Теорія ймовір. та матем. статист.
Вип. 85, 2011, стор. 39–49

Teor. Ĭmovir. ta Matem. Statyst.
No. 85, 2011, pp. 39–49

НЕПАРАМЕТРИЧНЕ ОЦІНЮВАННЯ ДЛЯ СКЛАДНОГО ПУАСОНІВСЬКОГО ПРОЦЕСУ, ЩО КЕРУЄТЬСЯ ЛАНЦЮГОМ МАРКОВА

УДК 519.21

О. А. ВОЙНА

Анотація. В роботі розглянуто задачу непараметричного оцінювання для пуасонівського процесу, що керується ланцюгом Маркова з неперервним часом, при неповних спостереженнях. Запропоновано метод непараметричного оцінювання, що базується на можливості представлення такого процесу у вигляді прихованої марковської моделі. Проведено детальний аналіз запропонованої статистичної процедури на прикладі оцінювання невідомих функцій розподілу при неповних спостереженнях складного пуасонівського процесу, що керується простим марковським процесом регенерації.

Аннотация. В работе рассмотрена задача непараметрического оценивания для пуассоновского процесса, управляемого цепью Маркова с непрерывным временем, в условиях неполных наблюдений. Предложен метод непараметрического оценивания, основанный на возможности представления такого процесса в виде скрытой модели Маркова. Сделан подробный анализ предложенной статистической процедуры на примере оценивания неизвестных функций распределения в условиях неполных наблюдений сложного пуассоновского процесса, управляемого простым Марковским процессом регенерации.

Abstract. The problem of nonparametric estimation for the Poisson process controlled by Markov process with the continuous time in the case of incomplete observations is considered. A method of the nonparametric estimation based on the possibility of the presentation of such process in the form of hidden Markov models is proposed. An example of application of this method for estimation of unknown distribution functions in the case of incomplete observations of a composite Poisson process controlled by a simple Markov regeneration process is considered.

1. Вступ

Розглянемо приклад багатовимірного стохастичного процесу з дискретним втручанням випадку, який можна умовно назвати "складним пуасонівським, процесом, що керується ланцюгом Маркова з неперервним часом". Подібного типу процеси завжди викликали зацікавлення з боку математиків, оскільки мають яскраво виражений прикладний характер. Конкретна інтерпретація елементів їх структури створює широкі можливості безпосереднього їх використання до дослідження різноманітних явищ. Так, наприклад, в роботі [2] показано, що процеси, які вивчаються в цій статті, можна ефективно використати до оптимального керування ризиком в моделях управління запасами, при вивченні ризику в багатовимірних моделях страхування, і т.п. Варто підкреслити, що бібліографія присвячена різноманітним класам процесів з дискретним втручанням випадку є практично необмеженою, тому наведені в статті посилання [1, 3, 4, 6, 7] стосуються головним чином праць, які мають безпосередній зв'язок з моделями, що в ній досліджуються. В цитованих та подібних до них працях предметом математичних досліджень були перш за все питання строгого формального визначення різноманітних типів процесів з дискретним втручанням

2000 *Mathematics Subject Classification.* Primary 62M05; Secondary 60J99, 93E11.

Ключові слова і фрази. Складний пуасонівський процес, неповні спостереження, непараметричне оцінювання, прихована марковська модель, модель Кокса.

випадку, умов їх існування, вивчення розподілу функціоналів від таких процесів та їх гранична поведінка, можливість апроксимації таких процесів більш простими, і т.п., при умові, що стохастична структура процесу, який досліджується, чітко визначена. Дослідження проведені в запропонованій роботі стосуються в певному розумінні протилежної проблеми, оскільки отримані в ній результати розв'язують задачу відновлення елементів "стохастичної структури" процесу з дискретним втручанням випадку на підставі спостережень функціоналів від його траєкторій. При цьому припускається, що значення самих процесів недоступні для безпосереднього спостереження і невідомі. Відповідний тип статистичних даних та пов'язаних з ними методів дослідження отримав назву " прихованих маркоських моделей" ([8, 9, 10]).

2. Основна частина

Нехай на тому самому імовірносному просторі $(\Omega, \Im, \mathsf{P})$ визначено наступні об'єкти:

- Ланцюг Маркова $x(t)$, $t \geq 0$, з неперервним часом, з множиною станів
$$E = \{1, 2, \ldots, m\},$$
інтенсивності переходів якого визначаються матрицею:
$$A = \|a_{ij}\|, \qquad i, j = 1, 2, \ldots, m.$$
- Для кожного стану $i \in E$ визначено випадкову величину ζ_i з функцією розподілу:
$$U_i(x) = \mathsf{P}\{\zeta_i < x\}, \qquad i \in E = \{1, 2, \ldots, m\}.$$
- Пуасонівський процес $\Pi(t)$, $t \geq 0$, з параметром λ, який не залежить ні від ланцюга $x(t)$, ні від сімейства випадкових величин $\{\zeta_i, i \in E\}$.

Припустимо, що точковий випадковий процес $S(n) \in [0, +\infty)$ з дискретним часом $n = 0, 1, 2, \ldots$ описує потік стрибків пуасонівського процесу $\Pi(t)$, тобто:
$$S(0) = 0, \qquad \Pi(t) = \max\{k \colon S(k) \leq t\}.$$
Введемо незалежну від випадкових процесів $x(t)$ та $\Pi(t)$ сукупність випадкових величин $[\{\zeta_i^{(k)}, k = 0, 1, 2, \ldots\}, i \in E]$, таких, що для різних значень індексу k величини $\zeta_i^{(k)}$ незалежні між собою, їх розподіл не залежить від індексу k і визначається функцією розподілу $U_i(x)$, $i = 1, 2, \ldots, m$.

Означення 2.1. Багатовимірним складним пуасонівським процесом, що керується ланцюгом Маркова з неперервним часом, будемо називати випадковий векторний процес:
$$\xi(t) = \{\xi_1(t), \xi_2(t), \ldots, \xi_m(t)\}, \qquad t \geq 0,$$
для якого:
$$\xi_{x(S(k))}(t) = \xi_{x(S(k))}(S(k)) + \zeta_{x(S(k))}^{(k)}, \qquad S(k) < t \leq S(k+1);$$
$$\xi_j(t) = \xi_j(S(k)), \qquad j \neq x(S(k)), \qquad S(k) < t \leq S(k+1).$$

Введемо також випадковий процес:
$$\zeta(k) = \zeta_{x(S(k))}^{(k)}, \qquad k = 1, 2, 3, \ldots.$$
Процес $\zeta(k)$, $k = 1, 2, 3, \ldots$, згідно з термінологією [8, 9], можна розглядати, як спеціальний приклад "прихованої маркоської моделі". В роботі [2] показано, що цей випадковий процес є однією з координат "прихованої маркоської моделі" $(\zeta(k), \eta(k))$, $k = 0, 1, 2, \ldots$, де $\eta(k)$, $k = 0, 1, 2, \ldots$ деякий ланцюг Маркова, та знайдено перехідні імовірності "прихованого" ланцюга.

Розглянемо задачу непараметричного оцінювання для описаної щойно моделі в умовах неповній статистичній інформації. А саме, припустимо, що безпосередньому спостереженню доступний лише випадковий процес $\zeta(k)$, $k = 1, 2, 3, \ldots$, а

$$\{\zeta(1), \zeta(2), \ldots, \zeta(N)\}$$

числовий вектор, що складається з N послідовних його значень. При цьому значення ланцюга Маркова $x(t)$, $t \geq 0$, та пуасонівського процесу $\Pi(t)$, $t \geq 0$, в інтервалі спостереження невідомі. Відома лише їх стохастична структура, тобто матриця $A = \|a_{ij}\|$, $i, j = 1, 2, \ldots, m$, а також вектор $\pi = \{\pi_1, \pi_2, \ldots, \pi_m\}$ початкового розподілу для ланцюга Маркова $x(t)$, $t \geq 0$, і параметр λ пуасонівського процесу $\Pi(t)$, $t \geq 0$.

Нехай a — деяке фіксоване дійсне число. Введемо вектор:

$$u^{(a)} = \{u_1^{(a)}, u_2^{(a)}, \ldots, u_m^{(a)}\},$$

де

$$u_i^{(a)} = U_i(a) = \mathsf{P}\{\zeta_i < a\}, \qquad i \in E = 1, 2, \ldots, m, \ -\infty < a < +\infty.$$

Необхідно на підставі вектора спостережень

$$\{\zeta(1), \zeta(2), \ldots, \zeta(n), \zeta(n+1), \ldots, \zeta(n+m)\}$$

збудувати спроможні та асимптотично нормальні оцінки $\hat{u}_n^{(a)}$ вектора $u(a)$.

В роботі автора [10] запропоновано метод статистичного оцінювання параметрів прихованих марковських моделей. Приклад його застосування наведено в [11]. Використаємо результати роботи [10], а також отримане в [2] зображення процесу $\zeta(k)$, $k = 1, 2, 3, \ldots$ за допомогою "прихованої маркоської моделі", щоб розв'язати сформульовану вище задачу непараметричного оцінювання.

Справедливе наступне твердження. Нехай A — деяка випадкова подія, а символ $\chi[A]$ означає її індикатор. Введемо сукупність статистик

$$\{Q_{sn}(y, z), s = 1, 2, \ldots, m\}, \qquad n = 1, 2, \ldots,$$

де y та z — довільні фіксовані дійсні числа, наступним чином:

$$Q_{sn}(y, z) = \frac{1}{n} \sum_{t=1}^{n} \chi[\zeta(t) < y, \zeta(t+s) < z], \qquad s = 1, 2, \ldots, m.$$

Нехай також $\{q_{sn}^{(a)}, s = 1, 2, \ldots, m\}$, $n = 1, 2, \ldots$, означає частковий випадок сімейства статистик $\{Q_{sn}(y, z)\}$, коли $y = z = a$, тобто:

$$q_{sn}^{(a)} = Q_{sn}(a, a), \qquad s = 1, 2, \ldots, m.$$

Визначимо на вихідному імовірносному просторі $(\Omega, \Im, \mathsf{P})$ дискретний, однорідний ланцюг Маркова $\eta(k)$, $k = 0, 1, 2, \ldots$, з множиною станів $E = \{1, 2, \ldots, m\}$, який не залежить від ланцюга $x(t)$, не залежить від пуасонівського процесу $\Pi(t)$, а також не залежить від сімейства випадкових величин $\{\zeta_i, i \in E\}$. Перехідні імовірності

$$P = \|p_{ij}\|, \qquad i, j = 1, 2, \ldots, m,$$

ланцюга $\eta(k)$ за один крок визначаються з системи лінійних рівнянь:

$$p_{ij} - \delta_{ij} = \sum_{k=1}^{m} p_{ik} \cdot \frac{a_{kj}}{\lambda},$$

$$i = 1, 2, \ldots, m, \qquad j = 1, 2, \ldots, m,$$

де

$$\delta_{ij} = 1, i = j, \qquad \delta_{ij} = 0, i \neq j.$$

Нехай

$$P(s) = \|p_{ij}^{(s)}\| = P^s$$

означає матрицю перехідних імовірностей ланцюга Маркова $\eta(k)$ за s кроків.

Теорема 2.1. *Припустимо, що виконані наступні умови:*

1. Ланцюг Маркова $\eta(k)$, $k = 0, 1, 2, \ldots$ є незвідним та неперіодичним.

2. Власні числа $\{b_1, b_2, \ldots, b_m\}$ матриці P — різні, відмінні від нуля, дійсні числа.

3. Якщо $\pi = \{\pi_1, \pi_2, \ldots, \pi_m\}$ — вектор стаціонарного розподілу ланцюга Маркова $\eta(k)$, то для довільних $i = 1, 2, \ldots, m$, та $j = 1, 2, \ldots, m$, виконується рівність:

$$\sum_{k=1}^{m} \frac{\pi_k}{\sqrt{\pi_i \cdot \pi_j}} \cdot p_{ki} \cdot p_{kj} = \sum_{k=1}^{m} \frac{\sqrt{\pi_i \cdot \pi_j}}{\pi_k} \cdot p_{ik} \cdot p_{jk} \tag{1}$$

Тоді з імовірністю, яка прямує до одиниці при умові, що $n \to \infty$, система рівнянь:

$$\sum_{i=1}^{m} \sum_{j=1}^{m} \pi_i \cdot p_{ij}^{(s)} \cdot u_i \cdot u_j = q_{sn}^{(a)}, \qquad s = 1, 2, \ldots, m, \tag{2}$$

має дійсні розв'язки.

Доведення. Доведення теореми випливає безпосередньо з результатів робіт [10] та [2]. □

В роботі [10] не тільки показано, що виконання умови (1) для ланцюга Маркова $\eta(k)$ в прихованій марковській моделі $(\zeta(k), \eta(k))$, $k = 0, 1, 2, \ldots$, гарантує існування розв'язків система рівнянь (2), але й запропоновано метод для їх знаходження. Як випливає з результатів [10], один серед тих розв'язків буде спроможною оцінкою вектора

$$u^{(a)} = \{u_1^{(a)}, u_2^{(a)}, \ldots, u_m^{(a)}\}.$$

Тому для розв'язування системи рівнянь (2) можна використати описаний в [10] метод, а для вибору серед можливих розв'язків спроможної оцінки

$$\hat{u}_n^{(a)} = \{\hat{u}_{n1}^{(a)}, \hat{u}_{n2}^{(a)}, \ldots, \hat{u}_{nm}^{(a)}\}$$

вектора $u^{(a)}$ застосувати запропоновану там відповідну статистичну процедуру.

В [10] досліджено також властивості оцінок, отриманих з системи рівнянь (2). На підставі доведених там загальних результатів можна стверджувати, що вектор оцінок $\hat{u}_n^{(a)}$ визначених теоремою 2.1 має асимптотично нормальний розподіл.

Велика кількість параметрів, що визначають модель, яка досліджується, в поєднанні з досить складним методом оцінювання, описаним в теоремі 2.1, створює ситуацію, коли детальний аналіз у загальному випадку запропонованої статистичної процедури та дослідження властивостей отриманих оцінок є практично неможливими. Теорема 2.1 вказує обгрунтований теоретичними результатами [10] підхід до розв'язку поставленої проблеми. З практичної точки зору більш доцільним буде проведення детального аналізу отриманих оцінок в конкретних ситуаціях, коли числові значення параметрів моделі відомі. Тому на закінчення статті, з метою ілюстрації описаного в теоремі 2.1 методу, застосуємо його до оцінювання при неповних спостереженнях невідомих функцій розподілу складного пуасонівського процесу, що керується простим марковським процесом регенерації. Іншими словами розглянемо частковий випадок моделі і проведемо детальний аналіз запропонованої статистичної процедури в найпростішій ситуації, коли $m = 2$.

Для цього на початку розглянемо задачу статистичного оцінювання невідомих функцій розподілу в моделі Кокса [5], а саме, нехай $(c(t), \tau(t))$ — випадковий процес з дискретним часом із значеннями в $\{1, 2\} \times (-\infty, \infty)$. При цьому $c(t) \in \{1, 2\}$ - однорідний ланцюг Маркова з матрицею C перехідних імовірностей за один крок:

$$C = \|c_{ij}\|, \qquad i \in \{1, 2\}, \qquad j \in \{1, 2\},$$

де

$$c_{11} = 1 - \alpha, \qquad c_{12} = \alpha,$$

$$c_{21} = \beta, \qquad c_{22} = 1 - \beta.$$
$$0 < \alpha < 1; \qquad 0 < \beta < 1,$$

а $\tau(t)$ — процес умовно незалежних випадкових величин, визначених на ланцюгу Маркова $c(t)$, тобто:

$$\tau(t) = \tau_{c(t)}^{(t)}, \qquad t = 1, 2, \dots,$$

де $\{\tau_i^{(t)}, i \in \{1, 2\}, t = 1, 2, \dots\}$ — сімейство незалежних від ланцюга $c(t)$ випадкових величин, незалежних для різних $t = 1, 2, \dots$ і таких, що розподіл:

$$T_i(x) = \mathsf{P}\{\tau_i^{(t)} < x\}, \qquad t = 1, 2, \dots,$$

не залежить від індексу $t = 1, 2, \dots$. Нехай $c = \{c_1, c_2\}$ — вектор стаціонарного розподілу ланцюга $c(t)$, де

$$c_1 = \frac{\beta}{\alpha + \beta}, \qquad c_2 = \frac{\alpha}{\alpha + \beta},$$

а $\{1, b\}$ — вектор власних чисел матриці C, де:

$$b = 1 - \alpha - \beta.$$

Введемо функції:

$$T^{(0)}(x) = c_1 \cdot T_1(x) + c_2 \cdot T_2(x);$$
$$T^{(1)}(x, y) = \sum_{i=1}^{2} \sum_{j=1}^{2} c_i \cdot c_{ij} \cdot T_i(x) \cdot T_j(y);$$
$$T^{(2)}(x, y) = \sum_{i=1}^{2} \sum_{j=1}^{2} c_i \cdot c_{ij}^{(2)} \cdot T_i(x) \cdot T_j(y);$$

де

$$C^{(2)} = C^2 = \|c_{ij}^{(2)}\|, \qquad i \in \{1, 2\}, \qquad j \in \{1, 2\},$$

перехідні імовірності ланцюга Маркова $c(t)$ за два кроки. Позначимо:

$$T^{(1)}(x) = T^{(1)}(x, x);$$
$$T^{(2)}(x) = T^{(2)}(x, x).$$

Нехай тепер

$$\{\tau(1), \tau(2), \dots, \tau(n), \tau(n+1), \tau(n+2)\}$$

— вектор спостережень процесу $\tau(t)$ на проміжку $t \in \{1, 2, \dots, n+2\}$. На підставі тих спостережень визначимо наступні статистики:

$$\hat{T}_n^{(1)}(x, y) = \frac{1}{n} \sum_{t=1}^{n} \chi[\tau(t) < x, \tau(t+1) < y];$$
$$\hat{T}_n^{(2)}(x, y) = \frac{1}{n} \sum_{t=1}^{n} \chi[\tau(t) < x, \tau(t+2) < y];$$
$$\hat{T}_n^{(0)}(x) = \frac{1}{n} \sum_{t=1}^{n} \chi[\tau(t) < x];$$

і позначимо відповідно:

$$\hat{T}_n^{(1)}(x) = \hat{T}_n^{(1)}(x, x); \qquad \hat{T}_n^{(2)}(x) = \hat{T}_n^{(2)}(x, x).$$

Легко переконатися, що для ланцюга Маркова $c(t) \in \{1, 2\}$ з двома станами завжди виконується умова (1) теорема 2.1. Нехай z означає вектор-стовбчик з координатами z_1 та z_2. Δ означає діагоналну матрицю з числами c_1 та c_2 на діагоналі:

$$\Delta = \mathrm{diag}\{c_1, c_2\}.$$

Тоді система рівнянь (2) для визначення спроможної оцінки вектора

$$T(x) = \{T_1(x), T_2(x)\}$$

приймає вигляд:

$$z^T \cdot \Delta \cdot C \cdot z = \hat{T}_n^{(1)}(x), \qquad z^T \cdot \Delta \cdot C^2 \cdot z = \hat{T}_n^{(2)}(x). \tag{3}$$

Використовуючи результати [10] можна довести, що для довільного дійсного числа x, а також для довільних $i = 1, 2$, має місце збіжність за імовірністю:

$$\hat{T}_n^{(i)}(x) \to T^{(i)}(x). \tag{4}$$

Тому дослідимо більш детально розв'язки системи рівнянь:

$$z^T \cdot \Delta \cdot C \cdot z = T_n^{(1)}(x), \qquad z^T \cdot \Delta \cdot C^2 \cdot z = T_n^{(2)}(x). \tag{5}$$

Система рівнянь (5) є сумісною, оскільки має розв'язок:

$$\{z_1, z_2\} = \{T_1(x), T_2(x)\}.$$

Нехай:

$$\hat{\Delta} = \mathrm{diag}\{\sqrt{c_1}, \sqrt{c_2}\},$$

а ортогональна матриця $U = \|u_{ij}\|$, $i \in \{1, 2\}$, $j \in \{1, 2\}$, визначається наступним чином:

$$u_{11} = \frac{\sqrt{\beta}}{\sqrt{\alpha + \beta}}, \qquad u_{12} = \frac{\sqrt{\alpha}}{\sqrt{\alpha + \beta}},$$
$$u_{21} = -\frac{\sqrt{\alpha}}{\sqrt{\alpha + \beta}}, \qquad u_{22} = \frac{\sqrt{\beta}}{\sqrt{\alpha + \beta}}.$$

Виконаємо послідовно в системі рівнянь (5) дві заміни змінних:

$$\hat{z} = \hat{\Delta} \cdot z,$$

а потім

$$y = U^{-1} \cdot \hat{z}.$$

Внаслідок чого отримаємо наступну систему рівнянь:

$$y_1^2 + b \cdot y_2^2 = T^{(1)}(x), \qquad y_1^2 + b^2 \cdot y_2^2 = T^{(2)}(x).$$

Для матриць C та $C^{(2)}$ можна записати наступні зображення:

$$C = G_1 + b \cdot G_2, \qquad C^{(2)} = G_1 + b^2 \cdot G_2.$$

При цьому $G_1 = \|g_{ij}^{(1)}\|$, $i \in \{1, 2\}$, $j \in \{1, 2\}$. де

$$g_{11}^{(1)} = c_1, \qquad g_{12}^{(1)} = c_2,$$
$$g_{21}^{(1)} = c_1, \qquad g_{22}^{(1)} = c_2,$$

а $G_2 = \|g_{ij}^{(2)}\|$, $i \in \{1, 2\}$, $j \in \{1, 2\}$, де

$$g_{11}^{(2)} = c_1, \qquad g_{12}^{(2)} = -c_2,$$
$$g_{21}^{(2)} = -c_1, \qquad g_{22}^{(2)} = c_2.$$

Тоді:

$$A = \hat{\Delta} \cdot C \cdot \hat{\Delta}^{-1} = \hat{G}_1 + b \cdot \hat{G}_2,$$

елементи матриці $\hat{G}_1 = \|\hat{g}_{ij}^{(1)}\|$, $i \in \{1, 2\}$, $j \in \{1, 2\}$. визначаються рівностями:

$$\hat{g}_{11}^{(1)} = c_1, \qquad \hat{g}_{12}^{(1)} = \sqrt{c_1 \cdot c_2},$$
$$\hat{g}_{21}^{(1)} = \sqrt{c_1 \cdot c_2}, \qquad \hat{g}_{22}^{(1)} = c_2,$$

а матриці $\hat{G}_2 = \|\hat{g}_{ij}^{(2)}\|$, $i \in \{1, 2\}$, $j \in \{1, 2\}$, відповідно:

$$\hat{g}_{11}^{(2)} = c_1, \qquad \hat{g}_{12}^{(2)} = -\sqrt{c_1 \cdot c_2},$$

$$\hat{g}_{21}^{(2)} = -\sqrt{c_1 \cdot c_2}, \qquad \hat{g}_{22}^{(2)} = c_2.$$

Введемо вектор-стовбчик $T(x)$ з координатами $T_1(x)$ та $T_2(x)$. Оскільки значення:

$$z_1 = T_1(x), \qquad z_2 = T_2(x)$$

є розв'язками системи рівнянь (5), то справедливі наступні рівності:

$$T(x)^T \cdot A \cdot T(x) = T^{(1)}(x), \qquad T(x)^T \cdot A^2 \cdot T(x) = T^{(2)}(x).$$

Позначимо:

$$v_1(x) = T(x)^T \cdot \hat{G}_1 \cdot T(x); \qquad v_2(x) = T(x)^T \cdot \hat{G}_2 \cdot T(x).$$

Оскільки координати вектора $T(x)$ невід'ємні, то очевидно виконується наступна нерівність:

$$v_1(x) > v_2(x).$$

З іншого боку:

$$v_1(x) + b \cdot v_2(x) = T^{(1)}(x), \qquad v_1(x) + b^2 \cdot v_2(x) = T^{(2)}(x),$$

тобто:

$$\frac{T^{(2)}(x) - b \cdot T^{(1)}(x)}{1-b} = g_1(x) > g_2(x) = \frac{T^{(1)}(x) - T^{(2)}(x)}{b \cdot (1-b)}. \tag{6}$$

Розв'язуючи систему рівнянь (5) і аналізуючи з використанням нерівності (6) її додатні розв'язки, одержуємо наступні висновки:

• У випадку, коли:

$$\beta < B_1 = \frac{T^{(1)}(x) - T^{(2)}(x)}{T^{(1)}(x) \cdot (1+b) - T^{(2)}(x)} \tag{7}$$

система рівнянь (5) має тільки один додатній розв'язок $z^{(1)} = \{z_1^{(1)}, z_2^{(1)}\}$, де:

$$z_1^{(1)} = \sqrt{\frac{T^{(2)}(x) - b \cdot T^{(1)}(x)}{\alpha + \beta}} + \sqrt{\frac{c_2}{c_1}} \cdot \sqrt{\frac{T^{(1)}(x) - T^{(2)}(x)}{b \cdot (\alpha + \beta)}}$$

$$z_2^{(1)} = \sqrt{\frac{T^{(2)}(x) - b \cdot T^{(1)}(x)}{\alpha + \beta}} - \sqrt{\frac{c_1}{c_2}} \cdot \sqrt{\frac{T^{(1)}(x) - T^{(2)}(x)}{b \cdot (\alpha + \beta)}}$$

• У випадку, коли:

$$\beta > B_2 = \frac{b \cdot (T^{(2)}(x) - b \cdot T^{(1)}(x))}{T^{(1)}(x) \cdot (1+b) - T^{(2)}(x)} \tag{8}$$

серед розв'язків системи рівнянь (5) буде тільки один додатній $z^{(2)} = \{z_1^{(2)}, z_2^{(2)}\}$, де:

$$z_1^{(2)} = \sqrt{\frac{T^{(2)}(x) - b \cdot T^{(1)}(x)}{\alpha + \beta}} - \sqrt{\frac{c_2}{c_1}} \cdot \sqrt{\frac{T^{(1)}(x) - T^{(2)}(x)}{b \cdot (\alpha + \beta)}}$$

$$z_2^{(2)} = \sqrt{\frac{T^{(2)}(x) - b \cdot T^{(1)}(x)}{\alpha + \beta}} + \sqrt{\frac{c_1}{c_2}} \cdot \sqrt{\frac{T^{(1)}(x) - T^{(2)}(x)}{b \cdot (\alpha + \beta)}}$$

• При виконанні умови

$$B_1 < \beta < B_2 \tag{9}$$

обидва ці розв'язки будуть додатні.

Введемо випадковий вектор $y_n(x) = \{y_{n1}(x), y_{n2}(x)\}$, координати якого визначаються рівностями:

$$y_{n1}(x) = \sqrt{\frac{\hat{T}_n^{(2)}(x) - b \cdot \hat{T}_n^{(1)}(x)}{\alpha + \beta}},$$

$$y_{n2}(x) = \sqrt{\frac{\hat{T}_n^{(1)}(x) - \hat{T}_n^{(2)}(x)}{b \cdot (\alpha + \beta)}}.$$

Враховуючи співвідношення (7), (8), (9), а також збіжність (4) статистик:

$$\{\hat{T}_n^{(1)}(x), \hat{T}_n^{(2)}(x)\}$$

можна довести наступне твердження.

Теорема 2.2. *Припустимо, що виконані наступні умови:*

1. $\alpha + \beta \neq 1$.
2. $\alpha \neq \beta$.

Тоді для будь якого $x \in (-\infty, \infty)$, такого, що:

$$T_1(x) > 0, \qquad T_2(x) > 0,$$

з імовірністю, яка прямує до одиниці при умові, що $n \to \infty$, координати випадкового вектора $y_n(x) = \{y_{n1}(x), y_{n2}(x)\}$ являються дійсними числами і при цьому справедливі наступні співвідношення:

- *Якщо $\beta < B_1$, то вектор:*

$$\hat{y}_n^{(1)} = \{y_{n1}(x) + \sqrt{\frac{\alpha}{\beta}} \cdot y_{n2}(x); y_{n1}(x) - \sqrt{\frac{\beta}{\alpha}} \cdot y_{n2}(x)\}$$

  буде спроможною оцінкою вектора $T(x) = \{T_1(x), T_2(x)\}$.
- *Якщо $\beta > B_2$, то вектор:*

$$\hat{y}_n^{(2)} = \{y_{n1}(x) - \sqrt{\frac{\alpha}{\beta}} \cdot y_{n2}(x); y_{n1}(x) + \sqrt{\frac{\beta}{\alpha}} \cdot y_{n2}(x)\}$$

  буде спроможною оцінкою вектора $T(x) = \{T_1(x), T_2(x)\}$.
- *У випадку, коли $B_1 < \beta < B_2$, то спроможною оцінкою вектора*

$$T(x) = \{T_1(x), T_2(x)\}$$

  буде один з векторів $\hat{y}_n^{(1)}$, чи $\hat{y}_n^{(2)}$.

Зауваження 2.1. Як виникає з теореми 2, не виключена ситуація, коли спроможна оцінка вектора $T(x)$ визначається неоднозначно, тобто декілька розв'язків системи рівнянь (3) можуть бути такою оцінкою. В роботі [10] показано, що у загальному випадку "прихованої маркоської моделі"

$$(\zeta(k), \eta(k)), \qquad k = 0, 1, 2, \ldots$$

з множиною станів $E = \{1, 2, \ldots, m\}$, при виконані умови:

$$p_{i_1} \cdot T_{i_1}^{(0)} + \cdots + p_{i_r} \cdot T_{i_r}^{(0)} \neq 0 \tag{10}$$

для довільних $1 \leq r \leq m$, $i_1, i_2, \ldots, i_r \in E = \{1, 2, \ldots, m\}$, де:

$$p = \{p_1, p_2, \ldots, p_m\} = U^T \cdot \hat{\Delta} \cdot c,$$

$$T^{(0)}(x) = \{T_1^{(0)}(x), T_2^{(0)}(x), \ldots T_2^{(0)}(x)\} = U^T \cdot \hat{\Delta} \cdot T(x),$$

виділити серед розв'язків системи рівнянь (3) спроможну оцінку вектора $T(x)$ можна за допомогою статистики $\hat{T}_n^{(0)}(x)$. У випадку, який розглядається в теоремі 2.2, умова (10) не виконується. Дійсно:

$$p = \{p_1, p_2\} = U^T \cdot \hat{\Delta} \cdot c = \{1, 0\},$$

$$\begin{aligned} T^{(0)}(x) &= \left\{T_1^{(0)}(x), T_2^{(0)}(x)\right\} = U^T \cdot \hat{\Delta} \cdot T(x) \\ &= \{c_1 \cdot T_1(x) + c_2 \cdot T_2(x), \sqrt{c_1 \cdot c_2} \cdot (T_2(x) - T_1(x))\}. \end{aligned}$$

А отже:

$$p_2 \cdot T_2^{(0)}(x) = 0.$$

Тому скористатися зі статистики $\hat{T}_n^{(0)}(x)$ до вибору відповідної оцінки неможливо. Спроможну оцінку вектора $T(x)$ серед двох можливих, які пропонує теорема 2, можна вибрати використовуючи одну із статистик:

$$\hat{T}_n^{(s,r)}(x) = \frac{1}{n}\sum_{t=1}^{n}\chi[\tau(t) < x, \tau(t+s) < x, \tau(t+s+r) < x];$$
$$1 \le s \le 2, \qquad 1 \le r \le 2.$$

Виникає це з леми 4 праці [10].

Зауваження 2.2. Якщо не виконується умова 1 теореми 2.2 і має місце рівність $\alpha + \beta = 1$, то випадковий процес $\tau(t)$, $t = 1, 2, \dots$ буде послідовністю незалежних випадкових величин і на підставі вектора спостережень $\{\tau(1), \dots, \tau(n+2)\}$ другої компоненти випадкового процесу $(c(t), \tau(t))$, $t = 1, 2, \dots$ можна збудувати оцінку лише для суміші:

$$c_1 \cdot T_1(x) + c_2 \cdot T_2(x)$$

функцій розподілу $\{T_1(x), T_2(x)\}$.

Зауваження 2.3. Якщо не виконується умова 2 теореми 2.2, тобто $\alpha = \beta$, то збудувати оцінки для невідомих функцій розподілу $\{T_1(x), T_2(x)\}$ на підставі лише вектора спостережень $\{\tau(1), \dots, \tau(n+2)\}$ другої компоненти випадкового процесу $(c(t), \tau(t))$, $t = 1, 2, \dots$ можна, але встановити їх відповідність до станів ланцюга Маркова $c(t)$ неможливо.

Використаємо теорему 2.2 до оцінювання складного пуасонівського процесу, що керується простим марковським процесом регенерації. А саме, припустимо, що на тому самому імовірносному просторі $(\Omega, \Im, P)$ визначено два сімейства випадкових величин:

$$\Xi = \{\xi_k, k = 1, 2, \dots\}; \qquad Z = \{\zeta_k, k = 1, 2, \dots\},$$

а також два випадкових процеси:

- Незалежний від випадкових величин Ξ та Z простий марковський процес регенерації $x(t)$, $t \ge 0$, з параметрами (μ, ν), тобто процес, що має два стани: $x(t) \in E = \{1, 2\}$, а час перебування в них є показниковим, відповідно з параметром μ в стані 1, та з параметром ν в стані 2.
- Незалежний від випадкових величин Ξ та Z, незалежний від випадкового процесу $x(t)$, $t \ge 0$ пуасонівський процес $\Pi(t)$, $t \ge 0$ з параметром λ.

Нехай $S(0) = 0$, а $S(k)$, $k = 1, 2, \dots$, означає випадкову величину, яка має розподіл Ерланга з параметрами (k, λ).

Означення 2.2. Складним пуасонівським процесом, що керується простим марковським процесом регенерації, будемо називати двовимірний випадковий процесс $(\xi(t), \zeta(t))$, $t \ge 0$, визначений наступним чином:

$$\xi(t) = \sum_{k=1}^{\Pi(t)} \xi_k \cdot \chi[x(S(k)) = 1]; \qquad \zeta(t) = \sum_{k=1}^{\Pi(t)} \zeta_k \cdot \chi[x(S(k)) = 2].$$

Визначимо випадковий процес:

$$\theta(k) = \xi_k \cdot \chi[x(S(k)) = 1] + \zeta_k \cdot \chi[x(S(k)) = 2], \qquad k = 1, 2, \dots. \tag{11}$$

Лема 2.1. *Випадковий процес з дискретним часом $\theta(k)$, $k = 1, 2, \dots$ описує наступну схему умовно незалежних випадкових величин, визначених на ланцюгу Маркова:*

$$\theta(k) = \theta_{\eta(k)}^{(k)}, \qquad k = 1, 2, \dots, \tag{12}$$

де $\eta(k)$, $k = 1, 2, \dots$ *-однорідний ланцюг Маркова, що має два стани:*

$$\eta(k) \in E = \{1, 2\},$$

а матриця перехідних імовірностей його за один крок $P = \|p_{ij}\|$, $i, j = 1, 2$, *визначається наступним чином:*

$$p_{11} = \frac{\lambda + \nu}{\lambda + \mu + \nu}, \qquad p_{12} = \frac{\mu}{\lambda + \mu + \nu}$$

$$p_{21} = \frac{\nu}{\lambda + \mu + \nu}, \qquad p_{22} = \frac{\lambda + \mu}{\lambda + \mu + \nu}.$$

$[\{\theta_i^{(k)}, k = 0, 1, 2, \dots\}, i = 1, 2]$ *— незалежна від ланцюга Маркова* $\eta(k)$, *сукупність незалежних між собою для різних значень індексу* k *випадкових величин, таких, що розподіл* $\theta_1^{(k)}$ *не залежить від індексу* k *і співпадає з розподілом* ξ_k, *а розподіл* $\theta_2^{(k)}$ *не залежить від індексу* k *і співпадає з розподілом* ζ_k.

Доведення. Дійсно, порівнюючи рівності (11) та (12) можна ствердити, що

$$\theta(k) = x(S(k)), \qquad k = 1, 2, \dots$$

Використовуючи властивості процесу Пуассона $\Pi(t)$, та стохастичну структуру процесу регенерації $x(t)$, отримуємо:

$$\begin{aligned} p_{ij} &= \mathsf{P}\{\eta(k+1) = j/\eta(k) = i\} = \mathsf{P}\{x(S(1)) = j/\eta(0) = i\} \\ &= \lambda \int_0^\infty e^{-\lambda \cdot t}\, \mathsf{P}\{x(t) = j/x(0) = i\}\, dt = \lambda \tilde{p}_{ij}(\lambda), \qquad i, j = 1, 2. \end{aligned}$$

При цьому $\tilde{p}_{ij}(s)$ означає перетворення Лапласа перехідної функції марковського процесу $x(t)$, тобто:

$$\tilde{p}_{ij}(s) = \int_0^\infty e^{-\lambda \cdot t} p_{ij}(t) dt,$$

де

$$p_{ij}(t) = \mathsf{P}\{x(t) = j/x(0) = i\}.$$

Перехідні функції $p_{ij}(t)$, $i, j = 1, 2$, марковського процесу $x(t)$ задовольняють наступну систему диференційних рівнянь Колмогорова:

$$\frac{dp_{11}(t)}{dt} = -\mu \cdot p_{11}(t) + \nu \cdot p_{12}(t);$$

$$\frac{dp_{22}(t)}{dt} = -\mu \cdot p_{22}(t) + \nu \cdot p_{21}(t).$$

Обчислюючи перетворення Лапласа для кожного з цих рівнянь, а також використовуючи рівності (12), отримаємо:

$$p_{11} - 1 = -\frac{\mu}{\lambda} \cdot p_{11} + \frac{\nu}{\lambda} \cdot p_{12};$$

$$p_{22} - 1 = -\frac{\mu}{\lambda} \cdot p_{22} + \frac{\nu}{\lambda} \cdot p_{21}.$$

Враховуючи те, що матриця $P = \|p_{ij}\|$, $i, j = 1, 2$, є стохастичною, отримаємо доведення леми 2.1. □

3. Висновки

Нехай
$$\{\theta(1), \theta(2), \dots, \theta(n+2)\}$$
вектор спостережень випадкового процесу $\theta(k)$, $k = 1, 2, \dots$ Припустимо, що функції розподілу
$$K(x) = \mathsf{P}\{\xi_k < x\}, \qquad L(x) = \mathsf{P}\{\zeta_i < x\},$$
відповідно сімейств Ξ та Z невідомі. Тоді на підставі леми 2.1 оцінити функції $K(x)$ та $L(x)$ можна безпосередньо використовую теорему 2.2, замінюючи при цьому статистики $\hat{T}_n^{(i)}(x)$, $i = 1, 2$, відповідно статистиками:
$$\hat{q}_n^{(i)}(x) = \frac{1}{n}\sum_{t=1}^{n} \chi[\theta(t) < x, \theta(t+i) < x], \qquad i = 1, 2,$$
та покладаючи:
$$\alpha = \frac{\mu}{\lambda + \mu + \nu}; \qquad \beta = \frac{\nu}{\lambda + \mu + \nu}.$$
Що стосується виконання умов теореми 2.2, то згідно з визначенням марковського процесу регенерації $x(t)$ рівність $\alpha = \beta$ виключається, оскільки така рівність означає, що $\mu = \nu$, тобто $x(t)$, $t \geq 0$ не буде процесом регенерації, а простим пуасонівським процесом з параметром μ. Оскільки
$$0 < \frac{\mu + \nu}{\lambda + \mu + \nu} < 1,$$
то умова $\alpha + \beta \neq 1$ теж виконується.

Література

1. В. В. Анисимов, *Случайные процесы с дискретной компонентой. Предельные теоремы*, “Вища школа”, Київ, 1988.
2. А. А. Война, *Асимптотическая оптимизация для стохастических моделей построенных на основании сложного пуассоновского процесса*, Кибернетика и системный анализ (2011), № 4, 11–23.
3. И. И. Гихман, А. В. Скороход, *Теория случайных процессов*, т. 1, “Наука”, Москва, 1971; т. 2, “Наука”, Москва, 1973.
4. И. Н. Коваленко, *Анализ редких событий при оценке эффективности и надежности систем*, Москва, 1980.
5. Д. Кокс, П. Льюис, *Статистический анализ последовательностей событий*, “Мир”, Москва, 1969.
6. В. С. Королюк, А. Ф. Турбин, *Математические основы фазового укрупнения сложных систем*, “Наукова думка”, Київ, 1978.
7. Д. С. Сильвестров, *Полумарковские процессы с дискретным множеством состояний*, Москва, 1980.
8. M. Tim Jones, *Al Application Programming*, Charles River Media, Inc, Hingham, Massachusetts, 2003.
9. L. R. Rabiner and B. H. Juang, *An introduction to Hidden Markov Models*, IEEE ASSP Mag. **3** (1986), no. 1, 4–16.
10. A. A. Voina, *Statistical estimation in a scheme of random variables on on Markov chains, with incomplete observations*, Theor. Prob. and Math. Statist. **37** (1988), 19–28.
11. A. A. Voina and E. Czapla, *An application of the correlation structure of a Markov chain for the estimation of shift parameters in queuing systems*, Theor. Prob. and Math. Statist. **71** (2005), 53–61.
12. A. Wojna, *Ryzyko w Procesach Finansowych Oraz Metody Badania Koniunktury*, PK, Koszalin, 2009.

Інститут економіі та управління, Кошалінський технічний університет, вул. Квятковського, 6E, Кошалін, 75-343, Польща

Адреса електронної пошти: avoina@hotmail.com

Надійшла 09/09/2010