

УДК 550.343

# Метод оценивания вероятностей катастроф в неоднородных потоках экстремальных событий и его применение к прогнозированию землетрясений в Арктике<sup>1</sup>

ISSN 1812-5220

© Проблемы анализа риска, 2016

**В. Ю. Королев,**

Московский государственный университет им. М. В. Ломоносова; Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, г. Москва

**Е. В. Арефьева,**

ФГБУ ВНИИ ГОЧС МЧС России (ФЦ), г. Москва

**Ю. С. Нефедова,**

Московский государственный университет им. М. В. Ломоносова; Institute for Advanced Studies, Vienna, Austria

**А. К. Горшенин,**

Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, г. Москва

**Р. А. Лазовский,**

Московский государственный университет им. М. В. Ломоносова

**Аннотация**

В статье рассматривается задача прогнозирования вероятностей катастроф в неоднородных потоках экстремальных событий. Эта задача решается с помощью метода, основанного на предельной теореме для геометрических случайных сумм независимых неодинаково распределенных случайных величин и теории Балкемы — Пикандса — Де Хаана. Рассмотрена конструкция, в рамках которой в качестве предельного распределения для геометрических случайных сумм независимых неодинаково распределенных случайных величин возникает распределение Вейбулла — Гнеденко. Эффективность метода иллюстрируется на примере его применения к прогнозированию вероятностных характеристик землетрясений в Арктике.

**Ключевые слова:** катастрофа, экстремальное событие, случайная сумма, геометрическая сумма, закон больших чисел, распределение Вейбулла — Гнеденко, теорема Балкемы — Пикандса — Де Хаана, обобщенное распределение Парето.

**Содержание**

Введение

1. Метод прогнозирования вероятностей катастроф в неоднородных потоках экстремальных событий
2. Пример применения метода к прогнозированию вероятностных характеристик землетрясений в Арктике

Заключение

Литература

**Введение**

В данной статье рассматривается задача прогнозирования вероятностных характеристик катастроф в неоднородных потоках экстремальных событий. Рассмотрим некоторую систему, подвергающуюся влиянию некоторого фактора. Предположим, что сила воздействия этого фактора на систему в каждый момент

<sup>1</sup> Работа поддержана Российским фондом фундаментальных исследований (проект 15-07-04040а).

времени характеризуется некоторым числом, причём это число изменяется во времени. Это может быть:

- финансовая система, которая характеризуется финансовым индексом, таким как DAX, NIKKEI, NASDAQ и т.п.; при этом резкие колебания индекса неблагоприятны и свидетельствуют о тех или иных кризисных явлениях;
- экологическая система, например среда обитания человека, в частности состояние жилых помещений в местностях, подверженных наводнениям или землетрясениям в сейсмоопасных зонах, которое зависит соответственно от силы подземных толчков и уровня подъёма воды;
- социальная система, которая подвержена политической неустойчивости или воздействию террористических организаций; при этом в качестве числовой характеристики активности неблагоприятных воздействий может выступить, к примеру, число упоминаний некоторых соответствующих ключевых слов или фраз в социальных информационных сетях или сетях связи;
- наконец, общее состояние планеты Земля, зависящее от расстояния, на которое подлетают к ней потенциально опасные космические объекты — астероиды или кометы.

При этом, естественно, возникает задача прогнозирования катастроф. Однако без применения специализированных методов, ориентированных на противодействие конкретным рискам, практически никогда нельзя абсолютно точно предсказать силу воздействия фактора на систему в каждый момент времени в будущем. Другими словами, будущее развитие фактора непредсказуемо, вследствие чего значение числа, характеризующего силу воздействия фактора на систему, рассматриваемое как функция времени, целесообразно рассматривать как *случайный процесс*. Поэтому задача прогнозирования самого момента катастрофы сводится к прогнозированию значения случайного процесса (т.е. его значения на вполне определенном элементе множества элементарных исходов) специальными методами, что чрезвычайно трудоемко и при рассмотрении современных сложных стохастических систем практически не реализуемо с приемлемой точностью. В то же время возникает вполне реальная и важная задача прогнозирования *распреде-*

*ния* указанного случайного процесса в те или иные моменты времени, т.е. задача прогнозирования его статистических свойств. В результате решения этой задачи появляется возможность правильно оценить уровни угрозы в каждой конкретной ситуации. Некоторым методам решения последней задачи и посвящена данная статья.

Предположим, что очень большие изменения случайного процесса, характеризующего воздействие фактора на систему, неблагоприятно влияют на систему и могут вызвать ее необратимые изменения. Вместе с тем малые флуктуации случайного процесса, характеризующего воздействие фактора на систему, вполне допустимы (в таких случаях говорят о «фоновом значении» фактора). Поэтому с целью предсказания катастроф разумно рассматривать не все изменения случайного процесса, а лишь такие, величина которых превышает некоторый *потенциально опасный порог*.

Будем говорить, что моменты превышений изменениями случайного процесса потенциально опасного порога в совокупности с самими значениями этих превышений образуют *экстремальный случайный процесс*. Другими словами, экстремальным процессом будем называть маркированный точечный процесс  $\{(\tau_i, X_i)\}_{i \geq 1}$ , где  $\{(\tau_i)\}_{i \geq 1}$  — точечный случайный процесс, а  $\{X_i\}_{i \geq 1}$  — случайные величины. Далее по смыслу задачи будет предполагаться, что  $X_i > 0, i = 1, 2, \dots$

Среди всех превышений случайным процессом потенциально опасного порога лишь некоторые очень большие влекут катастрофические последствия. Поэтому наряду с *потенциально опасным порогом* рассмотрим критический *порог*, превышение которого экстремальным процессом и будем считать *катастрофой*.

Для удобства точку отсчета (нуль временной шкалы) поместим в то время, которое будем считать «настоящим». Тем самым «настоящее» характеризуется значением  $t = 0$ .

Поскольку по условию экстремальный процесс считается случайным, то нельзя точно предсказать момент наступления очередной катастрофы. Однако можно вычислить или оценить *вероятности наступления катастрофы* в течение некоторого интервала времени  $[0, \tau)$ , где  $\tau > 0$ . Если  $T$  — момент наступления катастрофы, то событие «ка-

катастрофа наступила в течение интервала времени  $[0, \tau]$  эквивалентно тому, что  $T < \tau$ . В качестве исходных данных будем использовать информацию о развитии экстремального процесса на некотором интервале времени  $[t_0, t_1]$ , где  $t_0 < t_1 < 0$ .

Простейшее (примитивное) решение задачи об отыскании вероятности наступления катастрофы в течение интервала времени  $[0, \tau]$  при условии  $\tau < t_1 - t_0$  выглядит так.

Разобьем интервал времени  $[t_0, t_1]$  на непересекающиеся подынтервалы длиной  $\tau$ . Пусть внутри интервала  $[t_0, t_1]$  поместилось  $N_\tau$  подынтервалов длиной  $\tau$ . Подсчитаем количество подынтервалов, внутри каждого из которых наступила хотя бы одна катастрофа. Пусть таких подынтервалов оказалось ровно  $n_\tau$ . Тогда для вероятности наступления катастрофы в течение интервала времени  $[0, \tau]$  справедлива оценка

$$P(T < \tau) \approx \frac{n_\tau}{N_\tau}, \quad (1)$$

основанная на классическом определении вероятности как (предела) частоты.

Недостатки такой оценки очевидны. Например,  $n_\tau$  просто может оказаться равным нулю, что дает тривиально оптимистичную оценку. Далее, и  $N_\tau$ , и  $n_\tau$  могут быть (и, как правило, являются) слишком маленькими, чтобы обеспечить приемлемую точность оценки. Более того, одной-единственной катастрофы может оказаться достаточно для полного уничтожения системы, так что дальнейший сбор информации просто может оказаться невозможным.

## 1. Метод прогнозирования вероятностей катастроф в неоднородных потоках экстремальных событий

### 1.1. Особенности метода

К сожалению, именно оценками типа (1) зачастую пользуются на практике для расчетов, связанных с так называемыми большими рисками в страховании, например при страховании промышленных рисков, связанных с крупными авариями и экологическими катастрофами. В данной статье описан метод оценивания указанных вероятностей на-

ступления катастроф, основанный на довольно сложных математических моделях, но свободный от указанных недостатков. Особенность этого метода заключается в том, что для того, чтобы прогнозировать возможности наступления катастроф, не обязательно иметь статистику самих катастроф.

Простейший вариант этого метода описан в работах [1, 2, 3] и книгах [4, 5], где предполагалось, что экстремальный процесс является маркированным процессом восстановления, а именно: в указанных работах предполагалось, что моменты  $\tau_1, \tau_2, \dots$  превышений исходным процессом потенциально опасного порога образуют процесс восстановления. Это означает, что случайные величины

$$\zeta_i = \tau_i - \tau_{i-1}, \quad i = 1, 2, \dots, \tau_0 = 0 \quad (2)$$

независимы и имеют одинаковое распределение, то есть подчиняются одним и тем же статистическим закономерностям. Другими словами, интенсивность потока экстремальных событий считалась постоянной. В то же время в реальных сложных системах, которые в подавляющем большинстве случаев не являются информационно и/или энергетически замкнутыми и подвержены влиянию внешней среды, интенсивности потоков информативных событий не являются постоянными. Например, при адекватном прогнозировании поведения финансовых индексов ключевую роль играет представление о том, что интенсивности потоков информативных событий на финансовых рынках являются случайными [6]. Отказ от предположения о постоянстве интенсивности потока экстремальных событий, естественно, приводит к необходимости предположить, что случайные величины (2) имеют неодинаковое распределение. Такое обобщение методов, предложенных в работах [2—5], рассмотрено в статье [7]. Суть метода, описанного в [7], заключается в следующем.

Обозначим величину превышения исходным процессом потенциально опасного порога в момент  $\tau_i$  символом  $X_i$ ,  $i = 1, 2, \dots$  Будем считать, что  $X_1, X_2, \dots$  — независимые и одинаково распределенные случайные величины. Это означает, что значения этих случайных величин подчиняются одним и тем же статистическим закономерностям, характеризуемым функцией распределения

$F(x) = P(X_i < x)$ ,  $-\infty < x < \infty$ ,  $i = 1, 2, \dots$  Будем считать, что последовательность  $X_1, X_2, \dots$  статистически независима от последовательности  $\tau_1, \tau_2, \dots$

Пусть  $x_0$  — критический порог, превышение которого значением  $X_i$  и есть катастрофа (то есть катастрофическое событие формально записывается в виде неравенства  $X_i \geq x_0$ ).

Очевидно, что время  $T$  наступления катастрофы (то есть время первого превышения уровня  $x_0$  какой-либо из величин  $X_i$ ) можно представить в виде геометрической случайной суммы

$$T = \sum_{j=1}^N \zeta_j, \quad (3)$$

где случайные величины  $\zeta_i$  определены соотношением (2), а  $N$  — это случайная величина, имеющая геометрическое распределение с параметром  $P(X_i < x_0) = F(x_0)$ . Это означает, что  $P(N = k) = (F(x_0))^{k-1}(1 - F(x_0))$ ,  $k = 1, 2, \dots$  При этом в силу независимости последовательностей  $X_1, X_2, \dots$  и  $\tau_1, \tau_2, \dots$  число  $N$  слагаемых в сумме (3) независимо от самих слагаемых  $\zeta_1, \zeta_2, \dots$  При этом принципиальным отличием геометрических случайных сумм, рассматриваемых здесь, от геометрических сумм в традиционном понимании (см., например, [8, 9]) является то, что в данном случае слагаемые имеют *неодинаковое* распределение, тогда как в указанных классических книгах изучались геометрические суммы *одинаково* распределенных слагаемых и, соответственно, использовались методы, ориентированные именно на такую ситуацию.

В рамках подхода, рассматриваемого в данной статье, краеугольными камнями являются два теоретических результата. Первый из них — версия закона больших чисел для случайных сумм неодинаково распределенных случайных величин (теорема 1, ниже), обосновывающая использование распределения Вейбулла — Гнеденко в качестве модели распределения интервалов времени между катастрофами. Второй — теорема Балкема — Пикандса — Де Хаана (теорема 2, ниже), обосновывающая использование обобщенного распределения Парето в качестве модели распределения критических значений неблагоприятного фактора. Эти два общих результата являются основой предлагаемого метода.

## 1.2. Вспомогательные результаты

Пусть  $\xi_1, \xi_2, \dots$  — необязательно одинаково распределенные случайные величины. Для каждого натурального  $n \geq 1$  положим  $S_n = \xi_1 + \xi_2 + \dots + \xi_n$ . Рассмотрим последовательность целочисленных неотрицательных случайных величин  $\{N_n\}_{n \geq 1}$  и будем считать, что при каждом  $n$  случайные величины  $N_n, \xi_1, \xi_2, \dots$  независимы в совокупности. Более того, предположим, что

$$N_n \rightarrow \infty \text{ по вероятности при } n \rightarrow \infty. \quad (4)$$

Условие (4) означает, что  $P(N_n \leq m) \rightarrow 0$  при  $n \rightarrow \infty$  для любого  $m > 0$ . Везде далее символ  $\Rightarrow$  будет обозначать сходимость по распределению.

*Лемма 1. Пусть для некоторой последовательности положительных чисел  $\{b_n\}_{n \geq 1}$  выполнены условия  $b_n \rightarrow \infty$  при  $n \rightarrow \infty$  и*

$$\frac{S_n}{b_n} \Rightarrow 1, n \rightarrow \infty. \quad (5)$$

*Предположим, что выполнено условие (4). Для того чтобы при  $n \rightarrow \infty$  имела место сходимость случайных сумм  $S_{N_n}$ , нормированных некоторой последовательностью положительных чисел  $\{d_n\}_{n \geq 1}$ , такой, что  $d_n \rightarrow \infty$  при  $n \rightarrow \infty$ , к некоторой случайной величине  $Z$ :*

$$\frac{S_{N_n}}{d_n} \Rightarrow 1, \quad (6)$$

*необходимо и достаточно, чтобы*

$$\frac{b_{N_n}}{d_n} \Rightarrow 1, n \rightarrow \infty. \quad (7)$$

*Замечание 1.* В силу вырожденности распределения предельной случайной величины в (5) сходимость по распределению (5) оказывается эквивалентной сходимости по вероятности: для любого  $\varepsilon > 0$

$$\lim_{n \rightarrow \infty} P\left(\left|\frac{S_n}{b_n} - 1\right| > \varepsilon\right) = 0,$$

которую иногда легче проверять.

*Доказательство* леммы 1 приведено в статье [10].

*Замечание 2.* Лемма 1 является версией закона больших чисел для случайных сумм. Согласно классическим законам больших чисел при увели-

чении числа слагаемых в рассматриваемых средних арифметических информация о конкретном виде распределений слагаемых затухает, стягиваясь в информацию об одном лишь числе. Точно такой же эффект наблюдается в лемме 1: при рассмотрении случайных средних арифметических информация о распределениях слагаемых затухает, так что предельное распределение случайного среднего арифметического определяется видом предельного распределения для случайного индекса (числа слагаемых в сумме) при надлежащей нормировке.

Для общности пусть  $x_n = x_{0,n}$  — (возрастающая) последовательность критических порогов такая, что

$$p_n \equiv 1 - F(x_n) \rightarrow 0 \quad (n \rightarrow \infty). \quad (8)$$

Тогда в данном случае случайная величина  $N = N_n$  имеет геометрическое распределение с параметром  $q_n = 1 - p_n$ . При этом условие (8) гарантирует выполнение условия (4). Более того,  $EN_n = p_n^{-1}$  и, как хорошо известно:

$$\limsup_{n \rightarrow \infty} \left| P(p_n N_n \geq y) - e^{-y} \right| = 0. \quad (9)$$

Предположим, что постоянные  $b_n$ , обеспечивающие выполнение условия (5), имеют вид  $b_n = bn^\gamma$  при некоторых  $b > 0$  и  $\gamma > 0$ . При этом значения  $\gamma > 1$  соответствуют той ситуации, когда случайные величины  $\zeta_i$  «в среднем» возрастают, то есть экстремальные события происходят все реже и реже, значения  $\gamma < 1$  соответствуют той ситуации, когда случайные величины  $\zeta_i$  «в среднем» убывают, то есть экстремальные события происходят все чаще и чаще, а значение  $\gamma = 1$  соответствует той ситуации, когда интенсивность потока экстремальных событий «в среднем» постоянна, например, в поведении интенсивности наблюдаются проявления цикличности, причем периоды изменения интенсивности заметно меньше периода фиксации наблюдений.

Теперь выберем нормирующие постоянные  $d_n$  так, чтобы геометрическая случайная сумма  $S_{N_n}$  имела нетривиальное предельное распределение. Из леммы 1 вытекает, что если с учетом выбранной формы постоянных  $b_n$  и соотношения (7) постоянные  $d_n$  выбрать в виде  $d_n = bp_n^{-\gamma}$ , то для любого  $y > 0$

$$\begin{aligned} \lim_{n \rightarrow \infty} P\left(\frac{b_{N_n}}{d_n} < y\right) &= \lim_{n \rightarrow \infty} P\left(\left(p_n N_n\right)^\gamma < y\right) = \\ &= \lim_{n \rightarrow \infty} P\left(p_n N_n < y^{1/\gamma}\right) = 1 - \exp\{-y^{1/\gamma}\}. \end{aligned}$$

При этом согласно лемме 1 такое же распределение Вейбулла — Гнеденко с показателем  $1/\gamma$  является предельным и для геометрической случайной суммы независимых неодинаково распределенных случайных величин  $S_{N_n}$ , причем в силу непрерывности предельного распределения Вейбулла — Гнеденко сходимость (6) равномерна по  $x \in \mathbb{R}$ . Оформим сказанное в виде следующего утверждения.

**Теорема 1.** *Предположим, что случайная величина  $N_n$  имеет геометрическое распределение с параметром  $p_n$ , причем  $p_n \rightarrow \infty$  при  $n \rightarrow \infty$ . Предположим, что существуют конечные  $\gamma > 0$  и  $b > 0$  такие, что*

$$\frac{S_n}{bn^\gamma} \Rightarrow 1 \quad (n \rightarrow \infty).$$

Тогда

$$\limsup_{n \rightarrow \infty} \left| P\left(p_n^\gamma S_{N_n} \geq bx\right) - \exp\{-x^{1/\gamma}\} \right| = 0.$$

### 1.3. Описание метода

Итак, учитывая сделанные предположения о нормирующих постоянных, можно заключить, что при достаточно больших значениях  $x_0$

$$P(T < t) \approx 1 - \exp\left\{-[1 - F(x_0)]\left(\frac{t}{b}\right)^{1/\gamma}\right\}, \quad t > 0. \quad (10)$$

Применение описываемого метода вычисления временных характеристик катастроф в неоднородных потоках экстремальных событий заключается в следующем. Пусть  $\varepsilon \in (0, 1)$  — произвольное число. Решение уравнения

$$P(T < t) = \varepsilon$$

относительно  $t$  обозначим  $t(\varepsilon)$ . Если распределение случайной величины  $T$  имеет вид (10), то, очевидно,

$$t(\varepsilon) = b \left[ \frac{\ln(1-\varepsilon)}{F(x_0) - 1} \right]^\gamma.$$

Смысл значения  $t(\varepsilon)$  — это тот момент времени, вероятность наступления катастрофы до которого равна  $\varepsilon$ . Из соображений здравого смысла особый интерес представляют значения  $\varepsilon$ , близкие к нулю (соответствующее значение  $t(\varepsilon)$  — это то время, до которого катастрофа, скорее всего, не наступит), близкие к единице (соответствующее значение  $t(\varepsilon)$  — это то время, до которого катастрофа, скорее всего, наступит), а также  $\varepsilon = 1/2$  (соответствующее

значение  $t(1/2)$  — это «среднее» время до наступления катастрофы).

Особо следует сказать, что при прогнозировании «среднего» или «ожидаемого» времени до катастрофы можно использовать как медиану  $t(1/2)$  случайной величины  $T$ , которая определяется как решение уравнения

$$1 - \exp\left\{-[1 - F(x_0)]\left(\frac{t}{b}\right)^{1/\gamma}\right\} = \frac{1}{2}$$

относительно  $t$  и, очевидно, равна

$$t(1/2) = b \left[ \frac{\ln 2}{1 - F(x_0)} \right]^\gamma,$$

так и математическое ожидание

$$ET = \frac{b\Gamma(1 + \gamma)}{[1 - F(x_0)]^\gamma}.$$

При этом необходимо отметить, что, например, в случае  $\gamma = 1$  медиана  $t(1/2)$  случайной величины  $T$  почти в полтора раза (точнее, в  $(\ln 2)^{-1}$  раз) меньше математического ожидания  $ET$ .

При этом параметры  $b$  и  $\gamma$  легко оценить методом наименьших квадратов. Предположим, что в нашем распоряжении имеется выборка  $Z_1, Z_2, \dots, Z_n$  предыдущих значений случайных величин  $\zeta_i$ . Нормирующая функция  $b_k = bk^{\gamma}$  параметра  $k$  имеет смысл тренда, или основной тенденции поведения реализации  $R_k = Z_1 + \dots + Z_k$  случайной функции  $S_k$ . С целью линеаризации регрессионной задачи прологарифмируем  $b_k$  и  $R_k$ , обозначим  $\beta = \log b$  и получим приближенные равенства

$$\log R_k \approx \beta + \gamma \log k, \quad k = 1, \dots, n, \quad (11)$$

в правой части которых стоят линейные функции параметров  $\beta$  и  $\gamma$ . Используя стандартный метод наименьших квадратов оценивания параметров линейной регрессии (11), получим оценки

$$\gamma \approx \hat{\gamma} = \frac{n \sum_{k=1}^n (\log k \cdot \log R_k) - \log n! \sum_{k=1}^n \log R_k}{n \sum_{k=1}^n (\log k)^2 - (\log n!)^2},$$

$$b = \exp\{\beta\} \approx \exp\left\{\frac{1}{n} \left( \sum_{k=1}^n \log R_k - \hat{\gamma} \log n! \right)\right\}.$$

Чтобы получить оценку величины  $1 - F(x_0)$ , необходимо построить разумную и адекватную параметрическую математическую модель (приближе-

ние) для функции  $F(x)$ . С этой целью используем метод построения асимптотических аппроксимаций для  $F(x)$  при больших  $x_0$ , основанный на теореме Балкема — Пикандса — Де Хаана и называемый методом превышений порога (POT-метод, POT = Peaks Over Threshold).

Пусть случайная величина  $\zeta$  имеет функцию распределения  $F(x)$ . В рамках рассматриваемого метода прогнозирования катастроф как превышений экстремальным процессом критических уровней большой интерес представляет описание функции  $F_u(y) = P(\zeta - u < y \mid \zeta > u)$ , условного распределения превышения случайной величиной  $\zeta$  некоторого (большого) порога  $u$ ,  $0 \leq y \leq x_F - u$ , где  $y = x - u$  — превышение порога и  $x_F = \sup\{x \in \mathfrak{R} : F(x) < 1\} \leq \infty$ . Функция этого условного распределения  $F_u$  может быть выражена через  $F$ :

$$F_u(y) = \frac{F(u+y) - F(u)}{1 - F(u)} = \frac{F(x) - F(u)}{1 - F(u)}.$$

Если порог  $u$  достаточно велик, то большинство реализаций случайной величины  $\zeta$  лежит между  $0$  и  $u$ , так что оценить  $F$  в этом промежутке несложно. Но оценить  $F_u$  проблематично, так как соответствующих наблюдений мало. На помощь приходит следующая теорема.

**ТЕОРЕМА 2** [11, 12]. *Функция распределения  $F$  принадлежит области тах-притяжения распределения, предельного для экстремальных значений, тогда и только тогда, когда существует измеримая функция  $\sigma(u) > 0$ , такая, что*

$$\limsup_{u \rightarrow x_F} \sup_{0 \leq y < x_F - u} |F_u(y) - G_{\delta, \sigma(u)}(y)| = 0,$$

где

$$G_{\delta, \sigma}(y) = \begin{cases} 1 - \left(1 + \frac{\delta}{\sigma} y\right)^{-1/\delta}, & \delta \neq 0, \\ 1 - e^{-y/\sigma}, & \delta = 0, \end{cases}$$

функция обобщенного распределения Парето.

Условием теоремы удовлетворяет большинство используемых на практике распределений. Параметр  $\delta$  показывает, насколько тяжел хвост: чем больше  $\delta$ , тем тяжелее хвост. Например, при моделировании финансовых данных обычно используется  $\delta \geq 0$ .

## 2. Пример применения метода к прогнозированию вероятностных характеристик землетрясений в Арктике

Описанный выше метод был применен к прогнозированию вероятностных характеристик землетрясений в Арктике. Сильные землетрясения носят катастрофический характер, уступая по числу жертв только тайфунам и значительно (в десятки раз) опережая извержения вулканов. Материальный ущерб одного разрушительного землетрясения может составлять сотни миллионов долларов. Число слабых землетрясений гораздо больше, чем сильных. Так, из сотни тысяч землетрясений, ежегодно происходящих на Земле, только единицы катастрофических. Они высвобождают около 1020 Дж потенциальной сейсмической энергии, что составляет всего 0,01% тепловой энергии Земли, излучаемой в космическое пространство.

Северный морской путь, короткая дорога из Атлантики в Тихий океан, с каждым годом становится все более оживленным. Количество перевозимых по маршруту грузов за последние пять лет выросло в четыре раза. Оценка состояния льда и изменения климатических условий в Арктике важны не только для судов, но и для нефте- и газодобычи. Доказательство этого — самая масштабная за последние двадцать лет научно-исследовательская экспедиция в этот регион, организованная Роснефтью.

Для успешного освоения Арктики необходимо следить за поведением земной коры. В морях Северного Ледовитого океана случаются сильные землетрясения, их магнитуда может достигать семи баллов. Слабая сейсмическая активность отмечается практически на всем Арктическом шельфе. Даже незначительные землетрясения могут провоцировать большие последствия для инженерных сооружений. Например, известны случаи, когда небольшие землетрясения с магнитудой ниже трех баллов вызывали оползни.

Причина землетрясений в Арктике — так называемые метановые бомбы. Это выбросы большого количества газа из толщи Земли. Для кораблей встреча с метановыми бомбами равносильна попаданию на мину. Восточно-Сибирский Арктический шельф является самым крупным источником метана в мире. Объемы газа, выходящего здесь

на поверхность, сопоставимы с выбросами со дна всего Мирового океана. Основная часть метана сосредоточена в так называемых газогидратах, похожих на спрессованный в кристаллы снег, в состав которых входят молекулы воды и газа. Пока льда и на поверхности воды, и на дне много, газогидраты ведут себя тихо, но, когда его становится мало, кристаллы легко распадаются на составляющие, сотни тысяч кубометров метана вырываются со дна на поверхность огромными гейзерами, способные перевернуть даже буровую платформу. Из-за глобального потепления и таяния льдов в районе Арктического шельфа метановым пузырям становится легче прорваться на поверхность, и, как следствие, возрастает число и сила землетрясений в регионе. Более точные данные о рисках, связанных с сейсмической активностью, могут сделать добычу ископаемых в Арктике, а также проводку судов более безопасными.

В качестве исходных данных использованы данные, опубликованные на сайте геологической службы США <http://earthquake.usgs.gov/earthquakes/>, где приведен полный хронологический список арктических землетрясений начиная с 1950 г. Сила землетрясений измеряется в магнитудах. Магнитуда землетрясений — величина, характеризующая энергию, выделившуюся при землетрясении в виде сейсмических волн. Она обычно определяется по шкале, основанной на записях сейсмографов. Эта шкала известна под названием шкалы магнитуд, или шкалы Рихтера (по имени американского сейсмолога Ч. Ф. Рихтера, предложившего ее в 1935 г.). Магнитуда землетрясения — безразмерная величина, пропорциональная логарифму отношения максимальных амплитуд определенного типа волн данного землетрясения и некоторого стандартного землетрясения. Существуют различия в методах определения магнитуд близких, удаленных, мелкофокусных (неглубоких) и глубоких землетрясений.

Для удобства все данные из указанного списка арктических землетрясений сведены в диаграмму, приведенную на рис. 1.

По представленным данным можно сделать грубую частотную оценку: землетрясения с магнитудой не менее 6,5 происходят раз в 1475 дней  $\approx$  4,04 года. Но, во-первых, частотная оценка типа среднего арифметического никак не учитывает сис-

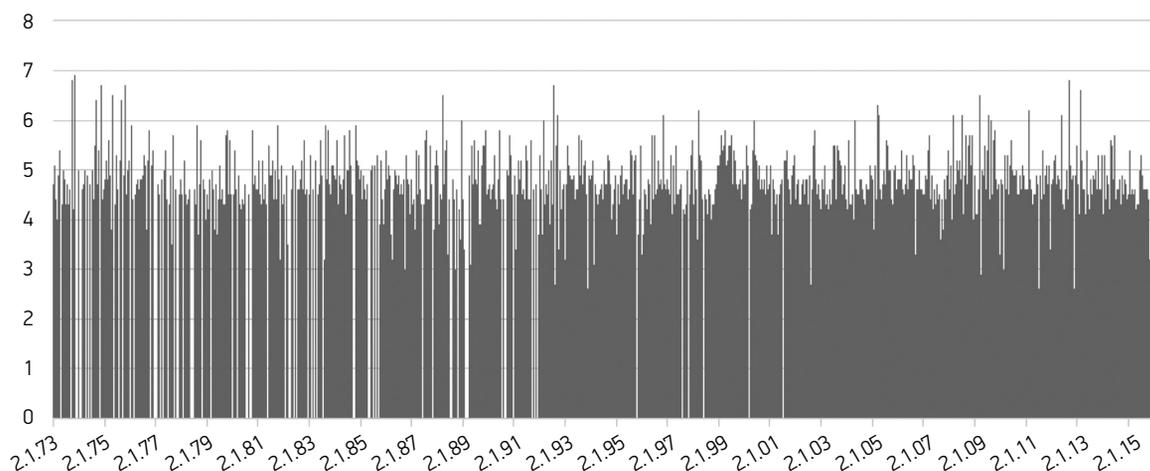


Рис. 1. Данные о землетрясениях в Арктике

тематические глобальные изменения и смещения горных масс, могущие повлиять на интенсивность землетрясений, и, во-вторых, слишком груба, поскольку объем выборки — число интервалов времени между соответствующими землетрясениями — очень мал (двенадцать интервалов для магнитуды, превышающей значение 6,5). Для землетрясений с магнитудой не менее 6,9 частотный подход уже не применим, т. к. имеется лишь одно наблюдение с такой характеристикой.

Землетрясения разной магнитуды (по шкале Рихтера) проявляются следующим образом (табл. 1).

На основании данной классификации был взят потенциально опасный порог  $\theta$ , равный 4,5 магнитуды. При таком значении порога получились следующие оценки максимального правдоподобия параметров обобщенного распределения Парето:

- параметр формы  $\delta = 0,184501$ ;
- параметр масштаба  $\sigma = 0,304574$ ;
- параметр сдвига  $\theta = 4,5$ .

Таблица 1

Почти неощутимые толчки	—
Толчки ощущают лишь немногие, особенно на верхних этажах	2
Толчки ощущают немногие, дребезжит стекло, раскачиваются всяческие предметы	2,5—3
Толчки ощущают все, кто находится внутри здания, трескаются потолки, звенит посуда	3,5
Толчки ощущают все, спящие люди просыпаются, в помещении раскачиваются всяческие предметы	4—4,5
Просыпаются спящие, люди покидают дома, останавливаются настенные часы с маятником, сильно раскачиваются деревья	5
Трескаются стены домов, осыпается штукатурка	5,5—6
Образуются обширные и глубокие трещины в стенах, сдвигается, иногда опрокидывается мебель, трещины в грунтах достигают нескольких сантиметров	6—6,5
В стенах возникают бреши, рушатся перегородки, трещины в грунтах достигают 10 см	7
Здания рушатся, реки выходят из берегов, трещины в грунтах несколько дециметров, иногда около 1 м	7,4
Повреждение большинства зданий, разрушение мостов, значительные деформации почвы, горные обвалы	8,0
Почти полное разрушение, радикальное изменение земной поверхности	8,9

Оценки наименьших квадратов для параметров  $b$  и  $\gamma$  оказались равными

$$b \approx 86,8341 \text{ дня} \approx 0,2379 \text{ года}, \gamma = 0,7185.$$

Необходимо особо отметить, что значение параметра  $\gamma$  оказалось заметно меньше единицы, что в соответствии со сказанным выше означает, что интенсивность землетрясений (число землетрясений в единицу времени) возрастает. Этот вывод наглядно иллюстрируется графиком, приведенным на рис. 2, где по горизонтальной оси располагаются порядковые номера, а по вертикальной оси — длины интервалов времени между последовательными землетрясениями, то есть высота очередного столбика на этом графике равна длине соответствующего временного промежутка между землетрясениями. На рис. 2 ясно видна тенденция к уменьшению высоты столбиков с ростом номера интервала. Другими словами, землетрясения в Арктике действительно происходят все чаще и чаще.

Гистограмма и плотность распределения Парето, подогнанного к данным с порогом  $\theta = 4,5$ , изображены на рис. 3 слева, а соответствующая кривая  $y(k) = bk^\gamma$ , подогнанная к  $R(k)$  по тем же данным, изображена на рис. 3 справа.

В табл. 2 приведены значения  $t(\varepsilon)$  — времени, вероятность наступления катастрофы до которого равна  $\varepsilon$ :

$$t(\varepsilon) = b \left[ \frac{\ln(1-\varepsilon)}{F(x_0)-1} \right]^\gamma,$$

для некоторых  $x_0$  и  $\varepsilon = 0,001, 0,999$  и  $0,5$ .

Из табл. 2 видно, что среднее время между катастрофическими землетрясениями с магнитудами не менее 6,5 равно примерно 4,06 года, а среднее время между катастрофическими землетрясениями с магнитудами не менее 7 равно примерно 6,69 года.

Однако ученые считают, что землетрясения с магнитудой более 9 произойти на Земле не могут. И вот почему. Известно, что каждое землетрясение представляет собой толчок или серию толчков, которые возникают в результате смещения горных масс по разлому. Расчеты показали, что размер очага землетрясения (то есть величина площади, на которой произошло смещение горных пород, которыми и определяется сила землетрясения и его энергия) при слабых, едва ощутимых человеком толчках измеряется в длину и по вертикали несколькими метрами. При землетрясениях средней силы, когда возникают трещины в каменных зданиях, размеры очага достигают уже километров. Очаги же при самых сильных, катастрофических землетрясениях имеют протяженность 500—1000 км и уходят на глубину до 50 км. У максимального из зарегистрированных на Земле землетрясений очаг равен  $1000 \times 100$  км, т.е. близок к максимальной длине

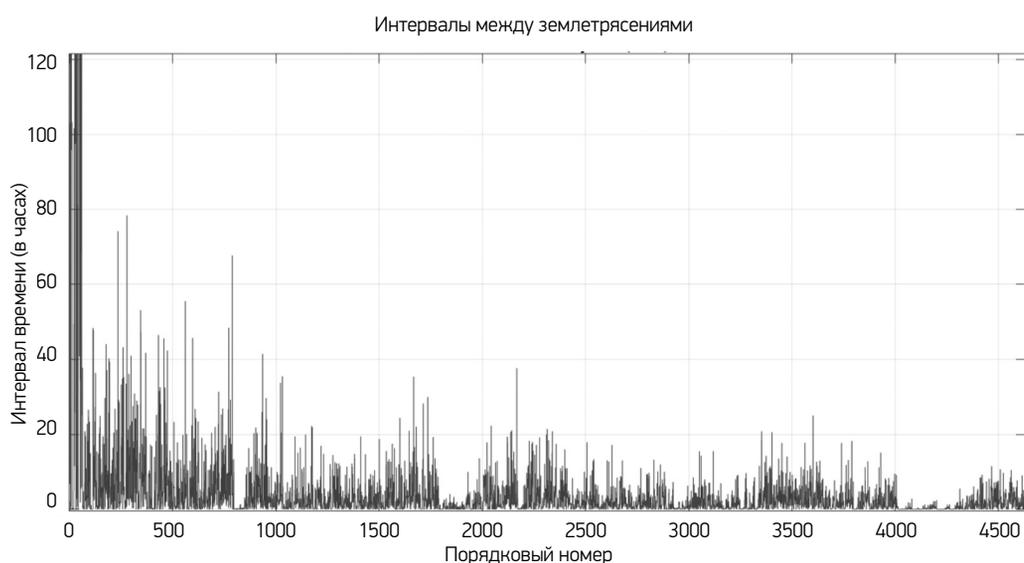


Рис. 2. Тенденция уменьшения интервалов времени между последовательными землетрясениями

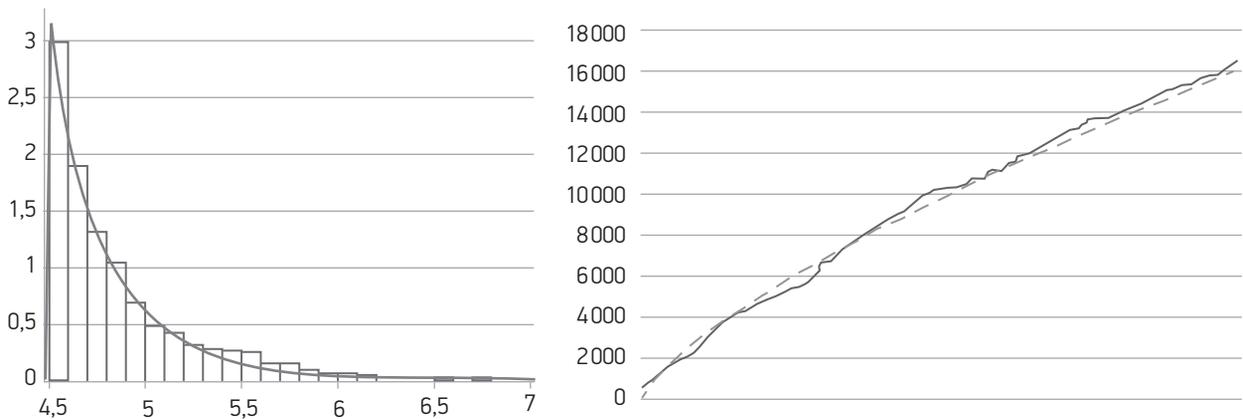


Рис. 3. Плотность распределения Парето, подогнанного к данным с порогом  $\theta = 4,5$  (слева), и кривая  $y(k) = bk^\gamma$ , подогнанная к  $R(k)$  по тем же данным (справа)

разломов, известных ученым. Невозможно и дальнейшее увеличение глубины очага, так как земное вещество на глубинах более 100 км переходит в состояние, близкое к плавлению. Следовательно, такое землетрясение, как Чилийское, можно считать близким к максимальному.

В связи с этим вместо аппроксимации неизвестной функции распределения  $F(x)$  распределением Парето для оценки  $1 - F(x_0)$  можно использовать эмпирическую функцию распределения, которая на имеющихся данных принимает вид

$$F(x) = \begin{cases} 0 & x \leq 4,5 \\ 0,7443 & 4,5 < x \leq 5 \\ 0,9194 & 5 < x \leq 5,5 \\ 0,9782 & 5,5 < x \leq 6 \\ 0,9922 & 6 < x \leq 6,5 \\ 1 & x > 6,5 \end{cases}$$

В табл. 3 приведены значения  $t(\epsilon)$  — времени, вероятность наступления катастрофы до которого равна  $\epsilon$ .

### Заключение

Представленные в статье задачи прогнозирования рисков экстремальных событий на основе использования метода, основанного на предельной теореме для геометрических случайных сумм независимых неодинаково распределенных случайных величин и теории Балкемы — Пикандса — Де Хаана

Значения  $t(\epsilon)$  (порог  $\theta = 4,5$ )

Таблица 2

$x_0$	$\epsilon = 0,001$		$\epsilon = 0,999$		$\epsilon = 0,5$	
	дни	годы	дни	годы	дни	годы
5	1.7325	0.0047	993.2877	2.7213	190.3902	0.5216
5,5	3.9000	0.0107	2234.1000	6.1208	428.2000	1.1732
6	7.6000	0.0208	4368.6000	11.9688	837.4000	2.2942
6,5	13.5000	0.0370	7739.6000	21.2044	1483.5000	4.0644
7	22.0000	0.0603	12742.0000	34.9096	2442.0000	6.6904

Значения  $t(\epsilon)$  (порог  $\theta$ ) для эмпирической функции распределения

Таблица 3

$x_0$	$\epsilon = 0,001$		$\epsilon = 0,999$		$\epsilon = 0,5$	
	дни	годы	дни	годы	дни	годы
5	1.6177	0.0044	927.4695	2.5410	177.7744	0.4871
5,5	3.7000	0.0101	2125.9000	5.8244	407.5000	1.1164
6	9.5000	0.0260	5439.7000	14.9033	1042.7000	2.8567
6,5	20.0000	0.0548	11384.0000	31.1890	2182.0000	5.9781

и получения распределения Вейбулла — Гнеденко как предельного распределения для рассматриваемых случайных величин, позволяют получать более адекватные оценки статистических параметров экстремальных событий. Разработанный метод позволяет оценить вероятность наступления катастрофического события, выполнять прогнозы редких

экстремальных событий, не обязательно основываясь на статистике самих этих катастрофических событий. При этом случайные величины имеют не обязательно одинаковое распределение, что является существенным моментом, повышающим адекватность прогнозирования рисков. Эффективность метода иллюстрируется на примере его применения к прогнозированию вероятностных характеристик землетрясений в Арктике.

Метод может найти широкое применение в прогнозировании рисков экстремальных чрезвычайных ситуаций как природного, так и техногенного характера.

## Литература

1. Королев В.Ю., Соколов И.А. Некоторые вопросы анализа катастрофических рисков, связанных с неоднородными потоками экстремальных событий // Системы и средства информатики. Спец. вып. Математические методы и модели информатики. Стохастические технологии и системы. М.: ИПИ РАН, 2005. С. 109—125.
2. Королев В.Ю., Соколов И.А., Гордеев А.С., Григорьева М.Е., Попов С.В., Чебоненко Н.А. Некоторые методы анализа временных характеристик и катастроф в неоднородных потоках экстремальных событий // Системы и средства информатики. Спец. вып. Математические методы в информационных технологиях. М.: ИПИ РАН, 2006. С. 5—23.
3. Королев В.Ю., Соколов И.А., Гордеев А.С., Григорьева М.Е., Попов С.В., Чебоненко Н.А. Некоторые методы прогнозирования временных характеристик рисков, связанных с катастрофическими событиями // Актуарий, 2007. № 1. С. 34—40.
4. Королев В.Ю., Соколов И.А. Математические модели неоднородных потоков экстремальных событий. М.: Торус Пресс, 2008. 200 с.
5. Королев В.Ю., Шоргин С.Я. Математические методы анализа стохастической структуры информационных потоков. М.: ИПИ РАН, 2011. 130 с.
6. Королев В.Ю., Черток А.В., Корчагин А.Ю., Горшенин А.К. Вероятностно-статистическое моделирование информационных потоков в сложных финансовых системах на основе высокочастотных данных // Информатика и ее применения, 2013. Т. 7. Вып. 1. С. 12—21.
7. Григорьева М.Е., Королев В.Ю., Соколов И.А. Предельная теорема для геометрических сумм независимых неодинаково распределенных случайных величин и ее применение к прогнозированию вероятности катастроф в неоднородных потоках экстремальных событий // Информатика и ее применения, 2013. Т. 7. Вып. 4. С. 66—74.
8. Kalashnikov V. Geometric Sums: Bounds for Rare Events with Applications. — Dordrecht-Boston-London: Kluwer Academic Publishers, 1997. 288 p.
9. Королев В.Ю., Бенинг В.Е., Шоргин С.Я. Математические основы теории риска. 2-е изд., перераб. и дополн. М.: ФИЗМАТЛИТ, 2011. 620 с.
10. Королев В.Ю. Сходимость случайных последовательностей с независимыми случайными индексами. I // Теория вероятностей и ее применения, 1994. Т. 39. Вып. 2. С. 313—333.
11. Balkema A., de Haan L. Residual life time at great age // Annals of Probability, 1974. Vol. 2. P. 792—804.
12. Pickands J. Statistical inference using extreme order statistics // Annals of Statistics, 1975. Vol. 3. P. 119—131.

## Сведения об авторах

**Королев Виктор Юрьевич:** факультет вычислительной математики и кибернетики Московского государственного университета им. М.В. Ломоносова; Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук  
Количество публикаций: более 280

Область научных интересов: математическая статистика, теория вероятностей

*Контактная информация:*

Адрес: 119991, Москва, Ленинские горы, МГУ, д. 1, стр. 52, факультет ВМК

Тел.: +7 (495) 939-53-94

E-mail: victoryukorolev@yandex.ru

**Арефьева Елена Валентиновна:** ФГБУ «Всероссийский научно-исследовательский институт по проблемам гражданской обороны и чрезвычайных ситуаций МЧС России» (федеральный центр науки и высоких технологий)  
Количество публикаций: более 80

Область научных интересов: анализ и управление рисками, математическое моделирование

*Контактная информация:*

Адрес: 121352, Москва, ул. Давыдовская д. 7

Тел.: +7 (495) 400-90-20

E-mail: elaref@mail.ru

**Нефедова Юлия Сергеевна:** факультет вычислительной математики и кибернетики Московского государственного университета им. М.В. Ломоносова; Institute for Advanced Studies, Vienna, Austria

Количество публикаций: более 40

Область научных интересов: математическая статистика, теория вероятностей

*Контактная информация:*

Адрес: 119991, Москва, Ленинские горы, МГУ, д. 1, стр. 52, факультет ВМК

Тел.: +7 (495) 939-53-94

E-mail: y.nefedova@gmail.com

**Лазовский Руслан Александрович:** факультет вычислительной математики и кибернетики Московского государственного университета им. М.В. Ломоносова

Количество публикаций: более 10

Область научных интересов: математическая статистика, теория вероятностей

*Контактная информация:*

Адрес: 119991, Москва, Ленинские горы, МГУ, д. 1, стр. 52, факультет ВМК

Тел.: +7 (495) 939-53-94

E-mail: gospbuff@gmail.com

**Горшенин Андрей Константинович:** Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Количество публикаций: более 50

Область научных интересов: математическая статистика, теория вероятностей, информационные технологии

*Контактная информация:*

Адрес: 119333, Москва, Вавилова, д. 44, кор. 2

Тел.: +7 (499) 135-62-60

E-mail: a.k.gorshenin@gmail.com