FEC 522:
Financial Econometrics II

Harald Schmidbauer

İSTANBUL BİLGİ ÜNİVERSİTESİ

About These Slides

- The present slides are not self-contained; they need to be explained and discussed.

- Even though being a "work in progress" and subject to revision, the slides constitute copyrighted material.
  If you want to reproduce or copy anything from the slides, please ask:

Harald Schmidbauer	**harald** at **hs-stat** dot **com**
Angi Rösch	**angi** at **angi-stat** dot **com**

- The slides were produced using LaTeX (www.latex-project.org) and R (the R project; www.R-project.org) on a GNU/Linux system.

- R files used for this course are available upon request.

Chapter 8:

Models for Return Distributions

8.1 Introduction

Return distributions.

- What are the characteristics of the distribution of returns?
- This question can be essential for financial decision making.
- Our goal in this chapter is to present reasonable stochastic models for returns on stock indices, stock prices, and commodity prices.

8.1 Introduction

The definition of return.

- Let

  P_t = closing price of a stock on day (week/month) t.

- A return for day (week/month) t can be defined as

$$R_t = \frac{P_t - P_{t-1}}{P_{t-1}} \qquad \text{or} \qquad r_t = \ln \frac{P_t}{P_{t-1}}.$$

- Using elementary calculus, it can be shown that $R_t \approx r_t$, provided that $P_t \approx P_{t-1}$.

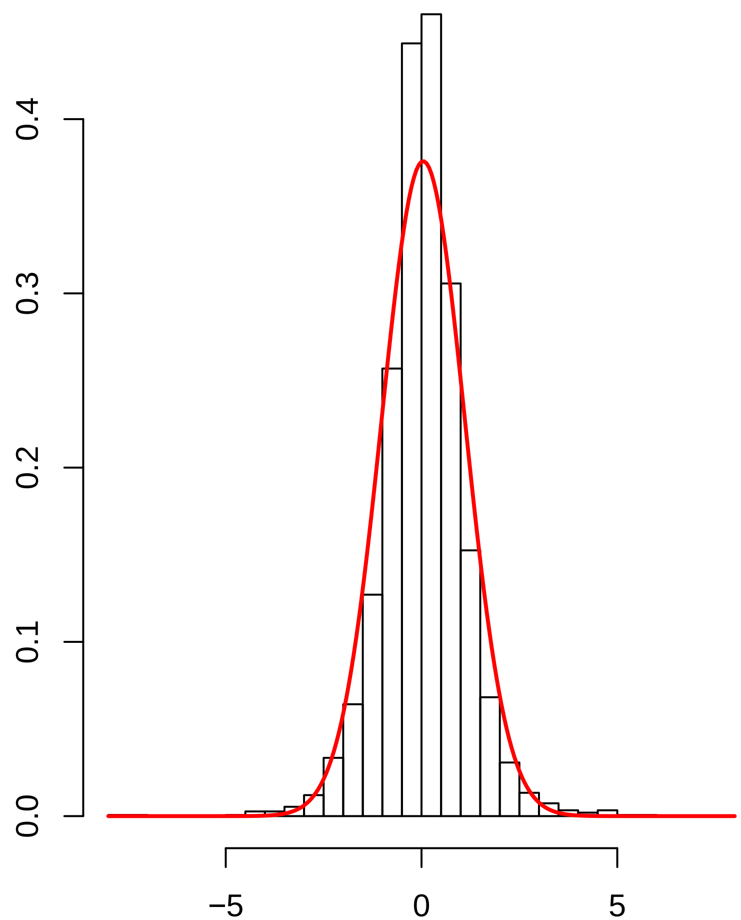
8.2 The Empirical Distribution of Returns

Empirical return distributions.

- We are now going to look at *empirical* distributions of returns on. . .
  - a stock index: Dow-Jones,
  - a stock: IBM,
  - a commodity: Brent crude oil.
- This will highlight the requirements on distributional models.

8.2 The Empirical Distribution of Returns

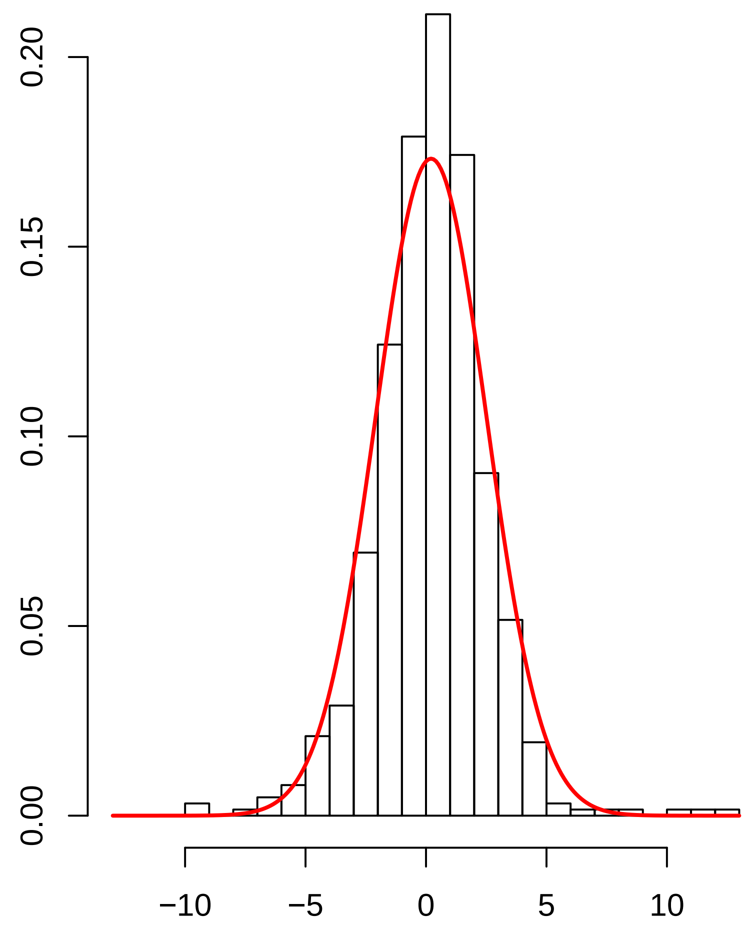
Dow-Jones: daily returns, Jan 1995 through Nov 2006.



first.day	1995-01-03
last.day	2006-11-30
observations	2990
NAs	118
mean.daily	0.04441
var.daily	1.12664
sd.daily	1.06143
skewness.daily	−0.11169
se.skewness.daily	0.18648
kurtosis.daily	4.36803
se.kurtosis.daily	0.76746
min.daily	−7.18380
max.daily	6.34875
day.of.min	1997-10-27
day.of.max	2002-07-24

8.2 The Empirical Distribution of Returns

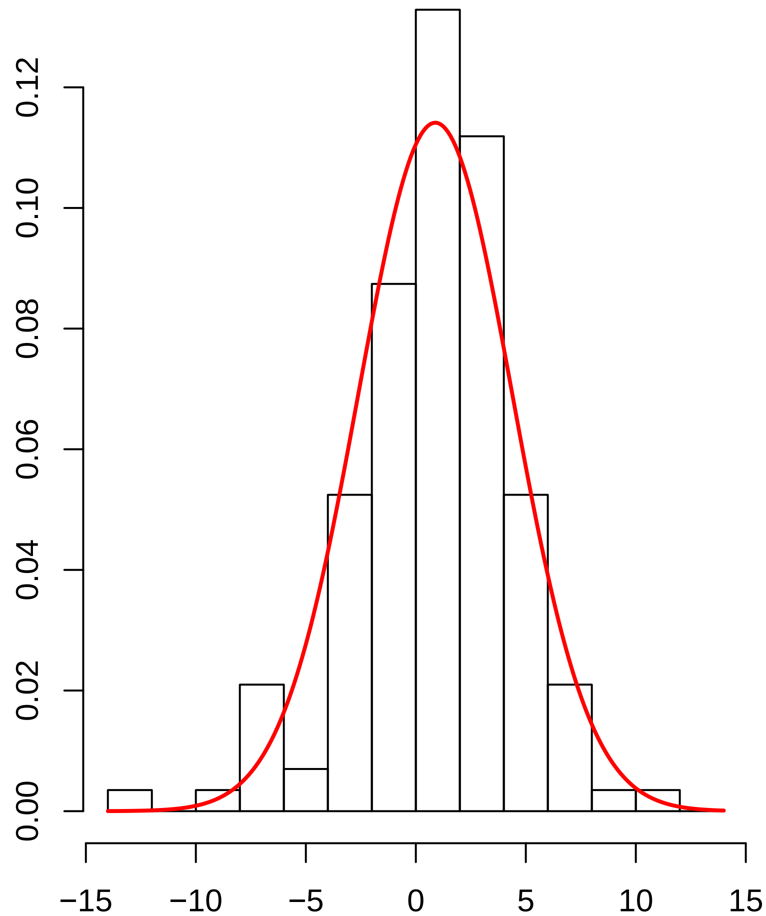
Dow-Jones: weekly returns, Jan 1995 through Nov 2006.



first.week	1995-01-10
last.week	2006-11-28
observations	621
NAs	0
mean.weekly	0.21190
var.weekly	5.29605
sd.weekly	2.30132
skewness.weekly	0.14744
se.skewness.weekly	0.32410
kurtosis.weekly	3.42075
se.kurtosis.weekly	0.99927
min.weekly	−9.09702
max.weekly	12.69361
week.of.min	2002-07-23
week.of.max	2002-07-30

8.2 The Empirical Distribution of Returns

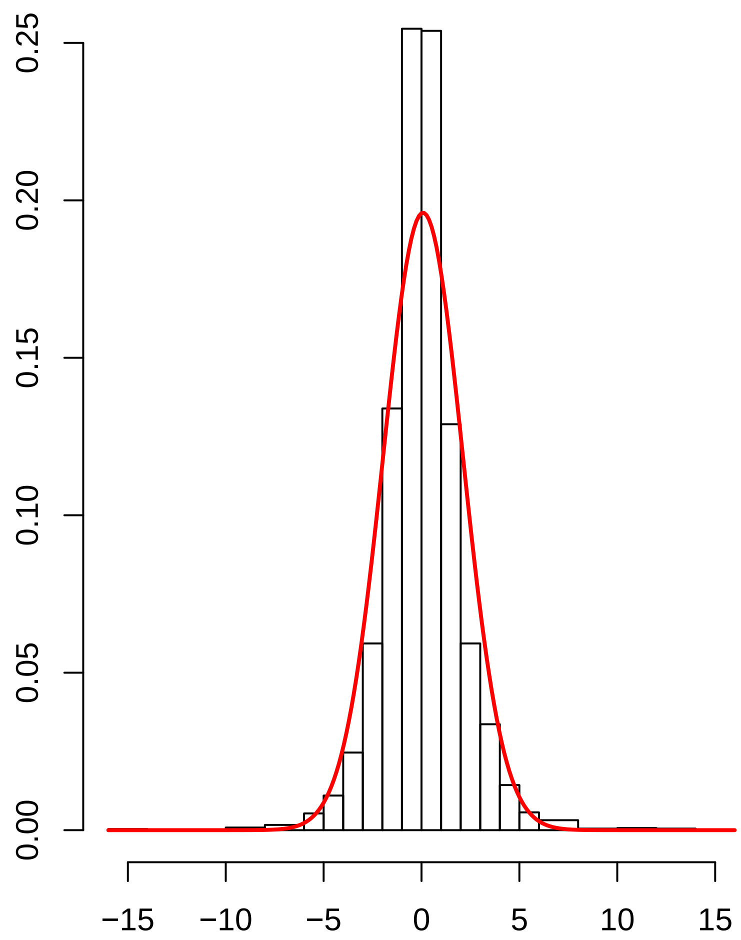
Dow-Jones: monthly returns, Jan 1995 through Nov 2006.



first.month	1995-01
last.month	2006-11
observations	143
NAs	0
mean.monthly	0.88474
var.monthly	12.21629
sd.monthly	3.49518
skewness.monthly	−0.55646
se.skewness.monthly	0.26924
kurtosis.monthly	1.21606
se.kurtosis.monthly	0.72587
min.monthly	−12.33317
max.monthly	10.39100
month.of.min	2001-09
month.of.max	1998-11

8.2 The Empirical Distribution of Returns

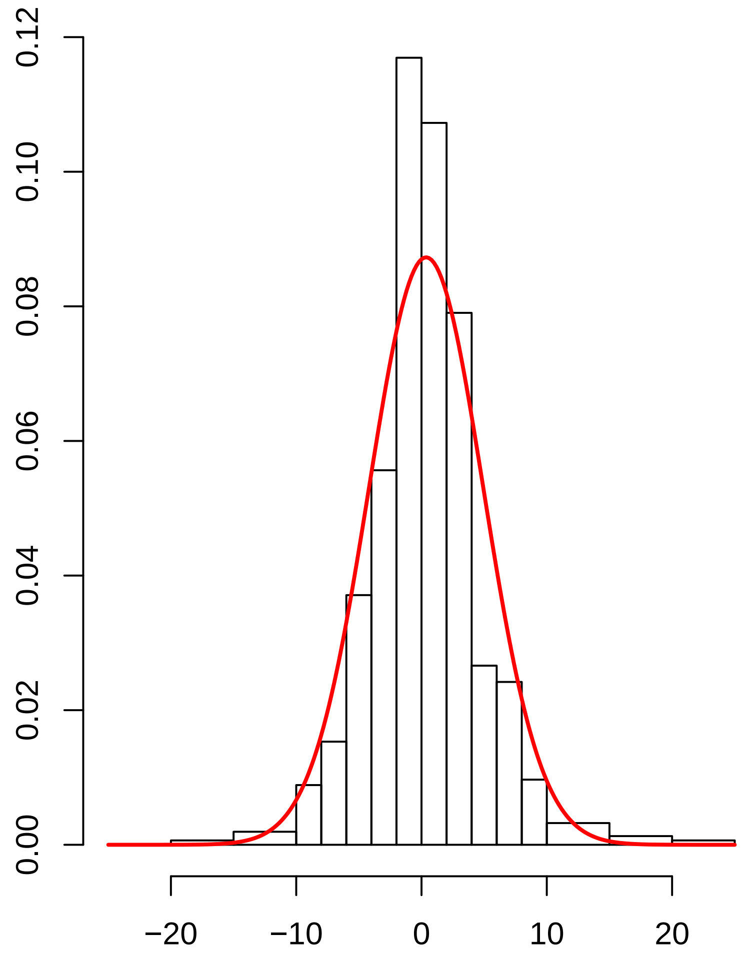
IBM: daily returns, Jan 1995 through Nov 2006.



first.day	1995-01-03
last.day	2006-11-30
observations	3002
NAs	106
mean.daily	0.07746
var.daily	4.14009
sd.daily	2.03472
skewness.daily	0.22746
se.skewness.daily	0.25993
kurtosis.daily	6.31412
se.kurtosis.daily	1.23160
min.daily	−15.54086
max.daily	13.16147
day.of.min	2000-10-18
day.of.max	1999-04-22

8.2 The Empirical Distribution of Returns

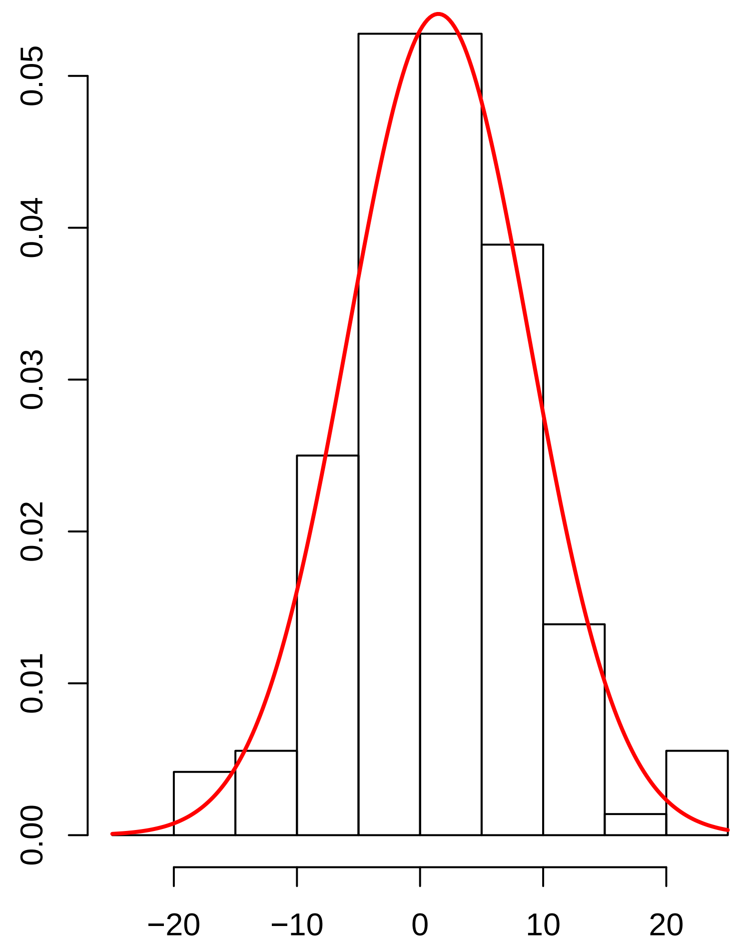
IBM: weekly returns, Jan 1995 through Nov 2006.



first.week	1995-01-10
last.week	2006-11-28
observations	621
NAs	0
mean.weekly	0.37610
var.weekly	20.87730
sd.weekly	4.56917
skewness.weekly	0.53280
se.skewness.weekly	0.26999
kurtosis.weekly	3.49849
se.kurtosis.weekly	0.74854
min.weekly	−19.08134
max.weekly	24.89072
week.of.min	2000-10-24
week.of.max	1999-04-27

8.2 The Empirical Distribution of Returns

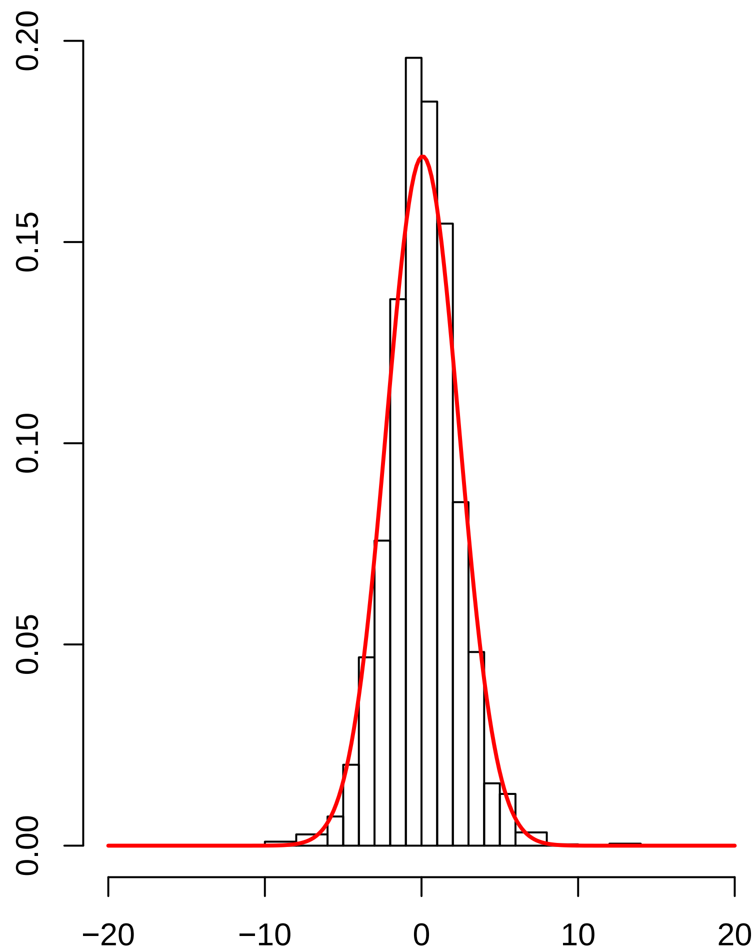
IBM: monthly returns, Jan 1995 through Nov 2006.



first.month	1995-01
last.month	2006-12
observations	144
NAs	0
mean.monthly	1.48554
var.monthly	54.41103
sd.monthly	7.37638
skewness.monthly	0.32133
se.skewness.monthly	0.21880
kurtosis.monthly	0.53411
se.kurtosis.monthly	0.37942
min.monthly	−16.76360
max.monthly	24.07360
month.of.min	1999-10
month.of.max	1996-02

8.2 The Empirical Distribution of Returns

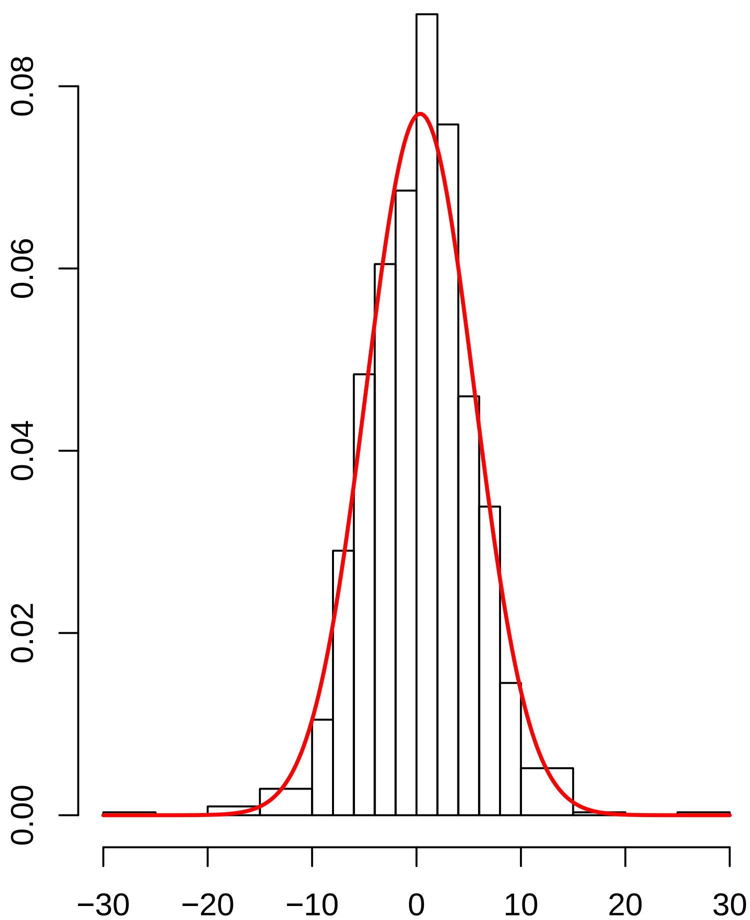
Brent crude oil: daily returns, Jan 1995 through Nov 2006.



first.day	1995-01-03
last.day	2006-11-30
observations	3034
NAs	74
mean.daily	0.07253
var.daily	5.42245
sd.daily	2.32862
skewness.daily	0.04323
se.skewness.daily	0.23085
kurtosis.daily	4.00435
se.kurtosis.daily	1.22182
min.daily	−18.03735
max.daily	17.65182
day.of.min	2001-09-24
day.of.max	1998-03-23

8.2 The Empirical Distribution of Returns

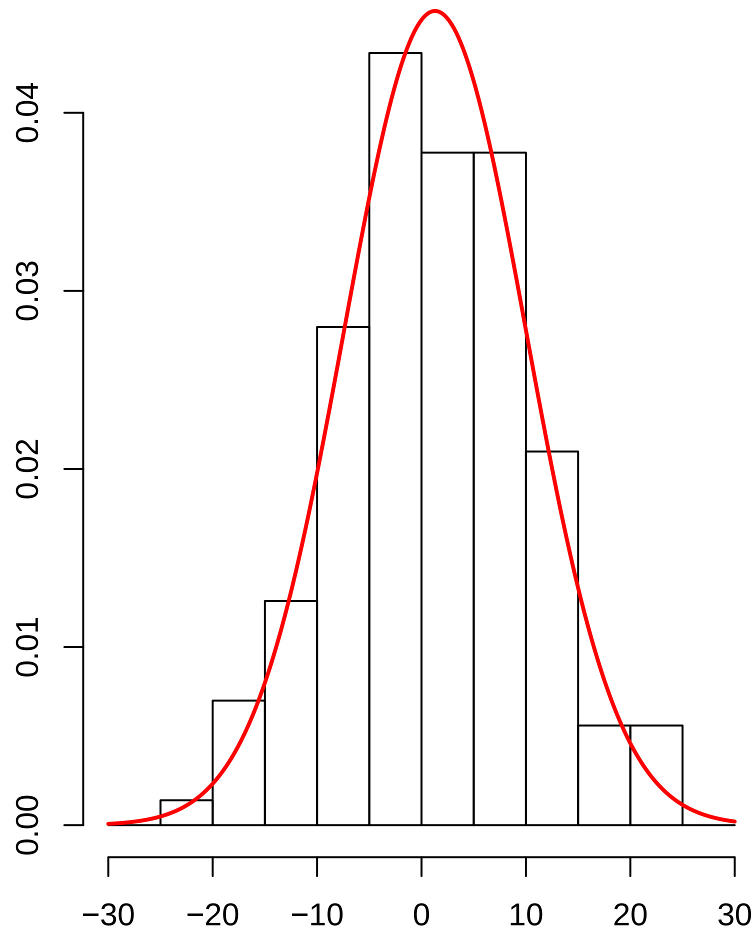
Brent crude oil: weekly returns, Jan 1995 through Nov 2006.



first.week	1995-01-10
last.week	2006-11-28
observations	621
NAs	0
mean.weekly	0.35201
var.weekly	26.80994
sd.weekly	5.17783
skewness.weekly	−0.05792
se.skewness.weekly	0.27250
kurtosis.weekly	2.05191
se.kurtosis.weekly	1.11301
min.weekly	−26.90632
max.weekly	26.87783
week.of.min	2001-09-25
week.of.max	1998-03-24

8.2 The Empirical Distribution of Returns

Brent crude oil: monthly returns, Jan 1995 through Nov 2006.



first.month	1995-01
last.month	2006-11
observations	143
NAs	0
mean.monthly	1.29530
var.monthly	76.14568
sd.monthly	8.72615
skewness.monthly	−0.08475
se.skewness.monthly	0.16689
kurtosis.monthly	−0.16898
se.kurtosis.monthly	0.23569
min.monthly	−21.17294
max.monthly	22.24545
month.of.min	2000-12
month.of.max	1999-04

8.2 The Empirical Distribution of Returns

Some conclusions from these examples.

- The shorter the time spacing, the heavier the tails of the return distribution.
- The longer the time spacing, the higher the standard deviation.
- The normal distribution is not an appropriate model as a return distribution — at least not for daily or weekly returns!

8.2 The Empirical Distribution of Returns

Some further remarks.

- It is not sufficient to look only at *empirical* return distributions:
  - They do not allow for a conditional approach.
  - They provide only shaky information (especially for the tails).
- Our further program is:
  - In 8.3, we shall see a GARCH approach to model the return distribution.
  - In 8.4 and 8.5, we shall look at efforts to model only the tails of a return distribution.
  - 8.5 is devoted to models for the distribution of stock prices.

8.3 A GARCH Approach

The conditional return distribution.

- The GARCH(1,1) was defined by

$$\epsilon_t = \nu_t \cdot \sqrt{h_t}, \quad h_t = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 h_{t-1},$$

  where (ν_t) is white noise.

- If returns follow a process (R_t) with $R_t = c + \epsilon_t$, the conditional distribution of R_t, given $R_{t-1}, R_{t-2}, \ldots$ is:

$$R_t | R_{t-1}, R_{t-2}, \ldots \ \sim \ \mathsf{N}(c, h_t)$$

  (if (ν_t) is Gaussian white noise).

8.3 A GARCH Approach

Example: Dow-Jones, 2002-01-02 through 2006-12-06.

- We had:

```
Coefficient(s):
    Estimate  Std. Error  t value Pr(>|t|)
a0  0.005481    0.002682    2.044    0.041 *
a1  0.056106    0.009821    5.713 1.11e-08 ***
b1  0.936337    0.011614   80.619  < 2e-16 ***
```

- Mean return: 0.022.

- The model equations are:

$$
\begin{aligned}
R_t &= 0.022 + \nu_t \cdot \sqrt{h_t}, \\
h_t &= 0.0055 + 0.0561\epsilon_{t-1}^2 + 0.9363 h_{t-1}.
\end{aligned}
$$

8.3 A GARCH Approach

Example: Dow-Jones, 2002-01-02 through 2006-12-06.

- Last day used: Wednesday, 2006-12-06 ($t = 1243$).
- For that day: $\epsilon_{1243} = -0.18124$, $h_{1243} = 0.54525$.
- The conditional distribution of R_{1244}, the return the next day (Thursday, 2006-12-07), is therefore: $\mathsf{N}(0.022, h_{1244})$ with

$$
\begin{aligned}
h_{1244} &= 0.0055 + 0.0561 \cdot (-0.18124)^2 + 0.9363 \cdot 0.54525 \\
&= 0.5179.
\end{aligned}
$$

8.3 A GARCH Approach

Example: Dow-Jones, 2002-01-02 through 2006-12-06.

- This result can also be used for VaR computations.
- The 1% VaR for Thursday, 2006-12-07, is given by:

$$-2.33 \cdot \sqrt{0.5179} + 0.022 \approx -1.65\%.$$

- The return on Thursday, 2006-12-07, was indeed about -0.25%.

8.3 A GARCH Approach

Example: Dow-Jones, 2002-01-02 through 2006-12-06.

- Let's compare this to an "average" approach, based on the normal distribution.
- The mean daily return for this period is 0.02, the empirical variance is 1.002.
- Using the normal distribution, this implies an average (absolute) 1% VaR for the next day of about 2.31%.
- However, the empirical 1% quantile is -2.48%!
- Using the normal distribution underestimates the risk.

8.3 A GARCH Approach

Final remarks.

- The GARCH approach yields the conditional distribution of returns.
- This approach can be used to determine the VaR for the next day.
- It can also be used for other purposes, for example: the determination of days with unexpectedly high gain or loss.
- We used Gaussian white noise. — The normal distribution has no heavy tails, but mixing normals (as in a GARCH process) makes tails heavier!

8.4 The Distribution of Extreme Returns

The distribution of extreme returns.

Let returns $R_1, \ldots, R_n$ from n days be iid. Define:

$$M_n = \max\{R_1, \ldots, R_n\}$$

Then, for sufficiently large n,

$$P(M_n \leq z) \approx G(z) = \exp\left\{ - \left[1 + \xi \left(\frac{z - \mu}{\sigma} \right) \right]^{-1/\xi} \right\}.$$

This is called the generalized extreme value (GEV) distribution.

8.5 The Distribution of Excess Returns

The distribution of excess returns.

$R_1, \ldots, R_n$ be as before, and R distributed like R_i. Then, for large n and u, the distribution function of the excess

$$R - u, \quad \text{conditional on} \quad R > u,$$

is approximately

$$H(y) = 1 - \left(1 + \xi \frac{y}{\sigma}\right)^{-1/\xi}$$

with suitable ξ and σ. This is called the generalized Pareto distribution (GPD).

8.5 The Distribution of Excess Returns

The "tail index".

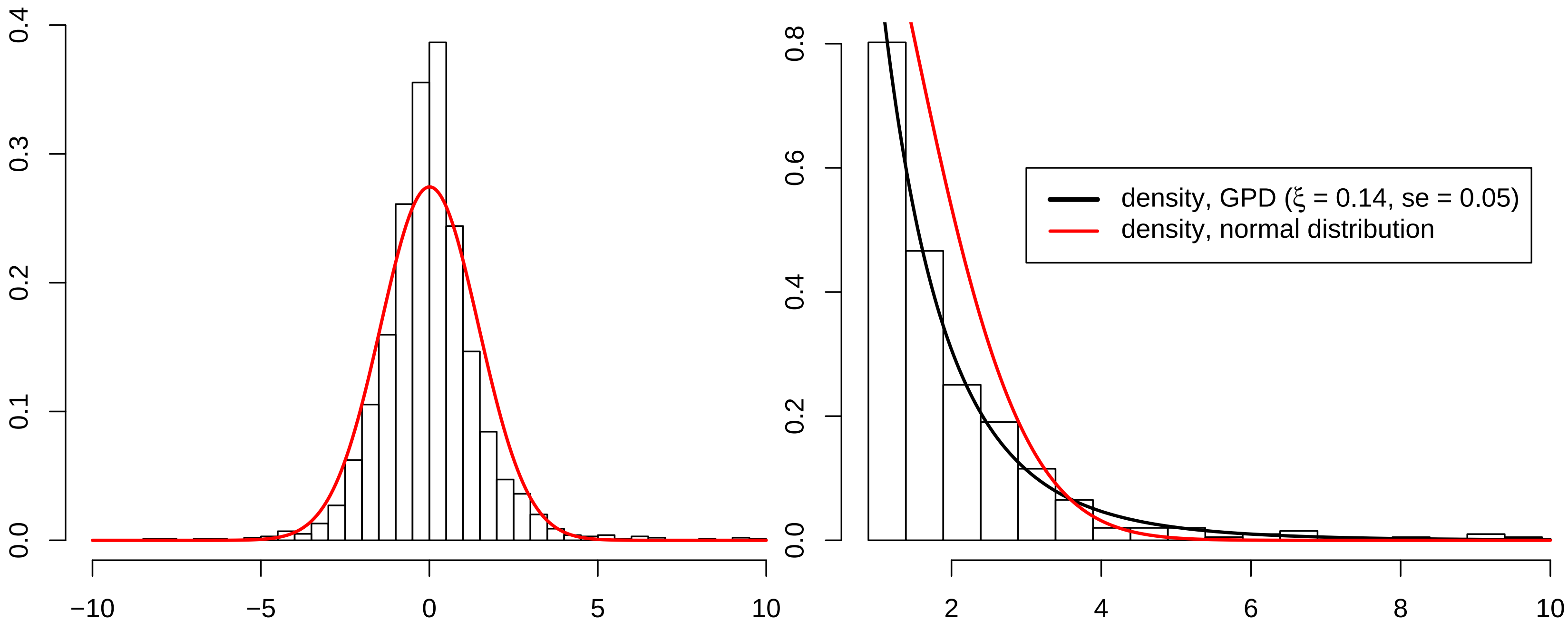
In both distributions, the parameter ξ characterizes the tail of the distribution.

- $\xi > 0$: distribution of Type II (Fréchet; the tail decreases polynomially fast — "heavy tails")
- $\xi = 0$: distribution of Type I (Gumbel; the tail decreases exponentially fast — "thin tails")
- $\xi < 0$: distribution of Type III (Weibull; there is no tail)

8.5 The Distribution of Excess Returns

Example: The SSEC (Shanghai Stock Exchange Composite).

The following figures compare the empirical return distribution, the normal distribution, and the generalized Pareto distribution.



8.5 The Distribution of Excess Returns

Example: The SSEC (Shanghai Stock Exchange Composite).

- Period: 1997-07-03 through 2005-10-20 (1992 observations)
- 80% quantile: $q_{0.8} = 0.8944$
- Estimated parameters:

parameter	estimated	s.e.
ξ	$\hat{\xi} = 0.136$	0.055
σ	$\hat{\sigma} = 0.924$	0.069

- Estimated distribution function of excesses:

$$P(R - q_{0.8} \leq y | R > q_{0.8}) = 1 - \left(1 + \frac{0.136}{0.924} y\right)^{-\frac{1}{0.136}} = 1 - (1 + 0.147y)^{-7.41}$$

8.5 The Distribution of Excess Returns

Example: The SSEC (Shanghai Stock Exchange Composite).

What is the (estimated) probability that the daily return on SSEC is larger than 4%?

- This is the probability that the excess is larger than $4 - 0.8944 = 3.1056$.
- On condition that R exceeds $q_{0.8}$, this probability is

$$1 - H(3.1056) = (1 + 0.147 \cdot 3.1056)^{-7.41} = 0.06.$$

- The unconditional probability is then

$$0.06 \cdot 0.2 = 0.012 \approx 1\%.$$

8.5 The Distribution of Excess Returns

Example: The SSEC (Shanghai Stock Exchange Composite).

What is the (estimated) probability that the daily return on SSEC is larger than 4%?

- Let us compare this probability with one computed using the normal distribution.
- Estimation yields: $R \sim \mathsf{N}(0.005, 2.113)$.
- Therefore:

$$P(R > 4) = P\left(\frac{R - 0.005}{\sqrt{2.113}} > \frac{4 - 0.005}{\sqrt{2.113}}\right) = 0.003 = 0.3\%.$$

- Conclusions?

8.6 The Distribution of Stock Prices

Lognormally distributed stock prices.

- Given P_{t-1}, the stock price at time t is

$$P_t = X \cdot P_{t-1}$$

  with a random factor X (the "gross return").

- What is the distribution of X?

- Some celebrated models assume that X is lognormally distributed. This means: $\ln X$ is normally distributed.

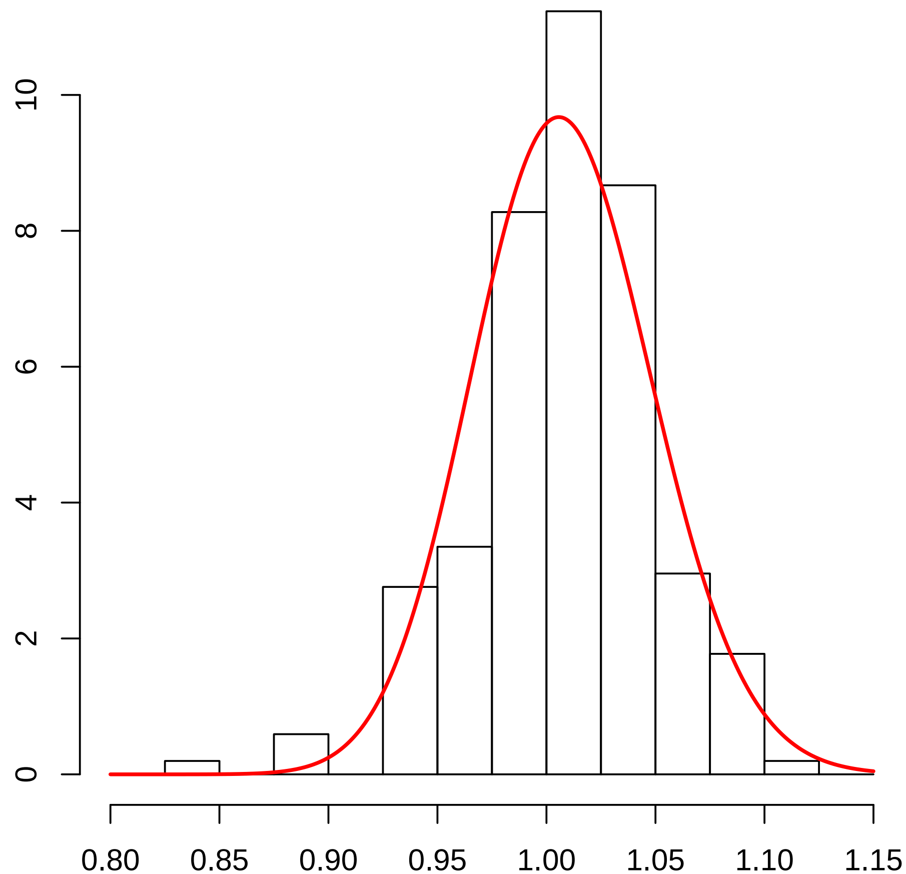
8.6 The Distribution of Stock Prices

Outlook. We shall now see:

- the empirical distribution of monthly and quarterly gross returns, for the period Jan 1990 through Nov 2006, on
  - Dow-Jones,
  - IBM,
  - Brent crude oil,
- together with the estimated LN density,
- and the p-values of the null hypothesis that the true distribution is a LN distribution (based on the Kolmogorov-Smirnov test).

8.6 The Distribution of Stock Prices

Distribution of monthly gross return: Dow-Jones.



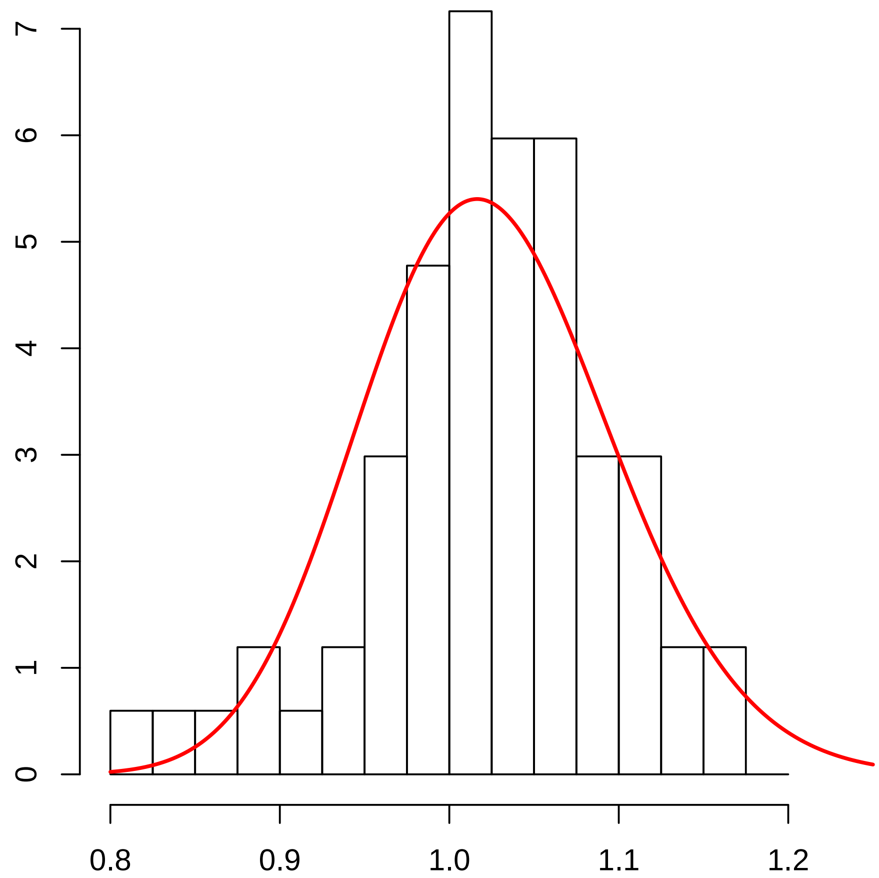
$H_0 : F = \text{LN}$

$H_1 : F \neq \text{LN}$

p-value: 0.3878

8.6 The Distribution of Stock Prices

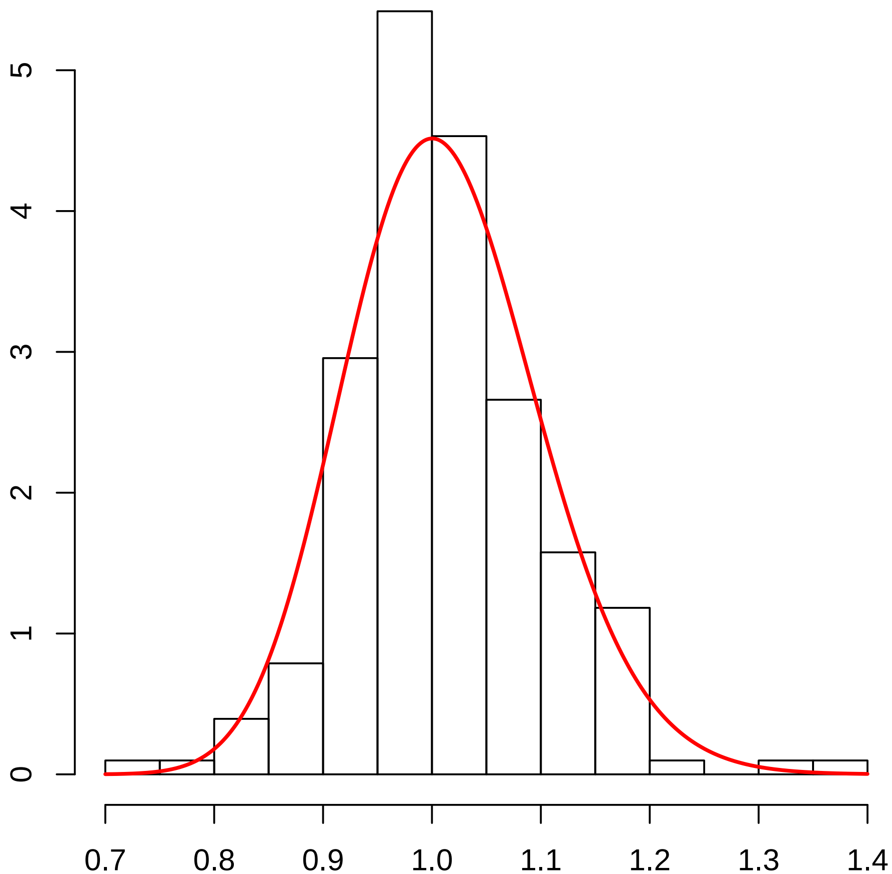
Distribution of quarterly gross return: Dow-Jones.



$H_0 : F = \text{LN}$
$H_1 : F \neq \text{LN}$
p-value: 0.4229

8.6 The Distribution of Stock Prices

Distribution of monthly gross return: IBM.



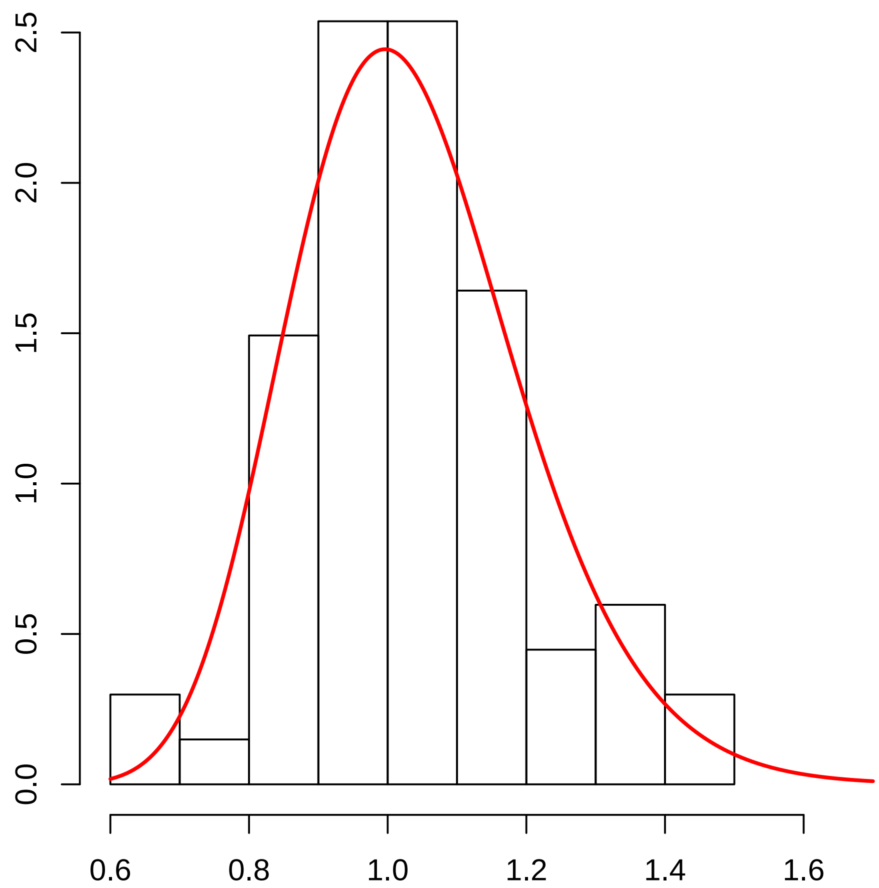
$H_0 : F = \text{LN}$

$H_1 : F \neq \text{LN}$

p-value: 0.7194

8.6 The Distribution of Stock Prices

Distribution of quarterly gross return: IBM.



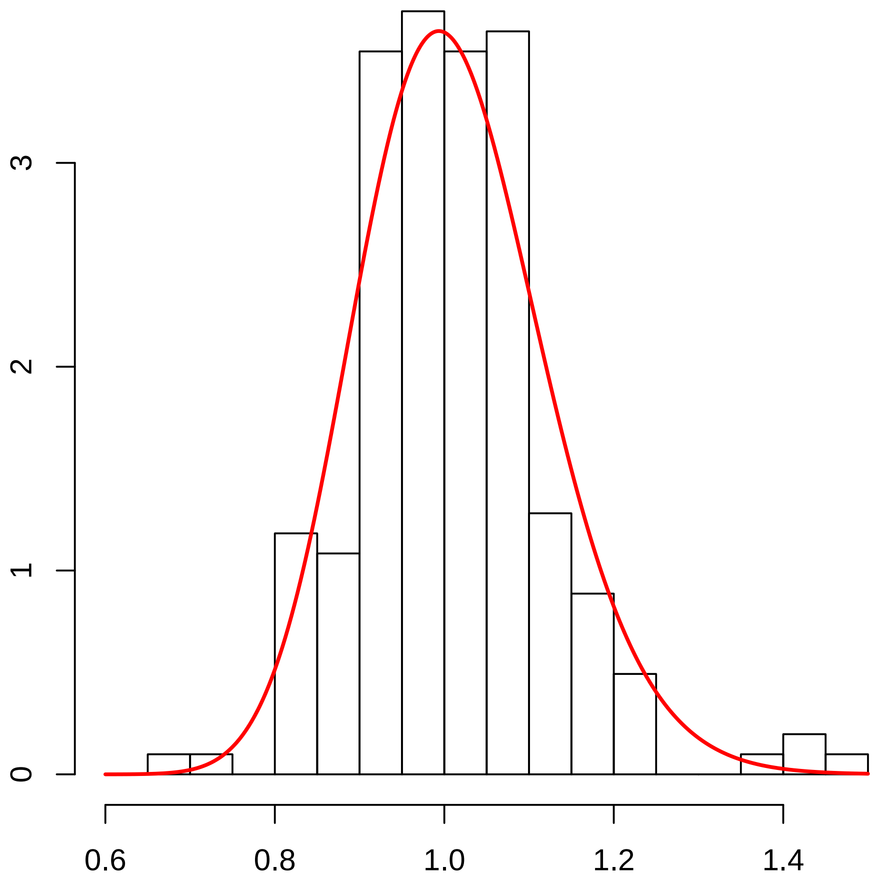
$H_0 : F =$ LN

$H_1 : F \neq$ LN

p-value: 0.9853

8.6 The Distribution of Stock Prices

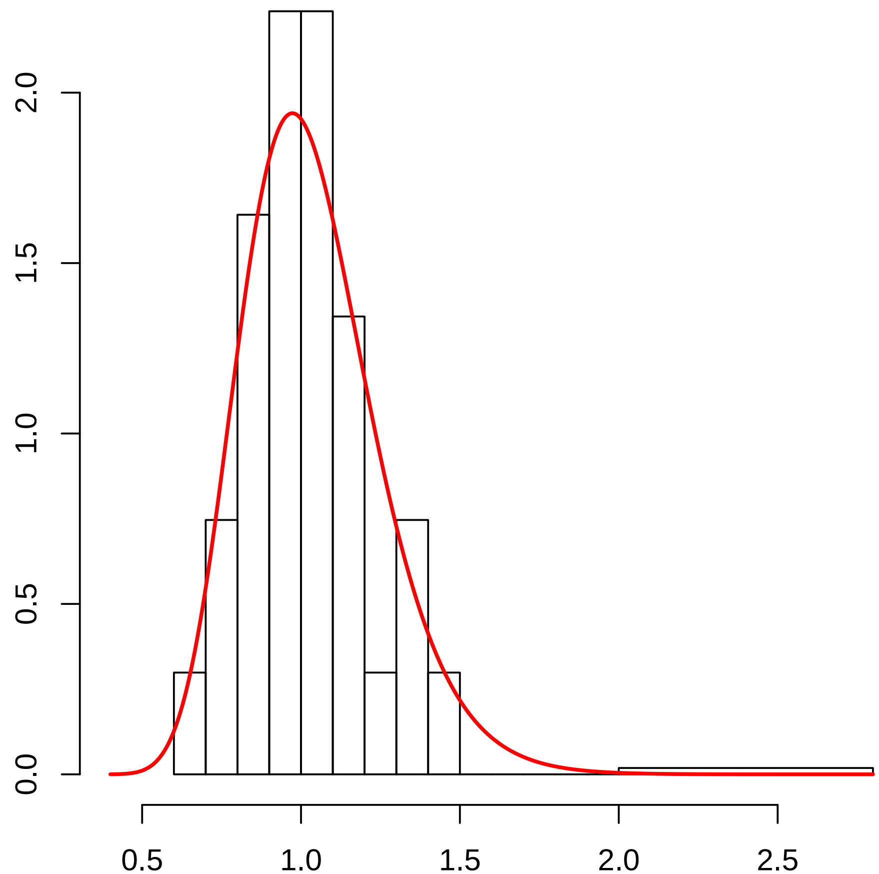
Distribution of monthly gross return: Brent.



$H_0 : F = \text{LN}$
$H_1 : F \neq \text{LN}$
p-value: 0.3792

8.6 The Distribution of Stock Prices

Distribution of quarterly gross return: Brent.



$H_0 : F = \text{LN}$

$H_1 : F \neq \text{LN}$

p-value: 0.2651

8.6 The Distribution of Stock Prices

Lognormally distributed gross returns.

Here is a justification for the LN assumption, based on the CLT.

With

$$R_t = \frac{P_t - P_{t-1}}{P_{t-1}} = \text{return on day } t,$$

it holds that

$$P_t = P_{t-1} \cdot (1 + R_t) \approx P_{t-1} \cdot \mathrm{e}^{R_t}$$

8.6 The Distribution of Stock Prices

Lognormally distributed gross returns.

Likewise, given the initial price P_0, the stock price is on. . .

day 1: $P_0 \cdot (1 + R_1) \approx P_0 \cdot \mathrm{e}^{R_1}$

day 2: $P_0 \cdot (1 + R_1)(1 + R_2) \approx P_0 \cdot \mathrm{e}^{R_1}\mathrm{e}^{R_2} = P_0 \cdot \mathrm{e}^{R_1 + R_2}$

. . .

day T: $P_0 \cdot \prod_{t=1}^{T}(1 + R_t) \approx P_0 \cdot \prod_{t=1}^{T} \mathrm{e}^{R_t} = P_0 \cdot \mathrm{e}^{\sum_{t=1}^{T} R_t}$

so that

$$P_T \approx P_0 \cdot X_T \quad \text{with} \quad X_T = \exp\left(\sum_{t=1}^{T} R_t\right)$$

and $X_T \sim \mathrm{LN}(0, T\sigma^2)$ approximately, if the R_t are an iid sequence.

8.6 The Distribution of Stock Prices

Brownian motion.

- There is another way to lognormally distributed returns: via Brownian motion.

- $(W_t)_{t \geq 0}$ is called Brownian motion if
  - $W_0 = 0$,
  - $W_t \sim \mathrm{N}(0, t)$,
  - $W_t - W_s$ and W_s are independent for $t > s$.

- A model for a stock price is then: $P_t = P_0 \cdot \mathrm{e}^{\mu + \sigma W_t}$.

8.6 The Distribution of Stock Prices

Brownian motion — a simulated sample path.