УДК 339.138(03)

ЭКОНОМЕТРИЧЕСКИЕ МОДЕЛИ КАЧЕСТВЕННЫХ ПЕРЕМЕННЫХ В ПРОГНОЗНЫХ ЗАДАЧАХ МАРКЕТИНГА

О. С. Воищева

Воронежский государственный университет

Рассмотрены основные аспекты эконометрического подхода к моделированию качественных переменных. Изложение аппарата проведено таким образом, что позволяет понять его преимущества перед методами дискриминантного анализа, составляющими альтернативу эконометрическим моделям дискретного выбора в практическом маркетинге. В частности, обсуждаемый аппарат обеспечивает проведение предельного анализа, позволяющего выяснить механизмы факторного воздействия на возможность наступления того или иного события.

ВВЕДЕНИЕ

Специфика прогнозных задач, возникающих при обосновании маркетинговых решений, заключается в том, что в предполагаемом образе будущего присутствуют не только процессы, но и события. И если аппарат прогнозирования экономических процессов хорошо известен и успешно применяется в маркетинговых исследованиях, то вопросы, связанные с предсказанием событий, еще не получили должного развития. Рекомендуемые для решения подобного рода задач экспертные методы, к сожалению, не всегда обеспечивают получение надежных результатов и, кроме того, упрощая формализацию задачи, они в то же время значительно усложняют процедуру получения прогнозных оценок.

В последние два десятилетия в эконометрике [2] интенсивно разрабатывается новый класс регрессионных моделей, позволяющий моделировать переменные с дискретным характером изменения. В моделях подобного типа заложена возможность получения вероятностных оценок, характеризующих реальность появления того или иного события. Однако этот аппарат практически не используется в маркетинговых исследованиях и происходит это, по нашему убеждению, из-за отсутствия достаточного количества разработок, иллюстрирующих его аналитические и прикладные возможности.

Наиболее подробное изложение на русском языке теоретических основ построения эконометрических моделей с дискретной зависимой переменной можно найти в [1]. Там же описаны отдель-ные маркетинговые ситуации, в которых целесообразно применять эти модели. В рамках исследования этих ситуаций продемонстрированы свойства моделей и их использование в аналитических целях. Кроме того изложен материал по применению отдельных моделей этого типа для прогнозирования в номинальных и ранговых шкалах.

Ниже приводится описание спецификации и методов построения эконометрических моделей с дискретной зависимой переменной, а также обсуждаются возможности их применения в маркетинговых исследованиях.

СПЕЦИФИКАЦИЯ И ОЦЕНКА ПАРАМЕТРОВ

Модель бинарного выбора. Обсуждение вопросов, связанных с природой закономерностей, объясняющих взаимосвязь между непрерывным и дискретным, начнем с модели бинарного выбора. Для этого введем в рассмотрение линейные функции полезности, характеризующие результат выбора альтернативы, в зависимости от условий, в которых этот выбор производился.

Пусть

u_1 – полезность, получаемая при выборе 1-й альтернативы;

u_2 – полезность, получаемая при выборе 2-й альтернативы.

Каждая из полезностей зависит от факторов, описывающих условия выбора, линейно, то есть

$$\mathrm{u}_1 = \mathbf{xd}_1,\ \mathrm{u}_2 = \mathbf{xd}_2. \tag{1}$$

Здесь $\mathbf{x} = (1, x_1, x_2, ..., x_m)$ расширенная вектор-строка, как это принято в эконометрике, а $\boldsymbol{d}_k = (d_{k0}, d_{k1}, ..., d_{km})'$ – вектор-столбец параметров k-й функции полезности.

К сожалению, в определении полезности преобладает субъективный фактор и она, как правило, является ненаблюдаемой величиной, что исключает возможность применения эконометрических методов для оценки коэффициентов этих линейных функций. В то же время, несмотря на отсутствие реальных функций, имеющих формализованное представление, в каждом конкретном случае в соответствии с этими гипотетическими полезностями выбор наиболее предпочтительной альтернативы все же был сделан. Предполагая, что выбор осуществлялся рационально с ориентацией на получение наибольшей выгоды, можно записать следующее правило

$$y = \begin{cases} 1, & \text{при} \quad u_1 - u_2 > 0 \\ 0, & \text{при} \quad u_1 - u_2 < 0 \end{cases} \cdot \tag{2}$$

Если теперь разность полезностей представить как

$$u1 - u2 = \mathbf{xb} + \varepsilon\,, \tag{3}$$

где ε – случайная ошибка, которая может быть допущена при оценке полезности выбора, то вероятность того, что предпочтение получит первая альтернатива, записывается следующим образом:

$$\mathrm{P}\,(y=1) = \mathrm{P}\,(\varepsilon > -\mathbf{xb}) = \mathrm{P}\,(\varepsilon < -\mathbf{xb}) = F(\mathbf{xb}). \tag{4}$$

В тех случаях, когда в качестве $F(\mathbf{xb})$ выбирается нормальное распределение

$$\mathrm{P}(y=1) = \int_{-\infty}^{\mathbf{xb}} \phi(t)dt = \Phi(\mathbf{xb})\,, \tag{5}$$

где $\phi(t) = \frac{1}{\sqrt{2\pi}} e^{\frac{t^2}{2}}$ является плотностью нормального распределения, то мы имеем дело с пробит-моделью. Если выбирается логистическое распределение

$$\mathrm{P}(y=1) = e^{\mathbf{xb}}(1+e^{\mathbf{xb}})^{-1} = \Lambda(\mathbf{xb})\,, \tag{6}$$

то выражение (6) называют логит-моделью.

Для построения и пробит- и логит-модели используются статистические наблюдения ситуаций бинарного выбора, то есть наблюдения, в которых значения зависимой переменной принимают всего два значения 0 и 1, а независимые являются непрерывными или категоризированными переменными. Фактически это ситуация, порождающая задачу деления выборочного множества на два класса наблюдений. В маркетинговых исследованиях для решения подобного рода задач обычно используется дискриминантный анализ. Сравнение между собой дискриминантной функции и модели бинарного выбора оказывается в пользу последней. Эконометрическая модель является более гибким инструментом, так как обеспечивает не только разделение выборочной совокупности на два класса, но и проведение предельного анализа факторов. Техника предельного анализа будет изложена ниже.

Модель бинарного выбора нелинейная, и поэтому ее параметры оцениваются с помощью метода максимального правдоподобия. Для выборочного множества наблюдений функция максимального правдоподобия записывается в виде

$$\mathrm{L}(\mathbf{y},\mathbf{b}) = \prod_{i=1}^{n} \mathrm{F}(\mathbf{x}_i\mathbf{b})^{y_i}\left[1-\mathrm{F}(\mathbf{x}_i\mathbf{b})\right]^{1-y_i}. \tag{7}$$

В данной форме записи множители произведения селектируются с помощью компонент вектора **y**, принимающих всего два значения 0 или 1. В произведении сохраняются те сомножители, у которых показатель степени оказался равным 1.

Математически проще максимизировать логарифмическую функцию правдоподобия

$$\ln \mathrm{L} = \sum_{i=1}^{n} [y_i \ln\mathrm{F}(\mathbf{x}_i\mathbf{b}) + (1-y_i)\ln(1-\mathrm{F}(\mathbf{x}_i\mathbf{b}))]. \tag{8}$$

Используя сокращенные записи $\mathrm{F}_i = \mathrm{F}\,(\mathbf{x}_\mathbf{i}\,\mathbf{b})$ и $\mathrm{F}_b'(\mathbf{x}_\mathbf{i}\,\mathbf{b}) = \mathrm{f}_\mathrm{i}$, выпишем для логарифмической функции правдоподобия условия максимизации первого порядка

$$\frac{\partial \ln \mathrm{L}}{\partial \mathbf{b}} = \sum_{i=1}^{n}\left[\frac{y_i\,\mathrm{f}_i}{\mathrm{F}_\mathrm{i}} + (1-y_i)\frac{-\mathrm{f}_i}{(1-\mathrm{F}_\mathrm{i})}\right]\mathbf{x}_i' = 0. \tag{9}$$

В случае логистической функции распределения эта система уравнений может быть упрощена

$$\sum_{i=1}^{n}(y_i - \Lambda_i)\,\mathbf{x}_i = 0. \tag{10}$$

В силу нелинейности для получения решения этих систем используются численные методы. Чаще других для этих целей используется метод Ньютона-Рафсона. Можно также использовать метод Берксона, представляющий собой итерационную схему обобщенного МНК. Подобного рода итерационные процедуры используются во многих статистических пакетах. В частности, возможность построения моделей бинарного выбора с использованием итерационной процедуры реализована, например, в пакетах Eviews и Statistica.

Мультиномиальная логит-модель множественного выбора. Мультиномиальная логит-модель является обобщением логит-модели бинарного выбора. Предварительно при ее построении все варианты возможного выбора нумеруются в произвольном порядке 0, 1, 2, ..., J. (Нумерация начата с 0 только для того, чтобы было единообразие с предыдущей кодировкой 0 и 1.) Вероятность наступления того или иного варианта описывается полиномиальной логит-моделью

$$\mathrm{P}(y_i = j) = \frac{e^{\mathbf{x}_i\mathbf{b}_j}}{\sum_{j=0}^{J} e^{\mathbf{x}_i\mathbf{b}_j}}, \quad j = 0,\ 1,\ 2, \ldots, J. \quad (11)$$

Вектор независимых переменных $\mathbf{x}_i = [\,\mathbf{z}_i, \mathbf{w}_i]$ составлен из двух подвекторов, каждый из которых имеет собственную смысловую нагрузку. Компоненты вектора $\mathbf{z}_i$ принято называть атрибутами и понимать их как показатели, по которым различаются альтернативы. В свою очередь компоненты вектора $\mathbf{w}_i$ называют характеристиками, понимая под ними описание индивидуальных черт тех лиц, которые осуществляли выбор альтернатив.

Оценка параметров модели (11) не дает однозначного результата, так вместе с вычисленными коэффициентами $\hat{\mathbf{b}}$ идентичные вероятности позволяет получить вектор $\hat{\mathbf{b}} + \mathbf{d}$. Избежать этой неоднозначности позволяет операция нормализации (стандартизации), смысл которой в том, чтобы для одного из вариантов, например $y_i = J$, положить $b_j = 0$. Тогда оценивается не $J + 1$ функция, а J функций одного вида

$$\mathrm{P}(y_i = j) = \frac{e^{\mathbf{x}_i\mathbf{b}_j}}{1 + \sum_{j=0}^{J-1} e^{\mathbf{x}_i\mathbf{b}_j}}, \qquad j = 1,\ 2, \ldots, J, \quad (12)$$

после чего определяется еще одна функция через значения этих J функций путем вычитания их суммы из единицы

$$\mathrm{P}(y_i = J) = \frac{1}{1 + \sum_{j=0}^{J-1} e^{\mathbf{x}_i\mathbf{b}_j}}. \quad (13)$$

Это одна из особенностей построения полиномиальной логит-модели. В соответствии с этой особенностью компьютерные пакеты рассчитывают только коэффициенты первых J зависимостей $\mathbf{b}_0\ \mathbf{b}_1, \ldots, \mathbf{b}_{J-1}$, по которым вычисляются в соответствии с (12) первые J вероятностей $\mathrm{P}(y_i = 0)$, $\mathrm{P}(y_i = 1)$, $\mathrm{P}(y_i = J - 1)$. Вероятность выбора последнего варианта $\mathrm{P}(y_i = J)$ компьютером не рассчитывается, а определяется отдельно с помощью (13).

Оценивание коэффициентов модели осуществляется путем численного решения уравнений правдоподобия. Для записи самого уравнения правдоподобия, а точнее его логарифмической формы, удобно ввести переменную d_{ij}, которая принимает значение 1, если в i-м наблюдении (i-м индивидуумом) был выбран j-й альтернативный вариант среди ($J + 1$)-го, и 0 – в противном случае. Тогда для каждого i только одно из будет d_{ij} равно 1.

Используя введенную переменную d_{ij}, запишем функцию логарифмического правдоподобия

$$\ln \mathrm{L} = \sum_{i=1}^{n} \sum_{j=0}^{J} d_{ij} \ln\left(\frac{e^{\mathbf{x}_i\mathbf{b}_j}}{\sum_{k=0}^{J} e^{\mathbf{x}_i\mathbf{b}_k}} \right). \quad (14)$$

После дифференцирования этого выражения по $\boldsymbol{b}_j$ и несложных преобразований, получаем систему уравнений максимального правдоподобия

$$\frac{\partial \ln \mathrm{L}}{\partial \mathbf{b}_j} = \sum_i [d_{ij} - \mathrm{P}_{ij}]\mathbf{x}_i' = \mathbf{0},\ j = 0, 1, 2, \ldots J - 1. \quad (15)$$

Заметим, что в j-м блоке уравнений суммирование идет по всем i, причем если в i-м случае был выбран j-й вариант, т.е. $y_1 = j$, то в квадратных скобках имеем $[1 - P_{ij}]$, в противном случае $[-P_{ij}]$.

Решение этой системы с учетом того, что $\mathbf{b}_J = \mathbf{0}$ осуществляется численно с помощью метода Ньютона – Рафсона. Компьютерная реализация устроена таким образом, и об этом уже говорилось, что нулевые значения получают параметры той модели, которая соответствует последней из указанных альтернатив. Другими словами, если бы мы захотели, чтобы $\boldsymbol{b}_0 = \mathbf{0}$, а не $\boldsymbol{b}_J = \mathbf{0}$, то данные, соответствующие альтернативе с номером, должны быть введены последними.

Мультиномиальная логит-модель, как и модель бинарного выбора, успешно используется в задачах дискриминантного анализа, но только в тех случаях, когда возникает необходимость разделения множества объектов более чем на два класса. С ее помощью можно оценить степень реальности ожидаемых событий в зависимости от условий, способствующих их появлению. Причем нет строгих ограничений на количество оцениваемых альтернатив, однако следует помнить, что каждая новая альтернатива требует дополнительного введения в модель $m + 1$ параметра.

Опыта применения мультиномиальной модели в маркетинговых исследованиях практически нет. Можно только надеяться, что знакомство с этим аппаратом эконометрического моделирования будет способствовать появлению специфических задач, решение которых ориентировано именно на этот тип моделей. Хорошая перспектива у этого аппарата просматривается при реализации комбинированных подходов. В [ДТ] предложена прогнозная модель, построенная на принципе комбинирования многовариантных прогнозных расчетов с оценкой реальности каждого варианта, рассчитываемой с помощью эконометрической модели множественного выбора.

Модели множественного выбора с упорядоченными альтернативами. В моделях, рассмотренных выше, предполагалось, что моделируемая величина (показатель) измеряется в номинальных шкалах. Ниже, продолжая развивать идею моделирования качественных переменных, мы рассмотрим схему построения пробит и логит-моделей множественного выбора для случая, когда моделируемая переменная измеряется в ранговой шкале.

При решении маркетинговых задач, ситуации, когда результаты моделирования в соответствии с содержательным смыслом должны быть представлены в ранговой шкале, встречаются довольно часто. Приведем несколько примеров.

1. Определение рейтинга банков по их надежности.
2. Моделирование результатов дегустационных тестов.
3. Определение приоритетов по результатам опроса потребителей.
4. Выяснение приоритетности инвестиционных проектов.

При моделировании подобных ситуаций будем предполагать, что есть переменная y^*, значения которой определяются некоторым набором объясняющих переменных в соответствии с зависимостью

$$y^* = \mathbf{xb} + \varepsilon\,. \qquad (16)$$

Сама переменная y^* – ненаблюдаемая величина, но известны значения дискретной переменной, которые в нашем представлении связаны с ненаблюдаемой следующими соотношениями:

$$\begin{array}{l} y = 0,\ \text{если } y^* \le 0 \\ y = 1,\ \text{если } 0 < y^*\ \mu_1 \\ y = 2,\ \text{если } \mu_1 < y^* \le \mu_2 \\ .\ \ .\ \ .\ \ .\ \ .\ \ .\ \ .\ \ .\ \ . \\ y = J,\ \text{если } \mu_{j-1} \le y^* \end{array} \qquad (17)$$

Неравенства реализуют некую форму цензурирования. Причем уровни цензурирования μ_j неизвестны и представляют собой параметры, оцениваемые вместе с коэффициентами **b**.

Заменим в неравенствах ненаблюдаемую переменную ее модельным представлением и вычтем из каждой части

$$\begin{array}{l} y = 0,\ \text{если } \varepsilon \le -\mathbf{xb} \\ y = 1,\ \text{если } -\mathbf{xb} \le \mu_1 - \mathbf{xb} \\ y = 2,\ \text{если } \mu_1 - \mathbf{xb}_1 < \varepsilon \le \mu_2 - \mathbf{xb} \\ .\ \ .\ \ .\ \ .\ \ .\ \ .\ \ .\ \ .\ \ . \\ y = J,\ \text{если } \mu_{j-1} - \mathbf{xb} \le \varepsilon. \end{array} \qquad (18)$$

Будем считать, что случайная величина ε нормально распределена по наблюдениям и, кроме того, нормирована таким образом, что имеет нулевое математическое ожидание и единичную дисперсию. В случае построения логит-модели предполагается, что случайная велична ε имеет логистическое распределение.

Для пробит-модели выписанные неравенства позволяют записать следующие вероятности:

$$\begin{array}{l} \mathrm{P}(y=0) = \Phi(-\mathbf{xb}) \\ \mathrm{P}(y=1) = \Phi(\mu_1 - \mathbf{xb}) - \Phi(-\mathbf{xb}) \\ \mathrm{P}(y=2) = \Phi(\mu_2 - \mathbf{xb}) - \Phi(\mu_1 - \mathbf{xb}) \\ .\ \ .\ \ .\ \ .\ \ .\ \ .\ \ .\ \ .\ \ .\ \ .\ \ . \\ \mathrm{P}(y=J) = 1 - \Phi(\mu_{J-1} - \mathbf{xb}) \end{array} \qquad (19)$$

Чтобы все вероятности были положительными, оцениваемые параметры положения должны удовлетворять неравенствам

$$0 < \mu_1 < \mu_2 < \ldots < \mu_{J-1}\,. \qquad (20)$$

Смысл параметров положения хорошо иллюстрирует рисунок 1.

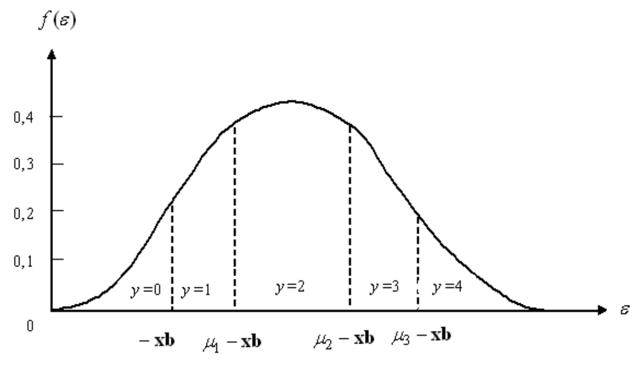

Рис. 1. Вероятности в упорядоченной пробит-модели

Если учесть, что

$$\Phi(\mu_k - \mathbf{xb}) - \Phi(\mu_{k-1} - \mathbf{xb}) = \frac{1}{\sqrt{2\pi}} \int_{z_1}^{z_2} e^{-\frac{t^2}{2}}\, dt\,, \qquad (21)$$

где $z_1 = \mu_{k-1} - \mathbf{xb}$, $z_2 = \mu_k - \mathbf{xb}$, то становится понятным, как построить логарифмическую функцию правдоподобия. Аналогично предыдущим моделям метод Ньютона–Рафсона позволяет оптимизировать функцию правдоподобия и получить искомые параметры положения и коэффициенты модели.

Модель Пуассона. Модель, основанную на распределении Пуассона целесообразно использовать в тех ситуациях, когда возникает необходимость в предсказании показателя, характеризующего число случаев, возникающих в течение заданного промежутка времени. Корректность построения данной модели, естественно, требует, чтобы выполнялись все предположения, лежащие в основе пуассоновского процесса. Модель можно применять для предсказания числа полетов, которое может быть достигнуто аэропортом в конкретный день в зависимости от даты, числа несчастных случаев в зависимости от условий труда, которые обеспечиваются фирмой, количество банкротств в зависимости от повышения ставки за банковский кредит и т.д.

Спецификация регрессионной модели Пуассона предполагает, что каждое наблюдаемое значение моделируемого показателя сгенерировано распределением Пуассона с параметром λ_i, связанным с вектором объясняющих переменных $\boldsymbol{x}_i$. Основное уравнение модели имеет вид

$$\mathrm{P}(Y_i = y_i) = \frac{e^{-\lambda_i} \lambda_i^{y_i}}{y_i!} \qquad (22)$$

Естественной и наиболее общей зависимостью, связывающей с объясняющими переменными, является логарифмическая линейная модель

$$ln\,\lambda_i = \mathbf{x}_i \mathbf{b}. \qquad (23)$$

В соответствии со свойствами распределения Пуассона можно в рамках данной модели определить ожидаемое число событий в зависимости от значений объясняющих переменных

$$\mathrm{E}[y_i | \mathbf{x}_i] = \lambda_i = e^{\mathbf{x}_i \mathbf{b}}. \qquad (24)$$

В принципе, модель Пуассона представляет собой нелинейную регрессию, параметры которой удобно оценить с помощью метода максимального правдоподобия. Логарифмическая функция правдоподобия в этом случае имеет вид

$$\ln L = \sum_{i=1}^{n} [-\lambda_i + y_i \mathbf{x}_i \mathbf{b} - \ln y_i!]. \qquad (25)$$

Ее дифференцирование позволяет получить уравнения правдоподобия

$$\frac{\partial \ln L}{\partial \mathbf{b}} = \sum_{i=1}^{n} (y_i - \lambda_i) \mathbf{x}_i = 0, \qquad (26)$$

численное решение которых дает оценки параметров модели $\hat{\mathbf{b}}$.

В отличие от предыдущих, данная модель обеспечивает моделирование переменных, множество допустимых значений которых не конечно, а счетно.

ПРЕДЕЛЬНЫЙ АНАЛИЗ ФАКТОРОВ

Все рассмотренные модели нелинейные и поэтому непосредственная содержательная интерпретация их параметров теряет всякий смысл. Выход из этой ситуации можно найти, если использовать аппарат предельного анализа. Этот аппарат, реализуя единый подход, основанный на вычислении производных, в то же время позволяет уловить специфику каждой отдельной модели. Отражение специфики помогает понять механизм факторного воздействия на моделируемый показатель в нелинейной модели и использовать эти знания при обосновании маркетинговых решений.

Предельный анализ модели бинарного выбора. Рассмотрим общий случай модели бинарного выбора, и запишем для события условное математическое ожидание

$$\mathrm{E}(y_i | \mathbf{x}_i) = 1 \cdot \mathrm{F}(\mathbf{x}_i \mathbf{b}) + 0 \cdot (1 - \mathrm{F}(\mathbf{x}_i \mathbf{b})). \qquad (27)$$

Предельный эффект *i*-го фактора вычисляется в виде первой производной

$$\frac{\partial \mathrm{E}(y_i | \mathbf{x}_i)}{\partial x_{ik}} = \left\{ \frac{\partial \mathrm{F}(\mathbf{x}_i \mathbf{b})}{\partial (\mathbf{x}_i \mathbf{b})} \frac{\partial (\mathbf{x}_i \mathbf{b})}{\partial x_{ik}} \right\} = \mathrm{f}(\mathbf{x}_i \mathbf{b}) b_k, \qquad (28)$$

где *f(•)* – функция плотности, связанная с соответствующим кумулятивным распределением *F(•)*.

Полученный предельный эффект можно интерпретировать как величину, на которую изменяется вероятность выбора при изменении фактора на единицу, то есть, по сути, как изменится неопределенность ситуации бинарного выбора. Однако механизм формирования этой величины не так прост, как в линейной модели, и представляет собой взаимодействие двух составляющих, каждая из которых имеет собственную интерпретацию.

Первая составляющая определяется плотностью распределения, которая в предельном эффекте является изменяемой характеристикой, зависящей от $\mathbf{x}_i$. Рассмотрим основные механизмы формирования этой составляющей и ее содержательную интерпретацию. Прежде всего, обратим внимание на то обстоятельство, что изменение *k*-й переменной, например, в сторону увеличения ее значения может привести как к увеличению, так и снижению плотности вероятности. Механизм реализации этих изменений начинает действовать с изменения значения линейной формы

$$z_i = \mathbf{x}_i\mathbf{b} = b_0 + b_1 x_{i1} + \ldots + b_k x_{ik} + \ldots + b_m x_{im}\,. \quad (29)$$

Изменение линейной формы зависит от величины и знака коэффициента регрессии b_k. В свою очередь, изменение плотности вероятности зависит не только от величины, на которую изменилось значение линейной формы, но и от того, где это значение расположено на оси z. Если оно расположено в левой половине распределения, то увеличение z приводит к возрастанию плотности, если в правой – то к снижению. Это положение хорошо иллюстрирует рисунок 2.

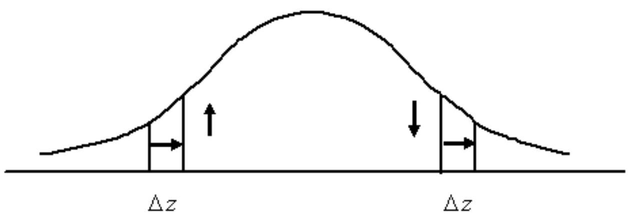

Рис. 2. Изменение плотности вероятности в зависимости от расположения значения линейной формы

Анализ предельной производительности факторов позволяет обнаружить, что максимально возможная производительность фактора достигается в тех точках, в которых плотность имеет наибольшее значение. Интересно, что именно в этих точках ситуация бинарного выбора обладает самым высоким уровнем неопределенности. Это становится совершенно очевидным для логит-модели, если вспомнить о выражении для плотности и предельную производительность ее k-го фактора записать в виде

$$\frac{\partial \mathrm{E}(y_i \mid \mathbf{x}_i)}{\partial x_{ik}} = \mathrm{F}(\mathbf{x}_i\mathbf{b})(1 - \mathrm{F}(\mathbf{x}_i\mathbf{b}))b_k\,. \quad (30)$$

Максимальное значение первой составляющей, которая в данном выражении представлена произведением вероятностей, достигается при $\mathrm{F}(\mathbf{x}_i\mathbf{b}) = 0{,}5$, т.е. когда имеет место самый высокий уровень неопределенности.

Вторая составляющая менее интересна для анализа. Она равна постоянной величине и в основном играет роль мультипликатора, усиливающего или снижающего вклад первой в предельную производительность. Геометрически при увеличении x_{ik} на единицу коэффициент b_k определяет ширину прямоугольника с высотой $\mathrm{F}(\mathbf{x}_i\mathbf{b})$, на величину площади которого изменяется вероятность бинарного выбора в условиях, описываемых вектором $\mathbf{x}_i$.

Так как события бинарного выбора несовместны, то при рассмотрении результатов предельного анализа нужно помнить, что увеличение вероятности возможного появления одного из событий влечет за собой уменьшение на ту же самую величину вероятности возможного появления альтернативного события. Поэтому, если из двух вероятностей увеличивается при изменении x_{ik} та, которая имеет большее значение, то неопределенность выбора снижается, если та, которая имеет меньшее значение, то неопределенность выбора увеличивается.

Предельный анализ мультиномиальной модели множественного выбора. Дифференцируя по l-му атрибуту в i-й точке j-ю вероятность получаем предельный эффект в виде

$$\delta_{ij} = \frac{\partial \mathrm{P_{ij}}}{\partial x_l} = \frac{\partial}{\partial x_l}\left[\frac{e^{\mathbf{x}_i\mathbf{b}_j}}{\sum_{k=0}^{J} e^{\mathbf{x}_i\mathbf{b}_k}}\right] =$$

$$= \frac{e^{\mathbf{x}_i\mathbf{b}_j}\left(\sum_{k=0}^{J} e^{\mathbf{x}_i\mathbf{b}_k}\right)b_{lj} - e^{\mathbf{x}_i\mathbf{b}_j}\left(e^{\mathbf{x}_i\mathbf{b}_0}b_{l0} + e^{\mathbf{x}_i\mathbf{b}_1}b_{l1} + \cdots + e^{\mathbf{x}_i\mathbf{b}_J}b_{lJ}\right)}{\left(\sum_{k=0}^{J} e^{\mathbf{x}_i\mathbf{b}_k}\right)^2} =$$

$$= \frac{e^{\mathbf{x}_i\mathbf{b}_j}}{\sum_{k=0}^{J} e^{\mathbf{x}_i\mathbf{b}_k}}\left[\frac{\sum_{k=0}^{J} e^{\mathbf{x}_i\mathbf{b}_k}}{\sum_{k=0}^{J} e^{\mathbf{x}_i\mathbf{b}_k}}b_{lj} - \left(\frac{e^{\mathbf{x}_i\mathbf{b}_0}}{\sum_{k=0}^{J} e^{\mathbf{x}_i\mathbf{b}_k}}b_{l0} + \frac{e^{\mathbf{x}_i\mathbf{b}_1}}{\sum_{k=0}^{J} e^{\mathbf{x}_i\mathbf{b}_k}}b_{l1} + \cdots + \frac{e^{\mathbf{x}_i\mathbf{b}_J}}{\sum_{k=0}^{J} e^{\mathbf{x}_i\mathbf{b}_k}}b_{lJ}\right)\right] =$$

$$= \mathrm{P}_{ij}[b_{lj} - \bar{b}_{il}]. \quad (31)$$

При расчетах по этой формуле нужно помнить, что вектор $\mathbf{b}_0$ в соответствии с принятым соглашением нулевой и поэтому первое слагаемое при определении математического ожидания коэффициента равно нулю.

Таким образом, предельный эффект, получаемый при изменении l-го атрибута (l-й независимой переменной) представляет собой произведение вероятности $\mathrm{P}(y_i = j)$ на разность коэффициента стоящего перед l-ой переменной и средней величиной этого коэффициента. Предельный эффект зависит от атрибута, причем механизм этой зависимости реализуется через вероятность и через среднюю величину коэффициента, при определении которой задействована та же самая вероятность. При высокой вероятности, также как и при малой, предельный эффект незначительный. Это объясняется тем, что при больших P_{ij}, в средней величине коэффициента $\overline{b}_l$ доминирует величина b_{lj} и разность между ними близка к нулю. При малых значениях P_{ij} разность большая, но значение самой вероятности близко к нулю, а следовательно, и величина предельного эффекта небольшая. Обобщая, можно утверждать, что $\delta_{ij} \to 0$ в двух случаях: когда $P_{ij} \to 0$ и когда $P_{ij} \to 1$. Своего максимального значения он достигает, когда вероятность близка к 0,5, то есть имеет место ситуация с самым большим уровнем неопределенности при выборе j-го варианта. Это естественно, так как именно в этой ситуации наиболее ценной является любая информация, уточняющая наше представление о выборе альтернатив.

Предельный анализ модели множественного выбора с упорядоченными альтернативами.

Для пояснений рассмотрим упрощенный пример, в котором моделируется с помощью нормального распределения ситуация из трех категорий с одним неизвестным параметром положения μ

$$\begin{aligned} &\mathrm{P}(y=0) = 1 - \Phi(\mathbf{xb}) \\ &\mathrm{P}(y=1) = \Phi(\mu - \mathbf{xb}) - \Phi(-\mathbf{xb}) \\ &\mathrm{P}(y=2) = 1 - \Phi(\mu - \mathbf{xb}). \end{aligned} \tag{32}$$

Модель нелинейная и поэтому предельные эффекты факторов не равны коэффициентам. Дифференцирование уравнений по любому из факторов приводит к соотношениям

$$\begin{aligned} &\frac{\partial\,\mathrm{P}(y=0)}{\partial x_k} = -\phi(\mathbf{xb})b_k \\ &\frac{\partial\,\mathrm{P}(y=1)}{\partial x_k} = [\phi(-\mathbf{xb}) - \phi(\mu - \mathbf{xb})]b_k \\ &\frac{\partial\,\mathrm{P}(y=2)}{\partial x_k} = \phi(\mu - \mathbf{xb})b_k. \end{aligned} \tag{33}$$

Таким образом, предельный эффект это величина, перераспределяемая между вероятностями полученного распределения. Причем, сумма всех изменений равна нулю. Действительно в рассматриваемом примере при $b_k > 0$ вероятность события $y = 0$ уменьшается на $\phi(\mathbf{xb})b_k$. Одновременно с этим вероятность $\mathrm{P}(y = 1)$ увеличивается на эту же величину и уменьшается на $\phi(\mu - \mathbf{xb})b_k$, а вероятность $\mathrm{P}(y = 1)$ – увеличивается на $\phi(\mu - \mathbf{xb})b_k$. Нетрудно заметить, что при положительном b_k смещение вероятности происходит вправо, а при отрицательном b_k – влево. Из этого следует, что увеличение факторной переменной, когда коэффициент при ней положителен, приводит к увеличению вероятностей тех событий, которые получили более низкие ранги. Происходит как бы деформирование кривой плотности вероятностей с перемещением некоторой вероятностной массы (площади под кривой) слева направо. В случае отрицательного коэффициента все происходит наоборот. Весь этот механизм хорошо иллюстрируется на рисунокм 3.

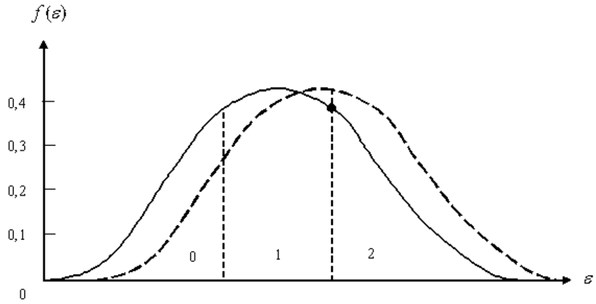

Рис. 3. Изменения распределения вероятностей под воздействием предельных эффектов

Предельный анализ модели Пуассона. Предельный эффект k-го фактора определяется дифференцированием (24)

$$\frac{\partial\,\mathrm{E}(y_i|/\mathbf{x}_i)}{\partial x_k} = \lambda_i b_k. \tag{34}$$

Полученное выражение показывает, что величина предельного эффекта равна произведению условного среднего, вычисленного в i-й точке (наблюдении) и k-го коэффициента модели. Он показывает, насколько в среднем увеличится ожидаемая величина моделируемого показателя, если соответствующий фактор изменился на единицу при условии, что эта единица достаточно мала. Это позволяет оценивать степень факторной угрозы, ориентированной на формирование реальности появления нежелательных событий.

ЗАКЛЮЧЕНИЕ

Изложение аппарата эконометрического моделирования качественных переменных проведено в ключе, позволяющем понять его альтернативность методам дискриминантного анализа, которые в настоящее время получили «постоянную прописку» в маркетинговых исследованиях. Показано, что в дополнение к прикладным возможностям, которыми наделен дискриминантный анализ, этот аппарат обеспечивает проведение предельного анализа, позволяющего выяснить механизмы факторного воздействия на ситуацию, способствующую или препятствующую появлению того или иного события. Естественно, это обеспечивает более высокий уровень обоснованности маркетинговых решений.

ЛИТЕРАТУРА

1. *Давнис В.В.* Прогнозные модели экспертных предпочтений / В.В. Давнис, В.И. Тинякова. – Воронеж : Изд-во Воронеж. гос. ун-та, 2005. – 248 с.

2. *Green W.H.* Econometric Analysis, 4th ed. / W.H. Green – New York : Macmillian Publishing Company, 2000. – 1004 p.