Теорія ймовір. та матем. статист.
Вип. 90, 2014, стор. 91–101

Teor. Ĭmovir. ta Matem. Statyst.
No. 90, 2014, pp. 91–101

ДВОКОМПОНЕНТНІ БІНАРНІ СТАТИСТИЧНІ ЕКСПЕРИМЕНТИ З НАПОЛЕГЛИВОЮ ЛІНІЙНОЮ РЕГРЕСІЄЮ

УДК 519.21

Д. В. КОРОЛЮК

Анотація. Вивчається послідовність бінарних статистичних експериментів, що задається вибіркою випадкових величин з наполегливою лінійною регресією. Будується стохастична апроксимація послідовності статистичних експериментів
процесом авторегресії з нормальними збуреннями, а також стохастична апроксимація послідовності експоненційних статистичних експериментів процесом, який задається експоненційною нормальною авторегресією.

Abstract. The object of consideration is a sequence of binary statistical experiments, given by a sample of random variables with persistent linear regression. Its stochastic approximation is constructed as a aprocess autoregression with normal perturbations. For exponential statistical experiments, the stochastic approximation sequence is constructed by mean of a process of exponential normal autoregression.

Аннотация. Изучается последовательность бинарных статистических экспериментов, задаваемых выборкой случайных величин с настойчивой линейной регрессией. Строится стохастическая аппроксимация последовательности статистических экспериментов процессом авторегрессии с нормальными возмущениями, а также стохастическая аппроксимация последовательности экспоненциальных статистических экспериментов процессом, заданным экспоненциальной нормальной авторегрессией.

1. Вступ

Бінарні статистичні експерименти задаються вибіркою випадкових величин, що приймають два значення. У практичних застосуваннях, значення бінарних випадкових величин фіксують наявність або відсутність певної ознаки. Базовою характеристикою є відносні частоти наявності або відсутності ознаки.

Динаміка бінарних статистичних експериментів задається наполегливою регресією, яка передбачає залежність (в середньому) відносних частот наступного експерименту (у момент часу $k + 1$) від середнього значення частот статистичного експерименту в даний момент часу k. При цьому функція регресії приростів відносних частот задається напрямною лінійною дією, що забезпечує динаміку частот в напрямку рівноважного стану. Представлення бінарних статистичних експериментів з використанням відносних частот наявності ознаки забезпечує симетрію в аналізі та можливість визначення рівноважного значення.

2. Постановка задачі

Розглядаються статистичні експерименти, які породжуються вибіркою $(\delta_r(k), 1 \leq r \leq N)$ незалежних у сукупності при кожному $k \geq 0$ і однаково розподілених випадкових величин, що приймають два бінарних вибіркових значення ± 1.

2010 *Mathematics Subject Classification.* Primary 60J70.

Ключові слова і фрази. Бінарний статистичний експеримент, наполеглива регресія, стабілізація, стохастична апроксимація, експоненціальний статистичний експеримент, експоненціальний процес нормальної авторегресії.

Бінарні статистичні експерименти (СЕ) задаються сумами вибіркових значень

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

Позначимо умовні ймовірності елементарних подій як

$$\mathsf{P}\{\delta_r(k+1) = \pm 1 \mid S_N(k) = s\} = P_\pm(s), \qquad |s| \leq 1, \ k \geq 0. \tag{2}$$

Ймовірність повної події дорівнює 1, тобто

$$P_+(s) + P_-(s) = 1, \qquad \forall s \in [-1, +1].$$

Таким чином досить вивчати одну з частот: $S_N^+(k)$ або $S_N^-(k)$. А результати для однієї частоти легко переформулювати для іншої частоти.

Має місце частотне представлення СЕ (1)

$$S_N(k) = S_N^+(k) - S_N^-(k). \tag{3}$$

Основним об'єктом вивчення є двокомпонентні СЕ

$$\overline{S_N}(k) := (S_N^+(k), S_N^-(k)), \qquad k \geq 0, \tag{4}$$

Нам знадобиться позначення

$$\bar{p} := (p_+, p_-), \qquad 0 \leq p_\pm \leq 1, \ p_+ + p_- = 1,$$

вектора числових значень відносних частот S_N^+ та S_N^- у поточному експерименті.

Важливим завданням аналізу СЕ служить вивчення їх приростів

$$\Delta S_N^\pm(k) := S_N^\pm(k+1) - S_N^\pm(k), \qquad k \geq 0.$$

Початковою основною умовою аналізу СЕ служить наступне

Припущення. *Прирости СЕ мають властивість наполегливій регресії, яка задається функцією лінійної регресії, що визначається напрямною лінійною дією:*

$$\mathsf{E}\left[\Delta S_N^\pm(k) \mid \overline{S_N}(k) = \bar{p}\right] =: C_0^\pm(\bar{p}) = \mp[V_+p_+ - V_-p_-], \qquad 0 < V_\pm < 1. \tag{5}$$

При цьому функція регресії двокомпонентного СЕ

$$C^\pm(\overline{p}) := \mathsf{E}\left[S_N^\pm(k+1) \mid \overline{S_N}(k) = \overline{p}\right], \qquad k \geq 0, \tag{6}$$

задається співвідношенням

$$C^\pm(\overline{p}) = p_\pm + C_0^\pm(\overline{p}). \tag{7}$$

Визначення напрямної дії (5) передбачає виконання наступної умови балансу

$$C_0^+(\overline{p}) + C_0^-(\overline{p}) = 0. \tag{8}$$

Має місце також очевидне співвідношення

$$C^+(\overline{p}) + C^-(\overline{p}) = 1. \tag{9}$$

Зокрема, наполеглива лінійна регресія для частот СЕ задається співвідношеннями

$$\begin{aligned} \mathsf{E}\left[S_N^+(k+1) \mid \overline{S_N}(k) = \bar{p}\,\right] &= p_+ - V_+p_+ + V_-p_-, \\ \mathsf{E}\left[S_N^-(k+1) \mid \overline{S_N}(k) = \bar{p}\,\right] &= p_- - V_-p_- + V_+p_+. \end{aligned} \tag{10}$$

Зауваження 2.1. Формули (10) ілюструють роль напрямних параметрів V_+, V_- у встановленні рівноважного режиму. З них видно, що частота даної ознаки ($\pm$) зменшується пропорційно його частоті з напрямним коефіцієнтом $V_\pm$, і збільшується пропорційно частоті протилежного знака з напрямним коефіцієнтом $V_\mp$.

Функція регресії визначального СЕ $S_N(k)$ (див. (1))

$$C(s) := \mathsf{E}[S_N(k+1) \mid S_N(k) = s]$$

виражається через функцію регресії приростів СЕ і також має властивість наполегливої лінійної регресії:

$$C(s) = s + C_0(s), \qquad C_0(s) := C_0^+(\bar{p}) - C_0^-(\bar{p}) = -V_+(1+s) + V_-(1-s), \tag{11}$$

де $p_\pm := \frac{1}{2}(1 \pm s)$, $|s| \le 1$.

Лема 2.1. *СЕ(1), що задаються функцією регресії* (5)–(7)*, характеризуються умовною дисперсією, що має наступний явний вираз*

$$\mathbf{D}\left[S_N^\pm(k+1) \mid \overline{S_N}(k) = \overline{p}\,\right] = \frac{1}{N} C^\pm(\overline{p})[1 - C^\pm(\overline{p})], \qquad k \ge 0. \tag{12}$$

Доведення леми 2.1. З урахуванням формули (6) маємо

$$\mathbf{D}[S_N^\pm(k+1) \mid \overline{S_N}(k) = \overline{p}] = \mathsf{E}\left[\left(S_N^\pm(k+1)\right)^2 \mid \overline{S_N}(k) = \overline{p}\,\right] - \left[C^\pm(\overline{p})\right]^2.$$

Позначимо

$$a_r^\pm := I\{\delta_r(k+1) = \pm 1\}.$$

Застосуємо формулу (2) для обчислення

$$(S_N^\pm(k+1))^2 = \frac{1}{N^2}\left[\sum_{r=1}^{N}[a_r^\pm]^2 + \sum_{r,r'=1}^{N,r\neq r'}[a_r^\pm][a_{r'}^\pm]\right].$$

Враховуючи, що $[a_r^\pm]^2 = a_r^\pm$, а також факт, що наведена вище подвійна сума містить $N(N-1)$ доданків, одержуємо

$$\mathsf{E}\left[(S_N^\pm(k+1))^2 \mid \overline{S_N}(k) = \overline{p}\right] = \frac{1}{N}\left[C^\pm(\overline{p}) - (C^\pm)^2(\overline{p})\right] + (C^\pm)^2(\overline{p}),$$

звідки випливає формула (12). Лема 2.1 доведена. □

3. Рівноважний стан (еквілібріум)

Рівноважний стан (еквілібріум) імовірносних частот ρ_+, ρ_- і ρ характеризується інваріантними значеннями відповідних функцій регресії

$$C^+(\rho_+) = \rho_+, \qquad C^-(\rho_-) = \rho_-, \qquad C(\rho) = \rho.$$

Для функцій регресії (5) і (11) це означає, що напрямні дії в точках рівноважного стану обертаються в нуль:

$$C_0^\pm(\bar{\rho}) = 0, \quad \text{тобто } V_+\rho_+ = V_-\rho_-, \tag{13}$$

де $\bar{\rho} := (\rho_+, \rho_-)$, $\rho_+ + \rho_- = 1$.

Відповідно до припущень (див. формулу (5)), середнє значення СЕ в наступний момент часу збігається в точці рівноважного стану зі значенням СЕ в даний момент часу:

$$\mathsf{E}\left[S_N^\pm(k+1) \mid \overline{S_N}(k) = \overline{\rho}\right] = \rho_\pm, \quad \text{для всіх } k \ge 0,$$

а також

$$\mathsf{E}\left[S_N(k+1) \mid \overline{S_N}(k) = \overline{\rho}\right] = \rho := \rho_+ - \rho_- \quad \text{для всіх } k \ge 0.$$

Значення рівноважного стану $\bar{\rho}$ визначається розв'язком рівняння (13):

$$\begin{cases} \rho_+ V_+ = \rho_- V_- = C = \text{const}, \\ \rho_+ + \rho_- = 1. \end{cases} \tag{14}$$

Чи інакше

$$\begin{cases} \rho_{\pm} = C/V_{\pm}, \\ C = [1/V_{+} + 1/V_{-}]^{-1}. \end{cases} \tag{15}$$

Зокрема

$$\rho_{\pm} = V_{\mp}/V, V = V_{+} + V_{-}. \tag{16}$$

Зауваження 3.1. Введення рівноважного стану $\bar{\rho}$ виявляє наступні властивості напрямної дії:

$$C_0^{\pm}(\bar{p}) = -V(p_{+} - \rho_{+}) = -V(p_{-} - \rho_{-}), \tag{17}$$

$$C_0(s) = C_0^{+}(\bar{p}) - C_0^{-}(\bar{p}) = -V(s - \rho), \tag{18}$$

де використані такі позначення:

$$s = p_{+} - p_{-}, \qquad \rho := \rho_{+} - \rho_{-}.$$

Зауваження 3.2. Як видно з формул (17)–(18), *напрямна дія є функція флуктуацій*, тобто залежить від відхилення поточного значення СЕ від рівноважного значення. При цьому напрямна дія для кожної частоти визначається тільки флуктуаціями відповідної частоти.

Введемо процеси флуктуацій частот СЕ

$$\begin{gathered} \widehat{S}_N^{\pm}(k) := S_N^{\pm}(k) - \rho_{\pm}, \qquad k \geq 0, \\ \widehat{p}_{\pm} := p_{\pm} - \rho_{\pm} \end{gathered} \tag{19}$$

та їх відповідні функції регресії:

$$\widehat{C}^{\pm}(\widehat{p}_{\pm}) = \mathsf{E}\left[\widehat{S}_N^{\pm}(k+1) \mid \widehat{S}_N^{\pm}(k) = \widehat{p}_{\pm}\right] = (1 - V)\widehat{p}_{\pm}, \qquad k \geq 0. \tag{20}$$

Флуктуації визначального СЕ (1)

$$\widehat{S}_N(k) := S_N(k) - \rho, \qquad \rho := \rho_{+} - \rho_{-}, \qquad k \geq 0, \tag{21}$$

мають наступну функцію регресії:

$$\widehat{C}(\widehat{s}) := \mathsf{E}\left[\widehat{S}_N(k+1) \mid \widehat{S}_N(k) = \widehat{s}\right] = (1 - V)\widehat{s}, \qquad k \geq 0,\ \widehat{s} := s - \rho. \tag{22}$$

Зауваження 3.3. Таким чином, у нас є дві можливості аналізу СЕ (1)–(2):

а) безпосередньо з урахуванням (17)–(18);
б) спочатку здійснюється аналіз двокомпонентних СЕ (2), що задаються частотами $S_N^{\pm}(k)$, $k \geq 0$. А потім формулюються результати аналізу для СЕ (1) з урахуванням формули (3).

Зауваження 3.4. В нашій попередній роботі [1] бінарні статистичні експерименти $S_N(k)$, $k \geq 0$, приймаючі два вибіркових значення ± 1, вивчалися безпосередньо із застосуванням функції лінійної регресії $C(s) = (1 - a)s + a\rho$ або, що те ж саме, $\widehat{C}(\widehat{s}) = (1 - a)\widehat{s}$.

У розглянутій нами проблемі параметр керування $a = V$, а завдання рівноважного стану $\rho = \widehat{\rho} := \rho_{+} - \rho_{-}$ очевидно.

4. Збіжність СЕ до рівноважного стану

Напрямні дії (10) і (11) забезпечують наявність рівноважних станів СЕ (1)–(2).

Теорема 4.1. *При збіжності з імовірністю 1 початкових умов:*

$$S_N^{\pm}(0) \xrightarrow{\mathsf{P}\,1} \rho_{\pm}, \qquad N \to \infty, \tag{23}$$

мають місце збіжності з імовірністю 1:

$$S_N^{\pm}(k) \xrightarrow{\mathsf{P}\,1} \rho_{\pm}, \qquad N \to \infty \quad \forall k > 0. \tag{24}$$

$$S_N(k) \xrightarrow{\mathsf{P}\,1} \rho, \qquad N \to \infty \quad \forall k > 0. \tag{25}$$

Доведення теореми 4.1. Досить довести збіжність (24) для позитивних частот, тобто з імовірністю 1:

$$S_N^{+}(k) \xrightarrow{\mathsf{P}\,1} \rho_{+}, \qquad N \to \infty \quad \forall k > 0. \tag{26}$$

Тоді в силу співвідношення (3) має місце збіжність (24) для негативних частот, а також збіжність (25) в силу частотного представлення (3).

Введемо мартингал у вигляді суми

$$\widehat{M}_N^{+}(n) := \sum_{k=0}^{n-1} \widehat{\mu}_N^{+}(k+1) \tag{27}$$

мартингал-різниць

$$\widehat{\mu}_N^{+}(k+1) := \widehat{S}_N^{+}(k+1) - (1-V)\widehat{S}_N^{+}(k). \tag{28}$$

Очевидно, що

$$\widehat{\mu}_N^{+}(k+1) = \mu_N^{+}(k+1),$$

де

$$\mu_N^{+}(k+1) := S_N^{+}(k+1) - C^{+}(S_N^{+}(k)). \tag{29}$$

Так що квадратичні характеристики мартингал-різниць (29) збігаються і мають, згідно формули (12), наступний вигляд:

$$\langle \widehat{\mu}_N^{+} \rangle_k = \langle \mu_N^{+} \rangle_k = C^{+}(S_N^{+}(k))\big(1 - C^{+}(S_N^{+}(k))\big)/N. \tag{30}$$

Отже, при кожному фіксованому $n \geq 0$ має місце збіжність

$$\langle M_N^{+} \rangle_n \to 0, \qquad N \to \infty. \tag{31}$$

Звідси [5] випливає збіжність з імовірністю 1 мартингалів (27)

$$M_N^{+}(n) \xrightarrow{\mathsf{P}\,1} 0 \quad \text{при} N \to \infty,\ n \geq 0. \tag{32}$$

Зокрема при $n = 1$ маємо

$$M_N^{+}(0) = \widehat{S}_N^{+}(1) - (1-V)\widehat{S}_N^{+}(0).$$

З урахуванням збіжності до нуля $M_N^{+}(0)$ і $\widehat{S}_N^{+}(0)$ при $N \to \infty$, маємо також

$$\widehat{S}_N^{+}(1) \xrightarrow{\mathsf{P}\,1} 0 \quad \text{при} N \to \infty.$$

За індукцією виводимо, що при кожному $k \geq 1$ має місце збіжність

$$\widehat{S}_N^{+}(k) \xrightarrow{\mathsf{P}\,1} 0 \quad \text{при} N \to \infty.$$

Теорема 4.1 доведена. □

5. Апроксимація СЕ процесом нормальної авторегресії

Наявність рівноважного стану СЕ (1)–(2), сформульованого в Теоремі 1, є важливою, але найпростішою властивістю СЕ. Основна проблема спрощеного опису поведінки СЕ (1)–(2) при зростаючому обсязі вибірки N полягає в асимптотичному аналізі непередбаченої (хаотичної) частини СЕ, що задається мартингал-різницями (28).

Асимптотичну властивість мартингал-різниці (28) встановлює наступна

Теорема 5.1. *В умовах теореми 1 мають місце збіжності за розподілом:*

$$\sqrt{N}\left[S_N^{\pm}(k+1) - C^{\pm}(S_N^{\pm}(k))\right] \xrightarrow{d} \breve{\sigma} W^{\pm}(k+1), \qquad N \to \infty, \tag{33}$$

$$\sqrt{N}\left[S_N(k+1) - C(S_N(k))\right] \xrightarrow{d} \sigma W(k+1), \qquad N \to \infty, \tag{34}$$

де

$$\sigma^2 = 1 - \rho^2, \qquad \breve{\sigma}^2 = \rho_+ \rho_-,$$

причому $\sigma^2 = 4\breve{\sigma}^2$, $W(k)^{\pm}$, $W(k)$ — нормально розподілені стандартні випадкові величини.

Граничні співвідношення (33), (34) теореми 5.1 дають підстави використовувати стохастичну апроксимацію для трьох процесів СЕ в такій формі:

Пропозиція 5.1. *Процеси нормальної авторегресії, апроксимуючі послідовність визначальних СЕ (1)–(2), (5)–(7) (див. також (10)) задаються співвідношеннями:*

$$\widetilde{S}_N^{\pm}(k+1) = C^{\pm}(\widetilde{S}_N^{\pm}(k)) + \frac{\breve{\sigma}}{\sqrt{N}} W^{\pm}(k+1), \tag{35}$$

$$\widetilde{S}_N(k+1) = C(\widetilde{S}_N(k)) + \frac{\sigma}{\sqrt{N}} W(k+1). \tag{36}$$

Тут функції регресії $C^{\pm}(\overline{p})$ і $C(s)$ задаються співвідношеннями (5) і (11) відповідно.

Доведення теореми 5.1. Як у доведенні теореми 4.1, достатньо довести збіжність (33) для позитивних частот:

$$\sqrt{N}\left[S_N^{+}(k+1) - C^{\pm}(S_N^{+}(k))\right] \xrightarrow{d} \breve{\sigma} W^{+}(k+1) \quad \text{при } N \to \infty, \tag{37}$$

звідки випливають інші збіжності (33)–(34) Теореми 5.1.

Згідно з позначеннями (19), введемо нормовані флуктуації позитивних частот

$$\zeta_N^{+}(k) := \sqrt{N}\widehat{S}_N^{+}(k).$$

Зазначимо, що їх умовні математичні сподівання, згідно (20), мають вигляд:

$$\mathsf{E}\left[\zeta_N^{+}(k+1) \mid \zeta_N^{+}(k)\right] = (1-V)\zeta_N^{+}(k). \tag{38}$$

А їх умовні дисперсії, згідно (12):

$$\mathbf{D}\left[\zeta_N^{+}(k+1) \mid \zeta_N^{+}(k)\right] = C^{+}(S_N^{+}(k))\left[1 - C^{+}(S_N^{+}(k))\right]. \tag{39}$$

Тепер введемо мартингал у вигляді суми мартингал-різниць

$$\begin{aligned} M_N^{+}(n) &:= \sum_{k=0}^{n-1} \left\{\zeta_N^{+}(k+1) - \mathsf{E}\left[\zeta_N^{+}(k+1) \mid \zeta_N^{+}(k)\right]\right\} \\ &= \sum_{k=0}^{n-1} \left[\zeta_N^{+}(k+1) - (1-V)\zeta_N^{+}(k)\right]. \end{aligned} \tag{40}$$

Квадратична характеристика мартингалу (40) (див. формулу (39)) має наступний вигляд:

$$\langle M_N^+ \rangle_n = \sum_{k=0}^{n-1} C^+(S_N^+(k)) \left[1 - C^+(S_N^+(k))\right], \qquad n \geq 0. \tag{41}$$

Згідно з теоремою 4.1, має місце збіжність з імовірністю 1 квадратичної характеристики

$$\langle M_N^+ \rangle_n \xrightarrow{\mathsf{P}\,1} n\rho_+\rho_- = n\breve{\sigma} \quad \text{при} N \to \infty, \ n \geq 0. \tag{42}$$

Згідно з центральною граничною теоремою для послідовностей, що утворюють мартингал-різниці (напр. [5], VII: 8, Т.4), мартингал (40) збігається за розподілом при $N \to \infty$ до суми нормально розподілених випадкових величин

$$\sum_{k=0}^{n-1} [\zeta_N^+(k+1) - (1-V)\zeta_N^+(k)] \xrightarrow{d} \breve{\sigma} \sum_{k=0}^{n-1} W(k+1), \qquad N \to \infty, \ n \geq 1. \tag{43}$$

Граничні нормально розподілені випадкові величини $W(k)$, $k \geq 1$, незалежні в сукупності з

$$\mathsf{E}\, W(k) = 0, \qquad \mathsf{E}\, W^2(k) = 1, \qquad k \geq 1,$$

оскільки, згідно з формулою (42), гранична дисперсія мартингалу (40) дорівнює сумі дисперсій мартингал-різниць.

Збіжність мартингалу (40) означає, що має місце збіжність за розподілом (37) при кожному скінченному $k \geq 0$. Теорема 5.1 доведена. □

Пропозиція 5.2. *Для нормованих флуктуацій СЕ процес нормальної авторегресії задається розв'язком різницевого стохастичного рівняння:*

$$\Delta\zeta_N(k) = -V\zeta_N(k) + \sigma\Delta W(k), \tag{44}$$

де

$$\Delta W(k) := W(k+1) - W(k), \qquad \sigma^2 = 1 - \rho^2.$$

Зауваження 5.1. Процеси нормальної авторегресії можуть бути використані в статистичному аналізі реальних СЕ. Характеризація визначальних процесів СЕ процесами нормальної авторегресії (35), (36) та (44) задається напрямними параметрами $V_\pm$, а також рівноважними значеннями $\rho_\pm$, ρ і дисперсіями $\breve{\sigma}$, σ нормально розподілених випадкових величин $W^\pm$ і W. Різницеве стохастичне рівняння (44) також є об'єктом статистичного аналізу (напр., [3, 10]).

Статистичні оцінки числових параметрів процесів нормальної авторегресії будуються за траєкторіями визначальних СЕ.

6. Експоненційні статистичні експерименти

Граничні теореми 4.1 і 5.1 для СЄ (1)–(2) можуть бути використані в асимптотичному аналізі *експоненційних статистичних експериментів* (ЕСЕ), які задаються симетричною експоненційною статистикою

$$\Pi_N(\lambda, k) := \prod_{r=1}^{N} [1 + \lambda\delta_r(k)], \qquad k \geq 0. \tag{45}$$

Параметр λ $(0 < \lambda < 1)$ відіграє важливу роль регулювання в практичних моделях, наприклад застосовується в економіці в якості дисконтного параметра (напр. [6], [7]).

У представленні ЕСЕ (45) вибіркові значення $(\delta_r(k), 1 \leq r \leq N)$, $k \geq 0$ розглядаються в трьох окремих варіантах:

$$\delta_r(k) = \pm 1, \qquad \delta_r^\pm(k) := I(\delta_r(k) = \pm 1). \tag{46}$$

Так що ЕСЕ породжуються одним з трьох СЕ, заданих формулами (1) або (2).

Асимптотична поведінка ЕСЕ, при $N \to \infty$, досліджується з використанням Теорем 4.1 і 5.1, так що досить сформулювати результати для однієї з ЕСЕ (45).

Будемо розглядати ЕСЕ (45), породжені СЕ (1). Аналогічні результати можуть бути отримані для частотних ЕСЕ.

7. Сталий режим ЕСЕ

У схемі осереднення ЕСЕ (45) розглядається в нормуванні $\lambda_N = \lambda/N$.

Введемо експоненційний мартингал

$$\mu_N^e(\lambda/N, k+1) = \Pi_N(\lambda/N, k+1)/\overline{\Pi_N}(\lambda/N, k), \qquad k \geq 0, \tag{47}$$

в якому за визначенням

$$\overline{\Pi_N}(\lambda, k) := \mathsf{E}\,[\Pi_N(\lambda, k+1) \mid S_N(k)] = [1 + \lambda C(S_N(k))]^N. \tag{48}$$

Очевидна його мартингальна властивість:

$$\mathsf{E}\,[\mu_N^e(\lambda/N, k+1) \mid S_N(k)] = 1, \qquad k \geq 0. \tag{49}$$

Сталий режим ЕСЕ встановлює наступна

Теорема 7.1. *За умови збіжності з імовірністю 1 початкових значень СЕ*

$$S_N(0) \xrightarrow{\mathsf{P}\,1} \rho, \qquad N \to \infty, \tag{50}$$

має місце збіжність за ймовірністю ЕСЕ (47)

$$\Pi_N(\lambda/N, k) \xrightarrow{\mathsf{P}} \exp(\lambda\rho), \qquad N \to \infty,\ k \geq 0, \tag{51}$$

і збіжність за ймовірністю його середнього значення (48)

$$\overline{\Pi_N}(\lambda/N, k) \xrightarrow{\mathsf{P}} \exp(\lambda\rho), \qquad N \to \infty,\ k \geq 0. \tag{52}$$

Наслідок 7.1. *За умовою* (50)

$$\mu_N^e(\lambda/N, k+1) \xrightarrow{\mathsf{P}} 1, \qquad N \to \infty. \tag{53}$$

Доведення теореми 7.1. Скористаємося формулою апроксимації типу Ле Кама в такій формі:

Лема 7.1 (ср. [4, Lemma 6.3.1])**.** *Нехай виконується збіжність за ймовірністю*

$$\max_{1 \leq r \leq N} |\delta_r(k+1)/N| \xrightarrow{\mathsf{P}} 0, \qquad N \to \infty,\ \forall k \geq 0. \tag{54}$$

Тоді має місце збіжність за ймовірністю

$$\left\{\sum_{r=1}^{N} \ln[1 + \lambda\delta_r(k+1)/N] - \lambda S_N(k+1)\right\} \xrightarrow{\mathsf{P}} 0, \qquad N \to \infty. \tag{55}$$

Тут

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

У нашому випадку умова (54) леми 7.1 виконана, згідно теореми 4.1

$$S_N(k) \xrightarrow{\mathsf{P}\,1} \rho, \qquad N \to \infty. \tag{56}$$

Перепишемо збіжність (55) з урахуванням очевидної тотожності $\Pi = \exp\ln\Pi$:

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

Або інакше з урахуванням збіжності за ймовірністю (56)

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

Останній результат збігається з твердженням (51) теореми 7.1. Аналогічним чином доводиться збіжність (52) для середнього значення (48). Збіжність (53) випливає з (51) і (52). Теорема 7.1 доведена. □

8. Апроксимація ЕСЕ процесом нормальної авторегресії

Тепер ЕСЕ розглядаються в схемі серій з параметром серії $\lambda_N=\lambda/\sqrt{N}$. Так що

$$\Pi_N(\lambda/\sqrt{N},k+1)=\prod_{r=1}^{N}[1+\lambda\delta_r(k+1)/\sqrt{N}],\qquad k\geq 0. \tag{57}$$

Крім того, середнє значення ЕСЕ (57) (пор. з (48))

$$\overline{\Pi_N}(\lambda/\sqrt{N},k)=[1+\lambda C(S_N(k))/\sqrt{N}]^N,\qquad k\geq 0. \tag{58}$$

Введемо експоненційний мартингал

$$\mu_N^e(\lambda_N,k+1):=\Pi_N(\lambda/\sqrt{N},k+1)/\overline{\Pi_N}(\lambda/\sqrt{N},k),\qquad k\geq 0. \tag{59}$$

Очевидна мартінгальна властивість:

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

Теорема 8.1. *(Апроксимація ЕСЕ) В умовах теореми 7.1 має місце збіжність за розподілом ($\sigma^2=1-\rho^2$):*

$$\mu_N^e(\lambda/\sqrt{N},k+1)\xrightarrow{d}\exp[\lambda\sigma W(k+1)-\lambda^2\sigma^2/2],\qquad N\to\infty,\ k\geq 0. \tag{60}$$

Зауваження 8.1. Граничний результат (60) є аналогом експоненційного мартингалу (59).

При цьому очевидно, що

$$\mathsf{E}\exp\left[\lambda\sigma W(k+1)-\lambda^2\sigma^2/2\right]=1,\qquad k\geq 0. \tag{61}$$

Доведення теореми 8.1. Так само, як і при доведенні теореми 7.1, скористаємося тотожністю $\Pi=\exp\ln\Pi$, а також лемою апроксимації Ле Кама.

Лема 8.1 (Апроксимація Ле Кама [4, Lemma 6.3.1])**.** *Нехай виконується збіжність за ймовірністю*

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

а також суми

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

обмежені за ймовірностю. Тоді має місце збіжність за ймовірністю

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

Отже має місце збіжність за ймовірністю (при $N\to\infty$) з додатковим множником:

$$\Pi_N(\lambda/\sqrt{N},k+1)\exp\left[-\lambda\sqrt{N}S_N(k+1)+\lambda^2V_N(k+1)/2\right]\xrightarrow{\mathsf{P}}1,\qquad N\to\infty,\qquad k\geq 0. \tag{62}$$

Зауважимо, що в даному випадку

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

Так що збіжність (62) еквівалентна збіжності

$$\left\{\Pi_N\left(\lambda/\sqrt{N}, k+1\right)\exp\left[-\lambda\sqrt{N}S_N(k+1)\right]\right\} \xrightarrow{\mathsf{P}} \exp[-\lambda^2/2], \\ N\to\infty, \qquad k\geq 0. \tag{63}$$

Аналогічно умови леми 8.1 дозволяють отримати збіжність за ймовірністю (при $N\to\infty$) для середнього значення (58):

$$\left\{\overline{\Pi_N}\left(\lambda/\sqrt{N}, k\right)\exp\left[-\lambda\sqrt{N}C(S_N(k))\right]\right\} \xrightarrow{\mathsf{P}} \exp[-\lambda^2\rho^2/2], \qquad N\to\infty,\ k\geq 0. \tag{64}$$

Тут ми скористалися збіжністю, встановленою в теоремі 4.1 :

$$C(S_N(k)) \xrightarrow{\mathsf{P}\,1} C(\rho) = \rho, \qquad N\to\infty,\ k\geq 0.$$

Так що

$$C^2(S_N(k)) \xrightarrow{\mathsf{P}\,1} \rho^2, \qquad N\to\infty,\ k\geq 0.$$

Визначальний експоненційний мартингал (59) може бути поданий наступним чином:

$$\mu_N^e(\lambda/\sqrt{N}, k+1) = \left[\Pi_N^0\left(\lambda/\sqrt{N}, k+1\right)\overline{\Pi_N^0}\left(\lambda/\sqrt{N}, k\right)\right] \\ \times\exp\left\{\lambda\sqrt{N}[S_N(k+1) - C(S_N(k))]\right\}, \qquad k\geq 0. \tag{65}$$

Тут, за визначенням,

$$\Pi_N^0\left(\lambda/\sqrt{N}, k+1\right) := \Pi_N\left(\lambda/\sqrt{N}, k+1\right)\exp\left[-\lambda\sqrt{N}S_N(k+1)\right], \qquad k\geq 0, \\ \overline{\Pi_N^0}\left(\lambda/\sqrt{N}, k\right) := \overline{\Pi_N}\left(\lambda/\sqrt{N}, k\right)\exp\left[-\lambda\sqrt{N}C(S_N(k))\right], \qquad k\geq 0. \tag{66}$$

За теоремою 5.1 має місце збіжність (за розподілом):

$$\sqrt{N}[S_N(k+1) - C(S_N(k))] \xrightarrow{d} \sigma W(k+1), \qquad N\to\infty,\ k\geq 0. \tag{67}$$

Так що збіжності (63), (64) та (67) забезпечують збіжність експоненційного мартингалу (65), а значить і (60).

Теорема 8.1 доведена. □

Гранична теорема 8.1, а також вираз (59) для експоненційного мартингалу дають можливість виразити ЕСЕ (57) у схемі серій в наступному вигляді:

$$\Pi_N\left(\lambda/\sqrt{N}, k+1\right) = \overline{\Pi_N}\left(\lambda/\sqrt{N}, k\right)\mu_N^e\left(\lambda/\sqrt{N}, k+1\right). \tag{68}$$

Згідно (66), представлення ЕСЕ (68) перетвориться до виду

$$\Pi_N\left(\lambda/\sqrt{N}, k+1\right) = \overline{\Pi_N^0}\left(\lambda/\sqrt{N}, k\right)\exp\left[\lambda C(S_N(k))\right]\mu_N^e\left(\lambda/\sqrt{N}, k+1\right). \tag{69}$$

З врахуванням (64), маємо:

$$\overline{\Pi_N^0}\left(\lambda/\sqrt{N}, k\right) = \exp\left[-\lambda^2\rho^2/2\right]e^{R_N}. \tag{70}$$

Тут і далі залишковий член R_N прямує до нуля за ймовірністю при $N\to\infty$.

Представлення ЕСЕ (69) з урахуванням теореми 8.1, а також асимптотичної формули (70) дають підставу апроксимації ЕСЕ (47) нормальним процесом авторегресії.

Пропозиція 8.1. *Експонентні статистичні експерименти* (47) *апроксимуються нормальним процесом авторегресії*

$$\begin{aligned}\widetilde{\Pi}_N\left(\lambda/\sqrt{N}, k+1\right) &:= \prod_{r=1}^{N}\left[1+\lambda\widetilde{\delta}_r(k+1)/\sqrt{N}\right] \\ &= \exp\left[\lambda\sqrt{N}C\left(\widetilde{S_N}(k)\right)-\lambda^2\rho^2/2\right]\cdot\exp\left[\lambda\sigma W(k+1)-\lambda^2\sigma^2/2\right],\end{aligned} \tag{71}$$

або, еквівалентно,

$$\widetilde{\Pi}_N\left(\lambda/\sqrt{N}, k+1\right) = \exp\left[\left[\sqrt{N}\lambda C\left(\widetilde{S}_N(k)\right)+\sigma W(k+1)\right]-\lambda^2/2\right]. \tag{72}$$

Зауваження 8.2. Важливою підставою для застосування апроксимації нормальним процесом авторегресії (71) або (72) служить той факт, що умовне математичне сподівання нормального процесу авторегресії (71) (або (72)) асимптотично збігається з *функцією регресії* (умовним математичним сподіванням) визначальних ЕСЕ (45), а саме (див. (64)–(66))

$$\mathsf{E}\left[\widetilde{\Pi}_N(\lambda/\sqrt{N}, k+1)\right] = \exp\left[\sqrt{N}\lambda C(\widetilde{S_N}(k))-\lambda^2\rho^2/2\right] = \overline{\Pi_N}\left(\lambda/\sqrt{N}, k\right)e^{R_N}. \tag{73}$$

9. Висновки

Апроксимація нормальним процесом авторегресії (71) (або (72)) експоненційних статистичних експериментів створює різноманітні перспективи для практичних моделей статистичних експериментів, що мають властивість наполегливої лінійної регресії (5) - (7), в економіці (напр. [6, 7]) та біології (напр. [8, 9]). Перш за все виникають проблеми статистичної оцінки основних параметрів СЕ, а саме V, ρ і σ.

Література

1. Д. В. Королюк, *Бинарные повторяющиеся статистические эксперименты с настойчивой линейной регрессией*, УМВ **10** (2013), № 4, 497–506.
2. S. N. Ethier and T. G. Kurtz, *Markov Processes: Characterization and Convergence*, Wiley, New York, 1986.
3. А. В. Скороход, *Асимптотические методы теории стохастических дифференциальных уравнений*, "Наукова Думка", Киев, 1987.
4. Yu. V. Borovskikh and V. S. Korolyuk, *Martingale Approximation*, VSP, 1997.
5. А. Н. Ширяев, *Вероятность-2*, МЦНМО, Москва, 2004.
6. А. Н. Ширяев, *Основы стохастической финансовой математики. Т. 1 Факты. Модели.* "ФАЗИС", Москва, 1998.
7. Ю. С. Мішура, Г. М. Шевченко, *Математика фінансів*, ВПЦ "Київський університет", 2011.
8. M. Abundo, L. Accardi, L. Stella, and N. Rosato, *A stochastic model for the cooperative relaxation of proteins, based on a hierarchy of interactions between amino acidic residues*, M3AS (Mathematical Models and Methods in Applied Sciences) **8** (1998), 327–358.
9. V. S. Korolyuk and D. Koroliouk, *Diffusion approximation of stochastic Markov models with persistent regression*, Ukr. Mat. J. **47** (1995), no. 7, 928–935.
10. А. Щербина, *Оцінювання у моделі суміші зі змінними концентраціями*, Теор. Ймовір. та матем. статист. **84** (2011), 142–154.

Інститут телекомунікацій і глобального інформаційного простору НАНУ, Чоколовский бульвар, 13, Київ 03110, Україна

Адреса електронної пошти: dimitri.koroliouk@ukr.net

Надійшла 26/04/2013