- What is quantitative finance?

- Some say it is the 'quest' to understand the 'laws' of financial markets, in a quantitative sense.

- Once 'understood' these laws can be used to aid in better financial decision making.

- Three discussion points:

  1. What do we mean by laws?

  2. What 'financial decisions'?

  3. Where and how do time series and forecasting fit?

At the end of this unit, students should be able to:

- Demonstrate an understanding of the principles of financial time series and forecasting in financial context.
- Demonstrate the ability to identify and fit an appropriate time series model to financial data.
- Deduce and discuss the relevant theoretical properties of financial time series models and their implications for real data.
- Demonstrate the ability to forecast risk measures and assess and compare their results between competing models.
- Demonstrate some ability to make informed financial decisions from data and statistical models.
- Demonstrate proficiency in the software Matlab.

Module 1 : Properties of Financial Data, Review of Statistics and Factor Models

Chapters 1 and 2 in Tsay

1.1 Financial data

- Most financial analysts and quant studies focus on returns.

- *Why?*

- Prices or financial index values in this unit are denoted with P_t.

- This is the ...

- This unit focuses on *synchronously* observed financial prices and returns, over time.

- This means the data are observed at constant time frequency: e.g. daily, weekly, etc.

- This is NOT the way real data occurs, and is somewhat artificial and convenient.

- Thus, here P_t typically refers to the price or index value of an asset at a fixed point in time on day (week or month) t.

- The *actual* or *simple* return on investment between period $t-1$ and period t, is measured by:
$$R_t = \frac{P_t - P_{t-1}}{P_{t-1}}$$
which is the proportional change in the value of the asset or portfolio.

- We will most often work with daily percentage returns, i.e. simply $100R_t$ in this unit.

- Some commonly used and useful relations:

$$
\begin{aligned}
P_t &= P_{t-1}(1 + R_t) \\
1 + R_t &= \frac{P_t}{P_{t-1}} \\
R_t &= \frac{P_t}{P_{t-1}} - 1
\end{aligned}
$$

- Note that $R_t > -1$ but is not bounded above. Any ramifications?

- Figure 1 shows the prices of CBA stock at close of trading from each day from January, 1999 to February, 2017.

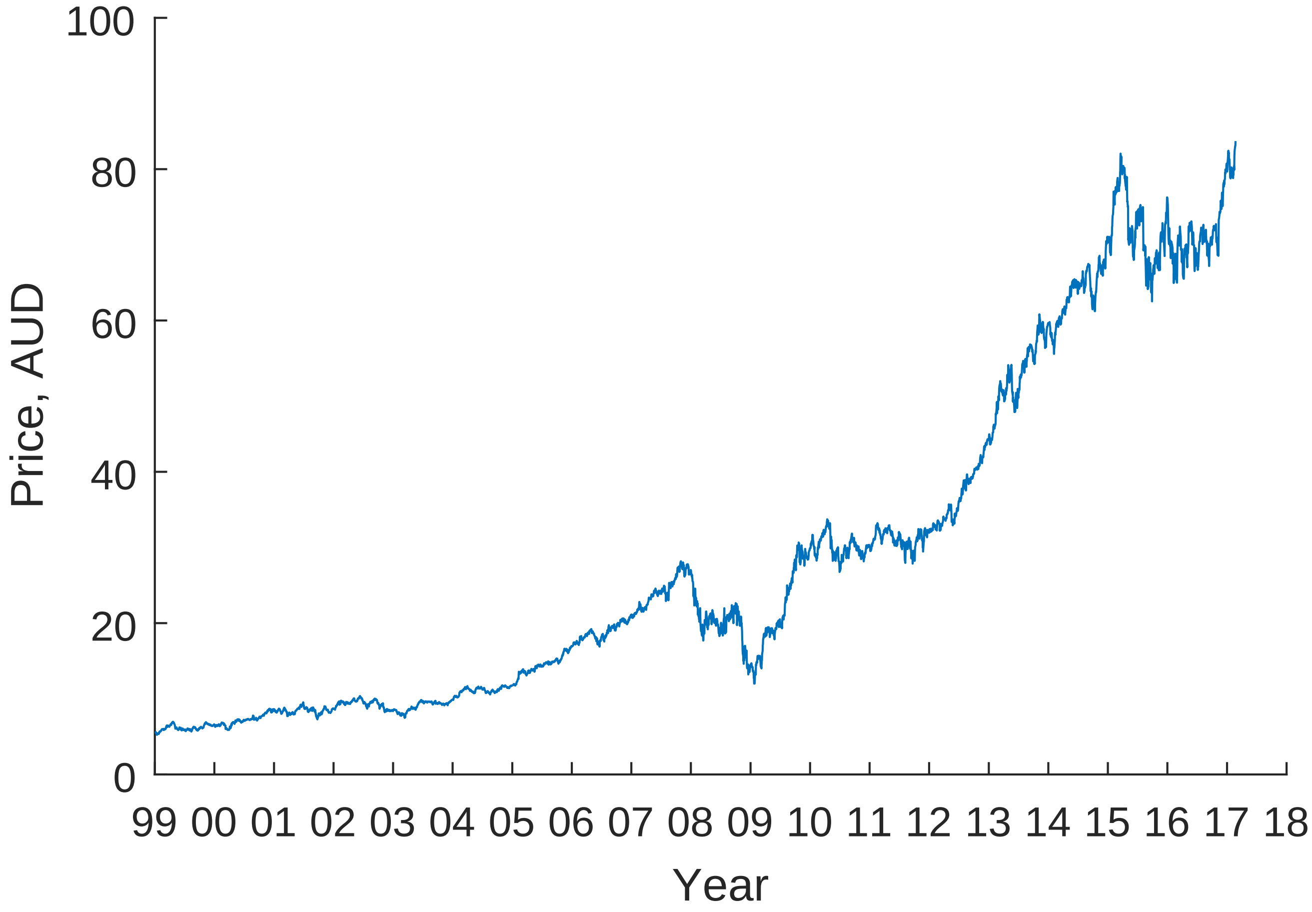

Figure 1: CBA prices 01/1999-02/2017.

- Figure 2 shows the actual returns on CBA stock at close of trading from each day from January, 1999 to February, 2017.

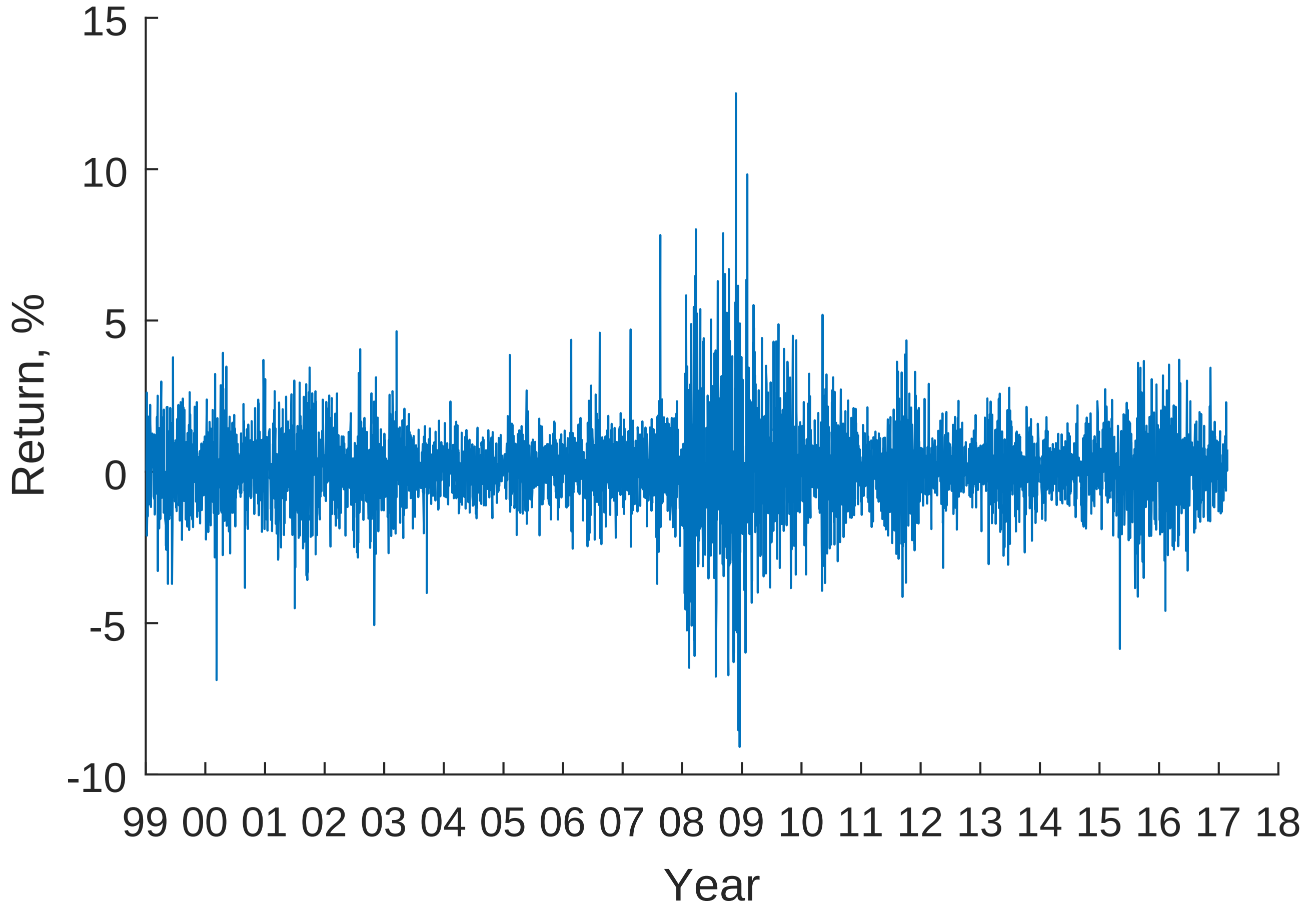

Figure 2: CBA simple percentage returns 01/1999-02/2017.

- Properties of prices? Returns?

- **Multi-period** returns are defined as:

$$R_t[k] = \frac{P_t - P_{t-k}}{P_{t-k}}$$

where the return is over k consecutive periods. Note that:

$$\begin{aligned} 1 + R_t[k] &= \frac{P_t}{P_{t-k}} = \frac{P_t}{P_{t-1}}\frac{P_{t-1}}{P_{t-2}} \dots \frac{P_{t-k+1}}{P_{t-k}} \\ &= (1 + R_t)(1 + R_{t-1}) \dots (1 + R_{t-k+1}) \\ &= \Pi_{i=1}^{k}(1 + R_{t-i+1}) \end{aligned}$$

- In practice, analysts often use log-returns, instead of simple returns.

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

- Why? :
  1. $r_t = \ln(1 + R_t) \approx R_t$ when R_t is small (by Taylor series expansion)
  2. $r_t[k] = \ln(P_t) - \ln(P_{t-k}) = r_t + r_{t-1} + \ldots + r_{t-k+1}$ (log returns are additive)
  3. Fits in with some financial theory regarding logarithms of prices (random walks)

- Figure 6 shows simple and percentage daily log-returns for CBA from 1999-2016.

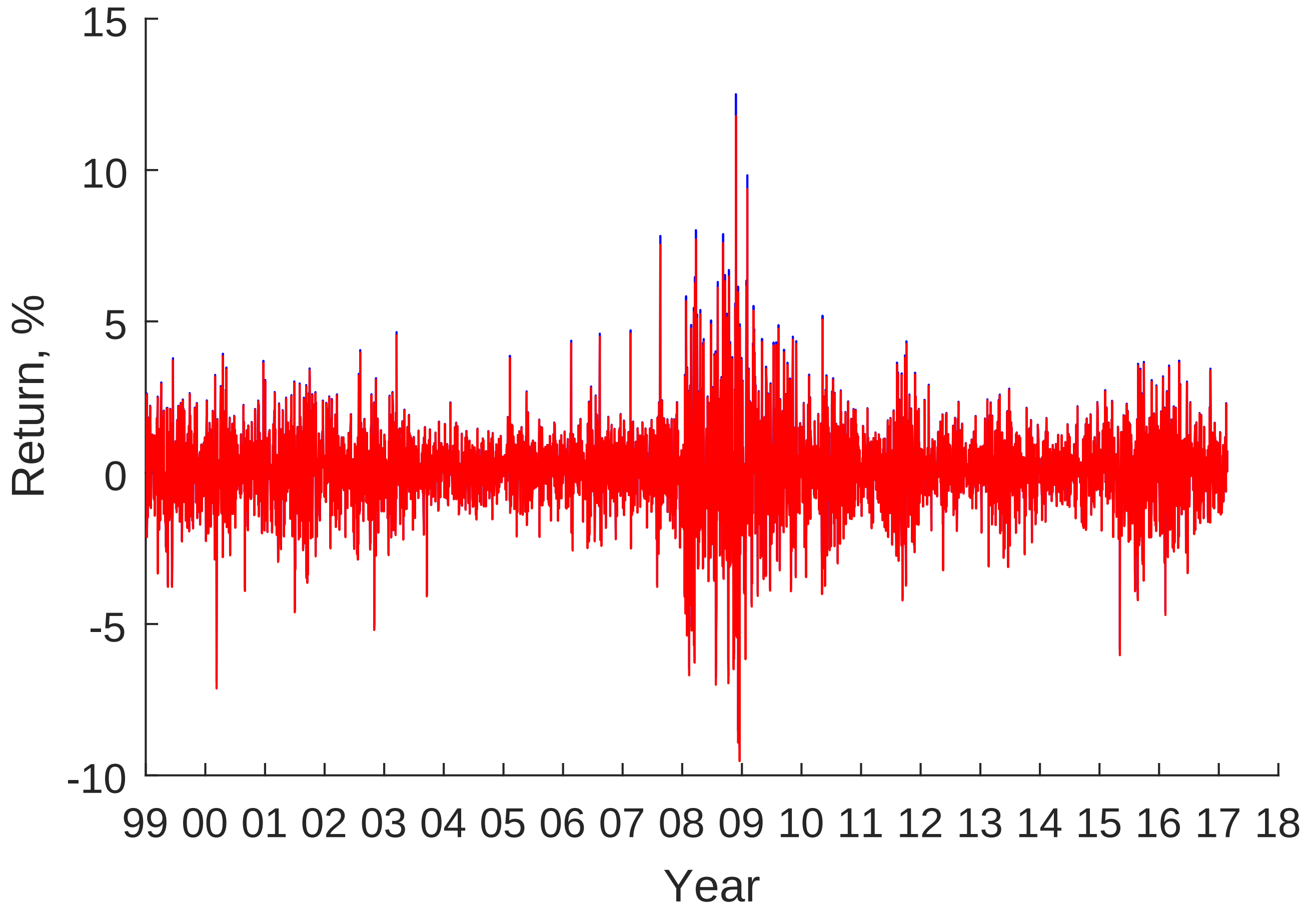

Figure 3: CBA simple (blue) and log-returns (red), in percentages, 01/1999-02/2016.

- Do log-returns always approximate simple returns well? Well enough?

- Figure 7 displays percentage log and simple returns on the same plot for small values of R_t where the differences are fairly small. Figure 8 shows the same plot for larger values of R_t. Comments?

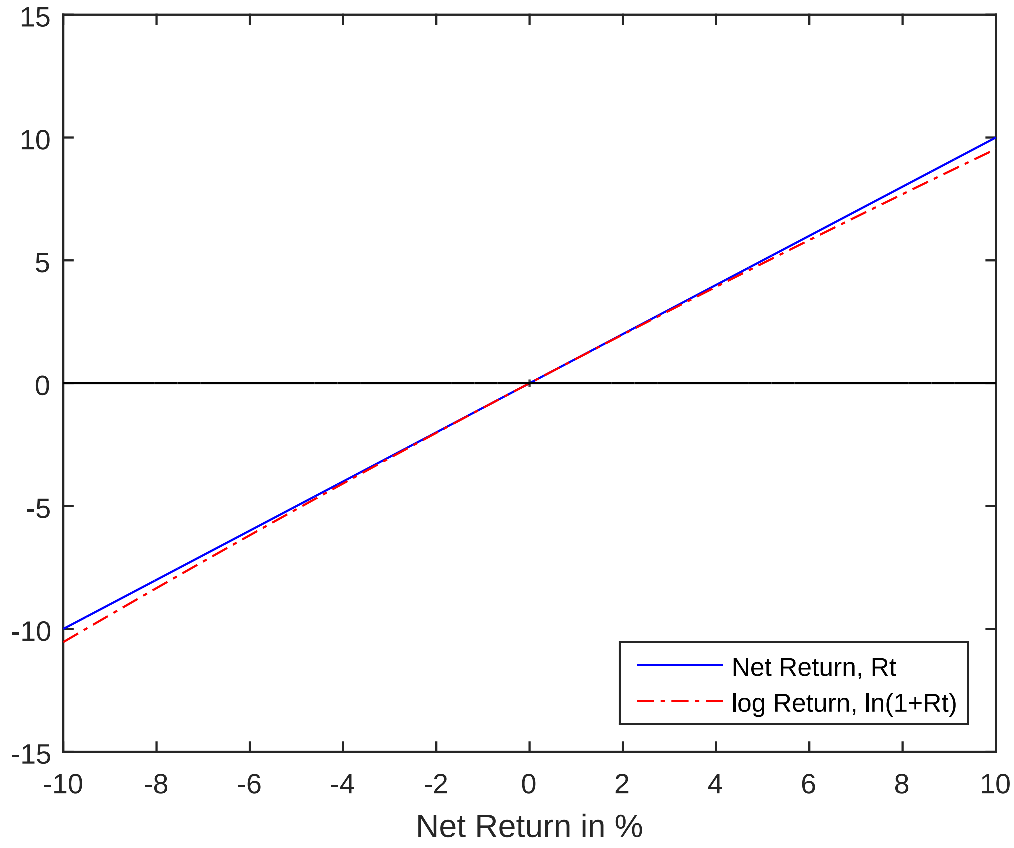

Figure 4: Percentage simple returns vs percentage log-returns for small returns.

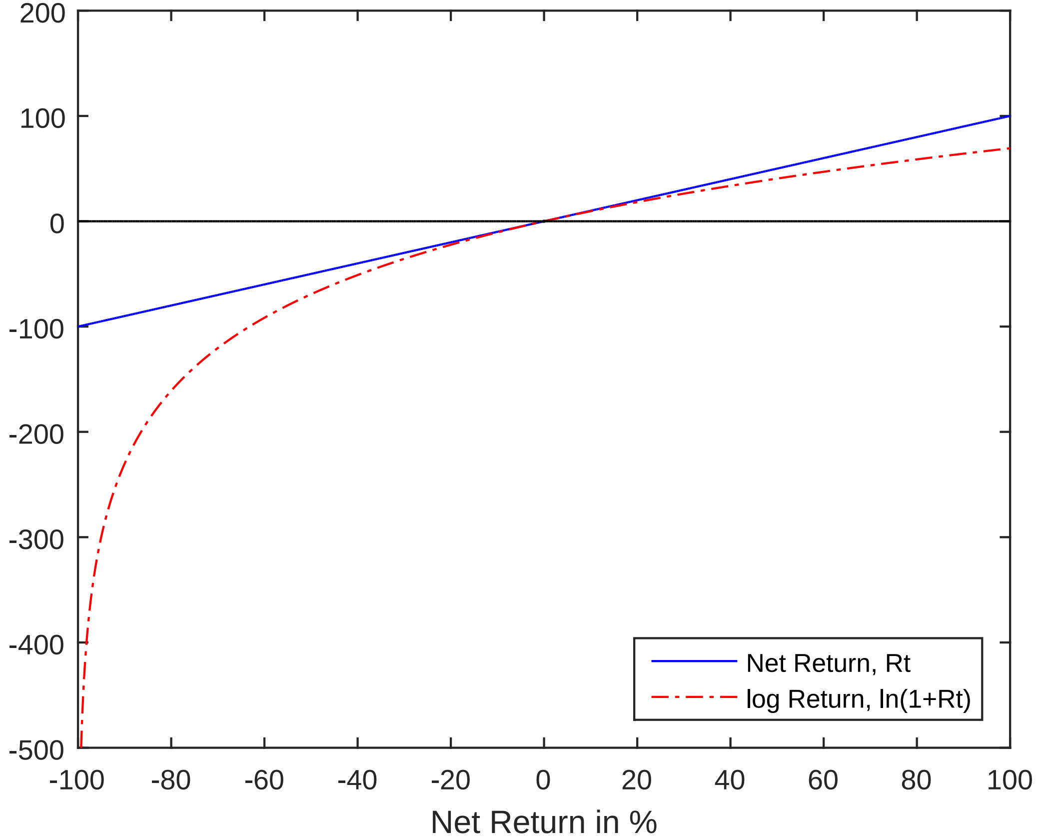

Figure 5: Percentage simple returns vs percentage log-returns for large returns.

1.2 Probability

- What is probability?

- Quantitative finance, asset pricing and risk management all depend on making probabilistic statements about financial instruments and returns.

- In this unit we will consider BOTH discrete and continuous probabilities.

- *Discrete* probability: X is a random variable (rv) and $p_i = Pr(X = X_i)$ for $i = 1, \ldots, D$ possible outcomes.

- Discrete probabilities have some basic rules:

  1. $0 \leq p_i \leq 1$
  2. $\sum_{i=1}^{D} p_i = 1$

- The collection $p_1, p_2, \ldots, p_D$ is called the discrete probability distribution for X.

- e.g. Consider the trade-by-trade changes in the price of IBM stock.

X	$\leq$ -3	-2	-1	0	1	2	$\geq$ 3
Pr(X)	0.013	0.029	0.169	0.577	0.169	0.027	0.014

- Are price changes really discrete?

- Figure 9 shows these probabilities in a bar chart.

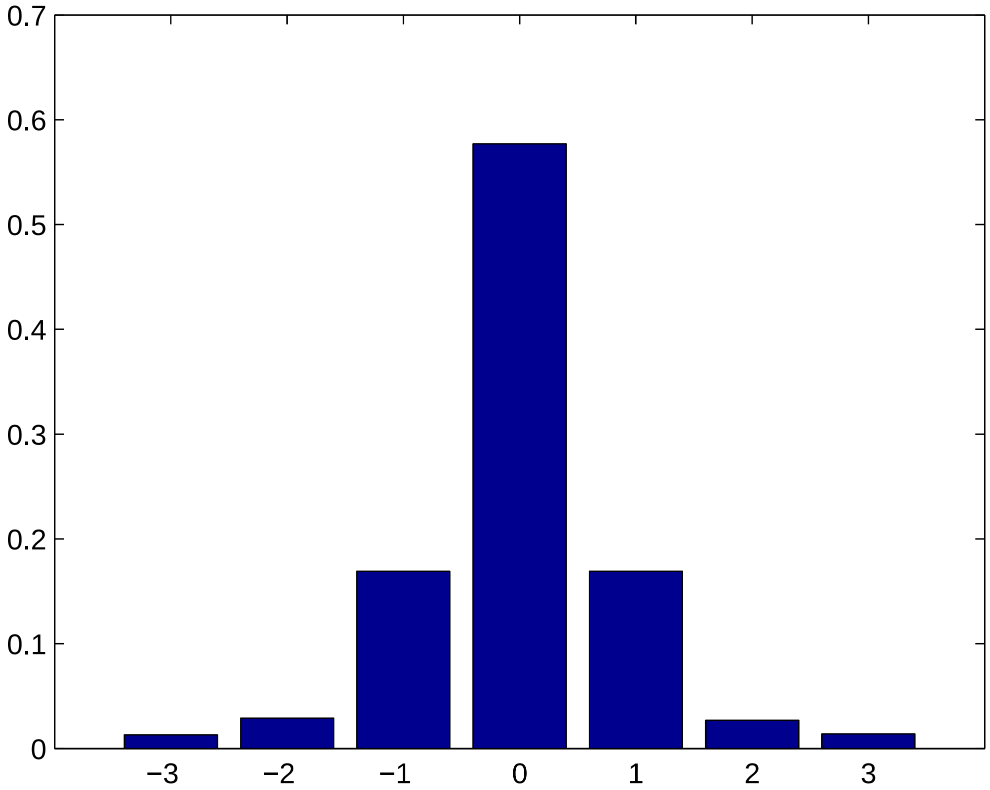

Figure 6: Trade by trade price changes in ticks of IBM stock.

- The probability that a price change is 0 or negative is $Pr(0)+Pr(-1)+Pr(-2)+Pr(\leq -3) = 0.577 + 0.169 + 0.029 + 0.013 = 0.788$ as shown in figure 10.

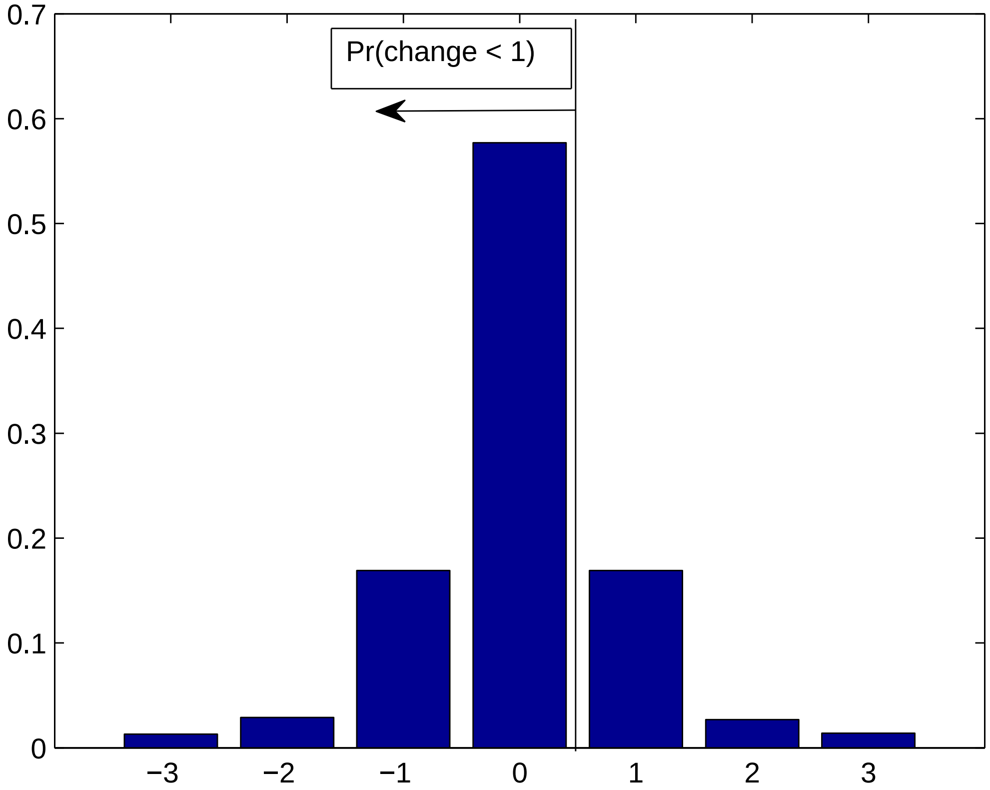

Figure 7: Trade by trade price changes in ticks of IBM stock.

- How do we know what these probabilities are?

- The mean or *expectation* of a discrete rv is defined as:

$$E(X) = \sum_{i=1}^{D} p_i X_i$$

- For a random sample from any distribution, the mean is estimated by:

$$\frac{1}{n} \sum_{t=1}^{n} x_t$$

- Why does this estimate $\sum_{i=1}^{D} p_i X_i$??

- *Continuous* probabilities are used for variables with a very large number of possible outcomes; e.g. ...

- Continuous probabilities are calculated for regions:

$$Pr(a < X < b) = \int_a^b p(x)dx$$

- $p(x)$ is called a probability density function (pdf).

- The exact probability of any value of X is ??
  e.g. $Pr(X = a) =?$

  why?

- The basic rules are:

  1. $0 \leq Pr(a < X < b) \leq 1$.
  2. $\int_{-\infty}^{\infty} p(x)dx = 1$.
  3. $p(x) \geq 0$.

- Figure 11 shows a continuous density function.

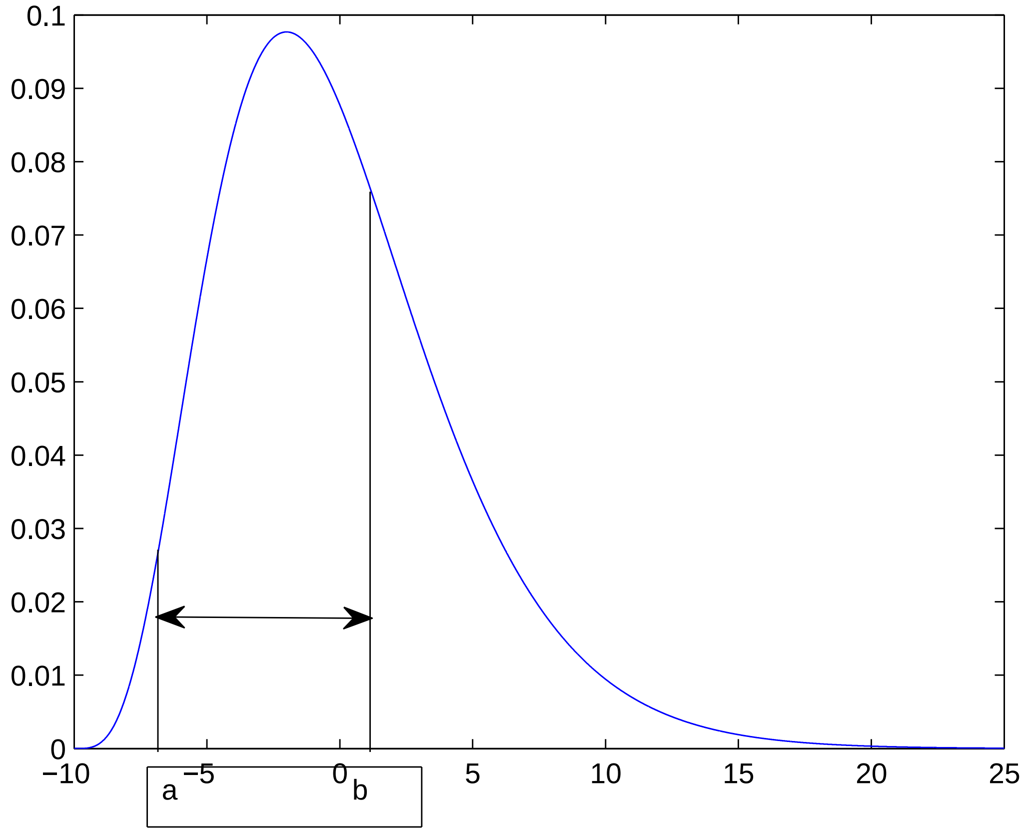

Figure 8: A continuous probability density function(pdf).

- The probability of that rv being between a and b is the area under the curve between the lines, which can be evaluated by $Pr(a < X < b) = \int_a^b p(x)dx$.

- The most common continuous distribution is the normal or Gaussian. See Figure 12

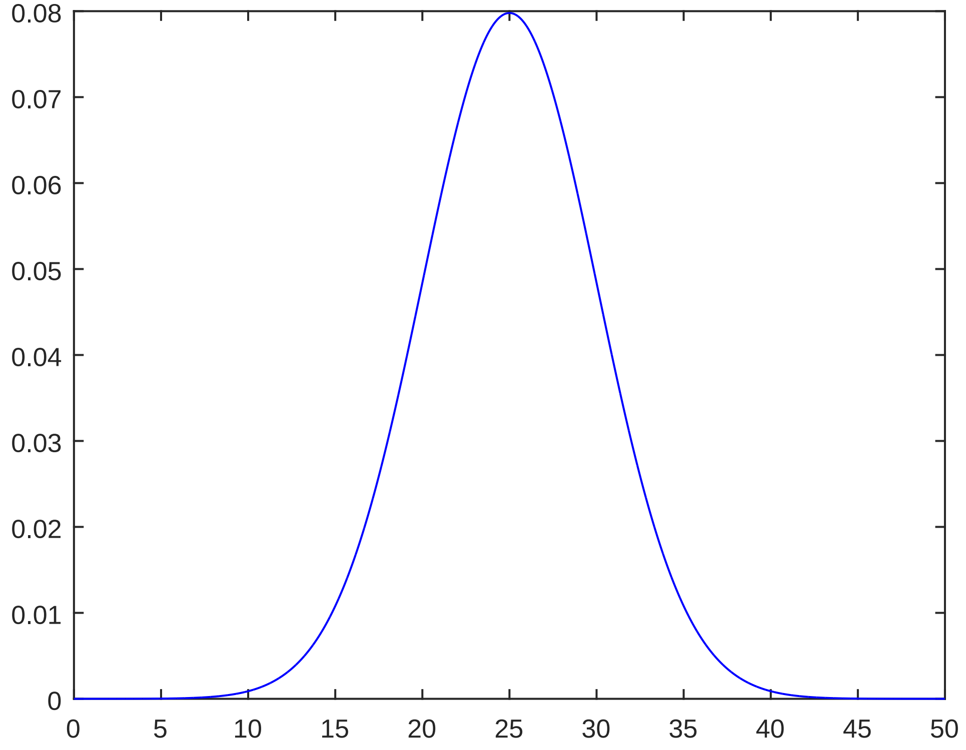

Figure 9: Normal probability density function when $\mu = 25 = \sigma^2$.

- Some Gaussian rules:

- The *expectation* of a continuous rv is defined as:

$$E(X) = \int xp(x)dx$$

  and is again interpreted as the average or mean of X.

- Some notation:

$$E(X) = \mu \; ; \; Var(X) = \sigma^2$$

- Variance is defined as:

$$\text{Var}(X) = E\left[(X - \mu)^2\right]$$

- Variance is the average squared distance of a rv X from its mean μ.

- For a random sample from any distribution, the variance can be estimated by:

$$\frac{1}{n}\sum_{t=1}^{n}(x_t - \bar{x})^2 \ OR \ s^2 = \frac{1}{n-1}\sum_{t=1}^{n}(x_t - \bar{x})^2$$

- The 2nd estimator is preferred and is called the sample variance.

- On average, over many samples, it will be correct; i.e. $E(s^2) = \sigma^2$

- If the average of an estimator across many samples is the true value it estimates, then the estimator is called **unbiased**.

- As the sample size becomes larger, the difference becomes negligible and both tend to the true value of σ^2.

- A Gaussian or normal distribution with mean μ and variance σ^2 is denoted
$$X \sim N(\mu, \sigma^2)$$
.

- The relative *likelihood* of each value of X occurring is given by the normal probability *density* function (pdf):
$$p(x) = (2\pi\sigma^2)^{-0.5} \exp\left[\frac{-(x-\mu)^2}{2\sigma^2}\right]$$

- The area under the Gaussian pdf determines the usual 'rules'.

- It also determines the CDF, i.e. $Pr(X < x) = \int_{-\infty}^{x} p(v)dv = F(x)$

- The height of the function in the plot (i.e. Figure 12) shows the *relative likelihood* for possible values of X.

Moments

- The kth *moment* of a distribution for rv X is defined as

$$\mu'_k = E(X^k)$$

where $E()$ stands for expectation, which is ...

- An expectation can be estimated from a sample $x_1, \ldots, x_n$ by:

$$E(X^k) \approx \frac{1}{n}\sum_{t=1}^{n} x_t^k$$

- The kth *central* moment is defined as

$$\mu_k = E\left((X - \mu'_1)^k\right) = E\left((X - \mu)^k\right)$$

which can be thus estimated by:

$$\frac{1}{n}\sum_{t=1}^{n}(x_t - \bar{x})^k$$

- For financial data and quant models in general, properties of moments can be highly important.

- First, if X is a rv with mean μ and variance σ^2, then $Z = \frac{X-\mu}{\sigma}$ is a rv with the SAME distribution as X and with $E(Z) = 0$, $\text{Var}(Z) = 1$. Z is a *standardised* variable.

- What is *skewness*? What is *kurtosis*? Intuitively ...

- Pearson's measure of *skewness* is defined as

$$\frac{\mu_3}{\text{Var}(X)^{3/2}} = E\left(\frac{(X-\mu)^3}{\sigma^3}\right) = E(Z^3)$$

where $Z = \frac{X-\mu}{\sigma}$.

- What does this measure ...? What does the 3rd central standardised moment represent about X

- Skewness has no units or scale. Thus we can compare skewness directly among ANY two rvs.

- Symmetric distributions have zero skewness. Why??

- A definition of symmetry is that, for a rv X, is $p(m-a) = p(m+a)$ where m is the median of X, i.e. $Pr(X < m) = 0.5$

- i.e. the pdf has the same height for all points equi-distant from the median.

- *Kurtosis* is defined as

$$\frac{\mu_4}{\text{Var}(X)^2} = E\left(\frac{(X-\mu)^4}{\sigma^4}\right) = E(Z^4)$$

- What does this measure ...?

- Kurtosis is also unit free and directly comparable between any two rvs.

- A normal or Gaussian distribution has skewness = 0 and kurtosis = 3.

- *Excess kurtosis* is defined as kurtosis - 3.

- Figure 13 shows some distributions with different skewness and kurtosis.

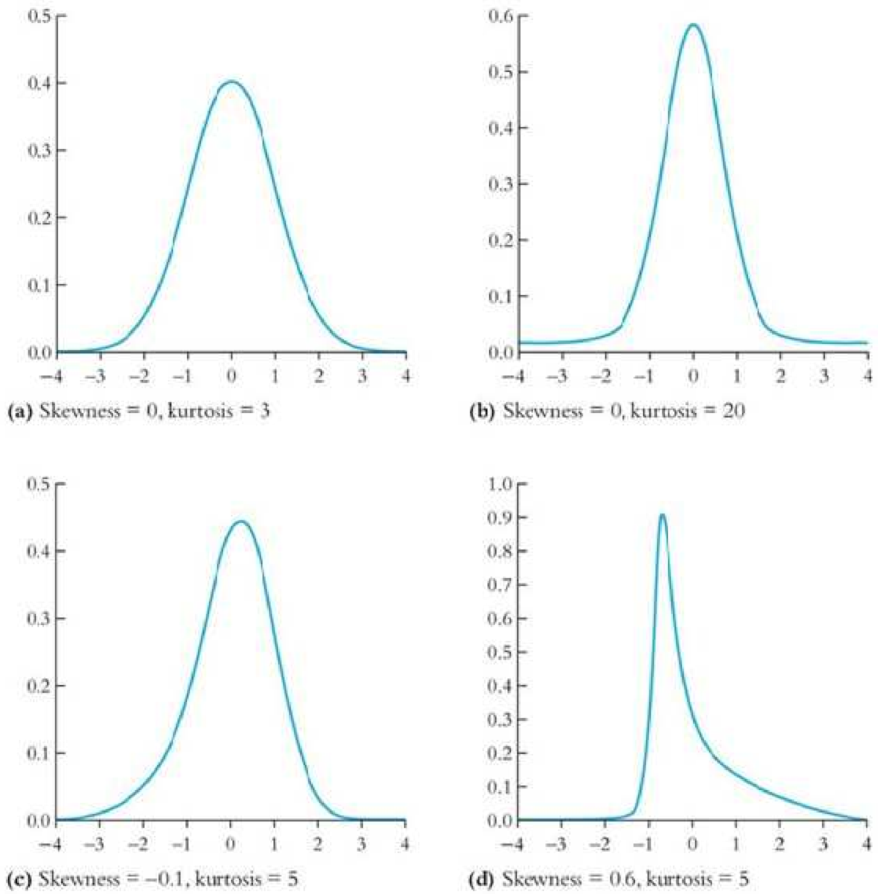

Figure 10: Four continuous distributions with varying skewness and kurtosis

- Log returns, i.e. $r_t = \ln(P_t/P_{t-1})$ are often assumed to follow a normal distribution in financial theory.

- For example, Black-Scholes option pricing theory, random walk price model, CAPM (later), etc.

- Figure 14 shows a histogram of daily percentage log-returns for CBA stock from January, 1999 to February, 2017.

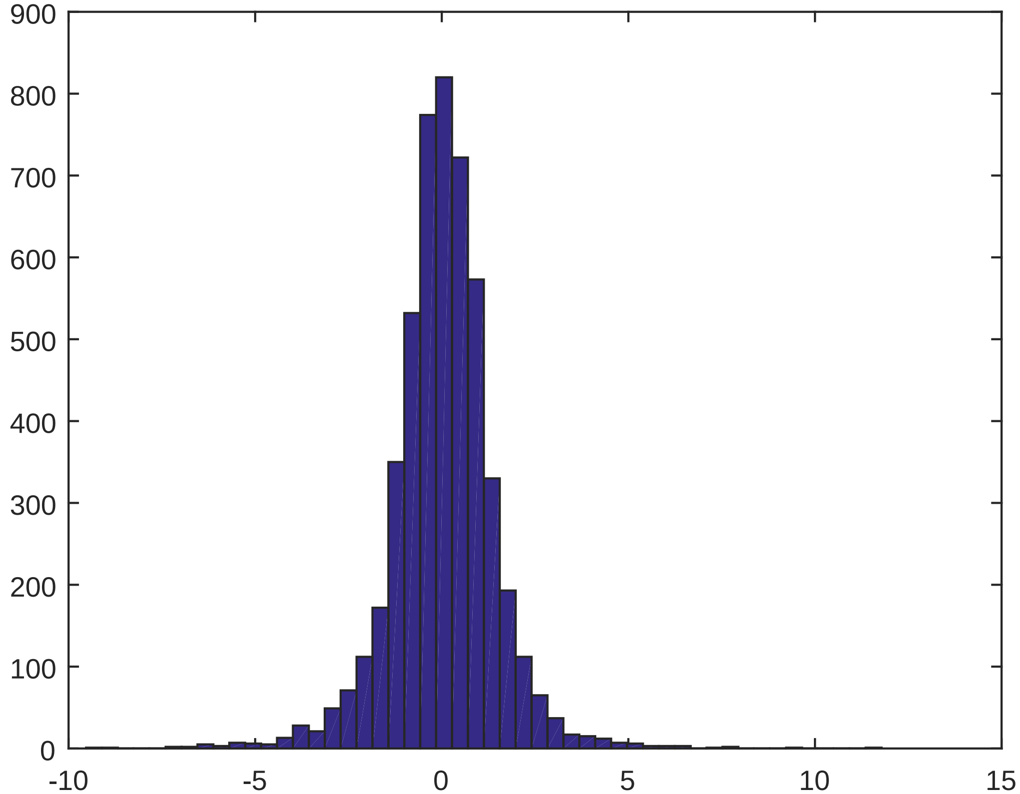

Figure 11: CBA returns 01/1999-02/2017.

- Figure 15 shows two estimated pdfs for CBA returns.

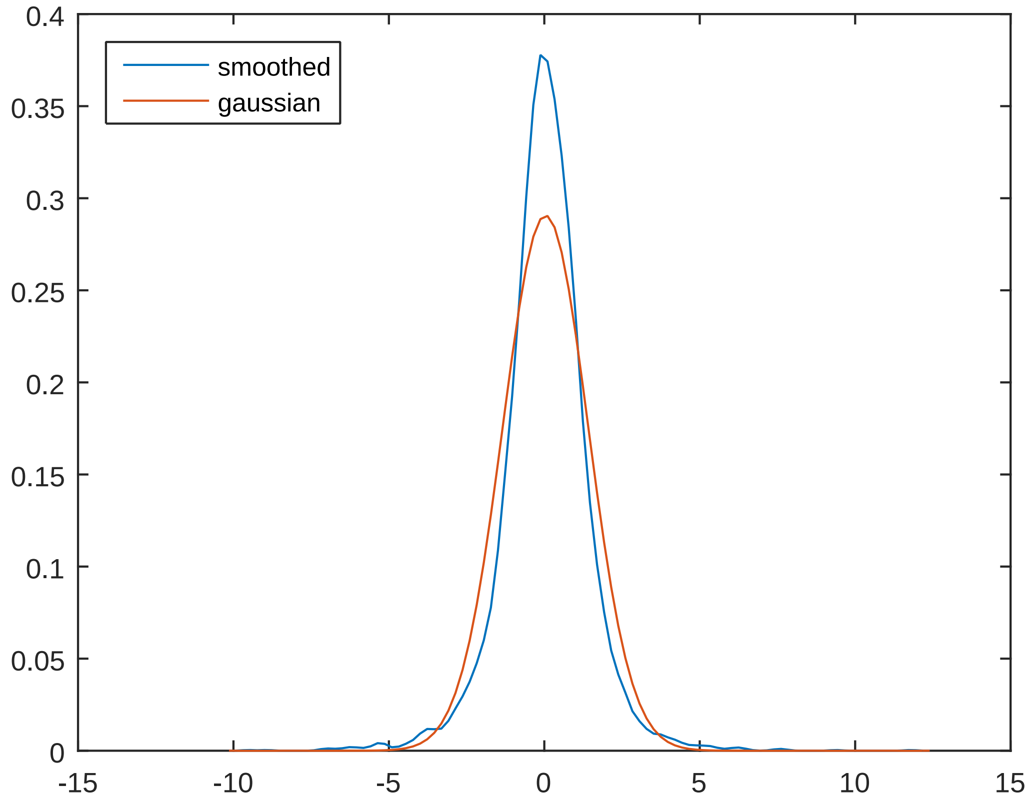

Figure 12: CBA returns estimated pdf and normal pdf with same mean and variance.

- One has been estimated assuming no distribution form and uses smoothing methods. The other estimates the mean and variance of the CBA returns and draws a normal pdf with that mean and variance.

- Compare the properties of CBA returns with those expected under a normal distribution.

- The Gaussian has roughly 68%, 95% and 99.7% within 1,2 and 3 standard deviations of the mean.

- CBA returns have 77.5%, 94.7% and 98.5% of values within 1,2 and 3 standard deviations of the mean. Thus, ...

-

%tile	0.1	0.5	1	2.5	5	25	50	75	95	97.5	99	99.5	99.9
return	-6.81	-5.14	-3.90	-2.80	-2.11	-0.65	0.04	0.77	2.06	2.66	3.75	4.65	6.89

- The mean CBA return is 0.038%; significantly different to 0 under a t-test (p-val = 0.046). *Does this result depend on which time period we choose?*

- Skewness is -0.06 ; Kurtosis is 8.28

- Note that:

$$\hat{S} = \frac{\sum_{t=1}^{n}(x_t - \bar{x})^3}{ns^3} \; ; \; \hat{\kappa} = \frac{\sum_{t=1}^{n}(x_t - \bar{x})^4}{ns^4}$$

- Discussion ...

- The Jarque-Bera test for normality is the most popular econometric test of a distributional assumption.

- The test focuses on whether the sample skewness and kurtosis of a dataset are consistent with normality (i.e. skewness =0, kurtosis =3)

- The test is a JOINT test for (the absence of) skewness and excess kurtosis.

- The null hypothesis is that:

$$H_0 : \text{skewness} = 0 \;\; \text{AND} \;\; \kappa = 3$$

- The test statistic is:

$$JB = \frac{n}{6}\left(S^2 + \frac{(\hat{\kappa} - 3)^2}{4}\right)$$

that (asymptotically) has a Chi-squared distribution with 2 df, under the null hypothesis, if the series is i.i.d. Gaussian.

- The test makes use of the fact that, if a time series is i.i.d. Gaussian, then, asymptotically:

$$S \sim N\left(0, \frac{6}{n}\right) \;\; ; \;\; \hat{\kappa} \sim N\left(3, \frac{24}{n}\right)$$

- The test is BEST performed on the standardised residuals from a model. *why??*

- For the CBA percentage log-returns, the p-value from this test is 0.001, so Gaussianity is clearly and strongly rejected.

Conditional and Joint Probability Distributions

- *Conditional* probability measures a distribution given some knowledge.
- Consider the linear regression model:

$$y_t = \beta_0 + \beta_1 X_t + \epsilon_t.$$

- If y is the excess return on a portfolio and X is the market portfolio excess return, then ... ?
- If we assume $\epsilon_t \sim N(0, \sigma^2)$, then the conditional distribution of y given X is: $y_t|x_t \sim N(\mu_t, \sigma^2)$, where $\mu_t = \beta_0 + \beta_1 X_t$.
- Regression and time series models are similarly based on conditional distributions.
- Here we have:

$$E(y_t|X_t) = \beta_0 + \beta_1 X_t \ ; \ Var(y_t|X_t) = \sigma^2$$

- The *joint* distribution for two random variables can be described by the joint CDF:
$$F(x, y) = Pr(X < x, Y < y) = \int_{-\infty}^{x} \int_{-\infty}^{y} p(u, v)dudv.$$

- The *conditional* probability density function (pdf) can be defined in terms of the *marginal* and joint pdfs:
$$p(y|x) = \frac{p(x, y)}{p(x)}$$

- This implies that:

- Two variables are *independent* iff
$$p(y|x) = p(y)$$
or equivalently
$$p(x, y) = p(x)p(y)$$

- Further, if two variables are independent, then:

$$E(y|X) = E(y); Var(y|X) = Var(y)$$

i.e. the conditional mean and variance of y ...

- However, the reverse conditions do NOT always apply.

1.3 Introduction to financial price/return modelling

- Time series data is simply data that is recorded over time.

- Time series models allow for dynamic patterns to be captured.

- An important property of time series is stationarity.

- Roughly, stationarity means that the properties of a time series remain the same over time.

- For example, a stationary in mean series has a constant long-mean that it continually reverts to and moves around.

- Are stock prices mean stationary? Stock returns?

The Random Walk Model

- A famous model for asset returns is:

$$p_t = p_{t-1} + \epsilon_t$$

where $p_t = \ln(P_t)$ and $\mathrm{E}(\epsilon_t) = 0$.

- This is the random walk (RW) model for (log-) stock prices.

- Why is it called a random walk? Why is it on log prices?

- Note that $r_t = \ln\left(\frac{P_t}{P_{t-1}}\right)$ implies that:

$$\ln(P_t) = \ln(P_{t-1}) + r_t \;\equiv\; P_t = P_{t-1}\exp(r_t)$$

and hence that $r_t = \epsilon_t$

- Thus, under the Gaussian RW model, we expect log-returns to be i.i.d. $N(0, \sigma^2)$

- This is the model first assumed by Bachelier (1905), then Black, Scholes, Merton, etc.

- Figure 16 shows some simulated log-price series from the Gaussian RW model.

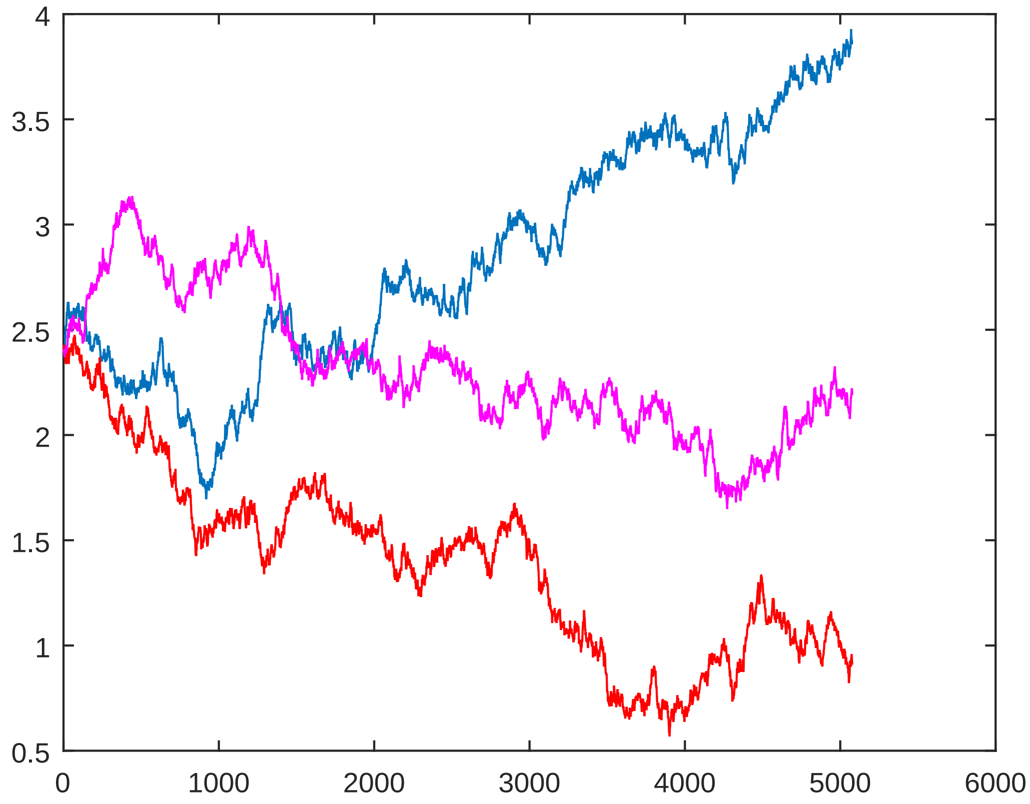

Figure 13: 3 simulated RW series

- What are some of the basic properties of log-price data from the RW model?

- Figure 17 shows the 1st differences from the simulated RW data. What are the properties of log-returns data under the Gaussian RW?

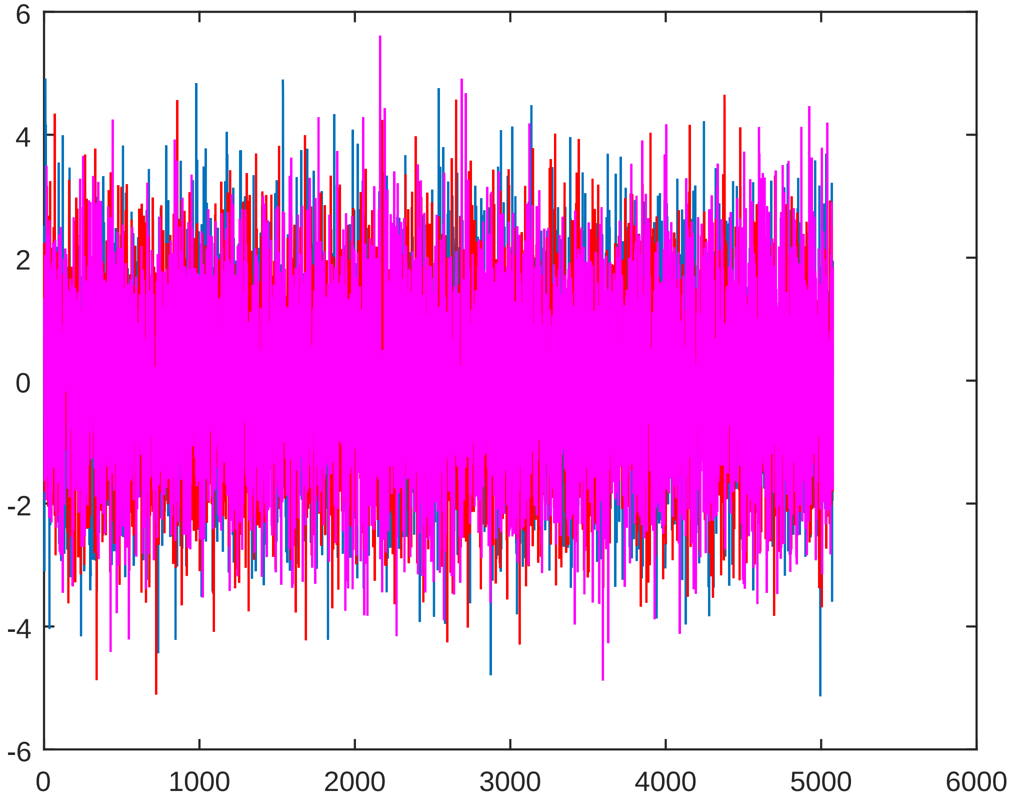

Figure 14: 3 differenced simulated RW series

- Figure 18 shows simulated data from the Gaussian RW model plus the series of log-prices from CBA

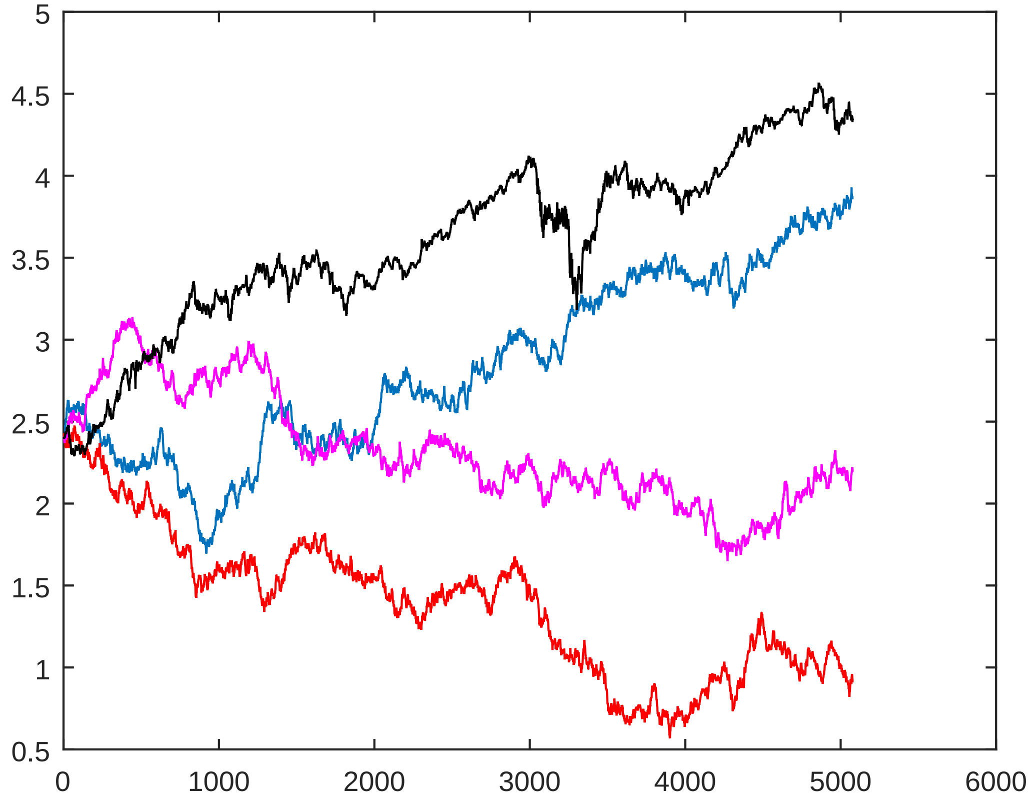

Figure 15: 3 simulated RW series plus the CBA log prices (black)

- Figure 19 shows 1st differences of the Gaussian RW model plus the series of log-returns from CBA

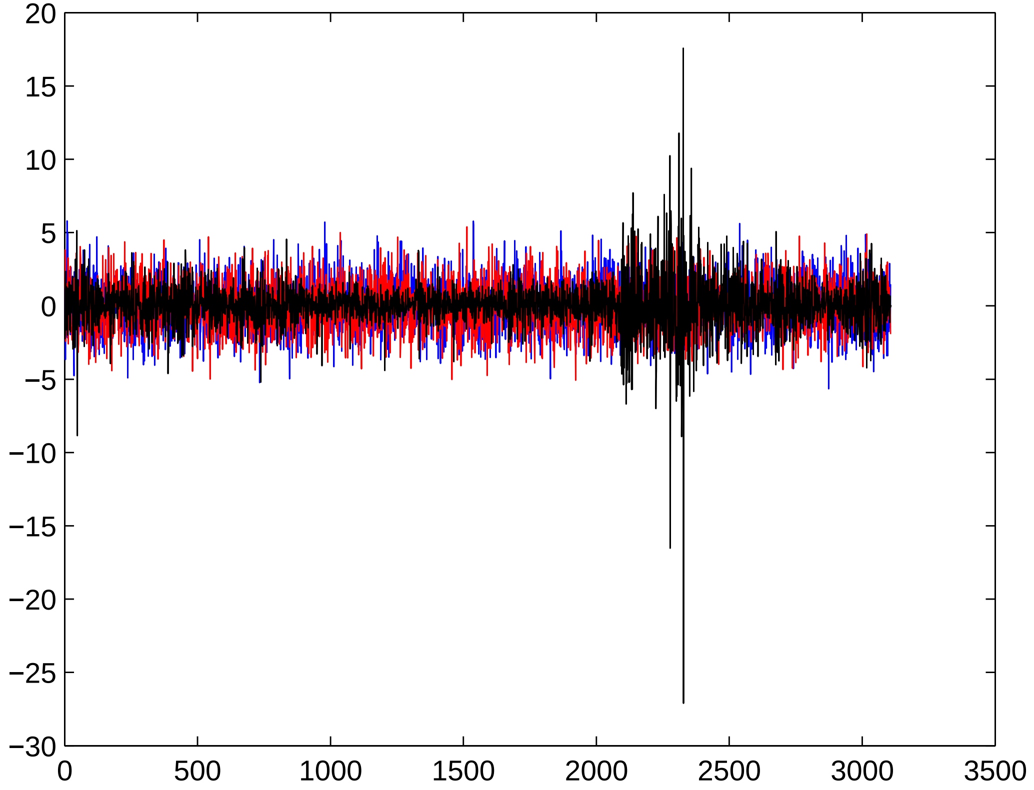

Figure 16: 1st differenced simulated RW data plus the CBA log returns (black)

- Figure 20 shows the histograms of these series

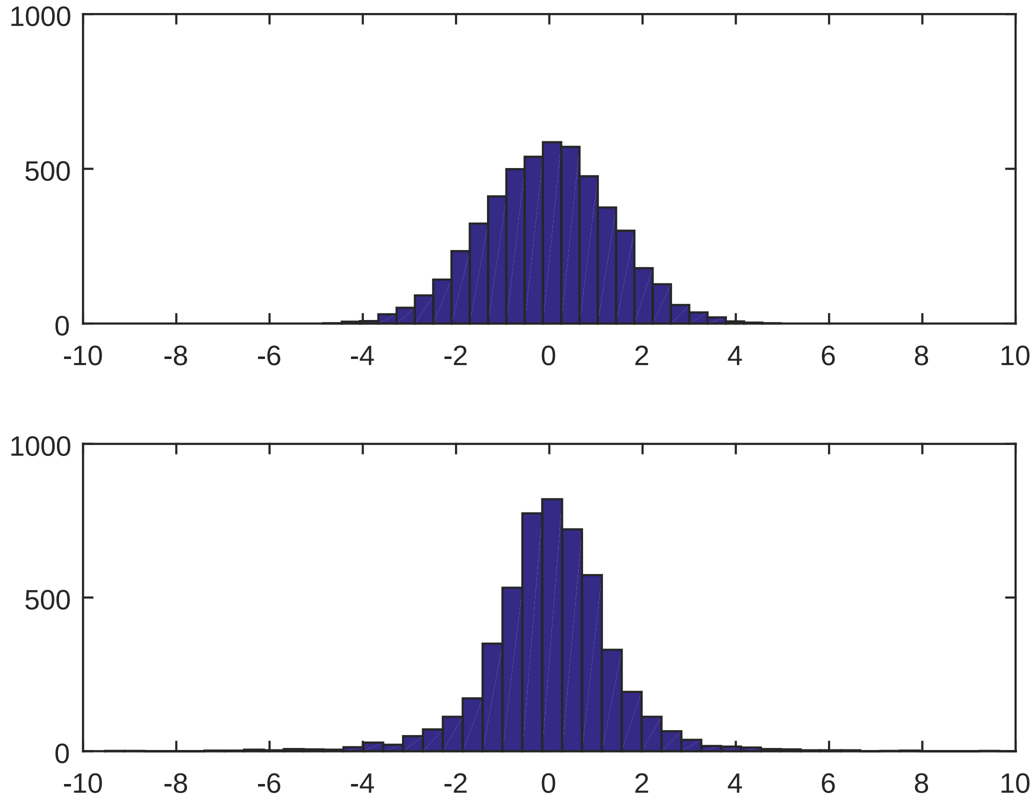

Figure 17: Top: Histogram of 1st differenced simulated RW data Bottom: Histogram of the CBA log returns

- What aspects seem apparent in the real (CBA) data but not in the simulated data and vice versa?

- Should we reject the simple RW model?