EF4822 Financial Econometrics

Week 3: Asset Returns and Statistics

Dr. LUO Ding

I. Definitions of Returns

Returns

- *Return*

$$R_{t\to t+1} = R_{t+1} = \frac{P_{t+1} + D_{t+1}}{P_t} - 1$$

- *Gross return*: $1 + R_{t+1} \geq 0$

- *Compound return* over k periods

$$\begin{aligned} R_{t\to t+k} &= (1 + R_{t+1})(1 + R_{t+2}) \cdots (1 + R_{t+k}) - 1 \\ &= \prod_{j=1}^{k} (1 + R_{t+j}) - 1 \end{aligned}$$

- *Log return*, also called *continuously compounded return*

$$r_{t+1} = \log(1 + R_{t+1}) = \log(P_{t+1} + D_{t+1}) - \log(P_t)$$

- *Log compound return* over k periods

$$r_{t \to t+k} = \log(1 + R_{t \to t+k}) = \log\left(\prod_{j=1}^{k}(1 + R_{t+j})\right)$$

$$= \sum_{j=1}^{k} \log(1 + R_{t+j}) = \sum_{j=1}^{k} r_{t+j}$$

Portfolio Returns

- *Portfolio return* (i indexes assets)

$$R_{pt+1} = \sum_{i=1}^{N} w_{it} R_{it+1}, \quad \text{where } \sum_{i=1}^{N} w_{it} = 1$$

  - Equally-weighted portfolio, $w_{it} = 1/N$
  - Value-weighted portfolio: $w_{it} = \frac{MV_{it}}{\sum_{i=1}^{N} MV_{it}}$

- Log portfolio returns

$$r_{pt+1} = \log(1+R_{pt+1}) = \log\left(\sum_{i=1}^{N} w_{it}(1+R_{it+1})\right) \neq \sum_{i=1}^{N} w_{it} r_{it+1}$$

Excess Returns

- *Excess return* is over a benchmark return (eg, Treasury bill)

$$R^e_{it} = R_{it} - R_{0t}$$

  - Corresponds to payoff on a zero-cost portfolio that goes long in asset i and short in the benchmark asset
  - You short one dollar in the benchmark and receive \$1, you use this \$1 to buy asset i

- *Log excess return*

$$r^e_{it} = r_{it} - r_{0t}$$

- *Compound excess return* over k periods

$$R^e_{it\to t+k} = (1 + R_{it\to t+k}) - (1 + R_{0t\to t+k})$$

- *Log excess return* over k periods

$$r^e_{it\to t+k} = \sum_{j=1}^{k} (r_{it+j} - r_{0t+j}) = \sum_{j=1}^{k} r^e_{it+j}$$

- Nominal (gross) return

$$1 + R_{t+1} = \frac{\$\text{received}_{t+1}}{\$\text{paid}_t}$$

- Real rates of returns are defined in terms of real \$'s or goods:

$$1 + R_{t+1}^{real} = \frac{\text{goods received}_{t+1}}{\text{goods paid}_t}$$

- Inflation:

$$1 + \Pi_{t+1} = \frac{\text{CPI}_{t+1}}{\text{CPI}_t}, \quad \text{where } \text{CPI}_t = \frac{\$_t}{\text{goods}_t}$$

- Therefore,

$$1 + R_{t+1}^{real} = \frac{1 + R_{t+1}}{1 + \Pi_{t+1}} \text{ or, in logs, } r_{t+1}^{real} = r_{t+1} - \pi_{t+1}$$

II. Review of Statistics

- Rather than get caught up in the math (probability and measure theory), we'll look at random variables, distributions, and statistics from a computational viewpoint

- More specifically, rather than talk about events directly, we will refer to random variables and (implicitly) the events determined by them

Definition

- The *sample space* is the set of all possible outcomes
- A *random variable* (RV), X, is a real-valued function whose values can be assigned a probability to any interval of the form $(-\infty, c]$.

Example: Log Stock Return

- Sample space is $\mathbb{R} = (-\infty, \infty)$
- Random variable here is $X = \ln(P_t + D_t) - \ln(P_{t-1})$

Example: Firm Bankruptcy

- Sample space is $\{\underline{O}perating, \underline{D}efault\}$
- Define X by $X(O) = 0$ and $X(D) = 1$, then X is a random variable

- A random variable is discrete if it takes on a *countable* number of values
- If X is discrete, it takes values $x_1, x_2, \dots$ with the associated probabilities $f(x_1), f(x_2), \dots$
- The points $x_1, x_2, \dots$ are called the *points of support*

Example: Firm Bankruptcy

- If outcomes $X(O) = 0$ and $X(D) = 1$ are equal, then $f(0) = f(1) = 1/2$
- The points of support are $x_1 = 0$ and $x_2 = 1$

- Let $F(c)$ give the probability that $X \leq c$ and call $F(\cdot)$ the (cumulative) distribution function (cdf)
  1. They are increasing and right-continuous with left limits
  2. $F(-\infty) = 0$ and $F(\infty) = 1$

- If a RV is discrete, then its cdf is a step function: $F(c) = \sum_{x_i < c} f(x_i)$

Example: Firm Bankruptcy

$$F(c) = \begin{cases} 0 & -\infty < c < 0 \\ 1/2 & 0 \leq c < 1 \\ 1 & 1 \leq c < \infty \end{cases}$$

- If a RV is continuous, then so is its cdf. We write $F(c) = \int_{-\infty}^{c} f(x)dx$, where we call $f(x) \equiv F'(x)$ the (probability) density function (pdf)

- Distributions involving one RV are called *univariate*; involving two RVs, *bivariate*; involving many, *multivariate*
- Suppose two discrete RVs, X and Y have this bivariate (joint) distribution:

Example

	$f(X,Y)$	X **-0.17**	**-0.07**	**0.02**	**0.12**
	-0.16	0.005	0.056	0.071	0
Y	**0.05**	0	0.109	0.685	0.045
	0.25	0	0	0.017	0.012

- Thus the probability of jointly observing $X = -0.07$ and $Y = 0.05$ is $f(-0.07, 0.05) = 0.109$ or 10.9 percent

- The probability that a discrete RV X takes on the values x_i whatever the value of Y is called the marginal distribution of X and will be denoted by $f_X(x)$. This holds analogously for $f_Y(y)$
- You can compute these via

$$f(x_i) = \sum_j f(x_i, y_j) \text{ and } f(y_j) = \sum_i f(x_i, y_j)$$

- This notation can be confusing because the function f is *overburdened*, but in most cases it will be clear from the context. When it isn't clear I'll use subscripts: f_X and f_Y

Example

- $f_X(0.02) = 0.773$

- Two random variables are *independent* iff $f(x, y) = f(x)f(y)$

Example

- $f(0.02, 0.05) = 0.685$ but $f_X(0.02)f_Y(0.05) = 0.773 \times 0.839 = 0.649$, so X and Y aren't independent

- For a discrete RV, the *expected value* or the *expectation* of X, when it exists, is defined by

$$\mathbb{E}[X] = \sum_i x_i f(x_i)$$

- For a continuous RV, we have $\mathbb{E}[X] = \int xf(x)dx$, where these unlabelled integrals are understood to be taken over $\mathbb{R}$ (and equivalently over the RV's support)
- Note $\mathbb{E}[X]$ is a number, not a RV. Some people denote with $E[X]$ or EX

Example

$$\mathbb{E}[X] = -0.17(0.005) + -0.07(0.165) + 0.02(0.773) + 0.12(0.057) = 0.0099$$
$$\mathbb{E}[Y] = -0.16(0.132) + 0.05(0.839) + 0.25(0.029) = 0.028$$

- The expected value of X might not exist (that is X is not an integrable random variable). For example,

1. It may be "infinite". For example, Bernoulli paradox: Suppose X takes on $2, 4, 8, \ldots$ with probabilities $1/2, 1/4, 1/8, \ldots$. We get

$$\begin{aligned} \mathbb{E}[X] &= 2(1/2) + 4(1/4) + 8(1/8) + \cdots \\ &= 1 + 1 + 1 + \ldots = \infty \end{aligned}$$

- If $a_1, a_2, \ldots, a_n$ are constants and $X_1, X_2, \ldots, X_n$ are integrable RVs, then

  $$\mathbb{E}[a_1 X_1 + a_2 X_2 + \cdots + a_n X_n] = a_1 \mathbb{E}[X_1] + a_2 \mathbb{E}[X_2] + \cdots + a_n \mathbb{E}[X_n]$$

- In words, the expectation operator is *linear*. This is very important and useful

- The covariance of two random variables is

$$\text{cov}(X, Y) = \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])]$$

- It is measure of the *linear* association between X and Y: $\text{cov}(X, Y) > 0$ implies that large (small) values of X tend to be associated with large (small) values of Y

Properties

- Can be written as $\text{cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$
- It is symmetric: $\text{cov}(X, Y) = \text{cov}(Y, X)$
- It is bilinear: $\text{cov}(a_1X_1 + a_2X_2, Y) = a_1\text{cov}(X_1, Y) + a_2\text{cov}(X_2, Y)$, for constants a_1, a_2
- If X and Y are independent, then $\text{cov}(X, Y) = 0$

Example

		X			
	XY	**-0.17**	**-0.07**	**0.02**	**0.12**
	-0.16	0.027	0.011	-0.003	-0.019
Y	**0.05**	-0.009	-0.004	0.001	0.006
	0.25	-0.043	-0.018	0.005	0.03

- Thus, the covariance is

$$
\begin{aligned}
\operatorname{cov}(X,Y) &= \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] \\
&= \sum_{i,j} x_i y_j f(x_i, y_j) - \sum_i x_i f(x_i) \times \sum_j y_j f(y_j) \\
&= 0.0015 - 0.0099 \times 0.028 = 0.0012
\end{aligned}
$$

- Variance is naturally defined as

$$\mathrm{cov}(X, X) = \mathrm{var}(X) = \mathbb{E}[(X - \mathbb{E}[X])^2]$$

- From bilinearity, $\mathrm{var}(aX) = a^2\mathrm{var}(X)$

- Further, if a and b are constants, and X and Y have finite variances, then

$$\mathrm{var}(aX + bY) = a^2\mathrm{var}(X) + b^2\mathrm{var}(Y) + 2ab\mathrm{cov}(X, Y)$$

- The correlation between two RVs X and Y is a measure of the *linear* association between them

$$\mathrm{corr}(X,Y) = \frac{\mathrm{cov}(X,Y)}{\sqrt{\mathrm{var}(X)}\sqrt{\mathrm{var}(Y)}}$$

Properties

- Correlation lies in the interval $[-1, 1]$
- If $\mathrm{corr}(X,Y) = 1$ then they're *perfectly (positively) correlated*
- If $\mathrm{corr}(X,Y) = 0$, then they're *uncorrelated*

Example

$$\begin{aligned}\mathrm{var}(X) &= \mathbb{E}[X^2] - \mathbb{E}[X]^2 \\ &= \sum_i x_i^2 f(x_i) - \left(\sum_i x_i f(x_i)\right)^2 \\ &= 0.0021 - 0.0099^2 = 0.0020\end{aligned}$$

$$\begin{aligned}\mathrm{var}(Y) &= 0.0065 \\ \mathrm{corr}(X, Y) &= \frac{0.0012}{\sqrt{0.0020}\sqrt{0.0065}} = 0.33\end{aligned}$$

Normal Distributions and Higher Moments

- By far, the most useful distribution is the normal distribution. For a given mean μ and variance σ^2 it is denoted

$$\mathcal{N}(\mu, \sigma^2)$$

- The univariate normal distribution with mean μ and variance σ^2 has the density with support $x \in \mathbb{R}$

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

- The *standard normal distribution* is standardized in the sense that its mean is zero and variance (and standard deviation) are one:

$$\mathcal{N}(0, 1)$$

- The *lth centered moment* of a continuous RV X with mean μ and variance σ^2 is defined as

$$m_l = \mathbb{E}[(X-\mu)^l]$$

- The third standardized central moment is called *Skewness* which measures the asymmetry of X with respect to its mean

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

- Normal distributions have zero skewness

- *Kurtosis* is the fourth standardized central moment and measures the tail thickness of X

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

- *Excess kurtosis* is $K(X) - 3$. Excess to what? The normal distribution, which has a kurtosis of 3 and therefore an excess kurtosis of zero

- A RV with positive excess kurtosis is said to be *leptokurtic* or have heavy tails, implying it puts more mass on the tails of its support than a normal distribution. That is, it more commonly takes on extreme values.

- Let $\{x_t\}_{t=1}^{T}$ be a random sample of X with T observations. From these data, we can estimate the first four moments

- Sample mean: $\widehat{\mu} = \frac{1}{T}\sum_{t=1}^{T} x_t$

- Sample (co)variance: $\widehat{\sigma^2} = \frac{1}{T-1}\sum_{t=1}^{T}(x_t - \widehat{\mu}_X)(y_t - \widehat{\mu}_Y)$

- Sample skewness: $\widehat{S(X)} = \frac{1}{(T-1)(\widehat{\sigma}^2)^{3/2}}\sum_{t=1}^{T}(x_t - \widehat{\mu})^3$

- Sample kurtosis: $\widehat{K(X)} = \frac{1}{(T-1)(\widehat{\sigma}^2)^2}\sum_{t=1}^{T}(x_t - \widehat{\mu})^4$

Distributional properties of returns

Key: What is the distribution of

$\{r_{it}; i = 1, \cdots, N; t = 1, \cdots, T\}$?

Some theoretical properties:

Moments of a random variable X with density $f(x)$: ℓ-th moment

$$m'_\ell = E(X^\ell) = \int_{-\infty}^{\infty} x^\ell f(x) dx$$

First moment: mean or expectation of X.

ℓ-th central moment

$$m_\ell = E[(X - \mu_x)^\ell] = \int_{-\infty}^{\infty} (x - \mu_x)^\ell f(x) dx,$$

2nd central moment: Variance of X.

standard deviation: square-root of variance

Skewness (symmetry) and kurtosis (fat-tails)

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

$K(x) - 3$: Excess kurtosis.

Q1: Why study the mean and variance of returns?

They are concerned with long-term return and risk, respectively.

Q2: Why is symmetry important?

Symmetry has important implications in holding short or long financial positions and in risk management.

Q3: Why is kurtosis important?

Related to volatility forecasting, efficiency in estimation and tests

High kurtosis implies heavy (or long) tails in distribution.

Estimation:

Data: $\{x_1, \cdots, x_T\}$

- sample mean:
$$\hat{\mu}_x = \frac{1}{T}\sum_{t=1}^{T} x_t,$$
- sample variance:
$$\hat{\sigma}_x^2 = \frac{1}{T-1}\sum_{t=1}^{T}(x_t - \hat{\mu}_x)^2,$$
- sample skewness:
$$\hat{S}(x) = \frac{1}{(T-1)\hat{\sigma}_x^3}\sum_{t=1}^{T}(x_t - \hat{\mu}_x)^3,$$
- sample kurtosis:
$$\hat{K}(x) = \frac{1}{(T-1)\hat{\sigma}_x^4}\sum_{t=1}^{T}(x_t - \hat{\mu}_x)^4.$$

Under normality assumption,
$$\hat{S}(x) \sim N(0, \frac{6}{T}), \quad \hat{K}(x) - 3 \sim N(0, \frac{24}{T}).$$

Some simple tests for normality (for large T).

1. Test for symmetry:
$$S^* = \frac{\hat{S}(x)}{\sqrt{6/T}} \sim N(0,1)$$
if normality holds.
Decision rule: Reject H_o of a symmetric distribution if $|S^*| > Z_{\alpha/2}$ or p-value is less than α.

2. Test for tail thickness:
$$K^* = \frac{\hat{K}(x) - 3}{\sqrt{24/T}} \sim N(0,1)$$

if normality holds.
Decision rule: Reject H_o of normal tails if $|K^*| > Z_{\alpha/2}$ or p-value is less than α.

3. A joint test (Jarque-Bera test):

$$JB = (K^*)^2 + (S^*)^2 \sim \chi_2^2$$

if normality holds, where χ_2^2 denotes a chi-squared distribution with 2 degrees of freedom.
Decision rule: Reject H_o of normality if $JB > \chi_2^2(\alpha)$ or p-value is less than α.

Empirical properties of returns

Data sources: Use packages, e.g. quantmod

- Yahoo Finance: https://finance.yahoo.com/
- CRSP: Center for Research in Security Prices (Wharton WRDS) https://wrds-web.wharton.upenn.edu/wrds/
- Various web sites, e.g. Federal Reserve Bank at St. Louis https://research.stlouisfed.org/fred2/
- Data sets of textbooks: http://faculty.chicagobooth.edu/ruey.tsay/teaching/fts3/

Empirical dist of asset returns tends to be skewed to the left with heavy tails and has a higher peak than normal dist.

Demonstration of Data Analysis

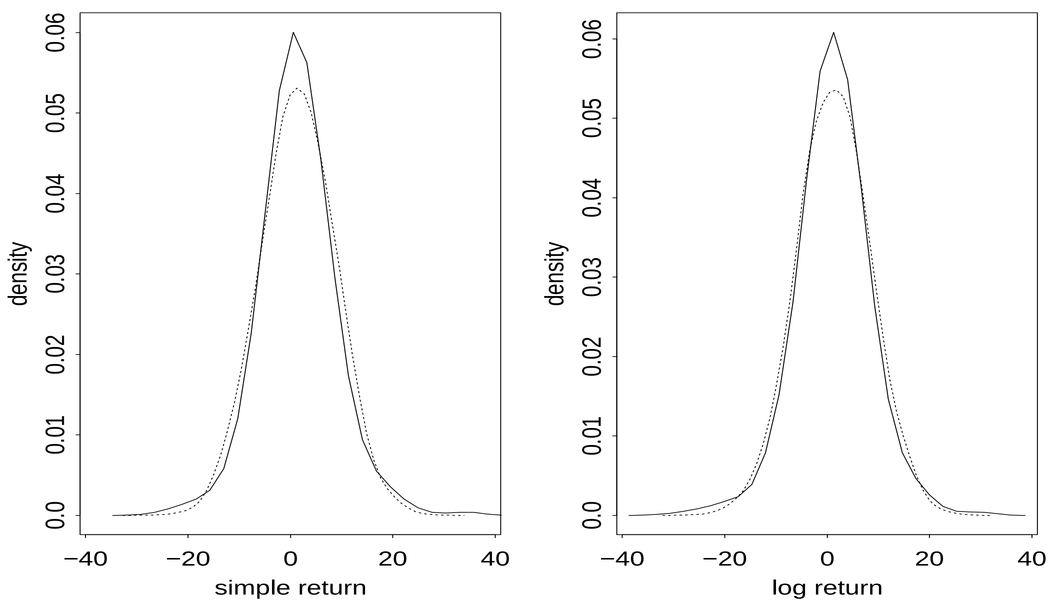

Figure 1: Comparison of empirical IBM return densities (solid) with Normal densities (dashed)

R demonstration: Use monthly IBM stock returns from 1967 to 2008.

```
**** Task: (a) Set the working directory
           (b) Load the library ``fBasics''.
           (c) Compute summary (or descriptive) statistics
           (d) Perform test for mean return being zero.
           (e) Perform normality test using the Jaque-Bera method.
           (f) Perform skewness and kurtosis tests.

> setwd("C:/Users/rst/teaching/bs41202/sp2017")   <== set working directory
> library(fBasics)    <== Load the library ``fBasics''.

> da=read.table("m-ibm-6815.txt",header=T)
> head(da)
  PERMNO     date    PRC ASKHI  BIDLO       RET    vwretd    ewretd    sprtrn
1  12490 19680131 594.50 623.0 588.75 -0.051834 -0.036330  0.023902 -0.043848
2  12490 19680229 580.00 599.5 571.00 -0.022204 -0.033624 -0.056118 -0.031223
3  12490 19680329 612.50 612.5 562.00  0.056034  0.005116 -0.011218  0.009400
4  12490 19680430 677.50 677.5 630.00  0.106122  0.094148  0.143031  0.081929
5  12490 19680531 357.00 696.0 329.50  0.055793  0.027041  0.091309  0.011169
6  12490 19680628 353.75 375.0 346.50 -0.009104  0.011527  0.016225  0.009120
> dim(da)
[1] 576   9
> ibm=da$RET  % Simple IBM return
> lnIBM <- log(ibm+1) % compute log return
> ts.plot(ibm,main="Monthly IBM simple returns: 1968-2015") % Time plot
> mean(ibm)
[1] 0.008255663
> var(ibm)
[1] 0.004909968
> skewness(ibm)
[1] 0.2687105
attr(,"method")
[1] "moment"
> kurtosis(ibm)
[1] 2.058484
attr(,"method")
[1] "excess"
> basicStats(ibm)
                   ibm
nobs        576.000000
NAs           0.000000
Minimum      -0.261905
Maximum       0.353799
1. Quartile  -0.034392
```

```
3. Quartile   0.048252
Mean          0.008256
Median        0.005600
Sum           4.755262
SE Mean       0.002920
LCL Mean      0.002521
UCL Mean      0.013990
Variance      0.004910
Stdev         0.070071
Skewness      0.268710
Kurtosis      2.058484
> basicStats(lnIBM)  % log return
                 lnIBM
nobs        576.000000
NAs           0.000000
Minimum      -0.303683
Maximum       0.302915
1. Quartile  -0.034997
3. Quartile   0.047124
Mean          0.005813
Median        0.005585
Sum           3.348008
SE Mean       0.002898
LCL Mean      0.000120
UCL Mean      0.011505
Variance      0.004839
Stdev         0.069560
Skewness     -0.137286
Kurtosis      1.910438
> t.test(lnIBM) %% Test mean=0 vs mean .not. zero

        One Sample t-test

data:  lnIBM
t = 2.0055, df = 575, p-value = 0.04538
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 0.0001199015 0.0115051252
sample estimates:
  mean of x
0.005812513

> normalTest(lnIBM,method='jb')
Title: Jarque - Bera Normalality Test

Test Results:
```

```
  STATISTIC:
    X-squared: 90.988
  P VALUE:
    Asymptotic p Value: < 2.2e-16

> s3=skewness(lnIBM); T <- length(lnIBM)
> tst <- s3/sqrt(6/T)  % test skewness
> tst
[1] -1.345125
> pv <- 2*pnorm(tst)
> pv
[1] 0.1785849
> k4 <- kurtosis(lnIBM)
> tst <- k4/sqrt(24/T)  % test excess kurtosis
> tst
[1] 9.359197
>q()   % quit R.
```