

ПРИМЕНЕНИЕ РЕГРЕССИОННОГО АНАЛИЗА В ЗАДАЧАХ ОПРЕДЕЛЕНИЯ СТОИМОСТНЫХ ХАРАКТЕРИСТИК ОБЪЕКТОВ НЕДВИЖИМОСТИ

Санкт-Петербург - Нижний Новгород - Петропавловск-Камчатский - Екатеринбург – Алматы –
Самара – Омск - Красноярск – Ростов-на-Дону

2014 - 2017

ЧТО, ГДЕ, КОГДА?

- **РЕГРЕССИЯ - ЗАВИСИМОСТЬ СРЕДНЕГО ЗНАЧЕНИЯ КАКОЙ-ЛИБО СЛУЧАЙНОЙ ВЕЛИЧИНЫ ОТ НЕКОТОРОЙ ДРУГОЙ ВЕЛИЧИНЫ ИЛИ ОТ НЕСКОЛЬКИХ ВЕЛИЧИН**
- **РЕГРЕССИОННЫЙ АНАЛИЗ – РАЗДЕЛ МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ, ОБЪЕДИНЯЮЩИЙ ПРАКТИЧЕСКИЕ МЕТОДЫ ИССЛЕДОВАНИЯ РЕГРЕССИОННОЙ ЗАВИСИМОСТИ МЕЖДУ ВЕЛИЧИНАМИ ПО СТАТИСТИЧЕСКИМ ДАННЫМ**
- **РЕГРЕССИОННЫЙ АНАЛИЗ РЕШАЕТ СЛЕДУЮЩИЕ ОСНОВНЫЕ ЗАДАЧИ:**
 1. **ВЫБОР МОДЕЛИ РЕГРЕССИИ, ЧТО ЗАКЛЮЧАЕТ В СЕБЕ ПРЕДПОЛОЖЕНИЯ О ЗАВИСИМОСТИ ФУНКЦИЙ РЕГРЕССИИ ОТ ВЛИЯЮЩИХ ПЕРЕМЕННЫХ И ПАРАМЕТРОВ РЕГРЕССИИ**
 2. **ОЦЕНКА ПАРАМЕТРОВ РЕГРЕССИИ В ВЫБРАННОЙ МОДЕЛИ МЕТОДОМ НАИМЕНЬШИХ КВАДРАТОВ**
 3. **ПРОВЕРКА СТАТИСТИЧЕСКИХ ГИПОТЕЗ О РЕГРЕССИИ**

МАТЕМАТИЧЕСКАЯ ЭНЦИКЛОПЕДИЯ: http://enc-dic.com/enc_math/Regressija-3920.html ,

ЧТО, ГДЕ, КОГДА?

- **МНОГОМЕРНЫЙ РЕГРЕССИОННЫЙ АНАЛИЗ – ЕДИНСТВЕННЫЙ ЗАЩИЩАЕМЫЙ КОЛИЧЕСТВЕННЫЙ МЕТОД В РАМКАХ СРАВНИТЕЛЬНОГО ПОДХОДА К ОЦЕНКЕ НЕДВИЖИМОСТИ**
 - СООТВЕТСТВУЕТ ПОНИМАНИЮ РЫНОЧНОЙ СТОИМОСТИ КАК СРЕДНЕЙ ПО ВСЕМУ РАССМАТРИВАЕМОМУ РЫНКУ ЦЕНЫ НА АНАЛОГИЧНЫЙ АКТИВ (ОБЪЕКТ)
 - ПОЗВОЛЯЕТ ПОЛУЧАТЬ ТОЧЕЧНУЮ ОЦЕНКУ СТОИМОСТИ И ИНТЕРВАЛ ЕЕ НЕОПРЕДЕЛЕННОСТИ
 - ПРИЗНАН ФЕДЕРАЛЬНЫМИ СТАНДАРТАМИ ОЦЕНКИ РФ (ФСО №7), ТРЕБОВАНИЯМИ КВАЛИФИКАЦИОННОГО ЭКЗАМЕНА ПО ОЦЕНКЕ НЕДВИЖИМОСТИ (2017), ЗАРУБЕЖНОЙ И ОТЕЧЕСТВЕННОЙ ПРАКТИКАМИ ОЦЕНКИ, А ТАКЖЕ СУДЕБНЫМИ СИСТЕМАМИ РФ И США
см. «Вопросы оценки», №03(77), 2014, с.11-21 http://srroo.ru/upload/iblock/b29/vo_03_14.pdf
- **ГЛАВНЫЙ НЕДОСТАТОК МЕТОДА – НЕВОЗМОЖНОСТЬ ПРИМЕНЕНИЯ НА ТРЕХ-ПЯТИ АНАЛОГАХ 😊**
- **ПРИМЕНЯЕТСЯ НА АКТИВНЫХ РЫНКАХ.**
В УСЛОВИЯХ «ЦЕНООБРАЗУЮЩЕЙ» ОЦЕНКИ (УНИКАЛЬНЫЕ ОБЪЕКТЫ, ПАССИВНЫЕ РЫНКИ) УПОВАТЬ НА МЕТОД НЕ СТОИТ.

ОСОБЕННОСТИ ИНДИВИДУАЛЬНОЙ ОЦЕНКИ НЕДВИЖИМОСТИ

- **СТРОЯТСЯ «ПРОСТРАНСТВЕННЫЕ» МОДЕЛИ:** ОПИСЫВАЕТСЯ ВЛИЯНИЕ ТЕРРИТОРИАЛЬНЫХ ФАКТОРОВ, КОТОРЫЕ СТАБИЛЬНЫ ВО ВРЕМЕНИ. ВРЕМЯ ВКЛЮЧАЕТСЯ В МОДЕЛИ, СКОРЕЕ, КАК ИСКЛЮЧЕНИЕ. ЭТО МНОГОЕ УПРОЩАЕТ
- **МОДЕЛИРУЕТСЯ ЦЕНООБРАЗОВАНИЕ НА МАЛЫХ ГЕНЕРАЛЬНЫХ СОВОКУПНОСТЯХ** (ЛОКАЛЬНЫХ ТЕРРИТОРИЯХ «ВОКРУГ» ОБЪЕКТА ОЦЕНКИ) ПРИ ОТНОСИТЕЛЬНО БОЛЬШИХ ДОЛЯХ ОТБОРА АНАЛОГОВ. НЕ «КЛАССИЧЕСКАЯ» СТАТИСТИКА
- **ИМЕЮТСЯ «КАЧЕСТВЕННЫЕ» ФАКТОРЫ, ПОДЛЕЖАЩИЕ УПОРЯДОЧЕНИЮ И ОЦИФРОВКЕ ДЛЯ УЧЕТА В МОДЕЛИ.** НУЖНА ОЦИФРОВКА В ШКАЛЕ ОТНОШЕНИЙ
- **ПРЕОБЛАДАЮТ НЕЛИНЕЙНЫЕ ЗАВИСИМОСТИ** МОДЕЛИРУЕМОЙ ВЕЛИЧИНЫ (СТОИМОСТИ, РЫНОЧНОЙ АРЕНДНОЙ ПЛАТЫ) ОТ ВЛИЯЮЩИХ ФАКТОРОВ. ЛИНЕЙНЫЕ ЗАВИСИМОСТИ «НЕ ОБЪЯСНЯЮТ» РЫНОК
- **ДЛЯ КАЖДОЙ ОЦЕНКИ СТРОИТСЯ СВОЯ МОДЕЛЬ.** «ЗДЕСЬ И СЕЙЧАС»

КАКИЕ АНАЛОГИ НУЖНЫ?

Лишь те, что образуют однородную выборку.

Однородной считается выборка, все элементы которой, включая оцениваемый объект, принадлежат одному сегменту рынка.

Все отобранные объекты сравнения характеризуются единым набором ценообразующих свойств (факторов) и влияние каждого из них на зависимую величину (стоимость, арендную плату) единообразно.

См. п.22б ФСО №7 - определение аналога

Одной моделью можно описать лишь одну зависимость от каждого фактора. Включение в модель объектов из разных сегментов затрудняет либо делает невозможным построение корректной модели.

На практике:

Не аналоги: встроенные офисные vs торговые помещения; ритейл-торговля vs форматы макси-мега-, гипер-; склады на производствах vs логистические комплексы А-класса, etc.

При формировании выборки стремимся отбирать аналоги «хуже» и «лучше» объекта оценки по свойствам. Это уменьшает неопределенность оценки и необходимость введения поправок вне модели.

СКОЛЬКО ФАКТОРОВ УЧИТЫВАТЬ?

В модели должны быть учтены все факторы, по которым различаются объекты сравнения и вариация которых способна объяснить наблюдаемую вариацию цен.

Речь идет о реальных факторах, проявляющихся на рынке

Пропуск (невключение в модель) существенно влияющего фактора – грубая ошибка спецификации модели, приводящая к смещению оценок модели.

Такие модели встречаются у оценщиков, «соблазненных» на простоту

Учет «лишнего» фактора – гораздо меньшее зло, да и не зло вовсе 😊 (смещения оценок нет, растет лишь неопределенность оценки, контролируемая расчетом доверительного интервала).

Значения учитываемых факторов должны быть представлены у всех аналогов (полная матрица свойств).

В жизни часто выполнить это требование сложнее, чем найти нужное число аналогов

На практике:

«Для гарантии» можно оставлять в модели один-другой фактор, не вполне удовлетворяющий статистическим тестам (невысокая значимость фактора м.б. обусловлена малым объемом выборки и/или ненормальностью остатков модели).

СКОЛЬКО НУЖНО АНАЛОГОВ?

Ровно столько, сколько нужно чтобы построить адекватную модель 😊

Чем больше аналогов – тем лучше проявляются связи и меньше неопределенность оценки.
Аналогов много не бывает!

Но! Число аналогов и их близость объекту оценки и друг другу (однородность выборки) в практике оценки, особенно, недвижимости – обменные показатели.

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 ХОТЯ БЫ ОДНУ НА ВСЕХ!**

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

КАК «ЭКОНОМИТЬ» ТРЕБУЕМЫЕ СТЕПЕНИ СВОБОДЫ?

Очевидный путь – отбирать аналоги, «такие же, как объект». Если они, конечно, есть 😊

На практике:

- ИСПОЛЬЗОВАТЬ **ТАБЛИЧНОЕ ПРЕДСТАВЛЕНИЕ** ФУНКЦИЙ ВЛИЯНИЯ ФАКТОРОВ НА МОДЕЛИРУЕМУЮ ВЕЛИЧИНУ ВМЕСТО АНАЛИТИЧЕСКОГО;
- **МИНИМИЗИРОВАТЬ** ПРИМЕНЕНИЕ **БИНАРНЫХ** ПЕРЕМЕННЫХ КАК САМОСТОЯТЕЛЬНЫХ ФАКТОРОВ;
- ПРИМЕНЯТЬ В КАЧЕСТВЕ ЦЕНООБРАЗУЮЩИХ ФАКТОРОВ **«ОБОБЩЕННЫЕ»** ПОКАЗАТЕЛИ (НАПРИМЕР, ПЛОЩАДЬ, ОБЪЕМ) ВМЕСТО НЕСКОЛЬКИХ **«ЭЛЕМЕНТАРНЫХ»** (длина, ширина, высота);
- РАССМАТРИВАТЬ ВОЗМОЖНОСТЬ **ОБЪЕДИНЕНИЯ** НЕСКОЛЬКИХ **ФАКТОРОВ** В ГРАДАЦИИ **ОДНОГО** (НАПРИМЕР, ВМЕСТО «удаленность от трассы + покрытие подъездных путей» - «условия подъезда»);
- ПРИ ОТБОРЕ АНАЛОГОВ СЛЕДИТЬ ЗА ТЕМ, ЧТОБЫ **РОСТ** ЧИСЛА **АНАЛОГОВ** **ОПЕРЕЖАЛ** **РОСТ** ЧИСЛА **ФАКТОРОВ**, ПО КОТОРЫМ ОНИ РАЗЛИЧАЮТСЯ

ВЫБОР ЕДИНИЦЫ СРАВНЕНИЯ

ОБЩАЯ РЕКОМЕНДАЦИЯ – В КАЧЕСТВЕ ЕДИНИЦЫ СРАВНЕНИЯ ВЫБИРАТЬ УДЕЛЬНЫЕ ПОКАЗАТЕЛИ (ЦЕНА /КВ.М, ЦЕНА/КУБ.М, ЦЕНА/МАШИНОМЕСТО, etc).

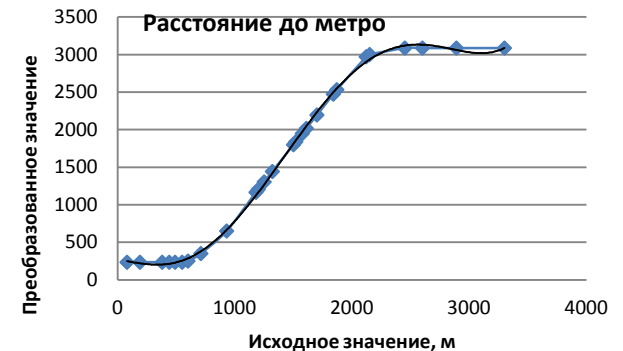
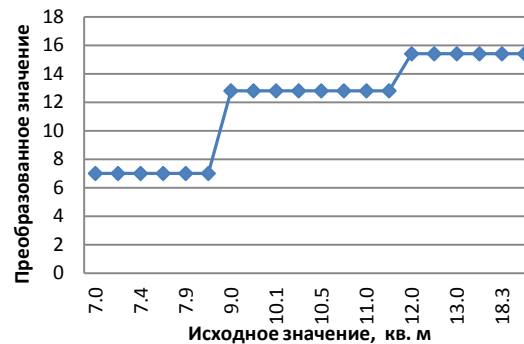
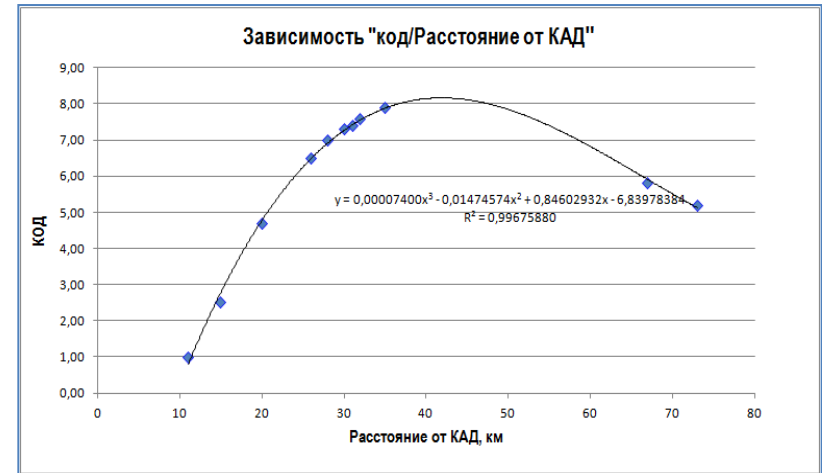
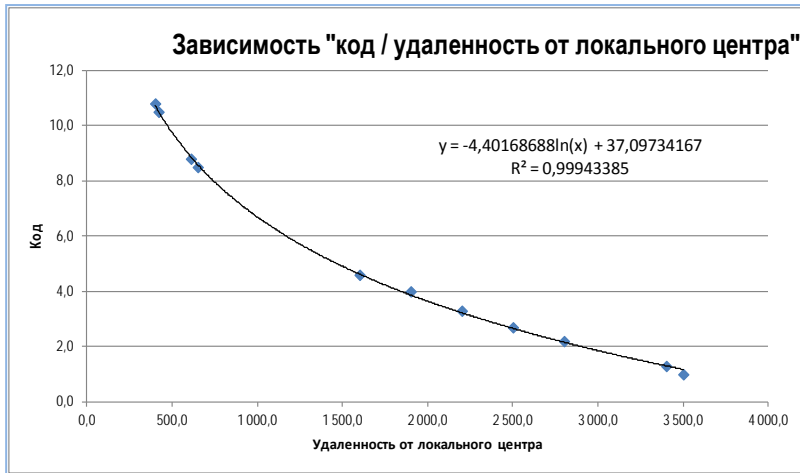
ПОЧЕМУ?

- УЧАСТНИКИ РЫНКА НЕДВИЖИМОСТИ ИСПОЛЬЗУЮТ ИМЕННО ИХ ДЛЯ СРАВНЕНИЯ ОБЪЕКТОВ
- ПЕРЕХОД К УДЕЛЬНЫМ ПОКАЗАТЕЛЯМ ЧАСТО УМЕНЬШАЕТ ГЕТЕРОСКЕДАСТИЧНОСТЬ (НЕРАВНОМЕРНОСТЬ ДИСПЕРСИИ) ОСТАТКОВ МОДЕЛИ
- ПРИ ИСПОЛЬЗОВАНИИ «АБСОЛЮТНЫХ» ЦЕН СИЛЬНАЯ «ПОЛОЖИТЕЛЬНАЯ» ЗАВИСИМОСТЬ ЦЕНЫ ОТ ФАКТОРА МАСШТАБА (ПЛОЩАДИ, ОБЪЕМА, ЧИСЛА МАШИНОМЕСТ, etc) ЗАТРУДНЯЕТ ВЫЯВЛЕНИЕ ЗАВИСИМОСТИ (КАК ПРАВИЛО, «ОТРИЦАТЕЛЬНОЙ») УДЕЛЬНОГО ПОКАЗАТЕЛЯ ОТ ЭТОГО ФАКТОРА МАСШТАБА
- ИНОГДА МОДЕЛЬ В «АБСОЛЮТНЫХ» ЦЕНАХ НЕ СТРОИТСЯ, ЗАВИСИМОСТИ ОТ ФАКТОРОВ НЕ ВИДНЫ, ОДНАКО С ПЕРЕХОДОМ К УДЕЛЬНЫМ ПОКАЗАТЕЛЯМ ОНИ ПРОЯВЛЯЮТСЯ

ЕСТЬ НЮАНСЫ:

ФАКТОР МАСШТАБА, ПО КОТОРОМУ ИСЧИСЛЯЕТСЯ УДЕЛЬНЫЙ ПОКАЗАТЕЛЬ (ПЛОЩАДЬ, ОБЪЕМ, ВМЕСТИМОСТЬ) – ОСОБЫЙ. ПРИ ВЫЯВЛЕНИИ ЗАВИСИМОСТИ ОТ НЕГО НУЖНО УЧИТЫВАТЬ ДОПОЛНИТЕЛЬНЫЕ ОГРАНИЧЕНИЯ – С ЕГО РОСТОМ ЦЕНА ОБЪЕКТА НЕ ДОЛЖНА УМЕНЬШАТЬСЯ

НЕЛИНЕЙНЫЕ СВЯЗИ НА РЫНКЕ



БОЛЬШИНСТВО ЗАВИСИМОСТЕЙ СТОИМОСТИ ОТ ВЛИЯЮЩИХ ФАКТОРОВ НА РЫНКЕ — НЕЛИНЕЙНЫЕ. ЧАСТЬ ИЗ НИХ — НЕМОНОТОННЫЕ.

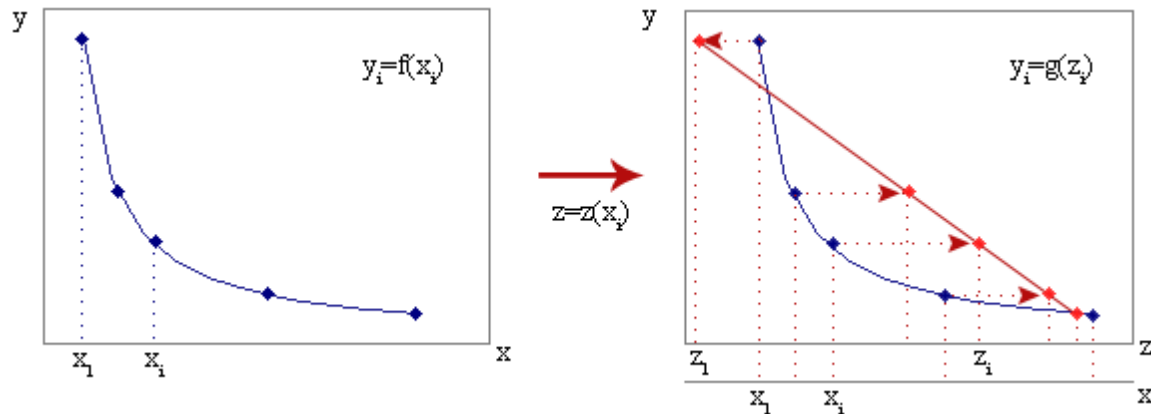
КАК УЧЕСТЬ НЕЛИНЕЙНЫЕ СВЯЗИ?

Рекомендации в литературе – строить **нелинейные** регрессионные модели с **аналитическим представлением зависимостей** (как правило - мультипликативные, приводимые к аддитивному виду логарифмированием и к линейному – заменой переменных).

Недостатки: **однотипность** нелинейности для большинства или всех факторов, сложности подбора вида и параметров аналитических зависимостей.

Альтернатива:

Для зависимостей, заданных дискретно (таблично), **всегда существует преобразование** координат (значений влияющего фактора), **линеаризирующее** связь зависимой переменной модели от фактора.



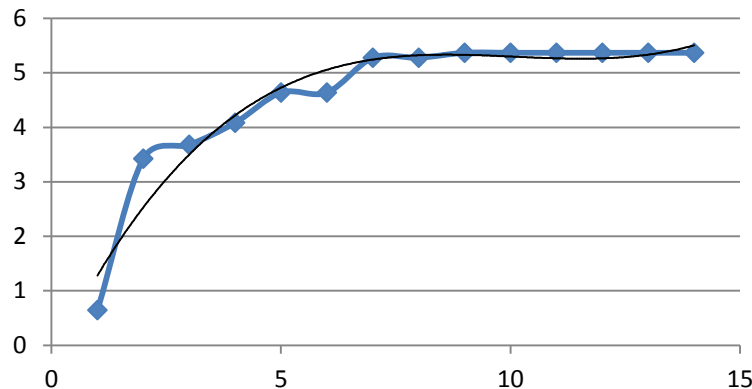
КАК УЧЕСТЬ НЕЛИНЕЙНЫЕ СВЯЗИ?

На практике: Преобразование **каждой** из переменных находится совместным применением инструментов *MS Excel*: функции «ЛИНЕЙН», табличной формой задания функций и макроса «Поиск решения». Максимизируем значение R^2

Монотонность зависимости (или ее отдельных участков) обеспечивается заданием **порядка следования** цифровых **меток** влияющего фактора.

Процесс ускоряется с помощью функции «ВПР»

При этом **контролируется вид** получаемых зависимостей на соответствие экономическим гипотезам, отражающим **ценообразование на рынке**.



«Гладкость» зависимости обеспечивается последующей «ручной доводкой»

Можно преобразовывать последовательно каждую из зависимостей, но **лучше** (и быстрее) – для всех переменных **одновременно**.

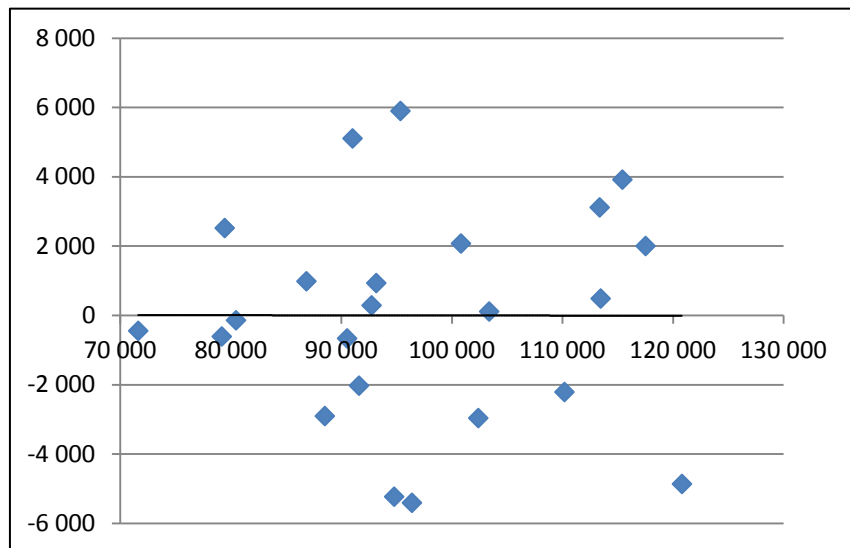
АНАЛИЗ ОСТАТКОВ МОДЕЛИ

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

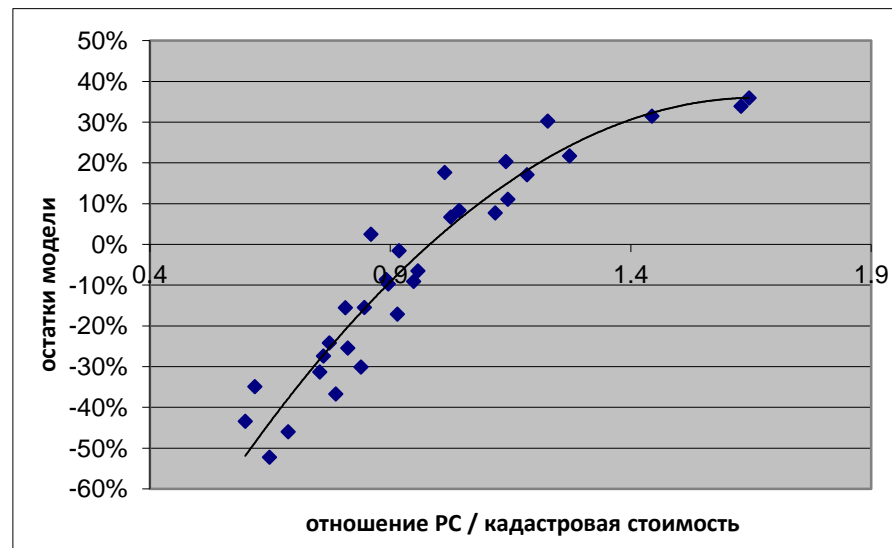
...Остатки, взятые по отдельности, указывают когда, где и в какой степени существующая **модель не смогла объяснить наблюдаемые события**».

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

Случайные остатки адекватной модели



Пропущенный неизвестный фактор



КРИТЕРИИ ОЦЕНКИ АДЕКВАТНОСТИ (КОРРЕКТНОСТИ) МОДЕЛИ

К нашему огорчению, единственного критерия адекватности (корректности) регрессионной модели не существует

Можно говорить о комплексе показателей, обеспечивающих адекватность модели рыночным данным и закономерностям:

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

На практике: Как правило, строятся несколько моделей и на последних этапах лучшая из них отбирается не столько по формальным показателям, сколько по соображениям соответствия рынку.

РАСЧЕТ ДОВЕРИТЕЛЬНОГО ИНТЕРВАЛА ДЛЯ СРЕДНЕГО ЗНАЧЕНИЯ ЦЕНЫ (СТОИМОСТИ)

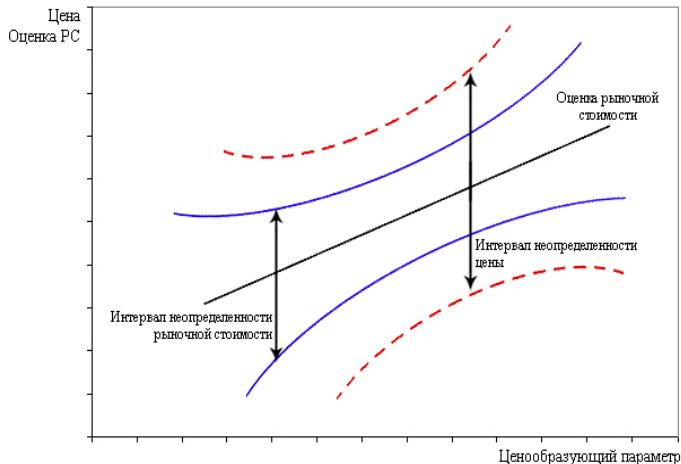
Однофакторная модель

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

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



<http://www.appraiser.ru/default.aspx?SectionId=32&q=posts&t=12164&p=1>

Многофакторная модель

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

Решение уравнения в матричном виде

$$B = (X^T X)^{-1} X^T Y$$

Доверительный интервал оценки **стоимости** в матричном виде

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

Это «классическая» формула для неограниченно больших генсовокупностей и малых долей отбора.

На реальных рынках (при конечных генеральных совокупностях цен и больших долях отбора) это соотношение дает **оценку сверху** для интервала неопределенности РС.

<http://www.appraiser.ru/UserFiles/File/Articles/leyfer/Leyfer-09-2013.pdf>

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 РЕГРЕССИИ

(УСЛОВИЯ ГАУССА-МАРКОВА)

Классический регрессионный анализ базируется на следующих предположениях:

А. На вектор неизвестных параметров (коэффициентов) регрессии не наложено никаких ограничений.

Б. Вектор случайных отклонений (остатков) ε – случайный. Отсюда следует, что и вектор зависимой переменной Y – также случайный.

В. Математическое ожидание отклонений равно нулю $E(\varepsilon_i)=0$. Выполняется автоматически в уравнении регрессии с постоянным членом.

Г. Для любых $i \neq j$ $E(\varepsilon_i * \varepsilon_j)=0$ и $E(\varepsilon_i^2) = \sigma^2$ Т.е. остатки независимы между собой и имеют постоянную дисперсию (гомоскедастичны).

Д. Матрица влияющих переменных X детерминирована, т.е. x_i не являются случайными переменными.

Е. $\text{rank}(X) = m$, где m – число влияющих переменных. Т.е. все влияющие переменные линейно независимы между собой.

Демиденко Е.З. Линейная и нелинейная регрессии, М., ФиС, 1981

«Требование нормальности распределения [зависимой величины] Y необходимо лишь для проверки значимости уравнения и его параметров, а также для интервального оценивания [коэффициентов]. Для получения точечных оценок [коэффициентов] этого условия не требуется».

Дубров А.М. Мхитарян В.С., Трошин Л.И. Многомерные статистические методы: Учебник. М., ФиС, 2003

ПРЕДПОСЫЛКИ РЕГРЕССИИ И ПОСЛЕДСТВИЯ ИХ НАРУШЕНИЙ

Нарушение условия построения модели либо 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 из другого сегмента рынка, 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, обеспечить 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

Увеличение дисперсии оценок. Оценки коэффициентов регрессии и зависимой величины не смещены.

Снижение эффективности оценок. Оценки коэффициентов регрессии смещены, **оценка зависимой величины** (уравнения в целом) **остаётся несмещенной**. Затруднена интерпретация модели в предметной области – оценка вклада каждого из факторов

ПРЕДПОСЫЛКИ РЕГРЕССИИ И ПОСЛЕДСТВИЯ ИХ НАРУШЕНИЙ

Нарушение условия построения модели либо предпосылки регрессии

Значения остатков неслучайны относительно зависимой переменной

Значения остатков зависят от влияющих переменных
(для пространственных моделей – нетипично)

Гетероскедастичность (различная дисперсия) остатков

Отсутствие нормальности распределения остатков модели

Наличие автокорреляции остатков (как правило, для временных рядов). Для пространственных моделей - нетипично

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.
При симметричных генеральных совокупностях **корректные оценки сверху** могут быть получены **в отсутствие нормальности** распределения остатков

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

ВЫВОДЫ

Реальных задач при построении регрессионных моделей для индивидуальной оценки недвижимости - всего 😊 шесть:

- **обеспечение однородности выборки аналогов**
На практике: объекты должны принадлежать одному сегменту (субсегменту) рынка
- **отбор требуемого числа аналогов (числа степеней свободы)**
Как минимум : каждому фактору – по степени свободы и еще одну на всех
- **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 для замены переменных 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*
- **финишный контроль чувствительности модели к изменению уровней факторов**
На практике: позволяет контролировать соответствие полученной модели рынку

РАСПРОСТРАНЕННЫЕ ОШИБКИ ПОСТРОЕНИЯ РЕГРЕССИОННЫХ МОДЕЛЕЙ

➤ на этапе отбора аналогов:

- использование объектов, принадлежащих разным сегментам (субсегментам) рынка;
- построение модели по выборке с объемом, меньшим минимально необходимого

➤ при определении состава ценообразующих факторов:

- отсутствие анализа ценообразующих свойств, по которым различаются объекты сравнения;
- предварительный отсев факторов по результатам построения корреляционной матрицы;
- невключение фактора, указанного в анализе рынка как влияющего, по которому объекты сравнения различаются;
- включение в состав факторов субъективных показателей (развитость соц.структуры, оптимальность местоположения, etc), ранжирование которых не поддается контролю;
- включение в состав факторов модели свойств объектов, влияние которых не подтверждено рынком;
- удаление факторов из модели по статистическим тестам без учета рыночных закономерностей;
- сохранение в модели факторов с алогичными знаками коэффициентов

➤ при оцифровке уровней ценообразующих факторов:

- использование числа градаций, превышающего возможность надежного разнесения по ним характеристик объектов сравнения;
- использование градаций, не соответствующих ценообразованию на рынке;
- изменение порядка следования цифровых меток уровней преобразованных факторов относительно первоначального, заданного согласно экономическим гипотезам

РАСПРОСТРАНЕННЫЕ ОШИБКИ ПОСТРОЕНИЯ РЕГРЕССИОННЫХ МОДЕЛЕЙ

➤ при определении функций влияния факторов:

- игнорирование нелинейности функций, неадекватное использование линейных зависимостей;
- выбор по наибольшему значению R^2 лучшей модели из нескольких простейших нелинейных;
 - произвольный (без обоснований) выбор вида нелинейной функции влияния фактора;
 - использование негладких (и локально немонотонных) связей между зависимой переменной и фактором;
- отсутствие графиков влияния факторов на зависимую переменную (преобразованных значений каждого фактора от исходных);
- изменение первоначального порядка следования цифровых меток при линейаризации модели (учете нелинейности) с помощью оптимизационных процедур

➤ при оценке корректности регрессионной модели:

- использование критерия R^2 (как самостоятельного либо в совокупности с F-критерием) в качестве доказательства адекватности (корректности) полученной модели;
- отсутствие статистических оценок значимости факторов, использованных в модели;
 - отсутствие контроля адекватности знаков коэффициентов регрессии;
- отсутствие графиков влияния факторов на зависимую величину (связи преобразованных и исходных значений факторов);
 - отсутствие оценок ошибок аппроксимации модели;
 - отсутствие анализа распределения остатков модели

ВМЕСТО РЕЗЮМЕ

**«Слова вы услышали, поиск пути – за вами.
Я надеюсь, что услышанное вами сегодня будет
преследовать вас всю оставшуюся жизнь.
В этом случае я сделал все, что мог»**

Эдвард Деминг

Наша тема несопоставима с созданной Э. Демингом новой теорией и практикой менеджмента, способствовавшей японскому производственному чуду второй половины XX века.

И у меня нет желания сниться вам в оставшуюся жизнь 😊

Но если вы поверили в регрессионный анализ как количественный метод анализа рынка и оценки недвижимости - я также сделал все, что мог.

СПАСИБО ЗА ВНИМАНИЕ

Николай Баринов

FRICS, к.т.н., доцент,
директор по научно-методической работе
ООО «Центр оценки «Аверс»

http: www.avg.ru

e-mail: N.Barinov@avg.ru

mobile: +7 921 941 9037