

構造変化を考慮した時系列トレンドに基づく売上傾向分析 Statistical Trend Analysis of Sales Data with Structural Changes

中山 雅仁[†]森山 卓[†]桑野 将司[†]

Masahito Nakayama Taku Moriyama Masashi Kuwano

1. はじめに

時系列データは景気や天候などの外的要因によってデータの時系列構造が変化することがある。これは構造変化と呼ばれる。時系列トレンドに関していえば、データが長期間であればあるほど、データ期間全体を通してトレンドが一定であるとは考えにくい。したがって、長期間の時系列データに線形トレンドを当てはめると誤った行動解釈を招く危険性がある。得られるデータの大規模化に伴い、データの鮮度が重要視されるなか、直近の傾向を把握するためには、構造変化を考慮したトレンドを抽出する必要がある。

本研究では、区分線形近似による構造変化を考慮したトレンド分析法を提案する。提案手法の有用性を確認するための実証分析として、実際にある小売業者のPOSデータを使用する。ここで、POSとは“Point of sales”の略称で、販売時点の情報管理を行うシステムのことである。POSシステムにより、お店のレジで支払いがされる際に商品のバーコードをスキャンすることによって日付、商品名、数量、金額などの情報が蓄積される¹⁾。多くの研究者によって、商品の売上確認だけにとどまらない、POSデータの様々な活用方法が議論されている²⁾。例えば、山根ら³⁾は時系列モデルを適用することで販促効果と購買意欲が商品購買数に与える影響を分析した。そして販促効果に基づいた顧客の分類を行い、各顧客群に販促施策の提案を行った。上田ら⁴⁾は、時系列モデルを用いて広告効果の分析を行い、商品ごとに性質が異なることを示した。

このように時系列モデルを用いたマーケティング研究は多数あるが、データ取得が長期間化するなか、どの時点までのデータを用いて今後のトレンドを予測するかについては研究されていない。本研究では、POSデータに含まれる売上データを対象に提案手法を適用し、売上傾向が変化する商品のトレンドを把握することを目的とする。

2. 使用データと分析対象の選択

本研究ではある小売業者のPOSデータにおける売上商品(以後、単に商品と呼ぶ)に注目する。データ取得期間は2015年4月1日から2020年2月29日までの約5年間である。ただし、店休日やレジシステムの不具合によりPOSデータが記録されていない日が45日存在する。

長期的な売上傾向を明らかにすることを目的とし、季節限定商品のような商品や売上個数が少ない商品は分析対象とせず、長期間継続的に売れている商品に着目する。その結果、125商品を分析対象とした。各商品の分析対象期間は「2020年の2月までの期間において毎月1個以上の売上がある」かつ「その期間が最大となる」期間とする。図1に分析対象期間別の商品数を示す。多くの商品の分析対象期間が240-260週(約5年間)である。

3. 構造変化に注目したトレンド分析

3.1 区分線形近似モデルの概要

構造変化を考慮したトレンドの把握を目的に、複数の構造変化点と区分線形トレンドを同時に推定する区分線形近似モデルを提案する。区分線形近似モデルは目的変数と説明変数の関係が区分的に線形である回帰モデルであり、区分線形関数を用いて近似する⁵⁾。区分線形関数は、各区分間が一次関数で表される関数である。目的変数 y 、説明変数 x 、切片 α 、傾き β_0, \dots, β_m 、 m 個の区分点を ψ_1, \dots, ψ_m とする区分線形関数は次の式(1)で表される。

$$y = \alpha + \beta_0 x + \sum_{s=1}^m \beta_s (x - \psi_s)_+ \quad (1)$$

ただし、 $m = 0$ の時、式(1)は単回帰式と一致する。

本研究では、説明変数 x を時間 t とみなし時系列データに区分線形近似モデルを応用する。時間間隔を1週間、目的変数 y_t を t 週目の総売上個数、 T を分析期間の長さとする。得られる区分線形関数を「トレンド」とみなし、直近のトレンドを「最新トレンド」と呼ぶことにする。さらに、売上個数 y_t は式(2)で表されるモデルに従うと仮定する。

$$y_t = \alpha + \beta_0 t + \sum_{s=1}^m \beta_s (t - \psi_s)_+ + \varepsilon_t \quad (2)$$

ここで、 ε_t は独立に平均0、分散 σ^2 の正規分布に従う誤差項とする。このとき、区分点は時系列データにおける構造変化点を表す。区分線形近似モデルのパラメータ $(\alpha, \beta_0, \dots, \beta_m, \psi_1, \dots, \psi_m)$ は式(3)で与えるAICの最小化で推定する。

$$AIC = -2\ln L + 2k \quad (3)$$

ここで、 k はパラメータの数、 L は式(4)で表される。

$$L = \prod_{t=1}^T \log f(\varepsilon_t) \quad (4)$$

f は平均0、分散 σ^2 の正規分布に従う確率密度関数である。構造変化点の数 m の決定に際しては、AICが1つ前の構造変化点を設定したときよりも高い値をとるまで、 m を0から1ずつ増加させる。そして、AICが増加した1つ前の値を最適な構造変化点の数とする。

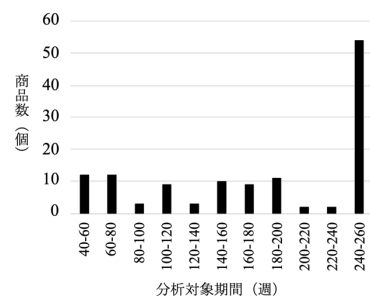


図1 分析対象期間別の商品数

[†] 鳥取大学 Tottori University

3.2 分析結果と考察

分析対象の125商品の売上時系列データに区分線形近似モデルを適用し、商品ごとに構造変化点を抽出した。表1に構造変化点数別の商品数の度数分布を示す。期間内に1回以上のトレンドが変化した商品は全体の約94%を占め、ほとんどの商品で構造変化点が検出された。商品ごとに分析対象期間が異なるため、単純に算出することはできないが、構造変化点の平均個数は約2個である。商品の分析対象期間は平均約179週であるため、平均的には約90週の間には構造変化が生じるといえる。

区分線形近似モデルを適用して得られた結果の例として、構造変化が1回も発生していない商品Aを図2、構造変化が4回発生した商品Bを図3に示す。図4より、201-210週にあたる2019年2月3日から2019年5月4日に構造変化の発生が最も多いことがわかる。この期間には店舗改装があった。これによりいくつかの商品の売上トレンドが変化したと推察される。

最新トレンドのみ着目し、トレンドの傾きについて有意水準5%としたt検定を125商品に対して行った結果を図5にまとめる。図5により、最終時点においてトレンドが正に有意、すなわち売上が増加傾向にある商品が23種類あることがわかる。一方、41種類の商品の売上が有意に減少傾向にあることが本分析より明らかとなった。

4. おわりに

構造変化を考慮しながら売れ行き傾向を商品別に明らかにできたことが本研究の成果である。分析の結果、ほとんどの商品の売上推移に構造変化が発生していたことがわかった。構造変化が生じた要因を考察することは、小売業者がマーケティング施策を考える上で重要な課題である。最新トレンドの活用例としては、商品の原価率と組み合わせで分類を行うことで、利益につながる商品と収益を減らす恐れのある商品を把握できる。原価率だけでなく売上個数や単価などと併せて考慮することでさらなる知見の獲得が期待できる。これは、今後のマーケティング戦略や品揃えなどを検討するうえで重要になりえる。

謝辞

本研究は鳥取県商工労働部兼農林水産部市場開拓局よりデータの提供を賜り実施した。また、分析において鳥取県庁の福山豊様、小林由加里様に多大なるご協力をいただいた。ここに記して謝意を表する。

参考文献

- 西本 章宏, “ビックデータ時代の消費者行動分析: 消費者行動データと消費者選択行動に関する一考察”, マーケティングジャーナル, Vol.34, No.2, pp.47-60 (2014).
- 小沢 佳奈, “近年のPOSデータ研究に関するレビュー”, 繊維機械学会誌, Vol.61, No.5, pp.357-363 (2008).
- 山根 智之, 菅原 光太郎, 西村 直樹, 小林 健, 吉田 祐輔, 高野 祐一, 中田 和秀, “時系列モデルによる商品販促効果の分析”, オペレーションズ・リサーチ, Vol.61, No.2, pp.65-70 (2016).
- 上田 隆徳, 小笠原 浩修, “スキャンパネルデータ利用の広告効果測定 (洗濯用重質洗剤のケース)”, 学習院大学経済論集, Vol.29, No.3,4 合併号, pp.175-203 (1993).
- Vito M. R. Muggeo, “Segmented: An R Package to Fit Regression Models With Broken-line Relationships”, R News, Vol.8, pp.20-25 (2008).

表1 構造変化点の度数分布

構造変化点の数	商品数 (個)
0	7
1	30
2	34
3	33
4	13
5	3
6	4
7	1

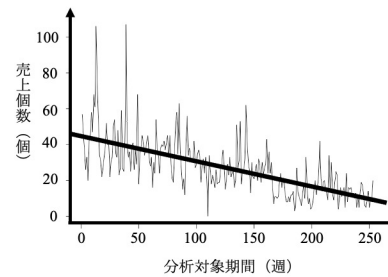


図2 構造変化点が0個の商品A

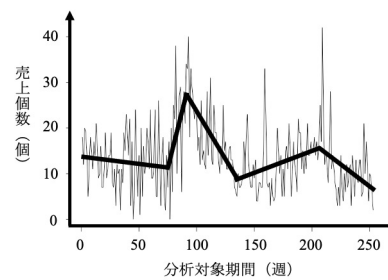


図3 構造変化点が4個の商品B

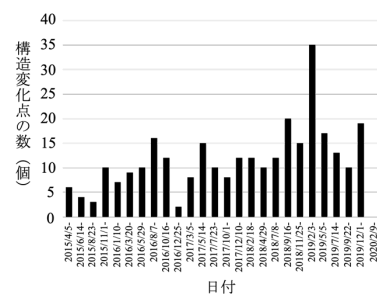


図4 125商品における構造変化の発生時期

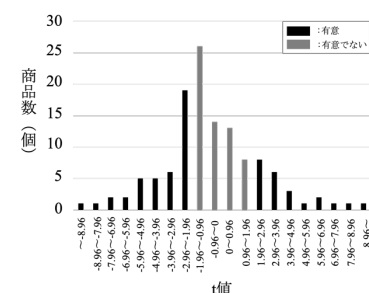


図5 125商品の最新トレンドのt値の分布