ПРОБЛЕМА ГЕНЕРИРОВАНИЯ СЛУЧАЙНЫХ ВРЕМЕННЫХ РЯДОВ, ОБЛАДАЮЩИХ АВТОКОРРЕЛЯЦИОННЫМИ СВОЙСТВАМИ, ПРИ РЕШЕНИИ ЗАДАЧ ИМИТАЦИОННОГО МОДЕЛИРОВАНИЯ

Е. П. Бочаров, О. Н. Алексенцева, С. В. Кузнецов (Саратов)

При необходимости учета в процессе имитационного моделирования каких-либо случайных факторов прибегают к процедурам генерирования случайных величин. Обычно внимание исследователей концентрируется на соответствии функции распределения плотности вероятности значений случайных величин некоторой искомой, определенной заранее, например, из эксперимента. При этом очень часто исследователи просто не обращают внимания на необходимость учитывать также свойства автокорреляции случайных величин, которая нередко присуща реальным задачам.

В качестве примера на рис. 1 приведен временной ряд $V(t_i)$ остатков на счетах «до востребования» (сумма остатков на расчетных счетах юридических лиц, остатков на карточных зарплатных счетах, депозитов «до востребования» физических и юридических лиц, корреспондентские счета других банков и т.п.) одного из коммерческих банков Поволжского региона ($i = 1,\ldots,N$ – номер банковского дня).

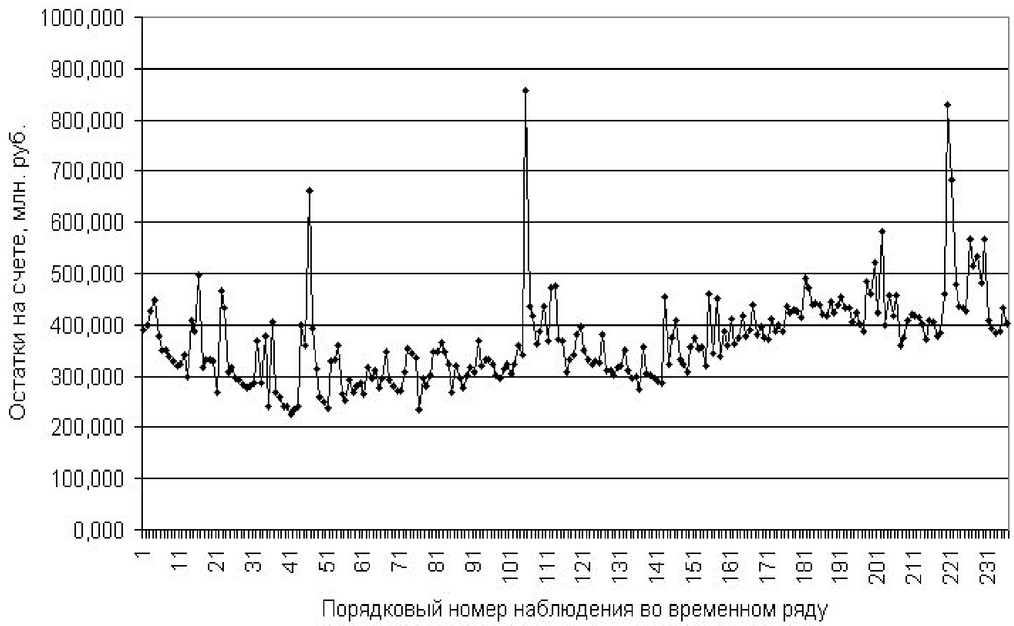

Рис 1. Временной ряд $V(t_i)$

Такие временные ряды необходимо учитывать в имитационных моделях коммерческих банков, поскольку остатки на счетах «до востребования» обеспечивают до 60%–70% мгновенной ликвидности малых и средних коммерческих банков.

Исследователи чаще всего рассматривают различные *параметрические семейства* распределений числовых случайных величин. Все такие распределения зависят от одного, двух или трех параметров (нормальное, логнормальное, экспоненциальное, гамма-распределение и ряд других). Это означает, что для полного описания распределения достаточно оценить одно, два или три числа. Поэтому широко развита параметрическая теория математической статистики, в которой предполагается, что распределения результатов наблюдений принадлежат тем или иным параметрическим семействам.

В реальных задачах параметрические распределения встречаются редко. Например, попытка «подгонки» ряда $V(t_i)$ к логнормальному распределению по критерию χ-квадрат (рис. 2) дала следующий результат: вероятность p ошибиться, если отвергнуть гипотезу о соответствии ряда логнормальному распределению составляет 0,3216. Попытки «подгонки» к другим видам параметрических распределений дали еще более худшие результаты.

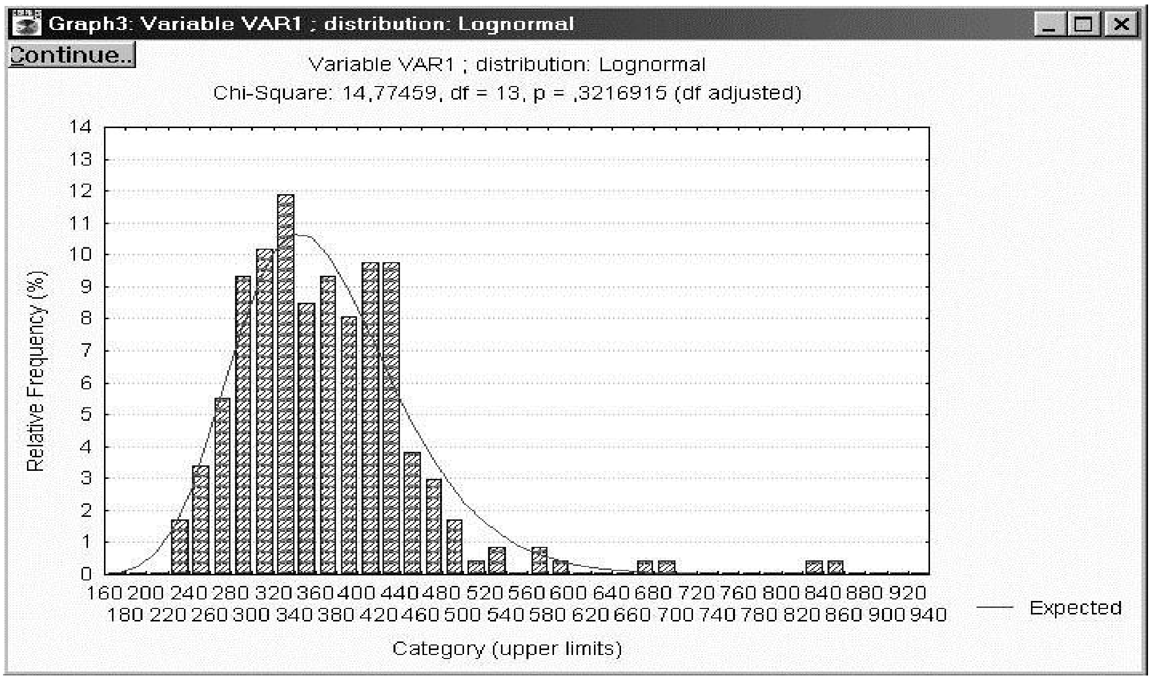

Рис. 2. Гистограмма ряда $V(t_i)$ и результаты попытки «подгонки» к логнормальному распределению

Таким образом, ряд $V(t_i)$ далек от какого-либо параметрического семейства распределений и к тому же обладает заметными автокорреляционными свойствами. Так, коэффициенты автокорреляции для первых трех временных лагов равны r_1= 0,566, r_2 = 0,497, r_3 =0,329.

В то же время развитые методы генерирования случайных коррелированных и автокоррелированных рядов разработаны в рамках параметрического подхода [1, 2].

Наиболее тесная автокорреляционная связь имеет место при лаге = 1. Если учитывать только эту ближнюю автокорреляционную связь, то соотношение для расчета сгенерированного ряда $W(t_i)$, имитирующего реальный ряд $V(t_i)$, в рамках параметрического подхода запишется в виде [1, 2]:

$$W(t_i) = \overline{V} + r_1(W(t_{i-1}) - \overline{V}) + \sqrt{1 - r_1^2}\,\sigma\,\eta\,, \quad (1)$$

где $\overline{V}$ – среднее значение временного ряда $V(t_i)$; r_1 – коэффициент автокорреляции членов ряда $V(t_i)$ для лага = 1; σ – среднеквадратичное отклонение значений ряда $V(t_i)$; η – случайное число; распределенное по нормальному закону ($\overline{\eta} = 0, \sigma_\eta = 1$).

Обозначим реальный ряд $V(t_i)$ как R-ряд, а ряд $W(t_i)$, сгенерированный по (1), – η-ряд. Чтобы использовать этот ряд в имитационной модели, необходимо показать, что R-ряд и η-ряд принадлежат одной генеральной совокупности. Поскольку мы имеем дело с рядами, распределение значений которых далеко от параметрических, следует использовать непараметрические критерии. Нами использовались **Sign test** и считающийся более мощным **Wilcoxon test** [3]. В табл. 1 приведены значения ***p*** **–** вероятности ошибиться при отвержении гипотезы «два ряда принадлежат к одной генеральной совокупности».

Таблица 1

Результаты применения тестов непараметрической статистики

	R-ряд, **Sign test**	R-ряд, **Wilcoxon test**	η-ряд, **Sign test**	η-ряд, **Wilcoxon test**	ξ-ряд, **Sign test**	ξ-ряд, **Wilcoxon test**
R-ряд			0,031	0,109	0,744	0,962
η-ряд	0,031	0,109			0,117	0,129
ξ-ряд	0,744	0,962	0,117	0,129		

Как видно из этой таблицы, η-ряд неудовлетворительно имитирует свойства реального ряда. Поэтому авторы решили полностью отказаться от парадигмы параметрической статистики, в рамках которой получено соотношение (1), и заменить в нем ряд нормально распределенных случайных чисел η на ряд случайных чисел, распределенных по тому же эмпирическому закону, что и R-ряд. Обозначим случайное число, распределенное по тому же закону, что и R-ряд, через ξ, а ряд, который получится в результате применения соотношения (1), в котором η заменено на ξ, через ξ-ряд. Как видно из табл. 1, ξ-ряд неплохо прошел испытания двумя тестами непараметрической статистики.

Аналогичные результаты получены с использованием других тестов непараметрической статистики: тест Колмогорова-Смирнова, тест Friedman and Kendall, тест Wald-Wolfowitz, тест Mann-Whitney, тест Kruskal-Wallis.

Как отмечается в [3], «к сожалению, параметрические семейства существуют лишь в головах авторов учебников по теории вероятностей и математической статистике». И это утверждение иллюстрируется в [3] многочисленными примерами из практики статистического анализа. Однако нельзя не отметить, что непараметрическая статистика еще далека от того, чтобы считаться строгой математической теорией. Так, многие выводы о «мощности непараметрических критериев» носят полуэмпирический характер даже в тех случаях, когда объемы выборок велики даже по меркам параметрической статистики.

Результаты исследования ξ-ряда представлены на рис. 3 и 4, а также в табл. 2.

Из сравнения рис. 1 и 3, рис. 2 и 4 (здесь следует обратить внимание на сравнение степени близости распределений значений R-ряда и ξ-ряда к логнормальному распределению), а также данных табл. 2, можно сделать вывод о достаточно удовлетворительной имитации сгенерированным по предложенному алгоритму ξ-рядом реального R-ряда.

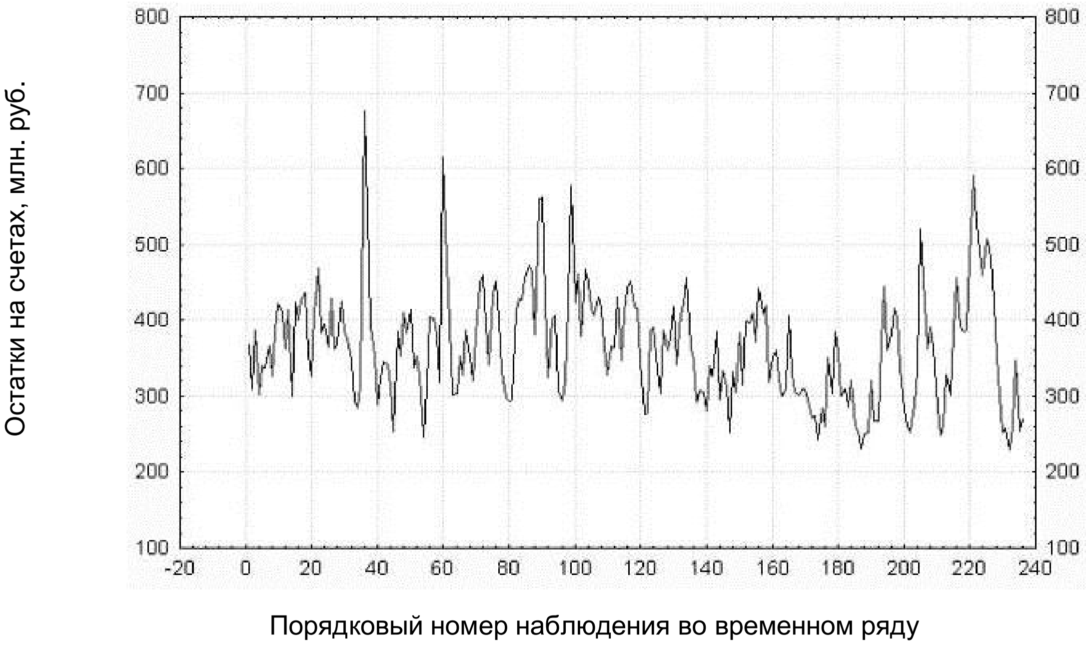

Рис. 3. Сгенерированный ξ-ряд

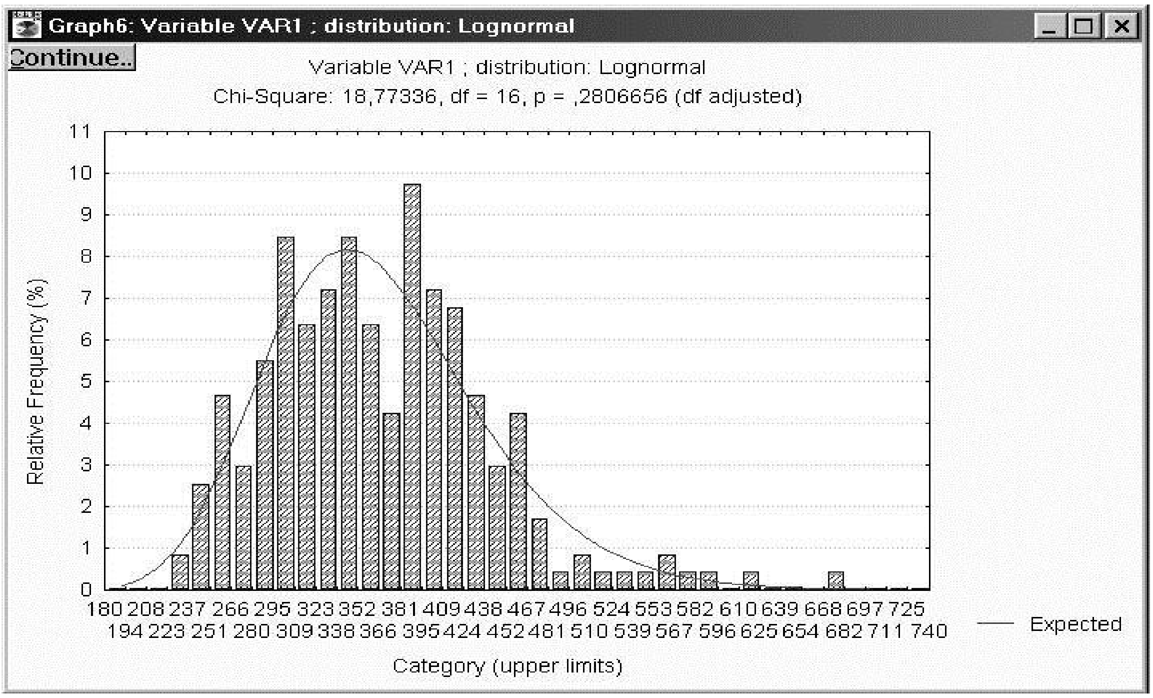

Рис. 4. Гистограмма ξ-ряда и результаты попытки «подгонки» к логнормальному распределению

Таблица 2

Сравнение параметров реального R-ряда и сгенерированного ξ -ряда

	Среднее значение, млн. руб.	Средн. квадр. откл., σ , млн. руб.	Коэф. автокорреляции $\mathbf{r_1}$	Коэф. автокорреля-ции $\mathbf{r_2}$	Коэф. автокорреля-ции $\mathbf{r_3}$
R-ряд	368,484	86,113	0,566	0,497	0,329
ξ -ряд	365,948	74,729	0,589	0,330	0,177

Литература

1. **Аверилл М. Лоу, Кельтон В. Дэвид.** Имитационное моделирование. СПб.: Питер, 2004. 846 с.
2. **Рыжиков Ю.И.** Имитационное моделирование. Теория и технологии. СПб.: Корона-принт, 2004. 380 с.
3. **Орлов А. И.** Эконометрика. М.: Экзамен, 2002. 576 с.