

МАРГАРИТА АКУЛИЧ

**Статистические методы,
используемые в маркетинговых
исследованиях**



Маргарита Акулич
Статистические методы,
используемые в маркетинговых
исследованиях

*http://www.litres.ru/pages/biblio_book/?art=26713246
ISBN 9785448579400*

Аннотация

В книге дано лаконичное описание большинства используемых в маркетинге методов и приведено пояснение, для чего они используются. Статистические методы являются для маркетинга одними из основных (если не самыми главными). Поэтому каждый маркетолог должен их знать хотя бы в общих чертах.

Содержание

Предисловие	5
I Применение в маркетинге Описательного анализа данных, графического, табличного и выборочного методов	6
1.1 Применение в маркетинге описательного анализа данных. Метод графический	7
1.2 Метод табличный. Метод выборочный	12
II Применение в маркетинге Корреляционного и регрессионного анализа	16
2.1 Анализ корреляционный. Регрессионный анализ линейный и нелинейный	16
2.2 Регрессионный анализ нелинейный. Регрессия категориальная	22
2.3 Регрессия логистическая. Регрессия мультиномиальная логистическая	27
2.4 Пробит-модель регресси. Регрессия Кокса. Анализ временных рядов	31
Конец ознакомительного фрагмента.	35

**Статистические
методы, используемые
в маркетинговых
исследованиях**

Маргарита Акулич

© Маргарита Акулич, 2020

ISBN 978-5-4485-7940-0

Создано в интеллектуальной издательской системе Ridero

Предисловие

В книге дано лаконичное описание большинства используемых в маркетинге методов и приведено пояснение, для чего они используются.

Статистические методы являются для маркетинга одними из основных (если не самыми главными). Поэтому каждый маркетолог должен их знать хотя бы в общих чертах.

I Применение в маркетинге Описательного анализа данных, графического, табличного и выборочного методов

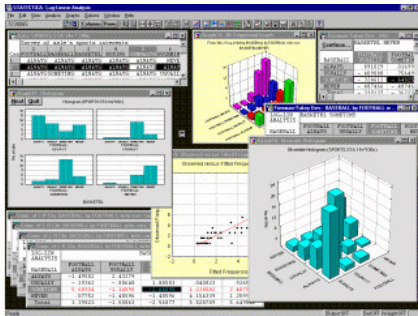


Фото из источника в списке литературы [2]

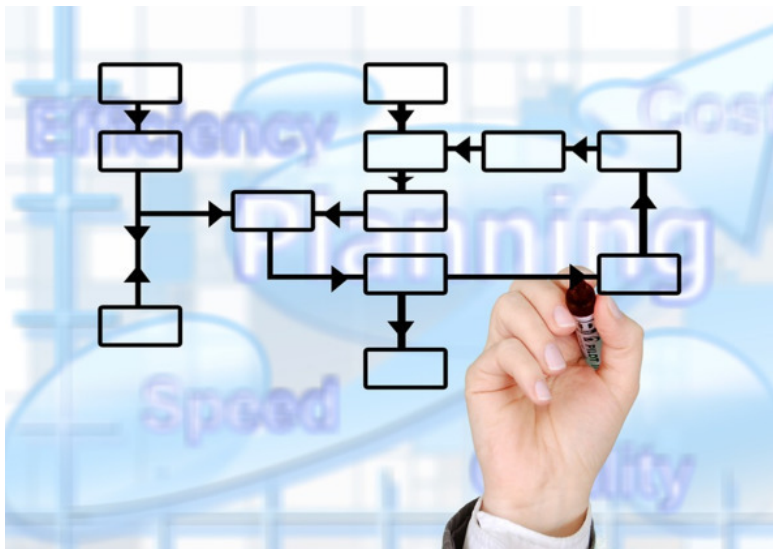
1.1 Применение в маркетинге описательного анализа данных. Метод графический

Применение в маркетинге описательного анализа данных

В маркетинге описательный метод широко применяется в целях проведения разных презентаций и подготовки аналитических маркетинговых отчетов.

Описательным анализом данных предусматривается, что при его реализации реально не только анализировать данные, но и выбирать аналитические методы их последующего анализа, к примеру, методы, которые позволяют проверять гипотезы и моделировать взаимосвязи.

Под описательной статистикой или разведочным анализом данных принято понимание статистических методов обработки данных, осуществления их систематизирования, представления их в наглядной форме – в виде графиков и разного рода таблиц, а также описания (количественного) данных посредством системы различных статистических показателей.



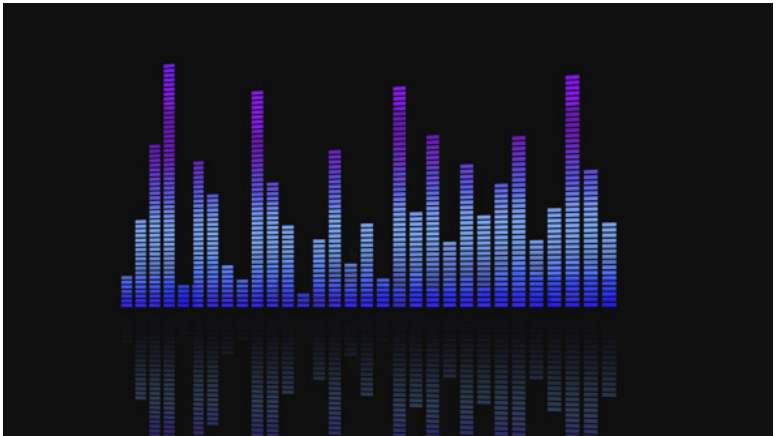
Представление информации в рассматриваемом анализе осуществляется в виде графиков и таблиц (частотных и сопряженности). Количественная информация обобщается посредством графиков, а также показателей: 1) среднего уровня (усредненного значения, медианы и моды), а также процентилей; 2) вариации (размаха, межквартильного размаха, дисперсии, коэффициента вариации, стандартного отклонения и др.); 3) формы распределения (асимметрии и эксцесса).

В анализе описательного вида принято представление информации об исследуемом явлении в агрегированном (обоб-

ценном) виде. В этих целях прибегают к использованию методов: исчисления статистических показателей; табличного; графического.

Исследование данных ориентируется на такие оценки показателей как интервальные и точечные.

Метод графический



Под графическим методом принято понимание метода, содействующего агрегированию данных при первичном описательном анализе. График представляется в виде чертежа, показывающего соотношение конкретных данных на основе применения таких средств, как изобразительные (средства) и геометрические образы. Графики содействуют представле-

нию статистических данных в наглядном виде.

Графики могут быть представлены в виде диаграмм и статистических карт. Диаграммы принято подразделять на виды согласно задачам осуществляемого анализа. Они могут выступать в качестве диаграмм: динамики, взаимосвязи сравнения и структуры. Благодаря диаграммам сравнения оказывается возможным изображение статистических данных, характеризующих разные территории (регионы, страны), объекты. Если применять диаграммы структуры, то можно наблюдать структуру и составные части изучаемой совокупности. Посредством диаграмм динамики анализируется развитие интересующих исследователя явлений во временном аспекте. На базе диаграмм взаимосвязи происходит изучение имеющихся между данными зависимостей. С помощью статистических карт отображают статистические данные на географических территориях.

Особенности графиков зависят от тех задач, которые благодаря им решаются. Строить графики в настоящее время необходимо с применением пакетов программ для статистического профессионального анализа, обладающих огромными возможностями. Они позволяют строить сотни модификаций диаграмм, которые бывают: столбиковыми; секторными; линейными и т. д.

Нередко применение графического метода происходит, когда этот метод дополняет методы аналитические (к примеру, метод, с помощью которого анализируются взаимосвязи,

или метод, выявляющий различия в группах). Но у графических средств имеются свои особенные преимущества, касающиеся выявления закономерностей, которые сложно описать количественно и проблематично обнаружить на базе прибегания к аналитическим процедурам.

1.2 Метод табличный. Метод выборочный

Метод табличный



Метод табличный – это метод универсальный. Его можно применять в разных и многих направлениях маркетинга. Прежде всего, он служит цели обеспечения удобства представления массивов данных об анализируемых объектах. Для этого прибегают к применению таблиц промежуточных (в которых представляют промежуточные итоги ис-

следований) и сводных (в которых дается представление финальных данных). Представление посредством таблиц результатов маркетинговых исследований, опросов, группировок и сводок данных делают отчет об исследовании более наглядным, профессиональным, убедительным.

Под табличным методом принято понимание метода агрегирования имеющихся данных, которое проводится на этапе осуществления первичного описательного анализа. В качестве статистической таблицы выступает система столбцов и строк. В данной системе логично и последовательно излагают разнообразную статистическую информацию об исследуемом процессе либо явлении. Табличное представление данных в маркетинге – это представление удобное и практичное, которое обеспечивает наглядность информации. Но таблица должна отличаться высоким качеством, для этого важно ее правильное построение и оформление.

Для построения разных (сложных и простых) таблиц целесообразно прибегать к использованию: профессиональных статистических пакетов, например таких, IBM SPSS Statistics; программы MS Excel. Благодаря пакетам и программам можно при минимуме временных затрат строить таблицы, осуществлять обработку данных, составлять отчеты по маркетинговым исследованиям.

Метод выборочный



В качестве выборочного метода в статистике принято понимание метода исследования общих свойств всей совокупности (генеральной совокупности или же выборки) каких-то объектов на базе анализа свойств только части данных объектов (выборочной совокупности). Применение этого метода может быть целесообразно из-за: большой обширности объекта исследования (к примеру, когда изучаются потребительские предпочтения на рынке); необходимости в сборе маркетинговой первичной информации в рамках «пилотных» исследований.

Выборочное обследование предусматривает, что нужно добиваться минимальности объема выборки при максимуме точности описания генеральной совокупности на базе выборочных данных. Поэтому выборке положено отличаться ре-

презентативностью (представительностью), чтобы объективно отражать свойства, имманентные генеральной совокупности.

При реализации выборочных обследований следует обеспечивать: количественную характеристику выборки либо определять минимальное количество объема выборки (наблюдений) для осуществления исследования; качественную характеристику выборки либо способов и методов, посредством которых происходит формирование совокупности выборочного вида.

Достижение точности итогов выборочных обследований возможно на основе прибегания к применению отличающихся сложностью методов, содействующих формирования выборок (это кластерный отбор, задание расслоения, использование отбора вероятностно-пропорционального вида, простой случайный либо неслучайный отбор, повторный либо бесповторный отбор).

Значение показателя минимального объема выборки обусловливается многими параметрами, принадлежащими оцениваемому показателю либо системе показателей (способ и методы формирования выборки, заданная надежность результатов, вариация исследуемых данных, максимально допустимая ошибка в оценивании показателей). Данный показатель определяют на базе статистических методов либо экспертно.

II Применение в маркетинге

Корреляционного и регрессионного анализа

2.1 Анализ корреляционный. Регрессионный анализ линейный и нелинейный

Анализ корреляционный

The screenshot shows a Microsoft Excel spreadsheet with the following data and formulas:

	A	B	C	D	E
1	Выручка	200,9	348,6	575,3	298,2
2	Оборотные средства	606,1	325,6	349,8	470,7
3	Оборот. средства расч	=B\$14+\$A\$14*B1	=B\$14+\$A\$14*C1	=B\$14+\$A\$14*D1	=B\$14+\$A\$14*E1
4	Отклонение (с.2-с.3)	=B2-B3	=C2-C3	=D2-D3	=E2-E3
5	У-У ср	=B2-\$B\$7	=C2-\$B\$7	=D2-\$B\$7	=E2-\$B\$7
6					
7	Уср	=CP3НАЧ(B2:E2)		(У-Уср) ²	=СУММКВ(B5:E5)
8					
9	Индекс корреляции	=КОРРЕЛ(B1:E1;B2:E2)			
10	Козф детерминации	=1-СУММКВ(B4:E4)/E7			
11					
12	коэффициенты линейн уравнени				
13	b		a		
14	=ЛИНЕЙН(B2:E2;B1:E1;1,1)	=ЛИНЕЙН(B2:E2;B1:E1;1,1)			
15					
16					
17					
18					
19					
20					
21					
22					
23					
24					

Фото из источника в списке литературы [3]

В качестве корреляционного анализа принято понимание статистического метода анализа взаимосвязи между случайными переменными-величинами в количестве двух и более. Случайные переменные-величины – это свойства изучаемых объектов наблюдения, являющиеся измеряемыми.

При проведении корреляционного анализа предусматривается исчисление коэффициентов корреляции, принимающих обычно либо отрицательные, либо положительные значения. По знаку коэффициента судят о направлении имеющейся связи, а по абсолютному значению – силу имеющейся связи.

Для оценивания направления связи между анализируемыми переменными, измерение которых производилось в шкале порядковой, прибегают к использованию ранговых непараметрических коэффициентов корреляции: коэффициентов ранговой корреляции Кендалла и коэффициента корреляции Спирмена. Применение нередко находит и исчисление коэффициентов: Фехнера (корреляция знаков), конкордации (множественная ранговая корреляция). Между переменными дихотомическими также измеряют связи посредством соответствующих метрик.

Для исчисления коэффициентов корреляции применяют способ расчета, зависящий от шкалы измерения тех из переменных, взаимосвязь между которыми исследуют.

Если переменные измерены в шкале количественного типа (шкала отношений либо шкала интервальная), то обеспечивается расчет ковариации или корреляционного момента, а на его базе – исчисление линейного коэффициента корреляции (коэффициента Пирсона).

Исчисление коэффициентов корреляции – относительно не трудное, к тому же их легко интерпретировать. Применять их могут даже люди специально не подготовленные. Однако, у рассматриваемого вида анализа – корреляционного – имеется собственная своя специфика и методика. Необходимо соблюдать предпосылки исчисления каждого из коэффициентов корреляции, а также проверять их значимость, которая базируется на принципе необходимости проверки статистических гипотез, нужно также правильно строить интервальные оценки коэффициентов. Надо также помнить, что иногда исследователям приходится сталкиваться с «ложными корреляциями», приводящими к ложным (обманчивым) выводам. Поэтому нужно практиковать расчет не только общих коэффициентов корреляции, но и частных.

На базе корреляционного анализа невозможно определение формы связи между переменными и предсказание значения одной зависимой переменной по одной либо ряду переменных независимых. Если говорить о переменных количественных, то для решения данной задачи можно прибегнуть к применению линейного регрессионного анализа.

В маркетинге корреляционный анализ в основном при-

меняют в альянсе с анализом регрессионным. Посредством первого вида анализа определяют наличие связи и степень ее тесноты, а регрессионный анализ служит для математического моделирования формы связи.

Регрессионный анализ линейный



Регрессионный анализ понимается в качестве метода исследования статистической взаимосвязи между одной

количественной зависимой переменной (результатирующей) и одной либо ряда количественных независимых переменных-факторов (предикторов или переменных объясняющих).

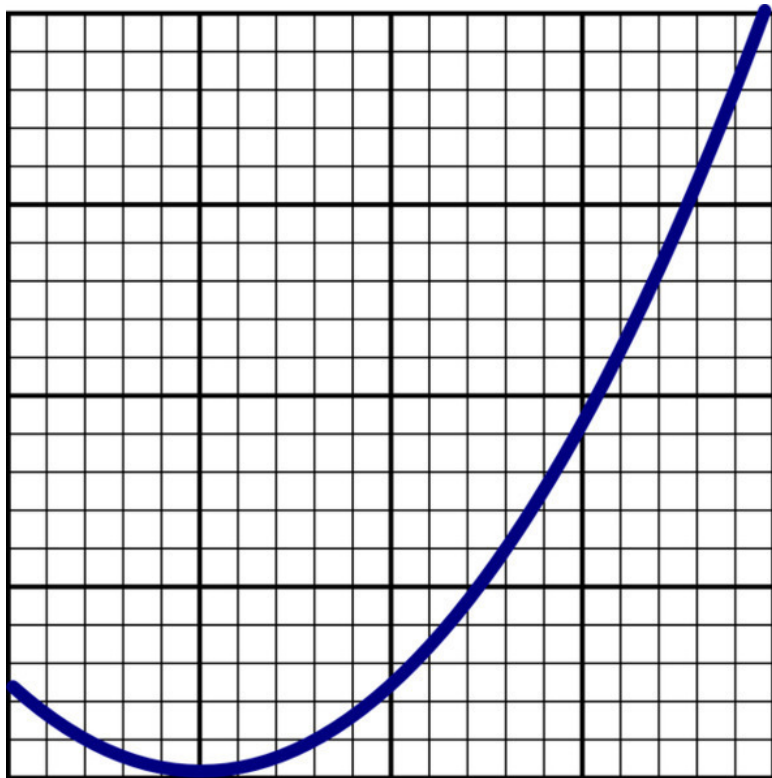
Уравнение регрессии, по сути, является выражением взаимосвязи между усредненным значением переменной результирующей и усредненными значениями переменных объясняющих. Оно представляется математической функцией, подбираемой на базе статистических исходных данных зависимой переменной и предикторов. Чаще всего прибегают к использованию функции линейного вида и осуществляют линейный регрессионный анализ.

Имеет место очень сильная связь регрессионного анализа с анализом корреляционным, предусматривающим изучение направления и тесноты связи между количественно выраженными переменными. В анализе регрессионном исследуют форму зависимости между данными переменными. То есть при применении обоих методов осуществляют исследование одной и той же взаимосвязи, но с различных сторон, эти методы, в принципе, являются взаимодополняющими. Причем практикуется выполнение корреляционного анализа производить перед проведением анализа регрессионного. После того, как посредством корреляционного анализа доказано наличие взаимосвязи, исследователь может обеспечить выражение формы данной связи, применив регрессионный анализ.

Целью регрессионного анализа является цель предсказания (прогнозирования) ожидаемого усредненного значения результирующей переменной посредством соответствующего уравнения. В маркетинге довольно часто необходимо прогнозировать разные важные показатели (к примеру, объема продаж или прибыли, или числа клиентов и т.д.).

2.2 Регрессионный анализ нелинейный. Регрессия категориальная

Регрессионный анализ нелинейный



$$y = ax^2 + bx + c$$

Под регрессией нелинейной принято понимать регрессионную модель зависимости переменной результативной от одной либо нескольких переменных объясняющих, вы-

ражаемую в виде нелинейной функции. Нелинейная модель (как и линейная) может быть парной и множественной.

Нелинейная регрессия согласно ее целям и задачам подобна регрессии линейной. Отличие обуславливается лишь формой связей и методами оценивания параметров. Выбрать форму связи зависимости нелинейного вида можно посредством: содержательного изучения исследуемого конкретного явления; опоры на итоги изучения взаимосвязи между переменными, к примеру, с применением графического метода.

Оценивание параметров нелинейной регрессии может базироваться: на линейаризации уравнения благодаря подходящим преобразованиям и оценки его параметров посредством применения метода наименьших квадратов; на оценке параметров на базе метода максимального правдоподобия и применения процедур оптимизационных методов.

Регрессии нелинейные различают согласно: включаемым в эти регрессии предикторам (такие нелинейного вида регрессии являются линейными по параметрам); включаемым в регрессии предикторам и подвергаемым оценке параметрам.

Если функции являются нелинейными по переменным объясняющим, возможно сведение их к линейным посредством замены переменных.

Если функции являются нелинейными по переменным-факторам и подвергаемым оцениванию параметрам, их сведение к линейным моделям происходит благодаря логарифмированию.

рифмированию и замене переменных.

Если подобрать линеаризующее преобразование невозможно, для оценивания параметров прибегают к использованию методов нелинейной оптимизации на базе исходных данных.

Наилучшая нелинейная модель обычно выбирается на базе наименьшей стандартной остаточной ошибки, исчисленной для разных моделей. Если имеет место наличие ряда нелинейных моделей с сопоставимой точностью, рекомендуется останавливать выбор на модели, отличающейся большей простотой.

Регрессия категориальная

Диаграмма рассеивания — график, оси которого соответствуют значениям двух переменных, а каждый испытуемый представляет собой

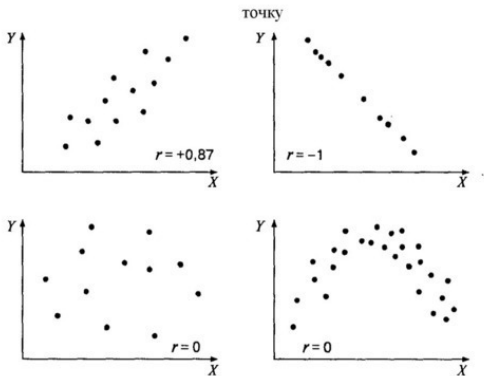


Фото из источника в списке литературы [4]

В качестве регрессии категориальной принято рассмотрение статистического метода моделирования взаимосвязи между категориальными переменными: зависимой и независимыми. Для построения модели рассматриваемой регрессии прибегают к шкалированию либо оцифровке переменных путем присвоения категориям числовых значений. После этого идет построение оптимального уравнения линейной регрессии относительно преобразованных новых переменных.

В данной модели и переменная и ее предикторы – категориальные.

Применение категориальной регрессии в маркетинге происходит, когда нужно описать покупательскую удовлетворенность в зависимости от таких факторов как простота совершения покупки, цена, качество товара. Посредством уравнения категориальной регрессии возможно прогнозирование уровня покупательской удовлетворенности в зависимости от всевозможных сочетаний значений категориальных независимых переменных.

2.3 Регрессия логистическая. Регрессия мультиномиальная логистическая

Регрессия логистическая

Под регрессией логистической или логит-регрессией понимают статистическую модель, используемую, чтобы предсказывать вероятности возникновения какого-то события посредством логистической функции. Эта регрессия относится к классу моделей бинарного выбора, в которых зависимая переменная является дихотомической (бинарной).

Зависимой переменной могут приниматься только значения в количестве двух и означать, к примеру, принадлежность к конкретной группе (скажем, к группе надежных клиентов либо ненадежных клиентов), совершаемое действие (покупка либо непокупка товара), ответы «да» либо «нет» (рекламный ролик нравится либо не нравится).

Построение обыкновенной регрессионной модели линейного типа с зависимыми бинарными переменными не допускается ввиду того, что если это построение будет иметь место, предсказанные значения зависимой переменной интерпретировать окажется практически невозможно.

Измерения значений факторов в моделях бинарного выбора – только количественные. В эти модели допускается

включение категориальных переменных (выступающих в качестве факторов). В данных моделях обеспечивается построение регрессионной модели зависимости с принятием во внимание вероятности, что результирующей дихотомической переменной будет принято значение 0 или 1, если значение факторов – заданное.

Для того чтобы смоделировать вероятность зависимой дихотомической переменной, нужно произвести подбор специальной монотонно возрастающей функции, значения которой могут варьироваться лишь от 0 до 1.

В моделях бинарного выбора в качестве специальной функции может быть выбрана функция: 1) логистическая; 2) стандартного нормального распределения.

Если модель бинарного выбора построена на базе логистической функции, то она рассматривается как логистическая регрессия или логит-модель. Если модель бинарного выбора построена на базе функции, стандартного нормального распределения, то ее рассматривают как пробит-модель.

Посредством логистической регрессии осуществляется прогнозирование вероятности отклика для зависимой переменной от переменных независимых, которые включены в модель. Прогнозные значения вероятности можно использовать для разделения наблюдений на две группы.

При построении модели регрессии логистической можно осуществить отдельный анализ – анализ Receiver Operator

Characteristic (ROC-кривых). Посредством данного анализа можно осуществить выбор оптимального порогового значения вероятности для классификации. ROC-кривую используют, чтобы представить результаты бинарной классификации и оценки уровня ее эффективности.

Использование логистической регрессии распространяется на решение задач, связанных с моделированием взаимосвязи и классификацией наблюдений. Она находит применение в скоринге: банковском (на ее основе возможно построение рейтинга заемщиков и управления кредитными рисками); потребительском (для моделирования потребительского поведения).

Регрессия мультиномиальная логистическая

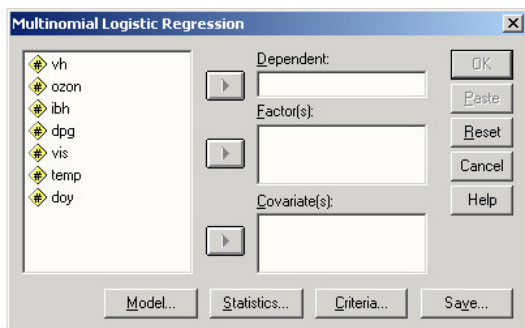


Фото из источника в списке литературы [5]

В качестве логистической регрессии мультиномиальной рассматривают общий случай модели логистической регрессии, в ней у зависимой переменной имеются категории в количестве более двух.

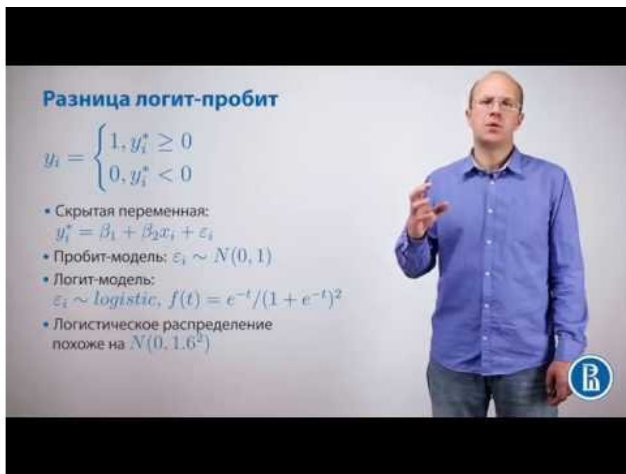
Измерение зависимой переменной (ковариаты) в рассматриваемой регрессии возможно в таких шкалах, как порядковая и номинальная. В качестве нее может выступать переменная потребительского выбора торговой марки. Переменные независимые (факторы) могут быть количественными либо категориальными.

В данной модели для каждой из категорий переменной зависимой предусматривается построение уравнения логистической бинарной регрессии. Причем одной из категорий переменной зависимой отводится роль переменной опорной, и происходит сравнение с ней всех других категорий.

Посредством уравнения мультиномиальной логистической регрессии прогнозируется показатель вероятности принадлежности к каждой категории зависимой переменной согласно значениям переменных независимых.

2.4 Пробит-модель регресси. Регрессия Кокса. Анализ временных рядов

Пробит-модель регрессии



Разница логит-пробит

$$y_i = \begin{cases} 1, & y_i^* \geq 0 \\ 0, & y_i^* < 0 \end{cases}$$

- Скрытая переменная:
 $y_i^* = \beta_1 + \beta_2 x_i + \varepsilon_i$
- Пробит-модель: $\varepsilon_i \sim N(0, 1)$
- Логит-модель:
 $\varepsilon_i \sim \text{logistic}, f(t) = e^{-t} / (1 + e^{-t})^2$
- Логистическое распределение похоже на $N(0, 1.6^2)$

В

Фото из источника в списке литературы [6]

Пробит-модель является статистической моделью бинарного выбора, используемой для того, чтобы предсказывать вероятность возникновения какого-то события на базе функции нормального стандартного распределения.

Модель пробит-регрессии, подобно модели логистической регрессии, относят к виду моделей бинарного выбора. По этой причине задачи ее построения и функции такие же, как в логит-модели.

В модели пробит-регрессии выражение расчетного значения зависимой переменной выступает в качестве значения функции нормального стандартного закона распределения. Пробит является значением, для которого исследователи вычисляют функцию нормального стандартного распределения. Имеет место зависимость значения пробита от комбинированных линейных значений факторных переменных. Для пробит-модели (также как и для логит-модели) зависимая переменная – дихотомическая. К факторам в пробит-модели предъявляется требование, чтобы они были количественно выраженными либо категориальными, но преобразованными в переменные дихотомические.

Применение пробит-модели относительно сферы аналогично применению логистической регрессии. Если осуществить моделирование и классификацию по пробит-модели и также по модели логистической регрессии, то результаты окажутся весьма сходными. Но в некоторых случаях результаты могут различаться.

Регрессия Кокса

Статины и/или Реваскуляризация (Регрессионная модель Кокса)

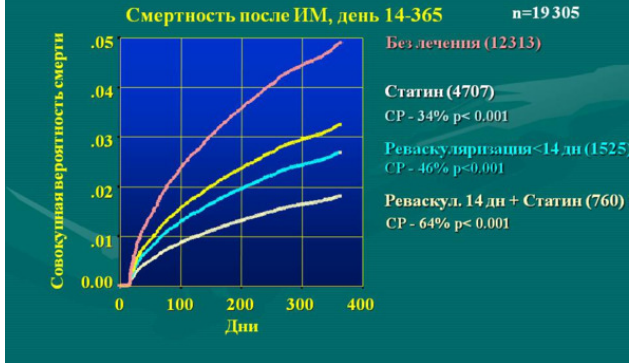


Фото из источника в списке литературы [7]

Регрессионную модель Кокса считают статистической моделью зависимости функции риска от переменных-факторов независимого вида.

Регрессию Кокса рассматривают в качестве модели отличающихся пропорциональностью рисков. Благодаря ей прогнозируют риск наступления события для какого-то объекта и оценивают влияние определенных заранее независимых предикторов (переменных) на данный риск. Риск рассматривают в качестве зависящей от времени функции. Риск не является вероятностью, поэтому его значения могут превышать единицу.

Объектом может быть клиент, для которого в маркетин-

ге практикуется прогнозирование риска наступления некоего события. Объект находится в поле зрения априори (то есть его постоянно наблюдают), в любой временной отрезок возможно наступление события, приводящего к его выбытию из группы риска. К примеру, таким событием может оказаться отказ клиента от товара либо услуги компании или его неспособность оплаты кредита.

Конец ознакомительного фрагмента.

Текст предоставлен ООО «ЛитРес».

Прочитайте эту книгу целиком, [купив полную легальную версию](#) на ЛитРес.

Безопасно оплатить книгу можно банковской картой Visa, MasterCard, Maestro, со счета мобильного телефона, с платежного терминала, в салоне МТС или Связной, через PayPal, WebMoney, Яндекс.Деньги, QIWI Кошелек, бонусными картами или другим удобным Вам способом.