

福祉と所得格差の国際比較

— クラスタ分析(2) —

吉 岡 慎 一

はじめに

2000年代の2時点における OECD 諸国の不平等度と貧困度によるいくつかの分類が、吉岡 (2014a) において行われた。本稿では1990年代半ばと2000年代末の2時点における OECD 諸国を福祉指標と不平等度によっていくつかクラスタリングする試みが続けられる¹⁾。福祉国家を特徴付ける指標として社会支出を用いることにたいする Esping-Andersen (1990, ch. 1) の批判はよく知られているが、1990年代半ば以降、先進国における福祉に関するデータの質が向上したことによって、福祉国家を特徴付ける福祉指標として社会支出を用いる研究例は多い²⁾。さらに、厚生経済学では不平等を福祉で測る伝統があり、両者は負の相関関係にあるからこの2つの指標を同時に用いて福祉レジームの観点から OECD30カ国の分類が試みられる。

我が国の1990年代半ばは、社会保障給付の総額も給付比（対国民所得）も1980年代から急上昇している時期であり、さらに1970年代から徐々に上昇していた所得不平等度の上昇傾向に拍車がかかった時期でもある。そして2000年代末まで、相対不平等度の上昇傾向³⁾と社会保障給付の総額および給付比（対国民所得）の上昇傾向は継続している。OECD 諸国の中には、この間不平等が上

1) クラスタ分析によって、OECD 諸国の消費パターンを国際比較した研究に Kovacs (1985) がある。

2) Atkinson et al. (1995, ch. 6), Korpi = Palme (1998), Castles (2008), Obinger = Wagshal (2012).

3) 吉岡 (2014).

表1 国名略語等

国名番号	国名略語	国名	国名番号	国名略語	国名
1	AUS	オーストラリア	16	ISL	アイスランド
2	AUT	オーストリア	17	ITA	イタリア
3	BEL	ベルギー	18	JPN	日本
4	CAN	カナダ	19	KOR	韓国
5	CHE	スイス	20	LUX	ルクセンブルク
6	CZE	チェコ	21	MEX	メキシコ
7	DEU	ドイツ	22	NLD	オランダ
8	DNK	デンマーク	23	NOR	ノルウェー
9	ESP	スペイン	24	NZL	ニュージーランド
10	FIN	フィンランド	25	POL	ポーランド
11	FRA	フランス	26	PRT	ポルトガル
12	GBR	イギリス	27	SVK	スロバキア
13	GRC	ギリシャ	28	SWE	スウェーデン
14	HUN	ハンガリー	29	TUR	トルコ
15	IRL	アイルランド	30	USA	アメリカ

昇した国家もあれば低下した国家もある。また、社会支出比（対 GDP）が上昇した国家もあれば低下した国家もある。そこで第1節において、社会支出比（対 GDP）による OECD30カ国のクラスタリング結果が1995年と2010年について各国の高齢化率との関連で比較され、第2節においては、ジニ係数による OECD 諸国のクラスタリング結果が1995年頃と2010年頃について比較される。第3節においては、社会支出比とジニ係数の2次元データによる OECD24カ国のクラスタリング結果が1995年頃と2010年頃について比較される。第3節まではこの種の研究分野の習慣に従って OECD 諸国が5つのクラスタに分割され、分析が行われるが、第4節では、比較的によく利用される4つの評価指標によって適切なクラスタ数の候補が探求され、2010年頃の OECD30カ国の2次元データによっていくつかの分類が試みられる。

1. OECD 諸国における社会支出比と日本の位置

本稿における分析対象は表1のように先進国としての OECD 諸国である。表2は1995年と2010年について OECD がネット上に公表している公的社会支出と GDP から社会支出比を作成し、各国のそれを降順にまとめたものである。

表2 OECD 諸国の社会支出比 (対 GDP)

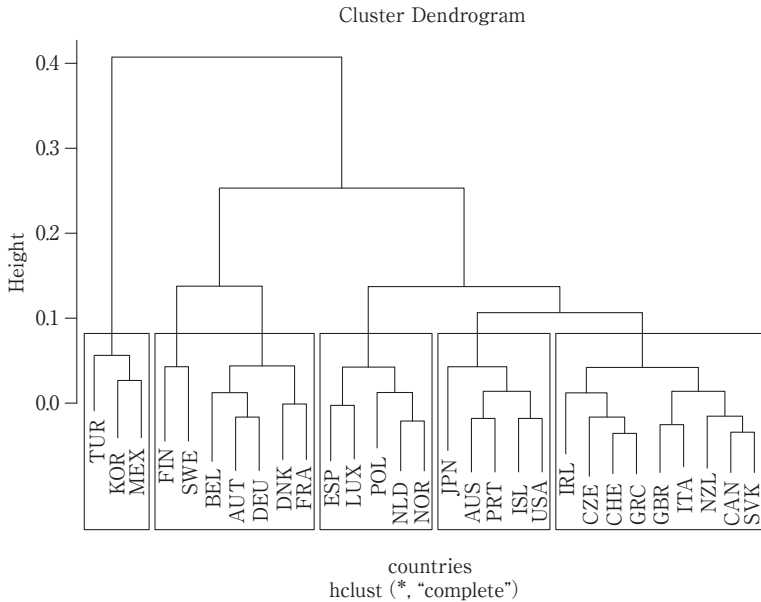
国名略語	1995年	国名略語	2010年
SWE	0.320	FRA	0.324
FIN	0.307	DNK	0.306
FRA	0.293	FIN	0.296
DNK	0.289	BEL	0.295
DEU	0.266	AUT	0.289
AUT	0.265	SWE	0.283
BEL	0.262	ITA	0.277
NLD	0.238	DEU	0.271
NOR	0.234	ESP	0.267
POL	0.226	PRT	0.254
ESP	0.214	GBR	0.238
LUX	0.208	IRL	0.237
GBR	0.199	NLD	0.234
ITA	0.198	GRC	0.233
CAN	0.189	LUX	0.230
SVK	0.188	NOR	0.230
NZL	0.186	HUN	0.229
IRL	0.181	JPN	0.223
GRC	0.175	POL	0.218
CHE	0.175	NZL	0.213
CZE	0.174	CZE	0.208
PRT	0.165	CHE	0.206
AUS	0.162	USA	0.198
USA	0.155	SVK	0.191
ISL	0.152	CAN	0.187
JPN	0.141	ISL	0.180
TUR	0.056	AUS	0.179
MEX	0.043	TUR	0.128
KOR	0.032	KOR	0.092
		MEX	0.081

(資料) OECD, StatExtracts より抜粋 (19 Oct 2014)。

(注) 国名略語は表1を参照のこと。

社会の高齢化に伴う年金給付費と医療給付費の急増のために1995年と2010年との間でほとんどの国の社会支出比は上昇しているが、スウェーデン、フィンランド、オランダ、ノルウェー、ポーランドおよびカナダのように低下した国(29カ国中6カ国)もある。低下を示した国々はカナダを除いて社会支出比が元々高い国である。社会支出比の上昇幅が狭いためにその順位自体は低下した国が6カ国(ドイツ、ルクセンブルク、スロバキア、ニュージーランド、スイス、オーストリア)ある。我が国の場合、1990年代中期に社会支出比が低かつ

図1 社会支出比の樹形図 (1995年)



(資料) 表2により計測・作成。

(注) 表2に同じ。

たために2000年代末までの上昇幅は広く、その順位は高いほうから26番目から18番目へ上昇しているが、未だに中位の位置までは達していない。

さて、表2を利用して1995年の社会支出比によるOECD諸国の階層的クラスタリング⁴⁾を行った結果が図1の樹形図(デンドログラム)である⁵⁾。分割数を一意的に決定する方法はないが、第4節で明らかになる分類数の候補に含まれこの研究分野で習慣的に採用される5分類をまず検討する⁶⁾。図1からOECD 29カ国の5分類が読み取れるが、クラスタ数 $k=5$ として非階層的クラスタリング法の1つである k 平均法⁷⁾を適用すると次のように前者と同一の5分類が

4) クラスタ分析の概要については、Kaufman = Rousseeuw (1990), Everitt et al. (2001), Hastie et al. (2009, ch.14)等を参照。

5) 以下、特に断らない限り、階層的クラスタリングにおいてユークリッド距離と完全連結法とが用いられた。国名略語は表1を参照のこと。

6) OECD (2008)。

得られる⁸⁾。

- 1) スウェーデン, フィンランド, フランス, デンマーク, ドイツ, オーストリア, ベルギー: 7カ国クラスタ平均=0.286
- 2) オランダ, ノルウェー, ポーランド, スペイン, ルクセンブルク: 5カ国クラスタ平均=0.224
- 3) イギリス, イタリア, カナダ, スロバキア, ニュージーランド, アイルランド, ギリシャ, スイス, チェコ: 9カ国クラスタ平均=0.185
- 4) ポルトガル, オーストラリア, アメリカ, アイスランド, 日本: 5カ国クラスタ平均=0.155
- 5) トルコ, メキシコ, 韓国: 3カ国クラスタ平均=0.044

このように1995年に社会支出比が一番高い第1クラスタには、北欧諸国と大陸欧州の主要国が属し、一番低い第5クラスタには非欧州かつ非英語圏の国々が属している。我が国は第4クラスタに属すが、そのクラスタ内では最後尾に位置している。高い社会支出比を持つ諸国をもう少し詳しく検討してみよう。社会民主主義レジームのスカンジナビア3国に属し高齢化率⁹⁾が高いのが、スウェーデン、ノルウェーおよびデンマークであり、保守主義レジームの大陸欧州に属し高齢化率が高いのが、フランス、ドイツ、オーストリアおよびベルギーである。フィンランドは社会民主主義レジームの北欧に属し高齢化率は中程度である(29カ国中第15位)。1990年代中期に社会支出比が相対的に高いスペインでは高齢化率が高いが、オランダ、ポーランドおよびルクセンブルクの高齢化率は高くない。1990年代中期に相対的高齢化社会になっているイギリスおよびイタリアの社会支出比は中位クラスタに属す。

次に、表2を利用して2010年の社会支出比によるOECD諸国の階層的クラスタリングを行った結果が図2の樹形図である。この図からOECD30カ国の5分類が読み取れるが、クラスタ数 $k=5$ として k 平均法を適用すると次のように階層的クラスタリングの結果に類似した5分類が得られる¹⁰⁾。

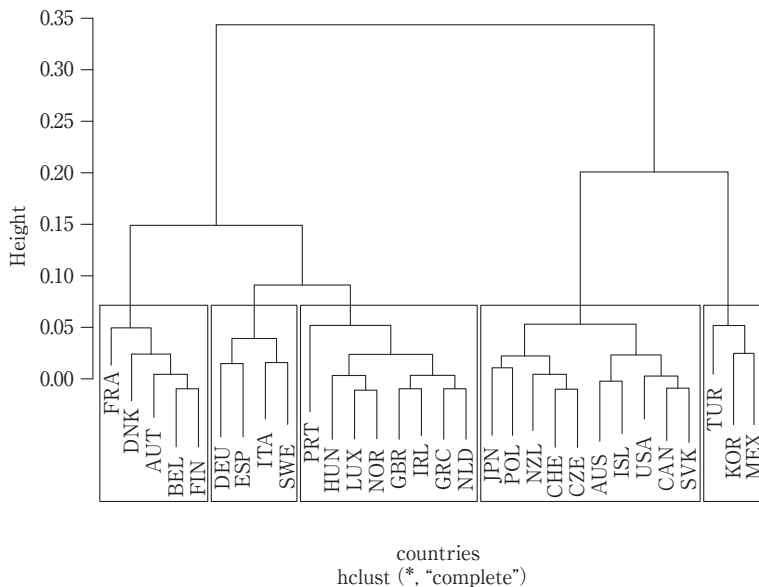
7) 以下、特に断らない限り、 k 平均法においてHartigan-Wong法が用いられた。

8) 階層的クラスタリングと非階層的クラスタリングとで分類結果は必ずしも同一ではない。

9) OECD諸国の高齢化率を比較した表3を参照。

図2 社会支出比の樹形図 (2010年)

Cluster Dendrogram



(資料) 表2により計測・作成。

(注) 表2に同じ。

- 1) フランス, デンマーク, フィンランド, ベルギー, オーストリア: 5カ国クラスタ平均=0.302
- 2) スウェーデン, イタリア, ドイツ, スペイン, ポルトガル: 5カ国クラスタ平均=0.270
- 3) イギリス, アイルランド, オランダ, ギリシャ, ルクセンブルク, ノルウェー, ハンガリー, 日本, ポーランド, ニュージーランド: 10カ国クラスタ平均=0.229

10) 階層的クラスタリングにおいては、k平均法において第2クラスタに属すポルトガルが第3クラスタへ、第3クラスタに属す日本、ポーランドおよびニュージーランドが第4クラスタへ移行している。このようにクラスタリング法によっては30カ国中4カ国で所属位置がずれる。例えばこのことは、先に指摘された我が国が2000年代末までに先進国の基準で社会支出比の中位の位置までには達していないという判断に対応する。

表3 OECD諸国の高齢化率(%)

国名略語	1995年	国名略語	2010年
SWE	17.47	JPN	23.02
ITA	16.72	DEU	20.98
DEU	16.14	ITA	20.41
NOR	15.94	GRC	19.10
BEL	15.89	PRT	18.48
GBR	15.85	SWE	18.28
ESP	15.29	AUT	17.67
DNK	15.23	FIN	17.27
FRA	15.17	BEL	17.14
AUT	15.13	ESP	16.95
GRC	15.12	FRA	16.85
PRT	14.92	CHE	16.85
CHE	14.73	HUN	16.67
JPN	14.56	DNK	16.55
FIN	14.21	GBR	15.98
HUN	14.21	NLD	15.45
LUX	14.01	CZE	15.37
NLD	13.24	NOR	14.97
CZE	13.22	CAN	14.16
USA	12.68	LUX	13.94
CAN	11.97	AUS	13.56
AUS	11.90	POL	13.45
NZL	11.53	USA	13.09
IRL	11.42	NZL	13.03
ISL	11.26	ISL	12.14
POL	11.19	IRL	11.31
KOR	5.89	KOR	11.04
TUR	5.45	TUR	7.12
MEX	4.32	MEX	6.18

(資料) OECD, StatExtracts より抜粋 (14 Sept 2015)。

(注) 表2に同じ。

4) チェコ, スイス, アメリカ, スロバキア, カナダ, アイスランド, オーストラリア: 7カ国クラスタ平均=0.193

5) トルコ, 韓国, メキシコ: 3カ国クラスタ平均=0.100

第1クラスタから第5クラスタまでの30カ国を降順に, 地理的あるいは言語的(人種的)に区分すると次のようになるが, 第1クラスタから第4クラスタまでの27カ国ですら福祉レジームで単純に区分することは困難である。

1) 北欧諸国および大陸欧州諸国¹¹⁾ :

2) 北欧1カ国, 大陸欧州1カ国および南欧諸国 :

- 3) アングロ・サクソン諸国, 大陸欧州諸国, 東欧諸国, 南欧1カ国, 北欧1カ国および日本:
- 4) アングロ・サクソン諸国, 東欧諸国, 大陸欧州1カ国および北欧1カ国:
- 5) 非欧州・非英語圏諸国:

1995年と2010年との間で社会支出比自体が低下した6カ国中スウェーデン, オランダ, ノルウェー, ポーランドおよびカナダにおいて, 社会支出順位が低下する方向へのクラスタ移行が起こっている。社会支出比の上昇幅が狭いためその順位自体が低下した6カ国中ドイツ, ルクセンブルク, スロバキアおよびスイスにおいて, 社会支出順位が低下する方向へのクラスタ移行が起こっている。

社会支出比が高い第1クラスタや第2クラスタに属す諸国を中心に高齢化率と関連させて2時点間の福祉動向をもう少し詳しく検討してみよう。デンマークでは高齢化率でみた場合, 2000年代末頃には高齢化傾向が弱く高齢化順位は低下しているが, 社会支出比自体が上昇しているため, 社会支出が一番高い第1クラスタに留まっている。フィンランドは2時点間で急速に高齢化したにもかかわらず社会支出比自体はやや低下しただけなので, 第1クラスタに留まっている。スウェーデンはやや高齢化したが高齢化順位は低下し, 社会支出比自体も低下しているため, 社会支出比の第2クラスタへの下方移行が起こっている。ノルウェーでは高齢化が止まり, 社会支出比自体がやや低下しているため, 社会支出比の第2クラスタから第3クラスタへの下方移行が起こっている。2時点に共通に社会支出比の第1クラスタに属すフランス, ベルギーおよびオーストリアの福祉動向はほぼ同じである。2時点間でやや高齢化したこの3カ国において, 社会支出比およびその順位が上昇したために社会支出が一番高い第1クラスタに留まっている。2000年代末頃に第2クラスタに属すドイツと南欧3カ国の福祉動向はいくぶん異なる。ドイツは1995年頃に相対的に高齢化社会になっていて, その後さらに高齢化の度合いを強めたが, 社会支出比の上昇が

11) 大陸欧州諸国とは, 本稿では北欧, 南欧及び東欧以外の大陸欧州諸国を指すことがある。

見られないために、社会支出比の第1クラスタから第2クラスタへの下方移行が起こっている。1990年代中期に相対的に高齢化社会になっていたイタリアはその後さらに高齢化の度合いを強め、社会支出比のかなりの上昇が見られるために、社会支出比の第3クラスタから第2クラスタへの上方移行が起こっている。スペインでは2時点間で高齢化の度合いが弱く、社会支出比の上昇傾向も弱いので、社会支出比の第2クラスタに留まっている。ポルトガルでは2時点間に高齢化率が急上昇し社会支出比も急上昇しているから、社会支出比の第4クラスタから第2クラスタへの上方移行が観察される。2時点間で高齢化率が上昇しているポーランドとカナダにおいては、社会支出比自体が低下しているために、社会支出比クラスタの下方移行が起こっている。

以上のように、社会の高齢化と社会支出との関係は先進諸国に限っても国ごとに異なるが、表2および表3を利用して高齢化率と社会支出比との相関係数を計算すると、1990年代中期の高齢化率と社会支出比との相関係数が0.80、2000年代末のそれが0.74となり、両者間には強い正の相関関係があるといえる。

2. OECD 諸国における不平等度と日本の位置

多数ある相対的不平等測度の1つのジニ係数（等価可処分所得）を1990年代中期と2000年代末ごとにまとめると表4が得られる。この2時点間にOECD25カ国中13カ国で不平等度の上昇が見られ、12カ国で低下が観察される。不平等度の上昇が見られたのは北欧4カ国、大陸欧州4カ国、アングロ・サクソン4カ国および日本においてであり、低下が見られたのは東欧2カ国、大陸欧州2カ国、アングロ・サクソン2カ国および南欧4カ国においてである。表4のOECD25カ国のジニ係数（1995年頃）に階層的クラスタリングが適用されて得られた樹形図が図3であり、これから国家の5分類が窺える。クラスタ数 $k=5$ として k 平均法を適用すると次のように階層的クラスタリングの結果に類似した5分類が得られる¹²⁾。

12) 階層的クラスタリングの分類結果と k 平均法による分類結果の違いは、後者におけるクラスタ4の先頭2国が前者においてクラスタ3へ移行している点である。

表4 OECD 諸国のジニ係数

国名略語	1995年頃	国名略語	2010年頃
SWE	0.211	DNK	0.248
DNK	0.215	NOR	0.250
FIN	0.218	CZE	0.256
AUT	0.238	SVK	0.257
NOR	0.243	BEL	0.259
CZE	0.257	FIN	0.259
LUX	0.259	SWE	0.259
DEU	0.266	AUT	0.261
FRA	0.277	HUN	0.272
BEL	0.287	LUX	0.288
CAN	0.289	FRA	0.293
HUN	0.294	IRL	0.293
NLD	0.297	NLD	0.294
AUS	0.309	DEU	0.295
JPN	0.323	ISL	0.301
IRL	0.324	CHE	0.303
NZL	0.335	POL	0.305
GBR	0.336	GRC	0.307
GRC	0.336	KOR	0.315
ESP	0.343	ESP	0.317
ITA	0.348	CAN	0.324
PRT	0.359	JPN	0.329
USA	0.361	NZL	0.330
TUR	0.490	AUS	0.336
MEX	0.519	ITA	0.337
		GBR	0.345
		PRT	0.353
		USA	0.378
		TUR	0.409
		MEX	0.476

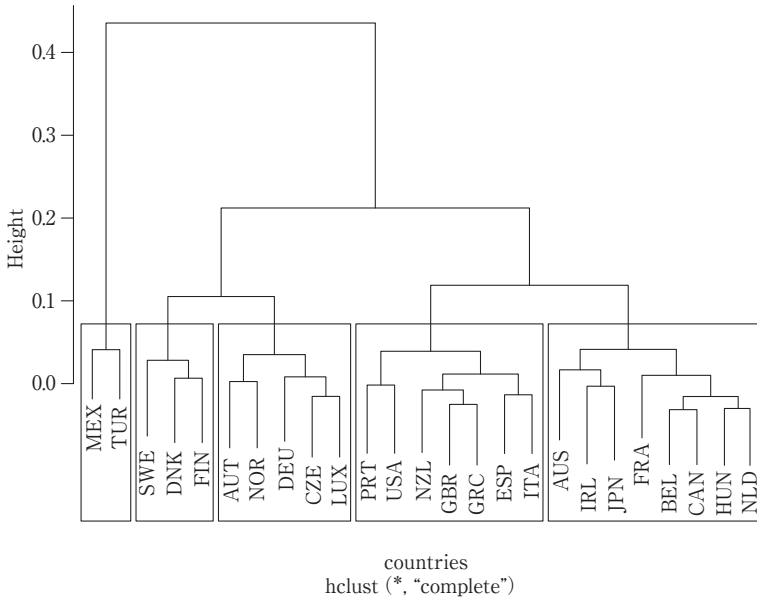
(資料) 1995年頃については OECD, StatExtracts より抜粋 (12 May 2012). 2010年頃については OECD (2011) Table A 1.1. (p.45) より抜粋。

(注) 表2に同じ。

- 1) スウェーデン, デンマーク, フィンランド: 3カ国クラスタ平均 = 0.215
- 2) オーストリア, ノルウェー, チェコ, ルクセンブルク, ドイツ: 5カ国クラスタ平均 = 0.253
- 3) フランス, ベルギー, カナダ, ハンガリー, オランダ, オーストラリア: 6カ国クラスタ平均 = 0.292

図3 ジニ係数の樹形図 (1995年頃)

Cluster Dendrogram



(資料) 表4により計測・作成。

(注) 表2に同じ。

4) 日本, アイルランド, ニューージーランド, イギリス, ギリシャ, スペイン, イタリア, ポルトガル, アメリカ: 9カ国クラスタ平均=0.341

5) トルコ, メキシコ: 2カ国クラスタ平均=0.505

1995年頃にジニ係数が一番低い第1クラスタには北欧諸国が属し, 次の第2クラスタには大陸欧州3カ国, 北欧1カ国および東欧1カ国が属している。ジニ係数が中程度の第3クラスタには大陸欧州3カ国, 東欧1カ国およびアングロ・サクソン2カ国が属している。第4クラスタには日本, アングロ・サクソン4カ国および南欧4カ国が属している。この時期には南欧諸国がアングロ・サクソン諸国よりも不平等な場合がある。表4のOECD30カ国のジニ係数(2010年頃)に階層的クラスタリングが適用されて得られた樹形図は吉岡(2014a, p.195)において公表されており, そこで明らかにされたk平均法に

よる国家の5分類を次に再掲する。

- 1) デンマーク, ノルウェー, チェコ, スロバキア, ベルギー, フィンランド, スウェーデン, オーストリア, ハンガリー: 9カ国クラスタ平均=0.258
- 2) ルクセンブルク, フランス, アイルランド, オランダ, ドイツ, アイスランド, スイス, ポーランド, ギリシャ, 韓国, スペイン: 11カ国クラスタ平均=0.301
- 3) カナダ, 日本, ニュージーランド, オーストラリア, イタリア, イギリス, ポルトガル: 7カ国クラスタ平均=0.336
- 4) アメリカ, トルコ: 2カ国クラスタ平均=0.394
- 5) メキシコ: 1カ国クラスタ平均=0.476

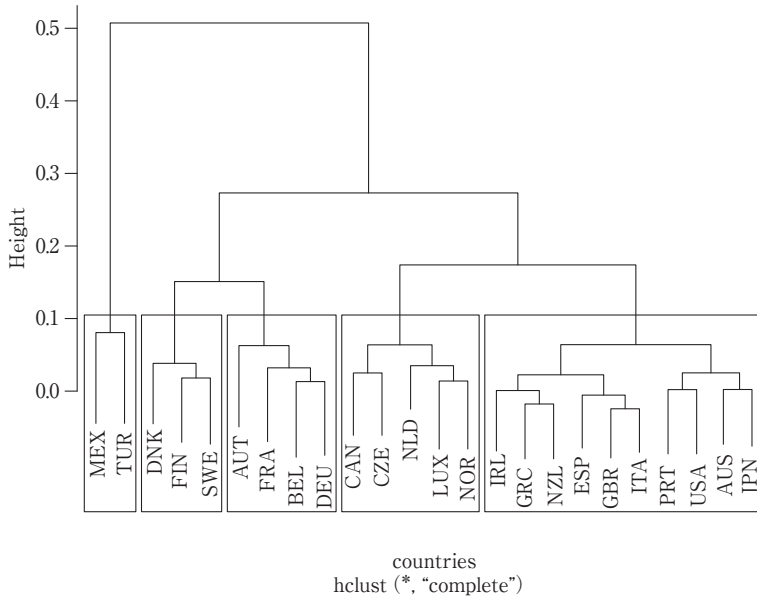
ジニ係数が一番低い第1クラスタには北欧4カ国, 東欧3カ国および大陸欧州2カ国が属し, 特に北欧4カ国とオーストリアではジニ係数自体の上昇が観察される。第2クラスタには大陸欧州5カ国, アングロ・サクソン1カ国, 北欧1カ国, 東欧1カ国, 南欧2カ国およびアジア1カ国が属し, 特に大陸欧州3カ国(ルクセンブルク, フランス, オランダ)ではジニ係数が上昇している。第3クラスタにはアングロ・サクソン4カ国, 南欧2カ国および日本が所属し, 特にアングロ・サクソン3カ国と日本ではジニ係数自体の上昇が観察される。第4クラスタにはジニ係数の上昇が見られたアメリカが属している。

3. OECD 諸国における社会支出比・不平等度と日本の位置

社会支出比とジニ係数を同時に評価して国々を分割するとどのような分類になるだろうか。2つの指標で2次元データを形成するなら, 1995年頃についてはジニ係数の25カ国データに制約を受け, かつ社会支出比の29カ国データの内前者にない1カ国(ハンガリー)データを除外する必要があるから, 24カ国データを利用することになる。この2次元データを用いて OECD 諸国の階層的クラスタリングを行った結果が図4の樹形図である。図4から OECD 24カ国の5分類が読み取れるが, クラスタ数 $k=5$ として k 平均法を適用すると次の

図4 ジニ係数・社会支出比の樹形図 (1995年頃)

Cluster Dendrogram



(資料) 表2及び表4により計測・作成。

(注) 表2に同じ。

ように前者と類似の5分類が得られる。

- 1) スウェーデン, デンマーク, フィンランド: 3カ国
- 2) オーストリア, ノルウェー, ドイツ, フランス, ベルギー, オランダ: 6カ国
- 3) チェコ, ルクセンブルク, カナダ: 3カ国
- 4) オーストラリア, 日本, アイルランド, ニュージーランド, イギリス, ギリシャ, スペイン, イタリア, ポルトガル, アメリカ: 10カ国
- 5) トルコ, メキシコ: 2カ国

第1クラスタ, 第4クラスタおよび第5クラスタは階層的クラスタリングによる分類結果と同じである。24カ国データにおけるジニ係数と社会支出比との相関係数が -0.89 と計算されるので, この2つの指標間には強い負の相関関係

がある。そこで、ジニ係数が低いほど、また社会支出比が高いほど第1クラスタに所属するように分類されていて、ジニ係数が高く社会支出比が低い国家は第5クラスタに所属する。第2クラスタと第3クラスタの所属国家はクラスタリング法によって少し異なるが、この2つのクラスタを合併した場合の所属国家は同一である。合併クラスタには北欧1カ国、大陸欧州6カ国、東欧1カ国およびアングロ・サクソン1カ国が属している。また第4クラスタにはアングロ・サクソン5カ国、日本および南欧4カ国が属している。このように我が国は福祉レジームに関して自由主義と家族主義の両面を持つといわれる位置に分類される。

次に、2010年頃に関する2つの指標の30カ国データから、1995年頃の2つの指標に対応する24カ国データを抽出し、それらに階層的クラスタリングを適用すると図5の樹形図を得る。また、クラスタ数 $k=5$ として k 平均法を適用すると次のように階層的クラスタリングの適用結果と類似の5分類が得られる。

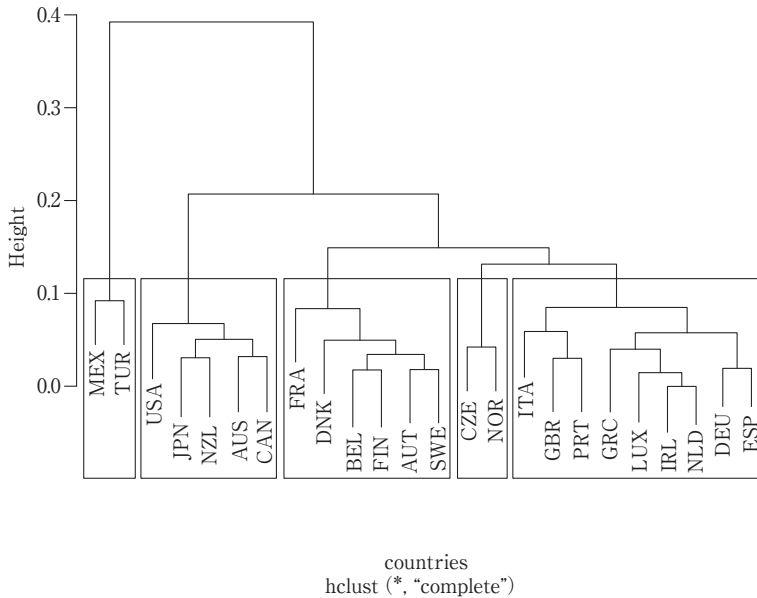
- 1) デンマーク、ベルギー、フィンランド、スウェーデン、オーストリア、フランス、ドイツ：7カ国
- 2) ノルウェー、チェコ、ルクセンブルク、アイルランド、オランダ、ギリシャ：6カ国
- 3) スペイン、イタリア、イギリス、ポルトガル：4カ国
- 4) カナダ、日本、ニュージーランド、オーストラリア、アメリカ：5カ国
- 5) トルコ、メキシコ：2カ国

第1クラスタには北欧3カ国に加え1995年頃には第2クラスタに属していた大陸欧州4カ国が入り、第2クラスタへは1995年頃には第3クラスタに属していた大陸欧州1カ国および東欧1カ国が移行し、第3クラスタへは第4クラスタに属していた南欧3カ国が移行している。つまりジニ係数の低下を福祉の上昇とみるならば¹³⁾、1995年頃に比べると2010年頃には大陸欧州諸国、東欧1カ国および南欧諸国では福祉が相対的に上昇していると解釈されるし、第4クラスタに留まった日本およびアングロ・サクソン4カ国では福祉が相対的に低下しているか停滞していると解釈されよう。

13) 不平等の測定論において、Atkinson (1970) や Sen (1973) のように社会的厚生で不平等を評価する立場がある。

図5 ジニ係数・社会支出比の樹形図 (2010年頃)

Cluster Dendrogram



(資料) 表2及び表4により計測・作成。

(注) 表2に同じ。

4. 適切なクラスタ数の候補とその適用

4.1 適切なクラスタ数の候補

前節まではこの種の研究分野の習慣に従って OECD 諸国を5つのクラスタに分割してきた。最適なクラスタ数 k を一意的に決める方法がないので、その目安を探るために数十の指標が提案されているが¹⁴⁾、この分野の応用においてほとんどすべての統計量が同一結果をもたらすことはめったにない。そこで、比較的よく利用される4つの評価指標によって適切なクラスタ数の候補を複数見出すことにする。クラスタリング結果の評価指標として、3つのクラスタリング法（階層的、 k 平均、 pam ¹⁵⁾）それぞれにたいし3つの評価指標（Connec-

14) Milligan = Cooper (1985), Rousseeuw (1987), Tibshirani et al. (2001).

tivity, Dunn index, Silhouette width)¹⁶⁾と k 平均法にたいする評価指標としての Pham 値¹⁷⁾とが採用される。Connectivity はクラスタ間の結合の程度を測る指標であり、データがその最近傍と同一のクラスタ内に位置している程度を表し、ゼロと無限大の間の値を取り無限大に近いほうが望ましい¹⁸⁾。Dunn 指標¹⁹⁾とシルエット幅²⁰⁾はクラスタ間の結合性と分離性とを統合した指標である。Dunn 指標はクラスタ内最大距離に対するそのクラスタにないデータ間の最小距離の比率であり、ゼロと無限大の間の値を取り無限大に近いほうが望ましい。シルエット幅はそれぞれのデータのシルエット値の平均であり、シルエット値は特定のクラスタリング結果の信頼度を測り、区間 $[-1, 1]$ 内の値を取ることで、シルエット幅が 1 に近いほどデータのクラスタリングに成功しており、 -1 に近いほど失敗に近いことを表す。

表 5 はそれぞれの評価指標の計算結果を基に導出された適切なクラスタ数の候補を表している。社会支出比による 1995 年頃の OECD 29 カ国の場合の分類以外は、5 分類は適切な分割の 1 つのようである。3 分類は本稿で取り上げたすべての場合に適切であるが、福祉国家の分類という観点からは大まか過ぎる分類かも知れない。表中のジニ係数・社会支出比に関する 2010 頃の OECD 諸国は 1995 年頃との比較のために 24 カ国しか採用されていない。ジニ係数・社会支出比の場合、2010 頃の OECD 30 カ国に関する適切なクラスタ数の候補は次のようになる ($3 \leq k \leq 8$)。

Connectivity : 3

Dunn 指標 : 3, 5, 6, 7

シルエット幅 : 3

Pham : 7

15) Kaufman = Rousseeuw (1990) において提案された k-medoids 法の 1 つ。

16) Brock et al. (2008).

17) Pham et al. (2005), 吉岡 (2014a).

18) Handl = Knowles (2005).

19) Dunn (1974).

20) Rousseeuw (1987).

表5 適切なクラスタ数 k の候補
 $3 \leq k \leq 8$

	1995年頃	2010年頃
評価指標	ジニ係数	ジニ係数
Connectivity	3	3
Dunn	5, 7, 8	5, 8
Silhouette	5	4, 5
Pham	3, 4, 7	5
評価指標	社会支出比	社会支出比
Connectivity	3	3
Dunn	6, 8	3, 4, 5
Silhouette	3	3
Pham	3	3
評価指標	ジニ係数・社会支出比	ジニ係数・社会支出比
Connectivity	3	3
Dunn	5, 6	8
Silhouette	3	5
Pham	3, 4	3, 4, 8

(資料) 表2及び表4により計測・作成。

(注) 評価指標については本文参照。

4.2 k 平均法の適用

そこで、2010年頃の OECD30カ国の2次元データによっていくつかの分類を試みる。クラスタ数 $k=3$ として OECD30カ国データに k 平均法を適用すると次のように3分類が得られる。第1クラスタからジニ係数の昇順に分類されており、それが社会支出比の降順にも対応している。2つの指標間には負の相関関係があるからで、3クラスタ分類のように大きな区分の場合にはその関係がクラスタ平均間でも成り立っている(表6)。

- 1) デンマーク, ベルギー, フィンランド, スウェーデン, オーストリア, フランス, ドイツ: 7カ国
- 2) ノルウェー, チェコ, ハンガリー, ルクセンブルク, アイルランド, オランダ, ギリシャ, アイスランド, スロバキア, スイス, ポーランド, スペイン, イタリア, カナダ, 日本, ニューージーランド, イギリス, オーストラリア, ポルトガル, アメリカ: 20カ国
- 3) 韓国, トルコ, メキシコ: 3カ国

表6 OECD30カ国のクラスタ平均

(2010年頃)

3分類	ジニ係数	社会支出比
第1クラスタ	0.2677	0.2949
第2クラスタ	0.3088	0.2216
第3クラスタ	0.4000	0.1003
4分類	ジニ係数	社会支出比
第1クラスタ	0.2677	0.2949
第2クラスタ	0.2842	0.2178
第3クラスタ	0.3388	0.2262
第4クラスタ	0.4000	0.1003
5分類	ジニ係数	社会支出比
第1クラスタ	0.2677	0.2949
第2クラスタ	0.2825	0.2216
第3クラスタ	0.3190	0.1595
第4クラスタ	0.3413	0.2386
第5クラスタ	0.4425	0.1045
6分類	ジニ係数	社会支出比
第1クラスタ	0.2588	0.2145
第2クラスタ	0.2677	0.2949
第3クラスタ	0.3061	0.2243
第4クラスタ	0.3308	0.1672
第5クラスタ	0.3380	0.2590
第6クラスタ	0.4425	0.1045
7分類	ジニ係数	社会支出比
第1クラスタ	0.2588	0.2145
第2クラスタ	0.2632	0.2988
第3クラスタ	0.3061	0.2243
第4クラスタ	0.3163	0.2717
第5クラスタ	0.3190	0.1595
第6クラスタ	0.3587	0.2300
第7クラスタ	0.4425	0.1045

(資料) 表5に同じ。

24カ国データの場合と同様に第1クラスタには北欧3カ国に加え大陸欧州4カ国が入り、第3クラスタには非英語圏かつ非欧州の3カ国が入るが、第2クラスタには多数の国家が入る。表6に見るようにクラスタの分類数が大きくなるにつれ、2つの指標のクラスタ平均間での負の関係が成り立つとは限らない点に注意を要するが²¹⁾、クラスタ数 $k=5$ として OECD30カ国データに k 平均法を適用すると次のように5分類が得られる。

- 1) デンマーク, ベルギー, フィンランド, スウェーデン, オーストリア, フランス, ドイツ: 7カ国
- 2) ノルウェー, チェコ, ハンガリー, ルクセンブルク, アイルランド, オランダ, ギリシャ, スロバキア, スイス, ポーランド: 10カ国
- 3) アイスランド, カナダ, オーストラリア, 韓国: 4カ国
- 4) スペイン, イタリア, 日本, ニュージーランド, イギリス, ポルトガル, アメリカ: 7カ国
- 5) トルコ, メキシコ: 2カ国

3分類の場合と同様に第1クラスタには北欧3カ国に加え大陸欧州4カ国が入り、第2クラスタには北欧1カ国、東欧4カ国、大陸欧州3カ国、南欧1カ国およびアングロ・サクソン1カ国が入る。第4クラスタはジニ係数が2番目に高いが社会支出比は第1クラスタに次いで2番目に高い。第3クラスタはジニ係数が3番目の高さだが社会支出比は4番目の高さなので、社会支出比を基準にすると第4クラスタの方が第3クラスタや第2クラスタよりも福祉が高いと解釈することが可能である。その第4クラスタにはジニ係数が2番目に高く社会支出比が2番目に高い南欧3カ国、日本およびアングロ・サクソン3カ国が属している。同じようにOECD30カ国データにk平均法を適用すると表6に見るように6分類も7分類も得られる。クラスタ数が増えるにつれOECD30カ国が篩に掛けられたように、各クラスタが少数の要素をもつようになりクラスタ順はあまり意味がなくなるが、2つの指標に関してより類似の国々が集まってくる。したがって、どのクラスタに様々なタイプの福祉国家としての個々の国々が所属するかというより、あるクラスタの個々の国々あるいは国々の集合がどのようなタイプの福祉国家の集合に属しているかという観点から検討する必要がある。OECD30カ国の3分類から7分類によって次の考察結果が得られる。

- (1) 大陸欧州の主要国としてのベルギー, オーストリア, フランスおよびドイツは、不平等が一番低く社会支出比が一番高い北欧3カ国（デンマーク,

21) 適切なクラスタ数ごとの各々のクラスタ順位とクラスタ内平均の順位との整合性の検討が必要になる。

フィンランド、スウェーデン)と類似の福祉構造を持つ。

- (2) 東欧3カ国(チェコ、ハンガリー、スロバキア)は、不平等は低く社会支出比が相対的に低い福祉構造を持ち、北欧ノルウェーは同じクラスタに属する。
- (3) 不平等が高く社会支出比が相対的に低い我が国は、5分類の場合スペイン、イタリアおよびポルトガルの南欧家族主義とニュージーランド、イギリスおよびアメリカの自由主義との混合福祉レジームに属するが、6分類や7分類の場合はルクセンブルク、オランダおよびスイスの保守主義とアイルランドおよびニュージーランドの自由主義との混合福祉レジームに属する。
- (4) 不平等が相対的に高く社会支出比が低い韓国は、カナダおよびオーストラリアのような自由主義国家の側面を持つ。

お わ り に²²⁾

第1節において社会支出比によるOECD30カ国のクラスタリング結果が1995年と2010年について各国の高齢化率との関連で比較された²³⁾。2010年の社会支出比によるOECD30カ国の第1クラスタから第5クラスタまでを降順に地理的あるいは言語的(人種的)に区分すると次のようになるが、第1クラスタから第4クラスタまでの27カ国ですら福祉レジームで単純に区分することは困難である。

- 1) 北欧2カ国および大陸欧州3カ国:
- 2) 北欧1カ国, 大陸欧州1カ国および南欧3カ国:
- 3) アングロ・サクソン3カ国, 大陸欧州2カ国, 南欧1カ国, 北欧1カ国, 東欧2カ国および日本:
- 4) アングロ・サクソン3カ国, 東欧2カ国, 大陸欧州1カ国および北欧1

22) 本稿における数値計算及びグラフ作成にはR言語・環境が利用された。

23) 1995年に高い社会支出比の国々(北欧3国, フランス, ドイツ, オーストリア, ベルギー)は概して高齢化率が高く, 1995年に社会支出比が低い国々(トルコ, メキシコ, 韓国)は高齢化率が低い。

カ国：

5) 非欧州・非英語圏 3 カ国：

第2節においてジニ係数による OECD 諸国のクラスタリング結果が1995年頃と2010年頃について比較された。この2時点間に OECD25カ国中13カ国で不平等度の上昇が見られ、12カ国で低下が観察され、不平等度の上昇が見られたのは北欧4カ国、大陸欧州4カ国、アングロ・サクソン4カ国および日本においてであり、低下が見られたのは東欧2カ国、大陸欧州2カ国、アングロ・サクソン2カ国および南欧4カ国においてである。2010年頃のジニ係数による OECD30カ国の第1クラスタから第5クラスタまでを昇順に地理的あるいは言語的（人種的）に区分すると次のようになる。

1) 北欧4カ国，東欧3カ国および大陸欧州2カ国：

2) 北欧1カ国，東欧1カ国，大陸欧州5カ国，アングロ・サクソン1カ国，南欧2カ国およびアジア1カ国：

3) アングロ・サクソン4カ国，南欧2カ国および日本：

4) アングロ・サクソン1カ国および非欧州・非英語圏1カ国：

5) 非欧州・非英語圏1カ国：

第3節において社会支出比とジニ係数の2次元データによる OECD24カ国のクラスタリング結果が1995年頃と2010年頃について比較された。5分類が採用された場合、ジニ係数の低下を福祉の上昇とみるならば、1995年頃に比べると2010年頃には大陸欧州諸国、東欧1カ国および南欧諸国では福祉が相対的に上昇していると解釈されるし、第4クラスタに留まった日本およびアングロ・サクソン4カ国では福祉が相対的に低下しているか停滞していると解釈されよう。

第3節まではこの種の研究分野の習慣に従って OECD 諸国が5つのクラスタに分割され、分析が行われた。そこで第4節では、比較的よく利用される4つの評価指標によって適切なクラスタ数の候補が探求された結果、3分類は本稿で取り上げたすべての場合に適切であり、また社会支出比による1995年頃の OECD29カ国の場合の分類以外は、5分類は適切な分割の1つであることが分かった。結局本稿で取り上げられた OECD 諸国の場合、その目的に応じて3分類から8分類の中から分類数を選択できる。しかし、2次元データとして社

会支出比とジニ係数を同時に評価する場合には、クラスタの分類数が大きくなるにつれ、2つの指標のクラスタ平均間での明確な関係が成り立つとは限らず、分類結果の解釈が曖昧になりがちになる点に注意を要する。ジニ係数・社会支出比による、2010頃のOECD24カ国の5分類の結果と大筋では変わらないが（第3クラスタと第4クラスタとの順位を問わねば）、最後にこの点を例示するために、2010頃のOECD30カ国の5分類が以下に報告される。

- 1) 北欧3カ国および大陸欧州4カ国：
- 2) 北欧1カ国，東欧4カ国，大陸欧州3カ国，南欧1カ国およびアングロ・サクソン1カ国：
- 3) 北欧1カ国，アングロ・サクソン2カ国およびアジア1カ国：
- 4) 南欧3カ国，アングロ・サクソン3カ国および日本：
- 5) 非欧州・非英語圏2カ国：

参考文献

- Atkinson, A. B. (1970). On the Measurement of Inequality, *Journal of Economic Theory*, 2, 244-263.
- Atkinson, A. B., Rainwater, L. and Smeeding, T. (1995). *Income Distribution in OECD Countries: The evidence from the Luxembourg Income Study*, Paris: OECD.
- Brock, G., Pihur, V., Datta, S. and Datta, S. (2008). clValid: An R package for cluster validation, *Journal of Statistical Software*, 25, 1-22.
- Castles, F. G. (2008). What Welfare States Do: A disaggregated expenditure approach, *Journal of Social Policy*, 38, 45-62.
- Dunn, J. C. (1974). Well Separated Clusters and Fuzzy Partitions, *Journal on Cybernetics*, 4, 95-104.
- Esping-Andersen, G. (1990). *The Three Worlds of Welfare Capitalism*, Cambridge: Polity Press.
- Everitt, B. S., Landau, S. and Leese, M. (2001). *Cluster Analysis, 4th ed.*, New York: Oxford University Press.
- Handl, J. and Knowles, J. (2005). Exploiting the Trade-off — The benefits of multiple objectives in data clustering, in Coello, L. A. et al. (eds), *Proceedings of the Third International Conference on Evolutionary Multicriterion Optimization*, Berlin: Springer-Verlag, 547-560.
- Hastie, T., Tibshirani, R. and Friedman, J. (2009). *The Elements of Statistical Learning*, 2nd ed., New York: Springer.
- Kaufman, L. and Rousseeuw, P. J. (1990). *Finding Groups in Data: An Introduction to Cluster Analysis*, New York: John Wiley.
- Korpi, W. and Palme, J. (1998). The Paradox of Redistribution and Strategies of Equality: Welfare State Institutions, Inequality, and Poverty in the Western Countries, *American Socio-*

- logical Review*, 63, 661-687.
- Kovacs, I. (1985). International Comparison of Consumption Patterns by Cluster Analysis, *Acta Oeconomica*, 35, 313-326.
- Obinger, H. and Wagshal, U. (2012). Social Expenditure and Revenues, in Castles et al. (eds), *The Oxford Handbook of the Welfare State*, Oxford : Oxford University Press, 333-352.
- Milligan, G. and Cooper, M. (1985). An Examination of Procedures for Determining the Number of Clusters in a Data Set, *Psychometrika*, 50, 159-179.
- OECD. (2008). *Growing Unequal? : Income Distribution and Poverty in OECD Countries*, Paris : OECD.
- OECD. (2011). *Divided We Stand : Why Inequality Keeps Rising*, Paris : OECD.
- Pham, D. T., Dimov, S. and Nguyen, C. D. (2005). Selection of k in K-means Clustering, *Mechanical Engineering Science*, 219, C, 103-119.
- Rousseeuw, P. J. (1987). Silhouettes : A Graphical Aid to the Interpretation and Validation of Cluster Analysis, *Computational and Applied Mathematics*, 20, 53-65.
- Sen, A. K. (1973). *On Economic Inequality*, Oxford : Oxford University Press.
- Tibshirani, R., Walther, G. and Hastie, T. (2001). Estimating the Number of Clusters in a Data Set via the Gap Statistic, *J. Roy. Statist. Soc. B*, 411-423.
- 吉岡慎一 (2014). 貧困及び不平等測度の要因分解と世帯構成の変化『西南学院大学経済学論集』48(3・4), 251-274.
- 吉岡慎一 (2014a). 所得格差の国際比較－クラスター分析－『西南学院大学経済学論集』49(2・3), 191-215.