

所得格差の国際比較 — クラスタ分析 —

吉 岡 慎 一

はじめに

OECD による所得不平等についての比較研究である OECD (2008, 2011) において、2008年までの約30年間 OECD 諸国における所得格差が拡大を続け、過去最大に達していることが報告された。長期のデータ系列が利用できる OECD22カ国中17カ国でジニ係数が上昇し続けているのである。近年の世界的経済危機の始まりまでに、OECD の大多数の国で家計所得の上位10%の所得上昇率が下位10%の所得上昇率を上回ったことが貧富の差の拡大の原因の一つであり、さらに、両報告書において不平等度の上昇の背景にあるこれ以外の主因についても明らかにされている。

しかし、OECD 諸国間で所得格差の水準に依然として大きな差異があることは確かなので、本稿の主目的は2000年代の OECD30カ国を所得格差によりいくつかのクラスタに分類することである。OECD 諸国を所得不平等度や所得貧困度で別々に分類することはそれほど困難ではない。しかし、所得格差は多次元概念だから例えば不平等度と貧困度とで同時に把握する必要がある。その場合の所得格差（不平等度・貧困度）を直接的に計測することは容易ではないが、それにより OECD 諸国をクラスタリングすることはできる。そこで、本稿では2000年代の2時点における OECD 諸国が所得格差（不平等度・貧困度）によりいくつかのクラスタに分割された結果の比較が試みられ、所得格差の数的特徴のみによってクラスタリングされた結果がどのような非数的特徴で分割されているのかが明らかにされる。

1. 先進国における所得格差

1.1 先進国における不平等度と日本の位置

我が国を含めた所得分配の国際比較ができる資料はデータの制約が厳しいのでほとんどない。国によって所得の定義やデータの作成法がいくぶん異なる。このような状況の中で、先進国に限定すれば OECD が提供する統計資料は比較的信頼性が高い。OECD が保有するデータは共通の手法と標準化された定義を用いて各国の専門家により作成されたものである。OECD (2008) は2000年代半ばにおける OECD30カ国を家計の等価可処分所得のジニ係数の小さいほうから大きい順に次のように5つにグループ化している。

- 1) デンマーク, スウェーデン: OECD30カ国平均 (0.31) を25%以上下回る。
- 2) ルクセンブルク, オーストリア, チェコ, スロバキア, フィンランド, ベルギー, オランダ, スイス, ノルウェー, アイスランド, フランス, ハンガリー, ドイツ, オーストラリア: OECD30カ国平均を3~17%下回る。
- 3) 韓国, カナダ, スペイン, 日本, ギリシャ, アイルランド, ニューゼーランド, イギリス: OECD30カ国平均を1~8%上回る。
- 4) イタリア, ポーランド, アメリカ, ポルトガル: OECD30カ国平均を13~24%上回る。
- 5) トルコ, メキシコ: OECD30カ国平均をそれぞれ38%, 52%上回る。

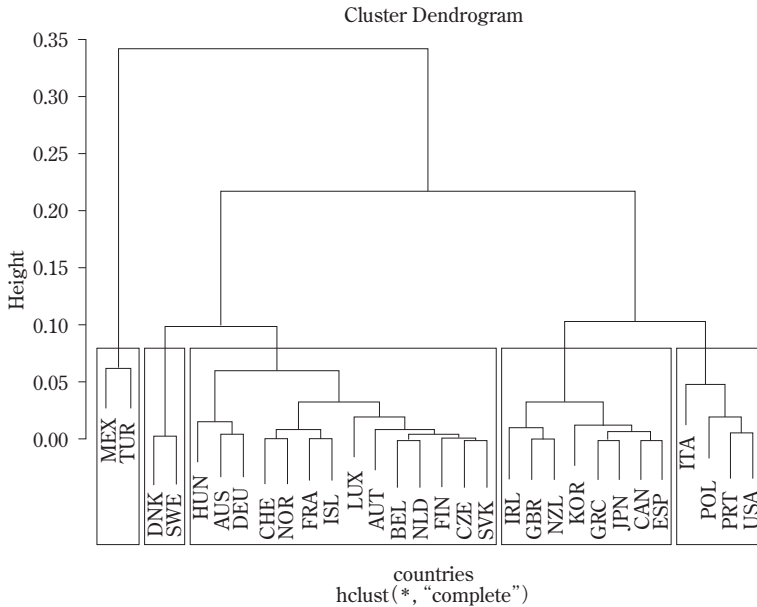
このように我が国のジニ係数 (0.32) は OECD 諸国の不平等度の平均値より3.2%以上高く、第3の不平等グループに属していて、OECD30カ国中、高いほうから11番目に位置している¹⁾。

OECD (2008)²⁾に掲載された等価可処分所得のジニ係数に階層的クラスタリングを適用すると³⁾、この OECD 報告による5分類が確認される。図1は OECD30カ国のジニ係数に階層的クラスタリングが適用されて得られた樹形図

1) 『国民生活基礎調査』(厚労省)の個票データが利用されている。

2) Table 1.A2.2. (p.51)。

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



(資料) OECD (2008) の Table 1. A2.2. (P.51) に掲載された等価可処分所得のジニ係数により計測・作成。

(注) 国名略語は表1を参照のこと。

図1 ジニ係数の樹形図 (2000年代半ば)

(デンドログラム) であり⁴⁾、この図から上と同一の5分類が抽出される。

上の分類結果との比較のためにクラスタ数 $k=5$ として非階層的クラスタリング法の一つである k 平均法を適用する。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が必ずしも同一ではないが、2000年代半ばにおける k 平均法による5分類は次のようになる⁵⁾。

- 1) デンマーク、スウェーデン、ルクセンブルク、オーストリア、チェコ、スロバキア、フィンランド、ベルギー、オランダ：9カ国クラスタ平均 = 0.260

4) 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。

5) 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。

表1 国名略語等

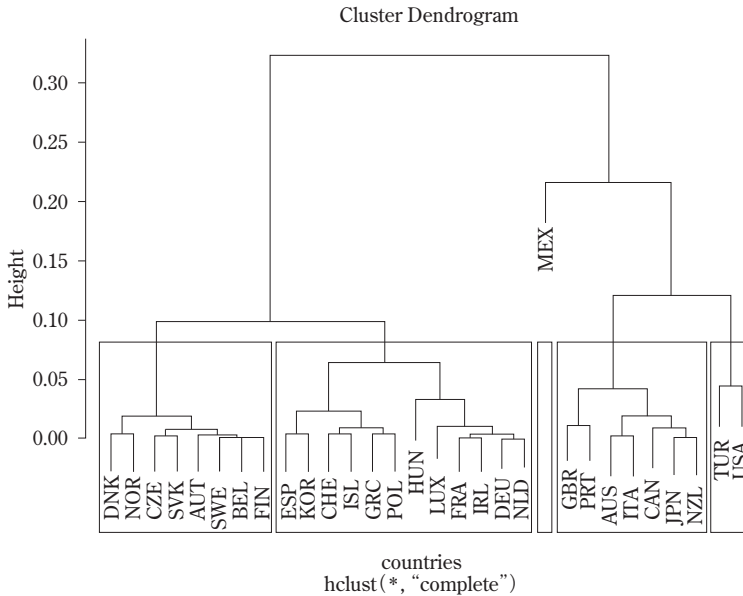
国名番号	国名略語	国名	国名番号	国名略語	国名
1	AUS	オーストラリア	16	ISL	アイスランド
2	AUT	オーストリア	17	ITA	イタリア
3	BEL	ベルギー	18	JPN	日本
4	CAN	カナダ	19	KOR	韓国
5	CHE	スイス	20	LUX	ルクセンブルク
6	CZE	チェコ	21	MEX	メキシコ
7	DEU	ドイツ	22	NLD	オランダ
8	DNK	デンマーク	23	NOR	ノルウェー
9	ESP	スペイン	24	NZL	ニュージーランド
10	FIN	フィンランド	25	POL	ポーランド
11	FRA	フランス	26	PRT	ポルトガル
12	GBR	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

- 2) スイス, ノルウェー, アイスランド, フランス, ハンガリー, ドイツ, オーストラリア: 7カ国クラスタ平均=0.286
- 3) 韓国, カナダ, スペイン, 日本, ギリシャ, アイルランド, ニュージーランド, イギリス: 8カ国クラスタ平均=0.324
- 4) イタリア, ポーランド, アメリカ, ポルトガル: 4カ国クラスタ平均=0.373
- 5) トルコ, メキシコ: 2カ国クラスタ平均=0.452

2000年代半ばにおけるk平均法による5分類と階層的クラスタリングによる5分類とはほとんど同じで、違いは最も平等なデンマークおよびスウェーデンを第1クラスタとして独立に分類するか否かという点にある。トルコおよびメキシコのジニ係数は外れ値のごとく異常に高いので、第5クラスタとして独立に分類されている。

2000年代末のOECD34カ国の等価可処分所得のジニ係数がOECD(2011)のTable A1.1.(p.45)に掲載されている。2000年代半ばの可処分所得のジニ係数との比較のために上と同じ30カ国を選び、階層的クラスタ分析によってジニ係数の小さいほうから昇順に5つにグループ化すると次のようになる⁶⁾。

- 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



(資料) OECD (2011) の Table A1.1. (P.45) に掲載された等価可処分所得のジニ係数により計測・作成。

(注) 図1に同じ。

図2 ジニ係数の樹形図 (2000年代末)

ド、スウェーデン、オーストリア：

2) ハンガリー、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：

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：

4) アメリカ、トルコ：

5) メキシコ：

2000年代半ばと2000年代末との短期間でも不平等順位の入替わりがかなりあるが、我が国はすくなくとも OECD30カ国平均のジニ係数 (0.308) を上回

6) ユークリッド距離を用い完全連結法で作成された階層的クラスターを示す図2を参照。

る第3クラスタに属し、2000年代末には OECD30カ国中、高いほうから9番目に位置するようになった。これは我が国のジニ係数が2000年代半ばの0.32から2000年代末の0.33へ上昇したことによる⁷⁾。

2000年代末における k 平均法による 5 分類も次のように、上記とほとんど同じである。

- 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

2000年代半ばから2000年代末にかけての短期間に、ジニ係数のクラスタ平均はどのクラスタにおいても上昇している。特にスウェーデン、ルクセンブルク、スイスおよびオーストラリアそれぞれのジニ係数の上昇が目立つようである。

1.2 先進国における貧困度と日本の位置

OECD (2008) の Figure 5.1⁸⁾を作成するために利用された相対的貧困率⁹⁾をまとめたのが表2である。この OECD30カ国の貧困率で作成されたのが図3の樹形図であり、これを基に2000年代半ばの貧困率によって昇順に5つにクラスタリングすると次のようになる。

7) 『国民生活基礎調査』の公表集計データが利用された吉岡 (2013) によると、2000年代中期から2000年代後期にかけての世帯総所得の不平等度の変動は、採用する不平等測度や所得分布の開端階級の処理の仕方によって異なる。しかし、開端階級の下限值を利用して計測された通常のジニ係数に従うならば、2000年代後期に不平等度が幾分上昇していると判断される。

8) Relative poverty rates for different income thresholds, mid-2000s.

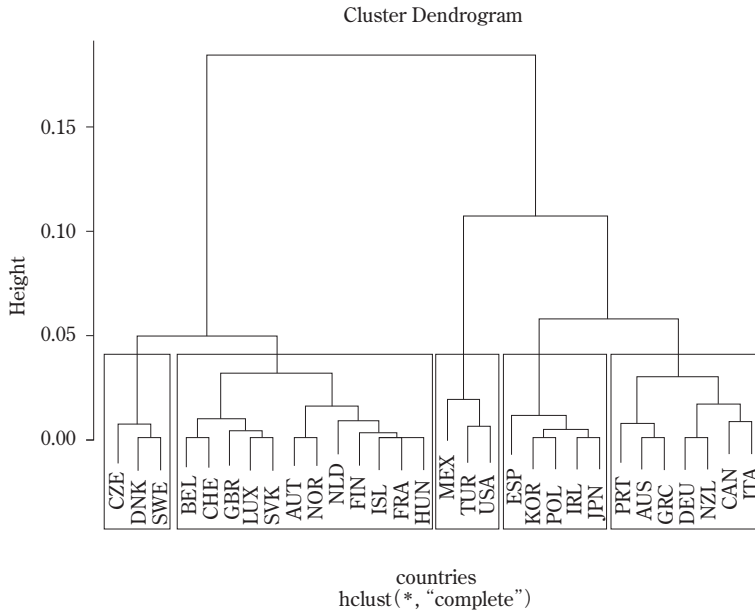
9) OECD, Stat, Last updated: 12-Sep-2008.

表2 OECD 諸国の貧困率（2000年代半ば）

国名	貧困率	国名	貧困率
オーストラリア	0.124	アイスランド	0.071
オーストリア	0.066	イタリア	0.114
ベルギー	0.088	日本	0.149
カナダ	0.120	韓国	0.146
スイス	0.087	ルクセンブルク	0.081
チェコ	0.058	メキシコ	0.184
ドイツ	0.110	オランダ	0.077
デンマーク	0.053	ノルウェー	0.068
スペイン	0.141	ニュージーランド	0.108
フィンランド	0.073	ポーランド	0.146
フランス	0.071	ポルトガル	0.129
イギリス	0.083	スロバキア	0.081
ギリシャ	0.126	スウェーデン	0.053
ハンガリー	0.071	トルコ	0.175
アイルランド	0.148	アメリカ	0.171

(資料) OECD, Growing Unequal? Income Distribution and Poverty in OECD Countries, 2008
により作成。

(注) 貧困線は可処分所得の中央値の50%。



(資料) 表2により計測・作成。

(注) 図1に同じ。

図3 貧困率の樹形図（2000年代半ば）

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

2000年代半ばにおける k 平均法による 5 分類も次のように, 上記とほとんど同じである。

- 1) デンマーク, スウェーデン, チェコ, オーストリア, ノルウェー: クラスタ平均=0.060
- 2) フランス, アイスランド, ハンガリー, フィンランド, オランダ, ルクセンブルク, スロバキア, イギリス, スイス, ベルギー: クラスタ平均=0.078
- 3) ニュージーランド, ドイツ, イタリア, カナダ, オーストラリア, ギリシャ, ポルトガル: クラスタ平均=0.119
- 4) スペイン, ポーランド, 韓国, アイルランド, 日本: クラスタ平均=0.146
- 5) アメリカ, トルコ, メキシコ: クラスタ平均=0.177

どちらの方法によっても, 我が国は貧困率が高い第 4 クラスタに属し, さらに OECD30カ国中, 高いほうから 4 番目に位置している。

表 3 は OECD がネット上に公表しているデータから 2000 年代末についての各国の貧困率を抜粋し, まとめたものである。この OECD30カ国の貧困率で作成されたのが図 4 の樹形図であり, これを基に 2000 年代末の OECD30カ国を貧困率によって昇順に 5 つにクラスタリングすると次のようになる。

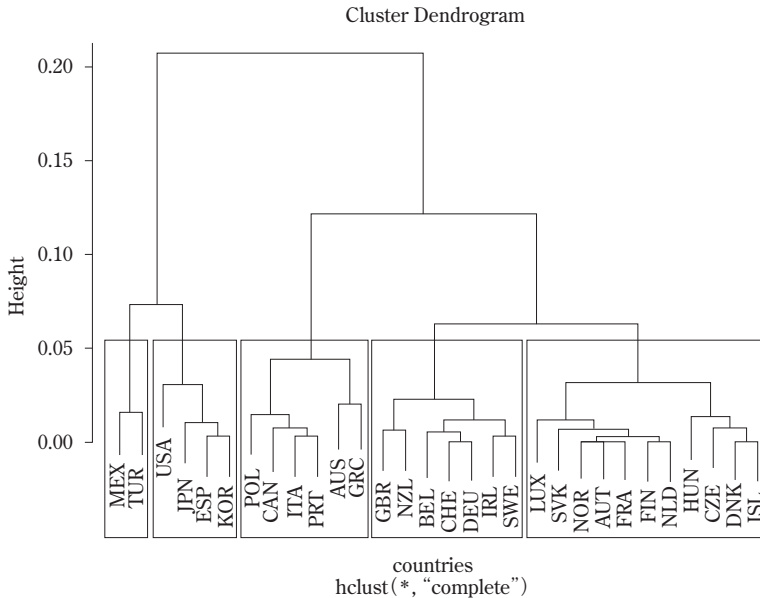
- 1) チェコ, アイスランド, デンマーク, ハンガリー, フィンランド, オランダ, オーストリア, フランス, ノルウェー, スロバキア, ルクセンブルク:

表3 OECD 諸国の貧困率（2000年代末）

国名	貧困率	国名	貧困率
オーストラリア	0.144	アイスランド	0.063
オーストリア	0.075	イタリア	0.121
ベルギー	0.094	日本	0.160
カナダ	0.123	韓国	0.153
スイス	0.095	ルクセンブルク	0.081
チェコ	0.059	メキシコ	0.204
ドイツ	0.095	オランダ	0.074
デンマーク	0.064	ノルウェー	0.075
スペイン	0.155	ニュージーランド	0.103
フィンランド	0.074	ポーランド	0.113
フランス	0.075	ポルトガル	0.120
イギリス	0.099	スロバキア	0.077
ギリシャ	0.130	スウェーデン	0.087
ハンガリー	0.068	トルコ	0.193
アイルランド	0.090	アメリカ	0.174

(資料) OECD, StatExtracts より抜粋 (30 Mar 2014).

(注) 貧困線は可処分所得の中央値の50%。



(資料) 表3により計測・作成。

(注) 図1に同じ。

図4 貧困率の樹形図（2000年代末）

- 2) スウェーデン, アイルランド, ベルギー, スイス, ドイツ, イギリス, ニュージーランド:
- 3) ポーランド, ポルトガル, イタリア, カナダ, ギリシャ, オーストラリア:
- 4) 韓国, スペイン, 日本, アメリカ:
- 5) トルコ, メキシコ:

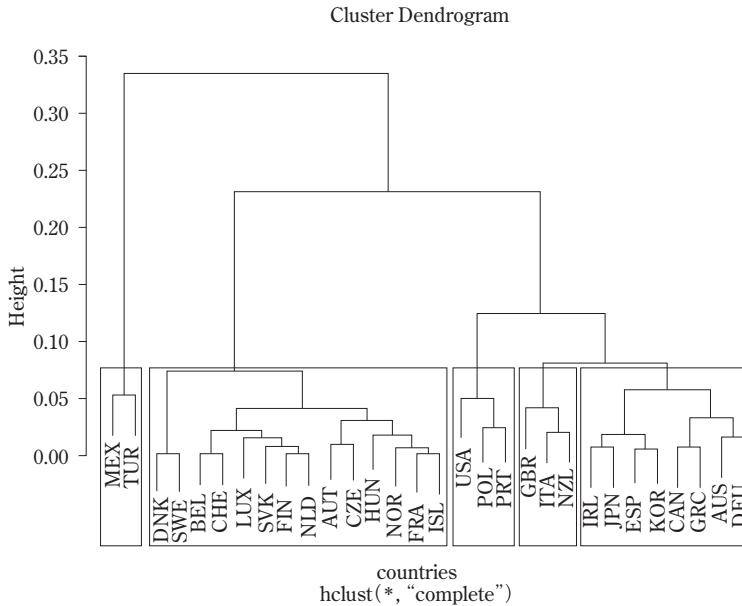
2000年代半ばから2000年代末にかけて、第2クラスタから第1クラスタへかなりの数の国々が移動している中で、我が国は貧困率が高い第4クラスタに属し、さらにOECD30カ国中、高いほうから4番目に依然として位置している。2000年代末におけるk平均法による5分類も次のように、上記とほとんど同じである。

- 1) チェコ, アイスランド, デンマーク, ハンガリー, フィンランド, オランダ, オーストリア, フランス, ノルウェー, スロバキア, ルクセンブルク: クラスタ平均=0.071
- 2) スウェーデン, アイルランド, ベルギー, スイス, ドイツ, イギリス, ニュージーランド: クラスタ平均=0.095
- 3) ポーランド, ポルトガル, イタリア, カナダ, ギリシャ: クラスタ平均=0.121
- 4) オーストラリア, 韓国, スペイン, 日本, アメリカ: クラスタ平均=0.157
- 5) トルコ, メキシコ: クラスタ平均=0.199

2000年代半ばから2000年代末の短期間に、ジニ係数の場合と同様に貧困率のクラスタ平均もどのクラスタにおいても上昇している。特にスウェーデンおよびオーストラリアのそれぞれの貧困率の上昇が目立つ。

1.3 先進国における所得格差（不平等度・貧困度）と日本の位置

所得不平等度は所得分配全体の不平等を捉え、所得貧困度は分配の下部に焦点を絞る。両者には強い関連があるが¹⁰⁾、それぞれ所得分配の異なる側面に光をあてている。所得格差は多次元概念なので、ここでは所得不平等度と貧困



(資料) 図1および図3に同じ。

(注) 図1に同じ。

図5 不平等度・貧困度の樹形図（2000年代半ば）

度とで同時に把握されると想定する。そこで、図1および図3を描くために利用されたデータに階層的クラスタリングを適用すると、2000年代半ばにおける所得格差のデンドログラムを表す図5を得る。この図によると、2000年代半ばにおける OECD 諸国は所得格差によって次のように5つにクラスタリングされる。

- 1) デンマーク、スウェーデン、オーストリア、ベルギー、スイス、チェコ、フィンランド、フランス、ハンガリー、アイスランド、ルクセンブルク、オランダ、ノルウェー、スロバキア：
- 2) オーストラリア、カナダ、ドイツ、スペイン、ギリシャ、アイルランド、日本、韓国：

10) 図1、図2、図3および図4を描くために利用されたデータで計測すると、ジニ係数と貧困率の相関係数は、2000年代半ばで0.86、2000年代末で0.85となる。

3) イギリス, イタリア, ニュージーランド:

4) ポーランド, ポルトガル, アメリカ:

5) メキシコ, トルコ:

2000年代半ばにおける k 平均法による 5 分類も次のように, 上記とほとんど同じで, 分類の違いは先頭のデンマークおよびスウェーデンの 2 カ国を他から分離するか, 最後尾のメキシコおよびトルコの 2 カ国を他から分離するかにある。5 分類は詳細すぎる分類かも知れないので次節で適切なクラスタ数の決定が試みられる。なお, 我が国の所得格差を 2 次元で捉えると OECD30カ国中, 高いほうから 9 番目ないし 10 番目あたりに位置するようであり, これにはジニ係数の結果が強く影響を及ぼしている。

1) デンマーク, スウェーデン:

2) オーストリア, ベルギー, スイス, チェコ, フィンランド, フランス, ハンガリー, アイスランド, ルクセンブルク, オランダ, ノルウェー, スロバキア:

3) オーストラリア, カナダ, ドイツ, スペイン, ギリシャ, アイルランド, 日本, 韓国:

4) イギリス, イタリア, ニュージーランド:

5) ポーランド, ポルトガル, アメリカ, メキシコ, トルコ:

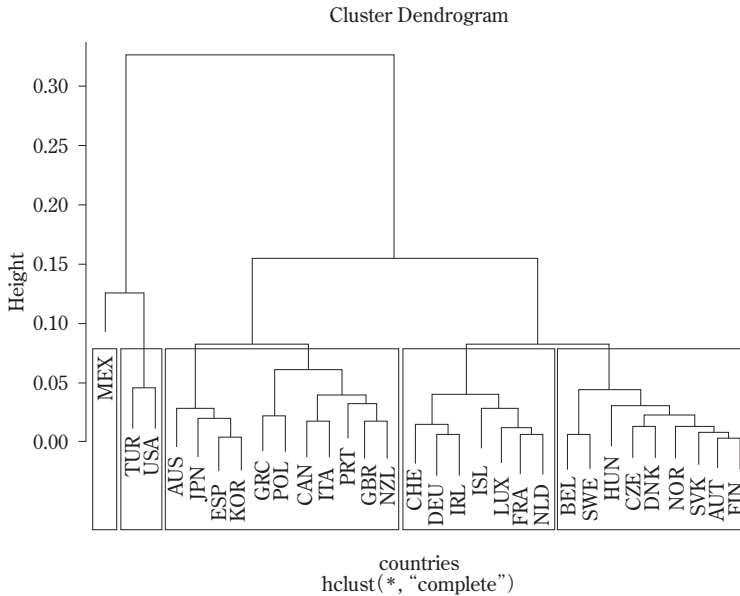
次に, 図 2 および図 4 を描くために利用されたデータに階層的クラスタリングを適用すると, 2000年代末における所得格差のデンドログラムを表す図 6 を得る。この図から 2000年代末における所得格差による OECD 諸国の 5 分類は次のようになる。

1) デンマーク, チェコ, ハンガリー, ノルウェー, 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:

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:

3) ポーランド, ギリシャ, 韓国, スペイン, 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:

4) アメリカ, トルコ:



(資料) 図2および図4に同じ。

(注) 図1に同じ。

図6 不平等度・貧困度の樹形図 (2000年代末)

5) メキシコ：

2000年代末におけるk平均法による5分類も次のように、上記とほとんど同じであり、分類の違いは先頭のデンマーク、チェコおよびハンガリーの3カ国を他から分離するか、最後尾のメキシコを他から分離するかにある。また、我が国は階層的クラスタリングによると第3クラスタに属し、k平均法によると第4クラスタに属す。ここでも、5分類は詳細すぎる分類かも知れないので次節で適切なクラスタ数の決定が検討される。

1) デンマーク、チェコ、ハンガリー：

2) ノルウェー、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：

3) 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：

- 4) ポーランド, ギリシャ, 韓国, スペイン, 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:
- 5) アメリカ, トルコ, メキシコ:

2. 適切なクラスタ数の決定 (k 平均法の利用)

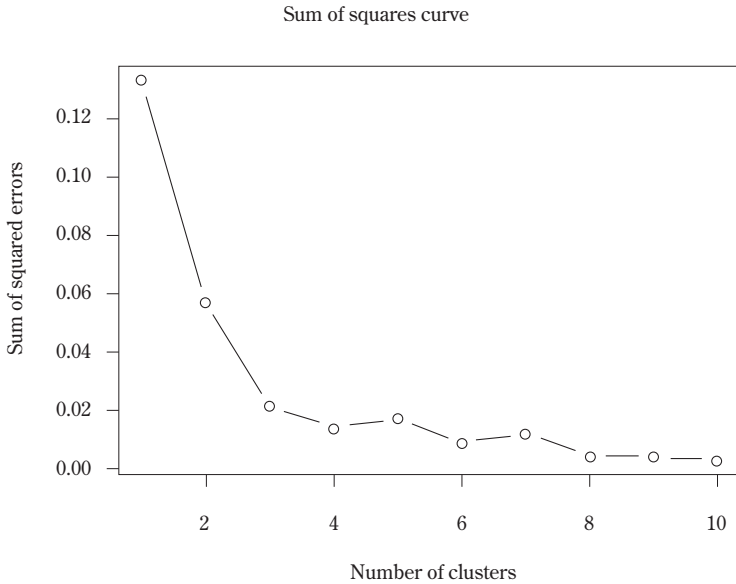
所得格差に関する樹形図を分析し k 平均法を適用するためには、事前に適切なクラスタ数 k を定める必要があるが、その方法は多数あり一意的には定まらないから、上では OECD30カ国をその不平等度で 5 分類する結果を報告した OECD (2008) に倣い、クラスタ数 $k=5$ として話が進められた。しかし、本稿のテーマについて 5 分類は詳細すぎる分類かも知れないので適切なクラスタ数を検討する必要がある。最適なクラスタ数を決定する基準となる指標が 30 種類以上も提案されているために¹¹⁾、この分野の応用においてほとんどすべての統計量が同一結果をもたらすことはめったにない。そこで、まず k 平均法をデータに適用し、よく知られた誤差の平方和曲線が利用される。図 7 は 2000 年代半ばの所得格差 (ジニ係数・貧困率) の 2 次元データに k 平均クラスタリングを適用した場合の総クラスタ内平方和とクラスタ数との関係を示している。このクラスタ内平方和曲線の屈曲部¹²⁾から適切なクラスタ数は 3 と推測される。図 8 は同様に 2000 年代末の所得格差についての総クラスタ内平方和曲線を示していて、この場合も適切なクラスタ数は 3 と推測される。ジニ係数および貧困率 (2000 年代半ば・末) それぞれについて同様の平方和曲線を描き適切なクラスタ数を推測した結果が表 4 の平方和曲線の行に示されている。

偏差平方和曲線による判断にはどうしても曖昧さが残るので、ギャップ統計量が提案されている¹³⁾。クラスタ内非類似度 (平方ユークリッド距離) W_k をクラスタ数 k の関数とするとき、ギャップ統計量は、

11) Milligan and Cooper (1985), Rousseeuw (1987), Tibshirani et al. (2001).

12) 膝, 肘とも呼ばれることがある。

13) Tibshirani et al. (2001).



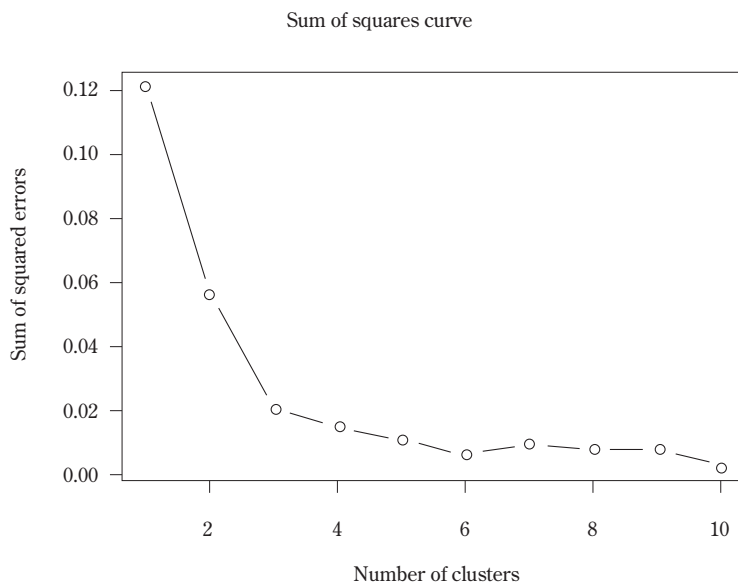
(資料) 図1の資料と表2により計測・作成。

図7 ジニ係数・貧困率に関するクラスタ内平方和 (2000年代半ば)

$$E^* \{ \log W_k \} - \log W_k$$

と表される。ここに、 $E^* \{ \log W_k \}$ はデータを含む矩形において一様分布を仮定して得られた $\log W_k$ の期待値である。ギャップ統計量によるクラスタ数 k の推定値は、 $k+1$ におけるギャップからそのギャップの標準偏差を引いた値よりも大きなギャップを生じる最小の k である。この方法によると、複数のクラスタ数だけでなく、単一のクラスタを検出できるのが特徴なのだが、ここで用いられたデータに関して表4にみるように、すべて単一のクラスタを報告しているので、OECD 諸国の分類という我々の目的にはそぐわない。

そこで、適切なクラスタ数の選択にクラスタリングの結果の評価関数を用いる新しい方法が、Pham et al.(2005) において提案されている。あるクラスタに属すデータとそのクラスタの中心との偏差から、クラスタごとに偏差平方和が



(資料) 図2の資料と表3により計測・作成。

図8 ジニ係数・貧困率に関するクラスタ内平方和 (2000年代末)

表4 最適クラスタ数の推定値

方 法	2000年代半ば	2000年代末	
	格差 (ジニ係数・貧困率)	格差 (ジニ係数・貧困率)	
平方和曲線 ギャップ値 Pham	3	3	
	1	1	
	2	3	
平方和曲線 ギャップ値 Pham	2000年代半ば ジニ係数	2000年代末 ジニ係数	
	2, 3	3	
	1	1	
平方和曲線 ギャップ値 Pham	9	5, 10	
	2000年代半ば 貧困率	2000年代末 貧困率	
			2, 3
1			1
Pham	7	5, 15	

(資料) 図7, 図8, 付表1及び付表2により作成。

得られるので、その k までの和を S_k と書く。すると、全体のデータ集合への各クラスタの偏差平方和の効果は、総偏差平方和に対する S_k の寄与度によって評価される。データに一樣分布を仮定し次元数の調整係数を α_k とすると、 S_k の推定値を $\alpha_k S_{k-1}$ と表すことができる。評価関数を

$$f(k) = S_k / \alpha_k S_{k-1}$$

で定義すると、この関数値はデータが一樣分布している場合、1 に近い値となる。データの分布に集中領域がある場合、 $f(k) < 1$ となり、この関数値が小さいほどデータの分布の集中度が高く、したがって最小の関数値を与える k 値が最適クラスタ数と判断される。Pham 関数値（付表 2）を利用して 2000 年代の格差等に関する適切なクラスタ数を判断した結果が表 4 の Pham の行に示されている。格差指標によっては大きなクラスタ数が報告されているが、少なくとも所得格差（ジニ係数・貧困率）の 2 次元データに関しては平方和曲線による推測とほぼ同一と捉えられる。

3. 先進国における所得格差の分析

3.1 先進国の所得不平等度の変動（GINI project）¹⁴⁾

本稿では先進国における所得格差の横断面的な分析が主な目的とされているから、所得の不平等度の長期変動は詳細には扱われない。その代わりこれ以降の分析や分類の参考のために、Tóth（2014）によって先進諸国の等価可処分所得の不平等度（ジニ係数）の変動（1980–2010）についてまとめる¹⁵⁾。この文献においては EU25 カ国を中心に先進 30 カ国が分析の対象であるが、この 30 カ国にメキシコおよびトルコは含まれず、オーストラリア、カナダ、日本、韓国およびアメリカが含まれている¹⁶⁾。この文献によると先進国の所得の不平等度の変動は次のように要約されよう。

14) Growing Inequalities' Impacts 研究プロジェクト。

15) 時系列データに関して GINI database のほうが LIS や OECD database より詳しい。

16) ブルガリア、エストニア、リトアニア、ラトビアおよびルーマニアも含まれている。

1) ジニ係数で測った先進諸国の不平等度は、期間1980-1984年 (0.201~0.331) に比べ期間2006-2010年 (0.228~0.373) のほうが総じて高い。

2) 不平等度の上昇は各国で一様ではない。

オーストリア、ベルギー、フランス、イタリア、アイルランドおよびスロベニアでは、不平等度はほとんど不変、あるいは同じ水準の回りを変動した。その他の国々では本質的な上昇が見られ、不平等における最も劇的な上昇を経験したのは、いくつかの社会主義からの移行国（ブルガリア、エストニア、リトアニア、ラトビア、ルーマニアおよびハンガリー）であり、この上昇率よりは低いがかなりの上昇を経験したのが北欧諸国、とりわけスウェーデンとフィンランドである。他の北欧グループやオランダでは穏やかな上昇が見られた。

3) 不平等度の急激な上昇の後に低下期間があったのが、ブルガリア、エストニアおよびハンガリーである。

4) 各国が比較に採用された期間の始めと終わりで、不平等区分間を移行することがあり得る。北欧諸国の不平等度は欧州で最低と思われていたが、もはやそうではない。バルト3国、ブルガリアおよびルーマニアのような社会主義からの移行国は不平等水準の区分において、低不平等区分（1980年代前期）から高不平等区分（1990年代後期）へ区分間の大きな移動を経験している¹⁷⁾。

3.2 所得格差による先進国の分類

a) 所得不平等度による分類

Tóth (2014) によると¹⁸⁾、2000年代後半（2006-2010年）の先進諸国のジニ係数による分類は、次のように4つに区分される¹⁹⁾。

1) ジニ係数0.250まで：チェコ、スロバキア、スロベニア；

2) ジニ係数0.251-0.300：オーストリア、ベルギー、デンマーク、フィン

17) 2000年代後半においては、リトアニア、ラトビアおよびルーマニアにおいては、より高い不平等区分への移行が見られる。

18) Table 2.2 (p.34).

19) この期間についての日本のデータが提出されていない。

ランド、フランス、ドイツ、ハンガリー、ルクセンブルク、オランダ、スウェーデン：

- 3) ジニ係数0.301-0.350：オーストラリア、ブルガリア、カナダ、エストニア、ギリシャ、アイルランド、イタリア、韓国、ポーランド、スペイン：
- 4) ジニ係数0.350超：リトアニア、ラトビア、ポルトガル、ルーマニア、イギリス、アメリカ：

Förster and Mira d'Ercole (2005) によると、1990年代後半の OECD27カ国が等価可処分所得についてのジニ係数によって、4グループ化されている。また、Hoeller et al.(2014) によると、2000年代末の OECD33カ国が等価可処分所得についてのジニ係数によって、4グループ化されている。第1節では2000年代の OECD30カ国が等価可処分所得についてのジニ係数と貧困率によって、5つのクラスタに分類される場合が示された。前節では適切なクラスタ数として単一クラスタから15クラスタまでが候補に上がった。そこで、所得分配研究における他の研究との比較という視点から3クラスタと4クラスタによる分析が適切と考えられ、k平均法が適用された結果、両者におおきな相違や矛盾がないから、k=4の場合のクラスタリング結果が以下に示される。

2000年代半ばにおける OECD30カ国はジニ係数によって次のように4つにクラスタリングされる。階層的クラスタリングの5分類において、第4クラスタと第5クラスタを統合した結果になっている。

- 1) デンマーク、スウェーデン：平均ジニ係数0.233
- 2) ルクセンブルク、オーストリア、チェコ、スロバキア、フィンランド、ベルギー、オランダ、スイス、ノルウェー、アイスランド、フランス、ハンガリー、ドイツ：平均ジニ係数0.275
- 3) オーストラリア、韓国、カナダ、スペイン、日本、ギリシャ、アイルランド、ニュージーランド、イギリス、イタリア：平均ジニ係数0.324
- 4) ポーランド、アメリカ、ポルトガル、トルコ、メキシコ：平均ジニ係数0.408

2000年代末における OECD 諸国はジニ係数によって次のように4つにクラスタリングされる。

- 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: 平均ジニ係数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: 平均ジニ係数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, アメリカ: 平均ジニ係数0.342
- 4) トルコ, メキシコ: 平均ジニ係数0.443

2000年代半ばから2000年代末にかけての短期間に、ジニ係数のクラスタ平均はどのクラスタにおいても上昇していて、OECD 諸国のほとんどで不平等が上昇している。特に、スウェーデン、ルクセンブルク、スイスおよびオーストラリアにおいてジニ係数の上昇幅が大きい、ノルウェー、ハンガリー、アイルランドおよびポーランドのように低下幅が大きな諸国もある。

b) 貧困度による分類

2000年代半ばにおける OECD30カ国は貧困率によって昇順に次のように4つにクラスタリングされる。

- 1) デンマーク, スウェーデン, チェコ, オーストリア, ノルウェー: 平均貧困率0.060
- 2) フランス, アイスランド, ハンガリー, フィンランド, オランダ, ルクセンブルク, スロバキア, イギリス, スイス, ベルギー: 平均貧困率0.078
- 3) ニュージーランド, ドイツ, イタリア, カナダ, オーストラリア, ギリシャ, ポルトガル: 平均貧困率0.119
- 4) スペイン, ポーランド, 韓国, アイルランド, 日本, アメリカ, トルコ, メキシコ: 平均貧困率0.158

2000年代末の OECD 諸国を貧困率によって昇順に4つにクラスタリングすると次のようになる。

- 1) チェコ, アイスランド, デンマーク, ハンガリー, フィンランド, オランダ, オーストリア, フランス, ノルウェー, スロバキア, ルクセンブルク

ク：クラスタ平均=0.071

2) スウェーデン, アイルランド, ベルギー, スイス, ドイツ, イギリス,
ニュージーランド: クラスタ平均=0.095

3) ポーランド, ポルトガル, イタリア, カナダ, ギリシャ, オーストラリア:
クラスタ平均=0.125

4) 韓国, スペイン, 日本, アメリカ, トルコ, メキシコ: クラスタ平均=
0.173

2000年代半ばから2000年代末にかけての短期間に、貧困率のクラスタ平均もどのクラスタにおいても上昇していて、OECD 諸国のほとんどで貧困率が上昇している。特に、スウェーデン、ノルウェー、イギリスおよびオーストラリアにおいて貧困率の上昇率が高いが、アイルランド、ポーランド、アイスランド、ハンガリーおよびオランダのように低下を経験した諸国もある。日本は最も貧困率の高い第4クラスタに属し、ジニ係数の場合は2番目に大きな第3クラスタに属し、2000年代半ばからの短期間では同一クラスタ内で両指標ともに上昇している。

c) 所得格差（不平等度・貧困度）による分類

2000年代半ばにおける OECD 諸国は、所得格差行列に k 平均法を適用すると、次のように4つにクラスタリングされる。

1) デンマーク, スウェーデン:

2) ルクセンブルク, オーストリア, チェコ, スロバキア, フィンランド,
ベルギー, オランダ, スイス, ノルウェー, アイスランド, フランス, ハンガリー:

3) ドイツ, オーストラリア, 韓国, カナダ, スペイン, 日本, ギリシャ,
アイルランド, ニュージーランド, イギリス, イタリア:

4) ポーランド, アメリカ, ポルトガル, トルコ, メキシコ:

2000年代半ばにおけるジニ係数によるクラスタリングとほぼ同じである。2000年代末における OECD 諸国も所得格差によって次のように4つにクラスタリングされる。

- 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:
- 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:
- 3) ポーランド, ギリシャ, 韓国, スペイン, 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:
- 4) アメリカ, トルコ, メキシコ:

2000年代末におけるジニ係数によるクラスタリングとほぼ同じであるが、第2クラスと第3クラスの構成員が幾分異なる。しかし、我が国はどちらの指標によっても不平等の高いほうから2番目の第3クラスに属している。2000年代半ばから2000年代末にかけての短期間に、所得格差が大きく拡大した国は、ジニ係数の場合と同様にスウェーデン、ルクセンブルク、スイスおよびオーストラリアであり、所得格差の順位を大きく低下した国は、やはりジニ係数の場合と同様にノルウェー、ハンガリー、アイルランドおよびポーランドである。

2000年代における OECD 諸国の所得格差による昇順区分は、次のように地理的密接さや言語文化的密接さに結びつくようである。

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

20) 第1クラスに属す。

南欧型やアジア型を追加する必要があるかもしれないが、福祉国家の3類型が第1クラスタから第4クラスタの中に分離されていることが読み取れる。社会主義の北欧諸国は第1クラスタに、保守主義の中欧諸国の一部は第2クラスタに、自由主義の英語圏諸国の一部は第3クラスタないし第4クラスタにそれぞれ属しているようである。社会主義体制から移行した東欧諸国の一部は第1クラスタに属している。このように、OECD諸国の所得格差によるクラスタリング結果は地理的密接さや言語文化的密接さだけでなく、歴史的経緯、政治体制、福祉レジーム等に関連していて、所得格差の数的特徴のみによってクラスタリングされた結果がどのような非数的特徴で分割されているのかが不十分ながら明らかにされた。

お わ り に

2000年代の2時点における先進国としてOECD諸国から、我が国を含め30カ国が選定され、2次元データとしての所得格差（不平等度・貧困度）によってクラスター分析が行われ、次のような結果が得られた²¹⁾。

- 1) 2000年代末の30カ国をジニ係数で階層的クラスタリングを行い、ジニ係数の昇順に5つに分類した結果、我が国はすくなくともOECD30カ国平均のジニ係数(0.308)を上回る第3クラスタに属し、2000年代末にはOECD30カ国中、高いほうから9番目に位置するようになった。
- 2) 2000年代末のOECD30カ国を貧困率によって昇順に5つにクラスタリングすると、我が国は貧困率が高い第4クラスタに属し、さらに2000年代半ばと同様にOECD30カ国中、高いほうから4番目に依然として位置している。
- 3) 2000年代末のOECD30カ国を所得格差（不平等度・貧困度）によって昇順に5つにクラスタリングすると、我が国は階層的クラスタリングによると第3クラスタに属し、k平均法によると第4クラスタに属す。

21) 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。

所得格差に関する樹形図を分析し k 平均法を適用するためには、事前に適切なクラスタ数 k を定める必要がある。その推定法は多数あり一意的には定まらないから、複数の方法が検討された結果、推定法が多ければ多いほどクラスタ数の候補数も多くなり、単一クラスタから15クラスタまで候補が上がった。そこで、所得分配研究における他の研究との比較という視点から3クラスタと4クラスタによる分析が行われ、両者の結果におおきな相違や矛盾がないから後者の結果が報告された。

2000年代の OECD30カ国がジニ係数、貧困率、所得格差(不平等度・貧困度)それぞれによって4つにクラスタリングされた結果は、5分類において第1クラスタと第2クラスタとが統合されるか、第4クラスタと第5クラスタとが統合された結果となることが多く、ここでも両者の結果におおきな相違や矛盾はない。そして、近年における OECD 諸国の所得格差によるクラスタリング結果は地理的密接さや言語文化的密接さだけでなく、歴史的経緯、政治体制、福祉レジーム等に関連していることが不十分ながら明らかにされた。

参 考 文 献

- Brandolini, A. and Smeeding, T. M. (2009). Income Inequality in Richer and OECD Countries, in W. Salverda, B. Nolan and T. M. Smeeding (eds.), *Oxford Handbook of Economic Inequality*, Oxford : Oxford University Press, 71-100.
- Everitt, B. S., Landau, S. and Leese, M. (2001). *Cluster Analysis*, 4th ed., New York : Oxford University Press.
- Förster, M. and Mira d'Ercole, M. (2005). Income Distribution and Poverty in OECD Countries in the Second Half of the 1990s, *OECD Social, Employment and Migration Working Papers*, 22.
- Hastie, T., Tibshirani, R. and Friedman, J. (2009). *The Elements of Statistical Learning*, 2nd ed., New York : Springer.
- Hoeller, P., Joumard, I., Pisu, M. and Bloch, D. (2014). Mapping Income Inequality Across the OECD, in P. Hoeller, I. Joumard and I. Koske (eds.), *Income Inequality in OECD Countries*, Singapore : World Scientific, 13-42.
- Kaufman, L. and Rousseeuw, P. J. (1990). *Finding Groups in Data : An Introduction to Cluster Analysis*, New York : John Wiley.
- 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, *J. 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.
- 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.
- Tóth, I. G. (2014) Revisiting Grand Narratives of Growing Inequalities : Lessons From 30 Country Studies, in B. Nolan, W. Salverda, D. Checchi, I. Marx, A. McKnight, I. Tóth and H. van de Werfhorst (eds.), *Changing Inequalities and Societal Impacts in Rich Countries*, Oxford : Oxford University Press, 11-38.
- 吉岡慎一 (2013). 日本の所得分配とパレート裾モデル —開端区間の平均値の推定— 『西南学院大学経済学論集』47(3・4), 91-108.

付表1 格差（ジニ係数・貧困率）に関するギャップ値

クラスタ数	ギャップ値	2000年代半ば 標準偏差	ギャップ値	2000年代末 標準偏差
1	0.2163	0.0942	0.2586	0.0690
2	0.1450	0.0722	0.1450	0.0647
3	0.2064	0.0771	0.0168	0.0717
4	0.1655	0.0807	0.2086	0.0766
5	0.0818	0.0850	0.1064	0.0946
6	0.1206	0.0917	0.1509	0.0962
7	0.1220	0.1012	0.1577	0.1001
8	0.2220	0.1015	0.1255	0.0936

(資料) 表2, 表3, OECD (2008, p.51) 及び OECD (2011, p.45) により計測。

付表2 格差（ジニ係数・貧困率）に関する Pham 関数値

クラスタ数	2000年代半ば	2000年代末
1	1.0000	1.0000
2	0.6094	0.6515
3	0.6130	0.6015
4	0.8773	0.9724
5	0.9424	0.8560
6	0.8016	0.7534
7	0.8990	0.8750
8	0.9749	0.9636

(資料) 付表1に同じ。