Основные направления научного руководства студентами

А.В. Лебедев

Московский государственный университет им. М.В.Ломоносова
Механико-математический факультет

Москва, МГУ, 1 апреля 2022 года

Стохастическая теория экстремумов

Теория экстремумов, теория экстремальных значений (extreme value theory)

Стохастическая теория экстремумов занимается изучением максимумов и минимумов (а также других порядковых статистик) систем случайных величин.

Фундаментальная работа: Б.В.Гнеденко (1941, 1943)
Предшественники: М.Фреше (1927), Р.Фишер и Л.Типпетт (1928), Р. фон Мизес (1936)
Классические монографии: Я.И.Галамбош (1984), М.Лидбеттер, Г.Линдгрен, Х.Ротсен (1989), P.Embrechts, C.Klüppelberg, T.Mikosh (2003), L. de Haan, A.Ferreira (2006)
Современные исследователи в России: В.И.Питербарг, В.Б.Невзоров, А.В.Степанов, С.Ю.Новак, Н.М.Маркович, А.В.Лебедев, А.А.Голдаева, И.В.Родионов и др. (+ ученики)

Классические результаты

Пусть $\xi_1, \dots, \xi_n, \dots$ — независимые одинаково распределенные случайные величины с ф.р. F и $M_n = \max\{\xi_1, \dots, \xi_n\}$.

Теорема об экстремальных типах (Б.В.Гнеденко)

Если существуют последовательности $a_n > 0$, b_n, $n \geq 1$, и невырожденное распределение G такие, что верно

$$\mathbf{P}(a_n(M_n - b_n) \leq x) \xrightarrow{w} G(x), \quad n \to \infty,$$

то G относится к одному из экстремальных типов, т.е. его ф.р. имеет вид $G_i(ax+b)$, $a > 0$,

$$G_1(x) = \exp\{-e^{-x}\},$$
$$G_2(x) = \begin{cases} 0, & x \leq 0 \\ \exp\{-x^{-\alpha}\}, & x > 0 \end{cases}, \quad \alpha > 0,$$
$$G_3(x) = \begin{cases} \exp\{-(-x)^{\alpha}\}, & x < 0 \\ 1, & x \geq 0 \end{cases}, \quad \alpha > 0.$$

И наоборот, любая такая ф.р. может быть предельной (при $F = G$).

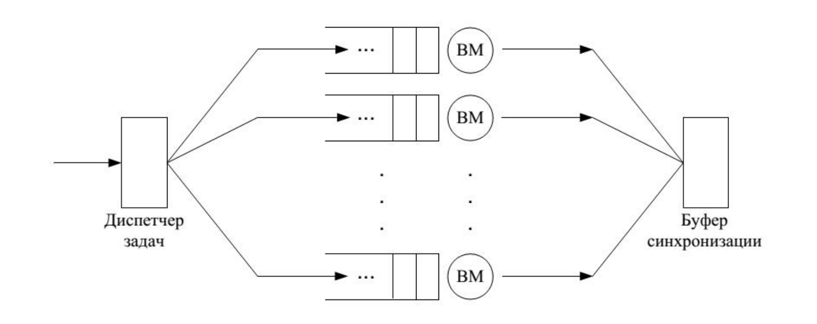
Экстремальный индекс (для стационарных последовательностей)

Определение

Пусть ξ_n, $n \geq 1$, имеют распределение F и $M_n = \vee_{k=1}^{n} \xi_k$. Если для каждого $\tau > 0$ существует такая числовая последовательность $u_n(\tau)$, что $n\bar{F}(u_n(\tau)) \to \tau$ и $\mathbf{P}(M_n \leq u_n(\tau)) \to e^{-\theta\tau}$, то θ называется экстремальным индексом.

Смысл: максимумы n зависимых величин асимптотически растут как максимумы $[\theta n]$ независимых; превышения высокого уровня образуют кластеры среднего размера $1/\theta$. Бывает $\theta \in [0, 1]$.
В приложениях это может означать природные катастрофы, отказы технических систем, потерю данных при передаче информации, финансовые потери и др. Понятно, что когда такие события происходят несколько раз подряд, это гораздо опаснее, чем единичные случаи, и должно учитываться в управлении рисками.

Системы массового обслуживания с разделением заявок (fork-join)



Заявка разделяется на $K \geq 2$ подзаявок. Время отклика (время пребывания заявки в системе) — максимум из времен пребывания подзаявок (они зависимы).

Классика: Нельсон, Тантави (1988)

Последние работы: Горбунова, Вишневский (2020), Горбунова, Лебедев (2022)

Копулы

Определение

Копулой (m-мерной) называется функция многомерного распределения на $[0,1]^m$ с равномерными частными (маргинальными) распределениями.
Копулой распределения F в R^m называется копула C, удовлетворяющая

$$F(x_1, \ldots, x_m) = C(F_1(x_1), \ldots, F_m(x_m)),$$

где $F_1, \ldots, F_m$ — частные функции распределения.

Такое представление существует по теореме Скляра (1959) и единственно в случае непрерывных частных распределений.
Учебник: R.Nelsen (2006).

Примеры (двумерных) копул

Копула Клейтона

$$C(u,v) = (u^{-\theta} + v^{-\theta} - 1)^{-1/\theta}, \quad \theta > 0.$$

Копула Гумбеля

$$C(u,v) = \exp\left\{-\left((-\ln u)^{\theta} + (-\ln v)^{\theta}\right)^{1/\theta}\right\}, \quad \theta \geq 1.$$

Копула Маршалла-Олкина

$$C(u,v) = \min\{u^{1-\alpha}v, uv^{1-\beta}\}, \quad 0 \leq \alpha, \beta \leq 1.$$

Границы Фреше-Хеффдинга

$$\max\{u+v-1, 0\} \leq C(u,v) \leq \min\{u,v\}.$$

Нетранзитивность стохастического предшествования (stochastic precedence)

Пусть заданы три независимые случайные величины X, Y и Z, такие, что

$$\mathbf{P}(X < Y) > \frac{1}{2} \tag{1}$$

и

$$\mathbf{P}(Y < Z) > \frac{1}{2}. \tag{2}$$

Казалось бы, из (1) и (2) должно следовать

$$\mathbf{P}(X < Z) > \frac{1}{2},$$

однако можно построить примеры, когда это не так, а напротив,

$$\mathbf{P}(Z < X) > \frac{1}{2}. \tag{3}$$

Максимум минимума вероятности в случае трех независимых величин

С.Трыбула (1959, с Г.Штейнгаузом; 1961)

$$\max_{X,Y,Z} \min\{\mathbf{P}(X<Y), \mathbf{P}(Y<Z), \mathbf{P}(Z<X)\} = \frac{\sqrt{5}-1}{2} \approx 0.618, \quad (4)$$

причем максимум достигается, например, на триплете

$$\begin{aligned} X &= \begin{cases} 1, & \text{с вероятностью } p, \\ 4, & \text{с вероятностью } 1-p, \end{cases} \\ Y &= 2, \\ Z &= \begin{cases} 0, & \text{с вероятностью } 1-p, \\ 3, & \text{с вероятностью } p, \end{cases} \end{aligned} \quad (5)$$

где

$$p = \frac{\sqrt{5}-1}{2},$$

тогда

$$\mathbf{P}(X<Y) = \mathbf{P}(Y<Z) = \mathbf{P}(Z<X) = \frac{\sqrt{5}-1}{2}.$$

Максимум минимума вероятности в случае n независимых величин

С.Трыбула (1965), З.Усыскин (1964), И.И.Богданов (2010)

$$\max_{X_1,\dots X_n} \min\{\mathbf{P}(X_1 < X_2), \dots, \mathbf{P}(X_{n-1} < X_n), \mathbf{P}(X_n < X_1)\} =$$
$$= 1 - \left(4\cos^2 \frac{\pi}{n+2}\right)^{-1}, \quad n \geq 3.$$

Приложение в теории прочности

С.Трыбула (1961). Пусть в лаборатории сравнивают попарно на прочность железные бруски одинакового размера и формы, помещая их в одну рамку и прилагая к ним одинаковую силу путем закручивания винта, пока один из брусков не сломается. Предположим, что бруски производятся на трех разных заводах (которые дают разное распределение прочности), и сравниваются бруски с первого и второго, второго и третьего, первого и третьего заводов. Тогда теоретически может сложиться парадоксальная ситуация, что бруски с первого завода «хуже» (т.е. чаще ломаются раньше) брусков со второго завода, бруски со второго «хуже» брусков с третьего, а бруски с третьего «хуже» брусков с первого.

Нетранзитивность в природе, технике и обществе

В биологии речь может идти о парных взаимодействиях животных в борьбе за пищу, территорию, размножение или доминирование в группе. Для людей речь может идти о сравнительных оценках различных товаров и услуг, голосовании и др. Проблема нетранзитивности предпочтений известна с XVIII века (*парадокс Кондорсе*).

Общим вопросам нетранзитивности отношения превосходства в природе, технике и обществе посвящен ряд научных и научно-популярных статей А.Н.Поддьякова.

Нетранзитивные кости

Кости Эфрона, изобретенные Б.Эфроном в 1960-е гг., представляют собой набор из четырех костей A, B, C, D, которые имеют на своих гранях следующие числа:

A: 4, 4, 4, 4, 0, 0;
B: 3, 3, 3, 3, 3, 3;
C: 6, 6, 2, 2, 2, 2;
D: 5, 5, 5, 1, 1, 1.

Тогда результат броска каждой кости из набора больше результата бросания следующей кости (по кругу) с вероятностью, большей $1/2$:

$$\mathbf{P}(A > B) = \mathbf{P}(B > C) = \mathbf{P}(C > D) = \mathbf{P}(D > A) = \frac{2}{3}.$$

Нетранзитивные кости популяризованы М.Гарднером (1970).

Нетранзитивность наборов непрерывных случайных величин

Пример

Пусть ф.р. заданы многочленами степени $N \geq 2$ на отрезке $[0,1]$:

$$F(x) = \sum_{k=1}^{N} f_k x^k, \quad \sum_{k=1}^{N} f_k = 1, \quad x \in [0,1],$$

при этом на коэффициенты f_k, $1 \leq k \leq N$, накладываются также ограничения, обусловленные неотрицательностью и возрастанием $F(x)$ от 0 до 1 на этом отрезке.

При $N = 2, 3$ нетранзитивности не может быть.
При $N = 4$ нетранзитивность может быть (получено численным поиском в пространстве коэффициентов).

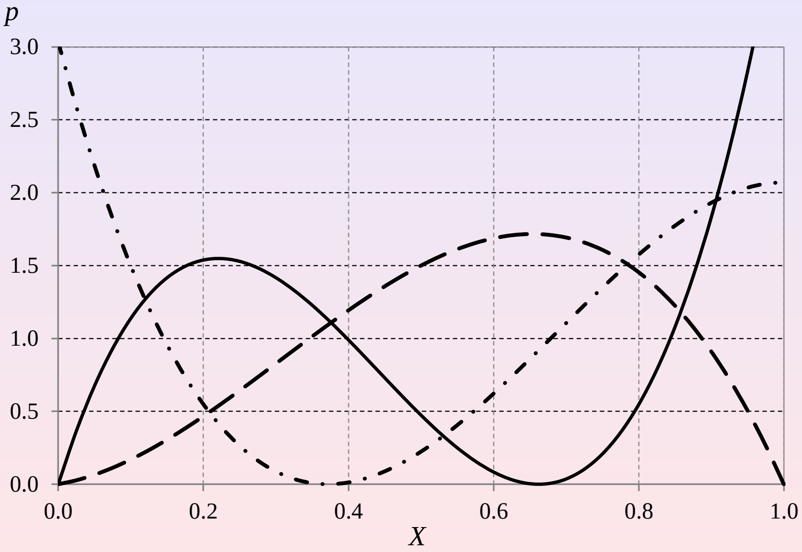
Когда нетранзитивность возможна: полиномиальные плотности на $[0,1]$.

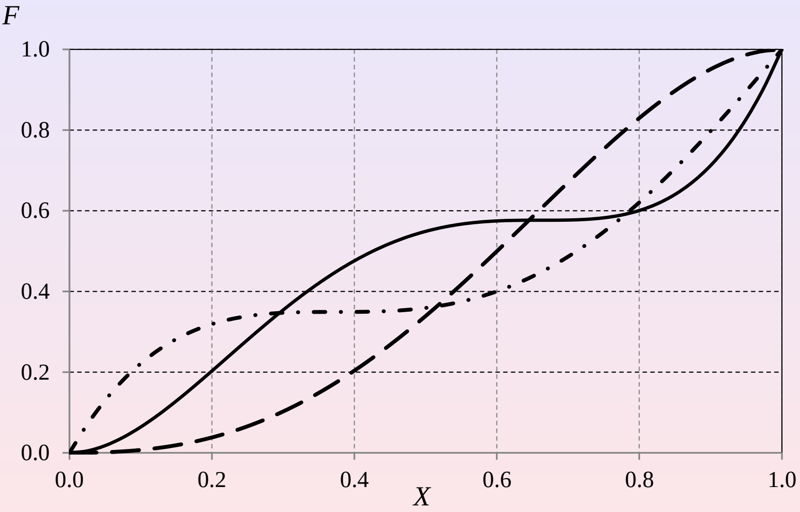
Горбунова, Лебедев (2021). Случай $n=3$:

$$p_{X_1}(x) \approx 15.783x - 47.665x^2 + 35.987x^3,$$
$$p_{X_2}(x) \approx 0.854x + 9.438x^2 - 10.292x^3,$$
$$p_{X_3}(x) \approx 3.034 - 18.631x + 34.374x^2 - 16.706x^3;$$

$$F_{X_1}(x) \approx 7.891x^2 - 15.888x^3 + 8.997x^4,$$
$$F_{X_2}(x) \approx 0.427x^2 + 3.146x^3 - 2.573x^4,$$
$$F_{X_3}(x) \approx 3.034x - 9.3154x^2 + 11.458x^3 - 4.1764x^4.$$

$$\widehat{\max_{X_1,X_2,X_3}} P_{X_1X_2X_3} \approx 0.5350.$$


p
3.0
2.5
2.0
1.5
1.0
0.5
0.0
0.0
0.2
0.4
0.6
0.8
1.0
X


F
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8
1.0
X

Лебедев Алексей Викторович, д.ф-м.н., доцент

E-mail: avlebed@yandex.ru

Семинар по нетранзитивности: goo.gl/bWdMJF

Обращаться студентам 2 курса без “троек”, присылать оценки (из личного кабинета) и желательное направление исследований.

- *Лебедев А.В.* Неклассические задачи стохастической теории экстремумов. Докт. дисс. Москва, 2016.
- *Лебедев А.В.* О возможности существования экстремальных индексов, превосходящих единицу // Вестник Московского университета. Сер. 1. Математика. Механика. 2022. N 1. С. 3–8.
- *Лебедев А.В.* Проблема нетранзитивности для трех непрерывных случайных величин // Автоматика и телемеханика. 2019. N 6. С. 91–103.
- *Лебедев А.В.* Нетранзитивные триплеты непрерывных случайных величин и их приложения // Информатика и ее применения. 2019. Т. 13. N 6. С. 20–26.
- *Горбунова А.В., Лебедев А.В.* Эффекты стохастической нетранзитивности в системах массового обслуживания // Управление большими системами. 2020. Вып. 85. С. 23–50.
- *Горбунова А.В., Лебедев А.В.* Нижние оценки нетранзитивности для трех и четырех случайных величин с полиномиальной плотностью на единичном отрезке // Материалы VI Всероссийской научно-практической конференции международным участием “Современные проблемы физико-математических наук”. Орел. 2020. С. 154–161.
- *Gorbunova A.V., Lebedev A.V.* Response Time Estimate for a Fork-join System with Pareto Distributed Service Time as a Model of a Cloud Computing System Using Neural Networks // Communications in Computer and Information Science. 2022. V. 1552, P. 318–332.
- *Gorbunova A.V., Lebedev A.V.* Nontransitivity of tuples of random variables with polynomial density and its effects in Bayesian models // Mathematics and Computers in Simulation (to appear)