Introduction to Financial Econometrics

Chapter 1: Describing Financial Series

Christophe Hurlin

Univ Orléans

May, 2019

1. Introduction

The outline of this chapter is the following:

1. Introduction

Notations: In this chapter, I will (try to...) follow some conventions of notation.

Y	random variable
y	realization
$f_Y(y)$	probability density or mass function
$F_Y(y)$	cumulative distribution function
$\Pr()$	probability
$\mathbf{y}$	vector
$\mathbf{Y}$	matrix

Problem: this system of notations does not allow to discriminate between a vector (matrix) of random elements and a vector (matrix) of non-stochastic elements (realization).

Abadir and Magnus (2002), Notation in econometrics: a proposal for a standard, Econometrics Journal.

Section 2

Financial Econometrics

2. Financial Econometrics

Objectives

1. Define **financial econometrics**
2. Define the notions of **sample** and **data generating process**
3. Distinguish **parametric** and **non-parametric** model
4. Define the **cross-sectional** data
5. Define the **time series** data
6. Define the **panel** data
7. Summarize the **general approach** of financial econometrics

2. Financial Econometrics

What is financial econometrics? Not an easy question!

The main objective of this course is to define this term!

Definition (financial econometrics)

Financial econometrics is the application of statistical methods to financial market data. [..] It differs from other forms of econometrics because the emphasis is usually on analyzing the prices of financial assets traded at competitive, liquid markets.

Source: Wikipedia.

Objectives of Financial Econometrics

- Combining finance theory with statistical theory.
- Transferring a theoretical (financial) model into a (financial) econometric model.
- Modelling financial data.
- Predicting financial variables as well as relations thereof.
- Developing specific model well-adapted to the stylized facts of financial data. Example: GARCH models

Is it useful?

"*People working in the finance industry or research in the finance sector often use* **econometric techniques** *in a range of activities – for example, in support of* **portfolio management** *and in the* **valuation of securities**. *Financial econometrics is essential for* **risk management** *when it is important to know how often 'bad' investment outcomes are expected to occur over future days, weeks, months and years"*.

Source: Wikipedia.

Academic development

Financial econometrics attracted substantial attention in recent years in both academia as well as financial practice.

- 2003: Nobel prize for **Robert F. Engle** for the GARCH models (cf chapter 6).
- 2003: Foundation of the **Journal of Financial Econometrics**.
- 2007: Foundation of the **Society for Financial Econometrics (SoFiE)**.

2. Financial Econometrics

Remark

There is sometimes a confusion between the following terms

1. **Financial econometrics** bridges the gap between financial economics, statistics and mathematical finance.
2. **Empirical finance** covers all the finance studies based on data analysis, e.g. *Journal of Empirical Finance*.
3. **Financial economics** is a "*highly empirical discipline, perhaps the most empirical among the branches of economics and even among the social sciences in general*" (Campbell, Lo, and MacKinlay, 1997)

2. Financial Econometrics

"Econometrics is the quantitative analysis of actual economic phenomena based on the concurrent development of theory and observation, related by appropriate methods of inference", P. A. Samuelson, T. C. Koopmans, and J. R. N. Stone (1954)

(Financial) econometrics is fundamentally based on four elements:

1. A **sample** of data
2. An econometric **model**
3. An **estimation** method
4. Some **inference** methods

Question: Why using a sample?

- Let us assume that we want to study a **characteristic / property** x of the **individuals** of a **population**.
- The individuals (unit) of the population are not necessarily some persons: it can be firms, assets, countries, time index etc..
- The characteristic x may be **quantitative** (price, returns, total asset, etc.) or **qualitative** (default, sector, etc.)
- The characteristic x may be **stochastic** or **deterministic**.

Definition (Population)

A **population** can be defined as including all people or items with the characteristic one wishes to understand.

1. In most of cases, it is impossible to observe the entire statistical population, due to cost constraints, time constraints, constraints of geographical accessibility.

2. A researcher would instead observe a statistical **sample** from the **population** in order to attempt to **learn something about the population as a whole.**

In most of cases, the sample is random:

Definition (Probability sampling)

A probability sampling is a **sampling method** in which every unit in the population has a chance (greater than zero) of being selected in the sample.

Consequence: a sample is a collection of **random variables** even the characteristic x is deterministic.

$$\text{sample: } \{X_1, X_2, ..., X_N\}$$

Example (random sample)

Let us consider a population of four firms and denote by $\widetilde{x}$ the characteristic (assumed to be non stochastic) of the firm with:

$$\widetilde{x}_A = 80 \quad \widetilde{x}_B = 50 \quad \widetilde{x}_C = 40 \quad \widetilde{x}_D = 90$$

Consider a random sample of $N = 2$ firms denoted by

$$\left\{ \underbrace{X_1}_{\text{characteristic. of the first firm selected in the sample}}, X_2 \right\}$$

So we can obtain a realization

$$\{x_1, x_2\} = \{50, 80\} \text{ or } \{x_1, x_2\} = \{90, 40\} \text{ or } \{x_1, x_2\} = \{90, 90\} \text{ etc.}$$

2. Financial Econometrics

Fact (random sample - data set)

The result of the probability sampling is a **random sample***, i.e. a collection of random variables* $X_1, X_2, .., X_N$*. In general, only* **one realization** *of the sample is available: this is your* **data set***!*

$$\{x_1, x_2, .., x_N\}$$

Remark When ones consider two or more variables, the notion of population is replaced by the concept of Data Generating Process (DGP)

Definition (Data Generating Process)

A **Data Generating Process (DGP)** is the joint probability distribution that is supposed to characterize the entire population from which the data set has been drawn.

Example (Data Generating Process)

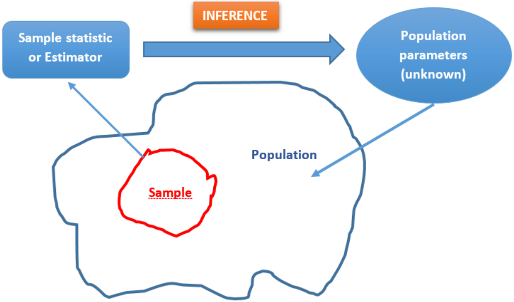
Let us assume that there is a linear relationship between the returns R_t and $R_{m,t}$ in the population, such that

$$\mathbb{E}\left(R_t \middle| R_{m,t} = r_m \right) = \alpha + \beta r_m$$

This relationship is the Data Generating Process for R_t.

2. Financial Econometrics

The **challenge** of econometrics is to draw conclusions about a population (or DGP) after observing only **one** realization $\{x_1, ..x_N\}$ of a random sample (your data set..).



2. Financial Econometrics

In econometrics, data sets can be mainly distinguished in three types:

1. Cross-sectional data
2. Time series data
3. Panel data

Cross-sectional data

- Data for different entities: assets, portfolios, firms, and so forth.
- No time dimension (even if date of data collection varies somewhat across units, it is ignored).
- Order of data does not matter!

2. Financial Econometrics

Time series data

- Data for a single entity (asset, firm, etc.) collected at multiple time periods. Repeated observations of the same variables (price, volume, etc.).
- Order of data is important!
- Observations are typically not independent over time;

Panel data or longitudinal data

- Data for multiple entities (asset, firm, etc.) in which outcomes and characteristics of each entity are observed at multiple points in time.
- Combine cross-sectional and time series issues.
- Present several advantages with respect to cross-sectional and time series data (depending on the question of interest!).

2. Financial Econometrics

Definition (Econometric model)

A **model** specifies the statistical relationship that is believed to hold between the various economic quantities pertaining to a particular economic phenomenon under study.

We can distinguish:

1. **Parametric model:** the relationship (**joint probability distribution**) between the dependent variable /vector Y and the explicative variables X is fully characterized by a set of parameters θ

$$Y = f(X;\theta) + \varepsilon$$

where link function $f(.)$ is assumed to be known.

2. **Non parametric and semi-parametric models:** the link function can not be described using a finite number of parameters. The link function is assumed to be unknown and has to be estimated.

2. Financial Econometrics

The general approach of (financial) econometrics is the following:

1. **Step 1:** Specification of the model (**Chapter 2**)
2. **Step 2:** Estimation of the parameters (**Chapter 2**)
3. **Step 3:** Diagnostic tests (**Chapter 3**)
    1. Significance tests;
    2. Specification tests;
    3. Backtesting tests;
    4. etc.
4. **Step 4:** Interpretation and use of the model (forecasting, historical study, etc.) (**Chapter 4**)

Key Concepts

1. Financial econometrics
2. Sample, population and data generating process
3. Parametric and non-parametric model
4. Cross-sectional data
5. Time series data
6. Panel data
7. General approach of financial econometrics

Section 3

Financial Data

3. Financial data

Objectives

1. Introduce the various notions of **asset prices**
2. Define the notion of **sampling frequency**
3. Distinguish **high frequency** data (intraday) vs. **low frequency** data
4. Define the **closing** and **opening** prices
5. Define the **adjusted** (closing) prices
6. Define the **simple (net) return**
7. Define the **continuously compounded** or **log-return**

Liquid markets

Quantitative financial research mainly focuses on "**liquid**" **financial markets**, i.e. organized markets where transactions are frequent and the number of actors is large.

Example ("liquid" markets)

Some example of "liquid" markets: foreign exchange market, organized futures markets, stock index markets and the market for large stocks

On these markets, prices are recorded several times a minutes or after each event (tick-by-tick data).

3. Financial data

Definition (prices)

Let P_t denotes the **price** of an asset at time t.

The prices may be observed (and recorded) at

1. An **irregular (sampling) frequency**: tick-by-tick price observations, volume-event observations (prices observed when the volume exceeds a given threshold), price-events (transactions associated with significant price changes), etc.

2. A **regular (sampling) frequency**: the prices are observed every m periods of time.

Banulescu D., Colletaz G., Hurlin C. and Tokpavi S. (2016), High Frequency Risk Measures, *Journal of Forecasting*, 35(3), 224-249.

Regular frequencies

- Most of the financial econometric studies are based on **regular frequencies**.
- In general, ones make a distinction between **intraday frequencies** (called **high frequencies**) and **lower frequencies** because the statistical properties of the corresponding data are different and the econometric tools used to model them, too.

Table 1: Some examples of sampling frequencies used for financial data

High frequency (intra-day) data	Low frequency data
10 seconds	**daily**
1 minute	weekly
5 minutes	**monthly**
10 minutes	quarterly
30 minutes	semi-annual
1 hour	annual
etc.	etc.

3. Financial data

Example (sampling frequency for prices)

Plot the (closing) prices for Intel Corporation (Ticker: INTC) from January 2015 to September 2018, sampled at the daily, monthly, yearly frequencies, respectively.

Note: The data are available at https://finance.yahoo.com

3. Financial data

Figure: Closing prices of Intel Corp, daily frequency (Jan 2013 - Sept 2018)

Source: Yahoo Finance

3. Financial data

Figure: Closing prices of Intel Corp, monthly frequency (Jan 2013 - Sept 2018)

Source: Yahoo Finance

3. Financial data

Figure: Closing prices of Intel Corp, yearly frequency (2010 - 2018)

Source: Yahoo Finance

Remark

For a given sampling frequency (e.g., daily), the prices can be recorded at different time within the observational period.

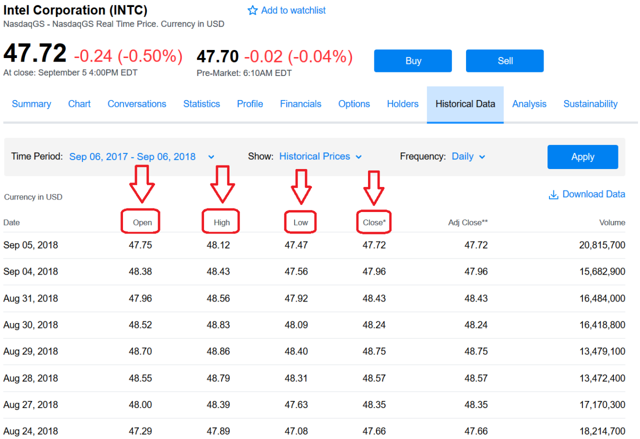
For instance, for a daily sampling frequency, we can consider:

- The **opening price** is the price at which a security first trades upon the opening of an exchange on a trading day.
- The **closing price** is the final price at which a security is traded on a given trading day, representing the most up-to-date valuation of a security
- The **highest** (or lowest) price of the trading day.
- The **adjusted closing prices** have been treated to correct for splits and dividends.

In general, we consider the **end-of-period price**, i.e. the closing price in the case of a daily frequency.

3. Financial data

Figure: Historical data available with Yahoo Finance (ticker: Intel Corp).



Intel Corporation (INTC) ☆ Add to watchlist

NasdaqGS - NasdaqGS Real Time Price. Currency in USD

47.72 -0.24 (-0.50%) **47.70** -0.02 (-0.04%) Buy Sell

At close: September 5 4:00PM EDT Pre-Market: 6:10AM EDT

Summary Chart Conversations Statistics Profile Financials Options Holders Historical Data Analysis Sustainability

Time Period: Sep 06, 2017 - Sep 06, 2018 Show: Historical Prices Frequency: Daily Apply

Currency in USD Download Data

Date	Open	High	Low	Close*	Adj Close**	Volume
Sep 05, 2018	47.75	48.12	47.47	47.72	47.72	20,815,700
Sep 04, 2018	48.38	48.43	47.56	47.96	47.96	15,682,900
Aug 31, 2018	47.96	48.56	47.92	48.43	48.43	16,484,000
Aug 30, 2018	48.52	48.83	48.09	48.24	48.24	16,418,800
Aug 29, 2018	48.70	48.86	48.40	48.75	48.75	13,479,100
Aug 28, 2018	48.55	48.79	48.31	48.57	48.57	13,472,400
Aug 27, 2018	48.00	48.39	47.63	48.35	48.35	17,170,300
Aug 24, 2018	47.29	47.89	47.08	47.66	47.66	18,214,700

Source: Yahoo Finance

Fact (prices vs returns)

Although ***prices*** *are what we observe in financial markets, most empirical studies are based on* ***returns****.*

- Returns are **scale-free** and have more attractive **statistical properties** than prices.
- Prices are, in general, **non-stationary** whereas returns are stationary.

We distinguish two types of return

1. The **simple (net) return**
2. The **continuously compounded** or **log-return**

3. Financial data

Definition (proportional return)

Holding a given asset for one period (from $t-1$ to t) yields the **one-period** (simple) **return** or **proportional return**:

$$\widetilde{R}_t \equiv \widetilde{R}_t[1] = \frac{P_t - P_{t-1}}{P_{t-1}}$$

where P_t is the (end-of-period) asset price at time t.

Definition (log-return)

Holding a given asset for one period (from $t-1$ to t) yields the **continuously compounded** or **log-return**:

$$R_t \equiv R_t[1] = \ln\left(\frac{P_t}{P_{t-1}}\right) = p_t - p_{t-1}$$

where p_t is the log-price of the asset at time t.

Remark

Simple return and log-return are very similar if $\widetilde{R}_t$ is close to zero;

$$\begin{aligned} R_t &= \ln\left(\frac{P_t}{P_{t-1}}\right) \\ &= \ln\left(1+\left(\frac{P_t}{P_{t-1}}-1\right)\right) \\ &= \ln\left(1+\widetilde{R}_t\right) \simeq \widetilde{R}_t \quad \text{if } \widetilde{R}_t \simeq 0 \end{aligned}$$

Example (simple and log-returns)

The closing price of the equity Intel Corporation (INTC) was equal to \$47.17 the August 16, 2018 and \$46.9 the August 17, then the daily return is equal to

$$\widetilde{R}_{17/08} = \frac{46.9 - 47.7}{47.7} = -0.0058 = -0.58\%$$

$$R_{17/08} = \ln\left(\frac{46.9}{47.7}\right) = -0.0057 = -0.57\%$$

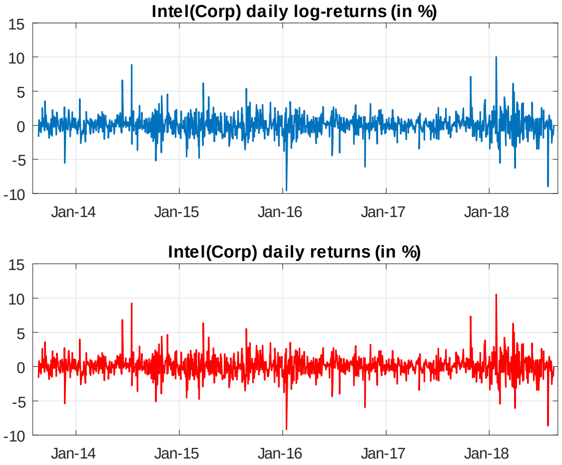
Example (Sample)

Consider a sample of **daily closing prices** for the equity Intel Corp. (ticker: INTC) from August 19, 2013 to August 17, 2018 (5 years). **Question:** compare the (simple) daily returns and the log-(daily) returns.

Note: The sample size T is equal to 1,259. Notice that each year corresponds to approximately 250 quotations, which implies about $250 \times 5 = 1,250$ observations for 5 years.

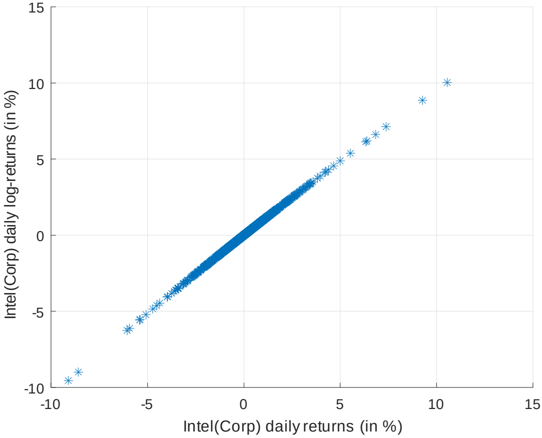
3. Financial data

Figure: Simple and log-returns for Intel(Corp), August 19, 2013 to August 17, 2018



3. Financial data

Figure: Scatter-plot of simple and log-returns for Intel(Corp), August 19, 2013 to August 17, 2018



Notations

For the rest of this chapter and the rest of the course, we will keep the following set-up:

- We observe the end-of-period **asset price** P_t at time t.
- We consider the corresponding one-period **(log-) returns** R_t.
- The **frequency** (periodicity) of the observations can be daily, weekly, monthly, etc.
- We collect a T-**sample** of prices $\{P_1, \ldots, P_T\}$, where T is the sample size.

Definition

The **k-period log-return** is given by:

$$R_t[k] = \sum_{j=0}^{k-1} R_{t-j}[1].$$

Proof: Log-returns are quite convenient since multiplication becomes addition!

$$\begin{aligned} R_t[k] &= \ln\left(\frac{P_t}{P_{t-k}}\right) \\ &= \ln\left(\frac{P_t}{P_{t-1}} \times \frac{P_{t-1}}{P_{t-2}} \times \ldots \times \frac{P_{t-k+2}}{P_{t-k+1}} \times \frac{P_{t-k+1}}{P_{t-k}}\right) \\ &= \sum_{j=0}^{k-1} \ln\left(\frac{P_{t-j}}{P_{t-j-1}}\right) = \sum_{j=0}^{k-1} R_{t-j}[1] \end{aligned}$$

3. Financial data

Definition (Excess return)

In many applications, it is convenient to use an **excess return** defined as

$$\text{excess return} = R_t - R_{f,t}$$

where $R_{f,t}$ is a reference rate.

Key Concepts

1. Prices and sampling frequency
2. High frequency data (intraday) vs low frequency data
3. Closing, opening, highest, lowest prices
4. Adjusted prices
5. Simple vs log-returns
6. k-periods (log-) return
7. Excess return

Section 4

Statistical Properties of Financial Data

4. Statistical properties of financial data

Objectives

1. Introduce the concept of **return distribution**
2. Define the central and non–central **moments** of the returns distribution
3. Define the mean, variance, **skewness**, and **kurtosis**
4. Define the notions of **sample**, **estimator**, and **estimate**
5. Study the **sampling distribution** of an estimator or a descriptive statistic

Notations

Definition (distribution of returns)

We represent the (log-) returns by the **real-valued random variable** R_t (or $R_t[k]$). The distribution of R_t is called the **distribution of returns**, or returns distribution

- The **probability density function (pdf)** of R_t is denoted $f_R(r)$.
- The **cumulative density function (cdf)** of R_t is denoted $F_R(r)$ such that

$$\Pr(R_t \leq r) \equiv F_R(r) = \int_{-\infty}^{r} f_R(x)\,\mathrm{d}x$$

Example (Normality assumption)

Let us assume that the return R_t has a normal distribution

$$R_t \sim \mathcal{N}\left(\mu, \sigma^2\right)$$

with μ the expected (mean of) return and σ^2 the variance of return. Then, the pdf and cdf of R_t are respectively defined by

$$f_R\left(r\right) = \frac{1}{\sigma}\phi\left(\frac{r-\mu}{\sigma}\right) = \frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{\left(r-\mu\right)^2}{2\sigma^2}\right)$$

$$F_R\left(r\right) = \Phi\left(\frac{r-\mu}{\sigma}\right) = \int_{-\infty}^{r} f_R\left(u\right)\mathrm{d}u$$

where $\phi\left(.\right)$ and $\Phi\left(.\right)$ respectively denote the pdf and cdf of the standard normal distribution.

The distribution of returns can be characterized either by its pdf or by its **moments**.

Definition (non-central and central moments)

The k^{th} **non-central moment** of R_t is given by

$$m_k = \mathbb{E}\left(R_t^k\right) = \int_{-\infty}^{\infty} x^k f_R(x)\,\mathrm{d}x$$

The k^{th} **central moment** of R_t is given by

$$\mu_k = \mathbb{E}\left(\left(R_t - m_1\right)^k\right) = \int_{-\infty}^{\infty} \left(x - m_1\right)^k f_R(x)\,\mathrm{d}x$$

with $m_1 = \mathbb{E}(R_t)$ the mean of the return distribution.

In financial econometrics, we are mainly interested by the fourth first moments:

Mean **Variance** **Skewness** **Kurtosis**

	Mean	Std. dev	Skew	Kurt	Max	Min	Q(15)	$Q_2(15)$
Durations								
RY	0.99	1.14	2.50	12.36	16.68	0.01	1325	387
PDG	0.93	1.50	7.68	170.59	63.33	0.01	5015	185
Returns								
RY	0.000	0.815	0.005	7.560	5.960	-6.422	8196	15417
PDG	0.001	0.890	-0.009	6.735	6.404	-6.421	2355	3118

The sample period runs from April 1st to June 30, 2001. It consists of 51,660 observations for the RY stock and 27,956 observations for the PDG stock. *Mean* is the sample mean, *Std. dev* is the sample standard deviation, *Skew* is the sample skewness coefficient, *Kurt* is the sample kurtosis, *Max* is the sample maximum, *Min* is the sample minimum. $Q(15)$ is the Ljung-Box test statistics with 15 lags, $Q_2(15)$ is the Ljung-Box test statistics applied to squared returns using 15 lags. The associated 95% critical value is 24.996.

Table 2: Descriptive Statistics of Deseasonalised Data for Royal Bank (RY) and Placer Dome (PDG) Stocks

Source: Dionne G., Duchesne, P., and Pacurar, M. (2005), Intraday Value at Risk (IVaR) using tick-by-tick data with application to the Toronto Stock Exchange, *Journal of Empirical Finance*, 16(5), 777-792.

Definition (mean and variance of return)

The **mean** (expected value) of the return R_t

$$\mu \equiv m_1 = \mathbb{E}\left(R_t\right) = \int_{-\infty}^{\infty} x \; f_R\left(x\right) \mathrm{d}x$$

The **variance** of the return R_t is

$$\sigma^2 \equiv \mathbb{V}\left(R_t\right) = \mathbb{E}\left(\left(R_t - \mu\right)^2\right) = \int_{-\infty}^{\infty} \left(x - \mu\right)^2 f_R\left(x\right) \mathrm{d}x$$

Definition (volatility)

The **volatility** corresponds to the standard deviation of the return distribution

$$\text{Volatility} \equiv \sigma = \sqrt{\mathbb{V}\left(R_t\right)}$$

Definition (skewness coefficient)

The third central moment measures the skewness of the distribution

$$\mu_3 = \mathbb{E}\left(\left(R_t - \mu\right)^3\right) = \int_{-\infty}^{\infty} \left(x - \mu\right)^3 f_R\left(x\right) \mathrm{d}x$$

The (standardized) **skewness coefficient** is defined as

$$\mathbb{S}\left(R_t\right) = \mathbb{E}\left(\left(\frac{R_t - \mu}{\sigma}\right)^3\right) = \frac{\mu_3}{\sigma^3}$$

Skewness

- When $\mathbb{S}(R_t) < 0$, large realizations of R_t are more often negative than positive (i.e., crashes are more likely than booms).
- The distribution is said to be **left-skewed**, **left-tailed**, or **skewed to the left**
- For **symmetric distributions** (e.g., normal, Student, etc.), the skewness coefficient is null

$$\mathbb{S}(R_t) = 0$$

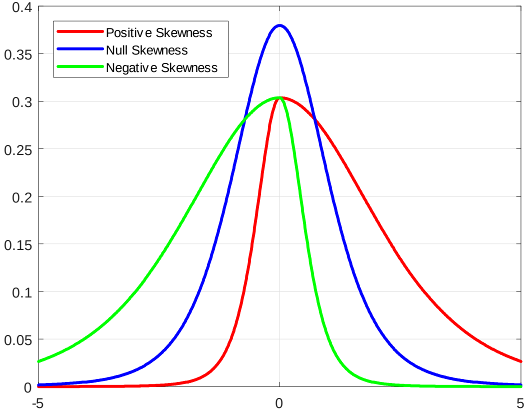


Figure: Illustration of the Skewness

Definition (kurtosis coefficient)

The fourth central moment measures the tail heaviness of the distribution

$$\mu_4 = \mathbb{E}\left(\left(R_t - \mu\right)^4\right) = \int_{-\infty}^{\infty} \left(x - \mu\right)^4 f_R\left(x\right) \mathrm{d}x$$

The (standardized) **kurtosis coefficient** is defined as

$$\mathbb{K}\left(R_t\right) = \mathbb{E}\left(\left(\frac{R_t - \mu}{\sigma}\right)^4\right) = \frac{\mu_4}{\sigma^4}$$

Kurtosis: properties

Fact (normal distribution)

For a **normal distribution** $R_t \sim \mathcal{N}\left(\mu, \sigma^2\right)$, *the kurtosis coefficient is equal to 3*

$$\mathbb{K}\left(R_t\right) = 3$$

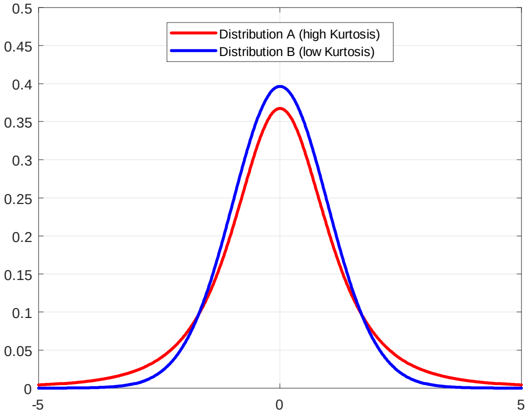
Note: the Gaussian distribution is not the only one to have a kurtosis equal to 3.

Kurtosis: properties (cont'd)

- Large $\mathbb{K}\left(R_t\right)$ implies that large realizations (positive or negative) are more likely to occur (compared to a normal distribution)
- If $\mathbb{K}\left(R_t\right) > 3$, the distribution of R_t is said to be **leptokurtic**.
- If $\mathbb{K}\left(R_t\right) = 3$, the distribution of R_t is said to be **mesokurtic**.
- If $\mathbb{K}\left(R_t\right) < 3$, the distribution of R_t is said to be **platykurtic**.
- The **excess kurtosis** coefficient is equal to $\mathbb{K}\left(R_t\right) - 3$.

4. Statistical properties of financial data

Figure: Illustration of the Kurtosis



Summary

Each (central or non-central) moment gives an information on the pdf of returns

	Formula	Interpretation
Mean	$\mathbb{E}(R_t) = \mu$	Indicator of central tendency
Variance	$\mathbb{V}(R_t) = \mathbb{E}\left((R_t - \mu)^2\right) = \sigma^2$	Indicator of dispersion around μ
Skewness	$\mathbb{S}(R_t) = \mathbb{E}\left((R_t - \mu)^3\right)/\sigma^3$	Indicator of symmetry
Kurtosis	$\mathbb{K}(R_t) = \mathbb{E}\left((R_t - \mu)^4\right)/\sigma^4$	Indicator of tail heaviness

Example (normal distribution)

If we assume that the returns are **normally** distributed with $R_t \sim \mathcal{N}\left(\mu, \sigma^2\right)$, then

$$\mathbb{E}\left(R_t\right) = \mu$$

$$\mathbb{V}\left(R_t\right) = \sigma^2$$

$$\mathbb{S}\left(R_t\right) = 0$$

$$\mathbb{K}\left(R_t\right) = 3$$

The normal distribution is **symmetric** and by definition, **mesokurtic**.

Example (Student distribution)

If we assume that the return R_t has a **Student distribution** with v degrees of freedom, i.e. $R_t \sim t(v)$, then we have

$$\mathbb{E}(R_t) = 0$$

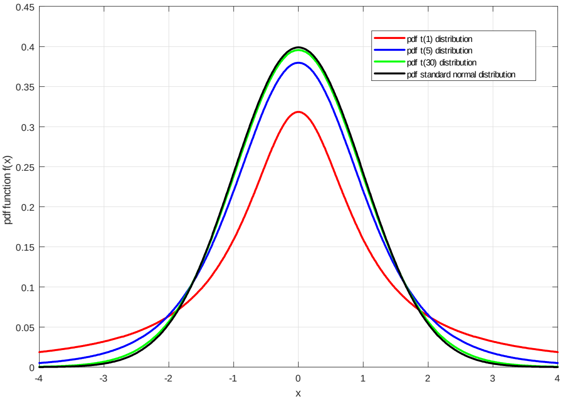
$$\mathbb{V}(R_t) = \frac{v}{v-2} \quad \text{if } v > 2$$

$$\mathbb{S}(R_t) = 0$$

$$\mathbb{K}(R_t) = 3 + \frac{6}{v-4} \quad \text{if } v > 4$$

The Student distribution is **symmetric** and **leptokurtic** (if v is "small"). Its excess of Kurtosis decreases with the number of degrees of freedom. When v tends to ∞, $t(v) \approx \mathcal{N}(0,1)$, and the distribution is mesokurtic.

Figure: Density function for a Student distribution



Remarks

- The moments allow to characterize the shape of the returns distribution.
- However, the (theoretical) moments are **unobservable** and we need to **estimate** them.
- Denote by $\{R_1, \ldots, R_T\}$ a **sample** of i.i.d. variables that have the same distribution as R_t
- Denote by $\{r_1, \ldots, r_T\}$ the realization of this sequence $=>$ your **dataset**.

Definition (estimator)

An **estimator** is any function $T\left(R_1, R_2, \ldots, R_T\right)$ of a sample. Any **descriptive statistic** is an estimator.

Example (sample mean)

Assume that $R_1, R_2, .., R_T$ are i.i.d. random variables. The sample mean (or average)

$$\widehat{\mu}_T \equiv \overline{R}_T = \frac{1}{T} \sum_{t=1}^{T} R_t$$

is an estimator of the (theoretical) mean $\mu = \mathbb{E}\left(R_t\right)$.

Example (sample variance)

Assume that $R_1, R_2, .., R_T$ are $i.i.d.$ random variables. The **sample variance**

$$\widehat{\sigma}_T^2 = \frac{1}{T-1} \sum_{t=1}^{T} \left(R_t - \overline{R}_T \right)^2$$

is an estimator of the (theoretical) variance $\sigma^2 = \mathbb{V}\left(R_t\right)$.

Note: the denominator is equal to $T-1$ as to define a sample variance corrected for the small sample bias (cf. chapter 2).

Example (sample skewness)

Assume that $R_1, R_2, .., R_T$ are *i.i.d.* random variables. The **sample skewness** coefficient

$$\widehat{S}_T = \frac{1}{T}\sum_{t=1}^{T}\left(\frac{R_t - \overline{R}_T}{\widehat{\sigma}}\right)^3 = \frac{T^{-1}\sum_{t=1}^{T}\left(R_t - \overline{R}_T\right)^3}{\left((T-1)^{-1}\sum_{t=1}^{T}\left(R_t - \overline{R}_T\right)^2\right)^{3/2}}$$

is an estimator of the (theoretical) skweness $\mathbb{S}\left(R_t\right)$.

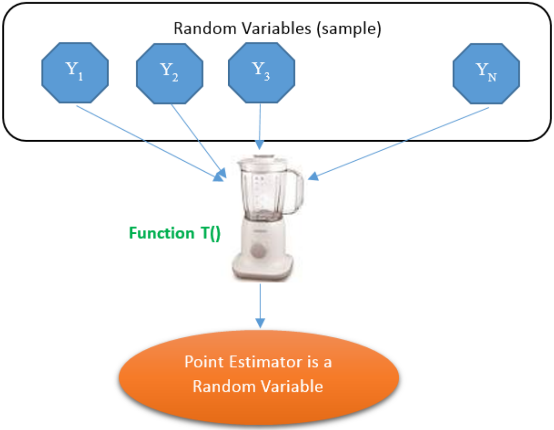
Example (sample kurtosis)

Assume that $R_1, R_2, .., R_T$ are $i.i.d.$ random variables. The **sample kurtosis** coefficient

$$\widehat{K}_T = \frac{1}{T}\sum_{t=1}^{T}\left(\frac{R_t - \overline{R}_T}{\widehat{\sigma}}\right)^4 = \frac{T^{-1}\sum_{t=1}^{T}\left(R_t - \overline{R}_T\right)^4}{\left((T-1)^{-1}\sum_{t=1}^{T}\left(R_t - \overline{R}_T\right)^2\right)^2}$$

is an estimator of the (theoretical) kurtosis $\mathbb{K}\left(R_t\right)$.

4. Statistical properties of financial data



4. Statistical properties of financial data

Fact

An estimator $\widehat{\theta}$ is a **random variable**.

Consequence: $\widehat{\theta}$ has a (marginal or conditional) **probability distribution**. This **sampling distribution** is characterized by a pdf $f_{\widehat{\theta}}\left(u\right)$

Definition (Sampling Distribution)

The probability distribution of an estimator (or a statistic) is called the **sampling distribution**.

Consequence: The sampling distribution of $\widehat{\theta}$ is characterized by **moments** such that the expectation $\mathbb{E}\left(\widehat{\theta}\right)$, the variance $\mathbb{V}\left(\widehat{\theta}\right)$, etc.

Definition (estimate)

A (point) **estimate** is the realized value of an estimator (i.e. a number) that is obtained when a sample is actually taken. For an estimator $\widehat{\theta}$ it can be denoted by $\widehat{\theta}(y)$.

Example (Point estimate)

For instance a realization of the estimator $\overline{R}_T$, denoted $\overline{r}_T$, is defined as

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

- If $T = 3$ and $\{r_1, r_2, r_3\} = \{3, -1, 2\}$ then $\overline{r}_T = 1.333$.
- If $T = 3$ and $\{r_1, r_2, r_3\} = \{4, -8, 1\}$ then $\overline{r}_T = -1$.
- etc..

Summary

Population moments (real nbs.)		Sample moments (random variables)	
Mean	$\mathbb{E}(R_t) = nb$	**Sample mean**	$\widehat{\mu}_T = \frac{1}{T}\sum_{t=1}^{T} R_t$
Variance	$\mathbb{V}(R_t) = nb$	**Sample variance**	$\widehat{\sigma}_T^2 = \frac{1}{T-1}\sum\limits_{t=1}^{T}\left(R_t - \overline{R}_T\right)^2$
Skewness	$\mathbb{S}(R_t) = nb$	**Sample Skewness**	$\widehat{S}_T = \frac{1}{T}\sum\limits_{t=1}^{T}\left(\frac{R_t - \overline{R}_T}{\widehat{\sigma}}\right)^3$
Kurtosis	$\mathbb{K}(R_t) = nb$	**Sample Kurtosis**	$\widehat{K}_T = \frac{1}{T}\sum\limits_{t=1}^{T}\left(\frac{R_t - \overline{R}_T}{\widehat{\sigma}}\right)^4$

Question: What constitutes a good estimator?

The idea is to study the **properties of the sampling distribution** and especially its **moments** such as

- $\mathbb{E}\left(\widehat{\theta}\right)$ for the bias,
- $\mathbb{V}\left(\widehat{\theta}\right)$ for the precision,
- $\mathbb{S}\left(\widehat{\theta}\right)$ for the symmetry,
- etc..

Estimators are compared on the basis of a variety of attributes.

1. **Finite sample properties** (or finite sample distribution) of estimators are those attributes that can be compared regardless of the sample size.
2. However, the finite sample distribution is known only in the case of specific distributional assumption on R_t (typically, normality).
3. When the normality assumption is no longer valid (and the finite sample distribution is unknown), estimators are evaluated on the basis on their **large sample**, or **asymptotic properties**

Theorem (finite sample distributions)

If we assume that the returns $R_1, R_2, .., R_T$ are i.i.d. with $R_t \sim \mathcal{N}\left(\mu, \sigma^2\right)$, then $\widehat{\mu}_T$ and $\left(T-1\right)\widehat{\sigma}_T^2/\sigma^2$ have a **finite sample distribution**

$$\widehat{\mu}_T \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{T}\right) \qquad \forall T \in \mathbb{N}$$

$$\frac{\left(T-1\right)}{\sigma^2}\widehat{\sigma}_T^2 \sim \chi^2\left(T-1\right) \qquad \forall T \geq 2$$

Example (finite sample distribution)

If we assume that the returns $R_1, R_2, .., R_T$ are $\mathcal{N}$.i.d.$\left(\mu, \sigma^2\right)$, if $T = 10$, then

$$\widehat{\mu}_{10} \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{10}\right) \qquad 9 \times \frac{\widehat{\sigma}_T^2}{\sigma^2} \sim \chi^2\left(9\right)$$

4. Statistical properties of financial data

In most of cases, it is impossible to derive the **exact/finite sample distribution** for the estimator (or a transformed variable).

Two reasons:

1. In some cases, the exact distribution of $R_1, R_2, \ldots, R_T$ is known, but the function $T(.)$ is too complicated to derive the distribution of $\widehat{\theta}$:

$$\widehat{\theta} = T(R_1, R_2, \ldots, R_T) \sim ??? \qquad \forall T \in \mathbb{N}$$

2. In most of cases, the distribution of the returns $R_1, R_2, \ldots, R_T$ is **unknown**...

$$\widehat{\theta} = T(R_1, R_2, \ldots, R_T) \sim ??? \qquad \forall T \in \mathbb{N}$$

Question: what is the behavior of the estimator $\widehat{\theta}_T$ when the sample size T tends to infinity?

Definition (Asymptotic theory)

Asymptotic or **large sample theory** consists in the study of the distribution of the estimator when the sample size is sufficiently large.

The asymptotic theory is fundamentally based on the notion of **convergence**.

We distinguish two types of convergence

1. **Convergence in probability** (or in mean squared, or almost sure):

$$\widehat{\theta}_T \text{ converges to a real number}$$

   This convergence types are used to derive the **consistency** property of the estimators.

2. **Convergence in distribution**:

$$\sqrt{T}\left(\widehat{\theta}_T - \theta_0\right) \text{ converges to a given distribution (say, normal)}$$

   This convergence is used to derive the **asymptotic distribution** of the estimators and to make **inference** (tests) about the true value of the parameters.

Definition (Convergence in probability)

Let X_T be a sequence random variable indexed by the sample size. X_T **converges in probability** to a constant c, if, for any $\varepsilon > 0$,

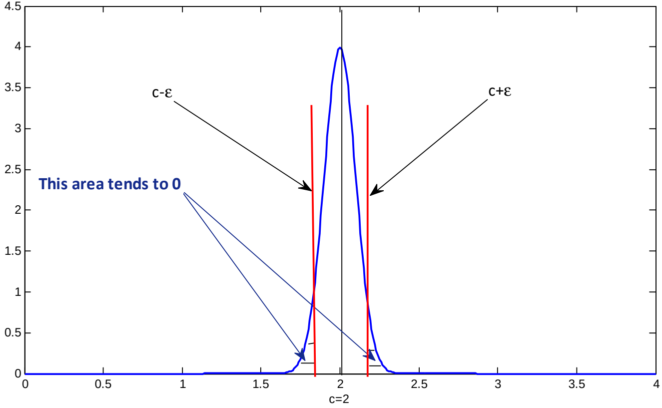
$$\lim_{T \to \infty} \Pr \left(\left| X_T - c \right| > \varepsilon \right) = 0$$

It is written

$$X_T \xrightarrow{p} c \qquad \text{or} \qquad \text{plim } X_T = c$$

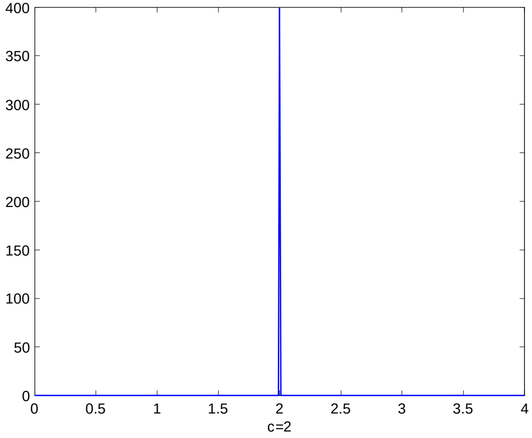
4. Statistical properties of financial data

$$X_T \xrightarrow{p} c \quad \text{if} \quad \lim_{t\to\infty} \Pr\left(|X_T - c| > \varepsilon\right) = 0$$



4. Statistical properties of financial data

$$X_t \xrightarrow{p} c \quad \text{if} \quad \lim_{T\to\infty} \Pr\left(|X_T - c| > \varepsilon\right) = 0 \quad \text{for a very small } \varepsilon...$$



Theorem (consistency)

If the returns $R_1, R_2, .., R_T$ are i.i.d. with finite moments, $\mathbb{E}(R_t) = \mu$, $\mathbb{V}(R_t) = \sigma^2$, $\mathbb{S}(R_t)$ and $\mathbb{K}(R_t)$, then the sample first four moments are **weakly consistent**

$$\widehat{\mu}_T \xrightarrow{p} \mathbb{E}(R_t) = \mu$$

$$\widehat{\sigma}^2_T \xrightarrow{p} \mathbb{V}(R_t) = \sigma^2$$

$$\widehat{S}_T \xrightarrow{p} \mathbb{S}(R_t)$$

$$\widehat{K}_T \xrightarrow{p} \mathbb{K}(R_t)$$

Interpretation

1. The distribution of the sample moments is highly concentrated around the true value (unknown) of the population moments of the returns when the sample size T tends to infinity.

2. The realization of the sample moments (mean, variance, skewness, etc.) are then "close" to the value of the population moment when the sample size T tends to infinity.

Definition (Convergence in distribution)

Let X_T be a sequence random variable indexed by the sample size with a cdf $F_T(.)$. X_T **converges in distribution** to a random variable X with cdf $F(.)$ if

$$\lim_{T\rightarrow\infty} F_T(x) = F(x) \quad \forall x$$

It is written:

$$X_t \xrightarrow{d} X$$

Theorem (asymptotic distributions for sample mean and variance)

If the returns $R_1, R_2, .., R_T$ are i.i.d. with finite mean $\mathbb{E}(R_t) = \mu$ and finite variance $\mathbb{V}(R_t) = \sigma^2$, then we can derive the following **asymptotic distributions** *when T tends to infinity*

$$\sqrt{T}\left(\widehat{\mu}_T - \mu\right) \xrightarrow{d} \mathcal{N}\left(0, \sigma^2\right)$$

$$\sqrt{T}\left(\widehat{\sigma}^2_T - \sigma^2\right) \xrightarrow{d} \mathcal{N}\left(0, 2\sigma^4\right)$$

Remark: The asymptotic results

$$\sqrt{T}\left(\widehat{\mu}_T - \mu\right) \xrightarrow{d} \mathcal{N}\left(0, \sigma^2\right)$$

$$\sqrt{T}\left(\widehat{\sigma}_T^2 - \sigma^2\right) \xrightarrow{d} \mathcal{N}\left(0, 2\sigma^4\right)$$

can be interpreted as

$$\widehat{\mu}_T \overset{asy}{\approx} \mathcal{N}\left(\mu, \frac{\sigma^2}{T}\right) \qquad \widehat{\sigma}_T \overset{asy}{\approx} \mathcal{N}\left(\sigma^2, \frac{2\sigma^4}{T}\right)$$

where the symbol $\overset{asy}{\approx}$ means "asymptotically distributed as".

Note: These asymptotic results for the sample moments can be used to perform **statistical tests** about the (population) moments.

Example (confidence interval)

Consider a sample of $T = 500$ i.i.d. returns $R_1, .., R_T$, and a realization (dataset) $\{r_1, .., r_T\}$ with

$$\widehat{\mu}_T = \overline{r}_T = -0.057 \quad \widehat{\sigma}^2_T = 0.0009$$

Question: compute a 95% confidence interval on the true value of the expected return $\mathbb{E}(R_t)$.

Answer

We known that $\widehat{\mu}_T \overset{asy}{\approx} \mathcal{N}\left(\mu, \sigma^2 / T\right)$, so we have

$$IC_{95\%} = \left[\mu \pm \frac{\sigma}{\sqrt{T}} \Phi^{-1}\left(0.975\right)\right]$$

Since μ and σ^2 are unobservable, we can use a consistent estimator (plug-in approach) and we get

$$\begin{aligned} IC_{95\%} &= \left[\widehat{\mu} \pm \frac{\widehat{\sigma}}{\sqrt{T}} \Phi^{-1}\left(0.975\right)\right] \\ &= \left[-0.057 \pm \sqrt{\frac{0.0009}{500}} \Phi^{-1}\left(0.975\right)\right] \\ &= \left[-0.0596; -0.0544\right] \end{aligned}$$

Theorem (asymptotic distributions for sample skewness and kurtosis)

If we assume that the returns $R_1, R_2, .., R_T$ are $\mathcal{N}.i.d.(\mu, \sigma^2)$, then we have

$$\sqrt{T}\left(\widehat{S}_T - 0\right) \xrightarrow{d} \mathcal{N}(0, 6)$$

$$\sqrt{T}\left(\widehat{K}_T - 3\right) \xrightarrow{d} \mathcal{N}(0, 24)$$

Note: These asymptotic results for the sample moments can be used to perform **statistical tests** about the returns distribution (Jarque-Bera test for instance).

Key Concepts

1. Returns distribution
2. Central and non–central moments
3. Mean, variance, skewness, and kurtosis
4. Leptokurtic or mesokurtic returns distribution
5. Symmetric or asymmetric returns distribution
6. Sample, estimator, estimate, and (descriptive) statistics
7. Independent and identically distributed (i.i.d.) random variables
8. Sampling distribution of an estimator
9. Finite sample vs asymptotic distribution

Section 5

Stylized Facts

5. Stylized Facts

Objectives

1. Present the main **stylized facts** of financial series
2. Define the **stationarity** property of asset returns
3. Show the **absence of autocorrelations** of asset returns
4. Show the **heavy tails** and the **asymmetry** of the distribution of returns
5. Show the **volatility clustering**
6. Define the **aggregational Gaussianity**
7. Introduce the **long range dependence**
8. Define the **leverage effect**

5. Stylized Facts

- The statistical properties of financial data revealed a wealth of **stylized facts**
- These stylized facts are common in a wide range of financial time series.

Cont, R. (2001), Empirical Properties of Asset Returns: Stylized Facts and Statistical Issues, *Quantitative Finance*, 1, 223– 236

Example (Intel Corp)

In order to illustrate some of these stylized facts, we consider a sample of 1,259 daily prices and (log-) returns for the S&P500 index (ticker: SPY) from August 19, 2013 to August 17, 2018 (5 years). The data are available in Data_SP500.xlsx.

5. Stylized Facts

Fan and Yao (2015) identify 8 main stylized facts

1. **Stationarity**
2. **Absence of autocorrelations**
3. **Heavy tails**
4. **Asymmetry**
5. **Volatility clustering**
6. **Aggregational Gaussianity**
7. **Long range dependence**
8. **Leverage effect**

Fan, J. and Yao, Q. The Elements of Financial Econometrics, Science Press, Beijing, 2015

5. Stylized Facts

Fan and Yao (2015) identify 8 main stylized facts

1. **Stationarity**
2. **Absence of autocorrelations**
3. **Heavy tails**
4. **Asymmetry**
5. **Volatility clustering**
6. **Aggregational Gaussianity**
7. **Long range dependence**
8. **Leverage effect**

Fan, J. and Yao, Q. The Elements of Financial Econometrics, Science Press, Beijing, 2015

Stylized Fact 1: Stationarity

Fact (stationarity)

In general, the prices are non-stationary whereas the returns are **stationary**.

- The prices of an asset recorded over times are often not stationary due to, for example, the steady expansion of economy, the increase of productivity, and financial crisis.
- However their returns, typically fluctuates around a constant level, suggesting a constant mean over time.
- Most return sequences can be modeled as a stochastic processes with at least time-invariant two first moments: **weak stationarity** (cf. chapter 5)

Definition (weak or second-order stationarity)

A time-series process $(x_t, \ t \in \mathbb{Z})$ is **weakly stationary** (second-order stationarity) if and only if:

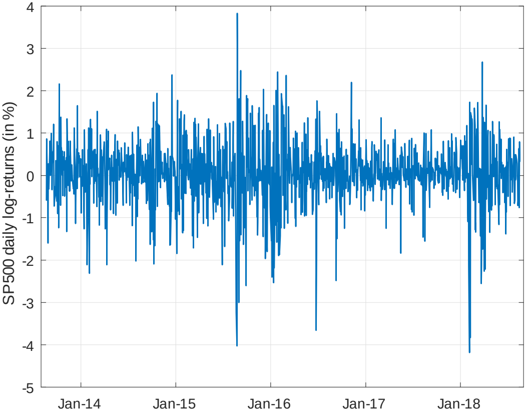
- $\forall t \in \mathbb{Z}, \ \mathbb{E}\left(x_t^2\right) < \infty$
- $\forall t \in \mathbb{Z}, \mathbb{E}\left(x_t\right) = \mu$ does not depend on t
- $\forall \left(t, h\right) \in \mathbb{Z}^2, \mathbb{C}ov\left(x_t, x_{t+h}\right) = \mathbb{E}\left(\left(x_{t+h} - m\right)\left(x_t - m\right)\right) = \gamma_h$ does not depend on t

5. Stylized Facts

Figure: Daily closing prices for the S&P500 index are non stationary

5. Stylized Facts

Figure: Daily returns for the S&P500 index are stationary



5. Stylized Facts

Fan and Yao (2015) identify 8 main stylized facts

1. **Stationarity**
2. **Absence of autocorrelations**
3. **Heavy tails**
4. **Asymmetry**
5. **Volatility clustering**
6. **Aggregational Gaussianity**
7. **Long range dependence**
8. **Leverage effect**

Fan, J. and Yao, Q. The Elements of Financial Econometrics, Science Press, Beijing, 2015

Stylized Fact 2: Absence of autocorrelations

Fact (absence of autocorrelations)

The **autocorrelations** *of asset returns R_t are often insignificant, except for very small intraday time scales ($\approx$ 20 minutes) for which microstructure effects come into play.*

Note: The fact that returns hardly show any serial correlation, does not mean that they are independent.

Definition (autocorrelation)

The **autocorrelation**, denoted $\rho(k)$, of a weak stationary process R_t is the correlation between values of the process at different times, defined as:

$$\rho_k = \mathbb{C}orr\left(R_t, R_{t-k}\right) = \frac{\mathbb{E}\left(\left(R_t - \mu\right)\left(R_{t-k} - \mu\right)\right)}{\mathbb{V}\left(R_t\right)} = \frac{\gamma_k}{\sigma^2}$$

with $\mu = \mathbb{E}\left(R_t\right)$, $\sigma^2 = \mathbb{V}\left(R_t\right)$, $\forall t$ and γ_k the autocovariance of order k.

Definition (sample autocorrelation)

The **sample autocorrelation**, denoted $\widehat{\rho}(k)$, of a weak stationary process R_t, is an estimator of $\rho(k)$ defined as

$$\widehat{\rho}_k = corr(R_t, R_{t-k}) = \frac{1}{(T-k)\,\widehat{\sigma}^2} \sum_{t=k+1}^{T} (R_t - \widehat{\mu})(R_{t-k} - \widehat{\mu})$$

where $\widehat{\sigma}^2$ and $\widehat{\mu}$ are consistent estimators of $\mu = \mathbb{E}(R_t)$ and $\sigma^2 = \mathbb{V}(R_t)$, $\forall t$.

The **AutoCorrelation function (ACF)** (or **correlogram**) represents the sample autocorrelation for different lags from $k = 1$ to a maximum lag order, say $k = 15$.

Figure: ACF for the S&P500 daily returns (Aug 2013 - Aug 2018). Source: Eviews 3

Correlogram of SP500_RETURNS

Date: 09/09/18 Time: 11:25
Sample: 8/19/2013 8/17/2018
Included observations: 1259

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.019	-0.019	0.4455	0.504
		2	-0.054	-0.054	4.1067	0.128
		3	0.024	0.022	4.8464	0.183
		4	-0.043	-0.045	7.1488	0.128
		5	-0.028	-0.027	8.1122	0.150
		6	-0.008	-0.014	8.1869	0.225
		7	0.013	0.011	8.3965	0.299
		8	-0.037	-0.039	10.165	0.254
		9	-0.054	-0.057	13.887	0.126
		10	-0.019	-0.028	14.324	0.159
		11	0.002	-0.003	14.330	0.215
		12	-0.001	-0.004	14.330	0.280
		13	-0.015	-0.022	14.617	0.332
		14	-0.020	-0.027	15.102	0.371
		15	-0.073	-0.079	21.926	0.110

Testing the nullity of autocorrelations

- It is possible to test the nullity of the autocorrelations through a **Q-test** (Box-Pierce test or Ljung-Box test).
- For more details on statistical tests: cf. Chapter 3.
- For a given lag-order K, the null hypothesis of the test is

$$\mathrm{H}_0 : \rho_1 = \ldots = \rho_K = 0$$

- The alternative hypothesis is

$$\mathrm{H}_1 : \exists j \in \{1, \ldots, K\} \ \ / \ \widehat{\rho}_j \neq 0$$

Definition (Box-Pierce test)

The **Box-Pierce** test statistic associated to the null $\mathrm{H}_0 : \rho_1 = \ldots = \rho_K = 0$, is defined as

$$Q_{BP} = T \sum_{k=1}^{K} \widehat{\rho}_k^2 \underset{\mathrm{H}_0}{\overset{d}{\to}} \mathcal{X}^2 \left(K \right)$$

Decision rule (cf. Chapter 3):

- If the realization of the test statistic Q_{BP} is larger than the quantile of a chi-squared distribution with K degrees of freedom at the probability $1 - \alpha\%$ (say, 95%), for a risk level $\alpha\%$ (say, 5%) we reject the null hypothesis H_0.

- If the p-value associated to the test statistic Q_{BP} is smaller the risk level $\alpha\%$ (say, 5%), we reject the null hypothesis H_0.

- If the null hypothesis H_0 is rejected, there is at least one autocorrelation (between lag 1 and K) which is non-null, i.e. the time series R_t is autocorrelated.

Figure: Q-statistics for the S&P500 daily returns (Aug 2013 - Aug 2018). Source: Eviews 3

Correlogram of SP500_RETURNS

Date: 09/09/18 Time: 11:25
Sample: 8/19/2013 8/17/2018
Included observations: 1259

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.019	-0.019	0.4455	0.504
		2	-0.054	-0.054	4.1067	0.128
		3	0.024	0.022	4.8464	0.183
		4	-0.043	-0.045	7.1488	0.128
		5	-0.028	-0.027	8.1122	0.150
		6	-0.008	-0.014	8.1869	0.225
		7	0.013	0.011	8.3965	0.299
		8	-0.037	-0.039	10.165	0.254
		9	-0.054	-0.057	13.887	0.126
		10	-0.019	-0.028	14.324	0.159
		11	0.002	-0.003	14.330	0.215
		12	-0.001	-0.004	14.330	0.280
		13	-0.015	-0.022	14.617	0.332
		14	-0.020	-0.027	15.102	0.371
		15	-0.073	-0.079	21.926	0.110

5. Stylized Facts

Fan and Yao (2015) identify 8 main stylized facts

1. **Stationarity**
2. **Absence of autocorrelations**
3. **Heavy tails**
4. **Asymmetry**
5. **Volatility clustering**
6. **Aggregational Gaussianity**
7. **Long range dependence**
8. **Leverage effect**

Fan, J. and Yao, Q. The Elements of Financial Econometrics, Science Press, Beijing, 2015

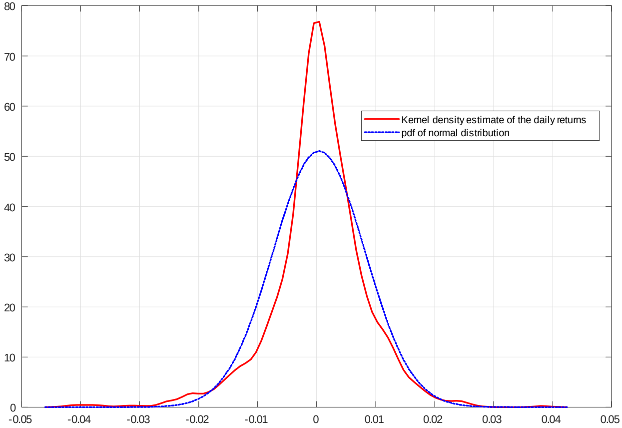
Stylized Fact 3: Heavy tails

Fact (heavy tails)

The probability distribution of return often exhibits **heavier tails** *than those of a normal distribution.*

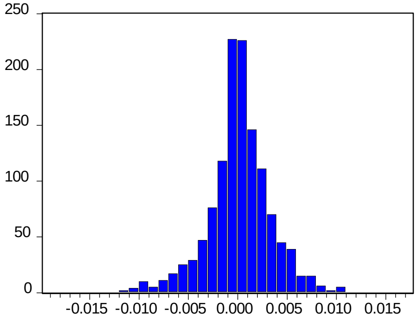
5. Stylized Facts

Figure: Kernel density estimate of the daily returns distribution (pdf) for the S&P500 index



5. Stylized Facts

Figure: Descriptive statistics for the daily returns of the S&P500 index (Eviews 3)



Series: SP500_RETURNS
Sample 8/20/2013 6/15/2018
Observations 1259

Mean	0.000189
Median	0.000231
Maximum	0.016630
Minimum	-0.018172
Std. Dev.	0.003390
Skewness	-0.566934
Kurtosis	6.405986
Jarque-Bera	675.9989
Probability	0.000000

5. Stylized Facts

Fan and Yao (2015) identify 8 main stylized facts

1. **Stationarity**
2. **Absence of autocorrelations**
3. **Heavy tails**
4. **Asymmetry**
5. **Volatility clustering**
6. **Aggregational Gaussianity**
7. **Long range dependence**
8. **Leverage effect**

Fan, J. and Yao, Q. The Elements of Financial Econometrics, Science Press, Beijing, 2015

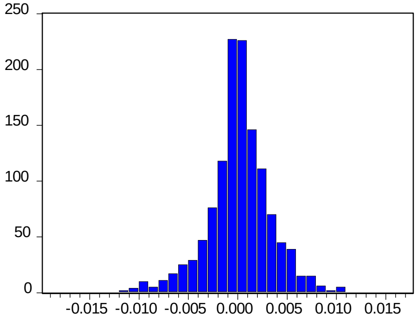
Stylized Fact 4: Asymmetry

Fact (asymmetry)

The distribution of return is **asymmetric** *and often* **negatively skewed***, reflecting the fact that the downturns of financial markets are often much steeper than the recoveries. Investors tend to react more strongly to negative news than to positive news*

5. Stylized Facts

Figure: Descriptive statistics for the daily returns of the S&P500 index (Eviews 3)



Series: SP500_RETURNS	
Sample 8/20/2013 6/15/2018	
Observations 1259	
Mean	0.000189
Median	0.000231
Maximum	0.016630
Minimum	-0.018172
Std. Dev.	0.003390
Skew ness	-0.566934
Kurtosis	6.405986
Jarque-Bera	675.9989
Probability	0.000000

5. Stylized Facts

Fan and Yao (2015) identify 8 main stylized facts

1. **Stationarity**
2. **Absence of autocorrelations**
3. **Heavy tails**
4. **Asymmetry**
5. **Volatility clustering**
6. **Aggregational Gaussianity**
7. **Long range dependence**
8. **Leverage effect**

Fan, J. and Yao, Q. The Elements of Financial Econometrics, Science Press, Beijing, 2015

Stylized Fact 5: Volatility clustering

Fact (volatility clustering)

The **volatility clustering** *means that large price changes (i.e. returns with large absolute values or large squares) occur in clusters. Indeed, large price changes tend to be followed by large price changes, and periods of tranquility alternate with periods of high volatility.*

Note: the volatility clustering is the consequence of the autocorrelation of the squared returns (cf. Stylized fact 7)

5. Stylized Facts

Figure: Daily returns for the S&P500 index

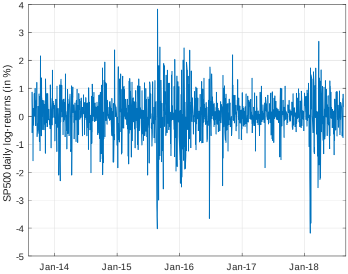
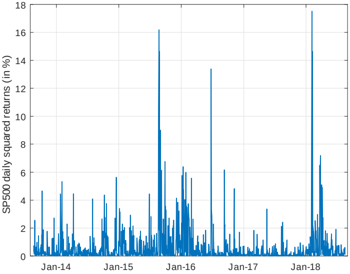


Figure: Daily squared-returns for the S&P500 index



5. Stylized Facts

Fan and Yao (2015) identify 8 main stylized facts

1. **Stationarity**
2. **Absence of autocorrelations**
3. **Heavy tails**
4. **Asymmetry**
5. **Volatility clustering**
6. **Aggregational Gaussianity**
7. **Long range dependence**
8. **Leverage effect**

Fan, J. and Yao, Q. The Elements of Financial Econometrics, Science Press, Beijing, 2015

Stylized Fact 6: Aggregational Gaussianity

Fact (aggregational Gaussianity)

A return over k days is simply the aggregation of k daily returns. When the time horizon k increases, the central limit law sets in and the distribution of the returns over a long time-horizon (such as a month) tends toward a **normal distribution***.*

Definition (Jarque-Bera test)

The **Jarque-Bera test** is

$$\mathrm{H}_0 : \mathbb{S}\left(R_t\right) = 0 \;\text{ and }\; \mathbb{K}\left(R_t\right) = 3$$

$$\mathrm{H}_1 : \mathbb{S}\left(R_t\right) \neq 0 \;\text{ and/or }\; \mathbb{K}\left(R_t\right) \neq 3$$

The Jarque-Bera test statistic is defined as

$$\mathrm{JB} = \frac{T}{6}\widehat{S}_T^2 + \frac{T}{24}\left(\widehat{K}_T - 3\right)^2 \underset{\mathrm{H}_0}{\overset{d}{\to}} \mathcal{X}^2\left(2\right)$$

where $\widehat{S}_T$ and $\widehat{K}_T$ denote the sample Skewness and Kurtosis coefficients, respectively.

Jarque-Bera test

Interpretation:

- The Jarque-Bera is not a normality test. However, if the null hypothesis is rejected the distribution of R_t **cannot** be normal.
- If one fails to reject the null hypothesis, it only implies that the distribution is symmetric and mesokurtic (the normal distribution is not the only one in this case).

Decision rule (cf. Chapter 3):

- If the JB test statistic is larger than the quantile of a chi-squared distribution with 2 degrees of freedom at the probability $1-\alpha\%$ (say, 95%), for a risk level $\alpha\%$ (say, 5%) we reject the null hypothesis H_0.
- If the p-value associated to the JB test statistic is smaller the risk level $\alpha\%$ (say, 5%), we reject the null hypothesis H_0.

Example (Jarque-Bera test statistic)

For the S&P500 index, the sample Skewness and Kurtosis coefficient are equal to -0.5669 and 6.4059, respectively for a sample size T equal to 1259. **Question:** compute the realization of the Jarque-Bera test statistic and conclude for a 5% significance level.

Answer: The realization of the Jarque-Bera test statistic equal to

$$\text{JB} = \frac{T}{6}\widehat{S}_T^2 + \frac{T}{24}\left(\widehat{K}_T - 3\right)^2 = \frac{1259}{6}\left(-0.5669\right)^2 + \frac{1259}{24}\left(6.4059 - 3\right)^2 = 675.9989$$

For a risk level $\alpha = 5\%$, the critical value is equal to

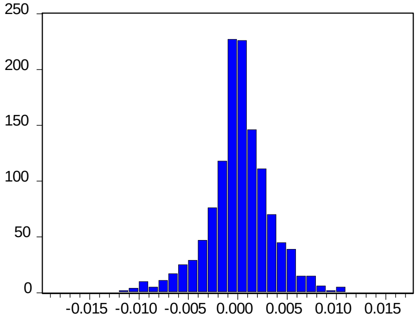
$$\text{critical value} = F_2^{-1}\left(\alpha\right) = F_2^{-1}\left(0.95\right) = 5.9915$$

where $F\left(.\right)$ is the cdf of a chi-squared distribution with 2 degrees of freedom.

Conclusion: For a risk level $\alpha = 5\%$, test statistic is larger than the critical value, so we reject the null hypothesis (symmetric and mesokurtic distribution). The distribution of the daily returns **cannot** be normal.

5. Stylized Facts

Figure: Descriptive statistics for the daily returns of the S&P500 index (Eviews 3)



Series: SP500_RETURNS
Sample 8/20/2013 6/15/2018
Observations 1259

Mean	0.000189
Median	0.000231
Maximum	0.016630
Minimum	-0.018172
Std. Dev.	0.003390
Skew ness	-0.566934
Kurtosis	6.405986
Jarque-Bera	675.9989
Probability	0.000000

5. Stylized Facts

Fan and Yao (2015) identify 8 main stylized facts

1. **Stationarity**
2. **Absence of autocorrelations**
3. **Heavy tails**
4. **Asymmetry**
5. **Volatility clustering**
6. **Aggregational Gaussianity**
7. **Long range dependence**
8. **Leverage effect**

Fan, J. and Yao, Q. The Elements of Financial Econometrics, Science Press, Beijing, 2015

Stylized Fact 7: Long range dependence

Fact (long range dependence)

Both daily squared and absolute returns often exhibit significant autocorrelations. Those autocorrelations are persistent, indicating possible **long-memory** *properties.*

Note: Those autocorrelations become weaker and less persistent when the sampling interval is increased from a day, to a week to a month.

5. Stylized Facts

Figure: ACF for the S&P500 returns

Date: 09/09/18 Time: 15:19
Sample: 8/19/2013 8/17/2018
Included observations: 1259

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.019	-0.019	0.4455	0.504
		2	-0.054	-0.054	4.1067	0.128
		3	0.024	0.022	4.8464	0.183
		4	-0.043	-0.045	7.1488	0.128
		5	-0.028	-0.027	8.1122	0.150
		6	-0.008	-0.014	8.1869	0.225
		7	0.013	0.011	8.3965	0.299
		8	-0.037	-0.039	10.165	0.254
		9	-0.054	-0.057	13.887	0.126
		10	-0.019	-0.028	14.324	0.159
		11	0.002	-0.003	14.330	0.215
		12	-0.001	-0.004	14.330	0.280
		13	-0.015	-0.022	14.617	0.332
		14	-0.020	-0.027	15.102	0.371
		15	-0.073	-0.079	21.926	0.110

Figure: ACF for the S&P500 squared returns

Date: 09/09/18 Time: 11:25
Sample: 8/19/2013 8/17/2018
Included observations: 1259

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.302	0.302	114.94	0.000
		2	0.251	0.176	194.81	0.000
		3	0.288	0.196	299.82	0.000
		4	0.251	0.116	379.50	0.000
		5	0.115	-0.052	396.30	0.000
		6	0.135	0.018	419.45	0.000
		7	0.095	-0.020	430.83	0.000
		8	0.108	0.044	445.63	0.000
		9	0.102	0.043	458.93	0.000
		10	0.123	0.060	478.14	0.000
		11	0.084	0.003	487.21	0.000
		12	0.050	-0.041	490.35	0.000
		13	0.048	-0.020	493.24	0.000
		14	0.040	-0.012	495.29	0.000
		15	0.031	0.010	496.51	0.000

Remarks

- The autocorrelation of the squared returns is called **ARCH effect** (cf. chapter 6).
- The ARCH effect depends on the sampling frequency.
- It is most important with daily returns, and less important with low frequency returns (monthly, quarterly, etc.).

5. Stylized Facts

Fan and Yao (2015) identify 8 main stylized facts

1. **Stationarity**
2. **Absence of autocorrelations**
3. **Heavy tails**
4. **Asymmetry**
5. **Volatility clustering**
6. **Aggregational Gaussianity**
7. **Long range dependence**
8. **Leverage effect**

Fan, J. and Yao, Q. The Elements of Financial Econometrics, Science Press, Beijing, 2015

Stylized Fact 8: Leverage effect

Fact (leverage effect)

Asset returns are negatively correlated with the changes of their volatilities: this negative correlation is called **leverage effect***.*

- As asset prices decline, companies become more leveraged (debt to equity ratios increase) and riskier, and hence their stock prices become more volatile.
- On the other hand, when stock prices become more volatile, investors demand high returns and hence stock prices go down.
- Volatilities caused by price decline are typically larger than the appreciations due to declined volatilities.

Key Concepts

1. Stationarity
2. Absence of autocorrelation
3. ARCH effect
4. Heavy tails and leptokurtic distribution
5. Asymmetric distribution
6. Volatility clustering
7. Aggregational Gaussianity
8. Long range dependence
9. Leverage effect

End of the Chapter 1

Christophe Hurlin