

Многомерное логарифмически нормальное распределение в оценке недвижимого имущества

М.Б. Ласкин 

E-mail: laskinmb@yahoo.com

Санкт-Петербургский институт информатики и автоматизации, Российская академия наук
Адрес: 199178, г. Санкт-Петербург, 14 линия, д. 39

Аннотация

Целью исследования является разработка методики оценки рыночной стоимости по совокупности ценообразующих факторов, подчиняющихся совместному логарифмически нормальному распределению. Под совместным логарифмически нормальным распределением понимается распределение случайного вектора, логарифмы компонент которого распределены совместно нормально. Предложен метод оценки рыночной стоимости по условному распределению цен при заданных значениях ценообразующих факторов, а также методы анализа цены предложения с точки зрения ее обоснованности, по условному распределению вектора ценообразующих факторов при заданной цене предложения. Рассмотрены особенности коэффициента застройки в зависимости от площади земельного участка. Приведены дополнительные аргументы в пользу оценки рыночной стоимости как моды условных законов распределения цен. Рассмотрен пример многомерного логарифмически нормального распределения цены и таких ценообразующих факторов, как площадь улучшений (устоявшийся в оценочном сообществе термин, означающий общую площадь возведенных на земельном участке зданий и сооружений) и площадь земельного участка на реальных данных, т.е. для случая трехмерного случайного вектора. Обоснована формула для определения точки абсолютного максимума плотности многомерного логарифмически нормального случайного вектора. Полученные результаты могут быть использованы при создании информационных систем поддержки принятия решений при оценке объектов недвижимого имущества.

Ключевые слова: рыночная стоимость; логарифмически нормальный закон распределения цен; многомерное логарифмически нормальное распределение; оценка недвижимого имущества.

Цитирование: Ласкин М.Б. Многомерное логарифмически нормальное распределение в оценке недвижимого имущества // Бизнес-информатика. 2020. Т. 14. № 2. С. 48–63. DOI: 10.17323/2587-814X.2020.2.48.63

Введение

Одним из распространенных методов оценки рыночной стоимости объектов недвижимого имущества является построение регрессионных зависимостей цены от значений ценообразующих факторов. При этом факторы могут быть как качественными (тип дома, обременение, этажность, вид из окна, состояние квартиры/помещения и т.д.), так и количественными (площадь объекта или земельного участка, расстояние до центра города, до метро, других объектов инфраструктуры и т.д.).

Существуют различные мнения относительно разбиения ценообразующих факторов на классы. В контексте настоящей статьи имеется в виду разбиение на качественные и количественные факторы с точки зрения возможности представления значений фактора вещественным числом (если такая возможность существует, то фактор является количественным). Совмещение в одной регрессионной модели количественных и качественных факторов представляет для аналитиков определенную трудность. Однако, эта проблема выходит за рамки настоящей статьи: здесь мы ограничимся только количественными (вещественными) факторами. Оказывается, что такие факторы, рассматривая их как случайные величины на множестве объектов сравнения, как правило, удается приблизить логарифмически нормальным распределением. Также имеются основания предполагать, что цены объектов сравнения также подчиняются логарифмически нормальному закону распределения.

Теоретическое обоснование формирования логнормальной генеральной совокупности для цен, образованных последовательными сравнениями, было приведено в работе [1]. На факт подчинения арендных ставок в недвижимости логарифмически нормальному распределению указывали еще Айтчинсон и Браун в 1963 году [2]. На логарифмическое распределение цен в недвижимости указывали и современные исследователи [3]. Такой подход пока не является традиционным с точки зрения существующей практики оценки недвижимого имущества, поскольку требует применения специальных прикладных статистических пакетов, которые не используются практикующими оценщиками, привыкшими работать с малым количеством объектов сравнения. В то же время быстро меняющаяся информационная среда побуждает исследователей искать новые, нетрадиционные подходы к оценке недвижимого имущества. В качестве примера можно привести работы [4–9], посвященные методу гедо-

нистического ценообразования, т.е. выявления статистической связи между средней или медианной стоимостью жилья, внутренними и внешними ценообразующими факторами.

Статистическая зависимость, как правило, оценивается через модели линейной, логарифмической или частично-логарифмической зависимости. В целом эта же идеология положена в основу отчета о кадастровой стоимости [10], выполненного Санкт-Петербургским ГУ «Кадастровая оценка» в 2018 году. В ряде работ используются нерегрессионные модели оценки объектов жилой недвижимости: например, в работах [11, 12] для прогнозирования стоимости жилых объектов используются нейронные сети, в работах [13, 14] – методы машинного обучения (случайный лес, метод опорных векторов), а в работе [15] сравниваются результаты применения таких методов, как деревья решений, наивный байесовский классификатор и AdaBoost. Такие методы требуют использования больших выборок данных. Другим подходом является использование индексов цен. Например, в работе [16] рассматривается индекс цен на жилье Case-Shiller. В работах [17–19] исследуется индекс повторной продажи, который прогнозирует изменение стоимости перепроданного объекта, исходя из разницы во времени и изменения его атрибутов между первоначальной продажей и последующей перепродажей. Авторы работ [20–23] рассматривают гибридный метод, сочетающий гедонистический подход и метод повторной продажи.

Основным подходом при исследовании ценпузырей является использование вариаций методов авторегрессии, примененных к усредненным ценам, например, в работах [24–28]. Таким образом, привлечение многомерных логарифмически нормальных распределений также находится в русле современных тенденций поиска нетрадиционных методов в оценке недвижимого имущества.

1. Оценка рыночной стоимости по условным распределениям при фиксированных значениях ценообразующих (количественных/вещественных) факторов

Пусть V – цена предложения (или сделки), а X_1, \dots, X_n – количественные (вещественные) ценообразующие факторы. Пусть $W = \ln(V)$, $Y_i = \ln(X_i)$, $i = 1, n$ (тогда $v = e^W$, $X_i = e^{Y_i}$).

Рассмотрим многомерный нормальный случайный вектор (W, Y_1, \dots, Y_n) с вектором средних

$(\mu_w, \mu_{Y_1}, \dots, \mu_{Y_n})$. Ковариационную матрицу запишем в блочном виде:

$$COV = \begin{pmatrix} \sigma_w^2 & cov(W, \vec{Y}) \\ cov(W, \vec{Y})^T & COV \end{pmatrix},$$

где COV – ковариационная матрица случайного вектора $\vec{Y} = (Y_1, \dots, Y_n)$;

$cov(W, \vec{Y})$ – вектор $(\rho_{wY_1} \sigma_w \sigma_{Y_1}, \dots, \rho_{wY_n} \sigma_w \sigma_{Y_n})$;

$\sigma_w^2, \sigma_{Y_1}^2, \dots, \sigma_{Y_n}^2$ – дисперсии случайных величин W, Y_1, \dots, Y_n ;

$\rho_{wY_1}, \dots, \rho_{wY_n}$ – соответствующие коэффициенты корреляции.

Тогда, условное математическое ожидание W , при условии $Y_1 = y_1, \dots, Y_n = y_n$ равно

$$E(W|Y_1 = y_1, \dots, Y_n = y_n) = \mu_w + \left(COV^{-1} \times cov(W, \vec{Y})^T, (\vec{Y} - \overline{\mu_Y}) \right),$$

где $\overline{\mu_Y} = (\mu_{Y_1}, \dots, \mu_{Y_n})$. Условная дисперсия W , при условии, что $Y_1 = y_1, \dots, Y_n = y_n$ равна

$$D(W|Y_1 = y_1, \dots, Y_n = y_n) = \sigma_w^2 - \left(COV^{-1} \times cov(W, \vec{Y})^T, cov(W, \vec{Y}) \right).$$

При фиксированных значениях ценообразующих факторов $X_1 = x_1, \dots, X_n = x_n$ наиболее вероятное значение цены предложения (или сделки, в зависимости от того, какие цены были в исходных данных) V рассчитывается по формуле условной моды

$$Mode(V|X_1 = x_1 = e^{y_1}, \dots, X_n = x_n = e^{y_n}) = \exp\left(\mu_w + (COV^{-1} \times cov(W, \vec{Y})^T, (\vec{Y} - \overline{\mu_Y}) - \sigma_w^2 + (COV^{-1} \times cov(W, \vec{Y})^T, cov(W, \vec{Y}))\right). \quad (1)$$

В соответствии с формулировкой ФЗ-135 [29], под рыночной стоимостью понимается наиболее вероятная цена, по которой объект оценки может быть отчужден на открытом рынке в условиях совершенной конкуренции. На практике оценщики, как правило, используют средние арифметические, реже средние геометрические значения. Такие оценки могут быть построены как оценки условного математического ожидания и условной медианы:

$$E(V|X_1 = x_1 = \exp(y_1), \dots, X_n = x_n = \exp(y_n)) = \exp(\mu_w + (COV^{-1} \times cov(W, \vec{Y})^T, (\vec{Y} - \overline{\mu_Y}))) + \frac{1}{2} \sigma_w^2 - \frac{1}{2} \left(COV^{-1} \times cov(W, \vec{Y})^T, cov(W, \vec{Y}) \right). \quad (2)$$

$$Median(V|X_1 = x_1 = \exp(y_1), \dots, X_n = x_n = \exp(y_n)) = \exp(\mu_w + COV^{-1} \times cov(W, \vec{Y})^T, (\vec{Y} - \overline{\mu_Y})). \quad (3)$$

Таким образом, если относительно некоторого ансамбля количественных ценообразующих факторов и цен объектов сравнения удастся принять в качестве рабочей гипотезу о совместном логарифмически нормальном распределении (совместном нормальном распределении логарифмов) компонент случайного вектора, то в качестве оценки рыночной стоимости могут быть приняты оценки по формуле (1). Могут быть приняты и оценки по формулам (2) и (3); но при этом следует помнить, что такие оценки не отвечают определению рыночной стоимости, закрепленной в ФЗ-135.

Рассмотрим пример, в котором использованы данные, подобранные известными российскими оценщиками и опубликованные на ресурсе [30]. Рассматривается 40 объектов недвижимости производственно-складского назначения с локацией в одном регионе (г. Санкт-Петербург), выставленные на продажу в одном периоде времени. Поскольку авторами примера обоснован отказ от ряда корректировок, в нашем примере мы также будем считать данные сопоставимыми и сравнимыми без дополнительных корректировок. Объекты производственно-складского назначения рассматриваются как единый комплекс, состоящий из земельного участка и улучшений (зданий). Данные представлены в *таблице 1*.

Объекты сравнения рассматриваются как действующие объекты производственно-складского назначения, выставленные на продажу в текущем использовании. Мы построим общий метод оценки рыночной стоимости (без учета скидки на торг), если для него заданы площади улучшений и земельного участка (полученная оценка, естественно, относится к тому же периоду времени, типу объектов и региону).

В данном случае имеют место случайные величины V – приведенная к 1 кв. м улучшений цена предложения, SB – площадь улучшений, SP – площадь земельного участка. Они образуют трехмерный случайный вектор (V, SB, SP) . Пусть $W = \ln(V)$, $Y = \ln(SB)$, $Z = \ln(SP)$, (тогда $v = e^W$, $SB = e^Y$, $SP = e^Z$). Для трехмерного нормального случайного вектора (W, Y, Z) вектор средних равен $(\mu_w, \mu_y, \dots, \mu_z)$. Ковариационная матрица выглядит следующим образом:

Таблица 1.

Исходные данные

Площадь зданий (кв. м)	Площадь земельных участков (кв. м)	Цена за весь объект (руб.)	Отношение цены к площади улучшений (руб./кв. м)
400	2 500	20 500 000	51 250
750	5 000	18 000 000	24 000
1 081	3 378	26 000 000	24 052
1 130	6 638	27 500 000	24 336
1 320	4 167	31 500 000	23 864
1 440	10 000	160 000 000	111 111
1 790	3 462	93 000 000	51 955
1 900	13 000	85 000 000	44 737
2 125	5 623	85 000 000	40 000
2 642	5 183	75 000 000	28 388
2 700	6 800	59 000 000	21 852
1 820	2 737	32 000 000	17 582
2 250	9 252	84 000 000	37 333
2 973	5 388	90 000 000	30 272
3 513	10 000	80 000 000	22 773
3 600	5 000	95 000 000	26 389
4 000	13 558	140 000 000	35 000
4 124	12 866	91 000 000	22 066
4 167	5 000	125 000 000	29 998
4 257	6 861	128 500 000	30 186
5 292	11 143	56 000 000	10 582
5 300	16 000	220 000 000	41 509
6 011	11 319	135 000 000	22 459
6 013	20 781	90 000 000	14 968
6 060	21 790	179 000 000	29 538
6 123	2 390	152 490 000	24 904
6 479	7 337	119 000 000	18 367
6 756	4 220	90 000 000	13 321
10 000	12 000	420 000 000	42 000
10 300	17 000	312 000 000	30 291
10 672	12 194	350 000 000	32 796
10 990	30 000	480 000 000	43 676
12 000	30 000	300 000 000	25 000
13 000	55 000	200 000 000	15 385
14 428	33 000	385 000 000	26 684
15 000	37 000	840 000 000	56 000
18 924	20 600	800 000 000	42 274
22 312	40 162	338 541 000	15 173
34 082	478 000	2 500 000 000	73 353
35 000	160 000	2 400 000 000	68 571

$$CV = \begin{pmatrix} \sigma_W^2 & \rho_{WY}\sigma_W\sigma_Y & \rho_{WZ}\sigma_W\sigma_Z \\ \rho_{YW}\sigma_W\sigma_Y & \sigma_Y^2 & \rho_{YZ}\sigma_Y\sigma_Z \\ \rho_{ZW}\sigma_W\sigma_Z & \rho_{ZY}\sigma_Y\sigma_Z & \sigma_Z^2 \end{pmatrix},$$

или в блочном виде:

$$CV = \begin{pmatrix} \sigma_W^2 & cov(W, \bar{Y}) \\ cov(W, \bar{Y})^T & COV \end{pmatrix},$$

где $COV = \begin{pmatrix} \sigma_Y^2 & \rho_{YZ}\sigma_Y\sigma_Z \\ \rho_{ZY}\sigma_Y\sigma_Z & \sigma_Z^2 \end{pmatrix}$; $\bar{Y} = (Y, Z)$;

$$cov(W, \bar{Y}) = (\rho_{WY}\sigma_W\sigma_Y, \rho_{WZ}\sigma_W\sigma_Z);$$

$\sigma_W^2, \sigma_Y^2, \sigma_Z^2$ – дисперсии случайных величин W, Y, Z ;

$\rho_{WY} = \rho_{YW}, \rho_{WZ} = \rho_{ZW}, \rho_{YZ} = \rho_{ZY}$ – соответствующие коэффициенты корреляции.

Условное математическое ожидание W , при условии, что $Y = y, Z = z$:

$$E(W|Y = y, Z = z) = \mu_W + \left(COV^{-1} \times cov(X, \bar{Y})^T, (y - \mu_Y, z - \mu_Z) \right).$$

Условная дисперсия W , при условии, что $Y = y, Z = z$:

$$D(W|Y = y, Z = z) = \sigma_W^2 - \left(COV^{-1} \times cov(X, \bar{Y})^T, cov(X, \bar{Y}) \right).$$

Пусть заданы значения площади улучшений $SB = sb$ и площади земельного участка $SP = sp$. В соответствии с введенными выше обозначениями $Y = \ln(SB), Z = \ln(SP), y = \ln(sb), z = \ln(sp)$. Наиболее вероятное значение цены предложения V при известных значениях площади улучшений и площади земельного участка рассчитывается по формуле [31]:

$$Mode(V|SB = sb, SP = sp) = \exp(\mu_W + (COV^{-1} \times cov(X, \bar{Y})^T, (y - \mu_Y, z - \mu_Z)) - \sigma_W^2 + (COV^{-1} \times cov(W, \bar{Y})^T, cov(W, \bar{Y}))). \quad (4)$$

Условное математическое ожидание рассчитывается по формуле:

$$E(V|SB = sb, SP = E) = \mu_W + (COV^{-1} \times cov(X, \bar{Y})^T, (y - \mu_Y, z - \mu_Z)) + \frac{1}{2} \sigma_W^2 - \frac{1}{2} (COV^{-1} \times cov(W, \bar{Y})^T, cov(W, \bar{Y})). \quad (5)$$

Условная медиана рассчитывается по формуле:

$$\text{Median}(V|SB = sb, SP = sp) = \exp(\mu_w + (COV^{-1} \times cov(W, \bar{Y}), (y - \mu_y, z - \mu_z))). \quad (6)$$

Прежде чем применить формулы (4)–(6) к данным *таблицы 1* проверим, есть ли основания предполагать логнормальность распределений компонент случайного вектора (V, SB, SP) (совместную нормальность логарифмов их компонент). Для проверки маргинальных распределений использовался параметрический тест Колмогорова–Смирнова. Получены следующие значения p -value:

V – цена 1 кв. м улучшений: при параметрах $\text{meanlog} = 10,3$ и $\text{sdlog} = 0,43$ p -value составляет 0,7016;

SB – площадь зданий: при параметрах $\text{meanlog} = 8,45$ и $\text{sdlog} = 1,02$ p -value равно 0,9761;

SP – площадь земельного участка: при параметрах $\text{meanlog} = 9,3$ и $\text{sdlog} = 1,01$ p -value составляет 0,8963.

Проверим три изучаемые случайные величины на совместную нормальность логарифмов. Для этого воспользуемся известным условием совместной нормальности: для того, чтобы многомерный случайный вектор имел многомерное нормальное распределение необходимо и достаточно, чтобы любая линейная комбинация его компонент была распределена нормально. В среде статистического пакета R была реализована следующая процедура:

- ◆ переменные прологарифмированы;
- ◆ полученные логарифмические значения переменных центрированы и нормированы, каждая на свое стандартное отклонение;

◆ с помощью стандартной функции R $\text{runif}(3,0,1)$ сгенерированы три весовых коэффициента, коэффициенты нормированы на их сумму;

◆ из центрированных и нормированных логарифмов составлена линейная комбинация со случайными положительными коэффициентами, в сумме равными единице;

◆ полученная линейная комбинация тестируется при помощи теста Колмогорова–Смирнова на нормальность, результат теста в виде значения p -value записывается в массив;

◆ процедура со случайной линейной комбинацией повторяется заданное количество раз, каждый раз записывается p -value. Затем общий массив p -value сравнивается с критическим уровнем (0,05).

На приведенном *рисунке 1* представлена гистограмма, показывающая, какие значения p -value были получены при повторении теста 100 000 раз.

Минимальное значение p -value равно 0,2867691, что выше критического уровня 0,05. Приведенная процедура 100 000-кратного повторения тестов для случайных линейных комбинаций компонент вектора (W, Y, Z) представляется достаточным основанием для того, чтобы сохранить в качестве рабочей гипотезу о совместной нормальности логарифмов изучаемых переменных.

Для логарифмов переменных «Отношение цены к площади улучшений», «Площадь зданий», «Площадь земельного участка», указанных в *таблице 1*, получены следующие значения вектора средних и ковариационная матрица (*таблица 2*).

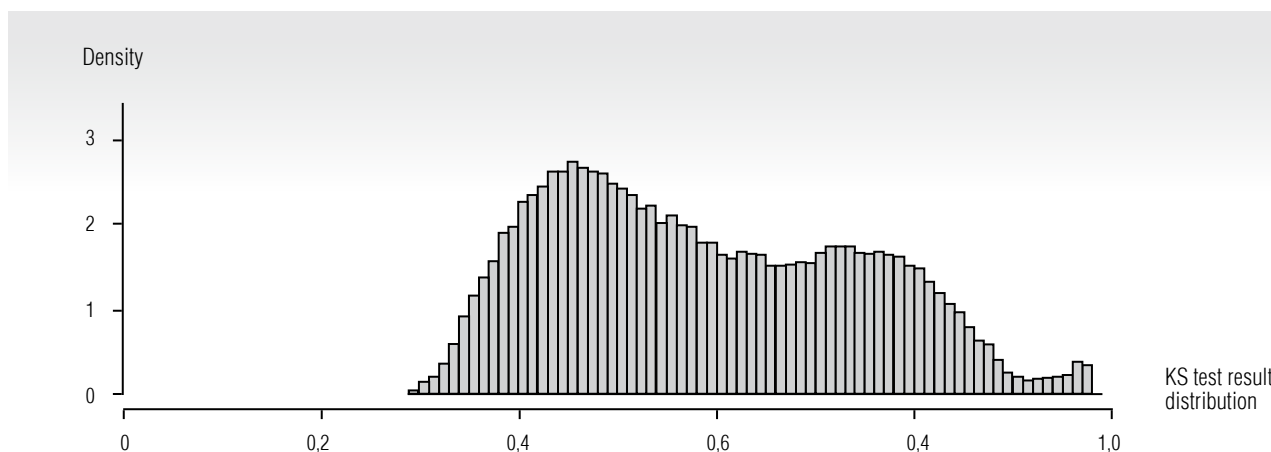


Рис. 1. Результаты тестирования случайных линейных комбинаций центрированных и нормированных компонент вектора $(W, Y, Z) = (\ln(V), \ln(SB), \ln(SP))$ на совместную нормальность тестом Колмогорова–Смирнова

Таблица 2.
Средние логарифмов переменных

Отношение цены к площади улучшений (V)	Площадь зданий (SB)	Площадь земельного участка (SP)
Вектор средних		
10,2993	8,4469	9,3506
Ковариационная матрица		
0,2381	0,0108	0,1467
0,0108	1,0635	0,8978
0,1467	0,8978	1,2140

Таким образом

$$\begin{aligned} \mu_W &= 10,2993; \mu_Y = 8,4469; \mu_Z = 9,3506, \sigma_W^2 = 0,2381, \\ \rho_{WY} \sigma_W \sigma_Y &= \rho_{YW} \sigma_W \sigma_Y = 0,0108; \\ \rho_{WZ} \sigma_W \sigma_Z &= \rho_{ZW} \sigma_W \sigma_Z = 0,1467; \sigma_Y^2 = 1,0635, \\ \rho_{YZ} \sigma_Y \sigma_Z &= \rho_{ZY} \sigma_Y \sigma_Z = 0,8978; \sigma_Z^2 = 1,2140. \end{aligned}$$

В среде статистического пакета R реализован программный код, позволяющий рассчитать оценку рыночной стоимости по заданным значениям параметров «Площадь зданий», «Площадь земельного участка» (формула (4)). Аналогичные расчеты могут быть выполнены для оценок по медианным значениям или по математическим ожиданиям (формулы (5) и (6)). Результаты представлены в таблице 3.

Из представленной таблицы видно следующее:

- ♦ оценка по моде всегда ниже оценки по медиане, а оценка по медиане всегда ниже оценки по математическому ожиданию (мнение автора: рыночную стоимость следовало бы определять как оценку по моде, в соответствии с формулировкой ФЗ-135, учитывая наблюдаемые на рынке несимметричные распределения цен, площадей и расстояний);

- ♦ при постоянной площади земельного участка рыночная стоимость в пересчете на 1 кв. м улучшений уменьшается при увеличении площади улучшений;

- ♦ при постоянной площади улучшений рыночная стоимость в пересчете на 1 кв. м улучшений увеличивается при увеличении площади земельного участка;

- ♦ по формуле (4) может быть рассчитана рыночная стоимость объекта недвижимости того же класса, что и объекты сравнения, для любых значений площадей улучшений и земельного участка (на ту же дату, для той же локации). Поскольку в оценочной среде нет общего единства мнений от-

носительно используемых для оценки рыночной стоимости числовых характеристик (мода, медиана, математическое ожидание), могут быть применены формулы (5) и (6), строго говоря, не отвечающие определению рыночной стоимости в соответствии с ФЗ-135.

2. Соотношение площади улучшений и площади земельного участка при известной цене предложения

В статье [30] рассматривается, в том числе, вопрос о ценовых тенденциях на рынке недвижимости производственно-складского назначения и зависимость рыночной стоимости от «коэффициента плотности застройки» земельного участка, под которым понимается отношение площади улучшений к площади всего земельного участка. Рассматриваемая в настоящей статье модель совместного логарифмически нормального распределения компонент случайного вектора (V, SB, SP) также позволяет посмотреть на проблему формирования ценовых тенденций. Отличие заключается в том, что все компоненты случайного вектора (V, SB, SP) распределены на положительной полуоси, т.е. для каждого заданного значения $V = v$ можно указать наиболее вероятные значения компонент SB (площади улучшений) и SP (площади земельного участка), соответствующие цене предложения. В отличие от предыдущего случая (оценка рыночной стоимости по заданным значениям SB и SP), область возможных отклонений от наиболее вероятного (медианного, среднего) значений находится не на числовой оси, а на плоскости и представляет из себя (как будет показано ниже) вложенные множества, получающиеся из эллипсов рассеяния логарифмированных значений SB и SP при обратном экспоненциальном преобразовании плоскости.

Пусть цена предложения $V = v$ известна. Требуется оценить соотношение площади улучшений и земельного участка для класса объектов с такой начальной ценой предложения, т.е. выделить объекты с нижней, средней и верхней ценовыми тенденциями [30]. Обозначения прежние: V – цена предложения, SB – площадь улучшений, SP – площадь земельного участка, $W = \ln(V)$, $Y = \ln(SB)$, $Z = \ln(SP)$ (т.е. $V = e^W$, $SB = e^Y$, $SP = e^Z$).

По прежнему, рассматриваем трехмерный нормальный случайный вектор (W, Y, Z) с вектором средних (μ_W, μ_Y, μ_Z) и ковариационной матрицей

Таблица 3.

Оценки рыночной стоимости, приведенной к 1 кв. м улучшений для различных значений площадей улучшений и земельных участков

Оценка РС по моде		Площадь земельного участка в метрах									
		2 000	7 000	12 000	17 000	22 000	27 000	32 000	37 000	42 000	47 000
Площадь улучшений в метрах	400	26 247	38 298	45 058	50 049	54 096	57 543	60 568	63 279	65 745	68 014
	2 400	16 938	24 714	29 076	32 297	34 909	37 133	39 085	40 835	42 426	43 890
	4 400	14 605	21 310	25 072	27 849	30 101	32 019	33 702	35 211	36 583	37 845
	6 400	13 327	19 445	22 877	25 411	27 466	29 216	30 752	32 129	33 381	34 533
	8 400	12 469	18 194	21 406	23 777	25 700	27 337	28 775	30 062	31 234	32 312
	10 400	11 835	17 269	20 317	22 567	24 392	25 947	27 311	28 533	29 645	30 668
	12 400	11 337	16 542	19 462	21 618	23 366	24 855	26 161	27 332	28 397	29 377
	14 400	10 930	15 948	18 763	20 842	22 527	23 962	25 222	26 351	27 378	28 323
	16 400	10 588	15 449	18 176	20 189	21 822	23 212	24 433	25 527	26 521	27 436
	18 400	10 294	15 021	17 672	19 629	21 217	22 569	23 755	24 818	25 786	26 675
Оценка РС по медиане		Площадь земельного участка в метрах									
		2 000	7 000	12 000	17 000	22 000	27 000	32 000	37 000	42 000	47 000
Площадь улучшений в метрах	400	31 947	46 615	54 843	60 918	65 844	70 039	73 722	77 021	80 023	82 784
	2 400	20 616	30 081	35 391	39 311	42 490	45 197	47 573	49 703	51 640	53 421
	4 400	17 777	25 938	30 517	33 897	36 638	38 972	41 021	42 858	44 528	46 064
	6 400	16 221	23 668	27 846	30 930	33 431	35 561	37 431	39 106	40 630	42 032
	8 400	15 177	22 146	26 055	28 941	31 281	33 274	35 023	36 591	38 017	39 329
	10 400	14 405	21 019	24 729	27 468	29 690	31 581	33 242	34 730	36 083	37 328
	12 400	13 799	20 134	23 688	26 312	28 440	30 252	31 843	33 268	34 564	35 757
	14 400	13 304	19 412	22 838	25 368	27 419	29 166	30 700	32 074	33 324	34 473
	16 400	12 887	18 804	22 123	24 574	26 561	28 253	29 739	31 070	32 281	33 395
	18 400	12 530	18 283	21 510	23 892	25 824	27 470	28 914	30 208	31 385	32 468
Оценка РС по математическому ожиданию		Площадь земельного участка в метрах									
		2 000	7 000	12 000	17 000	22 000	27 000	32 000	37 000	42 000	47 000
Площадь улучшений в метрах	400	35 246	51 428	60 506	67 208	72 643	77 271	81 334	84 974	88 285	91 332
	2 400	22 744	33 187	39 045	43 370	46 877	49 864	52 485	54 835	56 971	58 937
	4 400	19 612	28 616	33 668	37 397	40 421	42 996	45 257	47 283	49 125	50 820
	6 400	17 895	26 112	30 721	34 124	36 883	39 233	41 296	43 144	44 825	46 372
	8 400	16 744	24 432	28 745	31 929	34 511	36 710	38 640	40 369	41 942	43 389
	10 400	15 893	23 189	27 283	30 305	32 755	34 842	36 674	38 315	39 809	41 182
	12 400	15 224	22 213	26 134	29 029	31 377	33 376	35 130	36 703	38 133	39 449
	14 400	14 677	21 416	25 196	27 987	30 250	32 178	33 869	35 385	36 764	38 033
	16 400	14 218	20 746	24 408	27 111	29 304	31 171	32 810	34 278	35 614	36 843
	18 400	13 824	20 170	23 731	26 359	28 491	30 306	31 899	33 327	34 626	35 821

$$CV = \begin{pmatrix} \sigma_W^2 & \rho_{WY}\sigma_W\sigma_Y & \rho_{WZ}\sigma_W\sigma_Z \\ \rho_{YW}\sigma_W\sigma_Y & \sigma_Y^2 & \rho_{YZ}\sigma_Y\sigma_Z \\ \rho_{ZW}\sigma_W\sigma_Z & \rho_{ZY}\sigma_Y\sigma_Z & \sigma_Z^2 \end{pmatrix},$$

или в блочном виде:

$$CV = \begin{pmatrix} \sigma_W^2 & cov(W, \bar{Y}) \\ cov(W, \bar{Y})^T & COV \end{pmatrix},$$

где $COV = \begin{pmatrix} \sigma_Y^2 & \rho_{YZ}\sigma_Y\sigma_Z \\ \rho_{ZY}\sigma_Y\sigma_Z & \sigma_Z^2 \end{pmatrix}; \bar{Y} = (Y, Z);$

$cov(W, \bar{Y}) = (\rho_{WY}\sigma_W\sigma_Y, \rho_{WZ}\sigma_W\sigma_Z);$

$\sigma_W^2, \sigma_Y^2, \sigma_Z^2$ – дисперсии случайных величин $W, Y, Z;$

$\rho_{WY} = \rho_{YW}, \rho_{WZ} = \rho_{ZW}, \rho_{YZ} = \rho_{ZY}$ – соответствующие коэффициенты корреляции.

Условное математическое ожидание вектора $\bar{Y} = (Y, Z)$, при условии, что $W = w:$

$$E(\bar{Y}|W = w) = \bar{\mu} + \frac{cov(W, \bar{Y})^T}{\sigma_W^2} (w - \mu_W) = \begin{pmatrix} \mu_Y \\ \mu_Z \end{pmatrix} + \begin{pmatrix} \rho_{YW} \frac{\sigma_Y}{\sigma_W} (w - \mu_W) \\ \rho_{ZW} \frac{\sigma_Z}{\sigma_W} (w - \mu_W) \end{pmatrix} = \begin{pmatrix} \mu_Y + \rho_{YW} \frac{\sigma_Y}{\sigma_W} (w - \mu_W) \\ \mu_Z + \rho_{ZW} \frac{\sigma_Z}{\sigma_W} (w - \mu_W) \end{pmatrix} \quad (7)$$

Условная ковариационная матрица при условии, что $W = w:$

$$COV(\bar{Y}|W = w) = COV - \frac{cov(W\bar{Y})^T \times cov(W, \bar{Y})}{\sigma_W^2} = \begin{pmatrix} \sigma_Y^2 & \rho_{YZ}\sigma_Y\sigma_Z \\ \rho_{ZY}\sigma_Y\sigma_Z & \sigma_Z^2 \end{pmatrix} - \begin{pmatrix} \rho_{YW}^2\sigma_Y^2 & \rho_{YW}\rho_{ZW}\sigma_Y\sigma_Z \\ \rho_{YW}\rho_{ZW}\sigma_Y\sigma_Z & \rho_{ZW}^2\sigma_Z^2 \end{pmatrix} = \begin{pmatrix} \sigma_Y^2(1 - \rho_{YW}^2) & \sigma_Y\sigma_Z(\rho_{YZ} - \rho_{YW}\rho_{ZW}) \\ \sigma_Y\sigma_Z(\rho_{YZ} - \rho_{YW}\rho_{ZW}) & \sigma_Z^2(1 - \rho_{ZW}^2) \end{pmatrix} \quad (8)$$

Пусть известна цена предложения $V = v$. В соответствии с введенными выше обозначениями $W = \ln(V), w = \ln(v)$. Наиболее вероятная комбинация площади улучшений SB и площади земельного участка SP при фиксированной цене предложения рассчитывается по формуле:

$$Mode(\bar{Y}|V = v) = \exp \left(\bar{\mu} + \frac{cov(W, \bar{Y})^T}{\sigma_W^2} (w - \mu_W) - COV + \frac{cov(W\bar{Y})^T \times cov(W, \bar{Y})}{\sigma_W^2} \right).$$

В таблице 4 представлены расчетные наиболее вероятные комбинации площади улучшений SB и площади земельного участка SP для некоторых значений цены предложения.

Следует отметить, что при каждом значении цены предложения такая наиболее вероятная пара значе-

Таблица 4.

Наиболее вероятные комбинации площади улучшений SB и площади земельного участка SP , значения коэффициента плотности застройки земельного участка для некоторых значений цены предложения

Цена предложения в руб. за 1 кв.м. приведенной площади	7 000	12 000	21 000	28 000	40 000	60 000	80 000	100 000
Наиболее вероятная пара:								
площадь улучшений в кв.м.	619	634	650	659	669	682	691	698
площадь земельного участка в кв.м.	630	878	239	1 479	1 843	2 365	2 824	3 240
коэффициент плотности застройки	0,98	0,72	0,52	0,45	0,36	0,29	0,24	0,22

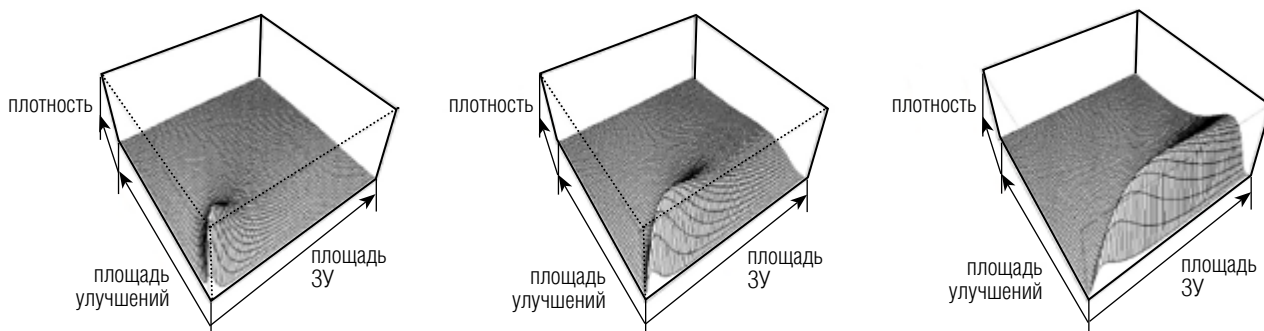


Рис. 2. Двумерные распределения SP (площадь земельного участка) и SB (площадь улучшений)

ний SB , SP является единственной (коэффициент плотности застройки в этом случае отвечает наиболее вероятной паре SB , SP). Попытки представить как наиболее удобную, с какой-либо точки зрения, другую пару компонент SB и SP означают выбор точки, для которой существует множество других равновероятных точек, с плотностью меньше максимальной, и для которых коэффициенты плотности застройки, очевидно, будут разными. На рисунке 2 представлены изображения двумерных распределений SB и SP для цены предложения 7 000 руб., 28 000 руб., 100 000 руб. (слева направо) за 1 кв. м. площади, приведенной к площади улучшений.

Хорошо видно, что любая другая точка плоскости SP , SB имеет множество равновероятных значений. Области таких значений при цене предложения 28 000 руб. за 1 кв. м улучшений единого комплекса (земельный участок плюс улучшения) представлены на рисунке 3.

3. Коэффициент плотности застройки (отношение площади улучшений к площади земельного участка)

Предположим, что цена предложения (сделки) известна. Требуется оценить, каким должен быть коэффициент плотности застройки при данной цене и известной площади земельного участка. Обратимся к формулам (7) и (8). При фиксированной цене предложения ($V = v$) формулы (7) и (8) дают расчетные значения условных математических ожиданий логарифмов площади улучшений ($SB|V = v$), площади земельного участка ($SP|V = v$) и условной ковариационной матрицы. Теперь предположим, что площадь земельного участка тоже известна. Введем новые обозначения для условных логарифмов площади улучшений ($SB|V = v$) и площади земельного участка ($SP|V = v$):

$$\mu_{condSB} = \mu_Y + \rho_{YW} \frac{\sigma_Y}{\sigma_W} (w - \mu_W),$$

$$\mu_{condSP} = \mu_Z + \rho_{ZW} \frac{\sigma_Z}{\sigma_W} (w - \mu_W),$$

$$\sigma_{condSB}^2 = \sigma_Y^2 (1 - \rho_{YW}^2),$$

$$\sigma_{condSP}^2 = \sigma_Z^2 (1 - \rho_{ZW}^2),$$

$$\rho = \sigma_Y \sigma_Z (\rho_{YZ} - \rho_{YW} \rho_{ZW})$$

(части нижних индексов “cond” означают «условное»). Рассмотрим двумерный случайный вектор ($SB|V = v$, $SP|V = v$) с указанными параметрами. При заданном значении площади земельного участка и заданном значении цены (в приведенном примере цены предложения) ($SP = sp$, $V = v$), условная мода величины SB (площади улучшений) равна (по аналогии с доказательством, приведенным в работе [32]):

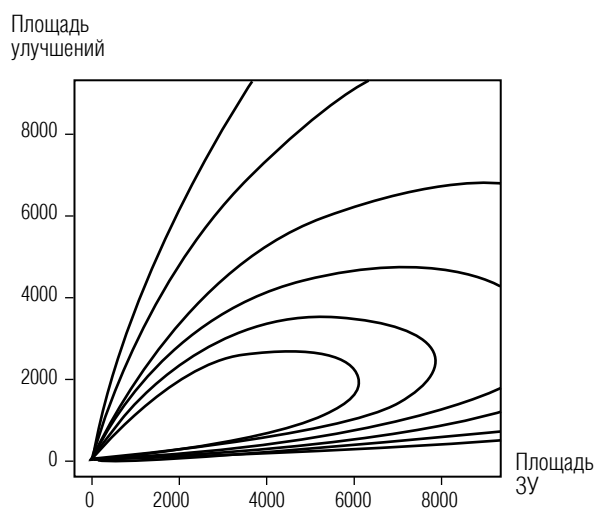


Рис. 3. Линии уровней равновероятных точек на плоскости с координатами SP (площадь земельного участка), SB (площадь улучшений)

$$Mode(SP|SP = sp, V = v) = \exp(\mu_{condSP} + \rho \times \frac{\sigma_{condSP}^2}{\sigma_{condSB}^2} (\ln(sb - \mu_{condSB})) - \sigma_{condSP}^2 (1 - \rho^2)). \quad (9)$$

Условная медиана величины SP равна:

$$Median(SP|SP = sp, V = v) = \exp(\mu_{condSP} + \rho \times \frac{\sigma_{condSP}^2}{\sigma_{condSB}^2} (\ln(sb - \mu_{condSB})).$$

Условное математическое ожидание величины SP равно:

$$E(SP|SP = sp, V = v) = \exp(\mu_{condSP} + \rho \times \frac{\sigma_{condSP}^2}{\sigma_{condSB}^2} (\ln(sb - \mu_{condSB})) + \frac{1}{2} \sigma_{condSP}^2 (1 - \rho^2)).$$

Предположим, что требуется оценить коэффициент застройки в группе объектов нижней, средней или верхней ценовой категории. Такие оценки могут быть построены в зависимости от площади земельного участка как по модальным, так и по медианным или средним значениям. Однако общий вид поверхностей, показанных на *рисунке 2*, говорит о том, что наиболее консервативными будут оценки по модальным значениям. Оценки по медианным или средним значениям выглядят завышенными (при $V = 28\,000$ руб./1 кв. м – примерно в 1,4 и 1,7 раза, *рисунок 4*). Предположим, что нас интересует следующий вопрос: если цена предложения 28 000 руб. за 1 кв. м приведенной площади, а площадь земельного участка равна 30 000 кв. м, то какую площадь улучшений (и, соответственно, какой коэффициент

плотности застройки) следовало бы считать адекватным такой цене и площади земельного участка. Под коэффициентом застройки будем понимать отношение оценочного значения площади улучшений к площади земельного участка, т.е.

$$\frac{Mode(SB|SP = sp)}{sp} \quad (\text{как вариант, } \frac{Median(SB|SP = sp)}{sp}, \frac{E(SB|SP = sp)}{sp}).$$

На *рисунке 4* показано, что оценки по модальным, медианным и средним значениям могут значительно отличаться. Применение формулы (9) к результатам расчета условных параметров при цене 28 000 руб./кв. м и значению площади земельного участка равному 30 000 кв. м дает результат 7 165 кв. м улучшений, и тогда коэффициент плотности застройки составляет $7\,165 / 30\,000 = 0,24$. Таким образом, исходя из данных примера (*таблица 1*), другой коэффициент площади застройки при цене 28 000 руб./кв. м, площади земельного участка 30 000 кв. м может пониматься как не отвечающий выставленной цене. Этот же результат можно было получить применением формулы, аналогичной формуле (1). В этом разделе последовательный учет условий (сначала цены $V = v$, затем площади земельного участка $SP = sp$) применен для того, чтобы показать, что коэффициент плотности застройки не является константой внутри одной ценовой группы и даже для одной конкретной цены, и имеет зависимость степенного характера от площади земельного участка. В левой части *рисунка 4* показаны линии модальных, медианных и средних значений площади улучшений, в зависимости от площади

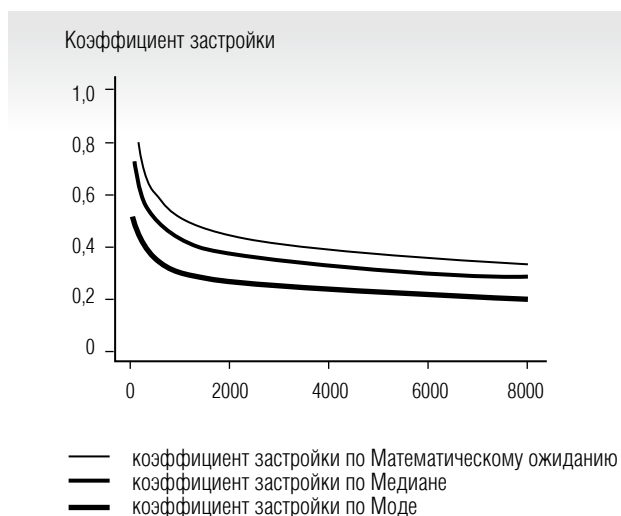
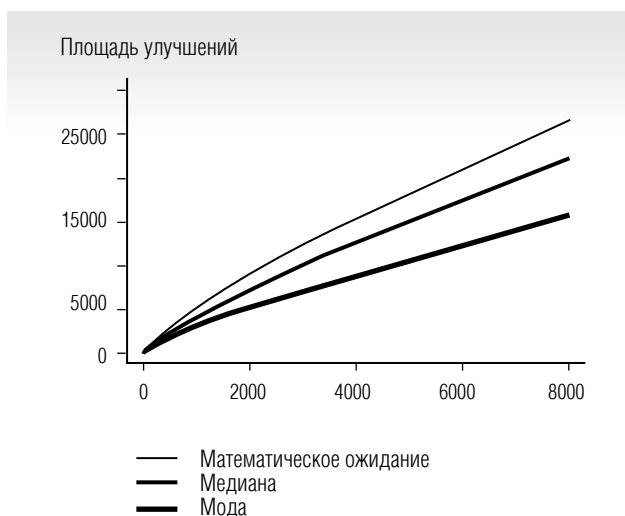


Рис. 4. Значения площади улучшений и коэффициентов застройки в зависимости от площади земельного участка

земельного участка для случая, когда цена предложения равна 28 000 руб./кв. м площади существующих улучшений, в правой части – линии коэффициентов застройки для соответствующих оценок площади улучшений. Из *рисунка 4* видно, что при заданной цене (ценовой группе) коэффициент застройки с приемлемой для целей оценки точностью может быть оценен константой только при достаточно большой площади земельного участка. Для участков с малой площадью коэффициент застройки не может быть оценен константой и должен изучаться индивидуально с учетом площади участка.

4. Замечание об общем виде совместного логарифмически нормального распределения вектора (V, SB, SP)

Многомерное распределение компонент вектора $(W, Y, Z) = (\ln(V), \ln(SB), \ln(SP))$ совместно нормально и обладает симметрией. Облака рассеяния эмпирических наблюдений будут приобретать вид трехмерных эллипсоидов. Точка максимума плотности имеет координаты, равные средним значениям компонент W, Y, Z . Распределение компонент вектора (V, SB, SP) ассиметрично, точка максимума плотности не является центром симметрии и может быть рассчитана (см. Приложение) по следующим формулам:

$$V_{max} = \exp(\mu_W - \sigma_W^2 - \rho_{WY}\sigma_W\sigma_Y - \rho_{WZ}\sigma_W\sigma_Z),$$

$$SB_{max} = \exp(\mu_Y - \sigma_Y^2 - \rho_{YW}\sigma_Y\sigma_W - \rho_{YZ}\sigma_Y\sigma_Z),$$

$$SP_{max} = \exp(\mu_Z - \sigma_Z^2 - \rho_{ZY}\sigma_Z\sigma_Y - \rho_{ZW}\sigma_Z\sigma_W).$$

На *рисунке 5* показаны: рассеяние исходных данных и рассеяние логарифмов исходных данных, точка максимальной плотности в пространстве (V, SB, SP) с координатами $V_{max} = 20\ 004$ руб./кв.м, $SB_{max} = 649$ кв.м, $SP_{max} = 1\ 202$ кв.м, а также точка максимальной плотности в логарифмированном пространстве $(W, Y, Z) = (\ln(V), \ln(SB), \ln(SP))$ с координатами $\mu_W = 10,30; \mu_Y = 8,45; \mu_Z = 9,35$. На рисунке черным цветом отмечены точки максимальной плотности: слева – в пространстве (V, SB, SP) , справа – в пространстве $(W, Y, Z) = (\ln(V), \ln(SB), \ln(SP))$.

Если сгенерировать 1000 трехмерных случайных векторов с теми же параметрами, то полученное рассеяние будет иметь вид, показанный на *рисунке 6*. Слева на рисунке показано рассеяние для 1000 сгенерированных случайных векторов, компоненты которых распределены совместно логарифмически нормально, с теми же параметрами, что и параметры исходной выборки (по которым проводилось тестирование на совместную нормальность логарифмов). Справа показано рассеяние логарифмов компонент 1000 сгенерированных случайных векторов. Черным отмечены точки максимальной плотности: слева – в пространстве (V, SB, SP) , справа – в пространстве $(W, Y, Z) = (\ln(V), \ln(SB), \ln(SP))$.

Очевидно, что (см. Приложение) точка максимума плотности многомерного вектора (мода), логарифмы компонент которого распределены совместно нормально – единственна. Всем остальным значениям плотности отвечают множества, описываемые в логарифмическом измерении полыми трехмерными эллипсоидами, а в исходных координатах множества, отвечающие одному значению плотности, представляют собой результат

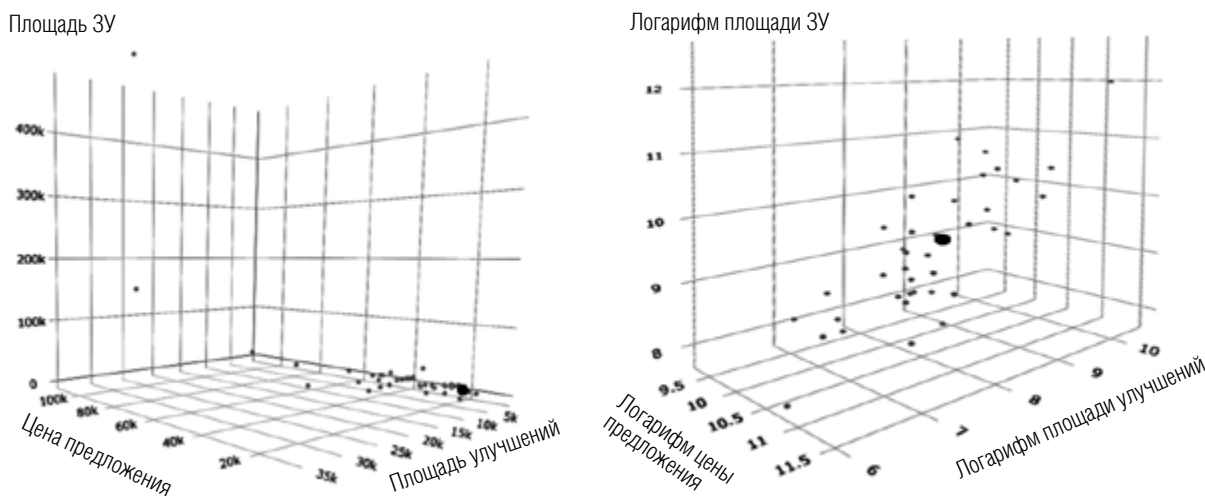


Рис. 5. Рассеяние исходных данных (слева), рассеяние логарифмов исходных данных (справа)

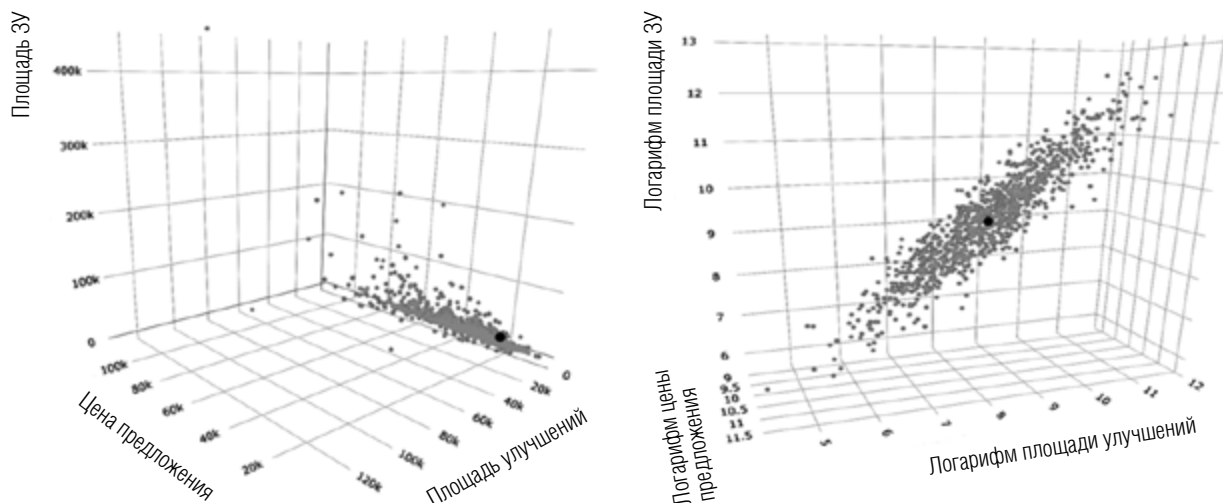


Рис. 6. Результат для 1000 сгенерированных случайных векторов

искажения (растяжения) полых эллипсоидов при обратном экспоненциальном преобразовании пространства. Таким образом, именно модальная оценка рыночной стоимости должна приводить к корректному, не создающему конфликтных ситуаций результату. Все остальные (не модальные) оценки рыночной стоимости потенциально являются источником постоянных споров относительно рыночной стоимости объекта оценки.

Заключение

Рассмотрение цен объектов сравнения и значений ценообразующих факторов как многомерных случайных величин открывает новые возможности в оценке недвижимого имущества. Часто оказывается, что эмпирические наблюдения цен и соответствующих им значений ценообразующих факторов хорошо приближаются логарифмически нормальным законом распределения, в том числе многомерным, что позволяет вывести расчетные формулы для различных задач оценки. Громоздкость этих формул компенсируется возможностями современных прикладных статистических пакетов (в частности, R). Кроме того возможность сведения расчетов к хорошо изученному многомерному нормальному закону логарифмированием компонент делает такой выбор модельного распределения предпочтительным.

Условные распределения цен при известных значениях ценообразующих факторов дают возможность оценить рыночную стоимость в полном соответствии с ее определением, закрепленным в российском законодательстве и зарубежных стан-

дартах, как точку максимума плотности условного распределения цен.

Условные распределения ценообразующих факторов при заданной цене предложения позволяют оценить адекватность цены предложения с точки зрения набора ценообразующих факторов.

Вряд ли следует ожидать, что практикующие оценщики готовы применять приведенные в настоящей статье формулы в ежедневной практике оценки и бизнес-анализа. Этого и не требуется. Однажды написанный и отлаженный скрипт (в статистическом пакете R или в других специализированных пакетах) позволит с легкостью решать подобные задачи практически в режиме реального времени. Следует признать, что в период цифровой трансформации экономики и бизнес-анализа оценочному бизнесу пора переходить к продвинутым статистическим пакетам и автоматической обработке данных. ■

Приложение

Утверждение. Абсолютный максимум (мода) плотности случайного логарифмически нормального вектора \vec{x} достигается в точке с координатами $\exp(\vec{\mu} - \Sigma \times \mathbf{1})$, где $\vec{\mu}$ – вектор математических ожиданий логарифмов компонент, Σ – ковариационная матрица логарифмов компонент, $\mathbf{1}$ – вектор, состоящий из единиц.

Доказательство. Рассмотрим плотность многомерного нормального распределения центрированного случайного вектора \vec{y} :

$$f(\bar{\mathbf{y}}) = \frac{1}{(2\pi)^{\frac{n}{2}} \sqrt{\det \Sigma}} \exp\left(-\frac{1}{2}(\Sigma^{-1}\bar{\mathbf{y}}, \bar{\mathbf{y}})\right).$$

При замене переменных $\bar{\mathbf{y}} = \ln(\bar{\mathbf{x}})$ плотность лог-нормального распределения случайного вектора $\bar{\mathbf{x}}$ будет:

$$f(\bar{\mathbf{x}}) = \frac{1}{(2\pi)^{\frac{n}{2}} \sqrt{\det \Sigma}} \times \frac{1}{\prod_{i=1}^n x_i} \times \exp\left(-\frac{1}{2}(\Sigma^{-1} \overline{\ln(\mathbf{x})}, \overline{\ln(\mathbf{x})})\right), \text{ где}$$

$\frac{1}{\prod_{i=1}^n x_i}$ – Якобиан преобразования координат,

Σ – ковариационная матрица, $\overline{\ln(\mathbf{x})}$ – центрированный случайный вектор. В точке абсолютного максимума плотности совместного логарифмически нормального распределения производная по любому направлению должна равняться нулю, что означает равенство нулю всех частных производных.

$$\frac{\partial f(\bar{\mathbf{x}})}{\partial x_j} = \frac{1}{(2\pi)^{\frac{n}{2}} \sqrt{\det \Sigma}} \times \frac{1}{\prod_{i=1}^n x_i} \left(-\frac{1}{x_j}\right) \times \exp\left(-\frac{1}{2}(\Sigma^{-1} \overline{\ln(\mathbf{x})}, \overline{\ln(\mathbf{x})})\right) + \frac{1}{(2\pi)^{\frac{n}{2}} \sqrt{\det \Sigma}} \times \frac{1}{\prod_{i=1}^n x_i} \exp\left(-\frac{1}{2}(\Sigma^{-1} \overline{\ln(\mathbf{x})}, \overline{\ln(\mathbf{x})})\right) \times \left(-\frac{2}{2} \Sigma^{-1} \overline{\ln(\mathbf{x})}\right) \times \frac{1}{x_j} = 0$$

После вынесения за скобки общих множителей остается условие:

$$-\mathbf{1} + (-\Sigma^{-1} \overline{\ln(\mathbf{x})}) = 0 \text{ или } (-\Sigma^{-1} \overline{\ln(\mathbf{x})}) = \mathbf{1},$$

где $\mathbf{1}$, $-\mathbf{1}$ – векторы размерности n , состоящие из единиц/минус единиц).

Умножим последнее равенство слева на Σ :

$$\Sigma \Sigma^{-1} \overline{\ln(\mathbf{x})} = -\Sigma \times \mathbf{1}, \\ \mathbf{E} \times \overline{\ln(\mathbf{x})} = -\Sigma \times \mathbf{1}.$$

Здесь \mathbf{E} – единичная матрица (на главной диагонали – единицы, остальные элементы равны нулю), $\mathbf{1}$ – вектор, состоящий из единиц. Т.е. значения вектора $\overline{\ln(\mathbf{x})}$, в котором все частные производные равны нулю, равны построчным суммам ковариационной матрицы, взятым с обратным знаком.

Остается вспомнить, что $\bar{\mathbf{y}} = \ln(\bar{\mathbf{x}})$ – центрированный случайный вектор. Если вектор математических ожиданий $\bar{\mu}$ содержит ненулевые значения, то окончательное решение:

$$\overline{\ln(\mathbf{x})} = \bar{\mu} - \Sigma \times \mathbf{1} \text{ или } \bar{\mathbf{x}} = \exp(\bar{\mu} - \Sigma \times \mathbf{1}).$$

С учетом отрицательной определенности квадратичной формы, составленной из вторых частных производных в точке $\bar{\mathbf{x}} = \exp(\bar{\mu} - \Sigma \times \mathbf{1})$ (автор опускает эту громоздкую запись, ее результат очевиден из геометрических соображений), точка $\bar{\mathbf{x}} = \exp(\bar{\mu} - \Sigma \times \mathbf{1})$ является точкой максимума плотности случайного логарифмически нормального вектора $\bar{\mathbf{x}}$.

Утверждение доказано.

Литература

1. Rusakov O.V., Laskin M.B., Jaksumbaeva O.I. Pricing in the real estate market as a stochastic limit. Log Normal approximation // International Journal of Mathematical Models and Methods in Applied Sciences. 2016. No 10. P. 229–236.
2. Aitchinson J., Brown J.A.C. The Lognormal distribution with special references to its uses in economics. Cambridge: University Press, 1963.
3. Ohnishi T., Mizuno T., Shimizu C., Watanabe T. On the evolution of the house price distribution // Columbia Business School. Center of Japanese Economy and Business. Working Paper Series. 2011. No 296.
4. Anselin L., Lozano-Gracia N. Errors in variables and spatial effects in hedonic house price models of ambient air quality // Empirical Economics. 2008. Vol. 34. No 1. P. 5–34. DOI: 10.1007/s00181-007-0152-3.
5. Benson E.D., Hansen J.L., Schwartz Jr. A.L., Smersh G.T. Pricing residential amenities: The value of a view // Journal of Real Estate Finance and Economics. 1998. Vol. 16. No 1. P. 55–73. DOI: 10.1023/A:1007785315925.
6. Debrezion G., Pels E., Rietveld P. The impact of rail transport on real estate prices: an empirical analysis of the Dutch housing market // Urban Studies. 2011. Vol. 48. No 5. P. 997–1015. DOI: 10.1177/0042098010371395.
7. Jim C.Y., Chen W.Y. Impacts of urban environmental elements on residential housing prices in Guangzhou (China) // Landscape and Urban Planning. 2006. Vol. 78. No 4. P. 422–434. DOI: 10.1016/j.landurbplan.2005.12.003.
8. Rivas R., Patil D., Hristidis V., Barr J.R., Srinivasan N. The impact of colleges and hospitals to local real estate markets // Journal of Big Data. 2019. Vol. 6. No 1. Article No 7 (2019). DOI: 10.1186/s40537-019-0174-7.
9. Wena H., Zhanga Y., Zhang L. Assessing amenity effects of urban landscapes on housing price in Hangzhou, China // Urban Forestry & Urban Greening. 2015. No 14. P. 1017–1026. DOI: 10.1016/j.ufug.2015.09.013.
10. Отчет об определении кадастровой стоимости объектов недвижимости на территории Санкт-Петербурга. 2018. №1. [Электронный ресурс]: <http://www.ko.spb.ru/interim-reports/> (дата обращения: 05.06.2019).

11. Peterson S., Flanagan A.B. Neural network hedonic pricing models in mass real estate appraisal // *Journal of Real Estate Research*. 2009. Vol. 31. No 2. P. 147–164.
12. Rafiei M.H., Adeli H. Novel machine-learning model for estimating construction costs considering economic variables and indexes // *Journal of Construction Engineering and Management*. 2018. Vol. 144. No 12. Article No 04018106. DOI: 10.1061/(asce)co.1943-7862.0001570.
13. Antipov E.A., Pokryshevskaya E.B. Mass appraisal of residential apartments: An application of Random forest for valuation and a CART-based approach for model diagnostics // *Expert Systems with Applications*. 2012. No 39. P. 1772–1778. DOI: 10.1016/j.eswa.2011.08.077.
14. Kontrimas V., Verikas A. The mass appraisal of the real estate by computational intelligence // *Applied Soft Computing*. 2011. No 11. P. 443–448. DOI: 10.1016/j.asoc.2009.12.003.
15. Park B., Baem J.K. Using machine learning algorithms for housing price prediction: The case of Fairfax County, Virginia housing data // *Expert Systems with Applications*. 2015. No 42. P. 2928–2934. DOI: 10.1016/j.eswa.2014.11.040.
16. Case K.E., Shiller R.J. Prices of single-family Homes since 1970: New indexes for four cities // *New England Economic Review*. 1987. September. P. 45–56. DOI: 10.3386/w2393.
17. Englund P., Quigley J.M., Redfearn C.L. The choice of methodology for computing housing price indexes: comparison of temporal aggregation and sample definition // *Journal of Real Estate Finance and Economics*. 1999. Vol. 19. No 2. P. 91–112. DOI: 10.1023/A:1007846404582.
18. Epley D. Assumptions and restrictions on the use of repeat sales to estimate residential price appreciation // *Journal of Real Estate Literature*. 2016. Vol. 24. No 2. P. 275–286. DOI: 10.5555/0927-7544.24.2.275.
19. Malpezzi S. Hedonic pricing models: A selective and applied review // *Housing economics and public policy: Essays in honor of Duncan MacLennan* (T. O’Sullivan, K. Gibb, eds.). Oxford, UK: Blackwell Science, 2002. P. 67–89. DOI: 10.1002/9780470690680.ch5.
20. Case B., Quigley J.M. The dynamics of real estate prices // *Review of Economics and Statistics*. 1991. Vol. 73. No 1. P. 50–58.
21. Englund P., Quigley J.M., Redfearn C.L. Improved price indexes for real estate: Measuring the course of Swedish housing prices // *Journal of Urban Economics*. 1998. Vol. 44. No 2. P. 171–196.
22. Jones C. House price measurement: The hybrid hedonic repeat-sales method // *Economic Record*. 2010. Vol. 86. No 272. P. 95–97. DOI: 10.1111/j.1475-4932.2009.00596.x.
23. Wang F., Zheng X. The comparison of the hedonic, repeat sales, and hybrid models: Evidence from the Chinese paintings // *Cogent Economics & Finance*. 2018. No 6. P. 1–19. DOI: 10.1080/23322039.2018.1443372.
24. Brunnermeier M.K. Bubbles // *The new Palgrave dictionary of Economics* (L.E. Blume, S.N. Durlauf, eds.). New York: Palgrave Macmillan, 2009.
25. Fabozzi F.J., Xiao K. The timeline estimation of bubbles: The case of real estate // *Real Estate Economics*. 2019. Vol. 47. No 2. P. 564–594. DOI: 10.1111/1540-6229.12246.
26. Fernandez-Kranz D., Hon M.T. A cross-section analysis of the income elasticity of housing demand in Spain: Is there a real estate bubble? // *Journal of Real Estate Finance and Economics*. 2006. Vol. 32. No 4. P. 449–470. DOI: 10.1007/s11146-006-6962-9.
27. Phillips P.C.B., Shi S. P., Yu J. Testing for multiple bubbles: Historical episodes of exuberance // *International Economic Review*. 2015. Vol. 56. No 4. P. 1043–1078. DOI: 10.1111/iere.12132.
28. Phillips P.C.B., Shi S. P., Yu J. Testing for multiple bubbles: Limit theory of real time detectors // *International Economic Review*. 2015. Vol. 56. No 4. P. 1079–1134. DOI: 10.1111/iere.12131.
29. Федеральный закон от 29.07.1998 № 135-ФЗ (ред. от 29.07.2017) «Об оценочной деятельности в Российской Федерации». [Электронный ресурс]: http://www.consultant.ru/document/cons_doc_LAW_19586/ (дата обращения: 14.03.2020).
30. Слуцкий А.А., Слуцкая И.А. Модифицированный метод выделения и обобщенный модифицированный метод выделения. Применение для анализа сегмента рынка, к которому относится объект оценки. 2020 [Электронный ресурс]: <http://tmpro.su/sluckij-a-a-sluckaya-i-a-mmv-i-ommv-primenenie-dlya-analiza-rynka-3/> (дата обращения: 14.03.2020).
31. Ласкин М.Б. Логарифмически нормальное распределение цен и рыночная стоимость на рынке недвижимости // *Известия Санкт-Петербургского государственного технологического института*. 2014. № 25 (51). С. 102–106.
32. Ласкин М.Б. Корректировка рыночной стоимости по ценообразующему фактору «площадь объекта» // *Имущественные отношения в Российской Федерации*. 2017. № 8 (191), с. 86–99.

Об авторе

Ласкин Михаил Борисович

кандидат физико-математических наук, доцент;
старший научный сотрудник, Санкт-Петербургский института информатики и автоматизации,
Российская академия наук (СПИИРАН),
199178, Россия, г. Санкт-Петербург, 14 линия, д. 39;
E-mail: laskinmb@yahoo.com
ORCID: 0000-0002-0143-4164

Multidimensional log-normal distribution in real estate appraisals

Michael B. Laskin

E-mail: laskinmb@yahoo.com

St. Petersburg Institute for Informatics and Automation, Russian Academy of Sciences
Address: 39, 14 Line, St. Petersburg 199178, Russia

Abstract

The purpose of the research was to develop a market value appraisal methodology based on a set of a joint logarithmically normal distribution of price-forming factors. Joint logarithmically normal distribution means random vector component logarithms are distributed together jointly normally. This article suggests a method for appraising the real estate market value based on the statistical hypothesis of a joint logarithmically normal distribution and conditional distribution of prices with fixed values of pricing factors. The article suggests a method of offer price analysis from the point of view of its relevance to pricing factor values. We consider the features of the coefficient of development depending on the area of the land plot. Additional arguments are given in favor of estimating market value as a mode of conditional laws of price distribution. An example of a multidimensional log-normal distribution of prices and pricing factors such as the area of the improvements (improvements mean buildings and constructions) area and the land area in real data, i.e. for the case of a three-dimensional random vector. We present a formula for determining the absolute maximum density point of a multidimensional logarithmically normal random vector. The proof is given in the Appendix. The results obtained can be used to create information systems to support decision-making in valuation activities for real estate properties.

Key words: market value; logarithmically normal law of price distribution; multidimensional logarithmically normal distribution, valuation of real estate.

Citation: Laskin M.B. (2020) Multidimensional log-normal distribution in real estate appraisals. *Business Informatics*, vol. 14, no 2, pp. 48–63. DOI: 10.17323/2587-814X.2020.2.48.63

References

1. Rusakov O.V., Laskin M.B., Jaksumbaeva O.I. (2016) Pricing in the real estate market as a stochastic limit. Log Normal approximation. *International Journal of Mathematical Models and Methods in Applied Sciences*, no 10, pp. 229–236.
2. Aitchinson J., Brown J.A.C. (1963) *The Lognormal distribution with special references to its uses in economics*. Cambridge: University Press.
3. Ohnishi T., Mizuno T., Shimizu C., Watanabe T. (2011) *On the evolution of the house price distribution*. Columbia Business School. Center of Japanese Economy and Business. Working Paper Series, no 296.
4. Anselin L., Lozano-Gracia N. (2008) Errors in variables and spatial effects in hedonic house price models of ambient air quality. *Empirical Economics*, vol. 34, no 1, pp. 5–34. DOI: 10.1007/s00181-007-0152-3.
5. Benson E.D., Hansen J.L., Schwartz Jr. A.L., Smersh G.T. (1998) Pricing residential amenities: The value of a view. *Journal of Real Estate Finance and Economics*, vol. 16, no 1, pp. 55–73. DOI: 10.1023/A:1007785315925.
6. Debrezion G., Pels E., Rietveld P. (2011) The impact of rail transport on real estate prices: an empirical analysis of the Dutch housing market. *Urban Studies*, vol. 48, no 5, pp. 997–1015. DOI: 10.1177/0042098010371395.
7. Jim C.Y., Chen W.Y. (2006) Impacts of urban environmental elements on residential housing prices in Guangzhou (China). *Landscape and Urban Planning*, vol. 78, no 4, pp. 422–434. DOI: 10.1016/j.landurbplan.2005.12.003.
8. Rivas R., Patil D., Hristidis V., Barr J.R., Srinivasan N. (2019) The impact of colleges and hospitals to local real estate markets. *Journal of Big Data*, vol. 6, no 1, article no 7 (2019). DOI: 10.1186/s40537-019-0174-7.
9. Wena H., Zhanga Y., Zhang L. (2015) Assessing amenity effects of urban landscapes on housing price in Hangzhou, China. *Urban Forestry & Urban Greening*, no 14, pp. 1017–1026. DOI: 10.1016/j.ufug.2015.09.013.
10. Saint Petersburg State Budget Department “Cadastral Valuation City Department” (2018) *Report on determining the cadastral value of real estate objects on the territory of Saint Petersburg*, no 1. Available at: <http://www.ko.spb.ru/interim-reports/> (accessed 05 June 2019).
11. Peterson S., Flanagan A.B. (2009) Neural network hedonic pricing models in mass real estate appraisal. *Journal of Real Estate Research*, vol. 31, no 2, pp. 147–164.

12. Raffei M.H., Adeli H. (2018) Novel machine-learning model for estimating construction costs considering economic variables and indexes. *Journal of Construction Engineering and Management*, vol. 144, no 12, article no 04018106. DOI: 10.1061/(asce)co.1943-7862.0001570.
13. Antipov E.A., Pokryshevskaya E.B. (2012) Mass appraisal of residential apartments: An application of Random forest for valuation and a CART-based approach for model diagnostics. *Expert Systems with Applications*, no 39, pp. 1772–1778. DOI: 10.1016/j.eswa.2011.08.077.
14. Kontrimas V., Verikas A. (2011) The mass appraisal of the real estate by computational intelligence. *Applied Soft Computing*, no 11, pp. 443–448. DOI: 10.1016/j.asoc.2009.12.003.
15. Park B., Baem J.K. (2015) Using machine learning algorithms for housing price prediction: The case of Fairfax County, Virginia housing data. *Expert Systems with Applications*, no 42, pp. 2928–2934. DOI: 10.1016/j.eswa.2014.11.040.
16. Case K.E., Shiller R.J. (1987) Prices of single-family Homes since 1970: New indexes for four cities. *New England Economic Review*, September, pp. 45–56. DOI: 10.3386/w2393.
17. Englund P., Quigley J.M., Redfearn C.L. (1999) The choice of methodology for computing housing price indexes: comparison of temporal aggregation and sample definition. *Journal of Real Estate Finance and Economics*, vol. 19, no 2, pp. 91–112. DOI: 10.1023/A:1007846404582.
18. Epley D. (2016) Assumptions and restrictions on the use of repeat sales to estimate residential price appreciation. *Journal of Real Estate Literature*, vol. 24, no 2, pp. 275–286. DOI: 10.5555/0927-7544.24.2.275.
19. Malpezzi S. (2002) Hedonic pricing models: A selective and applied review. *Housing economics and public policy: Essays in honor of Duncan MacLennan* (T. O'Sullivan, K. Gibb, eds.). Oxford, UK: Blackwell Science, pp. 67–89. DOI: 10.1002/9780470690680.ch5.
20. Case B., Quigley J.M. (1991) The dynamics of real estate prices. *Review of Economics and Statistics*, vol. 73, no 1, pp. 50–58.
21. Englund P., Quigley J.M., Redfearn C.L. (1998) Improved price indexes for real estate: Measuring the course of Swedish housing prices. *Journal of Urban Economics*, vol. 44, no 2, pp. 171–196.
22. Jones C. (2010) House price measurement: The hybrid hedonic repeat-sales method. *Economic Record*, vol. 86, no 272, pp. 95–97. DOI: 10.1111/j.1475-4932.2009.00596.x.
23. Wang F., Zheng X. (2018) The comparison of the hedonic, repeat sales, and hybrid models: Evidence from the Chinese paintings. *Cogent Economics & Finance*, no 6, pp. 1–19. DOI: 10.1080/23322039.2018.1443372.
24. Brunnermeier M.K. (2009) Bubbles. *The new Palgrave dictionary of Economics* (L.E. Blume, S.N. Durlauf, eds.). New York: Palgrave Macmillan.
25. Fabozzi F.J., Xiao K. (2019) The timeline estimation of bubbles: The case of real estate. *Real Estate Economics*, vol. 47, no 2, pp. 564–594. DOI: 10.1111/1540-6229.12246.
26. Fernandez-Kranz D., Hon M.T. (2006) A cross-section analysis of the income elasticity of housing demand in Spain: Is there a real estate bubble? *Journal of Real Estate Finance and Economics*, vol. 32, no 4, pp. 449–470. DOI: 10.1007/s11146-006-6962-9.
27. Phillips P.C.B., Shi S.-P., Yu J. (2015) Testing for multiple bubbles: Historical episodes of exuberance. *International Economic Review*, vol. 56, no 4, pp. 1043–1078. DOI: 10.1111/iere.12132.
28. Phillips P.C.B., Shi S. P., Yu J. (2015) Testing for multiple bubbles: Limit theory of real time detectors. *International Economic Review*, vol. 56, no 4, pp. 1079–1134. DOI: 10.1111/iere.12131.
29. The Federal Law of 29.07.1998 No 135-FZ (edition of 29.07.2017) “About assessment activity in the Russian Federation”. Available at: http://www.consultant.ru/document/cons_doc_LAW_19586/ (accessed 14 March 2020).
30. Slytsky A.A., Slytskaya I.A. (2020) *The modified extraction method and the generalized modified method of allocation. Use for analyzing the market segment that the item is being evaluated belongs to*. Available at: <http://tmpo.su/sluckij-a-a-sluckaya-i-a-mmv-i-ommv-primeneniye-dlya-analiza-rynka-3/> (accessed 14 March 2020).
31. Laskin M.B. (2014) Logarithmically normal distribution of prices and market value in the real estate market. *Saint Petersburg State Technological Institute Review*, no 25 (51), pp.102–106.
32. Laskin M.B. (2017) Market value adjustment for the pricing factor “square”. *Property Relations in the Russian Federation*, no 8 (191), pp. 86–99.

About the author

Michael B. Laskin

Cand. Sci. (Phys.-Math.), Associate Professor;

Senior Researcher, St. Petersburg Institute for Informatics and Automation,
Russian Academy of Sciences (SPIIRAS),
39, 14 Line, St. Petersburg 199178, Russia

E-mail: laskinmb@yahoo.com

ORCID: 0000-0002-0143-4164