

С.В. Грибовский, проф., д-р экон. наук;  
Н.П. Баринов, с.н.с., канд. техн. наук;  
И.Н. Анисимова, канд. физ-мат. наук,  
г. Санкт-Петербург

# О повышении достоверности оценки рыночной стоимости методом сравнительного анализа

## ВВЕДЕНИЕ

Канонические определения рыночной стоимости (РС) объекта оценки [1—3] трактуют ее как *наиболее вероятную величину цены, по которой данный объект оценки может быть отчужден на открытом рынке в условиях конкуренции*, указывая тем самым на статистическую ее природу.

С точки зрения математической статистики стоимость как случайная величина рассчитывается на основе значений цен объектов-аналогов  $x_i, i = 1, \dots, n$ , понимаемых как  $n$  ее независимых наблюдений. В качестве генеральной совокупности выступают цены всех объектов на рассматриваемом сегменте рынка. Стоимость объекта оценки получают в результате обработки доступной оценщику выборки значений из генеральной совокупности.

В теории и практике оценки в качестве показателя РС чаще всего используют ее *математическое ожидание*, оценку которого получают расчетом выборочного среднего, сопровождая его, как правило, оценкой точности в виде границ доверительного интервала [4, 5]. Однако упомянутые выше определения РС говорят о *наиболее вероятном* значении, которому в общем случае соответствует другая статистика случайной величины — *мода*. И лишь для симметричных одномодальных распределений случайных величин, каким, в частности, является и нормальное распределение, значения математического ожидания и моды совпадают.

Статистические оценки вычисляют исходя из конкретного закона распределения случайной величины, что подразумевает его априорное знание или эмпирическое определение по данным выборки. Однако в подавляющем большинстве практических случаев индивидуальной оценки недвижимости объем выборки оказывается недостаточно для надежного построения эмпирических функций распределения. Это объясняется как сложностью проведения полномасштабных исследований рынка, так и, зачастую, отсутствием на нем необходимой информации. На практике оценщики часто используют соотношения, полученные для нормального закона распределения, по умолчанию считая справедливой гипотезу о подчинении ему экспериментальных данных.

Это может быть объяснено общеизвестностью и хорошей «разработанностью» нормального распределения, а также наличием у него ряда уникальных свойств, в частности, свойства «нормализации» распределения случайной величины при возрастании числа независимо влияющих на нее факторов. Кроме того, свойства симметричности и одномодальности нормального распределения снимают проблему несоответствия значений оцениваемого на практике математического ожидания и требуемой стандартами моды.

Вместе с тем процедура подбора оценщиком аналогов оцениваемого объекта не является, строго говоря, процедурой случайного выбора из генеральной совокупности и не может гарантировать однородность выборки. Нельзя также исключать возможность ошибок оценщика — случайных или преднамеренных — при формировании выборки рыночных данных о ценах объектов-аналогов. Принимая это во внимание, следует признать, что гипотезу нормальности распределения выборки рыночных данных нельзя рассматривать как принимаемую автоматически, а значит, необходимо проводить ее проверку.

Подтверждение гипотезы нормальности распределения выборочных данных о ценах аналогов требуется также и для корректного применения корреляционно-регрессионных методов при определении стоимости объекта оценки с учетом отличий его от аналогов по одному или нескольким влияющим признакам. Известно [6, 7], что наличие оптимальных свойств у метода наименьших квадратов, применяемого при построении регрессионных зависимостей, тесно связано с нормальностью распределения результирующего параметра (выборки рыночных цен) и отсутствием в выборке грубых погрешностей. Еще одним условием обеспечения корректности построенной регрессии является нормальность распределения погрешностей, которая должна быть проверена на заключительной стадии регрессионного анализа.

Отсутствие проверки на нормальность распределения исходных данных в условиях малого объема выборки — типичнейшая ситуация практической оценки в современных российских условиях — ставит под сомнение надежность итоговых результатов расчета стоимости и корректность полученных оценок точности.

## **ПРЕДВАРИТЕЛЬНАЯ ПРОВЕРКА МАЛОЙ ВЫБОРКИ РЫНОЧНЫХ ДАННЫХ**

На наш взгляд, расчет стоимости объекта оценки в общем случае должен предваряться получением информации о законе распределения исследуемой случайной величины. В частных случаях, когда имеется достаточно оснований считать генеральную совокупность нормально распределенной, возникает необходимость проверки гипотезы о нормальности распределения имеющейся выборки. В рамках этой процедуры выделяют два этапа: отсева грубых погрешностей — аномальных, резко выделяющихся данных, наличие которых может существенно ухудшить статистические оценки результирующего значения стоимости, и собственно проверку на нормальность распределения выборки, полученной в результате отсева.

### **1. ОТСЕВ ГРУБЫХ ПОГРЕШНОСТЕЙ**

Появление резко выделяющихся наблюдений может быть вызвано прямой ошибкой или существенным искажением стандартных условий сбора статистических данных, при котором однородность выборки нарушается. Чаще всего речь может идти о наличии в выборке рыночных данных, принадлежащих другому, как правило, смежному сегменту рынка (другому классу объектов, региону и т.п.).

Известен ряд методов отсева грубых погрешностей [7—10]. Рассмотрим некоторые из них для выборок объемом  $n \geq 3$ . Общим моментом для всех методов является построение вариационного ряда  $x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n)}$  по имеющейся выборке значений  $x_1, x_2, \dots, x_n$ . Затем вычисляют значения специфических для каждого метода статистик, которые сравнивают с соответствующими критическими значениями.

#### **1.1. КРИТЕРИЙ СМИРНОВА — ГРАББСА**

Этот критерий основан на вычислении максимального относительного отклонения  $\frac{|x_i - \bar{x}|}{s_n}$ , где  $x_i$  — крайний (минимальный  $x_{(1)}$  или максимальный  $x_{(n)}$ ) элемент вариационного ряда,

$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$  — среднее по выборке,  $s_n = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2}$  — выборочное среднее квадратическое отклонение (СКО). Для малых выборок в вычисляемую статистику вводят уточняющий множитель  $\frac{1}{\sqrt{(n-1)/n}}$  [8], в результате чего она примет вид  $T_{(1)} = \frac{\bar{x} - x_{(1)}}{s}$  для минимального и  $T_{(n)} = \frac{x_{(n)} - \bar{x}}{s}$  для максимальных значений вариационного ряда;  $s = \sqrt{s^2}$ ,  $s^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$  — дисперсия эмпирического распределения [7].

Значения  $T_{(1)}$  и  $T_{(n)}$  сравнивают с критическим значением  $C_\alpha$  метода Смирнова — Граббса. Выборка не содержит грубых погрешностей, если  $T_{(i)} \leq C_\alpha$ ,  $i = 1, n$ .

Табличные значения  $C_\alpha$  для уровней значимости  $\alpha = 0,10$  (10%),  $0,05$  (5%) и  $0,025$  (2,5%) и  $n \leq 26$  приведены в [7, с. 302; 9, с. 283]. Наблюдения делят на три группы в зависимости от вычисленных относительных отклонений  $T_{(i)}$ ,  $i=1; n$ :

1)  $T_{(i)} \leq C_{10\%}$  — наблюдение не нарушает однородность выборки и не отсеивается ни в коем случае;

2)  $T_{(i)} > C_{2,5\%}$  — наблюдение значимо отклоняется от  $\bar{x}$ , а значит, является грубой погрешностью и отсеивается как нетипичное;

3)  $C_{10\%} < T_{(i)} \leq C_{2,5\%}$  — требуются дополнительные аргументы в пользу отсева наблюдения.

Таковыми аргументами могут послужить соображения эксперта-оценщика, связанные с методологией получения данных для выборки, или существенный выигрыш в снижении ошибки оценки — СКО с учетом «штрафа за уменьшение объема выборки» [5], выражающейся в ширине доверительного интервала. Напомним, что в случае нормального распределения математическое ожидание со статистической надежностью  $p$  лежит между доверительными границами  $\bar{x} \pm t_{n-1;1-p} \frac{s_n}{\sqrt{n}}$ , где  $t_{n-1;1-p}$  — значение  $t$ -распределения Стьюдента для уровня значимости  $\alpha = 1 - p$  и числа степеней свободы  $n - 1$ , см. таблицы [5, с. 325; 7, с. 325].

В случае сомнений целесообразно проверить наблюдение с помощью других критериев и сравнить полученные результаты.

## 1.2. КРИТЕРИЙ ГРАББСА

Критерий Граббса основан на сравнении сумм квадратов отклонений от среднего исходной и сокращенной (без крайнего элемента) выборки.

По вариационному ряду  $x_{(1)}, x_{(2)}, \dots, x_{(n)}$  получают оценки  $\bar{x}$ ,  $\bar{x}_1$  и  $\underline{x}_1$ , где  $\bar{x}_1 = \frac{1}{n-1} \sum_{i=1}^{n-1} x_{(i)}$  — среднее по выборке без  $x_{(n)}$ ,  $\underline{x}_1 = \frac{1}{n-1} \sum_{i=2}^n x_{(i)}$  — среднее без  $x_{(1)}$ , и проверяют критерии  $G_1$  для минимального  $x_{(1)}$  и  $G_n$  для максимального  $x_{(n)}$  значений ряда:

$$G_n = \frac{\sum_{i=1}^{n-1} (x_{(i)} - \bar{x}_1)^2}{\sum_{i=1}^n (x_{(i)} - \bar{x})^2}; \quad G_1 = \frac{\sum_{i=2}^n (x_{(i)} - \underline{x}_1)^2}{\sum_{i=1}^n (x_{(i)} - \bar{x})^2}.$$

Если значения статистик  $G_1$  или  $G_n$  окажутся меньше критического значения  $C'_\alpha$ , то соответствующие наблюдения  $x_1$  или  $x_n$  относят к грубым ошибкам. Критические значения  $C'_\alpha$  для  $n \leq 25$  приведены в [7, с. 304]. Для критерия Граббса наблюдение должно быть отсеено, если окажется  $G_i < C'_{2,5\%}$  ( $i = 1; n$ ), и признано типичным, если  $G_i \geq C'_{10\%}$ .

После отсева наблюдений, признанных нетипичными, проверку на грубые ошибки повторяют для сокращенной выборки.

Оба рассмотренных критерия применимы для проверки на аномальность единичных наблюдений (минимального или максимального), однако в ситуации, когда выборка содержит группу близких по значениям аномальных наблюдений, они могут не дать результата. В этом случае можно использовать критерий Титъена — Мура, являющийся обобщением критерия Граббса на несколько наблюдений.

### 1.3. КРИТЕРИЙ ТИТЬЕНА — МУРА

Для  $k$  максимальных наблюдений выборки проверяется статистика

$$L_{(k)} = \frac{\sum_{i=1}^{n-k} (x_{(i)} - \bar{x}_k)^2}{\sum_{i=1}^n (x_{(i)} - \bar{x})^2}; \quad \bar{x}_k = \frac{\sum_{i=1}^{n-k} x_{(i)}}{n-k}.$$

Модификация критерия Граббса для  $k$  минимальных значений выборки проводится аналогичным образом:

$$\underline{L}_{(k)} = \frac{\sum_{i=k+1}^n (x_{(i)} - \underline{x}_k)^2}{\sum_{i=1}^n (x_{(i)} - \bar{x})^2}; \quad \underline{x}_k = \frac{\sum_{i=k+1}^n x_{(i)}}{n-k}.$$

Критические значения СП<sub>5%</sub> статистик  $L_{(k)}$  для  $n \leq 20$  и  $k = 1, 2$  и  $3$  приведены в [7, с. 307]. Критериальные неравенства для статистик те же, что и для критерия Граббса.

**Пример.** Пусть у оценщика, проводящего расчет рыночной стоимости производственного здания, имеется выборка из девяти значений стоимости единицы площади аналогичных зданий, полученная после приведения цен зданий-аналогов к условиям оцениваемого (табл. 1).

Таблица 1. Исходная выборка рыночных данных

Порядковый номер объекта	1	2	3	4	5	6	7	8	9	Выборочное среднее $\bar{x}$	СКО $s_{\bar{x}}$
Стоимость единицы площади $x$ , у.е./м <sup>2</sup>	58,4	64,5	58,0	32,5	57,6	58,2	60,4	63,9	49,8	55,92	9,76
95%-ный доверительный интервал: 55,92 ± 7,5											

Вариационный ряд значений удельных стоимостей: 32,5; 49,8; 57,6; 58,0; 58,2; 58,4; 60,4; 63,9; 64,5. Среди значений ряда имеется одно — 32,5 у.е./м<sup>2</sup> (объект 4), явно выделяющееся из общей группы. Рассмотрим результаты вычислений в электронных таблицах Microsoft Excel описанных выше статистик (табл. 2).

Таблица 2. Проверка на наличие грубых погрешностей исходной выборки

$n = 9$		Минимальное значение 32,5	Максимальное значение 64,5		
Критерий Смирнова — Граббса	Статистики	$T_{(1)} = 2,546$	$T_{(n)} = 0,932$		
	Критические значения	$C_{10\%} = 2,10$	Отвергается	Принимается	
		$C_{5\%} = 2,24$	Отвергается	Принимается	
		$C_{2,5\%} = 2,46$	Отвергается	Принимается	
Критерий Граббса	Статистики	$G_{(1)} = 0,190$	$G_{(n)} = 0,891$		
	Критические значения	$C_{10\%} = 0,450$	Отвергается	Принимается	
		$C_{5\%} = 0,374$	Отвергается	Принимается	
		$C_{2,5\%} = 0,310$	Отвергается	Принимается	
Критерий Титъена — Мура		Пара минимальных значений 32,5 и 49,8		Пара максимальных значений 64,5 и 63,9	
	Статистики	$L_{(2)} = 0,067$		$L_{(2)} = 0,768$	
	Критические значения	$C_{5\%} = 0,194$	Отвергается	Принимается	

Все критерии свидетельствуют в пользу того, чтобы минимальное значение 32,5 было удалено из выборки как нетипичное, а максимальное значение — оставлено. По данным исходной выборки выборочное среднее, моделирующее удельную стоимость оцениваемого объекта, с 95% -ной статистической надежностью попадает в доверительный интервал (с учетом поправки на малость выборки)  $55,92 \pm 7,5$ ; после отсева значения 32,5 доверительный интервал примет вид  $58,85 \pm 3,8$ . Как видно, исключение из выборки одного аномального значения позволило уточнить оценку стоимости на 5% и вдвое уменьшить ее погрешность.

Результаты проверки по критерию Титъена — Мура для двух минимальных значений (32,5 и 49,8) ставят под сомнение еще одно значение — 49,8 (объект 9). Поскольку эти два значения трудно рассматривать как близкие друг другу, проверим сомнительное двумя оставшимися критериями по данным сокращенной выборки (табл. 3). Вариационный ряд: 49,8; 57,6; 58,0; 58,2; 58,4; 60,4; 63,9; 64,5.

Таблица 3. Проверка на наличие грубых погрешностей сокращенной выборки

$n = 8$		<i>Минимальное значение 49,8</i>	<i>Максимальное значение 64,5</i>
Критерий Смирнова — Граббса	Статистики	$\mathcal{T}_{(1)} = 2,128$	$\mathcal{T}_{(n)} = 1,329$
	Критические значения	$\mathcal{C}_{10\%} = 2,04$	Отвергается
		$\mathcal{C}_{5\%} = 2,17$	Принимается
	$\mathcal{C}_{2,5\%} = 2,37$	Принимается	Принимается
Критерий Граббса	Статистики	$\mathcal{G}_{(1)} = 0,353$	$\mathcal{G}_{(n)} = 0,748$
	Критические значения	$\mathcal{C}_{10\%} = 0,405$	Отвергается
		$\mathcal{C}_{5\%} = 0,326$	Принимается
	$\mathcal{C}_{2,5\%} = 0,262$	Принимается	Принимается
Критерий Титъена — Мура	Статистики	<i>Пара минимальных значений 49,8 и 57,6</i>	<i>Пара максимальных значений 64,5 и 63,9</i>
	Критические значения	$\mathcal{L}_{(2)} = 0,301$	$\mathcal{L}_{(2)} = 0,471$
	$\mathcal{C}_{5\%} = 0,146$	Принимается	Принимается

Минимальное значение 49,8 может быть признано нетипичным по критериям Граббса и Смирнова — Граббса только с весьма жестким 10% -ным уровнем значимости (доверительной вероятностью  $q = 0,90$ ). Этого нельзя утверждать с доверительной вероятностью  $q = 0,95$  и  $q = 0,975$  (на более низких 5% и 2,5% -ном уровнях значимости). В этом случае возникают серьезные сомнения в необходимости исключения из выборки значения 49,8. Рассмотрим возможные последствия отсева рассматриваемого значения.

Оценка среднего значения изменится еще на 2%, а ширина доверительного интервала сузится почти в полтора раза ( $60,14 \pm 2,7$ ), однако теперь критерий Титъена — Мура отвергнет пару максимальных значений выборки (табл. 4).

Таблица 4. Проверка последствий исключения из выборки значения 49,8

$n = 7$		<i>Пара минимальных значений 57,6 и 58</i>	<i>Пара максимальных значений 64,5 и 63,9</i>
Критерий Титъена — Мура	Статистики	$\mathcal{L}_{(2)} = 0,697$	$\mathcal{L}_{(2)} = 0,093$
	Критические значения	$\mathcal{C}_{5\%} = 0,146$	Принимается
		Принимается	Отвергается

Критерии для проверки одиночных крайних значений (Граббса, Смирнова — Граббса) не выявят значение 64,5 как ошибочное. Это объясняется тем, что оно весьма близко к предыдущему значению 63,9; однако вместе они отстоят от других значений выборки. Критерий Титъена — Мура специально предназначен для выявления подобных ситуаций и в данном случае требует удалить пару максимальных значений 63,9 и 64,5 как ошибку. Если сделать это, в выборке из пяти элементов грубой погрешностью будет признано значение 60,4. В результате объем выборки потребу-

ется сократить до четырех элементов, т.е. более чем в два раза по сравнению с исходной (с 9 до 4), что весьма нежелательно, а в случае дальнейшего применения методов регрессионного анализа — практически недопустимо.

Поэтому при необходимости дальнейшего проведения регрессионного анализа значение 49,8 целесообразно оставить в выборке, если не нарушается гипотеза о нормальности распределения.

## 2. ПРОВЕРКА ГИПОТЕЗЫ НОРМАЛЬНОСТИ РАСПРЕДЕЛЕНИЯ. ПРЕОБРАЗОВАНИЕ РАСПРЕДЕЛЕНИЙ К НОРМАЛЬНОМУ

При наличии достаточно большого объема рыночных данных проверка гипотезы нормальности осуществляется с помощью критериев согласия, основанных на сравнении функций плотности эмпирического и теоретического нормального распределений. Применение таких критериев подразумевает разбиение данных выборки на классы, построение гистограмм или полигонов распределения частот, что требует объема выборки, превышающего 30—50 элементов. Наиболее известными в этом ряду являются  $\chi^2$ -критерий и критерий Колмогорова — Смирнова [10—12].

Выборки малого объема не могут предоставить достаточного количества информации для применения подобных критериев. В этом случае приходится ограничиться применением более простых и грубых критериев, используемых в качестве «прикидочных» при большом объеме выборочных данных. Они основаны на использовании коэффициента вариации, среднего абсолютного отклонения, размаха варьирования, показателей асимметрии и эксцесса [11, 12]. Более или менее уверенный вывод о подтверждении гипотезы нормальности на основе данных малой выборки можно сделать только при получении положительных результатах проверки несколькими критериями этой группы.

### 2.1. ПРОВЕРКА ПО КОЭФФИЦИЕНТУ ВАРИАЦИИ

Проверку гипотезы нормальности начинают с вычисления коэффициента вариации  $v = \frac{s_n}{\bar{x}}$ . Если его значение превышает 33%, то гипотеза о нормальности распределения данных выборки не подтверждается. Дальнейшую проверку при этом не проводят, так как такие распределения должны преобразовываться с целью уменьшения коэффициента вариации. При обратном соотношении проверку продолжают по оставшимся критериям.

Проверим выполнение неравенства  $v < 33\%$  для выборок, сформированных в примере. Значения коэффициента вариации  $v$  для очищенных от грубых ошибок выборок с объемом  $n = 8$  (без значения 32,5) и  $n = 4$  (без 32,5; 49,8 и 60,4; 63,9; 64,5) составляют 8 и 1% соответственно, следовательно, проверку можно продолжать для обеих выборок.

### 2.2. КРИТЕРИЙ СРЕДНЕГО АБСОЛЮТНОГО ОТКЛОНЕНИЯ (САО)

Для не очень больших выборок ( $n \leq 120$ ) используют показатель САО, определяемый как

$$\text{САО} = \frac{\sum_{i=1}^n |x_i - \bar{x}|}{n}.$$

Известно, что для теоретического нормального распределения отношение САО/ $\sigma$  ( $\sigma^2$  — дисперсия теоретического распределения) равно  $\sqrt{2/\pi}$ . Для выборки, имеющей приближенно нор-

мальный закон распределения, должно выполняться соотношение  $\left| \frac{\text{САО}}{s_n} - \sqrt{\frac{2}{\pi}} \right| < \frac{0,4}{\sqrt{n}}$ , или в численном виде  $\left| \frac{\text{САО}}{s_n} - 0,7979 \right| < \frac{0,4}{\sqrt{n}}$ .

В нашем примере для выборки из восьми значений неравенство принимает вид  $|3,063/4,55 - 0,7979| < 0,4/2,83$ , или  $0,125 < 0,141$  — неравенство справедливо и гипотеза о нормальности распределения принимается. Для уменьшенной до четырех значений выборки неравенство также выполняется:  $|0,25/0,34 - 0,7979| < 0,4/2$ , или  $0,063 < 0,2$ . Таким образом, гипотеза о нормальности распределения по критерию САО может быть принята по отношению к обеим сокращенным выборкам.

### 2.3. ПРОВЕРКА ПО РАЗМАХУ ВАРЬИРОВАНИЯ

Данная проверка может быть проведена для широкого класса выборок с объемом  $3 \leq n \leq 1000$ . Размах варьирования  $R$  рассчитывают как разность между наибольшим и наименьшим элементами в выборке, или в терминах построенного по выборке  $x_1, x_2, \dots, x_n$  вариационного ряда  $x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n)}$ :  $R = x_{(n)} - x_{(1)}$ .

Рассчитывают критериальное отношение  $R/s_n$ , которое сопоставляют с критическими значениями верхней и нижней границ для различных уровней значимости [11, с. 299]. Если рассчитанное отношение лежит в пределах границ, гипотезу о нормальности распределения принимают, если же оно меньше нижней границы или больше верхней — гипотезу отвергают. Для уверенного принятия решения о нормальности данных выборки важно, чтобы условие нахождения внутри границ выполнялось на жестком 10% -ном уровне значимости ( $p = 0,1$ ), однако, если гипотеза нормальности подтверждается всеми остальными критериями, можно ограничиться выполнением условия и на 5% -ном уровне значимости ( $p = 0,05$ ).

Вновь обратимся к рассмотренному примеру. Для выборок объемом 8 и 4 элемента значения соотношения  $R/s_n$  равны соответственно 3,234 и 2,342. Сопоставление с табличными данными показывает, что условие нахождения в границах для обеих выборок выполняется на уровне значимости 10% -ные (интервалы [2, с. 59; 3, с. 308] и [2, с. 94; 2, с. 409] соответственно). Значит, по критерию размаха варьирования гипотеза о нормальности распределения может быть уверенно принята для обеих сокращенных выборок.

### 2.4. ПРОВЕРКА С ПОМОЩЬЮ ПОКАЗАТЕЛЕЙ АСИММЕТРИИ И ЭКСЦЕССА

О близости эмпирического распределения нормальному можно также судить, используя показатели асимметрии  $g_1$  и эксцесса  $g_2$ , которые позволяют делать качественные выводы о форме эмпирического распределения и возможности отнесения его к типу кривых нормального распределения. Для теоретического нормального распределения эти показатели равны нулю.

Величины асимметрии и эксцесса для выборки могут быть рассчитаны с помощью следующих формул:

$$g_1 = \frac{\sqrt{n} \sum_{i=1}^n (x_i - \bar{x})^3}{\left( \sum_{i=1}^n (x_i - \bar{x})^2 \right)^{3/2}}; \quad g_2 = \frac{n \sum_{i=1}^n (x_i - \bar{x})^4}{\left( \sum_{i=1}^n (x_i - \bar{x})^2 \right)^2} - 3.$$

Несмещенные оценки  $G_1$  и  $G_2$  для  $g_1$  и  $g_2$  соответственно рассчитываются как

$$G_1 = \frac{\sqrt{n(n-1)}}{n-2} g_1; \quad G_2 = \frac{n-1}{(n-2)(n-3)} ((n+1)g_2 + 6),$$

а их среднеквадратические отклонения  $S_{G_1}$  и  $S_{G_2}$

$$S_{G_1} = \sqrt{\frac{6n(n-1)}{(n-2)(n+1)(n+3)}}; \quad S_{G_2} = \sqrt{\frac{24n(n-1)^2}{(n-3)(n-2)(n+3)(n+5)}}.$$

Если показатель  $g_1$  положителен, то наблюдается правосторонняя асимметрия и значение среднего  $\bar{x}$  превышает значение моды. Если  $g_1$  отрицательно, то имеет место левосторонняя асимметрия и  $\bar{x}$  меньше моды. Для больших выборок асимметрия считается значимой, если значение  $g_1$  по модулю превосходит 0,5; значения, меньшие 0,25, во внимание не принимаются [12]. Выполнение условия симметричности распределения особенно важно, так как в этом случае даже при отсутствии нормальности можно уверенно использовать среднее  $\bar{x}$  в качестве оценки наиболее вероятного значения случайной величины, требуемого определениями РС. Показатель эксцесса для симметричного распределения характеризует его крутизну по сравнению со стандартным нормальным распределением.

В рассматриваемом критерии гипотеза нормальности исследуемого распределения может быть принята, если совместно выполняются условия  $|G_1| \leq 3S_{G_1}$  и  $|G_2| \leq 5S_{G_2}$ . В противном случае гипотеза должна быть отвергнута.

В нашем примере для выборки из восьми элементов:  $|G_1| = 0,94$ ;  $S_{G_1} = 0,75$  и  $|G_2| = 1,99$ ;  $S_{G_2} = 1,48$  — неравенства  $|G_1| \leq 3S_{G_1}$  и  $|G_2| \leq 5S_{G_2}$  выполняются одновременно. Для выборки из четырех элемен-

тов:  $|G_1| = 0,75$ ;  $S_{G_1} = 1,01$  и  $|G_2| = 0,34$ ;  $S_{G_2} = 2,82$  — оба неравенства также выполнены. Таким образом, гипотеза о нормальности распределения может быть принята по данному критерию для обеих выборок.

Резюмируя проверку на наличие грубых ошибок и нормальность, можно отметить, что выборка из восьми элементов, содержащая «сомнительное» значение 49,8, может считаться удовлетворяющей высказанным требованиям. На практике в условиях «информационного голода» такие значения целесообразно оставлять, сохраняя максимально возможный объем выборки, особенно если в дальнейшем предполагается использование процедур регрессионного анализа.

При наличии достаточного объема рыночных данных для повышения точности и «защищенности» полученных результатов оценки сомнительные значения лучше исключать из рассмотрения, так как оценка РС в виде среднего арифметического весьма чувствительна к наличию грубых погрешностей.

В тех случаях, когда оценщик убедился, что гипотеза о нормальности распределения должна быть отвергнута, а выборка не может быть дополнена однородными данными, следует предпринять попытку преобразования исходных данных с тем, чтобы их распределение подчинялось нормальному закону.

Например, на практике достаточно широко встречаются логарифмически нормальные распределения, обусловленные влиянием на исследуемый параметр мультипликативно действующих факторов. В таких распределениях наблюдается правосторонняя асимметрия, их можно нормализовать путем преобразования вида  $x' = \log(x \pm a)$ ,  $x' = 1/x$  или  $x' = 1/\sqrt{x}$ . Для нормализации распределения с левосторонней асимметрией используют тригонометрические преобразования и степенные вида  $x' = x^\alpha$ , где  $\alpha$  принимает значения от 1,5 (при умеренной асимметрии) до 2 (при сильно выраженной асимметрии). Тип и выраженность асимметрии указываются показателем  $g_1$ .

После преобразования данных вновь проверяют гипотезу нормальности, по результатам которой и принимают решение о продолжении поиска подходящего преобразования. После успешного проведения всей обработки данных и получения окончательного результата проводят обратное преобразование переменных.

Отметим, что вычислительные процедуры, реализующие проверку выборок малого объема на наличие грубых погрешностей и нормальность, достаточно просты. Программу проверок можно получить по электронной почте, направив запрос по адресу [bosy.of@spb.cityline.ru](mailto:bosy.of@spb.cityline.ru)

## ВЫВОДЫ

1. Обеспечение корректности статистических оценок РС в методе сравнительного анализа возможно лишь при соблюдении ряда требований, предъявляемых к выборке исходных рыночных данных:

- использование среднего арифметического значения в качестве оценки РС, определяемой как наиболее вероятное значение цены, корректно, когда значения моды и математического ожидания совпадают. Распределение выборки исходных данных при этом должно удовлетворять, по крайней мере, требованиям симметричности;
- сопровождающая значение среднего оценка точности его получения в виде доверительного интервала должна формироваться исходя из конкретного закона распределения выборочных данных. Использование известных соотношений нормального распределения допустимо в том случае, если по данным выборки гипотеза нормальности не отвергается. Кроме того, ввиду высокой чувствительности значения среднего к грубым погрешностям следует очистить от них выборку;
- при оценке РС с помощью более развитых по сравнению с оценкой среднего методов регрессионного анализа следует обеспечить отсутствие в выборке грубых погрешностей и нормальность ее распределения при максимально возможном объеме выборки.

2. Гипотезы симметричности, нормальности распределения выборки и наличия в выборке грубых погрешностей могут быть проверены с помощью критериев, действенных в условиях малого объема рыночных данных. Вычислительные процедуры, лежащие в основе этих критериев, просты и легко реализуемы на ЭВМ, в том числе в среде MS Excel.

3. В случае невыполнения предъявляемых требований выборка должна быть сокращена, дополнена однородными данными или преобразована в зависимости от применяемой в качестве оценки РС статистики и доступной оценщику информации.

## Литература

1. Федеральный закон «Об оценочной деятельности в Российской Федерации» от 29.07.98г. №35-ФЗ.
2. Постановление Правительства Российской Федерации «Об утверждении стандартов оценки» от 06.07.2001г. №519.
3. Международные стандарты оценки. Кн. 1: Перевод, комментарии, дополнения / Г.И. Микерин, М.И. Недужий, Н.В. Павлов, Н.Н. Яшина. М.: ОАО «Типография «НОВОСТИ», 2000.
4. Ковалев А.П. Оценка стоимости активной части основных фондов: Учебно-методическое пособие. М.: Финстатинформ, 1997.
5. Грибовский С.В. Оценка доходной недвижимости. СПб: Питер, 2001. (Сер. «Учебники для вузов»).
6. Линник Ю. В. Метод наименьших квадратов и основы теории обработки наблюдений. М.: Физматгиз, 1962.
7. Дубров А.М., Мхитарян В.С., Трошин Л.И. Многомерные статистические методы: Учебник. М.: Финансы и статистика, 2000.
8. Кассандрова О.Н., Лебедев В.В. Обработка результатов наблюдений. М.: Наука, 1970.
9. Пустыльник Е.И. Статистические методы анализа и обработка наблюдений. М.: Наука, 1968.
10. Айвазян С.А. Прикладная статистика: Основы моделирования и первичная обработка данных. Справ. изд. М.: Финансы и статистика, 1983.
11. Закс Л. Статистическое оценивание. М.: Статистика, 1976.
12. Теория статистики: Учебник / Под ред. Р.А. Шмойловой. М.: Финансы и статистика, 1999.