

Комбинированное оценивание распределений с правильно меняющимся правым хвостом

Г.А. Трифонов, П.С. Алексеева

1 Вступление

Анализ экстремальных значений распределений с тяжелыми хвостами является важным инструментом для различных приложений. В сейсмологии и климатологии, например, статистика экстремальных явлений используется для изучения землетрясений [1] или сильных осадков [2]. Еще одной важной областью исследований является анализ высоких финансовых потерь, который становится особенно интересным, если потери зависят от ковариации [3, 4].

Для исследования поведения тяжелых хвостов мы рассматриваем случайные величины из области максимального притяжения (MDA) распределения Фреше.

Пусть $\mathbf{X}_n = (X_1, \dots, X_n)$ – независимые одинаково распределенные (н.о.р.) случайные величины с функцией распределения F . Обозначим $M_n = \max\{X_1, \dots, X_n\}$. Теорема Гнеденко–Фишера–Типшета [5] утверждает, что если для некоторых последовательностей констант $\{a_n > 0\}$ и $\{b_n\}$ выполнено

$$\lim_{n \rightarrow \infty} P \left(\frac{M_n - b_n}{a_n} \leq x \right) = G(x) \quad (1)$$

для некоторой невырожденной ф.р. $G(x)$, то найдутся такие константы $a > 0$ и b , что $G(ax + b) = G_\gamma(x)$, где

$$G_\gamma(x) = \begin{cases} \exp(-(1 + \gamma x)^{-1/\gamma}), & 1 + \gamma x > 0, & \text{если } \gamma \neq 0, \\ \exp(-\exp(-x)), & x \in \mathbb{R}, & \text{если } \gamma = 0. \end{cases} \quad (2)$$

Если выполнены (1) и (2) для ф.р. F , то говорят, что F принадлежит области максимального притяжения $G_\gamma(x)$, пишем $F \in \mathcal{D}(G_\gamma)$. Параметр γ , называемый индексом экстремального значения, позволяет разделить распределения, удовлетворяющие условиям теоремы Гнеденко, на три класса.

Класс $\gamma < 0$ называется областью максимального притяжения Вейбулла, класс $\gamma = 0$ – областью максимального притяжения Гумбеля, а класс $\gamma > 0$, содержащий распределения с хвостами степенного типа, называется областью максимального притяжения Фреше.

Далее нас будет интересовать задача присоединения хвоста, оцененного классическими методами теории экстремумов, к телу распределения, оцененному с помощью непараметрических методов. Во

время комбинированного оценивания модели “тело + хвост” также возникает вопрос о нахождении оптимальной точки присоединения, ответ на который мы также попытаемся дать в этой работе.

2 Модель присоединенного хвоста

Для начала введем необходимые определения, опишем интересующую нас модель и рассмотрим уже имеющиеся результаты.

2.1 Постановка задачи

Пусть X_1, \dots, X_n – положительные н.о.р. случайные величины с неизвестной функцией плотности $f(x)$ и функцией распределения $F(x)$.

Определение 1. Говорят, что распределение F имеет тяжёлый правый хвост, если

$$\int_0^{\infty} e^{tx} dF(x) = \infty, \quad \forall t > 0$$

Примерами таких распределений являются распределения Парето, Бара, Коши и Вейбулла с параметром формы меньше 1.

Определение 2. Положительная измеримая по Лебегу функция $f(x)$ на $[0, \infty)$ называется правильно меняющейся функцией ($f(x) \in \mathbb{R}_\gamma$), если $\lim_{x \rightarrow \infty} f(tx)/f(x) = t^\gamma, \forall t > 0, \gamma \in \mathbb{R}$.

В случае $\gamma = 0$ функция называется медленно меняющейся и обозначается символом l .

Определение 3. Функция F называется функцией распределения с правильно меняющимся правым хвостом на бесконечности, если для достаточно больших x

$$1 - F(x) = x^{-\frac{1}{\gamma}} l(x), \quad \gamma > 0 \quad (1 - F(t) \in \mathbb{R}_{-\frac{1}{\gamma}}), \quad (3)$$

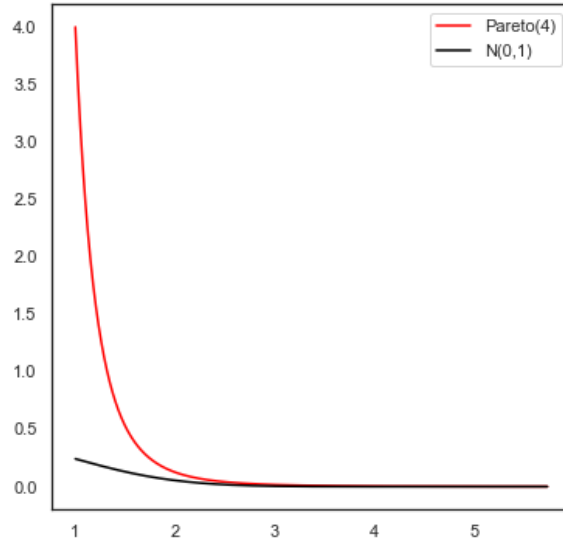
где $l(x)$ является медленно меняющейся функцией ($l(x) \in \mathbb{R}_0$).

В теории экстремальных значений параметр $\alpha = 1/\gamma > 0$ называется хвостовым индексом. Он показывает тяжесть правого хвоста распределения.

Определение 4. Будем называть функцию $\tilde{f}(x)$ комбинированной плотностью распределения $F(x)$ с тяжёлым правым хвостом, если

$$\tilde{f}(x) = \begin{cases} f(x), & x \in [0, \tau] \\ g(x), & x \in (\tau, \infty) \end{cases},$$

где f – функция плотности, отвечающая за тело распределения, g – плотность хвоста распределения, оцененная с помощью стандартных методов теории экстремумов, τ – точка присоединения тела к хвосту, оцененная или выбранная оптимальным путем.



Граф. 1: График плотности распределения Парето с параметром формы равным 4.

Основная наша задача состоит в комбинированном оценивании распределений с правильно меняющимся правым хвостом. Полученное в результате оценивания распределение должно обладать непрерывной плотностью \tilde{f} и хорошо приближать истинное распределение.

2.2 Предыдущие результаты

Предполагая, что $\mathbf{X}_n = (X_1, \dots, X_n)$ – положительные н.о.р. случайные величины с функцией распределения $F(x)$ и плотностью $f(x)$, Маркович и Кригер в статье [6] предложили следующую оценку для распределений с правильно меняющимся правым хвостом

$$\tilde{f} = \begin{cases} f^N(t), & t \in [0, X_{(k)}] \\ f_\gamma(t), & t \in (X_{(k)}, \infty) \end{cases},$$

где $f^N(t)$ – некоторая непараметрическая оценка $f(x)$ на отрезке $[0, X_{(k)}]$, отвечающая за тело распределения. В терминах тригонометрических функций $f^N(t)$ представляется в виде

$$f^N(t) = \frac{1}{X_{(k)}} \sum_{j=1}^N \lambda_j \phi_j\left(\frac{t}{X_{(k)}}\right),$$

а для оценивания хвоста рассматривается параметрическая модель Парето

$$f_\gamma(t) = \frac{1}{\gamma} t^{-1/\gamma-1} + \frac{2}{\gamma} t^{-2/\gamma-1}.$$

Граница между телом и хвостом в работе [6] определяется конкретной хвостовой квантилью $X_{(k)}$. Параметр k , ключевой для статистики экстремумов, определяется с помощью методов бутстрепа. Оценив хвостовое распределение с помощью классических методов теории экстремумов, авторы статьи, применяя метод минимизации структурного риска (S.R.M.M.), находят оптимальную оценку тела распределения. Однако авторы не отвечают на вопрос, почему в качестве оптимальной граничной точки тела выбирается пороговое значение, используемое для построения оценки Хилла. Кроме

того, при оценивании распределения хвоста не был учтен параметр масштаба.

Возможно свободное комбинирование непараметрических и параметрических оценок для построения \tilde{f} . Точность f_γ зависит от выбора соответствующей параметрической модели хвоста и точной оценки γ . На данный момент наиболее популярным и точным методом оценивания хвоста является модель GPD (Generalized Pareto distribution). Для оценки f^N , в частности, могут быть использованы методы ядерных и проекционных оценок.

В работе [7] предлагается ядерная оценка плотности, основанная на адаптивном определении параметра сглаживания по двум областям в соответствии с плотностью наблюдений и назначении адаптивных параметров полосы пропускания для каждого региона. Оценка имеет следующий вид

$$\hat{f}_h(x) = \frac{1}{n} \sum_{i=1}^n K_{x,h}(X_i),$$

где $K_{x,h}$ – непрерывная ядерная функция, $h = h(n) > 0$ – произвольная последовательность параметров, определяющих гладкость $\hat{f}_h(x)$.

Наблюдения разделяются на две области: область низкой плотности (LDR) и область высокой плотности (HDR), называемые **телом и хвостом** плотности соответственно. Такое разделение позволяет оценивать интересующую нас модель, используя следующий алгоритм поиска оптимальной точки склейки:

- 1) Стартуем с данных X_1, \dots, X_n , упорядочивая их как $X_{(1)} \leq \dots \leq X_{(n)}$.
- 2) Вычисляем расстояния d_i для каждой пары наблюдений (X_i, X_{i+1}) , где $d_i(X_i, X_{i+1}) = |X_i - X_{i+1}|$.
- 3) Определяем константу $\epsilon = S$, где S – корень из выборочной оценки дисперсии, построенной по наблюдениям d_1, \dots, d_{n-1} .
- 4) Группируем:

Область с высокой плотностью определяем как $S_{(HDR)} = \{x_i | d_i \leq \epsilon\}$.

Как только $d_i > \epsilon$, мы прекращаем проверку, оставшиеся наблюдения рассматриваем как область с низкой плотностью $S_{(LDR)}$.

Параметры адаптивной пропускной способности $h^{(0)}$ и $h^{(1)}$, отвечающие за области $S_{(HDR)}$ и $S_{(LDR)}$ соответственно, выбираются с помощью метода Markov chain Monte Carlo (MCMC), подробно рассмотренного, например, в [8].

Научных обоснований, почему в качестве параметра ϵ , обеспечивающего разбиение наблюдений на тело и хвост, выбирается стандартное отклонение, в данной работе не представлено.

Еще один метод комбинированного оценивания интересующей нас модели был предложен в работе [9]. В ней рассматривается новый метод оценки параметров GPD, позволяющий одновременно определить параметры формы, масштаба и сдвига путем объединения функций правдоподобия распределения тела и хвоста с использованием информационного критерия Акаике (AIC). Параметры распределений тела и хвоста оцениваются методом максимального правдоподобия. Параметр сдвига

делит распределение на тело и хвост. Более ранние методы могли оценить только параметры формы и масштаба, сдвиг же считался известным или находился опытным путем, то есть выбирался авторами, а не оценивался.

Для оценки плотности тела используется AIC. Он определяется следующим образом:

$$\varphi = -2(L_c(p_b^*) - n), \quad (4)$$

где $L_c(p_b^*)$ - логарифмическая функция максимального правдоподобия возможных распределений тела с оптимально подобранными параметрами p_b^* , а n - число параметров. Рассматриваться могут все распределения, имеющие непрерывную плотность. В статье исследование упрощено до 6 вариантов распределений: нормального, логнормального, гамма, вейбулловского, экспоненциального и максимально-устойчивых распределений (2). Если для разных распределений значения максимального правдоподобия одинаковы, то предпочтительнее распределение, имеющее наименьшее количество параметров, так как иначе при большом количестве параметров у распределений наблюдаются завышенные оценки.

Плотность, объединяющая тело и хвост, имеет следующий вид

$$f(y|p_b, p_t) = \begin{cases} f_b(y|p_b), & \text{если } y \leq \mu \\ \alpha f_t(y|p_t), & \text{иначе} \end{cases},$$

где f_b - плотность тела (рассчитывается по всем наблюдениям), f_t - плотность хвоста, а p_b и p_t - параметры f_b и f_t соответственно. Параметр α используется в качестве масштабного коэффициента и рассчитывается по формуле

$$0 \leq \alpha = 1 - F_b(\mu) \leq 1.$$

Параметры GPD находятся из решения оптимизационной задачи следующего вида

$$\max_{\xi, \sigma, \mu} = L(\xi, \sigma, \mu | y, p_b^*) \quad (5)$$

$$\sigma = \frac{1}{f_b(\mu | p_b^*)}, \quad (6)$$

где ξ, σ, μ - параметры формы, масштаба и сдвига распределения хвоста соответственно. Параметры распределения тела p_b^* оптимизируются предварительно с помощью (4). Ограничение (6) представляет собой непрерывность между PDF тела и хвоста распределения. Отдельных указаний на то, почему именно так выбирается параметр масштаба, в рассматриваемой статье не представлено.

Метод моделирования тела и хвоста сводится к следующему:

- 1) с помощью AIC в (4) выбирается наиболее подходящее распределение тела и его параметры;
- 2) на основе оцененного распределения тела находятся наилучшие параметры распределения хвоста с помощью оптимизации задачи (5).

Вышеизложенный метод позволяет оценить параметр сдвига в модели GPD. Он определяет границу между телом и хвостом. Сам параметр зависит только от заданных выборок. Если выборки не включают экстремальные события, то коэффициент α берется близким к нулю.

3 Комбинированное оценивание распределений

В данном разделе описан алгоритм оценки распределения с тяжелым правым хвостом, основанный на минимизации вероятностной метрики путем выбора различных точек склейки.

3.1 Разделение выборки

Предположим, что на вход мы получаем данные X_1, \dots, X_n , принадлежащие неизвестному классу распределений с тяжелым правым хвостом.

Пусть $X_{(1)} \leq \dots \leq X_{(n)}$ - вариационный ряд этой выборки.

Пусть τ - наблюдение, которое будет рассматриваться в качестве точки склейки оцененных функций плотности тела и хвоста неизвестного распределения. Далее из вариационного ряда извлекаются два набора наблюдений:

1) $\mathbf{X}_b = (X_{(1)} \leq \dots \leq X_{(d)}; d = \sup_{k \in \{1, \dots, n\}} \{k : X_{(k)} \leq \tau + p\})$ - отвечающий за оцениваемое тело распределения;

2) $\mathbf{X}_t = (X_{(s)} \leq \dots \leq X_{(n)}; s = \inf_{k \in \{1, \dots, n\}} \{k : X_{(k)} \geq \tau - p\})$ - отвечающий за оцениваемый хвост распределения.

Параметр p , отвечающий за расширение окна, в пределах которого необходимо брать данные, должен обеспечивать более точную аппроксимацию неизвестного распределения и непрерывность в точке склейки, но выбирается исследователем самостоятельно.

3.2 Оценка тела распределения

Для оценки тела распределения по имеющимся наблюдениям \mathbf{X}_b мы будем пользоваться ядерной оценкой плотности.

Напомним, что ядерная оценка плотности определяется как

$$\hat{f}_h(x) = \frac{1}{d} \sum_{i=1}^d K_h(x - x_i) = \frac{1}{dh} \sum_{i=1}^d K\left(\frac{x - x_i}{h}\right),$$

где d - объем данных, по которым оценивается тело распределения, K является ядром, а $h > 0$ сглаживающим параметром.

В качестве функции K можно использовать, например, Гауссово ядро:

$$K(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2}$$

Стоит отметить важность выбора полосы пропускания при ядерной оценке плотности. Ее оптимальный выбор позволяет повысить точность оценки. Проблема выбора наилучшей полосы пропускания изучалась многими авторами, но ни один из методов нельзя считать универсально лучшим. Стандартными эмпирическими методами считаются правило Скотта и правило Сильвермана (подробнее в работах [10], [11]). К классическим методам относят кросс-валидацию и АИС. Известен также метод plug-in, его описание приведено в статье [12]. Преимущества и недостатки классических

и plug-in методов обсуждались Лоудером в работе [13]. Некоторые авторы, например Холл в статье [14] и Хейзелтон в [15], применяли bootstrap для поиска полосы пропускания. Преимущества и недостатки различных методов были разобраны в работе [16].

3.3 Оценка хвоста распределения

Оценка хвостового распределения производится стандартными методами теории экстремальных значений. Для этого используется обобщенное распределение Парето, определяемое как

$$G_{\sigma,\gamma}(y) = \begin{cases} 1 - (1 + \frac{\gamma y}{\sigma})^{-1/\gamma}, & \gamma \neq 0, \sigma > 0 \\ 1 - \exp(-\frac{y}{\sigma}), & \gamma = 0, \sigma > 0 \end{cases},$$

где $0 < y < \infty$ ($\gamma \geq 0$) или $0 < y < -\sigma/\gamma$ ($\gamma < 0$).

Графики плотности двухпараметрического обобщенного распределения Парето (Generalized Pareto distribution) изображены на Граф. 2. Благодаря своей гибкости GPD может принимать множество различных форм и, следовательно, может использоваться в различных практических ситуациях. В реальности наблюдения, полученные из GPD, могут принимать очень большие значения при положительных значениях параметра γ (ситуации с тяжелым хвостом), тогда как очень легкие хвосты могут быть получены при отрицательных значениях γ .

Пикандс в работе [17], независимо с работой Балкема и де Хаана [18], показал, что GPD хорошо аппроксимирует условную функцию распределения превышения порога $F_u(y) = \frac{F(u+y)-F(u)}{1-F(u)}$, $u < x_0$, $0 < y < x_0 - u$, $x^* = \sup \{x : F(x) < 1\}$, тогда и только тогда, когда F принадлежит области притяжения одного из экстремальных типов распределений. То есть

$$\lim_{u \rightarrow x_0} \sup_{0 < y < x_0 - u} |F_u(y) - G(y; \sigma(u), \gamma)| = 0, \quad (7)$$

тогда и только тогда, когда $F \in \mathcal{D}(G_\gamma)$ для какого-то фиксированного значения γ .

Для того, чтобы оценить хвост распределения по данным \mathbf{X}_t , необходимо оценить параметры обобщенного распределения Парето.

В качестве оценки параметра формы γ используется оценка Хилла:

$$\hat{\gamma}_{n,k}^H := \frac{1}{k} \sum_{i=s}^{k-1} \ln X_{(n-i)} - \ln X_{(n-k)}$$

На практике точность оценки зависит от выбора k и эта проблема довольно затруднительна (см., например, [19, 20, 21]). На протяжении 25 лет было предложено множество работ, посвященных выбору числа k , а также некоторым свойствам и модификациям оценки Хилла [22, 23]. Чтобы оптимально выбрать порог k , мы используем Eye-Ball метод для оценки Хилла [24].

Существует также множество других методов оценки индекса экстремального значения, с кратким обзором которых можно познакомиться в работе [25].

Для оценки параметра σ в модели обобщенного распределения Парето мы используем **оценку масштаба**, описанную в разделе 4.2 монографии [26].

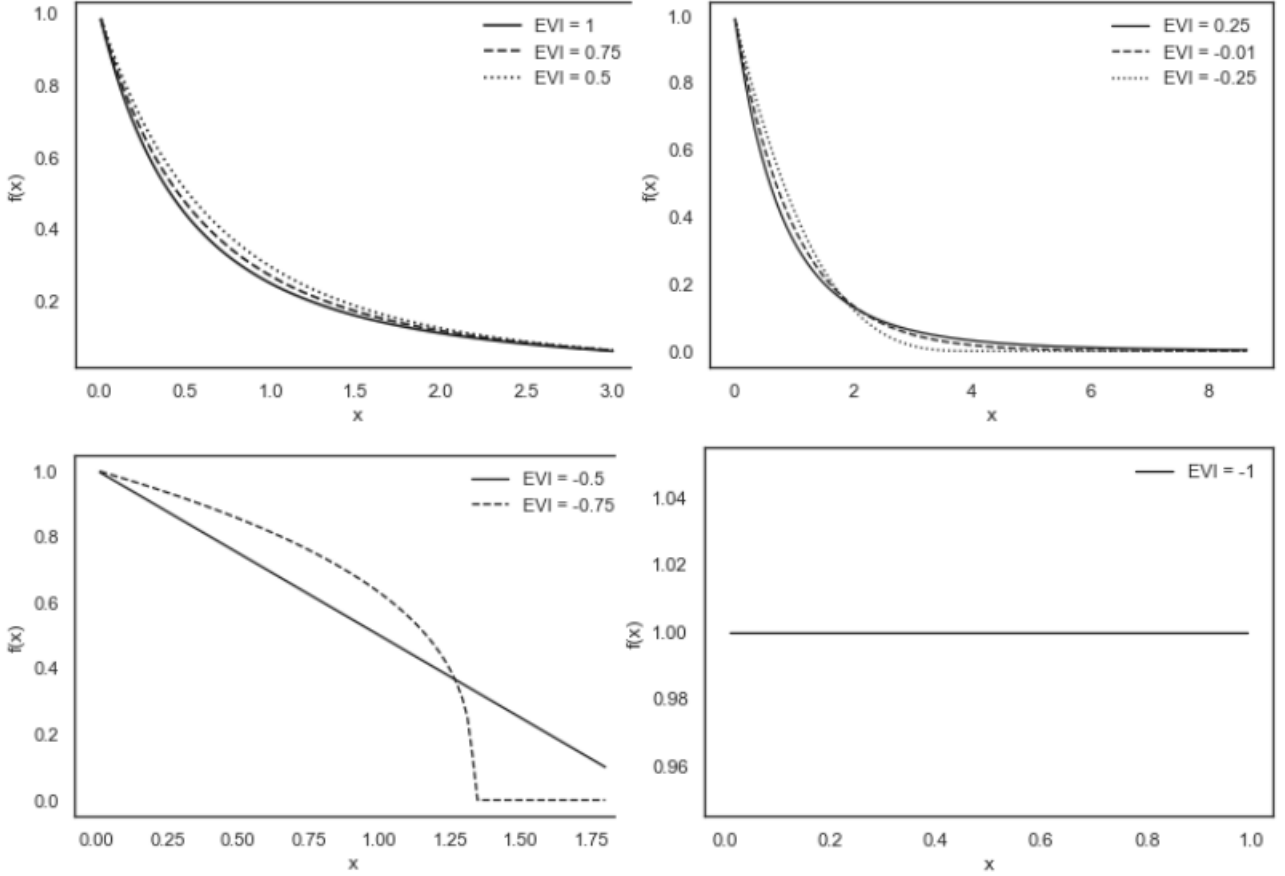
Используя соотношение (7), функция $F(x)$ переписывается как:

$$F(x) = P(X \leq x) = (1 - F(u))F_u(x - u) + F(u) = (1 - F(u))G_{\sigma,\gamma}(x - u) + F(u).$$

Оцененное распределение имеет вид:

$$\hat{F}(x) = (1 - F_n(u))G_{\hat{\sigma},\hat{\gamma}}(x - u) + F_n(u), \quad (8)$$

где $F_n(x)$ - эмпирическая функция распределения, а $\hat{\sigma}, \hat{\gamma}$ - оцененные параметры модели GPD.



Граф. 2: Графики плотности обобщенного распределения Парето с параметром формы $\sigma = 1$.

3.4 Расстояние между распределениями

В критериях типа ω^2 расстояние между гипотетическим и истинным распределениями рассматривают в квадратичной метрике:

$$\Omega_n^2 = n \int_{-\infty}^{\infty} [F_n(x) - F(x)]^2 \Psi(F(x)) dF(x), \quad (9)$$

где $F_n(x)$ - эмпирическая функция распределения, $\Psi(F(x))$ - некоторая неотрицательная весовая функция. При $\Psi(u) = 1$ получаем статистику критерия Крамера-Мизеса-Смирнова, а при $\Psi(u) = \frac{1}{u(1-u)}$ - статистику критерия Андерсона-Дарлинга.

Статистика критерия согласия Андерсона-Дарлинга в нашей задаче является более полезной мерой расстояния между распределениями по сравнению, например, со статистикой Колмогорова-Смиронова, **не уделяющей одинакового внимания подгонке распределения как в хвостах, так и в основной части распределения.**

Однако, нам бы хотелось придавать наблюдениям, расположенным в хвосте, еще большие веса при оценке точки склейки. В связи с этим введем понятия важности тела и хвоста.

Будем говорить, что важность i - числовая характеристика, определяемая нами самостоятельно, в зависимости от **решаемой** задачи. Например, при оценивании распределения величины убытков страховой компании, будем опираться на закон Парето. Этот закон говорит о том, что на 90% всех страховых случаев приходится только около 10% всех выплат. Исходя из этого закона можно сделать вывод, что при оценке распределения данные, расположенные в хвосте, в 9 раз важнее, чем данные, расположенные в теле.

Определим функцию $\Psi(u)$ из условий:

$$\int_0^{\tau^*} \Psi(u) du = i_b, \quad \int_{\tau^*}^1 \Psi(u) du = i_t, \quad (10)$$

где i_b, i_t - важности тела и хвоста распределения, соответственно, τ - точка склейки, $\tau^* = F(\tau)$.

Для начала предположим, что функция $\Psi(u)$ является постоянной на каждом из отрезков интегрирования (10). Тогда

$$\Psi(u) = \frac{i_b}{\tau^*} = \psi_b, \quad u \in [0, \tau^*],$$

$$\Psi(u) = \frac{i_t}{1 - \tau^*} = \psi_t, \quad u \in [\tau^*, 1],$$

и метрика (9) перепишется в следующем виде:

$$\begin{aligned} \Omega_n^2 &= n \int_{-\infty}^{\infty} [F_n(x) - F(x)]^2 \Psi(F(x)) dF(x) = n \int_{-\infty}^{X_{(1)}} F^2(x) \Psi(F(x)) dF(x) + \\ &+ n \sum_{i=1}^{n-1} \int_{X_{(i-1)}}^{X_{(i)}} \left[\frac{i}{n} - F(x) \right]^2 \Psi(F(x)) dF(x) + \int_{X_{(n)}}^{\infty} [1 - F(x)]^2 \Psi(F(x)) dF(x) = \\ &= ||u = F(x) \in [0, 1]|| = \\ &= n \int_0^{F(X_{(1)})} u^2 \Psi(u) du + n \sum_{i=1}^{n-1} \int_{F(X_{(i)})}^{F(X_{(i+1)})} \left[\frac{i}{n} - u \right]^2 \Psi(u) du + n \int_{F(X_{(n)})}^1 [1 - u]^2 \Psi(u) du. \end{aligned}$$

Пусть $\tau^* \in [F(X_{(j)}), F(X_{(j+1)})]$. Тогда

$$\int_{F(X_{(j-1)})}^{F(X_{(j)})} \left[\frac{j}{n} - u \right]^2 \Psi(u) du = \psi_b \int_{F(X_{(j)})}^{\tau^*} \left[\frac{j}{n} - u \right]^2 du + \psi_t \int_{\tau^*}^{F(X_{(j)})} \left[\frac{j}{n} - u \right]^2 du.$$

$$\begin{aligned}\Omega_n^2 = & n\psi_b \frac{[F(X_{(1)})]^3}{3} + n\psi_b \sum_{i=1}^{j-1} \frac{[i - nF(X_{(i)})]^3 - [i - nF(X_{(i+1)})]^3}{3n^3} + n\psi_b \frac{[j - nF(X_{(j)})]^3 - [j - n\tau^*]^3}{3n^3} + \\ & + n\psi_t \frac{[j - n\tau^*]^3 - [j - nF(X_{(j+1)})]^3}{3n^3} + n\psi_t \sum_{i=j+1}^{n-1} \frac{[i - nF(X_{(i)})]^3 - [i - nF(X_{(i+1)})]^3}{3n^3} - \\ & - n\psi_t \left(\frac{F^3(X_{(n)})}{3} - F^2(X_{(n)}) + F(X_{(n)}) - \frac{1}{3} \right).\end{aligned}$$

Для случаев $\tau^* \in [0, F(X_{(1)})]$ и $\tau^* \in [F(X_{(n)}), 1]$ формулы выводятся аналогичным способом.

Будем рассматривать комбинированную функцию распределения имеющую следующий вид:

$$F(x) = \begin{cases} F_b(x), & x \in [0, \tau - p) \\ F_c(x), & x \in [\tau - p, \tau + p) \\ F_t(x), & x \in [\tau + p, \infty). \end{cases},$$

где $F_b(x)$, $F_t(x)$ - оцененные функции распределения тела и хвоста,

$$F_c(x) = \frac{\tau + p - x}{2p} F_b(x) + \frac{x - \tau + p}{2p} F_t(x).$$

3.5 Оптимизационная задача

Перейдем к рассмотрению оптимизационной задачи:

$$\tau_{opt} = \arg \left\{ \min_{\tau \in [X_{(1)}, X_{(n)}]} \Omega_n^2 \right\}. \quad (11)$$

Решением задачи (11) является оптимальная (в смысле минимизации расстояния между теоретическим и оцененным распределениями) точка склейки хвоста и тела.

Для нахождения решения (11) возможно использование методов градиентного спуска, реализованных в функции `scipy.optimize.minimize` в Python, например, метода COBYLA (Constrained minimization by linear approximation) [27].

Список литературы

- [1] Beirlant, J., Kijko, A., Reynkens, T., Einmahl, J.H.J., Estimating the maximum possible earthquake magnitude using extreme value methodology: the Groningen case // Nat. Hazards 2018. № 169, P. 1–23.
- [2] Carreau, J., Naveau, P., Neppel, L., Partitioning into hazard subregions for regional peaks-over-threshold modeling of heavy precipitation // Water Resour. Res 2017. № 53, P. 4407–4426.
- [3] Chavez-Demoulin, V., Embrechts, P., Hofert, M., An extreme value approach for modeling operational risk losses depending on covariates // J. Risk. Insur. 2016. № 83, P. 735–776.

- [4] Hambuckers, J., Groll, A., Kneib, T., Understanding the economic determinants of the serverity of operational losses: a regularized generalized Pareto regression approach // Appl. Economet 2018. № 33, P. 898–935.
- [5] Gnedenko, B. V., Sur la distribution limite du terme maximum d’une serie aleatoire // Annals of Mathematics 1943. № 44, P. 423-453.
- [6] Markovich, N. M., Nonparametric Estimation of Heavy-Tailed Density by the Discrepancy Method // Nonparametric Statistics 2016. № 175, P. 1 - 14.
- [7] Ziane, Y., Zougab, N., Adjabi, S., Body tail adaptive kernel density estimation for nonnegative heavy-tailed data // Monte Carlo Methods Appl 2021. № 27, P. 57 - 69.
- [8] Andrieu, C., Jordan, M., Doucet, A., de Freitas, N., An Introduction to MCMC for Machine Learning // Machine Learning 2003. № 50(1), P. 5-43.
- [9] Lim, W., Lee, T.H., Kang, S., Cho, S., Estimation of body and tail distribution under extreme events for reliability analysis // Structural and Multidisciplinary Optimization 2016. № 54, P. 1631 - 1639.
- [10] Scott, D.W., Multivariate Density Estimation: Theory, Practice, and Visualization. John Wiley & Sons 1992.
- [11] Silverman, B.W., Density Estimation for Statistics and Data Analysis, Monographs on Statistics and Applied Probability 1986, Vol. 26.
- [12] Tenreiro, C., Asymptotic behaviour of multistage plug-in bandwidth selections for kernel distribution function estimators // Journal of Nonparametric Statistics 2006. № 18, P. 101– 116.
- [13] Loader, C. R., Bandwidth selection: Classical or plug-in? // The Annals of Statistics 1999. № 27(2), P. 415–438.
- [14] Hall, P., Using the bootstrap to estimate mean squared error and select smoothing parameter in nonparametric problems // Journal of Multivariate Analysis 1990. № 32, P. 177–203.
- [15] Hazelton, M., Bandwidth selection for local density estimators // Scandinavian Journal of Statistics 1996. № 23, P. 221–232.
- [16] Cañón-Tapia, E., Kernel Analyses of Volcanic Vent Distribution: How Accurate and Complete are the Objective Bandwidth Selectors? // Frontiers in Earth Science 2022. Vol. 10
- [17] Pickands, J. III, Statistical inference using extreme order statistics // Annals of Statistics 1975. № 3, P. 119-131.
- [18] Balkema, A., and de Haan, L., Residual life time at great age // Annals of Probability 1974. № 2, P. 792–804.
- [19] Danielsson, J., de Haan, L., Peng, L. and de Vries, C. G., Using a bootstrap method to choose the sample fraction in tail index estimation // J. Multivariate Anal. 2001. № 76, P. 226–248.

- [20] Dekkers, A. L. M. and de Haan, L., Optimal choice of sample fraction in extreme-value estimation // J. Multivariate Anal. 1993. № 47(2), P. 172–173.
- [21] Draisma, G., de Haan, L. and Peng, L., A Bootstrap-based method to achieve optimality in estimating the extreme-value index // Extremes 1999. № 2(4), P. 367–404.
- [22] Beirlant, J., Teugels, J. L. and Vynckier, P., Tail index estimation, Pareto quantile plots, and regression diagnostics // J. Amer. Statist. Assoc. 1996. № 91, P. 1659–1667.
- [23] Csörgö, S., Deheuvels, P. and Mason, D., Kernel estimates of the tail index of a distribution // Ann. Statist. 1985. № 13, P. 1050–1077.
- [24] Danielsson J., Ergun L. M., de Haan L., De Vries C., Tail Index Estimation: Quantile Driven Threshold Selection // SSRN Electronic Journal 2016.
- [25] Fedotenkov I., A Review of More than One Hundred Pareto-Tail Index Estimators // Statistica 2020. № 80, P. 245-299.
- [26] de Haan L., Ferreira A., Extreme Value Theory: An Introduction // Springer 2006.
- [27] Powell M.J.D., A Direct Search Optimization Method That Models the Objective and Constraint Functions by Linear Interpolation // Mathematics and Its Applications 1994. № 275, P. 51-67.