

社会生活基本調査から見た コロナ禍等の影響

～その2：クラスターサイズから見た影響～



木村 敦 | Kimura Atsushi

(独)統計センター 理事・CIO

■ NTT にて ICT 関連開発に長年携わり、(株)NTT ファシリティーズ総合研究所 取締役情報技術本部長を経て、2019 年 4 月から現職。1988 年 3 月名古屋大学大学院理学研究科博士課程（前期）修了、修士（理学）、専門統計調査士。

1. はじめに

前回の「社会生活基本調査から見たコロナ禍等の影響（その1）」^[1]では、コロナ禍前の2016年とコロナ禍の2021年の社会生活基本調査のデータ^{図1}に加えて2011年、2006年、2001年、1996年のデータを用いて、クラスター数2における非類似度について分析を行った。その結果、47都道府県を2つのクラスターに分割した際の非類似度が2021年調査データで有意に小さくなっていることが分かった。加えて、男性データのみの特徴として、2011年においてもコロナ禍と同様に非類似度が小さくなっていることが分かり、東日本大震災の影響の可能性を指摘した。

本稿では、社会生活基本統計の1日の20種類の行動種別に割り当てられる総平均時間を都道府県別かつ男女別に20次元の多変量データとして多変量分析を行った結果をクラスターサイズの観点から分析した。各都道府県におけるこの20次元データを、各都道府県の「生活スタイル」であるにとらえ、その「生活スタイル」全体の変化を定量的に分析してみることにする。

なお本稿の見解は筆者個人のものであり、所属する組織を代表するものではなく、本稿の誤

りなどは全て筆者の責に帰するものである。

2. クラスター数2での日本区分地図

今回は、クラスターサイズの観点から分析してみよう。クラスター数を2に固定して、各調査年次のデータをクラスター分析して日本区分地図で表示させてみる^{注1}。まず、男性のデータから見てみよう（図1～図6）^{注2}。2021年もクラスターサイズに変化が見られるが、2001年は「東京クラスター」が他の年次と異なり顕著に大きいことが分かる。2001年には、多くの都道府県が「東京クラスター」型にシフトしているということである。

女性データによる分析結果も見てみよう。図7～図12に示した結果を比較してみると、男性と同様に2001年の「東京クラスター」に含まれる都道府県が多いことが分かる。2001年を除けば、「東京クラスター」の傾向はおおむね安定的な類似したパターンを示している。これは、前回の（その1）で示した通り、コロナ禍における2021年のデータでも他の年次調査と類似しており、男性のようにクラスターサイズの影響は認められない。クラスターサイズの観点からは、コロナ禍の影響は女性の場合は見ることができない。

注1) 本分析では、東京を含むクラスターを「東京クラスター」、それ以外の都道府県で構成されるクラスターを「その他クラスター」と呼ぶ。詳細は木村 [1] 参照。



図 1 2021 年調査（男）による区分地図



図 4 2006 年調査（男）による区分地図



図 2 2016 年調査（男）による区分地図



図 5 2001 年調査（男）による区分地図



図 3 2011 年調査（男）による区分地図



図 6 1996 年調査（男）による区分地図

3. 景気動向などとの関連について

2001 年の「東京クラスター」が他の年次と大きく異なっていることについては、どのような解釈が考えられるであろうか？

図 13 は、内閣府 経済社会総合研究所 景気統計部が公表している「景気動向指標 令和 5（2023）年 12 月分（速報）」¹³ から「一致指数」に関する図を抜粋したものに筆者が一部加筆したもので

注 2) 図 1～図 26 の拡大版を(公財)統計情報研究開発センター HP (<https://www.sinfonica.or.jp/>) 内の [刊行物] > [エストレーラ] > [参考] に掲載した。





図7 2021年調査（女）による区分地図



図10 2006年調査（女）による区分地図



図8 2016年調査（女）による区分地図



図11 2001年調査（女）による区分地図



図9 2011年調査（女）による区分地図



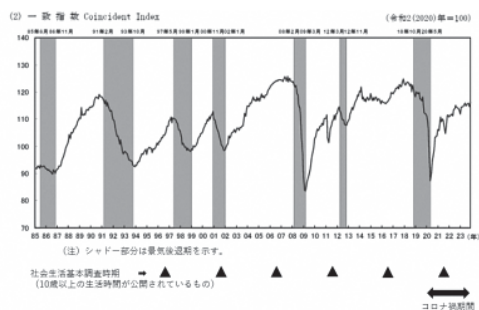
図12 1996年調査（女）による区分地図

ある。グラフのシャドー部分は景気後退期を示している。黒色の▲は過去6回の社会生活基本調査が実施されたタイミングを示している。これを見ると2001年調査の時期は、景気後退期の一つと

ピッタリと一致していることが分かる。他の調査年次では、どの景気後退期とも重なっていない。

2001年だけに特徴的な「東京クラスター」のサイズ拡大の要因として、この景気後退期が何ら

かの影響を与えている可能性はある。2001 年における非類似度が他の年次と大きく変わらないことを考慮すると、2001 年の景気後退期においては「東京クラスター」と「その他クラスター」間の非類似度自体には大きな変化がないが、多くの都道府県において「東京クラスター」型の生活スタイルにシフトしていたことを示している。



出典：「景気動向指数 令和 5 (2023) 年 12 月分 (速報)」内閣府経済社会総合研究所 景気統計部 P4 より抜粋し、筆者により社会生活基本調査時期等を追記

図 13 景気動向と社会生活基本調査年次の関係

景気後退期の生活スタイルとして、大都市型のスタイルにシフトした都道府県が複数あったということであり、景気後退期においては、非景気後退期と比較してより多くの都道府県において「東京クラスター」型の活動をせざるを得なかったということは十分考えることである。具体的にどのような行動変化が起きていたのかは本稿後半で行動種別ごとに詳細に見てみることにしよう。なお、女性データの特徴であった 2001 年及び 1996 年の非類似度が大きいことについては、何らかの別要因による影響と思われるが、現時点で要因は不明である。引き続き検討が必要である。

4. LIV 法などによるより詳細な分析

階層的クラスター分析の結果に関する多変量としての各変数（今回の場合であれば、行動種別 20 種類の各データ）の寄与率については、筆者が以前本誌で提案した「変数別寄与率算定法 (LIV 法)」[4] を用いて明らかにすることが

できる。「変数別寄与率算定法」についてここで簡単に説明しておこう。

Ward 法階層的クラスター分析法では、各凝集工程において凝集させる 2 つのクラスターを選定する際に「情報の損失 (loss of information)」が最も少ないペアを選ぶ。「情報の損失」は、クラスター p とクラスター q をクラスター r として 1 つに凝集させた場合には、

$$\{Np \cdot Nq / (Np + Nq)\} \times \|c_p - c_q\|^2$$

で表される。この「情報の損失」が Ward 法における「非類似度の定義」である [5]。ここで、 Np, Nq はそれぞれクラスター p , クラスター q が含むオブジェクトの数、 c_p, c_q は、それぞれクラスター p , クラスター q の平均値ベクトルである。「変数別寄与率算定法」では、ある凝集における「情報の損失」の大きさに変数 n が寄与する比率を LIV 値と名付け、変数 n の LIV 値 ($LIVn$) を、

$$LIVn = \{(c_{pn} - c_{qn})^2 / \|c_p - c_q\|^2\} \times 100$$

で定義している [4]。

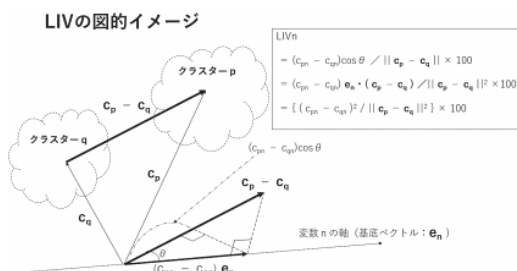


図 14 LIV の概念図

ここで c_{pn}, c_{qn} は、それぞれクラスター p , クラスター q の平均ベクトルにおける変数 n の値である。

LIV 値の大きな変数は、ある工程で凝集した 2 つのクラスター p と q において、違いが大きい変数を示す。つまり、凝集した 2 つのクラスターをそれぞれ特徴づける変数を明らかにすることができるわけである。

それでは、この LIV 値を用いて「東京クラスター」と「その他クラスター」における違いについて分析してみよう。

まず、男性について見てみよう。2021 年調査のデータをクラスター分析した時の「東京クラスター」と「その他クラスター」間の LIV 値を示したのが図 15 である。棒グラフは、各行動種別の LIV 値（左側縦軸）を示し、折れ線グラフは、LIV 値の累積値（右側縦軸）である。全ての行動種別の LIV 値を合計すると 100%となる。「東京クラスター」の値が大きい行動種別の棒グラフはグレーの網掛けで、「その他クラスター」の値が大きい行動種別の棒グラフは白抜きで表示している。

図の左寄りにある行動種別ほど両クラスター間の違いを生み出す寄与度が高い。図 15 の場合は、左から 4 つの行動で約 50%の累積寄与率となっていることが分かる。上位 3 行動について見ると、順に「学習・自己啓発・訓練（学業以外）」、「通勤・通学」、「睡眠」となっている。「東京クラスター」では、「学習・自己啓発・訓練（学業以外）」と「通勤・通学」に多くの時間を割いており、「睡眠」に割く時間が少ない。「その他クラスター」では逆の傾向になっている。

2016 年から 1996 年までの男性の調査データでも見てみよう。図 16 から図 20 までに示す。

これらの結果を見比べてみると、明らかに 2001 年の傾向が大きく異なっていることが分かる。他の年次では 1 位もしくは 2 位である「通勤・通学」が 2001 年だけ 4 位に下がっている。「通勤・通学」に関する違いよりも「睡眠」、「休憩・くつろぎ」、「買い物」の差が上位 3 位を占めている。

表 1 男性の行動種別順位

調査年	通勤・通学	睡眠	学習・自己啓発	(参考)
2021	2 位	3 位	1 位	
2016	2 位	3 位	4	1 位：移動
2011	1 位	3 位	2 位	
2006	1 位	3 位	2 位	
2001	4	1 位	7	2 位：休養・くつろぎ、3 位：買い物
1996	1 位	2 位	7	3 位：休養・くつろぎ

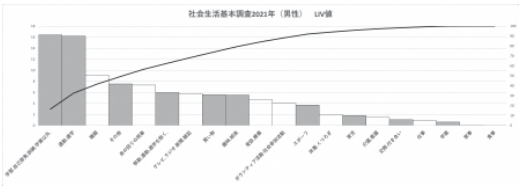


図 15 2021 年調査分析の LIV (男)

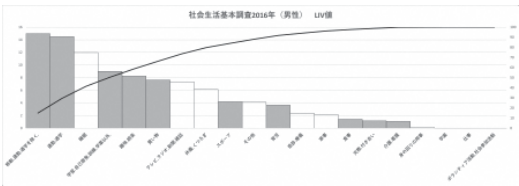


図 16 2016 年調査分析の LIV (男)

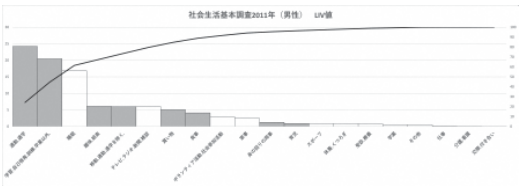


図 17 2011 年調査分析の LIV (男)

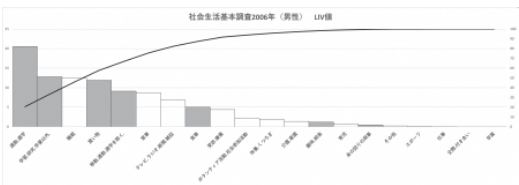


図 18 2006 年調査分析の LIV (男)

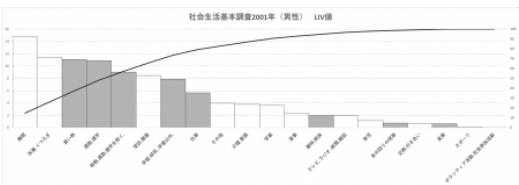


図 19 2001 年調査分析の LIV (男)

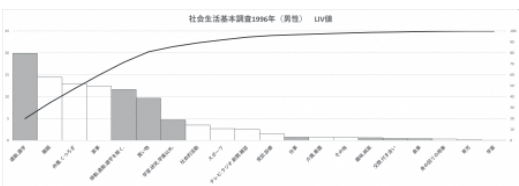


図 20 1996 年調査分析の LIV (男)

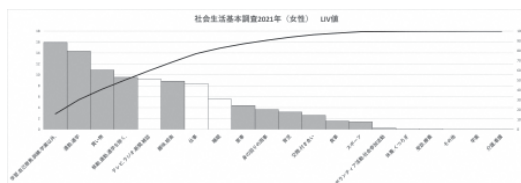


図 21 2021 年調査分析の LIV (女)

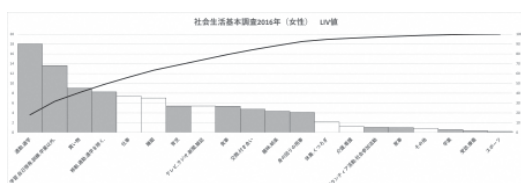


図 22 2016 年調査分析の LIV (女)

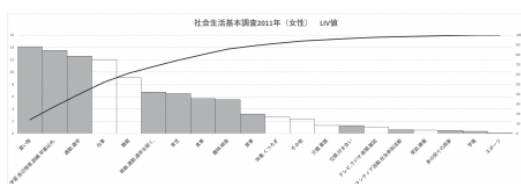


図 23 2011 年調査分析の LIV (女)

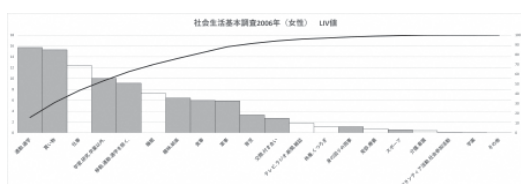


図 24 2006 年調査分析の LIV (女)

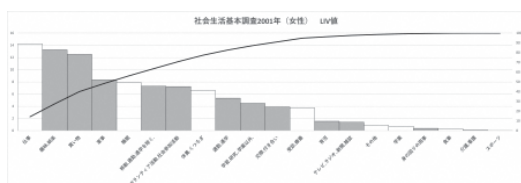


図 25 2001 年調査分析の LIV (女)

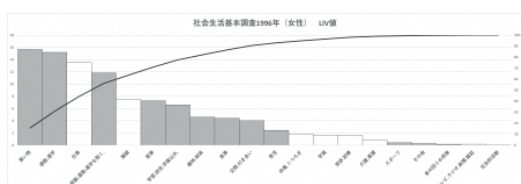


図 26 1996 年調査分析の LIV (女)

次に、女性のデータを図 21 ～図 26 に示す。女性の場合においても、他の年次では上位 3 位に入っている「通勤・通学」が 2001 年だけ 9 位に下がっている。また、2001 年では「趣味・娯楽」が 2 位になっている。

「趣味・娯楽」は他の年次ではかなり違いの少ない行動種類である。2001 年のクラスターサイズの違いは、これらの行動スタイルの変化に起因していることが分かる。

表 2 女性の行動種別順位

調査年	通勤・通学	買い物	学習・自己啓発	(参考)
2021	2 位	3 位	1 位	
2016	1 位	3 位	2 位	
2011	3 位	1 位	2 位	
2006	1 位	2 位	4	3 位：仕事
2001	9	3 位	10	1 位：仕事、2 位：趣味・娯楽
1996	2 位	1 位	7	3 位：仕事

5. おわりに

コロナ禍による「生活スタイル」の変化が大きく表れると想像していたが、予想外に景気後退期と重なった 2001 年調査のデータのほうがクラスターサイズの観点では、大きく影響を受けていたことが分かった。前稿 [1] では、クラスター数 2 における非類似度について分析を行った。コロナ禍の「生活スタイル」への特微的な影響として、この非類似度の縮小として表れることも分かった。

*参考文献

- [1] 木村敦 (2024) 「社会生活基本調査から見たコロナ禍等の影響～(その 1) クラスター非類似度から見た影響～」『ESTRELA』No.367 (2024 年 10 月号), pp.30-36, 統計情報研究開発センター。
- [2] 勤統計センター「SSDSE (教育用標準データセット)」
<https://www.nstac.go.jp/SSDSE/index.html>
- [3] 内閣府「平成 23 年度年次経済報告書」第 1 章第一節。
- [4] 木村敦 (2023) 「階層的クラスター分析結果の新たな解釈手法：「変数別寄与率算定法 (LIV 法)」の提案～家計調査クラスター分析結果への適用～」『ESTRELA』347 (2023 年 2 月号), pp.20-25, 統計情報研究開発センター。
- [5] Joe H. Ward, Jr.(1963) “Hierarchical Grouping to Optimize an Objective Function”, *Journal of the American Statistical Association*, Vol. 58, No. 301(Mar., 1963), pp.236-244.