

2025 年度 統計データ分析コンペティション

審査員奨励賞 [高校生の部]

合計特殊出生率の変動要因分析

松本 春飛

(京華高等学校)

合計特殊出生率の変動要因分析

1. 研究のテーマと目的

日本の少子化は、社会経済の持続可能性を揺るがす構造的課題として長年議論されてきた。その核心的指標である合計特殊出生率（以下、TFR）は、人口置換水準を大幅に下回り続け、かつその値には顕著な地域差が存在する。例えば、2022年において沖縄県のTFRが1.70である一方、東京都は1.04と大きな隔りがある¹。本研究の目的は、都道府県別の公表統計データを用い、TFRの変動を説明する社会・経済・人口構造的要因を多角的に分析することにある。特に、複数の要因が相互に影響し合う現実を考慮し、重回帰分析を主たる分析手法として採用する。

少子化の要因に関する研究は数多く蓄積されている。山田（1999）²は若年層の経済的基盤の脆弱性を指摘し、武田（2005）³は女性の就労と育児の両立の困難さに焦点を当てた。これらの研究は、個人や世帯レベルのミクロな視点から問題を分析した点で重要である。一方、マクロな地域レベルの分析として、稲葉（2024）⁴は都道府県データを用いた重回帰分析を行い、核家族世帯割合や保育所数がTFRや婚姻率に正の影響を与えうることを示した。この研究は、本稿と同様のアプローチを取る先行事例として重要である。本研究は、稲葉（2024）の分析を発展させ、所得や人口密度といったマクロ経済・人口構造変数に加え、新たに大学進学率（高学歴化の代理変数）や年平均気温といった変数を投入する。これにより、TFRの規定要因に関するモデルの精緻化を図り、より包括的な知見を得ることを目指す。

2. 研究の方法と手順

2-1. 使用データ

本研究では、統計データ分析コンペティションで提供されたSSDSEのデータを用いて分析を行なった。TFR（2022年）を目的変数とし、それに影響を与えると考えた以下の7つの指標を説明変数として設定した。

- 婚姻率: 人口1,000人あたりの婚姻件数
- 一人当たり県民所得: 県民の所得水準
- 高齢化率: 総人口に占める65歳以上人口の割合
- 人口密度: 都市化の度合いを示す指標
- 人口あたり保育所数: 人口10万人あたりの保育所等数
- 大学進学率: 高等学校卒業者に占める大学等進学者数の割合
- 年平均気温: 地理的・気候的要因

2-2. 分析手法

相関分析

各説明変数とTFRの間の単純な関係性を把握するため、相関係数を算出し、散布図を作成した。これにより、各変数が単独でどの程度TFRと関連しているかを視覚的・定量的に評価した。

重回帰分析

次に、複数の説明変数が相互に与える影響を統制した上で、各変数の大きさを評価するため、重回帰分析を実施した。これにより、相関分析だけでは見えない変数間の見せかけの相関を排除した。

3. データセットの加工

- 合計特殊出生率 (TFR)

データソース	列名	補足
SSDSE-E-2025	2022_合計特殊出生率	本分析で変動を説明したい対象

- 純移動率と過疎化進行地域の定義

データソース	列名	補足
SSDSE-E-2025	2023_総人口	その都道府県の総人口
SSDSE-E-2025	2023_転入者数_日本人移動者	他の都道府県から移り住んできた人の数
SSDSE-E-2025	2023_転出者数_日本人移動者	他の都道府県へ移り住んでいった人の数

$$\text{純移動率} = (\text{転入者数} - \text{転出者数}) \div \text{総人口} \times 1000$$

純移動率を 47 都道府県すべてで算出し、値が低い順に並び替えを行い、純移動率が最も低い下位 10 県を過疎化進行地域と定義する。

- 高齢化率の定義

データソース	列名	補足
SSDSE-E-2025	2023_総人口	その都道府県の総人口
SSDSE-E-2025	2023_65 歳以上人口	高齢化率_分子

$$\text{高齢化率} (\%) = (65 \text{ 歳以上人口} \div \text{総人口}) \times 100$$

- 婚姻率の定義 (人口 1,000 人あたり)

データソース	列名	補足
SSDSE-E-2025	2023_総人口	その都道府県の総人口
SSDSE-E-2025	2022_婚姻件数	その都道府県の婚姻件数

$$\text{婚姻率} (\%) = (\text{婚姻件数} \div \text{総人口}) \times 1,000$$

- 一人当たり県民所得

データソース	列名	補足
SSDSE-E-2025	2020_1 人あたり県民所得 _平成 27 年基準	データセットの値を直接使用

- 人口密度 (都市化度合い)

データソース	列名	補足
SSDSE-E-2025	2023_総人口	その都道府県の総人口
SSDSE-E-2025	2023_総面積 _北方地域及び竹島を除く	その都道府県の総面積

$$\text{人口密度} = \text{総人口} \div \text{総面積}$$

- 人口あたり保育所数 (子育て支援インフラの充実度)

データソース	列名	補足
SSDSE-E-2025	2023_総人口	その都道府県の総人口

SSDSE-E-2025	2022_保育所等数	その都道府県の保育所等数
--------------	------------	--------------

人口あたり保育所数 = (保育所等数 ÷ 総人口) × 100,000

● 大学進学率（教育水準が高い地域）

データソース	列名	補足
SSDSE-B-2025	高等学校卒業生数	大学進学率_分子
SSDSE-B-2025	高等学校卒業生のうち進学者数	大学進学率_分母

大学進学率 = 高等学校卒業生のうち進学者数 ÷ 高等学校卒業生数

● 年平均気温

データソース	列名	補足
SSDSE-B-2025	年平均気温	データセットの値を直接使用

4. データ分析の結果

4-1. 基礎集計

分析の第一歩として、全国の TFR（表 1）と出生数（表 2）を比較した。

（表 1）全国の TFR のランキング

上位10			下位10		
1	沖縄	1.70	1	東京	1.04
2	宮崎	1.63	2	宮城	1.09
3	鳥取	1.60	3	北海道	1.12
4	島根	1.57	4	埼玉	1.17
5	長崎	1.57	5	神奈川	1.17
6	鹿児島	1.54	6	秋田	1.18
7	佐賀	1.53	7	千葉	1.18
8	熊本	1.52	8	京府	1.18
9	福井	1.50	9	岩手	1.21
10	大分	1.49	10	大阪	1.22

（表 2）全国の出生数のランキング

上位10			下位10		
1	東京	91,097	1	高知	3,721
2	大阪	57,315	2	鳥取	3,752
3	神奈川	56,498	3	秋田	3,992
4	愛知	51,152	4	徳島	4,148
5	埼玉	43,451	5	島根	4,161
6	千葉	36,966	6	山梨	4,759
7	福岡	35,970	7	福井	4,861
8	兵庫	33,565	8	和歌山	5,238
9	北海道	26,407	9	佐賀	5,552
10	静岡	20,575	10	山形	5,674

東京都は、人口が非常に多いため生まれる赤ちゃんの総数（出生数）は全国で最も多いが、女性一人当たりが生む子供の数（TFR）は最も低い、という特徴的な結果になっていた。TFRは1人の女性が生涯に産む平均出生数を表す指標であるため、この数字だけを見ると、東京が最も少子化していると思われがちである。しかし、人口1000人あたりの普通出生率（表3）を見ると、東京は全国平均よりも高いことがわかった。つまり、一人あたりの出産確率は低くても、親になる可能性のある母数（分母）が圧倒的に多いため、地域全体として生まれる子供の数は多くなり、人口あたりの出生率も平均以上に押し上げられている。

（表3）人口1000人あたりの普通出生率のランキング

上位10			下位10		
1	沖縄	9.26	1	秋田	4.37
2	福岡	7.05	2	岩手	4.98
3	鳥取	6.99	3	青森	5.05
4	佐賀	6.98	4	北海道	5.19
5	熊本	6.95	5	福島	5.49
6	滋賀	6.94	6	新潟	5.52
7	宮崎	6.85	7	山形	5.53
8	愛知	6.84	8	栃木	5.54
9	鹿児島	6.80	9	高知	5.59
10	岡山	6.70	10	群馬	5.62
15	東京	6.47			
全国平均		6.20			

次に東京、過疎化進行地域、その他のエリアについて比較を行った。（表4）この比較表は、日本の少子化問題が単一ではなく、「東京」の問題と「地方」の問題という、性質の異なる2つの課題から成り立っていることを明確に示している。

（表4）東京、過疎化進行地域の比較

エリア	平均TFR	出生数	純移動率	高齢化率
東京	1.04	91,097	4.15	22.8
過疎化ワースト10	1.37	62,762	-3.9	34.3
その他のエリア	1.36	616,891	-1.22	31.1

東京は、若者は集まるが、子供を産みにくい「環境の問題」として捉えることができる。TFRは全国最の1.04である。これは、東京の生活環境（高い住居費・教育費、キャリア重視の価値観、核家族化など）が、一人ひとりの女性にとって子供を持つという選択を難しくしていることを示唆している。純移動率は圧倒的なプラス(+4.15)を示しており、全国から若者が集まり続けている。これにより、親世代となる人口の母数が巨大に保たれていることが理解できる。また、高齢化率は全国で最も低く(22.8)、若者が集まる結果、地域全体の人口構成は非常に若い。このことから、東京の問題は、子供を産む環境の厳しさであることがわかる。ポテンシャルのある親世代は大量に存在するにもかかわらず、そのポテンシャルを活かしきれない状態と言える。

過疎地域は、子供は産みやすいが、若者がいない「人口基盤の問題」として捉えることができる。TFRは非常に高い(1.37)く、全国平均よりも高い水準である。これは、もし若者が地域に残れば、東京よりも子供が生まれやすい社会環境である可能性を示唆している。純移動率は深刻なマイナス(-3.90)となっており、その親世代となる若者が圧倒的な勢いで流出している。高齢化率は高く(34.3)、若者が流出し、高齢者人口の割合が増えている。このことから、地方の問題は、「子供を産む親世代そのものの消滅」と言える。出生率自体は低くないのに、その担い手がいなくなることで、地域の出生数が激減し、人口減少が加速するという悪循環に陥っており、この10県を合計した出生数(約6.3万人)は、東京一都(約9.1万人)に遠く及ばない。

4-2. 相関係数と散布図

TFR と関係が強いと考えら得る次の7つの変数について、相関係数（表5）と散布図（図1～図7）を用い、分析を行なった。

（表5）TFR との相関係数

	相関係数
婚姻率	-0.10
一人当たり県民所得	-0.46
高齢化率	0.23
人口密度	-0.44
人口あたり保育所数	0.63
大学進学率	-0.49
年平均気温	0.53

婚姻率 ($r = -0.10$):婚姻率と TFR の間には、ほとんど相関が見られなかった。これは当初の仮説異なり、婚姻が出生の前提条件となっている日本の現状とは異なる結果となっている。

一人当たり県民所得 ($r = -0.46$):所得が高い都道府県ほど、TFR が低いという逆説的な関係が見られた。所得の高い都市部では、教育費をはじめとする子育てコストが高騰し、結果として持つ子供の数を抑制する傾向があることを示唆している。

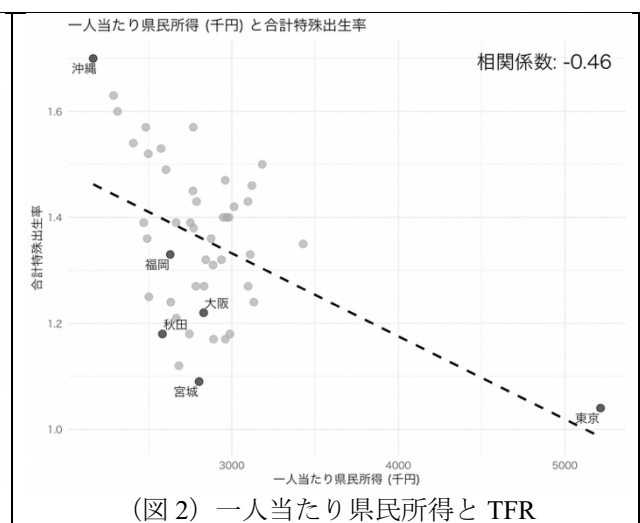
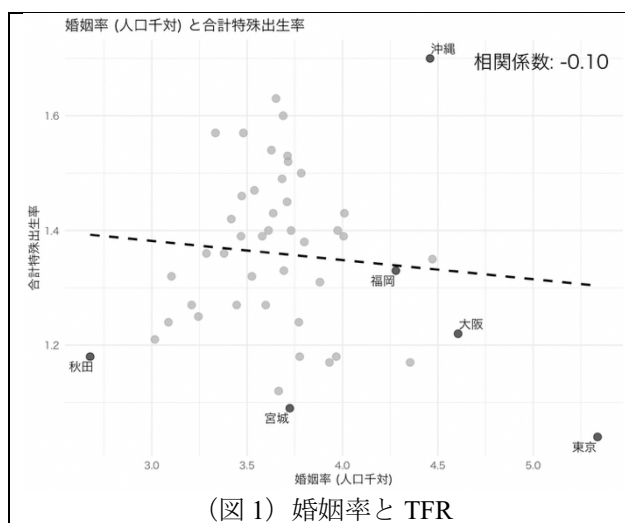
高齢化率 ($r = 0.23$):高齢化率と TFR には弱い正の相関が見られた。

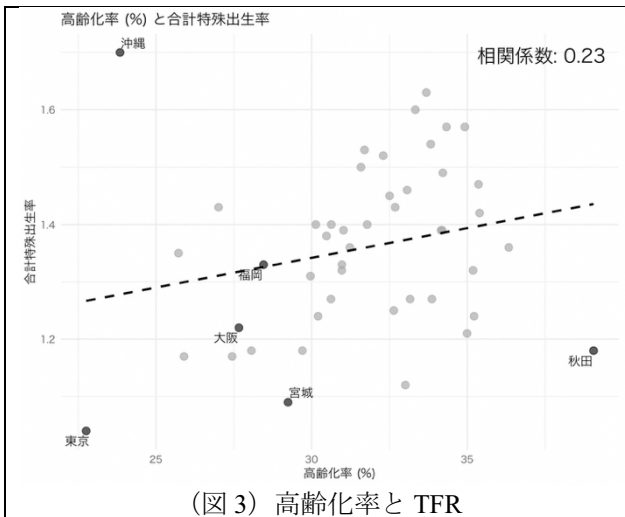
人口密度 ($r = -0.44$):人口密度が高い、すなわち都市部であるほど TFR が低い傾向にある。都市部特有の高い生活コスト、核家族化、待機児童問題などが複合的に影響していると考えられる。

人口あたり保育所数 ($r = 0.63$):人口あたりの保育所数が多い、つまり子育て支援インフラが充実している都道府県ほど、TFR が高い傾向にある。これは、保育サービスへのアクセスしやすさが、子供を持つという決断を後押しする重要な要因であることを示唆している。

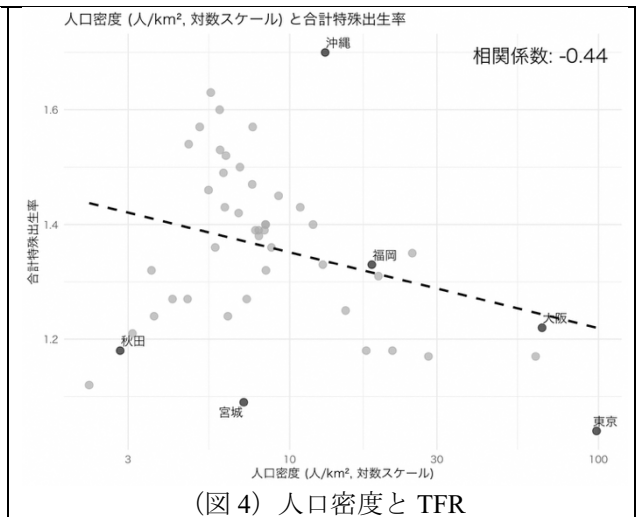
大学進学率 ($r = -0.49$):大学進学率が高い都道府県ほど、TFR が低い傾向にある。これは、教育期間が長くなることによる晩婚化・晩産化や、高学歴の女性が出産・育児のためにキャリアを中断するなどが背景にあると考えられる。

年平均気温 ($r = 0.53$):年平均気温が高い都道府県ほど、TFR が高いという中程度の相関が見られた。これは沖縄県や九州地方など、温暖な西日本の都道府県で出生率が比較的高く、寒冷な東北地方で低い。一方、気候が直接的な原因というよりは、温暖な地域と寒冷な地域における家族観、地域のコミュニティ、生活様式といった文化的な背景の違いが影響している可能性がある。

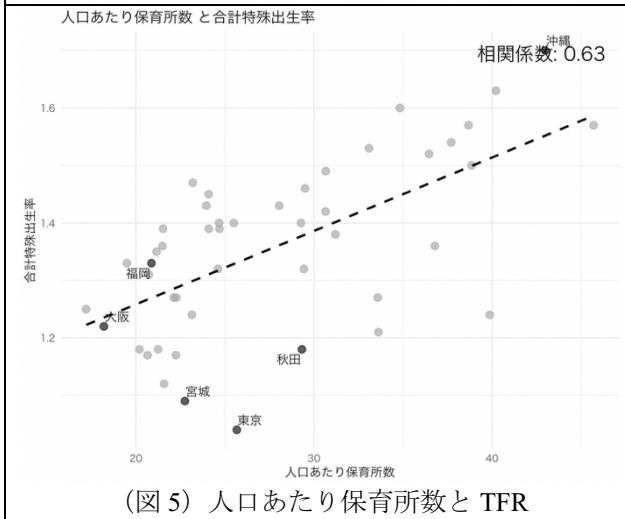




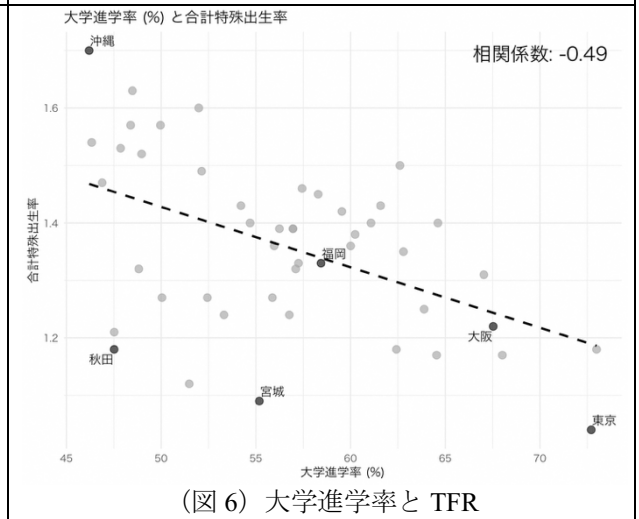
(図3) 高齢化率と TFR



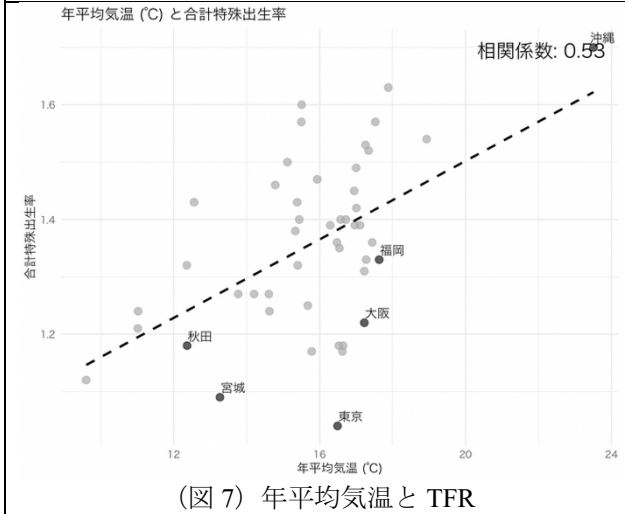
(図4) 人口密度と TFR



(図5) 人口あたり保育所数と TFR



(図6) 大学進学率と TFR



(図7) 年平均気温と TFR

4-3. 重回帰分析

7つの説明変数全てを投入した重回帰分析を行った結果、モデル全体として TFR のばらつきの約 77.5% を説明できるという非常に精度の高い結果が得られた（自由度調整済み $R^2=0.7746$ ）。各変数の係数と統計的有意性は（表6）の通りである。また、多重共線性の確認として VIF 値（表7）を計算した。VIF 値は最大でも 8.67 と基準値の 10 を下回っており、変数間の多重共線性はやや高いものの、分析結果の信頼性を大きく損なうレベルではないと判断した。

（表6）重回帰分析の結果

	係数	P値	評価
切片	-0.503	0.293	
婚姻率	0.209	0.004	**
一人当たり県民所得	0.000	0.926	
高齢化率	0.018	0.021	*
人口密度	-0.007	0.000	***
人口あたり保育所数	0.008	0.000	***
大学進学率	-0.002	0.312	
年平均気温	0.030	0.000	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

（表7）各説明変数の VIF 値

説明変数	VIF
婚姻率	8.67
一人当たり県民所得	2.29
高齢化率	5.61
人口密度	3.26
人口あたり保育所数	1.62
大学進学率	2.43
年平均気温	1.86

重回帰分析の結果、相関関係では見えなかった要因が明らかになった。相関分析では「一人当たり県民所得」と「大学進学率」が TFR と負の相関を示したが、他の変数の影響を取り除いた結果、これらの変数は統計的に有意ではなかった。人口密度がマイナスの値を持っていることから、人口密度という変数に集約されていることを示唆している。つまり、所得や学歴そのものが原因のではなく、都市に住むことに伴う複合的な要因（高い生活コスト、核家族化、通勤時間など）が、TFR を抑制するより要因である可能性が高い。

相関分析ではほぼ無関係に見えた婚姻率が、重回帰分析では有意な正の変数となった。これは、都市部で婚姻率と TFR が共に低いという見せかけの相関（疑似相関）が起こっていると考えられる。その影響を取り除くことで、婚姻件数の増加が TFR の上昇につながるという本来の関係性が明確になった。

5. 結果の解釈

重回帰分析の結果から、TFR に影響を与える要因は、①政策的に介入が困難な地理的・環境要因と、②政策的な介入が可能な要因の二つに大別できることが明らかになった。前者は TFR を取り巻く前提条件を、後者は具体的な政策の方向性を示唆する上で重要な意味を持つ。

1. 環境要因からの洞察：都市部における子育ての構造的困難性

人口密度や年平均気温といった変数は、政策で直接変えることはできないが、TFR に影響を与える「環境」や「背景」を理解する上で重要な示唆を与えてくれる。人口密度そのものを政策的に操作することは困難だが、それが TFR に対して強い負の影響を持つという事実は、都市部における子育ての構造的な困難さを浮き彫りにしている。高い人口密度は、単なる人の密集度ではなく、高騰する住居費や教育費、長い通勤時間、地域コミュニティの希薄化といった、子育て世帯にとっての複合的な障壁の代理変数と解釈できる。

したがって、求められる政策は、高密度な環境でも子供を育てやすい社会の設計であり、柔軟な働き方の推進による時間的制約の緩和や、都市公園・子育て支援施設の拡充といった物理的・社会的な環境整備が有効なアプローチと考えられる。

2. 施策要因からの提言：少子化対策の両輪

一方で、今回の分析では、政策的に介入可能かつ TFR に直接的な影響を与える二つの強力な要因が特定された。

保育インフラという直接的な解決策

他の要因の影響を考慮してもなお、人口あたり保育所数は TFR と極めて強い正の相関を示した。これは、出産後の就労継続を支える保育インフラへの公的投資が、依然として最も効果的な少子化対策の一つであることを統計的に裏付けている。

婚姻率という見過ごされがちな入り口

重回帰分析によって初めてその重要性が明らかになった婚姻率は、少子化対策が産み育てやすさだけでなく、その前段階である結婚のしやすさにも目を向けるべきであることを強く示唆している。若者が経済的な不安なく結婚を選択できる安定した雇用環境の整備や、新生活への支援は、TFR の基盤を支える上で不可欠な要素と言える。

参考文献

- 1 厚生労働省：「合計特殊出生率の年次推移」
- 2 山田昌弘：「パラサイト・シングル時代」(1999)
- 3 武田宏子：「結婚・出産と女性の労働参加」(2005)
- 4 稲葉拓真：「少子化進行抑止のための家庭・社会要因の探究」(2024)