

食品摂取の時代・年齢・世代効果

— 共変動最大化によるコウホート分析 —

藤 本 高 志¹

Period-Age-Cohort Model Maximizing Covariation Applied to Food Intake

Takashi FUJIMOTO (Osaka University of Economics)

The deviations of values on a cohort table can be defined as the total of period, age, and cohort effects. This study pays attention to the fact that the total variation (sum of squared deviations) consists of individual variations of the three factors' effects and covariation between them, and suggests decomposing the deviations into the three effects while maximizing the covariation. The maximization induces the behaviors of the three effects on the cohort table to be linked together as far as possible. The parts of the behaviors which are not capable of being linked are decomposable independent behaviors. The other parts which it is possible to link are considered as not decomposable linked behaviors. This study then shows that the estimated three factors' effects with this model are unbiased and reliable estimators. This model was applied to determine period, age, and cohort effects of 18 groups of food covered by The National Nutrition Survey in Japan.

Key words : period-age-cohort analysis, food intake, maximizing covariation, unbiasedness, reliability

1. は じ め に

人々の行動は時間の経過とともに変化する。その変化は、時代が移ることによる変化と、加齢による変化で構成されると考えられる。また、貧しい時代に育った人と豊かな時代に育った人では行動様式が違うように、人々の行動はその人が属する世代によって違うと考えられる。厳密には、時代や年齢による変化をコントロールした状態における、世代による違いがあると考えられる。例えば、1980年代生まれと1950年代生まれで、肉の摂取量を比較するとする。それぞれの世代が同じ年齢になる時点で比較すれば、何歳の時点で比較しても、1980年代生まれの方が多いとしよう。また、肉の摂取量が時代の影響を受けないとしよう。とすれば、肉の摂取量には、世代による違いがあると

言える。

以上のように、人々の行動の変化は、時代、年齢、世代による変化に分離できると考えられる。例えば、時代による変化は、各時代に固有の効果の差と言える。したがって、人々の行動は、各時代、各年齢、各世代に固有の効果に規定されると考えることもできる。

後掲の第1表は、厚生労働省『国民栄養調査』による、年次別・年齢階層別の肉類の摂取量である。この表のように、一定期間おきに調査された年齢別データの表をコウホート表と呼ぶ。本研究では、食品摂取のコウホート表を用い、摂取量を規定する、各時代、各年齢、各世代に固有の効果を計測する。

また、共変動最大化によるコウホート分析モデルを提案する。モデルの概要は以下のとおりである。各時代、各年齢、各世代に固有の効果と時代効果、年齢効果、世代効果と呼ぶことにしよう。コウホート表上の各セルの値は、全てのセルに共通の効果とそのセルの時代、年齢、世代に対応する時代効果、年齢効果、世

¹大阪経済大学
tfuji@osaka-ue.ac.jp

代効果で構成されるとする。例えば、1989年の行と40～49歳の列が交わるセルの摂取量61g/日/人は、共通の効果と1989年の時代効果、40～49歳の年齢効果、1940年代生まれの世代効果で構成されるとする。そして、各時代、各年齢、各世代に固有の効果のパラメータとして推定する。

本モデルは、コウホート表上の値の全変動が、3因子の各効果の変動に加え、それらの共変動により構成されることに注目する。そして、共変動を最大化することで、コウホート表上における3因子の効果の動きを可能な限り連動させるパラメータを特定する。ここで、連動させることができない動きの部分は、分離可能な独立した動きである。他方、連動させることができる動きの部分は、本モデルでは分離不可能な連動した動きとみなす。このように推定されたパラメータは、コウホート表上の値の動きを十分に分離するとは言えない。しかし、不偏性を満たすとともに、信頼できる推定量であることを示す。

2. 食品摂取のコウホート分析

食料消費を決定する要因の分析は、価格や所得など経済要因を重視してきたが、年齢や世代などデモグラフィック要因を重視してきたとは言い難い。しかし、年齢や世代の効果は、食料需要予測のための貴重な情報である。なぜなら、食料需要を決定する要因のうち、価格や所得など経済要因の将来予測は難しいが、社会の世代構成や年齢構成などデモグラフィック要因の将来予測は容易である。コウホート分析の特徴は、食料需要予測に有益な年齢や世代の効果を計測できる点にある。

また、コウホート分析は経済要因を無視するわけではない。時系列データを用い、社会の年齢構成や世代構成の変化を勘案せず、価格弾力性や所得弾力性を計測するとして。消費量が価格や所得に加えて年齢構成や世代構成に依存するにもかかわらず、消費量を価格と所得だけで説明するわけだから、推定値がバイアスするのは当然であろう。そこで、森他（2005）は、消費量の変化を時代・年齢・世代の3因子の効果に分離し、時代効果による消費量の変化を価格と所得の変化で説明すれば、推定値の信頼性が大幅に改善することを実証している。この結果は、経済要因の効果は、人々のライフスタイルなど時代要因の効果とともに、時代効果に集約されていることを示している。時代により価格や所得が変化し、価格や所得が変化すれば消費量が変化するので、価格や所得を時代効果の構成要素と考えることができる。

このようななかで、デモグラフィック要因に注目し、性別・年齢階層別の畜産物の需要を推定した森島（1985）、年齢階層別・世帯類型別の食料消費を推定した石橋（2006）がある。いずれも、総務省『家計調査年報』の個票を用い、各世帯の消費量を被説明変数、各世帯の年齢階層別世帯員数を説明変数とする重回帰分析により、年齢階層別の個人消費量を推定している。ただし、これらの研究では、世代効果が考慮されていない。

しかし、3因子の効果を知ることは容易でない。第1表において、1989年の40歳代の肉類摂取量61g/日/人を基準に、3因子の効果を推論してみよう。表を横に眺めれば、50歳代の摂取量が54gなので、摂取量は加齢により減ると推論できる。しかし、61gは1940年代生まれの摂取量、54gは1930年代生まれの摂取量であるので、古い世代ほど少ないとも推論できる。表を縦に眺めれば、40歳代の摂取量が1999年には83gに増えているので、摂取量は時代の変化により増えたとも推論できる。しかし、83gは1950年代生まれの摂取量なので、新しい世代ほど多いとも推論できる。表を斜めに眺めれば、1989年の40歳代のセルと1999年の50歳代のセルは、ともに1940年代生まれの摂取量である。したがって、1940年代生まれの摂取量は、10年を経過し、61gから54gに減ったと言える。しかし、時代の変化によって減ったのか、加齢によって減ったのか判断できない。

このようななかで、食料消費を時代、年齢、世代の3因子で説明した先駆的研究に、米を対象とし、後述するベイズ型コウホート分析モデルを用いた松田他（1993）がある。彼らも、家計調査年報を用いているが、世帯主年齢階層別世帯当たり消費量を世帯員数で割った消費量を用いている。すなわち、個人の消費量ではないという問題がある。しかし、家計調査年報の個票は、通常、利用できないので、やむをえない問題と言える。このようななかで、Mori et al.（2001, 2005）は、隣り合う年齢階層間の1人当たり消費量の差はゼロに近いという制約を課し、世帯主年齢階層別消費量から個人の年齢階層別消費量を推定している。そして、いくつかの食品の消費量を、ベイズ型モデルにより、3因子で説明している。さらに、田中他（2007）や森他（2008）は、ベイズ型モデルに加え後述するIEモデルを用い、両者の結果を比較し、これらモデルの妥当性を検証している。

また、森他（2009）は、米国における、食料消費のデモグラフィック要因に注目した研究をサーベイし、年齢効果は重視してきたが、時代効果や年齢効果との

第1表 年次別・年齢階層別の肉類摂取量

(1人1日当たりg)

	年	年 齢					
		20~	30~	40~	50~	60~	70歳~
		29歳	39歳	49歳	59歳	69歳	
1986	80	77	66	59	47	39	
1987	72	76	60	55	43	35	
1988	83	77	77	57	48	40	
1989	86	81	61	54	44	44	
1990	82	76	66	55	48	38	
1991	79	79	65	56	43	35	
1992	79	82	69	54	45	38	
1993	84	77	78	54	44	41	
1994	76	79	68	61	43	38	
1995	90	99	72	61	50	34	
1996	83	77	80	56	51	36	
1997	88	91	82	55	56	40	
1998	92	85	89	69	45	34	
1999	76	100	83	54	52	40	
2000	90	90	86	69	50	40	

註：厚生労働省『国民栄養調査』より。

分離が困難な世代効果は無視されてきたとしている。このようななかで、野菜への家計支出を、メディアン回帰により、時代、年齢、世代、価格、所得、その他個人属性の効果に分解した Stewart et al. (2007) を評価している。

本研究は、既存研究と異なり、共変動最大化によるコホート分析モデルを提案する。また、国民栄養調査を用いる。国民栄養調査は、全国の1人世帯の食品群別摂取量を、1986~2000年の間、年齢階層別に公表している。家計調査年報を用いた場合、個人の年齢階層別消費量を推定しなければならない。それに対して、国民栄養調査は個人の年齢階層別摂取量を公表している。しかし、1人世帯の摂取量なので、代表的な個人の摂取量とは言えないという問題がある。一般世

帯では、様々な年齢や世代から構成される家族員全ての口に合う食生活となり、1人世帯では、特定の年齢や世代の人の口に合う食生活となると考えられる。そのため、1人世帯の摂取量を用いた場合、年齢効果と世代効果が過大評価される危険性がある。国民栄養調査は19の食品群を対象としているが、その他の食品を除き、18群を分析対象とした。

3. コホート分析モデル

1) モデルの定式化

本研究が対象とするコホート表は、第1表のとおりである。年次は1年刻みだが、年齢階層は10歳刻みである。両者の刻みが一致する標準コホート表では、出生コホートが対角線上を移動する。しかし、この一般コホート表では、例えば1940年代生まれコホートは、1989年には40~49歳のセル、10年後の1999年には50~59歳のセルへという具合に移動する。そこで、標準コホート表形式とするため、10年おきのデータを1セットとする。データセットは、1986年と1996年、1987年と1997年、1988年と1998年、1989年と1999年、1990年と2000年の5セットとなる。第1図は、第1表の1989年と1999年のデータセットを模式的に描いている。第1図の1989年の行の $Q_{11} \sim Q_{16}$ は、第1表の1989年の行の86~44に相当し、第1図の1999年の行の $Q_{21} \sim Q_{26}$ は、第1表の1999年の行の76~40に相当する。各データセットを個々に分析し、得られた5セットのパラメータの値の平均を最終的な推定値とする。したがって、第1表で網かけした部分のデータは利用しない。

1989年と1999年のデータセットを用いコホート分析する場合を考えよう。摂取量とそれを規定する3因子の効果の関係を(1)式のように定式化した。

$$Q_{ij} = \mu + P_i + A_j + C_k \tag{1}$$

Q_{ij} は表の各セルの値である。 P_i は*i*期に固有の効果のパラメータである。どのデータセットにおいても、1986~1990年を P_1 、1996~2000年を P_2 とする。 A_j は*j*年齢階層に固有の効果のパラメータである。20~

	年次	年 齢					
		20~29 (j=1)	30~39 (j=2)	40~49 (j=3)	50~59 (j=4)	60~69 (j=5)	70~ (j=6)
1989 (i=1)	Q_{11} (k=6)	Q_{12} (k=5)	Q_{13} (k=4)	Q_{14} (k=3)	Q_{15} (k=2)	Q_{16} (k=1)	
1999 (i=2)	Q_{21} (k=7)	Q_{22} (k=6)	Q_{23} (k=5)	Q_{24} (k=4)	Q_{25} (k=3)	Q_{26} (k=2)	

第1図 模式的に描いた標準コホート表

29歳を A_1 、30～39歳を A_2 、以下同様に、70歳～の A_6 まで設定する。 C_k は k コウホートに固有の効果のパラメータである。1910年代生まれを C_1 、1920年代生まれを C_2 、以下同様に、1970年代生まれの C_7 まで設定する。第1図に示すように、コウホート表の対角に位置するセルの世代効果 C_k は同じである。 μ は、各セルの摂取量を総平均した値で、全てのセルに共通の総平均効果である。また、(2)(3)(4)式のゼロ和制約を課す(註1)。

$$\sum_{i=1}^2 P_i = 0 \quad (2)$$

$$\sum_{j=1}^6 A_j = 0 \quad (3)$$

$$C_1 + 2C_2 + 2C_3 + 2C_4 + 2C_5 + 2C_6 + C_7 = 0 \quad (4)$$

2) パラメータ推定方法の先行研究

ここで問題となるのが(1)式の推定方法である。回帰分析により推定できるように思える。しかし、説明変数間には、年次-年齢=世代、という線形依存関係がある。つまり、3因子のうちの2つが与えられれば、残りの1つは決まってしまう。例えば、2000年-40歳=1960年生まれ、という具合である。そのため、各セルの値を3因子で説明しようとしても、世代が年次と年齢で説明されるとすれば、各セルの値は年次と年齢だけで説明されてしまう。この線形依存関係により、回帰デザイン行列がフルランクとはならず、無数の解が存在する。これを識別問題と呼ぶ。そこで、伝統的には、パラメータベクトルの任意の部分に等値関係の制約を課し(例えば、20歳代と30歳代の年齢効果は等しいなど)、識別問題を回避してきた。しかし、このような制約は恣意性を免れない。

このようななかで、Nakamura (1986) はベイズ型コウホート分析モデルを提案した。例えば、20歳代と70歳代など、離れたパラメータ間には大差があるかもしれないが、20歳代と30歳代など、隣接するパラメータは近似するという漸進的条件を制約に課す。そして、無数に存在する最小二乗解から、漸進的条件を満たす解を特定する。しかし、隣接するパラメータが必ずしも近似するとは限らない。

また、Yang et al. (2008) は、IE (Intrinsic Estimator) と呼ばれる最小二乗解を提案し、「IEは、デ

ザイン行列の影響を除いた、主成分回帰の推定量の特別な形式」としている。田中他 (2007) は、IEを「回帰のデザイン行列を X とすれば、 $XB_0=0$ となる非負のベクトル B_0 が一個存在し、 $B_0b=0$ を制約条件とする b の最小二乗解」と解説している。すなわち、無数に存在するパラメータの最小二乗解から、 B_0 と直交する解を特定する。しかし、田中他 (2007) は、IEについて、「その統計的根拠は一応理解できるとしても、その現実的・経済学的な根拠が不明」としている。また、森他 (2007) による実験が示すように、ベイズ型モデルやIEモデルが、安定的な結果を保証するわけではない(註2)。

本研究が提案するモデルは、以下に示すように、仮定や外生的制約に頼らずにパラメータを推定する。

3) 共変動最大化によるパラメータの推定

(1)～(4)式を C_7 について整理すれば、(5)式を得る。パラメータを未知数とする連立方程式を解くのと同じ要領である。

$$\left. \begin{aligned} P_1 &= \frac{1}{6} C_7 + m_1 & P_2 &= -\frac{1}{6} C_7 - m_1 \\ A_1 &= -\frac{5}{6} C_7 + m_2 & A_2 &= -\frac{1}{2} C_7 + m_3 \\ A_3 &= -\frac{1}{6} C_7 + m_4 & A_4 &= \frac{1}{6} C_7 + m_5 \\ A_5 &= \frac{1}{2} C_7 + m_6 & A_6 &= \frac{5}{6} C_7 + m_7 \\ C_1 &= -C_7 + m_8 & C_2 &= -\frac{2}{3} C_7 + m_9 \\ C_3 &= -\frac{1}{3} C_7 + m_{10} & C_4 &= m_{11} \\ C_5 &= \frac{1}{3} C_7 + m_{12} & C_6 &= \frac{2}{3} C_7 + m_{13} \end{aligned} \right\} \quad (5)$$

$m_1 \sim m_{13}$ は、既知の定数で、既知の Q_{ij} の関数として表現できる(註3)。したがって、 C_7 を与えれば、全てのパラメータの値が決まる。(5)式は C_7 について整理しているが、 C_4 を除き、いずれのパラメータについて整理することもできる。

(1)式は(6)式のように変形できる。

$$(Q_{ij} - \mu) = P_{ij} + A_{ij} + C_{ij} \quad (6)$$

(註1) 世代効果のゼロ和制約が(4)式となるのは、コウホート表上において、 C_2 、 C_3 、 C_4 、 C_5 、 C_6 はそれぞれ2回現れるが、 C_1 と C_7 はそれぞれ1回しか現れないことによる。

(註2) 森他 (2007) は、3効果を予め想定して作成したコウホート表を用い、ベイズ型モデルとIEモデルが予め想定した3効果を復元できるかどうかを実験し、復元できないケースが少なくないとしている。

(註3) 例えば m_1 は以下のとおり。

$$m_1 = (11Q_{11} + 9Q_{12} + 7Q_{13} + 5Q_{14} + 3Q_{15} + Q_{16} - 11Q_{21} - 9Q_{22} - 7Q_{23} - 5Q_{24} - 3Q_{25} - Q_{26}) / 72$$

左辺は、各セルの値と平均値の差で、偏差である。 P_{ij} 、 A_{ij} 、 C_{ij} は、 Q_{ij} のセルにおける時代効果、年齢効果、世代効果である。 P_{ij} 、 A_{ij} 、 C_{ij} は、それらにゼロ和制約を課しているので、同様に、平均値(0)からの偏差である。両辺を二乗し、全てのセルについて合計すれば、(7)式を得る。

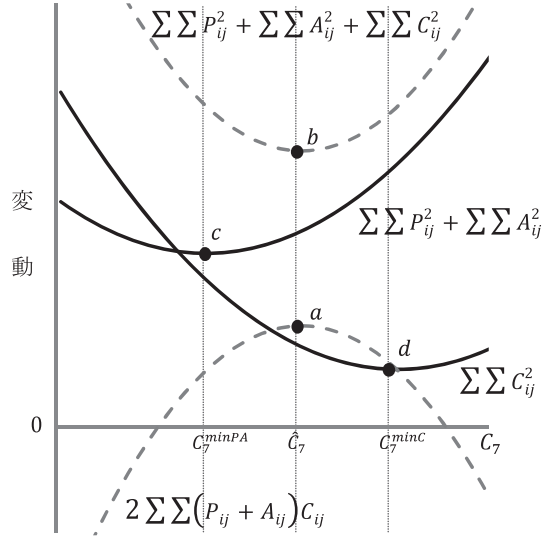
$$\sum_{i=1}^2 \sum_{j=1}^6 (Q_{ij} - \mu)^2 = \sum_{i=1}^2 \sum_{j=1}^6 P_{ij}^2 + \sum_{i=1}^2 \sum_{j=1}^6 A_{ij}^2 + \sum_{i=1}^2 \sum_{j=1}^6 C_{ij}^2 + 2 \sum_{i=1}^2 \sum_{j=1}^6 (P_{ij} + A_{ij}) C_{ij} \quad (7)$$

データの偏差の二乗和はデータの変動を示す。よって、左辺は全変動である。右辺の第1項は時代効果の変動、第2項は年齢効果の変動、第3項は世代効果の変動である。第4項は、時代+年齢効果と世代効果の共変動である。 $P_{ij}A_{ij}$ の項が消えるのは、コウホート表上において時代効果 P_i と年齢効果 A_j が相関せず、それらの相関係数がゼロだからである(註4)。共変動は、2つの変数の連動の程度を表し、2つの変数の標準偏差で基準化すれば、(8)式に示すように相関係数 $r_{PA \times C}$ となる。

$$r_{PA \times C} = \frac{\sum \sum (P_{ij} + A_{ij}) C_{ij}}{\sqrt{\sum \sum (P_{ij} + A_{ij})^2 \sum \sum C_{ij}^2}} \quad (8)$$

ここで、(5)式に任意の C_7 を外生的に与え、暫定的なパラメータを得て、それらパラメータを(7)式の右辺の各項に代入すれば、3因子の効果の各変動と共変動を得る。第2図は、代表的なコウホート表のデータを用い、 C_7 を連続的に変化させ、各変動の挙動を描いている。

$C_7 = \hat{C}_7$ のとき、共変動は a 点で最大、同時に3因子の効果の各変動の合計は b 点で最小となる。時代効果の変動と年齢効果の変動の合計は $C_7 = C_7^{minPA}$ のとき c 点で最小、世代効果の変動は $C_7 = C_7^{minC}$ のとき d 点で最小となる。 \hat{C}_7 は、(7)式右辺第4項に(5)



第2図 暫定的パラメータ C_7 と変動

式を代入し、 C_7 で微分し0と置けば求めることができる。 C_7^{minPA} と C_7^{minC} も、同様に求めることができる。 \hat{C}_7 は、 C_7^{minPA} と C_7^{minC} のちょうど中心に位置する(註5)。したがって、共変動最大化により特定されるパラメータは、全変動のうちの共変動を除く部分を $P_{ij} + A_{ij}$ の変動と C_{ij} の変動で説明するにあたり、いずれにも偏ったウェイトを置かないと考えられる。

本研究では、 $C_7 = \hat{C}_7$ を(5)式に代入し、パラメータを特定する。このように、共変動を最大化することは、(9)式に示すように、各セルの $P_{ij} + A_{ij}$ と C_{ij} の差の二乗和を最小化し、可能な限り $P_{ij} + A_{ij}$ と C_{ij} を近似することで、コウホート表上での両者の動きを連動させる。

(註4) P_i は、 A_j がどんな値をとろうが、 P_1 と P_2 の二値しかとらない。したがって、 P_{ij} と A_{ij} の相関係数はゼロとなる。また、 $\sum \sum P_{ij}A_{ij}$ は、次式のように変形でき、ゼロ和制約の $\sum P_i = 0$ と $\sum A_j = 0$ より、ゼロとなることが証明できる。

$$\sum_{i=1}^2 \sum_{j=1}^6 P_{ij}A_{ij} = \sum_{i=1}^2 P_i \sum_{j=1}^6 A_j = 0$$

(註5) \hat{C}_7 、 C_7^{minPA} 、 C_7^{minC} は次式により求めることができる。

$$\hat{C}_7 = \frac{1}{228} (-137Q_{11} - 36Q_{12} + 29Q_{13} + 58Q_{14} + 51Q_{15} + 8Q_{16} + 182Q_{21} + 63Q_{22} - 20Q_{23} - 67Q_{24} - 78Q_{25} - 53Q_{26})$$

$$C_7^{minPA} = \frac{1}{228} (-119Q_{11} - 27Q_{12} + 29Q_{13} + 49Q_{14} + 33Q_{15} - 19Q_{16} + 209Q_{21} + 81Q_{22} - 11Q_{23} - 67Q_{24} - 87Q_{25} - 71Q_{26})$$

$$C_7^{minC} = \frac{1}{228} (-155Q_{11} - 45Q_{12} + 29Q_{13} + 67Q_{14} + 69Q_{15} + 35Q_{16} + 155Q_{21} + 45Q_{22} - 29Q_{23} - 67Q_{24} - 69Q_{25} - 35Q_{26})$$

以上の3本の式を整理すれば、 $\hat{C}_7 = (C_7^{minPA} + C_7^{minC})/2$ 、を得る。

$$\sum_{i=1}^2 \sum_{j=1}^6 \{(P_{ij}+A_{ij})-C_{ij}\}^2 = \sum_{i=1}^2 \sum_{j=1}^6 P_{ij}^2 + \sum_{i=1}^2 \sum_{j=1}^6 A_{ij}^2 + \sum_{i=1}^2 \sum_{j=1}^6 C_{ij}^2 - 2 \sum_{i=1}^2 \sum_{j=1}^6 (P_{ij}+A_{ij})C_{ij} \quad (9)$$

ここで、連動させることができない動きの部分は、分離可能な独立した動きである。他方、連動させることができる動きの部分は、本モデルでは分離不可能な連動した動きとみなす。

4) パラメータの推定量の不偏性 (復元実験 1)

パラメータ P_i , A_j , C_k を仮想的に生成し、任意の総平均効果 μ とともに (1) 式に代入し、仮想的な Q_{ij} を得るとしよう。そして、共変動最大化により、それら Q_{ij} より、元のパラメータ、すなわち仮想的に生成したパラメータを復元する実験を考えよう。この実験における仮想的パラメータのような元のパラメータを真のパラメータと呼び、復元されたパラメータを復元パラメータと呼ぶことにしよう。

復元パラメータと真のパラメータは一致しない。なぜなら、真のパラメータによる共変動は、最大化された共変動よりも小さい。しかし、復元パラメータは、過大評価と過小評価が同様に起こりうるという意味で、不偏推定量と考えられる。なぜなら、 $P_{ij}+A_{ij}$ と C_{ij} を可能な限り近似しているということは、それぞれを大きくも小さくも特定していない。ということは、各パラメータに関しても、それぞれを大きくも小さくも

特定していない。そして、以下に示すように、真のパラメータは、復元パラメータの左右に均等に確率分布すると考えられるからである。

第 2 図において、真の C_7 に近づくため、共変動を、 $C_7 = \hat{C}_7$ のときの値よりも小さく特定するとしよう。とすれば、 C_7 は、 \hat{C}_7 よりも大きくか、あるいは小さく特定される。ここで注目したいのは、共変動が左右対称の関数ということである (註 6)。そのため、任意の共変動に対して 2 つの C_7 が特定され、それらは \hat{C}_7 を中心に左右対称に分布する。また、真の C_7 が \hat{C}_7 の左右いずれに位置するかは、それを判断する根拠がなく、同様に起こりうると考えられる。したがって、理論的には、真の C_7 は、 \hat{C}_7 の左右に均等に確率分布すると考えられる。また、(5) 式に示すように、 C_7 以外のパラメータは、 C_7 に定数 l_x を乗じ、定数 m_x を加えた値である。したがって、 C_7 以外のパラメータにも同じことが言える。以下では、パラメータの復元実験により、このことを実証しよう。

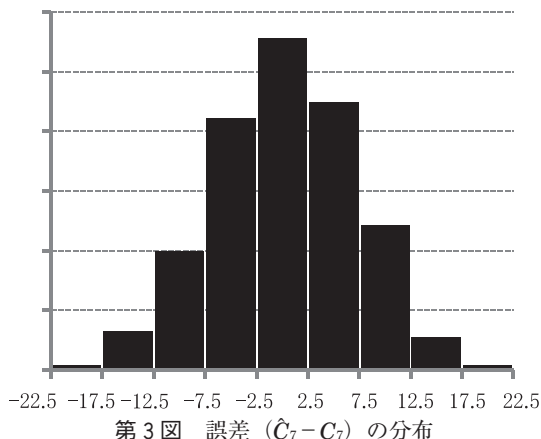
仮想的なパラメータの生成にあたり、総平均効果 $\mu = 100$ とし、(2)~(4) 式のゼロ和制約を課した。また、実際の推定値に近くなるように、各効果のパラメータのばらつき (例えば、 $A_1 \sim A_6$ のばらつき) に制約を課した (註 7)。以上の制約を課し、以下のよう、仮想的パラメータをランダムに発生させた。まず、 P_1 と P_2 , $A_1 \sim A_6$, $C_1 \sim C_7$ の値を、ばらつきの制

(註 6) (7) 式の右辺第 4 項に (5) 式を代入すれば、共変動は C_7 の関数となる。この関数を、 C_7 が共変動の関数となるように逆関数に変換すれば、次式を得る。

$$C_7 = \frac{\alpha \pm 72\sqrt{-684COV + \beta}}{5472}$$

COV は共変動。 α と β は、既知の Q_{ij} の関数で、既知の定数。

(註 7) 年齢効果のパラメータを例に、課した制約を示そう。本研究では、18 の食品群を対象とし、それぞれに 5 セット、合計 90 セットのデータがある。したがって、90 セットのパラメータが推定された。これらパラメータを総平均効果 μ が 100 となるように調整した。 A_1 から A_6 の 6 つのパラメータが推定されるが、 A_1 から A_6 のうちの最大値、最小値、そして 6 つパラメータの分散 ($= \sum_{j=1}^6 A_j^2 / 6$) を 90 の各パラメータセットについて求めた。そして、それら 90 の最大値、最小値、分散のメディアンをそれぞれ求め、それらメディアン の 0.5~1.5 倍の間に、仮想的パラメータの最大値、最小値、分散がそれぞれ収まるようにした (時代効果に関しては、最大値と最小値を制約とし、分散は制約に課していない)。以上のように、パラメータ自体に課した制約のほかに、それらパラメータを (1) 式に代入することで得られる仮想的 Q_{ij} の分布のばらつきにも制約を課した。90 の各データセットについて Q_{ij} の変動係数を求め、それら変動係数のメディアンを求めた。そして、そのメディアン の 0.5~1.5 倍の間に、仮想的 Q_{ij} の変動係数が収まるようにした。0.5~1.5 倍と範囲を持たせたのは、多様なばらつきを持つ仮想的パラメータを対象に復元精度を評価するためである。しかし、範囲を広げすぎると、非現実的なパラメータが生成される。そこで、範囲の決定にあたって注目したのが、後述するように、仮想的パラメータの相関係数 $r_{PA \infty C}$ ((8) 式) が復元精度に重要な影響を及ぼすことである。このことを考慮した復元精度の評価を行うため、仮想的パラメータによる Q_{ij} から復元されたパラメータの相関係数 $\hat{r}_{PA \infty C}$ の分布範囲が、後掲の第 3 表に示す実際の推定パラメータの相関係数 $r_{PA \infty C}$ の分布範囲 (-0.43~0.64) まで広がるように、仮想的パラメータのばらつきがとりうる範囲を試行的に広げた。その結果が 0.5~1.5 倍の範囲である。

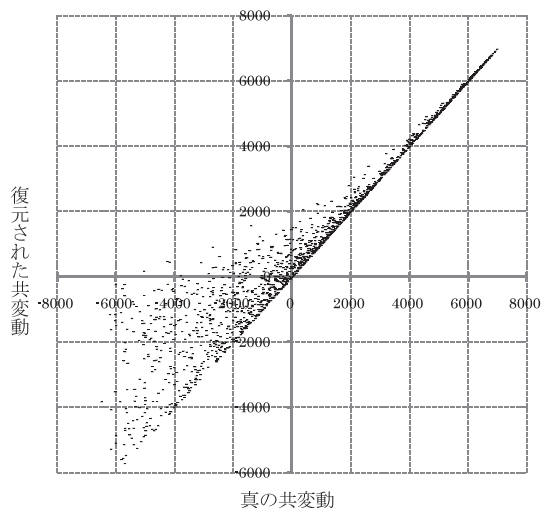


約（註8）を満たすように、ExcelのRANDBE-TWEENコマンドを用いてランダムに発生させた。そして、それらの値を初期値とし、Excelのソルバーを用い、ばらつきの制約に加えてゼロ和制約を満たす値をそれら初期値の近傍から探した。

そして、生成された仮想的パラメータを（1）式に代入して得た Q_{ij} から、共変動最大化により真のパラメータが復元できるかどうか実験した（復元実験1）。第3図は、1,000回の繰り返し実験から得た、 \hat{C}_7 と真の C_7 の差としての誤差（ $\hat{C}_7 - C_7$ ）の分布である。0を中心とする左右均等の分布となる。また、前述したように、 C_7 以外のパラメータの値は、 C_7 に定数 L_x を乗じ、定数 m_x を加えた値である。したがって、これらパラメータの誤差も、誤差（ $\hat{C}_7 - C_7$ ）と同様に分布する（ただし、 $|L_x| \leq 1$ なので、それらの誤差分散は誤差（ $\hat{C}_7 - C_7$ ）の分散よりも小さい）。したがって、共変動最大化により特定されるパラメータは、過大評価と過小評価が同様に起こりうるという意味で、不偏推定量と言える（註9）。

5) パラメータの推定量の信頼性（復元実験2）

次に、推定されたパラメータの信頼性を評価するための復元実験を行った。 $P_{ij} + A_{ij}$ と C_{ij} の共変動が最大となるようにパラメータを特定するので、その信頼性は $P_{ij} + A_{ij}$ と C_{ij} の相関関係により異なると考えられる。そこで、その相関関係を考慮するため、仮想的パラメータを生成する際、制約に相関係数 $r_{PA \leftrightarrow C}$



第4図 真の共変動と復元された共変動

（（8）式）を加えた。制約 $r_{PA \leftrightarrow C}$ は-0.77から0.90までの0.01刻みの値とし、各値について10セット、合計1,680セットの仮想的パラメータを生成した。

復元実験1と同様に、仮想的パラメータによる Q_{ij} から真のパラメータが復元できるかどうか実験した（復元実験2）。第4図は、真のパラメータによる共変動と復元パラメータによる共変動との散布図である。両者が一致するとき、散布図の点は45°線上に乗り、真のパラメータが完全復元される。ほぼ45°線上に乗る点が多いが、点は上方に偏り、共変動が過大評価されている。ここで注目したいのは、共変動が小さいほど、上方への点の偏りが深刻化し、共変動の過大評価が深刻化することである。すなわち、 $P_{ij} + A_{ij}$ と C_{ij} の相関が小さいほど、第3図に示すような誤差分散が裾の重い分布となり、推定値の信頼性が低下する。以下では、その理由を考えよう。

共変動を最大化するという事は、（9）式から理解できるように、各セルの $P_{ij} + A_{ij}$ と C_{ij} の差の二乗和を最小化している。したがって、共変動の過大評価は、各セルの $P_{ij} + A_{ij}$ と C_{ij} の差の過小評価と言える。ここで注目したいのは、 $P_{ij} + A_{ij}$ と C_{ij} の符号が異なるセルの存在である。各セルの値 Q_{ij} は $P_{ij} + A_{ij}$ と C_{ij} の合計なので、このようなセルでは、 $P_{ij} + A_{ij}$ と C_{ij}

（註8） 註7に示す仮想的パラメータの最大値、最小値、分散の制約。

（註9） 仮想的パラメータの生成に際して、それらのばらつきに制約を課した。しかし、第3図の分布が、この制約によって左あるいは右にずれるとは言えない。なぜなら、誤差（ $\hat{C}_7 - C_7$ ）を被説明変数、時代、年齢、世代の各効果の分散を説明変数とする重回帰モデルにおいて、帰無仮説「すべての係数がゼロである」を危険率0.70でも棄却できなかった（F検定；F値=0.463, n=1,000）。

が相殺される。相殺された効果は、 Q_{ij} の変化に顕在化せず、調査者が観察できない。そのため、 $P_{ij}+A_{ij}$ と C_{ij} の差が過小評価されてしまう。符号が異なるセルは、どのような仮想的パラメータによるコウホート表にも存在する。しかし、 $P_{ij}+A_{ij}$ と C_{ij} の相関が小さい仮想的パラメータによるコウホート表ほど、符号が異なるセルが多く、相殺される部分も大きい。これが、 $P_{ij}+A_{ij}$ と C_{ij} の相関が小さいほど、共変動の過大評価が深刻化し、推定値の信頼性が低下する原因と考えられる。

第2表は、復元実験2の結果を標本とし、真のパラメータと復元パラメータがどの程度一致するかを実証している。推定値の信頼性評価に利用するため、復元パラメータを(8)式に代入し、調査者が観測可能な相関係数 $\hat{r}_{PA\infty C}$ を得て、それらが第1列に示す区間に収まるように標本を分類した。後掲の第3表には、パラメータの最終的な推定値を示した。これら推定値による $\hat{r}_{PA\infty C}$ は $-0.43\sim 0.64$ の範囲にあり、それに対応するように、最小区間を $-0.49\sim -0.40$ 、最大区間を $0.61\sim 0.70$ に設定した。したがって、第2表において問題なく復元できることを確認できれば、第3表の推定値も信頼できる。第2列は、各区間に分類された標本数である。

まず、3因子の各効果において、真のパラメータ群と復元パラメータ群で、 $A_6>A_4>A_5>A_2>A_3>A_1$ など、パラメータの値のカテゴリー間での大小関係が一致するかどうかを実証した。第3列は、どの程度一致するかを表すKendallの一致係数(W ; 註10)を各標

本について求め、それらを平均した値(\bar{W})である。 W は、完全一致すれば1、全くランダムならば0となる。時代効果は、 $\hat{r}_{PA\infty C}$ にかかわらず \bar{W} が1に近く、大小関係が問題なく復元されている。年齢効果と世代効果は、 $\hat{r}_{PA\infty C}$ が小さいほど、 \bar{W} が小さく、復元精度が低下する。しかし、最小でも0.89である(註11)。したがって、パラメータの値のカテゴリー間での大小関係は、ほぼ問題なく復元できると言える(註12)。

真のパラメータ群と復元パラメータ群で、それぞれを構成するパラメータの値は、ともに平均をゼロとして分布する。なぜなら、パラメータにゼロ和制約を課している。パラメータの値のカテゴリー間での大小関係が両群でほぼ一致することが実証されているので、それらパラメータの値の分布のばらつきも一致すれば、分布の平均がともにゼロなので、両群はほぼ同一とみなせる。そこで、パラメータの値の分布のばらつきを、平均偏差を尺度に比較した(註13)。

第4列は、各標本について、(10)式に示すように、真のパラメータの平均偏差に対する復元パラメータの平均偏差の比(D)を求め、それらを平均した値(\bar{D})である。 D は、平均偏差の一致度で、完全復元されれば1となる。

$$D_P = \frac{\sum \sum |\hat{P}_{ij}|}{\sum \sum |P_{ij}|} \quad D_A = \frac{\sum \sum |\hat{A}_{ij}|}{\sum \sum |A_{ij}|} \quad D_C = \frac{\sum \sum |\hat{C}_{ij}|}{\sum \sum |C_{ij}|} \quad (10)$$

時代、年齢、世代のいずれの効果も、一部の例外を除き、 $\bar{D} \leq 1$ となり、分布のばらつきが過小評価される。また、年齢効果と世代効果では、 $\hat{r}_{PA\infty C}$ が小さいほど過小評価が深刻化する。しかし、 \bar{D} は最も小さくて

第2表 真のパラメータと復元パラメータの一致度の平均値

相関係数 ($\hat{r}_{PA\infty C}$)	標本数	順位 (\bar{W})			平均偏差 (\bar{D})		
		時代効果	年齢効果	世代効果	時代効果	年齢効果	世代効果
0.61~ 0.70	114	1.00	0.98	0.99	0.95	0.97	1.00
0.51~ 0.60	111	0.99	0.98	0.98	0.93	0.97	0.99
0.41~ 0.50	105	1.00	0.98	0.99	0.96	0.98	0.99
0.31~ 0.40	133	1.00	0.97	0.96	0.96	0.95	0.99
0.21~ 0.30	127	0.99	0.96	0.95	0.91	0.92	0.96
0.11~ 0.20	121	0.99	0.95	0.94	0.94	0.92	0.94
0.01~ 0.10	107	1.00	0.94	0.92	0.90	0.91	0.94
-0.09~ 0.00	111	1.00	0.94	0.94	1.03	0.91	0.94
-0.19~ -0.10	110	0.98	0.93	0.91	0.96	0.88	0.89
-0.29~ -0.20	102	0.99	0.93	0.92	0.98	0.88	0.92
-0.39~ -0.30	106	0.99	0.91	0.89	1.01	0.84	0.88
-0.49~ -0.40	92	0.98	0.92	0.90	0.92	0.86	0.92

も 0.84 で、パラメータの値の分布のばらつきが、最大でも 16% 過小評価されるにすぎない。直感的には、(10) 式から理解できるように、パラメータの絶対値が最大 16% 過小評価されるにすぎない。したがって、パラメータの値の分布のばらつきも、ほぼ問題なく復元できると言える (註 14)。

第 3 表の推定値は、仮想的パラメータの復元実験に基づく結論として、ほぼ信頼できると言える。なぜなら、各効果について、真のパラメータ群と推定パラメータ群において、パラメータの値のカテゴリー間での大小関係がほぼ一致し、それらの値の平均をゼロとする分布のばらつきもほぼ一致する。

6) パラメータの推定量の性質

パラメータの推定量が満たすべき性質は一致性である。しかし、共変動最大化による推定量は一致性を満たさない。なぜなら、真のパラメータを特定するのではなく、共変動が最大となるパラメータを特定する。したがって、標本数を増やしたとしても、推定値は、共変動が最大となる値に近づいても、真の値には近づかない。

この場合、推定値が真の値に一致するかどうかは、

必要条件としての不偏性と、十分条件としての信頼性から評価できると考えられる。パラメータの推定量は、復元実験 1 より、不偏性を満たすことが実証されている。また、復元実験 2 より、時代+年齢効果と世代効果に強い負の相関がない限り、各効果において真のパラメータ群と推定パラメータ群がほぼ一致し、信頼性もほぼ満たすことが実証されている。

問題となるのは以下の 2 点である。第 1 は、3 因子の効果の各パラメータ群内におけるパラメータの値の分布のばらつきが、不偏性を満たさず、過小評価される点である。この原因は、共変動が過大評価され、3 因子の効果の各変動が過小評価されることにある (註 15)。すなわち、コウホート表上の値の変化を説明するにあたり、3 因子の連動した効果が過大評価され、独立した効果が過小評価されることにある。

第 2 は、時代+年齢効果と世代効果の相関が小さいほど、パラメータの値の分布のばらつきの過小評価が深刻化し、推定値の信頼性も低下する点である。この原因は、第 4 図に示すように、時代+年齢効果と世代効果の相関が小さいほど、共変動が過大評価されることにある。

(註 10) W は次式より得ることができる。

$$W = \frac{12 \sum_{i=1}^k (R_i - \bar{R})}{n^2 (k^3 - k)}$$

n は標本数、 k はカテゴリー数 (例えば年齢効果の場合、 $A_1 \sim A_6$ の 6 カテゴリー) である。 R_i は、標本ごとに各カテゴリーに順位をつけ、カテゴリーごとに順位を合計した値である (年齢効果 A_1 の場合、 $n=2$ 、 A_1 の順位を 2 位と 3 位とすれば、 $R_1=2+3$)。 \bar{R} は R_i の平均値である。

(註 11) 世代効果を例に、 W の値が示す復元精度を感覚的に示そう。 $W=0.95$ および 0.91 となったパラメータセットを標本に (標本数 99 および 77)、真のパラメータの順位と復元パラメータの順位を照合した。 $C_1 \sim C_7$ 7 個のうち、順位が、一致、1 つ前後、2 つ前後するパラメータ数の平均は以下のとおりであった。 $W=0.95$ の場合、一致 3.4 個、1 つ前後 2.8 個、2 つ前後 0.8 個であった。 $W=0.91$ の場合、一致 2.7 個、1 つ前後 2.4 個、2 つ前後 1.9 個であった。順位が 3 つ以上前後するパラメータが存在する標本は、いずれの場合もなかった。

(註 12) 仮想的パラメータの生成に際して、それらのばらつきに制約を課した。この制約が順位一致係数 W に及ぼす影響を分析するため、時代、年齢、世代効果のそれぞれについて、仮想的パラメータの分散と W の相関係数を求めた。相関係数はそれぞれ 0.09、0.11、0.07 と低かったが、帰無仮説「相関係数はゼロである」は危険率 0.05 で棄却された ($n=1,680$)。そこで、仮想的パラメータの分散が W に及ぼす影響を、3 因子の効果のそれぞれについて、回帰モデルにより定量分析した。結果は、仮想的パラメータの分散が第 3 表の推定パラメータの分散のメディアンから半減したとしても、いずれの効果においても、 W は 0.01 低下するにすぎないというものであった。したがって、第 2 表の順位一致度 (\bar{W}) は、制約の影響を受けるが、その影響は非常に小さいと言える。

(註 13) 平均偏差は、データの偏差の和をデータ数で割り算した値である。標準偏差は、分布のばらつきの指標としてよく利用されるが、データを二乗するので、大きな異常値が存在すると、ばらつきが過大評価される。

(註 14) 仮想的パラメータの生成に際して、それらのばらつきに制約を課した。しかし、第 2 表の平均偏差一致度 (\bar{D}) が、この制約の影響を受けているとは言えない。なぜなら、仮想的パラメータの分散と平均偏差一致度 D の相関係数は、時代、年齢、世代効果のそれぞれについて、 -0.01 、 0.02 、 -0.04 で、帰無仮説「相関係数はゼロである」を危険率 0.30 でも棄却できなかった ($n=1,680$)。

(註 15) 3 因子の各効果の変動は、パラメータの値そのものの変動であり、それが過小評価されるということは、パラメータの値の分布のばらつきが過小評価されることを意味する。

7) パラメータの推定方法

各食品群について、データが5セットあり、それぞれをコウホート分析し、5セットのパラメータを得た。そして、それらの平均をとることで、最終的な推定値とした。

また、1989年と1999年のデータセット以外では、世代がモデルで定義した世代 k と一致しない。そこで、世代効果のパラメータ C_k は次のように推定した。1986年と1996年のデータセットより、1940年代生まれの C_4 を推定するとしよう。1986年、40～49歳のセルは1937～46年生まれ、30～39歳のセルは1947～56年生まれの摂取量である。まず、①1937～46年および1947～56年生まれのパラメータを $C_{37/46}$ および $C_{47/56}$ として、コウホート分析し、それぞれを推定する。そして、②1937～46年生まれに占める1940年代生まれの割合が0.7、1947～56年生まれに占める1940年代生まれの割合が0.3であるから、 $C_4 = 0.7C_{37/46} + 0.3C_{47/56}$ 、と加重平均して C_4 を推定した。

食品摂取に時代効果があるかどうかは、2標本の t 検定を用い、帰無仮説「 P_1 と P_2 に差はない」を棄却できるかどうかで検定した(註16)。また、年齢効果あるいは世代効果があるかどうかは、Kendallの一致係数(W ; 註10)を用い、帰無仮説「5つのパラメータセットにおける A_j あるいは C_k の大小関係に一貫性がない」を棄却できるかどうかで検定した(註17)。

8) 分析対象期間の摂取量変化とその要因の推定

また、分析対象期間における摂取量の変化とその要因を推定した。この10年間における時代効果による摂取量の変化(ΔQ_P)は(11)式より推定した。時代効果による変化は、時代が移ったことによる変化と解釈できる。

$$\Delta Q_P = P_2 - P_1 \quad (11)$$

年齢効果による変化(ΔQ_A)は(12)式より推定した。 α_j^2 と α_j^1 は、1998年(1996～2000年の中間年)と1988年(1986～1990年の中間年)の j 年齢階層の構成比($\sum \alpha_j^i = 1$)である。この間、社会の年齢構成が高齢化したので、高齢化が進んだことによる変化と解釈できる。

$$\Delta Q_A = \sum_{j=1}^6 A_j \alpha_j^2 - \sum_{j=1}^6 A_j \alpha_j^1 \quad (12)$$

世代効果による変化(ΔQ_C)は(13)式より推定

した。 c_k^2 と c_k^1 は、1998年と1988年の k コウホートの構成比($\sum c_k^i = 1$)である。この間、社会の世代構成が変わったので、世代交代が進んだことによる変化と解釈できる。

$$\Delta Q_C = \sum_{k=2}^7 C_k c_k^2 - \sum_{k=1}^6 C_k c_k^1 \quad (13)$$

そして、 ΔQ_P 、 ΔQ_A 、 ΔQ_C の合計 ΔQ は、分析対象期間における、代表的個人の摂取量変化の推定値である。

4. 結 果

パラメータの推定結果は第3表のとおりである。コウホート分析により得た5セットのパラメータおよび相関係数 $\hat{r}_{PA \times C}$ の平均値を示している。「***」は危険率1%、「**」は危険率5%、「*」は危険率10%で帰無仮説を棄却でき、「-」は危険率10%でも棄却できないことを示す。また、第4表は、分析対象期間(10年間)における、摂取量の変化とその要因の推定結果である。

米類の摂取量は、第3表より、時代の変化に伴って減少している。加齢とともに増加するが、60歳を超えると、老化により減少する。また、1940年代生まれが最も少なく、それ以降の新しい世代では、新しい世代ほど多く、それ以前の古い世代では、古い世代ほど多い。対象期間の摂取量は、第4表より、時代の変化により減少したと言える。小麦類の摂取量は、第3表より、時代の変化に伴って増加し、高齢層ほど少なく、新しい世代ほど多い。対象期間の摂取量は、第4表より、時代の変化と世代交代による増加が高齢化による減少を上回り、増加したと言える。

いも類、豆類、緑黄色野菜、その他の野菜は、ほぼ同様の傾向を示す。これら食品の摂取量は、第3表より、時代の変化に伴って増加している。加齢とともに増加するが、70歳を超えると、老化により減少する。また、新しい世代ほど少ない。対象期間の摂取量は、第4表より、時代の変化と高齢化による増加が世代交代による減少を上回り、増加したと言える。きのこ類の摂取量は、第3表より、時代の変化に伴って増加している。また、1940年代生まれが最も多く、それ以降の新しい世代では、新しい世代ほど少なく、それ以前の古い世代では、古い世代ほど少ない。しかし、年齢の影響を受けているとは言えない。対象期間の摂取

(註16) 統計量 $(\bar{P}_1 - \bar{P}_2) / s \sqrt{1/n_1 + 1/n_2}$ が、自由度 $df = n_1 + n_2 - 2$ の t 分布に従うことを利用した。 \bar{P}_i は平均値、 s は両群から合成した分散、 n_i はデータ数である。

(註17) 統計量 $W \cdot n(k-1)$ が、近似的に、自由度 $df = k-1$ の χ^2 分布に従うことを利用した。

第3表 時代・年齢・世代効果のパラメータの推定結果

(g/日/人)

		米類	小麦類	いも類	豆類	緑黄色野菜	その他の野菜
総平均		177.5	104.7	49.7	65.1	84.9	152.9
時代	1986~1990	19.0 ***	-2.5 *	-3.3 ***	-2.4 *	-9.6 ***	-4.9 **
	1996~2000	-19.0	2.5	3.3	2.4	9.6	4.9
年齢	20~29歳	-5.5 ***	4.1 ***	-6.2 ***	-18.8 ***	-18.7 ***	-14.8 ***
	30~39歳	6.3	11.7	-8.0	-8.5	-8.2	-14.5
	40~49歳	8.5	7.5	-2.8	0.4	-1.2	-6.1
	50~59歳	8.3	3.2	-1.8	13.4	7.6	12.1
	60~69歳	-2.1	-8.0	9.4	11.7	15.6	17.8
	70歳~	-15.4	-18.4	9.4	1.8	4.9	5.4
世代	~1919	8.6 **	-10.6 ***	3.0 ***	13.2 **	7.0 ***	2.9 ***
	1920~29	4.2	-10.2	6.8	8.6	5.7	9.1
	1930~39	-1.8	-7.6	7.0	0.3	6.6	5.7
	1940~49	-4.8	-0.8	-1.8	0.2	0.7	9.3
	1950~59	-2.8	10.6	-4.3	-3.9	-3.4	2.8
	1960~69	-2.6	10.6	-5.9	-9.9	-8.5	-16.8
	1970~	7.3	11.2	-7.9	-8.1	-13.0	-29.8
相関係数 $\hat{r}_{PA \llcorner C}$		-0.23	0.51	0.41	0.29	0.33	0.20
		きのこ類	海藻類	果実類	魚介類	肉類	卵類
総平均		9.5	5.7	137.2	88.0	64.4	38.1
時代	1986~1990	-1.0 ***	-0.2 -	1.7 -	-1.8 *	-1.9 ***	0.4 -
	1996~2000	1.0	0.2	-1.7	1.8	1.9	-0.4
年齢	20~29歳	-1.3 -	-0.7 ***	-24.3 ***	-14.8 ***	4.3 ***	-3.3 *
	30~39歳	-0.6	-1.1	-28.8	-7.2	8.3	-0.9
	40~49歳	-0.5	-0.7	-9.9	1.2	11.2	5.0
	50~59歳	1.2	0.4	9.9	11.0	2.1	2.0
	60~69歳	1.8	1.2	26.5	8.7	-6.9	0.7
	70歳~	-0.6	0.9	26.5	1.2	-18.9	-3.6
世代	~1919	-1.3 **	-0.3 ***	10.5 ***	-2.5 ***	-4.9 ***	1.9 -
	1920~29	-0.3	0.8	22.0	3.0	-8.9	0.3
	1930~39	0.6	0.7	24.9	5.4	-8.6	-1.3
	1940~49	1.9	0.8	10.7	11.6	-6.2	-2.2
	1950~59	0.7	0.1	-14.4	3.5	6.7	2.0
	1960~69	-0.5	-1.6	-34.6	-13.4	14.5	0.2
	1970~	-3.8	-1.8	-37.3	-22.7	15.0	0.2
相関係数 $\hat{r}_{PA \llcorner C}$		-0.23	0.17	0.64	0.45	0.52	-0.40
		乳類	油脂類	砂糖類	種実類	菓子類	調味嗜好飲料
総平均		116.8	17.5	10.7	1.9	19.0	229.7
時代	1986~1990	-7.6 **	0.5 ***	0.5 **	-0.5 ***	-1.3 ***	-35.6 ***
	1996~2000	7.6	-0.5	-0.5	0.5	1.3	35.6
年齢	20~29歳	-9.1 **	1.5 ***	-1.3 **	-0.2 **	-1.6 -	-62.3 ***
	30~39歳	4.2	2.4	-0.1	-0.1	-1.2	71.3
	40~49歳	-6.2	1.9	-0.2	-0.8	0.0	93.2
	50~59歳	-6.1	0.2	0.5	-0.1	1.2	27.2
	60~69歳	5.3	-1.8	0.4	0.6	2.9	-22.1
	70歳~	12.0	-4.1	0.6	0.7	-1.4	-107.3
世代	~1919	-14.8 *	-2.6 ***	0.5 **	-0.7 **	3.5 -	-12.0 ***
	1920~29	7.4	-2.0	0.6	-0.2	2.5	-54.9
	1930~39	8.8	-1.1	0.7	0.6	-2.0	-38.6
	1940~49	10.9	-0.4	0.3	0.6	-2.7	-6.4
	1950~59	-3.3	1.6	-0.3	0.7	-2.8	19.0
	1960~69	-11.7	2.7	-1.1	-0.7	0.6	61.4
	1970~	-7.8	1.6	-1.4	-1.6	6.0	69.6
相関係数 $\hat{r}_{PA \llcorner C}$		-0.07	0.57	0.08	-0.43	-0.43	0.16

註:「***」「**」「*」は、それぞれ危険水準1%、5%、10%で各効果が認められ、「-」は、危険水準10%でも効果を認めることができないことを意味する。

第4表 分析対象期間（10年間）における摂取量の変化とその要因

(g/日/人)						
摂取量変化	米類	小麦類	いも類	豆類	緑黄色 野菜	その他の 野菜
時代効果 (ΔQ_P)	-37.9	5.1	6.7	4.9	19.3	9.8
年齢効果 (ΔQ_A)	-1.0	-1.4	0.9	0.6	0.8	1.3
世代効果 (ΔQ_C)	0.4	3.6	-2.1	-3.2	-3.5	-6.4
合計 (ΔQ)	-38.5	7.2	5.6	2.3	16.6	4.7
摂取量変化	きのこ類	海草類	果実類	魚介類	肉類	卵類
時代効果 (ΔQ_P)	2.1	0.4	-3.3	3.6	3.7	-0.8
年齢効果 (ΔQ_A)	0.1	0.1	2.9	0.5	-1.4	-0.2
世代効果 (ΔQ_C)	-0.6	-0.4	-9.3	-4.4	3.8	-0.1
合計 (ΔQ)	1.5	0.1	-9.8	-0.3	6.2	-1.1
摂取量変化	乳類	油脂類	砂糖類	種実類	菓子類	調味嗜好 飲料
時代効果 (ΔQ_P)	15.1	-0.9	-1.1	1.1	2.6	71.2
年齢効果 (ΔQ_A)	0.4	-0.3	0.0	0.1	0.1	-9.7
世代効果 (ΔQ_C)	-0.4	0.7	-0.3	-0.2	0.8	16.5
合計 (ΔQ)	15.2	-0.6	-1.4	0.9	3.4	78.1

量は、第4表より、時代の変化による増加が世代交代による減少を上回り、増加したと言える。海草類の摂取量は、第3表より、時代の影響を受けているとは言えないが、加齢とともに増加し、新しい世代ほど少ない。対象期間の摂取量は、第4表より、世代交代による減少と、時代の変化あるいは高齢化による増加が相殺され、大きく変化しなかったと考えられる。果実類の摂取量は、第3表より、時代の影響を受けているとは言えないが、高齢層ほど多く、新しい世代ほど少ない。対象期間の摂取量は、第4表より、世代交代による減少が高齢化による増加を上回り、減少したと言える。

魚介類の摂取量は、第3表より、時代の変化に伴って増加している。加齢とともに増加するが、70歳を超えると、老化により減少する。また、1940年代生まれの摂取量が最も多く、それ以降の新しい世代では、新しい世代ほど極端に少なく、それ以前の古い世代では、古い世代ほど少ない傾向がみられる。対象期間の摂取量は、第4表より、時代の変化による増加と世代交代による減少が相殺され、大きく変化しなかったと言える。肉類の摂取量は、第3表より、時代の変化に伴って増加している。また、20歳代、30歳代、40歳

代で多く、50歳を超えると、加齢とともに減少する。また、新しい世代ほど多い。対象期間の摂取量は、第4表より、時代の変化と世代交代による増加が高齢化による減少を上回り、増加したと言える。卵類の摂取量は、第3表より、時代や世代の影響を受けているとは言えないが、加齢とともに増加し、40歳代をピークに減少に転じる。対象期間の摂取量は、第4表より、唯一有意性が認められた年齢効果による変化がゼロに近く、変化したとは言えない。乳類の摂取量は、第3表より、時代の変化に伴って増加している。年齢や世代の効果は、あると考えられるが、明確な傾向をつかむことができない。対象期間の摂取量は、第4表より、時代の変化により増加したと言える。

油脂類の摂取量は、第3表より、時代の変化に伴って減少している。また、高齢層ほど少なく、新しい世代ほど多い。対象期間の摂取量は、第4表より、時代の変化と高齢化による減少が世代交代による増加を上回り、減少したと言える。砂糖類の摂取量は、第3表より、時代の変化に伴って減少している。高齢層ほど多く、新しい世代ほど少ない。対象期間の摂取量は、第4表より、時代の変化と世代交代により減少したと言える。

種実類の摂取量は、第3表より、時代の変化に伴って増加している。60歳以上の高齢層で多い。また、1930年代、1940年代、1950年代生まれで多く、それ以降の新しい世代では少なく、それ以前の古い世代でも少ない。対象期間の摂取量は、第4表より、時代の変化による増加が世代交代による減少を上回り、増加したと言える。菓子類の摂取量は、第3表より、時代の変化に伴って増加している。しかし、年齢や世代の影響を受けているとは言えない。対象期間の摂取量は、第4表より、時代の変化により増加したと言える。調味嗜好飲料の摂取量は、第3表より、時代の変化に伴って増加している。また、高齢層ほど少なく（20歳代が少ないのは飲酒が関連していると考えられる；註18）、新しい世代ほど多い。対象期間の摂取量は、第4表より、時代の変化と世代交代による増加が高齢化による減少を上回り、増加したと言える。

5. おわりに

本研究では、国民栄養調査が対象とする18の食品群を対象に、摂取量を規定する、各時代、各年齢、各世代に固有の効果をパラメータとして推定した。パラメータの推定値は、仮想的パラメータの復元実験に基づく結論として、ほぼ信頼できると言える。なぜなら、各効果について、真のパラメータ群と推定パラメータ群において、パラメータの値のカテゴリー間での大小関係がほぼ一致し、それらの値の平均をゼロとする分布のばらつきもほぼ一致する。しかし、注意が必要な点の第1は、各効果におけるパラメータの値の分布のばらつきがやや過小評価されている点である。パラメータの絶対値が過小評価されていると言うこともできる。そのため、パラメータの推定値は、コウホート表上の摂取量の変化を説明するにあたり、3因子の独立した効果を過小評価し、連動した効果を過大評価している。第2は、コウホート表上における時代+年齢効果と世代効果の相関が小さいほど、その過小評価が深刻化し、推定値の信頼性も低下する点である。

そして、以上のように推定したパラメータを用い、分析対象期間（10年間）における摂取量の変化を、時代が移ったことによる変化、社会の高齢化が進んだことによる変化、世代交代が進んだことによる変化に分離した。

また、共変動最大化によるコウホート分析モデルを

提案した。本モデルの優れた点の第1は、仮定や外生的制約に頼らないという点である。第2は、推定されるパラメータが、過大評価と過小評価が同様に起こりうるという意味で、不偏推定量という点である。第3は、仮想的パラメータの復元実験による結論として、時代+年齢効果と世代効果に強い負の相関がない限り、パラメータの推定量がほぼ信頼できる点である。

しかし、弱点の第1は、パラメータの推定量が一致性を満たさない点である。そのため、標本数を増やしても、推定値が真の値に近づかない。第2は、3因子の効果の各パラメータ群内におけるパラメータの値の分布のばらつきが、不偏性を満たさず、過小評価される点である。第3は、コウホート表上における、時代+年齢効果と世代効果の相関が小さいほど、その過小評価が深刻化し、推定値の信頼性が低下する点である。

引用文献

- 石橋喜美子（2006）『家計における食料消費構造の解明：年齢階層別および世帯類型別アプローチによる』農林統計協会。
- 松田友義・中村隆（1993）「世帯主年齢階層別米消費量変化の分析」『農業経済研究』，64（4），213～220。
- Mori, H. and Gorman, W.D. (2001) A Cohort Analysis of Japanese Food Consumption: Old and New Generations, 森宏編『食料消費のコウホート分析：年齢・世代・時代』，専修大学出版，229～272。
- Mori, H. and Clason, D.A. (2005) Cohort Approach for Predicting Future Eating Habits: The Case of At-Home Consumption of Fresh Fish and Meat in an Aging Japanese Society, *International Food and Agribusiness Management Review*, 7（1），22-41。
- 森宏・石橋喜美子・田中正光・稲葉敏夫（2005）「年齢・世代効果を補正した需要弾力性の計測」『社会科学年報』，39，専修大学社会科学研究所，39～59。
- 森宏・三枝義清・川口雅正（2007）「コウホート分析における識別問題への対処：シミュレーションによる検定」『社会科学年報』，42，専修大学社会科学研究所，69～99。
- 森宏・三枝義清・石橋喜美子・華山宣胤（2008）「コウホート分析：食料消費（再訪）」『専修経済学論集』，43（2），1～44。
- 森宏・三枝義清・Dennis Clason（2009）「年齢・世代と食料消費：コウホート分析の経緯」『社会科学年報』，43，専修大学社会科学研究所，87～110。
- 森島賢（1985）「畜産物需要の世代別動向」崎浦誠司編『経済発展と農業開発』農林統計協会，54～69。
- Nakamura, T. (1986) Bayesian Cohort Models for General Cohort Tables, *Annals of the Institute of Statistical Mathematics*, 38, 353-370。

（註18）厚生労働省『国民健康・栄養調査』（平成17年）における飲酒の頻度と飲酒日の飲酒量より、1人1日当たり平均飲酒量を推定すれば、20歳代0.30合、30歳代0.57合、40歳代0.66合、50歳代0.59合、60歳代0.49合、70歳代0.30合であった。世代効果を考えないとすれば、70歳代とともに、20歳代の飲酒量は明らかに少ない。

- Stewart H. and Blisard, N. (2007) Are Younger Cohorts Demanding Less Fresh Vegetables? *Review of Agricultural Economics*, 30 (1), 43-60.
- 田中正光・三枝義清・森宏・川口雅正 (2007) 「コウホート分析における『識別問題』の克服：中村・IEモデルの比較検討」『専修経済学論集』, 42 (1), 1~44.
- Yang, Y., Schulhofer-Wohl, S., Fu, W.J. and Land, K. C. (2008) The Intrinsic Estimator for Age-Period-Cohort Analysis: What It Is and How To Use It, *American Journal of Sociology*, 113 (6), 1697-1736.

(2010年1月14日受付, 2010年9月27日受理)

要旨：コウホート表上の値の偏差は、時代、年齢、世代効果の合計と定義できる。本研究では、全変動（偏差の二乗和）は、3因子の効果の各変動とそれらの共変動により構成されることに注目し、共変動を最大化することで、その偏差を3つの効果に分離するコウホート分析モデルを提案した。共変動の最大化は、コウホート表上の3つの効果の動きを可能な限り連動させる。連動させることができない動きの部分は、分離可能な独立した動きである。連動させることができる他の部分は、分離できない連動した動きである。また、このように推定された3因子の効果が、不偏性ととも信頼性を満たす推定量であることを示した。そして、日本の国民栄養調査が対象とする18の食品群の時代、年齢、世代効果を本モデルにより計測した。

キーワード：コウホート分析, 食品摂取, 共変動最大化, 不偏性, 信頼性