На правах рукописи

Степанов Алексей Васильевич

Предельные теоремы и статистические процедуры для величин, связанных с рекордами и экстремальными порядковыми статистиками

01.01.05 – теория вероятностей и математическая статистика

АВТОРЕФЕРАТ

диссертации на соискание ученой степени

доктора физико-математических наук

Санкт-Петербург

2015

Работа выполнена на кафедре прикладной математики института прикладной математики и информационных технологий ФГАОУ ВПО «Балтийский федеральный университет имени Иммануила Канта»

Официальные оппоненты:

Грибкова Надежда Викторовна

доктор физико-математических наук, доцент,

профессор кафедры математики и моделирования ФГБОУ ВПО

«Петербургский государственный университет путей сообщения»

Егоров Владимир Алексеевич

доктор физико-математических наук, профессор,

профессор кафедры высшей математики №2 ФГБОУ ВПО

«Санкт-Петербургский государственный электротехнический

университет «ЛЭТИ» им. В.И. Ульянова (Ленина)»

Чупрунов Алексей Николаевич

доктор физико-математических наук, профессор,

профессор кафедры математического анализа ФГАОУ ВПО

«Казанский (Приволжский) федеральный университет»

Ведущая организация:

ФГБОУ ВПО «Московский государственный университет им. М.В. Ломоносова»

Защита состоится « ____ » ________ 2015 г. в « ____ » часов на заседании диссертационного совета Д 002.202.01 в ФГБУН Санкт-Петербургском отделении Математического института им. В.А. Стеклова Российской академии наук по адресу: 191023 Санкт-Петербург, наб. р. Фонтанки, 27, к. 311.

С диссертацией можно ознакомится в библиотеке и на сайте ФГБУН Санкт-Петербургского отделения Математического института им. В.А. Стеклова Российской академии наук, http://www.pdmi.ras.ru/

Автореферат разослан « ____ » ________ 2015 г.

Ученый секретарь

диссертационного совета Д 002.202.01,

доктор физико-математических наук А.Ю. Зайцев

Общая характеристика работы

Актуальность темы. Теория экстремальных порядковых статистик и рекордов, а эти понятия тесно связаны между собой, актуальна в связи с приложениями, возникающими в различных областях науки, техники, медицины, финансовой и актуарной математики. Верхние порядковые статистики и рекорды интересны, например, в прикладных задачах гидрологии при оценивании параметров. Нижние порядковые статистики находят применение в теории надежности при тестировании изделий, когда по первым вышедшим из строя изделиям делаются статистические выводы о всей испытуемой партии, при этом большинство испытуемых изделий сохраняется и используется в дальнейшем. Анализ появления времен рекордных величин позволяет обнаруживать в выборках нетипичные наблюдения. Наконец, рекордные величины и экстремальные порядковые статистики используются для сокращения выборок.

Порядковые статистики и рекорды индуцируют различные случайные величины. Такими величинами могут быть серии, образуемые порядковыми статистиками и рекордами, конкомитанты порядковых статистик и рекордов, числа величин, регистрируемых около порядковых статистик и рекордов. Серии, индуцированные порядковыми статистиками и рекордами, используются в биостатистике, в теории надежности, при проверке статистических гипотез, при исследовании форм вероятностных распределений. Интерес к конкомитантам порядковых статистик и рекордов проявляется в страховом бизнесе, когда конкомитанты наибольших страховых выплат также могут вызвать большие расходы для страховых компаний, в теории надежности, когда сложные системы состоят из нескольких компонент и выход из строя отдельных компонент системы приводит к прекращению работы всей системы, в биостатистике при исследовании побочных эффектов от действия лекарств. Числа величин, регистрируемых около рекордных величин или верхних порядковых статистик, а также их суммы находят применение в актуарной и финансовой математике. Анализ сумм "больших" страховых требований важен для страховых компаний, так как суммы выплат по таким требованиям могут достигать до 90% от суммы всех выплат.

Интерес к экстремальным порядковым статистикам проявился, по-видимому, впервые в работах Бернулли в начале восемнадцатого века и позднее, в неявной форме, в работах Пуассона, рассматривавшего теорию суммирования и теорию редких событий. В двадцатом веке работы фон Мизеса (von Mises 1923, 1936), Типпета (Tippett 1925), Фреше (Frechet 1927), Гнеденко (Gnedenko 1943), де Хаана (de Haan 1970, 1971) и др. о предельных распределениях экстремумов дали мощный толчок развитию теории экстремальных порядковых статистик. Несомненно, важным для развития теории порядковых статистик явилась возможность использования порядковых статистик в задачах оценивания параметров и проверки статистических гипотез. Частыми являются ситуации, когда границы носителей распределений зависят на одном или обоих концах от оцениваемых параметров (например, равномерные распределения). В этом случае экстремумы являются достаточными и состоятельными оценками, см., например, книги Кендалла и Стьюарта (1973) и Лемана (1979). Известно, что пользуясь методом наименьших квадратов, можно оценивать параметры сдвига и масштаба при помощи порядковых статистик, см., например, работу Ллойда (Lloyd 1952). Порядковые статистики также активно используются для характеризации распределений. Одними из первых работ в этом направлении были работы Россберга (Rossberg 1965), Говиндараджулу (Govindarajulu 1967), Чена (Chan 1967) и Фергюсона (Ferguson 1967).

В начале пятидесятых годов прошлого века в теории экстремальных порядковых статистик выделяется новое направление – теория рекордов. Теория рекордов бурно развивается и насчитывает в настоящее время значительное количество публикаций. Так же как и порядковые статистики, рекорды интересны асимптотическими результатами, статистическими процедурами, характеризациями. Они, наряду с порядковыми статистиками, активно используются в различного рода приложениях.

Значительную роль в развитии теории рекордов сыграли работы Реньи (Rényi 1962), Таты (Tata 1969) и Шоррока (Shorrock 1972), в которых были предложены методы, позволяющие переходить от зависимых рекордных времен и величин к суммам независимых слагаемых, работы Резника (Resnik 1973, 1974, 1975), в которых исследовались предельные распределения рекордов, работы Ахсануллаха (Ahsanullah 1978, 1979) и Нагараджи (Nagaraja 1977, 1978, 1988), в которых были получены первые характеризации, основанные на рекордах, работы Стьюарта (Stuart 1954, 1956), Фостера и Стьюарта (Foster and Stuart 1954), Фостера и Тейчроева (Foster and Teichroew 1955), Саманиего и Витакера (Samaniego and Whitaker 1986, 1988), в которых рекорды впервые использовались в статистических процедурах, работы Невзорова (1985, 1986, 1987, 1989, 1990), в которых изучались моментные свойства рекордов.

Тематика порядковых статистик и рекордов подробно обсуждалась в обзорных статьях Пайка (Pyke 1965), Невзорова (1987), Нагараджи (1988), а также в книгах Галамбоша (1984), Арнольда, Балакришнана и Нагараджи (Arnold, Balakrishnan and Nagaraja 1992, 1998), Невзорова (2000), Ахсануллаха и Невзорова (Ahsanullah and Nevzorov 2001), Дэйвида и Нагараджи (David and Nagaraja 2003) и Бейрланта, Гогебера, Тойгельса и Сегерса (Beirlant, Goegebeur, Teugels and Segers 2004).

В 1987 году В.Б. Невзоровым была защищена докторская диссертация на тему «Упорядоченные случайные величины и их суммы», в которой рассматривалась тематика экстремальных порядковых статистик и рекордов. С того момента прошло 28 лет и теория экстремальных порядковых статистик и рекордов сильно обогатилась новыми результатами, методами и подходами. В настоящей диссертации переосмысливаются и систематизируются эти методы и подходы, анализируются новые приложения.

Цель работы. Основная цель диссертационной работы – исследование асимптотических свойств рекордов, экстремальных порядковых статистик и индуцированных ими случайных величин. Целями данной работы также являются получение характеризационных теорем для рекордов и экстремальных порядковых статистик, анализ статистических гипотез, основанных на рекордных статистиках, и генерирование конкомитантов рекордов и верхних порядковых статистик.

Основные результаты работы и научная новизна. Все основные результаты являются новыми и состоят в следующем.

1) Получены обобщения леммы Бореля–Кантелли. Среди этих обобщений есть обобщения первой части леммы Бореля–Кантелли и обобщения, позволяющие находить необходимые и достаточные условия для сильной сходимости последовательностей, образующих цепи Маркова. Последние обобщения использовались в данной работе для получения сильных предельных теорем для рекордов и верхних порядковых статистик.

2) Предложен метод, позволяющий переходить от зависимых слабых рекордных величин к суммам независимых слагаемых. Этот метод позволил разработать асимптотическую теорию слабых рекордных величин. Теория правильно меняющихся функций использовалась

для исследования асимптотических свойств отношений слабых рекордных величин.

3) Предложены критерии проверки гипотезы однородности, основанные на рекордных величинах. Исследовалась и сравнивалась информация Фишера, содержащаяся в верхних и нижних рекордных величинах и временах. Предложен статистический критерий, основанный на рекордных временах с подтверждением, позволяющий определять, есть ли в выборке нетипичные наблюдения.

4) Разработана асимптотическая теория числа величин, регистрируемых около порядковых статистик и рекордных величин. Для вывода отдельных результатов данной тематики привлекалась теория правильно меняющихся функций. Обсуждалась возможность использования данных исследований в актуарной математике.

5) Выведены предельные теоремы для серий, основанных на порядковых статистиках и рекордах. Предложен статистический критерий, основанный на спейсингах порядковых статистик.

6) Изучены асимптотические свойства конкомитантов порядковых статистик и рекордов. Предложена классификация двумерных распределений, удобная для исследования асимптотических свойств конкомитантов порядковых статистик и рекордов. С помощью этой классификации выведены предельные теоремы для конкомитантов порядковых статистик и рекордов. Предложены методы генерирования конкомитантов верхних порядковых статистик и рекордов.

7) Получены характеризационные теоремы для порядковых статистик и рекордных величин.

Методы исследований. В работе используются классические методы доказательств предельных теорем, основанные на математическом анализе. В работе также применяются традиционные методы математической статистики и используется статистическое моделирование.

Теоретическая и практическая значимость работы. Диссертационная работа носит преимущественно теоретический характер. Она является вкладом в развитие теории рекордов и порядковых статистик. Разработанные в диссертации теоретические методы могут использоваться за рамками данной тематики. Теоретическая ценность работы определяется появлением новых лемм Бореля–Кантелли, удобных для работы с цепями Маркова, классификацией двумерных распределений, позволяющей получать предельные теоремы для конкомитантов рекордов и порядковых статистик, и предельными теоремами для величин, связанных с рекордами и экстремальными порядковыми статистиками. Практическая ценность работы состоит в новых методах генерации рекордов, конкомитантов рекордов и верхних порядковых статистик. Результаты работы могут использоваться в финансовой и актуарной математике.

Апробация работы. Результаты диссертации докладывались на различных международных конференциях и научных семинарах. Среди них были следующие конференции и семинары:

- международная конференция, посвященная 50-летию образования кафедры теории вероятностей и математической статистики СПбГУ, 1998;
- международная конференция "Baltic Sea Seminar," г. Карлскруне, Швеция, 2000;
- 4-ая (2001), 5-ая (2005) и 6-ая (2009) международные конференции по математическому моделированию, проводимые СПбГУ;
- семинар лаборатории прикладной математики университета г. Гавра, Франция, май 2002, руководитель – А. Берред;
- международная мини-конференция, посвященная "50-летию развития теории рекордов," проводимая университетом Мак-Мастер, Канада, 2002;
- международная мини-конференция по порядковым статистикам и рекордам, проводимая Польской академией наук, г. Варшава, Польша, 2003;
- международная конференция "Ordered Statistical Data: Approximations, Bounds and Characterizations (OSD 2004)," проводимая Польской академией наук, г. Варшава, Польша;
- семинар теории вероятностей и математической статистики Варшавского политехнического университета, апрель 2004, руководитель – Я. Веселовский;
- семинар лаборатории теории вероятностей МГУ, март 2006, руководитель – В.И. Питербарг;

- большом семинар кафедры теории вероятностей МГУ, 29 ноября 2006, 27 февраля 2008, руководитель – А.Н. Ширяев;
- семинар ПОМИ РАН им. В.А. Стеклова, 15 декабря 2006, 16 ноября 2007, 25 апреля 2014, руководитель – И.А. Ибрагимов;
- семинар кафедры актуарной математики университета г. Лозанна, Швейцария, 10 ноября 2010, руководитель – Е. Хашорва;
- международная конференция, посвященная 100-летию Б.В. Гнеденко, Москва, МГУ, июнь 2012;
- 2-ая международная конференция "Якоби-2013," проводимая БФУ им. И. Канта.

Публикации. По теме диссертации автором опубликовано 42 статьи, из них 33 статьи опубликованы в рецензируемых российских и международных журналах, входящих в список журналов ВАК. Список работ автора диссертации приведен в конце автореферата.

Личный вклад автора. Многие работы, опубликованные автором диссертации, были написаны в соавторстве с другими учеными. Результаты этих совместных работ, включенные в диссертацию, принадлежат диссертанту.

Структура и объем диссертации. Диссертация состоит из восьми глав и списка литературы из 228 наименований. Первая глава является введением. Общий объем диссертации – 260 страниц.

Содержание работы

В автореферате, при кратком изложении содержания диссертационной работы, нумерация формул, лемм, теорем и иных утверждений совпадает с нумерацией, используемой в соответствующих главах диссертации.

Глава 1 диссертации является введением. В главе 1 делается обзор тематики порядковых статистик и рекордов, описывается структура работы, вводятся основные обозначения и кратко излагается содержание диссертационной работы. Во введении рассматривается предел

$$\lim_{x \to r_F^-} \beta(x,a) = \beta(a) \in [0,1], \tag{1.2.1}$$

где $\lim_{x\to x_0^-} f(x)$ – предел функции $f(x)$ слева в точке x_0, $\beta(x,a) = \bar{F}(x+a)/\bar{F}(x)$, $\bar{F}(x) = 1 - F(x), a > 0$, $r_F = \inf\{x : F(x) = 1\}$ и F – функция распределения. Предел (1.2.1) активно используется в последующих главах диссертации для получения предельных теорем.

В главе 2 диссертационной работы рассматриваются обобщения леммы Бореля–Кантелли. Среди этих обобщений есть и обобщения, предложенные диссертантом. Современный вариант леммы Бореля–Кантелли (лемма 2.1.1) приводится в главе 2.

Лемма 2.1.1 *(1) Пусть $A_1, A_2, \ldots$ – последовательность событий, для которых выполняется неравенство $\sum_{n=1}^{\infty} P(A_n) < \infty$. Тогда $P(A_n$ б.ч.$) = 0$, где сокращение* б.ч. *используется для термина "бесконечно часто".*
(2) Пусть $A_1, A_2, \ldots$ – последовательность независимых событий и $\sum_{n=1}^{\infty} P(A_n) = \infty$. Тогда $P(A_n$ б.ч.$) = 1$.

Из результатов главы 2 приведем в этом параграфе две леммы автора диссертации. Данные результаты были получены в работах [21] и [41], соответственно.

Лемма 2.2.1 *Пусть A^c обозначает дополнение события A и $A_1, A_2, \ldots$ – последовательность событий, таких что $P(A_n) \to 0$. Если для некоторого $m \geq 0$ выполняется условие*

$$\sum_{n=1}^{\infty} P(A_n^c A_{n+1}^c \ldots A_{n+m-1}^c A_{n+m}) < \infty,$$

то $P(A_n$ б.ч.$) = 0$.

Определим последовательность событий, обладающих марковским свойством. Скажем, что последовательность событий A_n $(n \geq 1)$ обладает марковским свойством, если последовательность индикаторов I_{A_n} $(n \geq 1)$ является цепью Маркова.

Лемма 2.2.5 *Пусть $A_1, A_2, \ldots$ – последовательность событий, обладающих марковским свойством, такая что $P(A_n) \to 0$. Рассмотрим ряд*

$$\sum_{n=1}^{\infty} P(A_n A_{n+1}^c). \tag{2.2.6}$$

Если ряд (2.2.6) сходится, то справедливо равенство $P(A_n$ б.ч.$) = 0$. Если ряд (2.2.6) расходится, то $P(A_n$ б.ч.$) = 1$.

В последующих главах диссертации эти и другие обобщения леммы Бореля–Кантелли, приводимые в главе 2, используются для получения сильных предельных теорем.

В главе 3 диссертационной работы рассматриваются рекордные величины. Математическая теория рекордов берет свое начало с работы Чендлера[1]. На данный момент эта теория насчитывает большое количество работ, значительная часть из которых была опубликована в последние три или четыре десятилетия.

Определим последовательности верхних рекордных времен $L(n)$ и величин $X(n)$ следующим образом:

$$L(0) = 0, \quad L(1) = 1,$$
$$L(n+1) = \min\{j > L(n) : X_j > X_{L(n)}\}, \tag{3.1.1}$$
$$X(n) = X_{L(n)}, \quad n = 1, 2, \ldots$$

Если второй знак неравенства в (3.1.1) заменить на $<$, то вместо верхних рекордных времен и величин мы получим нижние рекордные времена $L^l(n)$ и величины $X^l(n)$. От верхних рекордных величин легко перейти к нижним, взяв вместо исходной последовательности величины $-X_1, -X_2, \ldots$ или, если $X_i > 0$ $(i \geq 1)$, случайные величины $1/X_1, 1/X_2, \ldots$ Таким образом, можно ограничиться изучением только верхних рекордных времен и величин, а само слово "верхние" в диссертационной работе мы будем опускать. Исключение составит параграф 4.3, в котором рассматриваются и верхние и нижние рекорды одновременно.

[1] Chandler, K.N. (1952). The distribution and frequency of record values, *J. Royal Statist. Soc.*, Ser. B, **14**, 220–228.

Для дискретных распределений и распределений, имеющих атомы, самостоятельный интерес представляют слабые рекордные времена и величины, введенные Верваатом[2]. Если второй знак неравенства в (3.1.1) заменить на $\geq$, то вместо рекордных времен и величин мы получим слабые рекордные времена $L^w(n)$ и величины $X^w(n)$. Вслед за работой Верваата, публикация диссертанта [4] явилась второй работой, посвященной слабым рекордам. В ней была предложена концепция перехода от зависимых слабых рекордных величин к суммам независимых величин, имеющих геометрические распределения.

Часто при проведении статистических экспериментов выборки содержат посторонние наблюдения. Такие нетипичные наблюдения могут сильно влиять на статистические выводы. Предположим, что выборка, состоящая из независимых одинаково распределенных X_i с непрерывной функцией распределения F, может быть "загрязнена" одним или несколькими наблюдениями, взятыми из выборки независимых одинаково распределенных Y_j с другой непрерывной функцией распределения G. Предположим, что $X_i \overset{st}{<} Y_j$, где $\overset{st}{<}$ означает стохастическое сравнение величин X_i и Y_j. Данное сравнение подразумевает, что величины Y_j имеют больше шансов стать рекордами в совместной выборке, состоящей из величин X_i и Y_j. Такие рассуждения указывают на необходимость введения нового понятия рекордов – рекордов с подтверждением. Понятие рекордных величин $\tilde{X}(n)$ и рекордных времен $\tilde{L}(n)$ с подтверждением было предложено Невзоровым[3]. Прежде чем определить рекорды с подтверждением, зафиксируем натуральные m, k, такие что $1 \leq m \leq k$. Пусть

$$\tilde{L}(1) = 1, \tilde{X}(1) = X_1$$

и для $n \geq 1$ справедливо:

$$\tilde{L}(n+1) = \min\left\{ j : j > \tilde{L}(n), \sum_{i=\tilde{L}(n)+1}^{j} I_{\{X_i > \tilde{X}(n)\}} = k \right\},$$

$$\tilde{X}(n+1) = X_{m,k}^{(n)},$$

где I_A – индикатор события A, а $X_{m,k}^{(n)}$ – порядковая статистика, полученная из выборки,

---

[2]Vervaat, W. (1973). Limit theorems for records from discrete distributions, *Stochast. Process. Appl.*, **1** (2), 317–334.

[3]Nevzorov, V.B. (2012). Records with confirmation. In: *Abstracts of International Conference Probability Theory and its Applications; In Commemoration of the Centennial of B.V. Gnedenko*, Moscow State University, 152–153.

состоящей из положительных наблюдений

$$I_{\{X_{\tilde{L}(n)+1}>\tilde{X}(n)\}}X_{\tilde{L}(n)+1},\ I_{\{X_{\tilde{L}(n)+2}>\tilde{X}(n)\}}X_{\tilde{L}(n)+2},\ \ldots,\ I_{\{X_j>\tilde{X}(n)\}}X_j.$$

В главе 3 излагаются предельные теоремы для слабых рекордов и рекордов с подтверждением. Важную роль для получения предельных теорем для слабых рекордов играют лемма 3.2.4 и представление 3.2.1. Определим величину ξ_i^w равенством:

$$\xi_i^w = k \quad (k \geq 0),$$

если k слабых рекордных величин зарегистрировано в точке i $(i = 0, \ldots, N)$.

Лемма 3.2.4 *Случайные величины* ξ_i^w $(i = 0, 1, \ldots, N)$ *независимы и*

$$P\{\xi_i^w = k\} = \beta_i(1-\beta_i)^k, \quad k = 0, 1, \ldots,$$

где $\beta_n = q_n/q_{n+1},\ q_n = P(X_1 \geq n)$, $\xi_N \overset{\text{п.н.}}{=} \infty$ *для* $N < \infty$.

Представление 3.2.1 *В условиях леммы 3.2.4 справедливо представление*

$$P\{X^w(n) > m\} = P\{\xi_0^w + \xi_1^w + \ldots + \xi_m^w < n\},$$

где $n \geq 1$ *и* $0 \leq m < N$.

Лемма 3.2.4 и представление 3.2.1 обобщаются в главе 3 на случай, когда функция распределения F имеет конечное число атомов (см. лемму 3.2.6). С помощью леммы 3.2.4 и представления 3.2.1 доказываются предельные теоремы для слабых рекордных величин. Приведем некоторые из них.

Пусть $X_1, X_2, \ldots$ – последовательность независимых одинаково распределенных случайных величин, принимающих значения $0, 1, \ldots$, таких что $F(n) < 1$ $(n \geq 0)$. Определим функции

$$R(n) = \sum_{i=0}^{n} \frac{1-\beta_i}{\beta_i}, \quad B(n) = \sum_{i=0}^{n} \frac{1-\beta_i}{\beta_i^2} \quad (n \geq 0).$$

Для слабых рекордных величин справедлив усиленный закон больших чисел Колмогорова.

Теорема 3.3.1 *Пусть распределение F удовлетворяет следующему условию:*

$$a = \inf_{n \geq 0} \beta_n > 0. \tag{3.3.1}$$

Тогда

$$\lim_{n \to \infty} \frac{R(X^w(n))}{n} \overset{п.н.}{=} 1.$$

Центральная предельная теорема и закон повторного логарифма для величины $R(X^w(n))$ приводятся в теореме 3.3.2.

Теорема 3.3.2 *Пусть распределение F удовлетворяет (3.3.1) и условию*

$$\lim_{n \to \infty} \frac{R(n)}{B(n)} = \epsilon \in [a, 1].$$

Тогда для слабых рекордных величин справедливы центральная предельная теорема и закон повторного логарифма, т.е.

$$\sqrt{\epsilon} \frac{R(X^w(n)) - n}{\sqrt{n}} \xrightarrow{d} Z \quad (n \to \infty),$$

$$\lim_{n \to \infty} \sup \sqrt{\epsilon} \frac{|R(X^w(n)) - n|}{\sqrt{2n \ln \ln n}} \overset{\text{п.н.}}{=} 1,$$

причем скорость сходимости в центральной предельной теореме для больших n имеет вид:

$$\sup_x \left| P\left\{ \sqrt{\epsilon} \frac{R(X^w(n)) - n}{\sqrt{n}} < x \right\} - \Phi(x) \right| = O\left(\frac{1}{\sqrt{n}} \right),$$

где Z и Φ – стандартная нормальная случайная величина и стандартное нормальное распределение, соответственно.

В параграфе 3.4 аппарат теории правильно меняющихся функций используется для получения предельных теорем как для отношений слабых рекордных величин $Z^w(n, m) = X^w(n + m)/X^w(n)$, так и для величин ξ_i^w. Приведем некоторые результаты этого параграфа.

Теорема 3.4.1 *Пусть $\bar{F}$ – быстро меняющаяся функция. Тогда для любого $m \geq 1$ справедливо:*

$$Z^w(n, m) \xrightarrow{p} 1 \quad (n \to \infty).$$

Теорема 3.4.2 *Пусть $\bar{F}$ – медленно меняющаяся функция. Тогда для любого $m \geq 1$ справедливо:*

$$Z^w(n,m) \xrightarrow{p} \infty \quad \textit{при} \quad n \to \infty.$$

Теорема 3.4.3 *Пусть $\bar{F}$ – правильно меняющаяся функция с показателем $\rho > 0$. Тогда для любого $m \geq 1$ справедливо:*

$$P\{Z^w(n,m) \geq x\} \to \frac{1 + (\rho \ln x) + \ldots + \frac{(\rho \ln x)^{m-1}}{(m-1)!}}{x^\rho}, \quad x \geq 1.$$

В параграфе 3.5 приводятся некоторые предельные теоремы для рекордных времен и величин с подтверждением. В частности, дается следующий результат.

Следствие 3.5.2 *Пусть F – непрерывная функция распределения. Тогда*

$$\frac{-\ln(1 - F(\tilde{X}(n)))}{n} \overset{\text{п.н.}}{\to} \frac{1}{k} + \ldots + \frac{1}{k-m+1}.$$

В главе 4 диссертационной работы обсуждаются некоторые статистические процедуры, связанные с рекордами. Через два года после публикации первой статьи по рекордам появились работы, в которых было предложено использовать рекорды для статистических выводов. См. работы Фостера и Стьюарта[4] и Фостера и Тейчроева[5]. На данный момент известно достаточное количество работ, в которых рекорды используются для оценки параметров и проверки статистических гипотез.

Естественно задаться вопросом – зачем нужно использовать рекорды в статистических процедурах. Почему бы не воспользоваться классической выборкой? Дело в том, что в некоторых случаях традиционная выборка может быть недоступна. В то же время рекордные величины (полученные из данной выборки) могут быть известны. Такие рекордные

[4]Foster, F.G. and Stuart, A. (1954). Distribution free tests in time-series band on the breaking records, *J. Royal Statist. Soc.*, Ser. B, **16** (1), 1–22.

[5]Foster, F.G. and Teichroew, D. (1955). A sampling experiment on the powers of the record tests for trend in a time series, *J. Royal Statist. Soc.*, Ser. B, **17** 115–121.

выборки рассматривались в работах Саманиего и Витакера[6,7] и назывались обратными выборками первого рода. Подобного типа выборки появляются в экспериментах теории сопротивления материалов, а также в теории надежности, где измерения проводятся последовательно и доступными оказываются величины, которые превосходят все предыдущие величины. Отметим, что использование выборок, состоящих только из рекордных величин, возможно также в случае, когда традиционные выборки содержат слишком много наблюдений.

В параграфе 4.2 рассматриваются критерии проверки гипотезы однородности, основанные на рекордных величинах. Пусть $X_1, X_2, \ldots$ и $Y_1, Y_2, \ldots$ – две независимые выборки независимых одинаково распределенных случайных величин с функциями распределения F_X и F_Y, соответственно. Пусть нулевая гипотеза состоит в том, что при всех значениях аргумента функции распределения равны между собой, то есть $H_0 : F_X = F_Y$. Пусть односторонняя альтернативная гипотеза H_1 имеет вид: существуют значения аргумента, для которых $F_X > F_Y$ (или $H_1' : F_X < F_Y$). Рассмотрим величины RM_i $(i = 1, 2, \ldots)$, определяемые равенствами

$$RM_i = \#\{j \in \{1, 2, \ldots\} : Y(i-1) < X(j) \leq Y(i)\}.$$

В последнем равенстве $Y(0) = -\infty$ и $X(i), Y(i)$ $(i = 1, 2, \ldots)$ – последовательности рекордных величин, полученных из выборок $X_1, X_2, \ldots$ и $Y_1, Y_2, \ldots$ Для натурального $r \geq 1$ определим три тестовые статистики. Первую статистику $RP_{(r)}$ назовем статистикой доминирования Y-рекордов:

$$RP_{(r)} = \sum_{i=1}^{r} RM_i.$$

Вторую статистику $RM_{(r)}$ назовем статистикой доминирования максимума Y-рекордов:

$$RM_{(r)} = \max_{1 \leq i \leq r} RM_i.$$

Наконец, третью статистику $RW_{(r)}$ назовем статистикой суммы рангов Y-рекордов:

$$RW_{(r)} = \sum_{i=1}^{r} \mathrm{Rank}(Y(i)).$$

[6]Samaniego, F.J. and Whitaker, L.R. (1986). On estimating popular characteristics from record breaking observations I. Parametric results, *Naval Res. Log. Quart.*, **33**, 531–543.

[7]Samaniego, F.J. and Whitaker, L.R.(1988). On estimating popular characteristics from record breaking observations II. Nonparametric results, *Naval Res. Log. Quart.*, **35**, 221–236.

Величины $\mathrm{Rank}(Y(i))$ являются рангами рекордных величин $Y(i)$ в последовательности $Y(1), \ldots, Y(r), X(1), \ldots, X(RP_{(r)})$. Тест, основанный на величине суммы рангов рекордов, аналогичен хорошо известному тесту Вилкоксона[8], основанному на сумме рангов порядковых статистик. В параграфе 4.2 рассматриваются критерии проверки гипотезы однородности, основанные на статистиках $RP_{(r)}, RM_{(r)}$ и $RW_{(r)}$. Эти критерии обсуждались ранее в работе диссертанта [13].

Информация Фишера (FI), содержащаяся в рекордах, слабых рекордах и в числах рекордных величин, исследуется в параграфе 4.3. Данные исследования используются при нахождении оценок неизвестных параметров. Приведем один из результатов параграфа 4.3 (теорема 4.3.1 ниже). Будем говорить, что распределение $F(x \mid \theta)$ (с плотностью $f(x \mid \theta)$) принадлежит семейству распределений $\mathrm{Exp}_1(\theta)$ (или $\mathrm{Exp}_2(\theta)$), если существуют измеримые функции $c(\theta)$ и $k(x)$, такие что $\frac{f(x|\theta)}{1-F(x|\theta)} = c(\theta)k(x)$ (или $\frac{f(x|\theta)}{F(x|\theta)} = c(\theta)k(x)$). Информация Фишера, содержащаяся в верхних и нижних рекордных величинах и временах, извлеченных из популяции, принадлежащей семействам $\mathrm{Exp}_1(\theta)$ и $\mathrm{Exp}_2(\theta)$, обсуждается в теореме 4.3.1 (см. также работу автора диссертации [18]).

Теорема 4.3.1 *При любом $n \geq 1$ справедливы равенства:*

$$I_n(R^u T^u) = \left(\frac{c'(\theta)}{c(\theta)}\right)^2 \sum_{i=1}^{n} 1/i, \quad F \in Exp_2(\theta),$$

$$I_n(R^l T^l) = \left(\frac{c'(\theta)}{c(\theta)}\right)^2 \sum_{i=1}^{n} 1/i, \quad F \in Exp_1(\theta),$$

где $I_n(R^l T^l)$ и $I_n(R^u T^u)$ – информация Фишера, содержащаяся, соответственно, во всех нижних рекордных величинах и временах и во всех верхних рекордных величинах и временах, принадлежащих данной выборке $X_1, \ldots, X_n$.

В параграфе 4.4 обсуждаются тесты, основанные на рекордных величинах с подтверждением. Подобного рода тесты могут быть полезны при обнаружении и удалении нетипичных наблюдений, имеющих "большие" значения. Предположим, что в результате статистического эксперимента получены две независимые между собой выборки, состоящие из независимых одинаково распределенных случайных величин. Первая выборка состоит из наблюдений X_i с непрерывной F, а вторая – из наблюдений X_i и Y_j с непрерывными F и G. Пусть $\tilde{L}(2)$ – второе рекордное время с подтверждением, полученное из какой-

[8]Wilcoxon, F. (1945). Individual comparisons by ranking methods, *Biometrics Bull.*, **1**, 80–83.

либо выборки. В параграфе 4.4 предлагается тест, основанный на рекордных временах с подтверждением, который позволяет идентифицировать выборку, то есть определить присутствуют ли в данной выборке нетипичные наблюдения Y_j. Данный тест основывается на статистике $T = \frac{1}{\tilde{L}(2)+1}$. Поскольку предполагается, что распределения F и G имеют один и тот же носитель и $X_i \overset{st}{<} Y_j$, где $\overset{st}{<}$ обозначает стохастическое сравнение величиин X_i и Y_j, наблюдения Y_j имеют больше шансов стать рекордными величинами в совместной выборке. Основная гипотеза H_0 задается следующим образом:

выборка состоит только из наблюдений X_i.

Альтернативная гипотеза H_1 определяется следующим образом:

первым наблюдением выборки является X_1,

среди остальных наблюдений выборки есть хотя бы одно наблюдение Y_j.

Поскольку $X_i \overset{st}{<} Y_j$, разумно предположить, что если справедлива альтернативная гипотеза H_1, то второе рекордное время появится раньше, чем это произойдет в случае, когда справедлива H_0. Данные рассуждения помогают определить критическую область: $\overline{T} > c + \frac{1}{2(k+1)}$, где $c > 0$ – надлежащее критическое значение и $\overline{T}$ – среднее значение выборки $T_1, \ldots, T_N$. Вышеописанный тест обсуждается в параграфе 4.4.

В главе 5 диссертационной работы обсуждаются асимптотические свойства чисел величин, регистрируемых около порядковых статистик и рекордов. Пусть $X_1, \ldots, X_n$ – выборка независимых одинаково распределенных величин с общей непрерывной функцией распределения F. Пэйкс и Стьютел[9] и Хмаладзе, Надареишвили и Никабадзе[10] предложили исследовать величину

$$K(n,a) = \#\{j : X_j \in (M_n - a, M_n)\} \quad (1 \leq j \leq n, a > 0),$$

которая описывает количество наблюдений выборки, попадающих в случайный интервал $(M_n - a, M_n)$. Оказалось, что величина $K(n,a)$ тесно связана с "верхними" спейсингами,

[9]Pakes, A. and Steutel, F.W. (1997). On the number of records near the maximum, *Austral. N. Z. J. Statist.*, **39**, 179–193.

[10]Khmaladze, E., Nadareishvili, M. and Nikabadze, A. (1997). Asymptotic behaviour of a number of repeated records, *Statist. Probab. Lett.*, **35**, 49–58.

построенными по порядковым статистикам, и может быть использована для получения результатов для спейсингов. Определим случайную величину $K_-(n,k,a)$:

$$K_-(n,k,a) = \#\{j : X_j \in (X_{k,n} - a, X_{k,n})\} \quad (0 \le j \le n, 1 \le k \le n).$$

Отметим, что $K_-(n,n,a) = K(n,a)$. В главе 5 также вводится случайная величина $K_+(n,k,a)$, равная числу наблюдений выборки, попадающих в случайный интервал $(X_{k,n}, X_{k,n} + a)$. Свойства величин K_- и K_+, изучавшиеся ранее во многих публикациях, рассматриваются в главе 5.

Напомним, что в главе 3 обсуждаются дискретные случайные величины ξ_i^w, равные числам слабых рекордных величин, попадающих в точки i $(i \ge 1)$. В главе 5 рассматривается непрерывный аналог этих величин – величины $\xi_n(a)$, задаваемые равенствами:

$$\xi_n(a) = \#\{j : L(n) < j < L(n+1),\ X_j \in (X(n) - a, X(n)]\} \quad (n \ge 1).$$

Величины $\xi_n(a)$ назовем числами величин, регистрируемых около рекордных величин, или числами околорекордных величин. Величины $\xi_n(a)$ и их суммы находят применение в актуарной и финансовой математике. Пусть X_i – положительные величины и $s_n(a)$ – сумма величин, регистрируемых около n-ой рекордной величины, т.е. $s_n(a) = \sum_{i=L(n)+1}^{L(n+1)-1} X_i$, где $X_i \in (X(n) - a, X(n)]$. Данная сумма может быть интерпретирована как сумма страховых требований, попадающих в случайный интервал $(X(n) - a, X(n)]$. Эти требования будут регистрироваться только после некоторого специального момента. Этот момент обозначен требованием, превосходящим все предыдущие требования, или, в наших терминах, является n-ым рекордным временем $L(n)$. В силу некоторых обстоятельств (текущего законодательства и т.п.) регистрация этих требований прекращается в момент появления следующего рекордного требования $L(n+1)$. Естественно, что для страховых компаний сумма "больших" страховых выплат $s_n(a)$ (а она может достигать до 90% от суммы всех выплат) представляет значительный интерес. Асимптотические свойства величин $\xi_n(a)$ и их сумм изучались во многих работах, среди которых были и работы диссертанта.

В параграфах 5.2–5.4 получены предельные результаты для величин $\xi_n(a)$, их сумм и величин K_- и K_+. Приведем некоторые из них.

Теорема 5.2.2 *Пусть $r_F = \infty$ и $\beta(a) = 1$. Для того чтобы*

$$K_-(n, n-k, a) \overset{\text{п.н.}}{\to} 0 \quad (n \to \infty)$$

(для некоторых фиксированных $k \geq 0$ и $a > 0$), необходимо и достаточно, чтобы

$$\int_{\mathbb{R}} \frac{F(x+a) - F(x-a)}{(1 - F(x-a))^2} dF(x) < \infty.$$

Теорема 5.4.10 *Предположим, что предел в (1.2.1) существует для $a, b > 0$, таких что a/b – иррационально. Тогда*

$$(\xi_n(a_1), \ldots, \xi_{n+m-1}(a_m)) \xrightarrow{d} (g_1(\beta(a_1)), \ldots, g_m(\beta(a_m)) \quad (n \to \infty),$$

где $g_i(p)$ – геометрическая случайная величина с параметром p и компоненты вектора, стоящего в правой части последнего предельного соотношения, независимы. В частности, $s^{\xi}_{n,m,\bar{a}} \xrightarrow{d} s^{g}_{m,\bar{a}}$, где

$$s^{\xi}_{n,m,\bar{a}} = \xi_n(a_1) + \ldots + \xi_{n+m-1}(a_m), \quad s^{g}_{m,\bar{a}} = g_1(\beta(a_1)) + \cdots + g_m(\beta(a_m)).$$

При этом производящая функция $s^{g}_{m,\bar{a}}$ имеет вид:

$$E\left[t^{s^{g}_{m,\bar{a}}}\right] = \prod_{i=1}^{m} \frac{\beta(a_i)}{1 - (1 - \beta(a_i))t}. \tag{5.4.16}$$

Распределение величины $s^{g}_{m,\bar{a}}$ является сверткой геометрических распределений. В частности, если все координаты вектора $\bar{a} = (a_1, \ldots, a_n)$ различны, то

$$P\left(s^{g}_{m,\bar{a}} = k\right) = \sum_{i=1}^{m} \alpha_i(m)\ \beta(a_i)\left(1 - \beta(a_i)\right)^k, \quad k = 0, 1, \ldots, \tag{5.4.17}$$

где

$$\alpha_i(m) = \frac{\beta(m)\ (1 - \beta(a_i))^m}{\beta(a_i) \prod_{\substack{j=1 \\ j \neq i}}^{m} (\beta(a_j) - \beta(a_i))}\ , \quad i = 1, \ldots, m, \tag{5.4.18}$$

$\beta(m) = \prod_{i=1}^{m} \beta(a_i)$. Если же $a_1 = \ldots = a_m$, то распределение $s^{g}_{m,\bar{a}}$ является отрицательным биномиальным. Получен также следующий результат.

Теорема 5.4.11 *Пусть $\bar{F}$ – быстро меняющаяся функция и предел в (1.2.1) существует для $a, b > 0$, таких что a/b – иррационально. Тогда $S^{s}_{n,m,\bar{a}}/X(n) \xrightarrow{d} s^{g}_{m,\bar{a}}$, где*

$S^s_{n,m,\bar a} = s_n(a_1) + \ldots + s_{n+m-1}(a_m)$ и величина $s^g_{m,\bar a}$ определяется производящей функцией (5.4.16).

(1) *(а) В частности, если $0 < \beta(a) < 1$ и все компоненты $\bar a$ различны, то распределение $s^g_{m,\bar a}$ задается (5.4.17) и (5.4.18).*

(б) Если же $a_1 = \cdots = a_m = a$, то

$$P\{s^g_{m,\bar a} = k\} = C^k_{m-1+k}\{\beta(a)\}^m\{1-\beta(a)\}^k \quad (k \geq 0).$$

(2) *Если $\beta(a) = 0$, то для любого вектора $\bar a$ с положительными компонентами верно: $S^s_{n,m,\bar a}/X(n) \xrightarrow{p} \infty$.*
(3) *Если $\beta(a) = 1$, то для любого вектора $\bar a$ с положительными компонентами верно: $S^s_{n,m,\bar a}/X(n) \xrightarrow{p} 0$.*

В главе 6 диссертационной работы обсуждаются предельные теоремы для серий, основанных на порядковых статистиках и рекордах. Понятия серий, образуемых спейсингами порядковых статистик и рекордов, были предложены, соответственно, в работах автора диссертации [22] и [28]. Свойства таких серий изучались далее во многих работах.

В главе 6 рассматриваются выборки двух типов. Выборка первого типа состоит из независимых и одинаково распределенных случайных величин $X_1, \ldots, X_n$ с непрерывной функцией распределения F. Из этих величин формируются соответствующие порядковые статистики $X_{1,n} \leq \ldots \leq X_{n,n}$. Выборки второго типа (в диссертационной работе они называются обратными выборками первого рода) – есть выборки, состоящие из рекордных величин $X(1), \ldots, X(n)$. Они используются тогда, когда классические выборки недоступны.

Серии, основанные на спейсингах порядковых статистик $X_{i,n} - X_{i-1,n}$ (спейсингах рекордных величин $X(j) - X(j-1)$), определяются в главе 6 (см. также [28] и [22]) следующим образом. Для $\varepsilon > 0$ определим величины $\nu^\varepsilon_{i,n}$ и ξ^ε_i соотношениями

$$\nu^\varepsilon_{1,n} = 1, \nu^\varepsilon_{n+1,n} = 1, \quad \xi^\varepsilon_1 = 1, \xi^\varepsilon_{n+1} = 1,$$

$$\nu^\varepsilon_{i,n} = \begin{cases} 1, & \text{если} \quad X_{i,n} - X_{i-1,n} > \varepsilon, \quad i = 2, \ldots, n, \\ 0, & \text{в противном случае,} \end{cases}$$

$$\xi^\varepsilon_n = \begin{cases} 1, & \text{если} \quad X(n) - X(n-1) > \varepsilon, \quad n \geq 2, \\ 0, & \text{в противном случае.} \end{cases}$$

Будем говорить, что последовательность спейсингов порядковых статистик

$$X_{i,n} - X_{i-1,n}, \dots, X_{j,n} - X_{j-1,n}$$

(спейсингов рекордных величин $X(i) - X(i-1), \dots, X(j) - X(j-1)$) формирует серию длиной $j - i - 1$ $(i + 1 \le j \le n + 1)$, если

$$\nu^{\varepsilon}_{i,n} = 1, \nu^{\varepsilon}_{i+1,n} = \ldots = \nu^{\varepsilon}_{j-1,n} = 0, \nu^{\varepsilon}_{j,n} = 1$$

$(\xi^{\varepsilon}_i = 1, \xi^{\varepsilon}_{i+1} = \ldots = \xi^{\varepsilon}_{j-1} = 0, \xi^{\varepsilon}_j = 1)$. Определим числа всех серий, образованных спейсингами данной выборки $R^{\varepsilon,\nu}_n$ и обратной выборки первого рода $R^{\varepsilon,\xi}_n$, равенствами

$$R^{\varepsilon,\nu}_n = \sum_{l=1}^{n} \nu^{\varepsilon}_{l,n}, \quad R^{\varepsilon,\xi}_n = \sum_{l=1}^{n} \xi^{\varepsilon}_l.$$

В главе 6 доказываются следующие предельные теоремы.

Теорема 6.4.2 *Пусть* $\beta(\varepsilon) = \gamma(\varepsilon) = 0$, *где* $\gamma(b) = \lim_{x\to-\infty} \frac{F(x)}{F(x+b)}$ $(l_F = -\infty, b > 0)$, *и*

$$\int_{\mathbb{R}} \frac{\bar{F}(x+\varepsilon)}{\bar{F}^2(x)}\, dF(x) < \infty, \quad \int_{\mathbb{R}} \frac{F(x-\varepsilon)}{F^2(x)}\, dF(x) < \infty.$$

Тогда

$$R^{\varepsilon,\nu}_n \overset{п.н.}{\to} 1.$$

Теорема 6.4.4 *Пусть* F *– абсолютно непрерывное распределение,* $\beta(\varepsilon) \in (0, 1]$ *и* $\beta(x, \varepsilon)$ *строго возрастает по переменной* x. *Тогда*

$$\frac{R^{\varepsilon,\xi}_n}{n} \overset{п.н.}{\to} \beta(\varepsilon).$$

В параграфе 6.5 рассматривается схема серий в случае равномерного распределения. Здесь величины $\nu^{\varepsilon}_{i,n}$ будут симметрично зависимыми. Пусть $X_1, \dots, X_{k_n}$ – независимые равномерно на $[0, 1]$ распределенные величины и $k_n \to \infty$ при $n \to \infty$. Пусть $\varepsilon = \varepsilon_n$ уменьшается по мере роста количества наблюдений в каждой серии, так что $0 < \varepsilon_n < \frac{1}{k_n-1}$. Справедлив следующий результат.

Теорема 6.5.1 *Пусть* $m_n < k_n$ – *последовательность натуральных чисел, таких что* $\lim_{n\to\infty} \frac{m_n}{k_n} = \alpha \in (0,1)$.

1) Пусть $\varepsilon_n = \frac{\lambda}{k_n - 1}$, *где* $k_n \to \infty$ *и* $\lambda \in (0,1)$. *Тогда*

$$P\left\{ \frac{R_{m_n}^{\varepsilon_n,\nu} - e^{-\lambda} m_n}{\sqrt{(e^{-\lambda} - e^{-2\lambda})\, m_n}} \le x \right\} \to \frac{1}{\sqrt{2\pi(1-\alpha)}} \int\limits_{-\infty}^{x} e^{-\frac{y^2}{2(1-\alpha)}} dy.$$

2) Пусть $\varepsilon = o(\frac{1}{k_n})$ *и* $\varepsilon_n k_n^{3/2} \to \infty$. *Тогда*

$$P\left\{ \frac{R_{m_n}^{\varepsilon_n,\nu} - m_n}{m_n \sqrt{\alpha^{-1}\varepsilon_n}} \le x \right\} \to \frac{1}{\sqrt{2\pi(1-\alpha)}} \int\limits_{-\infty}^{x} e^{-\frac{y^2}{2(1-\alpha)}} dy.$$

В главе 6 также предлагается статистический критерий, основанный на спейсингах, образованных порядковыми статистиками.

В главе 7 диссертационной работы рассматриваются предельные теоремы для конкомитантов порядковых статистик и рекордов. Интерес к конкомитантам порядковых статистик и рекордов проявляется в страховом бизнесе, когда конкомитанты наибольших страховых выплат также могут вызвать большие расходы для страховых компаний, в теории надежности, когда сложные системы состоят из нескольких компонент и выход из строя отдельных компонент системы приводит к прекращению работы всей системы, в биостатистике, при исследовании побочных эффектов от действия лекарств, и в некоторых других областях. Пусть $(X,Y),(X_1,Y_1),(X_2,Y_2),\ldots,(X_n,Y_n),\ldots$ – независимые одинаково распределенные векторы с двумерной непрерывной функцией распределения $F(x,y)$ и маргинальными функциями распределения $H(x)$ и $G(y)$. Пусть $X_{1,n} \le X_{2,n} \le \ldots \le X_{n,n}$ – порядковые статистики, образованные выборкой $X_1,\ldots,X_n$, а $X(n)$ $(n \ge 1)$ – рекордные величины, полученные из выборки $X_1, X_2, \ldots$ Порядковые статистики $X_{i,n}$ и рекордные величины $X(n)$ индуцируют конкомитанты порядковых статистик $Y_{[i,n]}$ и рекордных величин $Y[n]$ $(n \ge 1)$. Конкомитанты порядковых статистик были предложены в работах Дэйвида[11] и Бхаттачарии[12]. Конкомитанты рекордных величин, по-видимому, впервые обсуждались в работе Хоученза[13].

[11]David, H.A. (1973). Concomitants of order statistics, *Bull. Internat. Statist. Inst.*, **45** (1), 295–300.

[12]Bhattacharya, B.B. (1974). Convergence of sample paths of normalized sums of induced order statistics, *Ann. Statist.*, **2**, 1034–1039.

[13]Houchens, R.L. (1984). *Record value theory and inference*, Ph.D. Dissertation, University of California, Riverside, California.

Асимптотическое поведение вероятности $P(Y_{[n-k,n]} < y)$ в случае, когда k ($k \geq 0$) либо фиксировано, либо стремится к бесконечности (при $n \to \infty$), изучалось во многих работах, см., например, работы Егорова и Невзорова[14], Бхаттачарии[15], Дэйвида [16], Дэйвида и Галамбоша[17]. См. также ссылки, даваемые в этих работах. Стоит отметить, что работ, в которых рассматривались сильные предельные теоремы для конкомитантов порядковых статистик или рекордов, было немного. Упомянем здесь статью Гоела и Холла[18], в которой были получены сильные предельные теоремы для разностей между порядковыми статистиками и их конкомитантами, статью Сена[19], в которой рассматривались принципы инвариантности для конкомитантов порядковых статистик, и работы автора диссертации [10], [41], в которых обсуждались асимптотические свойства величин $Y_{[n-k,n]}$, в том числе, сильные предельные теоремы. Стоит также упомянуть работу диссертанта [11], в которой изучались асимптотические свойства конкомитантов рекордных величин.

Следующий предел рассматривается в главе 7 (см. также работы автора диссертации [10] и [11]):

$$\lim_{x \to r_H^-} \frac{G(y) - F(x,y)}{1 - H(x)} = \wp(y) \in [0,1]. \tag{7.2.6}$$

Предел в (7.2.6) используется в главе 7 для классификации двумерных распределений. Величина $\wp$ классифицирует распределения F следующим образом.

(i) Пусть для некоторого $c \in [-\infty, \infty]$

$$\wp(y) = 0 \quad (y < c) \quad \text{и} \quad \wp(y) = 1 \quad (y > c).$$

Тогда распределение F называется c-конкомитант-стабильным.

---

[14] Егоров В.А., Невзоров В.Б. (1981). О скорости сходимости к нормальному закону сумм индуцированных порядковых статистик, *Зап. научн. сем. ПОМИ*, **108**, 45—56.

[15] Bhattacharya, B.B. (1984). Induced order statistics: theory and applications, In *Handbook of Statistics 4*, Ed. Krishnaiah, P. R. and Sen, P. K., North Holland, Amsterdam, 383–403.

[16] David, H.A. (1994). Concomitants of extreme order statistics, In *Extreme Value Theory and Applications*, Proceedings of the Conference on Extreme Value Theory and Applications, **1**, Ed. Galambos, J., Lechner, J., and Simiu, E., Kluwer Academic Publishers, Boston, 211–224.

[17] David, H.A. and Galambos, J. (1974). The asymptotic theory of concomitants of order statistics, *J. Appl. Probab.*, **11**, 762–770.

[18] Goel, P. K. and Hall, P. (1994). On the average difference between concomitants and order statistics, *Ann. Probab.*, **22**, 126–144.

[19] Sen, P.K. (1981). Some invariance principles for mixed rank statistics and induced order statistics and some application, *Comm. Statist. Theory Methods*, **10** 1691–1718.

(ii) Если такой c не существует, то F называется конкомитант-нестабильным распределением. Например, если X и Y – независимые случайные величины, то $\wp(y) = G(y)$ и распределение F – конкомитант-нестабильное.

Приведем только один асимптотический результат главы 7. В диссертационной работе получена сильная предельная теорема для величины $Y[n]$.

Теорема 7.5.2 *Пусть для всех* $\varepsilon > 0$ *выполняются равенства* $\wp(c-\varepsilon) = 0$ *и* $\wp(c+\varepsilon) = 1$. *Для того чтобы*

$$Y[n] \stackrel{п.н.}{\to} c \quad (n \to \infty),$$

необходимо и достаточно, чтобы интегралы

$$\int_{\mathbb{R}} \frac{G(c+\varepsilon) - F(x, c+\varepsilon)}{(1 - H(x))^2} \, [dH(x) - F(dx, c+\varepsilon)]$$

и

$$\int_{\mathbb{R}} \frac{G(c-\varepsilon) - F(x, c-\varepsilon)}{(1 - H(x))^2} \, [dH(x) - F(dx, c-\varepsilon)]$$

сходились при всех $\varepsilon > 0$.

В главе 8 диссертационной работы обсуждаются характеризационные теоремы для порядковых и рекордных статистик. Тематика характеризации распределений свойствами рекордов и порядковых статистик обсуждалась в книгах Кагана, Линника и Рао[20], Галамбоша и Котца[21], Галамбоша[22], Рао и Шанбхага[23], Ахсануллаха[24], Арнольда, Балакришнана и Нагараджи[25,26], Невзорова[27], Дэйвида и Нагараджи[28].

С 1993 года слабые рекорды стали использоваться для характеризаций. Приводимая ниже теорема 8.2.1, опубликованная автором диссертационной работы в статье [5], явилась первым результатом, использующим свойства слабых рекордов для характеризации

[20] Каган А.М., Линник Ю.В., Рао С.Р. *Характеризационные задачи математической статистики.* – М. Наука, 1972.

[21] Galambos, J. and Kotz, S. (1978). *Characterizations of probability distributions*, Lecture Notes in Mathematics No. 657, Springer-Verlag, Berlin.

[22] Галамбош Я. *Асимптотическая теория экстремальных порядковых статистик.* – М. Наука, 1984.

[23] Rao, C.R. and Shanbhag, D.H. (1993). *Choquet–Deny type functional equations with applications to stochastic models*, John Wiley & Sons, Chichester, England.

[24] Ahsanullah, M. (1995). *Record statistics*, Nova Science Publishers, Commack, NY.

[25] Arnold, B.C., Balakrishnan, N. and Nagaraja, H.N. (1992). *A first course in order statistics*, John Wiley & Sons, NY.

[26] Arnold, B.C., Balakrishnan, N. and Nagaraja, H.N. (1998). *Records*, John Wiley & Sons, NY.

[27] Невзоров В.Б. *Рекорды. Математическая теория.* – М. ФАЗИС, 2000.

[28] David, H.A. and Nagaraja, H.N. (2003). *Order statistics*, Third edition, John Wiley & Sons, NY.

распределений.

Теорема 8.2.1 *Пусть $X_1, X_2, \ldots$ – последовательность независимых величин, принимающих значения $0, 1, \ldots,$ с общей функцией распределения F, такой что $r_F = \infty$. Для того чтобы случайная величина X_1 имела распределение вида*

$$P\{X_1 \geq m\} = \frac{\prod_{i=1}^{m}(\alpha + (i-1)\beta)}{\prod_{i=1}^{m}(1 + \alpha + i\beta)} \quad (m = 1, 2, \ldots)$$

(для некоторых $\alpha > 0$ и $\beta \geq 0$), необходимо и достаточно, чтобы

$$E\{X^w(n+1) - X^w(n) | X^w(n) = l\} = \alpha + \beta l$$

для всех $l = 0, 1, \ldots$ $(n \geq 1)$.

Обобщения этой теоремы и иные характеризации, основанные на слабых рекордах, были получены позднее во многих работах.

В параграфе 8.3 приводятся характеризации, основанные на порядковых статистиках. В параграфе 8.4 даны характеризации, основанные на второй рекордной величине и максимуме выборки. Пусть $X_1, X_2, \ldots$ – последовательность независимых величин с общей непрерывной функцией распределения F и $S_n = X(2) - M_n$.

Теорема 8.4.1 **(1)** *Случайные величины $S_n^- = \min\{0, S_n\}$ и M_n независимы для некоторого $n \geq 3$ тогда и только тогда, когда исходное непрерывное распределение F совпадает (с точностью до параметров сдвига и масштаба) с $F(x) = e^x$ $(x < 0)$.*

(2) *Случайные величины $S_n^+ = \max\{0, S_n\}$ и M_n независимы для некоторого $n \geq 1$ тогда и только тогда, когда исходное непрерывное F совпадает (с точностью до параметров сдвига и масштаба) с $F(x) = 1 - e^{-x}$ $(x > 0)$.*

3) *Величины S_n и M_n независимы для некоторого $n = 1, 2, \ldots$ тогда и только тогда, когда исходное непрерывное F совпадает (с точностью до параметров сдвига и масштаба) с $F(x) = 1 - e^{-x}$ $(x > 0)$ и $n = 1, 2$.*

Автор работы искренне благодарен Валерию Борисовичу Невзорову и Якову Юрьевичу Никитину за помощь, советы и поддержку на протяжении всей своей научной деятельности.

Список литературы

[1] Невзоров В.Б., Степанов А.В. (1988). Рекорды: мартингальный подход к нахождению моментов, В сб. *Кольца и модули. Предельные теоремы теории вероятностей*, Л. ЛГУ, 171–181.

[2] Степанов А.В. (1987). О логарифмических моментах для межрекордных времен, *Теория вероятн. и ее примен.*, **32**, 774–776.

[3] Степанов А.В. (1989). Характеризации геометрического класса распределений, *Теория вероятн. и матем. статист.*, Киев, **41**, 133–136.

[4] Степанов А.В. (1992). Предельные теоремы для слабых рекордов, *Теория вероятн. и ее примен.*, **37**, 586–590.

[5] Степанов А.В. (1993). Характеризационная теорема для слабых рекордных величин, *Теория вероятн. и ее примен.*, **38**, 762–764.

[6] Степанов А.В. (1996). Экстремальные порядковые статистики при изменении отношения порядка, *Теория вероятн. и ее примен.*, **41**, 896–900.

[7] Степанов А.В. (2002a). Условные распределения времен односторонних последовательных приближений, *Теория вероятн. и ее примен.*, **47**, 364–366.

[8] Степанов А.В. (2002b). Суммы случайных величин, имеющих геометрические распределения, *Известия КГТУ*, **1**, 179–183.

[9] Bairamov, I., Berred, A. and Stepanov, A. (2010). Limit results for ordered uniform spacings, *Statist. Pap.*, **51** (1), 227–240.

[10] Bairamov, I. and Stepanov, A. (2010). Numbers of near-maxima for the bivariate case, *Statist. Probab. Lett.*, **80**, 196–205.

[11] Bairamov, I. and Stepanov, A. (2011). Numbers of near bivariate record-concomitant observations, *J. Multivariate Anal.*, **102**, 908–917.

[12] Bairamov, I. and Stepanov, A. (2013). Numbers of near-maxima for F^{α}-scheme, *Statistics*, **47**, 191–201.

[13] Balakrishnan, N., Dembinska, A. and Stepanov, A. (2008). Precedence-type tests based on record values, *Metrika*, **68**, 233–255.

[14] Balakrishnan, N., Pakes, A. and Stepanov, A. (2005). On the number and sum of near record observations, *Adv. Appl. Probab.*, **37**, 1–16.

[15] Balakrishnan, N. and Stepanov, A. (2004a). A note on the paper of Khmaladze *et al.*, *Statist. Probab. Lett.*, **68**, 415–419.

[16] Balakrishnan, N and Stepanov, A. (2004b). Two characterizations based on order statistics and records, *J. Statist. Plann. Inference*, **124**, 273–287.

[17] Balakrishnan, N. and Stepanov, A. (2005). A note on the number of observations registered near an order statistic, *J. Statist. Plann. Inference*, **134**, 1–14.

[18] Balakrishnan, N. and Stepanov, A. (2006). On the Fisher information in record data, *Statist. Probab. Lett.*, **76**, 537–545.

[19] Balakrishnan, N. and Stepanov, A. (2008a). Asymptotic properties of the ratio of order statistics, *Statist. Probab. Lett.*, **78**, 301–310.

[20] Balakrishnan, N. and Stepanov, A. (2008b). Asymptotic properties of numbers of near minimum observations under progressive Type-II censoring, *J. Statist. Plann. Inference*, **38**, 1010–1020.

[21] Balakrishnan, N. and Stepanov, A. (2010). Generalization of the Borel-Cantelli lemma, *Math. Sci.*, **35**, 61–62.

[22] Balakrishnan, N. and Stepanov. A. (2013). Runs based on records: Their distributional properties and an application to testing for dispersive ordering, *Methodol. Comput. Appl. Probab.*, **15**, 583–594.

[23] Balakrishnan, N. and Stepanov, A. (2014). On the use of bivariate Mellin transform in bivariate random scaling and some applications, *Methodol. Comput. Appl. Probab.*, **16**, 235–244.

[24] Bayramoglu, I. and Stepanov, A. (2006). A note on large deviations for weak records, *Statist. Probab. Lett.*, **76**, 1449–1453.

[25] Berred, A. and Stepanov, A. (2005). Ties for the the second place, in book *Recent Developments in Ordered Random Variables*, Ed. Ahsanullah, M. and Raqab, M., Nova Science Publisher, NY, 171–185.

[26] Dembinska, A. and Stepanov, A. (2006). Limit theorems for the ratio of weak records, *Statist. Probab. Lett.*, **76**, 1454–1464.

[27] Dembinska, A. Stepanov, A. and Wesolowski, J. (2007). How many observations fall in a neighborhood of an order statistic, *Comm. Statist. Theory Methods*, **36**, 851–867.

[28] Eryilmaz, S. and Stepanov, A. (2008). Runs in an ordered sequence of random variables, *Metrika*, **67**, 299–313.

[29] Hashorva, E. and Stepanov, A. (2012). Limit theorems for the spacing of weak records, *Metrika*, **75**, 163–180.

[30] Nevzorov, V.B. and Stepanov, A. (2014). Records with confirmation, *Statist. Probab. Lett.*, **95**, 39–47.

[31] Stepanov, A. (1998). Limit behavior of the times of one-sided successive approximations, *Istatistik*, **1** (3), 43–46.

[32] Stepanov, A. (1999). The second record time in the case of arbitrary distribution, *Istatistik*, **2** (2), 65–70.

[33] Stepanov, A. (2001). Records when the last point of increase is an atom, *J. Appl. Statist. Sci.*, **10** (2), 161–167.

[34] Stepanov, A. (2003). Conditional moments of record times, *Statist. Pap.*, **44** (1), 131–140.

[35] Stepanov, A. (2004). Random intervals based on record values, *J. Statist. Plann. Inference*, **118**, 103–113.

[36] Stepanov, A. (2006). The number of records within a random interval of the current record value, *Statist. Pap.*, **48**, 63–79.

[37] Stepanov, A. (2007). Weak records, in book *Encyclopedia of Statistical Sciences*, Ed. Balakrishnan, N.

[38] Stepanov, A. (2011a). Runs based on discrete order statistics, *TWMS J. Appl. Engineering Math.*, **1** (2), 185–191.

[39] Stepanov, A. (2011b). Runs based on the ratios of consecutive order statistics, *Comm. Statist. Theory Methods*, **40** (18), 3252–3268.

[40] Stepanov, A. (2011c). Limit theorems for runs based on 'small' spacings, *Statist. Probab. Lett.*, **81**, 54–61.

[41] Stepanov, A. (2014). On the use of the Borel-Cantelli lemma in Markov chains, *Statist. Probab. Lett.*, **90**, 149–154.

[42] Stepanov, A.V., Balakrishnan, N. and Hofmann, G. (2003). Exact distribution and Fisher information of weak record values, *Statist. Probab. Lett.*, **64**, 69–81.