УДК 519.272

© 2002 г. **Н.М.Маркович, канд.тех.наук**

(Институт проблем управления им. В.А.Трапезникова РАН, Москва)

ОЦЕНИВАНИЕ КВАНТИЛЕЙ ВЫСОКИХ ПОРЯДКОВ ДЛЯ РАСПРЕДЕЛЕНИЙ С ТЯЖЕЛЫМИ ХВОСТАМИ

Статья посвящена оцениванию квантилей высоких порядков для распределений с тяжелыми хвостами. Предлагается оценка квантилей в предположении, что хвост распределения типа Парето. Доказывается, что распределение логарифма отношения предлагаемой оценки к истинной квантили асимптотически нормально. То же доказывается для оценки Вайсманна.

1. Введение

Пусть $X^n = \{x_1, x_2, \ldots, x_n\}$ - выборка положительных независимых случайных величин (сл.в.), одинаково распределенных с неизвестной функцией распределения (ФР) $F(x)$. Квантилью $(1-p)$-го порядка распределения F для заданной величины p называется решение $x = x_p$ уравнения

$$(1) \qquad 1 - F(x) = p.$$

Наша цель - оценить "высокую" квантиль (a high quantile), т.е. квантиль, соответствующую вероятности p, близкой к 0 в ситуации, когда $F(x)$ имеет тяжелый хвост. На практике довольно часто приходится оценивать риск больших, но, возможно, редких потерь. Анализ таких рисков играет важную роль в контроле работы Интернета и, в целом, при установке порогов изменения параметров технических систем. Многие наблюдаемые сл.в. в Интернете, как например длины пересылаемых файлов, временные промежутки между сессиями, продолжительности сессий, распределены с тяжелыми хвостами. Для управления сетями области определения наблюдаемых случайных параметров сужаются до конечных интервалов, верхние границы которых и должны быть определены с помощью высоких квантилей. Неверное принятие порогов, например игнорирование больших, но достаточно вероятных файлов, может привести к перегрузке сети. В таких ситуациях квантиль может располагаться на границе или за пределами разброса выборки.

Классический подход с использованием эмпирической ФР $F_n(x)$ в (1) не подходит для "высоких" квантилей также, как и взвешенные эмпирические квантильные оценки, [1] так как $F_n(x) = 1$ для $x \geq x_{(n)}$ ($x_{(n)}$ - максимальная порядковая статистика выборки).

Основная проблема, возникающая при оценивании таких квантилей, связана с отсутствием информации за пределами разброса выборки. Поэтому главная идея всех методов оценивания "высоких" квантилей состоит в выборе предварительной ("pilot") оценки квантили внутри разброса выборки (это может быть одна из порядковых статистик, близких к границе) и затем перемещение ее вправо. Для этой цели оцениваются параметры масштабирования по выборке. [2]

Очевидно, чтобы экстраполировать "промежуточную" квантиль к некоторой за пределами выборки, необходимо использовать какую-то модель хвоста распределения. Однако во многих приложениях такая модель недоступна. Поэтому приходится использовать асимптотические модели поведения хвостов, основанные на распределении наибольшей порядковой статистики. Из теории экстремальных величин [2] известно, что если $F(x)$ такова, что предельное распределение наибольшей порядковой статистики существует, то это предельное распределение может быть только следующим:

$$H_\gamma(x) = \begin{cases} \exp(-(1+\gamma x)^{-1/\gamma}), & \gamma \neq 0, \\ \exp(-e^{-x}), & \gamma = 0. \end{cases}$$

Если индекс экстремальной величины (extreme value index (EVI)) $\gamma > 0$, то распределение F тяжело-хвостовое. В [3] показано, что для величин X,

[1] Пусть $x_1 \leq x_2 \leq \ldots \leq x_n$ - порядковые статистики выборки. Обычно $x_{(k)}$, где $k = [np] + 1$ ($[\cdot]$ - наибольшее целое), используется в качестве оценки для x_p. Так как в общем $E[x_{(k)}] \neq x_p$ и $P\{x_{(k)} \leq x_{(p)}\} \neq 0,5$, то $x_{(k)}$ не является ни средним, ни медианой для x_p. Это ведет к оценкам, которые являются интерполяцией между порядковыми статистиками, например взвешенное среднее: $x_p = (1-g)x_{(j)} + gx_{(j+1)}$, где $j = [np]$ и $g = np - j$ [1].

[2] Тот же принцип может быть применен к оцениванию "малых" квантилей, когда p близко к 1.

превосходящих достаточно высокий уровень u, распределение $Y = X - u$ сходится к распределению Парето (the Generalized Pareto Distribution (GPD))

$$\Psi_\gamma(x) = \begin{cases} 1 - (1+\gamma x)^{-1/\gamma}, & x \geq 0, \\ 0, & x < 0. \end{cases}$$

Часто используют GPD или модели "типа Парето" (Pareto-type tail model) [4, 5]

$$(2) \qquad 1 - F(x) = cx^{-1/\gamma}\left(1 + dx^{-\beta} + o(x^{-\beta})\right),$$

где $\gamma, \beta > 0$, $c > 0$ и $-\infty < d < \infty$ для моделирования хвоста распределения F.

Различные оценки "высоких" квантилей (см., например, [6]-[8]) следуют из этих предположений относительно хвоста. Отметим наиболее используемые из них.

В POT-методе (the peaks-over-threshold (POT)) GPD используется как распределение выходов за пределы некоторого высокого порога u:

$$x_p^{POT} = u + \frac{\widehat{\sigma}}{\widehat{\gamma}}\left(\left(\frac{p}{1 - F_n(u)}\right)^{-\widehat{\gamma}} - 1\right),$$

где $\widehat{\sigma}$ и $\widehat{\gamma}$ - оценки параметров GPD [9].

Оценка [8] получена для модели хвоста "типа Парето":

$$(3) \qquad x_p^w = x_{(n-k)}\left(\frac{k+1}{(n+1)p}\right)^{\widehat{\gamma}}, k = 1, \ldots, n-1,$$

где $x_{(1)} \leq x_{(2)} \leq \ldots \leq x_{(n)}$ порядковые статистики, соответствующие выборке X^n.

В статье предлагается несколько иная оценка, чем x_p^w, основанная на идее комбинированной оценки для тяжело-хвостовых плотностей распределения вероятностей (ПР). Эта оценка ПР "склеивает" параметрическую оценку для хвоста распределения с непараметрической оценкой для остальной части ПР:

$$(4) \qquad f^*(t, \gamma, N) = \begin{cases} \frac{\widehat{f}(t,\gamma,N)}{\int_0^\infty \widehat{f}(t,\gamma,N)dt}, & t \in A, \\ 0, & t \notin A \end{cases}$$

для $A = \left\{t : \widehat{f}(t, \gamma, N) > 0\right\}$, где

$$(5) \qquad \widehat{f}(t, \gamma, N) = \begin{cases} f^N(t), & t \in [0, x_{(n-k)}], \\ f_\gamma(t), & t \in (x_{(n-k)}, \infty). \end{cases}$$

Здесь и далее без ограничения общности рассматриваются только положительные сл.в.

Используется модель хвоста "типа Парето":

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

для $x > x_{(n-k)}$ и разложение по некоторым базисным функциям $f^N(t)$ для точек $x \leq x_{(n-k)}$ такое, что $\int_0^{x_{(n-k)}} f^N(t)dt = 1$. Далее используется оценка для ФР $F(x)$, соответствующая $f^\star(t, \gamma, N)$. Заметим, что для $x_p > x_{(n-k)}$

$$p = P\{x > x_p\} = P\{x > x_p | x > x_{(n-k)}\} P\{x > x_{(n-k)}\}.$$

$P\{x > x_{(n-k)}\}$ может быть заменена на частоту превышения предела $x_{(n-k)}$: k/n. Из (4) и (5) следует оценка:

$$\widetilde{F}(x) = \frac{1}{c(\gamma)} \left(\int_0^{\min(x, x_{(n-k)})} f^N(t) \mathbf{1}(f^N(t) > 0)dt + \int_{\min(x, x_{(n-k)})}^{x} f_\gamma(t)dt \right)$$

для ФР $F(x)$, где

$$c(\gamma) = \int_0^\infty \widetilde{f}(t, \gamma, N) \mathbf{1}(\widetilde{f}(t, \gamma, N) > 0)dt$$

может быть аппроксимировано величиной $\int_0^\infty \widetilde{f}(t, \gamma, N)dt$, т.е.

(7) $$c(\gamma) \approx 1 + x_{(n-k)}^{-1/\gamma} + x_{(n-k)}^{-2/\gamma}.$$

Так как

$$\int_0^{x_{(n-k)}} f^N(t) \mathbf{1}(f^N(t) > 0)dt = c(\gamma) - \int_{x_{(n-k)}}^\infty f_\gamma(t)dt,$$

то из (6) имеем для $x > x_{(n-k)}$

$$\widetilde{F}(x) = \frac{1}{c(\gamma)} \left(\int_0^{x_{(n-k)}} f^N(t) \mathbf{1}(f^N(t) > 0)dt + \int_{x_{(n-k)}}^{x} f_\gamma(t)dt \right)$$
$$= 1 - \frac{1}{c(\gamma)} (x^{-1/\gamma} + x^{-2/\gamma}).$$

Обозначим

$$1 - F_\gamma(x) = x^{-1/\gamma} + x^{-2/\gamma}.$$

Тогда

$$P\{x > x_p | x > x_{(n-k)}\} = \frac{1 - F(x_p)}{1 - F(x_{(n-k)})}$$

$$\simeq \frac{1-\widetilde{F}(x_p)}{1-F_\gamma(x_{(n-k)})} < \frac{1}{c(\gamma)} \left(\frac{x_{(n-k)}}{x_p}\right)^{1/\gamma} \left(1+\left(\frac{x_{(n-k)}}{x_p}\right)^{1/\gamma}\right).$$

Отсюда, ввиду приближенного равенства

$$p \simeq \frac{k}{nc(\gamma)} \left(\left(\frac{x_{(n-k)}}{x_p}\right)^{1/\gamma} + \left(\frac{x_{(n-k)}}{x_p}\right)^{2/\gamma}\right),$$

статистика

$$x_p^c = x_{(n-k)} \left(-0,5+\sqrt{0,25+\frac{pnc(\widehat{\gamma})}{k}}\right)^{-\widehat{\gamma}} \qquad (8)$$

может быть использована в качестве оценки для x_p.

Оценка x_p^c отличается от оценки x_p^w нормализующим множителем $c(\widehat{\gamma})$, отражающим тот факт, что оценка ФР $\widetilde{F}(x)$ включает не только параметрическую оценку хвостовой области (как в x_p^w), но и оценку основной части распределения. Так как предполагается, что $\int_0^{x_{(n-k)}} f^N(t)dt = 1$, это позволяет использовать любую оценку (параметрическую или непараметрическую) для основной части ПР с таким свойством.

Общим недостатком оценок квантилей является их чувствительность к выбору порога (это, например, u в x_p^{POT}) или, что часто эквивалентно, к выбору k. Определение k также необходимо для оценки γ (extreme value index (EVI)). Известны различные методы оценивания γ, например оценка Хилла [11], момент-оценки [7], оценка отношения (the ratio estimate) [12], UH-оценка [6]. Когда k возрастает (т.е. $x_{(n-k)}$ становится малым), дисперсия оценки EVI уменьшается, но возрастает ее смещение. Однако, если k уменьшается (т.е. используется меньше данных), смещение стремится к 0, а дисперсия растет.

Теоретически оптимальное k должно минимизировать среднеквадратичную ошибку (MSE) $E\left(\widehat{x}_p(k)-x_p\right)^2$. Получить точное выражение для MSE трудно, поэтому обычно предлагается минимизировать по k асимптотическую MSE, т.е. $as.E\left(\widehat{x}_p(k)-x_p\right)^2$ (точнее, ее бутстреп (bootstrap)-оценку [13]) или $as.E\left(\log\frac{\widehat{x}_p(k)}{x_p}\right)^2$ [6]. В [5] было предложено минимизировать MSE оценки вероятности превышения некоторого высокого уровня: $E\left(\overline{F}_\theta(\overline{F}^{-1}(p))-p\right)^2$ ($\overline{F}_\theta$ - некоторая оценка хвоста).

2. Результаты моделирования

Новая оценка x_p^c сравнивалась с x_p^{POT} и x_p^w посредством моделирования. Для сравнения исследовались те же распределения, что и в [9]. А именно, распределение Парето с параметром $1/\gamma = \{1; 2\}$ и распределение Стьюдента с параметром $1/\gamma = \{1; 2\}$. Следуя [9] в качестве характеристик квантилей использовались эмпирические оценки смещения (the bias) и среднеквадратического отклонения (the mean squared error) оценок квантилей в процентах к истинной величине квантили x_p:

$$\%Bias = \frac{\frac{1}{NR}\sum_{i=1}^{NR}\widehat{x}_{p_i} - x_p}{x_p},$$

$$\%RMSE = \frac{\sqrt{\frac{1}{NR}\sum_{i=1}^{NR}(\widehat{x}_{p_i} - x_p)^2}}{x_p},$$

где $NR = 25$ число моделируемых выборок. В обеих оценках x_p^c и x_p^w γ оценивалась оценкой Хилла (10), а k находилось из минимума бутстреп-оценки среднеквадратичной ошибки $E\left(\widehat{x}_p(k) - x_p\right)^2$:

$$\widehat{MSE}(n_1, k_1) = E\{\left(\widehat{x}_p^\star(n_1, k_1) - \widehat{x}_p(n, k)\right)^2 |X^n\}$$

по k_1, где $\widehat{x}_p^\star(n_1, k_1)$ - оценка квантили, построенная по подвыборке (re-sample) объема n_1 с параметром k_1. Такая подвыборка формируется случайным выбором с возвращением из исходной выборки объема n. Для процедуры бутстреп использовалось 25 подвыборок. Величины вспомогательных параметров α и β для вычисления размера подвыборок $n_1 = n^\beta$ и соотношения между k и k_1: $k = k_1\left(\frac{n}{n_1}\right)^\alpha$ были взяты равными $\alpha = 2/3$, $\beta = 1/2$, как и в [10].

Моделирование показывает (см. Табл. 1), что предложенная квантильная оценка x_p^c лучше, чем x_p^{POT} и x_p^w, особенно для более высоких квантилей, и демонстрирует меньшие средне-квадратичные отклонения.

3. Распределения для оценок квантилей высоких порядков

Найдем распределения и доверительные интервалы для x_p^c и x_p^w (см. (3), (8)).

Очевидно, что

$$\frac{x_p^c}{x_p} = \frac{x_{(n-k)}}{x_p\left(-0,5 + \sqrt{0,25 + \frac{pnc(\widehat{\gamma})}{k}}\right)^{\widehat{\gamma}}}$$

и

$$\log(\frac{x_p^c}{x_p}) = \log x_{(n-k)} - \log x_p - \widehat{\gamma}\log\left(-0,5 + \sqrt{0,25 + \frac{pnc(\widehat{\gamma})}{k}}\right)$$

$$\sim \log x_{(n-k)} - \log x_p - \widehat{\gamma} \log \left(\frac{c(\widehat{\gamma})}{a_n} \right),$$

где $a_n = k/(pn)$ ($f(x) \sim g(x)$ для $x \to a$ означает, что $\lim_{x\to a} f(x)/g(x) = 1$). Заметим, что

$$\log(\frac{x_p^c}{x_p}) \sim \log(\frac{x_p^w}{x_p}) - \widehat{\gamma} \log c(\widehat{\gamma}). \tag{9}$$

Найдем распределение для $\log(\frac{x_p^c}{x_p})$ и $\log(\frac{x_p^w}{x_p})$. Используем оценку Хилла для γ [11]:

$$\widehat{\gamma} = \frac{1}{k} \sum_{i=1}^{k} \log x_{(n-i+1)} - \log x_{(n-k)}. \tag{10}$$

Т е о р е м а 1. *Пусть истинное распределение хвоста - "типа Парето" (2) и* $k, n \to \infty, \frac{k}{n} \to 0$. *Тогда* $\log(\frac{x_p^c}{x_p})$ *и* $\log(\frac{x_p^w}{x_p})$ *имеют следующие асимптотические распределения:*

$$P\{\log \left(\frac{x_p^w}{x_p} \right) > x\} < 1 - F_w(x, p, k)$$

и

$$P\{\log(\frac{x_p^c}{x_p}) > x\} < 1 - F_w \left(x + \gamma^\star \left(\frac{A}{e} \right)^{1/\gamma^*}, p, k \right) + R_c, \tag{11}$$

где

$$\gamma^\star = \frac{\gamma}{2} + 2B(k), \tag{12}$$

$$B(k) \sim -\gamma d c^{-\gamma\beta} \left(\frac{k+1}{n} \right)^{\gamma\beta} \frac{\gamma\beta}{\gamma\beta + 1},$$

$$A = \begin{cases} (c^\gamma(1 + \gamma d c^{-\gamma\beta}))^{-1}, & \text{если } c^\gamma(1 + \gamma d c^{-\gamma\beta}) \neq 0, \\ 0 & \text{в противном случае,} \end{cases} \tag{13}$$

$$1 - F_w(x, p, k) = 1 + \mathrm{I}_{\exp(\frac{1}{\gamma})\frac{k}{n}}(k+1, n-k) - \Phi \left(\frac{\sqrt{k}(x - \varepsilon)}{\gamma \log(\frac{k}{np})} \right)$$

$$- \frac{1}{3\sqrt{2\pi k}} \left(1 - \frac{k(x-\varepsilon)^2}{\gamma^2 (\log(\frac{k}{np}))^2} \right) \exp(-\frac{k(x-\varepsilon)^2}{2\gamma^2(\log(\frac{k}{np}))^2})$$

$$+ (\gamma\beta)^2 \left(\frac{1}{k+1} - \frac{1}{n} \right) \left(\frac{c^{\gamma\beta}}{\gamma d} \left(\frac{n}{k+1} \right)^{\gamma\beta} - 1 \right)^{-2} + O\left(\left(\frac{k}{n} \right)^{2\gamma\beta} \right)$$

для любого $\varepsilon > 0$,

$$R_c = \Phi(-\frac{\sqrt{k}}{2}) + \frac{1}{3\sqrt{2\pi k}}\left(1 - \frac{k}{4}\right)\exp(-\frac{k}{8}) + \mathrm{I}_{\exp(-1/\gamma)}(k+1, n-k)$$

$$+ \sum_{i=1}^{k} \mathrm{I}_{\frac{i}{n}\exp(1/\gamma\beta)}(i, n-i+1) + 1 - \mathrm{I}_{\frac{k+1}{n}\exp(\frac{1}{\gamma\beta})}(k+1, n-k) + o\left(\frac{1}{\sqrt{n}}\right),$$

где

$$\mathrm{I}_x(a,b) = \frac{1}{B(a,b)}\int_0^x t^{a-1}(1-t)^{b-1}dt$$

- неполная бета-функция, $B(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$.

Доказательство теоремы приведено в Приложении.

<u>З а м е ч а н и е</u>. *Из теоремы 1 следует, что* $\log(\frac{x_p^w}{x_p})$ *и* $\log(\frac{x_p^c}{x_p})$ *распределены асимптотически нормально:* $\mathrm{N}(\varepsilon, \frac{\gamma\log\frac{k}{np}}{\sqrt{k}})$ *и* $\mathrm{N}(\varepsilon - \gamma^*\left(\frac{A}{e}\right)^{1/\gamma^*}, \frac{\gamma\log\frac{k}{np}}{\sqrt{k}})$ *соответственно, если* $k, n \to \infty, \frac{k}{n} \to 0$.

4. Доверительные интервалы.

Запишем доверительный интервал для x_p^w:

$$P\{|\log\left(\frac{x_p^w}{x_p}\right)| < x\} = P\{x_p^w e^{-x} < x_p < x_p^w e^x\}$$

$$\sim F_w(x,p,k) - F_w(-x,p,k) = \psi_w(x,p,k) = 1 - \eta.$$

Пусть

(14) $$x^\star = \psi_w^{-1}(1-\eta, p, k).$$

Тогда с вероятностью $1-\eta$

(15) $$x_p^w e^{-x^\star} < x_p < x_p^w e^{x^\star}.$$

Так как $1 - \Phi(x) = \Phi(-x)$, то для $\varepsilon > 0$ имеем

$$\psi_w(x,p,k) \simeq 2\Phi\left(\frac{\sqrt{k}(x-\varepsilon)}{\gamma\log(\frac{k}{np})}\right) - 1$$

и для $\varepsilon \approx 0$

(16) $$x^\star \approx \frac{\gamma\log(\frac{k}{np})}{\sqrt{k}}\Phi^{-1}(1-\frac{\eta}{2}).$$

Из (11) доверительный интервал для x_p^c определяется как

$$P\{|\log(\frac{x_p^c}{x_p})| < x - \gamma^\star \left(\frac{A}{e}\right)^{1/\gamma^\star}\} \sim F_w(x,p,k) - F_w(-x,p,k) = 1 - \eta.$$

Из рассуждений выше следует теорема.

Т е о р е м а 2. *Пусть условия теоремы 1 выполняются, $x^\star$ удовлетворяет (14), $\gamma^\star$ и A определяются из (12) и (13). Тогда с вероятностью $1-\eta$ выполнено*

$$x_p^w e^{-x^\star} < x_p < x_p^w e^{x^\star},$$

$$x_p^c \exp(-x^\star + \gamma^\star \left(\frac{A}{e}\right)^{1/\gamma^\star}) < x_p < x_p^c \exp(x^\star - \gamma^\star \left(\frac{A}{e}\right)^{1/\gamma^\star}). \quad (17)$$

З а м е ч а н и е. *Последний доверительный интервал зависит от параметров неизвестного распределения. Однако, ввиду (9) и $\widehat{\gamma}\log c(\widehat{\gamma}) > 0$, распределение x_p^c может быть определено посредством распределения x_p^w, т.е.*

$$P\{\log(\frac{x_p^c}{x_p}) > x\} < P\{\log(\frac{x_p^w}{x_p}) > x\},$$

и может быть использован доверительный интервал

$$x_p^c e^{-x^\star} < x_p < x_p^c e^{x^\star}.$$

П р и м е р ы. Результаты иллюстративных примеров приведены в Табл. 2. Рассматриваются распределение Парето с ФР $F(x) = 1 - x^{-1/\gamma}$ $(x \geq 1)$ и $\gamma = 1/2$, для которого $c = 1, d = 0, \beta = 0$ в (2), и распределение Фреше (Fréchet) с ФР $F(x) = \exp(-x^{-2})$ и $\gamma = 1/2$, $c = 1$, $d = -1/2$, $\beta = 2$. Для $\eta = 0,05$ вычисляется $x^\star$ по формуле (16). Доверительный интервал для x_p^w вычисляется по (15), а для x_p^c по (17). Приведены точные значения квантилей x_p для указанных распределений. Примеры показывают, что интервалы (17) точнее интервалов (15).

5. Заключение

Предложена новая оценка для квантилей высоких порядков. Оценка сравнивается с известными оценками средствами моделирования для различных тяжело-хвостовых распределений. Моделирование показывает, что эта оценка имеет меньшую среднеквадратичную ошибку. Доказывается, что распределения двух оценок (предложенной и оценки Вайсманна) асимптотически нормальны. Приведены доверительные интервалы оценок.

Приложение

Для доказательства теоремы 1 нужна следующая лемма.

Л е м м а. *Для оценки Хилла (10) выполнено*

$$(П.1.)\quad P\{\log(\frac{k}{np})(\widehat{\gamma}-\gamma)>x\}<$$

$$1-\Phi\left(\frac{\sqrt{k}(x-\varepsilon)}{\gamma\log(\frac{k}{np})}\right)-\frac{1}{3\sqrt{2\pi k}}\left(1-\frac{k(x-\varepsilon)^2}{\gamma^2(\log(\frac{k}{np}))^2}\right)\exp(-\frac{k(x-\varepsilon)^2}{2\gamma^2(\log(\frac{k}{np}))^2})+O\left(\left(\frac{k}{n}\right)^{2\gamma\beta}\right).$$

Д о к а з а т е л ь с т в о л е м м ы.

Будем использовать следующее равенство по вероятности: $x_{(i)} =^d \overline{F}^{-1}(\exp(-S_i))$, $1\le i\le n$, где $\overline{F}(x)=1-F(x)$, S_i - порядковые статистики, соответствующие выборке объема n независимых сл.в., экспоненциально распределенных с математическим ожиданием, равным 1, $\exp(-S_i)$ - порядковые статистики из выборки независимых сл.в., распределенных равномерно на $(0,1)$.

Для распределений типа Парето (2) имеем

$$(П.2.)\quad x_p =^d \overline{F}^{-1}(p)=\left(\frac{c}{p}\right)^{\gamma}\{1+\gamma d c^{-\gamma\beta}p^{\gamma\beta}+o(p^{\gamma\beta})\},$$

$$(П.3.)\quad x_{(n-k)} =^d \left(\frac{c}{\exp(-S_{n-k})}\right)^{\gamma}\{1+\gamma d c^{-\gamma\beta}\exp(-\gamma\beta S_{n-k})+o(\exp(-\gamma\beta S_{n-k}))\},$$

$$(П.4.)\quad \log x_{(n-k)} =^d \gamma(\log c+S_{n-k})+\gamma d c^{-\gamma\beta}\exp(-\gamma\beta S_{n-k})+o(\exp(-\gamma\beta S_{n-k})).$$

Из представления Реньи (Rényi)

$$S_{n-k} =^d \sum_{j=1}^{n-k}\frac{z_j}{n-j+1}=\zeta_{n-k}+T_{n-k}$$

(равенство выполнено по вероятности), где z_j - независимые экспоненциально распределенные сл. в. с единичным средним,

$$T_{n-k} =^d \sum_{j=1}^{n-k}\frac{z_j-1}{n-j+1},$$

$$\zeta_{n-k} = \sum_{j=1}^{n-k} \frac{1}{n-j+1} = \sum_{j=k+1}^{n} \frac{1}{j} = \log\left(\frac{n}{k+1}\right) + O(\frac{1}{n})$$

для $1 \leq k < n$, имеем математическое ожидание и дисперсию S_{n-k}:

$$ES_{n-k} = \log\left(\frac{n}{k+1}\right) + O(\frac{1}{n})$$

и

$$(\text{П.5.}) \qquad VarS_{n-k} = \frac{1}{(k+1)^2} + \ldots + \frac{1}{n^2} \simeq \frac{1}{k+1} - \frac{1}{n} + O(\frac{1}{k}).$$

Далее имеем:

$$\exp(-\gamma\beta S_{n-k}) = \exp(-\gamma\beta E(S_{n-k}))\exp(-\gamma\beta(S_{n-k} - E(S_{n-k})))$$

$$\sim \left(\frac{n}{k+1}\right)^{-\gamma\beta} \exp\left(-\gamma\beta\left(S_{n-k} - \log\left(\frac{n}{k+1}\right)\right)\right)$$

$$\sim \left(\frac{k+1}{n}\right)^{\gamma\beta}\left(1 + \gamma\beta(\log\frac{n}{k+1} - S_{n-k}) + o(\log\frac{n}{k+1} - S_{n-k})\right)$$

$$(\text{П.6.}) \qquad = \left(\frac{k+1}{n}\right)^{\gamma\beta}(1 + V_k),$$

где для $V_k = \gamma\beta\left(\log\frac{n}{k+1} - S_{n-k}\right)$ из неравенства Чебышева[3] следует свойство:

$$(\text{П.7.}) \qquad P\{\sup_{1\leq k<n} |V_k| \geq 1\} \leq (\gamma\beta)^2\left(\frac{1}{k+1} - \frac{1}{n}\right).$$

Теперь для оценки Хилла (10) для индекса γ следует:

$$\widehat{\gamma} =^d \gamma\left(\frac{1}{k}\sum_{i=1}^{k}(S_{n-i+1} + dc^{-\gamma\beta}\exp(-\gamma\beta S_{n-i+1})) - S_{n-k} - dc^{-\gamma\beta}\exp(-\gamma\beta S_{n-k})\right)$$
$$+o(\exp(-\gamma\beta S_{n-k})).$$

Очевидно, что

$$\frac{1}{k}\sum_{i=1}^{k} S_{n-i+1} - S_{n-k} = \frac{1}{k}\sum_{j=n-k+1}^{n} z_j.$$

Представим $\widehat{\gamma}$ следующим образом:

$$(\text{П.8.}) \qquad \widehat{\gamma} = A(k) + B(k) + C(k),$$

[3]Если σ^2 - дисперсия сл.в. t, то $P\{|t - Et| \geq \sigma\chi\} \leq \frac{1}{\chi^2}, \qquad \forall\chi > 0.$

где

$$A(k) = \gamma \frac{1}{k} \sum_{j=n-k+1}^{n} (z_j - 1),$$

$$B(k) = \gamma d c^{-\gamma\beta} (\frac{1}{k} \sum_{i=1}^{k} \left(\frac{i}{n}\right)^{\gamma\beta} - \left(\frac{k+1}{n}\right)^{\gamma\beta})$$

$$= \gamma d c^{-\gamma\beta} \left(\frac{k+1}{n}\right)^{\gamma\beta} \left(\frac{1}{k(k+1)^{\gamma\beta}} \sum_{i=1}^{k} i^{\gamma\beta} - 1\right)$$

$$\text{(П.9.)} \qquad \sim -\gamma d c^{-\gamma\beta} \left(\frac{k+1}{n}\right)^{\gamma\beta} \frac{\gamma\beta}{\gamma\beta + 1},$$

$$C(k) = \gamma \left(1 + d c^{-\gamma\beta} \left(\frac{k+1}{n}\right)^{\gamma\beta} \left(\frac{1}{k(k+1)^{\gamma\beta}} \sum_{i=1}^{k} i^{\gamma\beta} V_{i-1} - V_k\right) + o(\exp(-\gamma\beta S_{n-k}))\right)$$

$$< \gamma \left(1 + d c^{-\gamma\beta} \frac{\max_{1 \le i \le k} V_i}{\gamma\beta + 1} \left(\frac{k}{n}\right)^{\gamma\beta} + o(\exp(-\gamma\beta S_{n-k}))\right).$$

Тогда имеем:

$$\text{(П.10.)} \quad \widehat{\gamma} - \gamma < \gamma \left(\frac{1}{k} \sum_{j=n-k+1}^{n} (z_j - 1) - d c^{-\gamma\beta} \left(\frac{k+1}{n}\right)^{\gamma\beta} \frac{\gamma\beta}{\gamma\beta + 1} + R(k)\right),$$

где

$$R(k) = d c^{-\gamma\beta} \left(\frac{k}{n}\right)^{\gamma\beta} \frac{\max_{1 \le i \le k} V_i}{\gamma\beta + 1} + o(\exp(-\gamma\beta S_{n-k})).$$

Из (П.10.) следует, что

$$\log(\frac{k}{np})(\widehat{\gamma} - \gamma) < \gamma \log(\frac{k}{np}) \left(\frac{1}{k} \sum_{j=n-k+1}^{n} (z_j - 1) + R(k)\right).$$

Предположим, что $\gamma \frac{\log(\frac{k}{np})}{k} \sum_{j=n-k+1}^{n} (z_j - 1) \le x - \varepsilon$ и $\gamma \log(\frac{k}{np}) R(k) \le \varepsilon$, $\forall \varepsilon > 0$. Тогда $\log(\frac{k}{np})(\widehat{\gamma} - \gamma) \le x$. Отсюда,

$$P\{\log(\frac{k}{np})(\widehat{\gamma} - \gamma) > x\} <$$

$$\text{(П.11.)} \quad P\{\gamma \frac{\log(\frac{k}{np})}{k} \sum_{j=n-k+1}^{n} (z_j - 1) > x - \varepsilon\} + P\{\gamma \log(\frac{k}{np}) R(k) > \varepsilon\}.$$

Оценим сначала второе слагаемое в правой части (П.11.). Для этого заметим, что из неравенства Чебышева и (П.5.), (П.6.) следует

$$\text{(П.12.)} \quad P\{\gamma \log(\frac{k}{np}) \exp(-\gamma\beta S_{n-k}) > \varepsilon/2\} = P\{\gamma \log(\frac{k}{np})(1 + V_k) \left(\frac{k+1}{n}\right)^{\gamma\beta} > \varepsilon/2\}$$

$$< P\{|V_k| > \frac{\varepsilon}{2\gamma \log(\frac{k}{np})} \left(\frac{k+1}{n}\right)^{-\gamma\beta} - 1\}$$

$$< (\gamma\beta)^2 \left(\frac{1}{k+1} - \frac{1}{n}\right) \left(\frac{\varepsilon}{2\gamma \log(\frac{k}{np})} \left(\frac{n}{k+1}\right)^{\gamma\beta} - 1\right)^{-2} \sim O\left(\frac{1}{k}\left(\frac{k}{n}\right)^{2\gamma\beta}\right).$$

Аналогично имеем:

$$\text{(П.13.)} \quad P\{dc^{-\gamma\beta}\frac{\gamma}{\gamma\beta+1}\log(\frac{k}{np})\max_{1\le i\le k} V_i \left(\frac{k}{n}\right)^{\gamma\beta} > \varepsilon/2\}$$

$$< kP\{|V_i| > \frac{\varepsilon c^{\gamma\beta}(1+\gamma\beta)}{2\log(\frac{k}{np})d\gamma}\left(\frac{n}{k}\right)^{\gamma\beta}\}$$

$$< k(\gamma\beta)^2\left(\frac{1}{k+1} - \frac{1}{n}\right)\frac{4(d\gamma)^2(\log(\frac{k}{np}))^2}{\varepsilon^2 c^{2\gamma\beta}(1+\gamma\beta)^2}\left(\frac{k}{n}\right)^{2\gamma\beta} \sim O\left(\left(\frac{k}{n}\right)^{2\gamma\beta}\right).$$

Тогда из (П.12.), (П.13.) имеем:

$$\text{(П.14.)} \quad P\{\gamma\log(\frac{k}{np})R(k) > \varepsilon\}$$

$$< P\{dc^{-\gamma\beta}\frac{\gamma}{\gamma\beta+1}\log(\frac{k}{np})\max_{1\le i\le k} V_i\left(\frac{k}{n}\right)^{\gamma\beta} > \frac{\varepsilon}{2}\} + P\{\gamma\log(\frac{k}{np})\exp(-\gamma\beta S_{n-k}) > \frac{\varepsilon}{2}\}$$

$$\sim O\left(\left(\frac{k}{n}\right)^{2\gamma\beta}\right).$$

Оценим теперь первое слагаемое в правой части (П.11.). Ввиду того, что

$$E(z_j - 1) = 0, \qquad \sigma^2 = E(z_j-1)^2 = 1, \qquad \alpha_3 = E(z_j-1)^3 = 2,$$

верно

$$P\{\frac{\gamma\log(\frac{k}{np})}{k}\sum_{j=n-k+1}^{n}(z_j-1) > x\} = 1 - P\{\frac{1}{\sqrt{k}}\sum_{j=n-k+1}^{n}(z_j-1) < \frac{\sqrt{k}x}{\gamma\log(\frac{k}{np})}\}$$

$$= 1 - \Phi\left(\frac{\sqrt{k}x}{\gamma\log(\frac{k}{np})}\right) - \frac{1}{3\sqrt{2\pi k}}\left(1 - \frac{kx^2}{\gamma^2(\log(\frac{k}{np}))^2}\right)\exp(-\frac{kx^2}{2\gamma^2(\log(\frac{k}{np}))^2})$$

с точностью до $o(1/\sqrt{k})$, где $\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}\exp(-\frac{t^2}{2})dt$ - ФР стандартного нормального распределения. [4] Отсюда и из (П.11.) и (П.14.) следует (П.1.).

[4] Из теоремы Петрова [14] следует: для независимых сл.в. $x_j, j = 1, ..., n$ с общей непрерывной ФР и $Ex_1 = 0$, $\sigma^2 = Ex_1^2$, $\alpha_3 = Ex_j^3 < \infty$ для $F_n(x) = P\{\frac{1}{\sigma\sqrt{n}}\sum_{j=1}^{n} x_j < x\}$ верно

$$\text{(П.15.)} \quad F_n(x) = \Phi(x) + \frac{\alpha_3}{6\sigma^3\sqrt{2\pi n}}(1-x^2)\exp(-x^2/2) + o(n^{-1/2}).$$

З а м е ч а н и е. *Из леммы вытекает, что статистика* $\log(\frac{k}{np})(\widehat{\gamma}-\gamma)$ *асимптотически нормально распределена со средним* ε *и дисперсией* $\frac{\gamma^2(\log(\frac{k}{np}))^2}{k}$ *при* $n,k\to\infty$, $\frac{k}{n}\to 0$. *Это совпадает с классическим результатом:* $\sqrt{k}(\widehat{\gamma}-\gamma)\to^d N(0,\gamma^2)$ *(если* $\varepsilon\simeq 0$*). Однако, (П.1.) показывает точность этой сходимости и позволяет построить неасимптотический доверительный интервал для* $\widehat{\gamma}$.

Д о к а з а т е л ь с т в о т е о р е м ы 1.

Найдем сначала распределение для $\log\left(\frac{x_p^w}{x_p}\right)$. Из (7) имеем:

$$(\text{П.16.})\quad \widehat{\gamma}\log c(\widehat{\gamma})=\widehat{\gamma}\log\{1+x_{(n-k)}^{-1/\widehat{\gamma}}+x_{(n-k)}^{-2/\widehat{\gamma}}\}\simeq\widehat{\gamma}x_{(n-k)}^{-1/\widehat{\gamma}}+o(x_{(n-k)}^{-1/\widehat{\gamma}}).$$

Из (П.2.-П.4.), (П.16.) следует, что

$$(\text{П.17.})\quad \log(\frac{x_p^w}{x_p})=\log x_{(n-k)}-\log x_p+\widehat{\gamma}\log a_n$$

$$=\gamma(S_{n-k}-\log\left(\frac{n}{k}\right))+(\gamma-\widehat{\gamma})\log\left(\frac{n}{k}\right)+(\gamma-\widehat{\gamma})\log p$$

$$+\gamma dc^{-\gamma\beta}(\exp(-\gamma\beta S_{n-k})-p^{\gamma\beta})+o(\exp(-\gamma\beta S_{n-k}))+o(p^{\gamma\beta}).$$

Из (9), (П.3.) и (П.16.) имеем

$$(\text{П.18.})\quad \log(\frac{x_p^c}{x_p})\sim\log(\frac{x_p^w}{x_p})$$

$$-\widehat{\gamma}\left(\frac{\exp(-S_{n-k})}{c}\right)^{\gamma/\widehat{\gamma}}\left(1+\gamma dc^{-\gamma\beta}\exp(-\gamma\beta S_{n-k})+o(\exp(-\gamma\beta S_{n-k}))\right)^{-1/\widehat{\gamma}}.$$

Заметим, что

$$(\text{П.19.})\quad (\gamma-\widehat{\gamma})\log p+(\gamma-\widehat{\gamma})\log\left(\frac{n}{k}\right)=(\widehat{\gamma}-\gamma)\log(\frac{k}{np}).$$

Можно найти ФР для $S_{n-k}-\log\left(\frac{n}{k}\right)$. Так как по вероятности $S_{n-k}=^d-\log U_{(k+1)}$, где $U_{(k)}$ - порядковая статистика из выборки равномерно распределенных сл.в., то

$$(\text{П.20.})\quad P\{S_{n-k}-\log\frac{n}{k}<x\}=1-P\{U_{(k+1)}\le\exp(-x)\frac{k}{n}\}$$

$$=1-\mathrm{I}_{\exp(-x)\frac{k}{n}}(k+1,n-k)$$

$$=1-\frac{1}{B(k+1,n-k)}\int_0^{\exp(-x)k/n}t^k(1-t)^{n-k-1}dt,$$

где $\mathrm{I}_x(a,b)$ - неполная бета функция, $B(k+1,n-k)=\frac{\Gamma(k+1)\Gamma(n-k)}{\Gamma(n+1)}$.

Рассмотрим равенство (П.17.), учитывая (П.19.). Предположим, что $\gamma(S_{n-k} - \log(\frac{n}{k})) \leq -1$, $\log(\frac{k}{np})(\widehat{\gamma} - \gamma) \leq x$ и $\gamma d c^{-\gamma\beta}\exp(-\gamma\beta S_{n-k}) \leq 1$. Тогда $\log\left(\frac{x_p^w}{x_p}\right) < x$. Отсюда и из (П.1.), (П.6.), (П.7.), (П.20.) получим формулу для ФР статистики $\log\left(\frac{x_p^w}{x_p}\right)$:

$$P\{\log\left(\frac{x_p^w}{x_p}\right) > x\} < P\{\gamma(S_{n-k} - \log(\frac{n}{k})) > -1\} + P\{\log(\frac{k}{np})(\widehat{\gamma} - \gamma) > x\}$$

$$+P\{\gamma d c^{-\gamma\beta}\exp(-\gamma\beta S_{n-k}) > 1\}$$

$$< 1 + \mathrm{I}_{\exp(\frac{1}{\gamma})\frac{k}{n}}(k+1, n-k) - \Phi\left(\frac{\sqrt{k}(x-\varepsilon)}{\gamma\log(\frac{k}{np})}\right)$$

$$-\frac{1}{3\sqrt{2\pi k}}\left(1 - \frac{k(x-\varepsilon)^2}{\gamma^2(\log(\frac{k}{np}))^2}\right)\exp(-\frac{k(x-\varepsilon)^2}{2\gamma^2(\log(\frac{k}{np}))^2})$$

$$+(\gamma\beta)^2\left(\frac{1}{k+1} - \frac{1}{n}\right)\left(\frac{c^{\gamma\beta}}{\gamma d}\left(\frac{n}{k+1}\right)^{\gamma\beta} - 1\right)^{-2} + O\left(\left(\frac{k}{n}\right)^{2\gamma\beta}\right) = 1 - F_w(x,p,k). \tag{П.21.}$$

Теперь найдем ФР для $\log\frac{x_p^c}{x_p}$. Из (П.18.) имеем

$$\log(\frac{x_p^c}{x_p}) \sim \log(\frac{x_p^w}{x_p}) - \widehat{\gamma}\left(\frac{\exp(-\gamma S_{n-k})}{c^\gamma(1 + \gamma d c^{-\gamma\beta}\exp(-\gamma\beta S_{n-k}))}\right)^{1/\widehat{\gamma}}.$$

Предположим, что $\gamma S_{n-k} \leq 1$, $\log(\frac{x_p^w}{x_p}) \leq x$, $\frac{1}{k}\sum_{j=n-k+1}^{n}(z_j - 1) \geq -0,5$, $V_{i-1} \geq 1$ для всех $i = 1, ..., k$ и $V_k < 1$. Из последних трех предположений и из (П.8.) следует:

$$\widehat{\gamma} > \gamma/2 + 2B(k) = \gamma^\star,$$

где $B(k)$ определяется из (П.9.). Отсюда следует:

$$\log(\frac{x_p^c}{x_p}) \leq x - \gamma^\star\left(\frac{A}{e}\right)^{1/\gamma^\star},$$

где

$$A = \begin{cases} (c^\gamma(1 + \gamma d c^{-\gamma\beta}))^{-1}, & \text{если } c^\gamma(1 + \gamma d c^{-\gamma\beta}) \neq 0, \\ 0 & \text{в противном случае .} \end{cases}$$

Тогда,

$$P\{\log(\frac{x_p^c}{x_p}) > x - \gamma^\star\left(\frac{A}{e}\right)^{1/\gamma^\star}\} < P\{\log(\frac{x_p^w}{x_p}) > x\} + P\{\gamma S_{n-k} > 1\} \tag{П.22.}$$

$$+P\{\frac{1}{k}\sum_{j=n-k+1}^{n}(z_j-1)<-1/2\}+\sum_{i=1}^{k}P\{V_{i-1}<1\}+P\{V_k\geq 1\}.$$

Из (П.15.) имеем:

$$P\{\frac{1}{k}\sum_{j=n-k+1}^{n}(z_j-1)<-1/2\}$$

$$(\text{П.23.})\quad =\Phi(-\frac{\sqrt{k}}{2})+\frac{1}{3\sqrt{2\pi k}}\left(1-\frac{k}{4}\right)\exp(-\frac{k}{8})+o\left(\frac{1}{\sqrt{k}}\right).$$

Из (П.20.) получим:

$$(\text{П.24.})\quad P\{V_{i-1}<1\}=P\{\log(n/i)-S_{n-i+1}<\frac{1}{\gamma\beta}\}=\mathrm{I}_{\frac{i}{n}\exp(1/\gamma\beta)}(i,n-i+1).$$

Аналогично можно получить:

$$(\text{П.25.})\quad P\{V_k\geq 1\}=1-P\{V_k<1\}.$$

Далее имеем:

$$(\text{П.26.})\quad P\{\gamma S_{n-k}>1\}=P\{-\log U_{(k+1)}>\frac{1}{\gamma}\}=\mathrm{I}_{\exp(-1/\gamma)}(k+1,n-k).$$

Из (П.21.)-(П.26.) следует (11) и теорема доказана.

СПИСОК ЛИТЕРАТУРЫ

1. *Dielman T., Lowry C., Pfaffenberger R.* A comparison of quantile estimators// Commun. Statist.-Simula. 1994. № 23(2). P.355-371.
2. *Gnedenko B.V.* Sur la Distribution Limite du Terme Maximum d'une Série Aléatoire// Ann. Math. 1943. № 44. P.423-453.
3. *Pickands J.* Statistical inference using extreme order statistics//Ann. Statist. 1975. № 3. P. 119-131.
4. *Hall P.* Using the Bootstrap to Estimate Mean Squared Error and Select Smoothing Parameter in Nonparametric Problems// J. Mult. Anal. 1990. № 32. P.177–203.
5. *Hall P., Weissman I.* On the estimation of exreme tail probabilities//Ann. Stat. 1997. № 25(3). P.1311-1326.
6. *Beirlant J.* Tail index estimation and exponential regression model//Extremes. 1999. № 2(2). P.177-200.
7. *Dekkers A.L., Einmahl J.H.J., de Haan L.* A moment estimator for the index of an extreme-value distribution //Ann. Stat. 1989. № 17(4). P.1833-1855.
8. *Weissman I.* Estimation of parameters and large quantiles based on the k largest observations// J. American Statist. Assoc. 1978. № 73. P.812-815.
9. *McNeil A.J.* Estimating the tails of loss severity distributions using extreme value theory// Astin Bulletin. 1997. № 27(1). P.117–137.
10. *Caers J., Van Dyck J.* Nonparametric tail estimation using a double bootstrap method//Computat. Statist. & Data Analysis. 1999. № 29. P. 191-211.
11. *Hill B.M.* A simple general approach to inference about the tail of a distribution// Ann.Statist. 1975. № 3. P. 1163-1174.
12. *Goldie C.M., Smith R.L.* Slow variation with remainder: theory and applications// Quart.J.Math. Oxford. 1987. № 38. P.45-71.
13. *Ferreira A., de Haan L., L.Peng.* Adaptive estimators for the endpoint and high quantiles of a probability distribution// Eurandom: Research Report. 2000. No. 99-042.
14. *Petrov V.V.* Sums of Independent Random Variables. Springer, N. Y. 1975.

Таблица 1. Точность оценивания квантилей

Квантильная оценка	Объем выборки	$\%Bias$		$\%RMSE$	
		$1-p=0,99$	$1-p=0,999$	$1-p=0,99$	$1-p=0,999$
		Распределение Парето ($1/\gamma=2$)			
x_p^{POT}	250	2,5	23,60	29,17	132,18
x_p^c		-9,7	-22,9	18,0	27,7
x_p^w		-9,3	-6,6	24,7	20,1
x_p^{POT}	500	2,58	19,16	20,88	91,93
x_p^c		-9,7	-9,6	15,7	17,4
x_p^w		-7,3	10,4	19,4	22,7
		Распределение Парето ($1/\gamma=1$)			
x_p^{POT}	250	10,52	207,49	83,28	1611,14
x_p^c		-27,5	-25,5	35,5	58,2
x_p^w		8,4	16,5	45,5	93,9
x_p^{POT}	500	7,69	51,70	50,25	254,02
x_p^c		-19,8	-25,0	32,9	44,7
x_p^w		12,0	12,1	40,5	65,5
		Распределение Стьюдента ($1/\gamma=2$)			
x_p^{POT}	250	1,12	4,61	25,17	81,24
x_p^c		-15,2	32	37,9	69,8
x_p^w		34,1	122	52,4	157,9
x_p^{POT}	500	0,95	3,22	20,72	65,16
x_p^c		-22,2	2,7	25,7	32,5
x_p^w		17,7	62,8	30,4	86,3
		Распределение Стьюдента ($1/\gamma=1$) (Коши)			
x_p^{POT}	250	14,75	131,52	69,82	495,95
x_p^c		-17,4	-28,9	40,4	57,7
x_p^w		37,7	55,7	62,3	132,7
x_p^{POT}	500	3,98	31,86	40,16	163,80
x_p^c		-27,7	-31,7	36,6	42,1
x_p^w		20,9	31,0	47,7	72,7

Таблица 2. Доверительные 95% интервалы оценок квантилей x_p^w и x_p^c.

Оценка Квантилей	Размер выборки	Доверительные интервалы $1-p=0,99$	Доверительные интервалы $1-p=0,999$
		Распределение Парето ($\gamma=1/2$)	
		$x_p=10$	$x_p=31,623$
x_p^c	500	$[7,269;12,286]$	$[14,247;34,307]$
x_p^w		$[8,785;14,984]$	$[17,532;42,607]$
x_p^c	1000	$[7,936;11,684]$	$[27,437;39,835]$
x_p^w		$[9,581;14,038]$	$[33,277;48,759]$
		Распределение Фреше ($\gamma=1/2$)	
		$x_p=9,975$	$x_p=31,615$
x_p^c	500	$[6,073;10,124]$	$[11,978,32,724]$
x_p^w		$[5,605;10,398]$	$[10,782;37,258]$
x_p^c	1000	$[6,478;9,351]$	$[13,678,29,219]$
x_p^w		$[5,681;8,946]$	$[10,883;36,715]$