ОТЗЫВ ОФИЦИАЛЬНОГО ОППОНЕНТА

о диссертационной работе Голдаевой Анны Алексеевны "**Тяжелые хвосты, экстремумы и кластеры линейных стохастических рекуррентных последовательностей**", представленной на соискание ученой степени кандидата физико-математических наук по специальности 01.01.05 - теория вероятностей и математическая статистика.

Содержание работы и основные результаты

Диссертация посвящена важной и актуальной теме: исследованию кластеров случайных процессов, возникающих вследствие наличия зависимостей и тяжелых хвостов распределения процессов. Под кластерами могут пониматься блоки данных с хотя бы одним превышением процессом какого-то уровня; блоки данных с превышениями, разделяемыми заданным числом наблюдений процесса, пробегающих под уровнем или конгломераты последовательных превышений процессом некоторого уровня. В силу возможности широкой трактовки и необходимости моделирования кластеров превышений уровней в различных прикладных областях таких, как сейсмология, климатология, финансовые, телекоммуникационные и социальные системы, исследование стохастических свойств таких кластеров играет важную роль в анализе экстремальных величин. Краеугольный камень в исследовании кластеров случайных последовательностей был заложен Лидбеттером, Линдгреном и Рутзеном в 1983г.

Одной из главных характеристик такой кластерной структуры является экстремальный индекс, имеющий множество интерпретаций. В частности, наиболее прагматичная интерпретация, лежащая в основе многих оценок, следующая: обратная величина экстремального индекса аппроксимирует предельный средний размер кластера. Экстремальный индекс позволяет оценить предельное распределение максимума из конечного числа зависимых случайных величин, а также его квантили; распределения размеров кластеров, т.е. числа превышений в кластере, и межкластерных расстояний. Последние важны в оценивании распределений так называемых периодов возврата, важных в сейсмологии и др. областях. Экстремальный индекс различных процессов, в том числе и нелинейных, рассматривался многими авторами, например, Микошем, Клюппельберг, Вайсманном, Смит, Зегерсом, Робинсоном, Бейрлантом, Гомес и многими другими.

В работе определяются экстремальный индекс или его границы, а также хвостовой индекс, показывающий тяжесть хвоста распределения, для линейных процессов со случайными коэфициентами с различными распределениями. Работа содержит большое количество примеров, которые не только улучшают понимание работы, но и представляют самостоятельный интерес для дальнейших исследований.

Диссертация, объемом 94 стр., содержит введение, четыре главы и список литературы из 51 наименования.

Содержание Главы 1.

Большое внимание в Главе 1 диссертации и далее уделяется авторегрессионному процессу вида

$$Y_n = A_n Y_{n-1} + B_n, \qquad n \geq 1, Y_0 \geq 0 \tag{1}$$

со случайными независимыми между собой коэффициентами (A_n, B_n). Для это процесса получаются индексы: хвостовой и экстремальный, при различных условиях на распределения коэффициентов. Результаты работы (не только Главы 1) во многом базируются на Теореме А классической работы de Haan et al (1983). В этой теореме показано, что экстремальный индекс авторегрессионного процесса определяется только коэффициентами A_n, а также дано простое представление экстремального индекса через kый максимум последовательности $S_n = \sum_{i=1}^{n} \ln A_i$, $n \geq 1$ (Замечание 1.1.1

диссертации). Это представление часто используется в диссертации для получения собственных результатов.

В диссертации результаты работы de Haan et al (1983) расширены. Дается ответ на вопрос: как различаются хвостовой и экстремальный индексы двух процессов с разными коэффициентами A_n, связанными некоторым соотношением. Приводятся примеры некоторых соотношений. В Теореме 1.1.1, принадлежащей автору, показано, что если $\ln A_n$ для какого-то процесса может быть представлен в виде суммы случайного числа одинаково распределенных независимых слагаемых такого же типа $\ln A_i$, то хвостовой индекс этого процесса будет такой же, как у процессов, соответствующих $\ln A_i$, а экстремальный индекс будет больше.

Далее доказывается Теорема 1.4.1. Из нее следует, что любой процесс Y_n, удовлетворяющий (1), у которого коэффициенты A_n представляют собой геометрическую броуновскую смесь, удовлетворяет и условиям Теоремы А. Для Y_n существует процесс $\widetilde{Y}_n$ с теми же коэффициентами A_n, представляющий собой последовательность наблюдений процесса (1.9) (в нумерации диссертации), полученные через случайные интервалы времени. Причем указанные процессы имеют одни и те же хвостовые и экстремальные индексы. Полученные в теореме результаты имеют важное практическое значение, поскольку показывают, что процесс непрерывного времени может быть заменен процессом дискретного времени, что всегда и имеет место на практике.

Содержание Главы 2.

Глава посвящена нахождению границ экстремального индекса для броуновских смесей. Для этого рассматривается процесс $(X_t)_{t\geq 0}$, удовлетворяющий стохастическому дифференциальному уравнению (1.9), и процесс Y_n, совпадающий с (X_t) только в некоторых независимых, одинаково распределенных моментах времени. Чтобы найти границы экстремального индекса используется метод, предложенный в работе Klüppelberg (2004). Теорема 2.1.1 содержит верхнюю оценку экстремального индекса, которая косвенно (через параметры) зависит от так называемой масштабной функции $s(x)$. Это согласуется, например, с работой Robinson and Tawn (2000), где показано, что экстремальный индекс зависит от масштаба. Таким образом, полученная оценка не является грубой только при $s(x)$ близкой к нулю. Это показано и на Рис. 2.1. Поэтому автор предлагает использовать комбинированную оценку: там, где $s(x)$ близко к нулю, используется так называя первая равномерная оценка экстремального индекса из Теоремы 2.1.1, в противном случае - вторая равномерная оценка (2.11).

Резюмируя результаты глав 1 и 2 можно сказать, что получены важные результаты по исследованию свойств экстремумов авторегрессионных процессов со случайными параметрами. Приведена коллекция примеров (не все принадлежат диссертантке), содержащие точные выражения для экстремального и хвостового индексов или границы экстремального индекса для ряда процессов. Это особенно важно для моделирования кластеров превышений процессов. Другой, практически важный, на мой взгляд, результат содержится в Теореме 1.4.1., обосновывающий существование процесса с дискретным временем, который может заменить процесс с непрерывным временем, не потеряв экстремальных свойств в том смысле, что экстремальные индексы обоих процессов совпадают.

Содержание Главы 3.

В §3.1 приводится теорема о непрерывности, где дается ответ на вопрос, сохраняются ли близкие значения экстремального и хвостового индексов, если распределения коэффициентов A_1 заменены на коэффициенты с близкими в каком-то смысле распределениями.

В §3.2 дается пример троичного распределения для величины $Z_1 = \ln A_1$, которая может принимать три значения $1, -1, 0$ с соответствующими вероятностями p, q, r. Для троичного распределения в Теореме 3.2.1 получены хвостовой и экстремальный индексы процесса (1), а также распределение размера кластера.

В Теореме 3.3.1 получены хвостовой и экстремальный индексы процесса (1) при $r > \alpha/(\alpha+\beta)$ в предположении, что величина Z_1 имеет обобщенное лапласовское распределение. Это является обобщением результатов работы Новицкой и Яцало 2008 в случае, если $\beta > \alpha$, поскольку $r = \beta/(\alpha+\beta)$ в этой работе. Распределения k-го максимума T_k и размера кластера получены в Теоремах 3.3.2 и 3.3.3. в условиях Теоремы 3.3.1.

Содержание Главы 4.

Глава 4 посвящена нахождению экстремального и хвостового индексов для многомерных последовательностей, которые с помощью линейных комбинаций $Y_n(z) = \sum_{i=1}^{n} z_i Y_n^{(i)}$ с положительными коэффициентами последовательностей случайных величин $\{Y_n^{(i)}\}_{1\le i\le l}$ сводятся к одномерному случаю. Рассматриваются только неотрицательные случайные величины $\{Y_n^{(i)}\}$, имеющие распределения с регулярно меняющимися хвостами и некоторыми хвостовыми индексами. В теореме 4.1.1 доказывается, что если эти индексы различны, то хвостовой индекс линейной комбинации совпадает с минимальным хвостовым индексом, а экстремальный индекс будет равен экстремальному индексу соответствующей одномерной последовательности $\{Y_n^{(i)}\}$. В §4.2 исследуется случай независимых случайных величин $Y_n^{(1)}, ..., Y_n^{(l)}$, когда все хвостовые индексы равны. В Теореме 4.2.1 доказывается, что экстремальные индексы $Y_n(z)$ и $Y_n^*(z) = \max(z_1 Y_n^{(1)}, ..., z_l Y_n^{(l)})$ совпадают и приводится их значение. В Теореме 4.3.1 доказывается, что при переходе к собственному базису экстремальный индекс не меняется. §4.4 посвящен исследованию многомерного процесса с непрерывным временем, удовлетворяющего векторно-матричному стохастическому дифференциальному уравнению (4.7). В Теореме 4.4.1 получено решение (4.7). Далее рассматриваются проблемы перехода к другому базису и как при этом изменяется экстремальный индекс.

Замечания

1. Не всегда четко обозначена мотивация и применение той или иной рассматриваемой задачи. Например, в §1 Главы 1 не объясняется, для чего нужно изучать совместно два линейных процесса типа (1) Y_n и $\tilde{Y}_n$. Только в Главе 2 это становится ясно.

2. Изложение не всегда удачно, какие-то вещи не поясняются или определяются нечетко. Например, приводя Теорему A из работы de Haan et al (1983), часто используемую в диссертации, не определяется, что $T_0 = \infty$, что важно дальше при определении θ_k, зависящего от $k-1$-го максимума T_{k-1}. Пояснение, что π_k в Теореме A - это вероятность иметь в кластере k превышений уровня, дается только перед Теоремой 3.3.3. Другие примеры: в теореме 2.1.1 не указано, что параметр $a = -(d+\sigma^2/2)$, а это важно для понимания, что верхняя оценка экстремального индекса зависит от масштабной функции $s(x)$ через параметры d и σ; в §2.1 мера скорости $m(x)$ процесса не определена.

3. В работе содержится много рисунков, но выводы из них, как правило, не делаются.

4. Название Главы 1 "Основные свойства индексов"неудачно. Надо было конкретно указать, какие индексы имеются ввиду.

5. Что касается обзора литературы, то в связи с доказательством распределения размера кластера в Теореме 3.3.3, было бы уместно упомянуть работы с этим связанные, например, среди них:
   Robert C.Y. 2009 Inference for the limiting cluster size distribution of extreme values, The Annals of Statistics, Vol.37 (1), 271-310
   Hsing T 1991 Estimating the parameters of rare events. Stochastic Process. Appl. 37, 117-139,
   а в связи с зависимостью экстремального индекса от масштаба данных - работу

Robinson, M.E., Tawn, J.A.: Extremal analysis of processes sampled at different frequences. Journal of the Royal Statistical Society Series B. 62(1), 117–135 (2000)

Результаты диссертации опубликованы в 5ти статьях российских журналов из списка ВАК, доложены на двух конференциях молодых ученых "Ломоносов-2010, 2012"и одна работа депонирована в ВИНИТИ. На мой взгляд, недостатком является, что результаты не были доложены ни на одной международной конференции.

Заключение.

Указанные недостатки касаются стилистики изложения и не умаляют достоинств диссертации. Результаты диссертации являются новыми, актуальными и соответствующими мировому уровню. Считаю, что диссертация полностью соответствует требованиям специальности 01.01.05 - теория вероятностей и математическая статистика, а ее автор Голдаева А.А. заслуживает присуждения ученой степени кандидата физико-математических наук.

Официальный оппонент,
доктор физ-мат наук, главный научный сотрудник Института проблем упраления им. В.А. Трапезникова РАН
Маркович Наталья Михайловна

Подпись Маркович Н.М.

17.02.2014

Подпись Маркович Н.М.
ЗАВЕРЯЮ
ЗАМ. ЗАВ. ОБЩИМ ОТДЕЛОМ
АФАНАСЬЕВА Н.С.