

ホーム・スキャン・データを用いた食卓メニュー分析

02401770 東京理科大学 永井 亮雄 NAGAI Fusao, 02401460 生田目 崇 NAMATAME Takashi
01404540 (株) NTT データ *中川慶一郎 NAKAGAWA Keiichiro, 高橋 浩之 TAKAHASHI Hiroyuki
早稲田大学 向井 健太 MUKAI Kenta, 田畑 智章 TABATA Tomoaki

1. はじめに

本稿は、日本オペレーションズ・リサーチ学会マーケティング・モデル研究部会においておこなわれた第5回データ解析コンペティションの報告である。本コンペは、購買/献立メニュー/使用食材に関するホーム・スキャン・データ(食MAPTM¹)を用いて献立メニューと購買行動の構造を把握するモデルを提案し、その独自性を競うものである。

我々は日々の夕食に着目し、献立メニューおよびその材料を考慮したモデルを提案する。

2. 使用データ

ある世帯のメニューおよび、使用材料に関する1年間の履歴を用いて解析をおこなった。なお、データに含まれる項目は以下の通りである。

モニタ世帯データ：

モニタ ID, 家族人数, 共働き状況, 合計世帯年収
世帯メンバ・データ：

モニタ ID, 続柄コード, 性別, 年齢, 職業

購買データ：

モニタ ID, JAN コード, 購入日, 購入価格,
購入数量, 購入先, JICFS コード, 材料コード

献立メニュー・材料・商品データ：

モニタ ID, 飲食種別, 飲食日付, メニュー ID,
メニュー通番, 材料コード, JAN コード,
使い切りフラグ

3. 概念モデル

本研究では、各世帯は以下に示す概念モデル(図. 1)に従って献立メニューを決定するものとして分析を行う。

Stage.1: その日の食卓のフレーム・ワークとして主食と主菜の組み合わせを決定する。

Stage.2: 選択された主食と主菜の組み合わせに該当する主菜のメニュー・アイテムを決定する。

Stage.3: 主食, 主菜との調和を考えた副菜のメニュー・アイテムを決定する。

Stage.4: 選択された主菜, 副菜のメニュー・アイテムに必要な材料を決定する。

また、メニューの履歴を考えると、日々のメニューの移り変わりは特に直前のメニューから影響を受けると考えられる。よって、Stage 3 までで決定されたメニュー・アイテムが将来のメニューに何らかの影響を及ぼすと考えられるが、このことは分析の対象外とする。

4. 分析モデル

本研究では、概念モデルの各段階について、以下の分析を行う。

4.1. Stage.1

Stage.1 では各夕食における主食と主菜の組み合わせが将来の組み合わせに対してどのような影響を及ぼすかという点に着目する。そこで、その影響を測定するために、同じ組み合わせの発生間隔について考察する。なお、データより主食, 主菜を以下のように分類した。

主食: ご飯, 丼, 麺類, その他, なし

主菜: 肉類, 魚類, その他, なし

これら主食・主菜がどのような間隔で選択されるかについてハザード率を考える。たとえば、主菜に関するハザード率は次のようになる。

図. 2 を見ると、肉類と魚類ではハザード率のグラフが大きく異なる。肉類の場合は間隔に寄らず高いハザード率を保っているが、魚類の場合はハザード率が増加する傾向が見られる。主食についても、上記の分類に関して異なったグラフを確認できる。これは、主食と主菜の組み合わせを考えてみても、例えば「ご飯と肉料理」の組み合わせの直後は健康的な理由や食事のバランスなど考え同じ組み合わせを避けるが、時間間隔とともにまた同じ組み合わせの食事をしたくなるというような現象に相当する。

本稿では主食と主菜の組み合わせにハザード・モデルを適用するが、上記の事実からハザード・モデルの密度関数にワイブル分布を用いることとする。ワイブル分布は指数分布を一般化したものであり、ハザード率が時間とともに増加または減少するような現象を説明できる。ワイブル分布の累積密度関数は以下のようになる。

$$F(t; \alpha, \gamma) = 1 - \exp \left\{ - \left(\frac{t}{\alpha} \right)^\gamma \right\} \quad (1)$$

¹食MAPTM は (株) NTT データの商標です。

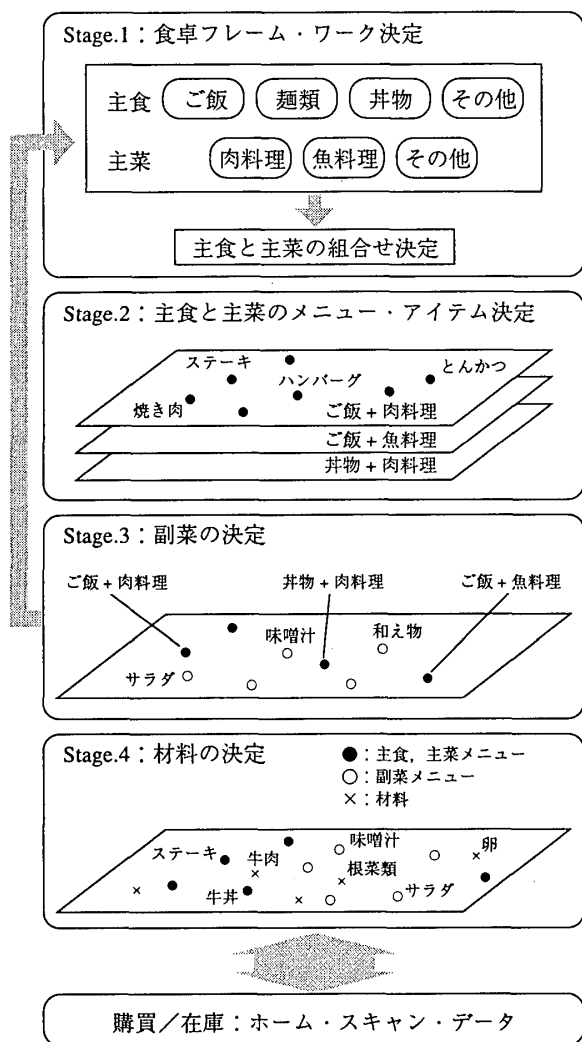


図. 1. 献立メニュー決定の概念モデル

ただし、 t は経過時間、 α は形状パラメータ、 γ は尺度パラメータである。 $0 < \gamma < 1$ の場合は DHR (Decreasing Hazard Rate)、 $\gamma > 1$ の場合は IHR (Increasing Hazard Rate) である。

4.2. Stage.2

Stage.2 では Stage.1 において決定された組み合わせの中で主菜の個別メニュー・アイテム間の競合状態を考察する。ここでいうメニュー・アイテムとは、肉類でいえば「牛肉の炒めもの」といった具体的なメニューである。このとき、あるメニュー・アイテムが選択された直後に同じアイテムが選択される確率は低く、時間間隔に応じて選択される確率が高くなるものとする。これより、個人の選択モデルとして多項ロジット・モデルを用いる。なお、説明変数は各メニュー・アイテムが選択されてからの時間間隔とする。

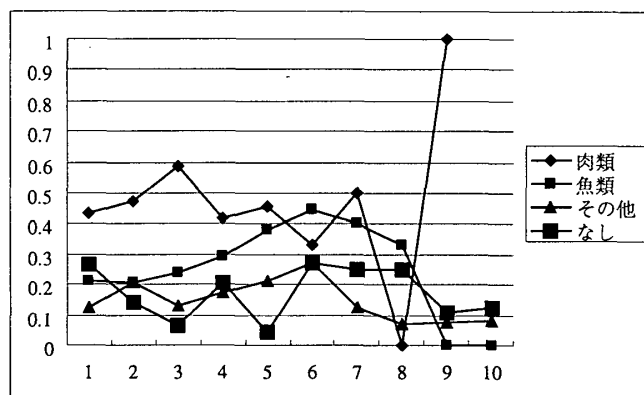


図. 2. ある世帯の主菜のハザード率

4.3. Stage.3

Stage.3 では、Stage.2 で選択された主食・主菜と副菜 (味噌汁, サラダ等) の組み合わせについて考える。それぞれの主食・主菜と副菜について数量化 III 類と数量化 IV 類を用いてそれぞれのアイテム同士の親近性を図示する。

4.4. Stage.4

Stage.4 では、Stage.3 までで選択された食卓メニューに対する材料の選択をおこなうと想定する。メニューと材料について、数量化 III 類を用いてそれらの親近性を図示する。

5. 分析

詳しい分析結果については、発表時に報告する。

6. おわりに

本稿では、ホーム・スキャン・データを用いた食卓メニュー分析モデルを提案した。本モデルでは、食卓メニューを決定する過程をモデル化し、その各段階においてメニュー決定までの行動を把握するような分析をおこなった。こういった分析を通して、食卓を意識したマーケティング活動の情報の一端を担えるのではないかと期待している。

謝辞

本研究に関して、立教大学の守口剛教授をはじめとするマーケティング・モデル研究部会の皆様から有益なコメントをいただきました。ここに感謝の意を表します。

参考文献

- [1] 田中豊, 垂水共之: 「Windows 版 統計解析ハンドブック 多変量解析」, 共立出版 (1995).
- [2] 竹内啓 (編): 「統計学辞典」, 東洋経済新報社 (1989).