



Николай Баринов FRICS, к.т.н.

директор по научно-методической работе

ПРАКТИКА ПРИМЕНЕНИЯ МНОГОМЕРНОГО РЕГРЕССИОННОГО АНАЛИЗА В ЗАДАЧАХ ИНДИВИДУАЛЬНОЙ ОЦЕНКИ НЕДВИЖИМОСТИ

Санкт-Петербург - Нижний Новгород

2014

ЧТО, ГДЕ, КОГДА?

- **МНОГОМЕРНЫЙ РЕГРЕССИОННЫЙ АНАЛИЗ – ЕДИНСТВЕННЫЙ ОБОСНОВАННЫЙ КОЛИЧЕСТВЕННЫЙ МЕТОД В РАМКАХ СРАВНИТЕЛЬНОГО ПОДХОДА К ОЦЕНКЕ НЕДВИЖИМОСТИ.**
 - ОРГАНИЧНО ВПИСЫВАЕТСЯ В ПОНИМАНИЕ СТОИМОСТИ КАК СРЕДНЕЙ ПО ВСЕМУ РЫНКУ ЦЕНЫ НА АНАЛОГИЧНЫЙ АКТИВ.
 - ПОЗВОЛЯЕТ ПОЛУЧАТЬ ТОЧЕЧНУЮ ОЦЕНКУ СТОИМОСТИ И ИНТЕРВАЛ ЕЕ НЕОПРЕДЕЛЕННОСТИ.
- **ГЛАВНЫЙ НЕДОСТАТОК МЕТОДА – НЕВОЗМОЖНОСТЬ ПРИМЕНЕНИЯ НА ТРЕХ-ПЯТИ АНАЛОГАХ. 😊**
- **РЕЧЬ ИДЕТ О ТРЕТЕЙСКОЙ (ИЗМЕРИТЕЛЬНОЙ) ОЦЕНКЕ НА АКТИВНЫХ РЫНКАХ. В УСЛОВИЯХ ЦЕНООБРАЗУЮЩЕЙ ОЦЕНКИ (УНИКАЛЬНЫЕ ОБЪЕКТЫ, ПАССИВНЫЕ РЫНКИ) НАДЕЯТЬСЯ НА МЕТОД НЕ СТОИТ.**



ОСОБЕННОСТИ ИНДИВИДУАЛЬНОЙ ОЦЕНКИ НЕДВИЖИМОСТИ

- «ПРОСТРАНСТВЕННЫЕ» МОДЕЛИ: ОПИСЫВАЕТСЯ ВЛИЯНИЕ **ПРОСТРАНСТВЕННЫХ ФАКТОРОВ**. ВРЕМЯ ВКЛЮЧАЕТСЯ В МОДЕЛИ, СКОРЕЕ, КАК ИСКЛЮЧЕНИЕ. ЭТО МНОГОЕ УПРОЩАЕТ.
- МОДЕЛИРОВАНИЕ ЦЕНООБРАЗОВАНИЯ НА **МАЛЫХ ГЕНЕРАЛЬНЫХ СОВОКУПНОСТЯХ** (ЛОКАЛЬНЫХ РЫНКАХ «ВОКРУГ» ОБЪЕКТА ОЦЕНКИ) ПРИ ОТНОСИТЕЛЬНО **БОЛЬШИХ ДОЛЯХ ОТБОРА** АНАЛОГОВ.
- НАЛИЧИЕ «КАЧЕСТВЕННЫХ» ФАКТОРОВ, ПОДЛЕЖАЩИХ ОЦИФРОВКЕ ДЛЯ УЧЕТА В МОДЕЛИ.
- **НЕЛИНЕЙНЫЕ ЗАВИСИМОСТИ** МОДЕЛИРУЕМОЙ ВЕЛИЧИНЫ (СТОИМОСТИ, РЫНОЧНОЙ АРЕНДНОЙ ПЛАТЫ) ОТ ВЛИЯЮЩИХ ФАКТОРОВ.
- ДЛЯ КАЖДОЙ ОЦЕНКИ - СВОЯ МОДЕЛЬ. «ЗДЕСЬ И СЕЙЧАС».



СКОЛЬКО НУЖНО АНАЛОГОВ?

Ровно столько, чтобы построить адекватную модель 😊

Чем больше однородных аналогов – тем лучше проявляются связи и меньше неопределенность оценки.

Но в реальной жизни - число аналогов и их близость объекту оценки и друг другу (однородность выборки) – обменные параметры.

На однородных выборках адекватные модели могут быть построены при числе аналогов $n \geq 2(k+1)$, где k – число учитываемых ценообразующих факторов.

$$n - k - 1 = \text{ЧСС} \quad \longrightarrow \quad n = \text{ЧСС} + (k + 1) \quad \longrightarrow \quad \text{ЧСС} \geq k + 1$$

Каждому фактору – по степени свободы, и еще одну на всех!

На практике: Ищем аналоги, начиная с «точно таких же». Считаем факторы, по которым различаются объекты, и набираем аналоги, пока, как минимум, не обеспечим требуемое число степеней свободы. Иначе - не стоит и начинать строить модель.



СКОЛЬКО ФАКТОРОВ УЧИТЫВАТЬ?

В модели должны быть учтены **ВСЕ ЦЕНООБРАЗУЮЩИЕ ФАКТОРЫ**, по которым отличаются объекты сравнения и **ВАРИАЦИЯ КОТОРЫХ СПОСОБНА ОБЪЯСНИТЬ НАБЛЮДАЕМУЮ ВАРИАЦИЮ ЦЕН.**

ПРОПУСК (НЕВКЛЮЧЕНИЕ В МОДЕЛЬ) **СУЩЕСТВЕННО ВЛИЯЮЩЕГО ФАКТОРА** — **ГРУБАЯ ОШИБКА СПЕЦИФИКАЦИИ МОДЕЛИ**, ПРИВОДЯЩАЯ К **СМЕЩЕНИЮ ОЦЕНОК МОДЕЛИ.**

УЧЕТ «ЛИШНЕГО» ФАКТОРА — ГОРАЗДО МЕНЬШЕЕ ЗЛО — **СМЕЩЕНИЯ ОЦЕНОК НЕ ПРОИСХОДИТ** (РАСТЕТ ЛИШЬ НЕОПРЕДЕЛЕННОСТЬ ОЦЕНКИ).

НА ПРАКТИКЕ:

1. При достаточном числе аналогов в модель включаются все потенциально влияющие факторы, по которым наблюдаются различия отобранных объектов сравнения. В процессе построения модели незначимые факторы удаляются из нее с учетом статистических показателей и данных рынка. Не включать факторы в модель по результатам предварительной оценки корреляции между фактором и зависимой переменной — недопустимо.
2. При недостатке аналогов выявляются основные влияющие факторы, строятся предварительные модели, затем в них добавляются «пропущенные» факторы. Невключение факторов, по которым объекты различаются, должно быть обосновано вне модели.
3. Для «ГАРАНТИИ» несмещенности - можно оставлять в модели один-два не вполне значимых фактора.



КАК ОЦИФРОВЫВАТЬ «КАЧЕСТВЕННЫЕ» ФАКТОРЫ?

Число градаций «качественных» факторов (состояние конструкций, окружающей среды и т.п.) должно выбираться из условия **НАДЕЖНОГО ОТНЕСЕНИЯ КАЖДОГО ОБЪЕКТА СРАВНЕНИЯ К ОПРЕДЕЛЕННОЙ ГРАДАЦИИ ФАКТОРА.**

НЕВЕРНОЕ ОТНЕСЕНИЕ ОБЪЕКТА К ДРУГОЙ (СМЕЖНОЙ) ГРАДАЦИИ ФАКТОРА — **ГРУБАЯ ОШИБКА** ИЗМЕРЕНИЯ, ЗАТРУДНЯЮЩАЯ ПОСТРОЕНИЕ АДЕКВАТНОЙ МОДЕЛИ.

ПОТЕРЯ ЧАСТИ ИНФОРМАЦИИ ПРИ УМЕНЬШЕНИИ ЧИСЛА ГРАДАЦИЙ ФАКТОРА В УСЛОВИЯХ ЧЕТКОГО ИХ РАЗЛИЧИЯ И, ТЕМ САМЫМ, ИСКЛЮЧЕНИЯ ОШИБОК ОТНЕСЕНИЯ К ГРАДАЦИИ - МЕНЬШЕЕ ЗЛО.

НА ПРАКТИКЕ: Рекомендуется описывать градации «качественных» факторов набором вспомогательных признаков (показателей, характеристик), поддающихся количественному измерению. Это позволяет надежно относить наблюдаемое значение фактора к одной из сформированных градаций.

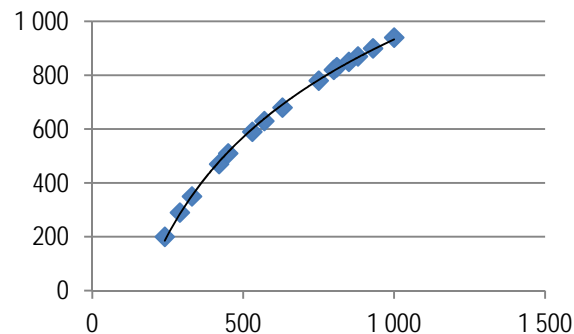
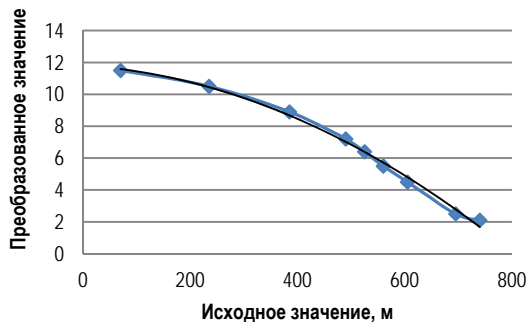
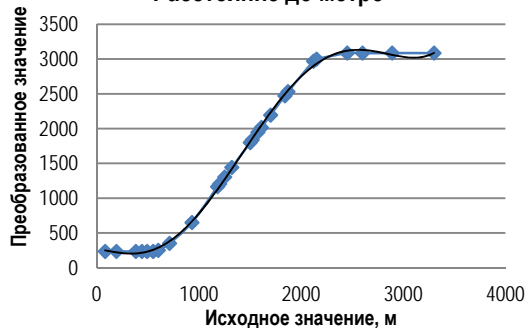
НЕ РЕКОМЕНДУЕТСЯ использовать факторы, не поддающиеся описанию количественно измеряемыми вспомогательными признаками (например, «престижность местоположения»)

НЕ РЕКОМЕНДУЕТСЯ использовать более 7 градаций «качественных» факторов.

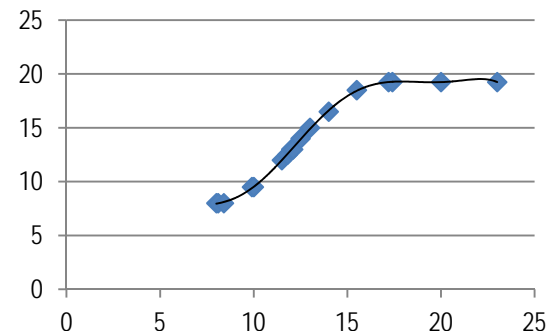
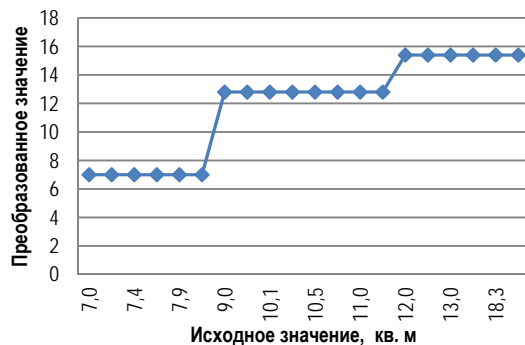
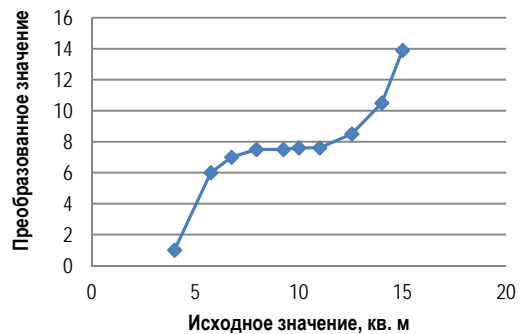


НЕЛИНЕЙНЫЕ СВЯЗИ НА РЫНКЕ

Расстояние до метро



Площадь кухни

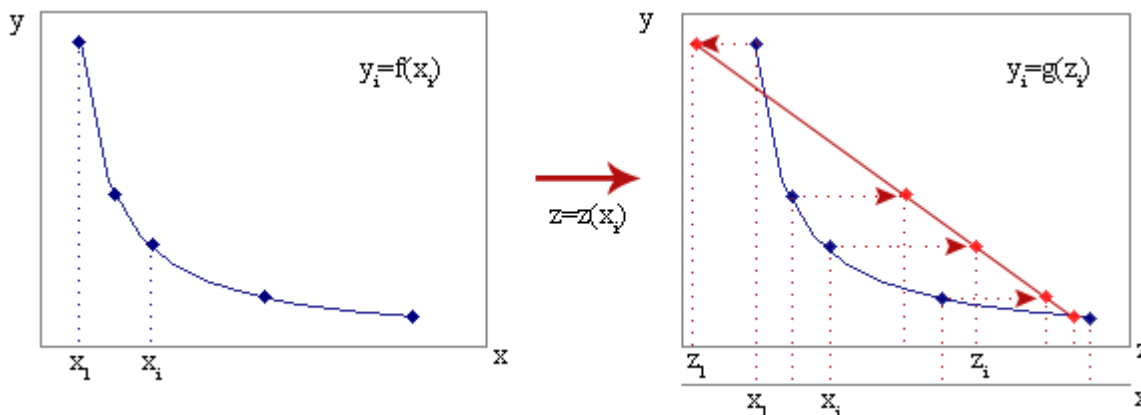


КАК УЧЕСТЬ НЕЛИНЕЙНЫЕ СВЯЗИ

Рекомендации в литературе – строить *нелинейные* регрессионные модели.

Недостатки этого подхода – *однотипность* нелинейности для всех или большинства факторов, сложности выбора вида и параметров зависимостей.

Для зависимостей, заданных дискретно, *всегда существует* линеаризующее преобразование



На практике - линеаризующее преобразование – *свое* для каждой из переменных - удобно проводить с помощью оптимизационной процедуры - макроса MS Excel – «Поиск решения» с контролем получаемых зависимостей на соответствие экономическим гипотезам.

Именно здесь проявляется *сужденческий потенциал* оценщика как эксперта рынка.

Ускорить процесс помогает применение функции MS Excel «Просмотр»

<http://www.appraiser.ru/default.aspx?SectionId=35&Id=3605>



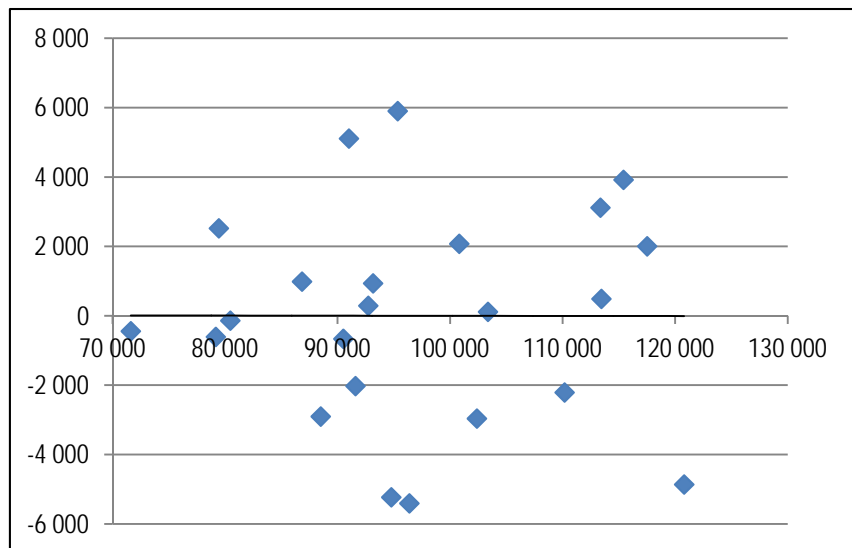
АНАЛИЗ ОСТАТКОВ МОДЕЛИ

«Вы будете видеть в этих остатках потенциально неограниченный источник для зарождения новых идей, а, возможно, и новых теорий.

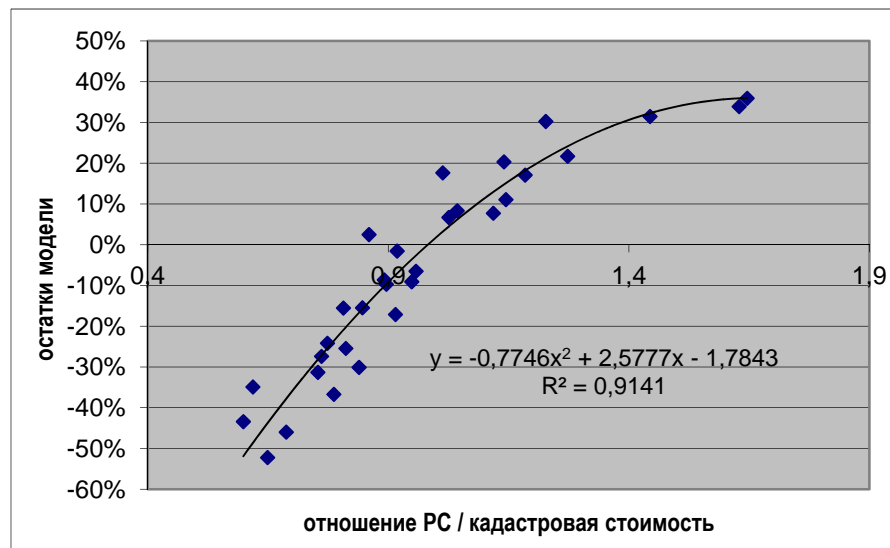
...Остатки, взятые по отдельности, указывают когда, где и в какой степени существующая модель не смогла объяснить наблюдаемые события».

Кристофер Доугерти. Введение в эконометрику. 3-е издание, 2009

Случайные остатки модели



Пропущенный неизвестный фактор



КРИТЕРИИ ОЦЕНКИ АДЕКВАТНОСТИ МОДЕЛИ

К нашему огорчению☹, **единственного критерия адекватности (соответствия моделируемому рынку) регрессионной модели не существует.**

Можно говорить о наборе показателей, обеспечивающем адекватность модели рыночным данным и закономерностям:

- **средняя и максимальная ошибки аппроксимации;**
- **статистическая значимость большинства коэффициентов модели;**
- **соответствие знаков коэффициентов модели экономическим гипотезам;**
- **соответствие вида функции влияния каждого из факторов на зависимую переменную экономическим гипотезам;**
- **случайность распределения остатков модели от значений зависимой переменной (модельных);**
- **ширина доверительного интервала для точечной оценки;**
- **устойчивость (балансировка) регрессионной модели.**

На практике – на последних этапах построения моделей лучшая из них отбирается не столько по формальным показателям, сколько по соображениям соответствия рынку.



РАСЧЕТ ДОВЕРИТЕЛЬНОГО ИНТЕРВАЛА ДЛЯ СРЕДНЕГО ЗНАЧЕНИЯ ЦЕНЫ (СТОИМОСТИ)

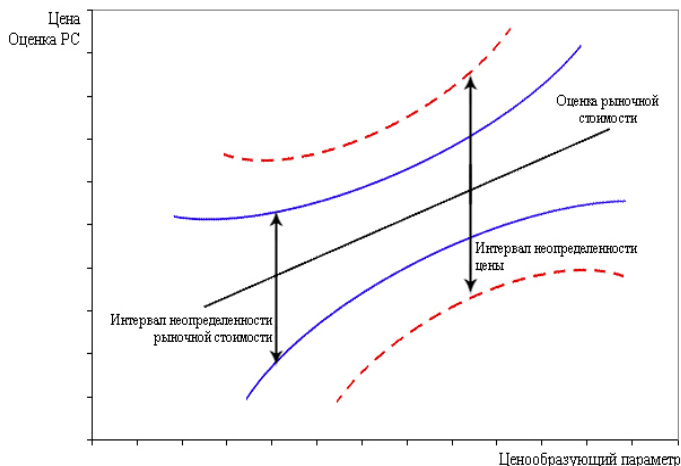
Однофакторная модель

$$\tilde{y} = a_0 + a_1 f(z) \rightarrow \tilde{y} = b_0 + b_1 x$$

Доверительный интервал для оценки стоимости

$$\tilde{y}_o \pm t_\alpha s \sqrt{\frac{1}{n} + \frac{(x_o - \bar{x})^2}{\sum_i (x_i - \bar{x})^2}}$$

Модель парной регрессии



Многофакторная модель

$$\tilde{y} = b_0 + b_1 x_{i1} + b_2 x_{i2} + \dots + b_k x_{ik}$$

Доверительный интервал оценки стоимости в матричном виде

$$\tilde{y}_o \pm t_\alpha s \sqrt{x_o^T (X^T X)^{-1} x_o}$$

На реальных рынках (при конечных генеральных совокупностях цен) и для корректных моделей это соотношение можно рассматривать как оценку сверху для интервала неопределенности РС.

Примеры расчета доверительного интервала инструментами MS Excel:

<http://www.appraiser.ru/default.aspx?SectionId=32&q=posts&t=5828&p=2>

<http://www.appraiser.ru/default.aspx?SectionId=32&q=posts&t=5350&p=16>



ПРЕДПОСЫЛКИ ЛИНЕЙНОЙ РЕГРЕССИИ

Регрессионные оценки корректны при выполнении некоторых гипотез (предпосылок):

- Между зависимой и влияющими переменными существует линейная зависимость.
Т.е. модель верно специфицирована (правильно определен набор факторов и функции их влияния)
- Факторные переменные детерминированы (неслучайны), наблюдаются без ошибок и линейно независимы
Наблюдаются без ошибок: измерены с достаточной для прикладной задачи точностью .
Линейно независимы: в модели не включаются факторы, являющиеся линейной комбинацией других.
- Остатки модели случайны (зависимая переменная не коррелирует с остатками)
- Среднее значение остатков равно нулю
Выполняется автоматически в моделях с постоянным членом
- Дисперсия остатков модели постоянна (гомоскедастичность остатков)

Дополнительно, для проведения статистических тестов и интервальных оценок:

- Остатки модели имеют многомерное нормальное распределение

Нормального (как и любого другого наперед заданного) распределения моделируемой величины (цен, арендных ставок и т.п.) НЕ ТРЕБУЕТСЯ.



ПРЕДПОСЫЛКИ РЕГРЕССИИ И ПОСЛЕДСТВИЯ ИХ НАРУШЕНИЙ

Нарушение условия построения /предпосылки регрессии

Объем выборки не превышает числа варьируемых переменных $n \leq k$, либо между переменными существует строгая линейная зависимость

Малый объем и малая доля отбора выборки

Выборка неоднородна (включены объекты, не подчиняющиеся закономерностям основной совокупности)

Неправильная спецификация модели – **недобор** влияющих переменных

Неправильная спецификация модели – **перебор** влияющих переменных

Сопряженность (мультиколлинеарность) влияющих переменных

Негативные последствия, возможные действия для устранения эффекта

Невозможно вычислить коэффициенты уравнения регрессии. Необходимо увеличить объем выборки, повышать ее однородность, а также **исключить** из рассмотрения строго зависимые переменные

Снижение эффективности, увеличение дисперсии и интервалов неопределенности оценок. Затруднено применение статистических критериев.
Необходимо увеличивать объем выборки.

Регрессионные оценки смещены. Также **смещена оценка моделируемой величины** (уравнения в целом). **Модель неадекватна.**
Необходимо **удалять из выборки** наблюдения других совокупностей

Смещение и несостоятельность оценок. Модель неадекватна.
Необходимо **пересмотреть спецификацию модели**

Снижение эффективности (увеличение дисперсии) оценок. Оценки не смещены.

Снижение эффективности оценок. Оценки коэффициентов регрессии смещены, оценка моделируемой величины (уравнения в целом) остается несмещенной
Затруднена интерпретация модели в предметной области



ПРЕДПОСЫЛКИ РЕГРЕССИИ И ПОСЛЕДСТВИЯ ИХ НАРУШЕНИЙ

Нарушение условия построения /предпосылки регрессии

Значения остатков неслучайны относительно зависимой переменной

Значения остатков зависят от влияющих переменных (для переменных со случайной составляющей)

Гетероскедастичность (различная дисперсия) остатков

Отсутствие нормальности распределения остатков модели

Наличие автокорреляции остатков (как правило, для временных рядов)

Негативные последствия, возможные действия для устранения эффекта

Значения зависимой переменной также неслучайны, оценки смещены, модель неадекватна.

Необходимо пересмотреть спецификацию модели - вид функциональной связи и/или состав влияющих факторных переменных

Регрессионные оценки смещены, модель неадекватна.

Необходимо пересмотреть спецификацию модели - вид функциональной связи и/или состав влияющих факторных переменных.

Коэффициенты регрессии остаются несмещенными, но становятся неэффективными. Затруднена проверка значимости коэффициентов по t-критерию, а также оценка доверительных интервалов моделируемой величины
В качестве первого шага борьбы рекомендуется переход к удельным показателям

Затруднено построение интервальных оценок, а также проверка значимости коэффициентов регрессии и уравнения в целом. При распределениях генеральных совокупностей, близких к симметричным, надежные интервальные оценки могут быть получены в отсутствие нормальности распределения остатков

Коэффициенты регрессии остаются несмещенными, но становятся неэффективными.
Рекомендуется пересмотреть спецификацию модели



ПРЕДПОСЫЛКИ РЕГРЕССИИ И ПОСЛЕДСТВИЯ ИХ НАРУШЕНИЙ

Итак, **реальных задач** при построении регрессионных моделей для индивидуальной оценки недвижимости **всего пять** 😊:

➤ **обеспечение требуемого объема выборки аналогов (числа степеней свободы)**

➤ **обеспечение однородности выборки аналогов**

Однородной признается выборка, в составе которой все объекты, включая оцениваемый, имеют единый набор ценообразующих факторов и ценообразование по каждому из факторов может быть описано единой функцией.

На практике - объекты должны принадлежать одному сегменту (субсегменту) рынка.

➤ **недопущение пропуска (неучета) существенно влияющих факторов**

Реальный прием – сохранение в модели одного-двух факторов, значимость которых недостаточна по формальным признакам

➤ **корректная спецификация модели – учет нелинейных связей факторов с моделируемой величиной**

Реальный прием - использование оптимизационных процедур линеаризации модели с табличным заданием функций и контролем вида полученных зависимостей по графикам.

➤ **контроль случайности остатков модели от значений моделируемой (зависимой) переменной и влияющих факторов**

Реальный прием – построение точечных диаграмм остатков модели от модельных значений зависимой переменной.

Для «пространственных» моделей зависимость остатков от влияющих факторов нетипична. Легко проверяется инструментами макроса «Регрессия» электронных таблиц MS Excel



ПРИМЕРЫ ПРАКТИЧЕСКОГО ПОСТРОЕНИЯ РЕГРЕССИОННЫХ МОДЕЛЕЙ В ОЦЕНКЕ

Примеры **пошагового построения** в среде электронных таблиц MS Excel моделей множественной регрессии для задач оценки активов (с необходимыми комментариями):

➤ **оценка рыночной стоимости квартиры:**

<http://www.appraiser.ru/default.aspx?SectionId=32&g=posts&t=5350&p=16>

<http://www.appraiser.ru/default.aspx?SectionId=32&g=posts&m=162883#162883>

➤ **оценка рыночной стоимости сухогруза:**

<http://www.appraiser.ru/default.aspx?SectionId=32&g=posts&t=5828&p=2>

Пример текста раздела **практического отчета** по оценке рыночной стоимости земельного участка методом регрессионного анализа :

<http://www.appraiser.ru/default.aspx?SectionId=32&g=posts&m=162883#162883>



СПАСИБО ЗА ВНИМАНИЕ

Николай Баринов

FRICS, к.т.н.

директор по научно-методической работе

ООО «ИНКО»

http: www.intconsalting.ru

e-mail: N.Barinov@intconsulting.ru

mobile: [+7 921 941 9037](tel:+79219419037)

