

日次データへの季節調整の適用

— 家計調査による実例 —

有田 帝馬[†]Application of Seasonal Adjustment to Daily Data
- Example of the Family Income and Expenditure Survey -

ARITA Tetsuma

コンピュータの能力向上や ICT サービスの普及を背景に、高頻度データがビッグデータの 1 つとして活用され始めている。この高頻度データの活用にあたり、週周期等の時系列的な特性への対処が課題となっている。

この課題に対処するため、本研究では日次データに適用できるように修正した季節調整プログラム (X-12-ARIMA) を構築し、家計調査の日次データに対し季節調整を行った。結果、Regression ARIMA モデルの推計等の事前調整、X-11 の手順によるトレンド、季節変動、不規則変動の分離、いずれも良好な結果が得られた。

本研究は、本邦の経済統計の日次データに対し X-12-ARIMA 等の季節調整法を適用した初めての分析事例と考えられる。そして本研究で用いた手法は、課題であった週周期等の時系列的な特性に対処する上で、有力な手法になると考えられる。

キーワード：時系列分析、季節調整、X-12-ARIMA、日次データ、高頻度データ

Reflecting the improvement of computer capabilities and the spread of ICT services, high-frequency data is beginning to be used as one type of big data. A key issue in utilizing high frequency data is to deal with time-series characteristics such as the weekly periodicity.

In order to address the issue, this study modified a seasonal adjustment program (X-12-ARIMA) to apply to daily data and applied the program to the daily data of family income and expenditure survey. The preliminary adjustments such as the estimation of the Regression ARIMA model, and the separation of trend, seasonal, and random fluctuations by the X-11 procedure both showed reliable results.

This study firstly successes to apply the seasonal adjustment method such as X-12-ARIMA to the daily data of Japanese economic statistics. The method used in this study is expected to be a viable method for addressing the issues to deal with time-series characteristics of high-frequency data such as the weekly periodicity.

Key Words: Time series analysis, Seasonal adjustment, X-12-ARIMA, Daily data, High-frequency data

[†] 日本銀行金融機構局 E-mail: tetsuma-arita@rio.odn.nc.jp

1. はじめに

コンピュータの計算処理や通信処理の能力向上、インターネットを中心とした ICT サービス普及は、日次や時次等の高頻度データが集計・公表され始めるなど、これまで四半期や月次データが中心だった公的統計の集計にも影響を及ぼしている。例えば家計調査や各種金融統計（外国為替、金利、株価）では、日次データが集計・公表されるようになってきている。また身近な例としては、新型コロナウイルスの関連データも日次で公表されており、感染者数等の集計値や7日前比等が報道されている。この他、電力使用実績データ、移動通信やインターネットのトラフィックデータ等では、より高頻度の時次データが公表されている。また民間でも、総務省（2020）によれば企業でも IoT の導入の進捗を映し POS データ、e コマースによる販売記録、アクセスログ等から得られたデータの収集・活用が大きく進展しており、データの質（多様性、粒度、頻度等）の整備への取り組みを検討する先が増加している。

このように収集・公表された高頻度データは、ビッグデータまたは非伝統的データの 1 つとして活用され始めている。活手法として小寺・藤田ほか（2018）は、POS データ等の速報性の高い日次データに機械学習の手法を活用した消費分析を加える手法により、経済分析の精度と速報性を向上できる可能性が高いことを示している。また藤代（2020）は、プラットフォームが公開する高頻度データが、リアルタイムで経済活動を把握するうえで有用であることを指摘し、人々の移動に関する日次データがリアルタイムに経済活動の動きを映じていることを示している。海外でも Aladangady（2019）が、米国の決済処理サービス会社（Fiserv）の匿名化された日次決済データを元に消費支出を推計する手法を提案している。また Lewis et al.（2020）は、小売売上高、消費者心理指数、給与所得源泉徴収額、鉄道輸送量、電力消費量などの 10 種類の日次または週次データを用い、米国の経済状況をほぼリアルタイムに把握できる指数（WBI: a weekly economic index）を提案し、このモデルに基づいて集計された指数が週次でアップデートされている¹。

このように国内外で高頻度データの活用が広がり始める中、データに含まれる季節性や曜日要因等、時系列的な特性への考慮の必要性が指摘されている。2020 年 11 月に開催された東京大学金融教育研究センター・日本銀行調査統計局共催ビッグデータフォーラムのパネルディスカッションにおいて、パネリストが非伝統的データの短所の 1 つとして「季節性・曜日要因などの時系列的な特性が明らかでない」ことを挙げ、「こういった短所は、ユーザーが考慮しなければならないが、その認識は十分に進んでいないのではないかと指摘している。また若田部（2020）は、高頻度データの利用における課題の一つとして季節調整がなされていないことを挙げている。こうした中、株式会社インテージリサーチ（2020）は経済産業省委託調査において、13 の主要国・機関²におけるビッグデータを公的統計へ適用する際の時系列的な特性への考慮（季節調整）に関する文献調査を行ったが、いずれの国・機関ともビッグデータを公的統計に活用する際にどのように季節調整を行っているかについては確認できない結果となった。

ビッグデータ、特に高頻度データの活用にあたり時系列的な特性への考慮が課題となる要因として季節調整法のプログラムの制約があげられる。本邦では経済統計を中心に米・センサス局で開発された X-12-ARIMA が季節調整法として利用されており、海外でも X-12-ARIMA や後継の X-13ARIMA-SEATS（以下、X-12-ARIMA 等）、またはスペイン中央銀行で開発

¹ WBI のアップデート値は、セントルイス連銀の FRED（Federal Reserve Economic Data）の <https://fred.stlouisfed.org/series/WBI> に掲載されている。

² 調査対象国・機関は、イギリス、オランダ、デンマーク、フィンランド、シンガポール、カナダ、アメリカ、オーストラリア、ニュージーランド、中国、ユーロスタット、OECD および国連。

された TRAMO-SEATS が季節調整法として主に利用されている。これらの季節調整法のプログラムは、年周期の季節性を持つ四半期または月次データを処理することを想定としているため、処理できる周期は 4 または 12 に限定されている。しかし高頻度データの多くが日次または時次であり、それぞれ週単位の季節性（周期 7）や日周期の季節性（周期 24）を有することが多い。これらの周期には X-12-ARIMA 等の季節調整のプログラムでは対処できない点が、高頻度データの時系列的な特性を考慮する上での制約になっている。

高頻度データに対する季節調整に関する先行研究として Ladiray et al. (2018) が挙げられる。この研究では、季節調整の中で用いられているノンパラメトリックおよびパラメトリックの手順を週周期や年周期の複数の季節性が含まれる高頻度データに適用するために必要となる改良について検討を行い、実際に 1996 年から 2016 年までのフランスの電力消費量データに対し改良した手順を適用することで、高頻度データに対しても主要な季節調整法（TRAMO-SEATS、X-12-ARIMA 等）が適用可能であることを示した。ただし X-12-ARIMA 等に関しては、本来、フィルター長（各種の移動平均の項数）をデータの性質により定めているところ、改良した手順では固定長とした点に改善余地があるとしている。また有田（2021）では、日次データである新型コロナウイルス感染者数に対し X-12-ARIMA 等と同じロジックを用いて季節調整を行っているが、この研究でもフィルター長を固定している。

本研究では、日次データを処理可能とした上でデータの性質に応じフィルター長を定める手順を組み入れる改良を加えた X-12-ARIMA プログラムを準備し、これを用いて高頻度データの一つである家計調査（日次データ）に季節調整を行う。これにより、高頻度データの週単位の季節性等の時系列的な特性を考慮する手法の適用事例を示すことを本研究の目的とする。先行研究との関係では、フィルター長の決定プロセスを改善している点に新規性がある。ただし Ladiray et al. (2018) では複数の季節性を同時に考慮しているのに対し、本研究では考慮対象を週周期の季節性に絞っている点には留意が必要である。また本邦の経済統計の日次データに X-12-ARIMA 等の季節調整法の適用を試みた先行研究は確認できず、本研究が初めての分析事例と考えられる。

本稿の構成は以下の通りである。まず「2. 本研究の手法」において、本研究で用いた日次データに対し季節調整を行う手法を紹介するとともに、本研究で用いた分析対象データ（家計調査の日次データ）についても概説する。そして「3. 分析の結果」では、本研究の手法により季節調整を行った結果を示し、得られた季節変動等の分離結果の検証を行う。最後に「4. 考察」では、本研究で得られた結果の位置づけと本研究の今後の課題について述べる。

2. 本研究の手法

2.1 季節調整用プログラムの修正

本研究では、有田（2021）が新型コロナウイルス感染者数の分析の際に用いた X-12-ARIMA 等のプログラムにフィルター長の決定に関する改良を加えたものを利用した。X-12-ARIMA 等は、事前調整パートと季節調整パートの 2 つのサブプログラムに分かれることから、以下、改良したプログラムの概要をサブプログラム別に示す。

2.1.1 事前調整パート

事前調整パートは、季節調整の対象データ（以下、対象データ）から以下の式(1)で表現される Regression ARIMA モデルを推計し、このモデルを用いて、対象データの異常値等（外れ値や休日に伴う変動）を一時的に除去すると同時に、対象データの端

の期間の季節調整の精度を上げるための予測値を推計するサブプログラムである。事前調整パートでは、このほか調整前後のデータの季節性有無を確認する機能³も有している。

$$\phi(B)\Phi(B^s)(1-B)^d(1-B^s)^D\left(y_t - \sum_i \beta_i x_{it}\right) = \theta(B)\Theta(B^s)a_t \quad (1)$$

B:バックシフト・オペレータ、s:周期、d:階差の次数、D:季節階差の次数、
 $\phi(B)$:自己相関オペレータ、 $\Phi(B^s)$:季節自己相関オペレータ、
 $\theta(B)$:移動平均オペレータ、 $\Theta(B^s)$:季節移動平均オペレータ、
 $a_t \sim i.i.d(0, \sigma)$ 、 y_t :対象データ、 x_{it} :回帰変数、 β_i :回帰パラメータ

事前調整パートは、データの周期が12以下の任意の整数であれば実行可能であることから、X-12-ARIMA等のサブプログラムをそのまま利用可能である。もっとも、高頻度データを取り扱うためデータ数が増大し、X-12-ARIMA等のプログラムの上限値(780)を超えてしまう。このため本研究ではX-12-ARIMA等のUnix版を準備し、サブプログラムのパラメータを修正し再コンパイルすることで、データ数の上限値を1120(日次データで約3年間)に引き上げている。ちなみに13以上の周期の季節性を有するデータ(例えば日単位の周期を持つ時次データ)を扱う場合も、Unix版を準備し周期を定めるパラメータを修正し再コンパイルすることで対処可能である。

2.1.2 季節調整パートの概要

季節調整パートは、X-11の手順により事前調整済の対象データをトレンド、季節変動および不規則変動に分解し、事前調整パートで一時的に除去した異常値等(外れ値や休日に伴う変動)を、分解した各変動に戻す⁴サブプログラムである。

通常のX-12-ARIMA等の季節調整パートのサブプログラムでは周期4または12のデータしか処理できない。そして周期7への対応には、以下に述べる移動平均等のロジックについてプログラムを作り直す必要がある。このため、本研究ではFindley et al. (1998)、Dagum (1999)およびScott et al. (2007)で示されているX-11の手順をベースに、周期7のデータに対応できるサブプログラムをFortran 90でゼロから構築し、これを季節調整パートとして利用した。本研究で用いた、周期7に対応したX-11の処理のフローチャートを表1に示す。

2.2 X-11の手順における移動平均の扱い

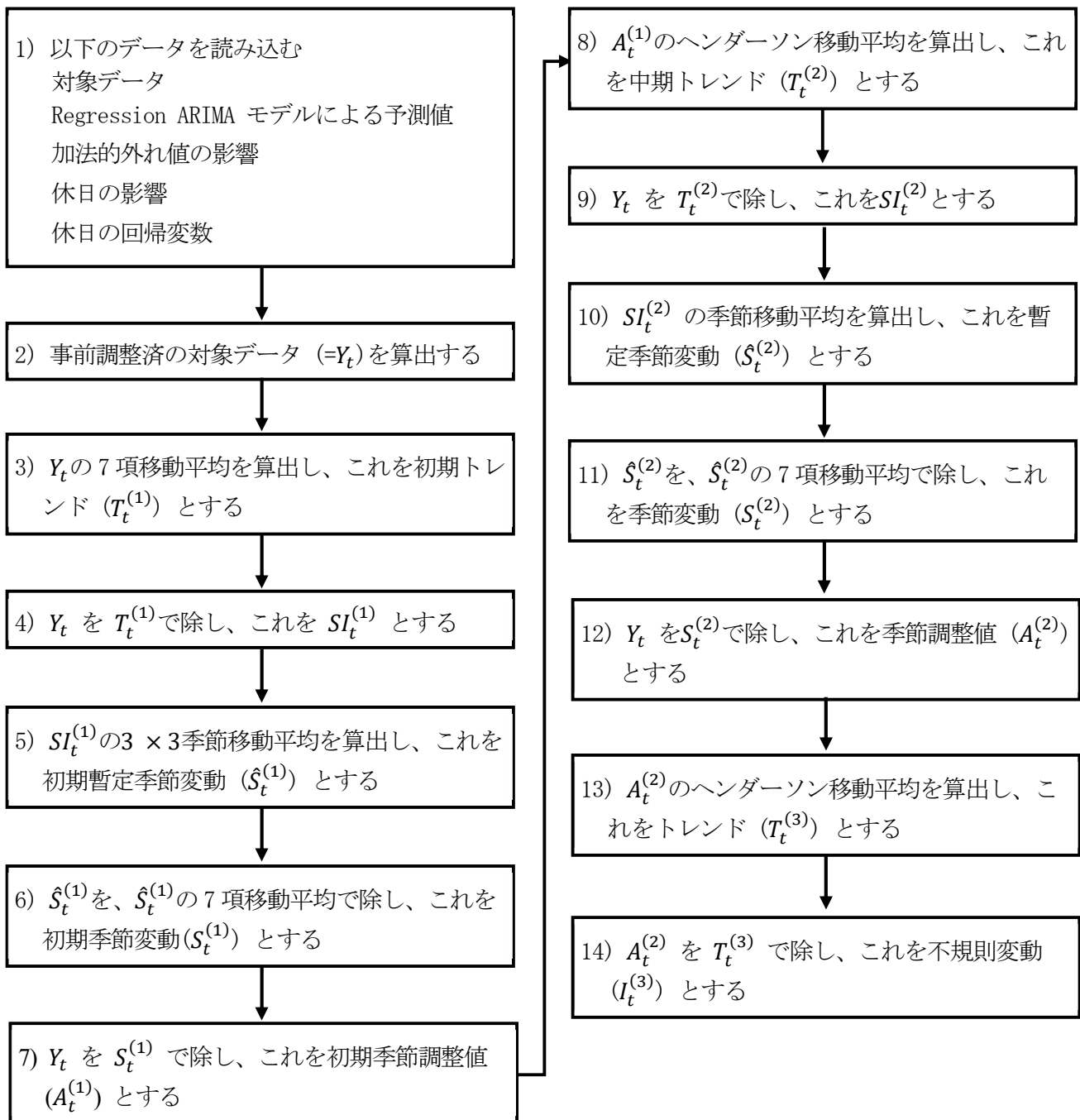
X-11の手順では、事前調整済の対象データをトレンド、季節変動、不規則変動に分解するために、(単純な)移動平均、季節移動平均、ヘンダーソン移動平均の3種類の移動平均を用いる。ここでは周期7への対応にあたり、これらの移動平均をどの様に扱ったかを示す。

³ QS統計検定(QS statistic for seasonality)と呼ばれる、帰無仮説を「時系列データに季節性がない」とする統計検定を行う機能が準備されている。

⁴ 一時的に除去した異常値等は、それぞれ以下の変動に戻される。

一時的に除去した異常値等	戻される変動
加法的な外れ値および減衰的な外れ値	不規則変動
水準変化	トレンド
休日に伴う変動	季節変動

表1 周期7に対応した季節調整パートのフローチャート



注1 変数の表記は Findley et al. (1998)に沿っている。

注2 8) と 13) で用いるヘンダーソン移動平均の項数は、Findley et al. (1998) と Dagum(1999) に沿い、不規則変動とトレンドについて、それぞれの1期前対比での変化率の絶対値の平均値を求め、その比を基準に定めている。

注3 10) で用いる季節移動平均の項数は、Bee Dagum(1999)、Scott et al.(2007) および U.S. Census Bureau (2017) に沿い、不規則変動と季節変動について、それぞれの1周期前対比での変化率の平均値を求め、その値の絶対値の比を基準に定めている。

2.2.1 (単純な) 移動平均 (表1の3)

X-11の手順では、季節性の周期を項数とする(単純な)移動平均を用い、事前調整済みの対象データ(Y_t)から季節変動と不規則変動を除去することで、初期トレンド($T_t^{(1)}$)を抽出する。ただし周期4または12は項数が偶数のため、代表時点を整数に修正するために2つの移動平均を重ねた、以下の式(2)または(3)で示される中心化を行った移動平均を用いている。

$$\text{周期4の中心化を行った移動平均: } T_t^{(1)} = \frac{Y_{t-2} + 2Y_{t-1} + 2Y_t + 2Y_{t+1} + Y_{t+2}}{4 \times 2} \quad (2)$$

$$\text{周期12の中心化を行った移動平均: } T_t^{(1)} = \frac{Y_{t-6} + 2Y_{t-5} + \dots + 2Y_{t+5} + Y_{t+6}}{12 \times 2} \quad (3)$$

周期7の場合、項数は奇数であり中心化を行う必要がないため、以下の式(4)で示される単純な移動平均を用いた。

$$\text{周期7の移動平均: } T_t^{(1)} = \frac{Y_{t-3} + Y_{t-2} + \dots + Y_{t+2} + Y_{t+3}}{7} \quad (4)$$

この修正の影響を考察する。まず季節変動の除去の機能は、季節性は周期単位の平均で均される性質を利用して季節性を除去する点で不変であることから、修正前と同程度の機能度が期待できる。次に不規則変動の除去の機能は、修正前後とも大数の法則を利用する点で共通であり、不規則変動の除去の機能度は移動平均の項数次第となる。この点、修正後の移動平均の項数は7であることから、周期12の中心化を行った移動平均(項数13)には劣るが、周期4の中心化を行った移動平均(項数4)以上の機能度は期待できる。

2.2.2 季節移動平均 (表1の5および10)

X-11の手順では、同一期⁵別のトレンド除去後の季節・不規則変動($SI_t^{(1)}$ や $SI_t^{(2)}$)の加重移動平均(季節移動平均)を算出することで、暫定季節変動($\hat{S}_t^{(1)}$ や $\hat{S}_t^{(2)}$)の抽出と不規則変動の除去を行う。周期7への対応にあたっては、周期12で用いる以下の式(5)、(6)、(7)で表される季節移動平均の k (周期)を周期7にしたものを用いた⁶。

$$\hat{S}_t = \frac{1}{3 \times 3} (SI_{t-k \times 2} + 2SI_{t-k} + 3SI_t + 2SI_{t+k} + SI_{t+k \times 2}) \quad (5)$$

$$\hat{S}_t = \frac{1}{3 \times 5} (SI_{t-k \times 3} + 2SI_{t-k \times 2} + 3SI_{t-k} + 3SI_t + 3SI_{t+k} + 2SI_{t+k \times 2} + SI_{t+k \times 3}) \quad (6)$$

$$\hat{S}_t = \frac{1}{3 \times 9} (SI_{t-k \times 5} + 2SI_{t-k \times 4} + 3SI_{t-k \times 3} + 3SI_{t-k \times 2} + 3SI_{t-k} + 3SI_t + 3SI_{t+k} + 3SI_{t+k \times 2} + 3SI_{t+k \times 3} + 2SI_{t+k \times 4} + SI_{t+k \times 5}) \quad (7)$$

\hat{S}_t : 暫定季節変動、 SI_t : 季節・不規則変動、 k : 周期

⁵ 例えば週周期の日次データだと同じ曜日、年周期の月次データだと同じ月のデータが同一期になる。

⁶ 季節移動平均の選択も周期12の扱いと同様、以下の様に行っている。詳細はU. S. Census Bureau (2017)およびDagum, E. B. (1999)参照。

初期暫定季節変動の抽出時：式(5)を選択

暫定季節変動の抽出時： $SI_t^{(2)}$ に式(5)を利用し仮の暫定季節変動と暫定不規則変動を抽出し、それぞれの1周期前対比での変化率の平均値を求め、その値の絶対値の比を基準に利用する季節移動平均を選択する。仮の暫定不規則変動の1周期前対比での変化率の平均値の割合が高い場合、項数の多い季節移動平均を選択する。逆に仮の暫定季節変動の割合が高い場合、季節変動の変化を捉えるため項数の少ない季節移動平均を選択する。

式(5)、(6)、(7)は、周期7への対応後も、季節・不規則変動のデータを同一期別に分けて処理しており、周期が12から7への置き換えに対応できていると評価できる。また加重移動平均の形式も同じであることから、季節変動の抽出と不規則変動の除去の機能度も、周期12と同程度を期待できると考えられる。

ただし季節移動平均の種類は、データの性質（季節変動と不規則変動の変化率の割合）によるが、この点、周期7のデータの性質の知見は限定的である。そこで本研究では、一旦周期12と同じ3種類の季節移動平均を用いて分析を行い、得られた季節変動と不規則変動が適切に分離されているか周波数分析で検証する。

2.2.3 ヘンダーソン移動平均（表1の8および13）

ヘンダーソン移動平均は、以下の式(8)で表される、季節性を含まないデータから3次曲線の形でトレンドを抽出するのに適するようにウェイト付け⁷した加重移動平均である。

$$T_t = \sum_{j=-H}^H h_j^{2H+1} A_{t+j} \quad (8)$$

T_t :トレンド、 $2H + 1$:ヘンダーソン移動平均の項数、

h_j^{2H+1} :ヘンダーソン移動平均のウェイト、 A_{t+j} :季節調整値

X-11の手順では、季節調整値($A_t^{(1)}$ や $A_t^{(2)}$)からトレンド($T_t^{(2)}$ や $T_t^{(3)}$)を抽出し不規則変動を除去するために、周期4の場合は5、7項、周期12の場合は9、13、23項のヘンダーソン移動平均を用いる。周期7への対応にあたっては、周期12と同様に9、13または23項のヘンダーソン移動平均を利用した⁸。

ヘンダーソン移動平均の対象とする季節調整値($A_t^{(1)}$ や $A_t^{(2)}$)は季節性が除去されており、式(8)には季節性の周期が含まれないことから、季節性の周期を12から7に置き換えた場合でも、式(8)はそのまま利用できる。

ただしヘンダーソン移動平均の項数は、データの性質（トレンドと不規則変動の変化率の割合）によるため、9、13、23項のヘンダーソン移動平均だけでは不十分である可能性がある。本研究では、まず周期12と同じ9、13、23項ヘンダーソン移動平均を用いて分析を行い、得られたトレンドと不規則変動が適切に分離されているか周波数分析で検証する。

2.3 使用データ

本研究では、家計調査（二人以上の世帯のうち勤労者世帯）の品目分類別・日別の支出データを用いた。データ期間は、新型コロナウイルス感染症およびそれへの対応が家計支出に与えた影響を確認するため、感染症が拡大する1年前の2019/1月から2021/5

⁷ ヘンダーソン移動平均のウェイトはShiskin et al. (1967)およびDagum(1999)を参照。

⁸ ヘンダーソン移動平均の項数の選択も周期12の扱いと同様、以下の様に行っている。詳細はFindley et al. (1998)およびDagum(1999)を参照。

季節調整値($A_t^{(1)}$ や $A_t^{(2)}$)に13項ヘンダーソン移動平均を利用し仮の暫定トレンドと暫定不規則変動を抽出し、それぞれの1期前対比での変化率の絶対値の平均値を求め、その比を基準に定める。仮の暫定不規則変動の割合が高い場合、項数の多いヘンダーソン移動平均を選択する。逆に仮の暫定トレンドの割合が高い場合、トレンドの変化を捉えるため項数の少ないヘンダーソン移動平均を選択する。

月までとした。なお品目分類別・日別の支出データは、月極め払いなどが多い品目⁹が除かれているほか、日付不祥のデータも本研究の対象から除いている。

本研究では、品目分類の支出データのうち消費支出と十大費目のうち食料およびその内訳の一部（一般外食、飲酒代）を分析対象とした。なお飲酒代については、2020年4月22日および同年5月7日のデータは0となっているが、0以下のデータが含まれると事前調整パート内で行う対数変換に支障が出ることから、本研究ではこの2データを0.01に置き換えている。このほか、事前調整パートで用いる休日の影響を一時的に除去するための回帰変数は、有田（2012）に沿って以下の手順で作成した。

- 1) 内閣府が公表する国民の祝日に関するデータ¹⁰を元に土日以外の休日を洗い出し、休日を1、それ以外の日を0とした時系列データを作成
- 2) 作成した時系列データの各曜日の平均値を算出し、週単位の周期性を抽出
- 3) 1)の時系列データから、2)の各曜日の平均値を減ずることで、週単位の周期性を除去。これを、休日の影響を除去するための回帰変数とした

3. 分析の結果

3.1 事前調整パートの結果

事前調整パートでは、はじめに前述の式(1)で表現される Regression ARIMA モデルの推計を行う。モデルの推計にあたっては、支出データに分散不均質性が存在することから、対数変換により、分散不均質性がモデル推計に及ぼす影響を緩和している。

一時的に除去する異常値等は、支出データの形状から洗い出した上で、X-12-ARIMA 内の outlier コマンドで洗い出し漏れがないか確認している。また ARIMA モデルの次数は、コレログラムを用いて暫定の次数を定めている。その上で、異常値等と ARIMA モデルの次数は相互に影響するため、両者の推計と修正を繰り返し Regression-ARIMA モデルを決定した。

表2に、支出データに対する週単位の季節性の有無の検定（以下、QS 統計検定）結果と、季節性を有する品目の Regression ARIMA モデルの推計結果を示す。

表2 Regression ARIMA モデルと季節性の有無の確認結果

	季節性 有無 (注1)	Regression ARIMA モデル(注2)				
		ARIMA モデル	休日	回帰変数		
				外れ値等(注3, 4)		
				AO	TC	LS
消費支出	あり (0.000)	(1 1 1) (0 1 1)	あり	あり	あり	あり
食料		(0 1 1) (0 1 1)				
一般外食		(0 1 2) (1 1 1)				
飲酒代		(1 1 1) (1 1 1)	なし			

注1 「季節性の有無」は、帰無仮説「時系列データに季節性がない」が5%の有意水準で棄却される場合を季節性「あり」とした。括弧は、帰無仮説棄却時のP値を示す。

注2 ARIMA モデルの次数と回帰変数はすべて、t 値が絶対値2以上で有意となっている。

注3 AO は加法的な外れ値、TC は減衰的な外れ値、LS は水準変化の回帰変数を示す。

⁹ 日別の支出データから月極め払いなどが多い品目（例：学校給食、公営家賃）が除かれている。

¹⁰ 内閣府「国民の祝日について」（<https://www8.cao.go.jp/chosei/shukujitsu/gaiyou.html>）に掲載されている。

注4 外れ値等の回帰変数を用いた年月日は「付表1 Regression ARIMA モデルにおける外れ値等の回帰変数」を参照。

表2の「季節性の有無」の列をみると、今回の対象データ（消費支出、食料、一般外食および飲酒代）の全てについて、週単位の季節性を有することがわかる。これにより対象データに季節調整を行うことの正当性が確認できる。

次に表2のRegression ARIMA モデルのうち「ARIMA モデル」の列をみると、ARIMA モデルの階差、季節階差は全て1、各モデルの次数は0~2であり、経済統計の月次データや四半期データでRegression ARIMA モデルを推計する場合に得られる階差や次数と同様の傾向がみられた。また「回帰変数」の列をみると、飲酒代を除き休日の回帰変数が有意となったほか、有意となる外れ値等の存在も確認された。付表1に具体的な外れ値等の回帰変数を示している。なお外れ値等の数については、データ数（882）とボラティリティの大きさを考えると違和感のない水準であった。

このRegression ARIMA モデルの推計結果を用い、対象データから異常値等（外れ値や休日に伴う変動）を一時的に除去し、対象データの端に21日分の予測値を付加したものを事前調整済のデータとした。

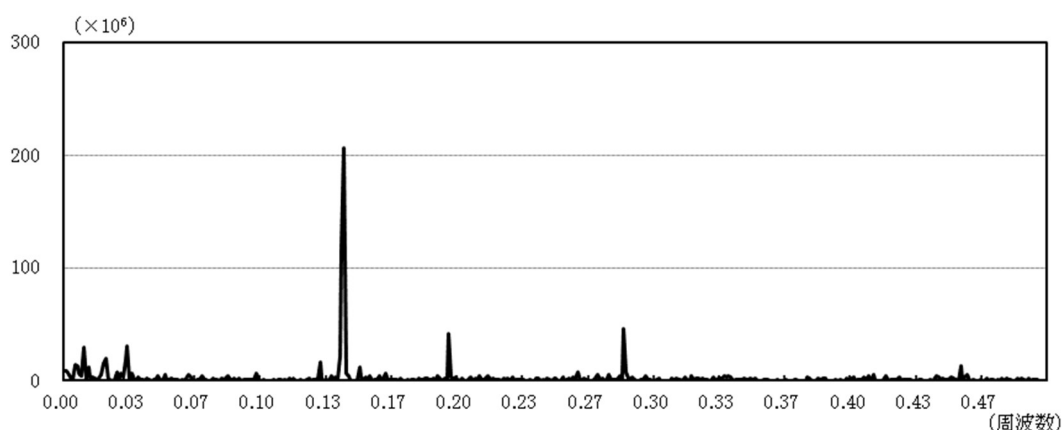
3.2 季節調整パートの結果

季節調整パートでは「2.1.2 季節調整パートの概要」で示したX-11の手順で事前調整済のデータをトレンド、季節変動、不規則変動に分離し、「3.1 事前調整パートの結果」で一時的に除去した異常値等を各変動に戻すことで最終的な分離結果を得た。以下では事前調整前のデータと分離結果に対し周波数分析等を行うことで、各変動の分離の妥当性の検証を行う。

3.2.1 事前調整前のデータ

はじめに、図1および付図1に事前調整前の消費支出および他の対象データ（食料、一般外食および飲酒代）にピリオドグラム法による周波数分析を行った結果を示す。

図1 消費支出の事前調整前のデータの周波数分析結果



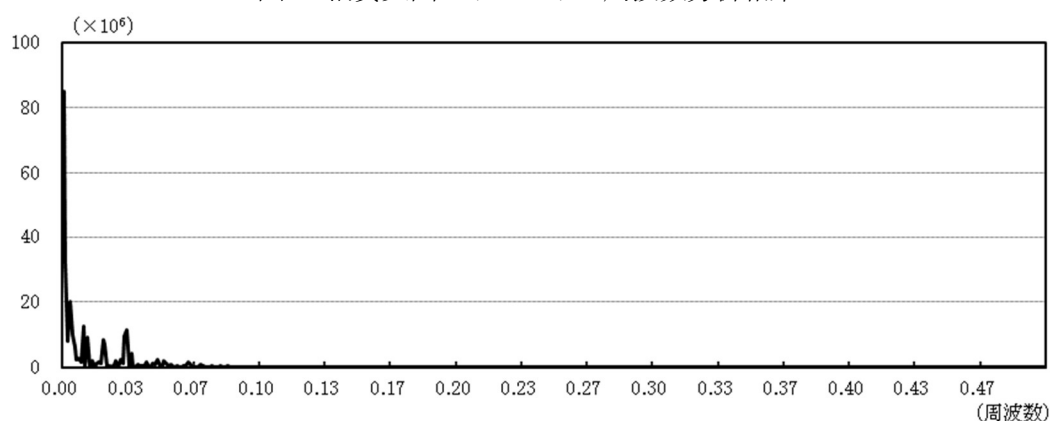
結果をみると、どの事前調整前のデータにも周波数0.14（周期7日）に最も強いパワースペクトルがみられ、週単位の周期変動が卓越していることが確認できる。また周波数が0.14未満（週より長周期）の領域では、飲酒代の周波数0.002~0.003（周期1年）に週周期の半分程度のパワースペクトルが存在する以外に、強いパワースペ

クトルはみられなかった。本研究は、改良を加えた X-12-ARIMA プログラムを用い、家計調査の支出データを、最も強い週の周期性、トレンド、不規則変動の3つの変動要素に分解するものと整理できる。

3.2.2 トレンド

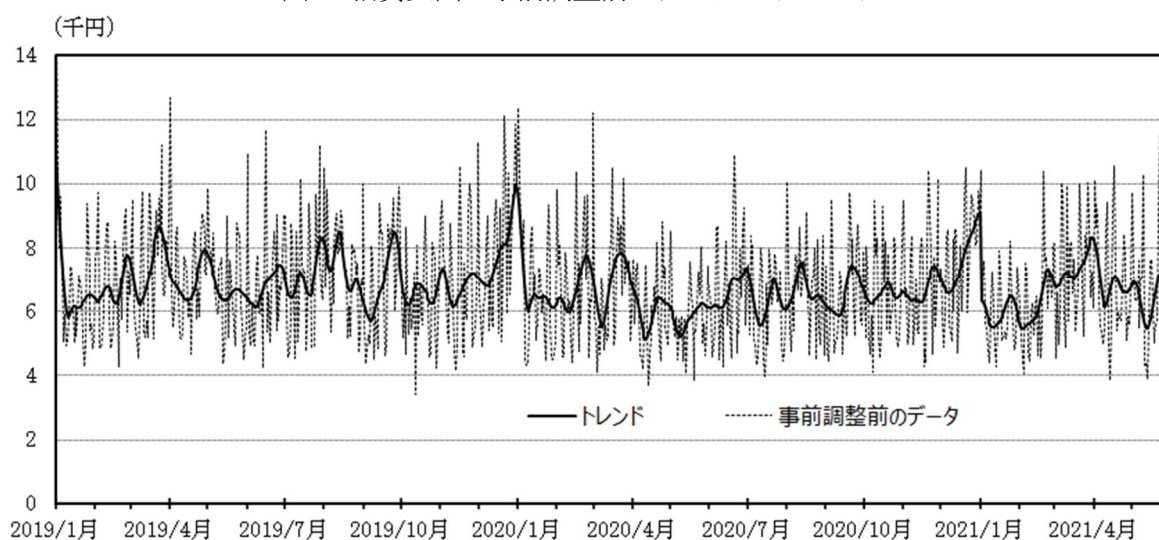
はじめに、図2および付図1に消費支出等のデータから分離したトレンドに対し周波数分析を行った結果を示す。いずれのデータとも周波数が小さい領域のパワースペクトルのみとなっており、トレンドの抽出結果は良好と考えられる。

図2 消費支出のトレンドの周波数分析結果



つぎに図3および付図2に、消費支出および他の対象データ（食料、一般外食および飲酒代）のトレンドと事前調整前のデータを比較した結果を示す。こちらも、視覚的に事前調整前のデータにみられる週単位の季節性が除去されていることが確認できる。さらに QS 統計検定を用いて、分離後のトレンドに週単位の季節性が残っていないか検証したが、こちらも視覚的な確認結果と同様、「時系列データに季節性はない」結果であった。

図3 消費支出の事前調整前のデータとトレンド

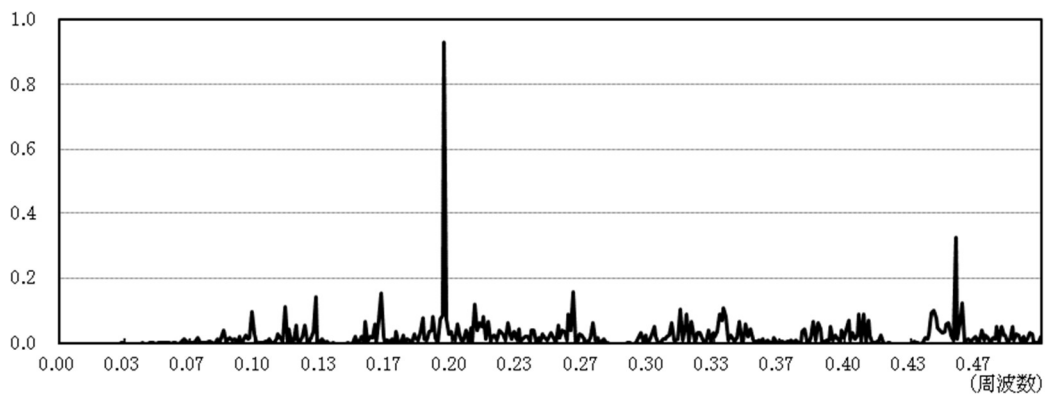


3.2.3 不規則変動

まず不規則変動に対し周波数分析を行った結果を図4および付図1に示す。結果をみると、食料等において周波数0.1近傍(周期10日前後)にパワースペクトルが残る以外は、周波数の小さい領域のパワースペクトルは除去されており、不規則変動の抽出結果は概ね良好だったと考えられる。

この他、消費支出の不規則変動の中に周波数0.2(周期5日)の領域に強いパワースペクトルがみられるが、これは五十日(ごとうび)が生み出す周期性の可能性が考えられる。

図4 消費支出の不規則変動の周波数分析結果



次に不規則変動について、横軸に不規則変動(各時点のトレンドを1とした時の比率)の大きさ、縦軸に頻度(日数)をとったヒストグラムを用いて確認する。図5に消費支出、付図3に他の対象データ(食料、一般外食および飲酒代)のヒストグラムを示す。ちなみに、外れ値等を戻す前のベースでの不規則変動についても同じ集計を行ったが、ほぼ同じ形状のヒストグラムが得られた。

図5 消費支出の不規則変動の分布

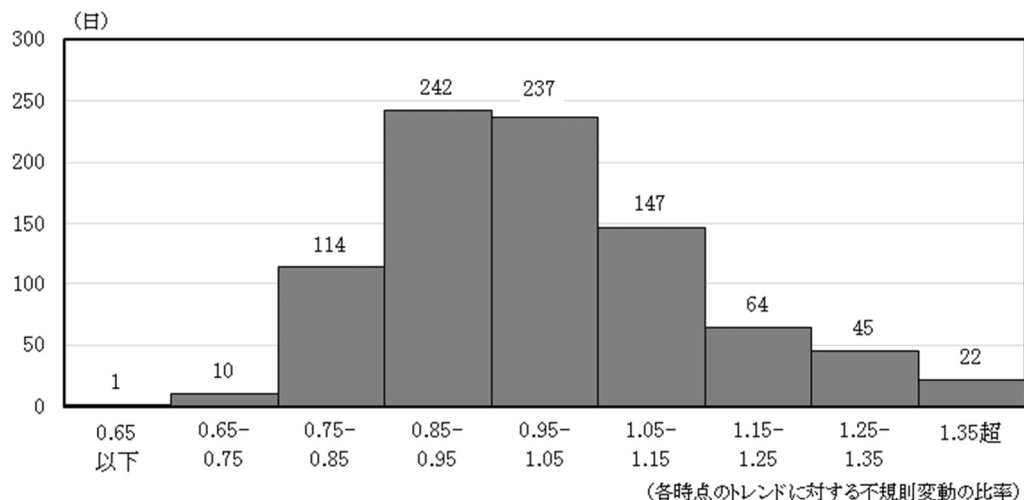


図5の消費支出の不規則変動のヒストグラムの形状をみると、右に長く裾を引く単峰型になっている。また付図3の他の対象データ(食料、一般外食および飲酒代)の

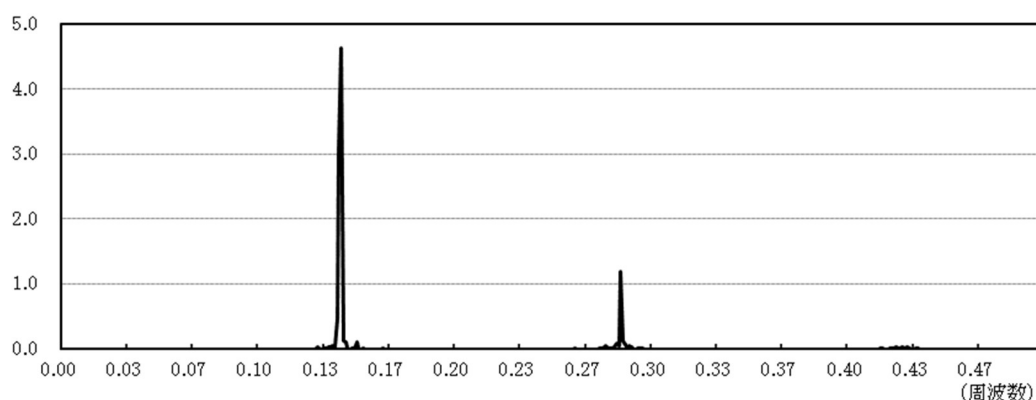
ヒストグラムの形状も、若干右に長く裾を引く単峰型である点は共通である。ただし仔細にみると、食料の尖度が大きく飲酒代は小さいなどの差異がみられる。

不規則変動の分布が右に裾を引く理由として、不規則変動の大きさが消費支出等の水準に比例する中、トレンドと季節変動を通じた消費支出等の水準の変動が不規則変動に分散不均一性を与え、不規則変動の分布が右に裾を引く分布である対数正規分布になったことが考えられる。よって右に裾を引いていることは問題ないと判断でき、不規則変動を分離できていると判断できる。

3.2.4 季節変動

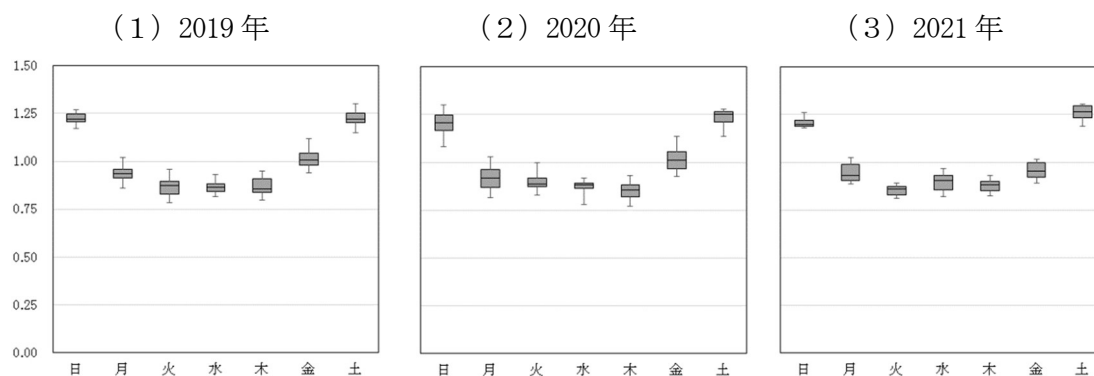
まず図6および付図1に季節変動に対し周波数分析を行った結果を示す。いずれも周波数0.14(周期7日)に強いパワースペクトルがみられ、その整数倍の周波数にそれに次ぐ強さのスペクトルが確認できることから、週単位の周期性の抽出は良好と評価できる。

図6 消費支出の不規則変動の周波数分析結果



次に、季節変動の時系列的な変化を抽出できる季節調整法の特徴を活かし、季節変動(各時点のトレンドを1とした時の値)について、2019年、2020年、2021年(5月末まで)の3期間でのボックスプロット(箱ひげ図)を用いて確認する。図7に消費支出、付図4に他の対象データ(食料、一般外食および飲酒代)のボックスプロットを示す。

図7 消費支出の季節変動のボックスプロット



消費支出を四分位範囲でみると、土日が高く（1.25程度）、他の日が低くなる週単位の季節性が確認された。そして季節性の傾向は、どの年も大きな違いは見られなかった。付図4の他の対象データ（食料、一般外食および飲酒代）のボックスプロットをみると、食料と一般外食は土日が高く他の平日が低い週単位の季節性、飲酒代は週末である金土が高く他の曜日が低い季節性がみられた。いずれも想定されるパターンを捕捉しており、週単位の季節変動も概ね分離できていると考えられる。

分離した季節変動を仔細に見ると、消費支出の金曜日の季節性は2019年と2020年は平均が1程度であったものが、2021年は第三四分位（全体の3/4の部分）で1と、やや低下しており金曜日の消費支出が弱くなっていることが分かる。この点、付図4のうち飲酒代の季節性の推移をみると、新型コロナウイルスが拡大する前である2019年は、金曜日が約1.75だったものが、2020年が約1.5、2021年は1.25弱にまで低下しており、これは新型コロナウイルスへの対応等による金曜日の外での飲酒機会の減少を映じているものと考えられる。

4. 考察

「2. 本研究の手法」で示した手法により、「3. 分析の結果」の通り、家計調査の消費支出等の日別の支出データをトレンド、季節変動、不規則変動に分離することに成功した。本邦の経済統計の日次データに対し季節調整法の適用に成功した初めての事例と考えられ、本研究の手法は高頻度データの活用における課題の一つであった、季節性・曜日要因などの時系列的な特性を考慮する上で、新たな有力な手法になり得ると考えられる。

もっとも、本研究および手法には以下に挙げる課題が残っている。まず、本研究では週単位の季節性への対応に着眼して分析を行ったが、この中で年周期等のメタの周期性が、週単位の周期性に与える影響について考慮できていない。このため本研究での分析では、「週単位の周期性」の変化について、その要因がメタの周期性なのか他の環境変化等なのか区分できていない。もし前者の要因による変化であれば、年等のメタの周期で「週単位の周期性」は元に戻る。他方、後者の要因の場合、「週単位の周期性」の変化は、それに影響を与える環境変化等の動向次第となる。本手法を通じて得られる「週単位の周期性」の変化を適切に評価する上で本課題の解決は重要と考えられる。解決に向けては、Ladiray et al. (2018)と本研究を組み合わせた、新しい手法の検討が考えられる。

次に、7以外の周期への対応も課題である。すなわち、高頻度データには、電力使用実績データや移動通信トラフィックデータ等、時次データの日単位の周期性、すなわち周期24のデータも存在し、これらを含めた多種多様な高頻度データの活用が広がり始めている状況を踏まえると、本研究の手法を周期24のデータ等、異なる周期へ拡張することも重要な課題と考えられる。

また月次統計との連携の検討も重要な課題である。本研究で利用した家計調査では月次データ等も公表されている。この月次データ等に対する季節調整では、既に月末の曜日の影響等、1年を周期としない要因による影響も精緻に分析し除去されているが、本研究の手法を活用することで日単位の変動を詳細に分析し、月次データの季節調整のさらなる精緻化につなげることも検討課題と考えられる。

このほか、本研究では季節移動平均の種類やヘンダーソン移動平均の項数の選択において、周期12と同じセットの中から選択している。周波数分析等の結果から、トレンド、季節変動、不規則変動は概ね適切に分離されていると評価できるが、本研究の手法の適用を他の日次データに拡大し、その結果を分析することで、本手法の一般性および頑健性を確認することも重要と考えられる。

5. その他

本論文は筆者個人の見解であり、筆者の所属する組織の公式見解ではなく、ありうべき誤りはすべて筆者個人に帰すべきものである。

謝辞

本稿については2名の査読者から有益な助言を賜り、重要な改善を行うことができました。ここに感謝申し上げます。最後に、本稿作成をサポートしてくれた妻と娘にも感謝します。

参考文献

- [1] 有田帝馬(2012), 『入門季節調整』, 東洋経済新報社.
- [2] 有田帝馬(2021), 「季節調整法を応用した新型コロナウイルス感染症の報告日別の新規陽性者数の変動分析」, 『厚生指標』, 2021年9月号(第68巻第11号), 一般財団法人厚生労働統計協会, 8-13,
<http://www.hws-kyokai.or.jp/paper/120-2016-02-15-03-07-32/2557-202109-2.html> .
- [3] 株式会社インテリサーチ(2020), 「経済産業省大臣官房調査統計グループ調査分析支援室委託調査『令和元年度 ビッグデータを活用した新指標開発事業(短期の販売・生産動向把握)』報告書」2020年3月,
https://www.meti.go.jp/meti_lib/report/2019FY/000493.pdf .
- [4] 小寺信也、藤田隼平、井上祐介、新田堯之(2020), 「POS・テキストデータを用いた消費分析—機械学習を活用して—」, 『経済財政分析ディスカッション・ペーパー』DP/18-1, 内閣府政策統括官(経済財政分析担当),
<https://www5.cao.go.jp/keizai3/discussion-paper/dp181.pdf> .
- [5] 総務省(2008), 「月末の曜日などの変動の要因を含めた季節調整値」, 『家計調査の結果を見る際のポイント』(No. 11), 総務省統計局,
<https://www.stat.go.jp/data/kakei/point/pdf/point11.pdf> .
- [6] 総務省(2009), 「季節調整法の変更について」, 『家計調査の結果を見る際のポイント』(No. 12), 総務省統計局,
<https://www.stat.go.jp/data/kakei/point/pdf/point12.pdf> .
- [7] 総務省(2020), 『令和2年版 情報通信白書』, 216-233,
<https://www.soumu.go.jp/johotsusintokei/whitepaper/ja/r02/pdf/index.html>.
- [8] 日本銀行(2020), 「東京大学金融教育研究センター・日本銀行調査統計局共催ビッグデータフォーラムの様式」, 『調査論文』2020年12月25日,
https://www.boj.or.jp/announcements/release_2020/rel201016a.htm/ .
- [9] 藤代宏一(2020), 「月次指標よりも高頻度データ」, 『Market Flash』, 2020年5月11日, 第一生命経済研究所,
<https://www.dlri.co.jp/report/macro/2020/fuji20200511mf.html> .
- [10] 若田部昌澄(2020), 「最近の金融経済情勢と金融政策運営—佐賀県金融経済懇談会における挨拶—」,
https://www.boj.or.jp/announcements/press/koen_2020/ko200902a.htm/ .
- [11] Aladangady, A., Aron-Dine, S., Dunn, W., Feiveson, L., Lengermann, P. and Sahm, C. (2019), From Transaction Data to Economic Statistics: Constructing Real-time, High-frequency, Geographic Measures of Consumer Spending, Big Data for 21st Century Economic Statistics, University of Chicago Press, <http://www.nber.org/chapters/c14267> .

- [12] Dagum, E. B. (1999), The X11ARIMA/88 seasonal adjustment method - foundations and user's manual, Statistics Canada technical report,
<https://www.census.gov/library/working-papers/1999/adrm/dagum-01.html> .
- [13] Findley, D. F., Monsell, B. C., Bell, W. R., Otto, M. C. and Chen, B-C. (1998), New Capabilities and Methods of the X-12-ARIMA Seasonal Adjustment Program, Journal of Business & Economic Statistics, Vol.16, Number 2, 127-152,
<https://www.tandfonline.com/doi/abs/10.1080/07350015.1998.10524743> .
- [14] Ladiray, D., Palate, J., Mazzi, G.L. and Proietti, T. (2018), Seasonal Adjustment of Daily and Weekly Data, Handbook on Seasonal Adjustment 2018 edition, Publications Office of the European Union, 757-783,
<https://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-/ks-gq-18-001> .
- [15] Lewis, D., Mertens, K. and Stock, J.H. (2020), U.S. ECONOMIC ACTIVITY DURING THE EARLY WEEKS OF THE SARS-COV-2 OUTBREAK, NBER WORKING PAPER SERIES 26954, <http://www.nber.org/papers/w26954/> .
- [16] Shiskin J, Young AH. and Musgrave JC.(1967), The X-11 variant of the census method II seasonal adjustment program, U.S. Census Bureau Technical Paper, No. 15,
<https://www.census.gov/library/working-papers/1967/adrm/shiskin-01.html> .
- [17] Scott, S., Tiller, R. and Chow, D.(2007), Empirical Evaluation of X-11 and Model-based Seasonal Adjustment Methods. Methods. 2007, Working Paper, Bureau of Labor Statistics,
<https://www.bls.gov/osmr/research-papers/2007/st070120.htm> .
- [18] U.S. Census Bureau (2017), X-13ARIMA-SEATS Reference Manual (Version 1.1), U.S. Census Bureau,
<https://www2.census.gov/software/x-13arima-seats/x13as/unix-linux/documentation/docx13as.pdf> .

付表1 Regression ARIMA モデルにおける外れ値等の回帰変数

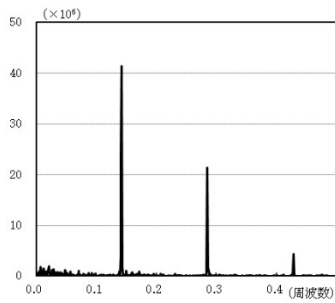
		外れ値等の回帰変数		
		AO	TC	LS
消費支出	2019年	4/1 10/12	1/1 12/30	—
	2020年	—	—	—
	2021年	5/10	—	1/2
食料	2019年	1/14 2/11 5/6 7/15 9/16 9/23 10/12 10/14 11/4 12/29	1/2 3/25 4/29 5/1 8/12 10/10 10/21 12/24 12/27 12/30	—
	2020年	1/1 1/13 2/3 2/24 11/23 12/28 12/29 12/30 12/31	3/19 3/25 5/4 7/20 8/10 8/13 9/21 9/23 12/24	1/4
	2021年	1/1 1/2 1/3 1/11 2/2 3/3	1/7 5/3	—
一般外食	2019年	1/1 1/2 1/3 1/14 1/25 2/11 2/25 3/25 4/29 5/2 5/3 5/6 6/25 7/15 9/16 9/23 10/12 10/14 11/4 11/25	4/25 4/30 8/12 9/25 10/21 12/30	—
	2020年	1/13 3/25 4/24 5/25 6/25 8/25 11/3 11/23 11/25 12/25	1/2 2/10 2/24 3/19 5/4 7/22 8/10 8/13 9/21 9/23 12/28	4/2
	2021年	1/11 1/25 2/11 2/23 2/25 4/29 5/25	5/3	—
飲酒代	2019年	—	—	—
	2020年	1/1 4/22 4/29 7/13 10/15 12/24	4/24 4/27 5/7 7/26 9/1 12/28 12/31	3/3 3/30 4/9 5/2 5/26
	2021年	1/13 2/15 2/24 4/14 5/12 5/17 5/21 5/26	1/18 5/3 5/23	1/7 2/4 4/25

注1 AOは加法的な外れ値、TCは減衰的な外れ値、LSは水準変化の回帰変数を示す。

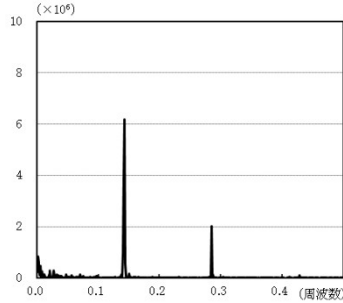
注2 回帰変数のt値はすべて絶対値2以上で有意となっている。

付図1 食料、一般外食および飲酒代の周波数分析結果

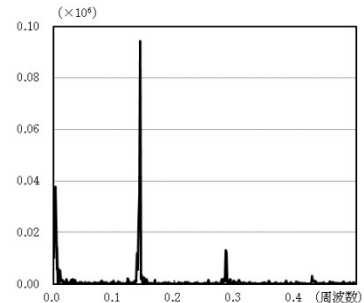
1. 事前調整前のデータ
(食料)



(一般外食)

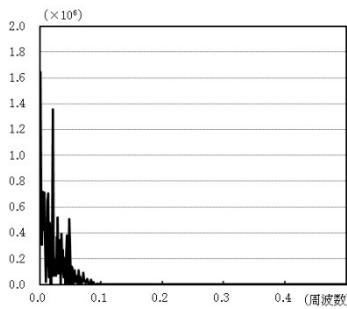


(飲酒代)

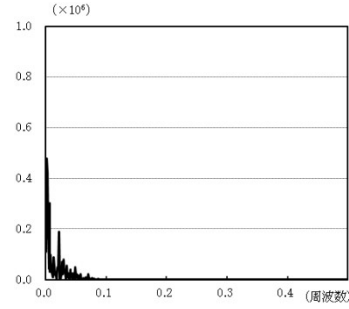


2. トレンド

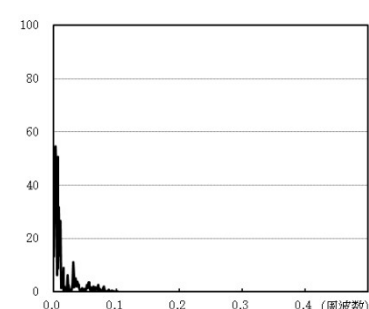
(食料)



(一般外食)

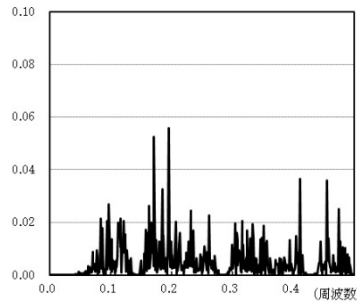


(飲酒代)

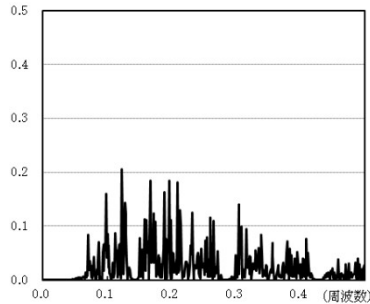


3. 不規則変動

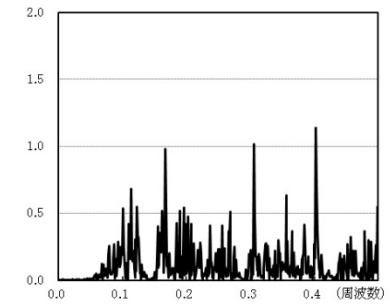
(食料)



(一般外食)

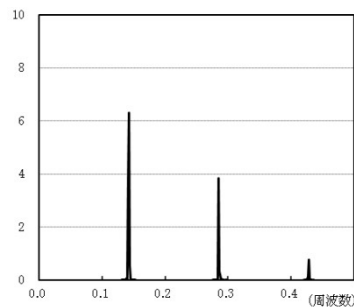


(飲酒代)

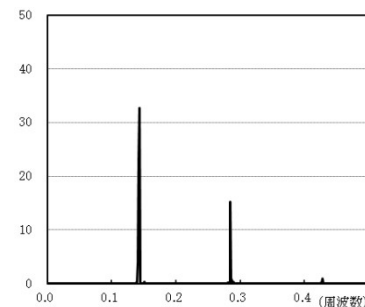


4. 季節変動

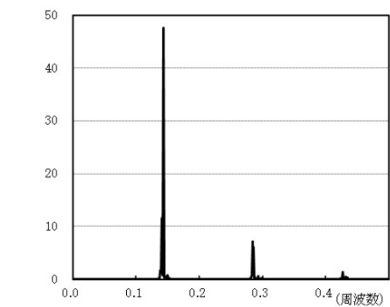
(食料)



(一般外食)



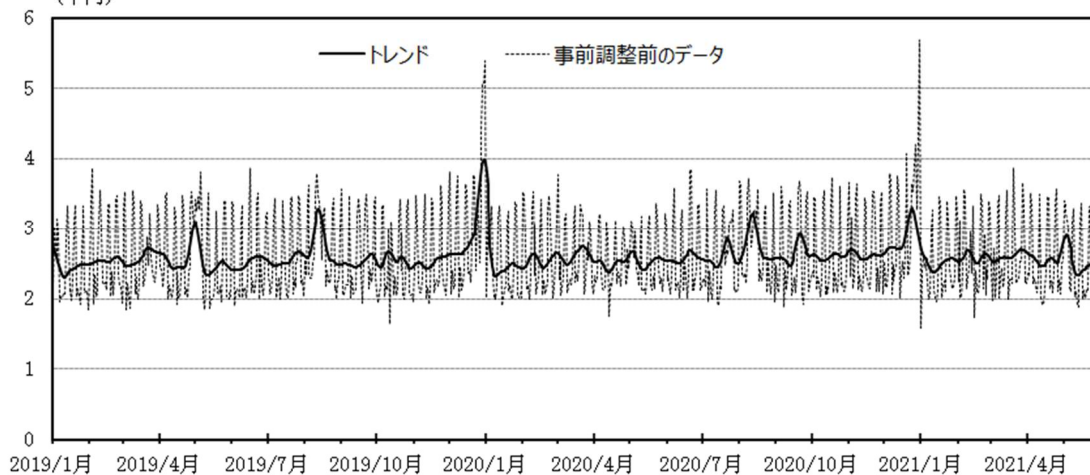
(飲酒代)



付図2 食料、一般外食および飲酒代の事前調整前のデータと分離されたトレンド

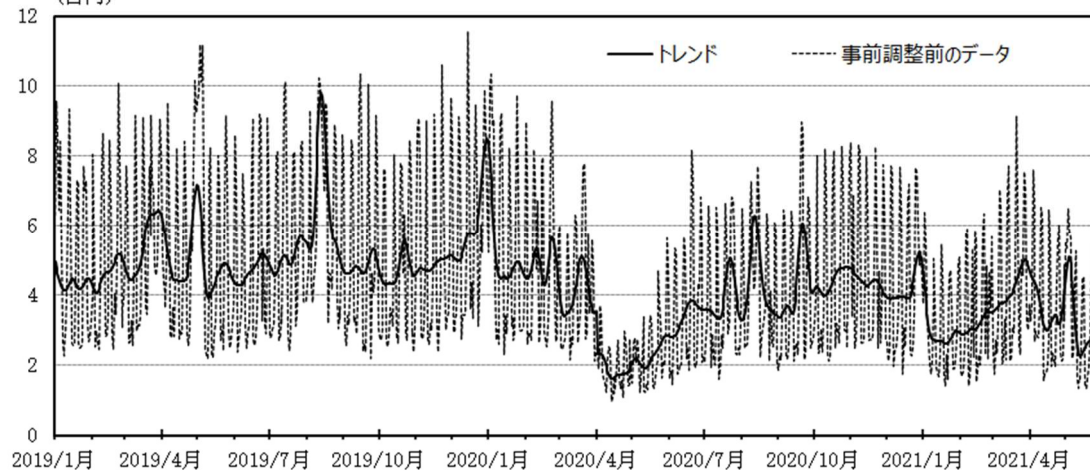
1. 食料

(千円)



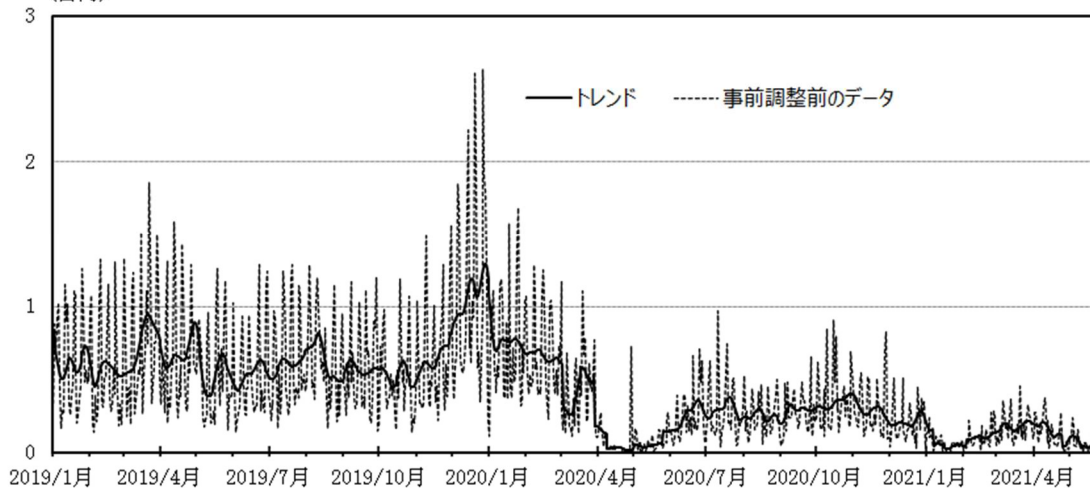
2. 一般外食

(百円)



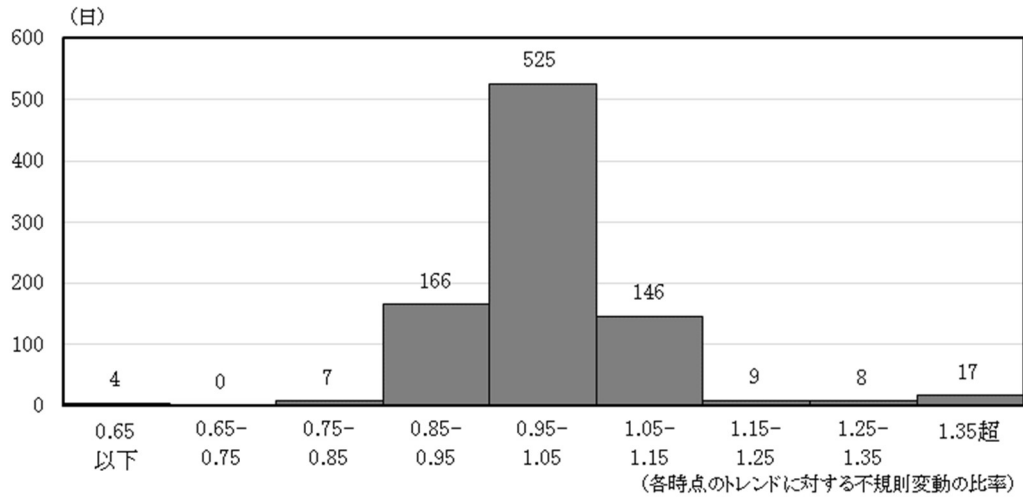
3. 飲酒代

(百円)

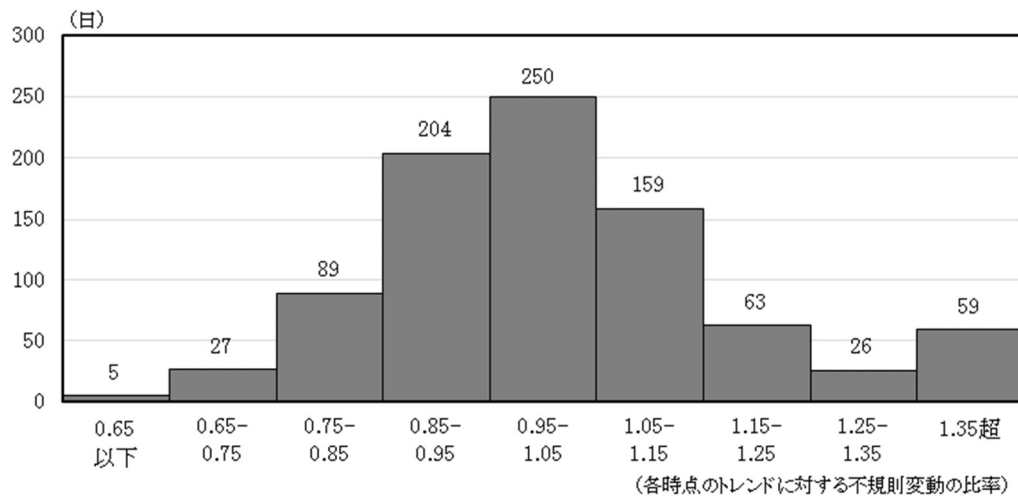


付図3 食料、一般外食および飲酒代の不規則変動の分布

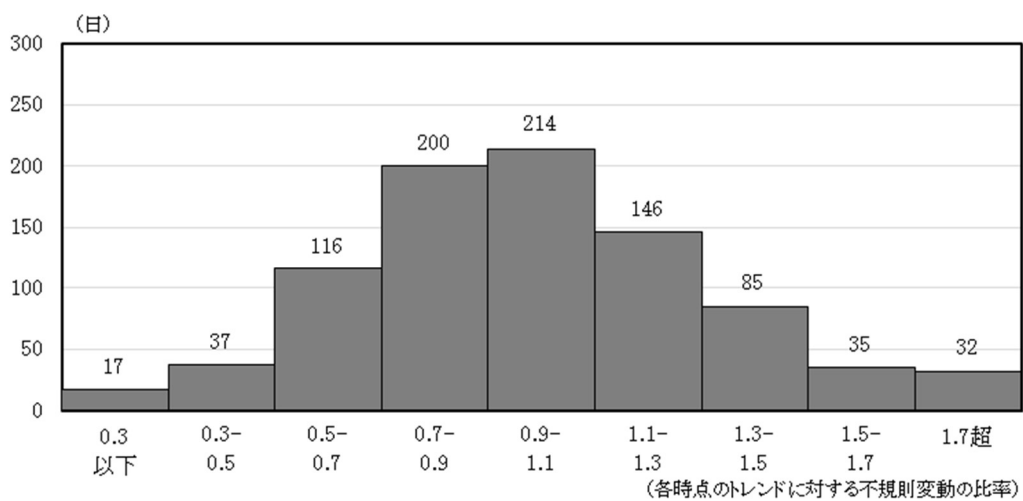
1. 食料



2. 一般外食

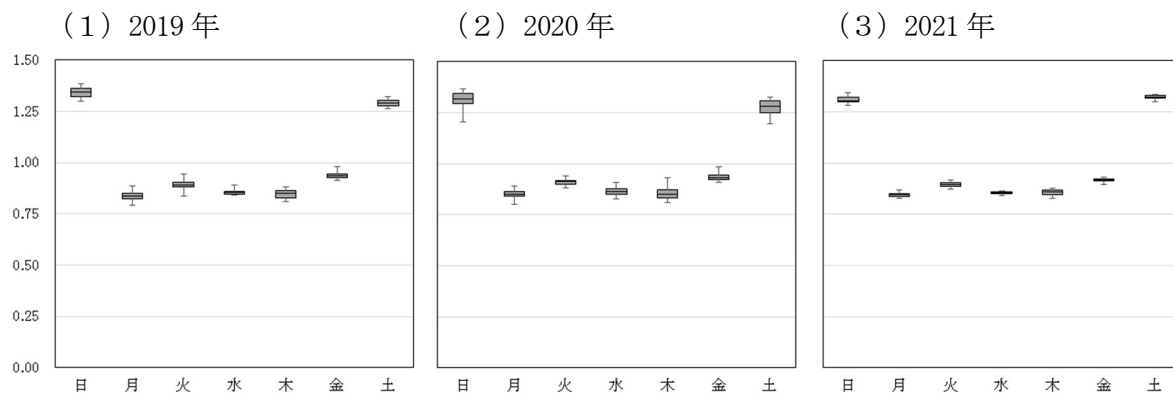


3. 飲酒代

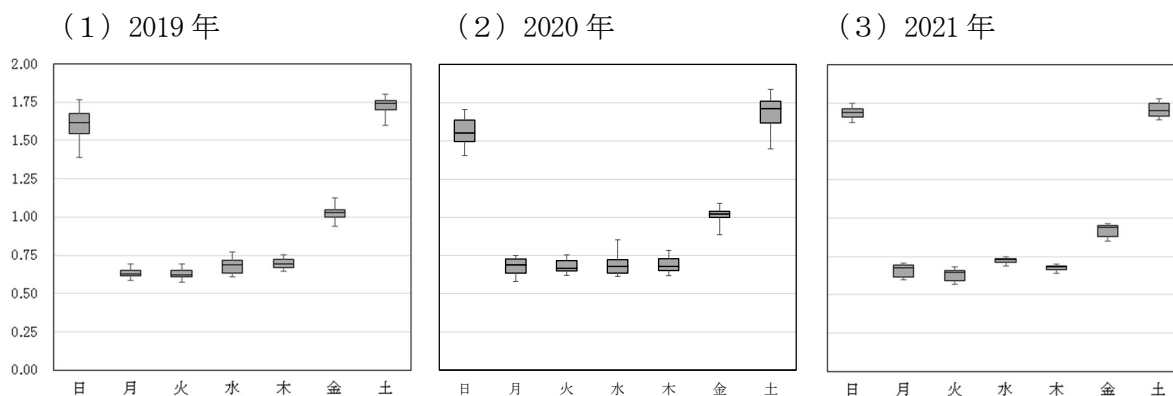


付図4 食料、一般外食および飲酒代の季節変動のボックスプロット

1. 食料



2. 一般外食



3. 飲酒代

