FINANCIAL RISK MANAGEMENT 2

Ta Quoc Bao

Department of Mathematics,
International University-VNUHCM

Course Description:

This course provides students notions and advanced tools of statistics and Mathematics for financial risk management e.g., time series models, copula theory and Extreme Value Theory.

References

1. A.J. MacNeil, R. Frey and P. Embrechts. Quantitative risk management. Princeton University press. 2015
2. J. Danielsson. Financial Risk Forecasting. Wiley Finance. 2011
3. Allan M. Malz, Financial Risk Management: Models, History, and Institutions, Willey, 2011.

1.1. Prices and Returns

Denote P_t price of a stock at time t. Usually we are interested in the return that we make on an investment

Definition 1

Return is the relative change in the price of a financial asset over a given time interval, often represented as a percentage.

There are two type of returns including

- Simple returns (R)
- Compound or log returns $\log(Y)$

Definition 2 (Simple returns)

A simple return is the percentage change in prices indicated by R

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

The term $1 + R_t$ is called simple gross return
If the stock payed dividends d_t then

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

Property: a multi-period (n-period) (daily or annually) return is given by

$$R_t(n) = \frac{P_t}{P_{t-n}} - 1 = (1 + R_t)(1 + R_{t-1})(1 + R_{t-2}) \cdots (1 + R_{t-n+1}) - 1$$

Definition 3

The logarithm of gross return is called continuously compounded return, indicated by Y or denoted by r

$$Y_t = \log(1 + R_t) = \log\left(\frac{P_t}{P_{t-1}}\right) = \log(P_t) - \log(P_{t-1}).$$

Property: The advantages of compound returns for multiperiod returns:

$$\begin{aligned} Y_t(n) = \log(1 + R_t(n)) &= \sum_{i=0}^{n-1} \log(1 + R_{t-j}) \\ &= Y_t + Y_{t-1} + \cdots Y_{t-n+1} \end{aligned}$$

Remark For small price changes the difference of simple return and log return is small (negligible). Indeed, from Taylor approximation we have

$$\log(1+x) = x - \frac{x^2}{2!} + \frac{x^3}{3!} + ... \approx x$$

As a general rule, with returns under 10% it does not matter if we can use simple or log-return.

Compare Simple and continuous compounded returns

- There is not large difference between R_t and Y_t
- As the time between observations goes to zero then

$$\lim_{\Delta t \to 0} Y_t = R_t$$

Indeed, for example:

(i) $\log(1000) - \log(995) = 0.005012 \approx \frac{1000}{995} - 1 = 0.005025$

(ii) $\log(1000) - \log(885) = 0.12216 \neq \frac{1000}{855} - 1 = 0.12994$

Symmetry property

Continuous compounded return is symmetry, but Simple return is not. For example

- $\log\left(\frac{1000}{500} = -\log\left(\frac{500}{1000}\right)\right)$
- $\frac{1000}{500} - 1 \neq -\left(\frac{500}{1000} - 1\right)$

Remark

Consider a portfolio of N stocks with simple returns $R_{t,i}$ at time t, respectively. Denote $R_{t,p}$ the return of portfolio at time t, $Y_{t,p}$ the continuously compounded return of the portfolio at time t.

(i) We have the simple return of the portfolio is (proof!!!)

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

(ii) For continuously compounded returns we do not have equality

$$Y_{t,p} = \log\left(\frac{P_{t,p}}{P_{t-1,p}}\right) \neq \sum_{i=1}^{n} \omega_i \left(\frac{P_{t,i}}{P_{t-1,i}}\right) = \sum_{i=1}^{n} \omega_i Y_{t,i}$$

However, the difference between compounded and simple returns may not be very significant for small returns, e.g., daily return

$$Y_p = \sum_{i=1}^{N} \omega_i R_i$$

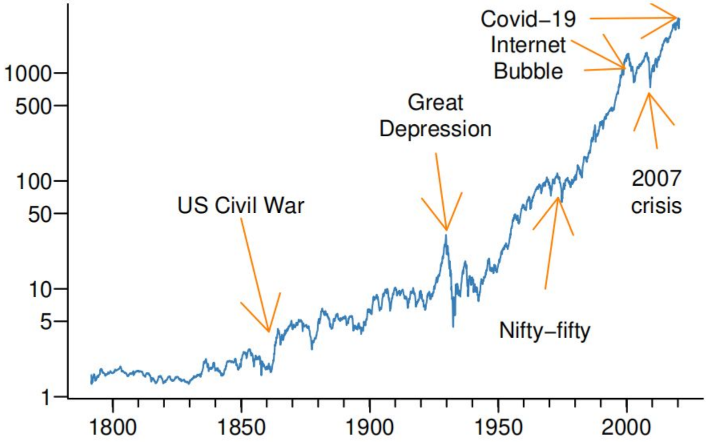
when time between observations goes to 0, then we have

$$\lim_{\Delta t \to 0} Y_{t,p} = R_{t,p}$$

So, in practice we note that

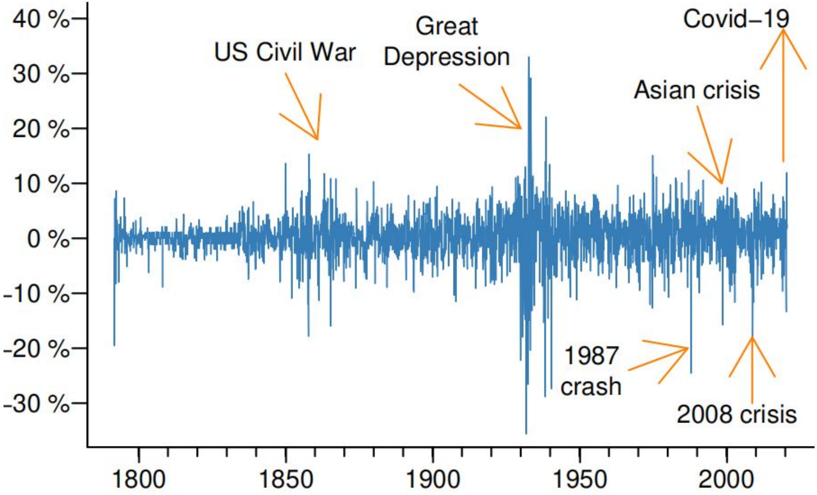
- Simple returns are
  - Used for accounting purposes
  - Investors are usually concerned with simple returns
- Continuously compounded returns have some advantages
  - Mathematics is easier, we will see later
  - Used in derivatives pricing, e.g. the Black–Scholes model

S&P 500 index



Covid–19
Internet
Bubble
Great
Depression
US Civil War
2007
crisis
Nifty–fifty
1000
500
100
50
10
5
1
1800
1850
1900
1950
2000

S&P 500 returns



1.2. The random walk model

Definition

Let $Z_1, Z-2, .., $ be i.i.d random variables with mean μ and standard deviation σ. Let S_0 be an arbitrary starting point and

$$S_t := S_0 + Z_1 + Z_2 + ... + Z_t$$

The process $(S_t)_{t\geq 0}$ is a called random walk and $Z_1, Z_2, .., $ are its steps. If the steps are normally distributed, then the random walk is called a normal random walk. We have

We write as

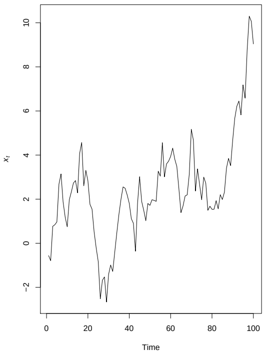
$$S_t = c + S_{t-1} + \epsilon_t$$

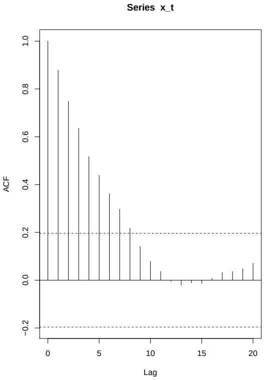
with c is a constant and ϵ_t is white noise, i.e., ϵ_t is a stochastic process with mean 0, variance σ^2 and uncorrelated.

$$\mathbb{E}(S_t \mid S_0) = S_0 + \mu t \quad \text{and} \quad Var(S_t \mid S_0) = \sigma^2 t$$

Parameter μ is called the drift determines the general directions of random walk, and σ is called the volatility and determines how much the random walk fluctuates about the conditional mean $S_0 + \mu t$.

We use R to simulate a normal random walk with 100 steps





Geometric random walks

We have

$$\log\left(\frac{P_t}{P_{t-n}}\right) = \log(1 + R_t(n)) = Y_t + Y_{t-1} + \cdots + Y_{t-n+1}$$

hence,

$$\frac{P_t}{P_{t-n}} = e^{Y_t + Y_{t-1} + \cdots + Y_{t-n+1}}$$

taking $n = t$ then

$$P_t = P_0 e^{Y_t + Y_{t-1} + \cdots + Y_1}$$

this process is called geometric random walks or exponential random walk. if $Y_1, Y_2, ...,$ are i.i.d, $N(\mu, \sigma^2)$ then P_t is lognormal for all t

Remark

The lognormal geometric random walk needs two assumptions

(1) the log returns are normally distributed.

(2) the log returns are mutually independent.

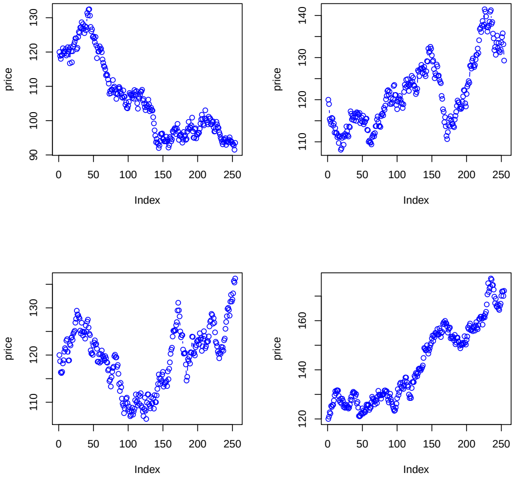
(3) in general, prices does not usually follow a lognormal geometric random walk or its continuous-time analog, geometric Brownian.

(4) The independence assumption can be also violated. Since,

   (i) there is is some correlation between returns.

   (ii) returns exhibit volatility clustering, i.e., if we see high volatility in current returns then we can expect this higher volatility to continue, at least for a while motion

Simulation of 4 geometric random walks



Practice: write your R code for simulation 6 geometric random walks

1.2. Volatility

Definition

There are two concepts of volatility:

(1) Unconditional volatility, or volatility for short, is volatility over an entire time period, denoted by σ

(2) Conditional volatility is volatility in a given time period, conditional on what happened before, denoted by σ_t

1.2. Volatility

Definition

There are two concepts of volatility:

(1) Unconditional volatility, or volatility for short, is volatility over an entire time period, denoted by σ

(2) Conditional volatility is volatility in a given time period, conditional on what happened before, denoted by σ_t

Remark

- The subscript t means that it is volatility on a particular time period, usually a day
- Clear evidence of cyclical patterns in volatility over time, both in the short run and the long run

Calculations

Consider a sample x_i with mean μ and sample size N. Then we have an estimation of volatility

- For daily volatility

$$\sigma = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(x_i - \mu)^2}$$

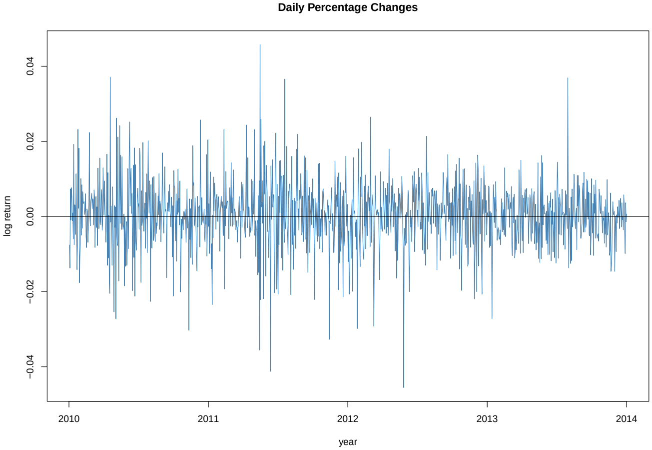
- For annually volatility

$$\sigma = \sqrt{250}\sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(x_i - \mu)^2}$$

why 250 but not 360?

Volatility clusters

The volatility over a decade, year and month, we see that it comes in many cycles we called these volatility clusters. The following figure describes the daily volatility of McDonald's stock from 2010-2014



1.3 Distributions for Financial Data

Note that under the Random Walk model, assuming independent Gaussian single-period returns, the distribution of both multi-period returns and the prices are derived. However, log returns are typically heavy tailed and thus these results are in question.

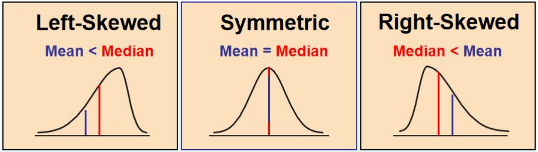
1.3.2. Skewness, Kurtosis, and Moments

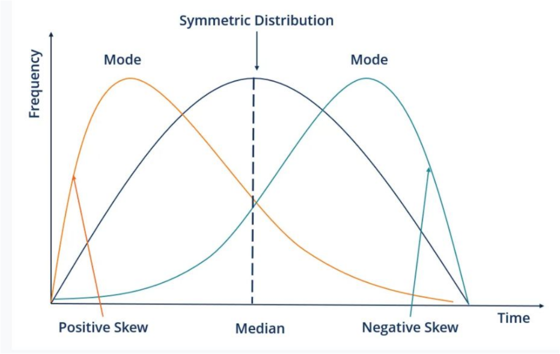
In basic statistic and probability theory, we almost exclusively deal with the first and second center moment of a random variable, namely expectation and variance.

- $k-th$ moment of r.v X: $m_k := \mathbb{E}(X^k)$, e.g., $\mu = \mathbb{E}(X)$
- $k-th$ center moment of r.v X: $\mu_k := \mathbb{E}(X-\mu)^k$, e.g., $Var(X) = \mathbb{E}(X-\mu)^2$

Skewness and Kurtosis help characterize the shape of a probability distribution.

- Skewness measures the degree of asymmetry.
  - zero skewness indicated a symmetric distribution
  - positive skewness (right-skewed) indicates a relatively long right tail compared to the left tail.
  - negative skewness (left-skewed) indicates a relatively long left tail compared to the right tail.





Definition. The skewness of a random variable X is

$$Sk = \mathbb{E}\left\{\frac{X - \mathbb{E}(X)}{\sigma}\right\}^3 = \frac{\mathbb{E}(X - \mathbb{E}(X))^3}{\sigma^3}$$

- If $Sk = 0$ then X has symmetric distribution
- If $Sk > 0$ then X has positive skewness, i.e., distribution has a heavy tail on the right hand side
- If $Sk < 0$ then X has negative skewness, i.e., distribution has a heavy tail on the left hand side

Financial situation: If two investments' return distributions have identical mean and variance, but different skewness parameters. Which one is to prefer? Typically, risk managers are wary of negative skew, in this situation, small gains are the norm, but big losses can occur, carrying risk of going bankruptcy.

- If a return distribution shows a positive skew, investors can expect recurrent small losses and few large returns from investment. Conversely, a negatively skewed distribution implies many small wins and a few large losses on the investment.
- Hence, a positively skewed investment return distribution should be preferred over a negatively skewed return distribution since the huge gains may cover the frequent – but small – losses. However, investors may prefer investments with a negatively skewed return distribution. It may be because they prefer frequent small wins and a few huge losses over frequent small losses and a few large gains.

- Kurtosis is a statistical measure that defines how heavily the tails of a distribution differ from the tails of a normal distribution. In other words, kurtosis identifies whether the tails of a given distribution contain extreme values.

- Kurtosis is a statistical measure that defines how heavily the tails of a distribution differ from the tails of a normal distribution. In other words, kurtosis identifies whether the tails of a given distribution contain extreme values.
- Along with skewness, kurtosis is an important descriptive statistic of data distribution. However, the two concepts must not be confused with each other. Skewness essentially measures the symmetry of the distribution, while kurtosis determines the heaviness of the distribution tails.

- Kurtosis is a statistical measure that defines how heavily the tails of a distribution differ from the tails of a normal distribution. In other words, kurtosis identifies whether the tails of a given distribution contain extreme values.
- Along with skewness, kurtosis is an important descriptive statistic of data distribution. However, the two concepts must not be confused with each other. Skewness essentially measures the symmetry of the distribution, while kurtosis determines the heaviness of the distribution tails.
- In finance, kurtosis is used as a measure of financial risk. A large kurtosis is associated with a high level of risk for an investment because it indicates that there are high probabilities of extremely large and extremely small returns. On the other hand, a small kurtosis signals a moderate level of risk because the probabilities of extreme returns are relatively low.

Definition. Kurtosis of a random variable X is

$$Kur(X) = \mathbb{E}\Big(\frac{X - \mathbb{E}(X)}{\sigma}\Big)^4 = \frac{(X - \mathbb{E}(X))^4}{\sigma^4.}$$

The kurtosis of a normal variable is 3.

Example

Let X have a binomial distribution $B(p, n)$. Then

$$Kur(X) = 3 + \frac{1 - 6p(1-p)}{np(1-p)}$$

Let X have t-distribution then

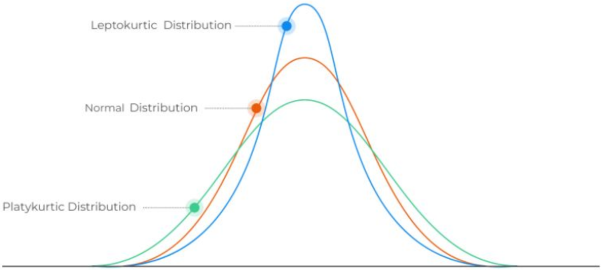
$$Kur(X) = 3 + \frac{6}{\nu - 4}$$

Type of Kurtosis

$$\kappa = \text{Excess Kurtosis } = \text{Kurtosis} - 3$$

1) Mesokurtic. Data that follows a mesokurtic distribution shows an excess kurtosis of zero or close to zero. This means that if the data follows a normal distribution.
2) Leptokurtic. Leptokurtic indicates a positive excess kurtosis $(\kappa > 0)$. The leptokurtic distribution shows heavy tails on either side, indicating large outliers. In finance, a leptokurtic distribution shows that the investment returns may be prone to extreme values on either side. Therefore, an investment whose returns follow a leptokurtic distribution is considered to be risky. It means that big losses (as well as big gains) can occur.

3) Platykurtic. A platykurtic distribution shows a negative excess kurtosis. The kurtosis reveals a distribution with flat tails. The flat tails indicate the small outliers in a distribution. In the finance context, the platykurtic distribution of the investment returns is desirable for investors because there is a small probability that the investment would experience extreme returns.



- Moments. Let X be a random variable. The kth moment of X is $\mathbb{E}(X^k)$. The kth absolute moment is $\mathbb{E} \mid X \mid^k$. The kth moment center is

$$\mu_k = \mathbb{E}\big(X - \mathbb{E}(X)\big)^k$$

The skewness coefficient of X is

$$Sk(X) = \frac{\mu_3}{(\mu_2)^{3/2}}$$

and Kurtosis of X is

$$Kur(X) = \frac{\mu_4}{(\mu_2)^2}$$

Let $X_1, X_2, ..., X_n$ be a observations of X with sample mean $\bar{X}$ and sample standard deviation s. Then the sample skewness denoted by $\widehat{Sk}$ is

$$\widehat{Sk} = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{X_i - \bar{X}}{s}\right)^3$$

and the sample kurtosis, denoted by $\widehat{Kur}$, is

$$\widehat{Kur} = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{X_i - \bar{X}}{s}\right)^4$$

1.3. Fat tails

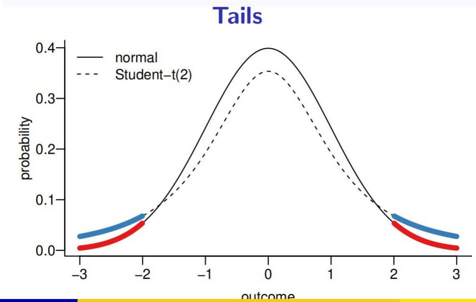
Definition of Fat tails (also known as heavy tails). A random variable is said to have fat tails if it exposes more extreme outcomes than a normal distributed random variable with the same mean and variance.

- On the other hand, fat tails describe the greater-than-expected probabilities of extreme values. Financiers have used the normal distribution to model the distribution of probabilities for the values of a quantity, such as price returns
- The mean–variance model assumes normality
- The tails are the extreme left and right parts of a distribution

Example

Consider the t-student X with degrees of freedom ν. The values of ν indicate how fat the tails are

- if $\nu = \infty$ then X is the normal
- if $\nu < 2$ then X has fat tails
- For a typical stock we have $3 < \nu < 5$
- The t-Student distribution is convenient for modeling a fat tailed distribution



1.3.1 Identification of fat tails

Two main approaches for identifying and analyzing tails of financial returns including:

(i) statistical tests

(ii) graphical methods

- The Jarque-Bera (JB) and the Kolmogorov-Smirnov (KS) tests are popular statistical methods to test for fat tails
- QQ plots is a common graphical method to analyze tails graphically by comparing quantiles of sample data with quantiles of reference distribution

The Jarque-Bera Test of normality compares the sample skewness and kurtosis to 0 and 3, their value under normality. The test statistic is

$$JB = n\left(\frac{\widehat{Sk}}{6} + \frac{(\widehat{Kur} - 3)^2}{24}\right) \sim \chi^2(2)$$

in R, the test can be computed with: $jarque.bera.test()$

Q-Q plots

- A QQ plot (quantile-quantile plot) compares the quantiles of sample data against quantiles of a reference distribution, like normal
- Used to assess whether a set of observations has a particular distribution
- Can also be used to determine whether two datasets have the same distribution

```
library(car)
qqPlot(y)
qqPlot(y,distribution="t",df=5)
```

Definition

the p^{th} quantile of CDF F of a random variable X is that the value x_p such that

$$F(x_p) = p \quad \text{or} \quad x_p = F^{-1}(p)$$

Suppose we want to compare two CDF: F and G

Definition

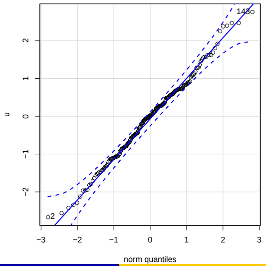
The theoretical Q-Q plot is the graph of the quantiles of a the CDF F, $x_p = F^{-1}(p)$, versus the corresponding quantiles of the CDF, G, $y_p = G^{-1}(p)$ that is the graph $(F^{-1}(p), G^{-1}(p))$ for $p \in (0, 1)$.

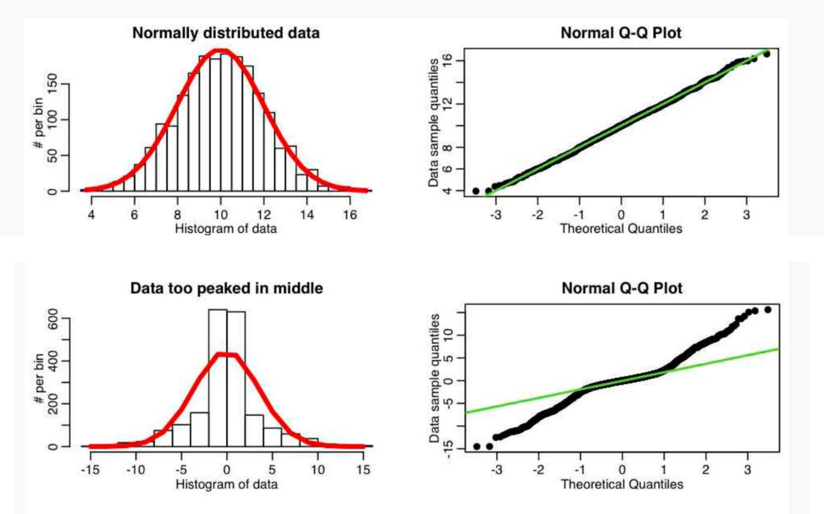
Property of Q-Q plots

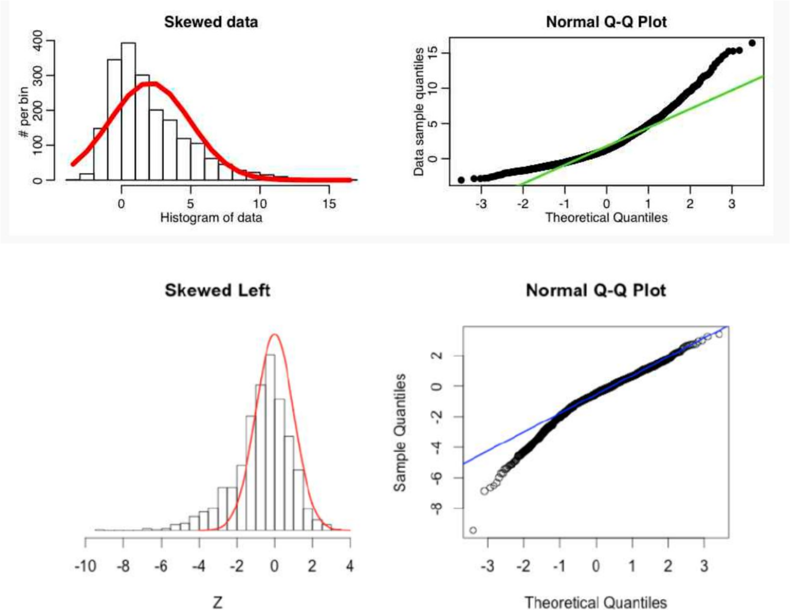
If $G(x) = F(\frac{x-\mu}{\sigma})$ for some constants μ and $\sigma \neq 0$ then

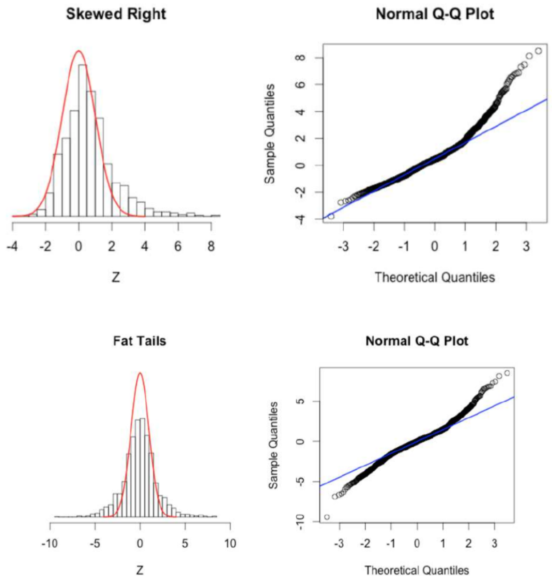
$$y_p = \mu + \sigma x_p$$

Example: Let $F \sim N(0,1)$ then $G(x) = F(\frac{x-1}{\sqrt{2}}) \sim N(1,2)$.
Generate a standard normal distribution from -10 to 10. We compare with $N(0,1)$, we get




Normally distributed data
per bin
0
50
100
150
4
6
8
10
12
14
16
Histogram of data
Normal Q-Q Plot
Data sample quantiles
4
6
8
12
16
-3
-2
-1
0
1
2
3
Theoretical Quantiles
Data too peaked in middle
per bin
0
200
400
600
-15
-10
-5
0
5
10
15
Histogram of data
Normal Q-Q Plot
Data sample quantiles
-15
-5
0
5
10
-3
-2
-1
0
1
2
3
Theoretical Quantiles


Skewed data
per bin
0
100
200
300
400
0
5
10
15
Histogram of data
Normal Q-Q Plot
Data sample quantiles
0
5
10
15
-3
-2
-1
0
1
2
3
Theoretical Quantiles
Skewed Left
-10
-8
-6
-4
-2
0
2
4
Z
Normal Q-Q Plot
Sample Quantiles
2
0
-2
-4
-6
-8
-3
-2
-1
0
1
2
3
Theoretical Quantiles


Skewed Right
-4
-2
0
2
4
6
8
Z
Normal Q-Q Plot
Sample Quantiles
-4
-2
0
2
4
6
8
-3
-2
-1
0
1
2
3
Theoretical Quantiles
Fat Tails
-10
-5
0
5
10
Z
Normal Q-Q Plot
Sample Quantiles
-10
-5
0
5
-3
-2
-1
0
1
2
3
Theoretical Quantiles

1.3.2 Mixture Distributions

Another class of heavy-tailed models is the set of mixture distributions. Consider a simple example made up of 90% N(0,1) and 10% N(0, 25), The density function of such a construct can be written as

$$f_{mix}(x) = 0.9f_{N(0,1)}(x) + 0.1f_{N(0,25)}(x)$$

To generate a random variable Y according to that distribution, we can do that by two-step process

- First, draw from uniform (0.1) random variable U and normal random variable $X \sim N(0,1)$.
- Second, if $U < 0.9$, then $Y = X$. If $U > 0.9$ then $Y = 5X$

Note that this model could be appropriate for a stock that for most of the time shows little variability, but occasionally, e.g., after some earning announcement or other events, make much bigger movements.

The mean of Y is zero and the variance is

$$Var(Y) = 0.9 * 1 + 0.1 * 0.25 = 3.4$$

However, the mixture model is different from $N(0, 3.4)$