Empirical Finance: Methods & Applications

Introduction & Basic Concepts

Pasquale Della Corte
Imperial College London

Week 1

Introduction

Why are empirical methods for finance (and economics more in general) important?

"Theorists develop models with testable predictions; empirical researchers document 'puzzles' – stylized facts that fail to fit established theories – and this stimulates the development of new theories. Such a process is part of the normal development of any science. Asset pricing, like the rest of economics, faces the special challenge that data are generated naturally rather than experimentally, and so researchers cannot control the quantity of data or the random shocks that affect the data."

John Y. Campbell (2000). *Asset Pricing at the Millemium*, Journal of Finance.

Introduction

But economics is not an experimental science and we face difficult problems of inference.

"The economic world is extremely complicated. There are millions of people and firms, thousands of prices and industries. One possible way of figuring out economic laws in such a setting is by controlled experiments ... like those done by chemists, physicists, and biologists ... Economists have no such luxury when testing economic laws. They cannot perform the controlled experiments of chemists or biologists because they cannot easily control other important factors. Just like astronomers or meteorologists, they usually have to solely use their observation."

Paul Samuelson & William Nordhaus (1985). *Economics*, McGraw Hill.

Simple (or Discrete) Returns

What is an Asset Return?

The return of an asset is an important concept of modern finance

- It is the benefit that an investor receives for holding an asset,
- It depends on the holding period of the asset (e.g., one day, one month, etc).

To calculate an asset return, we need the price of the asset

- Price are formed in different markets: auction markets, dealer markets, etc.
- We may observe transaction prices, indicative quotes, or effective quotes.

Simple Return: One Period

Let P_t be the price of an asset that pays no dividends at time t

- Suppose you hold the asset for a single period, meaning between times t and $t+1$,
- Start from simple (or discrete) returns as prices are observed at some discrete intervals.

The simple gross return between times t and $t+1$ is calculated as

I could have also written $R_{t,t+1}$ ⟵ $$R_{t+1} = \frac{P_{t+1}}{P_t}$$ ⟶ Why t+1 rather than t? Because the return is known ex-post at t+1

The simple net return (or simple return) between times t and $t+1$ is calculated as

I could have also written $r_{t,t+1}$ ⟵ $$r_{t+1} = \frac{P_{t+1}}{P_t} - 1$$ ⟶ $1 + r = R$

Simple Return: Two Periods

The simple gross return between times t and $t+2$ is calculated as

$$R_{t,t+2} = \frac{P_{t+2}}{P_t} = \frac{P_{t+2}}{P_{t+1}} \times \frac{P_{t+1}}{P_t}$$

one-period gross return $R_{t+1,t+2}$

one-period gross return $R_{t,t+1}$

The simple gross return between times t and $t+2$ can be written (drop the first subscript) as

$$R_{t,t+2} = R_{t+2} \times R_{t+1}$$

Simple Return: Two Periods

The simple net return between times t and $t+2$ is then obtained as

$$\begin{aligned} r_{t,t+2} &= (R_{t+2} \times R_{t+1}) - 1 \\ &= (1 + r_{t+2}) \times (1 + r_{t+1}) - 1 \end{aligned}$$

rewrite $R = 1 + r$

Recall that for ε and ζ near to zero ('small')

$$(1+\varepsilon)(1+\zeta) \approx 1 + \varepsilon + \zeta \longrightarrow \text{we drop the term } \zeta\varepsilon$$

If returns are small, the simple net return between times t and $t+2$ can be written as

$$r_{t,t+2} \approx r_{t+2} + r_{t+1}$$

Simple Return: Multi Periods

The simple gross return between times t and $t+k$ is calculated as

$$R_{t,t+k} = \frac{P_{t+k}}{P_t} = \underbrace{\frac{P_{t+k}}{P_{t+k-1}}}_{\text{one-period gross return } R_{t+k-1,t+k}} \times \underbrace{\frac{P_{t+k-1}}{P_{t+k-2}}}_{\text{one-period gross return } R_{t+k-1,t+k-2}} \times \ldots \times \underbrace{\frac{P_{t+1}}{P_t}}_{\text{one-period gross return } R_{t+1}}$$

The simple gross return between times t and $t+k$ can be written (drop the first subscript) as

$$R_{t,t+k} = R_{t+k} \times R_{t+k-1} \times \ldots \times R_{t+1} = \prod_{j=1}^{k} R_{t+j}$$

Simple Return: Multi Periods

The simple net return between times t and $t+2$ is then obtained as

$$r_{t,t+k} = \left[\prod_{j=1}^{k} R_{t+j}\right] - 1$$

rewrite $R = 1 + r$

$$= \left[\prod_{j=1}^{k} (1 + r_{t+j})\right] - 1$$

If returns are small, the simple net return between times t and $t+k$ can be written as

$$r_{t,t+k} \approx \sum_{j=1}^{k} r_{t+j}$$

Approximately, the multi-period discrete return is the sum of one-period discrete returns.

Average Simple Returns

What is the average return if you hold an asset over multiple periods?

- For example, you want to compare assets held over different periods.

The average simple gross return can be computed as follows

$$\prod_{j=1}^{k} \overline{R} = \prod_{j=1}^{k} R_{t+j}$$

one-period constant return

one-period actual return

The left-hand side can be simplified so that

$$\overline{R}^k = \prod_{j=1}^{k} R_{t+j}$$

$\overline{R}$ and R_t should deliver an identical payoff over the same holding period.

Average Simple Returns

The average simple gross return is thus given by

$$\overline{R} = \left[\prod_{j=1}^{k} R_{t+j}\right]^{1/k} \longrightarrow \text{Geometric Average}$$

The average simple net return follows as

$$\overline{r} = \left[\prod_{j=1}^{k} R_{t+j}\right]^{1/k} - 1$$

If returns are small, the average simple net return is approximately

$$\overline{r} \approx \frac{1}{k}\sum_{j=1}^{k} r_{t+j} \longrightarrow \text{Arithmetic Average}$$

Simple Returns: Dividend Payment

If there is a dividend payment, the one-period simple gross return is calculated as follows

$$R_{t+1} = \frac{P_{t+1} + \underbrace{D_{t+1}}_{\text{Dividend Payment}}}{P_t}$$

We can also decompose the gross return as follows

$$R_{t+1} = \underbrace{\frac{P_{t+1}}{P_t}}_{\text{Capital Gain}} + \underbrace{\frac{D_{t+1}}{P_t}}_{\text{Dividend Yield}}$$

In practice, dealing with dividend payments is quite complicated as they are paid infrequently. We typically rely on the data provider to make available prices adjusted for dividends.

Simple Returns: Dividend Payment

Nikkei 225 (^N225)
Osaka - Osaka Delayed Price. Currency in JPY

Follow

33,377.42 +89.12 (+0.27%)
At close: January 5 03:15PM JST

Summary · Chart · Conversations · **Historical Data** · Options · Components

Time Period: Dec 28, 2023 - Jan 08, 2024 · Show: Historical Prices · Frequency: Daily · Apply

Currency in JPY · Download

Date	Open	High	Low	Close*	Adj Close**	Volume
Jan 05, 2024	33,397.52	33,568.04	33,257.43	33,377.42	33,377.42	115,300,000
Jan 04, 2024	33,193.05	33,299.39	32,693.18	33,288.29	33,288.29	117,300,000
Dec 29, 2023	33,458.64	33,652.71	33,305.17	33,464.17	33,464.17	89,000,000
Dec 28, 2023	33,477.47	33,571.73	33,411.24	33,539.62	33,539.62	73,300,000

*Close price adjusted for splits. **Adjusted close price adjusted for splits and dividend and/or capital gain distributions.

Simple Returns in Real Terms

In some cases, we need the performance of an asset after accounting for inflation

- For example, there are different inflation regimes when using long-span data.

The real simple gross return is calculated as follows

$$R_{t+1}^r = \frac{P_{t+1} \,/\, CPI_{t+1}}{P_t \,/\, CPI_t}$$

or

$$R_{t+1}^r = \frac{P_{t+1}}{P_t} \times \frac{CPI_t}{CPI_{t+1}}$$

Consumer Price Index at time t

Consumer Price Index at time t+1

Simple Returns in Real Terms

Let's dig a bit more into it as follows

$$R^r_{t+1} = \frac{P_{t+1}}{P_t} \times \frac{CPI_t}{CPI_{t+1}}$$

$(1+\pi_{t+1})^{-1}$

Inverse of gross inflation rate

Let π_{t+1} be the inflation rate between t and $t+1$ and rewrite as

$$r^r_{t+1} = \frac{1+r_{t+1}}{1+\pi_{t+1}} - 1 = \frac{r_{t+1}-\pi_{t+1}}{1+\pi_{t+1}}$$

If π_{t+1} is reasonable small, we have that

$$r^r_{t+1} \approx r_{t+1} - \pi_{t+1}$$

Simple Riskless Return

The riskless return is the return on an investment with no risk

- It is a theoretical quantity that does not exists,
- We often use short-term government bond rate with almost no risk of default as a proxy.

Which proxy then?

- The US Treasury Bill (or simply T-Bill) rate is an example of short-term riskless rate
- The German Federal Treasury rate is another example of short-term riskless rate.
- One-month and three-month maturity instruments are generally used.

Simple Riskless Return

Let B_t be the price of a short-term government bond at time t

- A zero-coupon bond that expires after one period (e.g., one month),
- The liquidation price at time $t+1$ is known at time t.

The simple gross riskless rate between times t and $t+1$ is calculated as

I could have also written $R^f_{t,t+1}$ ← $R^f_t = \frac{B_{t+1}}{B_t}$ → Why t rather than $t+1$? Because the return is known ex-ante at t

The simple net riskless rate between times t and $t+1$ is calculated as

I could have also written $r^f_{t,t+1}$ ← $r^f_t = \frac{B_{t+1}}{B_t} - 1$ → $1 + r^f = R^f$

Simple Excess Return

The simple excess return between times t and $t+1$ is calculated as

$$R^e_{t+1} = R_{t+1} - R^f_t$$

$\longrightarrow$ R_{t+1} and R^f_t are both defined over the same investment horizon

or alternatively as

$$r^e_{t+1} = r_{t+1} - r^f_t$$

$\longrightarrow$ $r^e_{t+1} = R^e_{t+1}$

The excess return is always a net return, no matter if you use gross or net returns.

Simple Portfolio Return

Consider n assets and hold them for a single period ($i = 1, 2, \ldots, n$)

- $P_{i,t}$ is the price of the asset i,
- Q_i is the number of the asset i.

The value of the portfolio at time t is

$$V_t = Q_1 P_{1,t} + Q_2 P_{2,t} + \ldots + Q_n P_{n,t}$$

The portfolio weight of asset i at time t is

$$w_{i,t} = \frac{Q_i P_{i,t}}{V_t}$$

Simple Portfolio Return

The value of the portfolio at time $t+1$ is

$$V_{t+1} = Q_1 P_{1,t+1} + Q_2 P_{2,t+1} + \ldots + Q_n P_{n,t+1}$$

Replace

$$Q_i = V_t \frac{w_{i,t}}{P_{i,t}}$$

Rewrite as

$$V_{t+1} = \frac{w_{1,t} V_t P_{1,t+1}}{P_{1,t}} + \frac{w_{2,t} V_t P_{2,t+1}}{P_{2,t}} + \ldots + \frac{w_{n,t} V_t P_{n,t+1}}{P_{n,t}}$$

$$= V_t \left(w_{1,t} R_{1,t+1} + w_{2,t} R_{2,t+1} + \ldots + w_{n,t} R_{n,t+1} \right)$$

Simple Portfolio Return

The simple portfolio return at time $t+1$ is given by

$$R^p_{t+1} = \frac{V_{t+1}}{V_t}$$
$$= w_{1,t}R_{1,t+1} + w_{2,t}R_{2,t+1} + \ldots + w_{n,t}R_{n,t+1}$$

Simple returns have the property of portfolio additivity since

$$R^p_{t+1} = \sum_{i=1}^{n} w_{i,t}R_{i,t+1}$$

Log (or Continuously Compounded) Returns

Log Returns: One Period

Simple returns are calculated over a given interval

- Suppose you have an investment horizon of a single day,
- Buy the asset at time t (today) and sell it at time $t+1$ (tomorrow).

You realize the interest rate after a day and calculate it as

$$\frac{P_{t+1}}{P_t} = (1 + r_{t+1})$$

Suppose the interest rate is now paid over m sub-interval m (e.g., every hour)

$$\frac{P_{t+1}}{P_t} = \left(1 + \frac{r_{t+1}}{m}\right)^m$$

Log Returns: One Period

If the interest rate is cumulated continuously $(m \to \infty)$, we have

$$\frac{P_{t+1}}{P_t} = \lim_{m\to\infty}\left(1 + \frac{r_{t+1}}{m}\right)^m$$

$$= e^{r_{t+1}}$$

The one-period log return between times t and $t+1$ is thus given by

$$r_{t+1} = ln(P_{t+1}) - ln(P_t)$$

or alternatively from gross discrete returns

$$r_{t+1} = ln(R_{t+1})$$

Warning: to avoid additional notation, I have used the same notation to indicate both net gross and log returns, but they are not identical.

Log Returns: Two Periods

The log return between times t and $t+2$ is calculated as

$$r_{t,t+2} = ln\left(\frac{P_{t+2}}{P_t}\right) = ln\left(\underbrace{\frac{P_{t+2}}{P_{t+1}}}_{\text{one-period gross return } R_{t+2}} \times \underbrace{\frac{P_{t+1}}{P_t}}_{\text{one-period gross return } R_{t+1}}\right)$$

The log return between times t and $t+2$ can be then written as

$$r_{t,t+2} = ln(R_{t+2}) + ln(R_{t+1})$$

the sum of one-period gross returns in log terms.

Log Returns: Multi Periods

The log return between times t and $t+k$ is calculated as

$$r_{t,t+k} = ln\left(\frac{P_{t+k}}{P_t}\right) = ln\left(\frac{P_{t+k}}{P_{t+k-1}} \times \frac{P_{t+k-1}}{P_{t+k-2}} \times \ldots \times \frac{P_{t+1}}{P_t}\right)$$

one-period gross return R_{t+k}

one-period gross return R_{t+k-1}

one-period gros return $R_{t,t+1}$

The log return between times t and $t+k$ can be then written as

$$r_{t,t+k} = \sum_{i=1}^{k} ln(R_{t+i})$$

the sum of one-period gross returns in log terms.

Average Log Returns

The average log return between times t and $t+k$ can be then written as

$$\bar{r} = \frac{1}{k}\sum_{i=1}^{k} ln(R_{t+i})$$

the average of one-period gross returns in log terms being them time additive.

Log Returns: Other Definitions

The one-period log return between times t and $t+1$ with a dividend payment is

$$r_{t+1} = ln\left(\frac{P_{t+1} + D_{t+1}}{P_t}\right)$$

$$= ln(P_{t+1} + D_{t+1}) - ln(P_t)$$

The one-period log return between times t and $t+1$ in real terms

$$r^r_{t+1} = ln\left(\frac{P_{t+1}}{P_t} \times \frac{CPI_t}{CPI_{t+1}}\right)$$

$$= ln\left(\frac{P_{t+1}}{P_t}\right) - ln\left(\frac{CPI_{t+1}}{CPI_t}\right)$$

$$= r_{t+1} - \pi_{t+1}$$

Log Returns: Other Definitions

The one-period log excess return between times t and $t+1$ is

$$r^e_{t+1} = ln(R_{t+1}) - ln(R^f_t) \longrightarrow$$

Different from
$ln(R_{t+1} - R^f_t)$

or alternatively as

$$r^e_{t+1} = r_{t+1} - r^f_t \longrightarrow$$

This is the log
riskless return

The portfolio log excess return between times t and $t+1$ is calculated as

$$r^p_{t+1} = ln(R^p_{t+1}) = ln\left(\sum_{i=1}^n w_{i,t} R_{i,t+1}\right) \neq \left(\sum_{i=1}^n w_i r_{i,t+1}\right)$$

Log of weighted average
simple gross returns

weighted average
of log returns

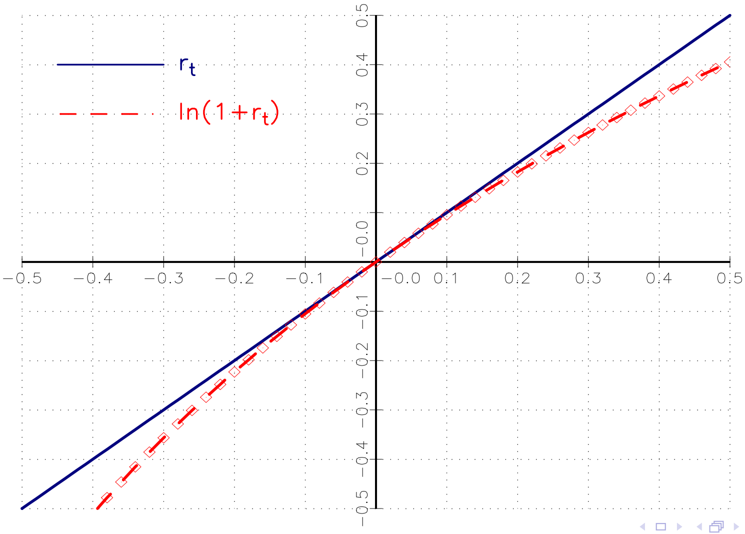
Log Returns: Other Definitions

When returns are small ('measured over short time intervals')

$$r_{t+1}^{p} \approx \left(\sum_{i=1}^{n} w_i r_{i,t+1} \right)$$

When do we use log returns? For time-series analysis to limit the impact of large outliers.

Simple vs Log Returns



Stylized Facts

Stylized Facts

Empirical studies have identified several stylized facts among financial data

1. Fat Tails

   The unconditional distribution of returns displays fatter tails than the normal distribution. By using a normal distribution to model financial returns, we underestimate both the number and the magnitude of crashes and booms.

2. Asymmetry

   The unconditional distribution is negatively skewed suggesting that extreme negative returns are more frequent than extreme positive returns. The asymmetry and the fat-tails persists even after adjusting for conditional heteroskedasticity. This means that the conditional distribution is also non-normal.

Stylized Facts

3. Aggregate Normality
   As the frequency of the returns diminishes, the distribution of financial returns approaches the normal distribution.
4. Absence of Serial Correlation
   Financial returns display little or no serial correlations, except at high frequency.
5. Volatility Clustering
   Volatility of returns is serially correlated, suggesting that a large positive (negative) return tends to be followed by another large positive (negative) return.
6. Time-varying Correlations
   Correlations between asset returns tends to increase during high-volatility periods, in particular during crashes.

Moments

Let X be a continuous random variable

- The probability density function (pdf) of X is a nonnegative function $f(x)$ such that for any interval $[a, b]$

$$Pr(a \leq X \leq b) = \int_a^b f(x)dx$$

  where $Pr(\cdot)$ is the *area under the probability curve over the interval* $[a, b]$

- The pdf $f(x)$ must satisfy

$$\int_{-\infty}^{\infty} f(x)dx = 1$$

Moments

The cumulative density function (cdf) of X is defined as

$$F(x) = Pr(X \leq x) = \int_{-\infty}^{x} f(u)du$$

and satisfies the following properties

1. the cdf is non-negative: $F(x) \geq 0$,
2. the cdf goes to zero on the far left: $lim_{x \to -\infty} F(x) = 0$,
3. the cdf goes to one on the far right: $lim_{x \to \infty} F(x) = 1$,

Moments

4. the *cdf* is non-decreasing: $F(b) \geq F(a)$ if $b \geq a$. If $b \geq a$, then $X \leq a$ is a sub-set of $X \leq b$, and sub-sets never have higher probabilities,
5. $Pr(X \geq x) = 1 - F(x)$,
6. $Pr(x_1 \leq X \leq x_2) = F(x_2) - F(x_1)$,
7. $\frac{d}{dx}F(x) = f(x)$.

Moments

The k^{th} uncentered moment of X is defined as

$$m_k = E\left[X^k\right] = \int_{-\infty}^{\infty} x^k f(x) dx$$

where k is an integer.

The k^{th} centered moment of X is defined as

$$\mu_k = E\left[(X - m_1)^k\right] = \int_{-\infty}^{\infty} (x - m_1)^k f(x) dx$$

where $m_1 = E\left[X\right]$ is the first uncentered moment.

Moments

Uncentered and centered moments (if they exist) are related via the binomial transformation as

$$\mu_k = \sum_{i=0}^{k} (-1)^{k-i} \binom{k}{i} m_i m_1^{k-i}$$

Moments

To sum up, we have the following

$$\begin{aligned}
\mu_1 &= 0 \\
\mu_2 &= m_2 - m_1^2 \\
\mu_3 &= m_3 - 3m_2 m_1 + 2m_1^3 \\
\mu_4 &= m_4 - 4m_3 m_1 + 6m_2 m_1^2 - 3m_1^4
\end{aligned}$$

where

$\mu_1 \longrightarrow$ zero by construction

$\mu_2 \longrightarrow$ this is the variance

$\mu_3 \longrightarrow$ non-standardized skewness

$\mu_4 \longrightarrow$ non-standardized kurtosis

Moments

The mean is the first non-centered moment

$$\mu = E[X] = m_1$$

and an indicator of central tendency.

The variance is the second centered moment

$$\sigma^2 = V[X] = m_2 - m_1^2$$

and an indicator of dispersion around μ.

Moments

The skewness is the standardized third centered moment

$$sk = SK[X] = E\left[\left(\frac{X-\mu}{\sigma}\right)^3\right]$$

and an indicator of symmetry.

The kurtosis is the standardized fourth centered moment

$$ku = KU[X] = E\left[\left(\frac{X-\mu}{\sigma}\right)^4\right]$$

and an indicator of tail heaviness.

Empirical Moments

How do practically estimate the moments?

- Let $\{x_1, \ldots, x_T\}$ be a random sample from X with T observations.

The sample mean is

$$\widehat{\mu} = \frac{1}{T} \sum_{i=1}^{T} x_t,$$

The sample variance is

$$\widehat{\sigma}^2 = \frac{1}{T-1} \sum_{i=1}^{T} (x_t - \widehat{\mu})^2$$

Empirical Moments

The sample skewness is

$$\widehat{s} = \frac{1}{T-1}\sum_{i=1}^{T}\frac{(x_t-\widehat{\mu})^3}{\widehat{\sigma}^3}$$

The sample kurtosis is

$$\widehat{\kappa} = \frac{1}{T-1}\sum_{i=1}^{T}\frac{(x_t-\widehat{\mu})^4}{\widehat{\sigma}^4}$$

Empirical Moments

Kendall & Stuart (1977) derive the following asymptotic distributions under normality

$$
\begin{aligned}
\sqrt{T}(\widehat{\mu} - \mu) &\overset{a}{\Rightarrow} N(0, \sigma^2) \\
\sqrt{T}(\widehat{\sigma}^2 - \sigma^2) &\overset{a}{\Rightarrow} N(0, 2\sigma^4) \\
\sqrt{T}\widehat{s} &\overset{a}{\Rightarrow} N(0, 6) \\
\sqrt{T}(\widehat{\kappa} - 3) &\overset{a}{\Rightarrow} N(0, 24)
\end{aligned}
$$

where $\overset{a}{\Rightarrow}$ denotes asymptotic convergence. Note that since the asymptotic variances are large, sample skewness and sample kurtosis are only informative for large samples.

Testing for Normality

Jarque & Bera (1987) show that under the null of normality ($s = 0$ and $\kappa - 3 = 0$)

$$JB = \frac{\widehat{s}^2}{6/T} + \frac{(\widehat{\kappa} - 3)^2}{24/T} \overset{a}{\backsim} \chi^2_{(2)}$$

where $\chi^2_{(2)}$ is a chi-square with two degree of freedom.

This test combines

$$\frac{\widehat{s}}{\sqrt{6/T}} \overset{a}{\backsim} N(0,1)$$

$$\frac{\widehat{\kappa} - 3}{\sqrt{24/T}} \overset{a}{\backsim} N(0,1)$$

knowing that the sum of two normals delivers a chi-square distribution.

Testing for Normality

The Jarque & Bera (1987) test presents a number of drawbacks

- the asymptotic distribution only holds for large samples
- $\widehat{s}$ and $\widehat{\kappa}$ are computed for given values of mean and variance,
- $\widehat{s}$ and $\widehat{\kappa}$ are both subject to sampling errors.

Empirical Moments

Fat tails and asymmetry are key aspects of financial time series

- Normality is often rejected because of high kurtosis: Extreme values are more likely than what the normal distribution implies (leptokurtic distribution),
- A mismatch between the unconditional mean and mode: A tail is longer than the other (skewed distribution),
- Negative skewness: the mean is at the left of the mode or the left tail is longer (left-skewed distribution),
- Positive skewness: the mean is at the right of the mode or the right tail is longer (right-skewed distribution).

Fat Tails

The Student's t-distribution is widely used in finance to accommodate excess kurtosis

- Y is a random variable from a t-distribution with location parameter 0, scale parameter 1, and the degree of freedom ν

$$Y \sim t\left(0, \frac{\nu}{\nu-2}, \nu\right)$$

The density of Y is given by

$$f(y|\nu) = \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})(\pi\nu)^{1/2}} \left(1 + \frac{y^2}{\nu}\right)^{-\frac{\nu+1}{2}}$$

The distribution has fat tails when ν is small

- we have a Cauchy distribution when $\nu = 1$,
- we approach a Normal distribution when $\nu > 30$.

Fat Tails

The key moments of a t-distribution

$$E(Y) = 0 \qquad \text{for } \nu > 1$$

$$V(Y) = \frac{\nu}{\nu - 2} \qquad \text{for } \nu > 2$$

$$Sk(Y) = 0 \qquad \text{for } \nu > 3$$

$$Ku(Y) = 3 + \frac{6}{\nu - 4} \qquad \text{for } \nu > 4$$

Fat Tails

How do we sample from a t-distribution?

$$x \sim N(0,1)$$

$$s \sim \chi^2_{(\nu)}$$

$$y = \frac{x}{\sqrt{\frac{s}{\nu}}}$$

$$y \sim t\left(0, \frac{\nu}{\nu-2}, \nu\right)$$

For empirical applications, it is convenient to redefine the t-distribution so to have unit variance.

Standardized Distributions

Transformation of Random Variables

Let X be a random variable with a given density

$$f_X(x)$$

Define a new random variable Y

$$Y = g(X)$$

The density of Y is obtained via the density function method

$$f_Y(y) = \left| \frac{d}{dy} g^{-1}(y) \right| f_x(g^{-1}(y))$$

Transformation of Random Variables

As an example, let $X \sim N(0,1)$ and define

$$Y = \mu + \sigma X$$

We can now compute the density of Y as follows

$$\begin{aligned} f_Y(y) &= \left| \frac{d}{dy}\left(\frac{y-\mu}{\sigma}\right) \right| f_X\left(\frac{y-\mu}{\sigma}\right) \\ &= \frac{1}{\sigma}\frac{1}{\sqrt{2\pi}} exp\left\{ -\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2 \right\} \end{aligned}$$

Transformation of Random Variables

We can use the same trick to have a t-student with unit variance

- If X has a t-distribution, we aim at Y with a standardized t-distribution.

Let X have the following distribution

$$X \sim t(0, \frac{\nu}{\nu - 2}, \nu)$$

and then define Y as

$$Y = \frac{X}{\sqrt{\frac{\nu}{\nu-2}}}$$

Transformation of Random Variables

The density of Y is then computed as

$$
\begin{aligned}
f_Y(y) &= \left|\frac{d}{dy}\left(y\sqrt{\frac{\nu}{\nu-2}}\right)\right| f_X\left(y\sqrt{\frac{\nu}{\nu-2}}\right) \\
&= \sqrt{\frac{\nu}{\nu-2}}\,\frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})(\pi\nu)^{1/2}}\left(1+\left(y\sqrt{\frac{\nu}{\nu-2}}\right)^2\frac{1}{\nu}\right)^{-\frac{\nu+1}{2}} \\
&= \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})(\pi(\nu-2))^{1/2}}\left(1+\frac{y^2}{\nu-2}\right)^{-\frac{\nu+1}{2}}
\end{aligned}
$$

$f_Y(y)$ is the density of a t-distribution with zero mean, unit variance, and dof $\nu > 2$.

Fat Tails

Let Y be a random variable from a t-distribution with zero mean, unit variance, and dof ν.

$$Y \sim t(0, 1, \nu)$$

Its density is described by

$$f(y|\nu) = \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})(\pi(\nu-2))^{1/2}} \left(1 + \frac{y^2}{\nu-2}\right)^{-\frac{\nu+1}{2}}$$

with (otherwise undefined)

$$\nu > 2$$

Fat Tails

The key moments are

$$Sk(Y) = 0 \qquad \text{for} \quad \nu > 3$$

$$Ku(Y) = 3\frac{\nu - 2}{\nu - 4} \qquad \text{for} \quad \nu > 4$$

Sampling

$$x \sim t(0, \frac{\nu}{\nu - 2}, \nu)$$

$$y = x / \sqrt{\frac{\nu}{\nu - 2}}$$

$$y \sim t(0, 1, \nu)$$

Fat Tails and Asymmetry

To accommodate both excess kurtosis and skewness, Hansen (1994) proposed a skewed Student distribution by combining a Student distribution with the asymmetry coefficient λ

- Although this density is quite easy to implement, he did not discuss the relation between λ and higher moments.

More recently, Fernandéz and Steel (1998) have developed a skewed t-distribution which has the advantage that all the parameters have a clear interpretation

- They introduce skewness in any univariate pdf $g(\cdot)$, which is unimodal and symmetric around 0, by changing the scale at each side of the mode.

Fat Tails and Asymmetry

Start from a symmetric distribution with zero mean and unit variance

$$X \sim g(0,1)$$

Define a Bernoulli random variable with probability of success $\xi^2/(1+\xi^2)$ as

$$W \sim B\left(\xi^2/(1+\xi^2)\right)$$

Consider the following mixture

$$Y = W\xi\,|X| - (1-W)\frac{1}{\xi}\,|X|$$

Fat Tails and Asymmetry

We have a class of skewed unimodal distributions indexed by skewness parameter

$$\xi \in (0, \infty)$$

The density of this new distribution is defined as

$$f(y|\xi) = \frac{2}{\xi + \frac{1}{\xi}} g(y\xi^{I})$$

where

$$I = \begin{cases} -1 & if \quad y \geq 0 \\ 1 & if \quad y < 0 \end{cases}$$

Fat Tails and Asymmetry

What happens?

- When y is positive, we divide y by a positive constant ξ.
- When y is negative, we multiply y by a positive constant ξ,

$f(y|\xi)$ maintains the same mode at 0 as $g(y)$ but loses the asymmetric shape whenever $\xi \neq 1$ such that the probability mass ratio above and below the mode is

$$\frac{\Pr(y \geq 0|\xi)}{\Pr(y < 0|\xi)} = \xi^2$$

Fat Tails and Asymmetry

The parameter ξ determines the direction of the skewness of the distribution $f(y|\xi)$

- Positive skewness corresponds to $\log \xi > 0$,
- Negative skewness corresponds to $\log \xi < 0$.

The Skewed t-distribution

Let $\nu > 2$ and consider

$$X \sim t(0, 1, \nu)$$

Following Fernandéz and Steel (1998), we have

$$Y \sim SKT(m, s^2, \xi, \nu)$$

is a skewed t-distribution with

- mean m,
- variance s^2,
- degree of freedom $\nu > 2$,
- skewness parameter $\xi > 0$.

The Skewed t-distribution

The density if defined as

$$f(y|\xi,\nu) = \frac{2}{\xi + \frac{1}{\xi}} \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})[\pi(\nu-2)]^{1/2}} \left(1 + \frac{(y\xi^{I})^2}{\nu-2}\right)^{-\frac{\nu+1}{2}}$$

where

$$E(y|\xi,\nu) = \frac{\Gamma\left(\frac{\nu-1}{2}\right)\sqrt{\nu-2}}{\Gamma\left(\frac{\nu}{2}\right)\sqrt{\pi}} \left(\xi - \frac{1}{\xi}\right) \equiv m$$

$$V(y|\xi,\nu) = \left(\xi^2 + \frac{1}{\xi^2} - 1\right) - m^2 \equiv s^2$$

The Skewed t-distribution

The Skewed-t distribution must be normalized to have zero mean and unit variance

- For example, to describe the residuals of a regression.

Consider the following transformation

$$Z = \frac{Y - m}{s}$$

Calculate the density using the usual trick

$$f_Z(z) \quad = \quad \left| \frac{d}{dz}(zs + m) \right| f_Y(zs + m)$$

The Skewed t-distribution

We then obtain

$$f_Z(z) = \frac{2s}{\xi + \frac{1}{\xi}} \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})[\pi(\nu-2)]^{1/2}} \left\{ 1 + \frac{[(sz+m)\xi^I]^2}{\nu - 2} \right\}^{-\frac{\nu+1}{2}}$$

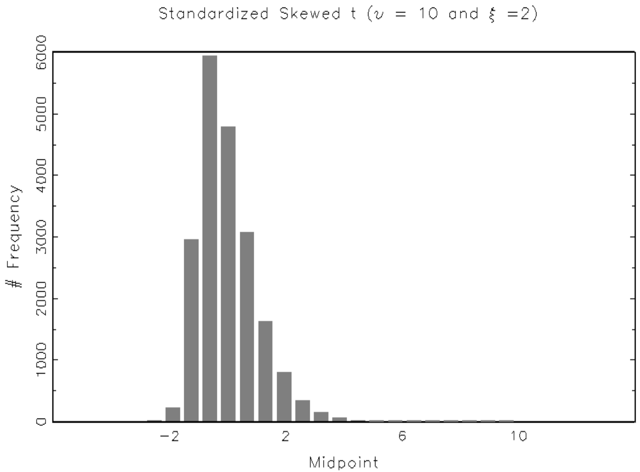
where

$$I = \begin{cases} -1 & \text{if} \quad z \geq -m/s \\ 1 & \text{if} \quad z < m/s \end{cases}$$

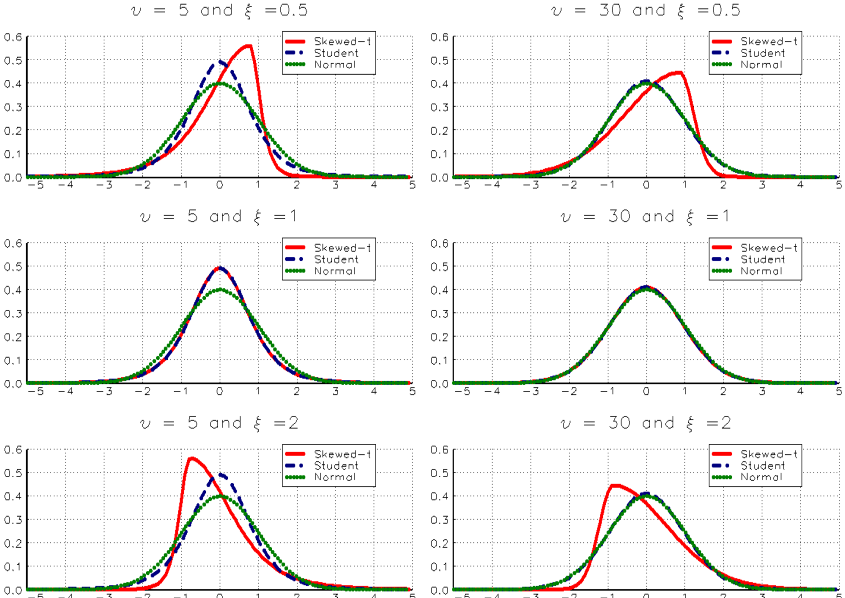
It follows that

$$Z \sim SKT(0, 1, \xi, \nu)$$

The Skewed t-distribution



The Skewed t-distribution



A Recent Application

Adrian, Boyarchenko, and Giannone (2019)

- Examine the conditional distribution of GDP growth as a function of economic and financial conditions,
- Find that deteriorating financial conditions are associated with an increase in the conditional volatility and a decline in the conditional mean of GDP growth,
- Find that the lower quantiles of GDP growth vary with financial conditions and the upper quantiles to be stable over time.

Upside risks to GDP growth are low in most periods while downside risks increase as financial conditions become tighter.

A Recent Application