

SNSにおける口コミとオンラインショップの売上に関するシミュレーション

○武蔵野大学 経済学部 星野雄介, 信州大学 繊維学部松村嘉之

要旨

SNSの普及のより消費者間で情報が増幅され、企業の需要予測に新たな困難が生まれている。本論文では、SNS上での情報拡散と購買行動を「感染症モデル」に当てはめてモデルを作成し、モンテカルロ・シミュレーションによる売上予測と実際の売上高と比較した。その結果、売上予測と実際の売上高にはかい離がありつつも、当モデルには一定の妥当性があることが示された。

1. イントロダクション

1990年代から普及が始まったインターネットは、2010年代のスマートフォンとSNSの登場によって、誰もがどこからでもアクセスされるようになり、我々の生活やビジネスを大きく変化させている。

その動きに対して、ビジネスにおいては、店舗で見てネットで購入するというショールーミングや、店舗・ネット・配送方法などを複数組み合わせるオムニチャネルという概念が生み出されてきた。

他方で、消費者がSNSから情報を収集し、またSNSに投稿することによって、人気が過剰となり短期間で店頭から商品が姿を消すという事例も見られるようになった。これは今後の流行の規模予測や生産管理にとっては、SNSにおける情報拡散を分析する必要があることを示している。

そこで、本論文では「SNSによる情報拡散は、実際の販売量に対し、どの程度のインパクトを持つのか」を問いにし、感染症モデル・モンテカルロ・シミュレーションをキー概念として分析していく。

2. 方法論

(1) 感染症モデル (SIRモデル)

本論文が活用するSIRモデルは、もともとKermackとMcKendricによる古典的な感染症流行モデルとして生み出された[1]。このモデルは局地的な人口における急速かつ短期的な流行に関するモデルであり、世界保健機関による病気の感染予測シミュレーションの基礎となっている。

このモデルでは感受性保持者(Susceptible)、感染者(Infections)、免疫保持者(Removed)の時系列に沿った人口変動によって成り立っている。これらの変動は微分方程式で表すことができ、また、3者の合計が母集団となる特徴がある。

このSIRモデルは、病気の感染以外にも活用可能である。例えば、岡田らは、2011年の東日本大震災におけるデマの拡散過程を、SIRモデルを利用し分析している[2]。このようにSIRモデルにおける母集団を「ある情報に触れたもの」、感染者を「その情報の影響を受けた者」というように見なすとき、本論文が対象とするSNS上でのマーケティングにも適用可能だと思われる。

そこで本研究では感受性保持者S、感染者I、免疫保持者Rを以下のように想定する。

【感染症の流行過程のモデル方程式】

$$\text{感受性保持者(Susceptible)} \quad \frac{d}{dt}S(t) = -\beta S(t)I(t)$$

$$\text{感染者(Infected)} \quad \frac{d}{dt}I(t) = \beta S(t)I(t) - \gamma I(t)$$

$$\text{免疫保持者(Recovered)} \quad \frac{d}{dt}R(t) = \gamma I(t)$$

(β : 感染率, γ : 回復率)

- 感受性保持者S: 購買の可能性のある消費者
- 感染者I: 購買行動に至った消費者
- 免疫保持者R: 購買経験のある消費者

母集団、交流率 λ 、感染率 β 、回復率 γ を以下のように設定することでTwitter上での情報拡散をとらえることとした。

- 母集団: 対象アカウントのフォロワー数 \times 対象となる集団の割合
- 交流率 λ : 対象となる期間のTweetに対するリツイート数の分布の特徴を抽出し、この分布を対数正規分布に近似させる。このときの対数正規乱数を交流率 λ とする
- 感染率 β : 購入リツイート数 \div 商品宣伝記事のリツイート数
- 回復率 γ : 標準正規乱数

また、URLアクセス数と商品購入者数を求めるために大学生16名にアンケートを取り、その数字を利用し以下のように設定した。

$$\text{URLアクセス数} = \sum_{k=5}^5 IP_k A_k = M$$

$$\text{商品購入者数} = \sum_{k=1}^5 MP_k B_k$$

I: ツイートに反応した集団 (SIRモデルで定義した感染者推移)

P1-P5: よくある～まったくない

A1-A5: アンケート「SNS上の記事にあるURL」にアクセスすることはあるか」の回答割合

(A1: 25%, A2: 43.8%, A3: 25%, A4: 0%, A5: 6.3%)

B1-B5: アンケート「SNSで見つけた商品を購入したことはあるか」の回答割合

(B1: 6.3%, B2: 18.8%, B3: 12.5%, B4: 43.8%, B5: 37.5%)

以上で得られたモデルをもとに、100000万回のモンテカルロ・シミュレーションを行い、実際の売上結果と比較していく。

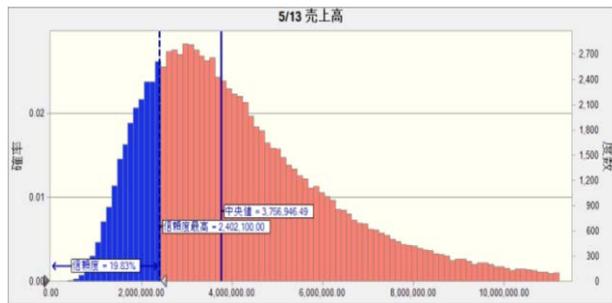
(2)事例

以上のモデルとシミュレーションを検証するための事例として、2016年4月14日九州震災を受けて販売された期間限定グッズを取り上げる。この期間限定グッズの販売期間は、2016年5月2日から6月1日であったが、最終的には6月12日まで販売が延長された。売上公表日は、5月13日、6月1日、6月12日の3回であったことから、この事例はシミュレーションの結果を検証するのにふさわしいと思われる。

3. 結果

以下に、モンテカルロ・シミュレーションの結果と実際の売上高を比較する。

図1 5月13日時点での売上予測分布と実際の売上高



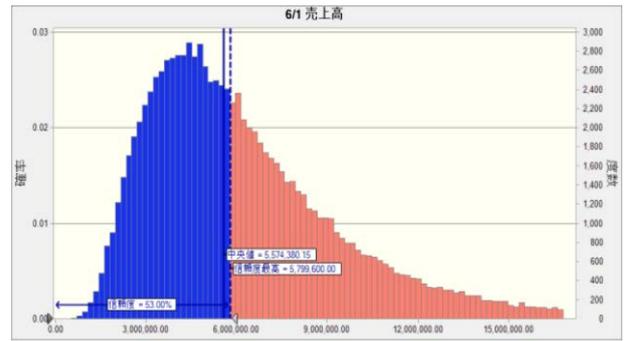
5月13日時点での実際の売上高は、2402100円であった(図1)。それに対し売上高予測分布 μ は4322967円、 σ は2459150円であった。この時期の実際の売上高は、シミュレーション結果の確率密度数分布の $-0.8\sigma < X < 0.8\sigma$ の範囲に収まっている。また、中央値に対する実際の売上高の比率は、63.9%であった。

図2 6月1日時点での売上予測分布と実際の売上高



6月1日時点での実際の売上高は、5799600円であった(図2)。それに対し売上高予測分布 μ は6422973円、 σ は3706126円であった。この時期の実際の売上高は、シミュレーション結果の確率密度数分布の $-0.2\sigma < X < 0.2\sigma$ の範囲に収まっている。また、中央値に対する実際の売上高の比率は、104%であった。

図3 6月12日時点での売上予測分布と実際の売上高



6月12日時点での実際の売上高は、6532500円であった(図3)。それに対し売上高予測分布 μ は6964091円、 σ は4024573円であった。この時期の実際の売上高は、シミュレーション結果の確率密度数分布の $-0.2\sigma < X < 0.2\sigma$ の範囲に収まっている。また、中央値に対する実際の売上高の比率は、108%であった。

4. 考察と結論

以上の結果によると、本論文が提示したモデルとシミュレーションによる売上予測には一定の妥当性があるといえる。しかしながら、かい離していることも事実である。第1に、初期に36.1%マイナスだった理由については、SNS 上での情報拡散から購買行動に至るまでにはタイムラグが生じるためだと考えられる。第2に、中期に売上予測が実際の売上高を4%上回った理由としては、前期から中期の期間中に SNS 上で商品の情報が十分に拡散され、顧客の購買意欲が増進されたのだと考えられる。第3に、終盤に売上予測が実際の売上高を8%上回った理由としては、期間限定販売の効果で駆け込み需要が高まったためだと考えられる。

以上の点を考慮に入れモデルを再構築すべきだと考えられるが、本研究は一定の成果をあげたと考えられるため、モデルの修正は今後の課題とする。

参考資料：

- [1] Kermack, W. O. and A. G. McKendrick. (1927) A contribution to the mathematical theory of epidemics. *Proc. R. Soc. Lond. A* 115, 700-721.
- [2] 岡田佳之, 榎剛史, 鳥海不二夫, 篠田孝祐, 風間一洋, 野田五十樹, & 栗原聡. (2013). マイクロブログにおけるデマの拡散過程の分類と拡張 SIR モデルに基づく解析. *SIG-SAI*, 16(1), 1-9.