

刘勇¹，王高²

摘要 尽管离散选择模型在 20 世纪 70 年代末开始应用于交通、能源、住房、通讯及营销等研究领域，但离散选择模型在国内营销研究方面的研究还处于起步阶段。本文采用离散选择模型中 MNL、MNP 以及 Mixed Logit 三种模型，研究妇女卫生巾的销售因素（品牌、价格和人员促销、其他促销）对顾客选择及市场份额的影响，发现品牌及各个促销因素在顾客购买决策时确实起到了积极的影响，尤其是礼品促销等给顾客带来具体利益的促销影响最大（对市场份额的影响呢？）。这为企业优化配置促销资源提供决策依据；同时，本文还根据三种离散选择模型的特点在同一批促销数据上的应用进行比较，发现 Mixed Logit 以及 MNP 比 MNL 模型拟合效果更优。

关键词 离散选择模型，MNL，MNP，Mixed Logit

离散选择模型在零售中的应用

—基于 MNL、MNP 和 Mixed Logit 的比较³

0 引言

“我知道广告费至少有一半被浪费掉了，但问题是，我不知道究竟是哪一半”，约翰·沃纳梅克的名言，一直在美国企业界广为流传。中国企业对此深有体会：随着市场竞争的日益激烈，企业为了赢得顾客，取得市场份额，除了提供高质量的产品，努力打造优秀品牌外，也更多地尝试采用价格折扣、端架促销、堆头促销、人员促销、礼品促销等多种促销手段来促进产品销售，扩大市场份额。众多促销手段对于产品的销售起到了积极的作用。这些不同的促销手段是如何影响产品的销售，如何影响市场份额的变化，则是企业非常关心的问题。例如，假如竞争对手品牌的产品降价 10% 的话，该市场份额会如何发生变化？假如某品牌产品采取人员促销的话，销售收入和市场份额会如何发生变化？

¹ 刘勇，北京工商大学商学，Email: liuyong@sem.tsinghua.edu.cn

² 王高，中欧国际工商学院教授，Email: wgao@ceibs.edu

³ 本论文为国家自然科学基金资助项目“营销变量对消费者品牌选择影响的模型研究（批准号 70572008）”部分研究成果。

企业的资源是有限的，企业总是希望能够通过合理地配置资源，达到收益的最大化。同时，由于不同的产品对应的市场具有不同的特点，自然相应的促销方式也各有不同，有些市

场适合这种促销方式，有些市场适合那种促销方式，企业如何用好这些促销组合，对大多数企业来说，是一件很难判断的事情。国内外学术界对此展开了深入研究，提出各种模型来探索价格、品牌以及各种促销因素对消费者品牌选择以及市场份额变化的影响。而消费者的品牌选择是一个离散的选择过程，离散选择模型（Discrete Choice Model）就得到了广泛的应用。

离散选择模型从决策者效用最大化角度出发来研究决策者的选择问题，广泛应用于包括在交通(Ben, 1972)、营销(Green, 1977)、能源(Train, 1986)、住房(Lee, 1999)、通讯(Train, McFadden 和 Goett, 1987)等研究领域，该模型能够揭示出决策者的微观决策过程，这是决策活动的基础，也是该模型区别于其他分析工具的优势所在。美国加州大学伯克利分校的MacFadden 教授以“对分析离散选择的原理和方法所做出的发展和贡献”获 2000 年诺贝尔经济学奖。

MNL (Multinomial Logit) 是离散选择模型中最基本的一种模型。随着对离散选择模型研究的深入，人们对离散选择模型的参数估计的能力不断增强，离散选择模型中更复杂的模型如嵌套多项式 Logit 模型 (Nested Multinomial Logit, NMNL)、MNP 模型 (Multinomial Probit) (Chintagunta 和 Pradee, 1992)、混合 Logit 模型 (Mixed Logit) 等模型也开始应用于市场营销研究中(Erdem, 1996)。

本文通过分析建模，将离散选择模型中的 MNL、MNP 和 Mixed Logit 模型运用到市场营销的研究中，通过模型的分析对比，选择合适的模型来计算品牌、价格以及各种促销活动对市场份额的影响。另外，本文还将研究不同品牌熟悉程度和不同购买欲望的消费者在购买时是如何受到价格、品牌以及各种促销因素的影响，以深入理解不同类别消费者购买决策时受到的各种因素影响情况，对企业的营销决策有直接的借鉴意义。

1 相关研究介绍

离散选择模型在营销研究方面的运用非常广泛。1977 年 Green, Carmone 和 Wachspress 首次系统地介绍了如何将离散选择模型中的 Logit 模型运用到市场营销研究中，受到了营销研究者的重视：Guadagni, Peter 和 John (1983) 在营销科学杂志上发表了利用 MNL 模型拟合扫描研究消费者对咖啡品牌的选择；Russell, Gary 和 Petersen (2000) 运用嵌套 Logit 模型分析消费者在商品种类的选择问题；Zsolt Sándor 和 Michel (2005) 运用蒙特卡罗方法和贝叶斯模型来研究在离散选择模型中消费者的异质性问题。

我国对离散选择模型的研究起步较晚（聂冲和贾生华，2005），主要是在交通规划领域应用较多。吴中和毛荣昌（2004）对 MNL 模型在交通方式划分中的应用作了描述性介绍；周建等（2006）介绍了嵌套 logit 模型的在交通方式划分的求解方法，并举例说明；刘振和周溪召（2006）对嵌套 logit 模型在交通方式选择行为中的应用做了探讨，并将 4 种交通方式分为私人交通和公共交通两大类，通过 20 个样本的事例来求解模型；王树盛等（2006）研

究了 Mixed Logit 模型在交通方式分担中的应用，并利用 Daganzo1979 年的调查数据，设定分布函数对参数进行估计。

离散选择模型在营销研究的应用文献寥寥无几，在 2008 年 5 月对中国期刊网的文献检索中，只有 7 篇和营销相关的 Logit 模型论文。这些文献偏重于介绍模型方法并利用国外数据进行模拟计算，在我国进行实地数据收集、建立模型进行实证分析方面的研究还非常有限，只有两篇（张群、曹丽和李纯青，2005；李德春和肖岳峰，2007），采用混合 Logit 模型来进行营销研究目前尚是空白。本文采用了 MNL、MNP 和 Mixed Logit 模型进行实证研究，不仅对于指导企业决策有现实意义，而且对我国理论界运用该模型也有积极的促进作用。

2 相关模型介绍

2.1 MNL (Multinomial Logit) 模型

某一决策者 i 对可供选择的选择集 C ，其选择选项 j 的效用 U_{ij} 可以表示为

$$U_{ij} = V_{ij} + \varepsilon_{ij} \quad (1)$$

其中， V_{ij} 为第 i 个决策者选择第 j 个选项的可观测效用，既可以包括选择项本身的属性（如在选择不同的交通出行模式所需要耗费的时间、金钱等），也包括决策者的个人特征；随机扰动项 ε_{ij} 是不可观测到的那部分效用，通常我们假设 ε_{ij} 为独立同分布的随机变量。按照效用最大理论，决策者 i 仅在第 j 选项的效用大于于其他选项的效用时才选择第 j 选项，即

$$V_{ij} + \varepsilon_{ij} > V_{ik} + \varepsilon_{ik}, \forall j \neq k \quad (2)$$

则第 i 决策者选择第 j 项的概率可以表示为

$$P_{ij} = \Pr(V_{ij} + \varepsilon_{ij} > V_{ik} + \varepsilon_{ik}, \forall j \neq k) = \Pr(\varepsilon_{ik} < \varepsilon_{ij} + V_{ij} - V_{ik}, \forall j \neq k) \quad (3)$$

McFadden (1974) 证明，当该随机扰动项独立并服从相同的 Weibull 分布时（又称极值分布或双指数分布），

$$P_{ij} = \frac{\exp(V_{ij})}{\sum_{k=1}^J \exp(V_{ik})} \quad (4)$$

MNL 模型存在三个局限性：（1）口味问题（Taste）。该模型假设选项的同一项属性的效用对于所有的决策者来说都是无差别的。假如不同的决策者对于选项的某一属性的效用权重不同，该模型无法解决；（2）跨期间的重复选择。如个人过去的选择对当前选择产生影响，或属性改变后的滞后反应，该模型均无法处理；（3）替代形式，即“不相关选项间的独立性”

(Independence from Irrelevant Alternative, IIA), 这是 MNL 模型的一个限制性假设, 不同选项之间的可以成比例替代。但是事实上许多情况并非如此, 如 Chipman (1960) 提出著名的“红巴士/蓝巴士”难题 (red bus/ blue bus), 就是 MNL 模型所无法解决的。

该模型参数的估计如今已经有不少软件可以调用程序模块直接处理¹。

2.2 MNP (Multinomial Probit) 模型

和 MNL 假设的效用误差项 ε_{nj} 服从独立同极值 (iid) 分布不同, MNP 模型假设 ε_{nj} 服从均值为 0 和协方差矩阵为 Ω 的正态分布, 则 ε_n 的密度为

$$\phi(\varepsilon_n) = \frac{1}{(2\pi)^{J/2} |\Omega|^{J/2}} e^{-\frac{1}{2}\varepsilon_n'\Omega^{-1}\varepsilon_n} \quad (5)$$

其中, ε' 是由 ε_{nj} 组成的向量, 可以写成 $\varepsilon' = (\varepsilon_{n1}, \varepsilon_{n2}, \dots, \varepsilon_{nJ})$ 。

上式中的协方差矩阵 Ω 依赖决策者所面对的向量, 这样就放松了对模型的独立不相关选项等约束要求 (Train, 2002)。决策者 n 选择选项 i 的概率为

$$\begin{aligned} P_{ni} &= \text{Prob}(V_{ni} + \varepsilon_{ni} > V_{nj} + \varepsilon_{nj}, \forall i \neq j) \\ &= \int I(V_{ni} + \varepsilon_{ni} > V_{nj} + \varepsilon_{nj}, \forall i \neq j) \phi(\varepsilon_n) d\varepsilon_n \end{aligned} \quad (6)$$

其中的 $I(\bullet)$ 为指示性函数, 当括号内的项为真时等于 1, 否则为 0; $\phi(\varepsilon_n)$ 是 ε_n 的概率密度函数, ε_n 是带有 0 均值和协方差矩阵为 Ω 的正态分布, 是不可观测的。这一积分涵盖了所有的值, 是一个非封闭型。该概率无法用积分算出来, 只能够通过模拟 (Simulation) 方法来估计。包括瞬间模拟 (Simulated Moments, Mcfadden, 1989) 极大似然模拟、马尔可夫-蒙特卡罗方法模拟等²。

由于 MNP 模型中假设选项之间的协方差矩阵 Ω 不为零, 选项之间的相似性直接在个人效用函数的随机部分得到体现, 各个选项之间可以存在依存性。因此, 该模型解决了困扰 MNL 模型的三大难题, 即口味问题、跨期间的重复选择和替代形式。

但是, MNP 模型也有一个限制局限性就是要求对所有效用的不可观测部分服从正态分布。在许多情况下, 不可观测变量并不服从正态分布, 比如消费者价格必然为负效用 (消费者不愿意多付钱), 消费者购买某种令他满意的商品时候其价格效用所获得的效用必然为正。因为正态分布密度函数分布在 0 的两侧, 这时候假设不可观测因素服从正态分布就与实际情况相冲突。

¹如 STATA 9 以上版本可以直接用 Conditional Logit 语句进行处理, SAS 可以调用 MDC 和 PHRG, 模块处理, R 软件中也有专门的软件包。

²该模型可以在 SAS9.1 以上版本中直接调用相关模块, 或者从 R2.3.0 统计软件中调用“MNP”软件包。

由于 MNP 模型比较难于估计, 在我国的实证研究的文献中非常少见。通过中国期刊网全文数据库经济类的文献中只查到两篇: 一篇是张群、曹丽和李纯青 (2005) 比较 Logit 和 Probit 在零售业应用的研究, 提出 Probit 更适合于零售业的动态客户关系管理; 另一篇是邹德强、王高、赵平、王燕研究功能性价值和象征性价值对品牌忠诚的影响 (2007)。

2.3 混合 Logit 模型 (Mixed Logit)

Mixed Logit 模型比 MNL 和 MNP 应用范围更加广泛, 约束条件放得更松。MNL 模型因为 IIA 的条件导致了口味问题、时序问题、替代形式等三个问题无法解决, 只能处理消费者的系统性偏好; MNP 模型避免了 MNL 的三大难题但却受限制于误差项分布必须是正态分布; 而 Mixed Logit 模型可以自由设置解释变量系数成为一定的随机分布形式, 从而可以处理消费者的随机性偏好。

Mixed Logit 模型中, 决策者选择 i 的概率可以表示为

$$P_{ni} = \int L_{ni}(\beta) f(\beta) d\beta \quad (7)$$

在式 (7) 中, $L_{ni}(\beta)$ 是在参数 β 下的 logit 概率, 表达式为

$$L_{ni}(\beta) = \frac{e^{V_{ni}(\beta)}}{\sum_{j=1}^J e^{V_{nj}(\beta)}}$$

式 (7 中的) $f(\beta)$ 是密度函数; $V_{ni}(\beta)$ 是在参数 β 基础上的被观察部分的效用; 如果效用和影响因素的关系是线性的, 其系数为 β , 即 $V_{ni}(\beta) = \beta' x_{ni}$ 。那么这时 Mixed Logit 概率为

$$P_{ni} = \int \left(\frac{e^{\beta' x_{ni}}}{\sum_{j=1}^J e^{\beta' x_{nj}}} \right) f(\beta) d\beta \quad (8)$$

混合分布 $f(\beta)$ 也可以是离散的。如果 β 可以是 M 中可能值 b_1, b_2, \dots, b_M , 在 $b = b_m$ 的概率是 s_m , 则 Mixed Logit 模型就成了潜分层 (Latent class) 模型, 该模型目前也广泛应用于心理学和营销学的研究。式 (8) 中概率可写成

$$P_{ni} = \sum_{m=1}^M s_m \left(\frac{e^{b_m' x_{ni}}}{\sum_{j=1}^J e^{b_m' x_{nj}}} \right)$$

Mixed Logit 模型在表达式 (8) 中没有封闭解, 需要通过模拟的方法来进行估计。一般是按照某一分布密度进行抽样生成随机数, 然后把这些随机数代入函数 $L_{ni}(\beta)$ 得到函数值, 最后对这些函数值进行平均计算从而得到选择概率 P_{ni} 的模拟解。

Mixed Logit 模型也是很早就被提出了, 但是一直到近期才得到广泛应用。Boyd 和 Mellman(1980)和 Cardellhe 和 Dunbar(1980)最早将该模型应用于汽车需求分析, 但是他们的模型没有具体到微观层面消费者的选择而只是市场份额这样的宏观数据; 后来, Train(1987)和 Ben-Akiva 等 (1993) 开始将该模型应用到消费者微观选择层面, 但也仅能够通过模拟方式的求解一维和二维的积分; 随着计算机运算速度的增加以及人们模拟估计参数能力的增强, 该模型开始得到充分的应用: Bhat(1998), Brownstone 和 Train 等 (1999) 将该模型应用于截面数据, Erdem(1996)、Revelt 和 Train(1998)、Bhat(2000)将该模型应用于面板数据。

我国对 Mixed Logit 模型的应用非常有限, 在中国优秀博士论文全文数据库以及中国期刊全文数据库中只查找到五篇相关论文, 其中四篇应用于交通规划领域, 一篇用于财务领域估计财务困境, 营销类一篇也没有。而且, 我国这些文献都是 2006 年以后的文献, 其中只有三篇是实证研究。

3 数据收集及编码

3.1 数据收集

本文研究的产品类别是妇女卫生巾, 零售终端是北京一家位于繁华地段的中大型超市, 数据的时间单位是天, 时间范围为共 31 天, POS 机显示的交易行为达 669 次。该数据的内容由两部分构成: 商店的实际销售数据 (销售量和销售额), 来源于商店的电子扫描数据 (POS 数据); 商品的价格和促销数据, 来源于调查数据¹。

各个卫生巾品牌在观察期内, 采用了价格促销、人员促销、广告促销、堆头促销、优惠促销、礼品促销等方式。因为有些促销方式采用的次数比较有限, 对模型的影响不大。我们通过前期的数据分析和整理, 删除某些对品牌影响不显著的因素并把有些促销因素进行合并², 得到影响消费者选择卫生巾品牌的几个变量。(1) 品牌: 品牌是影响消费者购买的重要因素之一, 消费者对品牌偏好程度不同, 将影响他们的购买行为。本文中该超市消费者可选择的卫生巾品牌包括娇爽 (高)、娇爽 (低)、苏菲、护舒宝 (高)、护舒宝 (低)、洁婷、舒而美、安而乐、佳期、安乐、佩安婷、一片云、永芳、ABC、康芙娅、乐芙爽、娜芳、伊倍爽、益母草等 19 种品牌。

¹郭永新、王高、齐二石 (2007) 在《管理科学学报》发表了“品牌、价格和促销对市场份额影响的模型研究”, 运用乘法竞争互动模型 (MCI) 来进行研究。本文数据和上文数据源于同一家超市, 但不是完全相同的一批数据。

²我们将礼品促销中的礼品 (包括卫生巾、卫生护垫、口红、唇膏、长筒袜等) 统称为礼品促销, 将端架、堆头、路演促销为陈列式促销等起到广告的合并为广告促销。

(2) 价格：这里的价格是指消费者实际支付的价格，由销售额除以销售量计算得到。该价格已经把价格促销因素包含进去。调查时卫生巾是以包为单位计量的，由于每包卫生巾装的片数不一样，而且日用、夜用等不一样，为了使价格有可比性，我们对原始数据进行处理时，以单品牌产品的销售总片数除销售总金额得到产品的平均价格。因为在模型研究中，本文使用实际价格，所以折扣促销就没有单独包括在模型中。

(3) 人员促销：指生产企业雇佣的促销人员在现场的促销。

(4) 广告类促销：指某品牌采用广告传单、堆头、端架和人员路演等形式吸引顾客注意力进行促销，起到广告类作用，这里归为广告类促销。

(5) 礼品促销，赠送的礼品包括卫生巾、卫生护垫、口红等，由于每一种具体的礼品采用的时间都比较短，数据比较稀，所以将所有这些促销方式概括为一个变量，礼品促销。

3.2 基本数据编码

在上述变量中，价格是连续变量，而人员促销、广告类促销、赠品促销等，我们采取哑变量的形式进行编码，即当有这种促销时取值为 1，否则为 0。

我们最初得到的数据格式用 excel 整理，一共调查了 669 次购买行为，即有效数据为 669 行。在每一行中的调查的内容包括：购买行为编号 (Subject)、选择的品牌 (Choice, 该行范围选择为 1-19)、各品牌的价格 (从 price1 一直到 price19, 一共 19 列)、各品牌是否人员促销 (promogirl1 一直到 promogirl19, 一共 19 列)、各品牌是否有广告类促销 (从 adpromol 一直到 adpromo19, 一共 19 列)、各品牌是否有礼品类促销 (从 valpromol 一直到 valpromol 一共 19 列)。这样，我们得到一个 669 行、82 列的 excel 的数据表。

3.3 市场特点

表 1 给出了各品牌卫生巾的价格和促销情况等指标。从这些数字中我们可以看出卫生巾市场在该超市有如下的特点：

(1) 市场分散

表 1 显示，在 669 次购买中，选择最多的次数舒而美达到 123 次，占选择次数的 18.4%，选择苏菲、护舒宝和娇爽次数都在 10% 左右，选择其他的 13 个品牌的次数只有 329 次，不到 50%。这说明这说明该店的卫生巾市场是一个比较分散，竞争非常激烈的市场。

(2) 价格差异大

表 1 显示，从单片卫生巾常规价格来看最高价 (1.21 元, ABC) 是最低价 (0.2 元, 永芳) 的 6 倍，价格差异非常大。这说明卫生巾市场是一个产品差异性较大的市场，价格不是唯一的竞争因素。

(3) 促销频繁

在所观察的 30 天内，只有永芳没有进行过任何促销，其它品牌都或多或少进行过促销活动。这说明该市场的促销竞争是非常激烈的，这也从另一个层面反映出了卫生巾市场竞争

的激烈程度。

表 1 各品牌卫生巾的购买频次和价格

品牌编号	品牌名称	购买次数	购买概率 (%)	平均常规价格 (元/片)	平均实际价格 (元/片)
1	娇爽(高)	34	5.08	0.82	0.78
2	娇爽(低)	13	1.94	0.44	0.44
3	苏菲	94	14.05	0.94	0.91
4	护舒宝(高)	42	6.28	1.08	1.08
5	护舒宝(低)	34	5.08	0.53	0.53
6	洁婷	23	3.44	0.87	0.86
7	舒而美	123	18.39	0.48	0.48
8	安而乐	29	4.33	0.64	0.64
9	佳期	31	4.63	0.5	0.5
10	安乐	26	3.89	0.32	0.32
11	佩安婷	29	4.33	0.48	0.48
12	一片云	33	4.93	0.28	0.28
13	永芳	10	1.49	0.2	0.2
14	ABC	11	1.64	1.21	1.21
15	康芙娅	44	6.58	0.41	0.4
16	乐芙爽	8	1.20	0.38	0.38
17	娜芳	22	3.29	0.47	0.46
18	伊倍爽	27	4.04	0.33	0.31
19	益母草	36	5.38	0.68	0.64

4 不同模型的应用

4.1 MNL (Multinomial Logit) 模型应用

我们可将影响顾客品牌选择的因素（品牌、价格、人员促销、广告类促销和礼品类促销）与顾客效用之间的关系表示成线性方程。对顾客 n 而言，他选择品牌 i 得到的可观测效用如式 (9) 和 (10)。

$$U_{ni} = V_{ni} + \varepsilon_{ni} \quad (9)$$

$$V_{ni} = a_1 \times \text{realprice}_{ni} + a_2 \times \text{promogirl}_{ni} + a_3 \times \text{adpromo}_{ni} + a_4 \times \text{vapromo}_{ni} + b_i \times \text{brand}_i \quad (10)$$

从(10)可见,我们需要估计价格、人员促销、广告类促销、礼品类促销和品牌等参数,一共是包括 a_1 、 a_2 、 a_3 、 a_4 和 $b_i(i=1到18)$ 等22个参数。 b_i 只估计到*i*=18是因为把第19个品牌(益母草)作为基础值(0), $b_i(i=1到18)$ 只是其他品牌相对于该品牌多出的效用。

对于MNL模型,可以采用SPSS软件中的Coxreg命令(用于生存分析),SAS软件中的PHREG(生存分析)或MDC程序,或者直接在STATA软件中调用conditional logit命令,估计参数。在本文中,我们利用SAS9.1版本中的MDC程序,采用Newton-Raphson方法估计参数。参数估计和模型拟合情况参见表2。

4.2 MNP模型的应用

在MNP模型中的效用函数和MNL模型中的效用函数表达式相同。在MNP模型中,我们设定的是22个变量,即realprice, promogirl, adpromo, vapromo和bi(*i*=1,2,...,18),以“益母草”为基准值。

MNP模型的选择概率表达式没有解析解,计算这些概率需要多重积分,因此估计起来比较复杂。对于MNP模型的估计,有些学者采取模拟矩(MSM)的方法(Chintagunta,1992; Murthi and Srinivasan,1998);有些学者采取贝叶斯分析的马尔可夫链蒙特卡罗(MCMC)算法(Imaia, Kosuke and Dyk, 2005)。我们利用SAS9.1版本来估计参数,全模型结果如表2。

由于MNP模型和MNL模型中的 $F(\bullet)$ 分别代表不同的累计函数,因此这两个模型的系数之间不能够直接比较。但是,它们可以相互转换。Ben-Akiva and Lerman(1985)指出MNP的系数是相同设定的MNL模型的 $\pi/\sqrt{3}$ 倍,即近似为1.814倍(Aldrich and Nelson,1984)。但是,有些学者(Amemiya,1981; Kmenta,1986)建议将MNP的系数乘以1.6更能够得到接近于MNL回归系数的近似值。如表2中MNP模型的realprice的系数为-7.4915*1.814=-13.59,与MNL模型中的realprice系数-15.35差别不大。

表2 MNL模型(全模型)的参数估计及模型拟合情况

		MNL			MNP			Mixed Logit			
		Estimate	Standard t Value	Approx Pr> t	Estimate	Standard t Value	Approx Pr> t	Parameter	Parameter	Standard t Value	Approx Pr> t
参数	realprice	-15.3588	-3.53	0.0004	-7.4915	12.8913	0.0003	Inrealprice_M	-23.4028	-4.09	<.0001
								Inrealprice_S	15.4538	2.76	0.0058

	promogirl	0.3970	3.29	0.0010	0.1903	11.5605	0.0007	promogirl_M	0.3272	1.58	0.1152
								promogirl_S	1.1488	2.04	0.0418
	VaPromo	4.1932	4.62	<.0001	1.9991	21.6196	<.0001	VaPromo_M	42.3900	2.82	0.0223
								VaPromo_S	-43.5182	2.29	0.0629
	adpromo	0.3853	3.17	0.0015	0.1937	11.0551	0.0009	adpromo	0.4463	-1.86	0.0047
	b1	1.8006	2.69	0.0071	0.8837	7.6442	0.0057	b1	2.9987	2.87	0.0041
	b2	0.000291			0.00639	0.1873	0.6652	b2	1E-6		
	b3	4.5373	3.87	0.0001	2.2860	16.6136	<.0001	b3	5.6461	3.35	0.0008
	b4	6.9535	3.67	0.0002	3.4098	13.9062	0.0002	b4	6.0105	2.26	0.0236
	b5	-1.6929	-3.05	0.0023	-0.8254	9.6923	0.0019	b5	-4.5544	-4.00	<.0001
	b6	2.1761	2.41	0.0158	1.0788	6.1330	0.0133	b6	3.3528	2.54	0.0111
	b7	-1.5839	-2.13	0.0335	-0.6520	3.2743	0.0704	b7	-6.6384	-3.72	0.0002
	b8	-0.1977	-0.79	0.4289	-0.0962	0.6704	0.4129	b8	-0.2126	-0.84	0.3993
	b9	-2.7394	-4.21	<.0001	-1.3360	18.1904	<.0001	b9	-6.8435	-4.77	<.0001
	b10	-0.9440	-1.45	0.1475	-0.5000	2.7384	0.0980	b10	-5.8798	-3.21	0.0013
	b11	-2.9979	-3.77	0.0002	-1.4636	14.9790	0.0001	b11	-9.2362	-4.23	<.0001
	b12	-1.3058	-1.64	0.1007	-0.6826	3.3328	0.0679	b12	-9.8632	-3.53	0.0004
	b13	-3.7797	-3.25	0.0012	-1.8236	11.1376	0.0008	b13	-27.1941	-3.81	0.0001
	b14	11.5075	3.45	0.0006	5.6197	12.3778	0.0004	b14	5.9539	0.96	0.3390
	b15	-3.4858	-3.35	0.0008	-1.7046	11.6740	0.0006	b15	-13.4088	-4.11	<.0001
	b16	-1.2468	-2.28	0.0227	-0.5662	6.0485	0.0139	b16	-3.0755	-3.17	0.0015
	b17	0.7730	2.06	0.0391	0.3419	4.5204	0.0335	b17	1.1739	2.47	0.0136
	b18	-1.4575	-2.01	0.0440	-0.7543	4.9430	0.0262	b18	-7.7205	-3.54	0.0004
相 关 拟 合 指 标	Log likelihoo		-1788			-2426				-1769	
	AIC		3621			4898.835				3587	
	Schwarz Criterion		3720			5070.190				3700	
	Estrella		0.4338			0.4277				0.4698	

4.3 Mixed Logit 模型的应用

Mixed Logit 模型和 MNP 模型一样，没有标准的解析解，只能用统计模拟的方法来加以计算，模拟求解的方法可以采用 Newton-Rapson 法和梯度法等。在 SAS9.1 版本中，计算 Mixed Logit 模型中默认的随机数法是 Halton 序列法（杨勇攀，陈锬，2007）。

但是在实际应用中，可能会根据需要把不同的变量系数设为对数正态分布、均匀分布和三角分布等。本研究中，对各促销变量的参数分布做了以下几种设置：（1）各促销变量的参数都为正态分布；（2）除广告类促销外其他促销变量都为正态分布；（3）仅人员促销和礼品促销为正态分布；（4）价格系数为对数正态分布、人员促销、礼品促销为正态分布。

我们运用 Mixed Logit 模型对这四种设置模式进行估计，结果比较如表 3。

表 3 不同参数分布设置的四种模型估计结果比较

	模型 1	模型 2	模型 3	模型 4
Log Likelihood	-1772	-1771	-1787	-1769
AIC	3595	3592	3622	3587
Schwarz Criterion	3712	3704	3731	3700
Estrella 值	0.4646	0.466	0.4361	0.4698

将模型 1、模型 2、模型 3 和模型 4 相比较，参见表 3，发现无论是最大对数似然函数 Log Likelihood、还是 AIC 值、Schwarz Criterion 或者类 R^2 指标 Estrella 值，模型 4 都是最优。因此我们在使用 Mixed Logit 模型时，可以将广告类促销和品牌设定为固定值¹，将价格系数设定为对数正态分布（或是取对数值设定为正态分布）、把人员促销和礼品促销设定为正态分布。模型四的参数估计和模型拟合结果参见表 2。

4.4 模型结果分析

(1) MNL 基本结论

从表 2 中，中我们可以看出，对购买者而言，价格为负效用，人员促销、广告类促销和礼品类促销为正效用，这和我们的常识是一致的。从人员促销、广告类促销和礼品类促销的系数可以看出，礼品类促销（4.1932）带来的效用要大于人员促销（0.3970）和广告类促销（0.3853）。

a. 价格的影响

价格是消费者为获得卫生巾必须付出的经济成本，价格越高则对消费者的效用越低，那么消费者选择该品牌的概率也越低，价格的效用系数应该是负数。

¹我们也将将品牌影响作为正态分布进行参数估计，结果发现大多数品牌参数的估计方差的 t 检验值没有落入原假设（假定方差为 0）的拒绝域，因此接受原假设（参数的方差为零），即可以将参数设为固定值。

在表 2 中，价格的效用系数为-15.3588。从全模型来看，价格对市场有显著的负影响，这个经济学中的价格理论相一致。实际上，当我们分别去掉人员促销、礼品促销等因素时，价格效用系数依然为负。

b. 人员促销的影响

人员促销是通过生产企业雇佣的促销人员在现场的促销，可以吸引消费者，加强和消费者的沟通，给消费者带来效用，促进消费者购买。在全模型中，人员促销的效用系数为 0.3970，即当有人员促销时能够给消费者带来 0.397 的效用。

c. 广告类促销的影响

本文中的广告促销指卫生巾品牌采用广告传单、堆头、端架和人员路演等形式吸引顾客注意力进行促销，起到广告类的作用，这里归为广告类促销。这种促销吸引了消费者的注意力，给消费者购买某种品牌的提示，将该品牌的显著特点等信息传递给消费者，给消费者带来了一定的效用。

d. 礼品促销的影响

本次调查中的礼品类促销指购买卫生巾时赠送的礼品(包括卫生巾、卫生护垫、口红等)，由于每一种具体的礼品采用的时间都比较短，数据比较稀，所以将所有这些促销方式概括为一个变量，礼品促销。在全模型中，礼品类促销的效用系数为 4.1932，大大高于人员促销(0.3970)和广告类促销(0.3853)的效果。

e. 品牌的影响

品牌是消费者进行卫生巾购买时的主要考虑因素之一，对市场份额会有影响。品牌对消费者带来的效用越大，则品牌价值越高。模型所估计的是 b_i ，实际上就是 i 品牌对消费者产生的效用，这一效用系数的绝对值是没有实际意义的，而该效用系数在不同品牌间的差距是有意义的，它反映的是两个品牌给消费者带来的效用大小。

在模型中，我们将品牌 19 “益母草”作为参照品牌，所以其参数为 0。在全模型中，品牌 2 的参数在统计上与“益母草”没有显著区别(5%的误差水平)；在模型中，除了 b8 “安而乐”(p 值为 0.4289)和 b10 “安乐”(p 值为 0.1475)没有通过 $p < 0.05$ 的检验外，其余的品牌效用系数都通过了 $p < 0.05$ 的检验。

表 2 中的品牌效用系数按从大到小的顺序排列分别为：“ABC”、“护舒宝(高)”、“苏菲”、“洁婷”、“娇爽(高)”、“娜芳”、“娇爽(低)”、“益母草”、“安而乐”、“安乐”、“乐芙爽”、“一片云”、“伊倍爽”、“舒而美”、“护舒宝(低)”、“佳期”、“佩安婷”、“康芙娅”、“永芳”，品牌效用系数也揭示了消费者对这些品牌的不同偏好。

f. 促销因素对消费者效用影响的大小

我们从效用系数上可以看出价格、人员促销、广告类促销和礼品类促销给消费者带来的效用是不同的，价格的效用最大、其次是礼品类促销、然后才是人员促销和广告类促销。

(2) MNP 和 Mixed Logit 模型

这两个模型的基本结论与 MN 基本相同。L

表 2 中的第 6 列和第 9 列为 MNP 和 Mixed Logit 模型中各个系数的估计值，可以看出 MNP 以及 Mixed Logit 模型从估计结果上看与 MNL 模型大致相同。并且可以采用同 MNL 模型的方式来对结果进行解释。

Mixed Logit 对价格对数、促销以及人员促销的估计不但估计了该参数的平均值，还估计了这几个参数的标准差 Inrealprice_S、promogirl_S、VaPromo_S。这些系数估计出来不是一个常数，而是某种分布。

4.5 模型比较

(1) 模型拟合参数比较

我们从表 2 相关拟合指标如最大对数似然函数 Log Likelihood、还是 AIC 值、Schwarz Criterion 或者类 R^2 指标 Estrella 的比较中发现，在这几个模型中，Mixed Logit 的拟合程度最好，其次是 MNL 模型。MNP 拟合效果稍微差一点，可能原因是 MNP 模型也要求对所有效用的不可观测部分服从正态分布。而实际上，消费者价格必然为负效用（消费者不愿意多付钱），消费者购买某种令他满意的商品时其价格效用其所获得的效用必然为正？。因为正态分布密度函数分布在 0 的两侧，这时候假设不可观测因素服从正态分布就与实际情况相冲突。在混合模型中我们是通过假设价格满足对数正态分布（即价格的对数符合正态分布）来规避这一问题，更加接近现实。

(2) 对市场份额的预测准确程度比较

假设每个顾客购买数量和购买频率相同这一假设前提下，我们可以将这 669 次购买中的消费者购买不同品牌的平均购买概率看成市场份额¹。通过将各个消费者对不同品牌的选择概率的叠加后，可以预测市场市场份额，见表 4。

表 4 中，第三列的值是商场宏观数据观测到的实际市场份额；A 列的值等于第 4 列减去第 3 列值的绝对值；B 列的值等于第 5 列减去第 3 列值的绝对值；C 列的值等于第 6 列减去第 3 列值的绝对值。A、B、C 三列值分别代表 MNL、MNP 和 Mixed Logit 模型对不同品牌的平均选择概率与实际市场份额的差。

表 4 不同模型预测市场份额及准确程度

¹商场的宏观数据得到 19 个品牌的市场份额分别为 5.90%、2.10%、11.40%、4.60%、5.20%、5.30%、24.60%、1.50%、4.20%、1.60%、2.40%、5.80%、0.40%、2.00%、6.10%、2.80%、2.80%、5.10% 和 6.30%，和我们的假设也比较接近。

品牌	品牌名称	实际市场份额	MNL 模型平均选择概率	MNP 模型平均选择概率	混合模型平均选择概率	A	B	C
b1	娇爽(高)	0.059	0.05081875	0.05114438	0.05005871	0.00818	0.00786	0.00894
b2	娇爽(低)	0.021	0.01943709	0.01954526	0.02004373	0.00156	0.00145	0.00096
b3	苏菲	0.114	0.14051114	0.14168562	0.14113752	0.02651	0.02769	0.02714
b4	护舒(高)	0.046	0.06277956	0.0640886	0.06256941	0.01678	0.01809	0.01657
b5	护舒(低)	0.052	0.05082021	0.05066277	0.0511894	0.00118	0.00134	0.00081
b6	洁婷	0.053	0.03437843	0.03368045	0.0337325	0.01862	0.01932	0.01927
b7	舒而美	0.246	0.18385988	0.18549578	0.18579732	0.06214	0.06050	0.06020
b8	安而乐	0.015	0.0433486	0.04345275	0.04342166	0.02835	0.02845	0.02842
b9	佳期	0.042	0.04633579	0.0456138	0.04554347	0.00434	0.00361	0.00354
b10	安乐	0.016	0.03886418	0.03887631	0.03864413	0.02286	0.02288	0.02264
b11	佩安婷	0.024	0.0433477	0.04109828	0.04095392	0.01935	0.01710	0.01695
b12	一片云	0.058	0.0493285	0.04908387	0.0495058	0.00867	0.00892	0.00849
b13	永芳	0.004	0.01494687	0.01499593	0.0161141	0.01095	0.01100	0.01211
b14	ABC	0.02	0.0164421	0.01649295	0.01884204	0.00356	0.00351	0.00116
b15	康芙娅	0.061	0.06576728	0.06428231	0.06364333	0.00477	0.00328	0.00264
b16	乐芙爽	0.028	0.0119598	0.0120017	0.01174602	0.01604	0.01600	0.01625
b17	娜芳	0.028	0.03288517	0.0331851	0.03281145	0.00489	0.00519	0.00481
b18	伊倍爽	0.051	0.04035956	0.04078281	0.04011716	0.01064	0.01022	0.01088
b19	益母草	0.063	0.05381105	0.05383133	0.05412834	0.00919	0.00917	0.00887
累计		1	1	1	1	0.27857	0.27556	0.27068
平均误差						0.01466	0.01450	0.01425

从最后一行平均误差来看，我们看到 C 列累计最小，其次是 B 和 A。因此，可以认为三种模型对市场份额的预测准确程度最好的是 Mixed Logit 模型，其次是 MNP 模型，最后是 MNL 模型。但从第 4、5、6 列数据看，三种模型预测的市场份额差别不大，都比较接近真实值。

5 结论与展望

本论文采用 MNL、MNP 和 Mixed Logit 模型三种模型，利用北京某百货公司卫生巾的销售和促销数据，对促销因素如何影响消费者品牌选择进行实证分析，同时对各个模型的结果加以比较，得到以下结论：

5.1 卫生巾市场的顾客品牌选择方面的结论

(1) 对消费者而言，价格为负效用，人员促销、广告类促销和礼品类促销为正效用，这我们的常识是一致的。

(2) 价格的效用最大、其次是礼品类促销、然后才是人员促销和广告类促销。而礼品类促销的效果大大高于人员促销和广告类促销的效果。说明在这一市场中，顾客相对较理性，更倾向于得到价格上的折让或礼品等比较实惠的好处。

(3) 顾客本身对不同品牌的偏好也相当能明显。MNL 模型显示，品牌效用系数按从大到小的顺序排列分别为：“ABC”、“护舒宝（高）”、“苏菲”、“洁婷”、“娇爽（高）”、“娜芳”、“娇爽（低）”、“益母草”、“安而乐”、“安乐”、“乐芙爽”、“一片云”、“伊倍爽”、“舒而美”、“护舒宝（低）”、“佳期”、“佩安婷”、“康芙娅”、“永芳”，品牌效用系数也揭示了消费者对这些品牌的不同偏好。其他两个模型消费者品牌偏好方面的结论也和此类似。

5.2 营销研究方面的结论

(1) MNL、MNP 和 Mixed Logit 模型在分析促销因素如何影响消费者的品牌选择都具有较好的效果。通过模型拟合效果来看，MNL 和 Mixed Logit 模型要优于 MNP，主要原因可能是 MNP 对所有参数都假设为正态分布造成的偏差。这一方面说明了加入能够较好地设置参数的分布形式的话，Mixed Logit 模型具有一定的优越性；另一方面也说明尽管 MNP 解决了 MNL 的几个约束条件，但该模型对所有参数的正态分布假设这一局限性又反过来影响该模型精确性。

(2) 预测效果进行比较，Mixed Logit 模型和 MNP 模型效果要优于 MNL 模型。这一结论和以往研究相同。

(3) 在使用 Mixed Logit 模型时，可以将广告类促销和品牌设定为固定值，把人员促销和礼品促销设定为正态分布、将价格系数设定为对数正态分布（或是取对数值设定为正态分布）时效果最好。

5.3 研究局限性与展望

但是，在本研究中，由于本人知识背景、理论水平、理论视角、时间精力等限制，目前关于本研究存在着一定的局限。这些局限性和值得继续深入之处如下：

(1) 本研究调研的样本仅限于北京某百货商店，持续观察时间只有 31 天，POS 机显示的交易行为 669 次，调查样本数量以及调查地点有一定局限性。在实际调研中，可以采用扩大样本量和调查地点的方式弥补此不足。

(2) 我们将广告传单、堆头、端架和人员路演等形式吸引顾客注意力、起到广告类作用的促销都归为广告类促销，以简化模型，而实际上这些促销方式的影响还是有一定差别，我们计算得到的只是一个平均效用。如果条件，我们可以在限定其他促销形式不变的情况下，单独整理数据，专门针对这几类广告类促销进行研究。

(3) 本文在做市场份额变化预测时，作了一个假设：即假定每个顾客每次购买的数量和购买频率是相同的，则从 669 次购买数据中顾客累计购买品牌的比例可以看作是市场份额的抽样统计。尽管我们的计算表明，采用这样假设得到的市场份额和实际的市场份额差距并不大，但实际上每个顾客每次购买数量并不相同，并且购买频率也不一样。我们可以在将来采集数据时收集到这些顾客的购买数量和购买频率，将其作为计算市场份额的依据之一，计算得到的市场份额将更加接近实际。

参考文献

- [1] 陈锬、王晓红、朱敏 (2006), “Logit 模型在市场营销定量研究中的应用” [J], 《商业研究》24: 196-199
- [2] 董加伟, 王方华 (2004), “基于离散选择模型 DCM 的产品差异市场分析” [J], 《运筹与管理》, 2: 79-84
- [3] 郭永新、王高、齐二石 (2007), “品牌、价格和促销对市场份额影响的模型研究” [J], 《管理科学学报》, 2: p60-65
- [4] 李德春、肖岳峰 (2007), “基于多项式 Logit 模型的快速消费品促销效果评估” [J], 《桂林电子科技大学学报》2: 74-77
- [5] 聂冲、贾生华 (2005), “离散选择模型的基本原理及其发展演进评价” [J], 《数量经济技术经济研究》11: 151-159
- [6] 王树盛, 黄卫, 陆振波 (2006), “Mixed Logit 模型及其在交通方式分担中的应用研究” [J], 《公路交通科技》, 2006, 5: 88-91
- [7] 王树盛, 黄卫, 葛华 (2006), “Mixed Logit 模型随机参数分布形式选择探讨” [J], 《公路交通科技》, 8: 122-125
- [8] 吴中, 毛荣昌 (2004), “MNL 模型在交通方式划分中的新发展” [J], 《交通标准化》, 7: 40-42
- [9] 袁诚 (2003), “离散选择模型在产品差别研究中的应用” [J], 《统计研究》, 1: 52-56
- [10] 张群、曹丽、李纯青 (2005), “离散选择模型的比较及其在零售业中的应用” [J], 《西安工业学院学报》3: 293-298
- [11] 邹德强、王高、赵平、王燕 (2007), “功能性价值和象征性价值对品牌忠诚的影响: 性别差异和品牌差异的调节作用” [J], 《南开管理评论》, 3: 4-13
- [12] Ben-Akiva, M. (1972), The Structure of Travel Demand Models [D], PHD Thesis
- [13] Ben-Akiva, Moshe, and Steven R. Lerman (1985), Discrete Choice Analysis: Theory and Application to Travel Demand, Cambridge, MA: The MIT Press
- [14] Bhat, C. (1998), Accommodating Variations in Responsiveness to Level-of Service Variables in Travel Mode Choice Models [J], Transportation Research A 32: p495-510
- [15] Bhat, C. (2001), Incorporating Observed and Unobserved Heterogeneity in Urban Work Mode Choice Modeling [J], Transportation Research A 34: p228-238
- [16] Buzzell, R. D. and B. T. Gale (1987), The PIMS Principles: Linking Strategy to Performance [M], New York, NY: The Free Press.
- [17] Chintagunta, Pradeep K (1992), Estimating A Multinomial Probit Model of Brand Choice Using the Method of Simulated Moments [J], Marketing Science, Fall, Vol. 11 Issue 4, p386, 22p
- [18] Erdem, Tülin (1996), A Dynamic Analysis of Market Structure Based on Panel Data [J], Marketing Science, Vol. 15 Issue 4, p359, 20p
- [19] Green, Paul E., Frank J. Carmone, and David P. Wachspress (1977), On the Analysis of Qualitative Data in Marketing Research, Journal of Marketing Research, Vol. 18 (February), p52-9
- [20] McFadden, Daniel (1974), Conditional logit analysis of qualitative choice behavior [M], Frontiers in Econometrics, New York: Academic Press, 105-142
- [21] Russell, Gary J. and Ann Petersen (2000), Analysis of Cross Category Dependence in Market Basket Selection, Journal of Retailing, 76 (Fall), p367-392

- [22] Train, K. , D. McFadden, and A. Goett (1987), “Consumer attitudes and voluntary rate schedules for public utilities”[J],Review of Economics and Statistics, LXIX(3):383-391
- [23] Train(1999), Mixed Logit Models for Recreation Demand[M], in J. Herriges and C. Kling, eds., Valuing Recreation and the Enviroment, Edward Elgar, Northampton
- [24] Train (2002) , Discrete Choice Methods with Simulation [M], Cambridge University Press, p15-16
- [25] Zsolt SÁndor, Michel Wedel (2005), Heterogeneous Conjoint Choice Designs. Journal of Marketing Research 42:2, 210-218

The Application of Discrete Choice Model on Retailing Research –Based on Comparation between MNL、 MNP and Mixed Logit

Liu Yong, Wang Gao

Abstract Discrete Choice model has been used widely in many research fields such as traffic, reas estate, tele-communication, energy and marketing in the world, however the model in marketing research is still in its infancy in China. This paper applies three discrete selection models MNL, MNP, MIXED LOGIT. It focuses on the impact of marketing devices on customers' brand selction and market share in Sanitary Pad market. The author found that the brand and various other promotional factors are playing positive impact when customers purchasing decisions are made. Particularly, some promotional means, such as Gift-promotion, which can bring customers some visible advantages, effect the greatest.

Key words Discrete Mode, MNL, MNP, Mixed Logit

专业主编: 王高

