

ある本屋における顧客に関する統計的分析

2000MM091 寺田 昌美

指導教員 木村 美善

1 はじめに

私がアルバイトをしている本屋は専門店街の中にあり、様々な年代の顧客が出入りする。映画館もあるので待ち時間に利用する人も多いようだ。アルバイトを始めて約2年経つが、本来16時から出勤の私も店側の事情によりいろいろな時間帯に出勤してきた。これにより時間帯によって客層が違い、買うものに違いがあるのではないかと思い始めた。データによって本当に違いがあるのかを明らかにすることが本研究の目的である。数量化法では[1]、[2]を参照し、クラスター分析では[2]を参照した。

2 使用データの説明

データは8月・11月の1日から7日までの1週間を開店の10時から閉店の21時まで採ったものである。サンプル数は8月は本6937(平4362, 休2575)、雑誌3457(平2434, 休1023)、11月は本6567(平2539, 休4028)、雑誌3544(平1483, 休2061)である。ただし、数冊買った人は冊数分同じ人がいることにした。時間は1時間毎に分け11カテゴリーとした。部門は本を21カテゴリー、雑誌を23カテゴリーに分けた。客層は数量I類・判別分析では男女別にして年齢を5段階に分け合計10カテゴリーとし、数量化II類・クラスター分析では男女で分け2カテゴリーとした。

3 数量化I類

1人当たりの購入金額を外的基準とし、アイテムを平日休日・購入時間・客層・部門の4つとした。本の分析結果を表1に示す。

表1: 本

項目	偏相関係数	範囲
客層	0.107	144.254
部門	0.598	761.130
平日休日	0.024	23.388
時間	0.033	43.198
寄与率: 0.392		定数項: 805.583

雑誌の寄与率は0.236を示した。考察は以下である。

3.1 部門

偏相関係数は雑誌では、0.445を示した。

本の部門をいくつか取り上げる。1(コミック)については子供から大人までの客層が見込まれるので小学生でも無理のない単価に設定してあると考えられる。2(文庫)や8(新書)については内容やカバーやサイズから考えてもこれ以上高くなると顧客が離れてしまう可能性が出てくると考えられる。正の値が大きい7(ビジネス)や15(パソコン)は専門的な知識が多く含まれているので単価が高くなっているのは妥当だと考えられる。16(インテリア)、17(ペット・園芸)も同じように専門的な知識が多く含まれているし、カラーの部分も多いので単価が高く

なってしまうと推測できる。

雑誌の部門をいくつか取り上げる。4(TV雑誌・週刊誌)は1週間分の情報しか載っていないのが主なので情報量から考えても妥当な単価だと言える。3(ガイド誌)に関してもある地区の情報のみやテーマごとの情報しか載っていないものが多いので、買いたいものだけ手軽に買えるという点でもこの単価で妥当だと言える。15(ホビー)については、ファンの人にとってはうれしい知識や情報が多く含まれていることや付録に豪華なものが多いことから単価が高いのは当然と考えられる。

3.2 客層

偏相関係数は雑誌では、0.164を示した。

本について10代男女が他と比べて負の値を大きくとっているのは、金銭的に余裕の無い年代であるからと考えられる。使えるお金に限られるので、単価の低い商品を好んで買い、高い商品は親に任せることが考えられるからである。また、20代女性が負の値をとっているのも夏休みなので学生が多くなり、金銭的に余裕の無い人が増えたからと考えられる。反対に20代男性が正の値をとっているのは女性に比べて働いている人が多いと思われるので、学生が多くなる夏休みでも正の値をとったと考えられる。男女で比べると男性の値が大きいのも男性の方が仕事関係の単価の高い専門的な本を買った率が高くなるからということが推測できる。また、男性の方が値段を気にせず内容で商品を買っているように感じた。

雑誌については10代20代が男女共に負の値をとっている。これは本で述べたことと同じように、この年代が金銭的に余裕が無い年代であると考えられるので、出版社側が無理のない単価を設定しており、購入しやすくしていると考えられる。また、まだ興味が無いのか単価の高いインテリア関係の雑誌や趣味関係の雑誌、そして、ビジネスやパソコン等の雑誌をあまり買わないので、自然と値が下がったと考えられる。男女の差はあまり見られなかった。

3.3 平日休日・時間

偏相関係数は本雑誌ともに0.1未満だった。

あまり差は見られなかった。

4 数量化II類

客層を外的基準とし、アイテムを平日休日・購入時間・部門の3つとした。

各結果の相関比は本は0.1未満、雑誌は約0.25を示したので雑誌のみを取り上げる。考察は以下のようである。

4.1 部門

偏相関係数は0.485を示した。

2(女性誌・ティーン誌)や6(奥様誌)、8(男性誌・ヤング誌)はそれぞれターゲットを絞っているため男性女性に

分かれるのは当然である。12(育児雑誌)も女性の方子供にかかわる時間が多いと考えられるのでこれも妥当だといえる。9(車・バイク雑誌)や15(ホビー誌)は明らかに男性の方に傾いている。これは幼いときから男女の興味が乗り物やロボットとままごとやぬいぐるみのように分かれることからの延長といえるかもしれない。17(パソコン誌)や18(ビジネス誌)が男性に傾いているのも男性の方が女性よりも仕事に身をおく時間が長いので仕事に対して積極的だからと考えられる。

4.2 平日休日・時間

偏相関係数が0.1未満ということからも明らかなように時間帯についてはあまり差は無い。8月のみ時間の偏相関係数が0.103を示しているのと言うならば、男性が遅い時間に集まっているのは会社帰りということが考えられる。

5 クラスタ分析

8月本、8月雑誌、11月本、11月雑誌をそれぞれ平日と休日を分けた。そして、購入者数の(時間・客層)×(部門)クロス集計表から時間・客層別でユークリッド距離をとった。男性は奇数、女性は偶数となっている。

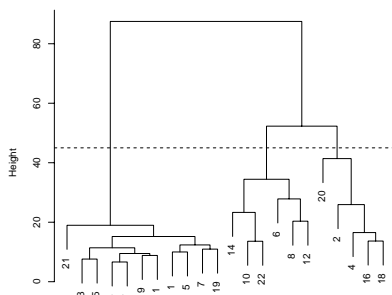


図1: 8月雑誌平日のデンドログラム

5.1 グループ分け

図1を距離45で切り左からそれぞれ1-3群に分けた。

第1群: 全時間帯の男性が集まっていることから、もともと男性の割合が低いので全部門を通して、購入数の少ない群と考えられる。また男性が好む商品の購入数が多くなっている群とも考えられる。

第2群: 昼から夕方にかけての時間帯なので食事や遊びに来る顧客の多い時間帯と考えられる。また、夏休みということで子供が家にいる割合が高くなっていると思われるので、普段より夕食の時間が早まっていることが考えられる。よって、買出しの時間も早まり、この時間帯に顧客が集中したと考えられる。また、20時台は閉店前ということで他の買い物や食事が終わった後に目をつけていたものを買いに来たり、映画待ちをしている人が集まる時間帯なので、全体的に顧客の多い時間帯といえる。よって、全体的によく売れている群だと考えられる。

第3群: 女性が少ない時間というよりも顧客が少ない時間帯が集まっていると思われる。平日の朝なので顧客の

多くを占める主婦にとっては忙しい時間帯だと考えられる。また、17時から20時台には夕食の時間帯を含んでいるので顧客の少なくなる時間帯になってくる。よって、女性のみでならば全体的に購入数の少ない群と推測できる。また、各部門の購入数にばらつきのある群だとも考えられる。

5.2 考察

2群に分けた場合ほぼ片方の群は女性のみとなった。男性よりも女性の割合の方が高くなっているの自然と各部門の購入数が高くなる。これがグループ分けに大きく影響したと考えられる。また、本については女性が好む部門、雑誌については男性が好む部門が多いように思われるので、これにより男女に分かれた可能性も考えられる。よって、部門に関しては多かれ少なかれ男女の違いはあると言える。

時間についてはパターンにより結果はまちまちだった。例えば、全部門がよく売れている群であり、顧客の多い時間帯の群では女性の割合が多いので女性中心になってしまうが、8月雑誌の休日、11月本と11月雑誌の休日では昼から夕方時間帯が集まった。これに対して、8月平日の本と雑誌は昼から夕方にかけてと最後の1時間が集まった。各部門に関してのグループ分けの説明はあまりうまくいかなかったが、部門に対しての購入数から考えて、平日か休日かによって違いが出てくるということがいえたのではないだろうか。

6 判別分析

クラスタ分析で平日休日の違いがみられたが、それをこの分析でも確認する。

部門・客層・時間を組み合わせて、分析を行ったところ(時間)×(客層)組み合わせがきれいに平日休日でわかれた。よって、平日休日に関して違いがあるということがわかった。また、部門に関しては8月と11月とでは平日と休日が逆になった部門が異なったので月によっても違いがあると言えるのではないだろうか。

7 おわりに

本研究では、購入金額に関しては年代や性別によって少しではあるが違いがあると言えたのではないだろうか。部門に関しては特に雑誌になるが、男女の購入商品に違いがあることがわかった。また、研究を始めた頃はあまり気にしていなかったが、幅のある商品単価に対して、およそではあるが顧客に1番買われている単価はどこかを知ることができた。時間に関してはあまりはっきり客層と関係があるとは言えなかったが、時間と購入金額に関係が無いことはわかってきた。また平日か休日かによって顧客の多い時間帯に違いが出てくるということもわかった。

参考文献

- [1] 田中豊・脇本和昌: 多変量統計解析法, 現代数学社, 1983.
- [2] 林知己夫: 数量化-理論と方法, 朝倉書店, 1993.