# Financial Time Series

## Lecture 1: Introduction

Financial time series (FTS) analysis is concerned with theory and practice of asset valuation over time.

What is the difference, if any, from traditional time series analysis?

Two topics are highly related, but FTS has added uncertainty, because it must deal with the ever-changing business & economic environment and the fact that volatility is not directly observed.

### Objective of the course

- to access financial data online and to process the embedded information
- to provide basic knowledge of FTS data such as skewness, heavy tails, and measure of dependence between asset returns
- to introduce statistical tools & econometric models useful for analyzing these series.
- to gain experience in analyzing FTS
- to introduce recent developments in financial econometrics and their applications, e.g., high-frequency finance
- to study methods for assessing market risk, credit risk, and expected loss. The methods discussed include Value at Risk, expected shortfall, and tail dependence.

- to analyze high-dimensional asset returns, including co-movement

**Examples of financial time series**

1. Daily log returns of Apple stock: 2007 to 2016 (10 years). Data downloaded using `quantmod`
2. The VIX index
3. CDS spreads: Daily 3-year CDS spreads of JP Morgan from July 20, 2004 to September 19, 2014.
4. Quarterly earnings of Coca-Cola Company: 1983-2009

   Seasonal time series useful in

   - earning forecasts
   - pricing weather related derivatives (e.g. energy)
   - modeling intraday behavior of asset returns
5. US monthly interest rates (3m & 6m Treasury bills)

   Relations between the two asset returns? Term structure of interest rates
6. Exchange rate between US Dollar vs Euro

   Fixed income, hedging, carry trade
7. Size of insurance claims

   Values of fire insurance claims ($\times 1000$ Krone) that exceeded 500 from 1972 to 1992.
8. High-frequency financial data:

   Tick-by-tick data of Caterpillars stock: January 04, 2010.

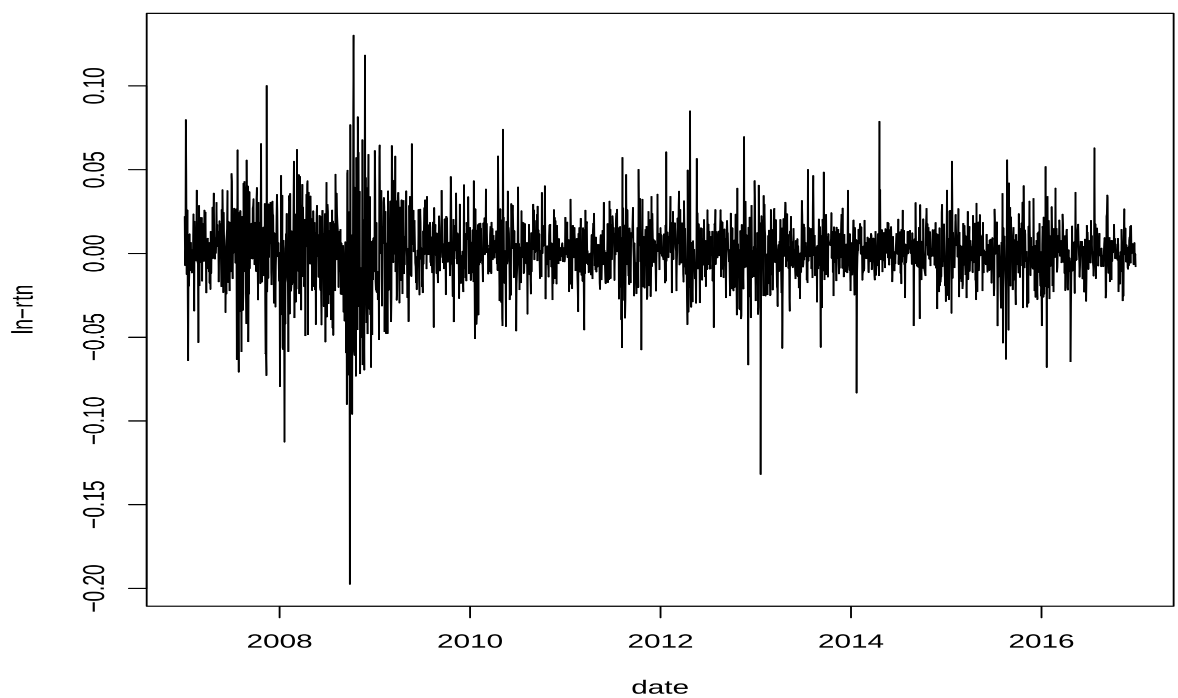


Figure 1: Daily log returns of Apple stock from 2007 to 2016

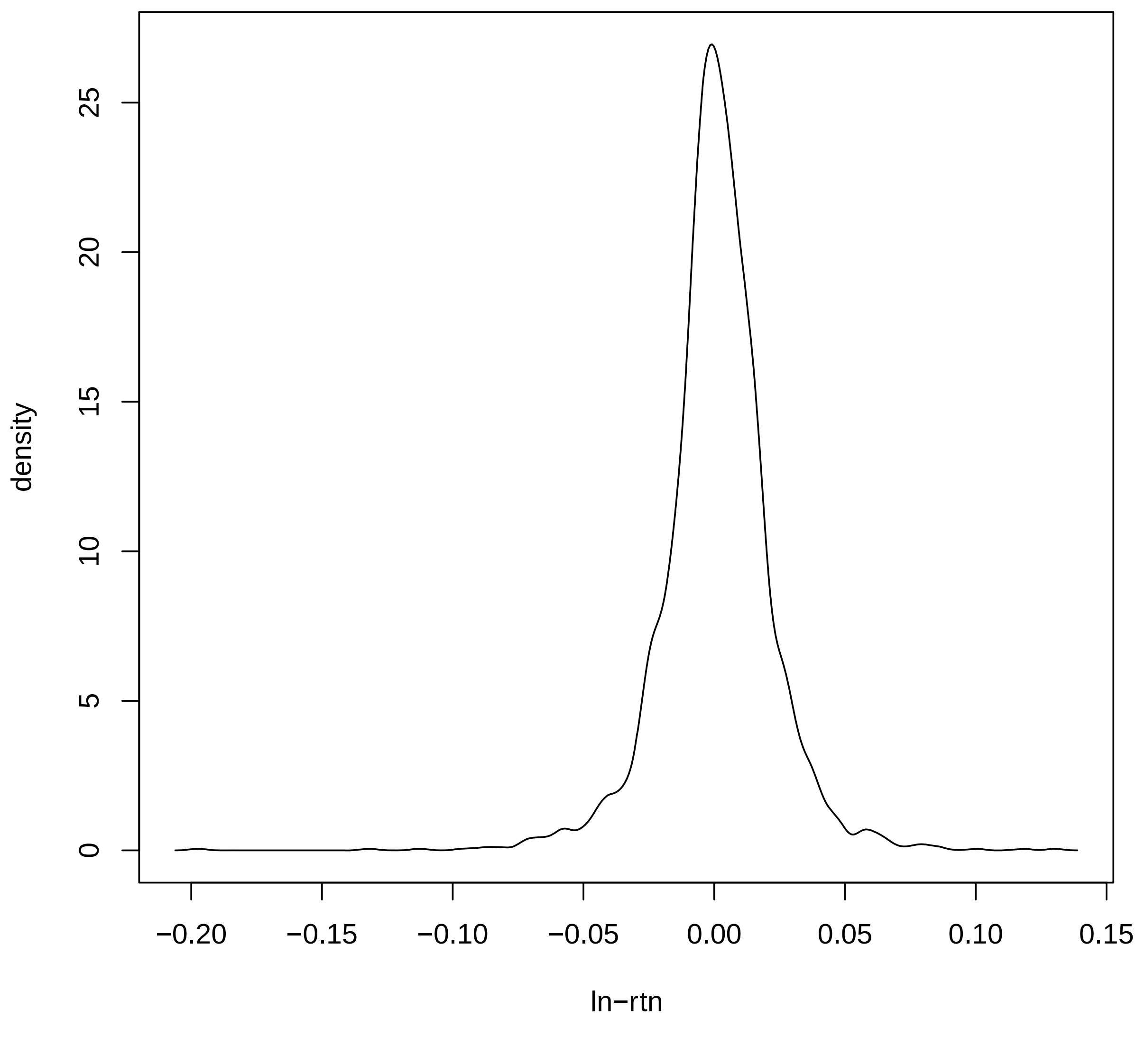


Figure 2: Empirical density function of daily log returns of Apple stock: 2007 to 2016

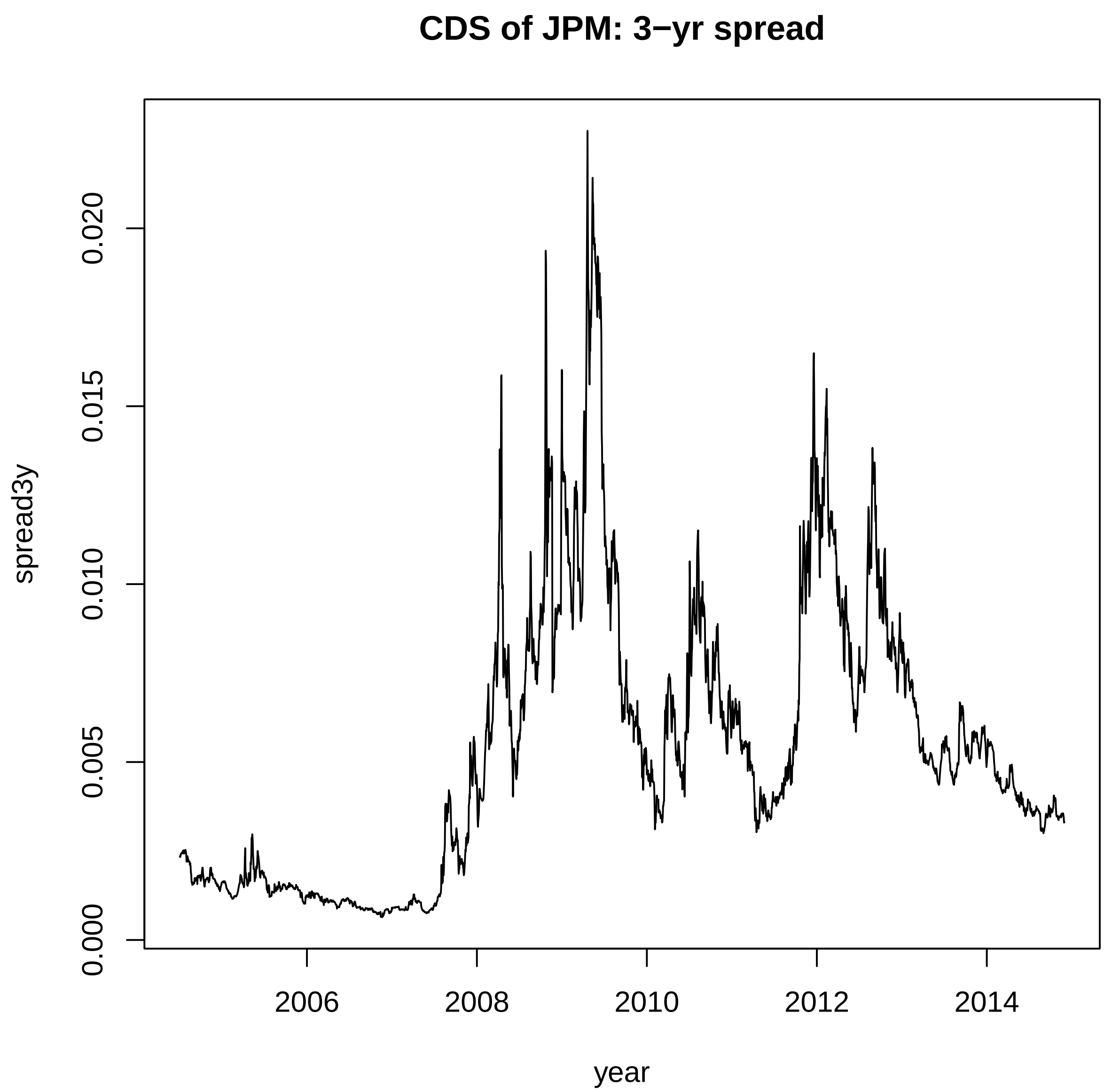


Figure 3: Time plot of daily 3-year CDS spreads of JPM: from July 20, 2004 to September 19, 2014.

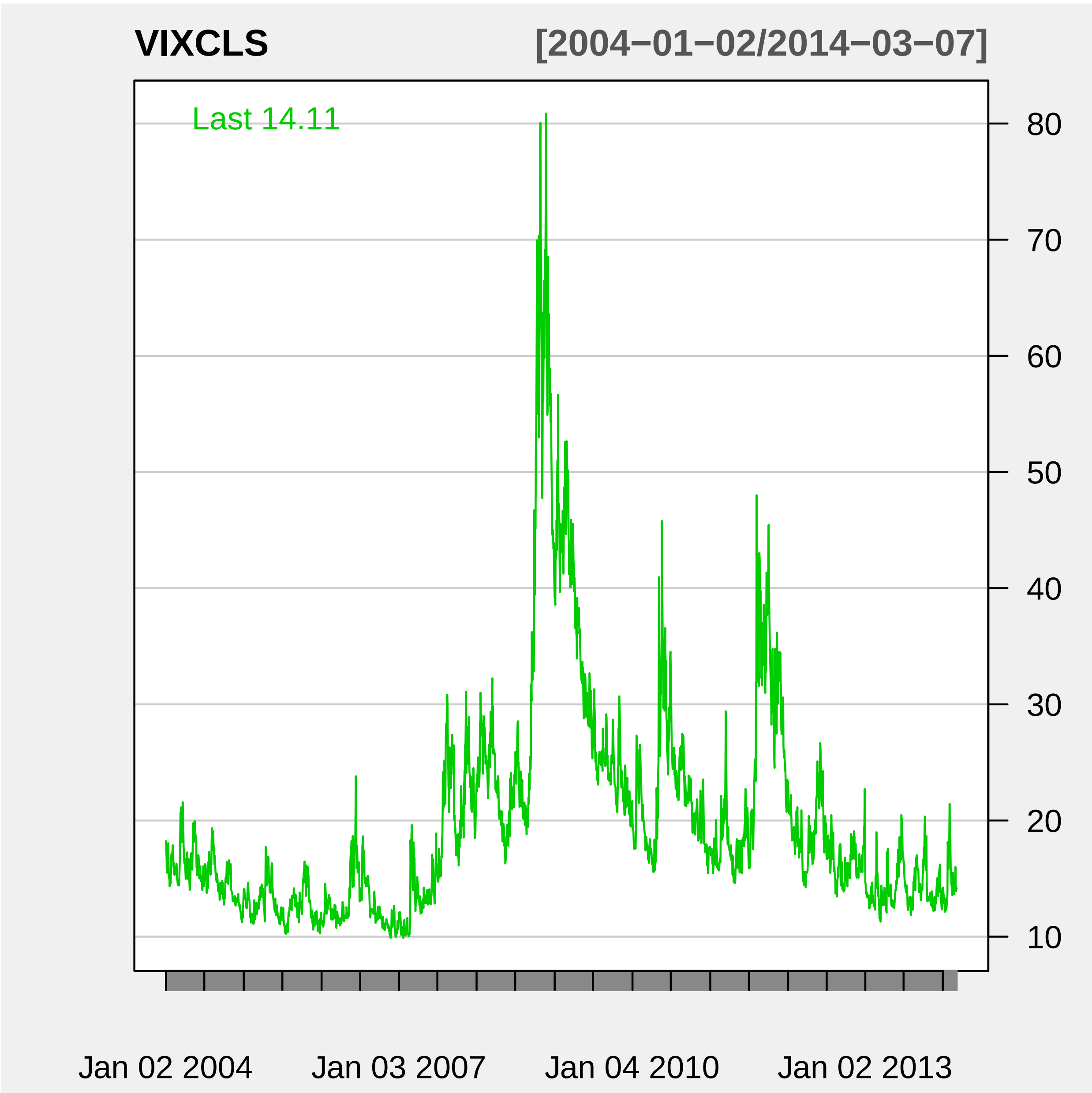


Figure 4: CBOE Vix index: January 2, 2004 to March 7, 2014.

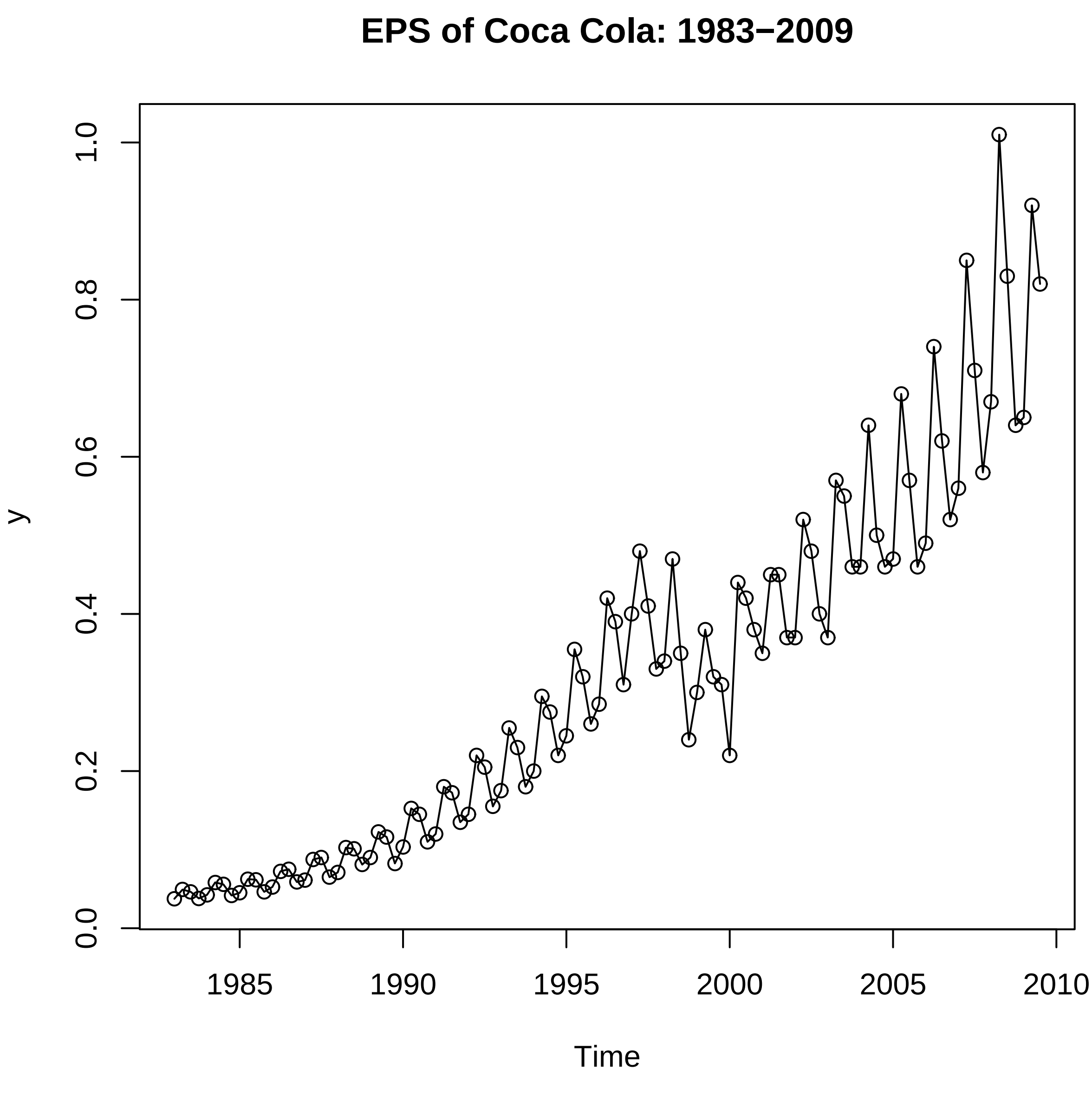


Figure 5: Quarterly earnings per share of Coca-Cola Company


Figure 6: Daily Exchange Rate: Dollars per Euro

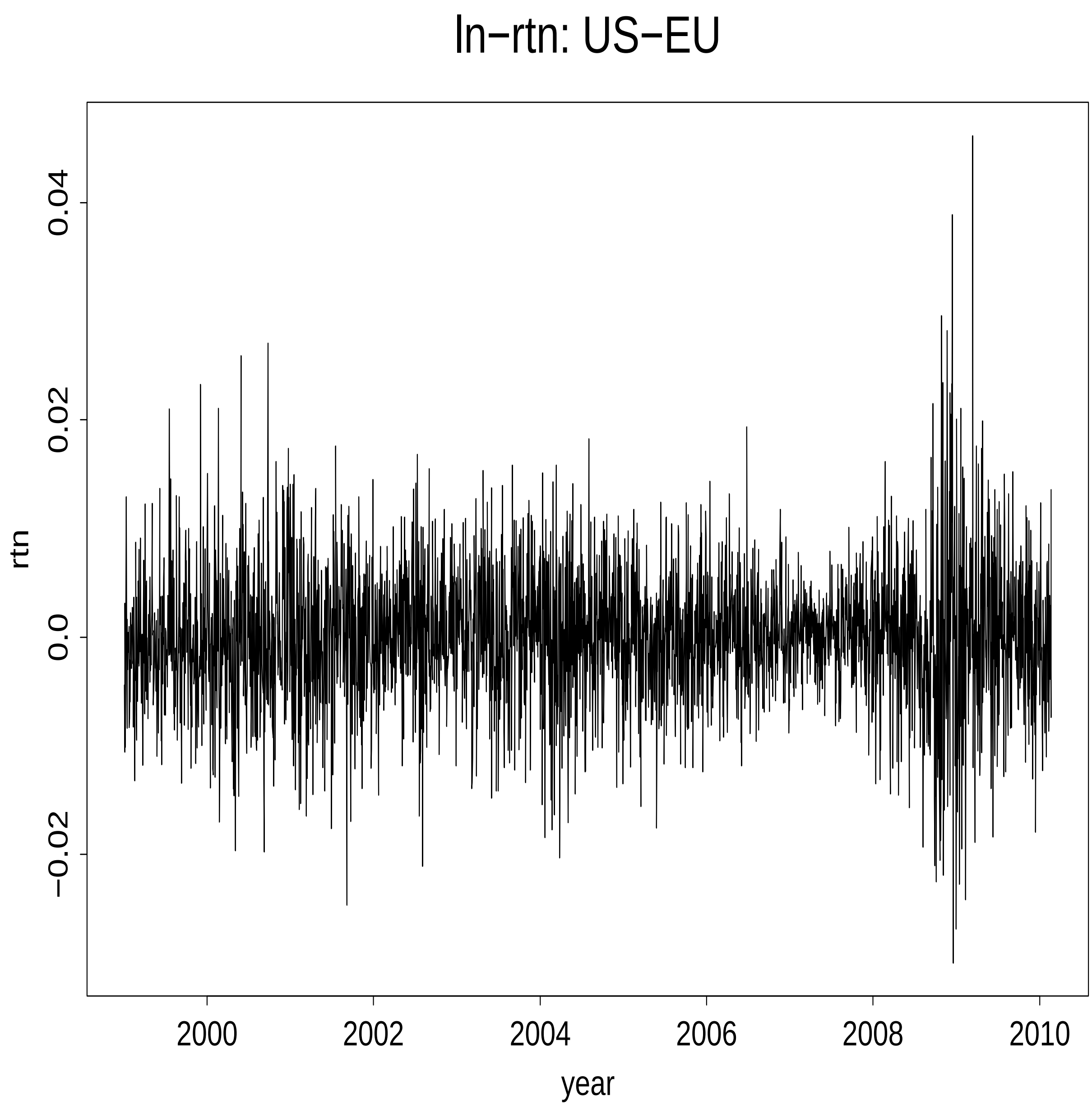


Figure 7: Daily log returns of FX (Dollar vs Euro)

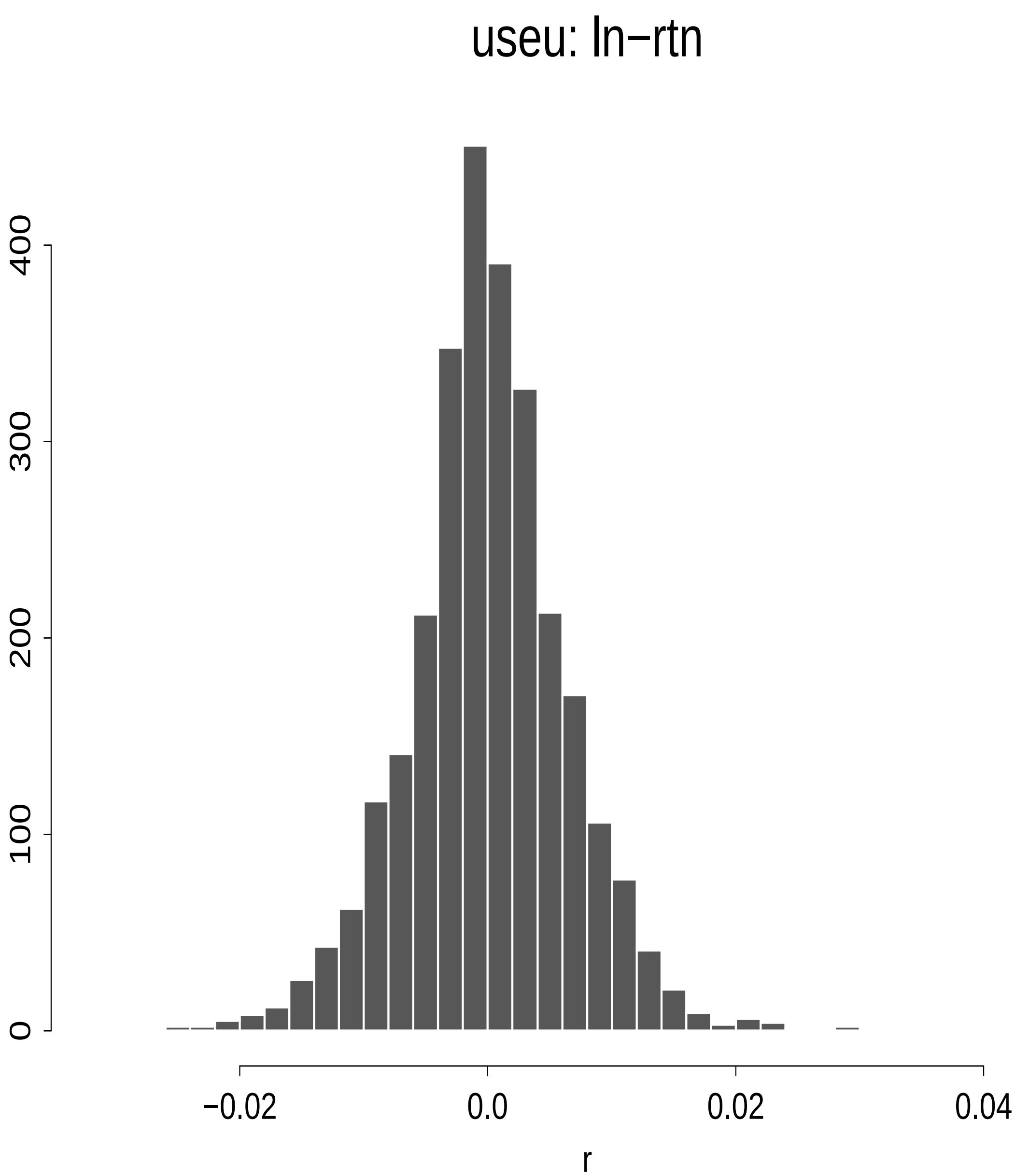


Figure 8: Histogram of daily log returns of FX (Dollar vs Euro)

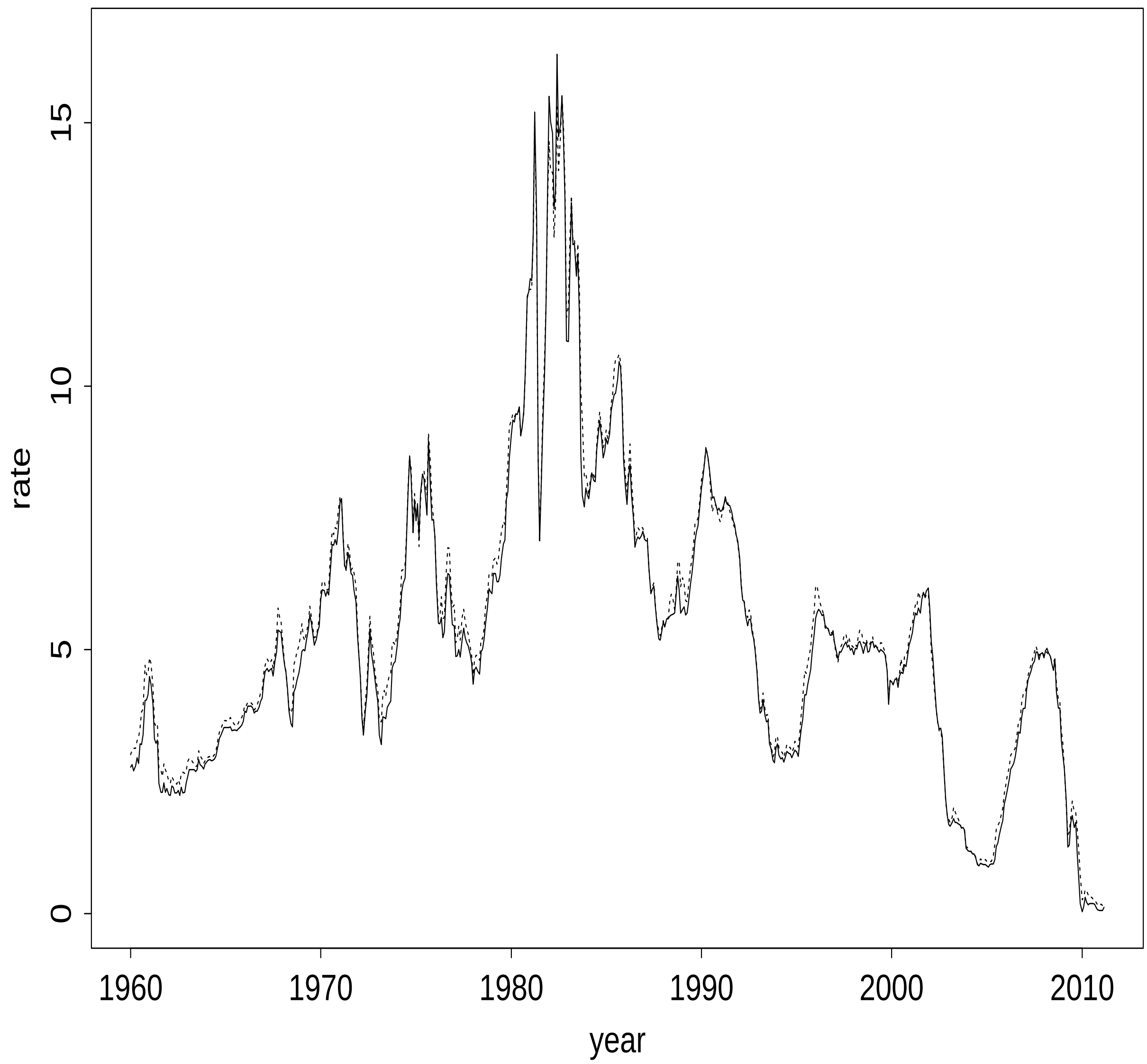


Figure 9: Monthly US interest rates: 3m & 6m TB

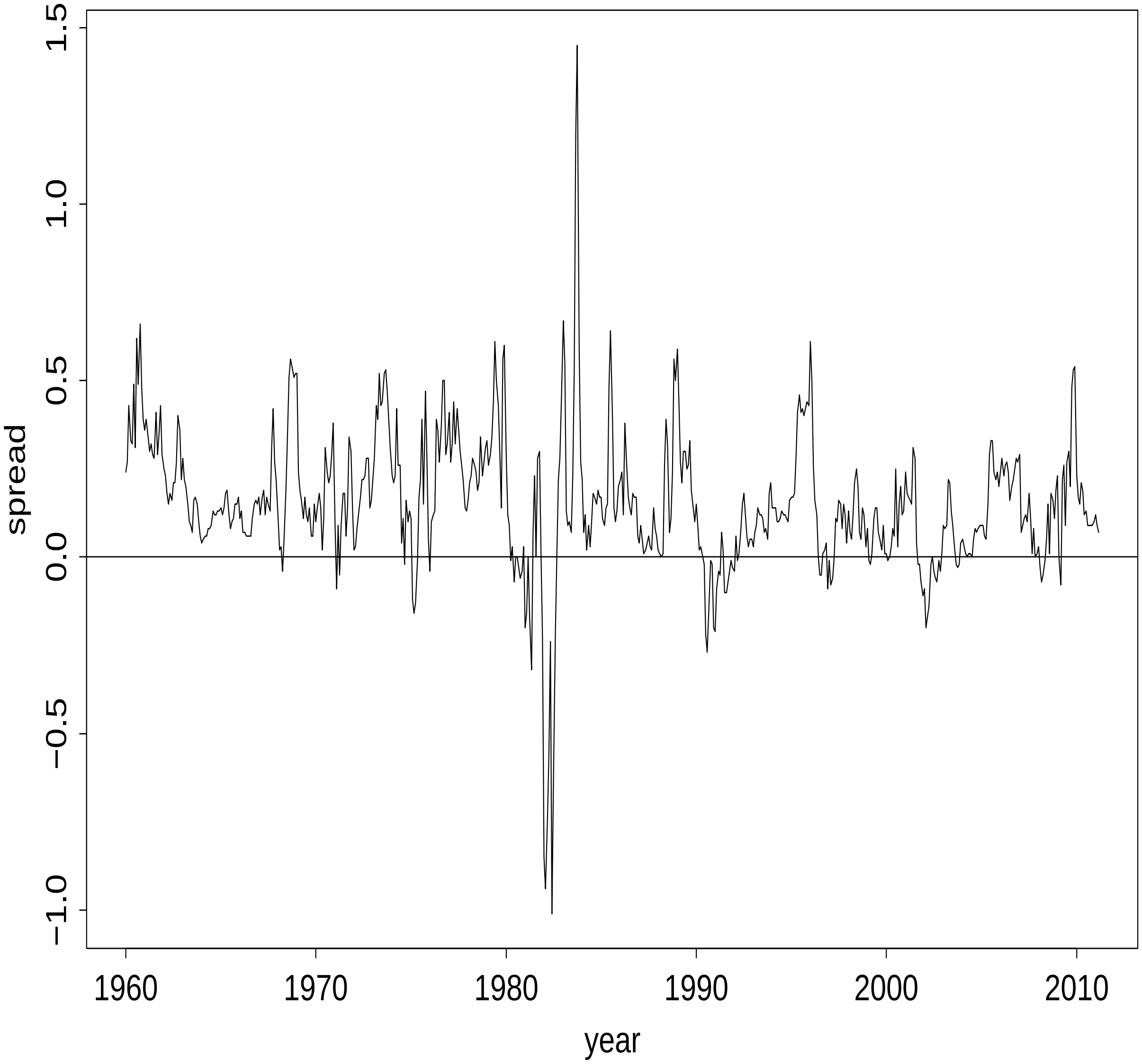


Figure 10: Spread of monthly US interest rates: 3m & 6m TB

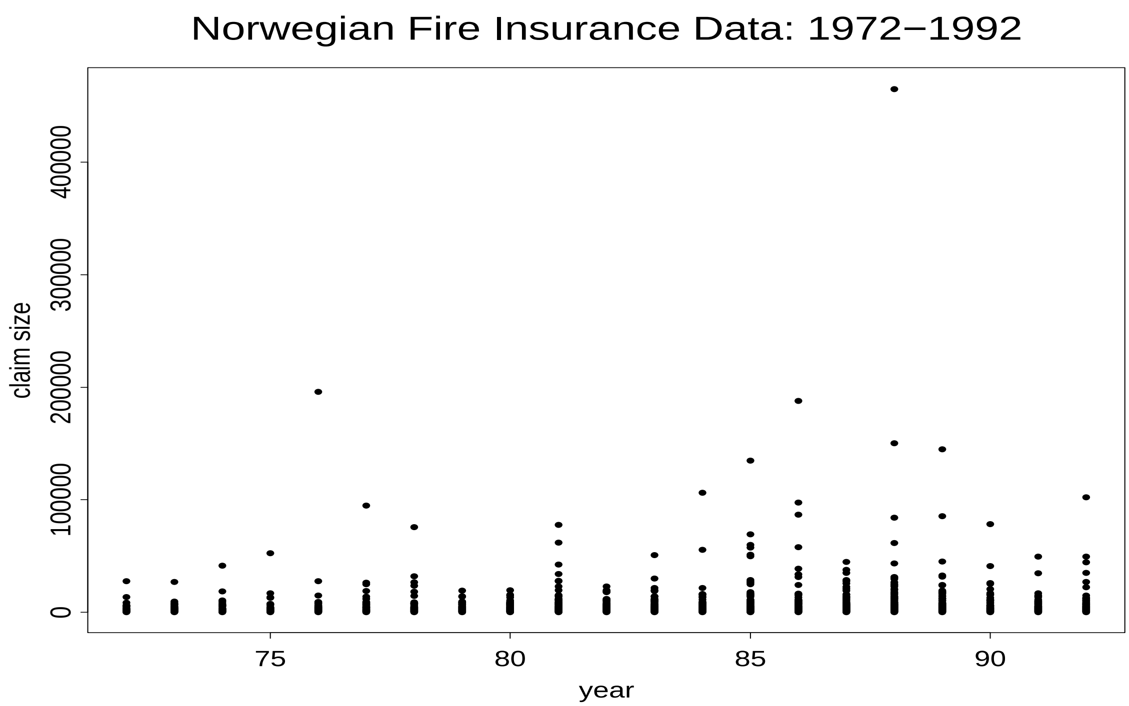


Figure 11: Claim sizes of the Norwegian fire insurance from 1972 to 1992, measured in 1000 Krone and exceeded 500.

CAT trade data on January 04, 2010.

```
date  hour minute second price size
20100104 9 30 0 57.65 3910
20100104 9 30 0 57.7 400
20100104 9 30 0 57.68 100
20100104 9 30 0 57.69 300
20100104 9 30 1 57.65 462
20100104 9 30 1 57.65 100
20100104 9 30 1 57.65 100
20100104 9 30 1 57.65 100
20100104 9 30 1 57.7 100
20100104 9 30 1 57.7 100
20100104 9 30 1 57.72 500
20100104 9 30 1 57.72 100
20100104 9 30 2 57.73 100
20100104 9 30 3 57.73 300
20100104 9 30 3 57.72 100
20100104 9 30 4 57.72 300
20100104 9 30 5 57.57 100
20100104 9 30 5 57.57 500
20100104 9 30 5 57.56 300
......
20100104 9 30 35 57.77 100
20100104 9 30 36 57.77 100
20100104 9 30 42 57.54 83600
20100104 9 30 42 57.57 100
.....
20100104 9 30 42 57.55 100
20100104 9 30 42 57.55 2400
20100104 9 30 42 57.56 100
20100104 9 30 42 57.55 100
20100104 9 30 42 57.55 100
20100104 9 30 42 57.55 100
20100104 9 30 42 57.54 170
20100104 9 30 42 57.54 200
```

## Outline of the course

- Returns & their characteristics: empirical analysis (summary statistics)
- Simple linear time series models & their applications
- Univariate volatility models & their implications

- Nonlinearity in level and volatility
- Neural network & non-parametric methods
- High-frequency financial data and market micro-structure
- Continuous-time models and derivative pricing
- Value at Risk, extreme value theory and expected shortfall (also known as conditional VaR)
- Analysis of multiple asset returns: factor models, dynamic and cross dependence, cross-section regression

**Asset Returns**

Let $P_t$ be the price of an asset at time $t$, and assume no dividend.
One-period simple return: Gross return

$$1 + R_t = \frac{P_t}{P_{t-1}} \quad \text{or} \quad P_t = P_{t-1}(1 + R_t)$$

**Simple return**:

$$R_t = \frac{P_t}{P_{t-1}} - 1 = \frac{P_t - P_{t-1}}{P_{t-1}}.$$

Multiperiod simple return: Gross return

$$\begin{aligned} 1 + R_t(k) &= \frac{P_t}{P_{t-k}} = \frac{P_t}{P_{t-1}} \times \frac{P_{t-1}}{P_{t-2}} \times \cdots \times \frac{P_{t-k+1}}{P_{t-k}} \\ &= (1 + R_t)(1 + R_{t-1}) \cdots (1 + R_{t-k+1}). \end{aligned}$$

The $k$-period simple net return is $R_t(k) = \frac{P_t}{P_{t-k}} - 1$.

**Example**: Table below gives six daily (adjusted) closing prices of Apple stock in December 2015. The 1-day gross return of holding the stock from 12/23 to 12/24 $1 + R_t = 107.45/108.02 \approx 0.9947$ so that the daily simple return is $-0.53\%$, which is $(107.45 - 108.02)/108.02$.

| Date | 12/23 | 12/24 | 12/28 | 12/29 | 12/30 | 12/31 |
|---|---|---|---|---|---|---|
| Price($) | 108.02 | 107.45 | 106.24 | 108.15 | 106.74 | 104.69 |

**Time interval** is important! Default is one year.
**Annualized (average) return**:

$$\text{Annualized}[R_t(k)] = \left[\prod_{j=0}^{k-1}(1 + R_{t-j})\right]^{1/k} - 1.$$

An approximation:

$$\text{Annualized}[R_t(k)] \approx \frac{1}{k}\sum_{j=0}^{k-1} R_{t-j}.$$

Continuously compounding: Illustration of the power of compounding (int. rate 10% per annum)

| Type | #(payment) | Int. | Net |
|---|---|---|---|
| Annual | 1 | 0.1 | $1.10000 |
| Semi-Annual | 2 | 0.05 | $1.10250 |
| Quarterly | 4 | 0.025 | $1.10381 |
| Monthly | 12 | 0.0083 | $1.10471 |
| Weekly | 52 | $\frac{0.1}{52}$ | $1.10506 |
| Daily | 365 | $\frac{0.1}{365}$ | $1.10516 |
| Continuously | $\infty$ | | $1.10517 |

$$A = C \exp[r \times n]$$

where $r$ is the interest rate per annum, $C$ is the initial capital, $n$ is the number of years, and exp is the exponential function.

**Present value**:

$$C = A \exp[-r \times n]$$

Continuously compounded (or log) return

$$r_t = \ln(1 + R_t) = \ln \frac{P_t}{P_{t-1}} = p_t - p_{t-1},$$

where $p_t = \ln(P_t)$.

Multiperiod log return:

$$\begin{aligned} r_t(k) &= \ln[1 + R_t(k)] \\ &= \ln[(1 + R_t)(1 + R_{t-1}) \cdots (1 + R_{t-k+1})] \\ &= \ln(1 + R_t) + \ln(1 + R_{t-1}) + \cdots + \ln(1 + R_{t-k+1}) \\ &= r_t + r_{t-1} + \cdots + r_{t-k+1}. \end{aligned}$$

**Example**. Consider again the Apple stock price.

1. What is the log return from 12/23 to 12/24:

   A: $r_t = \ln(107.45) - \ln(108.02) = -0.529\%$.

2. What is the log return from day 12/23 to 12/31?

   A: $r_t(6) = \ln(104.69) - \ln(108.02) = -3.13\%$.

Portfolio return: $N$ assets

$$R_{p,t} = \sum_{i=1}^{N} w_i R_{it}$$

**Example**: An investor holds stocks of IBM, Microsoft and Citi-Group. Assume that her capital allocation is 30%, 30% and 40%. Use the monthly simple returns in Table 1.2 of the text. What is the mean simple return of her stock portfolio?

**Answer**: $E(R_t) = 0.3 \times 1.35 + 0.3 \times 2.62 + 0.4 \times 1.17 = 1.66$.
Dividend payment:

$$R_t = \frac{P_t + D_t}{P_{t-1}} - 1, \quad r_t = \ln(P_t + D_t) - \ln(P_{t-1}).$$

Excess return: (adjusting for risk)

$$Z_t = R_t - R_{0t}, \quad z_t = r_t - r_{0t}$$

where $r_{0t}$ denotes the log return of a reference asset (e.g. risk-free interest rate).
Relationship:

$$r_t = \ln(1 + R_t), \quad R_t = e^{r_t} - 1.$$

If the returns are in **percentage**, then

$$r_t = 100 \times \ln(1 + \frac{R_t}{100}), \quad R_t = [\exp(r_t/100) - 1] \times 100.$$

Temporal aggregation of the returns produces

$$\begin{aligned} 1 + R_t(k) &= (1 + R_t)(1 + R_{t-1}) \cdots (1 + R_{t-k+1}), \\ r_t(k) &= r_t + r_{t-1} + \cdots + r_{t-k+1}. \end{aligned}$$

These two relations are important in practice, e.g. obtain annual returns from monthly returns.

**Example**. If the monthly log returns of an asset are 4.46%, −7.34% and 10.77%, then what is the corresponding quarterly log return?
**Answer**: $4.46 - 7.34 + 10.77 = 7.89\%$.
**Example**: If the monthly simple returns of an asset are 4.46%, −7.34% and 10.77%, then what is the corresponding quarterly simple return?
**Answer**: $R = (1+0.0446)(1-0.0734)(1+0.1077)-1 = 1.0721-1$
$= 0.0721 = 7.21\%$

**Distributional properties of returns**

Key: What is the distribution of

$\{r_{it}; i = 1, \cdots, N; t = 1, \cdots, T\}$?

**Some theoretical properties**:

Moments of a random variable $X$ with density $f(x)$: $\ell$-th moment

$$m'_\ell = E(X^\ell) = \int_{-\infty}^{\infty} x^\ell f(x) dx$$

First moment: mean or expectation of $X$.

$\ell$-th central moment

$$m_\ell = E[(X - \mu_x)^\ell] = \int_{-\infty}^{\infty} (x - \mu_x)^\ell f(x) dx,$$

2nd central moment: Variance of $X$.

standard deviation: square-root of variance

Skewness (symmetry) and kurtosis (fat-tails)

$$S(x) = E\left[\frac{(X - \mu_x)^3}{\sigma_x^3}\right], \quad K(x) = E\left[\frac{(X - \mu_x)^4}{\sigma_x^4}\right].$$

$K(x) - 3$: Excess kurtosis.

Q1: Why study the mean and variance of returns?

They are concerned with long-term return and risk, respectively.

Q2: Why is symmetry important?

Symmetry has important implications in holding short or long financial positions and in risk management.

Q3: Why is kurtosis important?

Related to volatility forecasting, efficiency in estimation and tests

High kurtosis implies heavy (or long) tails in distribution.

**Estimation**:

Data:$\{x_1, \cdots, x_T\}$

- sample mean:
$$\hat{\mu}_x = \frac{1}{T}\sum_{t=1}^{T} x_t,$$
- sample variance:
$$\hat{\sigma}_x^2 = \frac{1}{T-1}\sum_{t=1}^{T}(x_t - \hat{\mu}_x)^2,$$
- sample skewness:
$$\hat{S}(x) = \frac{1}{(T-1)\hat{\sigma}_x^3}\sum_{t=1}^{T}(x_t - \hat{\mu}_x)^3,$$
- sample kurtosis:
$$\hat{K}(x) = \frac{1}{(T-1)\hat{\sigma}_x^4}\sum_{t=1}^{T}(x_t - \hat{\mu}_x)^4.$$

Under normality assumption,
$$\hat{S}(x) \sim N(0, \frac{6}{T}), \quad \hat{K}(x) - 3 \sim N(0, \frac{24}{T}).$$

Some simple tests for normality (for large $T$).

1. Test for symmetry:
$$S^* = \frac{\hat{S}(x)}{\sqrt{6/T}} \sim N(0,1)$$
   if normality holds.
   **Decision rule:** Reject $H_o$ of a symmetric distribution if $|S^*| > Z_{\alpha/2}$ or p-value is less than $\alpha$.

2. Test for tail thickness:
$$K^* = \frac{\hat{K}(x) - 3}{\sqrt{24/T}} \sim N(0,1)$$

if normality holds.
**Decision rule:** Reject $H_o$ of normal tails if $|K^*| > Z_{\alpha/2}$ or p-value is less than $\alpha$.

3. A joint test (Jarque-Bera test):

$$JB = (K^*)^2 + (S^*)^2 \sim \chi^2_2$$

if normality holds, where $\chi^2_2$ denotes a chi-squared distribution with 2 degrees of freedom.
**Decision rule:** Reject $H_o$ of normality if $JB > \chi^2_2(\alpha)$ or p-value is less than $\alpha$.

### Empirical properties of returns

Data sources: Use packages, e.g. quantmod

- Course web:

- CRSP: Center for Research in Security Prices (Wharton WRDS) https://wrds-web.wharton.upenn.edu/wrds/

- Various web sites, e.g. Federal Reserve Bank at St. Louis https://research.stlouisfed.org/fred2/

- Data sets of the textbook: http://faculty.chicagobooth.edu/ruey.tsay/teaching/fts3/

Empirical dist of asset returns tends to be skewed to the left with heavy tails and has a higher peak than normal dist. See Table 1.2 of the text.

### Demonstration of Data Analysis

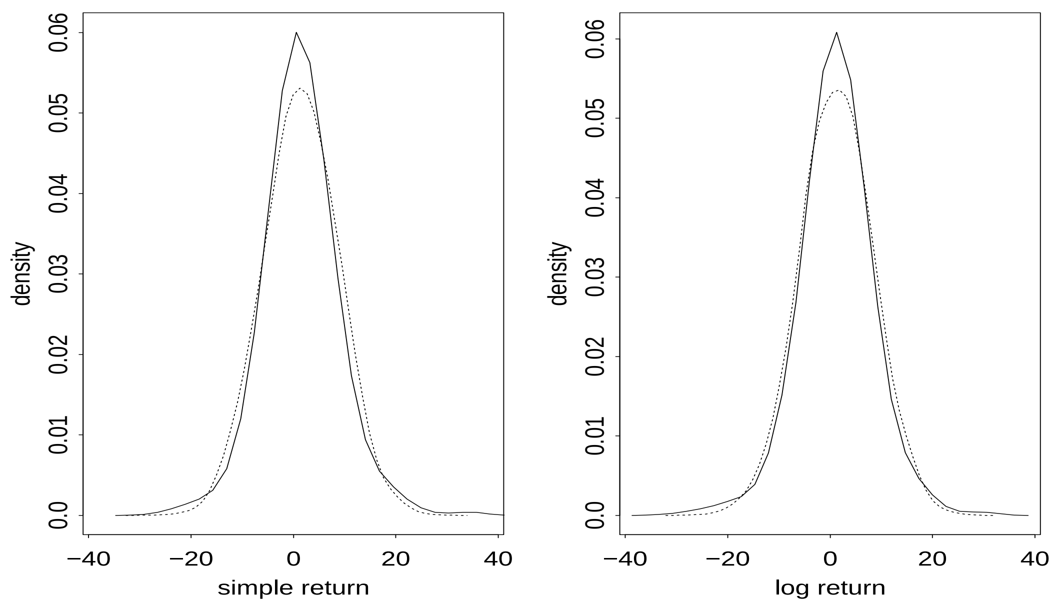


Figure 12: Comparison of empirical IBM return densities (solid) with Normal densities (dashed)

**R demonstration**: Use monthly IBM stock returns from 1967 to 2008.

```
**** Task: (a) Set the working directory
           (b) Load the library ``fBasics''.
           (c) Compute summary (or descriptive) statistics
           (d) Perform test for mean return being zero.
           (e) Perform normality test using the Jaque-Bera method.
           (f) Perform skewness and kurtosis tests.

> setwd("...")                                    <== set working directory
> library(fBasics)    <== Load the library ``fBasics''.

> da=read.table("m-ibm-6815.txt",header=T)
> head(da)
  PERMNO     date    PRC ASKHI  BIDLO       RET    vwretd    ewretd    sprtrn
1  12490 19680131 594.50 623.0 588.75 -0.051834 -0.036330  0.023902 -0.043848
2  12490 19680229 580.00 599.5 571.00 -0.022204 -0.033624 -0.056118 -0.031223
3  12490 19680329 612.50 612.5 562.00  0.056034  0.005116 -0.011218  0.009400
4  12490 19680430 677.50 677.5 630.00  0.106122  0.094148  0.143031  0.081929
5  12490 19680531 357.00 696.0 329.50  0.055793  0.027041  0.091309  0.011169
6  12490 19680628 353.75 375.0 346.50 -0.009104  0.011527  0.016225  0.009120
> dim(da)
[1] 576   9
> ibm=da$RET  % Simple IBM return
> lnIBM <- log(ibm+1) % compute log return
> ts.plot(ibm,main="Monthly IBM simple returns: 1968-2015") % Time plot
> mean(ibm)
[1] 0.008255663
> var(ibm)
[1] 0.004909968
> skewness(ibm)
[1] 0.2687105
attr(,"method")
[1] "moment"
> kurtosis(ibm)
[1] 2.058484
attr(,"method")
[1] "excess"
> basicStats(ibm)
                  ibm
nobs        576.000000
NAs           0.000000
Minimum      -0.261905
Maximum       0.353799
1. Quartile  -0.034392
```

```
3. Quartile   0.048252
Mean          0.008256
Median        0.005600
Sum           4.755262
SE Mean       0.002920
LCL Mean      0.002521
UCL Mean      0.013990
Variance      0.004910
Stdev         0.070071
Skewness      0.268710
Kurtosis      2.058484
> basicStats(lnIBM)  % log return
                 lnIBM
nobs        576.000000
NAs           0.000000
Minimum      -0.303683
Maximum       0.302915
1. Quartile  -0.034997
3. Quartile   0.047124
Mean          0.005813
Median        0.005585
Sum           3.348008
SE Mean       0.002898
LCL Mean      0.000120
UCL Mean      0.011505
Variance      0.004839
Stdev         0.069560
Skewness     -0.137286
Kurtosis      1.910438
> t.test(lnIBM) %% Test mean=0 vs mean .not. zero

        One Sample t-test

data:  lnIBM
t = 2.0055, df = 575, p-value = 0.04538
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 0.0001199015 0.0115051252
sample estimates:
  mean of x
0.005812513

> normalTest(lnIBM,method='jb')
Title: Jarque - Bera Normalality Test

Test Results:
```

```
  STATISTIC:
    X-squared: 90.988
  P VALUE:
    Asymptotic p Value: < 2.2e-16

> s3=skewness(lnIBM); T <- length(lnIBM)
> tst <- s3/sqrt(6/T)  % test skewness
> tst
[1] -1.345125
> pv <- 2*pnorm(tst)
> pv
[1] 0.1785849
> k4 <- kurtosis(lnIBM)
> tst <- k4/sqrt(24/T)  % test excess kurtosis
> tst
[1] 9.359197
>q()   % quit R.
```

**Normal and lognormal dists**

$Y$ is lognormal if $X = \ln(Y)$ is normal.
If $X \sim N(\mu, \sigma^2)$, then $Y = \exp(X)$ is lognormal with
Mean and variance:

$$E(Y) = \exp(\mu + \frac{\sigma^2}{2}), \quad V(Y) = \exp(2\mu + \sigma^2)[\exp(\sigma^2) - 1].$$

Conversely, if $Y$ is lognormal with mean $\mu_y$ and variance $\sigma_y^2$, then $X = \ln(Y)$ is normal with mean and variance

$$E(X) = \ln\left[\frac{\mu_y}{\sqrt{1 + \frac{\sigma_y^2}{\mu_y^2}}}\right], \quad V(X) = \ln\left[1 + \frac{\sigma_y^2}{\mu_y^2}\right].$$

**Application**: If the log return of an asset is normally distributed with mean 0.0119 and standard deviation 0.0663, then what is the mean and standard deviation of its simple return?
**Answer**: Solve this problem in two steps.

**Step 1**: Based on the prior results, the mean and variance of $Y_t = \exp(r_t)$ are

$$E(Y) = \exp\left[0.0119 + \frac{0.0663^2}{2}\right] = 1.014$$

$$V(Y) = \exp(2 \times 0.0119 + 0.0663^2)[\exp(0.0663^2) - 1] = 0.0045$$

**Step 2**: Simple return is $R_t = \exp(r_t) - 1 = Y_t - 1$. Therefore,

$$E(R) = E(Y) - 1 = 0.014$$

$$V(R) = V(Y) = 0.0045, \quad \text{standard dev} = \sqrt{V(R)} = 0.067$$

**Remark**: See the monthly IBM stock returns in Table 1.2.

**Processes considered**

- return series (e.g., ch. 1, 2, 5)
- volatility processes (e.g., ch. 3, 4, 10, 12)
- continuous-time processes (ch. 6)
- extreme events (ch. 7)
- multivariate series (ch. 8, 9, 10)

**Likelihood function** (for self study)
Finally, it pays to study the likelihood function of returns $\{r_1, \cdots, r_T\}$ discussed in Chapter 1.
**Basic concept**:
Joint dist = Conditional dist × Marginal dist, i.e.

$$f(x, y) = f(x|y)f(y)$$

For two consecutive returns $r_1$ and $r_2$, we have

$$f(r_2, r_1) = f(r_2|r_1)f(r_1).$$

For three returns $r_1$, $r_2$ and $r_3$, by repeated application,

$$\begin{aligned} f(r_3, r_2, r_1) &= f(r_3|r_2, r_1)f(r_2, r_1) \\ &= f(r_3|r_2, r_1)f(r_2|r_1)f(r_1). \end{aligned}$$

In general, we have

$$\begin{aligned} & f(r_T, r_{T-1}, \cdots, r_2, r_1) \\ = & f(r_T|r_{T-1}, \cdots, r_1)f(r_{T-1}, \cdots, r_1) \\ = & f(r_T|r_{T-1}, \cdots, r_1)f(r_{T-1}|r_{t-2}, \cdots, r_1)f(r_{T-2}, \cdots, r_1) \\ = & \vdots \\ = & \left[\prod_{t=2}^{T} f(r_t|r_{t-1}, \cdots, r_1)\right] f(r_1), \end{aligned}$$

where $\Pi_{t=2}^{T}$ denotes product.
If $r_t|r_{t-1}, \cdots, r_1$ is normal with mean $\mu_t$ and variance $\sigma_t^2$, then likelihood function becomes

$$f(r_T, r_{T-1}, \cdots, r_1) = \prod_{t=2}^{T} \frac{1}{\sqrt{2\pi}\sigma_t} \exp\left[\frac{-(r_t - \mu_t)^2}{2\sigma_t^2}\right] f(r_1).$$

For simplicity, if $f(r_1)$ is ignored, then the likelihood function becomes

$$f(r_T, r_{T-1}, \cdots, r_1) = \prod_{t=2}^{T} \frac{1}{\sqrt{2\pi}\sigma_t} \exp\left[\frac{-(r_t - \mu_t)^2}{2\sigma_t^2}\right].$$

This is the *conditional* likelihood function of the returns under normality.
Other dists, e.g. Student-$t$, can be used to handle heavy tails.

## Model specification

- $\mu_t$: discussed in Chapter 2
- $\sigma_t^2$: Chapters 3 and 4.

**Quantifying dependence**: Consider two variables $X$ and $Y$.

- Pearson's correlation coefficient:

$$\rho = \frac{\text{Cov}(X,Y)}{\text{std}(X)\text{std}(Y)}.$$

- Kendall's tau: Let $(\tilde{X}, \tilde{Y})$ be a random copy of $(X, Y)$.

$$\begin{aligned} \rho_\tau &= P[(X - \tilde{X})(Y - \tilde{Y}) > 0] - P[(X - \tilde{X})(Y - \tilde{Y}) < 0] \\ &= E[\text{sign}[(X - \tilde{X})(Y - \tilde{Y})]]. \end{aligned}$$

  This measure quantifies the probability of *concordant* over *discordant*. Here concordant means $(X - \tilde{X})(Y - \tilde{Y}) > 0$. For spherical distributions, e.g., normal, $\rho_\tau = \frac{2}{\pi}\sin^{-1}(\rho)$.

- Spearman's rho: rank correlation. Let $F_x(x)$ and $F_y(y)$ be the cumulative distribution function of $X$ and $Y$.

$$\rho_s = \rho(F_x(X), F_y(Y)).$$

  That is, the correlation coefficient of probability-transformed variables. It is just the correlation coefficient of the **ranks** of the data.

Q: Why do we consider different measures of dependence?

- Correlation coefficient encounters problems when the distributions are not normal (spherical, in general). This is particularly relevant in risk management.

- Correlation coefficient focuses no linear dependence and is not robust to outliers.
- The actual range of the correlation coefficient can be much smaller than $[-1, 1]$.

## R Demonstration

```
> head(da)
  PERMNO     date    PRC ASKHI  BIDLO       RET    vwretd    ewretd    sprtrn
1  12490 19680131 594.50 623.0 588.75 -0.051834 -0.036330  0.023902 -0.043848
2  12490 19680229 580.00 599.5 571.00 -0.022204 -0.033624 -0.056118 -0.031223
3  12490 19680329 612.50 612.5 562.00  0.056034  0.005116 -0.011218  0.009400
4  12490 19680430 677.50 677.5 630.00  0.106122  0.094148  0.143031  0.081929
5  12490 19680531 357.00 696.0 329.50  0.055793  0.027041  0.091309  0.011169
6  12490 19680628 353.75 375.0 346.50 -0.009104  0.011527  0.016225  0.009120
> ibm <- da$RET
> sp <- da$sprtrn
> plot(sp,ibm)
> cor(sp,ibm)
[1] 0.5785249
> cor(sp,ibm,method="kendall")
[1] 0.4172056
> cor(sp,ibm,method="spearman")
[1] 0.58267
> cor(rank(ibm),rank(sp))
[1] 0.58267

> z=rnorm(1000) %% Genreate 1000 random variates from N(0,1)
> x=exp(z)
> y=exp(20*z)
> cor(x,y)
[1] 0.3187030
> cor(x,y,method='kendall')
[1] 1
> cor(x,y,method='spearman')
[1] 1
```

## Takeaway

1. Understand the summary statistics of asset returns
2. Understand various definitions of returns & their relationships

3. Learn basic characteristics of FTS
4. Learn the basic **R** functions. (See Rcommands-lec1.txt on the course web.)

**R commands** used to produce plots in Lecture 1.

```
> x=read.table("d-aapl0413.txt",header=T)  <== Load Apple stock returns
> dim(x)         <== check the size of the data file
[1] 2517 3
> x[1,]         <== show the first row of the data
  Permno     date       rtn
1  14593 20040102 -0.004212
> y=ts(x[,3],frequency=252,start=c(2004,1))  <== Create a time-series object in R.
> plot(y,type='l',xlab='year',ylab='rtn')
> title(main='Daily returns of Apple stock: 2004 to 2013')

> par(mfcol=c(2,1))   <== To put two plots on a single page
> y=y*100             <== percentage returns
> hist(y,nclass=50)
> title(main='Percentage returns')
> d1=density(y)
> plot(d1$x,d1$y,xlab='returns',ylab='den',type='l')


> x=read.table("m-tb3ms.txt",header=T)   <== Load 3m-TB rates
> dim(x)
[1] 914   4

> y=read.table("m-tb6ms.txt",header=T)  <== Load 6m-TB rates
> dim(y)
[1] 615   4
> 914-615
[1] 299
> x[300,]   <== Check date of the 3m-TB
    year mon day value
300 1958  12   1  2.77
> y[1,]     <== Check date of the 1st observation of 6m-TB
  year mon day value
1 1958  12   1  3.01

> int=cbind(x[300:914,4],y[,4])  <== Line up the two TB rates
> tdx=(c(1:615)+11)/12+1959
> par(mfcol=c(1,1))
> max(int)
```

```
[1] 16.3
> plot(tdx,int[,1],xlab='year',ylab='rate',type='l',ylim=c(0,16.5))
> lines(tdx,int[,2],lty=2)   <== Plot the 6m-TB rate on the same frame.

> plot(tdx,int[,2]-int[,1],xlab='year',ylab='spread',type='l')
> abline(h=c(0))  <== Draw a horizontal like to ``zero''.

> x=read.table("q-ko-earns8309.txt",header=T)  <== Load KO data
> dim(x)
[1] 107   3
> x[1,]
     pends  anntime  value
1 19830331 19830426 0.0375
> tdx=c(1:107)/12+1983
> plot(tdx,x[,3],xlab='year',ylab='earnings',type='l')
> title(main='EPS of Coca Cola: 1983-2009')
> points(tdx,x[,3])
>
> y=read.table("d-exuseu.txt",header=T)  <== Load USEU exchange rates
> dim(y)
[1] 3567    4
> y[1,]
  year mon day  value
1 1999   1   4 1.1812
> tdx=c(1:3567)/252+1999
> plot(tdx,y[,4],xlab='year',ylab='eu',type='l')
> title(main='Dollars per Euro')

> r=diff(log(y[,4]))  <=== Compute log returns
> plot(tdx[2:3567],r,xlab='year',ylab='rtn',type='l')
> title(main='ln-rtn: US-EU')

> hist(r,nclass=50)
> title(main='useu: ln-rtn')
```

# Linear Time Series Models

Financial ts: a series of a financial instrument over time.
Example: log return $r_t$.
Data: $\{r_1, r_2, \ldots, r_T\}$ (sample size is $T$).
Goal: Construct a model for $\{r_t\}$?

## Basic Concepts

- Time series (ts): the ts $\{r_t,\ t = \ldots, -1, 0, 1, \ldots\}$ is a realization (sample path) of a stochastic process ($r_t$ is a random variable for each $t$). We define:
  - the mean function: $\mu(t) = \mathbb{E}(r_t)$
  - the autocovariance function at lag $h$: $\gamma(t+h, t) = \mathrm{Cov}(r_{t+h}, r_t)$.
- Stationary process: $r_t$ is a stationary ts if $\mu(t) \equiv \mu$ and $\gamma(t+h, t) = \gamma(h) \equiv \gamma_h$, i.e., the mean and autocovariance functions are independent of $t$. For stationary $r_t$, define the **autocorrelation function (ACF)** at lag $h$ as

$$\rho_h = \frac{\gamma_h}{\gamma_0} = \rho_{-h}, \qquad h = 1, 2, \ldots \ (\rho_0 = 1)$$

  Stationarity means that the first two moments (mean & ACF) do not change with time.

**Ex:** (White Noise = WN). If $r_t$ is *serially uncorrelated (efficient market)*, then:

$$\rho_h = 0, \qquad h = 1, 2, \ldots$$

**Exercise:** look back at the plots of the real series in the beginning and comment on the stationarity of each.

**Empirical estimates of first two moments of stationary series $r_t$**

- Mean:

$$\bar{r} = \frac{1}{T}\sum_{t=1}^{T} r_t$$

Asymptotic result:

$$\bar{r} \dot{\sim} \mathcal{N}(0, v^2/T), \qquad v^2 = \sum_{h=-\infty}^{\infty} \gamma_h$$

Note: if $r_t$ is WN then:

$$\bar{r} \dot{\sim} \mathcal{N}(0, \gamma_0/T), \qquad \gamma_0 = \mathbb{V}(r_t).$$

- ACF:

$$\hat{\rho}_h = \frac{\hat{\gamma}_h}{\hat{\gamma}_0}, \qquad \hat{\gamma}_h = \frac{1}{T}\sum_{t=1}^{T-h}(r_{t+h} - \bar{r})(r_t - \bar{r})$$

Note: if $r_t$ is WN then:

$$\hat{\rho}_h \dot{\sim} \mathcal{N}(0, 1/T), \qquad h = 1, 2, \ldots.$$

**Fundamental tests for stationary series $r_t$**

- $H_0 : \mu = 0 \quad vs \quad H_1 : \mu \neq 0$.
  Assuming that $r_t$ is WN, compute

$$t = \frac{\bar{r}}{\text{std}(\bar{r})} = \frac{\bar{r}}{\sqrt{\hat{\gamma}_0/T}}$$

p-value $= 2P(Z > |t|)$.

- $H_0 : \rho_1 = 0 \quad vs \quad H_1 : \rho_1 \neq 0$.
  Assuming that $r_t$ is WN, compute

$$t = \frac{\hat{\rho}_1}{\text{std}(\hat{\rho}_1)} = \frac{\hat{\rho}_1}{\sqrt{1/T}}$$

p-value $= 2P(Z > |t|)$.

Decision Rule: reject $H_0$ if p-value $< \alpha$ (the significance level).

- $H_0 : \rho_1 = \cdots = \rho_m = 0 \quad vs \quad H_1 : \rho_j \neq 0$ for some $j$.
  **Ljung-Box test**. Assuming that $r_t$ is WN, compute

$$Q(m) = T(T+2) \sum_{h=1}^{m} \frac{\hat{\rho}_h^2}{T-h}$$

  p-value $= P(\chi^2_{df} > Q(m))$, where $df = m$.
  Typical default values: $m \in \{5, 10, 15, 20\} \leq T - 1$.
  (Note: later when we fit a model with $g$ parameters $df = m - g$.)

- Sources of serial correlation

  - nonsynchronous trading (Ch 5)
  - bid-ask bounce (Ch 5)
  - risk premium (Ch 3)

  Thus: the presence of serial correlation does not necessarily imply market inefficiency.

**Ex:** (Monthly returns of IBM stock from 1926 to 1997).

- For $R_t$:

$$Q(5) = 5.4 \ (p = 0.37), \quad Q(10) = 14.1 \ (p = 0.17)$$

- For $r_t$:

$$Q(5) = 5.8 \ (p = 0.33), \quad Q(10) = 13.7 \ (p = 0.19)$$

Conclude: Monthly IBM stock returns are serially uncorrelated (WN).

**Ex:** (Monthly returns of CRSP value-weighted index from 1926 to 1997).

- For $R_t$:

$$Q(5) = 27.8 \ (p < 0.001), \quad Q(10) = 36.0 \ (p < 0.001)$$

- For $r_t$:

$$Q(5) = 26.9 \ (p < 0.001), \quad Q(10) = 32.7 \ (p < 0.001)$$

Conclude: Monthly CRSP returns are serially correlated (not WN).

## Back-shift or lag operator $(B)$

A useful notation in ts analysis.

- $Br_t = r_{t-1}$
- $B^2 r_t = B(Br_t) = B(r_{t-1}) = r_{t-2}$
- $B(3) = 3$, i.e., constants are unaffected.

## R demonstration:

IBM monthly simple returns from 1968 to 2015

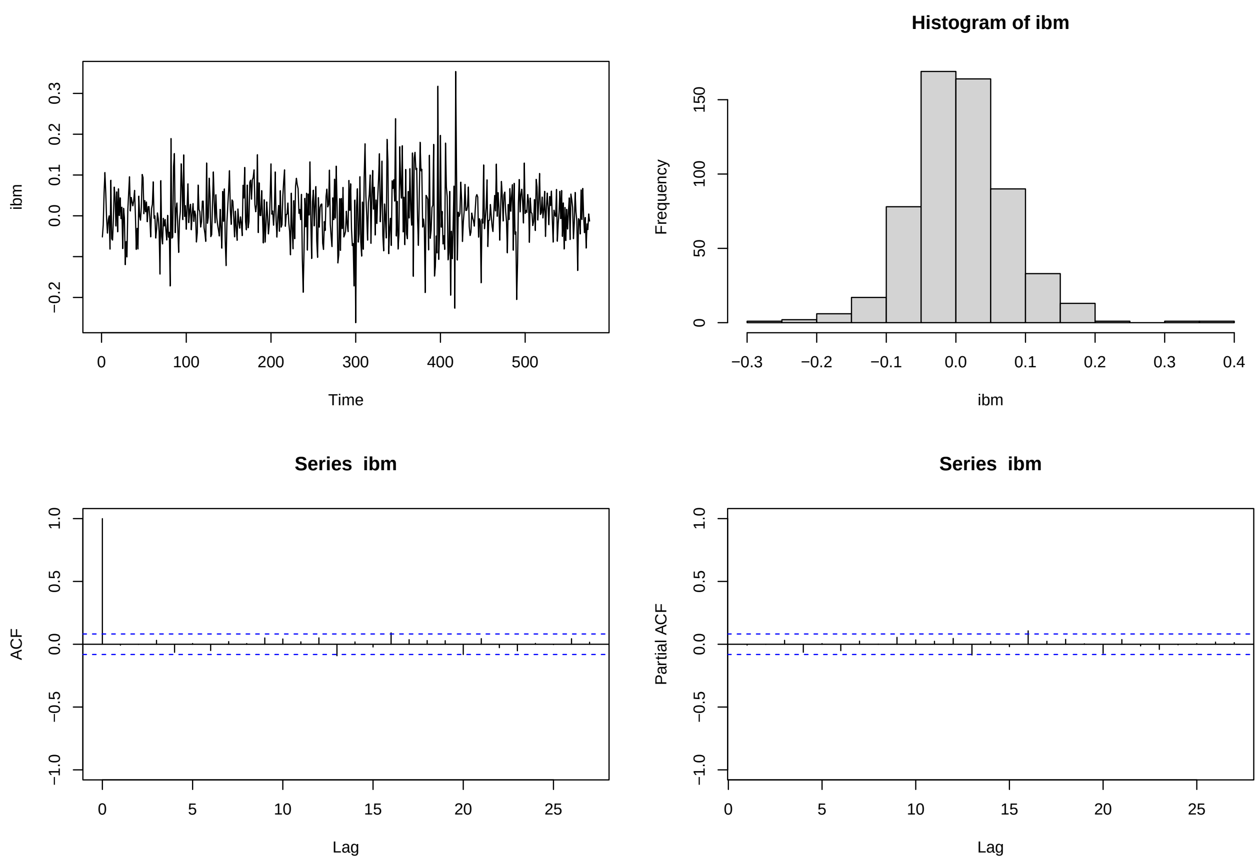


Figure 1: IBM monthly simple returns from 1968 to 2015.

```
> da=read.table("m-ibm-6815.txt",header=T)
> ibm=da$RET
> par(mfrow=c(2,2))
> plot.ts(ibm)
> hist(ibm)
> acf(ibm, ylim=c(-1,1))
> pacf(ibm, ylim=c(-1,1))

> m1 <- acf(ibm)
> names(m1)
[1] "acf"   "type"   "n.used"   "lag"   "series"   "snames"
> m1$acf
             [,1]
[1,] 1.0000000000 % lag 0
[2,] -0.0068713539 % lag 1
[3,] -0.0002212888
....
[28,] 0.0159729906

> m2 <- pacf(ibm) % Partial ACF (a variant of ACF)
> m2$pacf
             [,1]
[1,] 1.0000000000
[2,] -0.0068713539
[3,] -0.0002212888
....
[27,]  0.0127614307

> Box.test(ibm,lag=10) % Box-Pierce Q(m) test
Box-Pierce test
data: ibm
X-squared = 7.1714, df = 10, p-value = 0.7092

> Box.test(ibm,lag=10,type='Ljung') % Ljung-Box Q(m) test
Box-Ljung test
data: ibm
X-squared = 7.2759, df = 10, p-value = 0.6992
```

## Fundamental principle of financial time series modeling

At time point $t$, having observed: $\{r_1, \ldots, r_{t-1}\}$:

- The information set up to time $t-1$ is $F_{t-1} = \{r_1, \ldots, r_{t-1}\}$
- Decompose series into two parts as:

$$\begin{aligned} r_t &= \text{predictable part} + \text{unpredictable part} \\ &= \text{function of elements of } F_{t-1} + a_t \\ &= \mu_t + \sigma_t \epsilon_t \end{aligned}$$

  where

  - $\mu_t = \mathbb{E}(r_t|F_{t-1})$ = conditional mean of $r_t$
  - $a_t$ = shock or innovation at time $t$ (independent of $a_t$)
  - $\sigma_t^2 = \mathbb{V}(r_t|F_{t-1})$ = conditional variance of $r_t$ ($\sigma_t$ is called the *volatility*)
  - $\epsilon_t$ = an iid sequence with mean 0 and variance 1

- Typical models for $\mu_t$: ARMA/ARIMA (Lectures 2 & 3)
- Typical models for $\sigma_t$: ARCH/GARCH and variants (Lectures 4 & 5)

## Linear time series models

Stationary process $r_t$ is linear if:

$$r_t = \mu + \sum_{h=0}^{\infty} \psi_h a_{t-h}, \qquad a_t \sim \text{iid } (0, \sigma_a^2)$$

where $\mu = \mathbb{E}(r_t)$ is constant, the $\{a_t\}$ are the *shocks* or *innovations*, and the $\{\psi_h\}$ are the *impulse responses* of $r_t$.

Some examples:

- autoregressive (AR) models
- moving average (MA) models
- mixed ARMA models
- ARIMA and seasonal models
- regression models with time series errors
- fractionally differenced models (long-memory)

**Prediction of linear time series models**

Data: $\{r_1, \dots, r_n\}$ with $\mathbb{E}(r_t) = \mu$ and ACF $\rho_h$
Question: how to predict $r_{n+1}$?
Answer: use *Decision Theory*; we want

$$\widehat{r}_{n+1} = \text{function of } F_n \text{ (information set up to time } n\text{)}$$

such that (both):

(i) $\mathbb{E}(\widehat{r}_{n+1}) = \mu$

(ii) $\mathbb{E}(\widehat{r}_{n+1} - r_{n+1})^2$ is minimized.

Can show that the solution is the *Best Predictor* (BP):

$$\widehat{r}_{n+1} = \mathbb{E}(r_{n+1}|F_n)$$

For an AR($p$) this is simply:

$$\begin{aligned}
\widehat{r}_{n+1} &= \mathbb{E}(r_{n+1}|F_n) \\
&= \mathbb{E}\left(\phi_1 r_n + \cdots + \phi_p r_{n+1-p} + a_t|F_n\right) \\
&= \phi_1 r_n + \cdots + \phi_p r_{n+1-p} + \mathbb{E}\left(a_t|F_n\right) \\
&= \phi_1 r_n + \cdots + \phi_p r_{n+1-p}
\end{aligned}$$

For MA and ARMA models this is more complicated…(Lecture 2).

**Note:** This course uses statistical methods to find models that fit the data well for making inference, e.g. prediction. On the other hand, there exists economic theory that leads directly to time-series models. E.g., in the *real business-cycle theory* in macroeconomics, one can show that under some simplifying assumptions log(GDP) follows an AR(2). (See *Advanced Macroeconomics*, by David Romer, 2006).

**Ex:** (Quarterly growth rate of US real GNP, seasonally adjusted, from 1947 to 1991). An AR(3) model for the data is:

$$r_t = 0.005 + 0.35r_{t-1} + 0.18r_{t-2} - 0.14r_{t-3} + a_t, \qquad a_t \sim \text{WN}(0, \hat{\sigma}_a^2 = 0.01)$$

Given $\{r_n, r_{n-1}, r_{n-2}\}$ we can predict $r_{n+1}$ as:

$$\hat{r}_{n+1} = 0.005 + 0.35r_n + 0.18r_{n-1} - 0.14r_{n-2}$$

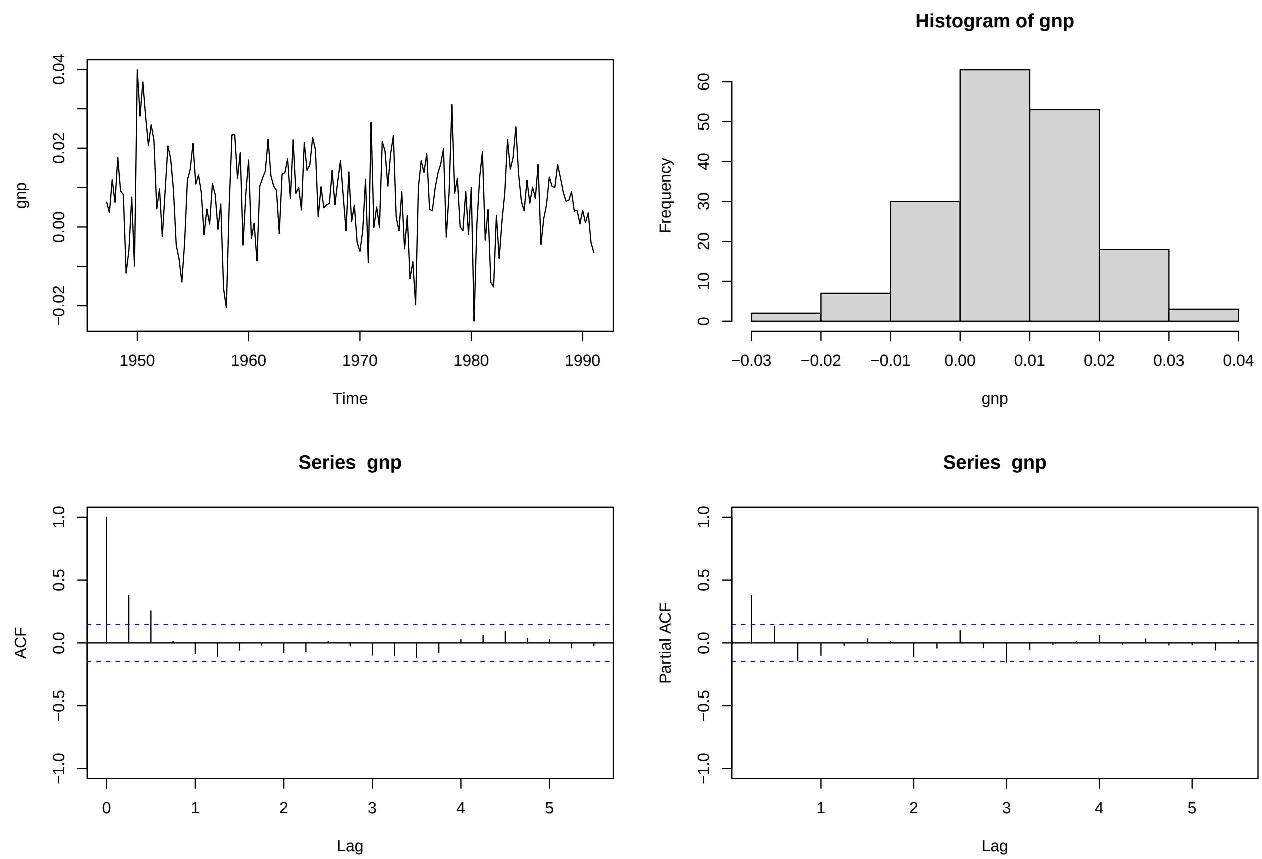


Figure 2: US real GNP quarterly growth rate: 1947 to 1991.

**Steps in time series model building**

- Check for stationarity, are there:
  - trends
  - cycles (seasonality is expected with quarterly and monthly data)
  - unit roots (Lecture 2)
- Estimate basic moments: mean, variance, ACF, PACF
- Model building:
  - specification (model/order selection)
  - estimation (maximum likelihood)
  - assessment (goodness-of-fit diagnostics for residuals)
- Make inference (e.g., forecasting).

**Ex:** (Monthly simple returns of CRSP equally-weighted index, from 1926 to 2008)
Fig 3: ACF of series suggests autocorrelation (tenuous).
A subset MA(9) model for the series is:

$$R_t = 0.012 + a_t + 0.189a_{t-1} - 0.121a_{t-3} + 0.122a_{t-9}, \qquad a_t \sim \mathrm{WN}(0, \hat{\sigma}_a^2 = 0.07)$$

Model checking for the residuals: $Q(10) = 11.4$ ($p = 0.122$), suggests model fits!
Fourth panel shows 10-step ahead forecasts based on model that leaves out last 10 data points (blue line); actual values shown as red line.

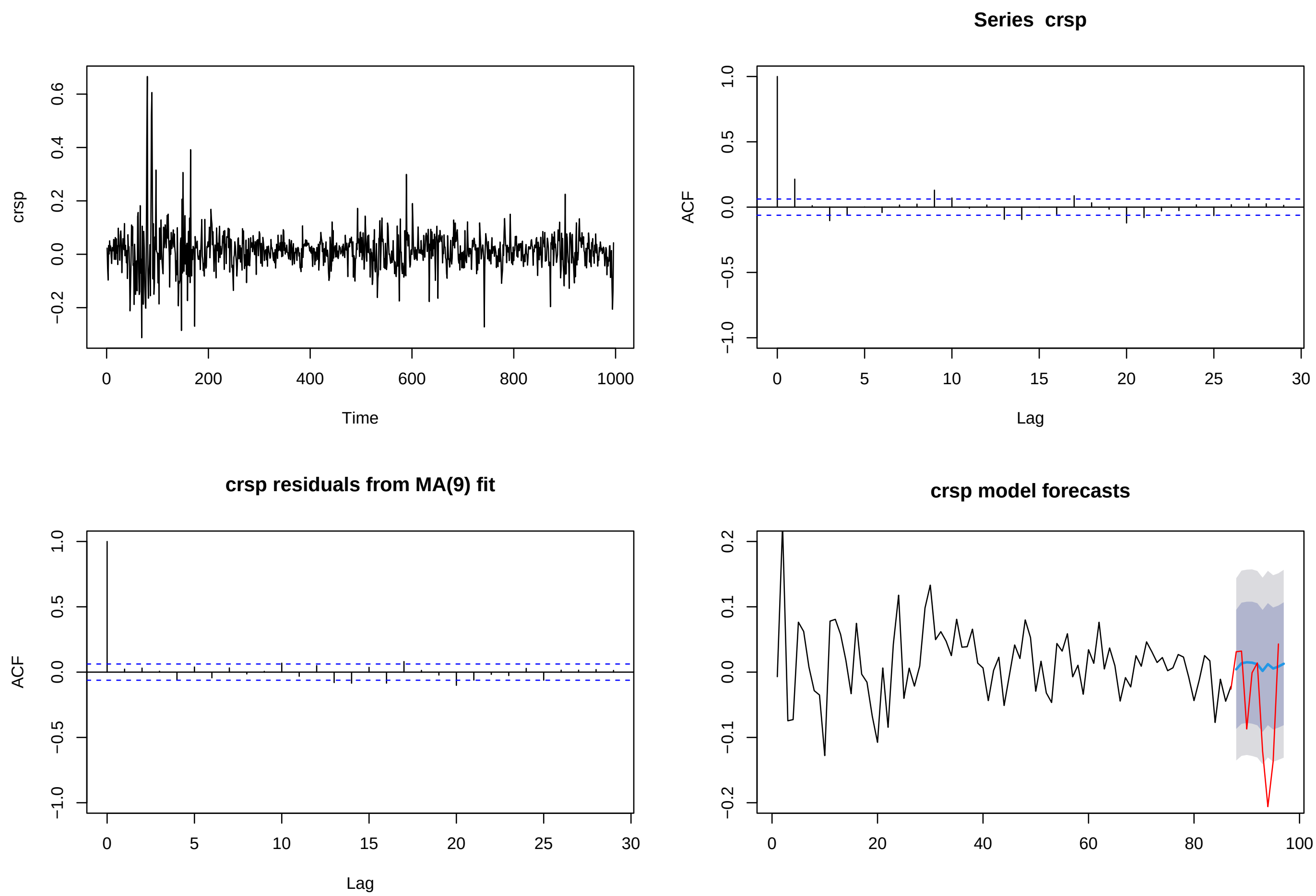


Figure 3: Monthly simple returns of CRSP equally-weighted index: 1926 to 2008.