Вопросы имущественной политики – практический опыт

Логарифмически нормальное распределение цен на объекты недвижимости

М.Б. Ласкин
директор ООО «ИНВЕСТ-СТРОЙ», доцент, кандидат физико-математических наук (г. Санкт-Петербург)
С.В. Пупенцова
доцент кафедры экономики и менеджмента недвижимости и технологий государственного образовательного учреждения «Санкт-Петербургский государственный политехнический университет», доцент, кандидат экономических наук (г. Санкт-Петербург)

Светлана Валентиновна Пупенцова, pupentsova@spbgpu-dreem.ru

В условиях активного рынка и общей доступности листингов предложения на рынке недвижимости применение статистических методов позволяет профессиональным консультантам определить цену, по которой с высокой степенью вероятности может состояться сделка на дату оценки. Напомним, что для корректного применения многих статистических методов требуется предварительная проверка гипотезы о том, что выборочные данные принадлежат генеральной совокупности, имеющей нормальное распределение. В работе [1, с. 65] отмечается типичное (за редким исключением) отсутствие в современной практической оценке этапа проверки исходных данных на нормальность распределения. Применение статистических методов без принятия гипотезы о нормальном законе распределения в выборке ставит под сомнение надежность итоговых результатов расчета. Более того, в настоящей статье будет показано, что логарифмически нормальная модель точнее описывает реальное распределение цен предложения на рынке недвижимости.

Пусть цена недвижимости V распределена по логарифмически нормальному закону, то есть случайная величина «натуральный логарифм цены» $\ln \boldsymbol{V}$ распределена по нормальному закону. Напомним, что математическое ожидание для логарифмически нормально распределенной случайной величины можно рассчитать по формуле:

$$M(V) = e^{m+\frac{\sigma^2}{2}}, \qquad (1)$$

где $M(V)$ – математическое ожидание случайной величины V;
V – стоимость 1 м^2 объекта недвижимости;
e – число Эйлера (здесь и далее e = 2,718281…);

m – математическое ожидание логарифма стоимости;

σ – стандартное отклонение логарифма стоимости.

Тогда медиану (0,5 квантиль) для логарифмически нормально распределенной случайной величины определим по формуле:

$$Me(V) = e^{m}, \qquad (2)$$

где $Me(V)$ – медиана для логарифмически нормально распределенной случайной величины;

m – математическое ожидание логарифма стоимости.

Моду для логарифмически нормально распределенной случайной величины определим по формуле:

$$Mo(V) = e^{m-\sigma^2}, \qquad (3)$$

где $Mo(V)$ – мода для логарифмически нормально распределенной случайной величины V;

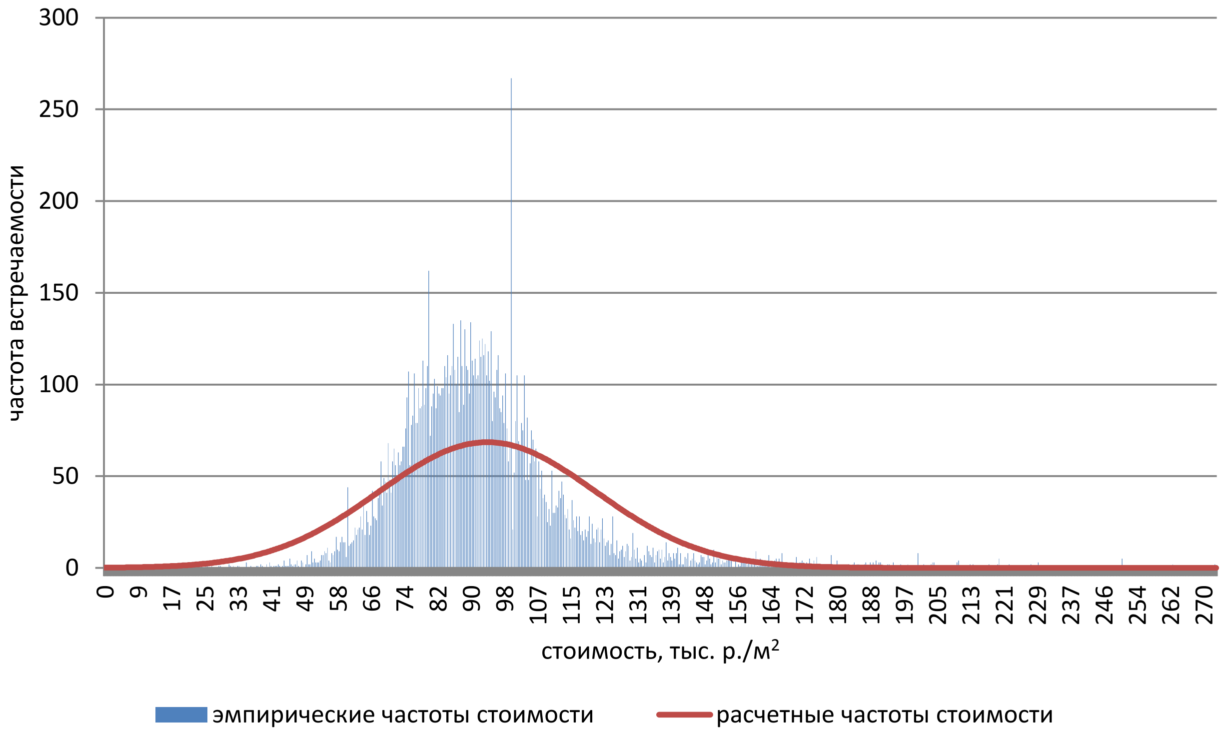
m – математическое ожидание логарифма стоимости;

σ – стандартное отклонение логарифма стоимости.

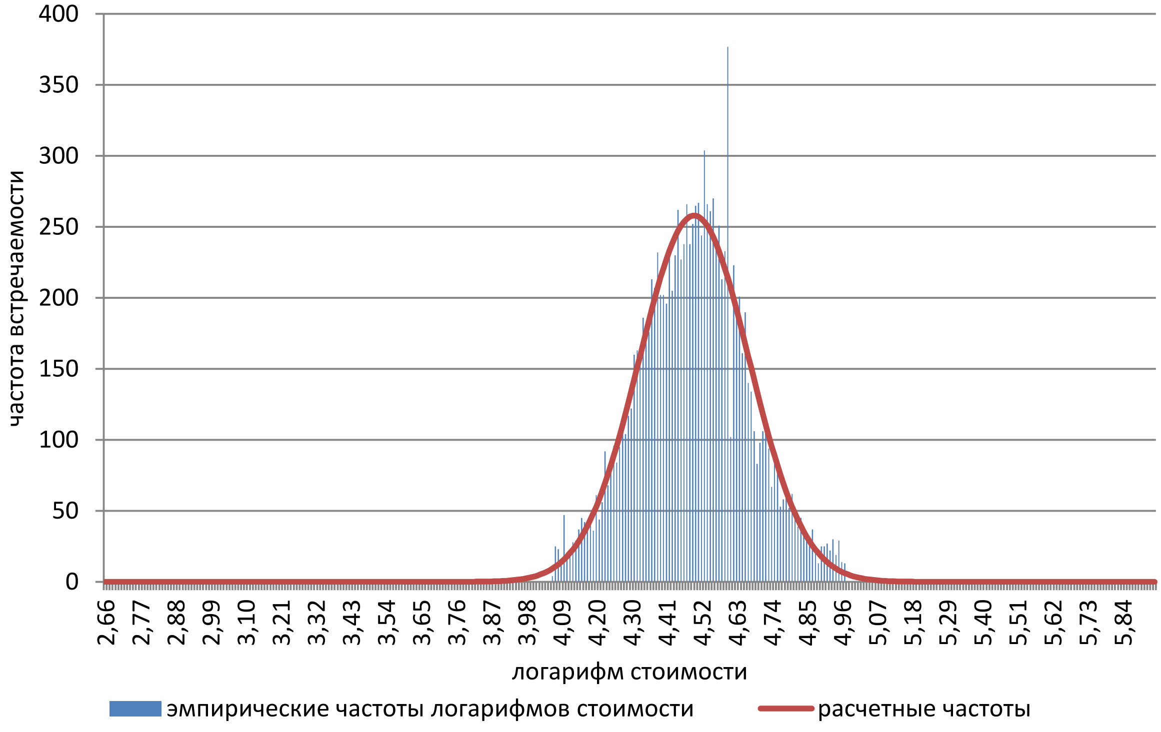
Для анализа выбран статистический материал из еженедельного журнала [2], содержащего объявления о продажах на вторичном рынке города Санкт-Петербурга. При формировании выборки удалены записи об объектах, принадлежащих другой генеральной совокупности (квартиры в Ленинградской области, комнаты в Санкт-Петербурге), а также некорректные объявления (не указаны цена или общая площадь). Таким образом, изучены все опубликованные в еженедельнике [2] предложения о продаже квартир в Санкт-Петербурге на 14 января 2013 года. Объем выборки n составил 13 759 объектов жилой недвижимости на вторичном рынке.

На рисунке 1 представлены эмпирические и расчетные частоты распределения стоимости в тысячах рублей за один квадратный метр (а) и логарифмов стоимости (б) всех объектов выборки.

На гистограмме стоимости объектов жилой недвижимости (а) хорошо видны асимметрия (правый «хвост» значительно длиннее и шире левого) и эксцесс эмпирического распределения в сравнении с теоретическим (нормальное с выборочными средним и стандартным отклонением). Общий характер эмпирического распределения подтверждает обоснованность выбора модели логарифмически нормального распределения, представленного на гистограмме (б).



а)



б)

Рис. 1. Гистограмма распределения стоимости жилой недвижимости на вторичном рынке

В таблице 1 представлены статистические характеристики выборки для моделей нормального распределения стоимости и логарифмически нормального распределения. При этом следует отметить, что по данным [3] (данные на 11 января 2013 года) средняя цена по городу (сектор вторичной недвижимости) составляла 95,225 тысячи рублей за один квадратный метр, хотя общее среднее по всей выборке оказалось равным 97,437 тысячи рублей за один квадратный метр. После удаления крайних выбросов из выборки на уровне значимости 0,9 среднее по выборке составило 94,221 тысячи рублей за один квадратный метр.

Таблица 1

Результаты анализа отобранных для анализа объектов

Показатель	***Выборка***		***Предельные значения для нормального распределения***
	распределение стоимости (рис. 1а)	*распределения логарифмов стоимости (рис. 1 б)*	
Среднее арифметическое *m*	94 221 р.	4,4967	–
Стандартное отклонение *σ*	26 613 р.	0,1719	–
Среднее отклонение *A*	16 896 р.	0,1369	–
Отношение стандартного отклонения к среднему отклонению, *λ = σ / A*	1,575	1,2556	~ 1,25
Коэффициент вариации, *Cv = σ / m*	0,283	0,0382	< 0,33
Коэффициент асимметрии	1,431	0,061	0
Коэффициент эксцесса	4,286	-0,1157	0

Полученные результаты не дают оснований отказаться от гипотезы нормального распределения логарифмов цен. Более того, они указывают на предпочтительность модели логарифмически нормального распределения цен. Следует отметить, что изучаемая выборка является отражением намерений участников рынка, а не информацией о законченных сделках, поэтому содержит большое количество выбросов субъективного характера (например «круглые цены» – 50, 55, 60…120, 125 тыс. р.), не влияющих, однако, на общий вид плотности распределения. По этой причине более строгая проверка гипотезы о нормальности распределения опиралась не на общепринятый критерий Пирсона χ^2 (чувствительный к таким выбросам частот), а на критерий согласия Гири (подробно см. [3]), не чувствительный к выбросам частот.

Статистика Гири d, определенная как отношение среднего отклонения к стандартному отклонению, распределена асимптотически нормально. Для проверки с помощью критерия согласия Гири рассчитаем среднее значение и дисперсию по следующим формулам:

$$M(d) \approx \left(1+\frac{2}{8n-9}\right)\sqrt{\frac{2}{\pi}} \approx 0,79788 \times \frac{n-0,875}{n-1,125}; \quad (4)$$

$$D(d) = \frac{1}{n}\left[\left(1-\frac{3}{\pi}\right)-\frac{1}{4\pi n}\right] \approx \frac{1}{n}\left(0,04507-\frac{0,0796}{n}\right), \quad (5)$$

где $M(d)$ – среднее значение;
$D(d)$ – дисперсия;
π – число Пи (3,141592…);
n – объем выборки;

и квантили по следующим формулам:

$$d\left(\frac{\alpha}{2}\right)=M(d)+U_{\frac{\alpha}{2}}\sqrt{D(d)};\qquad (6)$$

$$d\left(1-\frac{\alpha}{2}\right)=M(d)+U_{1-\frac{\alpha}{2}}\sqrt{D(d)},$$

где

$d\left(\frac{\alpha}{2}\right)$ и $d\left(1-\frac{\alpha}{2}\right)$ – квантили выборки;

α – уровень значимости,

$U_{\frac{\alpha}{2}}$, $U_{1-\frac{\alpha}{2}}$ – соответствующие квантили стандартного нормального распределения.

Напомним, что если $d\left(\frac{\alpha}{2}\right)<d<d\left(1-\frac{\alpha}{2}\right)$, то гипотеза нормальности не отклоняется. Для отобранных объектов d при принятом уровне значимости α на уровне 5 процентов равен 0,79645, тогда $d\left(\frac{\alpha}{2}\right)=0,79473$, $d\left(1-\frac{\alpha}{2}\right)=0,80106$, следовательно, условие Гири выполняется, оснований отклонить гипотезу о нормальном распределении логарифмов цен нет.

С точки зрения работы с подобными выборками (подверженными «субъективным» выбросам частот) критерий Гири представляется достаточно универсальным, так как он чувствителен к размаху выборки. То есть при подготовке выборки к анализу размах может быть отрегулирован различными уровнями значимости при исключении из выборки крайних значений, и при условии нормальности наиболее «плотной» части эмпирического распределения критерий Гири дает хорошие результаты.

Принимая модель логарифмически нормального распределения, получаем для распределения стоимости недвижимости в Санкт-Петербурге на 14 января 2013 года следующие значения:

- математическое ожидание величины стоимости $M(V)$ = 91,056 тыс. р.;
- медиана $Me(V)$ = 89,721 тыс. р.;
- мода (наивероятнейшее значение) $Mo(V)$ = 87,108 тыс. р.

Таким образом, принимая модель логарифмически нормального распределения, рыночная стоимость, принятая по наиболее вероятному модальному значению, составит 87,108 тысячи рублей за один квадратный метр. Полученное значение на 12 процентов ниже средней арифметической цены, равной 97,437 тысячи рублей за один квадратный метр и определенной распространенным способом в предположении нормального распределения цен.

В таблице 2 сведена информация о наиболее вероятных и средних ценах, полученных по сегментам в зависимости от количества комнат в квартире на 14 января 2013 года.

Таблица 2

Результаты средних арифметических и наиболее вероятных значений по сегментам (с округлением до целых)

Вид квартиры	Среднее значение, тыс. р.	Наиболее вероятное значение, тыс. руб.	Ошибка, %
Однокомнатная	102	96	6
Двухкомнатная	98	87	12
Трехкомнатная	92	81	14
Четырехкомнатная	97	83	17
Пятикомнатная	106	84	26

Авторами проведено исследование распределений цен и логарифмов цен вторичной недвижимости по всем данным бюллетеня недвижимости в период с 19 декабря 2011 года по 11 февраля 2013 года.

В течение всего периода прослеживается принцип логарифмически нормального распределения цен. Приведенная на рисунке 2 динамика средних и модальных цен показывает, что на протяжении всего периода наиболее вероятная стоимость одного квадратного метра недвижимости находилась заметно ниже средних цен.

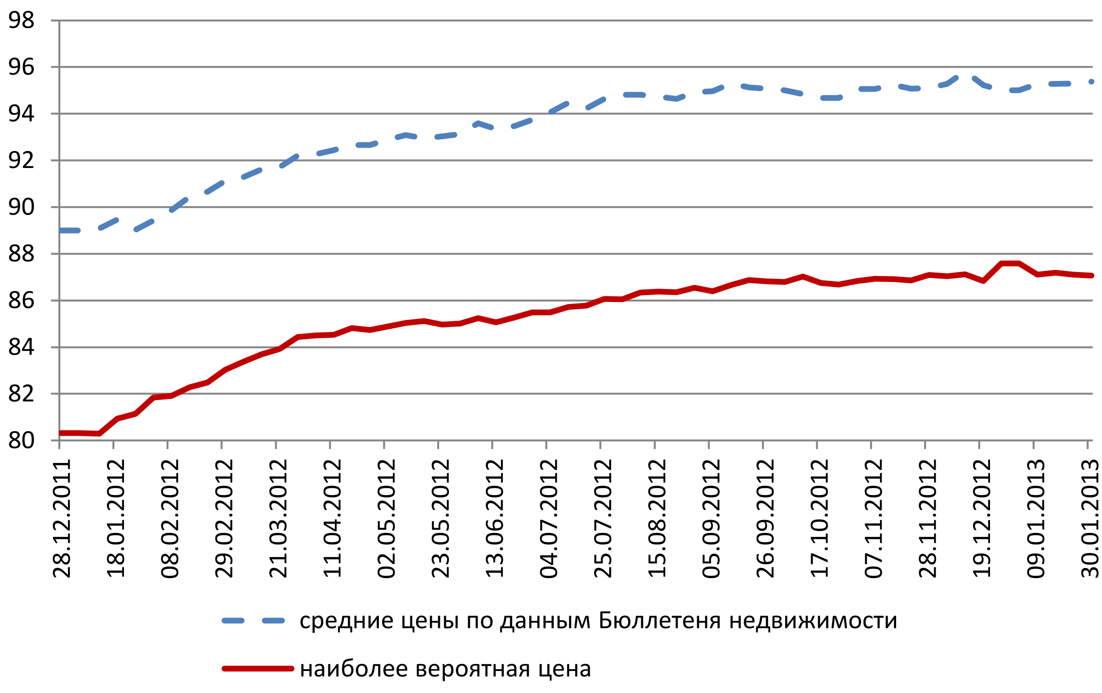


Рис. 2. Соотношение наивероятнейших и средних цен на недвижимость

Изменение во времени наиболее вероятных цен, показанное на рисунке 2, демонстрирует восходящий тренд с замедляющимся темпом роста.

Таким образом, использование логарифмически нормальной модели приводит к существенной корректировке рыночной стоимости по сравнению со средними выборочными значениями.

Кроме того, логарифмически нормальная модель, при некоторых допущениях, позволяет оценить размер скидки на торг. Предположим, что реальные стоимости (цены сделок) V_p распределены по такому же (логарифмически нормальному) закону, как и цены предложений $V_{пр}$, то есть подчиняются общему принципу формирования цен как логарифмически нормальному.

Рассмотрим случайную величину отношения V_p к $V_{пр}$.

Очевидно, что

$$\ln\left(\frac{V_p}{V_{пр}}\right) = \ln V_p - \ln V_{пр}. \tag{7}$$

В случае независимости случайных величин V_p, $V_{пр}$, логарифм отношения $\ln\left(\frac{V_p}{V_{пр}}\right)$ распределен по нормальному закону с параметрами математического ожидания, определенного как разница $m_p - m_{пр}$ и стандартного отклонения, равного $\sqrt{\sigma_p^2 + \sigma_{пр}^2}$, где m_p и $m_{пр}$ – математические ожидания, определенные соответственно для логарифмов цен сделок и цен предложения; σ_p^2 и $\sigma_{пр}^2$ – дисперсии, определенные соответственно для логарифмов цен сделок и цен предложения. Предположение о линейной зависимости V_p, $V_{пр}$ приводит к скидке на торг, равной константе для каждой сделки. Таким образом, предположение о независимости V_p, $V_{пр}$ позволяет оценить максимально достижимую наиболее вероятную скидку на торг. Тогда наивероятнейшее значение величины $V_p / V_{пр}$ равно моде, определенной по формуле:

$$Mo\left(\frac{V_p}{V_{пр}}\right) = e^{m_p - m_{пр} - \sigma_p^2 - \sigma_{пр}^2}. \tag{8}$$

Наиболее вероятное значение отношения $V_p / V_{пр}$ равно 1 при условии $m_p = m_{пр} + \sigma_p^2 + \sigma_{пр}^2$. Тогда при условии $m_p < m_{пр} + \sigma_p^2 + \sigma_{пр}^2$ величина C дает наивероятнейшее значение скидки от цены предложения**.**

Величину C определим по формуле:

$$C = 1 - Mo\left(\frac{V_p}{V_{пр}}\right), \tag{9}$$

где C – наивероятнейшее значение скидки от цены предложения;

$Mo\left(\frac{V_p}{V_{пр}}\right)$ – мода величины $V_p / V_{пр}$.

В условиях состояния рынка на 14 января 2013 г. сделаем предположение о том, что логарифмы реальных цен продаж и логарифмы цен предложений имеют одинаковые параметры нормального закона распределения m и σ.

Пусть $m_p = m_{пр} = m$ и $\sigma_p^2 = \sigma_{пр}^2 = \sigma^2$ (все перечисленные параметры относятся к нормальным законам распределений логарифмов цен предложений и сделок).

Тогда $\ln\left(\frac{V_p}{V_{np}}\right) \in N\left[0, \sqrt{2\sigma^2}\right]$, величина C будет определена по формуле:

$$C = 1 - Mo\left(\frac{V_p}{V_{np}}\right) = 1 - e^{0-2\sigma^2} = 1 - e^{-2\sigma^2}, \qquad (10)$$

где C – модальное значение скидки от цены предложения;

$Mo\left(\frac{V_p}{V_{np}}\right)$ – мода случайной величины V_p / V_{np};

σ^2 – квадрат стандартного отклонения логарифмов цен предложения.

Проведем расчет наиболее вероятной скидки на торг от цены предложений по распределению, приведенному на рисунке 1. Величина σ для этого распределения была получена на уровне 0,1719 (см. табл. 1).

Тогда наивероятнейшая скидка на торг от цены предложения до цены сделки для жилой недвижимости будет равна $C = 1 - e^{-2\times 0{,}1719^2} = 0{,}0574$, или 5,74 процента. Расчет скидки по сегментам подтвердил фактическую ситуацию на рынке: по ликвидным однокомнатным квартирам получена скидка на торг на уровне 2 процентов, для неликвидных многокомнатных квартир скидка увеличивается до 11 процентов.

Динамика скидки на торг, приведенная на рисунке 3, показывает, как менялось на протяжении всего периода модальное значение скидки от цены предложения до цены сделки для жилой недвижимости на вторичном рынке.

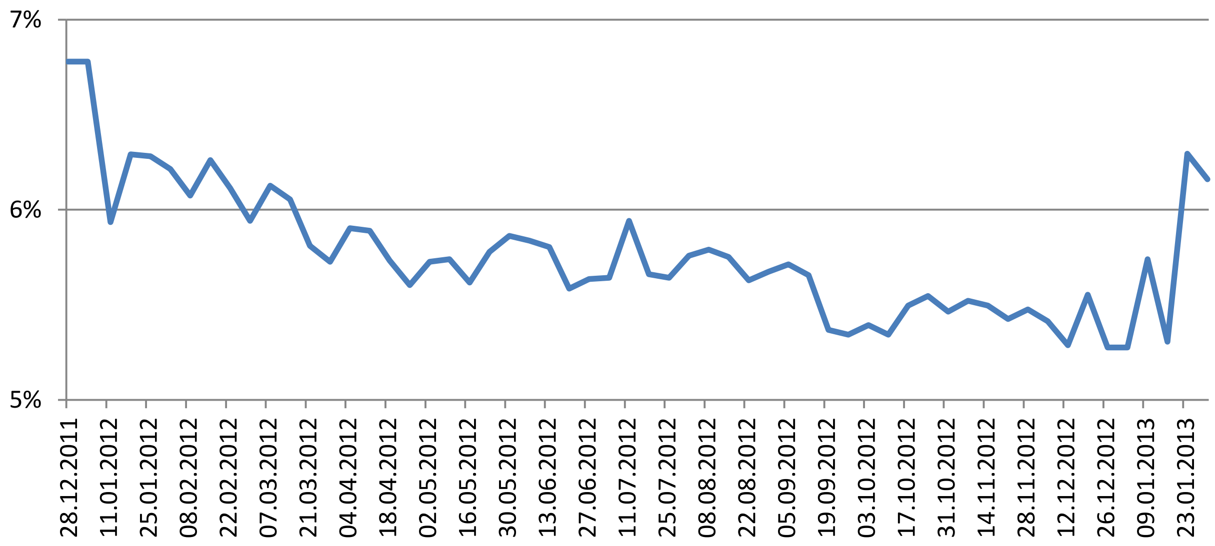


Рис. 3. Изменение скидки на торг

Кроме описания эмпирического и теоретического распределений цен, модель логарифмически нормального распределения позволяет решить задачу определения общего коэффициента капитализации R_o рынка (по районам, типам недвижимости, типам домов, типам квартир и т. д.). Если распределения цен независимы и логарифмически нормальны, то коэффициент капитализации R_o распределен логарифмически нормально [5] и может быть определен по формуле:

$$\ln R_o = \ln I_o - \ln V_o, \qquad (11)$$

где $\ln R_o$ – натуральный логарифм общего коэффициента капитализации;
$\ln I_o$ – натуральный логарифм чистого операционного дохода;
$\ln V_o$ – натуральный логарифм стоимости объекта.

Принимая модель логарифмически нормального распределения цен, можно исследовать распределения на непересекающихся подмножествах объектов недвижимости, таких как административные районы, количество комнат, типы домов и т.д. Если множество изучаемых объектов недвижимости можно разбить на непересекающиеся подмножества, на которых логарифмы цен распределены нормально, то и на объединении множеств можно, при определенных условиях, увидеть эмпирическое распределение, которое лучше приближается логарифмически нормальным распределением, чем нормальным.

Авторами статьи аналогичным образом получены подтверждения логарифмически нормальных распределений цен предложения на рынке жилой недвижимости, сгруппированных по административным районам, количеству комнат в квартире, типу домов.

Таким образом, проведенный анализ показал, что логарифмически нормальная модель точнее описывает реальное распределение цен на рынке предложения жилой недвижимости. Наивероятнейшее значение (мода) эмпирического распределения цен является более точной мерой рыночной стоимости. Выбор консультантами-оценщиками среднего значения цен предложения приводит к завышению стоимости, так как такой подход неявным образом использует нормальный закон, для которого математическое ожидание, мода, медиана совпадают. Кроме того, при отсутствии информации о реальных ценах сделок при использовании логарифмически нормальной модели может быть получена скидка на торг.

ЛИТЕРАТУРА И ИНФОРМАЦИОННЫЕ ИСТОЧНИКИ

1. *Грибовский С. В.* Математические методы оценки стоимости недвижимого имущества : учебое пособие. М. : Финансы и статистика, 2008.
2. Бюллетень недвижимости. 2013. № 1538.
3. *Кобзарь А. И.* Прикладная математическая статистика. М. : Физматлит, 2012.
4. База данных жилой недвижимости города Санкт-Петербурга. URL: на www.bn.ru
5. *Ласкин М. Б., Пупенцова С. В.* Использование коэффициента капитализации при изучении тенденций рынка недвижимости // Имущественные отношения в Российской Федерации. 2012. № 10 (133).