На правах рукописи

Швемлер Наталья Александровна

ОБНАРУЖЕНИЕ СКАЧКООБРАЗНОГО ИЗМЕНЕНИЯ В СТОХАСТИЧЕСКИХ МОДЕЛЯХ: НАБЛЮДЕНИЯ С РАЗРЫВНОЙ ПЛОТНОСТЬЮ ВЕРОЯТНОСТИ

Специальность 05.13.18 — Математическое моделирование, численные методы и комплексы программ

АВТОРЕФЕРАТ

диссертации на соискание учёной степени кандидата физико-математических наук

Тюмень — 2019

Работа выполнена в федеральном государственном автономном образовательном учреждении высшего образования «Тюменский государственный университет»

Научный руководитель:	**Мосягин Вячеслав Евгеньевич** кандидат физико-математических наук, доцент
Официальные оппоненты:	**Сабельфельд Карл Карлович** доктор физико-математических наук, профессор, Федеральное государственное бюджетное учреждение науки Институт вычислительной математики и математической геофизики Сибирского отделения Российской академии наук, главный научный сотрудник
	Авербух Юрий Владимирович кандидат физико-математических наук, Федеральное государственное бюджетное учреждение науки Институт математики и механики им. Н.Н. Красовского Уральского отделения Российской академии наук, старший научный сотрудник
Ведущая организация:	Федеральное государственное автономное образовательное учреждение высшего образования «Московский физико-технический институт (национальный исследовательский университет)»

Защита состоится 28 января 2020 г. в 15^{00} часов на заседании диссертационного совета Д 003.061.02 при Институте вычислительной математики и математической геофизики Сибирского отделения Российской академии наук по адресу: проспект академика Лаврентьева, 6, 630090, Новосибирск, Россия.

С диссертацией можно ознакомиться в библиотеке и на сайте ИВМиМГ СО РАН, http://icmmg.nsc.ru.

Автореферат разослан 03 октября 2019 года.

Ученый секретарь
диссертационного совета Д 003.061.02,
д.ф.-м.н. Сорокин Сергей Борисович

Общая характеристика работы

Актуальность темы. В практической деятельности приходится сталкиваться с процессами (системами), вероятностные характеристики которых могут измениться в случайный момент времени после вмешательства некоторого фактора. В результате воздействия этого фактора происходит сбой нормального режима функционирования процесса или, так называемая, разладка, приводящая к изменению вероятностных характеристик исходного процесса.

Впервые задача обнаружения изменения свойств случайной последовательности была поставлена Е. С. Пейджом в 1954 г. Сам термин «разладка» возник в теории контроля качества продукции. Впервые этот термин появился в докладе А.Н. Колмогорова и А.Н. Ширяева в 1960 г. В классической формулировке задача о разладке ставится следующим образом. Пусть дана совокупность наблюдений $(X_1, \ldots, X_\theta, X_{\theta+1}, \ldots, X_n)$, первые θ из которых имеют плотность распределение $g_1(x)$, а остальные – плотность распределение $g_2(x) \neq g_1(x)$. Требуется по данным наблюдениям определить параметр θ – момент разладки. Наиболее полное изложение теории и применения многочисленных задач о разладке, опирающихся на методы оптимальной остановки, можно найти в монографии А.Н. Ширяева 2017 г.

Возможен и несколько иной подход в постановке задачи о разладке. Пусть дана выборка из распределения с плотностью

$$f(x,\theta) = \begin{cases} f_1(x,\theta), & x < \theta, \\ f_2(x,\theta), & x > \theta, \end{cases} \tag{1}$$

имеющей разрыв первого рода в неизвестной точке $x = \theta$. Требуется по наблюдениям $(X_1, \ldots, X_n)$ оценить параметр θ или, что тоже, разделить упорядоченные наблюдения $X_{(1)} < \cdots < X_{(n)}$ на два вариационных ряда: $X_{(1)} < \cdots < X_{(i(\theta))}$ и $X_{(i(\theta)+1)} < \cdots < X_{(n)}$, относящиеся к условным распределениям $\frac{f_1(x,\theta)}{\int f_1(x,\theta)\,dx}$ и $\frac{f_2(x,\theta)}{\int f_2(x,\theta)\,dx}$ соответственно. В работе А.А. Боровкова 2018 г. показано, что задачу (1) можно трактовать, как версию задачи о разладке (в точке θ) для некоторого эмпирического процесса в условиях известной локализации параметра θ.

Наблюдения с разрывной плотностью вероятности возникают в теории обнаружения, обработке данных измерений, статистическом контроле, технической и медицинской диагностике. Например, хорошо известно, что случайное время X ремиссии у пациентов описывается экспоненциальным распределением, параметр α которого характеризует частоту происходящих рецидивов, а величина $1/\alpha$ – среднее время ремиссии. В реальной ситуации после начала лечения в случайный момент θ рецидивы могут продолжать происходить, но уже с другой постоянной частотой β. Таким образом происходит разрыв плотности экспоненциального распределения в некоторой точке θ, которую требуется обнаружить. Если мы имеем статистические данные о длительности ремиссий $(X_1, \ldots, X_n)$, то возникает задача обнаружения момента θ проявления лечебного эффекта.

Стохастическая модель с разрывной плотностью вероятности, возникает также и при обучении нейронных сетей. Последняя задача подробно изучена в диссертации.

Целью данной работы является обнаружение момента разладки в стохастических моделях, имеющих разрывную плотность вероятности.

Для достижения поставленной цели необходимо было решить следующие **задачи**:

1. найти предельное распределение нормированной оценки максимального правдоподобия для неизвестной точки разрыва плотности распределения;

2. разработать «ускоренный» алгоритм нахождения оценки максимального правдоподобия для неизвестной точки разрыва плотности распределения;

3. построить асимптотический доверительный интервал для оцениваемого параметра;

4. создать комплекс программ, реализующий разработанные алгоритмы, и провести серию вычислительных экспериментов.

Научная новизна заключается в следующем:

- впервые разработан метод нахождения функции распределения момента достижения максимума обобщенного процесса Пуассона с линейным сносом;

- впервые найдено предельное распределение последовательности нормированных оценок максимального правдоподобия для неизвестной точки разрыва плотности распределения;

- построена стохастическая модель, описывающая процесс обучения рекуррентных нейронных сетей с разрывной плотностью вероятности, позволяющая оценить момент перехода от «эффективного» периода обучения к «неэффективному»;

- впервые разработан метод построения асимптотического доверительного интервала для неизвестной точки θ разрыва плотности распределения и создан комплекс программ для его вычисления.

Методы исследования. Для решения поставленных задач и доказательства сформулированных утверждений применялись асимптотические методы теории вероятностей, математической статистики и случайных процессов.

Теоретическая и практическая значимость. В задачах об оценке неизвестной точки разрыва плотности распределения впервые найдено предельное распределение нормированных оценок максимального правдоподобия. Разработанные в диссертации методы позволяют обнаружить скачкообразные изменения в стохастических моделях прикладных задач, описываемых разрывной плотностью вероятности.

Достоверность полученных результатов подтверждается строгими математическими доказательствами и проведением численных экспериментов на тестовых моделях.

Основные положения, выносимые на защиту:

- метод нахождения функции распределения момента достижения максимума обобщенного процесса Пуассона с линейным сносом;
- предельное распределение последовательности нормированных оценок максимального правдоподобия для неизвестной точки разрыва плотности распределения;
- стохастическая модель, описывающая процесс обучения рекуррентных нейронных сетей с разрывной плотностью распределения, позволяющая оценить момент перехода от «эффективного» периода обучения к «неэффективному»;
- численный алгоритм нахождения состоятельной оценки и оценки максимального правдоподобия для неизвестной точки разрыва плотности распределения;
- метод нахождения точного асимптотического доверительного интервала для неизвестной точки разрыва плотности распределения;
- комплекс проблемно-ориентированных программ для моделирования и обнаружения момента изменения свойств стохастических систем, имеющих разрывную плотность вероятности.

Апробация работы. Основные результаты работы докладывались на:

VI International Conference, Modern Problems in Theoretical and Applied Probability (г. Новосибирск, РИЦ НГУ 22–25 августа 2016 г.); Всероссийская конференция молодых ученых «Математическое и информационное моделирование» (г. Тюмень, ТюмГУ, апрель 2017 г.); Международная конференция, посвященная 60–летию Института математики им. С. Л. Соболева, Математика в современном мире (г. Новосибирск, НГУ, 14–19 августа 2017); Международная конференция «Ломоносовские чтения на Алтае: фундаментальные проблемы науки и образования» (г. Барнаул, АлтГУ, 14–17 ноября 2017); 155 Городской научный семинар «Интеллектуальные информационные системы» (г. Тюмень, 15 января 2018); Международная (49-я, 50-я Всероссийская) молодежная школа-конференция «Современные проблемы математики и ее приложений» (г. Екатеринбург, ИМиМ УрО РАН, 4–10 февраля 2018, 3–9 февраля 2019); Объединенный семинар ИВМиМГ СО РАН (г. Новосибирск, 15 января 2019); Семинар отдела «Математическое моделирование экономических систем» ВЦ РАН ФИЦ ИУ РАН (г. Москва, 10 апреля 2019).

Личный вклад соискателя заключается в обсуждении постановки задачи, разработке и исследовании методов решения, численных алгоритмов, создании и тестировании программ, проведении серии вычислительных экспериментов и анализе полученных результатов.

Публикации. Основные результаты по теме диссертации изложены в 11 работах, 4 из которых опубликованы в журналах, рекомендованных ВАК (3 индексированны в базе данных Web of Science), 1 свидетельство о государственной регистрации программы для ЭВМ (РОСПАТЕНТ), 6 — в тезисах докладов.

Объем и структура работы. Диссертация состоит из введения, трех глав, заключения. Полный объем диссертации 77 страниц текста с 12 рисунками и 3 таблицами. В диссертации доказано 17 лемм и 9 теорем. Список литературы содержит 58 наименований.

Используемая в автореферате нумерация теорем и формул автономна от диссертации.

Содержание работы

Во введении обосновываются актуальность исследований, теоретическая и практическая значимость полученных результатов, формулируются цель, задачи работы, выносимые на защиту положения, дается краткая характеристика структуры диссертации.

Пусть $(X_1, \dots, X_n)$ — вещественная выборка состоящая из независимых случайных величин с общей плотностью распределения $f(x, \theta)$ (относительно меры Лебега). Параметр θ предполагается неизвестным с истинным значением θ_0 и подлежит оцениванию, а плотность – непрерывной по x всюду, за исключением точки $x = \theta$, в которой она имеет разрыв первого рода:

$$0 \neq q(\theta) = f(\theta - 0, \theta) \neq f(\theta + 0, \theta) = p(\theta) \neq 0, \quad \theta \in \Theta.$$

Пусть θ_n^* обозначает оценку максимального правдоподобия (ОМП) параметра θ, то есть одну из точек, удовлетворяющих уравнению:

$$\max\left\{\prod_{i=1}^{n} f(X_i, \theta_n^* - 0), \prod_{i=1}^{n} f(X_i, \theta_n^* + 0)\right\} = \sup_{\theta} \prod_{i=1}^{n} f(X_i, \theta).$$

Через $t_n^* = n(\theta_n^* - \theta)$ обозначим нормированные ОМП.

В монографии И.А. Ибрагимова, Р.З. Хасьминского 1979 г. установлен факт сходимости по распределению $t_n^* \Rightarrow t^*$, где предельная величина t^* является моментом достижения максимума процессом

$$Y(t) = at - \nu_+(pt) + \nu_-(-qt), \tag{2}$$

где

$$a = a(\theta) = \frac{p - q}{\ln(p/q)}, \quad p = p(\theta), \quad q = q(\theta),$$

$\nu_\pm(t)$ — независимые стандартные пуассоновские процессы при $t \geq 0$ и доопределенные нулем на отрицательной оси.

В главе 1 дано математическое описание стохастических моделей с наблюдениями, имеющими разрывную плотность вероятности в неизвестной точке, которую требуется оценить по результатам наблюдений. Найдены интегральные представления для функции распределения

$$G(x) = G(x, \theta) = \mathbf{P}\left(t^* < x\right)$$

на положительной и отрицательной осях.

При $x \geq 0$ справедливо представление

$$G(x) = 1 - \beta \int_0^\infty \varphi^+(x, z) e^{-\beta z}\, dz, \tag{3}$$

где

$$\varphi^+(x, z) = \mathbf{P}\left(\sup_{t<x} Y(t) \leq z + Y(x)\right),$$

а константа $\beta > 0$ является единственным положительным корнем уравнения

$$\left(1 - e^{-\beta}\right) / \beta = a/p. \tag{4}$$

При $x < 0$ справедливо представление:

$$G(x) = \int_0^\infty \varphi^-(x,z)\, d\psi(z), \tag{5}$$

где

$$\varphi^-(x,z) = \mathbf{P}\left(\sup_{t \geq x} Y(t) < z + Y(x)\right),$$

а функция $\psi(z)$ является функцией распределения и определена следующим образом

$$\psi(z) = (1 - q/a) \sum_{k=0}^{[z]} (-1)^k \left(\frac{q}{a}\right)^k \frac{(z-k)^k}{k!} e^{q(z-k)/a}, \tag{6}$$

где $[z]$ – целая часть числа $z \geq 0$.

В первой главе устанавливается непрерывность функций $\varphi^+(x,z)$ и $\varphi^-(x,z)$ по каждой из своих переменных в областях $x, z \geq 0$ и $x \leq 0, z \geq 0$ соответственно. Ключевым утверждением первой главы являются аналитические представления для функций $\varphi^+(x,z)$ и $\varphi^-(x,z)$ в точках $z = 0$ и $z = 1$ соответственно:

$$\varphi^+(x,0) = e^{-px} \sum_{k=0}^{[ax]} \psi(ax - k) \left(\frac{p}{a}\right)^k \left(\frac{(ax)^k}{k!} - \frac{(ax)^{k-1}}{(k-1)!}\right),$$

$$\varphi^-(x,1) = \sum_{k=[|x|]+1}^{\infty} \pi_k(q/a) - e^{b|x|} \sum_{k=[|x|]+1}^{\infty} \pi_k(q/a) e^{-bk},$$

где

$$\pi_k(q/a) = (qk/a)^{k-1} e^{-qk/a}/k!, \quad k = 1, 2, \ldots, \tag{7}$$

$$b = \beta(1 - q/p).$$

Доказано, что знание каждой из функций $\varphi^+(x,0)$ и $\varphi^-(x,1)$ достаточно для нахождения искомой функции распределения $G(x)$ несмотря на полученные представления (3) и (5). Также устанавливается существование п.в. частных производных $\varphi_x^\pm(x,z)$, $\varphi_z^\pm(x,z)$ у функций $\varphi^\pm(x,z)$, которые удовлетворяют дифференциальным уравнениям неклассического типа:

$$\varphi_x^+(x,z) = a\varphi_z^+(x,z) - p\varphi^+(x,z), \quad x > 0,\ 0 < z < 1,$$
$$\varphi_x^+(x,z) = a\varphi_z^+(x,z) - p\varphi^+(x,z) + p\varphi^+(x,z-1), \quad x > 0,\ z \geq 1,$$

с граничными условиями:

$$\varphi^+(0,z) = \psi(z),$$

$$\varphi^+(0,0) = 1 - q/a,$$

$$\varphi^+(x,+\infty) = 1,$$

и

$$\varphi_x^-(x,z) = a\varphi_z^-(x,z) + q\varphi^-(x,z) - q\varphi^-(x,z+1), \quad x < 0,\, z > 0,$$

с граничными условиями:

$$\varphi^-(0,z) = 1 - e^{-\beta z},$$

$$\varphi^-(x,0) = 0,$$

$$\varphi^-(x,+\infty) = 1.$$

В главе 2 из этих уравнений и интегральных представлений (3), (5) находятся обыкновенные дифференциальные уравнения, которым удовлетворяет функция распределения $G(x)$ на положительной и отрицательной осях:

$$G'(x) = a\beta\varphi^+(x,0), \quad x > 0, \tag{8}$$

$$G'(x) = q\,(1 - q/a)\,\varphi^-(x,1), \quad x \leq 0, \tag{9}$$

с граничным условием:

$$G(0) = \frac{(p-a)q}{(p-q)a}$$

.

Из (8) непосредственно вытекает

ТЕОРЕМА 1. *Если $x \geq 0$, то*

$$G(x) = \frac{(p-a)q}{(p-q)a} + \beta \int\limits_0^{ax} e^{-pt/a} \sum_{k=0}^{[t]} \psi(t-k) \left(\frac{p}{a}\right)^k \left(\frac{t^k}{k!} - \frac{t^{k-1}}{(k-1)!}\right) dt,$$

где $\psi(\cdot)$ определена в (6), а константа β в (4).

Из (9) выводится

ТЕОРЕМА 2. *Если $x \leq 0$, то*

$$G(x) = \frac{(p-a)q}{(p-q)a} - \frac{(a-q)q}{a^2} \int\limits_0^{a|x|} \sum_{k=[t]+1}^{\infty} \pi_k(q/a) - e^{bt} \sum_{k=[t]+1}^{\infty} \pi_k(q/a) e^{-bk}\, dt,$$

где $\pi_k(q/a)$ и b определены в (7).

Для положительных z введем обозначение:

$$\Lambda(z) = z - 1 - \ln z > 0.$$

Для функции плотности $g(x) = G'(x)$ случайной величины t^* получены неулучшаемые с точностью до константы экспоненциальные оценки:

$$g(x) \le (a-q)e^{-\Lambda(q/a)a|x|}, \quad x \le 0,$$
$$g(x) \le \beta a e^{-\Lambda(p/a)ax}, \qquad x \ge 0,$$

из которых вытекает существование асимптотического доверительного интервала для неизвестной точки разрыва плотности (1).

Используя этот результат, в конце второй главы выводятся так называемые интегро-локальные оценки для предельного распределения нормированной ОМП:

$$\mathbf{P}\left(x - \delta^- \le t^* \le x + \delta^+\right) \le \left(\delta^- + \delta^+\right) C e^{-c|x|},$$

где $\delta^{\pm} \ge 0$, а положительные константы C и c не зависят от x.

В главе 3 приводятся описание алгоритмов нахождения состоятельной оценки, ОМП, построения асимптотического доверительного интервала, а также комплекса программ реализующего все перечисленные алгоритмы. Излагаются прикладные аспекты полученных теоретических результатов. Проведены экспериментальные исследования по обработке реальных данных с целью апробации комплекса программ.

Пусть $(X_1, \ldots, X_n)$ — вещественная выборка большого объема из распределения с плотностью (1) и вариационным рядом $X_{(1)} \le X_{(2)} \le \cdots \le X_{(n)}$. Так как при $n \to \infty$ истинное (неизвестное) значение параметра θ_0 будет сколь угодно близко к одной из порядковых статистик, то состоятельную оценку следует искать среди членов вариационного ряда с порядковым номером t_n равным аргументу минимума следующей функции $Q(t)$:

$$Q(t) = \left(y_t^- - \frac{S_t}{t}\right)^2 + \left(y_t^+ - \frac{S_n - S_t}{n-t}\right)^2, \quad t = 1, 2, \ldots, n,$$

и определении значения $t_{\min}$ при котором достигается наименьшее значение $Q(t)$. Здесь

$$S_i = \sum_{k=1}^{i} X_{(k)},$$

сумма первых i порядковых статистик,

$$y_i^- = \int^{X_{(i)}} x\, d\mathbf{P}_{X_{(i)}}\left(X_1 < x \,|\, X_1 < X_{(i)}\right), \quad y_i^+ = \int_{X_{(i)}} x\, d\mathbf{P}_{X_{(i)}}\left(X_1 < x \,|\, X_1 > X_{(i)}\right),$$

условные математические ожидания в предположении, что $X_{(i)}$ фиксирована. В качестве состоятельной оценки $\widetilde{\theta}_n$ параметра θ_0 выбирается порядковая статистика $X_{(t_{\min})}$. Предложенный метод является модифицированным приемом одного из подходов описанных в статье Д. Л. Хокинса 1986 г., работе А.А. Боровкова, Ю.Ю. Линке 2005 г.

Трудность численного нахождения ОМП заключается в исследовании на максимум функции правдоподобия:

$$\prod_{i=1}^{n} f(X_i,\theta) = \begin{cases} \prod_{i=1}^{n} f_2(X_{(i)},\theta), \quad \theta \le X_{(1)}, \\ f_1(X_{(1)},\theta)\prod_{i=2}^{n} f_2(X_{(i)},\theta), \quad X_{(1)} < \theta \le X_{(2)}, \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ \prod_{i=1}^{k} f_1(X_{(i)},\theta)\prod_{i=k+1}^{n} f_2(X_{(i)},\theta), \quad X_{(k)} < \theta \le X_{(k+1)}, \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ \prod_{i=1}^{n} f_1(X_{(i)},\theta), \quad X_{(n)} < \theta, \end{cases} \tag{10}$$

которая задается кусочно на $n+1$ интервалах. При большом объеме выборки такой подход приводит к весьма громоздким вычислениям.

Объем вычислений можно значительно сократить, если использовать в качестве начальной оценки параметра θ_0 любую состоятельную оценку $\widetilde{\theta}_n$, $\widetilde{\theta}_n \to \theta_0$, о методе получения которой говорилось выше. По известной состоятельной оценке $\widetilde{\theta}_n$ поиск экстремума функции (10) можно вести только на $O(\ln n)$ интервалах $\big(X_{(k)}, X_{(k+1)}\big]$, расположенных слева и справа от точки $\widetilde{\theta}_n$. Такой прием позволяет значительно «ускорить» процесс численного нахождения оценки максимального правдоподобия, в случае разрывной плотности.

Замена неизвестного параметра θ, от которого зависит функция распределения $G(x,\theta)$, найденной ОМП позволяет найти квантили, необходимые при построении асимптотического доверительного интервала. Доказывается, что построенный доверительный интервал является асимптотически точным.

ТЕОРЕМА 3. *Пусть $\Theta \subset [d_1, d_2] \subset R$, а функции $p(\theta)$ и $q(\theta)$ непрерывны на отрезке $[d_1, d_2]$. Для произвольных положительных чисел ε_1, ε_2, таких что $\varepsilon_1+\varepsilon_2<1$, определим квантили $\lambda^{(n)}_{1-\varepsilon_1}$, $\lambda^{(n)}_{\varepsilon_2} > 0$ функции распределения $G(x,\theta^*_n)$ из уравнений:*

$$\begin{aligned} &G(\lambda^{(n)}_{1-\varepsilon_1}, \theta^*_n) = 1-\varepsilon_1, \\ &G(-\lambda^{(n)}_{\varepsilon_2}, \theta^*_n) = \varepsilon_2. \end{aligned} \tag{11}$$

Тогда интервал $[\theta^*_n - \lambda^{(n)}_{1-\varepsilon_1}/n;\, \theta^*_n + \lambda^{(n)}_{\varepsilon_2}/n]$ является точным асимптотическим доверительным интервалом уровня $1-\varepsilon_1-\varepsilon_2$:

$$\lim_{n\to\infty} \mathbf{P}\left(-\lambda^{(\mathbf{n})}_{\varepsilon_\mathbf{2}} \le \mathbf{n}(\theta^*_\mathbf{n} - \theta_\mathbf{0}) \le \lambda^{(\mathbf{n})}_{\mathbf{1}-\varepsilon_\mathbf{1}}\right) = \mathbf{1} - \varepsilon_\mathbf{1} - \varepsilon_\mathbf{2}. \tag{12}$$

Итак, сформулируем пошагово предлагаемый в диссертации

Алгоритм построения точного асимптотического доверительного интервала неизвестной точки разрыва плотности распределения:

Шаг 1. По выборке методом наименьших квадратов находится состоятельная оценка $\widetilde{\theta}_n$.

Шаг 2. С помощью полученной на предыдущем шаге состоятельной оценки $\widetilde{\theta}_n$ строится оценка максимального правдоподобия θ_n^*.

Шаг 3. В функции распределения $G(x, \theta)$, найденной во второй главе, значение неизвестного параметра θ заменяется оценкой максимального правдоподобия θ_n^*.

Щаг 4. При заданных ε_1, $\varepsilon_2 > 0$, $\varepsilon_1 + \varepsilon_2 < 1$ из уравнений

$$G(\lambda_{1-\varepsilon_1}^{(n)}, \theta_n^*) = 1 - \varepsilon_1,$$
$$G(-\lambda_{\varepsilon_2}^{(n)}, \theta_n^*) = \varepsilon_2.$$

находятся квантили $\lambda_{1-\varepsilon_1}^{(n)}, \lambda_{\varepsilon_2}^{(n)}$ уровней $1 - \varepsilon_1$ и ε_2 соответственно.

Шаг 5. По полученным на предыдущем шаге значениям $\lambda_{1-\varepsilon_1}^{(n)}, \lambda_{\varepsilon_2}^{(n)}$ и найденной на Шаге 2 оценке максимального правдоподобия θ_n^* строится точный асимптотический доверительный интервал $[\theta_n^* - \lambda_{1-\varepsilon_1}^{(n)}/n,\ \theta_n^* + \lambda_{\varepsilon_2}^{(n)}/n]$ уровня $1 - \varepsilon_1 - \varepsilon_2$.

Используя программный комплекс, в котором реализован вышеописанный алгоритм, проведены серии численных экспериментов. Результаты вычислений свидетельствуют о том, что с увеличением объема выборки оценки максимального правдоподобия и доверительные интервалы приближаются к истинному значению параметра, что позволяет судить об адекватности предложенных алгоритмов.

Результаты математического исследования применены к задаче обучения рекуррентных нейронных сетей. В работе построена стохастическая модель, описывающая процесс обучения рекуррентных нейронных сетей с разрывной плотностью вероятности, позволяющая оценить момент перехода от «эффективного» периода обучения к «неэффективному». В ходе экспериментов по тестированию модели были использованы базы данных результатов обучения рекуррентных нейронных сетей ООО «Объединение когнитивных ассоциативных систем». Представленная база содержит 2000 эпох (циклов) обучения с указанием процентов правильного ответа на каждой эпохе. В качестве выборочных данных $X_1, \ldots, X_n$ рассматривалось последовательное приращение процентов правильных ответов данных нейронной сетью. Анализ выборки показал, что сначала наблюдается быстрое обучение нейронной сети, описываемое экспоненциальной плотностью распределения, а с некоторого неизвестного момента характер обучения становится «неэффективным», что выражается нормальным распределением с нулевым средним. Предполагается, что в точке «перехода» $x = \theta$ плотность имеет разрыв первого рода. Это позволяет выдвинуть гипотезу, о том что выборка $X_1, \ldots, X_n$ имеет плотность распределения:

$$f(x,\theta) = \begin{cases} \dfrac{r\, e^{-\frac{x^2}{2\sigma^2}}}{\varPhi(\theta/\sigma)\sqrt{2\pi}\sigma}, & x < \theta, \\ (1-r)\,\alpha\, e^{-\alpha(x-\theta)}, & x > \theta, \end{cases}$$

где σ^2 – дисперсия «неэффективного» периода обучения, α – параметр «эффективного» периода обучения, $0 < r < 1$ предполагаются известными и определяются исходя из анализа выборочных данных. Оценка параметра θ дает возможность определить до какого момента происходит эффективное обучение нейронной сети, что в свою очередь позволяет оптимизировать временные затраты на регулярное переобучение рекуррентных нейронных сетей.

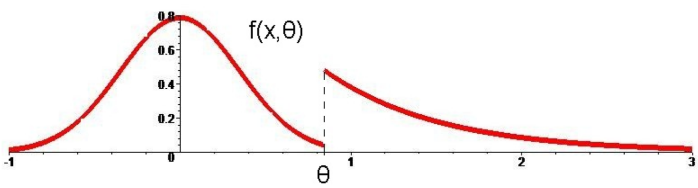

Рис. 1: Плотность $f(x,\theta)$

Параметры σ^2, α, r были определены экспериментально. Тогда соответствующая стохастическая модель обучения принимает вид (Рис. 1):

$$f(x,\theta) = \begin{cases} \dfrac{0,773\, e^{-\frac{x^2}{0,236}}}{\varPhi(\theta/0,344)}, & x < \theta, \\ 0,475\, e^{-1,428(x-\theta)}, & x > \theta, \end{cases}$$

После проведения серии вычислительных экспериментов по оценке неизвестного параметра θ, адекватность модели была подтверждена критерием хи-квадрат. Предложенная стохастическая модель позволяет сократить среднее время обучения при сохранении качества работы.

В заключении сформулированы основные результаты работы:

- построена стохастическая модель, описывающая процесс обучения рекуррентных нейронных сетей с разрывной плотностью вероятности, позволяющая оценить момент перехода от «эффективного» периода обучения к «неэффективному»;

- впервые разработан метод нахождения функции распределения момента достижения максимума обобщенного процесса Пуассона с линейным сносом;
- впервые найдено предельное распределение последовательности нормированных оценок максимального правдоподобия для неизвестной точки разрыва плотности распределения;
- разработан численный алгоритм нахождения состоятельной оценки и оценки максимального правдоподобия для неизвестной точки разрыва плотности распределения;
- впервые разработан метод нахождения асимптотического доверительного интервала для неизвестной точки разрыва плотности распределения и создан комплекс программ для его вычисления;
- проведены вычислительные эксперименты на тестовых моделях с целью проверки адекватности разработанных в диссертации методов и алгоритмов;
- разработанные в диссертации методы позволяют обнаружить скачкообразные изменения в стохастических моделях реальных прикладных задач по наблюдениям с разрывной плотностью вероятности.

В приложении А представлено свидетельство о государственной регистрации программы для ЭВМ в Федеральной службе по интеллектуальной собственности РОСПАТЕНТ.

В приложении Б приводятся копии документов о практической значимости результатов диссертационной работы.

Публикации автора по теме диссертации

Статьи в журналах, включенных в Перечень рецензируемых научных изданий, в которых должны быть опубликованы основные научные результаты диссертаций на соискание ученой степени кандидата наук:

1. Швемлер Н. А. Распределение момента максимума разности двух пуассоновских процессов с отрицательным линейным сносом / В. Е. Мосягин, Н. А. Швемлер // Сибирские электронные математические известия. – 2016. – Т.13. – С. 1229 – 1248. (Web of Science)

2. Швемлер Н. А. Локальные свойства предельного распределения статистической оценки точки разрыва плотности / В. Е. Мосягин, Н. А. Швемлер // Сибирские электронные математические известия. – 2017. – Т.14. – С. 1307 – 1316. (Web of Science)

3. Швемлер Н. А. Асимптотический доверительный интервал для точки разрыва плотности распределения /В. Е. Мосягин, Н. А. Швемлер // Тр. Ин-та математики и механики УрО РАН. – 2018. – Т.24. №2– С. 194 – 199. (Web of Science)

4. Швемлер Н. А. Вероятностный прием преобразования кратных интегралов к интегралам меньшей размерности / В. Е. Мосягин, Н. А. Швемлер // Вестник ТюмГУ. – 2014. №7. Физико-математические науки. Информатика. С. 192 – 198.

Свидетельство о регистрации программы для ЭВМ:

5. Швемлер Н. А. Программа для оценки неизвестного параметра точки разрыва плотности распределения / Н. А. Швемлер, В. Е. Мосягин // Федеральная служба по интеллектуальной собственности РОСПАТЕНТ. Свидетельство №2017660490 от 22.09.2017.

Публикации в других научных изданиях:

6. Швемлер Н. А. Стохастическая модель обучения рекуррентных нейронных сетей / В. Е. Мосягин, Н. А. Швемлер // Современные проблемы математики и ее приложений: тезисы Международной (50-я Всероссийской) молодежной школы-конференции. Екатеринбург: ИМиМ УрО РАН, 2019. – С. 123.

7. Швемлер Н. А. Статистическое оценивание точки разрыва плотности распределения / В. Е. Мосягин, Н. А. Швемлер // Современные проблемы математики и ее приложений: тезисы Международной (49-й Всероссийской) молодежной школы-конференции. Екатеринбург: ИМиМ УрО РАН, 2018. – С. 53

8. Shvemler N. A. Integro-local estimates for the time of attaining the maximum by the Poisson process with linear drift /V. E. Mosyagin, N. A. Shvemler // Математика в современном мире. Международная конференция, посвященная 60-летию Института математики им. С.Л. Соболева, Новосибирск, Россия, 2017 г. – С. 370.

9. Швемлер Н. А. Преобразование Лапласа момента достижения максимума процессом Ибрагимова–Хасьминского / В. Е. Мосягин, Н. А. Швем-

лер // Математическое и информационное моделирование: сборник научных трудов. Тюмень: Изд. ТюмГУ – 2017. – Вып.15. – С. 282 – 284.

10. Швемлер Н. А. Метод усредненного минимума для нахождения состоятельной оценки точки разрыва плотности распределения / В. Е. Мосягин, Н. А. Швемлер // Международная конференция «Ломоносовские чтения на Алтае: фундаментальные проблемы науки и образования». Барнаул: Изд. АлтГУ – 2017. – С. 453 – 455.

11. Shvemler N. A. Distribution on the time of attaining the maximum for a Poisson process with negative drift /V. E. Mosyagin, N. A. Shvemler // Modern Problems in Theoretical and Applied Probability, Novosibirsk, VI International Conference, Новосибирск, РИЦ НГУ – 2016. – P. 41 – 42