

Анализ доходности и волатильности финансовых активов

ЦМФ. Математические финансы

Первичный статистический анализ

Затабулированные распределения

Название	Обозначение в R	Параметры
Нормальное	norm	mean, sd
t-распределение	t	df
Равномерное	unif	min, max
Хи-квадрат	chisq	df
F-распределение	f	df1, df2
Гамма	gamma	shape, scale
...

пример со стандартным нормальным распределением

```
N <- 100; x <- seq(-5,5,by=0.1); alpha <- 0.95
```

```
rnorm(n=N,mean=0,sd=1)
```

генератор случайных чисел

```
qnorm(alpha,mean=0,sd=1)
```

квантиль

```
pnorm(x,mean=0,sd=1)
```

функция распределения

```
dnorm(x,mean=0,sd=1)
```

функция плотности

Гистограмма и эмпирическая плотность

```
y <- faithful$eruptions # исходные данные
```

```
# гистограмма
```

```
hist(y, nclass=20, prob=TRUE)
```

```
# добавление эмпирической плотности
```

```
y.pdf <- density(y, bw="ucv")
```

```
lines(y.pdf)
```

```
# добавление исходных данных
```

```
rug(y)
```

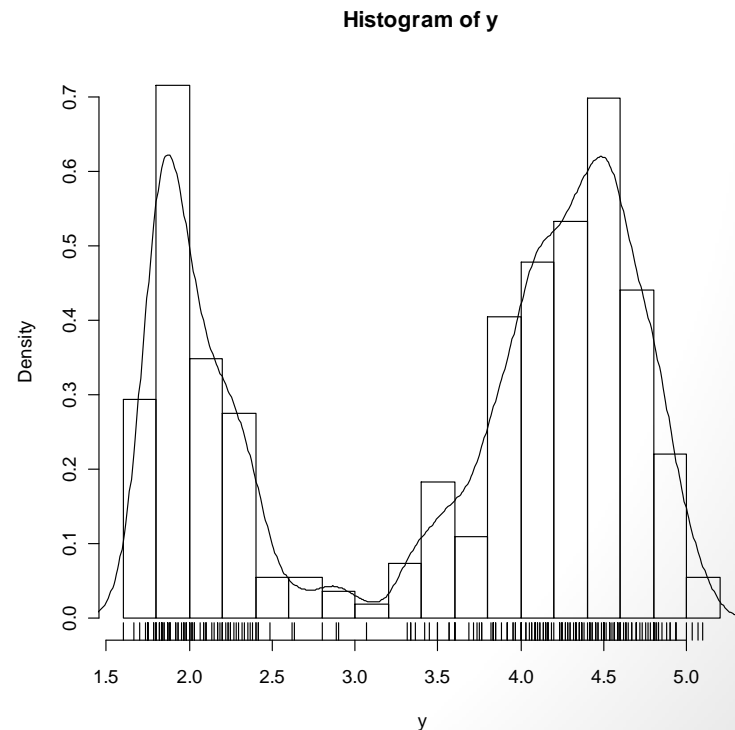


График «квантиль — квантиль»

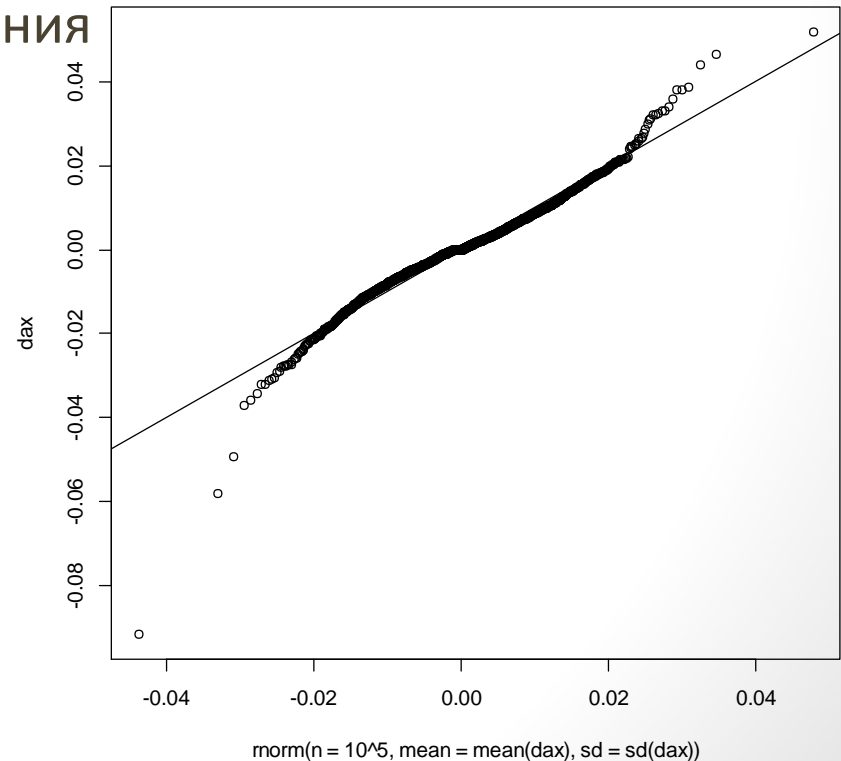
```
price <- EuStockMarkets[, "DAX"]; T <- length(price) - 1  
dax <- price[2:(T+1)] / price[1:T] - 1
```

в общем виде

```
qqplot(rnorm(n=10^5, mean=mean(dax), sd=sd(dax)), dax)  
abline(0, 1)
```

для нормального распределения

```
qqnorm(dax); qqline(dax)
```



Тесты на нормальность

Шапиро–Уилка

гипотеза: $H_0: x \sim N(\mu, \sigma)$

статистика: $W = \frac{(\sum a_i x_{(i)})^2}{\sum (x_i - \bar{x})^2}$, $(a_1, \dots, a_n) = \frac{m' V^{-1}}{(m' V^{-1} V^{-1} m)^{0.5}}$,

$m_i = E(x_{(i)} | x \sim N(0,1))$, $V = cov(m)$

`shapiro.test(dax)`

Колмогорова–Смирнова

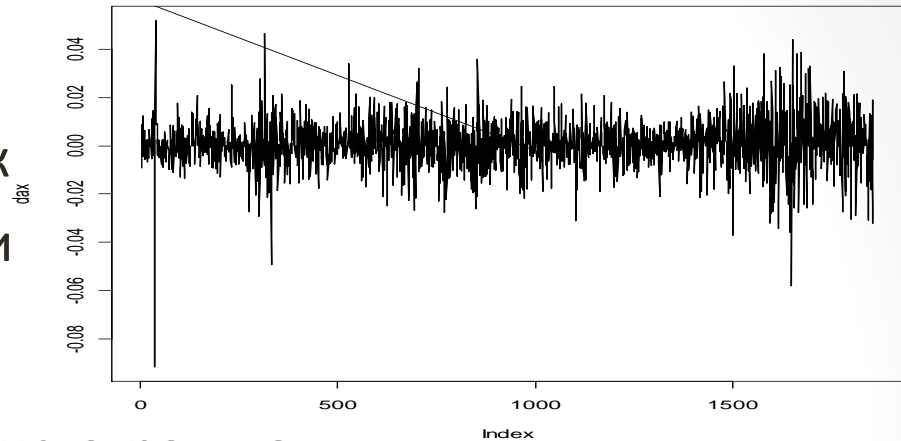
гипотеза: $H_0: x \sim F(x)$

статистика: $D = \sup_x |y.cdf(x) - F(x)|$

`ks.test(dax, "pnorm", mean=mean(dax), sd=var(dax)^0.5)`

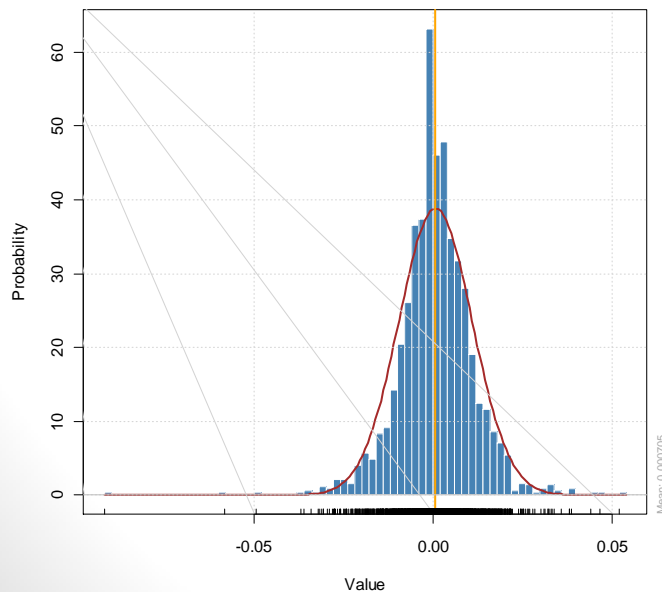
Сводная статистическая информация

```
library(fBasics)  
plot(dax,type="l") # график  
basicStats(dax) # статистики
```

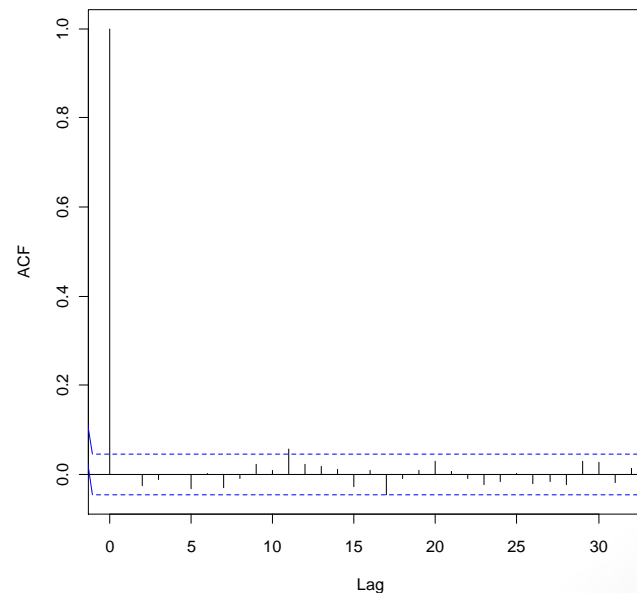


```
histPlot(timeSeries(dax)) # гистограмма  
acf(dax) # автокорреляционная функция
```

SS.1 Histogram



Series dax



Моделирование толстохвостых доходностей с помощью обобщённого гиперболического распределения

Обобщённое гиперболическое распределение (GHD)

$$f_{GHD}(x; \mu, \sigma, \gamma, \lambda, \chi, \psi) = \frac{(\psi\chi^2)\psi^\lambda\left(\psi+\frac{\gamma^2}{\sigma^2}\right)^{\frac{1}{2}-\lambda} K_{\lambda-\frac{1}{2}}\left(\sqrt{\left(\chi+\frac{(x-\mu)^2}{\sigma^2}\right)\left(\psi+\frac{\gamma^2}{\sigma^2}\right)}\right) e^{\frac{\gamma(x-\mu)}{\sigma^2}}}{(2\pi)^{\frac{1}{2}}\sigma K_\lambda(\psi\chi)^{\frac{1}{2}}\left(\sqrt{\left(\chi+\frac{(x-\mu)^2}{\sigma^2}\right)\left(\psi+\frac{\gamma^2}{\sigma^2}\right)}\right)^{\frac{1}{2}-\lambda}}$$

$K_u(v)$ — модифицированная функция Бесселя второго рода

Оценка параметров распределения

```
library(ghyp)
```

```
fit.[...]uv(dax,symmetric=FALSE,silent=TRUE)
```

если symmetric == FALSE, то оценивается скошенное

распределение, иначе — симметричное;

вместо [...] следует подставить название распределения:

ghyp — обобщённое гиперболическое

hyp — гиперболическое

NIG — нормально-обратное гауссовское

VG — Variance-Gamma

t — t-распределение Стьюдента

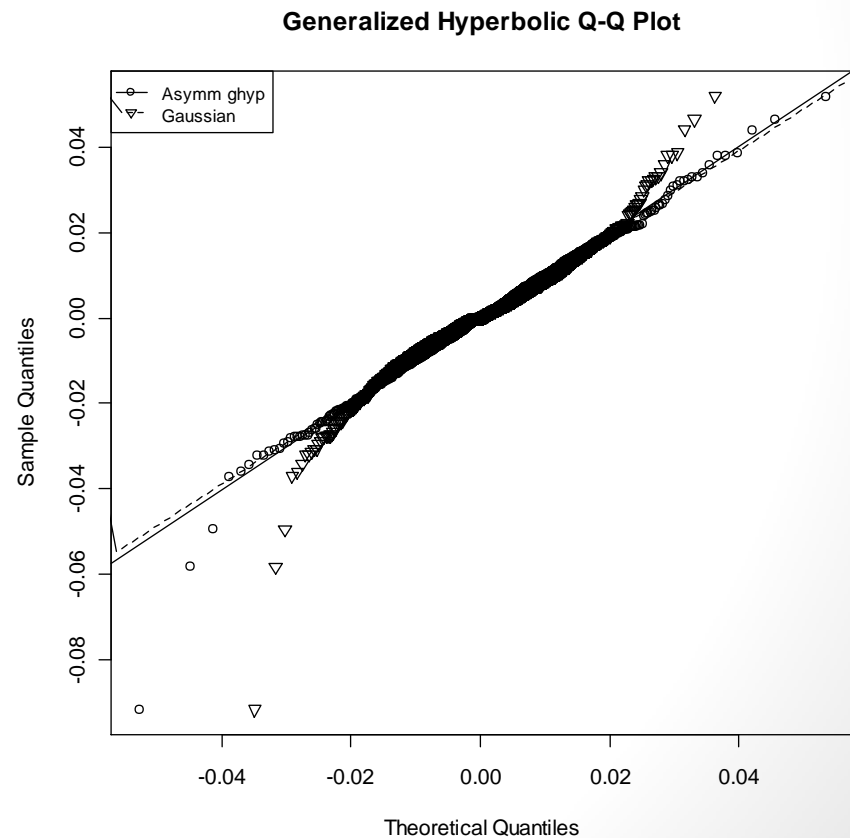
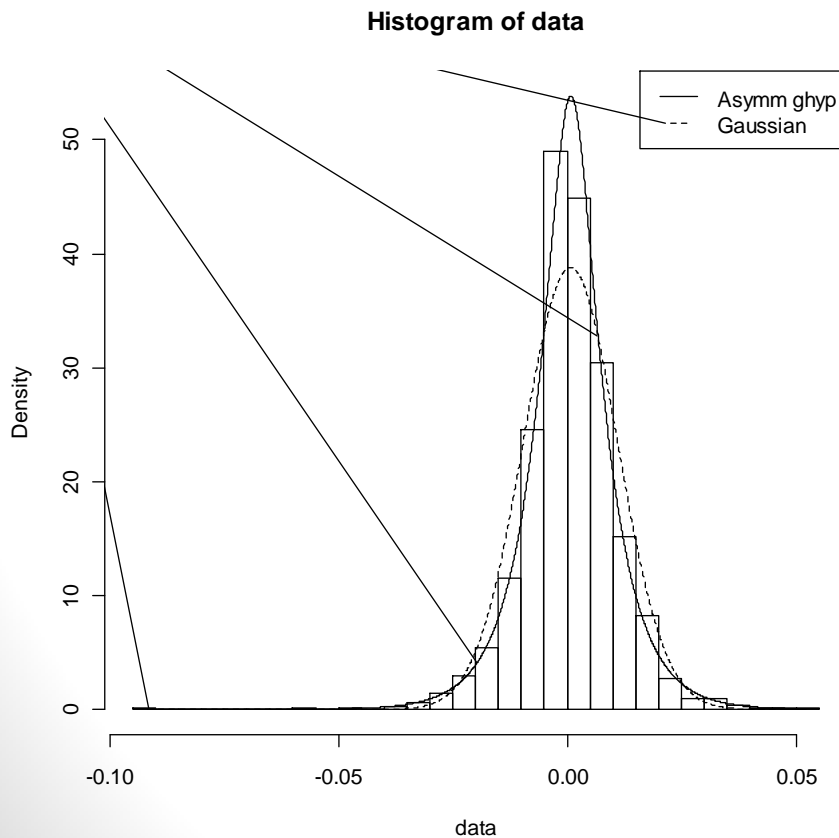
gauss — нормальное

Графический анализ модели

```
dax.ghyp <- fit.ghypuv(dax,symmetric=FALSE,silent=TRUE)
```

```
hist(dax.ghyp) # гистограмма
```

```
qqghyp(dax.ghyp) # график квантиль-квантиль
```



Выбор наилучшей модели

Отношение правдоподобия

H_0 : более общая модель обладает той же объясняющей силой, что и её частный случай

$$LR = -2 \ln \frac{L_{H_0}}{L_{H_{alt}}} \sim \chi^2(\nu), \quad \nu = df_{H_0} - df_{H_{alt}}$$

```
dax.t <- fit.tuv(dax, symmetric=FALSE, silent=TRUE)
lik.ratio.test(dax.ghyp, dax.t, conf.level=0.95)
```

Информационный критерий Акаике

$AIC = 2k - 2 \ln(L) \rightarrow \min$, k — количество параметров модели

```
aic.uv <- stepAIC.ghyp(dax, dist=c("gauss", "t", "ghyp"),
  symmetric=NULL, silent=TRUE)
summary(aic.uv$best.model) # статистики по модели
```

Оценка финансового риска

Меры риска:

- граница потерь (Value-at-Risk)

$$P(x < VaR_{1-\alpha}) = \alpha$$

- ожидаемые потери (Expected Shortfall)

$$ES_{1-\alpha} = E(x | x < VaR_{1-\alpha})$$

Метод Монте-Карло

```
alpha <- 0.1; N <- 10^6
```

```
dax.sim <- rghyp(n=N,object=aic.uv$best.model)
```

```
dax.sim <- sort(dax.sim)
```

```
VaR <- dax.sim[alpha*N]
```

```
# другой вариант: VaR <- qghyp(alpha,object=aic.uv$best.model)
```

```
ES <- mean(dax.sim[1:(alpha*N-1)])
```

VaR	-0.011
ES	-0.018

Кривая VaR

Используется для тестирования качества оценок риска

Кривая VaR — набор последовательных во времени значений VaR

разделим выборку на обучающую и экзаменующую

```
T1 <- 6*260; T2 <- T - T1
```

на пространстве экзаменующей выборки построим набор

последовательных значений VaR

```
VaR <- numeric()
```

```
h <- 0.5 * 260 # длина обучающей выборки
```

```
for (i in (T1+1):(T1+T2)) {  
  h.dax <- dax[(i-h):(i-1)]  
  dax.fit <- stepAIC.ghyp(h.dax,dist=c("gauss","t","ghyp"),  
    symmetric=NULL,silent=TRUE)  
  VaR[i-T1] <- qghyp(alpha,object=dax.fit$best.model)  
}
```

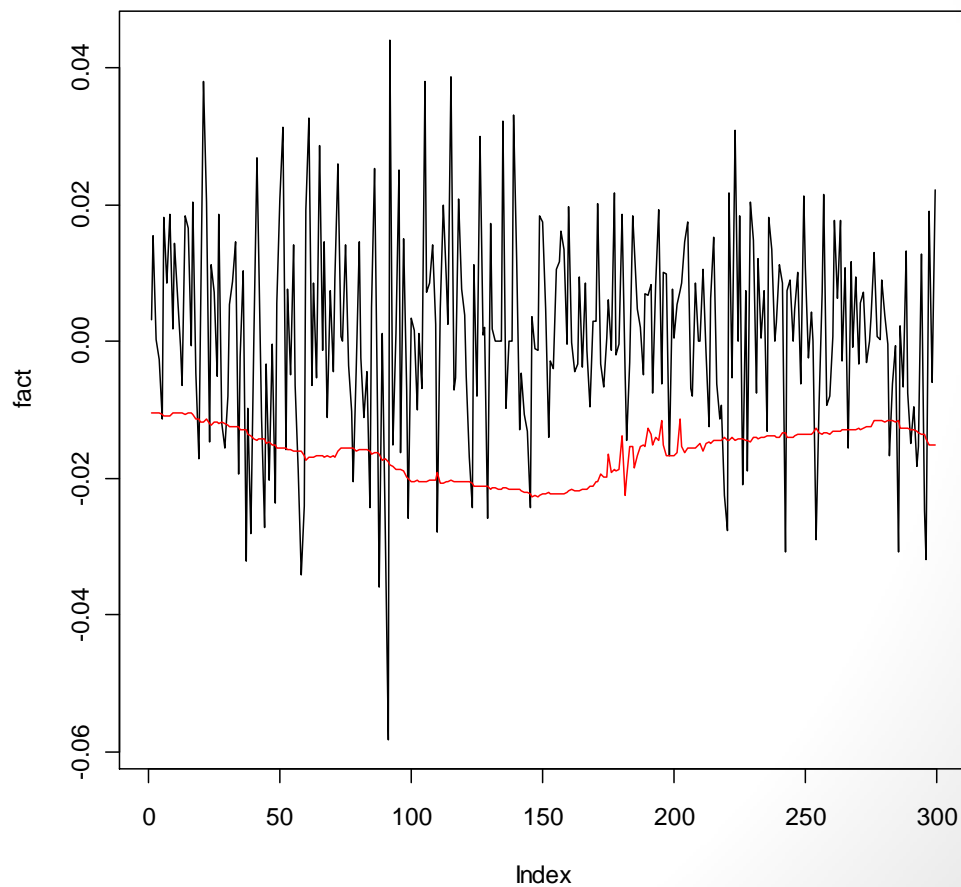
Кривая VaR

сравнение оценок риска с фактом

```
fact <- dax[(T1+1):(T1+T2)]
```

```
plot(fact,type="l")
```

```
lines(VaR,col="red")
```



Кривая VaR

Тест Купика

Идея состоит в сравнении модельной и эмпирической частот превышений фактическими убытками границы VaR

$$K = \sum I(x_t < VaR_t), \quad \alpha_0 = \frac{K}{T_2}$$

$$H_0: \alpha_0 = \alpha$$

Статистика:

$$S = -2 \ln((1 - \alpha)^{T_2 - K} \alpha^K) + 2 \ln((1 - \alpha_0)^{T_2 - K} \alpha_0^K) \sim \chi^2(1)$$

тест Купика в R:

```
K <- sum(fact<VaR); alpha0 <- K/T2  
S <- -2*log((1-alpha)^(T2-K)*alpha^K) +  
2*log((1-alpha0)^(T2-K)*alpha0^K)  
p.value <- 1-pchisq(S,df=1)
```

alpha	0.100
alpha0	0.130
p.value	0.092

Кривая VaR

Функции потерь

Величина функции потерь измеряет глубину пробоев кривой VaR и интерпретируется как размер понесённых потерь

Функция потерь Лопеса:

$$L_{Lo} = \frac{1}{K} \sum_{t=T_1+1}^{T_2} ((x_t - VaR_t)^2 \cdot I(x_t < VaR_t))$$

Функция потерь Бланко-Ила:

$$L_{BI} = \frac{1}{K} \sum_{t=T_1+1}^{T_2} \left(\frac{x_t - VaR_t}{VaR_t} \cdot I(x_t < VaR_t) \right)$$

функции потерь в R:

```
L.Lo <- sum((fact-VaR)^2*(fact<VaR))/K
```

```
L.BI <- sum((fact-VaR)/VaR*(fact<VaR))/K
```

L.Lo*10^4	1.399
L.BI	0.611

Моделирование хвостов распределения с помощью теории экстремальных значений

1. 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

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

тогда $F \in MDA(H)$

Generalized Extreme Value distribution (GEV)

Теорема Фишера-Типпетта-Гнеденко

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

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

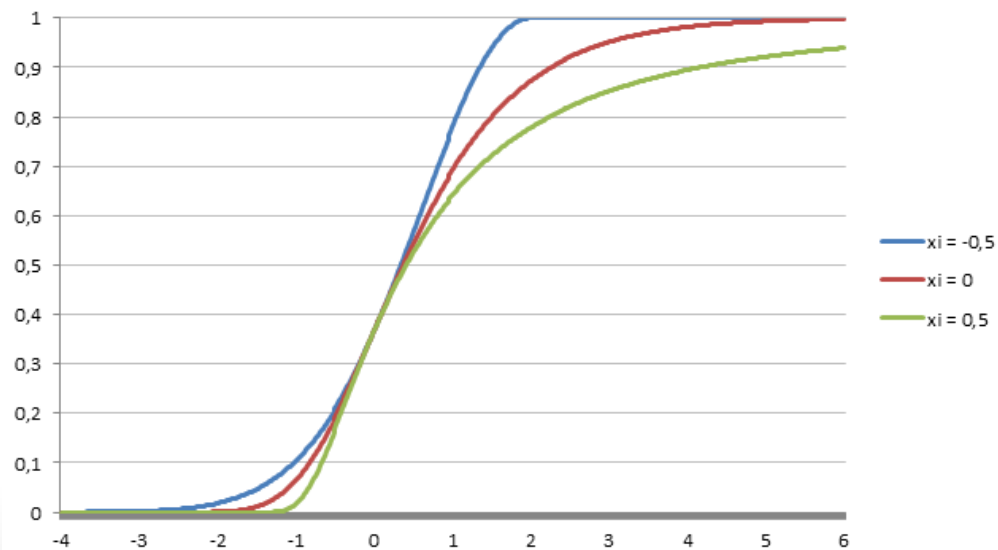
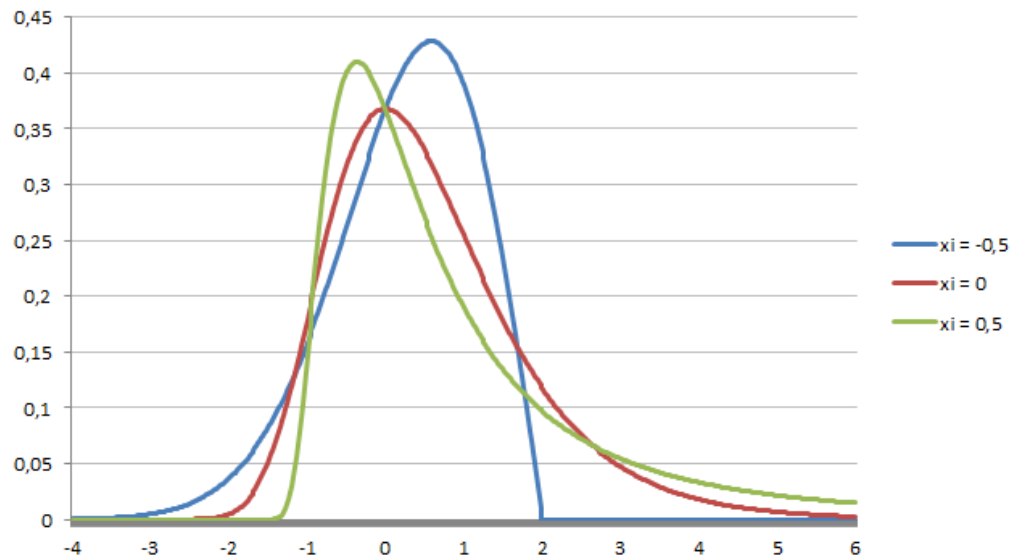
Частные случаи GEV:

- $\xi > 0$ — распределение Фрешè
- $\xi = 0$ — 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



Оценка параметров GEV

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

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

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

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

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

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

Оценка параметров GEV в R

Практический пример 1. Биржевой индекс DAX

загрузка данных

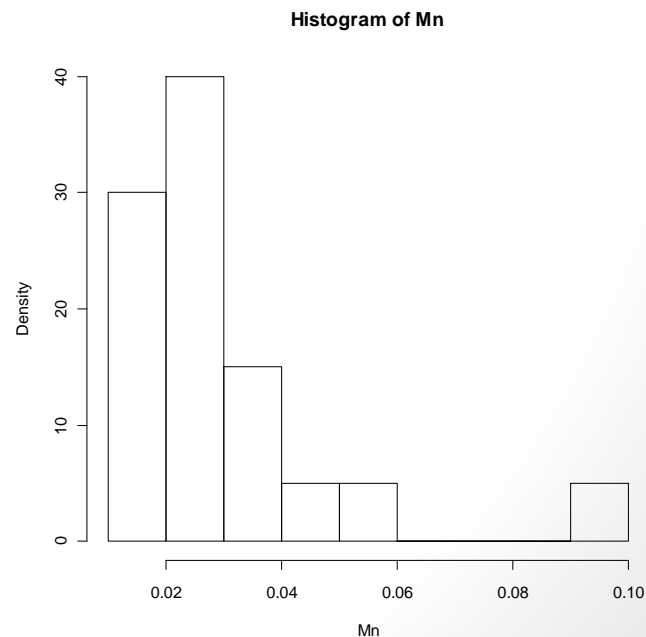
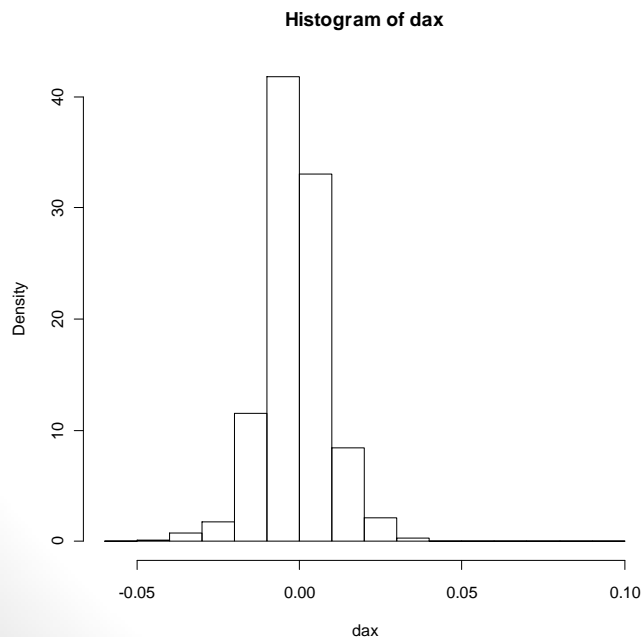
```
dax.loss <- -dax
```

расчёт максим

```
Mn <- numeric(0)
```

```
for (i in 1:m)
```

```
Mn[i] <- max(dax.loss[((i-1)*n+1):(i*n)])
```

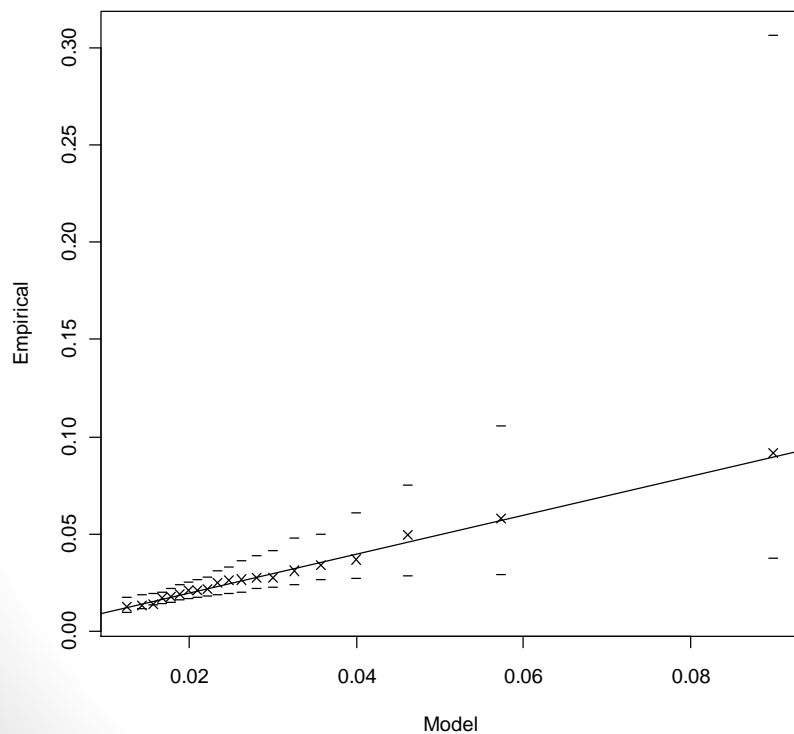


Оценка параметров GEV в R

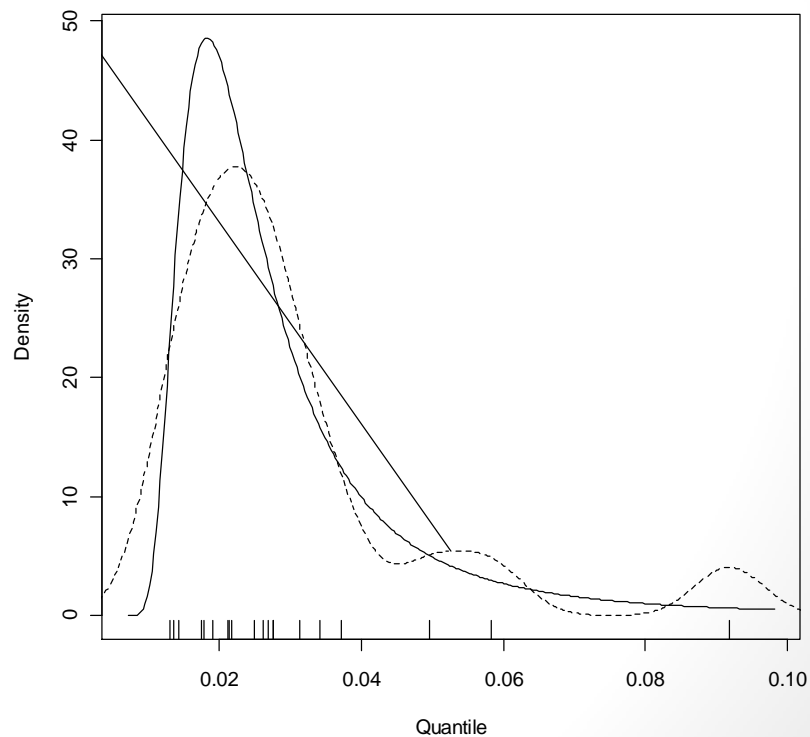
распределение максимум на основе GEV

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

Quantile Plot



Density Plot



Пороговый уровень и средний период наступления события

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

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

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

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

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))
```

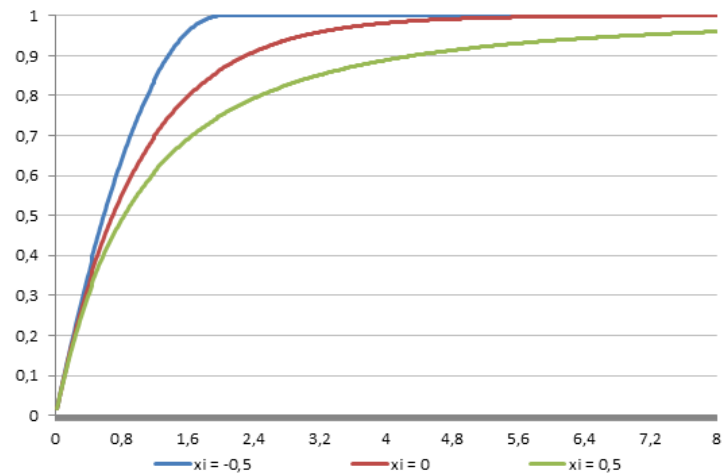
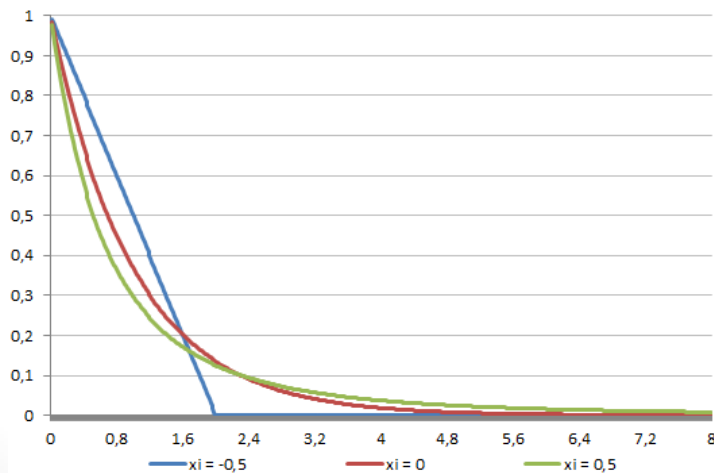
$r_{n,k}$	0.034
$k_{n,u}$	40.14

2. Generalized Pareto distribution (GPD)

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

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

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

$e(u) = E(x_i - u | x_i > u)$ — 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

Теорема Пикандса-Балкема-де Хаана

$$\exists \beta(u): \lim_{u \rightarrow x_F} \sup |F_u(x) - G_{\xi, \beta(u)}(x)| = 0, \quad 0 \leq x < x_F - u \Leftrightarrow$$

$$F \in MDA(H_\xi), \quad \xi \in R,$$

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

Моделирование превышений

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

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

Если $\vec{x} \sim iid$, то

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

Превышение более высокого порога:

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

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

$e(v)$ — 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:

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

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

$$ES_\alpha = VaR_\alpha + e(VaR_\alpha)$$

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

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

GPD в R

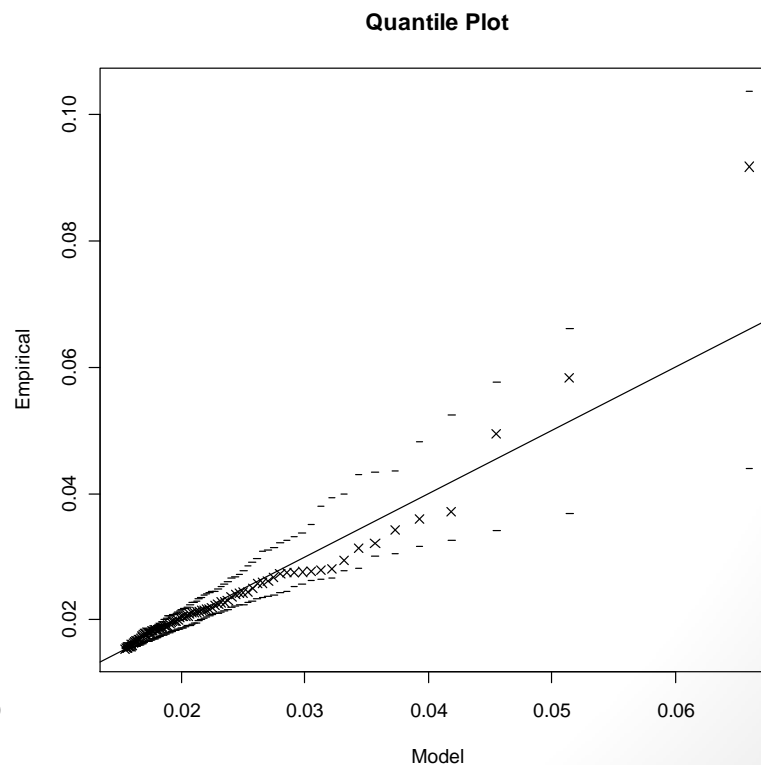
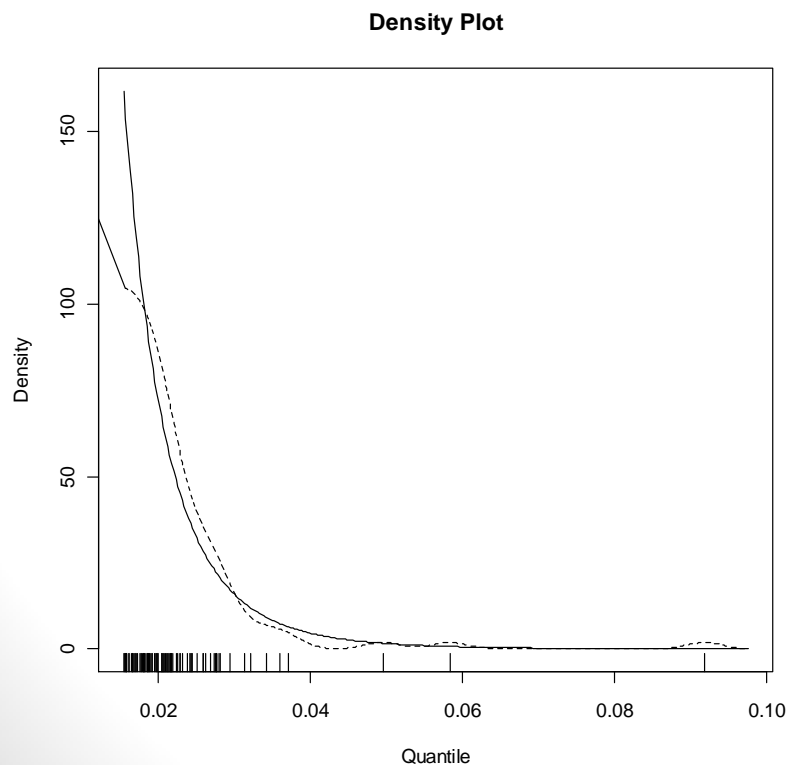
пороговое значение - 95% квантиль

```
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

Непараметрическое моделирование

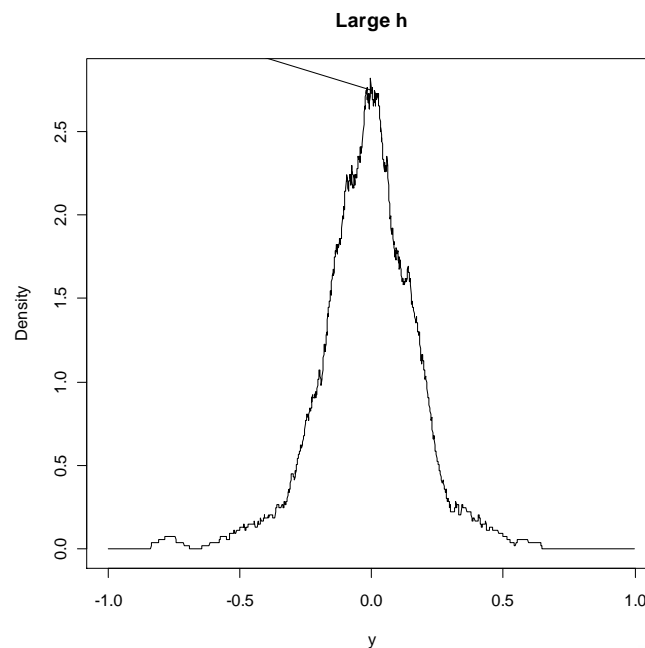
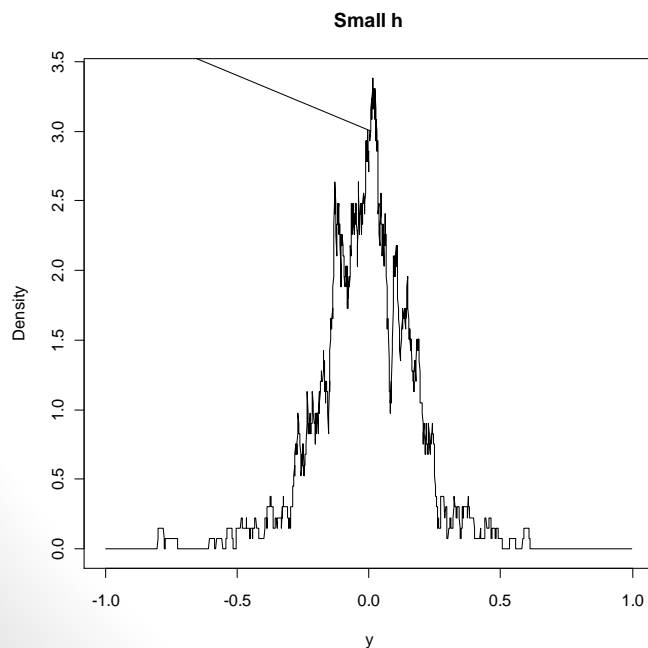
Простая непараметрическая оценка

Принцип построения простой (naïve) непараметрической оценки плотности в точке y состоит в подсчёте количества наблюдений, находящихся вблизи неё:

$$\hat{f}(y) = \frac{1}{nh} \sum_{i=1}^n I \left(y - \frac{h}{2} < y_i < y + \frac{h}{2} \right) \quad (3),$$

где h — длина интервала

Большие значения h дают более гладкие оценки:



Ядерная оценка

Простая оценка нигде не дифференцируема. Чтобы понять это перепишем формулу (3) в следующем виде:

$$\hat{f}(y) = \frac{1}{nh} \sum_{i=1}^n w\left(\frac{y-y_i}{h}\right), \text{ где } w(x) = I\left(|x| < \frac{1}{2}\right) \quad (4)$$

Проблема заключается в функции $w(x)$, которая придаёт наблюдениям дискретные веса (0 или 1)

Проблема решается с помощью замены функции $w(x)$ на ядерную функцию $K(x)$ с плавно изменяющимися весами:

$$\hat{f}(y) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{y-y_i}{h}\right) \quad (5)$$

Для того, чтобы оценка $\hat{f}(y)$ была функцией плотности, ядро должно удовлетворять условию $\int_{-\infty}^{+\infty} K(x)dx = 1$

Любая функция плотности удовлетворяет этому условию

Ядерные функции

В качестве ядерных функций обычно используются симметричные одномодальные функции плотности

Наиболее часто используемые на практике ядра:

$$K_G(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) \quad (6) \quad \text{— гауссовское ядро}$$

$$K_E(x) = \frac{3}{4\sqrt{5}} \left(1 - \frac{x^2}{5}\right) \cdot I(|x| < \sqrt{5}) \quad (7) \quad \text{— ядро Епанечникова}$$

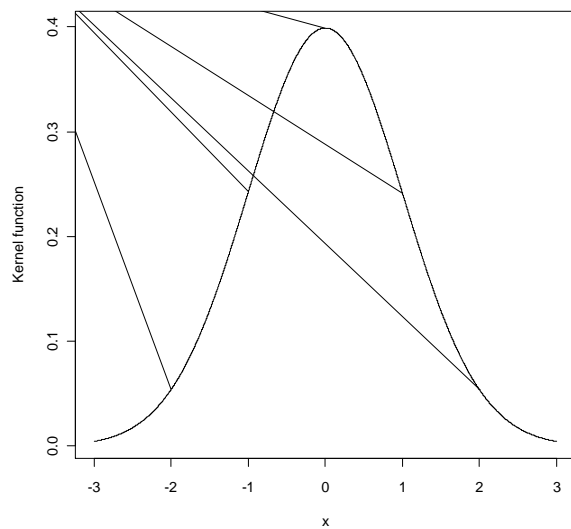
$$K_T(x) = (1 - |x|) \cdot I(|x| < 1) \quad (8) \quad \text{— треугольное ядро}$$

$$K_U(x) = \frac{1}{2} I(|x| < 1) \quad (9) \quad \text{— прямоугольное (равномерное) ядро}$$

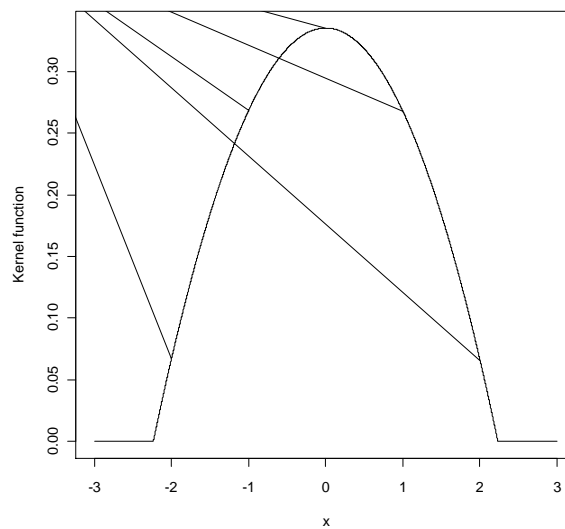
Вид этих функций представлен на следующем слайде

Ядерные функции

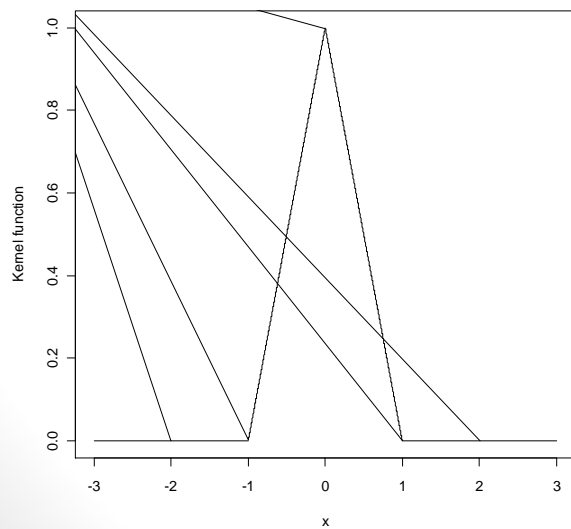
Gaussian kernel



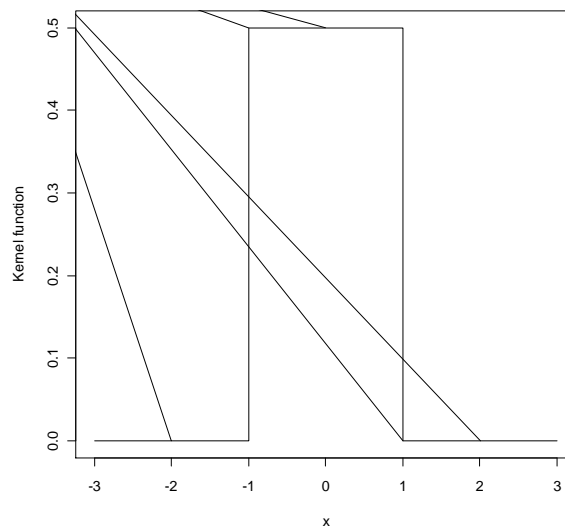
Epanechnikov kernel



Triangular kernel

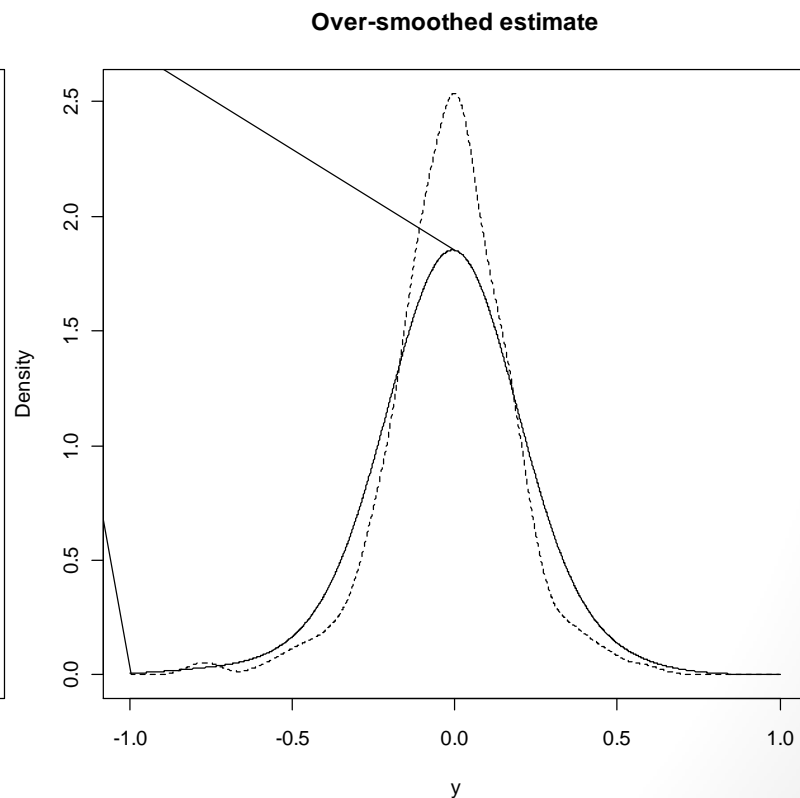
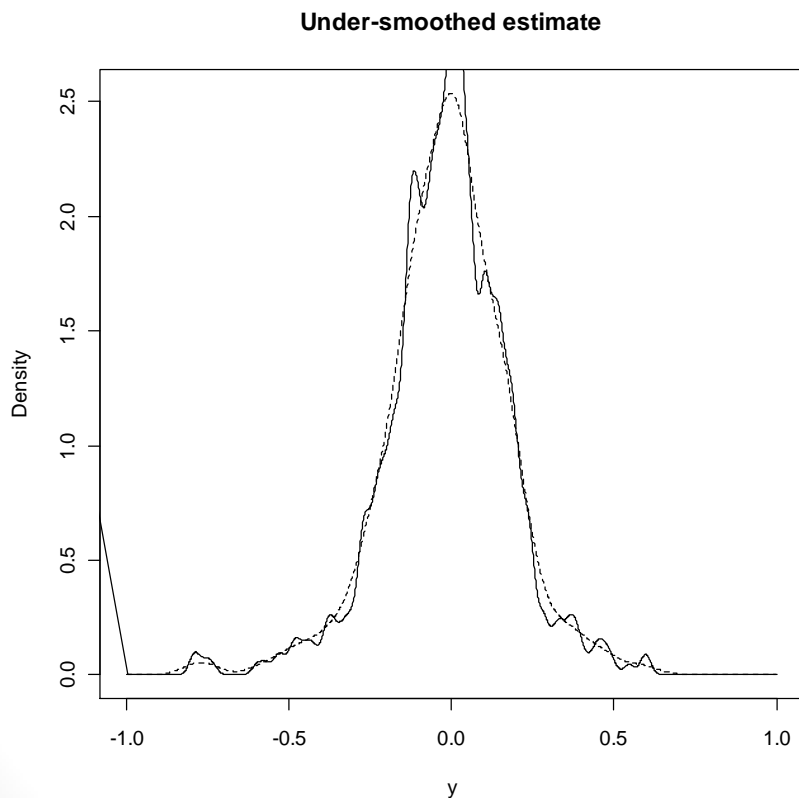


Uniform kernel



Влияние ширины интервала

Тогда как выбор ядра оказывает незначительное влияние на оценку плотности, выбор ширины интервала имеет решающее значение



Выбор ширины интервала

Существует два основных подхода к определению величины сглаживающего множителя (ширины интервала):

1. Фиксированная ширина интервала на всей выборке. В рамках этого подхода выделяют:
 - правило подстановки (rule of thumb);
 - метод перекрёстной проверки (cross-validation)
2. Ширина интервала меняется в зависимости от локальной концентрации наблюдений. Методы:
 - обобщённый метод ближайших соседей (generalized nearest neighbors);
 - адаптивный метод (adaptive nearest neighbors)

1. Фиксированная ширина интервала

Среднеквадратичная ошибка

Выбирать величину h следует так, чтобы оценка была как можно ближе к истинной плотности распределения, т.е. минимизировать разницу между $\hat{f}(y)$ и $f(y)$

Наиболее естественным кандидатом на эту разницу является среднеквадратичная ошибка (Mean Squared Error, MSE), рассчитываемая в конкретной точке y :

$$MSE(h, y) = E \left(\left(\hat{f}(y) - f(y) \right)^2 \right) \quad (10)$$

Распишем выражение (10) подробнее:

$$\begin{aligned} MSE(h, y) &= E \left(\left(\hat{f}(y) - f(y) \right)^2 \right) = E(\hat{f}^2(y)) - 2f(y)E(\hat{f}(y)) + f^2(y) \\ &= \left[E(\hat{f}^2(y)) - E^2(\hat{f}(y)) \right] + \left[E^2(\hat{f}(y)) - 2f(y)E(\hat{f}(y)) + f^2(y) \right] = \\ &= \text{var}(\hat{f}(y)) + \left(E(\hat{f}(y)) - f(y) \right)^2 \quad (11) \end{aligned}$$

Дисперсия и смещение оценки

$$MSE(h, y) = var(\hat{f}(y)) + \left(E(\hat{f}(y)) - f(y)\right)^2 \quad (11)$$

Первое слагаемое выражения (11) соответствует дисперсии оценки, второе — квадрату её смещения

Если ширина интервала слишком большая, то оценка оказывается пересглаженной, и растёт смещение

Если значение h слишком маленькое, то это увеличивает дисперсию

Минимальное смещение достигается при максимальной дисперсии ($h = 0$), а минимальная дисперсия — при максимальном смещении ($h \rightarrow +\infty$)

Нужно искать компромисс

Интегральная среднеквадратичная ошибка

Поскольку мы заинтересованы в минимизации отклонения между оценкой $\hat{f}(y)$ и плотностью $f(y)$ не только в конкретной точке y , рассмотрим интегральную среднеквадратичную ошибку (Mean Integrated Squared Error, MISE):

$$MISE(h) = E \left(\int_{-\infty}^{+\infty} \left(\hat{f}(y) - f(y) \right)^2 dy \right) \quad (12)^1$$

Мы можем переписать это так:

$$MISE(h) = \int E \left(\left(\hat{f}(y) - f(y) \right)^2 \right) dy = \int MSE(h, y) dy \quad (13), —$$

или в следующем виде:

$$MISE(h) = \int var \left(\hat{f}(y) \right) dy + \int \left(E \left(\hat{f}(y) \right) - f(y) \right)^2 dy \quad (14)$$

¹ Далее вместо определённого интеграла по всей числовой оси будет использоваться неопределённый

Оптимальная ширина интервала

Минимизируя аппроксимацию к критерию $MISE$, обозначаемую $AMISE$, можно найти оптимальное значение параметра сглаживания:

$$h_{opt} = \left(\int x^2 K(x) dx \right)^{-\frac{2}{5}} \left(\int K^2(x) dx \right)^{\frac{1}{5}} \left(\int f''^2(y) dy \right)^{-\frac{1}{5}} n^{-\frac{1}{5}} \quad (15)$$

Замечания к формуле (15):

- h_{opt} стремится к нулю по мере роста объема выборки, но сравнительно медленно (по степенному закону);
- h_{opt} уменьшается, если $f(y)$ сильно варьируется, и возрастает, если функция плотности варьируется слабо;
- наиболее подходящее ядро $K(x)$ можно определить, исходя из значения критерия $MISE$ (14)

Методы оценки оптимальной ширины интервала

В выражении для h_{opt} (15) остаётся неопределённость, связанная с незнанием истинной функции плотности $f(y)$

Мы рассмотрим два способа преодоления этой неопределённости:

1. Правило подстановки (Rule of Thumb);
2. Метод перекрёстной проверки (Cross-Validation)

Правило подстановки

Вместо $f(y)$ в выражение для оптимального интервала (15) подставляется какое-либо известное распределение

Если подставить нормальное распределение $N(\mu, \sigma^2)$ и использовать гауссовское ядро, то получим:

$$\hat{h}_{opt} \approx 1.059\sigma n^{-\frac{1}{5}} \quad (16)$$

В качестве оценки σ можно использовать выборочное

стандартное отклонение, $\hat{\sigma} = \sqrt{\frac{1}{n-1} \sum_{i=1}^n \left(y_i - \frac{1}{n} \sum_{j=1}^n y_j \right)^2}$,

или межквартильное расстояние, $\frac{\hat{q}_3 - \hat{q}_1}{1.349}$, где \hat{q}_i —

выборочное значение i -го квартиля, 1.349 —

межквартильное расстояние для стандартного нормального распределения

Модифицированное правило подстановки

Правило подстановки хорошо работает тогда, когда истинный закон распределения близок к подставляемому

Существует также модифицированное правило подстановки:

$$\hat{h}_{opt} = 0.9 \min \left(\hat{\sigma}, \frac{\hat{q}_3 - \hat{q}_1}{1.349} \right) n^{-\frac{1}{5}} \quad (17)$$

Модифицированное правило является более устойчивым к отклонениям истинного распределения от нормального закона

Метод перекрёстной проверки

Мы опишем вариацию метода, основанную на наименьших квадратах

Идея состоит в рассмотрении интегральной квадратической ошибки (Integrated Squared Error, ISE), аналогичной критерию *MISE* (12), но без математического ожидания:

$$ISE(h) = \int \left(\hat{f}(y) - f(y) \right)^2 dy = \\ \int \hat{f}^2(y) dy - 2 \int \hat{f}(y) f(y) dy + \int f^2(y) dy \quad (18)$$

Последнее слагаемое не зависит от h не играет роли в оптимизации

Величина $\int \hat{f}(y) f(y) dy$ есть матожидание оценки, которое приближённо равно $E \left(\hat{f}(y) \right) \approx \frac{1}{n} \sum_{i=1}^n \hat{f}_{-i}(y_i)$, где $\hat{f}_{-i}(y_i)$ — оценка плотности по всем наблюдениям, кроме y_i

Метод перекрёстной проверки

Таким образом, оптимизационная задача сводится к минимизации выражения

$$CV(h) = \int \hat{f}^2(y) dy - \frac{2}{n} \sum_{i=1}^n \hat{f}_{-i}(y_i) \quad (19)$$

Достоинства методов с фиксированной шириной интервала:

- простота вычислений;
- интуитивная понятность;
- оценки обладают известными статистическими свойствами

Недостатки:

- пересглаженный центр распределения;
- недосглаженные и тонкие хвосты

2. Адаптивные методы

Метод ближайших соседей

Идея состоит в таком определении интервала, чтобы в него всегда попадало фиксированное количество наблюдений k

Рассмотрим расстояние от i -го наблюдения в выборке до некоторой точки y :

$$d_i(y) = |y_i - y| \quad (20)$$

Отсортируем эти расстояния по возрастанию так, что

$$d_1(y) \leq d_2(y) \leq \dots \leq d_n(y) \quad (21)$$

k ближайших к точке y наблюдений находятся на расстоянии, не превышающем $d_k(y)$

Иными словами, отрезок $[y - d_k(y); y + d_k(y)]$ содержит k наблюдений из выборки

Оценка плотности

Положив $h = 2d_k(y)$, мы можем подставить эту величину в простую оценку плотности (3), которая примет вид

$$\hat{f}(y) = \frac{k}{2nd_k(y)} \quad (22)$$

В каждой отдельной точке y для любого значения k найдётся такое значение h , что оценки (3) и (22) дадут один и тот же результат, однако, рассматриваемая в целом, оценка по методу ближайших соседей отличается от простой оценки

Так же, как и в случае с простой оценкой, мы можем прибегнуть к помощи ядерных функций и получить обобщённую оценку по методу ближайших соседей:

$$\hat{f}(y) = \frac{1}{2nd_k} \sum_{i=1}^n K\left(\frac{y-y_i}{2nd_k(y)}\right) \quad (23)$$

Задача выбора оптимального значения k решается численно, сравнением критериев ISE (18) для различных значений k

Достоинства и недостатки метода

Достоинства:

- решается проблема недосглаженности и тонкости хвостов

Недостатки:

- оценка $\hat{f}(y)$ не дифференцируема там, где не дифференцируема функция $d_k(y)$;
- хвосты распределения могут казаться тяжелее, чем на самом деле, потому что $d_k(y)$ растёт очень медленно (как функция первой степени);
- в общем случае оценка $\hat{f}(y)$ не является функцией плотности

Адаптивный метод ближайших соседей

Оценка строится в следующем виде:

$$\hat{f}(y) = \frac{1}{n} \sum_{i=1}^n \frac{K\left(\frac{y-y_i}{hd_k(y_i)}\right)}{hd_k(y_i)} \quad (24)$$

Сглаживающий параметр разделяется на две части:

1. глобальная (h);
2. локальная концентрация наблюдений ($d_k(y_i)$)

Эта оценка лишена недостатков обобщённого метода ближайших соседей

Величину h определяют путём построения пилотной оценки плотности $\tilde{f}(y)$ с фиксированной шириной интервала

Часто вместо $d_k(y_i)$ используют показатель

$$\lambda_i = \left(\frac{g}{\tilde{f}(y_i)}\right)^\alpha, \quad g = \left(\prod_{i=1}^n \tilde{f}(y_i)\right)^{\frac{1}{n}}, \quad \alpha \in [0; 1] \quad (25)$$

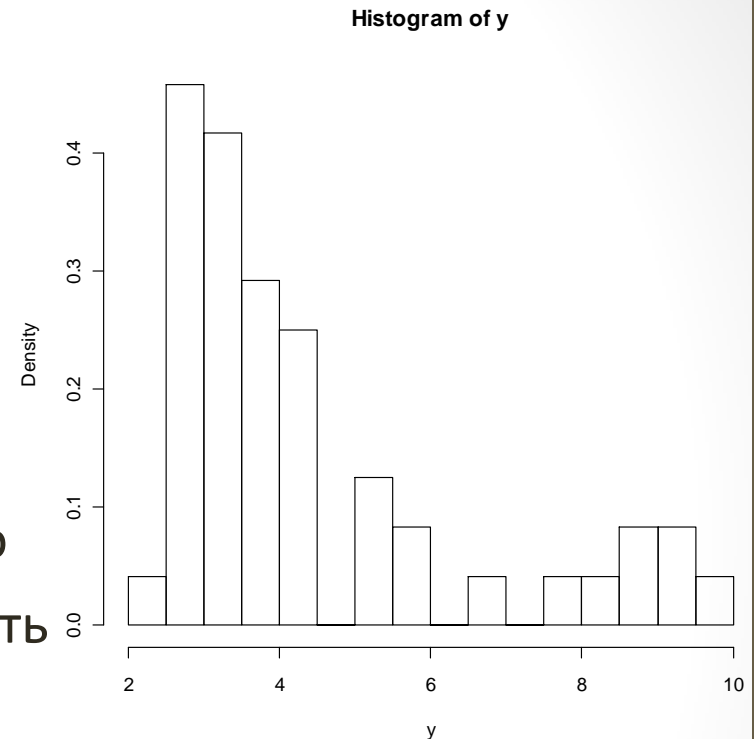
Пример 1. Острова

```
library(datasets)
y <- log(islands)
```

Построение гистограммы

```
hist(y, nclass=12, probability=TRUE)
```

- ***nclass*** определяет количество интервалов
- ***probability*** преобразует количество наблюдений в интервале в плотность распределения



Пример 1. Острова

Простая непараметрическая оценка плотности

```
L <- 10^4; N <- length(y)
```

```
h <- 2 # ширина интервала
```

в точках x будет оцениваться плотность

```
x <- seq(0,12,length=L) # последовательность 0 – 12 длиной L
```

```
f.naive <- numeric() # нулевой (пока) вектор оценок
```

считаем количество элементов в интервалах $x_i \pm h/2$

```
for (i in 1:L) f.naive[i] <- sum(1*((y>x[i]-h/2)&(y<x[i]+h/2)))
```

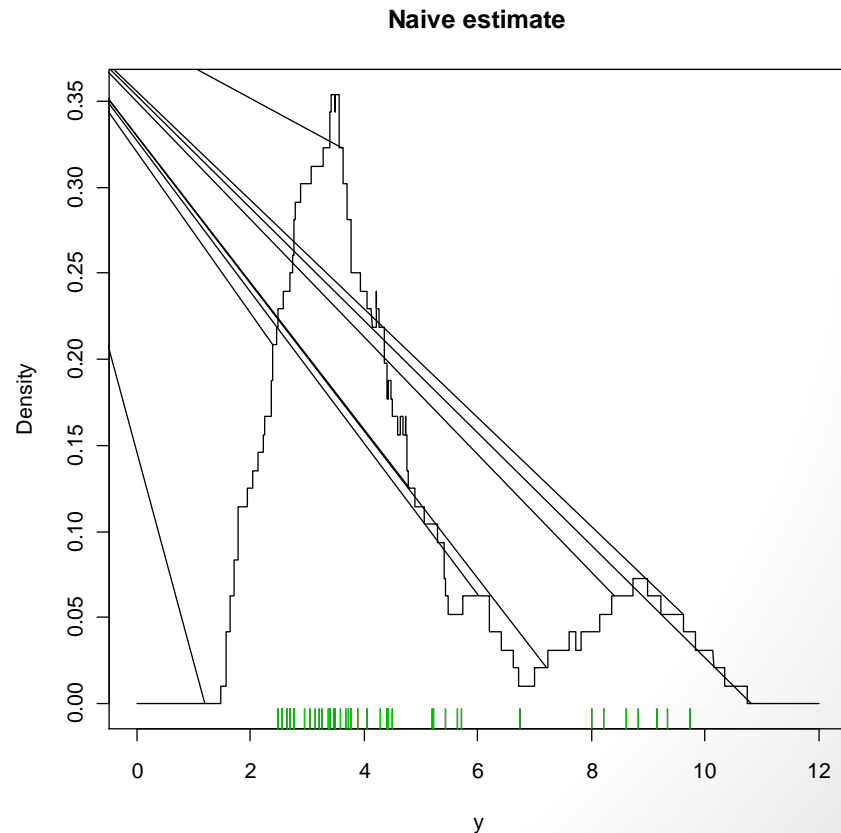
```
f.naive <- f.naive/(N*h) # нормируем оценку
```

Пример 1. Острова

График простой оценки

```
plot(x, f.naive, type="l", main="Naive estimate",  
xlab="y", ylab="Density")  
rug(y, col=3)
```

- **type** определяет вид графика
"l" — линии, "p" — точки, ...
- **main** — заголовок
- **xlab** — подпись на оси x
- **ylab** — подпись на оси y



Пример 1. Острова

Ядерные оценки

```
library(np)
f.fix <- npudens(tdat=y, edat=x,
  ckertype="gaussian", bwtype="fixed")
```

- ***tdat*** — обучающая выборка
- ***edat*** — точки, в которых рассчитывается оценка
- ***ckertype*** — вид ядерной функции
"gaussian", "epanechnikov", "uniform"
- ***bwtype*** определяет метод расчёта интервала h
"fixed", "generalized_nn", "adaptive_nn"
- ***f\$dens*** — искомые значения оценок

Пусть ***f.fix***, ***f.gen*** и ***f.ada*** — оценки плотности с фиксированным интервалом, по обобщённому методу ближайших соседей и по адаптивному методу ближайших соседей

Пример 1. Острова

Адаптивный метод с λ_i

```
pilot <- npudens(tdat=y, ckertype="gaussian", bwtype="fixed")
h <- pilot$bws$bw # оценка глобальной составляющей интервала

# среднегеометрическое пилотных оценок
g <- 1
for (i in 1:N) g <- g*pilot$dens[i]^(1/N)

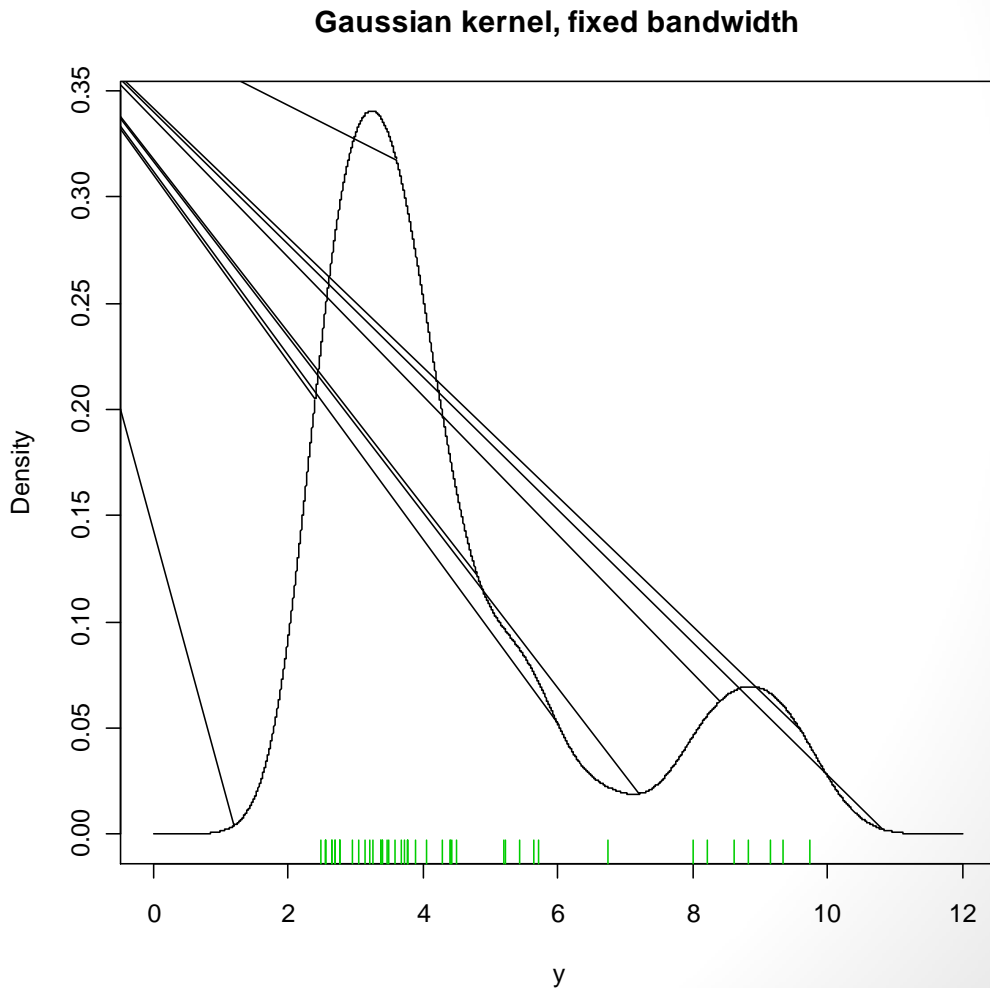
# расчёт локальной концентрации наблюдений
alpha <- 0.5
lambda <- (g/pilot$dens)^alpha

kern <- function(u) exp(-u^2/2)/sqrt(2*pi) # ядро Гаусса

# расчёт оценок плотности
f <- numeric(L)
for (i in 1:L) {
  f[i] <- sum(kern((x[i]-y)/(h*lambda)))/(h*lambda)
}
f <- f / N
```


Пример 1. Острова

```
plot(x, f.fix$dens, type="l",  
main="Gaussian kernel, fixed bandwidth",  
xlab="y", ylab="Density")
```



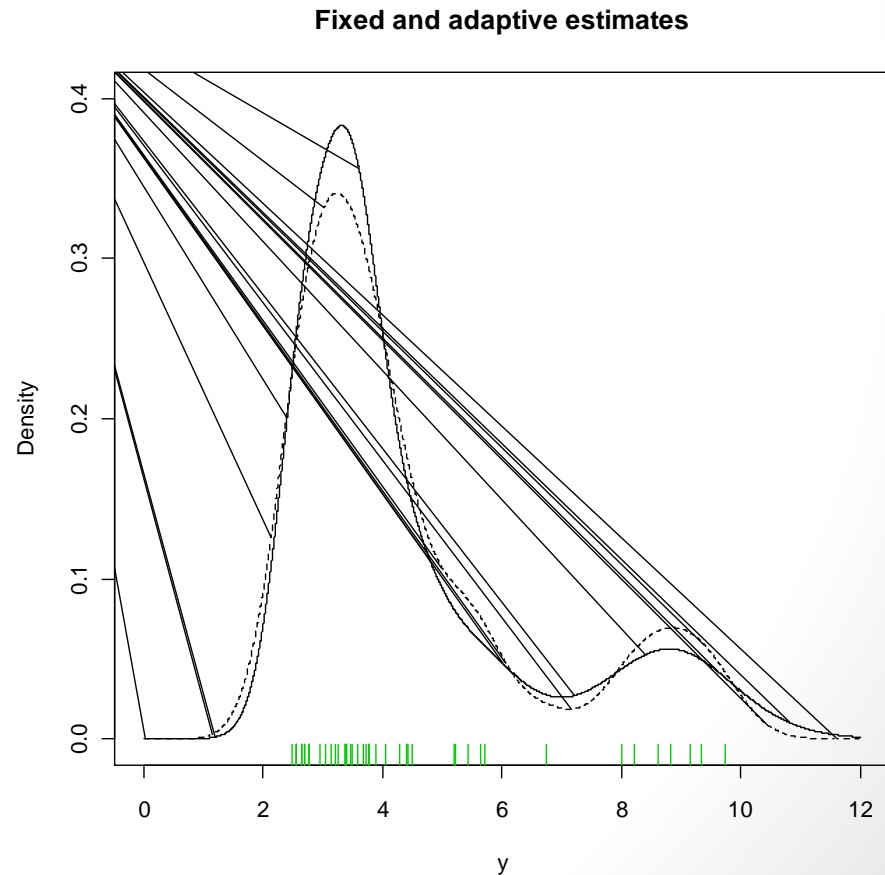
Пример 1. Острова

Сравнение адаптивной и фиксированной оценок

```
plot(x, f.fix$dens, type="l", lty="dashed", ylim=c(0, 0.4),  
main="Fixed and adaptive estimates",  
xlab="y", ylab="Density")
```

```
lines(x, f)
```

- ***lty*** — тип линии
"solid", "dashed", "dotted",
"dotdash", "longdash", ...
- ***ylim*** — границы по оси ординат
- ***lines*** — добавление кривых на существующий график



Пример 1. Острова

Значения логарифмической функции правдоподобия:

$\ln L = \sum_{i=1}^n \ln \hat{f}(y_i)$, — сумма логарифмов оценок в точках y_i

У нас есть оценки в точках x_j — последовательности 0 – 12

Пусть $dx = x_j - x_{j-1}$, найдём такой индекс j , что $x_j = y_i$:

$$x_1 + (j - 1)dx = y_i \Rightarrow j = \frac{y_i - x_1}{dx} + 1$$

```
dx <- x[2]-x[1]
```

```
llh.fix <- sum(log(f.fix$dens[round((y-x[1])/dx)+1]))
```

```
llh.ada <- sum(log(f[round((y-x[1])/dx)+1]))
```

```
llh.fix; llh.ada # вывод результатов на экран
```

llh.fix	-81.17
llh.ada	-81.29

Пример 1. Острова

Нахождение квантилей оценки распределения, $\hat{F}^{-1}(\alpha)$

оценка функции распределения

```
F.fix <- npudist(tdat=y, edat=x, ckertype="gaussian", bwtype="fixed")
```

для адаптивного варианта

```
F <- rep(0, times=L)
```

```
for (i in 1:L) F[i] <- sum(f[1:i])*dx
```

поиск квантиля методом деления пополам

```
alpha <- 0.99
```

```
a <- 1; b <- L; ab <- trunc((a+b)/2)
```

```
while ((b-a)>2) {
```

```
  if (F.fix$dist[ab]<=alpha) a <- ab
```

```
  if (F.fix$dist[ab]>=alpha) b <- ab
```

```
  ab <- trunc((a+b)/2)
```

```
}
```

```
q.fix <- x[ab]
```

q.fix	10.00
q.ada	10.49

Пример 1. Острова

Генератор случайных чисел

фиксированный интервал

```
M <- 10^6
```

```
y.fix.sim <- sample(x,prob=f.fix$dens,size=M,replace=TRUE)
```

```
q.fix <- sort(y.fix.sim)[alpha*M]
```

для адаптивного варианта

```
y.ada.sim <- sample(x,prob=f,size=M,replace=TRUE)
```

```
q.ada <- sort(y.ada.sim)[alpha*M]
```

q.fix	10.01
q.ada	10.46

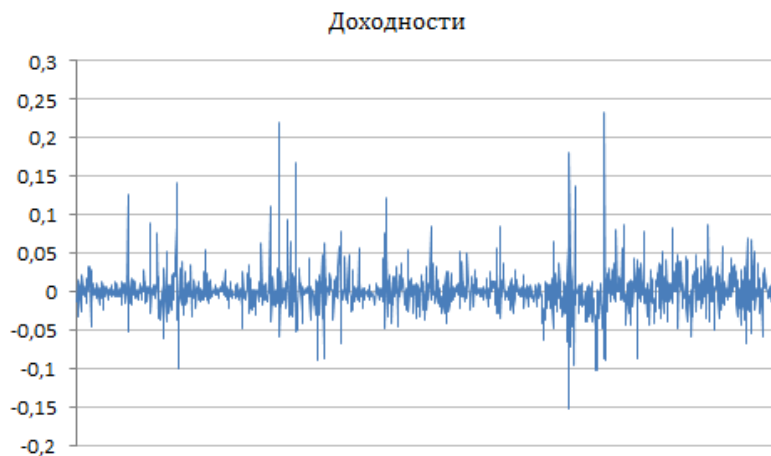
Моделирование волатильности финансовых активов с помощью GARCH

Моделирование средней доходности

Пусть y_t — доходность актива, тогда уравнение для средней доходности по модели ARMA(m,n) записывается так:

$$y_t = \mu + \sum_{i=1}^m a_i y_{t-i} + \sum_{j=1}^n b_j \varepsilon_{t-j} + \varepsilon_t = \mu + a(L)y_t + b(L)\varepsilon_t$$

Для финансовых временных рядов характерен эффект кластеризации волатильности



Возникает задача моделирования дисперсии доходности

Тест на ARCH-эффекты

Тест множителей Лагранжа (LM-тест)

Пусть $e_t = y_t - \hat{y}_t$. Рассмотрим регрессию:

$$e_t^2 = \delta_0 + \delta_1 e_{t-1}^2 + \dots + \delta_q e_{t-q}^2 + \varepsilon_t$$

$H_0: \delta_1 = \dots = \delta_q = 0$ (нет ARCH-эффектов)

$H_{alt}: \exists j \in \{1; \dots; q\} : \delta_j \neq 0$

Пусть $ESS_0 = \sum_{t=1+q}^T (e_t^2 - \overline{e^2})$, $\overline{e^2} = \frac{1}{T} \sum_{t=1}^T e_t^2$ и

$ESS_1 = \sum_{t=1+q}^T \hat{\varepsilon}_t^2$, тогда

$$S = \frac{(ESS_0 - ESS_1)/q}{ESS_1/(T-2q-1)} \sim_{H_0} \chi^2(q)$$

LM-тест в R

```
library(FinTS)  
ArchTest(dax, lags=12)
```

```
ARCH LM-test; Null hypothesis: no ARCH effects  
data:  dax  
Chi-squared = 85.4761, df = 12, p-value = 3.686e-13
```

Моделирование волатильности

- Уравнения для дисперсии по модели GARCH(p,q):

$$\varepsilon_t = z_t \sigma_t, \quad z_t \sim \text{idd}(0; 1)$$

$$\sigma_t^2 = \omega + \sum_{i=1}^p \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^q \beta_j \sigma_{t-j}^2 = \omega + \alpha(L) \varepsilon_t^2 + \beta(L) \sigma_t^2$$

Если $\forall i \beta_i = 0$, то GARCH(p,q) \sim ARCH(p,q)

- Степенное обобщение — модель APARCH(p,q):

$$\varepsilon_t = z_t \sigma_t, \quad z_t \sim \text{idd}(0; 1)$$

$$\sigma_t^\delta = \omega + \sum_{i=1}^p \alpha_i (|\varepsilon_{t-i}| - \gamma_i \varepsilon_{t-i})^\delta + \sum_{j=1}^q \beta_j \sigma_{t-j}^\delta,$$

где $\delta > 0$, $-1 < \gamma_i < 1$

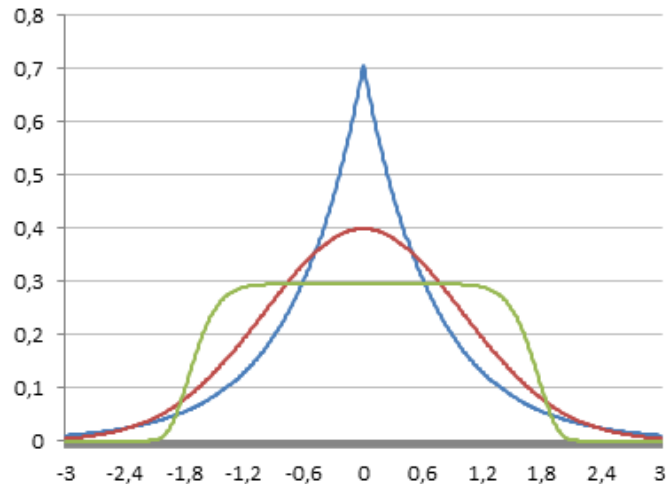
Если $\forall i \gamma_i = 0$ и $\delta = 2$, то APARCH(p,q) \sim GARCH(p,q)

Generalized Error Distribution (GED)

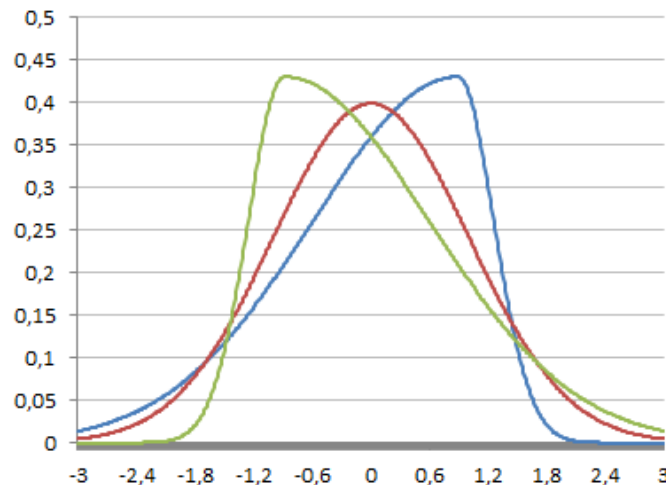
$$f_{GED}(x; \mu, \sigma, \nu) = \frac{\frac{1}{\nu} \exp\left(-\frac{1}{2} \left| \frac{x - \mu}{\sigma} \right|^\nu\right)}{2^{\frac{1}{\nu} + 1} \sigma \cdot \Gamma\left(\frac{1}{\nu} + 1\right)}$$

$\Gamma(.)$ – гамма-функция

Формы GED в зависимости от параметров



Параметр ν определяет
общую форму
распределения



Существует также
асимметричный вариант GED
с дополнительным
параметром $\xi > 0$

Общая схема расчёта модели APARCH

```
library(fGarch)
```

оценка параметров модели

```
dax.gfit <- garchFit(formula=~arma(m,n)+aparch(p,q), data=dax,  
cond.dist=[...], include.delta=T/F, leverage=T/F, trace=FALSE)
```

вместо [...] следует подставить распределение z_t :

“norm”, “snorm”, “std”, “sstd”, “ged”, “sged” или другие

используя комбинации степенного параметра и рычага

можно получать различные частные случаи модели

Расчёт частных случаев модели APARCH

GARCH(1,1)

$$\sigma_t^2 = \omega + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2$$

```
garchFit(formula=~aparch(1,1), data=dax, delta=2,  
include.delta=FALSE, leverage=FALSE, trace=FALSE)
```

TS-GARCH(1,1)

$$\sigma_t = \omega + \alpha_1 |\varepsilon_{t-1}| + \beta_1 \sigma_{t-1}$$

```
garchFit(formula=~aparch(1,1), data=dax, delta=1,  
include.delta=FALSE, leverage=FALSE, trace=FALSE)
```

T-GARCH(1,1) (GJR-GARCH)

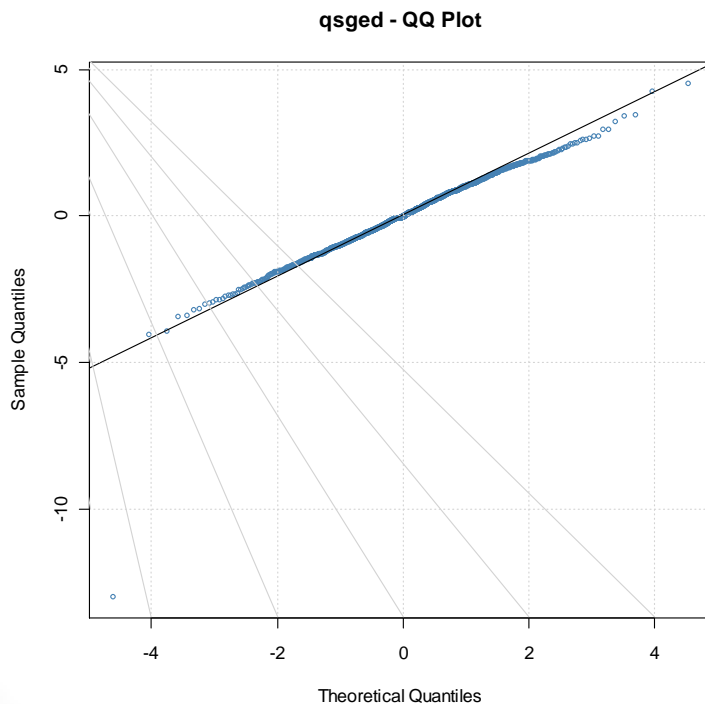
$$\sigma_t^2 = \omega + \alpha_1^* \varepsilon_{t-1}^2 + \gamma_1^* \cdot I(\varepsilon_{t-1} > 0) + \beta_1 \sigma_{t-1}^2$$

```
garchFit(formula=~aparch(1,1), data=dax, delta=2,  
include.delta=FALSE, leverage=TRUE, trace=FALSE)
```

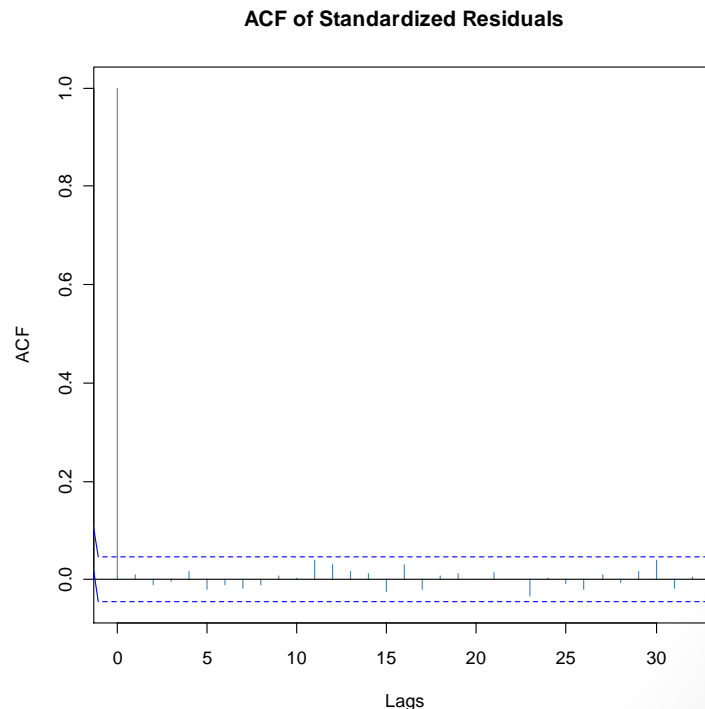
Графический анализ модели

```
dax.gfit <- garchFit(formula=~aparch(1,1),data=dax,delta=2,  
include.delta=FALSE,leverage=TRUE,cond.dist="sged",  
shape=1.25,include.shape=FALSE,trace=FALSE)
```

```
plot(dax.gfit,which=[...])
```



which=13



which=10

Тесты на единичный корень

Расширенный тест Дики-Фуллера (ADF)

$$\Delta y_t = \alpha + \beta t + \gamma y_{t-1} + \delta_1 \Delta y_{t-1} + \dots + \delta_{p-1} \Delta y_{t-p+1} + \varepsilon_t$$

$H_0: \gamma = 0$ (единичный корень)

$H_{alt}: \gamma < 0$

Тест Филлипса-Перрона (PP)

$$y_t = \delta y_{t-1} + \varepsilon_t$$

$H_0: |\delta| = 1$ (единичный корень)

$H_{alt}: |\delta| < 1$

Тест Квятковского-Филлипса-Шмидта-Шина (KPSS)

$$y_t = \alpha t + r_t + \varepsilon_t, \quad r_t \sim RW, \quad \varepsilon_t \sim I(0)$$

$$r_t = r_{t-1} + u_t, \quad u_t \sim iid(0; \sigma_u^2)$$

$H_0: \sigma_u^2 = 0$ (стационарность)

$H_{alt}: \sigma_u^2 > 0$

Тесты на единичный корень в R

```
library(tseries)
```

ADF-тест

```
adf.test(dax)
```

```
Dickey-Fuller = -11.1348, Lag order = 12, p-value = 0.01  
alternative hypothesis: stationary
```

PP-тест

```
pp.test(dax)
```

```
Dickey-Fuller Z(alpha) = -1759.696, Truncation lag parameter = 8,  
p-value = 0.01 alternative hypothesis: stationary
```

KPSS-тест

```
kpss.test(dax, null="Level")
```

```
KPSS Level = 0.4634, Truncation lag parameter = 9,  
p-value = 0.04991
```

Прогноз по модели ARMA-GARCH

прогноз среднего и дисперсии на i шагов вперёд

```
dax.frc <- predict(dax.gfit,n.ahead=i)
```

```
dax.frc[,1] # вектор средних
```

```
dax.frc[,3]^2 # вектор дисперсий
```

расчёт границы потерь

```
alpha <- 0.05
```

```
VaR <- dax.frc[1,1]+dax.frc[1,3]*qged(alpha,mean=0,sd=1,  
nu=dax.gfit@fit$par["shape"])
```

Домашнее задание

- рассчитать оценки риска для акции или биржевого индекса по всей совокупности наблюдений на основе моделей семейства GARCH, ОГР, а также с помощью инструментария ТЭЗ и непараметрического моделирования
- построить кривые VaR для указанных моделей и проверить качество оценок риска

Исходные данные — котировки с сайта finam.ru, finance.yahoo.com и др.