

统计过程控制在网络销售中的应用

廉惠然, 齐德全*

长春理工大学数学与统计学院, 吉林 长春

收稿日期: 2023年1月21日; 录用日期: 2023年2月16日; 发布日期: 2023年2月23日

摘要

现有网络销售的研究中, 人们非常关心订货量, 订货周期, 销售价格和利润等变量。从统计过程控制的角度, 给出了监控网络销售变量参数变化的一般框架。鉴于网络销售变量分布的多样性, 建立了多元的Max-EWMA控制图用来监控网络销售变量的均值向量的漂移。通过蒙特卡洛模拟比较了Max-EWMA控制图与MEWMA控制图和 T^2 控制图的性能。以马氏距离度量漂移的大小, 以AR(1)模型刻画变量间的相关性。模拟结果表明, Max-EWMA 控制图对监控中小漂移比较有效, 具有较理想的性能。

关键词

统计过程控制, 网络销售, 控制图, 订货周期

The Application of Statistical Process Control in Network Sales

Huiran Lian, Dequan Qi*

School of Mathematics and Statistics, Changchun University of Science and Technology, Changchun Jilin

Received: Jan. 21st, 2023; accepted: Feb. 16th, 2023; published: Feb. 23rd, 2023

Abstract

In the existing research of online sales, people are very concerned about the variables such as order quantity, order cycle, sales price and profit. From the perspective of statistical process control, a general framework for monitoring the change of network sales variable parameters is given. In view of the diversity of the distribution of online sales variables, a multivariate Max-EWMA control chart is established to monitor the drift of the mean vector of online sales variables. The performances of Max-EWMA control chart, MEWMA control chart and T^2 control chart are compared

*通讯作者。

by Monte Carlo simulation. Mahalanobis distance is used to measure the drift, and AR(1) model is used to describe the correlation between variables. The simulation results show that the Max-EWMA control chart is effective in monitoring small and medium drift, and has better performance.

Keywords

Statistical Process Control, Network Sales, Control Chart, Ordering Cycle

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

伴随着信息科技的发展与创新,网络销售应运而生。现在的网络销售问题,日益引起学者们的关注,从不同的视角展开研究。一方面,研究其特点、发展趋势及影响因素。陈金平(2022)论述了我国电子商务发展的特点和趋势[1]。戴先红(2022)分析了特色农产品网络营销发展的影响因素[2]。张函弛和张幸福(2023)分析了疫情时代对餐饮业网络销售的影响[3]。另一方面,进行预测、数学建模及数据分析。王斌等(2020)利用 ARIMA 模型对网络销售的订单进行了预测[4]。王长琼等(2021)研究了新零售供应链的订货模型及策略[5]。吴宇平和李磊(2022)研究了网络销售数据的稀疏性问题[6]。以上文献从宏观的角度论述了网络销售的特点及发展趋势,也从统计分析的角度进行了预测及数据分析。

在进行统计分析时,所得出的统计结论严重依赖于某些变量或统计量的分布以及所使用的概率或统计模型。当变量或模型的参数发生变化时可能会使得之前所得出的统计结论不再成立,甚至会相反的结果。因此,在网络销售的应用研究中,判断参数、分布或模型是否发生变化是非常有必要的。可以借助于统计过程控制(Statistical Process Control, SPC)中的控制图来解决这个问题。SPC 是保证产品或服务符合规定要求的一种质量管理技术,它包含一些用来降低产品质量波动以使产品质量保持稳定的诸多有效工具。朱建平等(2023)展示了统计学研究的 10 个重要发展方向[7]。在网络产品的质量和生产率统计方法及应用研究中,网络产品的统计过程控制研究是重要的一部分。

综上,本文从另一个视角研究网络销售问题,研究 SPC 在网络销售中的应用,仅给出监控网络销售变量参数是否变化的一般框架。在网络销售的研究中,所关注的变量往往是多元的,包括订货量、订货周期、销售价格等等。因此需要建立多元的控制图对网络销售的过程进行监控。多元控制图的研究成果已非常丰富,包括多元 Shewhart 形式的、累积和(Cumulative Sum, CUSUM)形式的、指数加权滑动平均(Exponentially Weighted Moving Average, EWMA)形式的,还包括多元非参数的控制图。Sanusi 等(2020)建立了二元的 Max-EWMA 控制图用来监控伽马和指数分布的数据[8]。鉴于网络销售变量分布的多样性(订货周期是两次订单的间隔时间,其分布是指数的;订货量是正态的),本文把 Sanusi 等的 Max-EWMA 控制图推广到多元的情况,然后用来监控网络销售的过程。通过蒙特卡洛模拟,Max-EWMA 控制图具有较好的表现。

2. Max-EWMA 控制图

在网络销售问题中,人们通常关注订货量、订货周期、利润、销售价格等多个变量。这里,不妨假设设有 p 元的网络销售指标 $\tilde{X} = (X_1, X_2, \dots, X_p)$ 。假设网络销售指标 \tilde{X} 的协方差矩阵 Σ 保持稳定不变。假

设在某个未知时刻 τ 之后, \tilde{X} 的均值向量由 τ 时刻之前的 μ_0 变为 μ_1 (称为发生了漂移)。我们的目标是对网络销售过程进行监控, 利用控制图来检测测量值分布的任何漂移。

记监控时刻为 t , t 时的样本为 $(X_{1t}, X_{2t}, \dots, X_{pt})$, 通过 Max-EWMA 控制图进行监控的过程如下:

首先计算 EWMA 序列

$$M_{1t} = \lambda_1 X_{1t} + (1 - \lambda_1) M_{1,t-1}, \dots, M_{pt} = \lambda_p X_{pt} + (1 - \lambda_p) M_{p,t-1}, t = 1, 2, \dots, \quad (1)$$

其中 $M_{10} = M_{p0} = 0$, 而 $\lambda_1, \dots, \lambda_p \in (0, 1)$ 称为光滑参数。

然后计算 Max-EWMA 序列 $\{Z_t\}$, 即

$$Z_t = \text{Max}(|M_{1t}|, \dots, |M_{pt}|). \quad (2)$$

当监控统计量 $Z_t > h$ 时, Max-EWMA 控制图发出销售过程失控的警报。 h 称为控制线, 通常使用二分法经蒙特卡洛随机模拟得到。

在线监控的流程图如图 1 所示。

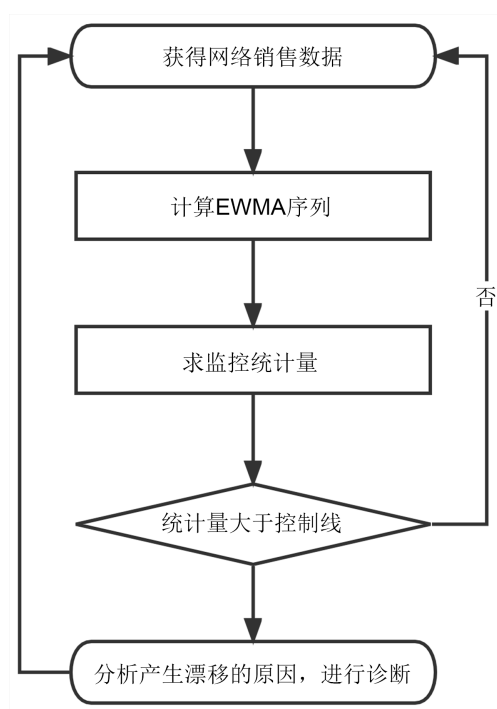


Figure 1. Flow chart of monitoring process

图 1. 监控过程的流程图

研究表明, 当光滑参数 λ 取比较小的数值时, EWMA 类型的控制图对小漂移更有效; 当 λ 取比较大的数值时, EWMA 类型的控制图对比较大的漂移更有效[9]。一般建议 λ 取 0.05, 0.1 或 0.2。如果想兼顾不同大小的漂移, 可以建立自适应的多元 EWMA 控制图[10]或自适应的多元 CUSUM 控制图[11]。

3. 统计模拟

通过统计模拟将 Max-EWMA 控制图与 MEWMA 控制图[12]和 Shewhart 形式的 T^2 控制图[9]进行性能比较。本节后面的所有结果都基于 10,000 次模拟。为简单起见, 假设所关心的网络销售指标是 p 元正态的。尽管对样本作了正态性假设, 事实上根据中心极限定理, 即使样本偏离正态假设, 3 个控制图的

结果仍然是近似可以用的。假设过程受控时, 网络销售指标的均值向量是零向量, 协方差矩阵满足 AR(1) 相关, 即

$$\text{cov}(X_i, X_j) = \begin{cases} \rho^{|j-i|}, & i \neq j \\ 1, & i = j \end{cases}$$

这里 $|\rho| < 1$ 。后面模拟时, ρ 被选为 0.5 和 0.8 两种情况。

在受控时的平均运行长度(ARL, 记为 ARL_0)相同或相近的情况下, 比较不同控制图的过程失控时的 ARL (记为 ARL_1)。哪个控制图的 ARL_1 越小, 其在过程失控时报警越早, 因此表现越好。为了比较的公平, Max-EWMA 控制图与 MEWMA 控制图的光滑参数 λ 都取为 0.05, 调整 3 个控制图的控制线使其 ARL_0 接近 370。分 $p = 5, \rho = 0.5$ (Scenario1)和 $p = 20, \rho = 0.8$ (Scenario2)两种情况, 模拟的控制线见表 1。

Table 1. Simulated h values of different charts

表 1. 模拟的三个控制图的控制线 h

	Max-EWMA	MEWMA	T^2
Scenario1	0.4893	0.3797	18.190
Scenario2	0.5442	0.9606	42.0491

通过马氏距离 $d = \sqrt{(\mu_1 - \mu_0)' \Sigma^{-1} (\mu_1 - \mu_0)}$ 来度量均值向量漂移的大小[13], 其中 μ_0 是过程受控时的均值向量($\mu_0 = 0$), μ_1 是过程失控时的均值向量(假定漂移大小 d 后, 由 d 算得)。这里, μ_1 取 $\delta \mathbf{1}$ 的简单形式, δ 是一个常数, $\mathbf{1}$ 是都为 1 的向量。表 2 给出了 $p = 5, \rho = 0.5$ 时各控制图的 ARL。 $d = 0$ 时表示过程是受控的, 其对应的 ARL 即为 ARL_0 , 其它情况对应的 ARL 是 ARL_1 。由表 2 可以看出, Max-EWMA 控制图对监控中小漂移比较有效, 而 T^2 控制图对监控大漂移比较有效。例如, $d = 0.3$ 时, Max-EWMA 控制图的 ARL_1 为 76.92, 而 T^2 控制图的 ARL_1 为 320.5; $d = 5$ 时, T^2 控制图的 ARL_1 为 1.13, 而 Max-EWMA 控制图的 ARL_1 为 3.16。可见, 对中小漂移, Max-EWMA 控制图能较早地报警; 对大漂移, T^2 控制图能较早地报警。

Table 2. Comparisons of ARL when $p = 5, \rho = 0.5$

表 2. $p = 5, \rho = 0.5$ 时, ARL 的比较结果

d	Max-EWMA	MEWMA	T^2
0	370.6	370.2	370.4
0.3	76.92	93.04	320.5
0.5	37.21	41.95	257.7
1	15.34	15.90	112.9
1.5	9.85	9.88	44.34
2	7.36	7.23	17.95
3	4.98	4.80	4.16
4	3.85	3.66	1.69
5	3.16	3.03	1.13

表 3 给出了 $p = 20, \rho = 0.8$ 时各控制图的 ARL, 与表 2 所得出的结论类似。因此, 在监控网络销售的 p 元指标时, 如果关心的是中小漂移可以建立 Max-EWMA 控制图, 如果关心的是大漂移可以建立 T^2 控

制图。进一步, 如果数据的分布未知, 想要建立稳健的控制图, 可以参考多元的非参数控制图[14]。如果只对某一侧的漂移感兴趣, 可以参考多元的单边控制图[15]。

Table 3. Comparisons of ARL when $p = 20, \rho = 0.8$

表 3. $p = 20, \rho = 0.8$ 时, ARL 的比较结果

d	Max-EWMA	MEWMA	T^2
0	370.4	370.4	370.5
0.3	84.71	149.9	350.9
0.5	41.56	67.48	316.3
1	17.46	23.98	205.5
1.5	11.31	14.66	110.4
2	8.48	10.71	54.01
3	5.79	7.08	12.85
4	4.47	5.36	3.88
5	3.69	4.35	1.75

为了便于理解控制图的使用, 图 2 给出了一个模拟的监控示例。在 $p = 5, \rho = 0.5$ 的情况下, 随机模拟了 50 个样本, 前 5 个样本是受控的, 后 45 个样本是在 d 取 0.5 的情况下产生的。计算各控制图在不同监控时刻对应的监控统计量, 画出散点图, 用虚线连接, 并用虚线在对应的图上分别画出每个控制图的控制线。从图 2 可以很直观地看出, Max-EWMA 控制图与 MEWMA 控制图发出了过程失控的警报, 而 T^2 控制图没有发现漂移。

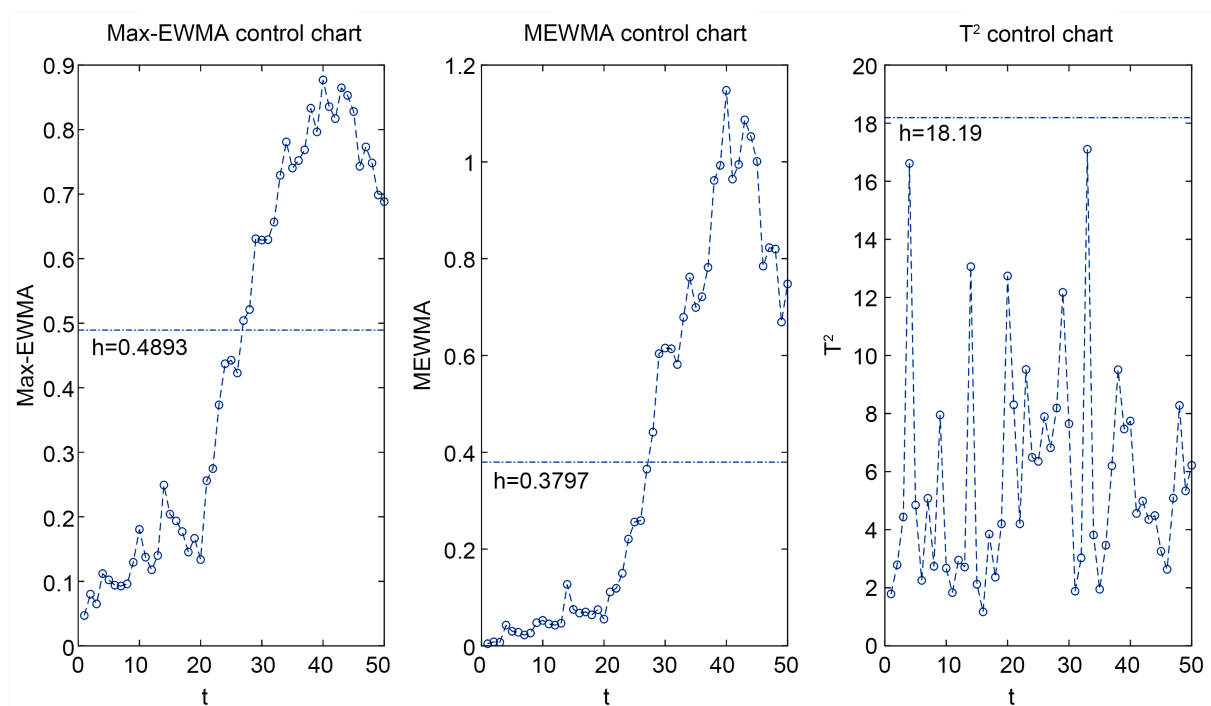


Figure 2. The Max-EWMA, MEWMA and T^2 control charts for the simulated data

图 2. 三个控制图的监控示例

4. 结论

提出了通过统计过程控制对网络销售中的订货量、订货周期、利润、销售价格等多变量进行在线监控的理论框架。建立了 Max-EWMA 控制图对网络销售过程进行监控, 利用该控制图来检测测量值分布的任何漂移。统计模拟表明该控制图对监控中小漂移有良好的表现。为了帮助我们适当地调整销售过程, 还需要检测到底是哪个测量分量发生了漂移, 这称为诊断, 可以参考 Zou 等 2011 年的工作[16]。

基金项目

吉林省教育厅项目(JJKH20210809KJ)、国家自然科学基金面上项目(12271271)。

参考文献

- [1] 陈金平. 我国电子商务发展的特点和趋势[J]. 上海商业, 2022(12): 28-30.
- [2] 戴先红. 基于 Logistic 分析的安徽特色农产品网络营销发展影响因素[J]. 安徽科技学院学报, 2022, 36(5): 102-108.
- [3] 张函弛, 张幸福. 后疫情时代下餐饮业的网络销售研究[J]. 商业经济, 2023(1): 64-66+70.
- [4] 王斌, 杨抒, 贾清, 等. ARIMA 模型在电商平台新疆红枣订单预测中的应用研究[J]. 统计与决策, 2020, 36(6): 35-38.
- [5] 王长琼, 田仁久, 邵明霞. 考虑转运的新零售供应链订货模型及策略[J]. 物流技术, 2021, 40(5): 91-97.
- [6] 吴宇平, 李磊. 易逝品在线销售数据的稀疏性问题及处理[J]. 统计理论与实践, 2022(12): 62-68.
- [7] 朱建平, 冯冲, 梁振杰. 交叉学科促进统计学的发展[J]. 统计研究, 2023, 40(1): 134-144.
- [8] Sanusi, R.A., Yin, T.S. and Khoo, M. (2020) Simultaneous Monitoring of Magnitude and Time-between-Events Data with a Max-EWMA Control Chart. *Computers & Industrial Engineering*, **142**, Article ID: 106378. <https://doi.org/10.1016/j.cie.2020.106378>
- [9] 王兆军, 邹长亮, 李忠华. 统计质量控制图理论与方法[M]. 北京: 科学出版社, 2013.
- [10] 齐德全, 陈实. 一种线性轮廓数据的在线监控方法[J]. 长春理工大学学报(自然科学版), 2022, 45(4): 129-134.
- [11] Dai, Y., Luo, Y., Li, Z. and Wang, Z. (2011) A New Adaptive CUSUM Control Chart for Detecting the Multivariate Process Mean. *Quality & Reliability Engineering International*, **27**, 877-884. <https://doi.org/10.1002/qre.1177>
- [12] Lowry, C.A., Woodall, W.H., Champ, C.W. and Rigdon, S.E. (1992) A Multivariate Exponentially Weighted Moving Average Control Chart. *Technometrics*, **34**, 46-53. <https://doi.org/10.2307/1269551>
- [13] Ghute, V.B. and Shirke, D.T. (2013) A Multivariate Moving Average Control Chart for Mean Vector. *Journal of Academia and Industrial Research*, **1**, 795-800.
- [14] Zou, C., Wang, Z. and Tsung, F. (2012) A Spatial Rank-Based Multivariate EWMA Control Chart. *Naval Research Logistics (NRL)*, **59**, 91-110. <https://doi.org/10.1002/nav.21475>
- [15] Qi, D., Li, Z. and Wang, Z. (2016) On-Line Monitoring Data Quality of High-Dimensional Data Streams. *Journal of Statistical Computation and Simulation*, **86**, 2204-2216. <https://doi.org/10.1080/00949655.2015.1106542>
- [16] Zou, C., Jiang, W. and Tsung, F. (2011) A LASSO-Based Diagnostic Framework for Multivariate Statistical Process Control. *Technometrics*, **53**, 297-309. <https://doi.org/10.1198/TECH.2011.10034>