【論文】

ロジスティック回帰モデルによる食味評価の分析

東北農業研究センター 下山 禎

Analysis of Food Taste Examination Using Logistic Regression Model

Tadashi SHIMOYAMA

National Agricultural Research Center for Tohoku Region

Abstract:

The purpose of this paper is examine what factors have working for consumers' attitudes for choice of Yuba. I approached this purpose by using Logistic Regression Model and Customer Satisfaction Analysis as preliminary consideration.

The Yuba that became an object of study is made from the soybean that is called Suzusayaka. Suzusayaka is soybean that is lacking all lipoxygenase isozymes developed in National Agricultural Research Center for Tohoku Region. So the soybeans' processed products do not smell green. We are expecting Suzusayaka to become the material of Yuba in the Tohoku region.

The finding were follows. (1) As a result of the CS analysis, the importance degree and the satisfaction rating of the "Thick" and "Sweetness" are high. And the satisfaction rating is high though the importance degree of "Smelling of Green" and "Over-Stewed Smell" is not high. Neither the importance degree nor the satisfaction rating of "Savor", "Looks", and "Texture" are too high. (2) The calculation process only shows externals of consumers no statistics cognitive construct by the CS analysis. (3) As a result of Logistic Regression Model, the confounding factors were observed in the "Looks" "Savor" and "Thick". (4) Finally, decided the model which consisted of three factors of "Looks" "Savor" and "Sweetness" by using backward method.

From these findings, Logistic Regression Model is more suitable for the analysis of a more accurate food taste examination than the CS analysis because of the above-mentioned result. And, through the decision process of this model, it is possible to use this method of searching for the major factor in a wide field.

1. はじめに

加工食品の分野における消費者ニーズ探索のためのマーケティング・リサーチでは、アンケート調査や食味・官能試験によるデータ解析がその中枢を担ってきた。それに適用される分析手法も有意差の検定から共分散構造分析まで幅広く、それ

ぞれの分野で研究が深められている。近年社会科学の分野では「選択型コンジョイント分析」を中心にロジスティック回帰モデルも広く利用されるようになってきた(註1)。

ロジスティック回帰モデルのルーツは1948年にアメリカのフラミンガムで開始された冠状動脈性疾患に関するコホート研究である。それまで病因の概念としては、疾病の原因としての「病原菌」、

結果としての疾病の「発症」という1:1の因果関係が定着していた(註2)。しかし、疾病の中には「病原菌」ではなく、喫煙、飲酒、高カロリー食等、種々のリスクファクターによって発症するものが少なくない。そこで医学分野で用いられるロジスティックモデルでは、異なるリスクファクターをその程度の大小も組み合わせて、発生状況を確率的な現象としてとらえようとする。それが医学分野で盛んに行われているケースコントロールである(註3)。

一方、選択型コンジョイント分析は数理心理学の理論から派生したコンジョイント分析を消費者選好の解析に応用したもので、今日では交通経済学や環境経済学にも応用されている。その計算方法として条件付きロジットが適用されるように、選択型コンジョイント分析はコンジョイント分析をロジスティック回帰モデルの理論的枠組みに取り込んだものである(註4)。どちらも、モデルの構成要因による発生確率や選択確率を明らかにしようとすることに変わりはないが、ケースコントロールではモデルの構築過程からモデルを構成する要因を統計学的にふるいにかけて発生要因を明確化しようとするのに対して、選択型コンジョイント分析はモデルの構築そのものが主要な目的ではなく、モデルを構成する各要因の影響力とその組み合わせによる限界支払意志額や需要関数を計測することを目的としている。

コンジョイント分析の他にも消費者行動、とりわけその評価要因を明らかにすることを目的としてロジスティック回帰モデルが適用された事例は少なくない。しかし、農業関連の分野においてケースコントロールのように各種の検定やデビアンスを適用してモデルを決定しオッズ比(効果の比)を明確に示した適用事例は希である(註5)。そこで、本稿ではケースコントロール研究の手順を湯葉の食味試験に適用し、その最適モデルの構築過程から評価要因の解析を試みる。また、その前段階として近年注目されつつある評価手法の一つであるCS(Customer Satisfaction)分析(註6)の結果を予備的に検討する。その結果と対照することでロジスティック回帰モデルの特質を明らかにし、食味試験への適応可能性を考察する。

本稿の構成は次の通りである。まず、アンケートの概要と分析対象となった大豆の新品種「すずさやか」について触れ(第2節)、CS分析による分析過程とその結果を提示する(第3節)。次いでロジスティック回帰モデルの概要(第4節1)と計量モデルを示し(第4節2)、食味試験への適用方法と分析の手順を明らかにし(第4節3)モデルについての各種の統計的検定方法を明らかにし、分析結果の検討を行う(第4節4・5)。その上で、湯葉の食味評価における最適モデルの決定過程を示し(第5節)課題の検討を行う。

2. 食味試験の概要とデータ

調査は2007年5月に、東北農業研究センター(以下「東北農研」)が募集した消費者モニターや協力者(盛岡市の20代~60代の主婦・加工業者・業務関係者)50名(20代6名、30代10名、40代13名、50代11名、60代10名)にサンプルを提供し、湯葉を対象とした食味試験の形式で実施した。調査の対象とした湯葉は東北農研が開発した大豆「すずさやか」で、コントロールは一般的に流通している大豆の「リュウホウ」である。「すずさやか」は青臭みの原因であるリポキシゲナーゼを欠失させた大豆(以下「リポ欠大豆」)であり、「リュウホウ」はリポ欠大豆ではないが食味、収量とも良く、広く利用されている。「すずさやか」は青臭みのなさがセールスポイントであるが、他の大豆品種と同様に採算性に問題を抱え普及が進んでいない。そのような状況の下で、近年東北地方においても業務用需要が拡大しつつあり、農家の生産意欲がより高い高級食材である湯葉への利用が期待されている。そこで、湯葉に対する加工適性とその食味に対する評価を得ることを目的に食味試験を実施した。試験に使用した湯葉は地元で広く業務用湯葉を手がけ、品種によって異なる湯葉をくみ上げる際の「呉」の濃度を適切に調整できる高い技術を持っている豆腐店に加工を依頼した。食味試験の回答項目は、大豆育種研究室が「すずさやか」の育種課程で実施した食味試験も参考にして①見た目、②食感、③風味、④青臭み、⑤えぐみ、⑥コク、⑦甘み、⑧総合評価の8項目を設

図1 回答結果(すずさやか)の相加平均

定した。

被験者には双方の湯葉を食べ比べてもらい、これらの8項目に対して、「リュウホウ」の湯葉を基準にして「すずさやか」の湯葉について答えてもらった。回答形式は「良い」5点、「やや良い」4点、「同じ」3点、「やや悪い」2点、「悪い」1点の5段階評価とした(「青臭み」「えぐみ」は「無い」5点～「ある」1点)、回答結果の平均をグラフにしたものが図1である。すべての項目で「すずさやか」が3点を上回っており、とりわけ「青臭み」「コク」「甘み」の3要因が「リュウホウ」よりも評価が高いという結果となった。この結果に、まずCS分析を適用してみることにしよう。(データの概要は付属データを参照)

3. CS分析による予備的検討

1) CS分析

CS (Customer Satisfaction) 分析とは、自社の製品・サービスの各種評価要因に対する顧客の重視度と満足度合いを基に顧客満足度を計測する手法で、製品やサービスの改善策を検討する際によく使われる。近年では民間企業のみならず公共団体においても多用されるようになってきている。その結果はCSポートフォリオや決定木分析、また改善度によって現場にフィードバックされる。すなわち、単純集計の結果を個々の要因ごとの重要度と満足度という指標に変換し、それをCSポートフォリオや決定木で視覚化し、さらに数量化された改善度で明示しようという点で分かりやすいと

表1 各項目の満足度・重要度と偏差値

項目	満足度	重要度	満足度 偏差値	重要度 偏差値
見た目	30.0	0.1807	35.51	47.98
食感	36.0	0.1569	40.26	46.40
風味	40.0	0.0785	43.43	41.19
青臭み	62.0	0.1273	60.87	44.43
えぐみ	56.0	0.0911	56.11	42.03
コク	56.0	0.4695	56.11	67.15
甘み	58.0	0.3740	57.70	60.81
平均	48.3	0.2111	50.00	50.00
標準偏差	12.619	0.1507		

されている(註7)。

CS分析は、次のように「満足度」を求めることから始まる。CS分析に用いるデータは回答データの「良い」と「やや良い」を足した累積回答数の全回答数に占める割合(%)を「満足度」とする。次に、「重要度」を求める。被験者50人の各評価項目(①見た目～⑦甘み)の回答データ(5点～1点)と⑧総合評価のデータ(5点～1点)の単相関係数を求め、これを「重要度」とする。これは、総合評価と相関の高い項目ほど総合評価に影響を与えている、すなわち重要と考えられるからである。

$$\text{満足度偏差値} = 10 \times \frac{\text{満足度} - \text{平均}}{\text{標準偏差}} + 50 \quad (1)$$

(重要度偏差値も同様)

以上の「満足度」と「重要度」の偏差値を数式(1)のように求めて、「満足度偏差値」と「重要度偏差値」とする。本項で用いたデータの「満足度」、「重要度」、また、各偏差値は表1に示しておく。

通常、求められた「満足度偏差値」と「重要度偏差値」を平面上にプロットして「CSポートフォリオ」が作成される。図2が「すずさやか」の湯葉の「CSポートフォリオ」である。

図2は、各項目の「満足度」と「重要度」を(50, 50)を原点として2次元上に表現したものであるから、「コク」と「甘み」は重要度も満足度も高く、「青臭み」と「えぐみ」は重要度は高くないが満足度は高く、「風味」「見た目」「食感」は重要度も満足度もそれほど高くないことが読み取れる。し

図2 CSポートフォリオ

たがって、この「CSポートフォリオ」の結果から直観的に判断すれば、「コク」「甘み」の2つの評価要因が総合評価に大きく影響を与えているように思われる。

2) 改善度とその計算結果

CS分析では、この満足度偏差値と重要度偏差値から各評価要因の改善度を求めていく。ここでは、「コク」を例に簡単に説明しておく。まず、図3のように「CSポートフォリオ」の原点(50、50)と「コク」の座標(56.11、67.15)との距離を求める。次に求められた距離18.21の直線と横軸の作る角度 θ を求める(註8)。

この角度 θ (19.58°)と「CSポートフォリオ」の原点(50、50)と(80、20)を結ぶ基準線(座標平面第4象限45°の直線)が作る角度 θ' から修正指数が計算される。この修正指数と先ほど求めた「コク」の座標と原点の距離を掛け合わせた5.13が改善度になる。改善度はマイナスで絶対値が大きいほど改善の必要がなく、プラスで絶対値が大きいほど改善の必要ありとされる(註9)。

以上が「CS分析」のあらましである。表2に各評価要因の修正指数と改善度を示しておく。表2にみるように改善度の+ (改善の必要あり)の要因は、大きな順に①見た目(6.02)、⑥コク

表2 各項目の改善度

	基準線から	原点との		
	の角度	修正指数	距離	改善度
見た目	52.94	0.4118	14.6	6.02
食感	65.29	0.2746	10.4	2.85
風味	98.29	-0.0921	11.0	-1.01
青臭み	162.12	-0.8013	12.2	-9.78
えぐみ	172.50	-0.9167	10.0	-9.21
コク	64.61	0.2820	18.2	5.13
甘み	80.46	0.1060	13.3	1.41

(5.13)、②食感(2.85)、⑦甘み(1.41)、改善度の- (改善の必要なし)を示した要因は④青臭み(-9.78)、⑤えぐみ(-9.21)、③風味(-1.01)、となり、CSポートフォリオの結果から直観的にはもっとも重要度が高いものの満足度も高いため改善の必要がないと思われた⑥コクと⑦甘み、また重要度低さの割には満足度が高いため改善の必要がないと考えられた①見た目に+の大きな改善度が示された。

これは計算方法を見てわかるように「改善度」は満足度と重要度を平面上にプロットした位置情報を基にしており、各評価要因を満足度と重要度との位置関係によって改善の必要性の有無を、その必要性の度合いをポートフォリオの原点からの距離によって表したものである。したがって、CS分析は調査全体における個々の評価要因の位

図3 角度 θ 、 θ' の求め方

置づけを確認することが可能であり、当該商品・サービスの消費者ニーズの現状把握やそれに対応したマーケティング方策の検討に役立てることができよう。しかし、CS分析そのものが1つのモデルを形成しているわけではないため、モデルとしての妥当性や個々の評価要因のウェイトを統計学的に明示することができない点に限界があるといえる。

4. ロジスティック回帰モデル

1) ロジスティックモデルの概要

ロジスティック回帰モデルは既述のように主に疫学研究や生化学の分野をルーツにし、リスクファクターの解明に用いられてきた。一方、我が国では類似した手法として判別分析や数量化理論・重回帰分析が多用されてきた。しかし、多くの場合、説明変数が正規分布をとる量的データであるという仮定のもとで計算が行われる。そのため、実際には理論に従わないデータが多少存在してもやむを得ず解析を行っているというのが現状である（註10）。その点、ロジスティックモデルは従属変数は「ある=1」「なし=0」の二値データを取り、独立変数にはあらゆるデータが適用できる。この「ある=1」「なし=0」の二値データは生死や

発症の有無といったイベントの発生を極めて明快に捉えデータとする。また、計算結果からその因子の独立した影響力（オッズ比）やイベントの発生確率を知ることができる。さらに、その過程におけるモデルのデビアンسや各種の検定による変数・モデルの選択は対数尤度関数を用いることによって統計学的に精緻なものとなっている。このことが医学分野においてロジスティックモデルが盛んに用いられていた理由であろう。本稿では以上のようなロジスティックモデルを湯葉の食味試験へ適用し結果を検討する。

すなわち、「総合評価において選択（評価）された=1」、「総合評価において選択（評価）されなかった=0」として有意な評価要因の抽出とモデルの構築を行い、CS分析の結果と照らし合わせて適用可能性を検証しようというものである。

2) 計量モデル

二値データの分析に一般的に用いられるのは次のような回帰分析を応用したモデルである。いま Y_i が0もしくは1をとる仮想的因子である Y_i^* が

$$Y_i^* = \beta_0 + \beta_i X_i + \varepsilon_i \quad (2)$$

で表すことができるとき、 Y_i^* は直接に値を観察

できない連続的な変数であるが、 X_i によって体系的に決まる部分である $\beta_0 + \beta_1 X_i$ と観察不能な ε_i （誤差項）の和で表すことができる。この ε_i （誤差項）は既知の確率分布に従う確率変数である。ロジスティック回帰モデルは、誤差項の確率分布がロジスティック分布であるとして、その累積分布関数

$$\Lambda(z) = \frac{e^z}{1+e^z} \quad (3)$$

を用い、 $Y_i = 1$ となる確率を

$$F^*(X_i) = \Lambda(\beta_0 + \beta_1 X_i) \quad (4)$$

とするものである。このとき、

$$P_{0i} = P(Y_i = 0 | X_i), P_{1i} = P(Y_i = 1 | X_i) \quad (5)$$

とすれば

$$P_{0i} = 1 - P_{1i} = 1 / [1 + \exp(\beta_0 + \beta_1 X_i)] \quad (6)$$

となり

$$\log(P_{1i}/P_{0i}) = \beta_0 + \beta_1 X_i \quad (7)$$

と表すことができる。すなわち、ロジスティック回帰モデルは Y_i が0となるか1となるかという2つの確率の比（odds）の対数（対数odds）が X_i の線形関数であるとしたものである。モデルの推定方法は今日では次のような最尤法が使われる。

$F(z)$ をロジスティック分布の累積分布関数 $\Lambda(z)$ であるとした時 $P(Y_i = 1 | X_i) = F(\beta_0 + \beta_1 X_i)$ および、 $P(Y_i = 0 | X_i) = 1 - F(\beta_0 + \beta_1 X_i)$ なので $Y_i = 1$ の時 $F(\beta_0 + \beta_1 X_i)$ を、 $Y_i = 0$ の時 $1 - F(\beta_0 + \beta_1 X_i)$ より、ある標本 $Y_1 Y_2 \cdots Y_n$ を得る同時確率は β_0 と β_1 の関数と考えれば次のような尤度関数が得られる。

$$L(\beta_0, \beta_1) = \prod_{y=1} (\beta_0 + \beta_1 X_i) \cdot \prod_{y=0} [1 - F(\beta_0 + \beta_1 X_i)] \quad (8)$$

$L(\beta_0, \beta_1)$ を最大にする $\hat{\beta}_0, \hat{\beta}_1$ はロジット最尤推

定量（logit maximum likelihood estimator）と呼ばれる。

（8）は積の形なので数学的に扱いやすくするために両辺の対数をとると

$$\ln L(\beta_0, \beta_1) = \sum_{i=1}^n \{Y_i \cdot \ln F(\beta_0 + \beta_1 X_i) + (1 - Y_i) \cdot \ln [1 - F(\beta_0 + \beta_1 X_i)]\} \quad (9)$$

モデルを推計する手段としては、通常SASやTSP、SPSS等が利用されるが、変数増加法や変数減少法等のオプションを用いるのではなく単に偏回帰係数や最尤推定量を求めるのであれば、エクセルのソルバー機能によって簡単に計算できる（註11）。

3）データと分析の手順

分析に用いたデータはCS分析に用いたデータを0と1の二値データに変換したものである。説明変数となる各評価項目（①見た目～⑦甘みと従属変数となる総合評価⑧の各回答データの「良い」と「やや良い」の回答を1、「同じ」「やや悪い」「悪い」の回答を0に変換した（註12）。本稿の分析の手順は、次の通りである。まず、総合評価で選択された+群を1、選択されなかった-群を0として従属変数とし、各評価項目（①見た目～⑦甘み）を説明変数 $X_n (n=1 \sim 7)$ とした $\ln\{P/(1-P)\} = \beta_0 + \beta_1 \cdot X_n (n=1 \sim 7)$ という X が0、1の2群のロジスティックモデルにより単変量モデルの推計を行う。

次に $\ln\{P/(1-P)\} = \beta_0 + \beta_n \cdot X_n (n=1 \sim 7)$ のすべての説明変数を組み込んだフルモデルの推計を行う。求められたパラメータからその対数オッズ比と信頼区間を計算する。この時の単変量モデルのオッズ比が「調整しないオッズ比」、フルモデルのオッズ比が「調整したオッズ比」となる。この2つのオッズ比に大きな差があれば「交絡」（見かけ上の相関）が疑われ、信頼区間の結果と併せて変数の検討が行われる（註13）。以上のようにして交絡の有無を確認した後、変数・因子をふるいにかけて最適なモデルを採用することになる。交絡因子が存在する場合は、交絡のリスクを評価した上でモデルへの組み込みを検討する。変

数の絞り込みには変数増加法や変数減少法などを適用し最適モデルを決定することになる。現在はSASやSPSS等にオプション機能として変数増加法や変数減少法を選択して利用できる環境が整っている。その結果採用されたモデルに対して最終的にモデルの帰無仮説の検定を係数のオムニバス検定を適用し、モデルの採否を確認する（註14）。

4) 対数オッズとモデルの検定

まず、対数オッズの意味するところを「見た目」の単変量モデルを例に見ていくことにする。数式(8)から「見た目」の単変量モデルは

$$\log(P_{1looks} / P_{0looks}) = \beta_0 + \beta_1 Xlooks$$

両辺の指数をとると $P_{1looks} / P_{0looks} = \exp[\beta_0 + \beta_1 Xlooks]$

したがって $Xlooks = 1$ の時のオッズは $\exp[\beta_0 + \beta_1]$

$Xlooks = 0$ の時のオッズは $\exp[\beta_0]$

よって「見た目」が良いと選択された時 ($Xlooks = 1$) と選択されなかった時 ($Xlooks = 0$) のオッズ比は

$$\begin{aligned} \text{オッズ比} &= \frac{\exp[\beta_0 + \beta_1]}{\exp[\beta_0]} \\ &= \exp[\beta_0 + \beta_1 - \beta_0] \\ &= \exp(\beta_1) \end{aligned}$$

すなわち、求められた偏回帰係数の指数をとったものがオッズ比となる。「見た目」の単変量モデルは $P_{1looks} = P_{0looks} = \exp[0.1719 + 0.8398 Xlooks]$ と求められるのでオッズ比 $= e^{0.8398} = 2.3158$ したがって、「見た目」が良いと評価された湯葉は、評価されなかった湯葉の2.3158倍総合評価に及ばず影響が強くなることを意味している。

次にこのオッズ比の意味について説明することにしよう。「見た目」が良い ($Xlooks = 1$) と選択された時の「総合評価」が良いとされる確率は (Y) は0.7334、「見た目」が良いと選択されなかった時の確率 (Y) は0.5429となる。この時の確率の差 (リスク差) は0.1905、確率の比 (リスク比) は1.3509に過ぎない。一方、オッズ比2.359は「見た目」が良いと評価された時と「見た目」が良い

と評価されなかった時の効果の差の大きさを表しており、検証すべきリスク要因（ここでは評価要因）そのものの効果の大きさを知ることができる。

次に行う作業は求められたオッズ比の信頼区間を示すことである。自然科学の分野では95%信頼区間を求めるが、ここでは一般消費者や業務関係者等の行った食味試験であることと、サンプル数が50と少ないことから90%信頼区間によるものとする。信頼区間は偏回帰係数と標準誤差 (SE) から簡単に求められる。偏回帰係数 β_1 の $(100 - \alpha)\%$ の信頼区間は正規分布の上側 $\alpha/2\%$ 点 ($Z_{\alpha/2}$) を用いて β_1 の推定値は $\pm Z_{\alpha/2} \cdot \text{SE}$ と表すことができる。したがって、 $\alpha = 10\%$ (信頼区間90%) とすれば、 $Z_{0.1/2} = 1.645$ なので「見た目」の単変量モデルの $\beta_1 = 0.8398$ の標準誤差SEの0.6753から

$$\begin{aligned} 90\% \text{信頼区間} &= 0.8398 \pm 1.645 \cdot 0.6753 \\ &= 0.8398 \pm 1.1109 \\ &= -0.2711, 1.9506 \end{aligned}$$

よって、オッズ比の信頼区間は $e^{-0.2711} = 0.7625 \sim e^{1.9506} = 7.033$ となる。ここで、求められたオッズ比の90%信頼区間0.7625~7.033にはその間に1を含んでいる。これはオッズ比が $e^0 = 1$ の可能性、言い換えれば $\beta_1 = 0$ である可能性がある。このことは10%水準で $\beta_1 = 0$ の検定 (β_1 は意味を持たない) が棄却できないことと対応している。したがって、 $\beta_1 = 0$ の可能性があり、「見た目」の単変量モデルの β_1 は10%水準で採択できないことになる（註15）。

次にモデル係数のオムニバス検定の適用例を説明する（註16）。これは尤度比検定とも呼ばれ、モデルの採用の可否の判定（帰無仮説の検定）に用いる。方法は定数項だけのモデルの対数尤度を $\ln L_0$ 、すべての説明変数を含むモデルの対数尤度を $\ln L_1$ とする時、 $T = -2(\ln L_0 - \ln L_1)$ は制約式に含まれる説明変数の数を q とすると、漸近的に自由度 q の χ^2 分布 $\chi^2(q)$ に従うことを利用している。「見た目」と「食感」の2変量モデルの場合の例を見てみよう。この場合の帰無仮説は、「見た目」と「食感」は「湯葉の総合評価」に影響しないことになり

表3 調整しないオッズ比 (A: 単変量解析) の90%信頼区間

変数	推定値	標準誤差	χ^2	p値	オッズ比	オッズ比の90%信頼区間
見た目	0.8398	0.6753	1.5463	0.2137	2.3158	0.7626 ~ 7.0324
食感	0.8303	0.6344	1.7134	0.1906	2.2941	0.8081 ~ 6.5129
風味	1.5198	0.6681	5.1739	0.0229 **	4.5713	1.5232 ~ 13.7194
青臭み	0.8473	0.5989	2.0012	0.1572	2.3333	0.8712 ~ 6.2493
えぐみ	1.0986	0.5986	3.3682	0.0665 *	3.0000	1.1207 ~ 8.0304
コク	1.4663	0.6152	5.6805	0.0172 **	4.3333	1.5752 ~ 11.9210
甘み	1.6306	0.6247	6.8138	0.0090 ***	5.1071	1.8278 ~ 14.2700

*: $p < 0.1$ **: $p < 0.05$ ***: $p < 0.01$

表4 調整したオッズ比 (B) とその90%信頼区間

変数	推定値	標準誤差	χ^2	p値	オッズ比	オッズ比の90%信頼区間
見た目	1.5686	0.9840	2.5414	0.1109	4.8003	0.9514 ~ 24.2211
食感	0.7080	0.9290	0.5808	0.4460	2.0301	0.4404 ~ 9.3575
風味	2.6321	1.0435	6.3623	0.0117 **	13.9029	2.4985 ~ 77.3630
青臭み	0.7817	0.8155	0.9188	0.3378	2.1853	0.5714 ~ 8.3581
えぐみ	1.0627	0.7974	1.7759	0.1826	2.8943	0.7796 ~ 10.7452
コク	0.3425	1.1000	0.0969	0.7555	1.4085	0.2306 ~ 8.6009
甘み	1.6674	1.2138	1.8872	0.1695	5.2988	0.7196 ~ 39.0182
定数	-3.2434	1.1762	7.6029	0.0058 ***	0.0390	

*: $p < 0.1$ **: $p < 0.05$ ***: $p < 0.01$

帰無仮説 $H_0: \beta_1 = 0$ かつ $\beta_2 = 0$
 対立仮説 $H_1: \beta_1 \neq 0$ または $\beta_2 \neq 0$

のようになる。

定数項だけのモデルを計算すると $\beta_0 = 0.4055$ と求められ、その対数尤度は $\ln L_0 = -33.6506$ となる (註17)。2つの説明変数を含むモデルの対数尤度は $\ln L_1 = -320.2426$ である。よって

$$T = -2(\ln L_0 - \ln L_1) \\ = 2.8159$$

となり、自由度2の χ^2 分布から、その有意確率は0.2446と求められ、モデル係数のオムニバス検定により帰無仮説は棄却できず、モデルは採用されない。

5) 分析結果

表3に①見た目、②食感、③風味、④青臭み、⑤えぐみ、⑥コク、⑦甘みの各変数ごとの単変量モデルの解析結果 (調整しないオッズ比)、表4に①見た目～⑦甘みのすべての変数を組み込んだフルモデルの解析結果 (調整したオッズ比) を示す。単変量モデルとフルモデルのオッズ比の比較

から、オッズ比に大きな差を生じない「食感」、「青臭み」、「えぐみ」、「甘み」については「すずさやか」を選択するに際して「見かけの相関」(交絡)を生じさせる因子ではないと思われる。一方、「見た目」、「風味」、「コク」は単変量モデルとフルモデルのオッズ比の差が大きく、「見た目」では「調整したオッズ比」4.8003に対して「調整しないオッズ比」は2.3158、同様に「風味」では13.9029に対して4.5713、「コク」では1.4085に対して4.3333である。このような場合、他の因子と「見かけの相関」(交絡)を起こしている可能性が否定できない。交絡の条件は、リスク因子 (例えば薬物投与) と結果 (反応) との関連をリスク因子と結果の双方に関連 (影響を及ぼす) する交絡因子 (Confounding factor) という第3の因子の存在が認められることである。ここでの具体的な「見かけの相関」の因子は、「見た目」と「食感」(見た目が良ければ「すずさやか」の評価が高まる+食感も良いと期待させる、食感の良さは「すずさやか」の評価を高める)、「風味」と「コク」(風味が良ければ「すずさやか」の評価が高まる+コクもあると感じさせる、コクの良さは「すずさやか」の評価を高める)、「コク」と「甘み」(豆腐関係の食味試験において「コク」と「甘み」はほぼ同じ因子と考えら

表5 変数減少法によるステップと最終採択モデル

	変数	推定値	標準誤差	χ^2	p値	オッズ比	オッズ比の90%信頼区間	
ステップ1	見た目	1.5687	0.9840	2.5415	0.1109	4.8003	0.9514	～ 24.2211
	食感	0.7081	0.9290	0.5809	0.4460	2.0301	0.4404	～ 9.3575
	風味	2.6321	1.0435	6.3623	0.0117	13.9029	2.4985	～ 77.3630
	青臭み	0.7818	0.8156	0.9188	0.3378	2.1853	0.5714	～ 8.3581
	えぐみ	1.0627	0.7975	1.7760	0.1826	2.8943	0.7796	～ 10.7452
	コク	0.3425	1.1000	0.0969	0.7555	1.4085	0.2306	～ 8.6009
	甘み	1.6675	1.2138	1.8872	0.1695	5.2988	0.7196	～ 39.0182
ステップ2	定数	-3.2434	1.1763	7.6029	0.0058	0.0390		
	見た目	1.5731	0.9911	2.5192	0.1125	4.8216	0.9445	～ 24.6150
	食感	0.6728	0.9175	0.5377	0.4634	1.9597	0.4333	～ 8.8633
	風味	2.6551	1.0346	6.5866	0.0103	14.2268	2.5946	～ 78.0086
	青臭み	0.8025	0.8099	0.9817	0.3218	2.2310	0.5888	～ 8.4541
	えぐみ	1.0809	0.7935	1.8556	0.1731	2.9472	0.7991	～ 10.8696
	甘み	1.9305	0.8897	4.7077	0.0300	6.8929	1.5952	～ 29.7845
ステップ3	定数	-3.2165	1.1656	7.6157	0.0058	0.0401		
	見た目	1.7624	0.9716	3.2901	0.0697	5.8266	1.1785	～ 28.8077
	風味	2.5717	1.0013	6.5959	0.0102	13.0878	2.5209	～ 67.9473
	青臭み	0.5923	0.7404	0.6398	0.4238	1.8081	0.5349	～ 6.1113
	えぐみ	1.0791	0.7852	1.8889	0.1693	2.9420	0.8087	～ 10.7033
	甘み	2.0923	0.8772	5.6896	0.0171	8.1033	1.9145	～ 34.2976
ステップ4	定数	-2.9767	1.0834	7.5496	0.0060	0.0510		
	見た目	1.7084	0.9604	3.1642	0.0753	5.5200	1.1373	～ 26.7915
	風味	2.5968	0.9929	6.8407	0.0089	13.4213	2.6214	～ 68.7155
	えぐみ	1.0871	0.7845	1.9203	0.1658	2.9655	0.8161	～ 10.7766
	甘み	2.1405	0.8750	5.9844	0.0144	8.5037	2.0163	～ 35.8640
ステップ5	定数	-2.6467	0.9663	7.5014	0.0062	0.0709		
	見た目	1.8879	0.9652	3.8260	0.0505	6.6056	1.3503	～ 32.3139
	風味	2.3064	0.9118	6.3982	0.0114	10.0385	2.2403	～ 44.9809
	甘み	2.3364	0.8521	7.5180	0.0061	10.3436	2.5467	～ 42.0113
	定数	-2.1494	0.8508	6.3822	0.0115	0.1165		

* : p<0.1 ** : p<0.05 *** : p<0.01

れていることから相関関係は強いと考えられる)のよう
に考えられる。以上のように「見かけの相関」
(交絡)の有無を確認したら、統計学的により良い
モデルを構築するために、すべての変数の組み合
わせの中から最適なモデルを探すことになる。こ
の場合の組み合わせの数 $\gamma C_n(n=0\sim 7)=128$
となる。本稿で採用した変数減少法(backward法)
の計算過程と結果は表5に示した通りである(註
18)。通常、p値の大きなものから順番に変数か
ら除外していく。第1ステップは7変数すべてを
組み込んだフルモデルで、その中で最もp値の大
きな「コク」を除外して第2ステップに進む。次
いで第2ステップで最もp値の大きな「食感」を
除外し、同様に第3ステップでは「青臭み」、第
4ステップでは「えぐみ」を除外し、最終第5ス
テップで「見た目」「風味」「甘み」の3変数からな
るモデルに決定する。有意確率は「見た目」10%、

「風味」5%、「甘み」1%で有意となり、そのオ
ッズ比もオッズ比の信頼区間の中に収まっており、
求められたオッズ比の90%信頼区間に1を含んで
いる変数はない。また、モデル係数のオムニバス
検定によってモデルの適合度を確認してみると、
定数項だけの対数尤度は $\ln L_0=-33.6506$ 、3変
数のモデルの対数尤度は $\ln L_1=-24.412$ となり、
その検定統計量のTは

$$T = -2(\ln L_0 - \ln L_1) \\ = 18.4772$$

自由度3の χ^2 分布から、その有意確率は
0.00035となりモデルは採用される。この結果から、
「甘い」と評価された場合、1%の有意水準で甘
みが評価されなかった場合の10.34倍、「見た目」
は10%と「風味」は5%に有意確率の精度を落と

すが、それぞれ6.6倍、10倍オッズ比が高くなる
ことが分かる。

このように、湯葉の食味試験でその総合評価に
強い影響を及ぼしていると統計学的に言えるのは
7変数の中では「見た目」と「風味」「甘み」の3
変数となった。「見た目」に関しては加工業者の
技術的な力量に左右されることから、加工業者へ
の貴重な情報と言えよう。「すずさやか」のセー
ルスポイントである「青臭み」のなさが有意な要
因とならなかった理由としては、大豆育種分野の
研究においてリポキシゲナーゼは加熱処理によっ
て「青臭み」の発生が抑えられることが指摘され
ていることから、「リュウホウ」の「青臭み」が
湯葉を揚げる前的大豆を加熱し呉を作る段階でか
なり減少していたと考えられ（註19）、そのため
リポ欠大豆の「すずさやか」と比較すれば「青臭
み」はあるものの、総合評価に大きな影響を及ぼ
すまでには至らなかったものと推測される。した
がって、湯葉に加工することによって「すずさや
か」のセールスポイントであるリポ欠の特徴は総
合評価に反映されることはなかったが、一般に広
く使われている「リュウホウ」に比べ「風味」と
「甘み」において極めて高い評価が得られたことは、
「すずさやか」湯葉加工適正が優れていることの
証左となり、今後生産を普及させる上で貴重な知
見として活用できよう。

5. まとめ

本研究ではリポ欠大豆「すずさやか」で作った
湯葉の食味試験のデータに、CS分析とロジステ
ィック回帰モデルを適用した。CS分析では、「CS
ポートフォリオ」から、「コク」と「甘さ」は重
要度も満足度も高く、「青臭み」と「えぐみ」は
重要度は高くないが満足度は高く、「風味」「見た
目」「食感」は重要度も満足度もそれほど高くない
ことが理解できた。その総合的な結果を表す「改
善度」では「見た目」「食感」「コク」「甘み」におい
て「改善の余地有り」と判定された。しかし、そ
れは被験者の回答データのへ平面上における位置
関係から各評価項目を可視化・数値化したもので
しかなく、要因ごとの総合評価への影響力や要因

間の関連については解析不能であった。

一方、ロジスティック回帰モデルでは単変量モ
デルとフルモデルの統計的解析の結果、「見た目」
と「風味」「コク」に「見かけの相関」が生じてい
ることが示唆された。単純集計では「甘み」と同
様の評価であった「コク」が変数減少法によりモ
デルから除外された。このことは、加工業者等大
豆製品に関わるプロの間では経験的に「甘み＝コ
ク」という認識が一般的であることから納得で
きる。また比較的影響力の大きいと思われた他の
変数も採用モデルに取り入れられることはなかつ
た。最終的にオッズ比やその信頼区間の検定、ま
たモデル係数のオムニバス検定により統計的に有
意な変数は「見た目」と「風味」「甘み」であるこ
とが判明した。このようにロジスティック回帰モ
デルでは、統計学的な検定やモデルのデビアン
スによって変数の選択と有意なモデルの構築が可
能となり「コク」のような必要性のない変数を除
外することができる。またオッズ比によってその影
響力を定量的に示すことができることも、情報の
提示の仕方として明快である。このように、食味
試験のようなアンケート分析にロジスティック回
帰モデルを適用することでより精緻な分析が可能
となることを示すことができた。今後さらに多
方面においてロジスティック回帰モデルの形成過
程を通じて、有意な要因を探索していく同様の研
究の蓄積が期待される。

（註1）農業分野における「選択型コンジョイント分
析」を用いた研究には〔1〕〔2〕〔3〕〔4〕の他多
数あり、近年は〔5〕のように環境評価手法とし
ても適用されている。ロジスティック回帰分析の
概要については〔6〕を参照のこと。

（註2）ロジスティック回帰モデルの概要については
〔7〕並びに〔8〕参照。

（註3）ケースコントロールはイベントの発生（例え
ば薬物投与による「死亡」）した集団case（ケース）
と発生しなかった集団control（コントロール）
について比較を行い、薬物投与とイベントの発生
（死亡）との関連を調べるものである。詳しくは
〔9〕第5章及び〔10〕〔11〕を参照されたい。

（註4）コンジョイント分析については片平〔6〕が
6.2において簡潔に、また〔5〕では環境評価手

法としての適用環境も述べられている。

(註5) 一般的な(選択型コンジョイント分析ではない)「ロジスティック回帰分析」を用いた研究には農業関連では[12][13] 他がある。松本ら[12]はロジットモデルとPOSA (Partial Order Scalogram Analysis) を用いて、生ゴミのリサイクル問題に対する住民意識とその構造を分析している。永木は[13]において酪農農家の自由化の下での意見の形成要因をロジットモデルを用いて解析している。

(註6) CS分析とは、民間企業のマーケティングに盛んに利用されている手法で、製品・サービスに対する顧客の評価要因をその重要度と満足度を基に定量調査するものである。その結果はCSポートフォリオの作成や改善度によってフィードバックされる。CS分析を用いた研究には[14][15]等がある。

(註7) 詳しい手法の適用方法と計算過程は[16]を参照。

(註8) 角度 θ は、図3のように「CSポートフォリオ」の原点(50, 50)と「コク」の座標(56.11, 67.15)から底辺と垂線の長さを求め逆三角関数を利用して次の式により簡単に計算でき、19.61度となる。

$$\text{角度 } \theta = \arctan\left(\frac{\text{垂線}}{\text{底辺}}\right) \cdot \left(-\frac{180}{\pi}\right) \quad (1)$$

「コク」の場合角度 θ は19.61度なので、基準線からの角度 θ' は $19.61 + 45 = 64.61$ のようにして求める。修正指数は以下の数式で求められ、0.282となる。

$$\text{修正指数 } r = \frac{90 - \text{角度 } \theta'}{90} \quad (2)$$

(註9) これは「CSポートフォリオ」の位置情報と角度情報を基準線を修正指数を最大値1とする指標に変換することを意味している。従って図3においては基準線から ± 90 度を超えるの領域が改善度－(改善の必要なし)、 ± 90 度より小さい領域が改善度＋(改善の必要あり)になる。

(註10) 数量化理論の問題点やロジスティック回帰モデルの利点については[17]を参照。

(註11) 合崎[18]はエクセルによる「選択コンジョイント分析」(条件付きロジットモデル)の計算方法を明らかにしている。本稿の場合「条件付きロジット」ではないため、より計算過程は簡潔になる。

(註12) 本研究では「同じ」は「良い」という評価ではないとして、またCS分析における取り扱いに準じてデータに変換した。「良い～同じ」= 1と変換した場合の計算結果は大きく異なってしまうので注意が必要である。

(註13) 交絡については[9]第5章C及び第6章3において浜田が詳しく論じている。交絡とは、得られた測定値の間には相互に相関が高いことがあり、そのことが解析結果に大きく影響を及ぼしてしまうことを意味している。したがって、もし個々の変数が独立の関係ならば、単変量モデルでもフルモデルでもそのオッズ比に大きな差は生じない筈である。この点について浜田は脳卒中におけるリスク因子の解析で、通常女性の方が脳卒中が多いにも関わらず喫煙や高血圧といったリスク因子によって性別と脳卒中の傾向が異なった傾向を示してしまう例を紹介している。そこで単変量モデルだけの解析ではなく、フルモデルの解析も行い、そのオッズ比を比較して「交絡」の有無を確認することが必要とされる。その上で、より良い統計モデルを構築するために変数選択を行う。ただし、本研究のような場合、「交絡は「見かけの相関」、「リスク因子」は「評価要因」等と言い換えた方が適切と思われる。

(註14) ロジスティック回帰モデルの変数やモデルの検定、モデルの採用にいたる計算過程については浜田[9]第6章、対馬[22]を参考にされたい。

(註15) オッズ比並びにオッズ比の信頼区間については、[9] pp.152-154、[7] pp.76-79に詳しい。

(註16) [19] pp.247-249参照のこと。

(註17) SPSSでは尤度比検定の結果のみが出力され、定数項のみの対数尤度は求めることができない。TSPで計算する場合は、エクセルのワークシートに総合評価のデータのみ入力したファイルを入れコマンドラインから(Aドライブにデータ名“YUBA.xls”のファイルがあるとした場合)

```
smpl 1 50;
read(file="a:YUBA.xls")Y;
logit Y C;
```

と入力すれば定数項のみの対数尤度を求めることができる。統計パッケージ等によらないで計算したい場合は[18]のようにエクセルのソルバー機能を用いて計算する方法がある。TSPについては[19][20]を参照。

(註18) 本稿ではSPSSのオプションから変数減少法(backward法)を選択して計算した。

(註19) [21] において「青臭み」=「不快臭」として食品工業的には主として加熱処理で不快臭発生を防止する方法が取られていることを紹介している。

引用・参考文献

- [1] 合崎英男・佐藤和夫・長利洋「選択実験による農業・農村の持つ多面的機能の個別機能評価に関する一提案」『農業経営通信』No.221、2004。
- [2] 澤田学・岩本博幸・山本康貴「消費者のHACCP及び環境対策認証表示牛乳に対する価値評価」『農業の与件変化と対応策』農林統計協会、pp.46-58、2002。
- [3] 大浦裕二他「選択型コンジョイント分析による青果物産地のブランド力の推定」『農業経営研究』第40巻1号、2002。
- [4] 下山禎・飯坂正弘・野中章久・川手督也・渋谷美紀由比進・山崎篤「イチゴに対する消費者ニーズの解明－国内産・超促成栽培イチゴと外国産イチゴを中心に－」『農業経営通信』No.214、2002。
- [5] 鷲田豊明・栗山浩一・竹内憲司編『環境評価ワークショップ 評価手法の現状』第2章、1999。
- [6] 片平秀貴『マーケティング・サイエンス』東京大学出版会、p.155、1994。
- [7] 丹後俊郎・山岡和枝・高木晴良『ロジスティック回帰分析－SASを利用した統計解析の実例－』朝倉書店、1996。
- [8] D. W., Hosmer, and S., Lemeshow, *Applied Logistic Regression*, John Wiley & Sons, 1989.
- [9] 浜田知久馬『学会・論文発表のための統計学』真興交易医書出版部、1999。
- [10] Jordan J. L., D. A. Hensher and J. D. Swait, *Stated Choice Methods: Analysis and Application*, Cambridge University Press, 2000.
- [11] W. O. Spitzer et al. "The use of β -agonists and the risk of death and near death from asthma," *New Eng. J. Med.* 326(8), pp.501-506, 1992.
- [12] 松本和子・霜浦森平・栗原伸一・大江靖雄「生ゴミ分別に対する住民意識と要因分析－栃木県茂木町を対象として－」『千葉大学園芸学部学術報告』Vol.59、pp.69-74、2003。
- [13] 永木正和「ロジット・モデルによる『自由化』に対する酪農経営者の意見形成要因の分析」『北海道農業経済研究』pp.36-47、1992。
- [14] 下山禎・三上理佳「CS（顧客満足度）分析とその適用方法」『農業経営通信』No.222、2004。
- [15] 原温久・熊谷宏「農業用水路に対する地域住民の評価と今後の整備の方向性－CS分析の適用－」『農村生活研究』Vol.51-1、pp.20-27、2007。
- [16] 菅民郎『Excelで学ぶ多変量解析入門』オーム社、2004。
- [17] 高橋善弥太『－医者のための－ロジスティック・COX回帰入門』日本医学館、第2章、1995。
- [18] 合崎英男「表計算ソフトを利用した選択実験の計測手順」『むらづくりテクダス』(独) 農業工学研究所農村計画部、2004。
- [19] 東京大学教養学部統計学教室編『自然科学の統計学』東京大学出版会、1992。
- [20] 和合肇・伴金美『TSPによる経済データの分析』東京大学出版会、1988。
- [21] 喜多村啓介・熊谷享・菊池彰夫「ダイズ種子中のリボキシゲナーゼ-2の遺伝様式、およびリボキシゲナーゼ-1、2、3アイソザイムの遺伝的関係」『育種学雑誌』vol.35、pp.413-420、1985。
- [22] 対馬栄輝『SPSSで学ぶ医療系データ解析』東京図書、2007。

[2008年6月27日受付、2009年3月4日受理]

ロジスティック回帰モデルによる食味評価の分析

付属データ

(食味試験のデータ)

No.	総合 評価	見た目	食感	風味	青臭み	えぐみ	コク	甘さ
1	3	4	3	3	3	4	4	3
2	5	5	2	4	5	5	3	4
3	2	1	4	3	3	4	3	3
4	2	3	3	3	5	3	3	3
5	5	3	4	5	3	4	4	4
6	2	4	3	3	4	4	3	2
7	4	3	5	4	5	5	4	4
.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.
43	4	3	4	2	5	2	4	4
44	5	2	2	4	2	2	4	4
45	4	5	3	2	4	4	5	5
46	3	3	3	3	5	3	4	4
47	4	3	3	4	4	5	5	5
48	4	3	3	2	2	2	3	3
49	5	2	3	3	5	5	5	5
50	5	4	3	5	3	3	4	5

(0,1に変換したデータ)

No.	総合 評価	見た目	食感	風味	青臭み	えぐみ	コク	甘さ
1	0	1	0	0	0	1	1	0
2	1	1	0	1	1	1	0	1
3	0	0	1	0	0	1	0	0
4	0	0	0	0	1	0	0	0
5	1	0	1	1	0	1	1	1
6	0	1	0	0	1	1	0	0
7	1	0	1	1	1	1	1	1
.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.
43	1	0	1	0	1	0	1	1
44	1	0	0	1	0	0	1	1
45	1	1	0	0	1	1	1	1
46	0	0	0	0	1	0	1	1
47	1	0	0	1	1	1	1	1
48	1	0	0	0	0	0	0	0
49	1	0	0	0	1	1	1	1
50	1	1	0	1	0	0	1	1