者的数量分别为 136、82 和 188。在这种情况下，净推荐值的点估计为 0.131，最高后验密度 95%区间为 $[0.072, 0.192]$。所有这些结果均通过 R 软件包获得。另一种简单的净推荐值点估计和区间估计方法，是按照第2节中的讨论，从这些数据计算出 (2) 和 (3)。

5 结论

文献中首次讨论了用于估计净推荐值的样本量确定问题。为解决这一问题，我们考虑一种贝叶斯

方法，即采用多项式/狄利克雷模型和平均长度准则。我们以封闭形式提供净推荐值的点估计量和区间估计量，或通过从净推荐值的后验分布中抽取样本并计算其汇总统计量来获得这些估计量。此外，贝叶斯方法使推断更新变得十分直接，如第4节所示，即采用序贯程序估计净推荐值。我们开发了计算工具，以便在实践中使用这些方法。

补充材料

Excel 电子表格可在 https://doi.org/10.5281/zenodo.7679211 获取。R 软件包可在 https://github.com/eliardocosta/BayesNPS 获取（DOI：10.5281/zenodo.7617770）。示例中使用的数据可在 https://www.kaggle.com/code/charlottetu/net-promoter-score/ 获取。

参考文献

Chen, M.-H. & Shao, Q.-M. (1999)。贝叶斯可信区间和最高后验密度区间的蒙特卡罗估计。计算与图形统计学杂志 8, 69–92。

Costa, E. G., Paulino, C. D. & Singer, J. M. (2021年)。估计压载水中生物体浓度的样本量：贝叶斯方法。巴西概率与统计学杂志 35, 158 – 171。

Eskildsen, J. K. & Kristensen, K. (2011)。不同分布假设下净推荐值的准确性。载于*2011*年国际质量、可靠性、风险、维护与安全工程会议。

Keiningham, T. L., Aksoy, L., Cooil, B., Andreassen, T. W.&Williams, L. (2008年)。对净推荐进行全面考察。数据库营销与客户战略管理杂志 15，79–90。

Kristensen, K. & Eskildsen, J. (2014年)。净推荐值（NPS）是一项值得信赖的绩效指标吗？《全面质量管理期刊》26，202–214。Markoulidakis, I., Rallis, I., Georgoulas, I., Kopsiaftis, G., Doulamis, A.& Doulamis, N. (2021年)。多分类混淆矩阵缩减方法及其在净推荐值分类问题中的应用。Technologies 9，81。R Core Team(2022年)。R：统计计算语言与环境。R 统计计算基金会，维也纳，奥地利。Reichheld, F. F. (2003年)。助你实现增长的关键数字。《哈佛商业评论》81，46–55。Rocks, B. (2016年)。“净推荐值”的区间估计。《美国统计学家》70，365–372。Rossi, P. E. & Allenby, G. M. (2003年)。贝叶斯统计与营销。《营销科学》22，304–328。Turkman, M. A. A., Paulino, C. D.& Müller, P. (2019年)。计算贝叶斯统计：入门。剑桥：剑桥大学出版社。