Український математичний вісник
Том 11 **(2014)**, № 4, 497 – 507

Экспоненциальные статистические эксперименты с настойчивой линейной регрессией с равновесием

Дмитрий В. Королюк

(*Представлена С. Я. Махно*)

Аннотация. Изучается асимптотическое поведение экспоненциальных статистических экспериментов, задаваемых экспоненциальной симметрической статистикой. При этом предполагается свойство настойчивой линейной регрессии, означающее зависимость отдельных результатов эксперимента от усреднения предыдущих экспериментов. Для экспоненциальных статистических экспериментов в схеме серий (при $N \to \infty$) установлен равновесный режим, а также построена аппроксимация экспоненциальным процессом нормальной авторегрессии.

2010 MSC. 40N30, 60J70, 62F05.

Ключевые слова и фразы. Экспоненциальные статистические эксперименты, настойчивая линейная регрессия, равновесный режим, стохастическая аппроксимация, экспоненциальный процесс нормальной авторегрессии.

Введение

В разнообразных задачах биологии и экономики возникают проблемы статистического анализа результатов наблюдений выборок, заданных совокупностью статистически независимых и одинаково распределённых случайных величин. При этом во многих динамических совокупностях (финансовые инструменты [4], взаимодействие молекул [5–8] и т.п.) наблюдается устойчивая зависимость динамики совокупности от усреднённых характеристик модели. Наиболее естественным и простым служит предположение о наличии условия

Статья поступила в редакцию 12.02.2013

ISSN 1810 – 3200.

"настойчивой регрессии", что означает зависимость усреднённых характеристик с постоянной функцией регрессии. В нашей предыдущей работе [1] исследовано асимптотическое поведение линейных статистических экспериментов, задаваемых суммами отдельных результатов испытаний при неограниченном возрастании объёма выборки $N \to \infty$.

В настоящей работе аналогичная проблема изучается для *экспоненциальных статистических экспериментов* (ЭСЭ), задаваемых экспоненциальной симметрической статистикой, выбор которой обосновывается широко известными фундаментальными результатами по стохастической финансовой математике (например, *(B,S)*-рынки [4, гл. 7, §11]). Устанавливаются условия равновесного режима (п. 3, теорема 3.1) и строится аппроксимация ЭСЭ (при $N \to \infty$) нормальным процессом авторегрессии (п. 4, предложение 4.1), в которой используется экспоненциальный процесс нормальной авторегрессии (п. 4, теорема 4.1). В первых двух пунктах излагаются (в конспективном виде) результаты асимптотического анализа линейных *статистических экспериментов* (СЭ) (теоремы 2.1, 2.2 и предложение 2.1).

1. Задание статистических экспериментов

Рассматриваются бинарные выборки $\delta(k) := (\delta_r(k),\ 1 \le r \le N)$, независимых в совокупности и одинаково распределённых при каждом $k \ge 0$ случайных величин $\delta_r(k)$, принимающих два значения ± 1.

Статистические Эксперименты (СЭ) задаются суммами выборочных значений

$$S_N(k) := \frac{1}{N} \sum_{r=1}^{N} \delta_r(k), \quad k \ge 0. \tag{1.1}$$

Настойчивая линейная регрессия с равновесием означает, что

$$E[S_N(k+1)|S_N(k) = s] = C(s), \quad k \ge 0. \tag{1.2}$$

в котором функция регрессии $C(s)$ имеет вид:

$$C(s) = (1-a)s + a\rho, \quad |s| \le 1. \tag{1.3}$$

Более того, предполагается, что случайные величины выборки $(\delta_r(k),\ 1 \le r \le N)$ условно независимы в совокупности при каждом $k \ge 0$: для любой борелевской функции $\varphi(s),\ |s| \le 1$ имеет место соотношение

$$E\left[\prod_{r=1}^{N}\varphi(\delta_r(k+1))|S_N(k)=s\right]=\prod_{r=1}^{N}E[\varphi(\delta_r(k+1))|S_N(k)=s]. \quad (1.4)$$

Направляющий параметр a и точка равновесия ρ удовлетворяют следующим соотношениям:

$$0<a<1, \quad |\rho|<1.$$

Замечание 1.1. Задание СЭ с помощью функции регрессии (1.2)–(1.4) означает, что вероятности выборочных значений определяются формулами:

$$P\{\delta_r(k+1)=\pm 1|S_N(k)=s\}=\frac{1}{2}[1\pm C(s)]. \quad (1.5)$$

Функция регрессии (1.3) преобразуется к следующему виду:

$$\begin{aligned} C(s)&=C_0(s)+s, \quad C_0(s)=-a(s-\rho),\\ C(s)&=C_\rho(s)+\rho, \quad C_\rho(s)=(1-a)(s-\rho). \end{aligned} \quad (1.6)$$

Заметим (и это будет использоваться в дальнейшем), что точка равновесия ρ удовлетворяет соотношениям

$$C_0(\rho)=C_\rho(\rho)=0. \quad (1.7)$$

Кроме того, задание СЭ (1.1)–(1.3) обеспечивает явный вид условной дисперсии

$$D[S_N(k+1)|S_N(k)=s]=\frac{1}{N}B(s), \quad B(s):=1-C^2(s), \quad k\geq 0. \quad (1.8)$$

Экспоненциальные Статистические Эксперименты (ЭСЭ) задаются симметрической экспоненциальной статистикой [2]

$$\Pi_N(\lambda,k+1):=\prod_{r=1}^{N}[1+\lambda\delta_r(k+1)]. \quad (1.9)$$

Легко вычисляется условное математическое ожидание (см. (1.4)):

$$\overline{\Pi_N}(\lambda,k):=E[\Pi_N(\lambda,k+1)|S_N(k)]=[1+\lambda C(S_N(k))]^N, \quad (1.10)$$

с функцией регрессии $C(s)$, определённой в (3).

Замечание 1.2. Так же, как и в работе [1], произвольные бинарные выборки со значениями $h_\pm$ могут быть сведены линейным преобразованием к симметрическим бинарным значениям ± 1.

2. Линейные статистические эксперименты

В дальнейшем нам понадобятся результаты усреднения и нормальной аппроксимации СЭ (1.1), полученные в работе [1], сформулированные для стандартных бинарных выборочных значений ± 1. Условная дисперсия СЭ задаётся формулой (1.8).

Установившийся режим СЭ реализуется следующим предельным соотношением

Теорема 2.1 (ср. [1, теорема 1]). *В условиях сходимости начальных условий:*

$$S_N(0) \xrightarrow{п.н.} \rho, \quad N \to \infty,$$

с вероятностью 1, *имеет место сходимость статистических экспериментов с вероятностью* 1*:*

$$S_N(k) \xrightarrow{п.н.} \rho, \quad N \to \infty,$$

при каждом конечном $k > 0$.

В работе [1] предложена аппроксимация СЭ (1.1) процессом нормальной авторегрессии.

Предложение 2.1 (ср. [1, предложение 1]). *Процесс нормальной авторегрессии в дискретном времени* $k \geq 0$, *задаётся соотношением*

$$\widetilde{S_N}(k+1) = C(\widetilde{S_N}(k)) + (\sigma/\sqrt{N})W(k+1), \quad k \geq 0, \tag{2.1}$$

в котором предельная дисперсия

$$\sigma^2 = 1 - \rho^2, \tag{2.2}$$

а $W(k+1)$, $k \geq 0$ *— нормально распределённые стандартные случайные величины, независимые при разных* $k \geq 1$.

Аппроксимация СЭ (1.1) процессом нормальной авторегрессии (1.6)–(1.7) обоснована следующей теоремой:

Теорема 2.2 (ср. [1, теорема 2]). *При выполнении условия теоремы* 2.1 *имеет место сходимость по распределению*

$$\sqrt{N}[S_N(k+1) - C(S_N(k))] \xrightarrow{d} \sigma W(k+1), \quad N \to \infty,$$

при каждом конечном $k \geq 0$.

Замечание 2.1. Очевидно, что процесс нормальной авторегрессии (2.1) сохраняет свойство настойчивой линейной регрессии

$$E[\widetilde{S_N}(k+1)|\widetilde{S_N}(k)] = C(\widetilde{S_N}(k)), \quad k \geq 0,$$

с той же функцией регрессии (1.2), что и исходный СЭ (1.1)–(1.3).

3. Установившийся режим ЭСЭ

Как и в предыдущей работе [1], ЭСЭ (1.2) в схеме усреднения рассматривается с $\lambda_N = \lambda/N$. Так что ЭСЭ в схеме усреднения задаётся формулой

$$\Pi_N(\lambda/N, k+1) = \prod_{r=1}^{N}[1 + \lambda\delta_r(k+1)/N], \quad k \geq 0. \tag{3.1}$$

Его условное среднее значение имеет вид (ср. (1.3)):

$$\overline{\Pi_N}(\lambda/N, k) = [1 + \lambda C(S_N(k))/N]^N, \quad k \geq 0. \tag{3.2}$$

Введём экспоненциальный мартингал

$$\mu_N^e(\lambda/N, k+1) = \Pi_N(\lambda/N, k+1)/\overline{\Pi_N}(\lambda/N, k), \quad k \geq 0.$$

Очевидно его мартингальное свойство:

$$E[\mu_N^e(\lambda/N, k+1)|S_N(k)] = 1, \quad k \geq 0.$$

Установившийся режим СЭ устанавливает следующая

Теорема 3.1. *При условии сходимости с вероятностью* 1 *начальных значений СЭ*

$$S_N(0) \xrightarrow{п.н.} \rho, \quad N \to \infty, \tag{3.3}$$

имеет место сходимость по вероятности ЭСЭ (1.9)

$$\Pi_N(\lambda/N, k) \xrightarrow{P} \exp(\lambda\rho), \quad N \to \infty,\ k \geq 0, \tag{3.4}$$

и сходимость по вероятности его среднего значения (1.10)

$$\overline{\Pi_N}(\lambda/N, k) \xrightarrow{P} \exp(\lambda\rho), \quad N \to \infty,\ k \geq 0. \tag{3.5}$$

Следствие 3.1. *При условии* (3.3)

$$\mu_N^e(\lambda/N, k+1) \xrightarrow{P} 1, \quad N \to \infty. \tag{3.6}$$

Доказательство теоремы 3.1. Воспользуемся формулой аппроксимации типа Ле Кама в следующей форме:

Лемма 3.1 (ср. [3, Lemma 6.3.1])**.** *Пусть выполняется сходимость по вероятности*

$$\max_{1\leq r\leq N} |\delta_r(k+1)/N| \xrightarrow{P} 0, \quad N \to \infty,\ \forall\, k \geq 0. \tag{3.7}$$

Тогда имеет место сходимость по вероятности

$$\sum_{r=1}^{N} \ln[1 + \lambda\delta_r(k+1)/N] - \lambda S_N(k+1) \xrightarrow{P} 0, \quad N \to \infty. \tag{3.8}$$

Здесь

$$S_N(k+1) := \frac{1}{N}\sum_{r=1}^{N} \delta_r(k+1), \quad N \to \infty.$$

В нашем случае условие (3.3) теоремы 3.1 выполнено. Согласно теореме 2.1 работы [1]

$$S_N(k+1) \xrightarrow{\text{п.н.}} \rho, \quad N \to \infty. \tag{3.9}$$

Перепишем сходимость (3.4) с учётом очевидного тождества $\Pi = \exp\ln\Pi$:

$$\exp\left\{\sum_{r=1}^{N} \ln[1 + \lambda\delta_r(k+1)/N] - \lambda S_N(k+1)\right\} \xrightarrow{P} 1, \quad N \to \infty.$$

Или иначе, с учётом сходимости (3.9), имеется предел по вероятности

$$P \cdot \lim_{N\to\infty} \Pi_N(\lambda/N, k+1)$$
$$= P \cdot \lim_{N\to\infty} \exp \sum_{r=1}^{N} \ln[1 + \lambda\delta_r(k+1)/N] = \exp(\lambda\rho).$$

Последняя сходимость совпадает с утверждением (3.4) теоремы 3.1. Аналогичным образом доказывается сходимость (3.5) для среднего значения (3.2). Сходимость (3.6) следует из (3.4) и (3.5). Теорема 3.1 доказана. $\square$

4. Аппроксимация процессом экспоненциальной нормальной авторегрессии

Так же, как и в предыдущей работе [1] ЭСЭ рассматриваются в схеме серий с параметром серии $\lambda_N = \lambda/\sqrt{N}$. Так что

$$\Pi_N(\lambda/\sqrt{N}, k+1) = \prod_{r=1}^{N} [1 + \lambda\delta_r(k+1)/\sqrt{N}], \quad k \geq 0. \tag{4.1}$$

Кроме того, среднее значение ЭСЭ (3.2) (ср. (1.10))

$$\overline{\Pi_N}(\lambda/\sqrt{N}, k) = [1 + \lambda C(S_N(k))/\sqrt{N}]^N, \quad k \geq 0. \tag{4.2}$$

Введём экспоненциальный мартингал

$$\mu_N^e(\lambda_N, k+1) := \Pi_N(\lambda/\sqrt{N}, k+1)/\overline{\Pi_N}(\lambda/\sqrt{N}, k), \quad k \geq 0. \tag{4.3}$$

Очевидно мартингальное свойство:

$$E[\mu_N^e(\lambda_N, k+1)|S_N(k)] = 1, \quad k \geq 0.$$

Теорема 4.1 (аппроксимация ЭСЭ)**.** *В условиях теоремы* 3.1 *имеет место сходимость по распределению*

$$\mu_N^e(\lambda/\sqrt{N}, k+1) \xrightarrow{d} \exp[\lambda\sigma W(k+1) - \lambda^2\sigma^2/2], \quad N \to \infty,\ k \geq 0, \tag{4.4}$$

$$\sigma^2 = 1 - \rho^2. \tag{4.5}$$

Замечание 4.1. Экспоненциальный мартингал (4.3), заданный ЭСЭ (4.1)–(4.2) в схеме серий, сходится по распределению, при $N \to \infty$, к экспоненциальному процессу нормальной авторегрессии (4.3). При этом, очевидно,

$$E \exp[\lambda\sigma W(k+1) - \lambda^2\sigma^2/2] = 1, \quad k \geq 0.$$

Доказательство теоремы 4.1. Так же, как и при доказательстве теоремы 3.1, воспользуемся тождеством $\Pi = \exp\ln\Pi$, а также леммой аппроксимации Ле Кама

Лемма 4.1 (аппроксимация Ле Кама [3, Lemma 6.3.1])**.** *Пусть выполняется сходимость по вероятности*

$$\max_{1\leq r\leq N} |\delta_r(k+1)/\sqrt{N}| \xrightarrow{P} 0, \quad N \to \infty,$$

а также суммы

$$V_N(k+1) := \frac{1}{N}\sum_{r=1}^{N}(\delta_r(k+1))^2,$$

ограничены по вероятности. Тогда имеет место сходимость по вероятности

$$\sum_{r=1}^{N} \ln[1+\lambda\delta_r(k+1)/\sqrt{N}] - \lambda\sqrt{N}S_N(k+1) + \lambda^2 V_N(k+1)/2 \xrightarrow{P} 0,\ N \to \infty.$$

Следовательно имеет место сходимость по вероятности с дополнительным множителем:

$$\Pi_N(\lambda/\sqrt{N}, k+1) \exp[-\lambda\sqrt{N} S_N(k+1) + \lambda^2 V_N(k+1)/2] \xrightarrow{P} 1,$$
$$N \to \infty, \quad k \geq 0. \quad (4.6)$$

Заметим, что в данном случае

$$V_N(k+1) = \frac{1}{N} \sum_{r=1}^{N} (\delta_r(k+1))^2 = 1.$$

Так что сходимость (4.6) эквивалентна сходимости

$$\Pi_N(\lambda/\sqrt{N}, k+1) \exp[-\lambda\sqrt{N} S_N(k+1)] \xrightarrow{P} \exp[-\lambda^2/2],$$
$$N \to \infty, \ k \geq 0. \quad (4.7)$$

Аналогично условия леммы 4.1 позволяют получить сходимость по вероятности (при $N \to \infty$) для среднего значения (4.2):

$$\overline{\Pi_N}(\lambda/\sqrt{N}, k) \exp[-\lambda\sqrt{N} C(S_N(k))] \xrightarrow{P} \exp[-\lambda^2 \rho^2/2],$$
$$N \to \infty, \ k \geq 0. \quad (4.8)$$

Здесь мы воспользовались сходимостью согласно теоремы 1 работы [1]:

$$C(S_N(k)) \xrightarrow{\text{п.н.}} C(\rho) = \rho, \quad N \to \infty, \ k \geq 0.$$

Так что

$$C^2(S_N(k)) \xrightarrow{\text{п.н.}} \rho^2, \quad N \to \infty, \ k \geq 0.$$

Исходный экспоненциальный мартингал (4.3) может быть представлен следующим образом:

$$\mu_N^e(\lambda/\sqrt{N}, k+1)$$
$$= [\Pi_N^0(\lambda/\sqrt{N}, k+1)/\overline{\Pi_N^0}(\lambda/\sqrt{N}, k)] \exp\{\lambda\sqrt{N}[S_N(k+1) - C(S_N(k))]\},$$
$$k \geq 0. \quad (4.9)$$

Здесь, по определению,

$$\Pi_N^0(\lambda/\sqrt{N}, k+1) := \Pi_N(\lambda/\sqrt{N}, k+1) \exp[-\lambda\sqrt{N} S_N(k+1)], \quad k \geq 0,$$
$$\overline{\Pi_N^0}(\lambda/\sqrt{N}, k) := \overline{\Pi_N}(\lambda/\sqrt{N}, k) \exp[-\lambda\sqrt{N} C(S_N(k))], \quad k \geq 0. \quad (4.10)$$

По теореме 2 работы [1] имеет место сходимость по распределению:

$$\sqrt{N}[S_N(k+1) - C(S_N(k))] \xrightarrow{d} \sigma W(k+1), \quad N \to \infty, \ k \geq 0. \quad (4.11)$$

Так что сходимости (4.7), (4.8) и (4.11) обеспечивают сходимость экспоненциального мартингала (4.4). Теорема доказана. □

Предельная теорема 4.1, а также выражение (4.3) для экспоненциального мартингала дают возможность представить ЭСЭ (3.1) в схеме серий в следующем виде:

$$\Pi_N(\lambda/N, k+1) = \overline{\Pi_N}(\lambda/N, k)\mu_N^e(\lambda/N, k+1). \tag{4.12}$$

Согласно (4.10), представление ЭСЭ (4.11) преобразуется к виду

$$\Pi_N(\lambda/N, k+1) = \overline{\Pi_N^0}(\lambda/N, k)\exp[\lambda C(S_N(k))]\mu_N^e(\lambda/N, k+1). \tag{4.13}$$

Теперь воспользуемся асимптотическими соотношениями для первого и третьего сомножителей в (4.13). Всюду далее в формулах остаточный член $R_N \to 0$ при $N \to \infty$.

$$\overline{\Pi_N^0}(\lambda/N, k) = \exp[-\lambda^2\rho^2/2N]e^{R_N},$$

а также (ср. (4.4))

$$\mu_N^e(\lambda/N, k+1) \overset{\mathrm{d}}{\approx} \exp[\lambda(\sigma/\sqrt{N})W(k+1) - \lambda^2\sigma^2/2N].$$

В результате получаем следующее асимптотическое представление ЭСЭ:

$$\Pi_N(\lambda/N, k+1) \overset{\mathrm{d}}{\approx} \exp[\lambda C(S_N(k)) - \lambda^2\rho^2/2N] \\ \times \exp[\lambda(\sigma/\sqrt{N})W(k+1) - \lambda^2\sigma^2/2N], \tag{4.14}$$

Или иначе с учётом (4.5)

$$\Pi_N(\lambda/N, k{+}1) \overset{\mathrm{d}}{\approx} \exp\{\lambda[C(S_N(k)){+}(\sigma/\sqrt{N})W(k{+}1)]{-}\lambda^2/2N\}. \tag{4.15}$$

Асимптотические формулы (4.14) и (4.15) дают основание для следующей аппроксимации ЭСЭ *нормальным процессом авторегрессии.*

Предложение 4.1. *Экспоненциальные статистические эксперименты* (3.1) *аппроксимируются нормальным процессом авторегрессии*

$$\prod_{r=1}^{N}[1 + \lambda\widetilde{\delta}_r(k+1)/N] = \exp[\lambda C(\widetilde{S_N}(k)) - \lambda^2\rho^2/2N] \\ \times \exp[\lambda(\sigma/\sqrt{N})W(k+1) - \lambda^2\sigma^2/2N], \tag{4.16}$$

или, эквивалентно,

$$\prod_{r=1}^{N}[1 + \lambda\widetilde{\delta}_r(k+1)/N] \\ = \exp\{\lambda[C(\widetilde{S_N}(k)) + (\sigma/\sqrt{N})W(k+1)] - \lambda^2/2N\}. \tag{4.17}$$

Замечание 4.2. Важным основанием для применения аппроксимации нормальным процессом авторегрессии (4.16) или (4.17) служит тот факт, что условное математическое ожидание нормального процесса авторегрессии (4.16) (или (4.17)) асимптотически совпадает с *функцией регрессии* (условным математическим ожиданием) исходных ЭСЭ (3.1), а именно (см. (4.15))

$$E\left[\prod_{r=1}^{N}[1+\lambda\widetilde{\delta}_r(k+1)/N]|\widetilde{S}_N(k)\right] = \exp[\lambda C(\widetilde{S_N}(k)) - \lambda^2\rho^2/2N] \\ = \overline{\Pi_N}(\lambda/N,k)e^{R_N}. \quad (4.18)$$

Заключение

Аппроксимация нормальным процессом авторегрессии (4.16) (или (4.17)) экспоненциальных статистических экспериментов с использованием процесса нормальной авторегрессии (4.4), полученного в теореме 4.1, создаёт разнообразные перспективы в приложениях моделей статистических экспериментов, обладающих свойством настойчивой линейной регрессии (1.4), в экономике и биологии. Прежде всего возникают проблемы статистической оценки основных параметров СЭ p и a, а также сопутствующих им ρ и σ.

Благодарности. Автор благодарит рецензента за его замечания, которые помогли улучшить изложение материала.

Литература

[1] Д. В. Королюк, *Бинарные повторяющиеся статистические эксперименты с настойчивой линейной регрессией* // УМВ **10** (2013), No. 4, 497–506.

[2] Yu. V. Borovskikh, V. S. Korolyuk, *Random Permanents*, VSP Utrecht, 1994.

[3] Yu. V. Borovskikh, V. S. Korolyuk, *Martingale Approximation*, VSP, 1997.

[4] А. Н. Ширяев, *Вероятность-2*, М., МЦНМО, 2004.

[5] M. Abundo, L. Accardi, N. Rosato, *A Markovian model for cooperative interaction in Proteins* // Mathematical Models and Methods in Applied Sciences, **5** (1995), No. 6 835–863.

[6] M. Abundo, L. Accardi, Agro'A. Finazzi, G. Mei, N. Rosato, *A stochastic model for the sigmoidal behaviour of cooperative biological systems* // Biophysical Chemistry, **58** (1996), 313–323.

[7] D. Koroliuk, V. S. Koroliuk, *Diffusion approximation of Markov chains for cooperative interaction in proteins* // Volterra Proc. UTV, (1993), No. 156.

[8] D. Koroliuk, V. S. Koroliuk, *Diffusion approximation of stochastic Markov models with persistent regression* // Ukr. Math. Journ., **47** (1995), No. 7, 928–935.

Сведения об авторах

Дмитрий Владимирович Королюк

Институт телекоммуникаций и глобального информационного пространства НАН Украины
Чоколовский бульвар, 13,
Киев, 03110
Украина
E-Mail: dimitri.koroliouk@ukr.net