Теория экстремальных значений (одномерный случай)

Финансовая эконометрика

Метод блочных максим

Распределение максимумов потерь

$\vec{x} = (x_1; \dots; x_n)$ — убытки, $\vec{x} \sim iid. F(x)$

$M_n = \max(x_1; \dots; x_n)$ — максима

$F_{M_n}(x) = P(M_n \leq x) = F^n(x)$ — распределение максимы

Пусть нормализованные максимумы сходятся к некоторому распределению $H(x)$, это означает, что

$\exists d_n, c_n > 0\colon \lim\limits_n P\left(\frac{M_n - d_n}{c_n} \leq x\right) = \lim\limits_n F^n(c_n x + d_n) = H(x),$

тогда $F \in MDA(H)$

Generalized Extreme Value distribution (GEV)

<u>Теорема</u> Фишера-Типпетта-Гнеденко

Если $F \in MDA(H)$ и H не сосредоточено в одной точке, то $H \sim GEV(\mu(c_n, d_n); \sigma(c_n, d_n); \xi)$

$GEV(0; 1; \xi)$: $H_\xi(x) = \begin{cases} e^{-(1+\xi x)^{-\frac{1}{\xi}}}, \xi \neq 0 \\ e^{-e^{-x}}, \xi = 0 \end{cases}$, где $1 + \xi x > 0$

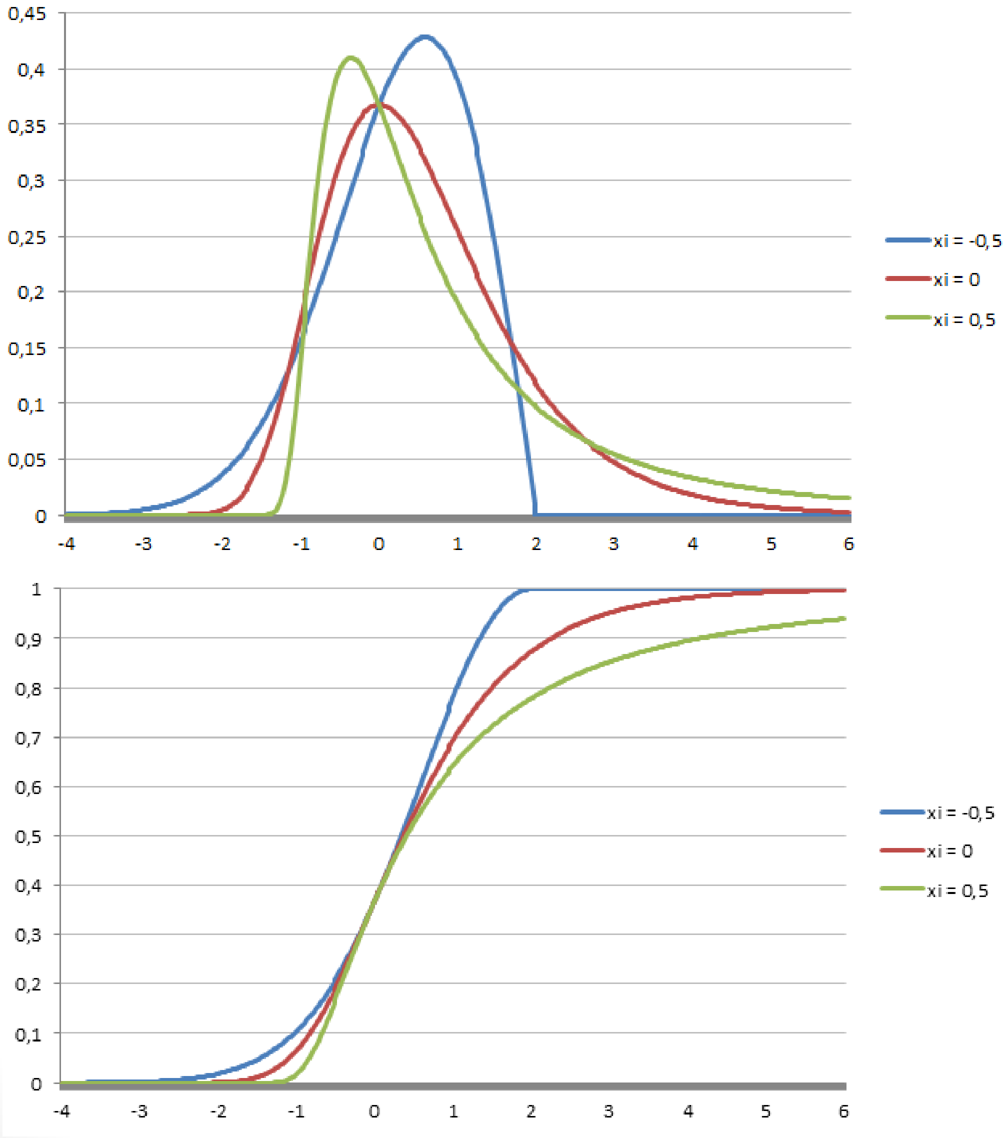
Частные случаи GEV:

- $\xi > 0$ — распределение Фрешѐ
- $\xi = 0$ — распределение Гумбеля
- $\xi < 0$ — распределение Вейбулла

Распределение Вейбулла имеет конечную правую точку $x_F = \sup\{x \in R: F(x) < 1\}$

Фреше и Гумбель не имеют конечных правых точек, но Фреше убывает значительно медленнее

Функции распределения и плотности GEV



Пример. Экспоненциальное распределение

$$F(x) = 1 - e^{-\beta x}, \qquad \beta > 0, x \geq 0$$

$$M_n \sim F^n(x) = \left(1 - e^{-\beta x}\right)^n$$

$$\frac{M_n - d_n}{c_n} \sim F^n(c_n x + d_n) = \left(1 - e^{-\beta(c_n x + d_n)}\right)^n$$

Пусть $c_n = \frac{1}{\beta}, d_n = \frac{1}{\beta}\ln n$, тогда

$$F^n(c_n x + d_n) = \left(1 - e^{-\beta\left(\frac{1}{\beta}x + \frac{1}{\beta}\ln n\right)}\right)^n = \left(1 - \frac{1}{n}e^{-x}\right)^n$$

$$\lim_n F^n(c_n x + d_n) = \lim_n \left(1 - \frac{1}{n}e^{-x}\right)^n = e^{-e^{-x}}$$, т.е.

$$\frac{M_n - d_n}{c_n} \xrightarrow{n} GEV(0; 1; 0)$$

Максима временного ряда

Пусть $(x_1, \dots, x_n) \sim F$ — стационарный временной ряд,
$(\bar{x}_1, \dots, \bar{x}_n) \sim WN(F)$ — соответствующий ему белый шум
Пусть $\bar{M}_n = \max(\bar{x}_1, \dots, \bar{x}_n)$, тогда
$\exists \theta \in (0; 1]$: $\lim\limits_n P\left(\frac{\bar{M}_n - d_n}{c_n} \leq x\right) = H(x) \Leftrightarrow \lim\limits_n P\left(\frac{M_n - d_n}{c_n} \leq x\right) = H^\theta(x)$, т.е.
если нормализованные максимы независимых величин сходятся к $H(x)$, то нормализованные максимы временного ряда сходятся к $H^\theta(x)$, причём
$H(x) \sim GEV(\mu; \sigma; \xi) \Longrightarrow H^\theta(x) \sim GEV(\mu(\theta); \sigma(\theta); \xi)$

θ — экстремальный индекс процесса $(x_1, \dots, x_n)$

Содержательная интерпретация экстремального индекса

Пусть $u = c_n x + d_n$, тогда при большом *n* имеем:

$P(M_n \leq u) \approx P^{\theta}(\bar{M}_n \leq u) = F^{n\theta}(u)$, таким образом распределение максимумов временного ряда длиной *n* может быть аппроксимировано распределением максимумов соответствующего ему белого шума длиной $n\theta < n$

При этом θ интерпретируется как количество относительно независимых кластеров во временном ряде

$\theta = 1 \Rightarrow$ экстремальные значения не кластеризуются,

$\theta < 1 \Rightarrow$ экстремумы имеют тенденцию кластеризоваться

- $\vec{x} \sim WN, ARMA(m; n) \Rightarrow \theta = 1$
- $\vec{x} \sim ARCH(q), GARCH(p; q) \Rightarrow \theta < 1$

Оценка параметров GEV

$\vec{x} = (x_1, \dots, x_T),\ T = m \cdot n$

$M_{n,j} = \max(x_{n(j-1)+1}, \dots, x_{nj})$

$M_n = \left(M_{n,1}, \dots, M_{n,m}\right) \sim GEV(\mu, \sigma, \xi)$

Пусть $h(x; \mu, \sigma, \xi)$ — плотность $GEV(\mu, \sigma, \xi)$, тогда

$$l\left(M_{n,1}, \dots, M_{n,m}; \mu, \sigma, \xi\right) = \sum_{i=1}^{m} \ln h\left(M_{n,i}; \mu, \sigma, \xi\right) = -m \cdot \ln \sigma -$$

$$-\left(1 + \frac{1}{\xi}\right) \sum_{i=1}^{m} \ln\left(1 + \xi \cdot \frac{M_{n,i} - \mu}{\sigma}\right) - \sum_{i=1}^{m} \ln\left(1 + \xi \cdot \frac{M_{n,i} - \mu}{\sigma}\right)^{-\frac{1}{\xi}} \to$$

$$\to \max_{\mu, \sigma > 0, \xi, 1 + \frac{\xi\left(M_{n,i} - \mu\right)}{\sigma} > 0}$$

Оценка параметров GEV в R

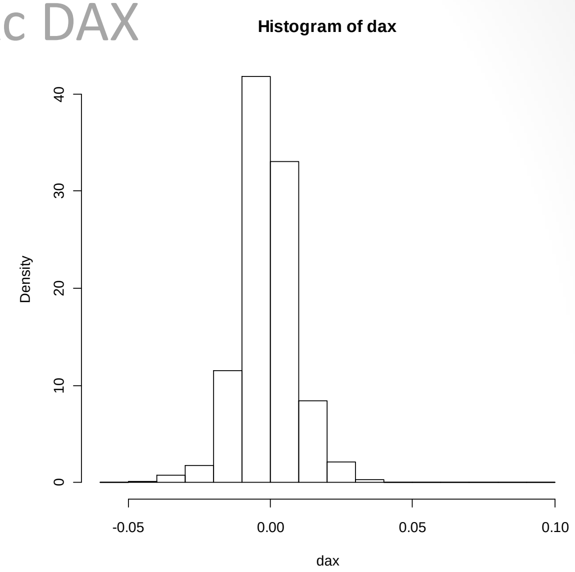
Практический пример 1. Биржевой индекс DAX

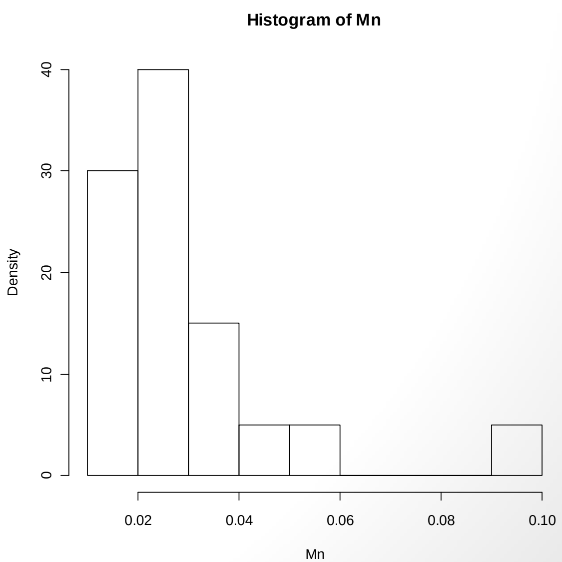
загрузка данных

```
library(datasets)
dax <- EuStockMarkets[,1]
n <- 90; m <- 20; T <- m*n
dax <- dax[2:(T+1)]/dax[1:T]-1
dax <- -dax
```

расчёт максим

```
Mn <- rep(0,times=m)
for (i in 1:m)
Mn[i] <- max(dax[((i-1)*n+1):(i*n)])
```

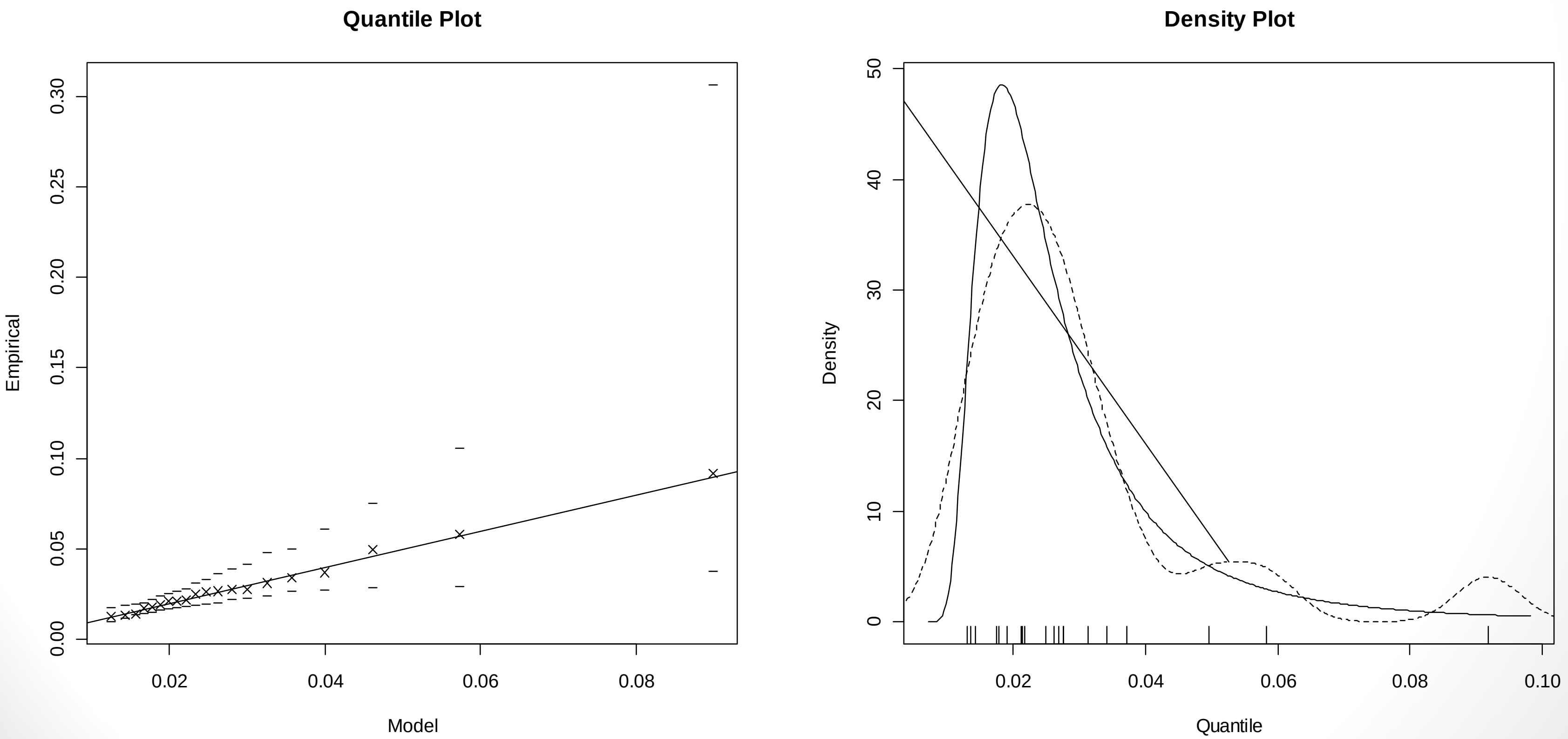




Оценка параметров GEV в R

распределение максим на основе GEV

```
library(evd)
Mn.fit <- fgev(Mn)
plot(Mn.fit)
```



Пороговый уровень и средний период наступления события

$$r_{n,k} = q_{1-\frac{1}{k}}(H) = H^{-1}_{\xi,\mu,\sigma}\left(1-\frac{1}{k}\right) = \mu + \frac{\sigma}{\xi}\left(\left(-\ln\left(1-\frac{1}{k}\right)\right)^{-\xi} - 1\right)$$

уровень, который будет превзойдён в среднем 1 раз за *k* блоков по *n* наблюдений

$k_{n,u} = \frac{1}{1-H(u)}$ — средний период наступления события $M_n > u$

$r_{n,k_{n,u}} = u$

расчёт этих показателей в R

```
mu <- Mn.fit$estimate[1]; sigma <- Mn.fit$estimate[2]
xi <- Mn.fit$estimate[3]; k <- 4; u <- 0.09
r.nk <- mu+sigma/xi*((-log(1-1/k))^(-xi)-1)
k.nr <- 1/(1-pgev(u,loc=mu,scale=sigma,shape=xi))
```

$r_{n,k}$	**0.034**
$k_{n,u}$	**40.14**

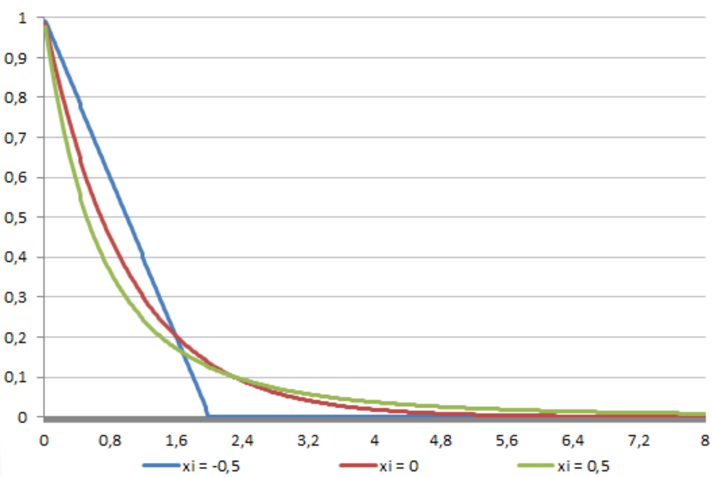
Обобщённое распределение Парето

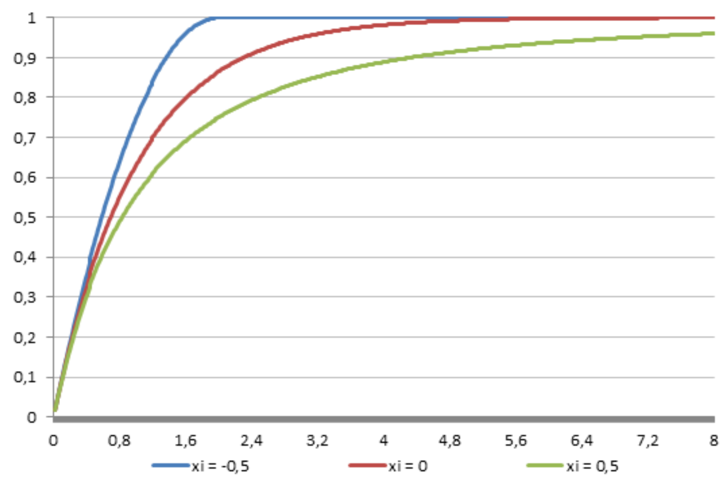
Generalized Pareto distribution (GPD)

$$G(x;\xi,\beta)=\begin{cases}1-\left(1+\frac{\xi x}{\beta}\right)^{-\frac{1}{\xi}}, & \xi\neq 0\\ 1-e^{-\frac{x}{\beta}}, & \xi=0\end{cases}, \text{где} \begin{cases}\begin{cases}0\le x\le-\frac{\beta}{\xi}, & \xi<0\\ x\ge 0, & \xi\ge 0\end{cases}\\ \beta>0\end{cases}$$

Частные случаи GPD:

- $\xi > 0$ — распределение Парето
- $\xi = 0$ — экспоненциальное распределение
- $\xi = 0$ — короткохвостое распределение Парето





Превышение порогового значения

Пусть $x_i \sim F$, тогда распределение превышений порога *u* равно

$$F_u(x) = P(x_i - u \le x | x > u) = \frac{F(x+u) - F(u)}{1 - F(u)},\ 0 \le x \le x_F - u$$

$e(u) = E(x_i - u | x_i > u)$ — среднее превышение порога

Если $F \equiv G_{\xi,\beta}$, то $F_u(x) \equiv G_{\xi,\beta(u)}(x),\ \beta(u) = \beta + \xi u$,

$e(u) = \frac{\beta(u)}{1-\xi} = \frac{\beta + \xi u}{1-\xi}$, т.е. распределение превышений для любого порога *u* остаётся GPD с тем же параметром формы ξ, а среднее превышение является линейной функцией относительно *u*

<u>Теорема</u> Пикандса-Балкема-де Хаана

$$\exists \beta(u):\ \lim_{u \to x_F} \sup \left| F_u(x) - G_{\xi,\beta(u)}(x) \right| = 0,\ 0 \le x < x_F - u \Leftrightarrow$$

$$F \in MDA\left(H_\xi\right),\ \xi \in R,$$

т.о. если распределение максимумов сходится к H_ξ, то превышения для высокого порога *u* описываются GPD

Оценка параметров GPD

Пусть $F_u(x) = G_{\xi,\beta}(x),\ 0 \le x \le x_F - u,\ \beta > 0,\ \xi \in R$

$\vec{x} = (x_1, \dots, x_n) \sim F,\ \sum_{i=1}^{n} I(x_i > u) = N_u \to \left(\bar{x}_1, \dots, \bar{x}_{N_u}\right)$: $\forall j \in \{1; \dots; N_u\}\ \bar{x}_j > u \to y_j = \bar{x}_j - u$

Если $\vec{x} \sim iid$, то

$$l\left(\xi, \beta; y_1, \dots, y_{N_u}\right) = \sum_{j=1}^{N_u} \ln g_{\xi,\beta}\left(y_j\right) =$$
$$- N_u \ln \beta - \left(1 + \frac{1}{\xi}\right) \sum_{j=1}^{N_u} \ln\left(1 + \xi \frac{y_j}{\beta}\right) \to max,\ \beta > 0, 1 + \xi \frac{y_j}{\beta} > 0$$

Превышение более высокого порога:

$$F_u(x) = G_{\xi,\beta}(x) \Rightarrow F_v(x) = G_{\xi,\beta+\xi(v-u)}(x),\ v \ge u$$

$$e(v) = \frac{\beta+\xi(v-u)}{1-\xi} = \frac{\xi v}{1-\xi} + \frac{\beta-\xi u}{1-\xi},\ \begin{cases} u \le v < +\infty, 0 \le \xi < 1 \\ u \le v \le u - \frac{\beta}{\xi}, \xi < 0 \end{cases}$$

$e(v)$ — линейно по *v*

Расчёт измерителей риска

Пусть $F_u(x) = G_{\xi,\beta}(x),\ 0 \leq x < x_F - u, \beta > 0, \xi \in R$, тогда для $x \geq u\ \ \tilde{F}_u(x) = P(x_i > u) \cdot P(x_i > x | x_i > u) = \bar{F}(u) \cdot P(x_i - u > x - u | x_i > u) = \bar{F}(u) \cdot \bar{F}_u(x-u) = \bar{F}(u) \cdot \left(1 + \frac{\xi(x-u)}{\beta}\right)^{-\frac{1}{\xi}}$ — распределение хвоста доходностей

Используя эту формулу, можно находить квантили убытков:

$$VaR_{1-\alpha} = q_\alpha(F) = u + \frac{\beta}{\xi}\left(\left(\frac{1-\alpha}{\bar{F}(u)}\right)^{-\xi} - 1\right),\ \alpha \geq F(u)$$

$ES_{1-\alpha} = \frac{1}{1-\alpha}\int_\alpha^1 q_x(F)dx = \frac{VaR_\alpha}{1-\xi} + \frac{\beta - \xi u}{1-\xi},\ \xi < 1$, также верно:

$$ES_{1-\alpha} = VaR_{1-\alpha} + e(VaR_{1-\alpha})$$

Smith (1987): $\hat{\bar{F}}(x) = \frac{N_u}{n}\left(1 + \frac{\hat{\xi}(x-u)}{\hat{\beta}}\right)^{-\frac{1}{\hat{\xi}}},\ x \geq u$

$$\alpha \geq 1 - \frac{N_u}{n} \to \widehat{VaR}_{1-\alpha}, \widehat{ES}_{1-\alpha}$$

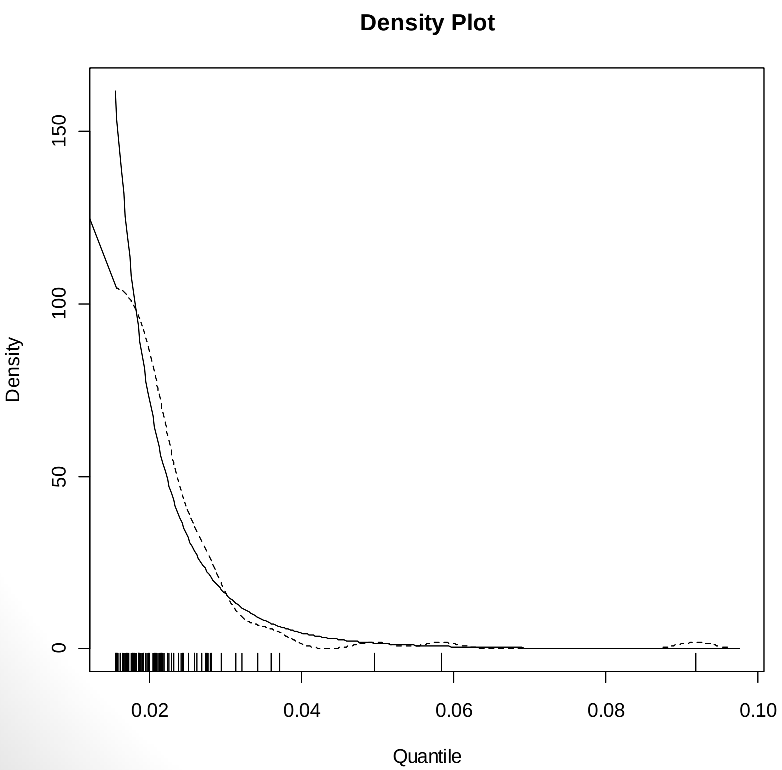
GPD в R

пороговое значение - 95% квантиль

```
u <- sort(dax)[0.95*T]
gpd.fit <- fpot(dax,threshold=u,model="gpd",method="SANN")
```

оценки параметров

```
beta <- gpd.fit$estimate[1]; xi <- gpd.fit$estimate[2]
```



GPD в R

расчёт мер риска

```
Fu <- gpd.fit$pat
alpha <- 1-1/260
```
соответствует k = 4

```
VaR <- u+beta/xi*(((1-alpha)/Fu)^(-xi)-1)
ES <- (VaR+beta-xi*u)/(1-xi)
```

VaR	0.036
ES	0.048

Домашнее задание

- рассчитать оценки риска для акции или биржевого индекса методом блочных максим и методом превышения порогового значения
- построить кривые VaR и протестировать качество полученных оценок
- написать комментарии

Исходные данные — котировки с сайтов finam.ru, finance.yahoo.org и др.