Econometrics

Slides-0 N PROPERTIES Part II

Rachida Ouysse
School of Economics[1]

UNSW SYDNEY Australia's Global University

UNSW AUSTRALIA AGSM Business School

Shape Characteristics: Population

Let X_t be a random variable with pdf $f(x)$

$$E[X_t] : \text{ center}$$

$$var(X_t) = E[(X_t - \mu)^2] : \text{ spread}$$

$$skewness(X_t) = S(X) = E\left[\frac{(X_t - \mu)^3}{\sigma^3}\right] : \text{ symmetry}$$

$$kustosis(X_t) = K(X) = E\left[\frac{(X_t - \mu)^4}{\sigma^4}\right] : \text{ tail thickness}$$

$$K(X) - 3 : \text{ Excess kurtosis}$$

Note: The k^{th} moment and central moment of X_t are:

$$m'_k = E[X_t^k]$$

$$m_k = E[(X_t - \mu)^k]$$

Shape Characteristics of Random Variable

- Why are the mean a[illegible]e of returns important?

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

- Why is return symmetry of interest in financial study?

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

- Why is kurtosis important?

  Related to volatility forecasting, efficiency in estimation and tests, etc.

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

Examle: Normal Random Variable

Normal Distri

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

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

$$\begin{aligned} E[X] &= \mu \\ \text{var}(X) &= \sigma^2 \\ \text{skew}(X) &= 0 \\ \text{kurt}(X) &= 3 \\ m_k &= 0 \text{ for } k \text{ odd} \end{aligned}$$

Shape Characteristics: Sample moments

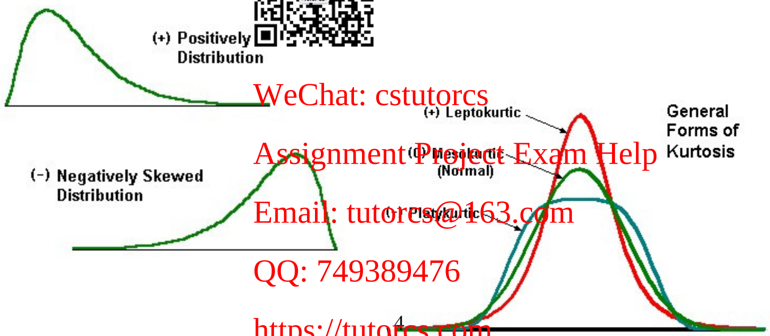
Sample moments

Let $\{r_t, \ldots, r_T\}$ denote a random sample of size T where r_t is a realization of the random variable r.

$$
\begin{aligned}
\hat{\mu} &= \frac{1}{T}\sum_{t=1}^{T} r_t, \quad \hat{\sigma}^2 = \frac{1}{T-1}\sum_{t=1}^{T}(r_t - \hat{\mu})^2 = \hat{m}_2 \\
\widehat{\text{skew}} &= \frac{\hat{m}_3}{\hat{\sigma}^3}, \quad \widehat{\text{kurt}} = \frac{\hat{m}_4}{\hat{\sigma}^3} \\
\hat{m}_k &= \frac{1}{T-1}\sum_{t=1}^{T}(r_t - \hat{\mu})^k,
\end{aligned}
$$

Note: we divide by $T-1$ to get unbiased estimates. Check software to see how moments are computed.

Shape Characteristics: Visually



Testing for normality

- **QQ-plot:** plot stan[illegible]pirical quantiles vs. theoretical quantiles from specified distri[illegible]e: Shapiro-Wilks (SW) test for normality: correlation coefficie[illegible]values used in QQ-plot
- **Jarque-Bera (JB)** test for normality

$$JB = \frac{T}{6}\left(\hat{skew}^2 + \frac{(\hat{kurt}-3)^2}{4}\right)$$

$$\sim_A \chi^2(2)$$

Note: if r_t is $N(\mu, \sigma^2)$ then:

$$\sqrt{T}\hat{skew} \sim N(0,6), \text{ and } \sqrt{T}(\hat{kurt}-3) \sim N(0,24)$$

Shape Characterirtics: Normality test

The null hypothesis:
H_0 : Data (the return) X_t is normally distributed.

1. **Skewness test:** $Z_{sk} = \frac{\hat{skew}}{\sqrt{6/T}} \sim N(0,1)$

   Reject H_0 if $|z_{sk}|$ is too large (> 1.96, at 5%).

2. **Kurtosis test:** $Z_{kt} = \frac{\hat{kurt}-3}{\sqrt{24/T}} \sim N(0,1)$

   Reject H_0 if $|z_{kt}|$ is too large (> 1.96, at 5%).

3. **Jaque-Bera test:** $JB = Z_{ks}^2 + Z_{kt}^2 \sim \chi_2^2$

   Reject JB is too large (> 5.99 at 5%)

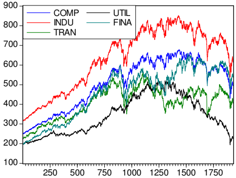
Example: Descriptive Statistics

- Descriptive

eg. NYSE index ...0103-20020830)

Composite ...

Trans, Utility, Finance.



Descriptive statistics of log returns.

	Composite	Industrial	Trans	Utility	Finance
Mean	0.035	0.034	0.031	0.007	0.052
Std. Dev.	1.006	1.009	1.320	1.087	1.310
Skewness	-0.316	-0.386	-1.044	-0.275	-0.042
Kurtosis	7.224	7.755	18.103	5.637	5.772

Portfolio variance and diversification:

$$Z = \frac{1}{2}(Y + X),$$

$$\mathrm{Var}(Z) = \frac{1}{4}[\mathrm{Var}(X) + \mathrm{Var}(Y) + 2\mathrm{Cov}(X, Y)]$$

Correlations of log returns

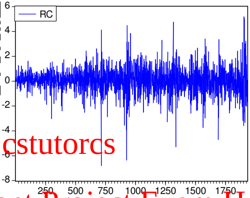
	Composite	Industrial	Trans	Utility	Finance
Composite	1				
Industrial	0.983	1			
Trans	0.731	0.708	1		
Utility	0.769	0.711	0.505	1	
Finance	0.885	0.800	0.668	0.623	1

Example: Descriptive Statistics

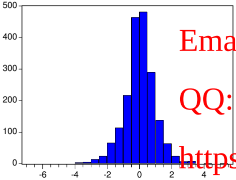
– Normality test

eg. Comp. index log

time series plot



histogram



Series: RC
Sample 1 1931
Observations 1930

Mean	0.035300
Median	0.052285
Maximum	5.178704
Minimum	-6.791142
Std. Dev.	1.006207
Skewness	-0.315728
Kurtosis	7.224376
Jarque-Bera	1467.129
Probability	0.000000

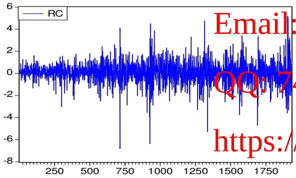
p-value = $P(\chi^2_{(2)} > JB)$

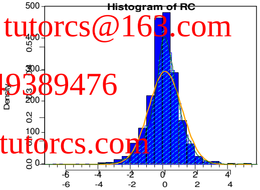
Stylized Fact: Large kurtosis

– Some stylise[…] out index return series

- concentrat[…] zero with a few large "outliers"
- large stand[…]ons (volatile)
- negative skewness (longer tail at the negative side)
- large kurtosis (tail probabilities larger than normal)
- large variations followed by large ones (clustering)

leptokurtic





Series: RC	
Sample 1 1931	
Observations 1930	
Mean	0.035300
Median	0.052285
Maximum	5.178704
Minimum	-6.791142
Std. Dev.	1.006207
Skewness	-0.315728
Kurtosis	7.224376
Jarque-Bera	1467.129
Probability	0.000000

Descriptive statistics: Autocorrelation

- **Predictability**
  - We say X_{t+1} is … information at t , eg. $\{X_t, X_{t-1} \cdots ,\}$, helps to improve our p… X_{t+1}.
  - In particular, X_t … able if X_{t+1} is correlated with X_{t-j} for some $j > 0$ (ie. $Cov($ … $\neq 0)$.
- **Autocorrelation Function (ACF)**
  - Autocovariance: $\gamma_j = Cov(X_t, X_{t-j}) = Cov(X_t, X_{t+j})$
    Sample autocovariance: $\hat{\gamma}_j = \frac{1}{T}\sum_{t=j+1}^{T}(X_t - \overline{X})(X_{t-j} - \overline{X})$
  - Autocorrelation: $\rho_j = \frac{\gamma_j}{\gamma_0}$
    Sample Autocorrelation: $\hat{\rho}_j = \frac{\hat{\gamma}_j}{\hat{\gamma}_0}$
- **Partial autocorrelation (PAC)**
  - PAC p_j is a measure of the direct relation between X_t and X_{t-j} for $j = 1, 2, \cdots$
  - p_j is the correlation between X_t and X_{t-j} after controlling for the effects of X_t and $X_{t-1} \cdots X_{t-j+1}$
  - $\hat{p}_1 = \hat{\phi}_{11}$ in $X_t = \phi_{10} + \phi_{11}X_{t-1} + e_{1t}$
  - $\hat{p}_2 = \hat{\phi}_{11}$ in $X_t = \phi_{20} + \phi_{21}X_{t-1} + \phi_{22}X_{t-2} + e_{2t}, \cdots$

Test for autocorrelation

The null hypothesis: H_0: There is no autocorrelation (White noise process)

1. **Autocorrelation test:** $\sqrt{T}\hat{\rho}_j \sim N(0,1)$ under the null hypothesis

   Reject if $|\hat{\rho}_j|$ is too large ($> 1.96/\sqrt{T}$, at 5% significance level)

Joint Hypothesis Tests

- We can also test the joint hypothesis that all m of the ρ_k correlation coefficients are simultaneously equal to zero using the Q-statistic developed by Box and Pierce:

$$Q = T \sum_{k=1}^{m} \hat{\rho}_k^2$$

where T=sample size, m=maximum lag length

- The Q-statistic is asymptotically distributed as a χ_m^2.

- However, the Box Pierce test has poor small sample properties, so a variant has been developed, called the Ljung-Box statistic:

$$Q^* = T(T+2) \sum_{k=1}^{m} \frac{\hat{\rho}_k^2}{T-k} \sim \chi_m^2$$

- This statistic is very useful as a portmanteau (general) test of linear dependence in time series.

An ACF Example

- Question:

  Suppose that a rese[illegible]estimated the first 5 autocorrelation coefficients using a [illegible]gth 100 observations, and found them to be (from 1 to 5): 0.[illegible], 0.086, 0.005, -0.022.

  Test each of the individual coefficient for significance, and use both the Box-Pierce and Ljung-Box tests to establish whether they are jointly significant.

Solution

A coefficient would be significant if it lies outside $(-0.196, +0.196)$ at the 5% level, so only the first autocorrelation coefficient is significant.

$Q = 5.09$ and $Q* = 5.26$

Compared with a tabulated $\chi^2(5)$=11.1 at the 5% level, so the 5 coefficients are jointly insignificant.

Example: ACF/PACF

- Descriptive st

  eg. NYSE composi

  AC test at 5% level:

  $1.96/\sqrt{T} = 0.04462$,

  H_0 is rejected at

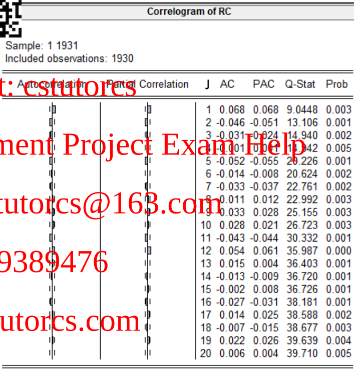
  $j = 1,2,5,12$

  LB test at 5% level:

  H_0 is rejected for

  all m, as all p-values

  are less than 0.05.



Correlogram of RC

Sample: 1 1931
Included observations: 1930

J	AC	PAC	Q-Stat	Prob
1	0.068	0.068	9.0448	0.003
2	-0.046	-0.051	13.106	0.001
3	-0.031	-0.024	14.940	0.002
4	-0.001	0.001	14.942	0.005
5	-0.052	-0.055	20.226	0.001
6	-0.014	-0.008	20.624	0.002
7	-0.033	-0.037	22.761	0.002
8	0.011	0.012	22.992	0.003
9	0.033	0.028	25.155	0.003
10	0.028	0.021	26.723	0.003
11	-0.043	-0.044	30.332	0.001
12	0.054	0.061	35.987	0.000
13	0.015	0.004	36.403	0.001
14	-0.013	-0.009	36.720	0.001
15	-0.002	0.008	36.726	0.001
16	-0.027	-0.031	38.181	0.001
17	0.014	0.025	38.588	0.002
18	-0.007	-0.015	38.677	0.003
19	0.022	0.026	39.639	0.004
20	0.006	0.004	39.710	0.005

Example: ACF/PACF of squared Returns

- Descriptive
  - What abou returns?

    Usually str elated.
  - Why squared returns?

    $$E(r_t^2) \approx Var(r_t)$$

eg. NYSE Composite

return squared

Correlogram of RC2

Sample: 1 1931
Included observations: 1930

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.179	0.179	62.044	0.000
		2	0.168	0.140	116.47	0.000
		3	0.175	0.131	176.00	0.000
		4	0.109	0.044	198.89	0.000
		5	0.184	0.130	264.80	0.000
		6	0.113	0.035	289.54	0.000
		7	0.136	0.067	325.24	0.000
		8	0.112	0.030	349.60	0.000
		9	0.108	0.039	372.11	0.000
		10	0.095	0.015	389.80	0.000
		11	0.107	0.042	411.90	0.000
		12	0.102	0.028	432.21	0.000
		13	0.066	-0.006	440.72	0.000
		14	0.053	-0.016	446.26	0.000
		15	0.041	-0.016	449.58	0.000
		16	0.097	0.052	467.86	0.000
		17	0.040	-0.018	470.96	0.000
		18	0.068	0.024	479.94	0.000
		19	0.063	0.014	487.77	0.000
		20	0.045	0.005	491.70	0.000

Summary of stylized Facts

KEY stylised facts abou [illegible] eturn series

1. the returns have small, often non-significant autocorrelations (no linear return predictability)
2. the squared returns have strong positive autocorrelations (predictability in volatility, volatility clustering)
3. large kurtosis (heavy tails, tail probabilities larger than normal)

Summary

- Characterizing Financial Time Series:
  - asset price and returns
  - stylised facts about index return series
- Normality tests Z_{ks}, Z_{kt}, JB
- Predictability in returns
  - Autocovariance and autocorrelation
  - Tests for autocorrelation: AC test and Q_m
- Next week: Application of linear regression in Finance (asset pricing)