Lecture 1. Financial Time Series and Their Characteristics

1. Asset Returns

Definition 1.1. Let P_t be the price of an asset at time t. A *simple return* from time $t-1$ to time t is

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

and a corresponding *simple gross return* is $1 + R_t$.

$\frac{P_t}{P_{t-1}} - 1$

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

Definition 1.2. The *continuously compounded return* or simply *log return* is

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

where $p_t = \ln P_t$.

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

- The k-period log return from time $t-k$ to time t, denoted by $r_{t-k\to t}$, is

$$\begin{aligned}
r_{t-k\to t} &= \ln\left(\frac{P_t}{P_{t-k}}\right) \\
&= \ln\left(\frac{P_t}{P_{t-1}} \times \frac{P_{t-1}}{P_{t-2}} \times \cdots \times \frac{P_{t-k+1}}{P_{t-k}}\right) \\
&= \ln(1+R_t) + \ln(1+R_{t-1}) + \cdots + \ln(1+R_{t-k+1}) \\
&= r_t + r_{t-1} + \cdots + r_{t-k+1},
\end{aligned}$$

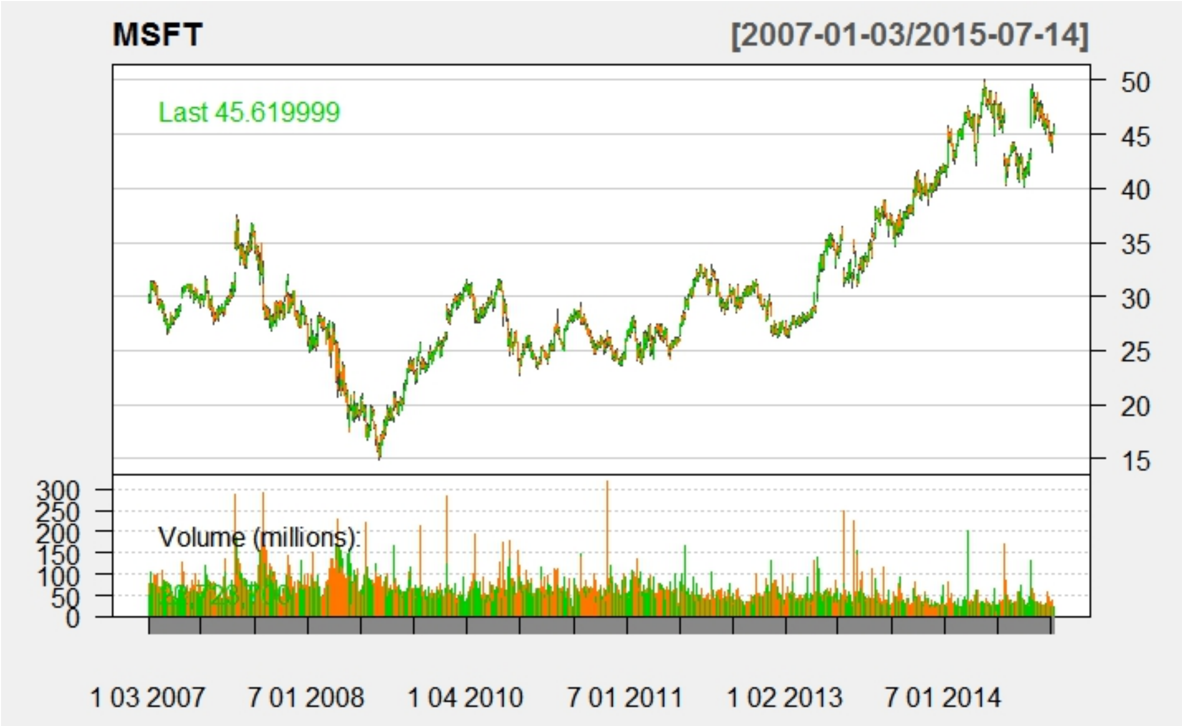
$= \ln P_t - \cancel{\ln P_{t-1}} + \cancel{\ln P_{t-1}} - \cancel{\ln P_{t-2}} + \cdots - \ln P_{t-k}$

which is the sum of continuously compounded one-period returns.

$= \ln P_t - \ln P_{t-k}$
$= p_t - p_{t-k}$

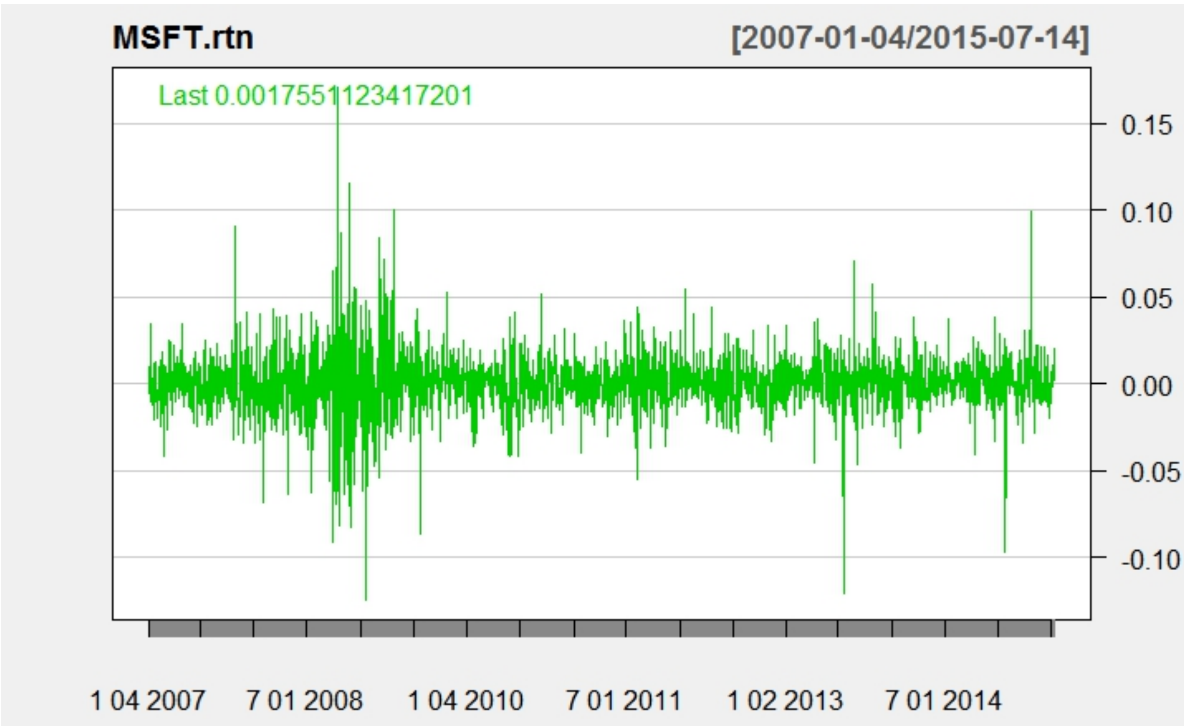
Example 1.3. One uses the `quantmod` package to download financial data from some open sources such as Yahoo Finance, Google Finance, and the Federal Reserve Economic Data (FRED).

```
> library(quantmod)
> getSymbols("MSFT")
[1] "MSFT"
> head(MSFT)
           MSFT.Open MSFT.High MSFT.Low MSFT.Close MSFT.Volume MSFT.Adjusted
2007-01-03     29.91     30.25    29.40      29.86    76935100      24.44541
2007-01-04     29.70     29.97    29.44      29.81    45774500      24.40448
2007-01-05     29.63     29.75    29.45      29.64    44607200      24.26530
2007-01-08     29.65     30.10    29.53      29.93    50220200      24.50272
2007-01-09     30.00     30.18    29.73      29.96    44636600      24.52727
2007-01-10     29.80     29.89    29.43      29.66    55017400      24.28167
> chartSeries(MSFT, theme = "white")
```




```
> MSFT.rtn <- diff(log(MSFT$MSFT.Adjusted))

> chartSeries(MSFT.rtn, theme = "white")
```



2. Reivew of Statistical Distribution and Their Moments

2.1. Probability Distributions

Definition 2.1. The *cumulative distribution function* (cdf) of a random variable X, denoted by $F(x)$, is defined by

$$F(x) = \Pr(X \leq x)$$

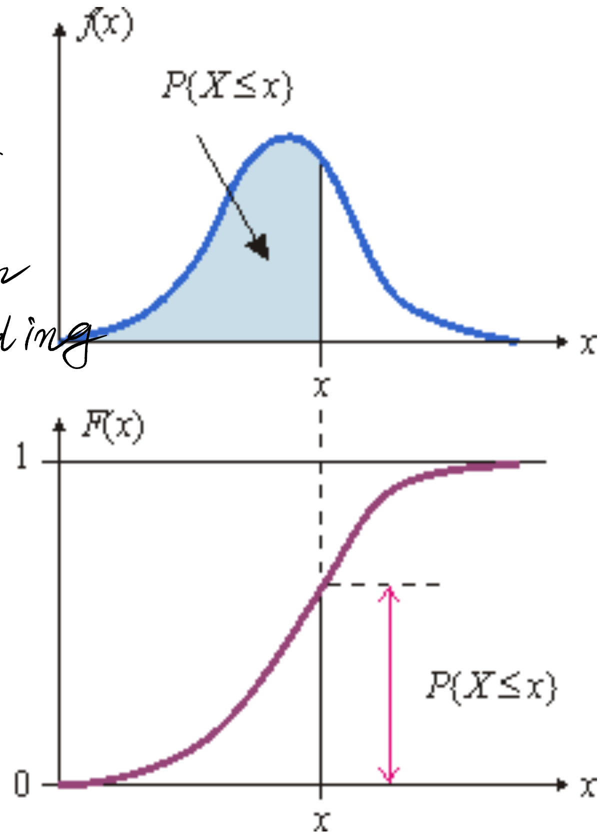
for all x. The *probability density function* (pdf), denoted by $f(x)$, of a continuous random variable X is the function that satisfies

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

for all x.

- The relationship between the pdf and cdf of a continuous random variable X is depicted as:

Same information
1-1 corresponding



Definition 2.2. The *expected value* of a continuous random variable X, denoted by $E[X]$ or μ_X, is

$$E[X] = \int_{-\infty}^{\infty} x f(x)dx.$$

The *variance* of X, denoted by either $Var[X]$ or σ_X^2, is defined as

$$Var[X] = E[(X - \mu_X)^2] = \int_{-\infty}^{\infty} (x - \mu_X)^2 f(x)dx.$$

Definition 2.3. A function $f(x,y)$ from $\mathscr{R}^2$ into $\mathscr{R}$ is called a *joint probability density* (or joint pdf) of the continuous bivariate random vector (X,Y) if, for every $A \subset \mathscr{R}^2$,

$$\Pr((X,Y) \in A) = \iint_A f(x,y)dxdy,$$

where the notation $\iint_A$ means that the limits of integration are set so that the function is integrated over all $(x,y) \in A$.

- Given the joint pdf, the *marginal pdfs* of X and Y are computed as

$$\begin{aligned} f(x) &= \int_{-\infty}^{\infty} f(x,y)dy \\ f(y) &= \int_{-\infty}^{\infty} f(x,y)dx. \end{aligned}$$

Definition 2.4. Let (X,Y) be a continuous random vector with joint pdf $f(x,y)$ and marginal pdfs $f(x)$ and $f(y)$. For any x such that $f(x) > 0$, the *conditional pdf of Y given that X* $= x$, denoted by $f(y|x)$, is defined by

$$f(y|x) = \frac{f(x,y)}{f(x)}. \tag{2.1}$$

- From (2.1), the relation among joint, marginal, and conditional distributions is

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

Definition 2.5. Let (X,Y) be a bivariate random vector with joint pdf $f(x,y)$ and marginal pdfs $f(x)$ and $f(y)$. Then X and Y are called *independent* random variable if, for every x and y,

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

- When X and Y are independent, the conditional pdf of Y given $X = x$ is

$$f(y|x) = \frac{f(x,y)}{f(x)} = f(y),$$

  meaning that the knowledge that $X = x$ does not give any more information about Y than what one already had.

Definition 2.6. The *covariance* of X and Y, denoted by either $Cov[X,Y]$ or σ_{XY}, is the number defined by

$$\begin{aligned} \sigma_{XY} &= E[(X - E[X])(Y - E[Y])] \\ &= E[XY] - \mu_X \mu_Y \end{aligned}$$

and the *correlation* of X and Y, denoted by ρ_{XY}, is the number defined by

$$\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}.$$

- The sign of $Cov[X,Y]$ gives information regarding the direction of of covariation of X and Y; specifically, positive values indicate that X and Y tend to move together, while negative values imply that X tends to move in the opposite direction to Y. If X and Y are independent random variables, then $Cov[X,Y] = 0$ and $\rho_{XY} = 0$.

Theorem 2.7. *For any random variables X and Y, it shows that*

1. $-1 \leq \rho_{XY} \leq 1$,
2. $|\rho_{XY}| = 1$ *if and only if there exists numbers* $a \neq 0$ *and* b *such that* $\Pr(Y = aX + b) = 1$. *If* $\rho_{XY} = 1$, *then* $a > 0$, *and if* $\rho_{XY} = -1$, *then* $a < 0$.

2.2. Chi-squared and Normal Distributions

Definition 2.8. The *gamma function* is defined as

$$\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t} dt$$

for $\alpha > 0$.

Definition 2.9. The pdf of the *gamma(α, β) distribution* is defined as

$$f(x) = \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-x/\beta}$$

for $0 < x < \infty$, $\alpha > 0$, and $\beta > 0$.

- The shape parameter α influences the peakedness of the distribution. The scale parameter β influences the spread of the distribution. For $X \sim gamma(\alpha, \beta)$ it shows that

$$E[X] = \alpha\beta$$
$$Var[X] = \alpha\beta^2.$$

$\chi^2(p)$
$=$
$Gamma(\frac{p}{2}, 2)$

Definition 2.10. The pdf of the *chi-squared distribution* with p degrees of freedom, denoted by $\chi^2_{(p)}$, is defined as

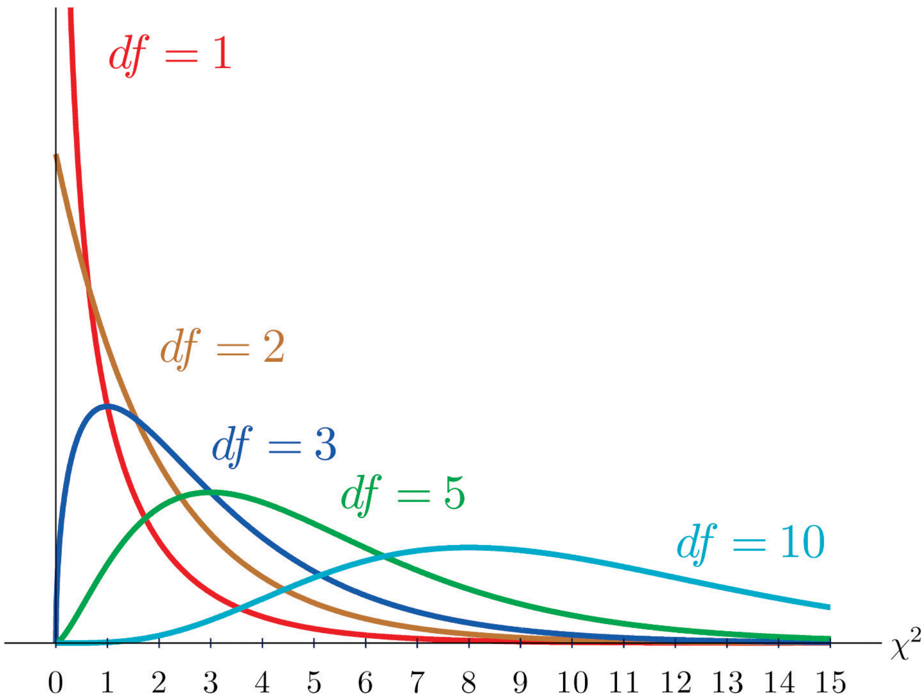
$$f(x) = \frac{1}{\Gamma(p/2)2^{p/2}} x^{(p/2)-1} e^{-x/2}$$

for $0 < x < \infty$.

- The chi-squared distribution is a special case of the gamma distribution when $\alpha = p/2$ and $\beta = 2$. So, its mean and variance is calculated using the gamma formulas: i.e., for $X \sim \chi^2_{(p)}$, one computes that

$$E[X] = \frac{p}{2} \times 2 = p$$
$$Var[X] = \frac{p}{2} \times 4 = 2p.$$

Example 2.11. Chi-squared distributions of various degrees of freedom



Definition 2.12. The pdf of the *normal distribution* with mean μ and variance σ^2, denoted by $N(\mu, \sigma^2)$, is given by

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

for $-\infty < x < \infty$.

Remark 2.13. The normal distribution (sometimes called the *Gaussian distribution*) plays a central role in statistics for several reasons. One, it has a tractable analytical form. Second, the normal distribution has the familiar bell shape, whose symmetry makes it an appealing choice for many applications. Last, the *central limit theorem* enables the normal distribution to be used to approximate a large variety of distributions in large samples.

Definition 2.14. For $Z \sim N(0,1)$, it is called that a random variable Z has the *standard normal distribution*.

- If $X \sim N(\mu, \sigma^2)$, one obtains
$$Z = \frac{X-\mu}{\sigma} \sim N(0,1).$$

2.3. Properties of a Random Sample

Definition 2.15. A sample of n observations $X_1, \ldots, X_n$ is called a *random sample* from the population $f(x)$ if $X_1, \ldots, X_n$ are mutually independent random variables and the marginal pdf of each X_i is the same function $f(x)$. Alternatively, $X_1, \ldots, X_n$ are called *independent and identically distributed*, or often abbreviated to *iid*, random variables with pdf $f(x)$.

Definition 2.16. Let $X_1, \ldots, X_n$ be a random sample of size n from a population and let $T(x_1, \ldots, x_n)$ be a real-valued function whose domain includes the sample space of $(X_1, \ldots, X_n)$. Then the random variable $Y = T(X_1, \ldots, X_n)$ is called a *statistic*. The probability distribution of a statistic Y is called the *sampling distribution of Y*.

Theorem 2.17. *Let $X_1, \ldots, X_n$ be a random sample from a population with mean μ and variance σ^2. Let $\bar{X} = (1/n)\sum_{i=1}^n X_i$ and $s^2 = [1/(n-1)]\sum_{i=1}^n (X_i - \bar{X})^2$. Then the followings are true:*

1. $E[\bar{X}] = \mu$,
2. $Var[\bar{X}] = \frac{\sigma^2}{n}$,
3. $E[s^2] = \sigma^2$.

Theorem 2.18. *Let $X_1, \ldots, X_n$ be a random sample from a $N(\mu, \sigma^2)$ distribution, and let $\bar{X} = (1/n)\sum_{i=1}^n X_i$ and $s^2 = [1/(n-1)]\sum_{i=1}^n (X_i - \bar{X})^2$. Then the followings are true:*

1. $\bar{X}$ *and* s^2 *are independent random variables,*
2. $\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$,
3. $(n-1)\frac{s^2}{\sigma^2} \sim \chi^2_{(n-1)}$.

$$T = \frac{Z}{\sqrt{U/r}} \sim t(r)$$

where $Z \sim N(0,1)$, $U \sim \chi^2(r)$

- Since $\bar{X} \sim N(\mu, \sigma^2/n)$, it shows

$$\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1).$$

Definition 2.19. Let $X_1, \ldots, X_n$ be a random sample from a $N(\mu, \sigma^2)$ distribution. The quantity

$$\frac{\bar{X} - \mu}{s/\sqrt{n}} \sim t(n-1)$$

has *Student's t distribution with $n-1$ degrees of freedom.*

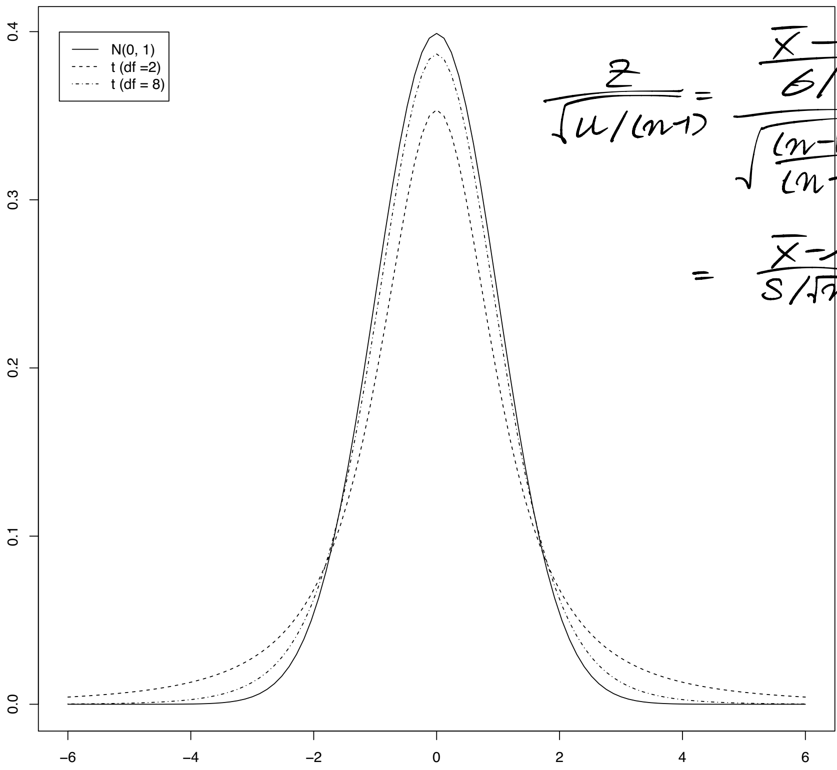
$X_i \sim N(\mu, \sigma^2)$

$\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$

$Z := \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$

$U := \frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$

Example 2.20. t-distributions with different degrees of freedom



$$\frac{Z}{\sqrt{U/(n-1)}} = \frac{\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}}{\sqrt{\frac{(n-1)S^2}{(n-1)\sigma^2}}} = \frac{\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}}{S/\sigma}$$

$$= \frac{\bar{X}-\mu}{S/\sqrt{n}} \sim t(n-1)$$

- Both t-distributions have fatter tails, but the tails of the t-distribution with eight degrees of freedom more closely resemble the normal distribution's tails. In general, as the degrees of freedom increase, the tails of the t-distribution becomes less fat.

2.4. Asymptotic Distribution Theory

Definition 2.21. Let $\{c_n\}_{n=1}^{\infty}$ be a sequence of deterministic numbers. Then, $\{c_n\}_{n=1}^{\infty}$ is said to *converge* to c if for any $\varepsilon > 0$, there exists a m such that $|c_n - c| < \varepsilon$ whenever $n \geq m$; in words, c_n will be close to c so long as n is sufficiently large.

Definition 2.22. A sequence of random variables $X_1, X_2, \ldots,$ *converges in probability* to a random variable X if, for every $\varepsilon > 0$,

$$P(|X_i - \mu| > k\sigma) < \frac{1}{k^2}$$

$$\lim_{n\to\infty} \Pr(|X_n - X| \geq \varepsilon) = 0 \qquad P(|X_i - \mu| > \varepsilon) < \frac{\sigma^2}{\varepsilon^2}$$

or equivalently,

$$P((X_i - \mu)^2 > \varepsilon^2) < \frac{\sigma^2}{\varepsilon^2}$$

$$\lim_{n\to\infty} \Pr(|X_n - X| < \varepsilon) = 1. \qquad P(g(X) > r) < \frac{E[g(X)]}{r}$$

- In words, when n is sufficiently large, the probability that X_n differs from X by more than ε can be arbitrarily small, or equally speaking, the probability that X_n is close to X can be very large. This is indicated as

$$X_n \xrightarrow{p} X.$$

Theorem 2.23. *(Chebychev's Inequality) Let X be a random variable and let $g(x)$ be a nonnegative function. Then, for any $r > 0$,*

$$\Pr(g(X) \geq r) \leq \frac{E[g(X)]}{r}.$$

Theorem 2.24. *(Weak Law of Large Numbers or WLLN) Let $X_1, X_2, \ldots$ be iid random variables with $E[X_i] = \mu$ and $Var[X_i] = \sigma^2$. Define $\bar{X}_n = (1/n)\sum_{i=1}^{n} X_i$. Then, for every $\varepsilon > 0$, we have*

$$\lim_{n\to\infty} \Pr(|\bar{X}_n - \mu| < \varepsilon) = 1;$$

that is, $\bar{X}_n$ converges in probability to μ. The WLLN states that, under general conditions, the sample mean approaches the population mean as the sample size increases.

Definition 2.25. Let $\{X_n\}_{n=1}^{\infty}$ be a sequence of random variables and let $F_{X_n}(x)$ be the cdf of X_n. Then, X_n is said to *converge in distribution* to X if there exists a cdf $F_X(x)$ such that

$$\lim_{n\to\infty} F_{X_n}(x) = F_X(x)$$

at any value x at which $F_X(\cdot)$ is continuous.

- In words, when n is sufficiently large, the distribution of X_n is approximately close to that of X. This is indicated as

$$X_n \xrightarrow{D} X.$$

Theorem 2.26. *(Central Limit Theorem or CLT) Let $X_1, X_2, \ldots$ be a sequence of iid random variables. Let $E[X_i] = \mu$ and $Var[X_i] = \sigma^2$. Define $\bar{X}_n = (1/n)\sum_{i=1}^{n} X_i$. Let $G_n(x)$ denote the cdf of $\sqrt{n}(\bar{X}_n - \mu)/\sigma$.*

$$G_n = \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}}$$

Then, for any x, we have

$$\lim_{n\to\infty} G_n(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy;$$

that is, $\sqrt{n}(\bar{X}_n - \mu)/\sigma$ *is a limiting standard normal distribution.*

- Equally speaking, the CLT implies that

$$\frac{\sqrt{n}(\overline{X}_n - \mu)}{\sigma} \xrightarrow{D} X \sim N(0,1)$$

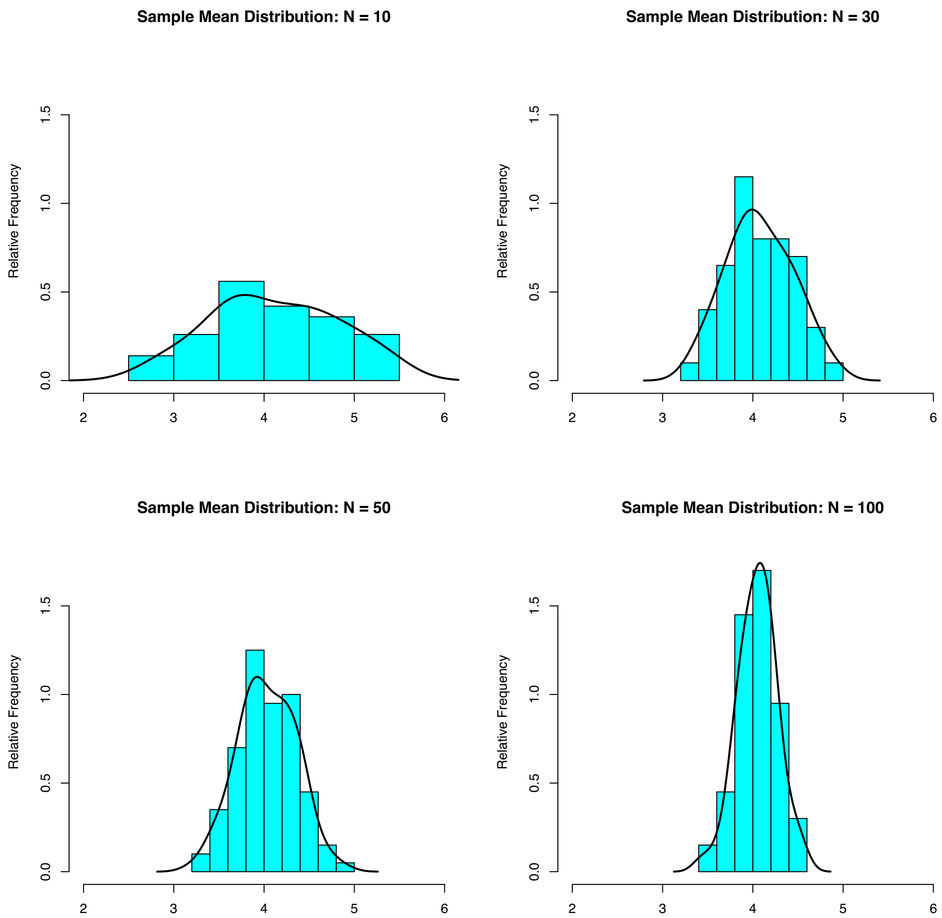
or

$$\overline{X}_n \approx N\left(\mu, \frac{\sigma^2}{n}\right),$$

where $\approx$ represents "is approximately distributed."

Remark 2.27. The powerful result of CLT is valid under any probability distribution. In general, when the sample size n is greater than or equal to 30, we can assume that the sample mean is approximately normally distributed.

Example 2.28. Suppose that a population is governed by a uniform distribution $U(0,8)$. The sample mean is computed from samples of size 10, 30, 50, and 100 respectively from this population.



2.5. Moments of a Random Variable

$m'_1 = E[X] = \mu_x$

$m_2 = E[(X-\mu_x)^2] = Var[x] = \sigma^2$

Definition 2.29. The kth *moment* of a random variable X is given by

$$m'_k = E[X^k].$$

- The first moment, denoted by $E[X] = \mu_x$, is called the mean or expectation of X. It measures the central location of the distribution.

Definition 2.30. The kth *central moment* of a random variable X is given by

$$m_k = E[(X-\mu_x)^k].$$

- The second central moment, denoted by σ_x^2, is called the variance of X. The positive square root, σ_x, of the variance is the standard deviation of X. The variance (or standard deviation) measures the variability or uncertainty of X.

Definition 2.31. The *skewness* and *kurtosis* of X are given by

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

That is, they are the normalized third and fourth central moments.

- The skewness measures the symmetry of X with respect to its mean. A symmetric distribution has $S(x) = 0$, while an asymmetric distribution has $S(x) \neq 0$.

- The quantity $K(x) - 3$ is called the *excess kurtosis*. The excess kurtosis of a Gaussian random variable is zero. A distribution with positive excess kurtosis is said to have heavy tails, meaning that the distribution puts more mass on the tails than a normal distribution does. Such a distribution is said to be *leptokurtic*.

Remark 2.32. Let $\{x_1, \ldots, x_T\}$ be a random sample of X with T observations. The moments are estimated by their sample counterparts:

estimator

$$\hat{\mu}_x = \frac{1}{T}\sum_{t=1}^{T} x_t$$
$$\hat{\sigma}_x^2 = \frac{1}{T-1}\sum_{t=1}^{T}(x_t - \hat{\mu}_x)^2$$
$$\hat{S}(x) = \frac{1}{(T-1)\hat{\sigma}_x^3}\sum_{t=1}^{T}(x_t - \hat{\mu}_x)^3$$
$$\hat{K}(x) = \frac{1}{(T-1)\hat{\sigma}_x^4}\sum_{t=1}^{T}(x_t - \hat{\mu}_x)^4.$$

Theorem 2.33. *In testing $H_0 : \mu_x = 0$ versus $H_1 : \mu_x \neq 0$, the test statistic is*

$$\frac{\hat{\mu}_x}{\hat{\sigma}_x/\sqrt{T}} \approx N(0,1)$$

for a sufficiently large T under the null hypothesis. $\left|\frac{\hat{\mu}_x}{\hat{\sigma}_x/\sqrt{n}}\right| > z_{0.025} \Rightarrow H_0$: reject

Remark 2.34. If a test statistic follows a standard normal distribution, H_0 is rejected at the 5% significance level if the absolute value of the test statistic is greater than $z_{0.025}$, a critical value representing the probability remains in the right tail for the standard normal distribution. Generally speaking, H_0 is rejected at the 5% level if the p-value is less than 0.05.

Theorem 2.35. *In testing $H_0 : S(x) = 0$ versus $H_1 : S(x) \neq 0$, the test statistic is*

$$\frac{\hat{S}(x)}{\sqrt{6/T}} \approx N(0,1)$$

under the null hypothesis. In testing $H_0 : K(x) - 3 = 0$ versus $H_1 : K(x) - 3 \neq 0$, the test statistic is

$$\frac{\hat{K}(x) - 3}{\sqrt{24/T}} \approx N(0,1)$$

under the null hypothesis.

fat tail
lepto kurtic

Theorem 2.36. *The Jarque-Bera test statistic is given by*

$$JB = \frac{T}{6}\left(\hat{S}^2(x) + \frac{(\hat{K}(x) - 3)^2}{4}\right).$$

Under the null hypothesis that x is normally distributed, it shows

$$JB \approx \chi^2_{(2)}.$$

Example 2.37. Consider the daily simple returns of the 3M stock from January 2, 2001 to September 30, 2011.

```
> mydat <- read.table("data1.txt", header = T)
> head(mydat)
      date      rtn
1 20010102 -0.010892
...
6 20010109 -0.015727
> rtn <- mydat[ , 2]
> mean(rtn)
[1] 0.000277767
> var(rtn)
[1] 0.0002398835
> t.test(rtn)
```

$\frac{0.00027-0}{\sqrt{0.000239}/\sqrt{n}} \sim t(n-1)$

```
One Sample t-test
data: rtn
t = 0.9326, df = 2703, p-value = 0.3511
alternative hypothesis: true mean is not equal to 0
```

p-value > 0.05
⇒ H_0 not reject

- One cannot reject $H_0 : \mu_x = 0$ at the 5% level.

```
> library(fBasics)
> (S <- skewness(rtn))
[1] 0.02794929
> T <- length(rtn)
> (t.stat <- S/sqrt(6/T))
[1] 0.593333
> (p.val <- 2*(1 - pnorm(t.stat)))
[1] 0.5529583
```

$\frac{\hat{S}_x}{\sqrt{\frac{6}{n}}}$

- One cannot reject zero skewness at the 5% level. H_0

```
> (K <- kurtosis(rtn))
[1] 4.630925
> (t.stat <- K/sqrt(24/T))
[1] 49.15475
> (p.val <- 2*(1 - pnorm(t.stat)))
[1] 0
```

$\frac{\hat{K}_x}{\sqrt{\frac{24}{n}}}$

- One rejects the null hypothesis of $K(x) = 3$. So, daily returns have heavy tails.

```
> normalTest(rtn, method = "jb")
Title: Jarque - Bera Normalality Test
Test Results:
STATISTIC: X-squared: 2422.4384
P VALUE: Asymptotic p Value: < 2.2e-16
```

- The normality assumption is rejected.

3. Normal and Lognormal Distributions

Definition 3.1. If X is a random variable whose logarithm is normally distributed (i.e., $\log(X) \sim N(\mu, \sigma^2)$), then X has a *lognormal distribution*. The pdf of X is given by

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma}\frac{1}{x}\exp\left(-\frac{(\log x - \mu)^2}{2\sigma^2}\right)$$

for $0 < x < \infty$.

- For $\log(X) \sim N(\mu, \sigma^2)$, it shows

$$E[X] = \exp\left(\mu + \frac{\sigma^2}{2}\right) \tag{3.1}$$

$$Var[X] = \exp\left(2(\mu+\sigma^2)\right) - \exp\left(2\mu+\sigma^2\right). \tag{3.2}$$

One may assume that the one-period simple return R_t is independently and identically distributed (iid) as normal, i.e.,

$$R_t \sim N(\mu, \sigma^2).$$

1) $-1 \leq R_t < \infty$

- The normality assumption induces several difficulties. First, the lower bound of a simple return is -1, while the normal distribution has no lower bound. Second, the multiperiod simple return is not normally distributed since it is a product of one-period simple gross returns minus one. Third, asset returns often exhibit positive excess kurtosis.

So, it is better to assume that the one-period log return r_t is iid as normal, i.e.,

2) $X_1 \sim N(\mu_1, \sigma_1^2)$
$X_2 \sim N(\mu_2, \sigma_2^2)$
$\Rightarrow X_1 \cdot X_2 \nsim N$

$$r_t = \ln(1+R_t) \sim N(\mu, \sigma^2).$$

Proposition 3.2. *Under the assumption, it shows that the one-period simple return is iid as lognormal with*

$r_t \sim N(\mu, \sigma^2)$

$$E[R_t] = \exp\left(\mu + \frac{\sigma^2}{2}\right) - 1$$

$$Var[R_t] = \exp(2\mu+\sigma^2)\left[\exp(\sigma^2) - 1\right].$$

simple return
vs
log return

Proof. From (3.1) and (3.2), one sees that

$$E[1+R_t] = 1 + E[R_t] = \exp\left(\mu + \frac{\sigma^2}{2}\right)$$

and

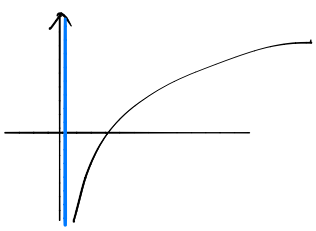
$$\begin{aligned} Var[1+R_t] &= Var[R_t] \\ &= \exp\left(2(\mu+\sigma^2)\right) - \exp\left(2\mu+\sigma^2\right) \\ &= \exp(2\mu+\sigma^2)\exp(\sigma^2) - \exp\left(2\mu+\sigma^2\right) \\ &= \exp(2\mu+\sigma^2)\left[\exp(\sigma^2)-1\right]. \end{aligned}$$

□

Remark 3.3. The log normal assumption has two advantages. First, r_t has no lower bound. Second, the multiperiod log return is also normally distributed since it is a sum of one-period returns. But, it is still inconsistent with the positive excess kurtosis.

$X_1 + X_2 \sim N$
where $X_1 \sim N(\mu_1, \sigma_1^2)$
$X_2 \sim N(\mu_2, \sigma_2^2)$

1)



$r_t + r_{t+k} \sim N$

$\log X \sim N(\mu, \sigma^2)$

$E[\log X] = \mu, \quad Var[\log X] = \sigma^2$

$E[X] = \exp(\mu + \frac{\sigma^2}{2}).$

$Var[X] = \exp(2(\mu+\sigma^2)) - \exp(2\mu+\sigma^2)$
$= \exp(2\mu+\sigma^2)[\exp(\sigma^2) - 1]$

1). $R_t \sim N(\mu, \sigma^2).$

1) $-1 \leq R_t$ exist lower bound.

2) multi-period simple return. $R_{t-k\to t} : \frac{P_t}{P_{t-k}} - 1$

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

= product of simple gross return $-1.$

It is not distribute normal.

$(\because R_t \sim N(\mu, \sigma^2). \quad (1+R_t)(1+R_{t-1}) \not\sim N(\mu, \sigma^2).$

3) positive excess kurtosis.

2) $r_t = \ln(1+R_t) \sim N(\mu, \sigma^2).$

advantage.

1) no lower bound

2) multi period simple return distributed N.

$r_{t-k\to t} = \ln(\frac{P_t}{P_{t-k}}) = \ln(\frac{P_t}{P_{t-1}} \cdot \frac{P_{t-1}}{P_{t-2}} \cdots \frac{P_{t-k+1}}{P_{t-k}})$

$= \ln(1+R_t) + \cdots + \ln(1+R_{t-k+1}).$

$= r_t + r_{t-1} + \cdots + r_{t-k+1}$

Sum of log return is also distribute normally.

$$r_t = \ln(1+R_t) \sim N(\mu, \sigma^2).$$

$$E[1+R_t] = \exp(\mu + \frac{\sigma^2}{2})$$

$$\therefore E[R_t] = \exp(\mu + \frac{\sigma^2}{2}) - 1$$

$$\begin{aligned} Var[1+R_t] &= Var[R_t] \\ &= \exp(2(\mu+\sigma^2)) - \exp(2\mu + \sigma^2) \\ &= \exp(2\mu+\sigma^2)[\exp(\sigma^2) - 1] \end{aligned}$$