

審査学位論文

新製品の試用購買時期と究極的普及率の測定

田島博和

1994年3月31日

学習院大学

①

学習院大学審査学位論文
新製品の試用購買時期と究極的普及率の測定

田島博和

1994年3月31日

謝辞

本論文を作成するに当たり、多くの先生方から、直接・間接的に多大なご指導を賜った。

今回の実証分析に用いたデータは、財団法人流通経済研究所 MID 部から頂戴した。また実証分析の一部は、同研究所所有のワークステーションを利用している。そしてスタッフの方々からは、いつも実務・理論の両面からご指導をいただいている。特に中村博主任研究員には、杉田・中村・田島の研究 [72] の際にも、ご指導いただいている。

理学部数学科の大津賀信教授には、大学 4 年から修士課程 2 年までの指導教官として、複素関数論をはじめとする数学の基礎をご指導いただいた。また小山昭雄教授には、博士課程 1,2 年の指導教官として、数学の社会科学への適用についてご指導いただいた。さらに新居玄武教授には、統計学についてご指導を頂戴した。

田島義博教授はマーケティングの科学化の必要性を唱えられ、筆者にマーケティング・サイエンスという学問の存在を教えてくださいました。

杉田善弘教授には、本論文の指導教官として、いつも厳しく、そしてあたたかいご指導をいただいている。

そして上田隆穂教授には、博士課程 3 年から現在までの指導教官として、いつもあたたかいご指導をいただいている。

この場を借りて、以上の方々に厚く御礼申し上げます。

目次

1	はじめに	1
2	耐久時間データとハザード関数	6
2.1	耐久時間データ	6
2.2	ハザード関数	8
2.2.1	定義およびその性質	8
2.2.2	スプリット・ハザード関数	10
2.2.3	ハザード関数を使用することの利点	12
3	既存研究	15
3.1	集計レベルの研究	16
3.1.1	Bass 1969	16
3.1.2	Fourt and Woodlock 1960	20
3.1.3	Horsky and Simon 1983	21
3.1.4	Nakanishi 1973	22
3.2	試用購買する過程を明示した研究	24
3.2.1	Kalish 1985	25
3.2.2	Roberts and Urban 1988	26
3.2.3	Chatterjee and Eliashberg 1990	28
3.3	ハザード関数を用いた非集計レベルの研究	34

3.3.1	Sinha and Chandrashekar 1992	35
3.3.2	杉田・中村・田島 1992	38
4	既存研究の問題点と本研究の目的	42
4.1	既存研究の問題点	42
4.1.1	集計レベルの研究	42
4.1.2	効用の概念を取り込んだ研究	43
4.1.3	ハザード関数を用いた非集計レベルの研究	44
4.2	本研究の目的	47
5	本研究のモデル	49
5.1	モデル	49
5.1.1	期待効用	49
5.1.2	潜在的試用購買者である確率	52
5.1.3	試用購買するタイミング	53
5.1.4	ハザード関数	53
5.1.5	究極的普及率	54
5.2	パラメータの推定方法	54
6	実証分析	56
6.1	データと共変量	56
6.1.1	家計	57
6.1.2	サンプル・スプリッティング	57
6.1.3	共変量	57
6.2	分析計画	61
6.2.1	モデル	61
6.2.2	パラメータの推定方法	64
6.2.3	説明力の検証	64

6.2.4	予測精度の比較	65
6.3	結論	67
6.3.1	パラメータ推定値	67
6.3.2	説明力	70
6.3.3	予測精度	71
6.3.4	まとめ	73
7	まとめと今後の課題	74
A	命題の証明	76
A.1	ハザード関数	76
A.2	杉田・中村・田島のモデル	79
A.3	本研究のモデル	80
B	EM アルゴリズム	86
B.1	原理	87
B.1.1	概略	87
B.1.2	EM 推定量の性質	88
B.1.3	命題の証明	90
B.2	杉田・中村・田島のモデルへの適用	93
B.3	本研究のモデルへの適用	94
C	プログラムおよびアウトプット	96
C.1	モデル 1 (本研究のモデル)	96
C.1.1	プログラム	96
C.1.2	アウトプット	99
C.2	モデル 0 (杉田・中村・田島のモデル)	100
C.2.1	プログラム	100
C.2.2	アウトプット	102

D 原データおよび家計別の対数尤度	105
参考文献	109

表 目 次

2.1	マーケティングにおける耐久時間データの例	7
2.2	密度関数とハザード関数の対応	9
3.1	Nakanishi モデルの推定結果	23
3.2	Chatterjee <i>et al.</i> モデルによる家計の分類	32
3.3	Chatterjee <i>et al.</i> モデルの予測結果	34
3.4	杉田・中村・田島モデルの推定結果	40
4.1	非集計レベルのハザードモデルにおける確率およびタイミングの定式化	45
4.2	非集計レベルのハザードモデルにおける共変量の取り込み方	45
6.1	店舗別試用購買実績	58
6.2	HEAVY 対 LOYAL の頻度表	59
6.3	LOYAL 対 TOKU の頻度表	59
6.4	TOKU 対 HEAVY の頻度表	59
6.5	店舗ダミーの定義	61
6.6	推定結果	68
6.7	説明力	70
6.8	モデル 1 とモデル 0 の比較	73
A.1	記号の凡例	76
B.1	不完全データ	87

目次

3.1 Bass モデルの普及曲線 ($q > r$)	18
3.2 Bass モデルの普及曲線 ($q < r$)	18
3.3 Kalish のモデル	27
3.4 Roberts and Urban のモデル	29
3.5 Sinha and Chandrashekarán のモデル	36
3.6 杉田・中村・田島のモデル	39
5.1 本研究のモデル	50
6.1 予測結果 (新規試用購買率)	72
6.2 予測結果 (累積試用購買率)	72
B.1 EM アルゴリズム	89

Chapter 1

はじめに

近年の景気後退および低価格志向による深刻な消費不況は周知の通りである。それは最寄品¹の市場でも例外ではなく、なかでも新製品の成功率が低下してきている [86]。そのため小売業者は店頭に並べる新製品数を絞り込み、在庫費用をはじめとするコストの削減や商品構成の適正化を行うことによって、業績の回復をはかっている [81, 82, 85]。また POS (Point Of Sales: 販売時点管理) データの普及によって商品別の細かな販売履歴データを低コストかつ迅速に収集できるようになったため、小売業者は、新製品数の絞り込みをはじめとするマーケティング・アクションをできるだけ早期に決定して市場のニーズにすばやく対応するために、新製品の診断情報を今まで以上に早く入手しなければならない [80]。特に大規模の小売業者はそのバイイング・パワーを背景にして、メーカーに対してより低価格で競争力のある商品の開発を要求しており、さらに EOS (Electric Order System: 自動補充発注システム) の導入によって、商品受発注の細分化・迅速化を進めている。そのためメーカーとしても、新製品の開発・生産計画を立てるうえで、診断情報を早期に入手する必要があると思われる [80]。

ところで上で述べた新製品の診断情報として、具体的にはどのようなものが考えられるのだろうか。最寄品の市場には強力な競争相手が多数存在し、規模の経済性が大きくかつ価

¹ 「顧客が頻繁・即座にほとんど努力なしで購買する商品」 [65, p.122] であり、たとえばスーパーマーケットなどの量販店で売られている加工食品や日用雑貨品がこれに該当する。

格に敏感である。従って新製品を導入する際のマーケティング・ミックス戦略としては、市場浸透戦略 (penetration strategy) を採用すべきだと一般に言われている [57, p.412]。すなわち、できるだけ多くの家計に「試し買い」 (trial purchase)²してもらう為のマーケティング・ミックス戦略を採用し、当該新製品の市場浸透³を高めるのである。特に競争力のある新製品の場合は試用購買者の反復購買率が高いという事が、実際の調査からも示されており [91, p.164]、この戦略の有効性を裏付けている。従って新製品の重要な診断情報のひとつとして、マーケティング・ミックスが市場浸透に与える影響を挙げることができよう。また家計の嗜好が多様化する今日では、全てのカテゴリー・ユーザーを対象にするよりも、市場をセグメント化して、当該新製品を試用購買する可能性が高い家計層を対象にしたターゲット・マーケティング戦略の方が効率的であることはいうまでもない [57, p.105]。従って購買のタイプをはじめとするマーケット・セグメンテーションの為の尺度が家計の試用購買に与える影響も、新製品の重要な診断情報のひとつであるといえよう。以上で述べたことを総合すると、マーケット・セグメンテーションのための尺度やマーケティング・ミックスが家計の試用購買に与える影響といった新製品の診断情報を市場投入後できるだけ早期に入手することが、メーカーや小売業者にとって重要であると考えられる。

ところでこの様な新製品の診断情報を早期に入手する試みは、実務からの要請もあって比較的古くから行われている [40]。そのなかでも、Bass の研究 [2] や Fourt and Woodlock の研究 [15] は、初期の代表的な研究であるといえる。そしてこれらのモデルを拡張した研究が、価格を考慮した Dolan and Jeuland の研究 [10] や広告を考慮した Horsky and Simon の研究 [27] をはじめとして、今日に至るまで数多く行われている。これらのモデルのパラメータを推定するために用いられるデータは、時点毎の販売履歴データまたは家計の購買履歴を時点毎に家計に関して集計して得られたデータなので、集計レベル (aggregate level) のモデルと呼ばれている。しかし集計レベルのモデルには次に述べるいくつかの問題点が

² 家計が当該新製品を「試しに」初めて購入することを意味している。なお本論文ではこれ以降「試用購買」と略称することにする。また「(新製品の) 採用」 (adoption) という用語も同じ意味で用いている場合がある [29, 31]。

³ 既に試用購買した家計がカテゴリー・ユーザー、すなわち当該新製品が含まれるカテゴリー (品目) を購買する家計、に占める割合である。なお本論文では市場浸透と市場普及 (diffusion) というふたつの用語を同じ意味で用いている。

ある。ひとつは、用いるデータが家計に関して集計されているため、試用購買する確率やそのタイミングに関する家計の異質性を検出できないことである。そしてもうひとつは、特に新製品を分析する際に、パラメータ推定値が統計的に安定するために十分な量のデータを確保できないという事である。モデルのパラメータ推定に用いるデータは一般に、最長でも市場導入時点から現在までの時点数分しか入手することができない。特に新製品は文字通り市場導入されてからまだ間もない製品なので、データは必然的に少なくなり、パラメータ推定値が統計的に安定しにくくなってしまふのである。

これらの問題点を解決すべく、非集計レベル (disaggregate level) のモデルを用いた研究も数多く行われている。これは家計毎の購買履歴のように家計に関して集計されていないデータを用いてパラメータを推定するモデルであり、試用購買する確率やタイミングに関する家計に異質性を検出することが可能である。また家計を増やすことによって、統計的に安定したパラメータ推定値を得るために十分な量のデータを確保することが可能である。

ところで新製品は一般に、全てのカテゴリー・ユーザーによって試用購買されるとは限らない。すなわちその新製品を認知しても決して試用購買しないカテゴリー・ユーザーが存在する可能性がある。従って各家計の試用購買する確率すなわち潜在的な試用購買者である確率を考慮する事によって、「当該新製品を決して試用購買しない家計」の存在を許したモデルの方が、現実を良く表しているといえよう。また価格やプロモーションをはじめとするマーケティング・ミックスはある期間ごとに変化するのが普通なので、共変量 (covariate)⁴のダイナミックな変化を許したモデルの方が現実を良く表しているのはもちろんである。しかし共変量のダイナミックな変化を試用購買する確率に取り込む (すなわち試用購買する確率が共変量の影響を受けながら時間に関して連続的に変化する) のは、解析的に困難である。したがって、「当該新製品を決して試用購買しない家計」の存在を許したモデルを実際に定式化するには、次のふたつのアプローチがあるといえよう。ひとつめは、共変量はダイナミックに変化し試用購買するタイミングに影響を与えるが、試用購買する確率には影響を与えないモデルである。これに該当する非集計レベルの研究としては、杉田・中村・

⁴ この用語は独立変数や説明変数と同義であるが [73, p.152]、本論文では便宜上、ハザード・モデル (2.2 節) に関する成書 [37] に従い、「共変量」に統一する。

田島の研究 [72] がある。しかし共変量が試用購買する確率に影響を与えないという仮定は制約的であり、さらに共変量がダイナミックに変化するので推定に時間がかかるという問題点がある。そしてもうひとつは、共変量はダイナミックに変化せず、試用購買する確率とタイミングの両方に影響を与えるモデルである。これに該当する非集計レベルの研究としては、試用購買する確率とタイミングに対する共変量の影響が独立に与えられた Sinha and Chandrashekarán の研究 [56] がある。しかし試用購買するタイミングは、試用購買する確率が与えられたという条件下で事後的に与えられると考えるのが自然である。そのため共変量が確率とタイミングのそれぞれに独立に影響すると仮定したこのモデルでは、整合性のとれたインプリケーションが期待できず、さらにパラメータの同時推定が不可能であるという問題が存在する。

そこで本研究では von Neumann and Morgenstern の効用理論 [35] を利用した期待効用を取り込むことによって、試用購買する確率とタイミングを同時に説明する非集計レベルの試用購買モデルを提案する。このモデルはダイナミックに変化しない共変量が試用購買する確率とタイミングを説明しているので、Sinha and Chandrashekarán の研究を踏襲していると考えられる。しかし Sinha and Chandrashekarán の研究では共変量が確率とタイミングを独立に説明しているのに対して、このモデルでは共変量が確率とタイミングを同時に説明しており、整合性のとれたインプリケーションとパラメータの識別性が確保され、さらに推定にかかる時間が短縮される。またこのモデルは、杉田・中村・田島や Sinha and Chandrashekarán の研究と同様に、ハザード関数によって定式化されている。ハザード関数は、「時点 t まで試用購買しないという条件下で、丁度、時点 t に試用購買する瞬間的な確率」を意味しており、医学や信頼性工学の分野では、センサリング（観測の打ち切り）が発生しているデータを分析する際によく用いられる関数である。

2 章ではハザード関数を紹介し、さらにハザード関数によって試用購買モデルを定式化することの利点を説明する。3 章では、新製品の試用購買に関するいくつかの既存研究を紹介する。4 章では、既存研究を整理し問題点を指摘することによって、本研究の目的および意図を明確化する。5 章では、本研究のモデルを詳しく紹介する。6 章では、朝食用シリアルを対象にした実証分析を行なう事によって、モデルの妥当性を検証し、さらに確

率やタイミングに対する共変量の影響や究極的普及率を測定する。そして7章では、本研究の総括を行う。

Chapter 2

耐久時間データとハザード関数

この章では、本研究で用いられるふたつの重要な概念について、あらかじめ解説を行う。ひとつは耐久時間データ (duration time data) である。耐久時間とは、関心の対象となる事象が発生するまでに要する時間であり、これを非負の確率変数とすると、その観測値は耐久時間データと呼ばれる。本研究が対象とする試用購買時期も、耐久時間のひとつの例である。2.1 節では、マーケティングにおける耐久時間の例や、耐久時間データの性質について説明する。

そしてもうひとつはハザード関数 (hazard function) である。ハザード関数は、「時点 t まで試用購買しないという条件下で、丁度、時点 t に試用購買する瞬間的な確率」であり、耐久時間データを分析する際によく用いられる。2.2 節では、その基本的性質や、回帰モデル等と比較した場合の利点について説明する。

2.1 耐久時間データ

耐久時間 (duration time) とは、関心の対象となる事象が発生するまでに要する時間であり、これを非負の確率変数とすると、その観測値は耐久時間データと呼ばれる。

人間が死亡するまでの時間や機械が故障するまでの時間などは、耐久時間データの例であり、それぞれ生存データ (survival data) とか寿命データ (life data) と呼ばれている [46]。

また本研究の対象となる新製品の試用購買時期も、耐久時間のひとつである。そこでマーケティングにおける耐久時間データの例を、表 2.1 にまとめておく。

表 2.1: マーケティングにおける耐久時間データの例

分野	問題	対象となる事象
広告	コピーテスト	想起までの時間
	メディア選択	露出時間
価格政策	価格の変更	反復購買間隔
マーケティング戦略	クーポンの実施	回収時期
新製品開発	成功率の予測	試用購買時期、反復購買間隔
市場調査	パネラーの選択	パネラーの脱落時期
	アンケートの実施	返答時期
流通チャネル	チャネル・デザイン	チャネル関係の持続期間

(Helsen and Schmittlein [26] より一部抜粋)

耐久時間データを分析する場合には、データのセンサリング (censoring) に注意しなければならない。センサリングとは、何らかの事情で、事象が発生したかしないかに関わらずに観測が打ち切られることを意味する。この事を試用購買の例を用いて説明しよう。家計 n がある新製品を試用購買するまでに要する時間を T_n とする。そして当該新製品の市場導入時点と実際の試用購買時点を、それぞれ C_S , C_n (暦上の時間) とすると、 $\hat{T}_n := C_n - C_S$ が T_n の実現値になる。しかし耐久時間データは様々な理由で、分析を行う以前に観測が打ち切られる場合が多い。そこで打ち切り時点を C_R とすると、 C_R 以前に試用購買を行わなかった家計 n については、 $\hat{T}_n > C_R - C_S$ という情報しか得られない。このとき家計 n は、右側センサリング (right censoring) されたと呼ばれる。右側センサリングを含むデータを、購買量データをはじめとする普通の連続データと同じように回帰モデルにかけると、パラメータ推定値にバイアスが発生する。これについては次の節で述べる。また C_S 以降に観測が開始された場合、データは左側センサリング (left censoring) されたという。ただし本研究の実証分析で用いるデータは、当該新製品の市場導入以前から観測を行っているため、左側センサリングは発生しない事に注意したい。

2.2 ハザード関数

家計 n の耐久時間 T_n (の確率分布) を定式化するには、いくつかのアプローチが存在する。ひとつは T_n の期待値を定式化する方法であり、回帰分析がこれに含まれる。回帰分析によって耐久時間を分析した例としては、プロモーションが購買間隔に与える影響を測定した Neslin, Henderson and Quelch の研究 [47] がある。もうひとつは T_n の確率分布を定式化する方法である。特に 3 章で紹介する集計レベルのモデルを用いた Bass の研究 [2] 等では、密度関数 $f(t)$ や分布関数 $F(t)$ ではなく、

$$h(t) = \frac{f(t)}{1 - F(t)} \quad (2.1)$$

を定式化している。ここで $h(t)$ はハザード関数 (hazard function) と呼ばれており、本研究のモデルも、ハザード関数を用いて定式化されている。

そこでこの節では、ハザード関数について解説し、さらにハザード関数を用いて耐久時間データを分析することの利点について述べる。

2.2.1 定義およびその性質

これ以降簡単のため、耐久時間データを新製品の試用購買に限って説明する。

定義

家計 n の試用購買時期を表す非負の確率変数を T_n とする。このとき T_n のハザード関数 $h_n(t)$ は、次のように定義される [30, 7]。

$$h_n(t) := \lim_{\delta \downarrow 0} \frac{\Pr(t \leq T_n < t + \delta | t \leq T_n)}{\delta} \quad (2.2)$$

ここで $h_n(t)$ は直観的には、「時点 t まで試用購買しないという条件下で、丁度、時点 t に試用購買する瞬間的な確率」と考えることができ、医学や信頼性工学の分野では、耐久時間データの分析に用いられている。またマーケティングでは、瞬間的購買生起率 (Instantaneous

Purchase Rate) と呼ばれており [43, p.281]、購買間隔モデル [19, 28, 26] や試用購買モデル [2, 45, 56, 72] で用いられている。

密度関数との対応関係

ところで簡単な計算によって次式が成立し、ハザード関数 $h_n(t)$ と密度関数 $f_n(t)$ が、一対一に対応する事が分かる¹。

$$h_n(t) = \frac{f_n(t)}{1 - F_n(t)} = \frac{f_n(t)}{S_n(t)} \quad (2.3)$$

$$f_n(t) = h_n(t) \exp\left(-\int_0^t h_n(\tau) d\tau\right) = h_n(t) \exp(-H_n(t)) \quad (2.4)$$

$$S_n(t) = \exp(-H_n(t)) \quad (2.5)$$

ただし $S_n(t) := 1 - F_n(t)$, $H_n(t) := \int_0^t h_n(\tau) d\tau$ はそれぞれ、生存関数 (survival function)、累積ハザード関数と呼ばれる関数である。さらに $h_n(t) \geq 0$ ならば、2.4 式によって計算される密度関数 $f_n(t)$ は、確率の公理を満足する。ここで密度関数とハザード関数の対応を、いくつかの分布に関して表 2.2 にまとめておく。

表 2.2: 密度関数とハザード関数の対応

分布	密度関数	ハザード関数
指数	$\lambda e^{-\lambda t}$	λ
Weibull	$k\lambda(\lambda t)^{k-1} e^{-(\lambda t)^k}$	$k\lambda(\lambda t)^{k-1}$
混合指数	$k(k/\lambda)^k / (t + k/\lambda)^{k+1}$	$k/(t + k/\lambda)$
対数ロジスティック	$k\lambda^k t^{k-1} / (1 + (t\lambda)^k)^2$	$kt^{k-1}\lambda^k / (1 + (t\lambda)^k)$

(Cox and Oakes [7, p.17] より一部抜粋)

たとえばハザード関数がダイナミックに変化せずに定数ならば、 T_n は指数分布に従がい、無記憶性 (memoryless property) [77, p.132] を持っている。そして Fourt and Woodlock

¹ 証明は Appendix A の命題 A.1.1 を参照のこと。

[15] の研究は離散的なモデルであるが、浸透戦略が採用された最寄り品の新製品の普及曲線に、この分布を当てはめている。この研究は 3.1.2 項で紹介する。

尤度関数

ハザード関数によって定式化されたモデルのパラメータを、(右側) センサリングを含むデータを用いて推定する方法を考える。いま左側センサリングは発生していないとし、観測開始時点をも $C_S = 0$ として観測の打ち切り時点をも $T_0 := C_R$ とする。このときオブザベーションは $\vec{T} := \{\hat{T}_n\}_{n=1}^N$ (ただしセンサリングされた家計 n については $\hat{T}_n := T_0$ と定義する) であるから、モデルの尤度関数 $L(\theta; \vec{T})$ および対数尤度関数 $\ell(\theta; \vec{T})$ は次のようになる²。

$$\begin{aligned} L(\theta; \vec{T}) &= \prod_{n=1}^N f_n(\hat{T}_n | \theta)^{d_n} S_n(T_0 | \theta)^{1-d_n} \\ &= \prod_{n=1}^N \left(h_n(\hat{T}_n | \theta) e^{-H_n(\hat{T}_n | \theta)} \right)^{d_n} e^{-H_n(T_0 | \theta) (1-d_n)} \end{aligned} \quad (2.6)$$

$$\ell(\theta; \vec{T}) = \sum_{n=1}^N \left\{ d_n \left(\log h_n(\hat{T}_n | \theta) - \int_0^{\hat{T}_n} h_n(t | \theta) dt \right) - (1-d_n) \int_0^{T_0} h_n(t | \theta) dt \right\} \quad (2.7)$$

ただし d_n は、家計 n がセンサリング時点以前に試用購買を行ったか否かを表すダミー変数であり、 $d_n = 0$ なる家計 n はセンサリングされたことを意味する。従って 2.7 式を用いて、パラメータの最尤推定を行うことが可能である。

2.2.2 スプリット・ハザード関数

本研究で提案するモデル、および 3 章で紹介する Sinha and Chandrashekarán や杉田・中村・田島のモデルの究極的普及率は、必ずしも 1 になるとは限らない。すなわち「当該新製品を決して試用購買しない家計」の存在を許している。更にこれらのモデルでは、潜

² 証明は Appendix A の命題 A.1.2 を参照のこと。

在的な試用購買者である確率がダイナミックに変化せず、一定である。特に杉田・中村・田島のモデルや本研究で提案するモデルでは、その確率が 0-1 の Bernoulli 分布で与えられている。

そこでここでは特に、確率が Bernoulli 分布で与えられているモデルを定式化する為に用いられる、スプリット・ハザード関数について説明する。

定義

家計 n が潜在的試用購買者であるか否かを表す確率的ダミー変数を v_n とする。上で述べたとおり v_n はダイナミックに変化しない。そして $v_n = 0$ は、家計 n が当該新製品を決して試用購買しないことを意味している。さらに v_n が家計に関して独立に Bernoulli 分布 $Bi(p_n)$ に従っているとす。

$$\Pr(v_n = v) = p_n^v (1 - p_n)^{1-v} \quad (2.8)$$

このとき家計 n の試用購買時期を表す非負の確率変数 T_n のハザード関数を $\tilde{h}_n(t)$ とすると、 v_n を与えた下での条件付きスプリット・ハザード関数 $h_n(t|v_n)$ は、次のように定義される。

$$h_n(t|v_n) := v_n \tilde{h}_n(t) = \begin{cases} \tilde{h}_n(t), & v_n = 1 \\ 0, & v_n = 0 \end{cases} \quad (2.9)$$

尤度関数

スプリット・ハザード関数を用いて定式化したモデルの尤度関数および対数尤度関数は、次のようになる。

$$L(\theta; \vec{T}) = \prod_{n=1}^N \left(\Pr(v_n = 1) f_n(\hat{T}_n | v_n = 1) \right)^{d_n} \left(\Pr(v_n = 1) S_n(T_0 | v_n = 1) + 1 - \Pr(v_n = 0) \right)^{1-d_n}$$

$$= \prod_{n=1}^N \left(p_n \tilde{h}_n(\hat{T}_n) e^{-\tilde{H}_n(\hat{T}_n)} \right)^{d_n} \left(p_n e^{-\tilde{H}_n(T_0)} + 1 - p_n \right)^{1-d_n} \quad (2.10)$$

$$\ell(\theta; \vec{T}) = \sum_{n=1}^N \left\{ d_n \left(\log p_n + \log \tilde{h}_n(\hat{T}_n) - \tilde{H}_n(\hat{T}_n) \right) + (1 - d_n) \log \left(p_n e^{-\tilde{H}_n(T_0)} + 1 - p_n \right) \right\} \quad (2.11)$$

2.2.3 ハザード関数を使用することの利点

ここでは特に、(右側) センサリングされたデータを分析する場合にハザード関数を用いることの利点を、密度関数を直接定式化する方法や回帰モデルと比較しながら説明する。

密度関数を直接定式化する方法との比較

密度関数 $f_n(t)$ を定式化する場合は、 $f_n(t) \geq 0$ かつ $\int_{-\infty}^{\infty} f_n(t) dt = 1$ なる条件を満足しなければならない。また分布関数 $F_n(t)$ を定式化する場合は、非負性や単調増加性の他に $\lim_{t \rightarrow \infty} F(t) = 1$ なる条件を満足しなければならない。

しかしハザード関数 $h_n(t)$ に科せられる条件は非負性のみであり、非負性を満足するハザード関数から 2.4 式によって計算される $f_n(t)$, $F_n(t)$ は、確率の公理を満足する。すなわちハザード関数 $h_n(t)$ を用いればより柔軟な定式化が可能であり、共変量の効果などを比較的自由に考慮することができる [40, p.19]。

回帰モデルとの比較

次に Helsen and Schmittlein の研究 [26] に従いながら、回帰モデルと比較した場合の利点を紹介する。センサリングを含むデータを回帰モデルによって分析すると、そのパラメータ推定値は、次のようにバイアスが発生する可能性がある。

1. サンプル・セレクションによるバイアス

観測期間内に実際に購買を行った家計のデータのみを回帰分析に用いる場合は、期間内に購買を行わなかった家計のデータは、たとえその後に購買を行ったとしても切り捨てられる。従っていわゆるセレクション・バイアスが発生し、パラメータは過剰に推定される傾向にある。

とくに試用購買や購買間隔が比較的長い商品を対象にした分析は、十分な量のデータを確保するのが困難なので、この問題は深刻である。

2. センサリングによるバイアス

いま耐久時間 T が独立かつ同一に期待値 $\beta'X$ の指数分布に従っており、さらに共変量 X がダイナミックに変化しないと仮定する。

$$f(t|X) = \frac{1}{\beta'X} e^{-\frac{t}{\beta'X}} \quad (2.12)$$

このとき線形回帰モデルのパラメータ推定量、すなわち共変量 X とセンサリング時点 T_0 を与えた下での条件付き期待値は、

$$E[t|X, T_0, t < T_0] = \beta'X \left(1 - \frac{\frac{T_0}{\beta'X} e^{-\frac{T_0}{\beta'X}}}{1 - e^{-\frac{T_0}{\beta'X}}} \right) \quad (2.13)$$

となり、パラメータ推定値にはバイアスがかかっていることが分かる。

3. 時間に依存する共変量を考慮することによるバイアス

回帰モデルのパラメータ推定値は、共変量が時間に依存することによってもバイアスがかかる。ここでも期待値 $\beta'X$ の指数分布に従う購買間隔データ T を例にとり、簡単のため、センサリングは発生していないとする。

さらに共変量は時間に依存しており、具体的には

$$X(t) := \begin{cases} X_1, & t < \tau \\ X_2, & t \geq \tau \end{cases} \quad (2.14)$$

であるとする、購買間隔の条件付き期待値は次のようになって、やはりバイアスがかかっていることが分かる。

$$E[t|X_1, X_2, \tau] = \beta' X_1 + \exp -\frac{\tau}{\beta' X_1} (\beta' X_2 - \beta' X_1) \quad (2.15)$$

したがってセンサリングが発生したデータを分析する場合は、パラメータ推定値にバイアスがかかる回帰モデルよりも、ハザード関数を用いた方が良いと思われる。

Chapter 3

既存研究

第1章で述べたとおり、新製品の普及 (diffusion of innovation) は、マネジリアルなニーズもあって、マーケティングでも比較的早く (1960 年代初頭) から研究されており、今日に至るまで膨大な量の論文が発表されている¹。そこでこの章では全体を大きく三つの節に分け、時代的な流れも考慮しながら、マーケティングに於ける普及モデルがどの様に発展していったかを、いくつかの研究を取り上げながら紹介したい。

最初の節では、Bass モデルをはじめとする集計レベル、すなわち集計レベルのデータを用いて普及曲線を近似する研究を紹介する。

次に、各家計が試用購買する過程を明示する事で試用購買する確率やタイミングに関する家計の異質性を考慮した研究を紹介する。これらは von Neumann-Morgenstern の期待効用理論をはじめとする、経済学に於ける家計の理論 [58, 8] を援用している。

ところで2章でも述べたとおり、ハザード関数を用いることによって柔軟なモデルを構築することが可能である。また試用購買データには一般に、センサリングが発生しているので、パラメータ推定値にバイアスがかかる回帰モデルよりも、ハザード関数の方が分析に適している。そこで最後に、ハザード関数によって定式化された非集計レベルの研究を紹介する。

¹ 1990 年頃までの研究を詳細に網羅したレビュー論文としては、Mahajan [40] 等がある。

3.1 集計レベルの研究

代表的な集計レベルの普及モデルとしては、まず最初に Bass の研究 [2] を挙げる事ができよう。

3.1.1 Bass 1969

モデル

彼は数多くの実証的研究を通じて、冷蔵庫、テレビ、エアコン、乾燥機、アイロン等の耐久消費財市場に於ける革新的新商品の普及に関して、次のような仮説を立てた。

1. 普及を促進する要因として、情報（の伝搬）が重要である。そして情報を伝搬する媒体としては、特に広告とクチコミ（word of mouth）の果たす役割が大きい。
2. 試用購買者は、革新的試用購買者（innovator）と模倣的試用購買者（imitator）の二つに大別される。ここで革新的試用購買者とは、広告の影響を受けて試用購買を行う者である。また模倣的試用購買者とは、既に試用購買した家計からのクチコミの影響を受けて試用購買を行う者である。

そして市場規模、すなわち潜在的な試用購買者数を M 、また時点 t までに試用購買した者の数を $N(t)$ とする時、時点 t まで試用購買しなかったという条件下で、時点 t に試用購買する確率（ハザード関数）を、次のように定義する。

$$h(t) = q + \frac{r}{M}N(t) \quad (3.1)$$

また潜在的な試用購買者の中で時点 t までに試用購買を行う者の割合を $F(t)$ とすると、 $F(t)$ は $N(t)/M$ に等しいので、上の定義は次のように書き直すことができる。

$$h(t) = q + rF(t) \quad (3.2)$$

すなわち耐久消費財の普及曲線を規定するハザード関数は、既に試用購買した家計からのクチコミと広告の影響によって決定されると考えられる。このため q, r はそれぞれ、革新係数 (coefficient of innovation)、模倣係数 (coefficient of imitation) と呼ばれている。

ここで 3.2 式を $F(t)$ について解くと

$$F(t) = \frac{1 - e^{-(q+r)t}}{\frac{r}{q}e^{-(q+r)t} + 1} \quad (3.3)$$

であるから非試用購買者、すなわち決して試用購買しない者を含めた全サンプル数を $N (\geq M)$ とするとき、Bass モデルの普及曲線

$$\frac{N(t)}{N} = p \frac{1 - e^{-(q+r)t}}{re^{-(q+r)t}/q + 1} \quad (3.4)$$

は、究極的普及率 $p := M/N$ に漸近してゆく。特にその微分 $f(t) := F'(t)$ に関して

$$f(t) = \frac{(q+r)^2}{q} \frac{e^{-(q+r)t}}{\left(\frac{r}{q}e^{-(q+r)t} + 1\right)^2} \quad (3.5)$$

が成立し、 $F(t)$ の変曲点は

$$t^* = \frac{1}{q+r} \log \frac{r}{q} \quad (3.6)$$

となるから、 $q > r$ すなわち広告が普及に与える影響の方がクチコミの影響よりも大きい場合は、その普及曲線は図 3.1 のように上に凸の急速浸透型になる。

また $q < r$ すなわちクチコミの方が広告等よりも普及に与える影響が大きい時は、その普及曲線は図 3.2 のように S 字型になる。

ところで $h = f/(1-F)$ だから、3.2 式を変形する事によって次式が得られる。

$$f(t) = (1 - F(t)) (q + rF(t)) \quad (3.7)$$

図 3.1: Bass モデルの普及曲線 ($q > r$)

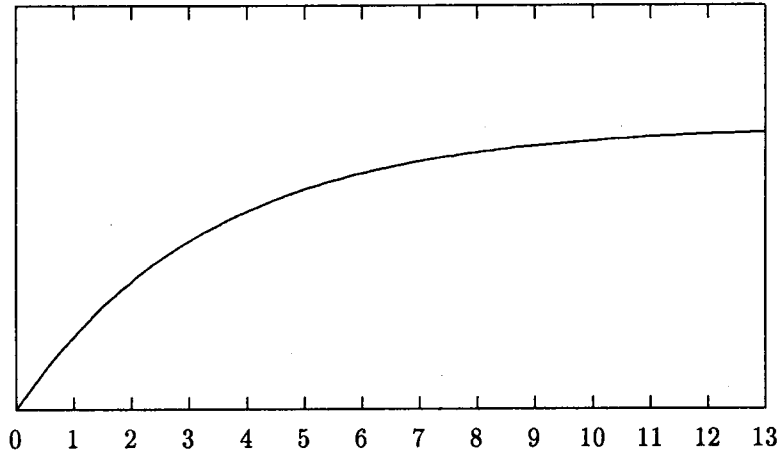
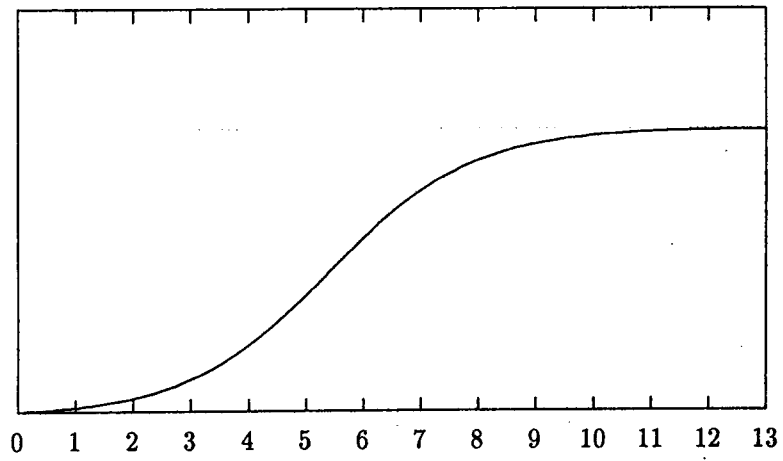


図 3.2: Bass モデルの普及曲線 ($q < r$)



これは $r \neq 0$ の時、潜在的な試用購買者の中で新たに試用購買する者の割合は、既に試用購買者した家計数に比例し、さらにその比例定数がクチコミや広告の影響力の関数になっている事を表している。その意味で Bass モデルは、伝染性 (epidemic) の普及モデルとも呼ばれる。

実証分析

なお 3.7 式から

$$n(t) := \frac{d}{dt}N(t) = q m + (r - q) N(t) - \frac{r}{m}N(t)^2 \quad (3.8)$$

が導かれ、その差分近似

$$n_t = q m + (r - q) N_{t-1} - \frac{r}{m}N_{t-1}^2 \quad (3.9)$$

は、時点 t の新規試用購買者数を表している。そこで Bass は 3.9 式に実際のデータを当てはめることによって、モデルのパラメータ m, q, r の最小自乗推定を行い、新規試用購買率のピーク t^* やその規模 $n(t^*)$ を計算し、さらに試用購買者数 $N(t)$ 長期予測を行っている。

この論文では、電気冷蔵庫やエアコンをはじめとする 11 の耐久消費財を対象にした実証分析を行っており、良好な予測結果が得られている。

まとめ

Bass モデルは、パラメータの推定が容易で、しかもその予測精度が比較的高いことから、新製品の予測モデルとして多くの企業で実際に用いられている。反面、このモデルでは共変量の影響を考慮していないので、定常な市場を対象としていると考えられる。そこで Bass モデルを拡張して共変量を明示的に取り込み、マーケティング戦略の立案等も可能にした研究が、数多く行われている。

3.1.2 Fournier and Woodlock 1960

Bass の研究の前後に行われた研究の多くは、その特殊型と考えることができる。例えば最寄り品を対象にした Fournier and Woodlock の研究 [15] は、模倣効果のない Bass モデル、すなわち潜在的な試用購買者は他人の影響を受けずに、広告だけで試用購買のタイミングを決定するモデルと考えることができる。

彼らはまず最寄り品の普及に関する実際の事例を検討した結果、その普及曲線 $\Phi(t)$ のグラフが次のような性質を持っていると指摘した。

1. 単調増加関数で次第に逓減する。すなわち試用購買する割合は市場導入時が最も高く、次第に低くなる傾向にある。
2. その極限である究極的普及率 p は一般に、1 を遥かに下回る²。すなわち全ての家計が試用購買する訳ではなく、むしろその殆んどが試用購買を行わない。

そしてこのような普及曲線を近似する為に、 $\Phi(t)$ の差分、すなわち各時点の新規試用購買率 $\phi(t) := \Phi(t) - \Phi(t-1)$ を、次のように定式化する方法を提案している。

$$\phi(t) = q(p - \Phi(t-1)) = q p(1-r)^{t-1} \quad (3.10)$$

ここで p は究極的普及率、すなわち潜在的な試用購買者が全サンプル中に占める割合を表している。

これは「まだ試用購買を行っていない潜在的な試用購買者のうちの一定の割合が、単位期間内に試用購買を行う」という事を仮定しており、一般に浸透戦略 [36, 65] 下での普及過程を表している。

なお 3.10 式は時間に関して離散的に定式化されているが、これを連続的なモデルに修正し、潜在的な試用購買者が時点 t に実際に試用購買する確率を $f(t) := \phi(t)/p$ とすると、次式が得られる。

² Fournier and Woodlock は究極的普及率を、‘the limiting ceiling penetration’ と呼んでいる。

$$f(t) = q(1 - F(t)), \quad F(t) := \int_0^t f(\tau) d\tau \quad (3.11)$$

これは 3.2 式において $r = 0$ としたものに他ならない。また最寄り品のシェアを予測するための Parfitt and Collins の研究 [49] も、試用購買のモデルはこれと全く同じである。

なおこのモデルは、ケーキミックスやマーガリン、シリアル等の最寄り品の普及データに適用されて、良好の成果を取めている。

3.1.3 Horsky and Simon 1983

Bass モデルや Fourt and Woodlock モデルはいわば、得られたデータの軌跡を忠実にトレースしながら予測を行うモデルであり、マーケティング・ミックスをはじめとする要因が時間とともに変動する市場では、正確な分析を行うことは難しい。また共変量 (covariate) を取り込んでいないので、シミュレーションによって最適なマーケティング戦略を立案することも不可能である。

そこで Bass モデルに価格や広告等の影響を明示的に取り込む試みが、数多くなされており [34]、価格を考慮した研究としては Dolan and Jeuland [10] や Kalish and Lilien [32] 等、また広告を考慮した研究としては Horsky and Simon [27] 等がある。ここでは Bass モデルに広告支出額の影響を取り込んだ Horsky and Simon の研究を紹介する。

彼等は 3.2 式で表される Bass モデルのハザード関数に、広告支出額を表す共変量 $A(t)$ を、次のように取り込んでいる。

$$h(t) := q + rF(t) + s \log A(t) \quad (3.12)$$

広告効果には、累積性と逓減性があることが先験的に知られている [38, p.265]。そこでこのモデルでは $A(t)$ を対数変換することによって逓減性を実現している。またハザード関数が累積試用購買率の関数となっているので累積性も実現されている。

なお電話による預貯金システム (telephonic banking system) の契約実績データに当ては

めた実証分析の結果、たとえば利率を考えない場合の最適広告戦略は単調減少型、すなわち市場導入と同時に積極的に広告を行い、その投入量を次第に減らしていく戦略であるとされている。

3.1.4 Nakanishi 1973

ここで紹介する Nakanishi の研究 [45] は、最寄り品を対象にした集計レベルの試用・反復購買モデルであり、ダイナミックに変化する共変量の影響を考慮している。そして試用・反復購買のタイミングを、それぞれスプリット・ハザード関数で定式化しているため、究極的な普及率や反復購買率を求めることが可能である。ここでは、試用購買の部分について説明する。

モデル

家計が潜在的な試用購買者である確率を p , $0 \leq p \leq 1$ とし、また潜在的な試用購買者が実際に試用購買を行うタイミングのハザード関数³を $h(t)$ とする。ここで彼は、ハザード関数 $h(t)$ を具体的に次のように定義している。

$$h(t) := \lambda h(t-1) + \beta' X(t) \quad (3.13)$$

ここで λ , $0 < \lambda < 1$ は、学習効果を表す分布ラグ係数である。また共変量 $X(t)$ としては、当期のテレビ広告（提供広告とスポット広告）や3期前のクーポンの支出額を採用している。

実証分析

モデルのパラメータ p , λ , β の推定には、日記式パネルを集計して得られた各時点の累積試用購買者数のデータ $\{N(t)\}_t$ と家計の総数 N が用いられる。

実際、潜在的な試用購買者が時点 t 以前に試用購買を行う確率を $F(t)$ とすると、

³ Nakanishi はハザード関数を、'conditional purchase propensity function' と呼んでいる。

表 3.1: Nakanishi モデルの推定結果

p	最小自乗推定値					R^2
	λ	定数項	提供広告	スポット広告	クーポン	
0.7	0.5939	0.00147	0.00141	0.00153	0.00149	0.8061
0.9	0.5891	0.00113	0.00110	0.00119	0.00116	0.8099
0.95	0.5882	0.00106	0.00104	0.00113	0.00110	0.8106
1.0	0.5874	0.00100	0.00099	0.00107	0.00104	0.8112

(Nakanishi [45, p.246] より抜粋)

$$p F(t) = p \left(1 - e^{-H(t)}\right) \simeq \frac{N(t)}{N} \quad (3.14)$$

は、時点 t 以前に試用購買を行う家計が全家計中にしめる割合を表している。そこで $h(t)$ の差分近似

$$h(t) = \frac{f(t)}{1 - F(t)} \simeq \frac{F(t) - F(t-1)}{1 - F(t-1)} \simeq \frac{N(t) - N(t-1)}{Np - N(t)} =: \hat{h}(t|p) \quad (3.15)$$

を、3.13 式における $h(t)$ のオブザベーションとする。ここでパラメータ p, λ, β の同時推定は行っていない。すなわち p を段階的に動かして、それぞれについて $\{\hat{h}(t|p)\}_t$ を計算し、3.13 式を用いて β の最小自乗推定値を求める。そして最も決定係数の高かった p と $\hat{\beta}$ の組を、モデルのパラメータ推定値として試用購買している。

26 週間にわたるトマトケチャップの試用購買データを用いた実証分析の結果を、表 3.1 にまとめておく。

この結果、究極的普及率を $p = 1$ とした場合の決定係数が最も高い。そしてその時、分布ラグ係数 λ の推定値は 1% 水準で有意になっており、学習効果が効いている事を示している。

まとめ

Nakanishi のモデルは、スプリット・ハザード関数を用いて定式化された試用・反復購買モデルとして、先駆的なものの一つである。そしてモデルの構造がシンプルであるにも関わらず、高い記述性と柔軟性を持っている。

しかしこのモデルの課題としては、次のふたつが挙げられよう。ひとつは、パラメータの同時推定が行われていない点である。3.13 式で表されるハザード関数は、潜在的な試用購買者であるという条件下で定義されているので、究極的普及率（潜在的な試用購買者である確率） p とハザード関数のパラメータ λ, β は、同時に推定されることが望ましい。

またもうひとつは、究極的普及率 p がアприオリに与えられているという事である。潜在的な試用購買者であるか否かという事が、アприオリに Bernoulli 分布に従っているという仮定は、極めて制約的である。なお 3.3.2 項で紹介する杉田・中村・田島のモデル [72] も、究極的普及率に関して同じ仮定を試用購買しているが、どちらのモデルも $\hat{p} = 1$ となっている事が大変興味深い。

3.2 試用購買する過程を明示した研究

集計データを用いた研究は、データの収集からモデルの構築そしてパラメータの推定やその解釈にいたるまで、分析手順が簡便でコストが低いという利点がある。しかしその反面、試用購買する確率やそのタイミングに関する家計の異質性を検出することができない。

そこで、経済学に於ける家計の理論 [58, 8] を援用することによって、各家計の試用購買する過程に関する異質性を考慮した研究が行われている。ここでは代表的なふたつの研究を紹介する。とくに後者の Roberts and Urban の研究は、非集計レベルのモデルである。

普及モデルはそのマネジリアルなニーズを考えると、市場導入後できるだけ早期に分析を行わなければならない。従って集計レベルのデータの場合は、分析に利用できるオブザベーション数が少なく、モデルのパラメータ推定値の頑健性を確保するのが困難である⁴。

⁴ 例えば Bass モデルの場合、市場導入後 4 年間の年次データ、すなわち 4 つのデータを用いて、パラメータの最小自乗推定量を求めている [34, p.92]

しかし非集計レベルのデータを用いることにより、市場導入後早期であっても対象となる家計数を増やすことによって、パラメータ推定に十分な量のオブザベーションを確保できるという利点を持つ。

3.2.1 Kalish 1985

価格と広告が普及に与える影響を考慮した Kalish の研究 [31] は、各家計の想起過程と試用購買過程をそれぞれ明示的にモデル化することによって、最終的に集計レベルの普及曲線を導出している。

彼は各家計の想起・試用購買過程を次のようにモデル化している。

1. 製品属性は、探索属性と経験属性に分類される。探索属性 (search attributes) は、色やサイズ等の、簡単に検証できる属性であって、広告やクチコミ等によって伝搬される。また経験属性 (experience attributes) は、信頼性などの、実際に製品を試用することによって得られる属性であって、クチコミによって伝搬される。
2. 想起 (awareness) とは、探索属性を取得する事を意味している。すなわち想起過程は、広告やクチコミに大きく影響される。特にここでは、家計は探索情報の取得に関して同質的であると仮定し、一般的な伝染型モデルを当てはめている。従って累積想起率 $I(t)$ は、広告支出額 $A(t)$ や累積試用購買者数 $N(t)$ 、市場規模 M の関数として、次式で表される。

$$\frac{d}{dt}I(t) = (1 - I(t)) \left\{ h(A(t)) + b \left(I(t) - \frac{N(t)}{M} \right) + c \frac{N(t)}{M} \right\} \quad (3.16)$$

すなわち想起率のハザード関数は、広告 $A(t)$ および、まだ試用購買していない想起者からのクチコミ $I(t) - N(t)/M$ と既に試用購買した家計からのクチコミ $N(t)/M$ によって決定され则认为る。

3. 当該新製品に対して形成する効用 u は家計毎に異なる。ここで u は、家計に関して

独立かつ同一に指数分布に従うと仮定している。

4. 試用購買前の想起者は、当該新製品の経験属性に対して不確実性（リスク）を持っており、それがリスク回避係数 r , $0 \leq r \leq 1$ という形で u に影響を与える。

そしてリスクはクチコミの流入とともに減少する。すなわち r は $N(t)/M$ の増加関数である。

5. 当該新製品の試用購買条件は、価格 $PR(t)$ に対して、

$$r\left(\frac{N(t)}{M}\right)u \geq PR(t) \quad (3.17)$$

である。

すなわちリスク修正済みの価値 $r\left(\frac{N(t)}{M}\right)u$ が実際の価格 $PR(t)$ を上回るときに、試用購買を決定する。

6. 試用購買条件 3.17 式を満足する家計が、その後の微少な時間 dt の間に実際に試用購買する確率は、 $k dt$ であると仮定する。

このとき集計レベルの普及曲線 $N(t)$ に関して、次式が成立する。

$$\frac{d}{dt}N(t) = \left\{ M \int_{ru \geq PR(t)} \Pr(v \leq e) de I(t) - N(t) \right\} k \quad (3.18)$$

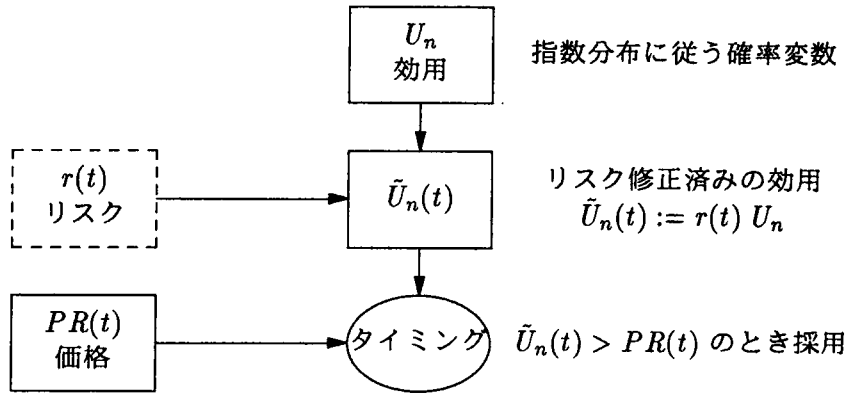
なお上式から計算された最寄り品の最適戦略は浸透戦略である。

3.2.2 Roberts and Urban 1988

Roberts and Urban の研究 [51] は、von Neumann-Morgenstern の期待効用理論 [35] を用いて、各家計が試用購買する過程を次のようにモデル化している。

1. 家計 n は当該新製品の属性 $X_n(t)$ を評価し、その客観的選好を $\beta'X_n(t)$ とする。ちなみにこの研究では、自家用車を対象にした実証分析を行っているが、そこで用いら

図 3.3: Kalish のモデル



れた属性は、スタイルや燃費、安全性等である。

2. 家計は情報の欠如や製品固有の変動により、当該新製品に対して、何らかの不確実性 $\eta_n(t)$ を抱く。ここで $\eta_n(t)$ は、家計に関して独立かつ同一に正規分布 $N(0, \sigma(t)^2)$ に従っている。そしてこの不確実性を考慮した主観的な選好を、 $Y_n(t)$ とする。

$$Y_n(t) := \beta' X_n(t) + \eta_n(t) \sim N(\beta' X_n, \sigma(t)^2) \quad (3.19)$$

なお耐久消費財市場における不確実性 $\eta_n(t)$ の分散 $\sigma(t)^2$ は、時間の経過や情報の流入とともに減少すると考えられ、Bayes 更新が行われる。

3. 主観的選好 $Y_n(t)$ から効用 $U_n(t)$ が、次のように形成される。ただし r は、リスク回避パラメータである。

$$U_n(t) = U(Y_n(t)) := -e^{-r Y_n(t)} \quad (3.20)$$

4. 各家計は期待効用値

$$E[U(Y_n(t)); \eta_n(t)] = -\exp\left(-r(\beta' X_n - \frac{r}{2}\sigma(t)^2)\right) \quad (3.21)$$

と status quo を比較して試用購買を決定するが、これは次のように書き直すことができる。

$$\chi_n(t) := \beta' X_n(t) - \frac{r}{2} \sigma(t)^2 > 0 \quad (3.22)$$

5. 家計の異質性などに起因する攪乱項 $\varepsilon_n(t)$ を考慮した場合、家計の試用購買基準は最終的に次のようになる。

$$\text{新製品を試用購買} \iff \tilde{\chi}_n(t) := \chi_n(t) + \varepsilon_n(t) > 0 \quad (3.23)$$

この時、 $\varepsilon_n(t)$ が独立かつ同一に第一種極値分布に従うと仮定すると、当該新製品の試用購買確率は良く知られたように [44]、次の二項ロジットモデルに帰着する。

$$p_n(t) = \frac{1}{1 + e^{-\tilde{\chi}_n(t)}} \quad (3.24)$$

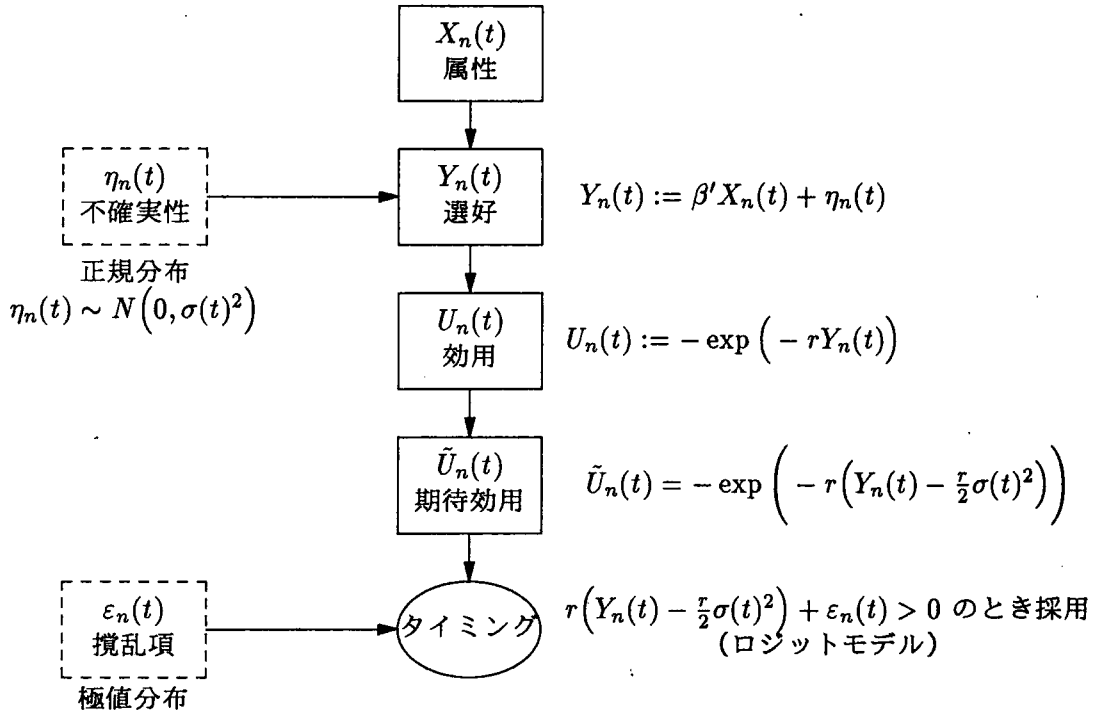
3.2.3 Chatterjee and Eliashberg 1990

Chatterjee and Eliashberg [4] は、von Neumann-Morgenstern の期待効用理論と情報（クチコミ）の流入による Bayes 更新を取り込んだ、個人レベルの試用購買モデルを構築した。このモデルは、選好構造や情報の信頼性に関する家計の異質性を考慮しているため、パラメータに関する期待値をとる事によって得られる普及曲線は、極めて柔軟である。さらに個人別のパラメータによって、試用購買時点に関する家計のセグメンテーションが可能である。

家計レベルの試用購買モデル

家計は当該新製品の価格とパフォーマンスを評価する。しかし真のパフォーマンスは未知であり、また選好や効用には不確実性がともなっているため、期待効用と価格を比較す

図 3.4: Roberts and Urban のモデル



ることによって試用購買を決定する。なお試用購買しない場合は、クチコミによる情報の流入の度に選好が Bayes 更新され、試用購買の意志決定が逐次行われる。すなわち試用購買のタイミングは、情報の流入量によって決定されるのである。

いま家計 n は既に、 i 単位の情報を受け取っているとす。この時、不確実性をともなう選好 $Y_{n,i}$

$$Y_{n,i} \sim N(m_{n,i}, s_{n,i}^2) \quad (3.25)$$

から、(確率的) 効用 $U_n(Y_{n,i})$ が次のように定義される。

$$U_n(Y_{n,i}) := 1 - e^{-r_n Y_{n,i}} \quad (3.26)$$

ここで $m_{n,i}, s_{n,i}^2$ はそれぞれ、当該新製品のパフォーマンスに対する認知とその不確実性を表している。また r_n は、リスク回避パラメータである。この時

$$\tilde{U}_{n,i} := E[U_n(Y_{n,i})] = 1 - \exp\left(-r_n m_{n,i} + \frac{r_n^2 + s_{n,i}^2}{2}\right) > k_n Price \quad (3.27)$$

すなわち

$$m_{n,i} > \underbrace{\frac{r_n + s_{n,i}^2}{2}}_{\text{リスク・ハードル}} - \underbrace{\frac{1}{r_n} \log(1 - k_n Price)}_{\text{価格ハードル}} \quad (3.28)$$

の場合に、当該新製品は試用購買される。ただし k_n は、価格の重要度を表すパラメータである。

また i 単位の情報を受け取った段階で試用購買を行わなかった時は、更に情報を受け取る事によって、選好が Bayes 更新される。家計 n が受け取る一単位の情報は、正規分布 $N(\mu, \sigma_n^2)$ に従っているとす。ただし μ は当該新製品の真のパフォーマンス、また σ_n^2 は情報の信頼性を表している。この時、実際に一単位の情報 z_n を受け取る事によって、選好 $Y_{n,i}$ の平均 $m_{n,i}$ および分散 $s_{n,i}^2$ が、次のように更新される。

$$m_{n,i+1} = \frac{m_{n,i}/s_{n,i}^2 + z_{i+1}/\sigma_n^2}{1/s_{n,i}^2 + 1/\sigma_n^2} \quad (3.29)$$

$$s_{n,i+1}^2 = \frac{1}{1/s_{n,i}^2 + 1/\sigma_n^2}$$

そして基本的には、実際に試用購買されるまで、試用購買の意志決定と Bayes 更新が繰り返されるのである。

家計の分類

ここで $y_{n,i}$, α_n , β_n を次のように定義する。

$$y_{n,i} := \frac{\sigma_n^2}{s_{n,i}^2} \left(\frac{r_n + s_{n,i}^2}{2} - \frac{1}{r_n} \log(1 - k_n Price) - m_{n,i} \right) \quad (3.30)$$

$$\alpha_n := y_{n,0} \quad (3.31)$$

$$\beta_n := -\frac{1}{r_n} \log(1 - k_n \text{Price}) \quad (3.32)$$

このとき 3.28, 3.29 式はそれぞれ、次のように書き直すことができる。

$$y_{n,i} < 0 \quad (3.33)$$

$$y_{n,i} = y_{n,i-1} + (z_i - \beta_n) \quad (3.34)$$

さらに漸化式 3.34 を解いて 3.33 式に代入することにより、試用購買基準は次のような形になる。

$$\sum_{i=1}^{i_n^*} (z_i - \beta_n) > \alpha_n \quad (3.35)$$

ただし i_n^* は、家計 n が当該新製品を試用購買するのに必要な情報量のイキ値である。

ここで特に z_n が確率的に与えられると考え、 $z_n \sim N(\mu, \delta^2)$ とすると

$$y_{n,i} - y_{n,i-1} = -(z_n - \beta_n) \sim N(-(\mu - \beta_n), \delta^2) \quad (3.36)$$

であるから、 $y_{n,i}$ はランダム・ウォークまたは Wiener 過程に従う。

このとき個人別のパラメータ α_n, β_n を用いて、試用購買のタイミングに関する家計のセグメンテーションが、次のように行われる。

1 $\alpha_n = y_{n,0} < 0$ なる家計

3.33 式より、この家計は当該新製品を直ちに試用購買し、革新的試用購買者に対応している。

2, 3 $\alpha_n \geq 0$ なる家計

この家計の場合は、市場導入と試用購買にタイムラグが存在する。また $i_n^* | \alpha_n, \beta_n$ が inverse Gauss 分布に従っており、特に $\beta_n < \mu$ の時の平均値は、次のようになって

いる。

$$\gamma_n := E[i_n^* | \alpha_n, \beta_n] = \frac{\alpha_n}{\mu - \beta_n} = \frac{\sigma_n^2}{s_{n,0}^2} \frac{1}{\mu - \beta_n} \left(\frac{r_n s_{n,0}^2}{2} + \beta_n - m_{n,0} \right) \quad (3.37)$$

ここで γ_n は革新度の複合測度になっており、広告に対する信頼度等のアンケート結果と負の相関がある。この結果は、模倣的試用購買者に対する外部情報の影響に関する過去の研究と整合性があり、 γ_n によって模倣的試用購買者を更にセグメンテーションする事の妥当性を示唆している。

また $\beta_n \geq \mu$ の時は、 $i_n^* | \alpha_n, \beta_n$ のモーメントは発散する。これは潜在的な試用購買者であるとは限らないことを示している。

家計のセグメンテーションに関する結果を、図 3.2 にまとめておく。

表 3.2: Chatterjee *et al.* モデルによる家計の分類

家計		究極的に試用購買する確率	$i_n^* \alpha_n, \beta_n$	
タイプ	特性		平均	分散
1	$\alpha_n < 0$	1	0	0
2	$\alpha_n \geq 0, \beta_n < \mu$	1	$\frac{\alpha_n}{\mu - \beta_n}$	$\frac{\alpha_n \delta^2}{(\mu - \beta_n)^3}$
3	$\alpha_n \geq 0, \beta_n \geq \mu$	$\exp\left(-\frac{2\alpha_n \mu - \beta_n }{\delta^2}\right)$	-	-

(Chatterjee *et al.*[4, p.1064] より抜粋)

集計レベルの普及曲線

ここで時点 t 迄に流入する情報量を $i(t)$ と表すことにする。この時タイプ 2, 3 の家計に関して、情報量のイキ値 $i_n^* | \alpha_n, \beta_n$ は inverse Gauss 分布に従っているので、試用購買時点を表す確率変数 T_n の密度関数 $f_n(t)$ は、次のようになる。

$$f_n(t) = f(t | \alpha_n, \beta_n)$$

$$= \Phi\left(\frac{-\alpha_n + (\mu - \beta_n)i(t)}{\delta\sqrt{i(t)}}\right) + \exp\left(\frac{2\alpha_n(\mu - \beta_n)}{\delta^2}\right) \Phi\left(\frac{-\alpha_n - (\mu - \beta_n)i(t)}{\delta\sqrt{i(t)}}\right) \quad (3.38)$$

したがって α_n, β_n に関する期待値を取ることによって、集計レベルの普及曲線 $F(t)$ が得られる。

$$F(t) = \int_0^\infty \int_{-\infty}^0 g_{\alpha,\beta}(\alpha, \beta) d\alpha d\beta + \int_0^\infty \int_0^\infty f(t|\alpha, \beta) g_{\alpha,\beta}(\alpha, \beta) d\alpha d\beta \quad (3.39)$$

さらに情報の流入が非確率的であるとすると、 $\delta = 0$ となるので、3.39 式で表される普及曲線は、次のように簡単になる。

$$F(t) = \psi_1 + \psi_2 G_\gamma(i(t)) \quad (3.40)$$

ただし ψ_1, ψ_2 はそれぞれ、タイプ 1, 2 の家計の占める割合である。Bass モデルをはじめとする多くの伝染型普及モデルは、3.40 式の特殊型として記述される。したがって Chatterjee and Eliashberg の普及曲線は、極めて柔軟性に富んでいることが分かる。

実証分析

この論文では実証分析として、実験によって得られたデータを用いて、3.28 式で表される個人レベルのモデルによる予測を行っている。

このモデルで推定すべきパラメータは、リスク回避度 r_n 、価格の重要度 k_n 、初期選好のパラメータ $m_{n,0}, s_{n,0}^2$ 、および情報 z_n とその信頼性 σ_n^2 である。まず富くじ (lottery) によって r_n を、そして価格とパフォーマンスが異なるいくつかの架空の商品からひとつを選択させる事によって k_n を推定する。次に模擬クチコミ実験で当該新製品を評価させることにより、 $m_{n,0}, s_{n,0}^2$ および $\{m_{n,i}, s_{n,i}^2\}_{i=1}^3$ を推定する。そして最後に、 z_n と σ_n^2 を 3.29 式から計算する。

上記の手順で推定されたパラメータを用いた予測結果を、表 3.3 にまとめておく。

なお上記の実験とは独立に調査されたアンケートによると、タイプ 2 の家計に関して、

表 3.3: Chatterjee *et al.* モデルの予測結果

データ	情報流入量				センサリング
	0	1	2	3	
実測値	12	8	3	2	24
予測値	13	4	3	1	28

(Chatterjee *et al.*[4, p.1074] より抜粋)

情報量のイキ値 $i_n^* | \alpha_n, \beta_n$ の期待値 γ_n と広告に関する信頼度等との間には、負の相関が見られた。これは、模倣的試用購買者は外部情報よりもクチコミ等の内部情報に大きく影響されるという、過去の研究成果と整合性を保っており、 γ_n による家計のセグメンテーションの有効性を示唆している。

3.3 ハザード関数を用いた非集計レベルの研究

上で紹介した研究は、各家計が試用購買する過程を取り込むことによって、家計の異質性を考慮した普及モデルを構築した。しかしこれらの研究は、実験やアンケートによってデータを収集する必要があったり、またパラメータの推定が複雑だったりして、分析に非常にコストがかかるのが問題である。

ところで非集計レベルのモデルをハザード関数で定式化する研究が、いくつか存在する。2章で述べたとおり、ハザード関数を用いる事によって、柔軟なモデル構築が可能である。さらにハザード関数は、耐久データを分析するのに適している。じっさいハザード関数のマーケティングへの適用は比較的早く、Bass [2] や Nakanishi [45] のモデルも、原論文ではハザード関数を用いて定式化されている。しかし非集計レベルのモデルはまだ研究例が少なく、ここで紹介する試用購買モデル以外では、購買間隔モデルの研究でいくつか存在する程度である [19, 28, 26]。

3.3.1 Sinha and Chandrashekar 1992

Sinha and Chandrashekar の研究 [56] では、スプリット・ハザード関数によって定式化された非集計レベルのモデルを用いて、現金自動支払機 (ATM: Automatic Teller Machine) の普及過程を分析している。とくにこのモデルでは、ATM を試用購入する確率を二項ロジットモデルで定式化することによって、確率とタイミングの両方を、ダイナミックに変化しない共変量で説明している点が興味深い。

モデル

まず潜在的な試用購入者であるか否かを表す確率的ダミー変数 v_n に対して、 $\Pr(v_n = 1)$ が次の二項ロジットモデルによって定式化されている。

$$\Pr(v_n = 1) = \frac{1}{1 + e^{\alpha' X_n}} \quad (3.41)$$

すなわち v_n は、 n に関して独立に、 $1/(1 + e^{\alpha' X_n})$ をパラメータに持つ Bernoulli 分布に従っているのである。

$$v_n \sim Bi\left(\frac{1}{1 + e^{\alpha' X_n}}\right) \quad (3.42)$$

また $v_n = 1$ すなわち、いつかは ATM を試用購入する銀行 n に対して、その条件付きハザード関数 $h_n(t|v_n)$ を、次の三通りに定義している。

1. 指数型

$$h_n(t|v_n = 1) := \lambda e^{\beta' X_n} \quad (3.43)$$

2. Weibull 型

$$h_n(t|v_n = 1) := (\lambda\gamma) (\lambda t)^{\gamma-1} e^{\beta' X_n} \quad (3.44)$$

3. 対数正規型

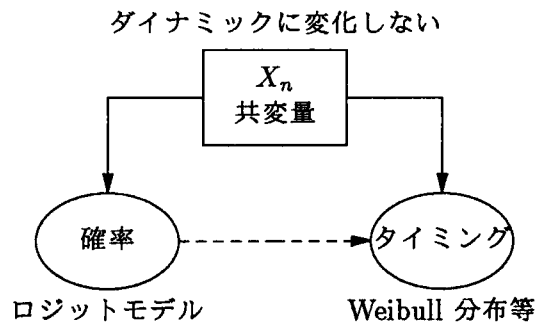
$$h_n(t|v_n = 1) := \frac{\phi(w_n)}{1 - \Phi(w_n)}, \quad w_n = \frac{\log t - \beta' X_n}{\sigma} \quad (3.45)$$

ここで ϕ, Φ はそれぞれ、正規分布の密度関数、分布関数を表している。

上で述べたとおり、このモデルでは確率とタイミングの両方に対して、共変量がそれぞれ独立に影響を与えている。したがって共変量および確率 v_n がダイナミックに変化した場合、その尤度関数を求めることは解析的に極めて困難である。なぜならば、このモデルの尤度関数は、試用購買以前の全ての $v_n(t)$ を与えた下での条件付き尤度関数の期待値によって定義されるからである。そこでここでは、各銀行の総資産や預金の成長率や、各銀行が担当する地域の平均賃金や銀行の集中度等の、ダイナミックに変化しない共変量を試用購買している。なおタイミングに影響を与える共変量と確率に影響を与える共変量は、同じ組み合わせである。

したがって本モデルの構造は、図 3.5 の様になっている。なおスプリット・ハザードモデルのパラメータの推定方法については、2.2.2 項を参照のこと。

図 3.5: Sinha and Chandrashekar のモデル



実証分析

この研究では、非集計レベルの ATM 普及データを用いて、ベースラインの確率分布や潜在的な試用購買者ダミー v_n を考慮する事の妥当性を検証し、さらに Bass モデル等との予測精度比較を行っている。

1. 確率分布の妥当性

まずベースラインの確率分布として、指数分布、Weibull 分布、そして対数正規分布のどれを仮定するのが最も妥当かを、 $v_n \equiv 1$ なる仮定の下で比較検討した。

それぞれの対数尤度を比較した結果、Weibull 分布および対数正規分布を仮定したモデルが、指数分布を仮定したモデルよりも当てはまりが良い事が分かった。しかし Weibull モデルおよび対数正規分布モデルの間には、有意な差が認められなかった。

次に v_n を二項ロジットモデルによって定式化した場合は、対数正規モデルの方が Weibull モデルよりも有意に優れていることが分かった。

2. 潜在的な試用購買者ダミー v_n を考慮する事の妥当性

対数正規分布を仮定したモデルを基にして、決して ATM を試用購買しない銀行の存在を考慮することの妥当性を検証した。

v_n を二項ロジットモデルによって定式化したモデルを、 v_n が独立かつ同一に Bernoulli 分布 $Bi(p)$ に従うモデルや、 $v_n \equiv 1$ としたモデルと、尤度比検定によって比較した結果、二項ロジットモデルによって定式化したモデルが有意に優れていることが分かった。

3. タイミングの予測精度

ATM を試用購買するタイミングの予測精度について、 v_n を二項ロジットモデルによって定式化した対数正規分布モデルを、 $v_n \sim Bi(p)$, (*i.i.d.*) とした対数正規分布モデルや Bass モデル等と比較した。

外挿予測値の平均自乗誤差 (MSE) 等を計算した結果、 v_n を二項ロジットモデルによって定式化したモデルの予測精度が最も高かった。

4. 試用購買確率の予測精度

潜在的な試用購買者である確率の予測精度について、対数正規分布モデルと Weibull モデルの比較を行った。ただし両方のモデルとも、 v_n は二項ロジットモデルによって定式化されている。

それぞれの的中率を計算した結果、対数正規分布モデルの方が優れていることが分かった。

まとめ

実証分析の結果を総合すると、二項ロジットモデルによって v_n を定式化した対数正規分布モデルが、モデルの当てはまりと予測精度の両面で最も優れている。

対数正規分布の分布関数は S 字型を描くので、確率分布に関するこの結論は、Bass のモデルをはじめとする耐久消費財の試用購買に関する既存研究の結果と一致している。また v_n の定式化に関する結論も、この研究の意図に沿ったものになっている。

しかしこのモデルの問題点として、確率とタイミングに対する共変量の影響が、独立かつアドホックに与えられていることが挙げられる。そのためパラメータ α, β の識別性が確保されず、また整合性のとれたインプリケーションが期待されない。じっさい試用購買するタイミングは、潜在的な試用購買者であるという条件下で定義されるので、ある構造下で同時に説明されるべきであろう。

3.3.2 杉田・中村・田島 1992

杉田・中村・田島の研究 [72] では、非集計レベルのスプリット・ハザードモデルを用いて、最寄り品の試用購買過程を分析している。Sinha and Chandrashekar のモデルでは、共変量がダイナミックに変化しないかわりに、確率とタイミングの両方を説明していた。しかしこのモデルでは、ダイナミックに変化する共変量がタイミングを説明するが、試用購買確率 v_n は Bernoulli 分布 $Bi(p)$ によってアприオリに与えられている。

モデル

まず、家計 n が潜在的な試用購買者であるか否かを表す確率的ダミー変数 v_n は、独立かつ同一に Bernoulli 分布 $Bi(p)$ に従っているとするとする。

$$v_n \sim Bi(p), \quad (i.i.d.) \quad (3.46)$$

したがって Nakanishi モデルと同様に、Bernoulli 分布のパラメータ p が、究極的普及

率を表している。

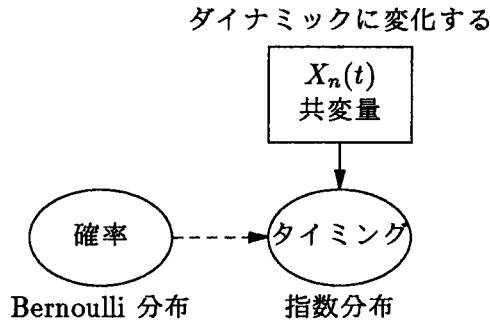
そして $v_n = 1$ なる家計 n に対して、ハザード関数 $h_n(t|v_n = 1)$ は次のように定義されている。

$$h_n(t|v_n = 1) := \lambda e^{\beta' X_n(t)} = e^{-a + \beta' X_n(t)}, \quad \lambda > 0 \quad (3.47)$$

すなわち試用購買のタイミングに指数分布を仮定しているので、このモデルは、浸透戦略を試用購買した場合のモデルと考えることができる。また共変量 $X_n(t)$ は、ヘビーユーザー一度をはじめとする家計の購買性向 (6.1.3 項参照) や価格・エンド陳列ダミーおよび店舗ダミーから成っている。

このモデルの構造は、図 3.6 のようになっている。

図 3.6: 杉田・中村・田島のモデル



ところで、打ち切り時点である T_0 以前に試用購買しなかった家計については、潜在的な試用購買者であるかどうか分からない。すなわちセンサリングされた家計の v_n が欠測している。したがってこの論文では、EM アルゴリズム [42, 88] によるパラメータ推定を試みている。なおアルゴリズムの詳細については Appendix B を参照されたい。

実証分析

この論文では、最寄り品のいくつかの新製品を対象にした、カテゴリー横断的な実証分析を行っている。その推定結果の一部を、表 3.4 にまとめておく。ただしパラメータ推定値の右隣にある a, b はそれぞれ、パラメータ推定値が 1%, 5% で有意になっていることを

表している。

表 3.4: 杉田・中村・田島モデルの推定結果

パラメータ	アイテム			
	ふりかけ	アイス	せんべい	重質洗剤
<i>a</i>	1.69	-2.94	-4.52	-1.01
ヘビーユーザー度	1.14 b	1.50 a	1.57 a	0.65
ロイヤルユーザー度	-0.70	-0.81	-3.50	-1.90 a
特売性向	0.44	-0.61	-0.70	1.64 a
店舗 a ダミー	0.76	1.00	0.91	2.06 b
店舗 b ダミー	0.31	0.70	0.51	1.25
エンド陳列ダミー	0.70 a	1.77 a	2.17	8.91
価格掛率	-6.24	-0.94	-2.90	-13.50
究極的普及率	0.98 a	0.99 a	0.94 a	0.28 a
対数尤度	-43.67	-65.62	-37.73	-51.70
尤度比検定量	19.40 b	35.08 a	36.26 a	35.09 a

(杉田・中村・田島 [72] より一部抜粋)

いずれのモデルも、共変量を考慮しないナルモデルとの尤度比検定量が有意になっており、またパラメータ推定値に符号も期待されたものであった。これは、このモデルが頑健であることを示唆している。また各共変量の平均値における弾力性を計算したところ、試用購買するタイミングに対する弾力性は、エンド陳列よりも価格の方が高いという、マネジリアルな含意が得られた。

まとめ

このモデルは、当該新製品を決して試用購買しない家計の存在を許し、ダイナミックに変化する共変量がタイミングを説明している。その結果、家計の異質性やマーケティング・ミックス等の共変量が試用購買のタイミングに与える影響を、短期間のうちに測定することができる。

またこのモデルでは、潜在的な試用購買者である確率がアプリアリに与えられている。この仮定は Nakanishi モデルと同様であるが、どちらのモデルでも究極的普及率の推定値

が 1 に近かったことは、大変興味深い。なおこの値に関する考察は、6.3.3 項で行われている。

しかし今後の課題として、パラメータ推定に要する時間の短縮が挙げられる。このモデルは共編量がダイナミックに変化するので、対数尤度関数の解析的構造が複雑である。従ってこのモデルのパラメータ推定には、非常に時間がかかる事が予想される。また今回はセンサリングされた家計の v_n が欠測している事に着目して、EM アルゴリズムによるパラメータ推定を行っている。しかし EM アルゴリズムによるパラメータの収束は、一般に非常に遅いことが知られている [88]。[72] では、実証分析の計算に unix ワークステーションを用いているが、それでもパラメータの推定に 20 時間以上かかっており、そこで実用性の観点からも改良の余地が残されている。

Chapter 4

既存研究の問題点と本研究の目的

この章では、3章でレビューした既存研究の総括および問題点の指摘を行い、さらに本研究の目的を明らかにする。

4.1 既存研究の問題点

4.1.1 集計レベルの研究

集計レベルの研究は、販売実績データ等の集計データを用いている。そして「潜在的な試用購買者の中で新たに試用購買する者の割合が既に試用購買した家計の数に比例する」という考えにたった Bass の研究や、Bass モデルに価格や広告等の共変量の影響を明示的に取り込んだ諸研究に代表される。

このアプローチの利点は、データ収集やパラメータ推定が比較的容易なので、分析コストが低いという点である。従ってこれらのモデルは、実際に多くの企業で普及率の推移や究極的普及率の予測、さらには最適戦略の立案の為に用いられている。そして今日に至るまでに膨大な研究例や実証例が報告されている [40]。

しかし集計レベルの研究は、ふたつの欠点を持つ。まず最初の問題点は、試用購買する確率やタイミングに関する家計が検出できないという事である。そのため実際の普及曲線

に現れる違いは、純粹にランダムなものであると解釈されるので、この仮定はかなり制約的であると思われる。

もうひとつの問題点は、安定したパラメータ推定値を得るために必要な量のデータを確保するのが困難な事である。なぜならば普及モデルはその性格上、市場導入後できるだけ早い時点で分析を行なう必要があるので、少数のデータしか用いることはできない。しかし集計レベルのデータは、期間当たりひとつのオブザベーションしか得られないので、分析に用いられるデータは極めて限られてしまう。実際 Bass の実証分析では、市場導入後 4 年間の年次データすなわち 4 つのデータを用いて、パラメータの最小自乗推定量を求めており、その統計的妥当性に関して疑問の余地が残る。

4.1.2 効用の概念を取り込んだ研究

これらの研究に共通する特徴は、経済学に於ける消費者の理論で説明される効用の概念を取り込むことにより、家計が試用購買する過程を明示的にモデル化している点である。すなわち家計の選好を効用関数で表現し、効用最大化ルールに従って新製品の試用購買過程を定式化している。さらにいくつかの研究では、情報の欠如や製品固有の変動等による効用の不確実性を考慮し、von Neumann-Morgenstern の期待効用理論を採用している。そして実際に効用を取り込むことによって、例えば Kalish の研究 [31] では試用購買する確率やタイミングに関する家計の異質性を考慮しており、また Roberts and Urban の研究 [51] や Sinha and Chandrashekar の研究 [56] では、当該新製品に対する選好と情報の伝達をそれぞれ確率的に与えている。さらに、このモデルは非集計レベルのデータを利用しているので、家計数を増やすことによってパラメータ推定に十分な量のデータを確保できるという利点もある。またマーケティング戦略上の含意を、理論的に導くことができる。

しかしこれらの研究に共通する問題点は、パラメータ推定に用いるデータの収集が困難なことである。じっさい多くの研究では、実験やアンケートによってデータを集めている。

4.1.3 ハザード関数を用いた非集計レベルの研究

普及過程をハザード関数を用いて定式化する試みは、比較的古くから行われてきた。例えば Bass の研究では、集計レベルのハザード関数を、累積試用購買者数の線形関数として定義し、それによって普及曲線を導出している。また Nakanishi の研究 [45] では、各家計の試用購買および反復購買のハザード関数を共変量の線形結合で定義し、最寄り品の売り上げ予測を行っている。

ハザード関数を用いて普及モデルを定式化することの利点としては、2 章で述べたとおり、次のふたつが挙げられる。まず最初に、ハザード関数が非負値である限り、対応する密度関数や分布関数はしかるべき確率の公理を満足する。したがってハザード関数を用いることによって、モデルの柔軟な定式化が可能であり、共変量の影響や家計の異質性を自由に考慮する事ができる。

ところで、試用購買データをはじめとするセンサリングされた耐久データを分析する場合に、一般的に用いられる回帰モデルを用いると、そのパラメータ推定値にバイアスが発生する。したがってハザードモデルを用いる利点のふたつめとしては、センサリングされた耐久データとの相性の良さがあげられよう。

このような利点があるにも関わらず、ハザード関数によって定式化された非集計レベルの研究は、ごくわずかしか報告されていない。そこでこれらのモデルを詳しく検討するために、それぞれの特徴を、表 4.1, 4.2 にまとめておく。なお既存製品の購買間隔をモデル化した Vilcassim and Jain [59] や Gupta [19] および Helsen and Schmittlein [26] の研究も、参考のために併記しておく。表 4.1 では確率とタイミングの定式化について、また表 4.2 では、共変量のダイナミズムとその取り込み方についてまとめてある。

杉田・中村・田島のモデル

最寄り品の普及過程を分析したこの研究は、

$$h_n(t|v_n) = v_n \lambda e^{\beta' X_n(t)} \quad (4.1)$$

表 4.1: 非集計レベルのハザードモデルにおける確率およびタイミングの定式化

	データ	モデルの定式化	
		確率	タイミング
Vilcassim <i>et al.</i>	クラッカー	ノンパラメトリック	セミパラメトリック
Gupta <i>et al.</i>	コーヒー	ガンマ分布	Erlang-2 分布
Helsen <i>et al.</i>	クラッカー	(なし)	Weibull 分布
Sinha <i>et al.</i>	ATM	ロジットモデル	Weibull 分布等
杉田ほか	最寄り品	Bernoulli 分布	指数分布
本研究のモデル	最寄り品	Bernoulli 分布	指数分布

表 4.2: 非集計レベルのハザードモデルにおける共変量の取り込み方

	共変量			
	ダイナミズム	確率	タイミング	確率とタイミング
Vilcassim <i>et al.</i>	×	×	○	
Gupta <i>et al.</i>	○	×	○	
Helsen <i>et al.</i>	○	—	○	
Sinha <i>et al.</i>	×	○	○	別々に説明
杉田ほか	○	×	○	
本研究のモデル	×	○	○	同時に説明

なる条件付きスプリット・ハザード関数によって定式化されている。ここで潜在的な試用購買者であるか否かを表す確率的ダミー変数 v_n は、独立かつ同一に Bernoulli 分布 $Bi(p)$ に従っているので、そのパラメータ p が究極的普及率を表している。

すなわちこのモデルは図 3.6 で表されるとおり、潜在的な試用購買者である確率はア priori に決っており、ダイナミックに変化する共変量 $X_n(t)$ が、試用購買するタイミングだけに影響を与えるモデルである。

このモデルにはふたつの問題点がある。ひとつは潜在的な試用購買者であるか否かという異質性に、共変量の影響がまったく考慮されていないという点である。家計のデモグラフィックスやマーケティング・ミックス等の共変量が試用購買する確率に影響を与えるという事は、多くの実証研究によって立証されている [16]。従って各家計の確率が純粋にランダムに決定されるという仮定は、極めて制約的であるといえる。

もうひとつは共変量がダイナミックに変化するので、推定に時間がかかるという点である。実際モデルのパラメータ推定に要した時間は、EM アルゴリズムを用いた場合は unix ワークステーション上で 20 時間以上、最尤推定を行った場合でもパソコン上で 10 時間以上であり、実用的であるとは言い難い。

Sinha and Chandrashekar のモデル

現金自動支払機 (ATM) の普及過程を分析したこの研究は、

$$h_n(t|v_n) = v_n h_0(t) e^{\beta' X_n} \quad (4.2)$$

なる条件付きスプリット・ハザード関数によって定式化されている。ここで v_n は銀行 n がいつかは ATM を採用する確率を表しており、つぎの二項ロジットモデルによって定義されている。

$$v_n = \frac{1}{1 + e^{\alpha' X_n}} \quad (4.3)$$

従って銀行の総数を N とするとき、 $E[v_n] \simeq \frac{1}{N} \sum_{n=1}^N \hat{v}_n$ が、究極的普及率になっている。

また $h_0(t)$ はベースハザードと呼ばれ、共変量が影響しなかった場合のハザード関数である¹。従って $h_0(t) \exp(\beta' X_n)$ が、共変量の影響を考慮した場合のハザード関数を表している。すなわちこのモデルでは図 3.5 で示されるとおり、ダイナミックに変化しない共変量 X_n が、試用購買確率とタイミングの両方に影響を与えている。

このモデルの問題点は、試用購買確率とタイミングに対する共変量の影響が、独立でしかもアドホックに与えられているという点である。したがってパラメータ α, β の識別性が確保されず、また整合性のとれたインプリケーションも保証されない。

4.2 本研究の目的

ところで既存研究や先験的知識を総合すると、潜在的な試用購買者である確率 v_n を取り込む事によって、「当該新製品を決して試用購買しない家計」の存在を許した方が、現実を良く表していると思われる。また Sinha and Chandrashekaran のところでも述べたように、共変量のダイナミックな変化を v_n に取り込むのは大変困難である。

したがって潜在的な試用購買者である確率を考慮したモデルを実際に定式化するには、次のふたつのアプローチが現実的である。ひとつは杉田・中村・田島のモデルのように、共変量はダイナミックに変化しタイミングに影響を与えるが、確率に影響を与えないアプローチである。そしてもうひとつは Sinha and Chandrashekaran のモデルのように、共変量はダイナミックに変化しないが、確率とタイミングの両方に影響を与えるアプローチである。

そこで本研究では Sinha and Chandrashekaran の研究を踏襲した、非集計レベルの試用購買モデルを提案する。特に彼らのモデルとの違いは、効用理論を利用した期待効用を取り込むことによって、確率とタイミングを同時に説明している点である。このモデルでは確率とタイミングを同時に説明する事によって、整合性のとれたインプリケーションとパラメータの識別性が確保され、さらに推定にかかる時間が短縮される。そしてこのモデル

¹ ここでは $h_0(t)$ に指数分布や Weibull 分布そして対数正規分布を当てはめて、それぞれの比較を行っている。

の定式化はハザード関数によって行い、またパネルデータを用いてパラメータを推定している。なお詳しくは 5 章で説明する。

さらに 6 章では、実証分析によって試用購買する確率やタイミングに対する共変量の影響を測定し、市場導入後 13 週間で究極的普及率の測定を試みる。

Chapter 5

本研究のモデル

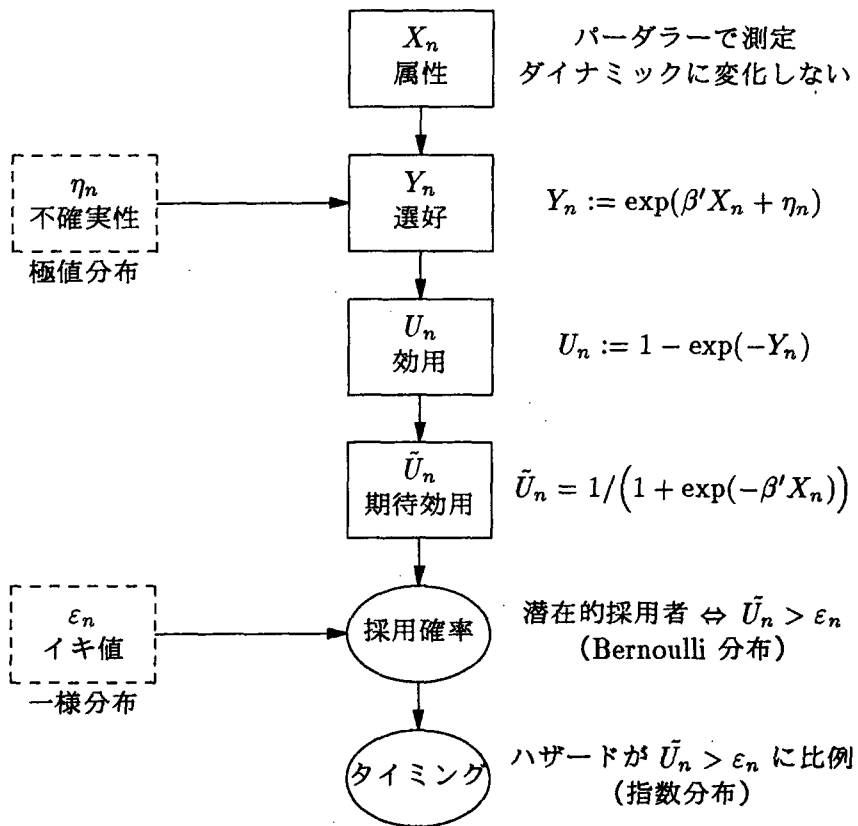
5.1 モデル

本研究のモデルでは、潜在的試用購買者である確率や実際に試用購買するタイミングが、共変量によって説明される。これは基本的に、Sinha and Chandrashekaran の研究を踏襲している。しかし彼らの研究では、確率やタイミングに対する共変量の影響がそれぞれ独立に与えられていたが、本研究では、潜在的試用購買者である確率が高いほど実際に試用購買するタイミングが早いと考え、効用理論に基づく期待効用を用いて確率とタイミングを同時に説明している。それによって整合性のとれたインプリケーションとパラメータの識別性が確保され、推定にかかる時間も短縮される。なおモデルの定式化はハザードモデルを用いて行っている。

5.1.1 期待効用

期待効用は von Neumann-Morgenstern の効用理論 [35] に基づいている。すなわち各家計は共変量を評価して効用を形成するが、情報の欠如や製品固有の不確実性を含んでいるので、期待効用値に従って試用購買の意志決定を行う。

図 5.1: 本研究のモデル



共変量

価格は家計の購買行動に対して大きな影響を与えるが、その取り扱い方法にはふたつのアプローチがある。ひとつは他の共変量と同じく線形に組み込む方法である。これは、Guadagni and Little のブランド選択モデル [17] をはじめとして、多くのマーケティングモデルで採用されている方法である。そしてもうひとつは、共変量を価格で割ることによって、単位価格当たりの値 (per dollar) に変換する方法である。これは Hauser and Gaskin の Defender モデル [22] 等で用いられている。本研究では、家計 n の共変量 X_n を、単位価格当たりの値で測定することにする¹。

不確実性

家計 n が当該新製品を評価する際に、不確実性に基づくリスク η_n を抱く。これは主に情報の欠如や製品固有の変動等によって生じ、潜在的な変数として扱われる。本研究では、 η_n は独立かつ同一に極値分布に従うと仮定する。

$$\Pr(\eta_n < \eta) := 1 - \exp(-e^{-\eta}) \quad (5.1)$$

選好

家計 n の選好 (preference) Y_n は、共変量 X_n と不確実性 η_n によって定義される。特にここでは、Chintagunta [6] や Hanemann [20] の研究と同様に、不確実性が乗法的に作用すると考え、選好を次のように定式化する。

$$Y_n := e^{\beta' X_n + \eta_n} \quad (5.2)$$

¹ 価格を線形に取り込んだモデルは、実証分析の結果、パラメータ推定値が安定しなかった。なお Hauser and Urban [23] では、価格を線形に取り込むアプローチ (the value priority hypothesis) と共変量を単位価格当たりの値で測定するアプローチ (the net value priority hypothesis) の理論的および実証的な比較を耐久消費財に関して行っており、その結果、共変量を単位価格当たりの値で測定するアプローチの方が優れているという結論を得ている。

効用・期待効用

家計 n が当該新製品に対して形成する効用 U_n を、Roberts and Urban [51] と同様に、選好 Y_n を用いて次のように定義する。

$$U_n := 1 - \exp(-Y_n) = 1 - \exp(-e^{\beta'X_n + \eta_n}) \quad (5.3)$$

このとき期待効用 \tilde{U}_n は、簡単な計算によって次のような二項ロジット型になることが分かる²。

$$\tilde{U}_n := E[U_n; \eta_n] = \frac{1}{1 + e^{-\beta'X_n}} \quad (5.4)$$

5.1.2 潜在的試用購買者である確率

本研究の対象である最寄り品市場は、頻繁に購買そして使用（飲食）されているので、各家計はそれぞれ効用のイキ値 ε_n を持っている [65, p.122]。そして新製品を試用購買するか否かは、その期待効用をイキ値と比較する事によって決定される。すなわち $\tilde{U}_n > \varepsilon_n$ なる家計 n は、潜在的試用購買者であると考ええる。

イキ値

上で述べた効用のイキ値 ε_n は、各家計の購買履歴等に大きく依存しており、家計は ε_n に関して異質的である。そこで本研究では ε_n は、家計に関して独立かつ同一に一様分布 $U(0,1)$ に従っていると仮定する。

$$\Pr(\varepsilon_n < \varepsilon) := \varepsilon, \quad 0 \leq \varepsilon \leq 1 \quad (5.5)$$

² 証明は Appendix A の命題 A.3.1 を参照のこと。

潜在的試用購買者である確率

家計 n が潜在的試用購買者であるか否かは、当該新製品の期待効用 \tilde{U}_n をイキ値 ε_n と比較することによって決定される。したがって潜在的試用購買者であるか否かを表す確率的ダミー変数 v_n は、次式によって定義される。

$$v_n := \begin{cases} 1, & \tilde{U}_n \geq \varepsilon_n \\ 0, & \text{otherwise} \end{cases} \quad (5.6)$$

すなわち v_n は、 $\Pr(\varepsilon_n < \tilde{U}_n) = \tilde{U}_n$ をパラメータに持つ Bernoulli 分布 $Bi(\tilde{U}_n)$ に独立に従っている。

5.1.3 試用購買するタイミング

浸透戦略を採用した場合の最寄り品の普及曲線は、指数分布を描く場合が多い [16]。そこで本研究では、 $v_n = 1$ なる潜在的試用購買者の条件付きハザード関数が、 $\tilde{U}_n - \varepsilon_n (> 0)$ に比例すると考える。

$$h_n(t|v_n = 1, \varepsilon_n) := \lambda(\tilde{U}_n - \varepsilon_n) \quad (5.7)$$

ただし $\lambda > 0$ は比例定数である。

5.1.4 ハザード関数

以上の議論により、攪乱項 ε_n と潜在的試用購買者ダミー v_n を与えた下での、家計 n の条件付きスプリット・ハザード関数 $h_n(t|v_n, \varepsilon_n)$ は、次式で表される。

$$h_n(t|v_n, \varepsilon_n) := v_n \lambda (\tilde{U}_n - \varepsilon_n), \quad \lambda > 0 \quad (5.8)$$

5.1.5 究極的普及率

杉田・中村・田島のモデルの v_n は、家計に関して独立かつ同一に Bernoulli 分布に従っているため、そのパラメータ p が究極的普及率を表していた。

しかし本研究のモデルでは、究極的普及率は次式によって計算される。

$$p = E_n[E[v_n; \eta_n]] = E_n[\tilde{U}_n] \simeq \frac{1}{N} \sum_{n=1}^N \tilde{U}_n(X_n) \quad (5.9)$$

すなわち本研究のモデルによって計算された究極的普及率は、マーケティング・ミックス等の共変量の状態がこれからも同じであったと仮定した場合の、究極的普及率を表している。したがって共変量 X_n を変化させることによって、市場導入前に究極的普及率のシミュレーションを行うことが可能である。

5.2 パラメータの推定方法

本研究のモデルでは、家計 n の試用購買時点を表す確率変数 T_n の密度関数 $f_n(t)$ および生存関数 $S_n(t) := 1 - F_n(t)$ は、次のようになる³。

$$f_n(t) = \frac{1 - (1 + \lambda \tilde{U}_n t) e^{-\lambda \tilde{U}_n t}}{\lambda t^2} \quad (5.10)$$

$$S_n(t) = \frac{1 - e^{-\lambda \tilde{U}_n t}}{\lambda t} + 1 - \tilde{U}_n \quad (5.11)$$

したがってその対数尤度関数は

$$\ell(\theta; \vec{T}) := \sum_{n=1}^N \left\{ d_n \log f_n(\hat{T}_n) + (1 - d_n) \log S_n(T_0) \right\}$$

³ 証明は Appendix A の命題 A.3.2 を参照のこと。

$$\begin{aligned}
&= \sum_{n=1}^N \left[d_n \left\{ \log \left(1 + (1 + \lambda \tilde{U}_n \hat{T}_n) e^{-\lambda \tilde{U}_n \hat{T}_n} \right) - \log \lambda - 2 \log \hat{T}_n \right\} \right. \\
&\quad \left. + (1 - d_n) \log \left(\frac{1 - e^{-\lambda \tilde{U}_n T_0}}{\lambda T_0} + 1 - \tilde{U}_n \right) \right] \tag{5.12}
\end{aligned}$$

となるから⁴、最尤法によるパラメータの同時推定が可能である。

⁴ 証明は Appendix A の命題 A.3.3 を参照のこと。

Chapter 6

実証分析

この章では実証分析によって、本研究のモデルの説明力および予測精度等のパフォーマンスを調べる。

4.2 節で述べたとおり、潜在的試用購買者である確率を考慮したモデルを実際に定式化するには、ふたつのアプローチがある。本研究のモデルは、「共変量はダイナミックに変化せず、確率とタイミングの両方に影響を与える」アプローチを採用している。

そこでここでは、もう一方のアプローチ、すなわち「共変量はダイナミックに変化しタイミングに影響を与えるが、確率に影響を与えない」アプローチを採用している杉田・中村・田島のモデルと比較しながら、分析を行うことにする。

6.1 データと共変量

(財) 流通経済研究所が収集した週別スキャンパネルデータを用いる。対象となる新製品は、中堅スナックメーカーの朝食用シリアルである。ただしこのメーカーは、シリアル市場のトップメーカーではない。またこのデータは、関東地方の互いに商圈が異なるスーパーマーケット三店舗で収集されているので、ひとつの家計が複数の店舗を利用することはない。さらにこれらの商圈は閉鎖的なので、他店舗の影響は無視しうる。なお収集期間は 91 年 4 月の第一週から 38 週間である。

6.1.1 家計

データに含まれる家計のうちで、週一回以上シリアルをする購買するものは、全部で 1483 家計存在した。今回は特に、その中から無作為に抽出した 200 家計を分析に用いる。

なお試用購買実績データを表 6.1 に、また分析に用いた家計の購買性向 (HEAVY, LOYAL, TOKU) 別頻度表を、表 6.2, 6.3, 6.4 にまとめておく。なお変数の意味については、6.1.3 項を参照のこと。

また 38 週間を過ぎた時点で、当該新製品を試用購買した家計は 45 世帯、すなわち普及率は 22.5% であった。

6.1.2 サンプル・スプリッティング

試用購買モデルはそのマネジリアルな意義を考えた場合、市場導入後できるだけ早い時点で分析を行いたい。従って今回は、センサリング時点 T_0 を人為的に第 13 週と定め、最初の 13 週間分のデータを用いて各モデルのパラメータを推定する¹。そして残りの 26 週間分のデータをホールドアウトして、予測精度の測定に用いる。

6.1.3 共変量

分析に用いる共変量は全部で 7 つある。最初の 3 つは家計の購買性向を表し、ダイナミックに変化しない。次の 2 つは店舗ダミーで、同じくダイナミックに変化しない。最後の 2 つは価格掛率とエンド陳列ダミーで、ダイナミックに変化する。これらの共変量を本研究のモデルで用いる場合は、バーダラーに変換する必要がある (6.2.1 項参照)。

ヘビーユーザー度 $X_n^1 = HEAVY_n$

家計 n のヘビーユーザー度を表すダミー変数 $X_n^1 = HEAVY_n$ は、シリアルの年間購買金額によって定義される。シリアルの年間総購買金額が家計全体の平均額を越える時は

¹ 究極的普及率は市場導入後 13 週間の普及率の高低で決定されるという実証研究が、いくつか報告されている (相原 [61, p.13]、流通経済研究所 [90, p.39])。

表 6.1: 店舖別試用購買実績

t	store a			store b			store c		
	$KAKE(t)$	$ISLE(t)$	$N(t)$	$KAKE(t)$	$ISLE(t)$	$N(t)$	$KAKE(t)$	$ISLE(t)$	$N(t)$
1	0.93			1.00			0.75	1	
2	0.87		1	1.00			0.75	1	2
3	0.87		2	1.00			0.75		
4	0.87		1	1.00			0.75		
5	0.87		2	0.88	1	1	0.75		
6	0.91			0.87	1	1	0.75		
7	1.00			1.00			0.75		
8	1.00			0.92			0.75		
9	1.00			0.87	1	3	0.75		
10	0.89		1	0.87		1	0.75		
11	0.87			0.87			0.75		
12	0.87	1	5	0.87		1	0.75		
13	0.87		1	1.00			0.75		
14	0.87			1.00			0.75		
15	0.98			1.00			0.93		
16	1.00			1.00		2	1.00		
17	1.00			0.88			1.00		
18	0.78	1	3	0.87			1.00		
19	0.78	1	2	0.87			0.86		
20	0.87			0.86			0.80		
21	0.87			0.87			0.80		
22	0.87			1.00			0.80		
23	0.88	1	3	1.00			0.81		1
24	1.00			1.00			1.00		
25	1.00			1.00			1.00	1	
26	1.00		1	0.95			1.00	1	
27	0.80	1	4	1.00			1.00	1	
28	0.86	1	1	0.99			1.00	1	
29	1.00			1.00			1.00	1	
30	0.88			1.00			1.00		
31	0.87	1		1.00			1.00		
32	0.87	1	2	1.00			1.00		1
33	0.87	1		0.99			1.00		
34	0.87	1	2	1.00			1.00		
35	0.87	1		1.00			1.00		
36	0.89	1		1.00			1.00		
37	1.00		1	1.00			1.00		
38	1.00			1.00			1.00		

表 6.2: HEAVY 対 LOYAL の頻度表

Heavy	Loyal			Total
	-1	0	1	
0	41	48	30	119
1	53	24	4	81
Total	94	72	34	200

表 6.3: LOYAL 対 TOKU の頻度表

Loyal	Toku		Total
	0	1	
-1	61	33	94
0	41	31	72
1	18	16	34
Total	120	80	200

表 6.4: TOKU 対 HEAVY の頻度表

Toku	Heavy		Total
	0	1	
0	65	55	72
1	54	26	34
Total	119	81	200

$HEAVY_n = 1$ 、そうでないときは $HEAVY_n = 0$ とする。

一般に、購買量が多いほど試用購買する確率が高いと言われる [55, p.540]。従ってヘビーユーザーの影響は正であると期待される。

ロイヤルユーザー度 $X_n^2 = LOYAL_n$

家計 n のロイヤルユーザー度を表すダミー変数 $X_n^2 = LOYAL_n$ は、ロイヤルブランド（最も多く購買しているブランド）への依存度を表している。シリアルの年間総購買金額に占めるロイヤルブランドの割合が 0.7 以上、0.4 ~ 0.7、0.4 以下の時、それぞれ $LOYAL_n = 1, 0, -1$ とする。

一般に、ロイヤルブランドへの依存度が高いほど試用購買する確率が低いと言われる [55, p.540]。従ってロイヤルユーザーの影響は負であると期待される。

特売反応度 $X_n^3 = TOKU_n$

家計 n の特売反応度を表すダミー変数 $X_n^3 = TOKU_n$ は、シリアルの特売時の年間総購買金額が家計全体の平均額以上の時は 1、そうでないときは 0 と定義される。

一般に、特売に対する感度が高いほど試用購買する確率が低いと言われる [55, p.540]。従って特売反応度の影響は負であると期待される。

店舗ダミー $X_n^4 = STORE_n^a, X_n^5 = STORE_n^b$

店舗間の異質性を考慮するため、家計 n が利用する店舗に関して、店舗ダミーを表 6.5 のように定義する。これらは一種のストア・ロイヤルティと考えることができる。

エンド陳列ダミー $ISLE_n(t)$

これは家計 n が利用する店舗における、当該新製品の第 t 週のエンド陳列ダミーである。なお各店舗における当該新製品の導入週を原点としている。

表 6.5: 店舗ダミーの定義

店舗	店舗ダミー	
	$STORE^a$	$STORE^b$
a	1	0
b	0	1
c	0	0

価格掛率 $KAKE_n(t)$

これは家計 n が利用する店舗における、当該新製品の第 t 週の価格掛率（＝販売価格／通常価格）である。エンド陳列ダミーと同様に、各店舗における当該新製品の導入週を原点としている。

6.2 分析計画

6.2.1 モデル

この章の冒頭でも述べたとおり、共変量がダイナミックに変化するか否かと、潜在的試用購買者である確率に対して共変量が影響を与えるか否かというふたつの点に関して、本研究のモデルと杉田・中村・田島のモデルは対照的である。従って比較の対象として、以下で述べるふたつのナルモデルの他に、杉田・中村・田島のモデルも採用する。

本研究のモデル（モデル1）

これ以降簡単のため、本研究のモデルを「モデル1」と呼ぶことにする。

1. 推定すべきパラメータ

モデル1は5.8式で定義されている。ここで $\lambda > 0$ なる制約条件が課せられているので、 λ を直接推定するには、制約付き最尤推定法を用いなければならない。そこで

実証分析の際には $\lambda = e^{-a}$ と変換して、 λ の代わりに a を推定する。また共変量は、パーダラーで測定されている。

$$h_n(t|v_n, \varepsilon_n) := v_n e^{-a(\tilde{U}_n - \varepsilon_n)}, \quad v_n \sim Bi(\tilde{U}_n) \quad (6.1)$$

$$\tilde{U}_n = \frac{1}{1 + e^{-\beta' X_n}} \quad (6.2)$$

$$\beta' X_n := \beta_0 + \sum_{k=1}^5 \beta_k \frac{X_n^k}{KAKEN} + \beta_6 \frac{ISLE_n}{KAKEN} \quad (6.3)$$

なおモデル 1 で推定すべきパラメータは、 $a, \beta_0, \beta_1, \dots, \beta_6$ の 8 個である。

2. ダイナミックに変化しない価格掛率とエンド陳列ダミー

モデル 1 の共変量はダイナミックに変化しないので、価格掛率 $KAKEN$ とエンド陳列ダミー $ISLE_n$ を次のように定義する²。

- センサリング時点 $T_0 = 13$ 週以前に試用購買を行った家計に対しては、試用購買した時の値を採用する。
- センサリングされた家計に対しては、13 週間の平均値を採用する。

3. 究極的普及率

究極的普及率 p に関して 5.9 式が成立するので、実際には次のように計算される。

$$\hat{p} = \sum_{n=1}^N \hat{U}_n = \sum_{n=1}^N \frac{1}{1 + \exp(-\hat{\beta}' X_n)}, \quad N = 200 \quad (6.4)$$

杉田・中村・田島のモデル (モデル 0)

これ以降簡単のため、杉田・中村・田島のモデルを「モデル 0」と呼ぶことにする。

² Sinha and Chandrashekar の研究では、センサリングされない家計については試用購買時期までの平均値を採用している [56, p.125]。本研究でもいくつかの変換方法を検討したが、本文中で述べた方法を採用した場合が、最もパラメータ推定値が安定した。

モデル0は3.47式で定義されている。ここで $\lambda > 0$ なる条件が課せられているので、モデル1と同様に $\lambda = e^{-a}$ と変換する。また v_n が従うBernoulli分布のパラメータ p に対しても、 $0 < p < 1$ なる制約条件が課せられているので、 $p = 1 - \exp(-e^q)$ と変換する。

$$h_n(t|v_n) := v_n e^{-a} e^{\beta' X_n(t)}, \quad v_n \sim Bi(1 - \exp(-e^q)) \quad (6.5)$$

$$\beta' X_n(t) := \sum_{k=1}^5 \beta_k X_n^k + \beta_6 ISLE_n(t) + \beta_7 KAKE_n(t) \quad (6.6)$$

なおこのモデルで推定すべきパラメータは、 $a, q, \beta_1, \dots, \beta_7$ の9個である³。

ナルモデル1

これはモデル1において $KAKE_n \equiv 1$ としたモデルである。

$$\beta' X_n := \beta_0 + \sum_{k=1}^5 \beta_k X_n^k + \beta_6 ISLE_n \quad (6.7)$$

モデル1の共変量はバーダラーで測定されているので、価格の効果を直接知ることができない。したがってこのナルモデル1と比較することによって、価格の効果を検証する。

なおナルモデル1で推定すべきパラメータは、モデル1と同じく $a, \beta_0, \beta_1, \dots, \beta_6$ の8個である。

ナルモデル0

これはモデル1において共変量の影響を全く考慮しないモデルである。

$$\tilde{U}_n \equiv \beta_0 \quad (6.8)$$

なおナルモデル0で推定すべきパラメータは、 a, β_0 の2個である。

³ パラメータ a との識別性を確保するため、 $\beta' X_n(t)$ は定数項を持たない。

6.2.2 パラメータの推定方法

上で述べた4つのモデルのパラメータはそれぞれ、最尤法によって同時推定される。尤度最適化のアルゴリズムは、Newton-Raphson法を採用する。そして最適化計算の初期値は、 $[0, 1]$ 上 (β_7 については $[-1, 0]$ 上) の一様乱数とする。

なお実際の計算は、i486 CPU を搭載したパーソナル・コンピュータ上で動く汎用統計パッケージ SAS に含まれる非線形最適化プロシジャ NLP [21, 67] を用いて行なう。

6.2.3 説明力の検証

説明力の検証は、共変量や価格の効果、そして共変量のダイナミックな変化をそれぞれ取り込んだ場合について行う。

共変量の取り込み

モデル1とナルモデル0を比較することによって、共変量を取り込む事による効果を検証する。なおこれらふたつのモデルはネストしているので、尤度比検定 [3, p.166] によって比較を行う。

価格の効果

前述の通り、モデル1の共変量はバーダラーで測定されているので、価格の効果を直接知ることができない。したがってナルモデル1と比較することによって、価格の効果を検証する。ただしモデル1とナルモデル1はネストしていないので、赤池情報量規準 (AIC: Akaike's Information Criterion)

$$AIC := -2 (\text{最大対数尤度} - \text{自由パラメータ数}) \quad (6.9)$$

を用いて比較する。なお AIC の差が 1 未満の時は、どちらのモデルの良さも同程度であることを意味する。またその差が 1~2 程度以上ならば、AIC が小さいモデルの方が有意

に優れていると考えられる [70, p.63]。

共変量のダイナミックな変化

モデル1の共変量はダイナミックに変化しないが、モデル0の共変量はダイナミックに変化する。そこでモデル1とモデル0を比較することによって、ダイナミックに変化する共変量の効果を検証する。

ただしモデル1とモデル0はネストしていないので、その比較はAICを用いて行うことにする。

6.2.4 予測精度の比較

13時点までのデータを用いて推定されたパラメータを用いて、14～34時点までの普及曲線と究極的普及率を予測させることによって、モデル1とモデル0の外挿予測精度を比較する。

新規試用購買率

普及曲線（累積試用購買率）

$$p(t) := \text{時点 } t \text{ までに試用購買する家計の割合} \quad (6.10)$$

は自分自身の過去の値の影響を受けるので、厳密な意味での外挿予測を行うことができない。そこで今回は普及曲線の差分である新規試用購買率

$$dp(t) := \text{時点 } t \text{ に試用購買する家計の割合} \quad (6.11)$$

を用いて外挿予測精度を測定する。

具体的には、 $dp(t)$ の予測値である

$$\hat{d}_p(t) := \frac{1}{N} \sum_{n=1}^N f_n(t|\hat{\theta}), \quad N = 200 \quad (6.12)$$

と実測値

$$\bar{d}_p(t) := \frac{1}{N} (\text{時点 } t \text{ に試用購買した家計数}), \quad N = 200 \quad (6.13)$$

を、ホールドアウト期間である 14 ~ 38 時点に渡って計算し、それを平均自乗誤差 (MSE: Mean Square Error)

$$MSE := \frac{1}{25} \sum_{t=14}^{38} (\hat{d}_p(t) - \bar{d}_p(t))^2 \quad (6.14)$$

や平均絶対誤差 (MAE: Mean Absolute Error)

$$MAE := \frac{1}{25} \sum_{t=14}^{38} |\hat{d}_p(t) - \bar{d}_p(t)| \quad (6.15)$$

によって比較する。

究極的普及率

モデル 1 の究極的普及率は、6.4 式によって計算される。またモデル 0 の究極的普及率は、Bernoulli 分布のパラメータとして推定される。

また 6.1 節の冒頭で述べたとおり、38 週間を過ぎた時点での実際の普及率は 22.5% であった。したがって両モデルの究極的普及率とこの値を比較することによって、究極的普及率に関する予測精度を測定する。

6.3 結論

6.3.1 パラメータ推定値

推定結果

モデル1 およびモデル0 の推定結果を、表 6.6 にまとめておく。なお括弧内の値は、パラメータ推定値の t 値である。またモデル1 の究極的普及率 p の t 値は、正確には $p = 1 - \exp(-e^q)$ なる q の t 値である。

モデル1、モデル0 は構造が全く異なるにもかかわらず、パラメータの符号が全て一致した。しかもナルモデル0 と比較した場合に、モデル1 の尤度の改善度が良好であった(6.3.2 項参照)。したがってモデル1 に致命的な miss-specification が発生していないと解釈することができる。

なおパラメータ推定に要した CPU 時間は、モデル0 が約 10 時間、そのほかのモデルが 15 分以下であった。

パラメータ推定値の解釈

1. ヘビーユーザー一度

パラメータの符号は負で、しかも 5% 水準で有意になっている。これは「ヘビーユーザーであるほど、当該新製品を試用購買する確率が低い」事を意味しており、先験的な知識と矛盾しているように思われる。

しかしシリアル市場は、トップメーカーの金額シェアが約 50% を占め、しかもヘビーユーザーは主にトップメーカーのブランドを購入している [89, p.128]。したがってこの結果は実際の市場特性と矛盾していない。

2. ロイヤルユーザー一度

パラメータの符号は正であり、これは「ロイヤルユーザーであるほど、当該新製品を試用購買する確率が高い」事を意味している。しかしモデル1 は 10% 水準で有意、

表 6.6: 推定結果

パラメータ	モデル	
	モデル 1	モデル 0
a	1.63 (4.35)	36.14 (1.28)
β_0	-4.81 (-2.63)	
ヘビーユーザー度	-2.22 (-2.01)	-1.45 (-2.25)
ロイヤルユーザー度	0.47 (1.41)	0.36 (1.23)
特売性向	0.44 (0.90)	0.78 (1.79)
店舗 a ダミー	3.49 (2.19)	9.08 (1.97)
店舗 b ダミー	3.33 (2.08)	9.41 (2.01)
エンド陳列ダミー	4.81 (3.36)	1.53 (3.27)
価格掛率		-57.00 (-1.52)
究極的普及率	0.189	1.000 (0.02)
対数尤度	-98.69	-98.00
AIC	213.38	214.00
10000*MSE	0.491	0.560
100*MAE	0.504	0.456

(括弧内は t 値)

またモデル0の方は有意にならなかった。

3. 特売感度

パラメータの符号は正で、これは先験的知識と合致している。モデル0は5%水準で有意、またモデル1は有意にならなかった。

4. 店舗ダミー

$STORE^a$, $STORE^b$ ともに、パラメータの符号が正になっている。しかもどのパラメータも5%水準で有意になっている。

これは、店舗 a,b が家計の効用の形成に与える影響（ストア・ロイヤルティ）が、店舗 c のそれに比べて有意に大きいことを表している。すなわち、店舗 a,b の家計は店舗 c の家計に比べて、もともと試用購買する確率が高いか、または他の共変量の正の影響にヨリ敏感であることを意味する。じっさい原データをみる限り、この推論は正しいと思われる。

5. 価格掛率

モデル0のパラメータの符号は負で、10%水準で有意になっている。これは先験的知識と合致している。

なおモデル1の共変量はバーダラーで測定されているので、価格のパラメータは存在しない。

6. エンド陳列

パラメータの符号は正で、これは先験的知識と合致している。さらに、どちらのモデルも1%で有意になっている。したがって当該新製品の試用購買を促進させるには、エンド陳列が有効であるといえる。

6.3.2 説明力

ここでは 6.2.3 で説明した手順に従って、モデルの説明力の検証を行う。モデル 1,2、およびナルモデル 0,1 の対数尤度と AIC を、表 6.7 にまとめておく。

表 6.7: 説明力

モデル	指標	
	対数尤度	AIC
モデル 1	-98.69	213.38
モデル 0	-98.00	214.00
ナルモデル 1	-101.20	220.40
ナルモデル 0	-126.07	256.14

共変量の取り込み

まずモデル 1 とナルモデル 0 の尤度比検定を行った結果、共変量を取り込む事によってモデルの当てはまりは著しく向上した。

$$-2(\ell_{null0} - \ell_{model0}) = 54.76 > 22.46 = \chi^2(df. = 6, p = 0.001) \quad (6.16)$$

価格の効果

つぎにモデル 1 における価格の効果を検証する。モデル 1 の共変量はパーダラーで測定されているので、価格掛率を一定としたナルモデル 1 と比較することによって、間接的に価格の効果を検証する。

$$|AIC_{model0} - AIC_{null1}| = 7.02 \quad (6.17)$$

このときモデル 1 とナルモデル 1 はネストしていないので、それぞれの AIC を計算した結果、モデル 1 が有意に優れていることが分かった。これは先験的知識や、モデル 0 において価格掛率のパラメータが 10% 水準で有意になっている事と合致する。

共変量のダイナミクス

最後にモデル1とモデル0の比較を行う。

$$|AIC_{model0} - AIC_{model2}| = 0.62 \quad (6.18)$$

このふたつのモデルはネストしていないので、それぞれのAICを計算した結果、その差が認められなかった。

6.3.3 予測精度

新規試用購買率

1～13週までのデータによって推定されたパラメータを用いて、14～38週の新規試用購買率(6.12式)を求めた。そのMSE, MAEが、表6.6の最後の二行である。

これによるとMSEではモデル1の方が優れており、またMAEではモデル0の方が優れている。したがって両モデルは、新規試用購買率の外挿予測精度に関してはあまり変わらないと考えられる。

究極的普及率

38週間を過ぎた時点での当該新製品の実際の普及率は22.5%であったが、モデル1から得られる究極的普及率は18.9%であった。モデル1の究極的普及率は6.4によって計算され、これはマーケティング・ミックス等の共変量および共変量に対する家計の反応が、これからも同じであったと仮定した場合の値である。したがってモデル1の究極的普及率が実測値を下回った理由は、センサリング時点である13週以降のプロモーションがそれ以前よりも消極的であるか、またはプロモーションに対する家計の反応が鈍化したためだと考えられる。じっさい店舗aでは14週以降に、それ以前よりも積極的にプロモーション(特にエンド陳列)を行っているが、13週間毎の累積試用購買者数はむしろ減少しており、この推論は正しいと思われる。

図 6.1: 予測結果 (新規試用購買率)

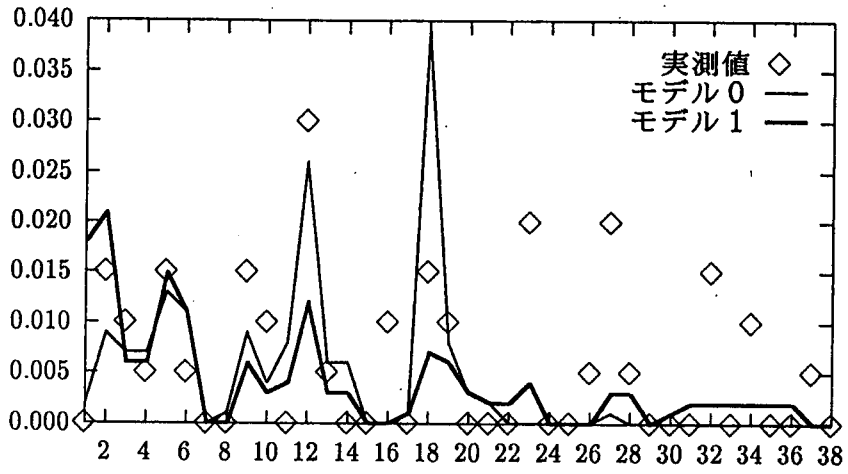
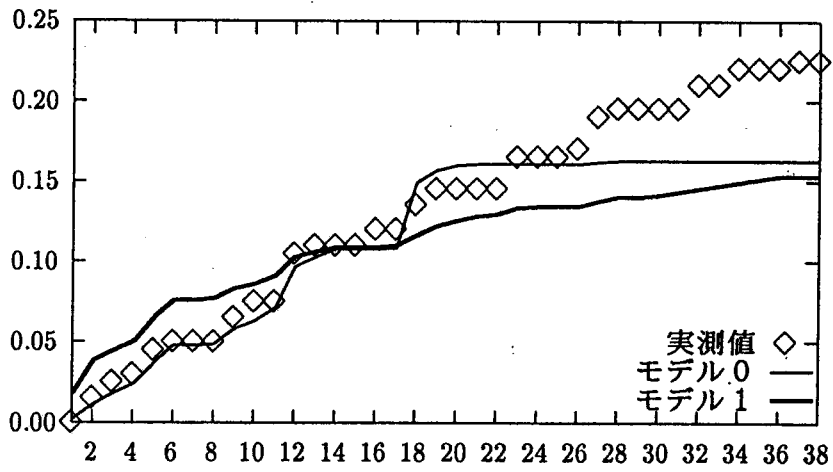


図 6.2: 予測結果 (累積試用購買率)



またモデル0から得られる究極的普及率は100%であった。モデル0の究極的普及率はBernoulli分布のパラメータであり、あらかじめ独立に決まっている。したがってこの値は、「製品ライフサイクルや家計の寿命を考慮しなければ、全ての家計がいつかは試用購買する」ということを意味しており、それなりに妥当性があると思われる。

6.3.4 まとめ

実証分析から得られた結果（表6.8）に基づき、モデル1（本研究のモデル）とモデル0（杉田・中村・田島のモデル）の比較をまとめておく。

表 6.8: モデル1 とモデル0 の比較

	モデル	
	モデル1	モデル0
共変量のダイナミクス	ない	ある
究極的普及率が共変量に依存するか	依存する	依存しない
推定時間（パソコン）	15分	10時間
説明力	あまり変わらない	
新規試用購買率の予測精度	あまり変わらない	

モデル0は図3.6で示されるとおり、共変量がダイナミックに変化するモデルである。しかしそのため密度関数等の計算が複雑になり、パラメータを推定するのに時間が非常にかかる。また共変量が潜在的試用購買者である確率に影響しないので、究極的普及率の解釈が困難な場合がある。

一方モデル1は図5.1で示されるとおり、共変量がダイナミックに変化しないが、確率とタイミングを同時に説明するモデルである。共変量が時間に依存しないので計算コストが低く、しかも予測精度等のパフォーマンスは、モデル0とあまり変わらなかった。さらに究極的普及率の解釈も自然である。

したがってモデル1は、理論的な構造のみならず、実用性の観点からも優れているといえよう。

Chapter 7

まとめと今後の課題

本研究では最寄り品の新製品の試用購買を対象にして、潜在的試用購買者である確率と実際に試用購買するタイミングの両方を考慮した非集計レベルのモデルを提案した。

ここで Sinha and Chandrashekar の研究は、ダイナミックに変化しない共変量が確率とタイミングの両方に影響を与えていたが、確率とタイミングに対する共変量の影響をそれぞれ独立に定式化していた。しかし本研究では彼らの研究を踏襲しながらも、von Neumann and Morgenstern の効用理論を利用した期待効用を取り込む事によって、確率とタイミングを同時に説明している。それによって、パラメータの識別性と整合性のとれたインプリケーションが確保され、推定に要する時間も短縮された。

更にこのモデルは、ハザード関数によって定式化されている。試用購買モデルをハザード関数によって定式化する利点はふたつある。ひとつはハザード関数の数学的制約条件は非負性のみなので、柔軟に定式化できる点である。そしてもうひとつは、試用購買データのようにセンサリングが発生しているデータを分析しても、回帰モデルとは異なり、パラメータ推定値にバイアスがかからないという事である。

朝食用シリアルのデータを用いた実証分析では、共変量がダイナミックに変化しながらタイミングに影響を与えるが、確率には影響を与えない杉田・中村・田島のモデルとの比較を行った。その結果、計算コストが極めて低いにもかかわらず、予測精度等のパフォーマンスはあまり変わらなかった。したがって本研究のモデルは、理論的な構造のみならず、

実用性の観点からも優れているといえよう。

最後に今後の課題をふたつ挙げておく。ひとつは、より本格的な実証分析の必要性である。今回の実証分析で得られた結果は、あくまでも朝食用シリアルに限定されている。従ってより普遍的な結果を得るためにも、カテゴリー横断的な分析を行う必要がある。また今回の実証分析では、共変量として価格とエンド陳列ダミー、および家計の購買性向を採用した。しかしこれ以外にも広告 [34]、競合関係 [16]、店内露出 [92]、家庭内在庫 [57]、配荷量 [53] をはじめとして、試用購買に影響する変数は数多く存在するので、これらの有効性を吟味することも必要である。さらに本モデルの定式化で用いられる変換関数や確率変数の分布について、他の可能性も吟味する必要がある。

もうひとつの課題は、極めて本質的である。本モデルは Sinha and Chandrashekar の研究を踏襲し、共変量は試用購買確率にも影響を与えているので、共変量はダイナミックに変化していないと仮定している。そのため本来はダイナミックに変化する価格やエンド陳列ダミーも、ダイナミックに変化しない形に変換しなければならなかった。しかしこの仮定は、先験的にも非常に制約的であり、ダイナミックに変化する共変量が確率に影響を与えるモデルを開発することが、今後の重要な課題である。

Appendix A

命題の証明

ここでは、本論文中に現れた基本的な命題について、その証明をつける。ただし各命題で用いられる記号の意味は、それぞれ次の通りである。

表 A.1: 記号の凡例

記号	凡例
T_0	センサリング（観測の打ち切り）時点
n	家計を表す添字 ($n = 1, 2, \dots, N$)
d_n	家計 n が T_0 以前に試用購買したか否かを表すダミー変数
T_n	家計 n の試用購買時点を表す非負の確率変数
$f_n(t)$	T_n の密度関数
$F_n(t)$	T_n の分布関数
$S_n(t)$	T_n の生存関数 $S_n(t) := 1 - F_n(t)$
$h_n(t)$	T_n のハザード関数
$H_n(t)$	T_n の累積ハザード関数 $H_n(t) := \int_0^t h_n(\tau) d\tau$
v_n	家計 n が潜在的試用購買者であるか否かを表す確率的ダミー変数

A.1 ハザード関数

ハザード関数 $h_n(t)$ の定義が

$$h_n(t) := \lim_{\delta \downarrow 0} \frac{\Pr(t \leq T_n < t + \delta | t \leq T_n)}{\delta}$$

であることに注意する。

命題 A.1.1 ハザード関数 $h_n(t)$ と密度関数 $f_n(t)$ は、次のように一対一に対応する。

$$\begin{aligned} h_n(t) &= \frac{f_n(t)}{1 - F_n(t)} = \frac{f_n(t)}{S_n(t)} \\ f_n(t) &= h_n(t)e^{-H_n(t)} \end{aligned}$$

(証明) まず条件付き確率の定義から

$$\begin{aligned} h_n(t) &= \lim_{\delta \downarrow 0} \frac{\Pr(t \leq T_n < t + \delta | t \leq T_n)}{\delta} \\ &= \lim_{\delta \downarrow 0} \frac{\Pr(t \leq T_n < t + \delta)}{\delta \Pr(t \leq T_n)} \\ &= \lim_{\delta \downarrow 0} \frac{\Pr(T_n \leq t + \delta) - \Pr(T_n \leq t)}{\delta \Pr(t \leq T_n)} \\ &= \frac{1}{S_n(t)} \lim_{\delta \downarrow 0} \frac{F_n(t + \delta) - F_n(t)}{\delta} = \frac{f_n(t)}{S_n(t)} \end{aligned}$$

が成立する。さらに

$$h_n(t) = \frac{f_n(t)}{S_n(t)} = -\frac{S'_n(t)}{S_n(t)} = -\frac{d}{dt} \log S_n(t)$$

となるから

$$\log S_n(t) = \int_0^t \frac{d}{d\tau} \log S_n(\tau) d\tau = -\int_0^t h_n(\tau) d\tau = -H_n(t)$$

従って

$$S_n(t) = e^{-H_n(t)}$$

すなわち

$$f_n(t) = h_n(t)S_n(t) = h_n(t)e^{-H_n(t)}$$

が成立する。

命題 A.1.2 ハザードモデル (2.3) の尤度関数 $L(\theta; \vec{T})$ および対数尤度関数 $\ell(\theta; \vec{T})$ は、それぞれ次のようになる。

$$L(\theta; \vec{T}) = \prod_{n=1}^N \left(h_n(\hat{T}_n | \theta) e^{-H_n(\hat{T}_n | \theta)} \right)^{d_n} e^{-H_n(T_0 | \theta) (1-d_n)} \quad (\text{A.1})$$

$$\ell(\theta; \vec{T}) = \sum_{n=1}^N \left\{ d_n \left(\log h_n(\hat{T}_n | \theta) - \int_0^{\hat{T}_n} h_n(t | \theta) dt \right) - (1-d_n) \int_0^{T_0} h_n(t | \theta) dt \right\} \quad (\text{A.2})$$

(証明) (A.1) 式の成立は、尤度関数の定義から自明であるから、(A.2) 式を証明する。対数尤度関数の定義により、

$$\begin{aligned} \ell(\theta; \vec{T}) &:= \log L(\theta; \vec{T}) \\ &= \log \prod_{n=1}^N f_n(\hat{T}_n | \theta)^{d_n} S_n(T_0 | \theta)^{1-d_n} \\ &= \sum_{n=1}^N \left\{ d_n \log f_n(\hat{T}_n | \theta) + (1-d_n) \log S_n(T_0 | \theta) \right\} \\ &= \sum_{n=1}^N \left\{ d_n \log \left(h_n(\hat{T}_n | \theta) e^{-H_n(\hat{T}_n | \theta)} \right) + (1-d_n) \log e^{-H_n(T_0 | \theta)} \right\} \\ &= \sum_{n=1}^N \left\{ d_n \left(\log h_n(\hat{T}_n | \theta) - H_n(\hat{T}_n | \theta) \right) - (1-d_n) H_n(T_0 | \theta) \right\} \end{aligned}$$

$$= \sum_{n=1}^N \left\{ d_n \left(\log h_n(\hat{T}_n | \theta) - \int_0^{\hat{T}_n} h_n(t | \theta) dt \right) - (1 - d_n) \int_0^{T_0} h_n(t | \theta) dt \right\}$$

A.2 杉田・中村・田島のモデル

杉田・中村・田島のモデル (モデル 0) の条件付きスプリット・ハザード関数 $h_n(t|v_n)$ が

$$h_n(t|v_n) := v_n \lambda e^{U_n(t)}, \quad U_n(t) := \beta' X_n(t)$$

であることに注意する。ここで潜在的試用購買者であるか否かを表す確率的ダミー変数 v_n は、 p をパラメータに持つ Bernoulli 分布 $Bi(p)$ に、独立かつ同一に従っている。

$$v_n \sim Bi(p), \quad (i.i.d.)$$

命題 A.2.1 密度関数 $f_n(t)$ と生存関数 $S_n(t)$ は、それぞれ次の通りである。

$$f_n(t) = p \lambda e^{U_n(t) - H_n(t|v_n=1)}$$

$$S_n(t) = 1 - p \left(1 - e^{-H_n(t|v_n=1)} \right)$$

(注意) ダイナミックに変化する共変量 $X_n(t)$ は、階段関数であると仮定している (6.1.3 項参照)。したがって条件付き累積ハザード関数 $H_n(t|v_n=1)$ は、次のように近似される。

$$\begin{aligned} H_n(t|v_n=1) &:= \int_0^t h_n(\tau|v_n=1) d\tau \\ &= \lambda \int_0^t e^{U_n(\tau)} d\tau \simeq \lambda \sum_{s=1}^t e^{U_n(s)} \end{aligned}$$

(証明) まず

$$F_n(t|v_n = 1) = 1 - S_n(t|v_n = 1) = 1 - e^{-H_n(t|v_n=1)}$$

である。従って $F_n(t|v_n = 0) = 0$ に注意すると

$$\begin{aligned} F_n(t) &= F_n(t|v_n = 1) \Pr(v_n = 1) + F_n(t|v_n = 0) \Pr(v_n = 0) \\ &= p(1 - e^{-H_n(t|v_n=1)}) \end{aligned}$$

すなわち

$$S_n(t) = 1 - F_n(t) = 1 - p(1 - e^{-H_n(t|v_n=1)})$$

が成立する。さらに

$$\begin{aligned} f_n(t) &= \frac{d}{dt} F_n(t) \\ &= -p \frac{d}{dt} e^{-H_n(t|v_n=1)} \\ &= p h_n(t|v_n = 1) e^{-H_n(t|v_n=1)} \\ &= p \lambda e^{U_n(t)} e^{-H_n(t|v_n=1)} = p \lambda e^{U_n(t) - H_n(t|v_n=1)} \end{aligned}$$

である。

A.3 本研究のモデル

本研究のモデル (モデル 1) の条件付きスプリット・ハザード関数 $h_n(t|v_n, \varepsilon_n)$ が

$$h_n(t|v_n, \varepsilon_n) := v_n \lambda(\tilde{U}_n - \varepsilon_n), \quad \tilde{U}_n := E[U_n; \eta_n]$$

であることに注意する。ただし v_n はそれぞれ、 $\Pr(\tilde{U}_n \geq \varepsilon_n) = \tilde{U}_n$ をパラメータを持つ Bernoulli 分布 $Bi(\tilde{U}_n)$ に、独立に従っている。

$$v_n \sim Bi(\tilde{U}_n)$$

命題 A.3.1 期待効用 \tilde{U}_n は、次の二項ロジットモデルに帰着する。

$$\tilde{U}_n = \frac{1}{1 + e^{-\beta' X_n}}$$

(証明) 仮定により、不確実性を表す確率変数 η_n は、極値分布に独立かつ同一に従っている。したがって η_n の密度関数 $p(\eta)$ は、次式で定義される。

$$p(\eta) := e^\eta \exp(-e^\eta)$$

このとき

$$\begin{aligned} \tilde{U}_n &:= E[U_n; \eta_n] \\ &= \int_{-\infty}^{\infty} U_n(\eta_n) p(\eta_n) d\eta_n \\ &= \int_{-\infty}^{\infty} \left(1 - \exp(-e^{\beta' X_n + \eta_n})\right) e^{\eta_n} \exp(-e^{\eta_n}) d\eta_n \\ &= 1 - \int_{-\infty}^{\infty} \exp\left(-\left(1 + e^{\beta' X_n}\right) e^{\eta_n}\right) e^{\eta_n} d\eta_n \\ &= 1 - \int_0^{\infty} \exp\left(-\left(1 + e^{\beta' X_n}\right) \delta_n\right) d\delta_n, \quad \delta_n := e^{\eta_n} \\ &= 1 - \int_0^{\infty} e^{-\alpha_n \delta_n} d\delta_n, \quad \alpha_n := 1 + e^{\beta' X_n} \\ &= 1 + \frac{1}{\alpha_n} e^{-\alpha_n \delta_n} \Big|_{\alpha_n=0}^{\infty} = 1 - \frac{1}{\alpha_n} = \frac{e^{\beta' X_n}}{1 + e^{\beta' X_n}} = \frac{1}{1 + e^{-\beta' X_n}} \quad \blacksquare \end{aligned}$$

命題 A.3.2 密度関数 $f_n(t)$ と生存関数 $S_n(t)$ は、それぞれ次の通りである。

$$f_n(t) = \frac{1 - (1 + \lambda \tilde{U}_n t) e^{-\lambda \tilde{U}_n t}}{\lambda t^2}$$

$$S_n(t) = \frac{1 - e^{-\lambda \tilde{U}_n t}}{\lambda t} + (1 - \tilde{U}_n)$$

(証明) 攪乱項 ε_n は仮定により、一様分布 $U(0, 1)$ に独立かつ同一に従っている。このとき密度関数を $p(\varepsilon)$ とすると

$$p(\varepsilon|v_n = 1) = \begin{cases} 1/\tilde{U}_n, & 0 \leq \varepsilon < \tilde{U}_n \\ 0, & \tilde{U}_n \leq \varepsilon \leq 1 \end{cases}$$

であることに注意する。実際 $0 \leq \varepsilon < \tilde{U}_n$ に対して

$$\begin{aligned} \Pr(\varepsilon_n < \varepsilon | v_n = 1) &= \Pr(\varepsilon_n < \varepsilon | \tilde{U}_n > \varepsilon_n) \\ &= \frac{\Pr(\varepsilon_n < \varepsilon, \varepsilon_n < \tilde{U}_n)}{\Pr(\varepsilon_n < \tilde{U}_n)} \\ &= \frac{\Pr(\varepsilon_n < \varepsilon)}{\Pr(\varepsilon_n < \tilde{U}_n)} = \frac{\varepsilon}{\tilde{U}_n} \end{aligned}$$

だから

$$p(\varepsilon|v_n = 1) = \frac{d}{d\varepsilon} \Pr(\varepsilon_n < \varepsilon | v_n = 1) = \frac{1}{\tilde{U}_n}, \quad 0 \leq \varepsilon < \tilde{U}_n$$

さらに $\int_0^{\tilde{U}_n} \Pr(\varepsilon_n < \varepsilon | v_n = 1) d\varepsilon = 1$ だから

$$p(\varepsilon|v_n = 1) = 0, \quad \tilde{U}_n \leq \varepsilon \leq 1$$

従って

$$\begin{aligned}
F_n(t|v_n = 1, \varepsilon_n) &= 1 - S_n(t|v_n = 1, \varepsilon_n) \\
&= 1 - \exp\left(-\int_0^t h_n(\tau|v_n = 1, \varepsilon_n) d\tau\right) = 1 - e^{-\lambda(\tilde{U}_n - \varepsilon_n)t}
\end{aligned}$$

に注意すると

$$\begin{aligned}
F_n(t|v_n = 1) &= \int_0^1 F_n(t|v_n = 1, \varepsilon) p(\varepsilon|v_n = 1) d\varepsilon \\
&= \int_0^{\tilde{U}_n} (1 - e^{-\lambda\tilde{U}_n t} e^{\lambda\varepsilon t}) \frac{d\varepsilon}{\tilde{U}_n} \\
&= 1 - \frac{e^{-\lambda\tilde{U}_n t}}{\tilde{U}_n} \int_0^{\tilde{U}_n} e^{\lambda\varepsilon t} d\varepsilon \\
&= 1 - \frac{e^{-\lambda\tilde{U}_n t}}{\tilde{U}_n} \frac{1}{\lambda t} e^{\lambda\varepsilon t} \Big|_{\varepsilon=0}^{\tilde{U}_n} \\
&= 1 - \frac{e^{-\lambda\tilde{U}_n t}}{\lambda\tilde{U}_n t} (e^{\lambda\tilde{U}_n t} - 1) = 1 - \frac{1 - e^{-\lambda\tilde{U}_n t}}{\lambda\tilde{U}_n t}
\end{aligned}$$

となる。さらに $F_n(t|v_n = 0)$ に注意すると

$$\begin{aligned}
F_n(t) &= F_n(t|v_n = 1) \Pr(v_n = 1) + F_n(t|v_n = 0) \Pr(v_n = 0) \\
&= \left(1 - \frac{1 - e^{-\lambda\tilde{U}_n t}}{\lambda\tilde{U}_n t}\right) \tilde{U}_n = \tilde{U}_n - \frac{1 - e^{-\lambda\tilde{U}_n t}}{\lambda t}
\end{aligned}$$

すなわち

$$S_n(t) = 1 - F_n(t) = \frac{1 - e^{-\lambda\tilde{U}_n t}}{\lambda t} + (1 - \tilde{U}_n)$$

が成立する。また密度関数 $f_n(t)$ は次のようになる。

$$\begin{aligned}
f_n(t) &= \frac{d}{dt} S_n(t) = -\frac{d}{dt} \frac{1 - e^{-\lambda \tilde{U}_n t}}{\lambda t} \\
&= -\frac{\lambda \tilde{U}_n t e^{-\lambda \tilde{U}_n t} - (1 - e^{-\lambda \tilde{U}_n t})}{\lambda t^2} \\
&= \frac{1 - 1 + (1 + \lambda \tilde{U}_n t) e^{-\lambda \tilde{U}_n t}}{\lambda t^2} = \frac{1 - (1 + \lambda \tilde{U}_n t) e^{-\lambda \tilde{U}_n t}}{\lambda t^2}
\end{aligned}$$

命題 A.3.3 本研究のモデルの対数尤度関数 $\ell(\theta; \vec{T})$ は、次式の通り。

$$\begin{aligned}
\ell(\theta; \vec{T}) &= \sum_{n=1}^N \left[d_n \left\{ \log \left(1 + (1 + \lambda \tilde{U}_n \hat{T}_n) e^{-\lambda \tilde{U}_n \hat{T}_n} \right) - \log \lambda - 2 \log \hat{T}_n \right\} \right. \\
&\quad \left. + (1 - d_n) \log \left(\frac{1 - e^{-\lambda \tilde{U}_n T_0}}{\lambda T_0} + 1 - \tilde{U}_n \right) \right]
\end{aligned}$$

(証明) 対数尤度関数の定義より、

$$\begin{aligned}
\ell(\theta; \vec{T}) &:= \log L(\theta; \vec{T}) \\
&= \log \prod_{n=1}^N f_n(\hat{T}_n | \theta)^{d_n} S_n(T_0 | \theta)^{1-d_n} \\
&= \sum_{n=1}^N \left[d_n \log f_n(\hat{T}_n | \theta) + (1 - d_n) \log S_n(T_0 | \theta) \right]
\end{aligned}$$

である。更に命題 A.3.2 より、

$$\log f_n(t) = \log \frac{1 - (1 + \lambda \tilde{U}_n t) e^{-\lambda \tilde{U}_n t}}{\lambda t^2}$$

$$\begin{aligned}
&= \log \left(1 - (1 + \lambda \tilde{U}_n t) e^{-\lambda \tilde{U}_n t} \right) - \log \lambda - 2 \log t \\
\log S_n(t) &= \left(\frac{1 - e^{-\lambda \tilde{U}_n t}}{\lambda t} + (1 - \tilde{U}_n) \right)
\end{aligned}$$

であるから、

$$\begin{aligned}
\ell(\theta; \vec{T}) &= \sum_{n=1}^N \left[d_n \left\{ \log \left(1 + (1 + \lambda \tilde{U}_n \hat{T}_n) e^{-\lambda \tilde{U}_n \hat{T}_n} \right) - \log \lambda - 2 \log \hat{T}_n \right\} \right. \\
&\quad \left. + (1 - d_n) \log \left(\frac{1 - e^{-\lambda \tilde{U}_n T_0}}{\lambda T_0} + 1 - \tilde{U}_n \right) \right]
\end{aligned}$$

となる。

Appendix B

EM アルゴリズム

2.2.2 項でも述べたとおり、各家計が潜在的試用購買者であるか否かを表す v_n が欠測しているので、条件付きスプリット・ハザードモデル

$$h_n(t|v_n) := v_n \tilde{h}_n(t), \quad v_n \sim Bi(p_n)$$

の（周辺）対数尤度関数は、条件付き対数尤度関数の v_n に関する期待値によって定義されている。

$$\ell(\theta; \vec{T}) = \sum_{n=1}^N \left\{ d_n \left(\log p_n + \log \tilde{h}_n(\hat{T}_n | v_n = 1) - \tilde{H}_n(\hat{T}_n | v_n = 1) \right) + (1 - d_n) \log \left(p_n e^{-\tilde{H}_n(T_0 | v_n = 1)} + 1 - p_n \right) \right\}$$

しかしこの式は解析的に複雑なので、Hesse 行列等の導出やパラメータの最尤推定が困難である。したがって杉田・中村・田島の研究 [72] では、その解析的な複雑さを避けるために、EM アルゴリズムによるパラメータ推定を行っている。

そこでここでは EM アルゴリズムによるスプリット・ハザードモデルのパラメータ推定、および杉田・中村・田島のモデルや本研究のモデルへの具体的な適用について述べる。なお EM アルゴリズムの一般論は [88]、ハザードモデルへの適用は [7, 37]、そしてマーケティング

ング・モデルへの適用については [42] を参照のこと。

B.1 原理

B.1.1 概略

EM アルゴリズムとは、欠測値や打ち切り標本等を含む「不完全データ」を用いてパラメータの最尤推定を行うアルゴリズムである [88]。

スプリット・ハザードモデルの場合は、家計 n が潜在的試用購買者であるか否かを表す確率的ダミー変数 v_n が、センサリングされた家計に関して欠測しているため、オブザベーションは不完全である (表 B.1 参照)。したがって (対数) 尤度関数は、条件付き尤度関数の v_n に関する期待値によって計算された。

表 B.1: 不完全データ

n	\hat{T}_n	\hat{v}_n
1	\hat{T}_1	1
\vdots	\vdots	\vdots
M	\hat{T}_M	1
$M+1$	T_0	.
\vdots	\vdots	\vdots
N	T_0	.

いま全ての家計に関して v_n が既知であったとする。このとき v_n のオブザベーションを \hat{v}_n 、また $\vec{v} := \{\hat{v}_n\}_{n=1}^N$ 、 $\vec{T} := \{\hat{T}_n\}_{n=1}^N$ すると、対数尤度関数は次のようになる筈である。

$$\begin{aligned} \ell(\theta; \vec{T}, \vec{v}) &= \sum_{n=1}^M \left\{ \log \tilde{h}_n(\hat{T}_n) - \tilde{H}_n(\hat{T}_n) + \log p_n \right\} \\ &\quad + \sum_{n=M+1}^N \left\{ -\hat{v}_n (\tilde{H}_n(T_0) - \log p_n) + (1 - \hat{v}_n) \log(1 - p_n) \right\} \end{aligned} \quad (\text{B.1})$$

しかし実際には v_n は欠測しているので、適当なパラメータの初期値 θ_0 と \vec{T} を与えた下での、 $\vec{v}^0 := \{\hat{v}_n^0 = \hat{v}_n(\theta_0)\}_{n=1}^N$ に関する条件付き期待値 $E[\cdot; \vec{v}^0 | \vec{T}, \theta_0]$ をとると、次のようになる。

$$\begin{aligned} Q(\theta, \theta_0) &:= E[\ell(\theta; \vec{T}, \vec{v}^0); \vec{v}^0 | \vec{T}, \theta_0] \\ &= \sum_{n=1}^M \left\{ \log \tilde{h}_n(\hat{T}_n) - \tilde{H}_n(\hat{T}_n) + \log p_n \right\} \\ &\quad + \sum_{n=M+1}^N \left\{ -\vec{v}_n^0 (\tilde{H}_n(T_0) - \log p_n) + (1 - \vec{v}_n^0) \log(1 - p_n) \right\} \end{aligned} \quad (\text{B.2})$$

ただしセンサリングされた家計 n , $M+1 \leq n \leq N$ に対して

$$\vec{v}_n^0 := E[\hat{v}_n^0 | \hat{T}_n, \theta_0] = \frac{p_n e^{-\tilde{H}_n(T_0 | \theta_0)}}{p_n e^{-\tilde{H}_n(T_0 | \theta_0)} + 1 - p_n} \quad (\text{B.3})$$

である。

従って EM アルゴリズムは

Expectation ステップ B.3 式から v_n の期待値を求め、さらに B.2 式から $Q(\theta, \theta_0)$ を計算する。

Maximization ステップ $Q(\theta, \theta_0)$ を最大化するパラメータ推定値 $\hat{\theta}$ を求める。

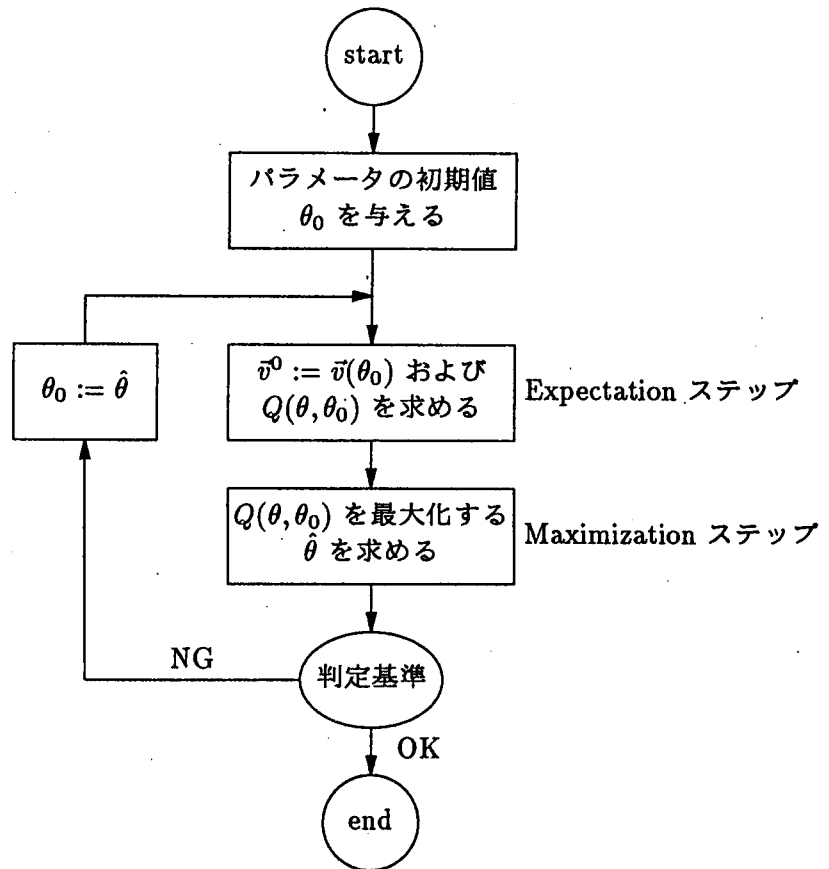
という手順を、所与の判定基準が満たされるまで繰り返すのである (図 B.1 参照)。

一般に対数尤度関数 $\ell(\theta; \vec{T})$ を最適化するよりも EM アルゴリズムを用いて $Q(\theta, \theta_0)$ を最適化する方が、パラメータ推定時の解析的労力が軽減される。

B.1.2 EM 推定量の性質

ところで

図 B.1: EM アルゴリズム



$$Q(\theta, \theta_0) \geq Q(\theta_0, \theta_0) \Rightarrow \ell(\theta; \vec{T}) \geq \ell(\theta_0; \vec{T}) \quad (\text{B.4})$$

$$\text{grad}_Q(\hat{\theta}) = 0 \Rightarrow \text{grad}_\ell(\hat{\theta}) = 0 \quad (\text{B.5})$$

が成立するので、EM アルゴリズムによるパラメータ推定値が、2.11 式で表される尤度関数 $L(\theta; \vec{T})$ の極大値になっていることが分かる。

しかしその反復計算の収束は、一般にかなり遅い事が報告されている [88]。たとえば杉田・中村・田島のモデルの EM 推定に要する時間は、unix ワークステーション上でも 20 時間以上である [72]。

B.1.3 命題の証明

B.1, B.2 式の証明

1. $1 \leq n \leq M$ すなわち $d_n = 1$ の時、 $\hat{v}_n^0 = 1$ に注意すると

$$L_n(\theta; \hat{T}_n, \hat{v}_n^0 = 1) = f_n(\hat{T}_n | v_n = \hat{v}_n^0 = 1) \Pr(v_n = \hat{v}_n^0 = 1) = \tilde{h}_n(\hat{T}_n) e^{-\tilde{H}_n(\hat{T}_n)} p_n$$

$$\therefore \ell_n(\theta; \hat{T}_n, \hat{v}_n^0 = 1) = \log \tilde{h}_n(\hat{T}_n) - \tilde{H}_n(\hat{T}_n) + \log p_n$$

$$\therefore Q_n(\theta, \theta_0) = \log \tilde{h}_n(\hat{T}_n) - \tilde{H}_n(\hat{T}_n) + \log p_n$$

2. $M + 1 \leq n \leq N$ すなわち $d_n = 0$ の時

$$L_n(\theta; \hat{T}_n, \hat{v}_n^0) = S_n(T_0 | v_n = \hat{v}_n^0) \Pr(v_n = \hat{v}_n^0) = e^{-\hat{v}_n^0 \tilde{H}_n(T_0)} p_n^{\hat{v}_n^0} (1 - p_n)^{1 - \hat{v}_n^0}$$

$$\begin{aligned}\therefore l_n(\theta; \hat{T}_n, \hat{v}_n^0) &= -\hat{v}_n^0 + \hat{v}_n^0 \log p_n + (1 - \hat{v}_n^0) \log(1 - p_n) \\ &= -\hat{v}_n^0 (\tilde{H}_n(T_0) - \log p_n) + (1 - \hat{v}_n^0) \log(1 - p_n)\end{aligned}$$

このとき $\bar{v}_n^0 := E[\hat{v}_n^0 | \hat{T}_n, \theta_0]$ とすると

$$Q_n(\theta, \theta_0) = -\bar{v}_n^0 (\tilde{H}_n(T_0) - \log p_n) + (1 - \bar{v}_n^0) \log(1 - p_n)$$

B.3 式の証明

$M + 1 \leq n \leq N$ すなわち $d_n = 0$ の時、Bayes の定理により

$$\begin{aligned}\bar{v}_n^0 &:= E[\hat{v}_n^0 | \hat{T}_n, \theta_0] \\ &= \Pr(\hat{v}_n^0 = 1 | \theta_0) \\ &= \frac{S_n(T_0 | \hat{v}_n^0 = 1, \theta_0) \Pr(\hat{v}_n^0 = 1 | \theta_0)}{S_n(T_0 | \hat{v}_n^0 = 1, \theta_0) \Pr(\hat{v}_n^0 = 1 | \theta_0) + S_n(T_0 | \hat{v}_n^0 = 0, \theta_0) \Pr(\hat{v}_n^0 = 0 | \theta_0)} \\ &= \frac{p_n S_n(T_0 | \hat{v}_n^0 = 1, \theta_0)}{p_n S_n(T_0 | \hat{v}_n^0 = 1, \theta_0) + 1 - p_n} = \frac{p_n e^{-\tilde{H}_n(T_0)}}{p_n e^{-\tilde{H}_n(T_0)} + 1 - p_n}\end{aligned}$$

B.4, B.5 式の証明

記号の簡略化のため、関数 R_n, R を次のように定義しておく。

$$R_n(\theta, \theta_0) := E[l_n(\theta; \hat{v}_n^0 | \hat{T}_n); \hat{v}_n^0 | \hat{T}_n, \theta_0] \quad (\text{B.6})$$

$$R(\theta, \theta_0) := E[l(\theta; \hat{v}_n^0 | \hat{T}_n); \hat{v}_n^0 | \hat{T}_n, \theta_0] \quad (\text{B.7})$$

1. $l(\theta; \vec{T}) = Q(\theta, \theta_0) - R(\theta, \theta_0)$ の証明

$1 \leq n \leq M$ すなわち $d_n = 1$ の時、

$$\begin{aligned}
 L_n(\theta; \hat{T}_n) &= f_n(\hat{T}_n | \theta) \\
 &= \frac{\Pr(v_n = \hat{v}_n^0 | \hat{T}_n, \theta) f_n(\hat{T}_n | \theta)}{\Pr(v_n = \hat{v}_n^0 | \hat{T}_n, \theta)} \\
 &= \frac{f_n(\hat{T}_n, \hat{v}_n^0 | \theta)}{\Pr(v_n = \hat{v}_n^0 | \hat{T}_n, \theta)} = \frac{L_n(\theta; \hat{T}_n, \hat{v}_n^0)}{L_n(\theta; \hat{v}_n^0 | \hat{T}_n)}
 \end{aligned}$$

$$\therefore l_n(\theta; \hat{T}_n) = l_n(\theta; \hat{T}_n, \hat{v}_n^0) - l_n(\theta; \hat{v}_n^0 | \hat{T}_n)$$

$$\therefore l_n(\theta; \hat{T}_n) = Q_n(\theta, \theta_0) - R_n(\theta, \theta_0)$$

なおこの式はあきらかに、 $1 \leq n \leq M$ すなわち $d_n = 1$ の時も成立する。

$$\therefore l(\theta; \vec{T}) = Q(\theta, \theta_0) - R(\theta, \theta_0) \tag{B.8}$$

2. $Q(\theta, \theta_0) \geq Q(\theta_0, \theta_0) \Rightarrow l(\theta; \vec{T}) \geq l(\theta_0; \vec{T})$ の証明

まず Jensen の不等式 [87, p.326] より

$$\begin{aligned}
 R_n(\theta, \theta_0) - R_n(\theta_0, \theta_0) &= E \left[l_n(\theta; \hat{v}_n^0 | \hat{T}_n) - l_n(\theta_0; \hat{v}_n^0 | \hat{T}_n); \hat{v}_n^0 | \hat{T}_n, \theta_0 \right] \\
 &= E \left[\log \frac{\Pr(\hat{v}_n^0 | \hat{T}_n, \theta)}{\Pr(\hat{v}_n^0 | \hat{T}_n, \theta_0)}; \hat{v}_n^0 | \hat{T}_n, \theta_0 \right] \\
 &\leq \log E \left[\frac{\Pr(\hat{v}_n^0 | \hat{T}_n, \theta)}{\Pr(\hat{v}_n^0 | \hat{T}_n, \theta_0)}; \hat{v}_n^0 | \hat{T}_n, \theta_0 \right] \\
 &= \log \int \frac{\Pr(\hat{v}_n^0 | \hat{T}_n, \theta)}{\Pr(\hat{v}_n^0 | \hat{T}_n, \theta_0)} \Pr(\hat{v}_n^0 | \hat{T}_n, \theta_0) d\hat{v}_n^0 = 0
 \end{aligned}$$

$$\therefore R_n(\theta, \theta_0) \leq R_n(\theta_0, \theta_0)$$

また、逆の不等号も同様に成立する。

$$\therefore R_n(\theta, \theta_0) \geq R_n(\theta_0, \theta_0)$$

$$\therefore R_n(\theta, \theta_0) = R_n(\theta_0, \theta_0), \quad n = 1, 2, \dots, N$$

$$\therefore R(\theta, \theta_0) = R(\theta_0, \theta_0) \tag{B.9}$$

したがって与式の成立は、B.8 式より明らかである。

3. $\text{grad}_Q(\hat{\theta}) = 0 \Rightarrow \text{grad}_\ell(\hat{\theta}) = 0$ の証明

B.9 式より $R(\theta, \theta_0) = R(\theta_0, \theta_0)$ だから、 $\text{grad}_R(\theta) = 0$ である。したがって与式の成立は、B.8 式より明らかである。

B.2 杉田・中村・田島のモデルへの適用

杉田・中村・田島のモデル (モデル 0) は次式で定義される。

$$h_n(t|v_n) = v_n \tilde{h}_n(t|v_n) = v_n \lambda e^{U_n(t)}$$

また潜在的試用購買者であるか否かを表す確率的ダミー変数 v_n は、 p をパラメータに持つ Bernoulli 分布 $Bi(p)$ に、独立かつ同一に従っている。

$$v_n \sim Bi(p), \quad (i.i.d.)$$

したがって、次式の成立は明らかである。

1. $d_n = 1$ の時

$$Q_n(\theta, \theta_0) = \log \lambda + U_n(\hat{T}_n) - \lambda \int_0^{\hat{T}_n} U_n(t) dt + \log p$$

2. $d_n = 0$ の時

$$Q_n(\theta, \theta_0) = -\bar{v}_n^0 \left(\lambda \int_0^{T_0} U_n(t) dt - \log p \right) + (1 - \bar{v}_n^0) \log(1 - p)$$

$$\bar{v}_n = \frac{p \exp \left(-\lambda \int_0^{T_0} U_n(t) dt \right)}{p \exp \left(-\lambda \int_0^{T_0} U_n(t) dt \right) + 1 - p}$$

B.3 本研究のモデルへの適用

本研究のモデル（モデル 1）は次式で定義される。

$$h_n(t|v_n, \varepsilon_n) = v_n \tilde{h}_n(t|v_n, \varepsilon_n) = v_n \lambda (\tilde{U}_n - \varepsilon_n)$$

ただし攪乱項 ε_n は、一様分布 $U(0, 1)$ に独立かつ同一に従っている。そして v_n はそれぞれ、 $\Pr(\tilde{U}_n \geq \varepsilon_n) = \tilde{U}_n$ をパラメータに持つ Bernoulli 分布 $Bi(\tilde{U}_n)$ に、独立に従っている。

$$v_n \sim Bi(\tilde{U}_n)$$

したがって、次式の成立は明らかである。

1. $d_n = 1$ の時

$$Q_n(\theta, \theta_0) = \log \lambda + \log(\tilde{U}_n - \varepsilon_n) - \lambda(\tilde{U}_n - \varepsilon_n)\hat{T}_n + \log \tilde{U}_n$$

2. $d_n = 0$ の時

$$Q_n(\theta, \theta_0) = -\bar{v}_n^0 \left(\lambda(\tilde{U}_n - \varepsilon_n)T_0 - \log \tilde{U}_n \right) + (1 - \bar{v}_n^0) \log(1 - \tilde{U}_n)$$
$$\bar{v}_n = \frac{\tilde{U}_n e^{-\lambda(\tilde{U}_n - \varepsilon_n)T_0}}{\tilde{U}_n e^{-\lambda(\tilde{U}_n - \varepsilon_n)T_0} + 1 - \tilde{U}_n}$$

Appendix C

プログラムおよびアウトプット

実証分析のコンピュータ・プログラムおよびアウトプットの一部を掲載する。6.2.2 項で述べたとおり、本研究の実証分析は、i486 CPU を搭載したパーソナル・コンピュータ上の汎用統計パッケージ SAS を用いて行っている。特にこのプログラムは、SAS に含まれる IML (Interactive Matrix Language) [54] という行列演算言語によって記述されている。

C.1 モデル 1 (本研究のモデル)

C.1.1 プログラム

```

/*****
/*
/*      ML-Estimation of SNT2 model using NLP subroutine
/*      -----
/*      ( with output formatted for Lotus-123 )
/*
/*      bXn = b0 + ( b1*heavy2 + ... + b5*jk + b7*islem ) / kakem
/*
/*      Uflag=1  Un = 1 - exp(-bXn)
/*      Uflag=2  Un = 1 - exp( -exp(bXn) )
/*      Uflag=3  Un = 1 / ( 1 + exp(-bXn) )
/*
/*
/*                               programmed by TAJIMA, Hirokazu (^_^;)
/*
/*
/*****
%macro snt2(itemno, Uflag);
title "SNT2: itemno=&itemno. Uflag=&Uflag.";
data kake; set dei.hzdata&itemno.; keep kake1-kake38;
data isle; set dei.hzdata&itemno.; keep isle1-isle38;

proc iml; reset noprint linesize=180;
```

```

/***** exchange SAS-dataset into matrix *****/
use dei.hzdata&itemno.;
read all var{heavy2 loyal2 toku2 dy jk} into behav;
read all var{kakem} into kakem;
read all var{islem} into islem;
read all var{t} into Tn;
read all var{d} into dn;
use kake; read all into kake;
use isle; read all into isle;
use dei.actual&itemno.; read all into actual;

/***** define constant variables *****/
Uflag = &Uflag.;
M = 1 + ( 1 + ncol( behav || islem ) );
N = nrow(behav);
TO = 13;
Tmax = 38;

/***** initialize parameters ( a = para[1] ) *****/
para_ini = j(1, M, .);
do mm = 1 to M;
    para_ini[mm] = uniform(0);
end;

/***** define utility function "Un" *****/
start Un(kakem, islem, para) global(N, M, behav, Uflag);
    bXn = para[2] + ( ( behav || islem ) * para[3:M] ) # kakem ## (-1);
    if Uflag = 1 then Un = 1 - exp(-bXn);
    if Uflag = 2 then Un = 1 - exp( -exp(bXn) );
    if Uflag = 3 then Un = 1 / ( 1 + exp(-bXn) );
return(Un);
finish;

/***** define density function "fn" *****/
start fn(t, Un, para);
    lmd = exp( -para[1] );
    fn = ( 1 - (1+lmd#Un#t) # exp(-lmd#Un#t) ) # ( lmd # t ## 2 ) ## (-1);
return(fn);
finish;

/***** define disaggregate log-likelihood function "logLn" *****/
start logLn(para) global(kakem, islem, dn, TO, Tn);
    lmd = exp( -para[1] );
    Un = Un(kakem, islem, para);
    Sn = ( 1 - exp(-lmd#Un#TO) ) # ( lmd # TO ) ## (-1) + 1 - Un;
logLn = log( fn(Tn, Un, para) ) # ( dn = 1 ) + log( Sn ) # ( dn = 0 );
return(logLn);
finish;

/***** define aggregate log-likelihood function "logL" *****/
start logL(para);
    logL = sum( logLn(para) );
return(logL);
finish;

/***** optimize parameters *****/
call nlpnrr(rc, para_est, "logL", para_ini, {1 2});
print rc;

/***** compute Hesse matrix *****/
call nlpfdd(logL, grad, Hesse, "logL", para_est);

```

```

/***** compute t-value of parameters *****/
if det(Hesse) = 0
  then do;
    print "ERROR: Hesse 行列が特異";
    tvalue = j(1, M, .);
    end;
  else do;
    diaginVH = diag( inv(Hesse) )[+, ];
    if any( diaginVH >= 0)
      then do;
        print "ERROR: 逆 Hesse 行列の或る対角成分が非負";
        tvalue = j(1, M, .);
        end;
      else tvalue = para_est # sqrt(-diaginVH)##(-1);
    end;

/***** compute forecasts "dpt" and ultimate penetration "p" *****/
dpt = j(Tmax,1,.);
do tt = 1 to Tmax;
  kakem2 = kake[ ,1:tt][ ,:];
  islem2 = isle[ ,1:tt][ ,:];
  Un2 = Un(kakem2, islem2, para_est);
  if tt = 13 then p = Un2[:];
  fnt = fn(tt, Un2, para_est);
  dpt[tt] = fnt[:];
end;

/***** compute mse, mae and Theil's U *****/
mse0 = ( actual ## 2 )[(T0+1):Tmax][:];
mse1 = ( dpt ## 2 )[(T0+1):Tmax][:];
mse = ( ( actual - dpt ) ## 2 )[(T0+1):Tmax][:];
mae = abs( actual - dpt )[(T0+1):Tmax][:];
Theil = sqrt(mse) / ( sqrt(mse0) + sqrt(mse1) );

/***** print parameters and forecasts *****/
output = j( 2*M+5, 1, .);
do mm = 1 to M; output[2*mm-1] = para_est[mm]; end;
do mm = 1 to M; output[2*mm] = tvalue[mm]; end;
output[2*M+1] = logL;
output[2*M+2] = p;
output[2*M+3] = Theil;
output[2*M+4] = mse;
output[2*M+5] = mae;
print output;
print dpt;

/***** print disaggregate log-likelihood *****/
logL_ini = logLn(para_ini);
logL_est = logLn(para_est);
print dn[format=2.] Tn[format=3.] behav[format=3.]
      kakem[format=5.2] islem[format=5.2] logL_ini logL_est;

quit;
%mend snt2;
%snt2(7,3);

```

C.1.2 アウトプット

Optimization Start Parameter Estimates

Parameter	Estimate	Gradient
1 X1	0.451885	74.137377
2 X2	0.660310	-78.812648
3 X3	0.187416	-41.913977
4 X4	0.718735	31.545539
5 X5	0.411923	-34.652494
6 X6	0.907743	-41.110923
7 X7	0.278676	-15.401842
8 X8	0.074218	-10.421058

Value of Objective Function = -264.9110817

Newton-Raphson Ridge Optimization
Without Parameter Scaling

Number of Parameter Estimates 302907400

Optimization Start: Active Constraints= 2097152 Criterion= -9.943619121E60
Maximum Gradient Element= 0.000

Iter	rest	nfun	act	optcrit	difcrit	maxgrad	ridge	rho
1*	0	2	0	-174.3422	90.5689	52.661	128.0	1.691
2	0	3	0	-123.7257	50.6165	8.863	32.000	1.291
3	0	4	0	-112.4493	11.2764	6.007	8.000	1.564
4	0	5	0	-105.0373	7.4120	2.753	2.000	1.433
5	0	6	0	-101.0341	4.0032	1.061	0.500	1.376
6	0	7	0	-99.3175	1.7166	0.464	0.125	1.304
7	0	8	0	-98.8082	0.5093	0.246	0.0313	1.260
8	0	9	0	-98.6985	0.1096	0.0740	0	1.134
9	0	10	0	-98.6912	0.00736	0.0073	0	1.082
10	0	11	0	-98.6911	0.000107	0.00027	0	1.014
11	0	12	0	-98.6911	5.711E-8	0.00002	0	0.944

Optimization Results: Iterations= 851979 Function Calls= 12
Hessian Calls= 1352073248
Active Constraints= -1405446712X Criterion= 0.0000243
Maximum Gradient Element= 0 Ridge= 1.02E-202

NOTE: GCONV convergence criterion satisfied.

Optimization Results Parameter Estimates

Parameter	Estimate	Gradient
1 X1	1.632944	-0.000017345
2 X2	-4.811144	-0.000005607
3 X3	-2.222856	-0.000005461
4 X4	0.465846	-0.000024277
5 X5	0.438357	0.000009441
6 X6	3.489630	-0.000003092
7 X7	3.332718	-0.000004047
8 X8	4.813195	-0.000007846

Value of Objective Function = -98.69107575

C.2 モデル 0 (杉田・中村・田島のモデル)

C.2.1 プログラム

```
/*-----*/
/*
/*      ML-Estimation of SNT1 model using NLP subroutine
/*      -----
/*      ( with output formatted for Lotus-123 )
/*
/*      Un   = b0 + b1*heavy2 + ... + b5*jk
/*      Wn(t) = - b6*kake(t) + b7*isle(t)
/*
/*
/*                      programmed by TAJIMA, Hirokazu (^_^;)
/*
/*-----*/
%macro SNT1(itemno);
title "SNT1 model: itemno=&itemno.";
data kake; set dei.hzdata&itemno.; keep kake1-kake38;
data isle; set dei.hzdata&itemno.; keep isle1-isle38;

proc iml; reset noprint linesize=180;

/*----- exchange SAS-dataset into matrix -----*/
use dei.hzdata&itemno.;
  read all var{heavy2 loyal2 toku2 dy jk} into behav;
  read all var{t} into Tn;
  read all var{d} into dn;
use kake; read all into kake;
use isle; read all into isle;
use dei.actual&itemno.; read all into actual;

/*----- define constant variables -----*/
M = 1 + ( 1 + ncol(behav) + 2 );
N = nrow(behav);
T0 = 13;
Tmax = 39;

/*----- initialize parameters ( p = para[1] ) -----*/
para_ini = j(1, M, .);
do mm = 1 to M;
  para_ini[mm] = uniform(0);
end;

/*----- define time-independent utility function "Un" -----*/
start Un(para) global(N, M, behav);
  Un = ( j(N, 1, 1) || behav ) * para'[2:(M-2)];
  return(Un);
finish;

/*----- define time-dependent utility function "Wn(t)" -----*/
start Wn(t, para) global(N, M, kake, isle);
  _isp_ = j(N, 2, .);
do nn = 1 to N;
  _isp_[nn, 1] = kake[nn, t[nn]];
  _isp_[nn, 2] = isle[nn, t[nn]];
end;
  Wn = _isp_ * para'[(M-1):M];
  return(Wn);
```

```

finish;

/***** define cumulative of exponential-utility "Zn(t)" *****/
start Zn(t, para) global(N M, kake, isle);
* if nrow(t) = 1 then t = j(N, 1, t);
_cumWn_ = j(N, 1, 0);
do nn = 1 to N;
do tt = 1 to t[nn];
_cumWn_[nn] = _cumWn_[nn]
+ ( kake[nn, tt] || isle[nn, tt] ) *para[M-1:M];
end;
end;
Zn = exp( Un(para) ) # exp( _cumWn_ );
return(Zn);
finish;

/***** define density function "fn" *****/
start fn(t, para);
fn = para[1] * exp( Un(para) ) # exp( Wn(t, para) ) # exp( -Zn(t, para) );
return(fn);
finish;

/***** define disaggregate log-likelihood function "logLn" *****/
start logLn(para) global(dn, Tn);
logfn = log( para[1] ) + Un(para) + Wn(Tn, para) - Zn(Tn, para);
Sn = para[1] * exp( -Zn(Tn, para) ) + 1 - para[1];
logLn = logfn # (dn=1) + log(Sn) # (dn=0);
return(logLn);
finish;

/***** define aggregate log-likelihood function "logL" *****/
start logL(para);
logL = sum( logLn(para) );
return(logL);
finish;

/***** optimize parameters *****/
blc = { 0, 1} || j(2, M-1, .); /* 0 <= p = para[1] <= 1; constraint */
call nlpnrr(rc, para_est, "logL", para_ini, {1 2}, blc);
print rc;

/***** compute Hesse matrix *****/
call nlpfdd(logL, grad, Hesse, "logL", para_est);

/***** compute t-value of parameters *****/
if det(Hesse) = 0
then do;
print "ERROR: Hesse 行列が特異";
tvalue = j(1, M, .);
end;
else do;
diaginvH = diag( inv(Hesse) )[+, ];
if any( diaginvH >= 0)
then do;
print "ERROR: 逆 Hesse 行列の或る対角成分が非負";
tvalue = j(1, M, .);
end;
else tvalue = para_est # sqrt(-diaginvH)##(-1);
end;

/***** compute forecasts "dpt" and ultimate penetration "p" *****/
dpt = j(Tmax, 1, .);

```

```

do tt = 1 to Tmax;
    dpt[tt] = fn(tt, para_est)[:];
end;

/***** compute mse, mae and Theil's U *****/
mse0 = ( actual ## 2 )[(T0+1):Tmax][:];
mse1 = ( dpt ## 2 )[(T0+1):Tmax][:];
mse = ( ( actual - dpt ) ## 2 )[(T0+1):Tmax][:];
mae = abs( actual - dpt )[(T0+1):Tmax][:];
Theil = sqrt(mse) / ( sqrt(mse0) + sqrt(mse1) );

/***** print parameters and forecasts *****/
output = j( 2*M+4, 1, .);
do mm = 2 to M; output[2*mm-3] = para_est[mm]; end;
do mm = 2 to M; output[2*mm-2] = tvalue[mm]; end;
output[2*M-1] = para_est[1];
output[2*M] = tvalue[1];
output[2*M+1] = logL;
output[2*M+2] = Theil;
output[2*M+3] = mse;
output[2*M+4] = mae;
print output;
print dpt;

/***** print disaggregate log-likelihood *****/
logL_ini = logLn(para_ini);
logL_est = logLn(para_est);
print dn[format=2.] Tn[format=3.] behav[format=3.] logL_ini logL_est;

quit;
%mend SNT1;

%SNT1(1);
%SNT1(3);
%SNT1(7);

```

C.2.2 アウトプット

Optimization Start Parameter Estimates		
Parameter	Estimate	Gradient
1 B0	0.476612	-318.764226
2 B1	0.083165	-200.188981
3 B2	0.208518	-63.955535
4 B3	0.090848	23.316753
5 B4	0.127931	-143.300128
6 B5	0.789078	-171.777486
7 B6	0.319021	295.698032
8 B7	0.114532	-62.876944
9 P	0.355479	-214.285608

Value of Objective Function = -427.0813413

Newton-Raphson Ridge Optimization

*** Termination Criteria ***

Minimum Iterations 6553600
Maximum Iterations 660734077

Maximum Function Calls. 658907719
 ABSGTOL Gradient Criterion 0.0000100
 GTOL Gradient Criterion 1E-8
 GTOL2 Gradient Criterion 0
 ABSFTOL Function Criterion 0
 FTOL Function Criterion 0
 FTOL2 Function Criterion 0
 FSIZE Parameter 0
 ABSXTOL Parameter Change Criterion 0
 XTOL Parameter Change Criterion 0
 XSIZE Parameter 0

Start Diagonal for Hessian Iteration 1.483E-143

Singularity Tolerance SINGULAR 1E-8
 Infinite Step Size 1.483E-143
 Constraint Precision CEPSEILON 1E-8
 Linearly Dependent Constraints (LDCTOL) 1E-8
 Releasing Active Constraints (DEACTOL) 1.483E-143

Newton-Raphson Ridge Optimization
Without Parameter Scaling

Number of Parameter Estimates 678428681
 Number of Functions (Observations) 200
 Number of Lower Bounds 65537
 Number of Upper Bounds 575088556

Optimization Start: Active Constraints= 20971529 Criterion= -0.000
 Maximum Gradient Element= 0.000

Iter	rest	nfun	act	optcrit	difcrit	maxgrad	ridge	rho
1	0	5	0	-329.5726	97.5087	130685	0.500	0.519
2	0	6	0	-256.3513	73.2213	72607	0.250	1.276
3	0	7	0	-224.1481	32.2032	37374	0.0625	1.310
4	0	8	0	-203.1391	21.0090	18798	0	1.355
5	0	9	0	-185.6731	17.4659	9381.8	0	1.370
6	0	10	0	-169.8095	15.8636	4656.9	0	1.376
7	0	11	0	-155.0418	14.7677	2292.9	0	1.375
8	0	12	0	-141.5185	13.5233	1112.7	0	1.371
9	0	13	0	-129.6466	11.8719	529.5	0	1.364
10	0	14	0	-119.5957	10.0509	255.4	0	1.359
11	0	15	0	-111.5600	8.0356	119.4	0	1.341
12	0	16	0	-105.9657	5.5944	52.818	0	1.326
13	0	17	0	-102.4660	3.4997	24.151	0	1.323
14	0	18	0	-100.3314	2.1346	12.127	0	1.308
15	0	19	0	-99.0703	1.2611	6.191	0	1.314
16	0	20	0	-98.3170	0.7533	3.467	0	1.311
17	0	25	1	-98.2719	0.0452	3.238	0.200	0.133
18	0	26	1	-98.0061	0.2658	0.301	0	0.938
19	0	27	1	-98.0039	0.00214	0.00212	0	1.005
20	0	28	1	-98.0039	1.224E-7	6.33E-6	0	1.034

Optimization Results: Iterations= 1900564 Function Calls= 65557
 Hessian Calls= -1235288032
 Active Constraints= -2143280251X Criterion= 6.33E-6
 Maximum Gradient Element= 0 Ridge= 0

NOTE: GCONV convergence criterion satisfied.

Optimization Results
Parameter Estimates

Parameter	Estimate	Gradient	Active BC
1 BO	36.143501	-0.000000146	

2	B1	0.783387	-0.000006333	
3	B2	-1.453128	0.000000644	
4	B3	0.364646	-0.000003383	
5	B4	9.083099	-0.000000118	
6	B5	9.414834	2.9238071E-8	
7	B6	56.998452	0.000000110	
8	B7	1.531302	-0.000000253	
9	P	1.000000	1.479461	Upper BC

Value of Objective Function = -98.00391472

Appendix D

原データおよび家計別の対数尤度

実証分析で用いた原データおよび各モデルの家計別の対数尤度を掲載する。なお表頭の意味は次の通りである（詳しくは 6.1.3, 6.2.1 項を参照のこと）。

表頭	意味
n	家計の通し番号 (1 ~ 200)
Tn	試用購買時期 (38 週以内に試用購買しなかった時は欠損値)
dn	13 週以内に試用購買を行ったか否かを表すダミー変数
X1	ヘビーユーザー度
X2	ロイヤルユーザー度
X3	特売反応度
X4, X5	店舗ダミー
KAKE	価格掛率
ISLE	エンド陳列ダミー
model1	モデル 1 (本研究のモデル) の対数尤度
model0	モデル 0 (杉田・中村・田島のモデル) の対数尤度

n	Tn	dn	X1	X2	X3	X4	X5	KAKE	ISLE	model1	model0
1	.	0	0	1	0	1	0	0.91	0.08	-0.232278	-0.203480
2	.	0	0	-1	0	1	0	0.91	0.08	-0.069068	-0.098128
3	.	0	1	-1	0	1	0	0.91	0.08	-0.001019	-0.022946
4	.	0	1	-1	1	1	0	0.91	0.08	-0.002546	-0.050226
5	.	0	0	0	1	1	0	0.91	0.08	-0.225572	-0.309299
6	27	0	1	0	0	1	0	0.91	0.08	-0.002694	-0.033042
7	.	0	1	0	0	1	0	0.91	0.08	-0.002694	-0.033042
8	5	1	0	0	0	1	0	0.87	0.00	-4.867687	-4.413479
9	34	0	1	-1	0	1	0	0.91	0.08	-0.001019	-0.022946
10	27	0	1	0	1	1	0	0.91	0.08	-0.006538	-0.072326

11	.	0	1	-1	1	1	0	0.91	0.08	-0.002546	-0.050226
12	.	0	1	0	1	1	0	0.91	0.08	-0.006538	-0.072326
13	.	0	0	0	1	1	0	0.91	0.08	-0.225572	-0.309299
14	.	0	1	-1	0	1	0	0.91	0.08	-0.001019	-0.022946
15	.	0	1	-1	0	1	0	0.91	0.08	-0.001019	-0.022946
16	.	0	0	0	0	1	0	0.91	0.08	-0.133953	-0.141305
17	12	1	0	0	1	1	0	0.87	1.00	-3.727808	-2.328750
18	.	0	1	-1	0	1	0	0.91	0.08	-0.001019	-0.022946
19	.	0	1	-1	0	1	0	0.91	0.08	-0.001019	-0.022946
20	.	0	1	-1	0	1	0	0.91	0.08	-0.001019	-0.022946
21	19	0	1	-1	1	1	0	0.91	0.08	-0.002546	-0.050226
22	.	0	0	-1	0	1	0	0.91	0.08	-0.069068	-0.098128
23	.	0	0	-1	0	1	0	0.91	0.08	-0.069068	-0.098128
24	.	0	0	1	0	1	0	0.91	0.08	-0.232278	-0.203480
25	.	0	0	0	0	1	0	0.91	0.08	-0.133953	-0.141305
26	.	0	1	-1	0	1	0	0.91	0.08	-0.001019	-0.022946
27	.	0	0	1	0	1	0	0.91	0.08	-0.232278	-0.203480
28	.	0	1	-1	0	1	0	0.91	0.08	-0.001019	-0.022946
29	.	0	0	1	0	1	0	0.91	0.08	-0.232278	-0.203480
30	.	0	1	-1	0	1	0	0.91	0.08	-0.001019	-0.022946
31	.	0	0	1	0	1	0	0.91	0.08	-0.232278	-0.203480
32	.	0	1	-1	1	1	0	0.91	0.08	-0.002546	-0.050226
33	26	0	0	0	0	1	0	0.91	0.08	-0.133953	-0.141305
34	34	0	0	-1	1	1	0	0.91	0.08	-0.129211	-0.214790
35	.	0	0	-1	0	1	0	0.91	0.08	-0.069068	-0.098128
36	2	1	0	1	0	1	0	0.87	0.00	-4.106759	-4.016372
37	23	0	0	0	1	1	0	0.91	0.08	-0.225572	-0.309299
38	.	0	1	-1	0	1	0	0.91	0.08	-0.001019	-0.022946
39	.	0	0	0	1	1	0	0.91	0.08	-0.225572	-0.309299
40	.	0	1	-1	0	1	0	0.91	0.08	-0.001019	-0.022946
41	.	0	1	0	1	1	0	0.91	0.08	-0.006538	-0.072326
42	.	0	0	-1	0	1	0	0.91	0.08	-0.069068	-0.098128
43	23	0	0	-1	1	1	0	0.91	0.08	-0.129211	-0.214790
44	4	1	0	1	1	1	0	0.87	0.00	-3.773492	-3.335920
45	.	0	0	-1	1	1	0	0.91	0.08	-0.129211	-0.214790
46	.	0	1	-1	0	1	0	0.91	0.08	-0.001019	-0.022946
47	.	0	0	-1	1	1	0	0.91	0.08	-0.129211	-0.214790
48	27	0	1	-1	0	1	0	0.91	0.08	-0.001019	-0.022946
49	12	1	0	1	1	1	0	0.87	1.00	-3.726096	-2.087915
50	.	0	0	1	1	1	0	0.91	0.08	-0.351493	-0.445391
51	12	1	0	0	0	1	0	0.87	1.00	-3.730516	-2.959304
52	.	0	1	-1	0	1	0	0.91	0.08	-0.001019	-0.022946
53	.	0	1	-1	1	1	0	0.91	0.08	-0.002546	-0.050226
54	28	0	0	-1	0	1	0	0.91	0.08	-0.069068	-0.098128
55	23	0	0	-1	0	1	0	0.91	0.08	-0.069068	-0.098128
56	19	0	0	0	1	1	0	0.91	0.08	-0.225572	-0.309299
57	.	0	0	1	0	1	0	0.91	0.08	-0.232278	-0.203480
58	.	0	1	-1	1	1	0	0.91	0.08	-0.002546	-0.050226
59	.	0	0	-1	0	1	0	0.91	0.08	-0.069068	-0.098128
60	.	0	1	1	0	1	0	0.91	0.08	-0.006904	-0.047581
61	.	0	0	-1	0	1	0	0.91	0.08	-0.069068	-0.098128
62	.	0	0	0	1	1	0	0.91	0.08	-0.225572	-0.309299
63	27	0	1	-1	0	1	0	0.91	0.08	-0.001019	-0.022946
64	32	0	1	-1	0	1	0	0.91	0.08	-0.001019	-0.022946
65	.	0	0	0	0	1	0	0.91	0.08	-0.133953	-0.141305
66	.	0	0	0	1	1	0	0.91	0.08	-0.225572	-0.309299
67	.	0	0	0	0	1	0	0.91	0.08	-0.133953	-0.141305
68	10	1	0	1	1	1	0	0.89	0.00	-4.251447	-4.533123
69	12	1	0	-1	1	1	0	0.87	1.00	-3.730735	-2.607416
70	.	0	0	-1	0	1	0	0.91	0.08	-0.069068	-0.098128
71	.	0	0	0	0	1	0	0.91	0.08	-0.133953	-0.141305
72	13	1	0	-1	0	1	0	0.87	0.00	-5.808790	-4.824827
73	37	0	0	-1	0	1	0	0.91	0.08	-0.069068	-0.098128
74	5	1	0	-1	1	1	0	0.87	0.00	-4.907239	-4.021482
75	.	0	0	1	1	1	0	0.91	0.08	-0.351493	-0.445391
76	.	0	0	0	1	1	0	0.91	0.08	-0.225572	-0.309299

77	.	0	1	-1	0	1	0	0.91	0.08	-0.001019	-0.022946
78	.	0	0	0	0	1	0	0.91	0.08	-0.133953	-0.141305
79	12	1	1	0	0	1	0	0.87	1.00	-3.811974	-4.313939
80	.	0	0	1	0	1	0	0.91	0.08	-0.232278	-0.203480
81	.	0	0	0	0	1	0	0.91	0.08	-0.133953	-0.141305
82	.	0	0	0	1	1	0	0.91	0.08	-0.225572	-0.309299
83	3	1	0	1	1	1	0	0.87	0.00	-3.703003	-3.295725
84	.	0	0	1	1	1	0	0.91	0.08	-0.351493	-0.445391
85	.	0	0	0	0	1	0	0.91	0.08	-0.133953	-0.141305
86	.	0	0	-1	0	1	0	0.91	0.08	-0.069068	-0.098128
87	.	0	0	-1	1	1	0	0.91	0.08	-0.129211	-0.214790
88	18	0	0	-1	0	1	0	0.91	0.08	-0.069068	-0.098128
89	32	0	0	0	0	1	0	0.91	0.08	-0.133953	-0.141305
90	18	0	0	0	1	1	0	0.91	0.08	-0.225572	-0.309299
91	.	0	0	-1	1	1	0	0.91	0.08	-0.129211	-0.214790
92	.	0	0	1	0	1	0	0.91	0.08	-0.232278	-0.203480
93	.	0	1	0	0	1	0	0.91	0.08	-0.002694	-0.033042
94	3	1	0	0	0	1	0	0.87	0.00	-4.788595	-4.387975
95	18	0	0	0	0	1	0	0.91	0.08	-0.133953	-0.141305
96	.	0	1	-1	1	1	0	0.91	0.08	-0.002546	-0.050226
97	.	0	1	-1	0	0	1	0.93	0.23	-0.003111	-0.043065
98	.	0	0	-1	0	0	1	0.93	0.23	-0.135655	-0.184165
99	16	0	0	-1	0	0	1	0.93	0.23	-0.135655	-0.184165
100	.	0	0	1	0	0	1	0.93	0.23	-0.354982	-0.381886
101	5	1	1	-1	1	0	1	0.88	1.00	-3.152800	-4.119918
102	.	0	0	0	0	0	1	0.93	0.23	-0.231553	-0.265198
103	.	0	1	-1	0	0	1	0.93	0.23	-0.003111	-0.043065
104	.	0	1	0	1	0	1	0.93	0.23	-0.017456	-0.135739
105	9	1	0	0	1	0	1	0.87	1.00	-3.413080	-2.179409
106	.	0	0	-1	1	0	1	0.93	0.23	-0.225038	-0.403113
107	.	0	0	0	0	0	1	0.93	0.23	-0.231553	-0.265198
108	12	1	0	0	1	0	1	0.87	0.00	-4.820636	-3.827392
109	.	0	1	-1	1	0	1	0.93	0.23	-0.007334	-0.094263
110	9	1	0	0	1	0	1	0.87	1.00	-3.413080	-2.179409
111	.	0	0	1	0	0	1	0.93	0.23	-0.354982	-0.381886
112	.	0	1	-1	0	0	1	0.93	0.23	-0.003111	-0.043065
113	.	0	1	-1	0	0	1	0.93	0.23	-0.003111	-0.043065
114	.	0	1	0	0	0	1	0.93	0.23	-0.007730	-0.062013
115	.	0	1	0	1	0	1	0.93	0.23	-0.017456	-0.135739
116	16	0	1	0	0	0	1	0.93	0.23	-0.007730	-0.062013
117	.	0	0	0	1	0	1	0.93	0.23	-0.347145	-0.580484
118	10	1	0	0	0	0	1	0.87	0.00	-5.271302	-4.259968
119	.	0	1	0	0	0	1	0.93	0.23	-0.007730	-0.062013
120	.	0	0	-1	1	0	1	0.93	0.23	-0.225038	-0.403113
121	6	1	0	-1	1	0	1	0.87	1.00	-3.081644	-2.275868
122	.	0	1	0	0	0	1	0.93	0.23	-0.007730	-0.062013
123	9	1	1	0	0	0	1	0.87	1.00	-3.542122	-4.001689
124	.	0	1	0	1	0	1	0.93	0.23	-0.017456	-0.135739
125	.	0	1	-1	0	0	1	0.93	0.23	-0.003111	-0.043065
126	.	0	1	-1	0	0	1	0.93	0.23	-0.003111	-0.043065
127	.	0	1	-1	0	0	1	0.93	0.23	-0.003111	-0.043065
128	.	0	0	-1	0	0	1	0.93	0.23	-0.135655	-0.184165
129	.	0	1	-1	1	0	1	0.93	0.23	-0.007334	-0.094263
130	.	0	1	1	1	0	1	0.93	0.23	-0.038995	-0.195465
131	.	0	1	0	0	0	1	0.93	0.23	-0.007730	-0.062013
132	.	0	1	-1	1	0	1	0.93	0.23	-0.007334	-0.094263
133	.	0	1	-1	0	0	1	0.93	0.23	-0.003111	-0.043065
134	.	0	1	-1	0	0	1	0.93	0.23	-0.003111	-0.043065
135	.	0	0	-1	0	0	1	0.93	0.23	-0.135655	-0.184165
136	2	1	0	-1	0	0	0	0.75	1.00	-3.147942	-5.447372
137	.	0	0	0	1	0	0	0.75	0.15	-0.001751	-0.059972
138	.	0	0	-1	0	0	0	0.75	0.15	-0.000169	-0.019027
139	.	0	0	-1	0	0	0	0.75	0.15	-0.000169	-0.019027
140	.	0	0	-1	0	0	0	0.75	0.15	-0.000169	-0.019027
141	.	0	1	0	0	0	0	0.75	0.15	-1.609E-6	-0.006407
142	.	0	0	0	1	0	0	0.75	0.15	-0.001751	-0.059972

143	.	0	0	0	0	0	0	0.75	0.15	-0.000570	-0.027399
144	.	0	0	-1	1	0	0	0.75	0.15	-0.000531	-0.041647
145	.	0	0	1	0	0	0	0.75	0.15	-0.001877	-0.039454
146	.	0	1	0	0	0	0	0.75	0.15	-1.609E-6	-0.006407
147	.	0	1	-1	0	0	0	0.75	0.15	-4.654E-7	-0.004449
148	.	0	1	0	0	0	0	0.75	0.15	-1.609E-6	-0.006407
149	.	0	0	-1	1	0	0	0.75	0.15	-0.000531	-0.041647
150	.	0	1	-1	1	0	0	0.75	0.15	-1.496E-6	-0.009739
151	.	0	1	-1	0	0	0	0.75	0.15	-4.654E-7	-0.004449
152	.	0	0	-1	1	0	0	0.75	0.15	-0.000531	-0.041647
153	.	0	0	0	0	0	0	0.75	0.15	-0.000570	-0.027399
154	.	0	0	-1	1	0	0	0.75	0.15	-0.000531	-0.041647
155	.	0	0	0	1	0	0	0.75	0.15	-0.001751	-0.059972
156	.	0	0	0	0	0	0	0.75	0.15	-0.000570	-0.027399
157	.	0	0	0	0	0	0	0.75	0.15	-0.000570	-0.027399
158	.	0	1	0	1	0	0	0.75	0.15	-5.167E-6	-0.014024
159	.	0	1	-1	0	0	0	0.75	0.15	-4.654E-7	-0.004449
160	.	0	0	1	1	0	0	0.75	0.15	-0.005559	-0.086360
161	.	0	1	0	0	0	0	0.75	0.15	-1.609E-6	-0.006407
162	.	0	0	0	0	0	0	0.75	0.15	-0.000570	-0.027399
163	.	0	1	-1	1	0	0	0.75	0.15	-1.496E-6	-0.009739
164	23	0	0	0	0	0	0	0.75	0.15	-0.000570	-0.027399
165	.	0	0	1	1	0	0	0.75	0.15	-0.005559	-0.086360
166	.	0	0	0	1	0	0	0.75	0.15	-0.001751	-0.059972
167	.	0	1	-1	0	0	0	0.75	0.15	-4.654E-7	-0.004449
168	.	0	0	-1	0	0	0	0.75	0.15	-0.000169	-0.019027
169	.	0	0	-1	1	0	0	0.75	0.15	-0.000531	-0.041647
170	.	0	0	1	1	0	0	0.75	0.15	-0.005559	-0.086360
171	.	0	1	-1	0	0	0	0.75	0.15	-4.654E-7	-0.004449
172	2	1	0	0	1	0	0	0.75	1.00	-2.768869	-4.318041
173	.	0	0	1	1	0	0	0.75	0.15	-0.005559	-0.086360
174	.	0	1	0	0	0	0	0.75	0.15	-1.609E-6	-0.006407
175	32	0	0	-1	1	0	0	0.75	0.15	-0.000531	-0.041647
176	.	0	1	-1	1	0	0	0.75	0.15	-1.496E-6	-0.009739
177	.	0	0	0	1	0	0	0.75	0.15	-0.001751	-0.059972
178	.	0	1	0	0	0	0	0.75	0.15	-1.609E-6	-0.006407
179	.	0	0	1	0	0	0	0.75	0.15	-0.001877	-0.039454
180	.	0	1	1	1	0	0	0.75	0.15	-0.000018	-0.020194
181	.	0	0	1	1	0	0	0.75	0.15	-0.005559	-0.086360
182	.	0	0	1	0	0	0	0.75	0.15	-0.001877	-0.039454
183	.	0	0	0	0	0	0	0.75	0.15	-0.000570	-0.027399
184	.	0	0	0	1	0	0	0.75	0.15	-0.001751	-0.059972
185	.	0	1	-1	1	0	0	0.75	0.15	-1.496E-6	-0.009739
186	.	0	1	0	0	0	0	0.75	0.15	-1.609E-6	-0.006407
187	.	0	1	-1	0	0	0	0.75	0.15	-4.654E-7	-0.004449
188	.	0	0	1	0	0	0	0.75	0.15	-0.001877	-0.039454
189	.	0	1	-1	1	0	0	0.75	0.15	-1.496E-6	-0.009739
190	.	0	0	0	1	0	0	0.75	0.15	-0.001751	-0.059972
191	.	0	1	-1	0	0	0	0.75	0.15	-4.654E-7	-0.004449
192	.	0	0	-1	0	0	0	0.75	0.15	-0.000169	-0.019027
193	.	0	1	-1	0	0	0	0.75	0.15	-4.654E-7	-0.004449
194	.	0	0	1	0	0	0	0.75	0.15	-0.001877	-0.039454
195	.	0	1	1	1	0	0	0.75	0.15	-0.000018	-0.020194
196	.	0	1	-1	0	0	0	0.75	0.15	-4.654E-7	-0.004449
197	.	0	1	-1	0	0	0	0.75	0.15	-4.654E-7	-0.004449
198	.	0	0	1	1	0	0	0.75	0.15	-0.005559	-0.086360
199	.	0	0	1	0	0	0	0.75	0.15	-0.001877	-0.039454
200	.	0	0	0	1	0	0	0.75	0.15	-0.001751	-0.059972

Bibliography

- [1] Amemiya, T. (1985), *Advanced Econometrics*, Oxford: Basil Blackwell.
- [2] Bass, F. M. (1969), "A New Product Growth Model for Consumer Durables," *Management Science*, 15 (January), 215–227.
- [3] Ben-Akiva, M. E. and S. R. Lerman (1985), *Discrete Choice Analysis: Theory and Application to Travel Demand*, MIT Press Series in Transportation Studies, 9, Cambridge, MA: The MIT Press.
- [4] Chatterjee, R. and J. Eliashberg (1990), "The Innovation Diffusion Process in a Heterogeneous Population: A Micromodeling Approach," *Management Science*, 36 (September), 1057–1079.
- [5] Chiang, J. (1991), "A Simultaneous Approach to the Whether, What and How Much to Buy Questions," *Marketing Science*, 10 (Fall), 297–315.
- [6] Chintagunta, P. K. (1993), "Investigating Purchase Incidence, Brand Choice and Purchase Quantity Decisions of Households," *Marketing Science*, 12 (Spring), 184–207.
- [7] Cox, D. R. and D. Oaks (1984), *Analysis of Survival Data*, Monographs on Statistics and Applied Probability, 21, London: Chapman and Hall.
- [8] Deaton, A. and J. Muellbauer (1980), *Economics and Consumer Behavior*, Cambridge: Cambridge University Press.
- [9] Dennis, J. E., Jr. and R. B. Schnabel (1989), "A View of Unconstrained Optimization," in *Optimization* (Nemhauser, G. L., A. H. G. R. Kan and M. J. Todd, eds.), Handbooks in Operations Research and Management Science, 1, Amsterdam: Elsevier Science Publishers.
- [10] Dolan, R. J. and A. P. Jeuland (1981), "Experience Curves and Dynamic Demand Models: Implications for Optimal Pricing Strategies," *Journal of Marketing*, 45 (Winter), 52–62.
- [11] Dolan, R. J. and A. P. Jeuland (1986), "Models of New Product Diffusion: Extension to Competition against Existing and Potential Firms over Time," in *Innovation Diffusion Models of New Product Acceptance* (Mahajan, V. and Y. Wind, eds.), Series on Econometrics and Management Sciences, 5, Cambridge, MA: Ballinger Publishing Company.
- [12] Ehrenberg, A. S. C. (1988), *Repeat-Buying: Facts, Theory and Applications*, London: Charles Griffin and Company, new edition.

- [13] Eliashberg, J. and R. Chatterjee (1986), "Stochastic Issues in Innovation Diffusion Models," in *Innovation Diffusion Models of New Product Acceptance* (Mahajan, V. and Y. Wind, eds.), Series on Econometrics and Management Sciences, 5, Cambridge, MA: Ballinger Publishing Company.
- [14] Eskin, G. J. (1973), "Dynamic Forecasts of New Product Demand Using a Depth of Repeat Model," *Journal of Marketing Research*, 10 (May), 115-129.
- [15] Fourt, L. A. and J. W. Woodlock (1960), "Early Prediction of Market Success for New Grocery Products," *Journal of Marketing*, 25 (October), 30-38.
- [16] Gatignon, H. A. and T. S. Robertson (1986), "Integration of Consumer Diffusion Theory and Diffusion Models: New Research Directions," in *Innovation Diffusion Models of New Product Acceptance* (Mahajan, V. and Y. Wind, eds.), Series on Econometrics and Management Sciences, 5, Cambridge, MA: Ballinger Publishing Company.
- [17] Guadagni, P. M. and J. D. C. Little (1983), "A Logit Model of Brand Choice Calibrated on Scanner Data," *Marketing Science*, 2 (Summer), 203-238.
- [18] Gupta, S. (1988), "Impact of Sales Promotions on When, What, and How Much to Buy," *Journal of Marketing Research*, 25 (November), 342-355.
- [19] Gupta, S. (1991), "Stochastic Models of Interpurchase Time with Time-Dependent Covariates," *Journal of Marketing Research*, 28 (February), 1-15.
- [20] Hanemann, W. M. (1984), "Discrete/Continuous Models of Consumer Demand," *Econometrica*, 52 (May), 541-561.
- [21] Hartmann, W. M. (1990), "Nonlinear Optimization with the SAS System," Working Paper, SAS Institute, Cary, NC.
- [22] Hauser, J. R. and S. P. Gaskin (1984), "Application of the Defender Consumer Model," *Marketing Science*, 3 (Fall), 327-351.
- [23] Hauser, J. R. and G. L. Urban (1986), "The Value Priority Hypotheses for Consumer Budget Plans," *Journal of Consumer Research*, 12 (March), 446-462.
- [24] Hauser, J. R. and K. Wisniewski (1982), "Dynamic Analysis of Consumer Response to Marketing Strategies," *Management Science*, 28 (May), 455-486.
- [25] Heckman, J. J. and B. Singer (1984), "Econometric Duration Analysis," *Journal of Econometrics*, 24, 63-132.
- [26] Helsen, K. and D. Schmittlein (1993), "Analyzing Duration Times in Marketing: Evidence for the Effectiveness of Hazard Rate Models," *Marketing Science*, 11 (Fall), 395-414.
- [27] Horsky, D. and L. Simon (1983), "Advertising and the Diffusion of New Products," *Marketing Science*, 2 (Winter), 1-17.
- [28] Jain, D. C. and N. J. Vilcassim (1991), "Investigating Household Purchase Timing Decisions: A Conditional Hazard Function Approach," *Marketing Science*, 10 (Winter), 1-23.

- [29] Jensen, R. (1982), "Adoption and Diffusion of an Innovation of Uncertain Profitability," *Journal of Economic Theory*, 27, 182–193.
- [30] Kalbfleisch, J. D. and R. L. Prentice (1980), *The Statistical Analysis of Failure Time Data*, Wiley Series in Probability and Mathematical Statistics, NY: John Wiley and Sons.
- [31] Kalish, S. (1985), "A New Product Adoption Model with Price, Advertising, and Uncertainty," *Management Science*, 31 (December), 1569–1585.
- [32] Kalish, S. and G. L. Lilien (1983), "Optimal Price Subsidy for Accelerating the Diffusion of Innovations," *Marketing Science*, 2 (Fall), 407–420.
- [33] Kalish, S. and G. L. Lilien (1986), "Applications of Innovation Diffusion Models in Marketing," in *Innovation Diffusion Models of New Product Acceptance* (Mahajan, V. and Y. Wind, eds.), Series on Econometrics and Management Sciences, 5, Cambridge, MA: Ballinger Publishing Company.
- [34] Kalish, S. and S. K. Sen (1986), "Diffusion Models and the Marketing Mix for Single Products," in *Innovation Diffusion Models of New Product Acceptance* (Mahajan, V. and Y. Wind, eds.), Series on Econometrics and Management Sciences, 5, Cambridge, MA: Ballinger Publishing Company.
- [35] Keeney, R. L. and H. Raiffa (1976), *Decision Making with Multiple Objectives: Preferences and Value Tradeoffs*, NY: John Wiley and Sons.
- [36] Kotler, P. (1980), *Marketing Management: Analysis, Planning, and Control*, Englewood Cliffs, NJ: Prentice Hall, fourth edition. 村田昭治監訳 (1983), 『マーケティング・マネジメント: 競争的戦略時代の発想と転換』, プレジデント社, 第4版.
- [37] Lancaster, T. (1990), *The Econometric Analysis of Transition Data*, Econometric Society Monographs, 17, Cambridge: Cambridge University Press.
- [38] Lilien, G. L., P. Kotler and K. S. Moorthy (1992), *Marketing Models*, Englewood Cliffs, NJ: Prentice Hall.
- [39] Maddala, G. S. (1983), *Limited-Dependent and Qualitative Variables in Econometrics*, Econometric Society Monographs in Quantitative Economics, 3, Cambridge: Cambridge University Press.
- [40] Mahajan, V., E. Muller and F. M. Bass (1990), "New Product Diffusion Models in Marketing: A Review and Directions for Research," *Journal of Marketing*, 54 (January), 1–26.
- [41] Mahajan, V. and Y. Wind (1986), "Innovation Diffusion Models of New Product Acceptance: A Reexamination," in *Innovation Diffusion Models of New Product Acceptance* (Mahajan, V. and Y. Wind, eds.), Series on Econometrics and Management Sciences, 5, Cambridge, MA: Ballinger Publishing Company.
- [42] Malhotra, N. K. (1984), "Analyzing Marketing Research Data with Incomplete Information on the Dependent Variable," *Journal of Marketing Research*, 24 (February), 74–84.
- [43] Massy, W. F., D. B. Montgomery and D. G. Morrison (1970), *Stochastic Models of Buying Behavior*, Cambridge, MA: The MIT Press.

- [44] McFadden, D. (1974), "Conditional Logit Analysis of Qualitative Choice Behavior," in *Frontiers in Econometrics* (Zarembka, P., ed.), NY: Academic Press.
- [45] Nakanishi, M. (1973), "Advertising and Promotion Effects on Consumer Response to New Products," *Journal of Marketing Research*, 10 (August), 242-249.
- [46] Nelson, W. (1982), *Applied Life Data Analysis*, NY: John Wiley and Sons. 奥野忠一監訳 (1988), 『寿命データの解析』, 日科技連.
- [47] Neslin, S. A., C. Henderson and J. Quelch (1985), "Consumer Promotions and the Acceleration of Product Purchases," *Marketing Science*, 4 (Spring), 147-165.
- [48] Oren, S. S. and R. G. Schwartz (1988), "Diffusion of New Products in Risk-Sensitive Markets," *Journal of Forecasting*, 7, 273-287.
- [49] Parfitt, J. H. and B. J. K. Collins (1968), "Use of Consumer Panels for Brand Share Prediction," *Journal of Marketing Research*, 5 (May), 131-146.
- [50] Richard, M. D. and A. W. Allaway (1993), "A Discrete-Time Hazard Model of Shopping Event Attendance," *Journal of Hospitality and Leisure Marketing*, 1 (4), 61-79.
- [51] Roberts, J. H. and G. L. Urban (1988), "Modeling Multiattribute Utility, Risk, and Belief Dynamics for New Consumer Durable Brand Choice," *Management Science*, 34 (February), 167-185.
- [52] Robinson, B. and C. Lakhani (1975), "Dynamic Price Models for New Product Planning," *Management Science*, 21 (June), 1113-1122.
- [53] Rosenbloom, B. (1990), *Marketing Channels: A Management View*, Hinsdale, IL: The Dryden Press, fourth edition.
- [54] SAS Institute (1988), *SAS/IML User's Guide*, Cary, NC: SAS Institute, release 6.03 edition.
- [55] Schiffman, L. G. and L. L. Kanuk (1991), *Consumer Behavior*, Englewood Cliffs, NJ: Prentice Hall, fourth edition.
- [56] Sinha, R. K. and M. Chandrashekar (1992), "A Split Hazard Model for Analyzing the Diffusion of Innovations," *Journal of Marketing Research*, 29 (February), 116-127.
- [57] Urban, G. L., J. R. Hauser and N. Dholakia (1987), *Essentials of New Product Management*, Englewood Cliffs, NJ: Prentice Hall. 林廣茂・中島望・小川孔輔・山中正彦訳 (1989), 『プロダクトマネジメント: 新製品開発のための戦略的マーケティング』, プレジデント社.
- [58] Varian, H. R. (1984), *Microeconomic Analysis*, NY: W. W. Norton and Company, second edition. 佐藤隆三・三野和雄訳 (1986), 『ミクロ経済分析』, 勁草書房.
- [59] Vilcassim, N. J. and D. C. Jain (1991), "Modeling Purchase-Timing and Brand-Switching Behavior Incorporating Explanatory Variables and Unobserved Heterogeneity," *Journal of Marketing Research*, 28 (February), 29-40.
- [60] Wheat, R. D. and D. G. Morrison (1990), "Assessing Purchase Timing Models: Whether or Not is Preferable to When," *Marketing Science*, 9 (Spring), 162-170.

- [61] 相原博之(1992), 「VRホームスキャン・データ分析事例(20):新製品市場浸透状況の概観」, 『VRダイジェスト』, 9-13.
- [62] 青木幸弘(1993), 「スキャナー・パネル・データと消費者行動分析」, 『POSとマーケティング戦略』(小川孔輔編), 有斐閣.
- [63] 阿部俊一(1987), 『システム信頼性解析法』, 日科技連.
- [64] 伊理正夫・藤野和建(1985), 『数値計算の常識』, 共立出版.
- [65] 上田隆穂・江原淳(1992), 『マーケティング』, 新経営学ライブラリ, 8, 新世社.
- [66] 片平秀貴(1987), 『マーケティング・サイエンス』, 東京大学出版会.
- [67] 岸本淳司(1993), 「非線形最適化を行うSASの新プロシジャNLPの紹介」, 『オペレーションズ・リサーチ』, 38(3), 142-145.
- [68] 小山昭雄・蔵田久作・岡本哲治(1975), 『経済数学』, 有斐閣ブックス, 314, 有斐閣.
- [69] 今野浩・山下浩(1978), 『非線形計画法』, ORライブラリー, 6, 日科技連.
- [70] 坂本慶行・石黒真木夫・北川源四郎(1983), 『情報量統計学』, 情報科学講座, A-5-4, 共立出版.
- [71] 佐和隆光(1979), 『回帰分析』, 統計ライブラリー, 朝倉書店.
- [72] 杉田善弘・中村博・田島博和(1993), 「非集計レベルのデータを用いた新製品トライアル購買モデル」, 『マーケティング・サイエンス』, 2(1-2), 32-45.
- [73] 竹内啓ほか編(1989), 『統計学辞典』, 東洋経済新報社.
- [74] 竹村彰通(1991), 『現代数理統計学』, 創文社現代経済学選書, 8, 創文社.
- [75] 田島博和(1993), 「新製品のトライアル購買時期と究極的普及率の測定」, 『日本SASユーザー会第12回論文集』, SASインスティテュートジャパン.
- [76] 東京大学教養学部統計学教室編(1991), 『統計学入門』, 基礎統計学, 1, 東京大学出版会.
- [77] 東京大学教養学部統計学教室編(1992), 『自然科学の統計学』, 基礎統計学, 3, 東京大学出版会.
- [78] 中林三平(1993), 「SAS/IMLのNLP機能を利用したポートフォリオ最適化」, 『日本SASユーザー会第12回論文集』, SASインスティテュートジャパン.
- [79] 『日経流通新聞』, 1992年10月29日.
- [80] 『日本経済新聞』, 1992年2月6日.
- [81] 『日本経済新聞』, 1992年3月20日.
- [82] 『日本経済新聞』, 1992年8月24日.
- [83] 『日本経済新聞』, 1992年10月8日.
- [84] 『日本経済新聞』, 1992年12月1日.
- [85] 『日本経済新聞』, 1993年1月16日.

- [86] 『日本経済新聞』, 1993年10月8日.
- [87] 畠中道雄 (1991), 『計量経済学の方法』, 創文社現代経済学選書, 4, 創文社.
- [88] 宮川雅巳 (1987), 「EM アルゴリズムとその周辺」, 『応用統計学』, 16 (1), 1-19.
- [89] 流通情報開発共同研究機構 (1993), 「CATEGORY FACTS BOOK 本編」, (財)流通経済研究所.
- [90] 流通情報開発共同研究機構 (1993), 「初年度報告書」, (財)流通経済研究所.
- [91] 流通ソフト連合研究開発機構 (1992), 「最終年度報告書」, (財)流通経済研究所.
- [92] 渡辺隆之 (1989), 「購買意志決定時における広告と店頭刺激の効果」, 『店頭研究と消費者行動分析: 店舗内購買行動分析とその周辺』 (田島義博・青木幸弘編), 誠文堂新光社.