ОТЗЫВ

официального оппонента о диссертации А.В.Степанова «Предельные теоремы и статистические процедуры для величин, связанных с рекордами и экстремальными порядковыми статистиками», представленной на соискание ученой степени доктора физико-математических наук по специальности 01.01.05 - теория вероятностей и математическая статистика

Диссертация относится к области теории вероятностей и предельных теорем математической статистики. В диссертации рассмотрен широкий круг проблем, связанных с рекордами, экстремальными порядковыми статистиками, конкомитантами. Исследуются асимптотические свойства такого рода статистик, при этом значительное внимание уделяется приложениям полученных результатов: построен ряд статистических процедур, основанных на порядковых статистиках и рекордах, изучены свойства тестовых статистик.

Представленное исследование находится в русле одного из важных и интенсивно развиваемых направлений современной теории вероятностей. Тема диссертации, безусловно, является актуальной, как в теоретическом отношении, так и ввиду многочисленных возможных приложений в различных областях науки и техники, таких как теория надежности, финансовая и актуарная математика, медицина, биология, метеорология и др. Интерес для приложений связан с необходимостью при принятии практических решений располагать данными об асимптотических свойствах времен наступления исключительных и редких событий, моделируемых рекордами и экстремальными порядковыми статистиками (отказов оборудования, наводнений, страховых случаев, связанных с крупными выплатами страховых компаний и т.п.)

Диссертация состоит из восьми глав и списка литературы из 228-ми наименований. Общий объем работы – 260 страниц.

Перейдем к краткому обзору содержания диссертации. Глава 1 представляет собой введение, в котором приводятся основные понятия и обозначения, делается краткий обзор основных результатов работы.

В главе 2 приводятся результаты, обобщающие классическую лемму Бореля - Кантелли. В частности, получены полезные обобщения этой леммы на случай, когда события последовательности связаны марковской зависимостью. Доказанные леммы 2.2.1 и 2.2.2 обобщают утверждение первой части леммы Бореля - Кантелли, которое, как показано, остается справедливым при несколько более слабом, чем в классической лемме, условии. Леммы 2.2.3 – 2.2.5 обобщают утверждение второй части леммы Бореля - Кантелли (формулируемой для последовательности независимых событий), на случай, когда события последовательности связаны марковской зависимостью. Найденные обобщения леммы Бореля - Кантелли используются в § 2.3 для исследования асимптотических свойств максимумов в случае F^α-схемы. Однако главным образом эти обобщения находят применение при доказательстве предельных теорем главы 5.

В главе 3 рассматриваются рекордные величины и рекордные времена. Основные результаты, представленные в диссертации, касаются свойств слабых рекордов и рекордов с подтверждением. В § 3.2 рассматривается последовательность независимых одинаково распределенных случайных величин $X_1, X_2, \ldots$, принимающих значения $0, 1 \ldots$. Вводятся случайные величины ξ_i^w, равные числу слабых рекордов, зарегистрированных в точке i. Ключевую роль при доказательстве асимптотических свойств слабых рекордов играют доказанная здесь лемма 3.2.4, в которой утверждается, что случайные величины ξ_i^w независимы между собой и имеют геометрические распределения, и представление 3.2.1, которое связывает функцию распределения n-го слабого рекорда $X^w(n)$ с функцией распределения суммы независимых случайных величин $\xi_0^w + \cdots + \xi_m^w$ (количества зарегистрированных слабых рекордов). Лемма 3.2.4 и представление 3.2.1 позволяют в дальнейшем при доказательстве предельных теорем для слабых рекордов рассматривать вместо зависимых слабых рекордных величин суммы независимых геометрически распределенных случайных величин. В § 3.3 доказываются предельные теоремы для слабых рекордов в последовательности независимых одинаково распределенных целочисленных случайных величин $X_1, X_2, \ldots$. На самом деле предельные теоремы (закон больших чисел, центральная предельная теорема , закон повторного логарифма) доказываются для величины $R(X^w(n))$ (случайного числа неслучайных слагаемых), где

$$R(n) = \sum_{i=0}^{n} \mathbf{E}\xi_i^w, \quad \mathbf{E}\xi_i^w = \frac{1-\beta_i}{\beta_i} \text{ и } \beta_i = \mathbf{P}(X_1 > i | X_1 \geq i).$$

В теореме 3.3.1 доказывается, что если $\inf_{n\geq 0} \beta_i > 0$, то $\frac{1}{n} R^w(X(n)) \to 1$ с вероятностью 1. В теореме 3.3.2 утверждается справедливость центральной предельной теоремы и закона повторного логарифма для величины $R(X^w(n))$. Теоремы 3.3.1 и 3.3.2 доказываются путем сведения к известным результатам того же типа для сумм независимых случайных величин, существенную роль при этом играет представление 3.2.1. Теорема 3.3.3 касается вероятностей больших уклонений для слабых рекордов $X^w(n)$, эта теорема также выводится по сути как следствие известных результатов В.В.Петрова о вероятностях больших уклонений для сумм независимых случайных величин.

В § 3.4 изучается асимптотическое поведение отношений слабых рекордных величин $Z^w(n, m) = X^w(n + m)/X^w(n)$ при фиксированном m и при $n \to \infty$ (в п. 3.4.2) и индикаторных величин ξ_i^w (в п. 3.4.3). В этом параграфе используется теория правильно меняющихся функций.

Глава 4 посвящена построению статистических процедур, базирующихся на выборках, состоящих из рекордных значений. Интерес к таким процедурам обусловлен тем, что обычные классические выборки в ряде случаев могут быть недоступны, как, например, в случае испытаний устройств на надежность или материалов на прочность. В диссертации предложены статистические критерии проверки гипотезы однородности против альтернативы стохастического доминирования, основанные на рекордных величинах и рангах рекордов. Эти критерии аналогичны хорошо известному критерию Вилкоксона, использующему обычные ранги наблюдений в классической выборке. В главе 4 находятся распределения тестовых статистик, приводятся таблицы вычисленных критических точек этих распределений, вычисляются мощности тестов при альтернативах Лемана. В § 4.3 обсуждаются вопросы, связанные с информацией Фишера, содержащейся в рекордах, слабых рекордах и числах рекордных величин. В одном из основных результатов этого параграфа, теореме 4.3.1, вычисляется информация Фишера, содержащаяся в верхних и нижних рекордных величинах и временах в случае, когда выборка получена из распределений, принадлежащих семействам экспоненциального типа. В § 4.4 обсуждаются тесты, основанные на рекордных величинах с подтверждениями. Предложен тест, основанный на рекордных временах с подтверждением, который позволяет детектировать присутствие в данной выборке нетипичных "посторонних" наблюдений, что очень важно, например, в теории и практике робастной статистики.

В главе 5 исследуются асимптотические свойства случайных величин, определяемых как числа элементов выборки, находящихся в окрестностях порядковых статистик, а также вблизи рекордных значений. В отличие от ситуации, рассмотренной в главе 3, исходное распределение предполагается теперь непрерывным. Вводятся величины $K_{\pm}(n,k,a)$ – числа наблюдений, находящихся в левой и правой окрестностях k-й порядковой статистики, и величины $\xi_n(a)$ – числа наблюдений, регистрируемых в левой окрестности рекордных значений $X(n)$, последние являются непрерывными аналогами числа повторений слабых рекордов ξ_i из главы 3.

В § 5.2-5.4 получен ряд результатов об асимптотическом поведении величин $\xi_n(a)$, их сумм, а также величин $K_{\pm}(n,k,a)$. В теореме 5.2.1 исследовано предельное распределение числа величин, находящихся в окрестности порядковой статистики $X_{k:n}$, в зависимости от "толщины" хвоста исходного распределения. В теореме 5.2.2 даны необходимые и достаточные условия для того, чтобы это число сходилось к нулю с вероятностью 1. В теоремах 5.2.3-5.2.4 представлены асимптотические свойства числа наблюдений, регистрируемых в правой окрестности $X_{k:n}$, эти свойства, очевидно, оказываются зависящими от того, является ли носитель распределения ограниченным. В § 5.4 исследуются асимптотика $\xi_n(a)$ – числа наблюдений, регистрируемых вблизи рекордных значений. Находится распределение числа околорекордных величин. В частности, в теореме 5.4.1 доказывается, что если носитель распределения ограничен, то число околорекордных величин стремится к ∞ с вероятностью 1, причем его рост экспоненциален по n. В п. 5.4.4 изучается асимптотическое поведение сумм околорекордных величин. В § 5.5 основные результаты главы 5 иллюстрируются большим числом различных примеров.

В главе 6 изучается предельное поведение серий, образованных спейсингами порядковых статистик и рекордов. Рассматриваются выборки двух типов: обычная классическая выборка и выборка, составленная из рекордов, которую предлагается использовать для статистических выводов в случае, когда первая выборка недоступна. Для фиксированного $\varepsilon > 0$ вводятся индикаторные величины, связанные со спейсингами порядковых статистик (рекордов), которые полагаются равными 1, если соответствующий спейсинг $> \varepsilon$, и 0 в противном случае. Последовательность спейсингов порядковых статистик (рекордов) образует серию длиной k, если в соответствующей последовательности индикаторов между двумя последовательными единицами находится k нулей, т.е. k последовательных спейсингов не превосходят величины ε. Изучается предельное поведение самой длинной серии, а также числа всех серий.

Глава 7 посвящена изучению асимптотических свойств конкомитантов порядковых статистик и рекордов. В § 7.1 вводятся основные понятия и определения, в § 7.2 рассматриваются распределения конкомитантов порядковых статистик и рекордов, предлагается классификация двумерных распределений с точки зрения асимптотических свойств конкомитантов верхних порядковых статистик и рекордных величин, вводится понятие с-конкоминант-стабильного распределения. В § 7.3 доказываются сильные (т.е. утверждающие сходимость п.н.) предельные теоремы для конкомитентов верхних порядковых статистик в случае, когда F-с-конкоминант-стабильное распределение. В теоремах 7.3.1 и 7.3.2 даются необходимые и достаточные условия сходимости конкомитентов верхних порядковых статистик к пределу с вероятностью 1, дальнейшие асимптотические результаты для конкомитентов верхних порядковых статистик приводятся в § 7.4. В § 7.5 аналогичные результаты приводятся для конкомитентов рекордов. В § 7.6 приведены примеры, иллюстрирующие результаты главы 7, предложены методы компьютерного генерирования конкомитантов верхних порядковых статистик и рекордов, основанные на результатах главы 7.

В главе 8 диссертации доказываются характеризационные теоремы. Известно, что свойства порядковых статистик и рекордов могут быть использованы для однозначной идентификации (характеризации) ряда вероятностных распределений. Диссертант получает новые характеризационные теоремы, основанные на свойствах слабых рекордов и свойствах порядковых статистик.

Имеются некоторые критические замечания.

1. Стр. 13 и 34: не вполне точно сформулирована лемма 2.2.1: "если для некоторого $m \geq 0$, то". Однако при $m = 0$ и $n = 1$ в выделенной формуле появляется событие A_0, которое не определено. Кроме того, при $m = 0$ условие $\sum_{n=1}^{\infty} \mathrm{P}(A_n^c A_{n-1}^c A_n) < \infty$ выполнено тривиально, но отсюда не следует, что $\mathrm{P}(A_n \, б.ч.) = 0$.

2. § 3.3, стр. 64. Асимптотические результаты, названные предельными теоремами для слабых рекордов, на самом деле относятся к случайному числу неслучайных слагаемых $R(X^w(n))$ – сумме математических ожиданий чисел повторений слабых рекордов в точках от 0 до $X^w(n)$. Не совсем ясно, с чем связан интерес к изучению асимптотики таких сумм, а не собственно сумм слабых рекордов. В диссертации это никак не комментируется.

3. Теорема 3.3.2: утверждения этой теоремы – центральная предельная теорема и закон повторного логарифма для $R(X^w(n))$ – доказываются сначала для сумм независимых геометрически распределенных случайных величин ξ_i (чисел повторений слабых рекордов), затем путем ссылки на представление 3.2.1, связывающее между собой распределение слабого рекорда $X^w(n)$ и распределение суммы ξ_i, одной фразой утверждается справедливость центральной предельной теоремы (причем, со скоростью сходимости в ней порядка $n^{-1/2}$) и закона повторного логарифма для величин $R(X^w(n))$. Данный переход не очевиден и требует более детального доказательства.

4. Стр 64, строка 8 снизу. В доказательстве теоремы 3.3.1 утверждается, что из того, что $\frac{S^w(n)}{R(n)} \to 1$ п.н. следует, что с вероятностью 1 при всех достаточно больших m и всех $\delta > 0$ справедливы неравенства $(1-\delta)R(m) < S^w(m) < (1+\delta)R(m)$. Здесь необходимо уточнение, поскольку номер m, начиная с которого выполняются указанные двойные неравенства, вообще говоря, зависит от события ω.

5. Имеется несколько опечаток на стр. 64. 1) строка 6 сверху. Напечатано: $\sum_{i=0}^{n}(1-\beta_i) = \infty$, то есть конечная сумма равна бесконечности (должен быть ряд). 2) Строка 8 сверху – опечатка в нижнем индексе в выделенной формуле, напечатано $k + 0$, должно быть $k = 0$.

6. Стр. 71, строка 2 снизу. Интервал $0 < x < o(\sqrt{n})$ назван в примере зоной нормальной сходимости. Однако зоной нормальной сходимости принято считать область значений аргумента x, в которой отношение хвоста распределения к соответствующему хвосту нормального распределения равномерно по x сходится к 1 при $n \to \infty$. В данном примере, как следует из теоремы 3.3.3 автора, эта зона имеет вид $0 < x < o(n^{1/6})$.

7. § 4.3, стр. 90, формула (4.2.2): неточность в формулировке альтернатив, которыми названы ситуации, когда существуют значения аргументов, при которых $F_X > F_Y$. Из контекста ясно, что под альтернативой здесь понимается стохастическое доминирование, при котором $F_X \geq F_Y$ для

всех значений аргумента, и существует хотя бы одно значение аргумента, при котором неравенство строгое.

8. Стр. 233, строка 3 снизу. Напечатано: aeyrwbz $M_{i,n}(x)$. Должно быть: функция $M_{i,n}(x)$.

Указанные замечания носят в основном редакционный характер и не влияют на общую положительную оценку диссертации.

Вообще, приведенные в диссертации научные результаты позволяют квалифицировать их как значимое научное достижение. Работа носит теоретический характер, ее результаты могут найти применение в теории предельных теорем математической статистики и в различных прикладных областях, связанных с регистрацией и изучением экстремальных и рекордных наблюдений.

Все результаты являются новыми, они неоднократно докладывались на российских и международных научных конференциях и семинарах. Результаты диссертации опубликованы в 42 статьях автора (частью совместных), около 40 статей опубликованы в изданиях, рекомендованных ВАК, причем около 30 из них – в международных журналах, имеющих высокий индекс цитирования (импакт-фактор). Автореферат правильно отражает содержание диссертации.

Диссертация Алексея Васильевича Степанова соответствует всем требованиям "Положения о порядке присуждения ученых степеней" ВАК РФ, а ее автор, А.В.Степанов, безусловно, заслуживает присуждения ему ученой степени доктора физико-математических наук по специальности 01.01.05 – теория вероятностей и математическая статистика.

Официальный оппонент

профессор кафедры "Математика и моделирование"
ФГБОУ ВПО "Петербургский государственный университет
путей сообщения Императора Александра I",

доктор физ.-мат. наук Грибкова Надежда Викторовна

29.04.2015

190031, Санкт-Петербург, Московский пр., 9, ПГУПС
Телефоны: 8 (812)572-61-17; 457-87-73
nv.gribkova@gmail.com

Подпись руки Н. В. Грибковой
удостоверяю.
Начальник Службы управления персоналом
университета Г.Е. Егоров
29 апреля 2015 г.