

曜日・祝日・降水量の影響を考慮した飲食店売上の時系列分析*

千葉 賢^{*1}

Time Series Analysis of Daily Restaurant Sales: Extraction of Trend, Day-of-the-Week Effect, and Rain Effect

Masaru CHIBA ^{*1}

^{*1} Faculty of Environmental and Information Sciences, Department of Management and Information Sciences

Daily sales of restaurants are affected by several factors such as a day of the week, holidays, and the weather. Therefore, a model decomposing a time series of sales into such factors would definitely be useful. The present study proposes a quantitative model for daily sales of a restaurant based on a state-space model framework, which helps in objectively understanding the characteristics of restaurant sales that were previously discussed only through on-site intuitions. In addition, this model was applied for three years to the daily sales data of a restaurant located adjacent to Fukui station. As a result, the trend was found to be significantly upward during the summer and year-end but declined during the autumn. It was also found that weather may significantly depress a restaurant's sales.

Key words: State-space models, Kalman Filter, Seasonality, Marketing mix modeling

1. 序 論

外食産業を取り巻く環境は、人口減や高齢化による食需要の減少、節約志向の高まりによる低価格化の進行、コンビニやスーパーなどの品揃え拡充に伴う中食市場の拡大、キャッシュレス対応や消費増税などにより、厳しさを増している。帝国データバンク⁽¹⁾は、“2018年度の飲食店の倒産、休廃業・解散の合計は1,180件(前年度比7.1%増)となり、東日本大震災発生後の2011年度(1,134件)、リーマン・ショックが発生した2008年度(1,113件)を上回り、2000年度以降で最多を更新した。”と報告している。そのため、多くの飲食店が、ファミリーレストランやファストフード、居酒屋などで実施されている画一的・均一的な運営から脱却し、他店との差別化を図ろうとする動きが活発になっている。実際、地域限定メニューや内装や食器、制服などによって違いを明確化することで、細分化・多様化した消費者の嗜好に対応している。つまり、店舗固有の状況や消費者の嗜好を考慮した運営が求められている。

前段のような発想に基づく運営を遂行するためには、各店舗売上の長期的傾向(トレンド)や日々の売上の的確に把握することが重要となる。特に後者は仕入れるものの大半が食材で日持ちしないことから、他の産業より正確かつ迅速に予測する必要がある。しかし、飲食店の日々の売上は曜日、祝日、天候、立地などさまざまな要因に左右され、各要因の売上への寄与の仕方や度合も店舗毎に異なる。そのため、各店舗の売上時系列データを複数の要因に分解し、各要因のパラメータの特徴を把握することは、食材発注や人材配置などといった飲食店経営において有益であることは疑いようもない。そこで本研究では、状態空間モデルの柔軟な表現力を利用した要因分解モデルを開発し、福井駅に隣接する飲食店の日次売上データに適用することで、当該飲食店にとって新たな知見や有益な示唆を得ることを目標に分析を行う。

本稿の構成は以下の通りである。第2章では分析対象店舗で収集された売上や来客数データ、福井市で観測された降水量データの特徴について考察する。第3章では、前章で獲得された知見を元に、売上金額を複数の要因に分解し、売上に寄与する要因の特定やその影響の度合いを測定する計量モデルを提案する。第4章では前章で提案されたモデルを観測データに適用し、得られた結果について考察する。最後に、第5章で結論と今後の課題について述べる。

* 原稿受付 2020年5月29日

^{*1} 環境情報学部 経営情報学科, E-mail: chiba-masaru@fukui-ut.ac.jp

2. 分析データの特徴

2.1 分析対象店舗、データに関する基本情報

本稿でとり上げるのは、JR 福井駅から徒歩 3 分ほどの商業ビル地下一階に店舗を構える寿司とアルコール飲料を主に提供する飲食店である。当該店舗は「ハピリン」と呼ばれる複合商業施設や「片町」と呼ばれる歓楽街の近隣に位置している。営業時間はランチ帯が 12:00 から 13:30、ディナー帯が 18:00 から 23:00 だが、ランチは予約が必要で、ディナーが主体となっている。定休日は日・月曜日となっているものの、予約状況によっては営業する場合もある。当該店舗ではさまざまな価格帯のメニューが用意されているが、実際はコース料理が主体となっている。なお、店舗面積は約 20 平米、席数は 13 席（カウンター席のみ）、従業員は経営者夫妻 2 名のみであることから、小規模な飲食店であることが伺える。

本稿で分析するのは、上記飲食店で収集された売上・来客数^{#1}と、気象庁⁽²⁾ が公表している福井市降水量に関するデータである。計測期間は 2016 年 1 月 1 日から 2018 年 12 月 31 日で、観測頻度は日次、計測期間数は 834 である。なお、福井市降水量を分析対象とした理由は、山口他⁽³⁾ や近藤⁽⁴⁾ などで降水量や気温などといった環境変数が小売店の売上に統計的に有意な効果をもたらすと指摘されていること、分析対象店舗は降水量の多い北陸地方にあるため、売上と天候の関係を定量的に把握することは店舗経営者にとって有益と判断したためである。

2.2 記述統計分析

最初に、記述統計量と（偏自己）相関係数を用いてデータの特徴を概説する。Table 1 は売上、来客数、降水量^{#2}の記述統計量だが、売上の中央値や平均の結果から分析対象店舗の 1 営業日当たりの売上は 50,000 円程度であることがわかる。しかし、最大値が 188,250 円、標準偏差が 33,989 円とかなり大きな値であることから、当該店舗の売上はバラつきが大きいことが伺える。また、1 営業日当たりの来客数の中央値、平均はともに約 9 人だが、この結果は分析対象店舗の席数が 13、ディナー主体の高価格帯店舗であることを考慮すれば、妥当な結果といえる。また、すべての変数で歪度は統計的に有意な正の値を示したことから、どの変数の分布も右に歪んでいる（より大きな値が観測されやすい）と判断できる。さらに、降水量の尖度は統計的に有意な正の値であることから、分析地域は豪雨豪雪が観測されやすい地域であることが統計的に確認できる。Fig. 1 上段は各変数の時系列推移である。売上、来客数については大きな特徴は見られないものの、クラスタリングにより、降水量が増加する時期には偏りがあることが確認できる。この結果は、福井市が典型的な日本海側の気候であることを鑑みれば、自然な結果である。Fig. 1 下段は各変数のヒストグラムだが、どの変数の分布も右に歪んでいること、降水量の分布では外れ値が観測されやすいことが視覚的に確認できる。

Table 1 売上、来客数、降水量の記述統計量

変数	最小値	最大値	中位値	平均	標準偏差	変動係数	歪度	尖度
売上 (円)	0	188,250	49,825	53,275* (1.177)	33,989	0.64	0.60* (0.08)	3.10 (0.17)
来客数 (人)	0	28	9	9.21* (0.19)	5.36	0.58	0.56* (0.08)	3.28 (0.17)
降水量 (mm)	0	106.5	0.5	7.03* (0.45)	13.11	1.87	3.05* (0.08)	15.08* (0.17)

注) 各変数の変動係数・歪度・尖度は単位に依存しない。計測期間数、変数 y の標本標準偏差を、それぞれ T 、 $\hat{\sigma}_y$ とすると、平均・歪度・尖度の標準誤差は、それぞれ $\hat{\sigma}_y / \sqrt{T}$ 、 $\sqrt{6/T}$ 、 $\sqrt{24/T}$ より計算できる。なお、当該値は各統計値の括弧内に記載している。また、* が付与されている値は、有意水準 5% で統計的に有意 (5 パーセンタイルは ± 1.96) であることを意味している。

^{#1} 売上・来客数データは店舗経営者より直接提供を受けた。記して感謝申し上げます。

^{#2} 降水量は、雨・雪・霰・雹・霜などをすべて水に換算した際の量を指す。一方、降雨量は雨だけが降った量を意味する。同様に、雪の降った量だけを表す場合は降雪量を用いる。なお、降雪量は雨量計で観測される水に換算した時の水の深さで、積もった雪の量を表すものではない。

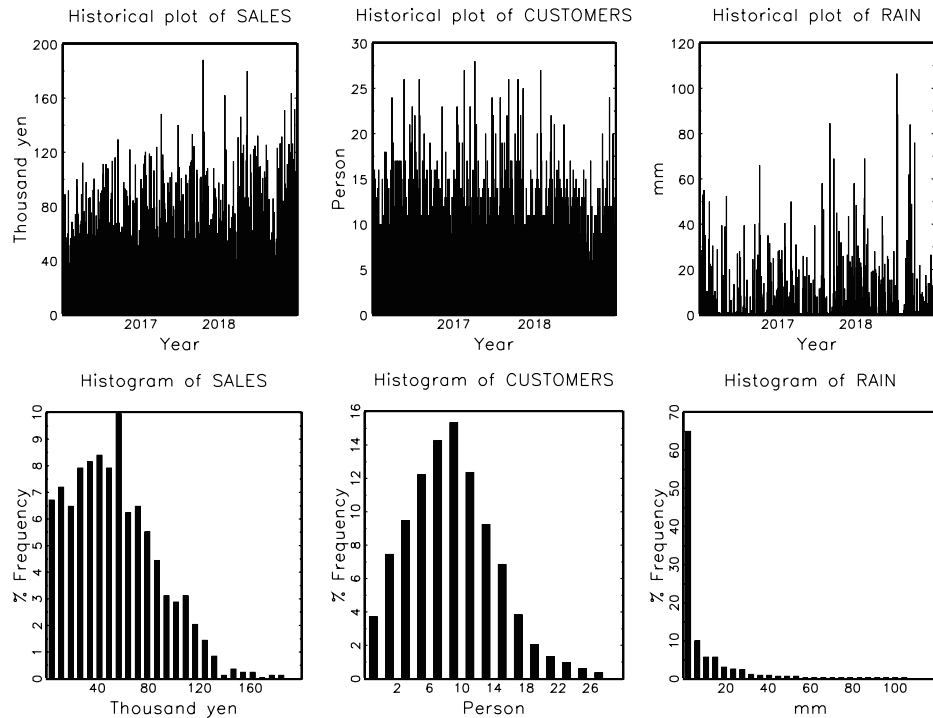


Fig. 1 売上 (左), 来客数 (中央), 降水量 (右) の時系列推移 (上段), ヒストグラム (下段)

Table 2 左側は変数間の相関係数だが, どの組み合わせにおいても統計的に有意な相関関係があることがわかる. ただし, 売上は訪問客の支払金額の総和であることから, 来客数と売上との間に正の相関関係があることは当然といえる. 確認しておく必要があるのは, 降水量と売上・来客数の関係である. 一般的な店舗では, 降水量が増えると来客数は減少し, 結果として売上は下がる場合が多いが, 商業ビルに入っている店舗などでは, 雨に濡れることを避けようとする同一商業ビルに勤務する社会人などがビル内での飲食を選択するため, 売上が上がる場合もある. 分析対象店舗では, 降水量と売上・来客数はともに負の相関関係であることから, 雨が降っても近隣の商業ビルに勤務する社会人が来訪する傾向は低いことが確認された.

Fig. 2 上段は変数間の散布図である. 左の散布図 (横軸: 来客数, 縦軸: 売上) では右上がり 45 度線付近に点が密集しているものの, 来客数が 20 人以上になると直線からの逸脱の度合いが大きくなることからわかる. また, 中央および右の散布図 (横軸: 降水量, 縦軸: 売上, 来客数) では右下がり 45 度線下部領域に点が密集しているが, この結果は降水量が増加するほど売上・来客数が減少する傾向にあることを示唆していることから, 相関関係の結果と整合的である.

Table 2 売上, 来客数, 降水量の相関係数, 偏自己相関係数 (PACF)

相関係数	売上	来客数	降水量	偏自己相関係数	1 次	2 次	3 次	4 次	5 次	6 次	7 次
売上	1			売上	0.01 (3.51)	-0.04 (3.49)	0.03 (3.45)	0.06 (3.45)	0.16* (3.45)	0.10* (3.49)	-0.02 (3.51)
来客数	0.84* (1.40)	1		来客数	-0.03 (3.49)	-0.01 (3.46)	0.00 (3.42)	0.03 (3.42)	0.15* (3.42)	0.12* (3.46)	-0.06 (3.48)
降水量	-0.09* (3.61)	-0.07* (3.59)	1	降水量	0.19 (3.50)	0.06 (3.55)	-0.04 (3.54)	0.02 (3.54)	0.09* (3.54)	-0.05 (3.55)	0.01 (3.49)

注) 括弧内の数値は相関係数, 偏自己相関係数の標準誤差で, 単位はともに % である. なお, 変数 x, y 間の相関係数 ρ_{xy} の標準誤差 $SE(\rho_{xy})$ は, $\sqrt{(1-\hat{\rho}_{xy}^2)/(T-2)}$ より推定できる. ここで, $\hat{\rho}_{xy}, T$ は, それぞれ変数 x, y 間の標本相関係数, 計測期間数である. また, 観測変数 y の p 次の偏自己相関係数 ϕ_p とその標準誤差は, p 次の自己回帰モデル $y_t - \hat{\mu} = \sum_{p=1}^p \phi_p(y_{t-p} - \hat{\mu}) + u_t$ における回帰係数の最小二乗推定値 $\hat{\phi}_p$ 及びその標準誤差 $SE(\hat{\phi}_p)$ より推定できる. ここで, $y_t, u_t, \hat{\mu}$ は, それぞれ t 期の観測変数, 擾乱項, 観測変数の標本平均である. なお, * が付与されている値は, 有意水準 5% で統計的に有意 (5 パーセンタイルは ± 1.96) であることを意味する.

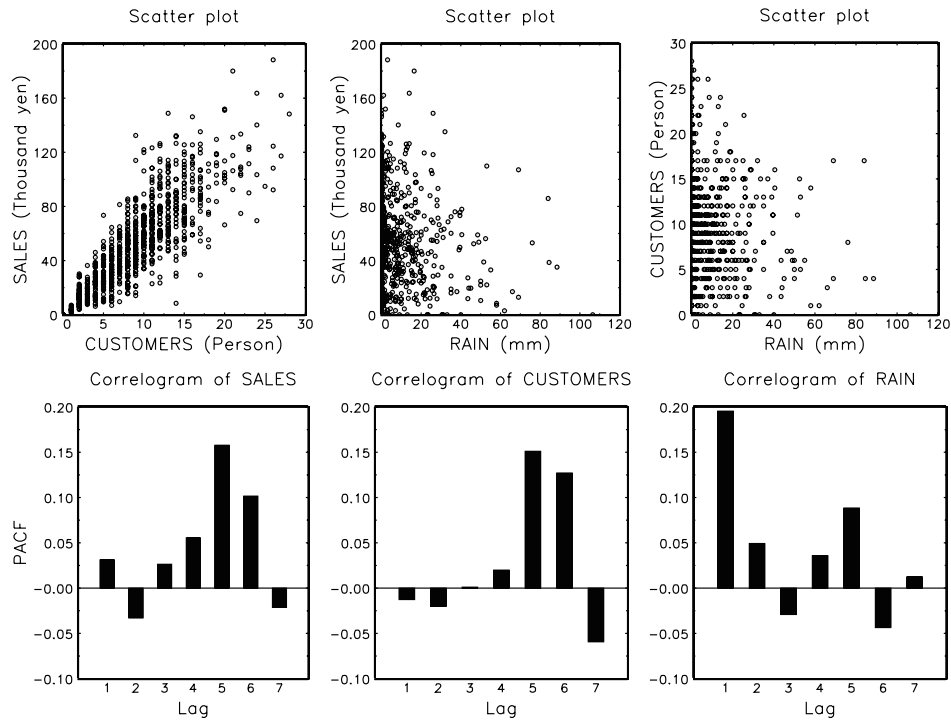


Fig. 2 売上, 来客数, 降水量の散布図 (上段), コレログラム (下段)

Table 2 右側は各変数の偏自己相関係数だが, 売上・来客数ではともに 5 次・6 次の偏自己相関係数が統計的に有意であることがわかる. この結果は, 今日の売上 (来客数) が 5, 6 営業日前の売上 (来客数) と関係性がある, つまり当該データに週次の周期性があることを示唆している. 通常の飲食店では, 金・土曜日もしくは日曜日は来客数が増加するが平日は減少する傾向があるため, この結果は一般的といえる. 一方, 降水量でもっとも高い値を示したのは 1 次の偏自己相関係数 (0.19) であるが, これは降水量が増加 (減少) する時期に偏りがあることを統計的に表している.

2.3 月別・曜日別分析

前節の分析から, 分析対象店舗の売上・来客数データには週次の周期性があることが確認された. なお, 経済・経営活動を通じて観測される時系列データが毎年類似したパターンを示しながら変動することは, 高岡⁽⁵⁾などでも指摘されている. このような変動は季節性と呼ばれ, (1) 季節的な自然条件の変化, (2) カレンダーの周期性, (3) 社会的慣習などによって生じるとされている.

(1) は四季の寒暖変化などによる変動を指している. たとえば, 農業生産は季節の変化に直接的な影響を受ける. 需要面においても, 夏季に清涼飲料水や酒類などに対する需要が高まり, 冬期に灯油の需要が高まるといった変化がもたらされることが知られている. (2) はカレンダーにおける曜日の配置や周期性が原因となって生じる変動である. 1 月から 12 月までの各月の日数は異なることから, 各月の営業日数にも差が生じる. 一般的な店舗では, 各年の売上のうち 2 月が常に低い値になるが, これは主に営業日数の差によって生じることが知られている. さらに, 営業日数は周期的な変化を伴う. カレンダー上では 1 年ごとに曜日が 1 日ずれるため, 同じ月でも年によって各曜日の日数に差異が生じる. また, 経済活動が祝日の影響を受けることは周知の事実であろう. (3) は, 正月, ゴールデンウィーク, 盆, クリスマスといった, 社会的慣習に伴う国民の経済活動の変動を指している. 以上の要因による変動は, 分析対象店舗でも同様に観察されると思われる. そこで本節では, 分析対象データを月・曜日ごとに集計することで, 月・曜日ごとにどのような変化がみられるか確認する.

Table 3 は, 分析対象期間 (36 ヶ月) に計測された来客数, 売上, 降水量を月別に集計したものだが, 1 営業日来客数 (売上), 客単価が高いのは 3, 4, 6, 12 月であることがわかる. 3 月は送別会, 4 月は歓迎会, 12 月は忘年会やクリスマスといったイベントにより来客数・売上が増加すると考えられるが, 特段目立った社会的慣習行事がない 6 月の売上が高い

のは興味深い。夏季賞与による売上増加も考えられるが、公務員、民間企業の夏季賞与支給日はそれぞれ6月30日、7月初旬が一般的であることから、その可能性は低いと思われる。代替理由として、当該月の降水量の少なさが考えられる。実際、6月の平均降水量は3.74mmと、もっとも低い値を記録している。降水量が少なく、気温も適度に上昇するこの時期に、友人や同僚、恋人や家族などと外食しようとするのは自然な意思決定といえよう。

一方、1営業日来客数(売上)、客単価が低いのは1, 2, 9, 10, 11月であることが確認できる。1, 2月の売上が低い理由としては、年末年始の出費による反動や営業日数の少なさ、年度末の出費を考慮した消費控えなどさまざまな要因が考えられる。9, 10, 11月の売上低迷の理由としては、社会的慣習行事の少なさ、年末の出費を考慮した消費控えなどが挙げられる。また、降水量の増加も大きな要因と考えられる。北陸地方では1月から2月にかけて降雪量が増加すること、9月に台風の襲来が増えることは、よく知られた事実である。

Table 4は来客数, 売上を曜日別に集計したものだが、1営業日来客数, 売上はともに土、金、水、木、火曜日の順で多く、通常の飲食店と同様に金・土曜日の集客力が高いことがわかる。金曜日以外の平日では水曜日がもっとも来客数や売上が高いが、これは水曜日にノー残業デーを実施する企業が多いため、余暇時間が増えた社会人が来訪していると考えられる。

Table 3 月別来客数, 売上, 降水量

月	営業日数	来客数	売上	1 営業日来客数	1 営業日売上	客単価	降水量
1	62	541	2,916,500	8.73 ⁽⁸⁾	47,040 ⁽¹⁰⁾	5,391 ⁽¹¹⁾	12.44 ⁽¹²⁾
2	55	469	2,328,460	8.53 ⁽⁹⁾	42,336 ⁽¹²⁾	4,965 ⁽¹²⁾	11.10 ⁽¹⁰⁾
3	75	757	4,245,150	10.09 ⁽²⁾	56,602 ⁽³⁾	5,608 ⁽¹⁰⁾	4.29 ⁽³⁾
4	69	672	3,862,920	9.74 ⁽⁴⁾	55,984 ⁽⁴⁾	5,748 ⁽⁸⁾	6.07 ⁽⁶⁾
5	73	664	3,733,400	9.10 ⁽⁷⁾	51,142 ⁽⁸⁾	5,623 ⁽⁹⁾	4.24 ⁽²⁾
6	71	745	4,426,650	10.49 ⁽¹⁾	62,347 ⁽²⁾	5,942 ⁽³⁾	3.74 ⁽¹⁾
7	71	672	3,926,800	9.46 ⁽⁵⁾	55,307 ⁽⁵⁾	5,843 ⁽⁵⁾	6.45 ⁽⁸⁾
8	71	661	3,914,300	9.31 ⁽⁶⁾	55,131 ⁽⁶⁾	5,922 ⁽⁴⁾	4.90 ⁽⁴⁾
9	70	582	3,390,400	8.31 ⁽¹¹⁾	48,434 ⁽⁹⁾	5,825 ⁽⁶⁾	11.18 ⁽¹¹⁾
10	73	585	3,384,750	8.01 ⁽¹²⁾	46,366 ⁽¹¹⁾	5,786 ⁽⁷⁾	6.14 ⁽⁷⁾
11	71	600	3,717,550	8.45 ⁽¹⁰⁾	52,360 ⁽⁷⁾	6,196 ⁽²⁾	5.77 ⁽⁵⁾
12	73	732	4,584,390	10.03 ⁽³⁾	62,800 ⁽¹⁾	6,263 ⁽¹⁾	9.70 ⁽³⁾
計	834	7,680	44,431,270	9.21	53,275	5,785	7.14

Table 4 曜日別来客数, 売上

曜日	営業日数	来客数	売上	1 営業日来客数	1 営業日売上	客単価
日 (SUN)	74	667	3,256,300	9.01	44,004	4,882
月 (MON)	6	32	255,450	5.33	42,575	7,983
火 (TUE)	146	851	5,464,650	5.83 ⁽⁵⁾	37,429 ⁽⁵⁾	6,421 ⁽¹⁾
水 (WED)	150	1,388	7,824,200	9.25 ⁽³⁾	52,161 ⁽³⁾	5,637 ⁽⁴⁾
木 (THU)	151	1,265	7,107,050	8.38 ⁽⁴⁾	47,067 ⁽⁴⁾	5,618 ⁽⁵⁾
金 (FRI)	153	1,606	9,507,710	10.50 ⁽²⁾	62,142 ⁽²⁾	5,920 ⁽²⁾
土 (SAT)	154	1,871	11,015,910	12.15 ⁽¹⁾	71,532 ⁽¹⁾	5,888 ⁽³⁾
合計	834	7,680	44,431,270	9.21	53,275	5,785

注) 来客数, 1 営業日来客数の単位は「人」、売上, 1 営業日売上, 客単価の単位は「円」、降水量の単位は「mm」である。なお、1 営業日来客数(売上)は、各月の来客数(売上)合計を対応する月の営業日数で除すことで計算した。客単価は、各月の売上合計を対応する月の来客数で除すことで計算した。また、降水量は、各月に観測された総降水量を対応する月の日数(営業日数ではない)で除すことで計算した。さらに、上付き文字括弧内の数字は各項目を降順(降水量は昇順)に並び替えた時の順位を表している。なお、曜日別集計値では、イレギュラーに営業する日・月曜日は順位付け対象外とした。

3. 要因分解モデル

前章の分析結果から、分析対象店舗の売上には季節性があること、曜日の違いや降水量の増減によって変化する可能性があることが示された。しかしながら、前章の分析だけでは、売上は上記要因からどのような形で影響を受けるのか、長期的にはどのような傾向で推移するのか、といった飲食店経営にとって重要な問いに答えることはできない。そこで本章では、観測された売上時系列データから各要因によって変動する成分を分解、抽出する方法を提案する。

3.1 時系列モデルの特定

観測された売上データから要因の異なる複数の成分を分解するためには、売上や各成分が従う時系列モデルを明示的に特定する必要がある。そこで本節では、売上データに内在する各成分が識別されるために必要となる条件を仮定する。本稿では、 t 期 (日目) に観測された売上 (千円) y_t は、以下の 4 つの観測不可能な成分の和によって表されると仮定する。

$$y_t = \mu_t + w_t + r_t + u_t, \quad t = 1, 2, \dots, T \quad (1)$$

ここで、 μ_t , w_t , r_t , u_t はそれぞれ t 期の傾向成分 (トレンド)、曜日効果 (パターン) 成分、天候成分、不規則成分である。因みに、式 (1) のような形式のモデルは加法型モデルと呼ばれる。

さらに、各成分がそれぞれ異なる統計的性質をもった確率過程であると考え、各性質に応じたモデルを明示的に仮定する。まず、傾向成分は滑らかに推移すると考え、以下のような酔歩 (ランダムウォーク) 過程に従うと仮定する。

$$\mu_t = \mu_{t-1} + e_t, \quad e_t \sim \text{i.i.d.} N(0, \omega^2) \quad (2)$$

ここで、 e_t は正規乱数である。次に、来客行動の曜日パターンについて、祝日は直前の日曜日のパターンに、祝日前日はそれぞれ直前の金・土曜日のパターンに類似していると仮定し、以下のように表現する。

$$w_t = s_t + d_{1t} [\alpha(s_{\text{SUN},t} - s_t)] + d_{2t} [\beta(s_{\text{FRI},t} - s_t) + \gamma(s_{\text{SAT},t} - s_t)] \quad (3)$$

ここで、 s_t は曜日効果の基本パターンであり、以下のように表される。

$$s_t = - \sum_{p=1}^6 s_{t-p} + \epsilon_t, \quad \epsilon_t \sim \text{i.i.d.} N(0, \varpi^2) \quad (4)$$

ここで、 ϵ_t は正規乱数である。なお、式 (4) は「曜日効果は 7 日間の周期性を持つ」という特徴を、「(確率的に変動するものの、) 曜日効果の和はほぼ一定 (0) となる」という形式で数学的に表現したものである。また、 d_{1t} , d_{2t} は、それぞれ祝日・祝日前日を示すダミー変数で、Table 5 のように表される。

Table 5 祝日・祝日前日を示すダミー変数の定義

d_{1t}	d_{2t}	値
t 日目が月曜日から金曜日の祝日	t 日目が祝日でない月曜日から木曜日で、かつ翌日が祝日	1
それ以外	それ以外	0

Table 6 t 日目の曜日効果と直前の金、土、日曜日の曜日効果の関係

t 日目の曜日	$s_{\text{FRI},t}$	$s_{\text{SAT},t}$	$s_{\text{SUN},t}$
日 (SUN)	s_{t-2}	s_{t-1}	s_t
月 (MON)	s_{t-3}	s_{t-2}	s_{t-1}
火 (TUE)	s_{t-4}	s_{t-3}	s_{t-2}
水 (WED)	s_{t-5}	s_{t-4}	s_{t-3}
木 (THU)	$-\sum_{p=0}^5 s_{t-p}$	s_{t-5}	s_{t-4}
金 (FRI)	s_t	$-\sum_{p=0}^5 s_{t-p}$	s_{t-5}
土 (SAT)	s_{t-1}	s_t	$-\sum_{p=0}^5 s_{t-p}$

さらに、 $s_{\text{SUN},t}$ は直前の日曜日、つまり s_t に一番近い過去の日曜日の成分とする。同様に、 $s_{\text{FRI},t}$ 、 $s_{\text{SAT},t}$ は直前の金、土曜日の成分とする。これらの成分は、 t 日目が該当する曜日によって、対応する成分が Table 6 のように変化する。また、 α は $s_{\text{FRI},t}$ と s_t との類似度を、 β 、 γ は、それぞれ $s_{\text{SAT},t}$ 、 $s_{\text{SUN},t}$ と s_t との類似度を表す媒介変数 (パラメータ) である。ただし、 $0 \leq \alpha \leq 1$ 、 $0 \leq \beta \leq 1$ 、 $0 \leq \gamma \leq 1$ 、 $0 \leq \beta + \gamma \leq 1$ が仮定される。

具体例として、 t 日目が火曜日で、かつ祝日であった場合、 w_t がどのように表されるか確認する。この場合、Table 5, 6 より、 $d_{1t} = 1$ 、 $d_{2t} = 0$ 、 $s_{\text{FRI},t} = s_{t-4}$ 、 $s_{\text{SAT},t} = s_{t-3}$ 、 $s_{\text{SUN},t} = s_{t-2}$ となるため、これらを式 (3) に代入すると、

$$w_t = s_t + 1 \cdot [\alpha(s_{t-2} - s_t)] + 0 \cdot [\beta(s_{t-4} - s_t) + \gamma(s_{t-3} - s_t)] = (1 - \alpha)s_t + \alpha s_{t-2}$$

となる。今度は、 t 日目が祝日でない水曜日で、かつ翌日が祝日であった場合を考える。この場合、 $d_{1t} = 0$ 、 $d_{2t} = 1$ 、 $s_{\text{FRI},t} = s_{t-5}$ 、 $s_{\text{SAT},t} = s_{t-4}$ 、 $s_{\text{SUN},t} = s_{t-3}$ となるため、これらを式 (3) に代入すると、

$$w_t = s_t + 0 \cdot [\alpha(s_{t-3} - s_t)] + 1 \cdot [\beta(s_{t-5} - s_t) + \gamma(s_{t-4} - s_t)] = (1 - \beta - \gamma)s_t + \gamma s_{t-4} + \beta s_{t-5}$$

となる。このように、式 (3) は曜日効果のみならず、祝日や祝日前日が売上に与える効果も柔軟に表現することができる。

次に、天候成分は福井市降水量が当該店舗売上に与える影響と考え、以下のような単純な線形関数を仮定した。

$$r_t = \delta z_t \quad (5)$$

ここで、 z_t は t 日目に福井市で観測された一日の降水量合計 (mm)、 δ は降水量に対する売上 (千円) の影響の度合いを表すパラメータで、単位は「千円/mm」である。なお、 δ が正 (負) の場合は「雨が降れば売上が上 (下) がる」ことを意味する。最後に、不規則成分は以下のような定常 2 次自己回帰過程 (AR(2)) に従うと仮定した。

$$u_t = \phi_1 u_{t-1} + \phi_2 u_{t-2} + \zeta_t, \quad \zeta_t \sim \text{i.i.d.} N(0, \sigma^2) \quad (6)$$

ここで、 ϕ_1 、 ϕ_2 は自己回帰係数、 ζ_t は正規乱数である。なお、自己回帰過程の次数を増やすことは技術的に可能だが、Kim⁽⁶⁾ や Clark⁽⁷⁾ など、多くの先行研究で次数は 2 次に設定されていることから、本稿でも 2 次とした。

3.2 状態空間モデルを用いた定式化

前節で説明したように、本稿では売上に観測不可能な 4 つの成分の和によって表されると仮定している。そのため、何らかの手法を用いてこれらを推定する必要がある。そこで本稿では、売上や各成分が従う確率過程を状態空間モデルで定式化し、カルマンフィルタ、スムーザと呼ばれる逐次アルゴリズムを用いて各成分を抽出する。

状態空間モデルは観測方程式、遷移方程式の 2 つの方程式によって構成されるが、観測不可能な変数 (潜在変数) を 1 階の確率差分方程式で表す必要がある。本稿では、傾向成分は式 (2) によって既に 1 階の確率差分方程式で表されていることから、本節では曜日効果成分と不規則成分の定式化について議論する^{#3}。

曜日効果成分の基本パターンである s_t は式 (4) によって表されることから、以下のように定式化できる^{#4}。

$$\begin{bmatrix} s_t \\ s_{t-1} \\ s_{t-2} \\ s_{t-3} \\ s_{t-4} \\ s_{t-5} \end{bmatrix} = \begin{bmatrix} -1 & -1 & -1 & -1 & -1 & -1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} s_{t-1} \\ s_{t-2} \\ s_{t-3} \\ s_{t-4} \\ s_{t-5} \\ s_{t-6} \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \cdot \epsilon_t \quad (7)$$

なお、式 (7) の 1 行目は式 (4)、2 行目以降は恒等式 ($s_{t-1} = s_{t-1}$ 、 \dots 、 $s_{t-5} = s_{t-5}$) となっている点に注意されたい。

同様に、不規則成分 u_t は式 (6) によって表されることから、以下のように定式化できる。

$$\begin{bmatrix} u_t \\ u_{t-1} \end{bmatrix} = \begin{bmatrix} \phi_1 & \phi_2 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} u_{t-1} \\ u_{t-2} \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} \cdot \zeta_t \quad (8)$$

^{#3} 天候成分は、一部が福井市降水量という形で観測可能であること、パラメータ δ が時刻 t に依存しないことから、式 (11) のように表現する。

^{#4} Table 6 において、なぜ t 日目が該当する曜日が木 (金、土) 曜日の時、 $s_{\text{FRI},t}$ ($s_{\text{SAT},t}$ 、 $s_{\text{SUN},t}$) が対応する成分が s_{t-6} でないかという点、式 (7)、(9) の定式化において s_{t-6} が各式の左辺に含まれていないからである。そこで、曜日効果の和はほぼ 0 という式 (4) の条件から $s_{t-6} = -\sum_{p=0}^5 s_{t-p}$ を導出し、 s_{t-6} を $-\sum_{p=0}^5 s_{t-p}$ に置き換えた。なお、当該処理は樋口⁽⁸⁾ の第 8 章を参考に行った。

ここで、(7) と同様に、式 (8) の 1 行目は式 (6), 2 行目は恒等式 ($u_{t-1} = u_{t-1}$) となっている。

次に、式 (2), (7), (8) の左辺に登場する観測不可能な成分を、以下のような潜在変数ベクトル \mathbf{x}_t でまとめて表現する。

$$\mathbf{x}_t = (\mu_t, s_t, s_{t-1}, s_{t-2}, s_{t-3}, s_{t-4}, s_{t-5}, u_t, u_{t-1})'$$

上記定式化のもと、左辺に \mathbf{x}_t , 右辺に 1 期前の潜在変数 \mathbf{x}_{t-1} を用いて、式 (2), (7), (8) の関係を以下のように表す。

$$\begin{aligned} \begin{bmatrix} \mu_t \\ s_t \\ s_{t-1} \\ s_{t-2} \\ s_{t-3} \\ s_{t-4} \\ s_{t-5} \\ u_t \\ u_{t-1} \end{bmatrix} &= \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & -1 & -1 & -1 & -1 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \phi_1 & \phi_2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \mu_{t-1} \\ s_{t-1} \\ s_{t-2} \\ s_{t-3} \\ s_{t-4} \\ s_{t-5} \\ s_{t-6} \\ u_{t-1} \\ u_{t-2} \end{bmatrix} + \begin{bmatrix} e_t \\ \epsilon_t \\ \zeta_t \end{bmatrix} \\ \Downarrow \\ \mathbf{x}_t &= \mathbf{F}\mathbf{x}_{t-1} + \mathbf{R}\mathbf{v}_t \end{aligned} \quad (9)$$

ここで、式 (9) の 1, 2, 8 行目は、それぞれ式 (2), (4), (6) に、それ以外の行は恒等式になっている。また、

$$\begin{aligned} \begin{bmatrix} e_t \\ \epsilon_t \\ \zeta_t \end{bmatrix} &\sim \text{i.i.d.} N \left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \omega^2 & 0 & 0 \\ 0 & \varpi^2 & 0 \\ 0 & 0 & \sigma^2 \end{bmatrix} \right) \\ \Downarrow \\ \mathbf{v}_t &\sim \text{i.i.d.} N(\mathbf{0}, \mathbf{Q}) \end{aligned} \quad (10)$$

となるが、各成分の正規乱数 e_t, ϵ_t, ζ_t は同時点及び異時点間で互いに独立と仮定されている点に注意されたい。なお、状態空間モデルでは、式 (9) のような 1 階の確率差分方程式を遷移方程式と呼ぶ。

次に、左辺に売上 (観測変数) y_t , 右辺に潜在変数 \mathbf{x}_t , 外生的に与えられる福井市降水量 (外生変数) z_t を置き、両者の関係を記述する。 t 日目が祝日や祝日前日でない場合、 $d_{1t} = 0, d_{2t} = 0$ より、 $w_t = s_t$ となることから、

$$\begin{aligned} y_t &= \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \mu_t \\ s_t \\ s_{t-1} \\ s_{t-2} \\ s_{t-3} \\ s_{t-4} \\ s_{t-5} \\ u_t \\ u_{t-1} \end{bmatrix} + \delta z_t \\ \Downarrow \\ y_t &= \mathbf{Z}_t \mathbf{x}_t + \mathbf{M} z_t \end{aligned} \quad (11)$$

となる。なお、式 (11) は状態空間モデルでは観測方程式と呼ばれる。

ここで注意しなければならないのは、式 (11) は、あくまでも「 t 日目が祝日や祝日前日でない」場合であり、そうでなかった場合、式 (3), Table 5, 6 にしたがって \mathbf{Z}_t を変更する必要があるという点である。一見、 \mathbf{Z}_t の変更は複雑に見えるが、Table 7 に示したようにそのパターンは高々 10 種類であり、 t 日目がどのパターンになるかは事前に確定しているため、実際は各時点で \mathbf{Z}_t を入れ替えるだけでよい。

以上の議論から、本稿で提案された飲食店売上の要因分解モデルは、式 (9), (10), (11) によって定式化される状態空間モデルで表すことができる。さらに、式 (9), (11) より遷移方程式・観測方程式では線形性が成立し、式 (10) より攪乱項は正規分布に従うことから、Kalman⁽⁹⁾ の濾波・平滑化アルゴリズムと最尤法を用いて潜在変数 \mathbf{x}_t や未知パラメータ $\boldsymbol{\theta} = (\alpha, \beta, \gamma, \delta, \phi_1, \phi_2, \omega, \varpi, \sigma)'$ の推定が可能となる^{#5}。

^{#5} 濾波・平滑化アルゴリズムと最尤法については付録参照。

Table 7 t 日目のカレンダー状況と Z_t の関係

No.	t 日目のカレンダー状況	Z_t
1	祝日, 祝日前日いずれでもない	$[1 \ 1 \ 0 \ 0 \ 0 \ 0 \ 0 \ 1 \ 0]$
2	月曜日の祝日	$[1 \ 1-\alpha \ \alpha \ 0 \ 0 \ 0 \ 0 \ 1 \ 0]$
3	火曜日の祝日	$[1 \ 1-\alpha \ 0 \ \alpha \ 0 \ 0 \ 0 \ 1 \ 0]$
4	水曜日の祝日	$[1 \ 1-\alpha \ 0 \ 0 \ \alpha \ 0 \ 0 \ 1 \ 0]$
5	木曜日の祝日	$[1 \ 1-\alpha \ 0 \ 0 \ 0 \ \alpha \ 0 \ 1 \ 0]$
6	金曜日の祝日	$[1 \ 1-\alpha \ 0 \ 0 \ 0 \ 0 \ \alpha \ 1 \ 0]$
7	祝日でない月曜日で, かつ翌日が祝日	$[1 \ 1-\beta-\gamma \ 0 \ \gamma \ \beta \ 0 \ 0 \ 1 \ 0]$
8	祝日でない火曜日で, かつ翌日が祝日	$[1 \ 1-\beta-\gamma \ 0 \ 0 \ \gamma \ \beta \ 0 \ 1 \ 0]$
9	祝日でない水曜日で, かつ翌日が祝日	$[1 \ 1-\beta-\gamma \ 0 \ 0 \ 0 \ \gamma \ \beta \ 1 \ 0]$
10	祝日でない木曜日で, かつ翌日が祝日	$[1 \ 1-2\beta-\gamma \ -\beta \ -\beta \ -\beta \ -\beta \ \gamma-\beta \ 1 \ 0]$

4. 推定結果

本章では, 第 3 章で提案された要因分解モデルを第 2 章で紹介した店舗の売上データ, 福井市降水量データに適用した結果, 得られたパラメータと潜在変数の推定結果について考察する.

4.1 パラメータ

Table 8 はパラメータの最尤推定値 $\hat{\theta}$ をまとめたものだが, 祝日効果・祝日前日効果の類似度を表すパラメータ $\hat{\alpha}, \hat{\beta}, \hat{\gamma}$ はそれぞれ 0.622, 0.374, 0.348 で, いずれも統計的に有意であることがわかる. この結果から, 平日の祝日は直前の日曜日の曜日効果に約 62% 類似していること, 平日の祝日前日は直前の金(土)曜日の曜日効果に約 37 (35)% 類似していること, 祝日効果より祝日前日効果の方が類似性が高い (62% < 72 (= 37 + 35)%) ことがわかる. なお, 祝日効果・祝日前日効果の類似度が際立って高くないのは, 当該効果はあくまで「平日における効果」であるため, 祝日後の仕事を考慮して外食や飲酒を控える社会人が多いためと考えられる.

また, 降水量に対する売上金額の回帰係数 $\hat{\delta}$ は -0.216 で統計的に有意であることから, 降水量が 1mm 増加すると (平均的には) 売上は約 216 円減少することがわかる. この値は一見小さいように思われるが, 1 日の降水量が 50mm であれば売上は約 10,800 円減少すること, 分析対象店舗の売上平均が約 50,000 円であることを鑑みれば, 決して軽視できない値といえる. さらに, 不規則成分の自己回帰係数 $\hat{\phi}_1, \hat{\phi}_2$ はどちらも統計的に有意でないことから, 不規則成分が売上に与える効果は当日以降の売上には影響しないと判断できる.

最後に, 傾向成分, 曜日効果成分, 不規則成分に付随する正規乱数の標準誤差 $\hat{\omega}, \hat{\varpi}, \hat{\sigma}$ を確認すると, それぞれ 0.787, 0.009, 33.412 であることから, もっとも変動が大きいのは不規則成分で, その次が傾向成分, もっとも変動が小さいのが曜日効果成分であることがわかる.

Table 8 パラメータ推定結果

$\hat{\alpha}$	$\hat{\beta}$	$\hat{\gamma}$	$\hat{\delta}$	$\hat{\phi}_1$	$\hat{\phi}_2$	$\hat{\omega}$	$\hat{\varpi}$	$\hat{\sigma}$
0.622* (0.004)	0.374* (0.004)	0.348* (0.004)	-0.216* (0.080)	0.001 (0.040)	0.000 (0.000)	0.787* (0.459)	0.009* (0.000)	33.412* (0.797)

注) 括弧内の数値は各パラメータの標準誤差である. また, * が付与されている値は, 有意水準 5% で統計的に有意 (5 パーセントailは ± 1.96) であることを意味している. また, $\hat{\delta}$ の単位は「千円/mm」, $\hat{\omega}, \hat{\varpi}, \hat{\sigma}$ の単位は「千円」, $\hat{\alpha}, \hat{\beta}, \hat{\gamma}, \hat{\phi}_1, \hat{\phi}_2$ は単位に依存しない. なお, 各パラメータの標準誤差の単位は, 対応するパラメータの単位と同一である.

4.2 潜在変数

次に、カルマンフィルタによって抽出された各成分について考察する。Fig. 3 上段は式 (2) で定義された傾向成分の時系列推移だが、当該成分は全体的には緩やかな上昇傾向を示していたことがわかる。具体的には、2016 年前期は 5 万円以下で推移することが多かったものの、夏に入ると 1 万円近く上昇し、2017 年には大半の時期で 5 万 5 千円程度の水準で推移していた。さらに、2018 年の夏には 6 万円から 6 万 5 千円近くまで上昇し、年末には 7 万円に手が届く所まで上昇していた。なお、2018 年 2 月上旬に傾向成分が急落しているが、これは 2018 年 2 月 5 日から 8 日にかけて福井県嶺北地方、石川県加賀地方を中心に降り積もった大雪^{#6}が主な原因と考えられる。この大雪により、学校の休校、企業の休業、灯油やガソリンなど生活物資の不足など、県民生活や観光、農林水産業をはじめとする経済活動全般が大きな打撃を受けた。分析対象店舗も 2 月 6 日から 16 日まで休業を余儀なくされるなど、深刻な被害を受けた。

Fig. 3 下段は式 (5) で定義された天候成分の時系列推移だが、時期によっては 1 万円から 2 万円近く売上を押し下げていたことがわかる。具体的には、2016 年の北陸地方は 3 月下旬から 5 月中旬を中心に高気圧に覆われて晴れた日が多かったため売上の低下は小さかったが、9 月は秋雨前線や気圧の谷の影響で曇りや雨の日が多かったため 1 万 5 千円近く売上を押し下げていた。一方、12 月は寒気の影響は少なく、降雪量が少なかったため、天候成分は 5 千円程度しか売上を押し下げなかった^{#7}。

2017 年 1 月から 2 月の北陸地方は、冬型の気圧配置が長続きせず暖冬・少雪だったこと、3 月から 5 月は高気圧に覆われて晴れた日が多かったため、天候成分による売上低下は小さかった。一方、7 月は日本海に停滞した活発な梅雨前線で、8 月は台風第 5 号の通過や湿った空気の影響で大雨となった日が多かったことから、1 万 5 千円近く売上が低下する時期が頻繁に発生していた。さらに、12 月は冬型の気圧配置の日が多く、雨または雪の日が多かったことから、長期的に 1 万円近く売上が低下する時期が続いた^{#8}。

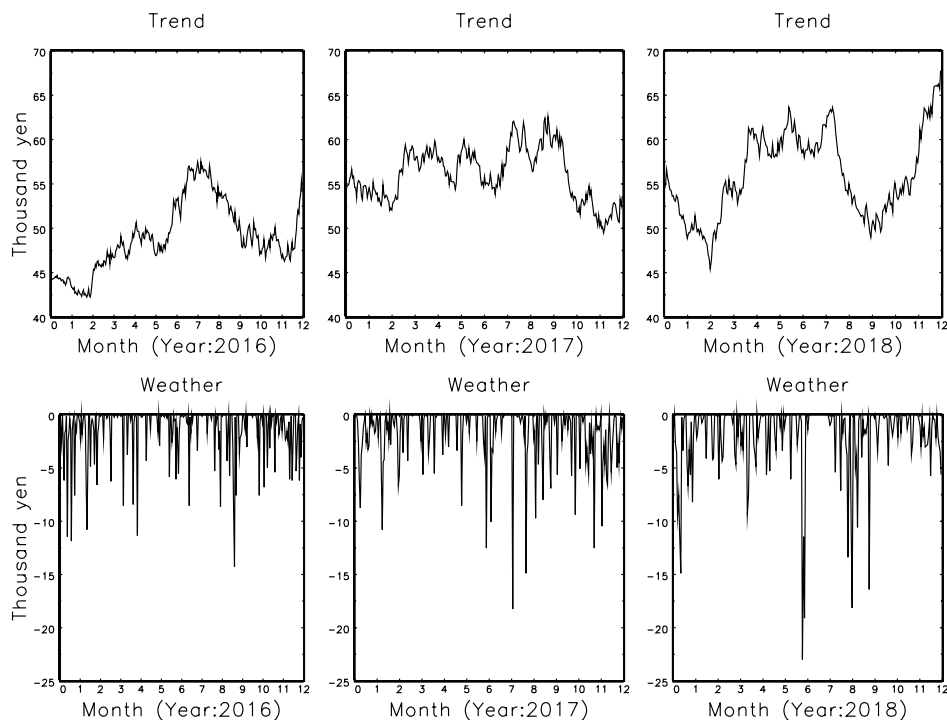


Fig. 3 推定結果 (上段：傾向成分, 下段：天候成分)

^{#6} 福井県内では 2 月 4 日から冬型の強い気圧配置となり、嶺北を中心に 5 日から 13 日にかけて大雪となった。福井市でも積雪が 147cm となり、昭和 56 年豪雪」以来 37 年ぶりに積雪 130cm を超える記録的な大雪となった。これにより、JR 北陸本線や越美北線などの公共交通機関は終日運休、北陸自動車道や中部縦貫自動車道は通行止めとなり、国道 8 号線では坂井市からあわら市にかけて約 1,500 台の車両が長時間にわたり滞留するなど県内特に嶺北において交通網が麻痺状態となった。なお、「平成 30 年福井豪雪」については松岡⁽¹⁰⁾を参照。

^{#7} 2016 年の北陸地方の天候経過の詳細は、新潟地方気象台⁽¹¹⁾を参照。

^{#8} 2017 年の北陸地方の天候経過の詳細は、新潟地方気象台⁽¹²⁾を参照。

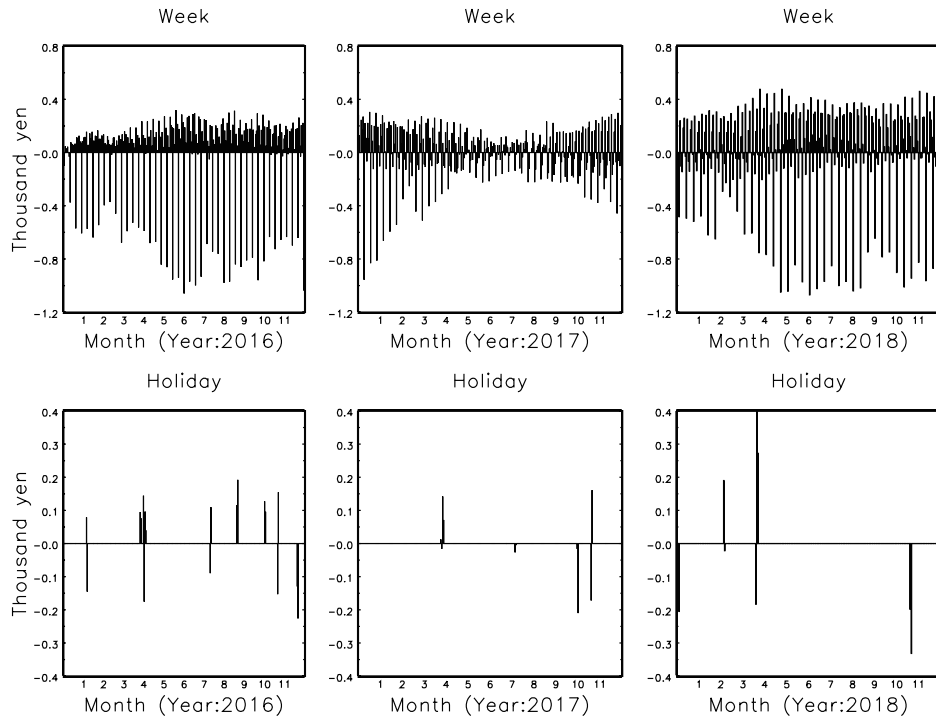


Fig. 4 推定結果 (上段：曜日効果成分, 下段：祝日および祝日前日効果成分)

2018 年 3 月から 6 月の北陸地方は、低気圧通過時に南から湿った空気が流れ込みやすかったため、かなりの多雨となり、天候成分が 2 万円近く売上を押し下げる時期があった。さらに、この年は北陸地方への台風の接近数が 6 個 (平年は 2 個から 3 個) と非常に多かったために 8 月から 9 月にかけて降水量が増加し、結果として 1 万 5 千円近く売上を押し下げていた^{#9}。

Fig. 4 上段は式 (4) で定義された曜日効果の時系列推移だが、千円ほど売上が低下する時期があるものの全体的に変動幅は小さく、その効果は限定的であることがわかる。Fig. 4 下段は式 (3) の第 2 項、第 3 項で定義された祝日および祝日前日効果の時系列推移だが、どちらの変動幅も極めて小さく、最大でも 400 円程度であることから、平日の祝日・祝日前日といった暦が分析対象店舗の売上に与える影響は極めて軽微であることが確認された^{#10}。

5. 結 論

本研究では、売上に寄与する要因の特定やその影響の度合いを測定する計量モデルを開発し、福井駅に隣接する飲食店の日次売上データに適用した。その結果、モデルを構成するパラメータや傾向成分、曜日効果成分、祝日・祝日前日成分、天候成分、不規則成分の推定に成功した。パラメータの推定結果から、平日の祝日は直前の日曜日の曜日効果に約 62% 類似していること、平日の祝日前日は直前の金 (土) 曜日の曜日効果に約 37 (35)% 類似しているなど、比較的高い類似性があること、降水量が 1mm 増加すると売上は約 216 円減少することなどが明らかとなった。また、分析対象店舗の売上トレンドは概して緩やかな上昇基調にあること、夏季 (6 月から 8 月) や年末は大きく上昇するが、秋季 (9 月から 11 月) は低下する傾向があることが確認された。さらに、天候成分は状況によっては 2 万円近く売上を押し下げることが判明した。一方、曜日や平日の祝日・祝日前日といった暦が売上に与える影響は極めて軽微であることが明らかとなった。このように、本研究で提案されたモデルを用いることで、従来は現場の勘といった類でしか語られてこなかった飲食店売上の特徴を客観的に把握することが可能となった。

一方、本研究で解決できなかった課題が残されていることも事実である。分析の結果、不規則成分の変動が売上変動の大部分を占めていること、当該成分には系列相関がみられないことが確認された。この結果は、本稿で提案された要因以

^{#9} 2018 年の北陸地方の天候経過の詳細は、新潟地方気象台⁽¹³⁾を参照。

^{#10} 紙幅の都合上、不規則成分の時系列推移を記載したグラフやその考察については割愛する。なお、当該箇所は著者に連絡することで入手できる。

外にも、分析対象店舗が提供している飲食物およびサービスの質や立地条件、近隣の競合店舗状況や商業施設でのイベント参加人数などといった要因も加味する必要があることを示唆している。しかしながら、そのような研究は質問紙調査やエリアマーケティング調査が必要となるため、今後の課題としたい。

謝 辞

本稿の執筆にあたり、野際大介先生(福島大学)、分析対象店舗経営者夫妻、1名の匿名レフェリーの方々より数多くの有益なコメントを頂きました。ここに記して感謝申し上げます。もちろん、本稿中にありうべき誤りはすべて筆者に帰するものです。なお、本研究は JSPS 科研費 JP18K01550 の助成を受けたものです。

文 献

- (1) 帝国データバンク“特別企画：「飲食店」の倒産、休廃業・解散動向調査(2018年度)”
<https://www.tdb.co.jp/report/watching/press/pdf/p191204.pdf> (参照日 2020 年 1 月 6 日)。
- (2) 気象庁“過去の気象データ・ダウンロード” <http://www.data.jma.go.jp/gmd/risk/obsdl/index.php> (参照日 2019 年 12 月 9 日)。
- (3) 山口類, 土屋映子, 樋口知之, “状態空間モデルを用いた飲食店売上の要因分解”, オペレーションズ・リサーチ: 経営の科学, Vol.49, No.5 (2004), pp.52-60。
- (4) 近藤文代, “デイリー POS データにおける曜日変動および値下げ効果の抽出”, オペレーションズ・リサーチ: 経営の科学, Vol.44, No.3 (1999), pp.154-163。
- (5) 高岡慎, 経済時系列と季節調整法, (2015), 朝倉書店。
- (6) C. Kim, “Dynamic linear models with Markov-switching”, *Journal of Econometrics*, Vol.60, No.1-2 (1994), pp.1-22。
- (7) P. K. Clark, “The Cyclical Component of U.S. Economic Activity”, *Quarterly Journal of Economics*, Vol.102, No.4 (1987), pp.797-814。
- (8) 樋口知之, 予測にいかす統計モデリングの基本—ベイズ統計入門から応用まで—, (2011), 講談社。
- (9) R. E. Kalman, “A New Approach to Linear Filtering and Prediction Problems”, *Journal of Basic Engineering*, Vol.82, No.1 (1960), pp.35-45。
- (10) 松岡孝恭, “平成 30 年 2 月の大雪が地域の鋳工業生産に与えた影響”, 福井県立大学経済経営研究, Vol.39, No.1, (2018), pp.1-27。
- (11) 新潟地方気象台, 2016 年の北陸地方の天候経過, <https://www.jma-net.go.jp/niigata/menu/kisetsu/2016year.pdf> (参照日 2019 年 12 月 14 日)。
- (12) 新潟地方気象台, 2017 年の北陸地方の天候経過, <https://www.jma-net.go.jp/niigata/menu/kisetsu/2017year.pdf> (参照日 2019 年 12 月 14 日)。
- (13) 新潟地方気象台, 2018 年の北陸地方の天候経過, <https://www.jma-net.go.jp/niigata/menu/kisetsu/2018year.pdf> (参照日 2019 年 12 月 14 日)。
- (14) 千葉賢, “レジーム転換を考慮した価格調整過程の分析”, 経営財務研究, Vol.39, No.1-2 (2019), pp.64-85。

付 録

本論でも示されているように、本稿で提案された要因分解モデルは、線形ガウス型状態空間モデル (9), (10), (11) によって定式化できる。これにより、Kalman⁽⁹⁾ が開発した逐次アルゴリズム (カルマンフィルタ, カルマンスムーザ) を用いて潜在変数の推定が可能となる。本研究では、以下の手順にしたがって潜在変数 \mathbf{x}_t と未知パラメータ $\boldsymbol{\theta}$ を推計した。

手順 1: カルマンフィルタ (濾波アルゴリズム)

カルマンフィルタは、1 期から t 期までの観測値集合 $\mathbf{y}_{1:t} = (\mathbf{y}'_1, \mathbf{y}'_2, \dots, \mathbf{y}'_t)'$ を用いて潜在変数 \mathbf{x}_t の濾波期待値 $\mathbf{x}_{t|t} = E(\mathbf{x}_t | \mathbf{y}_{1:t})$ 、濾波分散 $\mathbf{P}_{t|t} = E[(\mathbf{x}_t - \mathbf{x}_{t|t})(\mathbf{x}_t - \mathbf{x}_{t|t})' | \mathbf{y}_{1:t}]$ を計算することを目的としている。カルマンフィルタは、1 期から T 期まで以下の計算式を繰り返し実行する^{#11}。

$$\begin{aligned} \boldsymbol{\eta}_{t|t-1} &= \mathbf{y}_t - \mathbf{Z}_t \mathbf{x}_{t|t-1} - \mathbf{M}_t \mathbf{z}_t, & \mathbf{F}_{t|t-1} &= \mathbf{Z}_t \mathbf{P}_{t|t-1} \mathbf{Z}_t', & \mathbf{K}_t &= \mathbf{F} \mathbf{P}_{t|t-1} \mathbf{Z}_t' \mathbf{F}_{t|t-1}^{-1}, \\ \mathbf{x}_{t|t} &= \mathbf{x}_{t|t-1} + \mathbf{P}_{t|t-1} \mathbf{Z}_t' \mathbf{F}_{t|t-1}^{-1} \boldsymbol{\eta}_{t|t-1}, & \mathbf{P}_{t|t} &= \mathbf{P}_{t|t-1} - \mathbf{P}_{t|t-1} \mathbf{Z}_t' \mathbf{F}_{t|t-1}^{-1} \mathbf{Z}_t \mathbf{P}_{t|t-1}, & \mathbf{x}_{t+1|t} &= \mathbf{F} \mathbf{x}_{t|t}, & \mathbf{P}_{t+1|t} &= \mathbf{F} \mathbf{P}_{t|t} \mathbf{F}' + \mathbf{R} \mathbf{Q} \mathbf{R}' \end{aligned}$$

ここで、 $\boldsymbol{\eta}_{t|t-1}$, $\mathbf{F}_{t|t-1}$ は、 \mathbf{y}_t の予測誤差、予測誤差分散をそれぞれ意味する。また、 \mathbf{K}_t はカルマンゲインと呼ばれる。なお、上記アルゴリズム実行時に計算された成分は、最尤推定やカルマンスムーザ実行時に使用するため、計算機に保存しておく必要がある。

本稿では、対数尤度関数が予測誤差分解の形で表されることを利用して尤度を計算し、最尤法によりパラメータを推定する。対数尤度関数は、カルマンフィルタ実行時に計算される $\boldsymbol{\eta}_{t|t-1}$, $\mathbf{F}_{t|t-1}$ を用いて以下のように表すことができる。

$$\mathcal{L}(\boldsymbol{\theta}) = -\frac{T}{2} \log 2\pi - \frac{1}{2} \sum_{t=1}^T \log |\mathbf{F}_{t|t-1}| - \frac{1}{2} \sum_{t=1}^T \boldsymbol{\eta}_{t|t-1}' \mathbf{F}_{t|t-1}^{-1} \boldsymbol{\eta}_{t|t-1} \quad (12)$$

本稿では、非線形関数最適化アルゴリズム^{#12}を用いて対数尤度関数 (12) を最大化し、式 (12) が最大となる $\boldsymbol{\theta}$ を最尤推定値 $\hat{\boldsymbol{\theta}}$ とする。なお、最尤推定値 $\hat{\boldsymbol{\theta}}$ は、カルマンスムーザ実行時に使用するため、計算機に保存しておく必要がある。

手順 2: カルマンスムーザ (平滑化アルゴリズム)

カルマンスムーザは、1 期から T 期までの観測値集合 $\mathbf{y}_{1:T}$ を元に、 \mathbf{x}_t , \mathbf{v}_t の平滑化期待値 $\mathbf{x}_{t|T} = E(\mathbf{x}_t | \mathbf{y}_{1:T})$, $\mathbf{v}_{t|T} = E(\mathbf{v}_t | \mathbf{y}_{1:T})$ や平滑化分散 $\mathbf{P}_{t|T} = E[(\mathbf{x}_t - \mathbf{x}_{t|T})(\mathbf{x}_t - \mathbf{x}_{t|T})' | \mathbf{y}_{1:T}]$, $\mathbf{Q}_{t|T} = E[(\mathbf{v}_t - \mathbf{v}_{t|T})(\mathbf{v}_t - \mathbf{v}_{t|T})' | \mathbf{y}_{1:T}]$ を計算することを目的としている。カルマンスムーザは、カルマンフィルタ実行時に計算された $\boldsymbol{\eta}_{t|t-1}$, $\mathbf{F}_{t|t-1}$, \mathbf{K}_t , $\mathbf{x}_{t|t-1}$, $\mathbf{P}_{t|t-1}$ と最尤推定値 $\hat{\boldsymbol{\theta}}$ を用いて、 T 期から 1 期まで以下の計算式を繰り返し実行する^{#13}。

$$\begin{aligned} \mathbf{b}_t &= \mathbf{F}_{t|t-1}^{-1} \boldsymbol{\eta}_{t|t-1} - \mathbf{K}_t' \mathbf{r}_t, & \mathbf{D}_t &= \mathbf{F}_{t|t-1}^{-1} + \mathbf{K}_t' \mathbf{N}_t \mathbf{K}_t, & \mathbf{v}_{t|T} &= \mathbf{Q} \mathbf{R}' \mathbf{r}_t, & \mathbf{Q}_{t|T} &= \mathbf{Q} - \mathbf{Q} \mathbf{R}' \mathbf{N}_t \mathbf{R} \mathbf{Q}, \\ \mathbf{r}_{t-1} &= \mathbf{Z}_t' \mathbf{b}_t + \mathbf{F}' \mathbf{r}_t, & \mathbf{N}_{t-1} &= \mathbf{Z}_t' \mathbf{D}_t \mathbf{Z}_t + \mathbf{F}' \mathbf{N}_t \mathbf{F} - \mathbf{Z}_t' \mathbf{K}_t \mathbf{N}_t \mathbf{F} - \mathbf{F}' \mathbf{N}_t \mathbf{K}_t \mathbf{Z}_t, \\ \mathbf{x}_{t|T} &= \mathbf{x}_{t|t-1} + \mathbf{P}_{t|t-1} \mathbf{r}_{t-1}, & \mathbf{P}_{t|T} &= \mathbf{P}_{t|t-1} - \mathbf{P}_{t|t-1} \mathbf{N}_{t-1} \mathbf{P}_{t|t-1} \end{aligned}$$

カルマンスムーザによって計算された $\mathbf{x}_{t|T}$, $\mathbf{P}_{t|T}$ は、カルマンフィルタによって計算された $\mathbf{x}_{t|t}$, $\mathbf{P}_{t|t}$ より多くの情報を用いて計算されるため、推定精度が高いことが知られている。そのため、本稿ではカルマンスムーザの結果を用いて分析を行うこととした。

^{#11} カルマンフィルタを実行する際の初期値は、基本的には $\mathbf{x}_{1|0} = \mathbf{0}_9$, $\mathbf{P}_{1|0} = \mathbf{I}_9$ とした。ただし、 $\mathbf{x}_{1|0}$ の第 1 成分は分析対象期間初日 (2016 年 1 月 8 日) の売上金額、 $\mathbf{P}_{1|0}$ の第 1 行第 1 列成分は分析対象期間の売上金額の分散とした。

^{#12} 最尤推定に使用したソフトウェアおよびライブラリは GAUSS 19 と MAXLIK 5.0 である。また、非線形関数最適化アルゴリズムは BFGS 法を使用した。なお、最尤法を実行する際の初期値は、 $\alpha^{(0)} = 0.5$, $\beta^{(0)} = 0.25$, $\gamma^{(0)} = 0.25$, $\delta^{(0)} = 0$, $\phi_1^{(0)} = 0$, $\phi_2^{(0)} = 0$, $\omega^{(0)} = 1$, $\varpi^{(0)} = 1$, $\sigma^{(0)} = 33.989$ とした。また、パラメータ空間 $\boldsymbol{\Theta}$ に制約がある場合はデルタ法を用いて最適化を行っている。デルタ法については千葉⁽¹⁴⁾ の Web 補遺 (<http://jfa.main.jp/journal/journal-backnumber.html>) を参照。

^{#13} カルマンスムーザを実行する際の初期値は、 $\mathbf{r}_T = \mathbf{0}_9$, $\mathbf{N}_T = \mathbf{O}_{9,9}$ とした。