УДК 519.24

Д.В. КОРОЛЮК

КЛАССИФИКАЦИЯ БИНАРНЫХ ДЕТЕРМИНИРОВАННЫХ СТАТИСТИЧЕСКИХ ЭКСПЕРИМЕНТОВ С НАСТОЙЧИВОЙ РЕГРЕССИЕЙ

Аннотация. Для моделей бинарных детерминированных статистических экспериментов, задаваемых рекуррентно решениями разностных детерминированных уравнений для вероятностей альтернатив бинарных состояний, изучается классификация в зависимости от значений параметров направляющего действия, которые определяют функции регрессии приращений вероятностей альтернатив. Классификация обоснована предельными свойствами решений детерминированных разностных уравнений, порождающих вероятности альтернатив.

Ключевые слова: статистические эксперименты, разностные детерминированные уравнения для вероятностей альтернатив бинарних состояний, настойчивая регрессия, эквилибриум, притягивающее состояние, отталкивающее состояние, доминирование альтернатив.

В работе [1] исследована динамика рекуррентных статистических экспериментов с настойчивой регрессией в схеме серий с параметром серии N (объем выборки), $N \to \infty$. При этом существенно использовано условие, обеспечивающее наличие равновесного состояния, которое задается эквилибриумом функции регрессии.

В настоящей статье рассматривается динамика статистических экспериментов (СЭ) при возрастании времени $k \to \infty$. С учетом математических моделей популяционной генетики [2, 3] функция регрессии приращений вероятностей альтернатив задается произведением линейной компоненты, определяемой параметрами направляющего действия $V_{\pm}$, и нелинейной компоненты, обеспечивающей корректность определения функции регрессии:

$$C_0(P) = P_+ P_- [V_- P_- - V_+ P_+], \; 0 \le P_{\pm} \le 1. \tag{1}$$

При этом возникает проблема классификации моделей СЭ в зависимости от значений параметров направляющего действия $V_{\pm}$, определяющих функцию регрессии приращений СЭ.

Аналогичная задача классификации детерминированных моделей СЭ исследовалась в популяционной генетике [2, 4, 5] с использованием дифференциальных уравнений для вероятностей альтернатив генотипов.

ПОСТАНОВКА ЗАДАЧИ

Последовательность СЭ с настойчивой регрессией задается средними значениями выборки $\delta(k) := (\delta_r(k), 1 \le r \le N)$, $k \ge 0$, независимых в совокупности при фиксированных k и одинаково распределенных по r случайных величин $\delta_r(k)$, принимающих два бинарных значения — 1 или 0:

$$S_N(k) := \frac{1}{N} \sum_{r=1}^{N} \delta_r(k), \; k \ge 0. \tag{2}$$

Значения бинарных величин $\delta_r(k)$ интерпретируются как выбор между двумя альтернативами, которые обозначим символом + (положительная альтернатива: $\delta_r(k) = 1$) или – (отрицательная альтернатива: $\delta_r(k) = 0$).

Очевидно, что СЭ (2) задает среднее значение выбора положительной альтернативы на k-м этапе. Аналогично, среднее значение отрицательной альтернативы определяется как

$$S_N^-(k) := \frac{1}{N} \sum_{r=1}^{N} I\{\delta_r(k) = 0\}, \; k \ge 0.$$

Здесь по определению $I(A)$ является индикатором случайного события A. Имеет место очевидное тождество

$$S_N(k) + S_N^-(k) = 1 .$$

Введем вероятности альтернатив:

$$P_+(k) := P\{\delta_r(k) = 1\} = 1 - P\{\delta_r(k) = 0\} = 1 - P_-(k), \ k \ge 0 . \quad (3)$$

По определению имеем

$$P_+(k) + P_-(k) = 1, \ k \ge 0 .$$

Динамика вероятностей альтернатив определяется функцией регрессии приращений вероятностей (1).

Предположение 1. Вероятности альтернатив задаются решениями разностных детерминированных уравнений (РДУ):

$$\Delta P_\pm(k+1) = \pm P_+(k) P_-(k) [V_- P_-(k) - V_+ P_+(k)], \ k \ge 0 ,$$
$$\Delta P_\pm(k+1) := P_\pm(k+1) - P_\pm(k) . \quad (4)$$

Числовые параметры функций регрессии приращений вероятностей удовлетворяют дополнительным условиям

$$|V_\pm| < 1, \ |V_+ + V_-| > 0 , \quad (5)$$

обеспечивающим корректность задания вероятностей альтернатив РДУ (4).

Отметим, что функция регрессии в (4) имеет два поглощающих состояния: $\rho_\pm^0 = 0$, $\rho_\mp^0 = 1$, а также равновесное состояние (эквилибриум) (ρ_+, ρ_-), которое задается соотношением

$$V_- \rho_- = V_+ \rho_+ . \quad (6)$$

Таким образом, основной эквилибриум определяется соотношениями

$$\rho_\pm = V_\mp / V, \ V = V_+ + V_-, \ |V| > 0 . \quad (7)$$

Введем флуктуации вероятностей альтернатив

$$\hat{P}_\pm(k) := P_\pm(k) - \rho_\pm, \ k \ge 0. \quad (8)$$

Предположение 2. Флуктуации вероятностей альтернатив СЭ (3) задаются решениями РДУ

$$\hat{P}_\pm(k+1) - \hat{P}_\pm(k) = -V P_+(k) P_-(k) \hat{P}_\pm(k), \ k \ge 0. \quad (9)$$

Как и в предположении 1, вероятности альтернатив также определяются двумя параметрами (см. (6), (7)): V и $\rho_\pm$ $(\rho_+ + \rho_- = 1)$.

КЛАССИФИКАЦИЯ ДЕТЕРМИНИРОВАННЫХ МОДЕЛЕЙ СЭ

Поведение вероятностей альтернатив, определяемых решениями РДУ (4) или (9) с ограничениями (5), существенно зависит от значений параметров направляющего действия $V_\pm$.

Классификация моделей СЭ с вероятностями альтернатив $P_\pm(k)$, $k \ge 0$, которые задаются решениями РДУ (4), приведена в следующем предложении.

Предложение 1. Параметры направляющего действия $V_\pm$, удовлетворяющие допустимым значениям (5), задают классификацию моделей СЭ:

$\mathbb{M}1$ притягивающая: $V > 0$, $V_{\pm} \geq 0$;
$\mathbb{M}2$ отталкивающая: $V < 0$, $V_{\pm} \leq 0$;
$\mathbb{M}3+$ доминирование $+$: $V \neq 0$, $V_{+} \leq 0 \leq V_{-}$;
$\mathbb{M}3-$ доминирование $-$: $V \neq 0$, $V_{-} \leq 0 \leq V_{+}$.

Аналогичная классификация моделей СЭ, флуктуации вероятностей альтернатив которых (8) задаются решениями РДУ (9), имеет место с учетом значений эквилибриумов $\rho_{\pm}$, а также параметра $V = V_{+} + V_{-}$.

Предложение 2. Равновесные значения альтернатив (эквилибриумы) $\rho_{\pm}$ определяют классификацию моделей СЭ, которые задаются решениями РДУ (9):

$\mathbb{M}1$ притягивающая: $V > 0$, $0 < \rho_{\pm} < 1$;
$\mathbb{M}2$ отталкивающая: $V < 0$, $0 < \rho_{\pm} < 1$;
$\mathbb{M}3+$ доминирование $+$: $V \neq 0$, $V\rho_{-} \leq 0 \leq V\rho_{+}$;
$\mathbb{M}3-$ доминирование $-$: $V \neq 0$, $V\rho_{+} \leq 0 \leq V\rho_{-}$.

Вместе с тем возможна классификация моделей СЭ с учетом значений эквилибриума разности альтернатив $\rho = \rho_{+} - \rho_{-}$.

Предложение 3. Равновесные значения разности альтернатив $\rho = \rho_{+} - \rho_{-}$ задают следующую классификацию моделей СЭ:

$\mathbb{M}1$ притягивающая: $|\rho| \leq 1$, $V > 0$;
$\mathbb{M}2$ отталкивающая: $|\rho| \leq 1$, $V < 0$;
$\mathbb{M}3+$ доминирование $+$: $|\rho| \geq 1$, $V\rho > 0$;
$\mathbb{M}3-$ доминирование $-$: $|\rho| \geq 1$, $V\rho < 0$.

Замечание. Классификация моделей СЭ, задаваемых вероятностями альтернатив $P_{\pm}(t)$, $t \geq 0$ (в непрерывном времени), с функцией регрессии $Q(x) = x(1-x)(ax+b)$, содержится в [5] (см. также [2]) и соответствует классификации, приведенной в предложении 3.

ОБОСНОВАНИЕ КЛАССИФИКАЦИИ МОДЕЛЕЙ СЭ

Основанием классификации СЭ, приведенной в предложениях 1–3, является предельное поведение вероятностей альтернатив СЭ, которое формулируется в теореме.

Теорема 1. Вероятности альтернатив $P_{\pm}(k)$, $k \geq 0$, задаваемых решениями РДУ (4) или (9), имеют следующее асимптотическое поведение:

- в модели $\mathbb{M}1$ (притягивающая)

$$\lim_{k\to\infty} P_{\pm}(k) = \rho_{\pm}; \tag{10}$$

- в модели $\mathbb{M}2$ (отталкивающая)

$$\lim_{k\to\infty} P_{\pm}(k) = 1 \quad \text{или} \quad \lim_{k\to\infty} P_{\mp}(k) = 0 \tag{11}$$

при начальных условиях

$$P_{\pm}(0) > \rho_{\pm} \quad \text{или} \quad P_{\mp}(0) < \rho_{\mp}; \tag{12}$$

- в модели $\mathbb{M}3\pm$ (доминирование $\pm$)

$$\lim_{k\to\infty} P_{\pm}(k) = 1, \quad \lim_{k\to\infty} P_{\mp}(k) = 0. \tag{13}$$

Представляет интерес интерпретация, а также мотивация разных моделей СЭ в практических применениях, например в популяционной генетике [4, 5], экономике и моделях поведения [6, 7].

Рассмотрим интерпретацию классификации моделей СЭ с точки зрения моделей поведения.

С учетом функции регрессии приращений альтернатив (1) классификация моделей СЭ определяется значениями параметров направляющего действия $V_{\pm}$, характеризующих стимулы и сдерживания приращений вероятностей альтернатив [7].

В модели $\mathbb{M}1$ (притягивающая): $V > 0$, $V \pm \geq 0$, вероятности положительной альтернативы $P_+(k)$, $k \geq 0$, возрастают (стимулируются) пропорционально вероятности отрицательной альтернативы с параметром V_- и уменьшаются (сдерживаются) пропорционально вероятности положительной альтернативы с параметром V_+. Такая характеризация стимулов и сдерживаний приводит к наличию стационарного режима, определяемого эквилибриумом $\rho_{\pm}$ функции регрессии приращений.

В модели $\mathbb{M}2$ (отталкивающая): $V < 0$, $V \pm \leq 0$, характеризация стимулов и сдерживаний имеет обратное действие на вероятности альтернатив. Вследствие этого эквилибриум функции регрессии приращений становится отталкивающим, т.е. вероятности альтернатив стремятся к поглощающим состояниям $\rho_{\pm} = 0$ или 1. Отталкивающее равновесное состояние $\rho_{\pm}$ выполняет функцию ограничительного порога. При начальной вероятности альтернатив меньше этого порога следующие вероятности уменьшаются вплоть до поглощающего состояния $\rho = 0$, а больше порогового значения $\rho_{\pm}$ — возрастают вплоть до поглощающего состояния $\rho = 1$.

В модели $\mathbb{M}3\pm$ (доминирование $\pm$), $|\rho_+ - \rho_-| \geq 1$, эквилибриум функции регрессии приращений $\rho_{\pm}$ находится вне пределов отрезка $(0, 1)$. Таким образом, вероятности альтернатив, стремясь (притягиваясь или отталкиваясь) к эквилибриумам, в результате достигают поглощающих состояний $\rho_{\pm} = 0$ или 1. Причем в модели $\mathbb{M}3+$ (доминирование $+$) вероятности положительных альтернатив стремятся к 1, тогда как отрицательных — к 0, в модели $\mathbb{M}3-$ (доминирование $-$) вероятности отрицательных альтернатив стремятся к 1.

В связи с этим возникают статистическая задача проверки гипотез той или иной модели СЭ в реальных экспериментах, а также задачи оценивания параметров направляющего действия СЭ, задаваемых РДУ (4) или (9).

ДОКАЗАТЕЛЬСТВО ТЕОРЕМЫ 1

Основная идея доказательства теоремы 1 заключается в том, что вероятности альтернатив СЭ, задаваемых решениями РДУ (4) или (9), являются ограниченными и монотонными по $k \to \infty$. Это обусловливает существование пределов. Необходимо убедиться в том, что предельные соотношения для вероятностей частот, получаемых из РДУ (4) или (9), обеспечивают утверждение теоремы.

Доказательство основывается на базовом РДУ (см. (9)):

$$\Delta \hat{P}_{\pm}(k+1) = -VP_+(k)P_-(k)\hat{P}_{\pm}(k), \quad k \geq 0. \tag{14}$$

Модель $\mathbb{M}1$. Рассмотрим сначала случай $\mathbb{M}1+$, т.е. $\rho_+ < P_+(0) < 1$. Тогда согласно (8)

$$0 < \hat{P}_+(0) < \rho_-. \tag{15}$$

Из базового РДУ (14) при $V > 0$ вытекает монотонность последовательности $\hat{P}_+(k)$:

$$\hat{P}_+(k+1) < \hat{P}_+(k), \quad k \geq 0. \tag{16}$$

Перепишем РДУ (14) в следующем виде:

$$\hat{P}_+(k+1) = \hat{P}_+(k)[1 - VP_+(k)P_-(k)], \quad k \geq 0. \tag{17}$$

С учетом (5) имеет место неравенство

$$|VP_+ P_-| \leq \frac{1}{2}. \tag{18}$$

Согласно (15)–(18) получаем оценку (14)

$$\hat{P}_+(k+1) \ge \frac{1}{2}\hat{P}_+(k) \ge 0, \; k \ge 0. \tag{19}$$

Монотонность вероятностей (16) с оценкой (19) обеспечивают существование пределов

$$\hat{P}_\pm^* := \lim_{k\to\infty} \hat{P}_\pm(k), \; P_\pm^* := \lim_{k\to\infty} P_\pm(k), \tag{20}$$

для которых имеет место уравнение

$$P_+^* P_-^* \hat{P}_\pm^* = 0, \tag{21}$$

откуда следует (10):

$$\hat{P}_\pm = 0, \text{ т.е. } P_+^* = \rho_+, \; P_-^* = \rho_-.$$

В случае $\mathbb{M}1-$, т.е. $0 < P_+(0) < \rho_+$, в силу соотношения

$$\hat{P}_+(k) + \hat{P}_-(k) = 0, \; k \ge 0, \tag{22}$$

возникает двойственная задача для $\hat{P}_-(k)$: имеет место неравенство (ср. с (15))

$$\rho_- < P_-(0) < 1, \text{ т.е. } 0 < \hat{P}_-(0) < \rho_+.$$

Таким образом, приведенные выше рассуждения справедливы для $\hat{P}_-(0)$ вместо $\hat{P}_+(0)$. В этом случае получаем следующие предельные значення:

$$\hat{P}_-^* = 0, \text{ т.е. } P_-^* = \rho_-, \quad P_+^* = \rho_+,$$

что завершает доказательство (10) теоремы 1 для модели $\mathbb{M}1$.

Модель $\mathbb{M}2$. Рассмотрим сначала случай $\mathbb{M}2+$, т.е. $\rho_+ < P_+(0) < 1$. Исходя из РДУ (9) получаем

$$0 < \hat{P}_+(k) < \rho_-.$$

Согласно (14) при $V < 0$ имеет место монотонность последовательности $\hat{P}_+(k)$:

$$\hat{P}_+(k+1) > \hat{P}_+(k), \; k \ge 0.$$

Соотношения (14), (18) с учетом $V < 0$ дают неравенство

$$\hat{P}_+(k+1) \le \frac{3}{2}\hat{P}_+(k), \; k \ge 0. \tag{23}$$

По индукции из (23) получаем равномерную оценку

$$\hat{P}_+(k+1) \le \frac{3}{2}\rho_-.$$

Следовательно, существуют пределы (20), для которых имеет место уравнение (21), откуда следует (11):

$$P_+^* = 1, \quad P_-^* = 0.$$

В случае $\mathbb{M}2-$, т.е. $0 < P_+(0) < \rho_+$, в силу (22) возникает двойственная задача для $P_-(0)$: имеет место неравенство

$$\rho_- < P_-(0) < 1, \quad \text{т.е.} \quad 0 < \hat{P}_-(0) < \rho_+.$$

Приведенные выше рассуждения для модели $\mathbb{M}2+$ справедливы также для $\hat{P}_-(0)$. Поэтому в данном случае получаем предельные значення (12):

$$P_-^* = 1, \quad P_+^* = 0,$$

что завершает доказательство теоремы 1 для модели $\mathbb{M}2$.

• **Модель** $\mathbb{M}3$. Рассмотрим сначала случай $\mathbb{M}3+\Pi$, т.е. $V>0$, $\rho_+ \geq 1$. Имеет место оценка

$$-\rho_+ < \hat{P}_+(k) < \rho_- < 0, \quad k \geq 0. \tag{24}$$

Монотонное возрастание последовательности

$$\hat{P}_+(k+1) > \hat{P}_+(k), \quad k \geq 0,$$

вытекает из базового уравнения (14), а также неравенства для флуктуаций (24).

Ограниченность сверху

$$\hat{P}_+(k+1) \leq \frac{3}{2}\rho_-$$

следует из неравенств (23) и (24). Следовательно, существуют пределы (20) и имеет место предельное уравнение (21).

Отметим, что $\hat{P}_+^* \neq 0$, поскольку является строго отрицательным в силу (24). Также $P_+^* \neq 0$, так как $P_+(k+1) > P_+(k) > 0$. Следовательно, $P_-^* = 0$, т.е. $\lim_{k\to\infty} P_+(k) = 1$, что завершает доказательство (13) для $\mathbb{M}3+\Pi$.

В случае $\mathbb{M}3-\text{O}$, т.е. $V<0$, $\rho_+ \geq 1$, возникает двойственная задача (см. 22): имеет место неравенство

$$-\rho_- \leq \hat{P}_-(k) \leq \rho_+ .$$

Таким образом, остаются в силе все приведенные выше рассуждения для вероятностей отрицательной альтернативы, откуда следуют соотношения

$$P_-^* = 1, \qquad P_+^* = 0,$$

что завершает доказательство (13) теоремы 1 для модели $\mathbb{M}3-\text{O}$.

Аналогично теорема 1 доказывается в случае $\mathbb{M}3+\text{O}$ и в двойственном случае $\mathbb{M}3-\Pi$.

Теорема 1 доказана.

СПИСОК ЛИТЕРАТУРЫ

1. Королюк Д.В. Бінарні статистичні експерименти з наполегливою нелінійною регресією // Теорія ймовірностей та мат. статистики. — 2014. — № 91. — С. 64–73.
2. Hoppensteadt F. Mathematical methods of population biology. — Cambridge: Cambridge Univ. Press, 1982. — 149 p.
3. Ethier S.N., Kurtz T.G. Markov processes: Characterization and convergence. — New York: Wiley, 1986. — 534 p.
4. Свирежев Ю.М., Пасеков В.П. Основы математической генетики. — М.: Наука, 1982. — 511 с.
5. Skorokhod A.V., Hoppensteadt F., Salehi H. Random perturbation methods with applications in science and engineering. — New York: Springer–Verlag, 2000. — 488 p.
6. Bush R. and Mosteller F. A stochastic model with applications to learning // The Annals of Mathematical Statistics. — 1953. — **24**, N 4. — P. 559–585.
7. Амосов Н.М. Алгоритмы разума. — К.: Наук. думка, 1979. — 224 с.

Поступила 25.07.2014