Financial Econometrics

Wang, Yanchu

Shanghai University of Finance and Economics

September 2024

Returns

- Denote by P_t the price of an asset at date t with no dividend.
- The simple net return R_t between dates $t-1$ and t is defined as

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

- The simple gross return between date $t-1$ and t is defined as $1+R_t$.
- The gross return over k periods from date $t-k$ to date t is

$$\begin{aligned} 1+R_t(k) &= (1+R_t)(1+R_{t-1})\cdots(1+R_{t-k+1}) \\ &= \frac{P_t}{P_{t-1}}\frac{P_{t-1}}{P_{t-2}}\cdots\frac{P_{t-k+1}}{P_{t-k}} = \frac{P_t}{P_{t-k}}. \end{aligned} \tag{2}$$

- The simple net return over k periods is just $R_t(k)$.
- These multiperiod returns $R_t(k)$ are called compound return.

Returns

- We need to specify the return horizon in order to compare different returns.
- Annualized return:

$$\text{Annualized } R_t(k)_g = [\prod_{j=0}^{k-1}(1+R_{t-j})]^{1/k} - 1. \tag{3}$$

$R_t(k)_g$ is called the geometric average.

- When returns $R_{t \ j}$, $j = 0, \ldots, k-1$, are small, the linear approximation holds:

$$\text{Annualized } \overline{R_t}(k) \approx \frac{1}{k}\sum_{j=0}^{k-1} R_{t-j}. \tag{4}$$

$\overline{R_t}(k)$ is called the arithmetic average.

- We can show that $R_t(k)_g \leqslant \overline{R_t}(k)$.

Jensen Inequality

- If $f(x)$ is a convex function, then

$$f(E[x]) \leqslant E[f(x)] \tag{5}$$

- If $f(x)$ is a concave function, then

$$f(E[x]) \geqslant E[f(x)] \tag{6}$$

- Example, $f(x) = e^x$ is a convex function, thus we have

$$e^{E[x]} \leqslant E[e^x].$$

Returns

- Compounding frequency
  Suppose the annualized return is R, the initial investment value is A, after n years, the terminal value of the investment is
  $$A(1+R)^n$$
- Now suppose the rate is compounded m times per year, then the terminal value is
  $$A(1+\frac{R}{m})^{mn},$$
- Let $m \longrightarrow \infty$, thus the continuous compounding return is
  $$\lim_{m\to\infty} A(1+\frac{R}{m})^{mn} = Ae^{R_c n}.$$

Returns

- Continuous compounding (log return):

$$r_t = \log(1 + R_t) = \log \frac{P_t}{P_{t-1}} = p_t - p_{t-1}. \tag{7}$$

- The advantage of log return:

  1. It is easy to calculate multiperiod return:

$$\begin{aligned} r_t(k) &= \log(1 + R_t(k)) = \log((1 + R_t)(1 + R_{t-1}) \cdots (1 + R_{t-k+1})) \\ &= \log(1 + R_t) + \log(1 + R_{t-1}) + \cdots + \log(1 + R_{t-k+1}) \\ &= r_t + r_{t-1} + \cdots + r_{t-k+1}. \end{aligned} \tag{8}$$

  2. r_t has no constrained lower limit - make it easier to apply to statistical analysis.

- The disadvantage of log return:

  For a portfolio, it holds that $R_{pt} = \sum_{i=1}^{N} w_i R_{it}$, but it only holds approximately that $r_{pt} \approx \sum_{i=1}^{N} w_i r_{it}$.

Returns

- For assets with dividend, the simple net return at date t is

$$R_t = \frac{P_t + D_t}{P_{t-1}} - 1. \tag{9}$$

  The log return is defined as

$$r_t = \log(P_t + D_t) - \log(P_{t-1}). \tag{10}$$

- Excess returns, defined as the difference between the asset's return and the return on some reference asset:

$$Z_{it} = R_{it} - R_{0t}. \tag{11}$$

- The log excess return:

$$z_{it} = r_{it} - r_{0t}. \tag{12}$$

Skewness and Kurtosis

- The first two moments can uniquely determine a normal distribution.
- The third central moment measures the symmetry of X with respect to its mean. The standardized third central moment is called skewness:
$$S(x) = E[\frac{(X-\mu)^3}{\sigma_x^3}]. \tag{13}$$
- The standardized fourth central moment is called kurtosis, it measures the tail behavior of X:
$$K(x) = E[\frac{(X-\mu)^4}{\sigma_x^4}]. \tag{14}$$
- The quantity $K(x) - 3$ is called the excess kurtosis. For normal distributions, $K(x) = 3$.
- The positive excess kurtosis means that the random sample from the distribution tends to contain more extreme values compared to the normal distribution.

Estimation

- Estimation of skewness and kurtosis. Let $\{x_1, ..., x_T\}$ be a random sample of X with T observations.
- The sample mean:

$$\widehat{\mu} = \frac{1}{T} \sum_{t=1}^{T} x_t. \tag{15}$$

- The sample variance:

$$\widehat{\sigma}^2 = \frac{1}{T} \sum_{t=1}^{T} (x_t - \widehat{\mu})^2. \tag{16}$$

- The sample skewness:

$$\widehat{S} = \frac{1}{T\widehat{\sigma}^3} \sum_{t=1}^{T} (x_t - \widehat{\mu})^3. \tag{17}$$

- The sample kurtosis:

$$\widehat{K} = \frac{1}{T\widehat{\sigma}^4} \sum_{t=1}^{T} (x_t - \widehat{\mu})^4. \tag{18}$$

Normality Test

- The estimated $\widehat{S}$ and $\widehat{K}-3$ are distributed normally and have variances $6/T$ and $24/T$ respectively, thus the t-statistics are $t=\frac{\widehat{S}}{\sqrt{6/T}}$ and $t=\frac{\widehat{K}-3}{\sqrt{24/T}}$.
- The Jarque-Bera normality test statistic:

$$JB=\frac{\widehat{S}^2}{6/T}+\frac{(\widehat{K}-3)^2}{24/T}. \tag{19}$$

- The JB test statistic is distributed as a χ^2-distribution with 2 degrees of freedom under the null.

Distribution of Returns

- Consider a collection of N assets at date t, each with return R_{it}, where $t = 1, ..., T$. The most general model is its joint distribution function:

$$F(R_{11}, ..., R_{N1;} R_{12,} ... R_{N2}; ...; R_{N1,} ..., R_{NT}; \mathbf{x}; \boldsymbol{\theta}), \tag{20}$$

  where $\mathbf{x}$ denotes state variables and $\boldsymbol{\theta}$ represents the parameter vector.
- In practice, the model of (20) is too general.
- The CAPM considers the joint distribution of the cross section of returns, $\{R_{1t}, ..., R_{Nt}\}$.
- Other models focus on the dynamic process of individual asset returns, $\{R_{i1}, ..., R_{iT}\}$, the time series of returns.

Distribution of Returns

- Now consider a joint distribution function for $\{R_{i1}, ..., R_{iT}\}$, $F(R_{i1}, ...R_{iT}; \boldsymbol{\theta})$.
- We may rewrite $F(R_{i1}, ...R_{iT}; \boldsymbol{\theta})$ as the product of conditional distributions:

$$\begin{aligned} F(R_{i1}, ...R_{iT}; \boldsymbol{\theta}) &= F(R_{i1})F(R_{i2}|R_{i1}) \cdots F(R_{iT}|R_{i,T-1}, ..., R_{i1}) \\ &= F(R_{i1}) \prod_{t=2}^{T} F(R_{it}|R_{i,t-1}, ..., R_{i1}). \end{aligned} \tag{21}$$

- If R_{it} is a continuously random variable, then (21) implies the joint density function is:

$$\begin{aligned} f(R_{i1}, ...R_{iT}; \boldsymbol{\theta}) &= f(R_{i1})f(R_{i2}|R_{i1}) \cdots f(R_{iT}|R_{i,T-1}, ..., R_{i1}) \\ &= f(R_{i1}) \prod_{t=2}^{T} f(R_{it}|R_{i,t-1}, ..., R_{i1}). \end{aligned} \tag{22}$$

Distribution of Returns

- Normal distribution:
  1. Easy to handle.
  2. R_{it} has a lower bound of -1, violating the normal assumption.
  3. The product of R_{it} will not be normally distributed.
- Lognormal distribution, $r_{it} \sim N(\mu, \sigma^2)$, then:

$$E[R_{it}] = e^{\mu + \frac{1}{2}\sigma^2} - 1, \tag{23}$$

and

$$Var[R_{it}] = e^{2\mu + \sigma^2}[e^{\sigma^2} - 1]. \tag{24}$$

- For R_{it} close to zero, $E[R_{it}] \approx \mu + \frac{1}{2}\sigma^2$.

Distribution of Returns

- Alternatively, let m and s^2 be the mean and the variance for the simple return R_{it}, then for the log return r_{it}:

$$E[r_{it}] = \log \frac{m+1}{\sqrt{1+(\frac{s}{m+1})^2}}, \tag{25}$$

  and

$$Var[r_{it}] = \log[1+(\frac{s}{m+1})^2]. \tag{26}$$

- The multiperiod log return $r_{it}(k)$ is also a normal distribution.
- The lognormal assumption is not consistent with all the properties of historical stock returns.

Likelihood Function of Returns

- Consider (22) with log return r_{it}. If $f(r_{i1},...r_{iT};\boldsymbol{\theta})$ is normal with mean μ_t and variance σ_t^2, then

$$f(r_{i1},...r_{iT};\boldsymbol{\theta}) = f(r_1;\boldsymbol{\theta})\prod_{t=2}^{T}\frac{1}{\sqrt{2\pi}\sigma_t}\exp(-\frac{(r_t-\mu_t)^2}{2\sigma_t^2}), \tag{27}$$

- The maximum likelihood estimate of $\boldsymbol{\theta}$ is obtained by maximizing the likelihood function (27). It equals to maximize the log likelihood function:

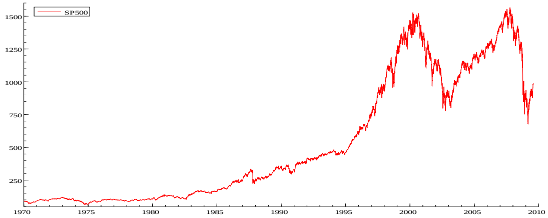
$$\begin{aligned}\boldsymbol{\theta} &= \max_{\boldsymbol{\theta}} \ln f(r_{i1},...r_{iT};\boldsymbol{\theta}) \\ &= \max_{\boldsymbol{\theta}} \ln f(r_1;\boldsymbol{\theta}) - \frac{1}{2}\sum_{t=2}^{T}(\ln(2\pi)+\ln(\sigma_t^2)+\frac{(r_t-\mu_t)^2}{\sigma_t^2}) \end{aligned} \tag{28}$$

Empirical Properties of Returns

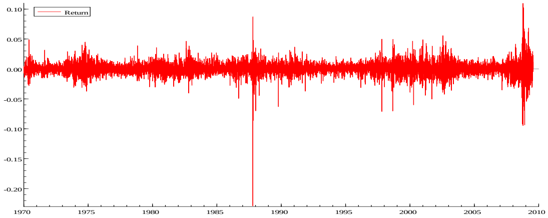
- The data are obtained from Yahoo!Finance.
- The daily closing price adjusted for dividends of SP500 Index are collected from 1970/01/05 $\sim$ 2009/07/31.

Empirical Properties of Returns

- The actual price process:



- The log return process:



Empirical Properties of Returns

- Descriptive statistics:

Start Date	End Date	Observations	Mean	Std. Deviation
70/01/05	09/07/31	9990	0.0002365	0.01084

Skewness	Excess Kurtosis	Maximum	Minimum	JB test
-1.083	27.91	0.1096	-0.2290	326268 (0.0)

- Returns are skewed and have fat tails, failed to follow normal distributions.
- Returns have more negative extreme values than positive extreme values.

Market Efficiency Hypothesis

- A market is said to be informationally efficient if it incorporates relative information to all market participants.
- Three forms of efficiency:
  1. Weak-form efficiency: I_t includes only the history of prices or returns.
  2. Semistrong-form efficiency: I_t includes all information known to *all* market participants (*publicly available information*).
  3. Strong-form efficiency: I_t includes all information known to *any* market participant (*private information*).
- If market is efficient, then price changes must be unforecastable given they are properly anticipated, i.e., the returns are random!
- Excess return is equal to zero on average - expected abnormal return $= 0$.

Conditional Expectation

- Suppose that y is a random scaler, and $\mathbf{x}$ is a vector of random variables, and aslo assume $E(|y| < \infty)$, we define

$$E[y|\mathbf{x}] = \mu(\mathbf{x})$$

  as the conditional expectation of y given the condition of $\mathbf{x}$.
- For example, if $\mathbf{x} = (x_1, x_2)$, $E[y|\mathbf{x}]$ could be

$$\begin{aligned}
E[y|x_1, x_2] &= \beta_0 + \beta_1 x_1 + \beta_2 x_2, \\
E[y|x_1, x_2] &= \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_2^2, \\
E[y|x_1, x_2] &= \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 x_2, \text{ and so on.}
\end{aligned}$$

- In financial econometrics, usually $\mathbf{x}$ is an information set. I_t denotes the information set which includes all relevatn information up to time t.

Conditional Expectation

- Suppose y is a function of $\mathbf{x}$, $y = f(\mathbf{x})$, we are interested in how y changes due the change of $\mathbf{x}$.
- However, there always exists unobservable errors, and thus usually impossible to precisely describe the change of y on the change of $\mathbf{x}$.
- We may instead investigate the marginal change of y on a certain element in $\mathbf{x}$,x_i, given all other elements are unchanged.
- The partial effect is defined as

$$\Delta E(y|\mathbf{x}) \approx \frac{\partial \mu(\mathbf{x})}{\partial x_i} \Delta x_i.$$

- When the conditional expecatation exists, we can always express the random scaler y as follows

$$y = E[y|\mathbf{x}] + u,$$

where $E[u|\mathbf{x}] = 0$.

Conditional Expectation

- $E[u|\mathbf{x}] = 0$ means:
  1. $E[u] = 0$.
  2. u is uncorrelated with any function of $\mathbf{x}$.
- The first property is derived from the law of iterated expectation

$$E[u] = E[E[u|\mathbf{x}]] = E[0] = 0.$$

Conditional Expectation

- The Law of Iterated Expectation.
  Define information set I_t and J_t, for which I_t is dominated by J_t, i.e. $I_t \subset J_t$. Consider the conditional expectation of a random variable X under these information sets we have:

$$E[X|I_t] = E[E[X|J_t]|I_t]. \tag{29}$$

  and

$$E[X|I_t] = E[E[X|I_t]|J_t].$$

- The smaller information set always dominates.
- For the unconditional expectation, we have

$$E[X] = E[E[X|Y]]. \tag{30}$$

- Proof for the discrete case:

$$
\begin{aligned}
E[E[X|Y]] &= \sum_y E[X|Y] \cdot P(Y = y) \\
&= \sum_y \sum_x (xP(X = x|Y = y)) \cdot P(Y = y) \\
&= \sum_y \sum_x x \cdot P(X = x, Y = y) \\
&= \sum_x x \cdot \sum_y P(X = x, Y = y) \\
&= \sum_x x \cdot P(X = x) \\
&= E[X].
\end{aligned}
$$

- Example: Let X be the schooling of the person, and Y be the monthly income of the person. Find expectation of Y, $E[Y]$, from the following Table.

X	Income Expecation	Probability of X_i
1	$E(Y\|X=1)=1000$	$P(X=1)=0.5$
2	$E(Y\|X=2)=500$	$P(X=2)=0.5$

$X=1$, have a university degree or above.
$X=2$, don't have a university degree.

- Solution:

$$\begin{aligned} E[Y] &= E[Y|X=1]P(X=1)+E[Y|X=2]P(X=2) \\ &= \sum E[Y|X=i]P(X=i) \\ &= E[E[Y|X=i]] \end{aligned}$$

Conditional Variance

- Define
$$Var(y|\mathbf{x}) = E[(y - E(y|\mathbf{x}))^2|\mathbf{x}] = E[y^2|\mathbf{x}] - [E(y|\mathbf{x})]^2.$$
- Properties:
  1. $Var(a(\mathbf{x})y + b(\mathbf{x})|\mathbf{x}) = [a(\mathbf{x})]^2 Var(y|\mathbf{x})$.
  2. $Var(y) = E[Var(y|\mathbf{x})] + Var(E[y|\mathbf{x}])$.
  3. $Var(y|\mathbf{x}) = E[Var(y|\mathbf{x}, \mathbf{z})|\mathbf{x}] + Var(E[y|\mathbf{x}, \mathbf{z}]|\mathbf{x})$.
  4. $E[Var(y|\mathbf{x})] \geqslant E[Var(y|\mathbf{x}, \mathbf{z})]$.

Market Efficiency Hypothesis

- Now consider a security price P_t can be a rational expectation of some fundamental value V^* conditional on information I_t available at time t.

$$P_t = E[V^*|I_t] = E_t V^*. \tag{31}$$

The same equation holds on one period ahead:

$$P_{t+1} = E[V^*|I_{t+1}] = E_{t+1} V^*. \tag{32}$$

- By the Law of Iterated Expectation:

$$E_t[P_{t+1} - P_t] = E_t[E_{t+1} V^* - E_t V^*] = 0.$$

Thus price changes are really unforecastable given information I_t available at time t.

Market Efficiency Hypothesis

- Is market efficiency testable?
- The test must assume an equilibrium model for normal security returns. If efficiency is rejected, it could be:
  1. The market is truly inefficient.
  2. The equilibrium model is incorrect.
- We test the joint hypothesis of market efficiency and market equilibrium.

Exercises 1

- Go to Yahoo!Finance and download daily, weekly and monthly data for Dow Jones Industrial Average Index from 1992/01/02 - 2014/12/31. Only use adjusted closing prices.
  1) For monthly prices, compute the one-month holding period returns as $R_t = \frac{P_t}{P_{t-1}} - 1$ over the whole period, and compute its arithmetic mean.
  2) Compute the geometric mean for monthly returns, compare the result with the arithmetic mean from 1).
  3) Compute the one-month log returns for monthly prices, calculate the mean μ and the variance σ^2 for the log returns.
  4) Check whether the approximation $E[R_t] \approx \mu + \frac{1}{2}\sigma^2$ holds or not.
  5) Provide the descriptive statistics for log returns on daily, weekly and monthly returns. Comment on the results.