原 著

欠測値がある基本味質データの相関構造の推定 Estimation of Correlation Structures of Four Basic Tastes Data with Missing Measurements

中村永友^{*1}・金子真紀子^{*2}・于 秉柯^{*3}

Nagatomo Nakamura^{*1}, Makiko Kaneko^{*2} and Yu Bing Ke^{*3}

^{*1} 札幌学院大学 経済学部経済学科, 総合教育センター

^{*2} 九州女子大学 家政学部栄養学科

^{*3} 哈爾濱醫科大學 (ハルビン医科大学)

^{*1} Department of Economics and General Education Center, Sapporo Gakuin University

^{*2} Department of Nutrition, Kyushu Women's University

^{*3} Harbin Medical University

e-mail:nagatomo@sgu.ac.jp

The means and correlation structures of four basic tastes data for Japanese and Chinese men and women, are compared. Tasting data were measured by a filter paper disk method, and recorded as missing values when subjects were not able to recognize a highest concentration of a taste substance solution. The missing values are recorded as counts, so the tasting data is expressed as censored data. A few statistics, i.e. mean vectors, variance-covariance matrices, and correlation matrices, are estimated by a multivariate statistical model for censored data to avoid biases on the estimation. Moreover, a correlative map among four tastes for men and women of Japan and China is drawn by using a principal components analysis. These results suggest that: (i) An appropriate statistical model is needed for analysing data with missing values. (ii) Positive correlations among the four basic tastes were estimated. (iii) The differences in experience and custom of eating appear as taste sensitivities and mean and correlation structures, such that the correlation between men and women is positive, while between Japanese and Chinese is negative. (iv) The correlation structures suggest a kind of an ordering of the tastes for Japanese and Chinese, and for men and women.

Key words: sensitivity of tasting, a filter paper disk method, multivariate censored normal distribution model, principal components analysis.

1. はじめに

ヒトの味覚感度は、それが生死に関わることはないことから、その検査は機能が正常か否かを測定することを主眼にしたものであった。味覚感度の低下と肥満の関係が指摘されたり(水沼他

(1998), 金子他 (1999)), QOL との関係も研究されつつあり, 味覚感度の研究はその重要性が認識されつつある.

本論文は, 日本人 (20~70 歳代) と中国人 (30~40 歳代) の男性と女性を対象として, 4 基本味質 (甘味, 塩味, 酸味, 苦味) の認知閾値を測定したデータから, その味質間の相関構造を分析した. ここで「認知」とは当該呈味物質として認識できたことをさす. 呈味物質は, 蔗糖 (甘味), 塩化ナトリウム (塩味), 酒石酸 (酸味), 塩酸キニーネ (苦味) を用いた. 各物質に対して蒸留水による一定の比率による希釈で 30 段階の溶液を作成し, 測定を行った (金子他 (1999), 金子他 (2003), 水沼他 (1998), 中村他 (2003)). これらの測定値の中には最高濃度でも測定できないデータがあり, その味質だけが欠測値となっている. 欠測したデータがある場合, その個体を取り除くことは貴重なデータを無駄にすることになり, 適当な値を代入して平均や分散, 味質間の相関係数を求めた場合には, 明らかに偏った推定がされてしまう. 欠測値を無視して分析した結果の信頼性はかなり低く, その解釈は慎重でなければならない. 手元のデータを十分生かすために, 欠測を考慮した統計モデルを用いてできるだけ偏りのないパラメータの推定を行いたい.

本研究の目的は, このような欠測を伴う味覚感度を測定したデータに対して, 多変量打ち切り正規分布モデルにより偏りのないパラメータを推定し, 基本味質の平均構造と味質間の相関構造を探ることである. さらに, 推定された相関係数行列を基礎として主成分分析を行い, 地図を描くことで日中男女の 4 味質の関係を明らかにしていく. ただし, 綿密な実験計画の下で日本全域, 中国全域で行われた調査データではないので, 一事例研究であり, 本論文の結論によって日中間の普遍的なことは断言できない. しかし, 味覚感度の測定において, このように非常に細かい区分で測定され, 高度な統計処理が行われた論文は前例がない.

過去の類似の研究として, 山内他 (1995) は主成分分析により基本味質の認知閾値と検知閾値の関係を探っている. 中村他 (2003) は, 欠測値を含む 4 つの各味質に対して, 1 次元の打ち切り正規分布により平均や分散を求めるなどのデータの基本的な分析を行っているが, これ以外に高度な統計処理を用いた報告や相関構造に關した分析は見あたらない.

2. 分析対象のデータ

本論文で用いたデータセットは, 男性は金子他 (1999), 女性は水沼他 (1998) で用いたものである. 日本では佐賀県在住の学生, 公務員, 銀行員, 会社員, 無職者の男女から調査対象者を抽出した. 中国の調査場所は, 東北部の黒龍江省の都市部 (ハルビン市) と農村部 (富裕県前進村) である. この地域はいくつかの民族が混在しており, 民族構成として, ハルビン市には漢族 (多数民族), 満州族, 朝鮮族, 回族 (少数民族) などが, 前進村には漢族 (多数民族), 満州族, キルギス族, 蒙古族, ダオール族 (少数民族) などが居住している. これらの地域に住む男女から調査対象者を抽出した. また, 日中共に基本的に口腔, 中耳などに疾患がなく, 味覚障害や臭覚障害のない対象者である. 日本では 1996 年から 1998 年, 中国では 1998 年から 2004 年にかけて段階的な調査を行った.

測定は次の方法で行った. 4 つの基本味質の味覚感度を測定するため, 各溶液を蒸留水により表 1 の最高濃度から一定の比率による希釈を行い, 30 段階の味質溶液を作製した. 測定の順番に

表 1. 呈味物質

基本味質	呈味物質	濃度範囲 (g/dl)	第 k 段階目の濃度計算式 ($k = 1, \dots, 30$)
甘味	蔗糖	80.0~0.0190	$80.0 \times 0.75^{30-k}$
塩味	塩化ナトリウム	30.0~0.0105	$30.0 \times 0.76^{30-k}$
酸味	酒石酸	8.0~0.0019	$8.0 \times 0.75^{30-k}$
苦味	塩酸キニーネ	4.0~0.0003	$4.0 \times 0.72^{30-k}$

準じて最低濃度を第 1 段階目, 最高濃度を第 30 段階目とした. 測定方法は, 舌に対する味質の刺激量や刺激範囲を一定に保つための最適な方法とされている, ろ紙ディスク法を用いた. 舌尖部における認知閾値を低濃度から上昇法により調べ, 味質を認知するまで続けた. また味質の測定順序は甘味, 塩味, 酸味, 苦味の順で, 次の味質の測定の際には蒸留水でよく含嗽させた(各味質溶液の各段階の詳細な濃度系列は中村他 (2003) を, 調査に関する詳細は金子他 (1999) と水沼他 (1998) を参照).

図 1 に日本人男性の各味質の濃度番号によるヒストグラムを示す. 他の日本人女性, 中国人男女も同様な傾向があるため, データの打ち切りを考慮した正規分布の仮定は妥当であると考ええる. 味質の濃度系列番号は対数配列しているので, 測定値の番号を濃度に変換した値は対数正規分布をしていると見なせる. 以上より本論文では, 濃度系列の番号を連続的な値(連続型変数)として扱って分析する. また, 測定時に第 30 段階目の最高濃度でも味質を認知できなかった場合は欠測値として扱い, 欠測数も同時に記録する.

図 1. 日本人男性の各味質のヒストグラム

横軸は各呈味物質の濃度番号. 第 30 段階が最高濃度でそれより大きな値は個体数のみが記録されている. 分布の切断を考慮すれば, 正規分布のあてはめは妥当である.

表 2 に分析対象データの基本的なデータ数を示す. 表の下半分はいずれかの味質で欠測がある個体数である. 各味質でかなりの数が欠測しており, とくに 4 味質がすべてそろっている割合は, 中国人は有意に低いことがわかる. また中国人女性で 4 味質すべてが欠測している人数の 11 名 (14.2%) は割合としてかなり高い.

3. 統計モデル

測定したデータには欠測値が多数存在するため, 偏りのないパラメータ(味質の平均や分散, 味質間の共分散, 相関係数)の推定には, 多変量版の打ち切り正規分布モデルを用いる. 欠測値の統計モデルに関する基本的考え方は Dempster, et al. (1977) や Little and Rubin(2002) にある.

表 2. 分析データのデータ数と欠測の状況

		日本人男性 20~70歳代	日本人女性 20~70歳代	中国人男性 30~40歳代	中国人女性 30~40歳代	日本人男性 30~40歳代	日本人女性 30~40歳代
標本の数 (N)		472	343	245	290	102	113
4 味質がそろっている標本の数 (完全データ数の割合)		308 (65.3%)	237 (69.1%)	77 (31.4%)	137 (47.2%)	69 (67.6%)	73 (64.6%)
いずれかの味質で 欠測がある 標本の数	甘味	12	11	6	6	1	3
	塩味	33	27	15	22	9	13
	酸味	45	32	57	71	9	13
	苦味	11	10	8	8	1	0
	甘, 塩	7	3	7	4	1	1
	甘, 酸	3	2	9	4	1	2
	甘, 苦	1	1	4	3	0	0
	塩, 酸	16	10	12	13	3	4
	塩, 苦	13	2	4	0	0	0
	酸, 苦	6	2	10	8	3	1
	甘, 塩, 酸	5	1	8	7	3	0
	甘, 塩, 苦	3	0	3	1	0	0
	甘, 酸, 苦	2	1	6	1	0	0
	塩, 酸, 苦	3	2	8	3	1	2
	甘, 塩, 酸, 苦	4	2	11	2	1	1

例えば 2 つの味質 (変数) で, 第 1 の味質を x , 第 2 の味質を y , そして 2 次元の正規分布モデルを仮定し, 図 2 に示すような状況を統計モデルとして表現して, それに基づいて平均ベクトル μ と分散共分散行列 Σ を推定することができる.

4 味質の 4 次元空間においては, c_x, c_y, c_z, c_w をそれぞれ各味質の打ち切り点とすると, データが存在する区間は $R_x = [-\infty, c_x]$, $R_y = [-\infty, c_y]$, $R_z = [-\infty, c_z]$, $R_w = [-\infty, c_w]$ と書け, 欠測区間は $\bar{R}_x, \bar{R}_y, \bar{R}_z, \bar{R}_w$ となる. このとき, 4 次元空間内の領域は $2^4 = 16$ 分割される. 4 味質のデータがそろっている領域を R_0 , 欠測区間を含む領域を $R_1, R_2, \dots, R_{15}$, 領域 R_k のデータの個数を m_k ($k = 0, \dots, 15$) とすると, 対数尤度関数は

$$\log L = \sum_{i=1}^{m_0} \log f(x_i, y_i, z_i, w_i \mid \mu, \Sigma) + \sum_{k=1}^{15} m_k \log \int_{R_k} f(t \mid \mu, \Sigma) dt \quad (1)$$

となる. ここで, f は 4 次元の正規分布の密度関数, $t = (x, y, z, w)^T$, 積分は 4 つの変数による多重積分である. この対数尤度関数を目的のパラメータで偏微分してゼロとおいて解くと, 推定したいパラメータである平均 μ と分散共分散行列 Σ の推定値は, 以下のように与えられる (中村他, 2005).

$$\hat{\mu} = \frac{1}{n+m} \left\{ n\tilde{\mu} + \sum_{k=1}^{15} m_k \frac{\int_{R_k} t f(t \mid \mu, \Sigma) dt}{\int_{R_k} f(t \mid \mu, \Sigma) dt} \right\}, \quad (2)$$

図 2. データの存在領域の模式図

2次元平面は c_x, c_y の2つの境界点により、4つの領域 R_0, R_1, R_2, R_3 に分割される。 P_k は領域 R_k における正規分布の確率 ($k = 0, 1, 2, 3$)。2次元平面で分割された領域は次の通り： $R_0 = R_x \times R_y$, $R_1 = \bar{R}_x \times R_y$, $R_2 = R_x \times \bar{R}_y$, $R_3 = \bar{R}_x \times \bar{R}_y$ 。それぞれの領域 R_k におけるデータ数を m_k ($k = 0, 1, 2, 3$) とする。2次元の場合は $2^2 = 4$ 分割であるが、4味質を同時に考えた場合は、4次元空間が $2^4 = 16$ 分割される。

$$\hat{\Sigma} = \frac{1}{n+m} \left\{ n\tilde{\Sigma} + \sum_{k=1}^{15} m_k \frac{\int_{R_k} (t-\mu)(t-\mu)^T f(t|\mu, \Sigma) dt}{\int_{R_k} f(t|\mu, \Sigma) dt} \right\}. \quad (3)$$

ここで、 $\tilde{\mu} = n^{-1} \sum_{i=1}^n x_i$ と $\tilde{\Sigma} = n^{-1} \sum_{i=1}^n (x_i - \tilde{\mu})(x_i - \tilde{\mu})^T$ は欠測のないデータから計算される擬平均と擬分散共分散行列、 $n = m_0$, $m = \sum_{k=1}^{15} m_k$ である。

(2) 式と (3) 式の各右辺には推定したいパラメータが含まれるので、この式から直接的に推定値を計算することはできない。数値的に推定値を得るためには EM 法 (Expectation-Maximization Algorithm; Dempster et al. 1977) による反復計算で求めることができる。この推定アルゴリズムは、欠測情報の期待値推定である E ステップと、対数尤度の最大化の M ステップからなる。具体的には以下の手順で行う。

欠損分布モデルのパラメータ推定法

S1 (E-ステップ: 欠測領域の確率の推定) 現在のパラメータ値を固定して、欠測領域の確率や関連する積分 ((1) 式, (2) 式, (3) 式の積分記号の部分) を計算する。

S2 (M-ステップ: 対数尤度の最大化) パラメータ値を (2) 式と (3) 式により更新する。

S3 収束条件を満足するまで S1 と S2 を繰り返す。

この推定アルゴリズムは EM 法の理論により収束性が保証されている。しかし、データ数が少ない、正規分布が仮定できないなど条件が悪い場合には解釈が困難な値に収束することがある。

4. 分 析

分析ではまず多変量打ち切り正規分布モデルによって偏りのない平均ベクトルや分散共分散行列を推定する．そして推定された相関係数行列を使って各味質間の相関構造を主成分分析により探る．分析の対象は日本人と中国人のすべてのデータであるが，中国人が 30～40 歳代のみしかないので，これと対比する意味で日本人も 30～40 歳代の分析も付加した．

4.1 パラメータの推定

多変量打ち切り正規分布モデルを使ってパラメータ推定した結果を表 3 に示す．表の左側の補正前の平均は，欠測のない完全データのみで計算したものである．また補正前の分散共分散行列や相関係数行列は，2 つの味質の組み合わせごとに欠測のないデータを用いて計算している．各群の補正された平均値を比較したものが図 3 で，4 群の平均構造の比較ができる．

各味質の平均について以下のことが観察される．(1) 補正後の値が大きくなっていることから，欠測値を無視した場合に各値を過小評価することになる．(2) 補正された相対量を比較すると，中国の方が大きい．とくに酸味は補正前の 20% 以上も大きくなっている．(3) 図 3 より，相対的に味覚の感度が一番よいのは日本人女性で，続いて日本人男性，中国人女性，中国人男性という傾向がある(日本人女性 > 日本人男性 > 中国人女性 > 中国人男性)．(4) 日本人男女 30～40 歳代は，女性の苦味の感度がよい(値が小さい)ことを除けば，日本人全体の平均とほぼ同じである．

次に，分散共分散行列と相関係数行列について，補正後の推定値から以下のことが観察される．(1) ほぼすべてにおいて，4 味質の間には何らかの正の相関が認められる，(2) 日本人の相関係数は，そのほとんどにおいて補正後の値が大きくなっている．このことから欠測値を無視すると各

図 3. 日中男女の推定された平均のプロフィール

各群の平均構造の比較．縦軸は各味質溶液の濃度番号であることに注意．一定の比率による希釈で検査溶液を作成していることと，データが対数正規性があるので，濃度番号を連続変数として扱うことが可能である．

表 3. パラメータの推定値

		補正前の推定値 (欠測のないデータによる)				補正後の推定値 (統計モデル適用)			
		甘味	塩味	酸味	苦味	甘味	塩味	酸味	苦味
日本人男性									
平均		22.09	19.75	21.69	19.26	22.92	21.94	24.20	20.71
標準誤差		0.246	0.259	0.264	0.312	0.255	0.308	0.300	0.322
標準偏差		5.344	5.637	5.735	6.784	5.549	6.682	6.517	6.993
分散共分散行列 相関係数行列	甘味	28.55	10.61	8.85	13.28	30.79	13.14	10.67	14.44
	塩味	0.352	31.78	15.66	18.35	0.354	44.65	20.19	22.90
	酸味	0.289	0.484	32.89	18.95	0.295	0.464	42.48	23.44
	苦味	0.366	0.480	0.487	46.02	0.372	0.490	0.514	48.90
日本人女性									
平均		19.98	17.58	19.54	16.83	21.04	20.54	21.84	18.28
標準誤差		0.281	0.324	0.337	0.361	0.300	0.376	0.382	0.389
標準偏差		5.212	5.998	6.239	6.690	5.555	6.972	7.081	7.213
分散共分散行列 相関係数行列	甘味	27.16	9.22	8.93	8.53	30.85	14.98	10.03	12.40
	塩味	0.295	35.97	22.55	16.68	0.387	48.61	31.70	22.57
	酸味	0.275	0.603	38.92	19.65	0.255	0.642	50.14	25.41
	苦味	0.245	0.416	0.471	44.76	0.310	0.489	0.497	52.03
中国人男性									
平均		23.54	22.42	23.50	21.33	26.09	25.37	28.85	24.08
標準誤差		0.325	0.321	0.371	0.418	0.320	0.388	0.362	0.439
標準偏差		5.085	5.031	5.803	6.548	5.013	6.079	5.674	6.878
分散共分散行列 相関係数行列	甘味	25.86	8.66	12.10	8.87	25.13	11.81	4.73	13.12
	塩味	0.338	25.31	11.05	12.17	0.388	36.95	4.30	13.69
	酸味	0.410	0.379	33.67	18.16	0.166	0.125	32.19	12.80
	苦味	0.266	0.370	0.478	42.88	0.380	0.327	0.328	47.31
中国人女性									
平均		21.86	19.47	22.26	18.35	23.31	22.35	26.82	19.91
標準誤差		0.304	0.318	0.317	0.374	0.315	0.393	0.366	0.376
標準偏差		5.181	5.418	5.407	6.375	5.365	6.694	6.228	6.411
分散共分散行列 相関係数行列	甘味	26.85	14.13	8.25	11.58	28.78	18.69	16.23	12.19
	塩味	0.503	29.36	9.27	14.47	0.520	44.80	8.53	14.03
	酸味	0.294	0.317	29.23	14.05	0.186	0.205	38.79	14.61
	苦味	0.350	0.419	0.407	40.65	0.354	0.327	0.366	41.10
日本人男性 30～40 歳代									
平均		22.01	20.13	20.81	19.57	23.19	23.13	23.97	21.20
標準誤差		0.592	0.515	0.560	0.637	0.595	0.618	0.656	0.650
標準偏差		5.979	5.203	5.660	6.437	6.006	6.239	6.623	6.561
分散共分散行列 相関係数行列	甘味	35.75	10.45	9.858	17.14	36.07	13.87	13.07	16.77
	塩味	0.336	27.07	18.17	14.62	0.370	38.92	25.83	17.15
	酸味	0.291	0.617	32.04	13.56	0.329	0.625	43.87	18.40
	苦味	0.445	0.437	0.372	41.43	0.426	0.419	0.430	43.04
日本人女性 30～40 歳代									
平均		20.58	16.89	18.23	13.55	21.93	21.15	22.30	15.93
標準誤差		0.397	0.482	0.545	0.548	0.453	0.662	0.693	0.630
標準偏差		4.220	5.125	5.792	5.824	4.818	7.037	7.362	6.702
分散共分散行列 相関係数行列	甘味	17.81	5.87	-0.81	2.73	23.22	12.13	3.00	6.84
	塩味	0.272	26.26	16.33	4.87	0.358	49.52	30.84	16.98
	酸味	-0.033	0.550	33.55	9.04	0.085	0.595	54.19	23.22
	苦味	0.111	0.163	0.268	33.92	0.212	0.360	0.471	44.92

1) 補正後の相関係数は、日本人女性 30～40 歳代の甘味-酸味の 0.085 をのぞいて 5%有意であり、中国人男性の塩味-酸味の 0.125 をのぞいて 1%有意である。2) 表中の平均値から各味質の濃度への変換式は、中村他 (2003) の表 3 に示してある。3) 「分散共分散行列/相関係数行列」は、下三角行列が相関係数、対角成分と上三角行列が分散共分散行列である。

値を過小評価することになる。(3) 一方, 中国人の相関係数はむしろそうでない場合があり, ほぼおなじ味質の組み合わせ(甘酸, 塩酸, 塩苦, 酸苦)の男女で値が小さくなっている。これは日本人に比べて相対的に欠測数が多いためであると考えられる。(4) 同じ国の男女の味質間の相関係数はほぼ同じ値の傾向がある, (5) 日本人の甘味と酸味の間の相関は低めであるが, 中国人はさらに低い, (6) 日本人の塩味と酸味の相関に比べて中国人はかなり低い。(7) 日本人男性とその30~40歳代はほぼ同じであるが, 日本人女性30~40歳代は日本人女性全体に比べて小さく(相関が低く)なっている。

また, 参考のため仮説検定を行い, その結果を表4に示す。

表4. 各種検定の結果 *1

番号	群の組合せ	検定の種類	帰無仮説 *3	検定の結果
T1	日中男女	平均の平行性	$H_{01}: \mathbf{C}\boldsymbol{\mu}_{\text{日男}} = \mathbf{C}\boldsymbol{\mu}_{\text{日女}} = \mathbf{C}\boldsymbol{\mu}_{\text{中男}} = \mathbf{C}\boldsymbol{\mu}_{\text{中女}}$	0.01% 有意 (p 値 = 0.0)
T2	日本人男性と女性	平均の平行性	$H_{02}: \mathbf{C}\boldsymbol{\mu}_{\text{日男}} = \mathbf{C}\boldsymbol{\mu}_{\text{日女}}$	有意差なし (p 値 = 0.18)
T3	中国人男性と女性	平均の平行性	$H_{03}: \mathbf{C}\boldsymbol{\mu}_{\text{中男}} = \mathbf{C}\boldsymbol{\mu}_{\text{中女}}$	1% 有意 (p 値 = 0.007)
T4	日本人男女	分散共分散行列の同等性	$H_{04}: \boldsymbol{\Sigma}_{\text{日男}} = \boldsymbol{\Sigma}_{\text{日女}}$	0.01% 有意 (p 値 = 0.0)
T5	中国人男女	分散共分散行列の同等性	$H_{05}: \boldsymbol{\Sigma}_{\text{中男}} = \boldsymbol{\Sigma}_{\text{中女}}$	有意差なし (p 値 = 0.35)
T6	日本人と中国人 *2	分散共分散行列の同等性	$H_{06}: \boldsymbol{\Sigma}_{\text{日}} = \boldsymbol{\Sigma}_{\text{中}}$	0.01% 有意 (p 値 = 0.0)
T7	日本人男女	平均の同等性	$H_{07}: \boldsymbol{\mu}_{\text{日男}} = \boldsymbol{\mu}_{\text{日女}}$	0.01% 有意 (p 値 = 0.0)
T8	中国人男女	平均の同等性	$H_{08}: \boldsymbol{\mu}_{\text{中男}} = \boldsymbol{\mu}_{\text{中女}}$	0.01% 有意 (p 値 = 0.0)
T9	日本人と中国人 *2	平均の同等性	$H_{09}: \boldsymbol{\mu}_{\text{日}} = \boldsymbol{\mu}_{\text{中}}$	0.01% 有意 (p 値 = 0.0)
T10	日本人男女	相関係数行列の同等性	$H_{0a}: \mathbf{R}_{\text{日男}} = \mathbf{R}_{\text{日女}}$	0.01% 有意 (p 値 = 0.0)
T11	中国人男女	相関係数行列の同等性	$H_{0b}: \mathbf{R}_{\text{中男}} = \mathbf{R}_{\text{中女}}$	有意差なし (p 値 = 0.80)
T12	日本人と中国人 *3	相関係数行列の同等性	$H_{0c}: \mathbf{R}_{\text{日}} = \mathbf{R}_{\text{中}}$	0.01% 有意 (p 値 = 0.0)

*1 すべて多変量の検定を行っている。検定の種類は以下の通りである(Seber(1984)の該当する節を示す)。T1~T3: プロフィール分析としてホテリングの T^2 検定(3.6.4節), T4~T6, T10~T12: 等分散共分散性(相関係数)の検定として尤度比検定(3.6.1節), T7~T9: 等分散共分散仮定下の等平均性の検定としてホテリングの T^2 検定(3.6.2節)。*2 日本人と中国人の検定は, 男女のパラメータを合併して比較を行った。*3 $\boldsymbol{\mu}_X$: X の平均ベクトル, $\boldsymbol{\Sigma}_X$: X の分散共分散行列, $\mathbf{R}_X$: X の相関係数

行列, $\mathbf{C}$ は差分を求める行列で, $\mathbf{C} = \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix}$ である。

4.2 相関構造の探索

多変量打ち切り正規分布モデルにより推定された分散共分散行列と相関係数行列から, 日中男女それぞれの4味質の相関構造を主成分分析で探る。主成分分析により得られた日中男女別の固有値表を表5に, 因子負荷量を表6に示す。

表5からわかることは, (1) 日本人男女は固有値の寄与率がほぼ同じ値であり, 同時に中国人男女でも同じことがいえる, (2) しかし日本人と中国人ではその寄与率の構成が異なる, (3) 第1主成分の寄与率を大きい順に並べると, 「日本人女性 > 日本人男性 > 中国人女性 > 中国人男性」となり, 前節で記述した味覚の平均構造(平均値の相対的な感度の良さ)の序列と一致している。一般には平均と分散の間には相関はなく, 独立である。主成分分析は対象となるデータは相関構造を表す相関係数行列や分散共分散行列であり, 平均構造を分析対象としていないが, 平均構造

表 5. 日中男女の主成分分析の結果：固有値表 *¹

		第 1 主成分	第 2 主成分	第 3 主成分	第 4 主成分
日本人男性	固有値	2.256	0.734	0.535	0.475
	寄与率	56.4%	18.4%	13.4%	11.9%
	累積寄与率	56.4%	74.8%	88.1%	100.0%
日本人女性	固有値	2.318	0.790	0.553	0.339
	寄与率	58.0%	19.7%	13.8%	8.5%
	累積寄与率	58.0%	77.7%	91.5%	100.0%
中国人男性	固有値	1.878	0.933	0.617	0.572
	寄与率	47.0%	23.3%	15.4%	14.3%
	累積寄与率	47.0%	70.3%	85.7%	100.0%
中国人女性	固有値	1.994	0.931	0.600	0.475
	寄与率	49.8%	23.3%	15.0%	11.9%
	累積寄与率	49.8%	73.1%	88.1%	100.0%
日本人男性 30-40	固有値	2.308	0.759	0.562	0.371
	寄与率	57.7%	19.0%	14.1%	9.3%
	累積寄与率	57.7%	76.7%	90.7%	100.0%
日本人女性 30-40	固有値	2.088	0.941	0.653	0.318
	寄与率	52.2%	23.5%	16.3%	8.0%
	累積寄与率	52.2%	75.7%	92.0%	100.0%

*¹ 日中男女ごとに補正した相関係数行列を基にして主成分分析を行った結果.表 6. 日中男女の因子負荷量 *¹

		第 1 主成分	第 2 主成分	第 3 主成分	第 4 主成分
日本人男性	甘味	0.639	0.755	0.103	0.104
	塩味	0.779	-0.118	-0.610	0.085
	酸味	0.768	-0.361	0.346	0.400
	苦味	0.807	-0.140	0.178	-0.545
日本人女性	甘味	0.587	0.801	0.034	0.111
	塩味	0.852	-0.097	0.301	-0.416
	酸味	0.816	-0.344	0.250	0.390
	苦味	0.762	-0.140	-0.631	-0.038
中国人男性	甘味	0.823	0.181	-0.334	0.423
	塩味	0.629	0.661	0.399	-0.090
	酸味	0.630	-0.663	0.376	0.151
	苦味	0.827	-0.178	-0.257	-0.468
中国人女性	甘味	0.760	0.416	0.035	0.498
	塩味	0.754	0.406	0.237	-0.460
	酸味	0.565	-0.708	0.418	0.073
	苦味	0.727	-0.305	-0.607	-0.101
日本人男性 30-40 歳代	甘味	0.673	0.612	0.411	0.062
	塩味	0.812	-0.359	0.169	-0.428
	酸味	0.802	-0.417	0.042	0.426
	苦味	0.743	0.287	-0.602	-0.049
日本人女性 30-40 歳代	甘味	0.477	0.857	-0.074	0.180
	塩味	0.835	0.063	0.414	-0.356
	酸味	0.805	-0.411	0.202	0.378
	苦味	0.718	-0.182	-0.660	-0.128

*¹ 日中男女ごとに補正した相関係数行列を基にして主成分分析を行った結果.

と同じ傾向が出たことは興味深い。

主成分分析の一般論として、すべての相関係数が正で相対的に大きい(相関が高い)ときには、第1主成分の寄与率が比較的大きくなり、因子負荷量の第1主成分がすべて正になる。今回分析している味覚データはこの状況と一致している。つまり、表6の第1主成分は各味質と比較的大きな相関を示している。このことから、各群の第1主成分は全体的な味覚感度を表すと解釈でき、表5の第1主成分の寄与率の大きさは、味覚間相関の総合的な大きさ(相関の高さ)を表す一つの指標になる。よって、第1主成分の寄与率の大きさの比較は、各国男女別の味覚感度の序列を表すものと言える。

また、寄与率は行列の固有値であるので、その構成が類似していることは、相関係数行列が類似していることを意味する。したがって、各国の男女の味覚の相関構造は似ていると言える。ただし、検定の結果(表4)からは、日本人男女の分散共分散行列や相関係数行列はデータ数が十分多いので差異があると判断される。

表6をさらにわかりやすくするために、多少の厳密さを欠くが、同一の主成分軸内にこれらの因子負荷量を表示すると図4のようなになる。累積寄与率から、第1、第2主成分で全体の約75%の情報量がある。このようなことができる理由は以下の通りである：因子負荷量の第2主成分をみ

図 4. 因子負荷量の散布図

■は日本人男性, △は日本人女性, ●は中国人男性, ○は中国人女性, ×は日本人男性 30~40 歳代, ◇は日本人女性 30~40 歳代, 実線楕円は日本人, 細い実線楕円は日本人 30~40 歳代, 破線楕円は中国人の同じ味質を表す。

ると各味質の正負の符号は完全に一致しないが、その絶対値の大きさや符号等を比較すれば、日中男女の第2主成分はほぼ同じ方向の主成分軸であると見なすことができ、全体の傾向を読み取ることが目的であれば、同一平面への表示は許容されるであろう。

この図から以下のことが観察される。(1) 第1主成分は主に日本人甘味、中国人酸味とこれ以外を分けることができる。(2) 第1、第2主成分を組み合わせると、日本人甘味と中国人酸味を分けることができる。すなわち第2主成分軸は、正方向が日本人甘味を、負方向が中国人酸味の情報を説明する軸と解釈できる。(3) 中国人の塩味を除いて同じ国の男女の同じ味質は空間内で近いところに付置している。このことは味質間に類似の相関構造があることを示している。(4) 日本人の塩味、酸味、苦味と中国人の苦味が1つの大きな塊を形成しているが、これらの特徴は他の主成分に出ている。(5) 各味質で日本人と中国人の傾向の差異が見て取れる。

5. 考 察

分析の結論から、データの偏り、相関構造の類似性、酸味の食経験、喫煙・加齢・肥満について考察する。

分析対象は日本人が佐賀県在住の20~70歳代、中国人は黒竜江省の30~40歳代であることから、単純な日本人と中国人の比較とはならない。本論文の結果からは一般論は言えず、事例研究の域を出ない。しかし、このような規模の詳細な味覚感度調査は非常に珍しく、データの入手も簡単ではないこと、値打ち切りされているデータの分布に対して多変量打ち切り正規分布を適用して偏りのない推定値を求めたことは、類似調査の基礎的研究となるであろう。

次に、検定の結果(表4)はあくまでも補助的に行ったものである。データ数が比較的多いのでほとんどが棄却され、分散共分散行列や相関係数行列が同等と見なせるケースは中国人男女以外はない。多数のデータ数の下で行った検定の結果、棄却されないものがあったことは、むしろ、違いが際だっていないと考えられる。差異を積極的に示すのであれば、むしろ相応の統計モデルを立てて情報量規準により判断する方が明快であろうが、本論文ではここでとどめておく。主成分分析で得られた各数値の近さ(地図上の付置)や寄与率の構成比率やその類似性から、各国男女間の差異や序列が結論付けられる。

酸味について中国人が日本人に比べてかなり低いことや、表2の酸味の欠測数(男性57名=23%、女性71名=24%)が他の味質に比べて飛び抜けて多い事実は、酸味の食習慣と食経験の違いを考慮する必要があると考えられる。日本人は主成分が酢酸の醸造酢を主に食しているが、中国人はモチ米から作られる黒酢を長時間熟成させた「香醋」をはじめとして、各地で独自に醸造したものを食している。これはアミノ酸が多量に含まれ、日本のお酢に比べて酸味がきついという決定的な違いがある。また、健康食品的に日常的に愛飲している地域や民族などがあることから、酸味に関する味覚が麻痺しているということも考えられる(本調査のデータではこの件の聞き取りはしていない)。直接的な酸味を測定するために酸味の呈味物質は酒石酸を用いているが、中国人はこのような「酸っぱさ」の食経験がないために、味覚感度が低いとも考えられる。以上のことは科学的に証明されていることではないので、今後の研究課題としたい。

味覚感度の低下に加齢、喫煙、肥満が関係していることは、経験的にも知られていて、近年そ

れが社会的にはっきり意識されるようになってきた。本報告で用いたデータは、日本では1996年から1998年、中国では1998年から2004年の調査で、これに基いて水沼他(1998)、金子他(1999)、金子他(2003)が女性の肥満と加齢による味覚感度の低下のあることが報告されている。本論文のデータ解析でも年齢、喫煙、BMIの情報により層別して分析することが望ましいが、当初からこれらのことを意識した調査ではなかったため、すべての調査でこれらのデータを取っていなかった。このため、本論文は味覚の全体の傾向を分析し報告した。

結果と以上の考察より、本論文の主たる結論は、(1) 欠測値を含むデータに対しては相応の統計モデルを用いて分析する必要がある、(2) 4つの基本味質間には何らかの正の相関がある、(3) 食習慣・食経験の違いが味覚感度やその平均・相関構造として現れていて、日中間で味質間の相関構造は異なり、各国男女間で相関構造が類似している、(4) 味覚感度の平均構造と相関構造の両方の推定・分析結果から、データを巨視的に眺めると「日本人女性 > 日本人男性 > 中国人女性 > 中国人男性」という序列が示唆された、である。

6. まとめ

日本人(20~70歳代)と中国人(30~40歳代)の4つの基本味質の味覚感度を調査したデータの分析を行った。データにはいずれかの味質で欠測値があるため、できるだけ偏りのない平均や相関係数を求めるために、欠測値を考慮した統計モデルを提示してパラメータ推定を行った。データは綿密な実験計画による調査で得られたものではないので、本論文は一事例解析という位置づけとなるが、このようなデータに対する分析の方法論の提示と味覚における日中間の傾向を探ることができた。

適切な統計モデルにより推定した結果、(1) 欠測を無視すると各味質の平均は過小評価することになる、(2) 日本人の方が味覚感度が高い傾向にある、(3) 日中男女すべてで4つの味質間に正の相関が認められる、等が示された。

さらに、推定した相関係数行列を用いて主成分分析を行い、日中男女の4味質の関連を探った。その結果、(1) 同じ国の男女は主成分寄与率が類似しており、日中間では異なる、(2) 第2主成分までの軸の解釈は日本と中国でほぼ同じである、(3) 第2主成分までの情報を使って日中男女の4味質の関係を地図で表示したところ、同じ国の男女はほぼ近い位置に付置され、日中の違いも現れていた。

2つの分析を通して以下のことが示唆された。(1) 同じ国の男女は味覚の相関構造が似ている、(2) 日本人と中国人の味覚の相関構造に差異がある、(3) 味覚感度の平均・相関構造の分析から、「日本人女性 > 日本人男性 > 中国人女性 > 中国人男性」という味覚感度(味覚構造)に関するある種の順序をつけられる。

謝 辞

本論文を完成するにあたり、匿名の2名の査読者から不備をご指摘して頂きました。この場を借りて御礼申し上げます。また、有意義な議論と適切なコメントを頂いた佐賀大学 文化教育学部 教授 水沼俊美氏に、データ取得・整理に関して協力していただいた哈爾濱医科大学(ハルビン医科

大学) 于 維漢氏に感謝いたします。本研究は札幌学院大学研究促進奨励金 (SGU-S05-198007-02) の補助を受け、2008 年度札幌学院大学長期国内研究員の期間内に完成させました。

参考文献

- Dempster A., N. Laird, and D. Rubin (1977). Maximum likelihood from incomplete data via the EM algorithm, *Journal of the Royal Statistical Society, Series B*, **39**, 1-38.
- 金子真紀子, 水沼俊美, 久野(永田)一恵, 荒尾恵介, 堀尾拓之, 久藤麻子, 坂井堅太郎, 于 秉柯, 于 維漢 (1999). 肥満における味覚の変化について, *肥満研究*, **5**, 30-34.
- 金子真紀子, 于 秉柯, 于 維漢, 水沼俊美, 中島 滋, 辻 眞紀子, 久野一恵, 久木野憲司, 坂井堅太郎 (2003). 中国東北部黒竜江省の農村と都市にすむ中年女性の食生活状況, *肥満研究*, **9**, 45-50.
- Little R. J. A. and D. Rubin (2002). *Satatistical Analysis with Missing Data*, 2nd eds., Wiley, New York.
- 水沼俊美, 金子真紀子, 久野(永田)一恵, 荒尾恵介, 堀尾拓之, 久藤麻子, 坂井堅太郎, 真鍋祐之, 久木野憲司 (1998). 女性の味覚感度は加齢で低下し肥満では酸味が低下する, *肥満研究*, **4**, 297-301.
- 中村永友, 金子真紀子, 水沼俊美, 于 秉柯, 于 維漢, 坂井堅太郎, 久野建夫, 久木野憲司 (2003). 日本人女性と中国人女性の味覚感度の比較—調査データの基本分析について—, *肥満研究*, **9**, 308-316.
- 中村永友, 上野玄太, 樋口知之, 小西貞則 (2005). 欠損混合分布モデルとその応用, *応用統計学*, **34**, 57-75.
- Seber, G. A. F. (1984). *Multivariate Observations*, Wiley, New York.
- 山内由紀, 遠藤壮平, 酒井文隆, 吉村 功 (1995). 全口腔法味覚検査(第1報)—基礎的検討および主成分分析— *日本耳鼻咽喉科学会会報*, **95**, 98-129.