

# クラスター分析による食物消費構造解明の試み

The trial of food consumption structure elucidation by the cluster analysis

中島 順一

Junichi NAKAJIMA

## Abstract

Until now, factor analysis, etc. are used as an analysis procedure for clarifying the food consumption structure, and the food consumption structure has been clarified from the side of the extraction of the main factor child which regulates the caloric intake. In this report, it tried to be new using K-means method of cluster analysis. There were small eclipse type, conventional modernity the modernity, four meal types of the gluttony type as a result of the cluster analysis. Scales of those meal types were small eclipses type =36%, conventional =19%, modernity =34%, gluttony type =11%.

Keywords : K-means method, cluster analysis, food consumption structure

## 1. はじめに

食物摂取における食品群間の相互関係を包括的に捉える食物消費構造の解明には因子分析法や主成分分析法が専ら用いられてきた(豊川ら 1975、飯泉ら 1980、豊川ら 1981、中島ら 1983)。そこでは食品群間の相関行列を基盤としてデータの内的構造が分析され、食物摂取を規定する主要因子が抽出され、食生活分析における計量栄養学分野の発展に大きく貢献してきた。しかしこの因子分析や主成分分析を利用して明らかにできる食物消費構造にも限界があり、新たな知見を生み出すことが少なくなってきた。このような現状を踏まえ、クラスター分析の一手法である K-means 法を利用して全く別の視点から食物小構造を解明することができるかどうか試行し若干の知見を得たので報告する。

て昭和 52 年 11 月に実施された食物摂取状況調査結果の世帯別一人一日当たりの食品群別摂取量である。この調査は国民栄養調査の方式に準じた連続 3 日間の世帯調査で県環境衛生部および保健所の協力の元に行われたもので、食品群別摂取量は大分類された 20 食品群である。食品群の内訳は、米、パン、めん、いも、砂糖、菓子、油脂、味噌、豆類、果実、緑色野菜、他の野菜、漬物、海草、調味料、魚介類、獣鳥肉類、卵類、牛乳、加工食品である。

698 世帯の 20 食品群からなるデータ行列を基にクラスター分析の一手法である K-means 法を適用してケース(対象)が指定したクラスター個数になるよう操作し、クラスター間における食物摂取パターンの相違から食物消費構造を明らかにしていく。

表 1 K-means法でクラスター数 2 と設定した分析結果(各クラスターの基本統計量)

## 2. 分析資料と統計分析の方法

本報告では、クラスター分析の一手法である K-means 法を利用した解析方法が食物消費構造の解明に有効かどうかを試行するものなので、研究に用いる資料は従来の手法による分析結果がすでに報告されているものを利用する。今回、分析に用いた資料は、岐阜県下 14 保健所管内の 15 地域に在住する 698 世帯を対象とし

	クラスター 1 n=299			クラスター 2 n=469			全体 n=698	
	平均値	標準偏差	変動係数	標準偏差	変動係数	変動係数	平均	標準偏差
米	288.1	89.3	31.0	291.0	94.6	32.5	290.0	92.8
パン	27.1	39.6	145.8	18.0	30.4	169.1	21.0	33.9
めん	34.7	46.7	134.5	26.1	31.7	121.4	28.9	37.5
いも	104.7	79.8	76.2	82.6	66.3	80.2	89.9	71.7
砂糖	21.0	19.4	92.4	13.9	12.3	88.0	16.2	15.3
菓子	49.1	46.0	93.8	30.6	35.9	117.2	36.7	40.4
油脂	11.4	10.9	96.1	8.4	7.6	90.5	9.4	9.0
味噌	25.0	15.7	62.6	23.3	14.5	62.1	23.9	14.9
豆類	74.7	54.5	72.9	66.0	49.5	74.9	68.9	51.3
果実	354.4	143.7	40.5	115.2	76.3	66.2	193.7	152.6
緑色野菜	88.0	74.9	85.1	64.9	53.5	82.5	72.5	62.2
他の野菜	267.2	135.8	50.8	183.8	91.1	49.5	211.2	114.6
漬物	52.0	63.9	123.0	49.6	52.0	104.7	50.4	56.1
海草	6.3	5.9	94.8	4.5	5.7	127.8	5.1	5.8
調味料	49.8	28.8	57.8	39.1	23.7	60.6	42.6	25.9
魚介類	103.4	60.3	58.3	76.8	44.8	58.4	85.6	51.9
獣鳥肉類	81.5	43.4	53.2	58.1	38.1	65.6	65.8	41.4
卵類	56.9	34.2	60.2	41.2	26.3	63.9	46.3	30.0
牛乳	105.4	100.3	95.2	51.5	67.1	130.2	69.2	83.4
加工食品	8.7	20.2	231.6	9.6	18.2	190.1	9.3	18.9

## クラスター分析による食物消費構造解明の試み

表2 クラスタ数2の時の分散分析の結果

グループ	間平方和	df	グループ	内平方和	df	F	有意確率
米	1272.7	1		6003223	696	0.1476	0.7010
パン	12949.0	1		788541.2	696	11.4294	0.0008
めん	11514.2	1		967158.5	696	8.2860	0.0041
いも	74824.8	1		3506464	696	14.8520	0.0001
砂糖	7677.8	1		158230.8	696	34.2043	0.0000
菓子	52400.2	1		1086245	696	33.5749	0.0000
油脂	1314.9	1		54516.94	696	16.7874	0.0000
味噌	442.1	1		153921.4	696	1.9992	0.1378
豆類	11666.6	1		1821121	696	4.4587	0.0351
果実	8808433.0	1		7430818	696	823.0329	0.0000
緑色野菜	82397.6	1		2618016	696	21.9054	0.0000
他の野菜	1069566.0	1		8083847	696	92.0871	0.0000
漬物	863.1	1		2194669	696	0.2737	0.6010
海草	500.6	1		23279.35	696	14.9663	0.0001
調味料	17798.0	1		451535.7	696	27.4339	0.0000
魚介類	109108.6	1		1769962	696	42.9046	0.0000
獣鳥肉類	84363.6	1		1109712	696	52.9120	0.0000
卵類	37775.1	1		591248.8	696	44.4677	0.0000
牛乳	447547.2	1		4400136	696	70.7916	0.0000
加工食品	116.6	1		247544.5	696	0.3278	0.5671

### 3. 結果と考察

すでに報告済み(中島1996)のものであるが、図1に主成分分析における第1主成分と第2主成分からなる食品群の因子負荷量の散布図すなわち食物消費の2次元空間図(食品)を示す。これらの主成分の寄与率はそれぞれ13.9%、11.4%で、累積寄与率は25.4%である。第1主成分の正領域には獣鳥肉類、果実、卵類、牛乳、油脂、パンをはじめ多くの食品群が配置している。一方、負の領域には米、漬物、味噌がある。後者は主食偏重の食事を連想させ、前者はおかずの多いつまり副食品多食の食事を想起させるので、この主成分を副食品多食型対主食偏重型食事の因子と読み取った。第2主成分ではパン、油脂、牛乳などが正値をとり、逆に米、みそ、いも、豆類などが負値をとっている。前者は近代的食品といわれるものであり、後者は伝統的食品といわれるものなので、この主成分を近代型食事対伝統型食事の因子と読み取った。これらの2因子は、これまでの食物消費構造を分析した報告と細部では異なるが大筋では一致しているもので、日本人の食生活の根底にある基礎的な因子と考え

られた。

さてK-means法によるクラスター分析ではクラスターの個数を指定して分析を行う。K-means法によるクラスター分析ではデンドログラムで象徴され凝集法と違い、多変量データを基に指定されたクラスター個数に対し、クラスター内のばらつきが最小になり、クラスター間のばらつきが最大になるように、クラスター間でケースを移動させる。いしかえるとK-means法では、クラスター間の分散分析の結果が最も有意になるようにケースを移動させる。クラスター数を2と指定して分析した結果を表1および表2に示す。表1はクラスターにおける平均値、標準偏差、変動係数を示したもので、表2は分散分析の結果である。米、油脂、漬物、加工食品以外の食品群ではクラスター間に有意差が認められ、いずれもクラスター1の平均値がクラスター2の平均値を上回っている。このことからクラスター1は、いろいろな食品をたくさん食べるタイプ、クラスター2はあまり食べないが米、油脂、漬物、加工食品ではクラスター1と差がないタイプと読み取れる。これは主成分分析の第1主成分で副食品多食型対主食偏重型食事の因子と名づけたものと符合して興味深い。

次にクラスター個数を4と指定して分析した。4と指定した理由は、通常行われる因子分析や主成分分析での解釈で成果を上げているのは第1因子と第2因子からなる2次元であり、これは4つの象限なのでクラスター個数の仮説として現況の食物消費構造との比較における解釈可能性などから妥当であると考えた。表3に各クラスターの基本統計量、表4に分散分析の結果、図2に各クラスターの摂取プロフィールを示す。加工食品を除くすべての食品群で摂取量の統計的な有意差が認められる。クラスター1は、ほとんどの食品の摂取量が少ない少食型の食事パターンを示している。このクラスターのケース数は249で

表3 K-means法でクラスター数4と設定した分析結果(各クラスターの基本統計量)

	クラスター1 n=249			クラスター2 n=130			クラスター3 n=240			クラスター4 n=79			全体 n=698	
	平均値	標準偏差	変動係数	平均値	標準偏差	変動係数	平均値	標準偏差	変動係数	平均値	標準偏差	変動係数	平均	標準偏差
米	280.1	86.8	31.0	373.2	89.7	24.0	257.3	74.7	29.0	283.9	86.3	30.4	290.0	92.8
パン	14.3	26.6	186.6	12.1	24.6	203.5	31.1	39.4	126.7	25.8	40.5	156.8	21.0	33.9
めん	26.6	31.3	117.7	24.5	36.7	149.7	29.8	35.0	117.6	40.8	56.9	139.5	28.9	37.5
いも	70.6	55.3	78.3	139.7	85.7	61.3	78.2	50.7	64.8	103.8	101.1	97.5	89.9	71.7
砂糖	12.3	11.7	95.1	20.2	18.7	92.9	17.6	14.9	84.8	18.1	17.9	98.9	16.2	15.3
菓子	23.1	30.6	132.7	35.2	36.4	103.5	45.8	42.7	93.4	54.5	51.6	94.7	36.7	40.4
油脂	7.9	7.5	94.3	6.6	6.2	95.1	11.6	9.6	82.8	12.2	12.3	100.4	9.4	9.0
味噌	21.6	12.0	55.8	29.5	16.3	55.2	22.1	14.2	64.3	27.1	19.2	71.0	23.9	14.9
豆類	53.7	39.6	73.8	93.7	61.7	65.9	66.9	47.0	70.3	81.6	59.1	72.4	68.9	51.3
果実	82.4	60.0	72.9	131.3	85.9	65.4	242.7	67.0	27.6	498.0	153.4	30.8	193.7	152.6
緑色野菜	54.0	47.2	87.5	85.3	67.2	78.8	79.3	67.7	85.3	88.8	65.1	73.2	72.5	62.2
他の野菜	139.6	60.3	43.2	337.3	134.2	39.8	197.6	67.1	34.0	270.2	123.6	45.7	211.2	114.6
漬物	52.3	51.8	99.1	65.8	78.4	119.2	41.3	46.4	112.4	46.6	46.5	99.8	50.4	56.1
海草	3.9	4.9	126.7	6.1	7.0	114.7	5.3	5.8	109.6	6.4	6.1	95.2	5.1	5.8
調味料	35.2	22.9	64.9	47.5	26.5	55.7	45.5	25.4	55.7	48.8	30.7	63.0	42.6	25.9
魚介類	71.7	43.3	60.4	96.5	58.8	60.9	83.7	43.9	52.5	117.0	68.4	58.5	85.6	51.9
獣鳥肉類	51.6	35.2	68.2	64.0	42.3	66.0	75.3	42.2	56.1	84.9	40.8	48.1	65.8	41.4
卵類	36.5	24.1	66.1	41.8	27.3	65.3	55.0	30.5	55.4	58.2	37.6	64.6	46.3	30.0
牛乳	37.8	53.4	141.3	41.1	51.6	125.6	100.0	94.3	94.3	120.9	106.6	88.2	69.2	83.4
加工食品	9.9	19.0	192.0	8.1	17.0	209.6	9.3	16.8	181.4	9.3	26.0	279.1	9.3	18.9

## クラスター分析による食物消費構造解明の試み

表4 クラスター数4の時の分散分析の結果

グループ	間平方和	df	グループ	内平方和	df	F	有意確率
米	1184791.0	3		4819705	694	56.8669	0.0000
パン	48146.5	3		753343.7	694	14.7846	0.0000
めん	15152.9	3		963519.8	694	3.6381	0.0126
いも	462807.5	3		3118481	694	34.3317	0.0000
砂糖	6473.6	3		157435	694	9.5123	0.0000
菓子	91211.1	3		1047434	694	20.1446	0.0000
油脂	3358.7	3		52473.14	694	14.8073	0.0000
味噌	7030.6	3		147333	694	11.0390	0.0000
豆類	150858.2	3		1681929	694	20.7491	0.0000
果実	11486410.0	3		4752842	694	559.0738	0.0000
緑色野菜	139191.0	3		2561223	694	12.5719	0.0000
他の野菜	3661367.0	3		5492047	694	154.2223	0.0000
漬物	52780.7	3		2142751	694	5.6983	0.0007
海草	633.9	3		23144.02	694	6.3562	0.0003
調味料	21714.1	3		447619.6	694	11.2220	0.0000
魚介類	142202.1	3		1736868	694	18.9399	0.0000
獣鳥肉類	101140.5	3		1092935	694	21.4076	0.0000
卵類	56063.8	3		572960.1	694	22.6358	0.0000
牛乳	787483.1	3		4060200	694	44.8675	0.0000
加工食品	277.9	3		217383.2	694	0.2599	0.8543

全体の36%を占めている。クラスター2は米、いも、味噌、豆類、他の野菜、漬物を多く摂取し、パン、油脂、果実、牛乳が少い伝統型食事を営んでいるタイプである。このクラスターのケース数は130で、伝統型食事の世帯割合は19%である。クラスター3は、パン、油脂、果実、獣鳥肉類、卵類、牛乳が多く、逆に少ないのは米、いも、味噌、豆類、他の野菜、漬物などである。これはいわゆる近代的食事を示すパターンである。ケース数は240で、全体の34%がこの食事パターンを営んでいる。クラスター4は、米や漬物を除くほとんどの食品群を多く摂るパターンで、多食型の食事を営んでいるタイプと読み取れる。このクラスターのケース数は79で、全体の11%の世帯がこのタイプである。また各クラスター間の関係を捉えるためにクラスター間の距離行列に対し多次元尺度法で2次元空間上に圧縮して表示したものが図3である。クラスター1の少食型食事とクラスター2の近代型食事を営むグループがやや近い距離にあって全体の70%がここにいる。少数派の伝統型食事や多食型食事をしていない世帯は異なった方向に距離を置いており興味深い様相を呈している。

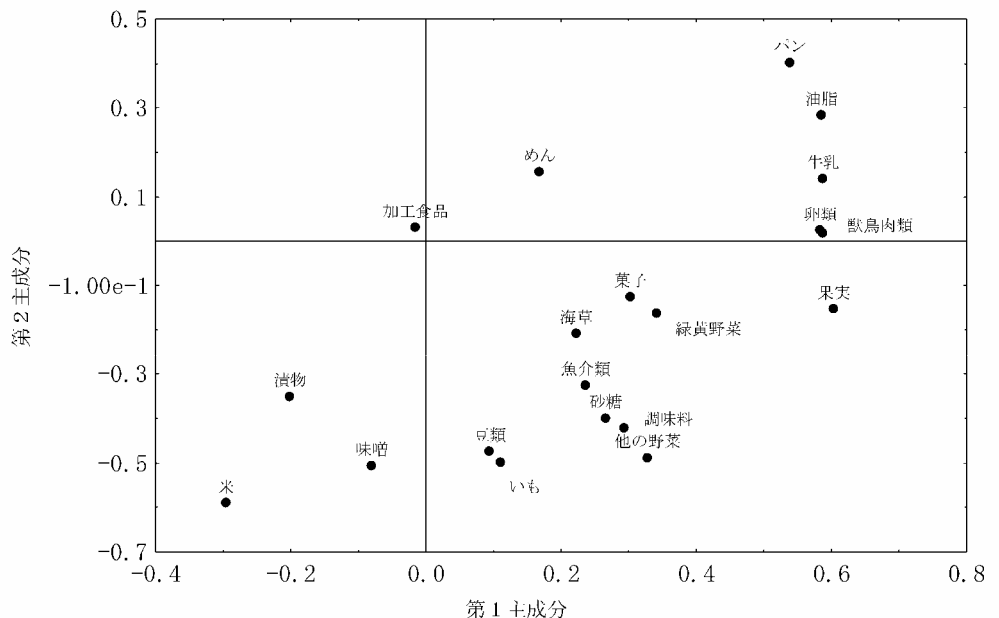
このようにクラスター数を4に増やすことによって主成分分析での第2主成分に相当する近代型食事対伝統型食事の因子と符合するクラスターも露呈してきた。このことからK-means法によるクラスター分析でも主成分分析や因子分析で得られる食事パターンと合致する食事パターンを入手できることが示唆される。このことは基となるデータは同一であるので主成分分析や

因子分析でスポットを当てられた側面がK-means法によっても同様にスポットを当てられて取り出すことができるということである。また従来の因子分析や主成分分析による分析結果では得ることのできなかった各食事パターンを営んでいるグループの規模がK-means法によるクラスター分析では明らかになる点、さらに分散分析によりどの食品群がクラスター間で差があるかを統計的に示すことができる点が利点として認められると考えられる。したがって従来の主成分分析法や因子分析法に併用することによって、より食物消費構造の新たな情報を獲得することができ、食物消費構造分析の新たな計量的アプローチとして貢献することが期待される。

## 4. 要約

従来、食物消費構造を解明するための分析手法として主成分

図1 食物消費の2次元空間図(食品)



分析法や因子分析法が使われ食物摂取を規定する主要因子の抽出という側面から食物消費構造を明らかにしてきた。この方法による研究成果は多く報告されているが分析手法に基づく制限は自らあり、それを越えた知見を得ることは不可能である。本報告では、多変量解析法の一手法であるクラスター分析のK-means法を利用して新たな試みを行い次のことが明らかとなった。

- 1) クラスター数を2と設定した分析では、いろいろな食品をたくさん食べるタイプと米、油脂、漬物、加工食品以外はあまり食べないタイプの2つの食事パターンを読み取った。これは主成分分析法による第1主成分の因子構造と類似する結果であった。
- 2) クラスター数を4と設定した分析結果からは、少食

図2 各クラスターの摂取プロフィール(標準化値使用)

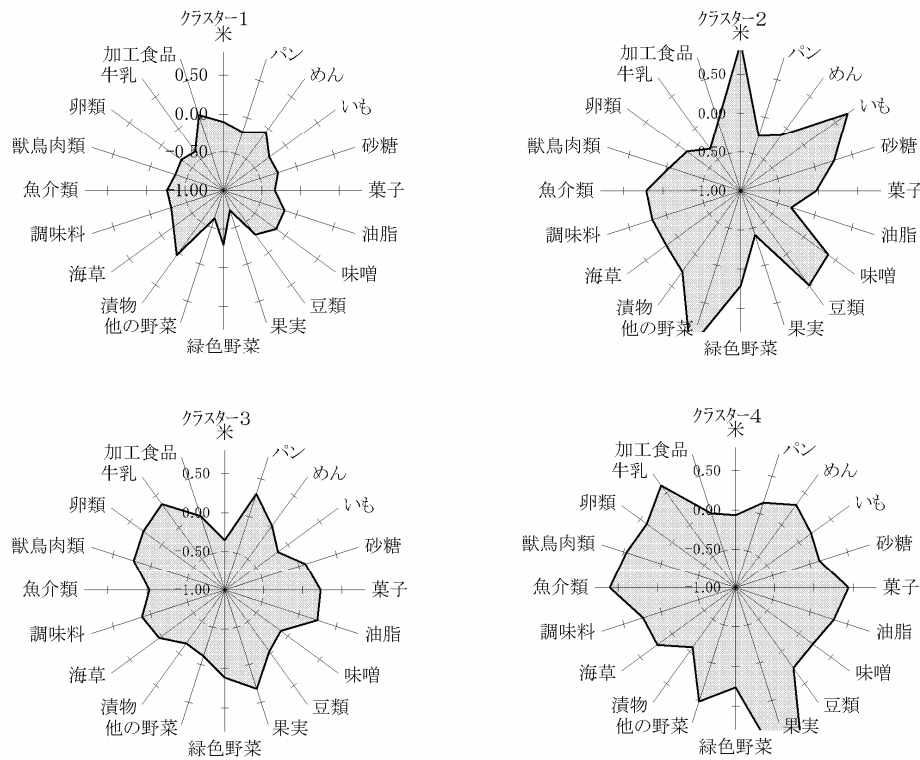
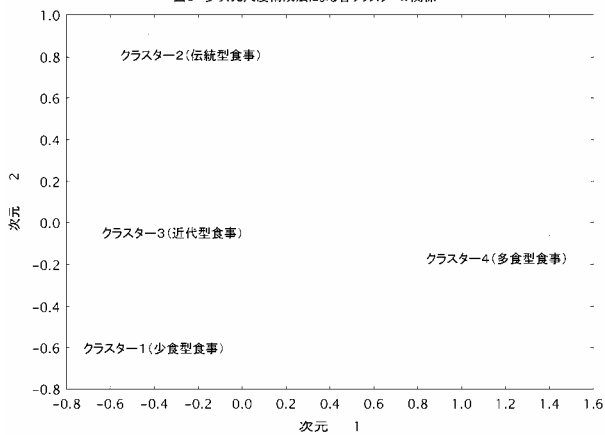


図3 多次元尺度構成法による各クラスターの関係



型、伝統型、近代型、多食型の4つの食事タイプと読み取られた。これらの食事タイプは、主成分分析で抽出された副食品多食型対主食偏重型食事の因子(第1主成分)と近代型食事対伝統型食事(第2主成分)での食事パターンと符合するものと考えられた。

3) 各食事タイプの規模は、少食型=36%、伝統型=19%、近代型=34%、多食型=11%で多勢を占める少食型と近代型はやや近い関係にあり、伝統型や多食型を営むものは少数でそれぞれ異なった方向に距離を置いていた。

4) K-means法によるクラスター分析でも従来の手法による食事パターンと同様の情報を手に入れることができ、さら

にグループの規模やグループ全体の構成状況など従来の方では知りえることのできなかった情報を得ることができることが示された。したがって従来の方法とK-means法によるクラスター分析を併用することによってより多くの食物消費構造の情報を得られることが期待された。

5. 参考文献

飯泉久子, 阪部京子, 古武彌三, 田崎武信, 雨宮武彦 (1980):ある集団における食物消費構造の多変量解析 食事バタンの抽出と個人の類型化, 日本公衛誌, 27, 543 - 551  
 豊川裕之, 三宅由子, 伊藤雅治(1975):わが国の食物

摂取に関する研究(第1報)全国規模の分析, 日本公衛誌, 22, 571-578

豊川裕之, 丸井英二, 小野田博一, 木村信子, 吉田節子, 金子俊, 山上雅子, 永山育子, 赤羽正之, 西岡葉子, 石井荘子, 森口覚 (1981): 東京近郊農村婦人(30~69歳)の食物摂取パターン, 栄養と食料, 34, 531-543

中島順一, 吉川周子(1983): 食物消費構造における多食性因子の除外, 民族衛生, 49, 59 - 66

中島順一(1983): 食物消費におよぼす地域特性と世帯主職業の影響, 民族衛生, 49, 164 - 171

中島順一(1996): 昭和52年の岐阜県の食物消費構造, 岐阜短大研究紀要, 45, 37 - 42

(提出期日 平成16年11月26日)