Quantitative Finance.

2. 条件期望：conditional expectation：

(条件分布的期望)

one random variable takes a specific value

$f_{Y|X}(y|x=2)$

· conditional variance;

3. Expected Value:

$E(X)=\sum P(x_i)x_i$

$\to [E(x)]^2 \neq E(X^2)$

$\to$ Variance: $\sigma^2(x) = E(X^2)-[E(x)]^2$

$\to E(X+Y)=E(X)+E(Y)$

$\to$ generally, $E(XY)\neq E(X)E(Y)$

if X and Y are independent, $E(XY)=E(X)E(Y)$

4. Median: (50% Quantile)

· odd $median(x) = x_{(\frac{n+1}{2})}$

· even $median(x) = \frac{1}{2}[x_{(\frac{n}{2})} + x_{(\frac{n}{2}+1)}]$

5. 25% Quantile, 75% Quantiles:

① sorted

② unit interval $= \frac{100\%}{(n-1)}$

③ Interpretation. (linear)

$\widehat{IQR} = \hat{q}_{75} - \hat{q}_{25}$

inter-quartile range

6. Variance and Moments

二阶矩：数据的离散程度

(measure dispersion)

定义式: $\sigma_X^2 = E(X-\mu)^2 = E[X-E(X)]^2$

计算式: $\sigma^2(x) = E(X^2)-[E(x)]^2$

$\sigma^2(aX+bY) = a^2\sigma^2(X)+b^2\sigma^2(Y)+2\rho a\cdot b\cdot\sigma_x\cdot\sigma_y$

1. Expectation of a function $g(X,Y)$:

$E[g(X,Y)] = \sum\sum g(X,Y) f_{X,Y}(X,Y)$

7. Covariance:

<一个变量随另一变量的变化>

$Covariance = E[(X-E(X))(Y-E(Y))]$

$= E(XY)-E(X)E(Y)$

• 定义式: $E[(X-E(x))(Y-E(Y))]$

• 计算式: $E(XY)-E(X)E(Y)$

• $(-\infty,+\infty)$

• relationship between covariance and variance:

$\to \sigma^2_{X\pm Y} = \sigma_X^2+\sigma_Y^2 + 2Cov(X,Y)$

$\to Cov(a+bX, Y) = Cov(a, cY) + Cov(bX, cY)$

$= b\times c\times Cov(X,Y)$

8. Correlation coefficient:

<标准化后的协方差>

$\rho_{X,Y} = \frac{Cov(X,Y)}{\sigma_x\cdot\sigma_y} = \frac{\sigma_{xy}}{\sigma_x\cdot\sigma_y} = \frac{E[(X-\mu_x)(Y-\mu_Y)]}{\sigma_x\cdot\sigma_y}$

• no unit.

• from $(-1,+1)$

9. Skewness and Moments:

• Skewness 三阶矩

mean < median < mode.

• 对称性:

→ 左偏: →极端大的损失.

右偏: (v) right tail

极端大的收益

mean > median > mode

→ 正偏: positive-skewed.

负偏: Negative-skewed.

• $S = \frac{[E(X-\mu_x)]^3}{\sigma_x^3} \Rightarrow S = \frac{E(X-\mu)^3}{\sigma_x^3}$

10. kurtosis 四阶矩

<陡峭程度 (尾巴薄厚)>

measure of tallness or flatness of a PDF

$$k = \frac{E(X-\mu_x)^4}{[E(X-\mu_x)^2]^2}$$

		Excess
• Leptokurtic	>3	>0
Mesokurtic	$=3$	$=0$
Platykurtic	<3	<0

11. Moments and Linear Transformation:

$Y = a + bX$, a,b constant.

- $E(Y) = E(a+bX) = a + bE(X)$
- $V(Y) = V(a+bX) = b^2V(X) = b^2\sigma^2$
- $b>0$, $\text{Skewness}(X) = \text{Skewness}(Y)$
  $\text{kurtosis}(x) = \text{kurtosis}(x)$

  $b<0$, $\text{Skewness}(X) = -\text{Skewness}(Y)$
  $\text{kurtosis}(x) = \text{kurtosis}(x)$

12. the BLUE mean estimator:

$\begin{cases} \text{estimator: 样本数据特点 } (\bar{x}) \\ \text{parameter: 总体数据特点 } (\mu) \end{cases}$

① Linear

② Unbiased: $E(\bar{x}) = \mu$
$E(S_x^2) = \sigma_x^2$

③ Best: minimum variance (Best)

- The mean estimator is the "Best Linear Unbiased estimator (BLUE) of the population mean when the data are iid.

$\boxed{\text{Summary:}}$

一、Random variable:

①定义: 取值不确定的量

②分类: $\begin{cases} \text{discrete} \\ \text{continuous.} \end{cases}$

③多元: probability matrix

④ marginal / conditional distribution. (条件期望)

⑤ Independence: $P(XY) = P(X)\cdot P(Y)$

二、矩 moment:

① 一阶矩: 中心趋势

1) mean
2) median $\begin{cases} x_{(\frac{n+1}{2})} \\ \frac{1}{2}(x_{(\frac{n}{2})} + x_{(\frac{n}{2}+1)}) \end{cases}$

3) quantile

50% → median
25% → $Q_{25\%}$
75% → $Q_{75\%}$
$IQR = Q_{75\%} - Q_{25\%}$

② 二阶矩: 离散程度

1) Variance: $\begin{cases} Var(x) = E(x-\mu)^2 \\ Var(x) = E(x^2) - E(x)^2 \end{cases}$

2) Covariance: $\begin{cases} Cov(X,Y) = E[(X-E(X))(Y-E(Y))] \\ Cov(X,Y) = E(XY) - E(X)E(Y) \end{cases}$

3) Correlation coefficient:

→ $\rho_{X,Y} = \frac{Cov(X,Y)}{\sigma_x \cdot \sigma_y} = \frac{\sigma_{X,Y}}{\sigma_X \cdot \sigma_Y}$

→ linear relationship.

③ 三阶矩: 对称关系:

1) Skewness

定义
→ 左 left tail. / 右 right tail
→ 正: postive tail (好) 极端大的收益 / 负: negative tail (极端大的损失)

④ 四阶矩：陡峭程度
(尾巴的薄厚)

1) kurtosis：

- 定义 { Leptokurtic (尖) $k>3$
  Mesokurtic (平) $k=3$
  platykurtic (矮) $k<3$
- 性质：

⑤ BLUE 标准：

- An estimator is BLUE means the true parameter can be approached through infinite sampling.
- $\hat{u}$ and S^2 are BLUE
- $\hat{\sigma}^2$ is not BLUE
- { law of large number.
  central limit theory.
- Best Linear Unbiased Estimator
  → Best: minimum variance.
  → Linear: linear function
  → Unbiased: $E(\bar{x})=\mu$
  $E(S_x^2)=\sigma_x^2$

Part 3: Distribution

一、Discrete distribution:

1) Binomial
2) Poisson

二、Continuous distribution:

1) normal.
2) lognormal.
3) t-distribution.
4) Chi-square
5) F-distribution.

三、两个定理：

1) 大数定理：LLN Law of Large Number
2) 中心极限定理：CLT Central Limit Theory.
3) 区别

① Binomial distribution:

1) Bernoulli distribution

$P(X=1)=P$
$P(X=0)=1-P$

2) 独立、重复 N次 Bernoulli:

$P(x)=P(X=x)=C_n^x\,p^x(1-p)^{(n-x)}$

3) 性质：

	Bernoulli	Binomial
Expectation:	p	np
Variance:	$p(1-p)$	$np(1-p)$

② Poisson Distribution:

$$P(k)=P(X=k)=\frac{\lambda^k}{k!}e^{-\lambda} \quad (\lambda=np)$$

1) $n\to\infty$
2) $P\to 0$
3) $np=\lambda$ (constant)
λ = mean = variance
mean = np

- λ means average.

Discrete distribution:

Continuous distribution:

③ Normal distribution: density function:

$$f(x)=\frac{1}{\sqrt{2\pi\sigma^2}}e^{[-\frac{1}{2\sigma^2}(x-\mu)^2]}$$

1) 性质
2) 标准化.
3) confidence interval.

1) 性质：

Skewness = 0, bell-shaped
kurtosis = 3
$X\sim N(\mu,\sigma^2)$
tail get thin and go to zero, but extend infinitely

① $X\sim N(\mu_1,\sigma_1^2)$ $Y\sim N(\mu_2,\sigma_2^2)$ independent.

- $aX+bY\sim N(a\mu_1+b\mu_2,\ a^2\cdot\sigma_1^2+b^2\cdot\sigma_2^2)$
- $\mu_p = w_1\cdot\mu_1 + w_2\cdot\mu_2$
- $\sigma_p^2 = w_1^2\cdot\sigma_1^2 + w_2^2\cdot\sigma_2^2 + 2\rho w_1 w_2 \sigma_1\sigma_2$

  if independent. $\rho=0$

② 标准化: standardization

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

  $Z=\frac{X-\mu}{\sigma}\sim N(0,1)$
- check Z-table to find CDF. $\Phi(\frac{X-\mu}{\sigma})$

③ key quantiles in normal distribution:

$$\begin{cases} 68\% & \mu\pm\sigma \\ 90\% & \mu\pm1.65\sigma \\ 95\% & \mu\pm1.96\sigma \\ 99\% & \mu\pm2.58\sigma \end{cases}$$

④ Approximating discrete random variables to normal random variable.

if $\begin{cases} \cdot\ np\geq10 \\ \cdot\ n(1-p)\geq10 \end{cases}$

⇒ binomial → normal random variable.

if $\lambda\geq1000$

⇒ Poisson → normal random variable $N(\lambda,\lambda)$

正态分布是一切分布的终极

四. t-distribution (矮峰, 肥尾)

1)性质
2)公式
3)查表.

1)性质: • symmetrical
- mean $=0$
- if variance known → normal distribution.

  variance unknown → t-distribution.

• from $\frac{X-\mu}{\sigma}$

↓ to

$\frac{\bar{X}-\mu_x}{S_x/\sqrt{n}}$ ($\frac{S_x}{\sqrt{n}}$ 为$\bar{x}$的标准差)

五. Lognormal distribution:

1)性质:

- $X\sim$ lognormal ↔ $\ln x$ → normal

  $\ln x$ → normal ↔ x → lognormal
- never take negative value.
- Right tail. 

⇒ Right tail $\begin{cases} \cdot\ \text{lognormal} \\ \cdot\ \text{Chi-square} \\ \cdot\ \text{F-distribution.} \end{cases}$ 右偏, >0

六. LLN and CLT

1) LLN: Strong Law of Large Number:

$E[X_i]\equiv\mu$ (恒等) (iid random variable)

⇒ $\hat{\mu}_n=\frac{1}{n}\sum_{i=1}^{n}X_i \xrightarrow{a.s.} \mu$ (收敛于)

- the mean in large sample $\hat{\mu}_n$ $(n\to\infty)$ is approximate the true population mean μ.

2) CLT: Central Limit theorem

中心极限定理

① 于什么: 未知分布 → 正态分布

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

③ 性质:

→ $E(\bar{X}_m)=\mu$

→ $Var(\bar{X}_m)=\frac{\sigma^2}{n}$

→ n large enough. $(n\geq30)$

→ Standard error:

$\sigma_{\bar{X}_m}=\frac{\sigma}{\sqrt{n}}$

样本均值的标准差

3) Compare LLN and CLT

- LLN: mean
- CLT: mean & Variance.

Summary: Distribution { discrete / continous.

一、Discrete

① Binomial
② Possion
} 性质, mean, variance
公式,

二. Continuous

① Normal distribution:
→性质
→标准化
→key value.

② t 分布
→性质: fat tail
→公式: $\frac{\bar{x}-\mu_x}{S_x/\sqrt{n}}$
→查表

③ lognormal:
→性质 • $\ln x \sim$ normal $\Downarrow$ $x \sim$ lognormal • $x \sim$ lognormal $\Downarrow$ $\ln x \sim$ normal
• right-skewed

三. CCL & LLN

① CCL → 目的 → 内容 → 性质
② LLN → 内容
③ 对比

Part 3: Hypothesis Testing

一. Confidence interval:

① spot estimate
② interval estimate

二. Hypothesis testing

① 步骤
② 种类
③ p-value
④ Type I, II, error.

一、Statistical inference:

population $\xrightarrow[\text{sampling}]{}$ sample.

sample statistic $\xrightarrow[\text{estimation}]{}$ population parameter
→mean
→variance
→skew
→kurtosis

二. 点估计: $\bar{x}\to\mu_x$, $S^2\to\sigma_x^2$

三. 区间估计:

confidence interval $=\bar{x}\pm k\cdot S_E$

k → reliability factor.
S_E → standard error.

• α – level of confidence.
• $(1-\alpha)$ – degree of confidence.

① Normal distribution with known variance.

$\bar{x}\pm Z_{(\frac{\alpha}{2})}\cdot\frac{\sigma}{\sqrt{n}}$ (Z-distribution)

② Normal distribution with unknown variance.

$\bar{x}\pm t_{(\frac{\alpha}{2})}\cdot\frac{s}{\sqrt{n}}$

③ Normal Distribution
- Yes → { Var known: Z ; Var unknown: t (df = n-1) }
- No → { $N\geq 30$: { Var known: Z ; Var unknow: t. } ; $N<30$: × }

非正态, 小样本不可估计 → 未知分布

• step 1:

> 原假设: $H_0: \mu=170$ Null

> 备择假设: $H_a: \mu\neq170$ Alternative

> $\begin{cases} \text{one tail: } \begin{cases} H_0: \mu\geq0 & H_a: \mu<0 \\ H_0: \mu\leq0 & H_a: \mu>0 \end{cases} \\ \text{two tails: } H_0: \mu=0 \quad H_a: \mu\neq0. \end{cases}$

• step2: Identify the test statistics

映射到 normal distribution

$$\frac{\bar{x}-\mu}{\sigma/\sqrt{n}} \quad \text{or} \quad \frac{\bar{x}-\mu}{S/\sqrt{n}}$$

• step3: select a level of significance: α

显著性水平

• step4: Formulate a decision rule:

if |test statistic| > critical value ⇒ Reject.

if |test statistic| < critical value ⇒ Fail to reject.

• step5: Arrive at decision.

⇒ significantly different from μ_0.

② Hypothesis test 种类

1> $H_0: \mu=\mu_0$ (1-组均值)

→ z-statistic: $z=\frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}$

→ t-statistic: $t=\frac{\bar{x}-\mu_0}{S/\sqrt{n}}$

2> 两组均值是否相等:

• $H_0: \mu_X=\mu_Y$

$\mu_X-\mu_Y=0$ $(\mu_Z=0)$

• $$T=\frac{(\hat{\mu}_X-\hat{\mu}_Y)-0}{\sqrt{\frac{\sigma_X^2+\sigma_Y^2-2\rho\sigma_X\cdot\sigma_Y}{n}}}=\frac{\hat{\mu}_Z-0}{\sqrt{\hat{\sigma}_Z^2/n}}$$

(因为 $\sigma_Z=\sigma_{X-Y}=\sigma_X^2+\sigma_Y^2-2\rho\sigma_X\cdot\sigma_Y$)

if X, Y iid & mutually independent.

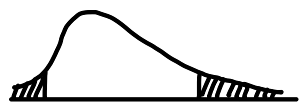
$$T=\frac{\hat{\mu}_X-\hat{\mu}_Y}{\sqrt{\frac{\sigma_X^2}{n_X}+\frac{\sigma_Y^2}{n_Y}}}$$

3> 方差一组是否为常数:

① $H_0: \sigma=$ constant

② Chi-square distribution.

$$\chi^2=\frac{S^2(B)\ (n-1)}{\sigma_0^2(A)}=\frac{(n-1)S^2}{\sigma_0^2} \quad df=n-1$$



(右偏)

4> 两组数据方差是否相关.

• $H_0: \sigma_1^2=\sigma_2^2$

• F-test: $F=\frac{S_1^2}{S_2^2}$ $\begin{cases} df_1=n_1-1 \\ df_2=n_2-1 \end{cases}$

• Always put the large variance in the numerator ($S_1^2>S_2^2$)

大的方差放分子上. $\left(\frac{S_1^2}{S_2^2}>1\right)$

5> Summary of hypothesis Testing:

→ Mean:

① $H_0: \mu=0$ (normal distribution) (known variance)

$$z=\frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}$$

② $H_0: \mu=0$ (normal distribution) (unknown variance)

$$t=\frac{\bar{x}-\mu_0}{S/\sqrt{n}}.$$

③ Two means: $\mu_1=\mu_2$ (t-distribution)

- $T=\dfrac{\hat{\mu}_z}{\sqrt{\dfrac{\sigma_X^2+\sigma_Y^2-2\hat{\sigma}_{XY}}{n}}}$
- $t(n-1)$

④ Variance: $H_0: \sigma^2=\sigma_0^2$

- Chi-square: $\chi^2=\dfrac{(n-1)S^2}{\sigma_0^2}$ $\quad \chi^2(n-1)$

⑤ Variance: $H_0: \sigma_1^2-\sigma_2^2=0$

- $F=S_1^2/S_2^2 \quad F(n_1-1, n_2-1)$

▲ Summary: hypothesis test:

⓪ H_0: H_a:

② $t=\dfrac{\bar{x}-\mu}{S/\sqrt{n}}$

③ $\alpha \rightarrow$ critical value

④ $|t| \gtrless |\text{critical}|$

四. P-value (概率)

- P-value 与 α 相比 (与单双尾无关)
- 越小越拒绝.

五: Type I and Type II Errors:

- Type I: 拒真 $P(\text{type I})$ = significance level (α)
- Type II: 取伪 $1-P(\text{type II})$ = power of a test.

①

	H_0: true	H_0: false
reject	type I, α	power of test
not reject	confidence level	type II $=1-$power of test

② 性质: type I/II 此消彼长

$n\uparrow$, type I, II 都下降.

▲ Summary:

一、Confidence Interval.

① spot estimate.

② 区间估计: 概念:

$\bar{x} \pm k \cdot SE$, 表格 ※ {normal, variance, sample.

二、假设检查:

① 步骤: 1) 2) 3) 4) 5)

② 种类: mean. variance

t. z. χ^2 F

③ P-value: 越小越拒绝.

④ type I / type II: { 1) 定义 2) 性质

Linear Regression.

一、假设 X Y

二、计算参数 a. b. OLS. $Y=a+bX$

三、检验模型：置信区间 假设检验

四、评价模型：R^2. adj.R^2

五、违反假设.（后果）

一、Elements of linear regression.

$Y = b_0 + b_1 X_1 + \underset{\downarrow}{\varepsilon}$

error term（凡是不能由 X 解释 Y 的部分，通通由 error term 解释）

- assumption 中没有一条件与 y 相关.

- Old assumption:
  - → linear relationship between X and Y
  - → X are not random,
    - $\rho_{x,\varepsilon}=0$
    - $\rho_{x_i,x_j}=0$
  - → $E(\varepsilon_i)=0$ $Var(\varepsilon_i)=$ constant.
  - → $\rho_{\varepsilon,\varepsilon}=0$ across observation.
  - → $\varepsilon \sim N(\mu, \sigma^2)$ normal distribution.

- New assumption：(5)
  - → $E(\varepsilon_i)=0$
  - → $Var(\varepsilon_i)=$ constant.
  - → $\sigma_x^2>0$ x 是变量
  - → Large outliers are unlikely.
  - → sample observations are i.i.d.
  - → the explanatory variables are not perfectly linear dependent. ($\rho_{x_i,x_j}\neq 1$)

二、OLS. Ordinary Least Squares：

① 原理：得到的 b_0, b_1, 可以使回归方程的 $\sum \varepsilon_i^2$ minimize $\sum \varepsilon_i^2 = \sum[Y_i-(\hat{b}_0+\hat{b}_1 \times X_1)]^2$

$\hat{b}_1 = \frac{Cov(X,Y)}{\sigma_x^2} = \hat{\rho}_{x,Y} \cdot \frac{\sigma_Y}{\sigma_X}$

$\hat{b}_0 = \overline{Y} - \overline{X}\hat{b}_1$ 代入点 $(\overline{X}, \overline{Y})$

② Confidence interval for the regression coefficient b_1 and b_0

$\hat{b}_1 \pm$ (Critical Value $\times S_{\hat{b}_1}$)

$\hat{b}_0 \pm$ (Critical Value $\times S_{\hat{b}_0}$)

③ Regression coefficient hypothesis testing：

$\hat{b}_1$ 的检验.

1) $H_0: \hat{b}_1 = 0$ $H_a: b_1 \neq 0$

2) test statistic:

$T = \frac{\hat{b}_1 - b_1}{S_{\hat{b}_1}} = \frac{\hat{b}_1}{S_{\hat{b}_1}}$ $\sim$ $S/\sqrt{n}$, $\sigma/\sqrt{n}$.

3) $p\text{-value} = \underset{\downarrow}{2}(1-\Phi|T|)$

two tail. $\Phi(T)$：累计概率

④ Joint hypothesis testing（多元线性回归）

- $H_0: b_1 = b_2 = b_3 \cdots = b_k = 0$
- $F = \frac{ESS/k}{RSS/(n-k-1)}$ E - explained R - residual
- if F(test) > F(critical value) ⇒ reject.

⑥三个 tables $\begin{cases} \text{• Regression statistics} \\ \text{• ANOVA} \\ \text{•} \end{cases}$

→ $R^2 = \frac{ESS}{TSS}$

Multiple $R = \sqrt{R^2} = R$

→ Adjusted R^2 : $1 - \frac{RSS/(n-k-1)}{TSS/(n-1)}$

→ Standard error (regression) $= \sqrt{\frac{RSS}{n-k-1}}$

- Regression / explained $df = k$

Residual $df = n-k-1$

Total $df = n-1$

$F = \frac{ESS/k}{RSS/(n-k-1)}$

三. Measure model fit.

① R^2 (the coefficient of determination)

$R^2 = \frac{ESS}{TSS} = 1 - \frac{RSS}{TSS}$ ↑ X解释y力度大.

→ simple two variable regression:

$r^2 = R^2 \rightarrow r = \pm\sqrt{R^2}$

$\underset{\text{correlation coefficient}}{\underset{\downarrow}{\rho}}{}^2 = R^2$

② Adjusted R^2

- adding a new variable to the model always increases the R^2
- Adj. $R^2 = 1 - \frac{RSS/(n-k-1)}{TSS/(n-1)}$
- R^2 can not be compared across models with different dependent variables.

四、Violation of the assumption.

- Assumption: ① $E(\varepsilon_i) = 0$

② $Var(\varepsilon_i) = constant$. (异方差)

③ $Var(X) > 0$

④ i.i.d.

⑤ no outliers.

* ⑥ $\rho_{x,x} \neq 1$

- conditional heteroskedasticity. 异方差.

1) 定义: $\sigma_\varepsilon^2 \neq constant$

$\begin{cases} \text{• } \sigma_\varepsilon^2 \text{ 与 X 有关} \\ \text{• } \sigma_\varepsilon^2 \text{ 与 X 无关} \end{cases}$

2) 影响:

- $\frac{\hat{b}}{\sqrt{n}}$ ↑ ⇒ Z↓ 更容易 fail to reject.

3) test heteroskedasticity:

- White Test.

- 多重共线性: Multi-collinearity (Multiple-regression only)

① 定义: $Y = b_0 + b_1\underline{X_1} + b_2\underline{X_2} + \varepsilon$

(高度相关)

→ T-test fail to reject.

→ $|r| > 0.7$

→ common sense.

重大遗漏变量: omitted variable Bias

① 定义: $\begin{cases} \text{与 at least one of the included regressors correlated} \\ \text{determinant} \end{cases}$

② implication:

→ the larger the correlation, the larger is the bias.

- Outliers:

① 定义: remove 后, 影响较大,

Cook's distance: $D_j = \frac{\sum_{i=1}^{n}(\hat{Y}_{ij} - \hat{Y}_i)^2}{kS^2}$

- cook's distance > 1 ⇒ outlier.

Summary:

一、假设:老的新的.

二. OLS: $\hat{b}_1 = \frac{Cov(X,Y)}{\sigma_X \cdot \sigma_Y} = \rho \cdot \frac{\sigma_Y}{\sigma_X}$

$\hat{b}_0 = \bar{Y} - \hat{b}_1 \bar{X}$

三. ANOVA Table.

- R^2:拟合程度好坏.
- Adj. R^2:拟合程度好坏的指标.
- Ad $R^2 = 1 - \frac{RSS/(n-k-1)}{TSS/(n-1)}$

$= 1 - \left(\frac{n-1}{n-k-1}\right)(1-R^2)$

四. 假设检测:

① t-test:单个系数检验.

$H_0: b_1 = 0$ (常规)

② F-test:对所有系数检验.

$H_0: b_1 = b_2 = b_3 \cdots = 1$

(ANOVA table)

$F = \frac{ESS/k}{RSS/(n-k-1)}$

五. 违反假设:

① 异方差 $Var_\varepsilon \neq$ constant

- 定义
- 图形
- 影响.
- 检验. white test.

② 多重共线性: 1) 定义 $X_1 . X_2$ 高度相关

2) 方法

3) 矫正:去掉一个X,重新做回归.

③ 重大遗漏变量: ① 定义

④ Outliers:

cook's distance $= \frac{\sum(\hat{Y}_i^{-j} - \hat{Y}_i)^2}{kS^2} > 1$ (outlier)

Time Series

一. Trend

二. seasonality

三. random walk.

四. noise

1. covariance stationary
2. white noise
3. 模型 $\begin{cases} AR \\ MA \\ ARMA \end{cases}$

一. Introduction to time series analysis:

- time series $\xrightarrow[\text{decompose}]{}$
  - trend
  - seasonal component
  - cyclical component (noise)

二. Non-Stationary time series:

- covariance-stationary time series.

mean. var, covariances do not depend on time.

if depend on time ⇒ Non-Stationary

三. Trend.

- Polynomial trends.

1) linear time trend.

mean depends on time.

$E[Y_t] = \delta_0 + \delta_1 t$.

2). Nonlinear trend:

including higher powers of time

$Y_t = \delta_0 + \delta_1 t + \delta_2 t^2 + \cdots + \delta_m t^m + \varepsilon_t$.

- Log-linear trend:

→ constant growth rate.

$\ln Y_t = \delta_0 + \delta_1 t + \varepsilon_t$.

四. Seasonality

- relating the mean of the process to the month or quarter of the year.
- seasonal dummy variables:
  - → $l_{it} = 1$ or 0.
  - → $Y_t = \delta + \gamma_1 l_{1t} + \gamma_2 l_{2t} + \cdots + \gamma_{s-1} l_{s-1t} + \varepsilon_t$
  - → dummy variable trap.

五. Random Walk:

$Y_t = Y_0 + \sum \varepsilon_i$

$V(Y_t) = V(\sum \varepsilon_i)$

$V^2(\varepsilon_1+\varepsilon_2) = \sqrt{V^2(\varepsilon_1) + V^2(\varepsilon_2) + 2\rho V(\varepsilon_1)V(\varepsilon_2)}$

- $V(Y_t) = t\sigma^2$ (Variance is not constant anymore)
- 如何判断 random walk.
  - → 引入 L 滞后算子 (算法)
  - → $Y_t = L Y_{t-1}$
    $L^2 Y_{t-2} = Y_t$
  - → $Y_t(1 - 1.8L + 0.8L^2) = \varepsilon_t$

  if unit roots (单位根, 说明有 random walk)

在AR模型中确定是否平稳?

AR
- AR(1) $\varphi = 1$ (√)
- AR(p) unit roots (√)

- Testing for unit Roots:

  Augmented Dickey-Fuller (ADF) Test.

  1) $\Delta Y_t = \gamma Y_{t-1} + \varepsilon_t$
  2) Null: $H_0: \gamma = 0$. $H_1: \gamma < 0$. (but 查 two-tail)
  3) t-test (two tail)

  if reject H_0, 不是 random walk.

Stationary Time Series

一. Sources of non-stationary:
1. Time trends
2. Seasonalities
3. Unit roots (random walk)

- Stationary time series:
  1) time trend ✗
  2) seasonalities ✗
  3) cyclical component ✓
     - shock
     - memory

- Why is stationary so important?

  <the ability of a model to forecast a time series depends crucially on whether the past resembles future>

  <历史会重演>

  <if probabilistic structure changes over time (not stationary) ⇒ no way to predict.

- Covariance stationary

  (满足协方差平稳)

  → mean and autocovariance 不随时间的改变而改变

  → mean, variance 满足以下条件:
  1) $E[Y_t] = \mu$ for all t
  2) $Cov(Y_t, Y_{t-h}) = \gamma_h$ for all t

     auto covariance.

     only depends on the distance between observation h
  3) $V[Y_t] = \gamma_0 < \infty$ for all t.

     $cov[Y_t, Y_t] = cov[Y_{t-h}, Y_{t-h}]$

二. Autocovariance:

- as the covariance between a stationary time series at different points in time.
- h^{th} autocovariance:

  $\gamma_{t,h} = E[(Y_t - E(Y_t))(Y_{t-h} - E(Y_{t-h}))]$

- $h=0$. $\gamma_{t,0}$ is the variance Y_t:

  $\gamma_{t,0} = E[(Y_t - E(Y_t))^2]$

- 滞后期相同,协方差相同.

三. Autocovariance function:

- A function of h that returns the autocovariance between Y_t and Y_{t-h}:
- $\mathrm{cov}[Y_t, Y_{t-h}] = \mathrm{Cov}[Y_{t+h}, Y_t]$

四、Autocorrelation function (ACF)

$$\rho_h = \frac{\mathrm{Cov}(Y_t, Y_{t-h})}{\sqrt{\mathrm{Var}(Y_t)}\sqrt{\mathrm{Var}(Y_{t-h})}} = \frac{\gamma(h)}{\sqrt{\gamma(0)}\sqrt{\gamma(0)}} = \frac{\gamma(h)}{\gamma(0)}$$

五. Partial Autocorrelation function (PACF)

- measure relationship between Y_t and Y_{t-h} removed the effects of $Y_{t-1} \cdots Y_{t-h+1}$

六. Stationary Time Series — White Noise:

- fundamental building block / simple WN / weak WN

  all the stationary time series models are established on white noise.

  $\varepsilon_t \sim WN(0, 6^2)$

White noise

- 1) mean=0

  2) var=constant

  3) Autocorrelation=0

  $\gamma(h) = \rho(h) = 0 \quad (h \neq 0)$

Covariance stationary

- mean=constant.
- var=constant
- autocovariance=constant

- independent white noise:

  $\varepsilon_t \overset{iid}{\sim} (0, 6^2)$

  ① mean=0

  ② variance=6^2

  ③ independent

- Gaussian white noise (normal white noise)

  ① mean=0

  ② Variance=constant

  ③ $\varepsilon_i \overset{iid}{\sim} N(0, 6^2)$

七. Stationary Time Series - AR. MA. ARMA

- model stationary time series process

① AR(p) model:

autoregressive model

$Y_t = \delta + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + \varepsilon_t$

$\varepsilon_t \sim WN(0, 6^2)$

- detect an AR(p) model to be stationary?
  (是否存在 random walk)

  ① lag polynomial

  引入L(滞后算子) $Y_t = L \cdot Y_{t-1}$

  引入滞后多项式: $1 - \phi_1 L - \phi_2 L^2$

  ②特征等式:

  $(z^2 - \phi z_1 - \phi_2)$

  ③特征根: z_1 and z_2

  if $z_1 = 1 \Rightarrow$ random walk.

Example:

Test the stationarity of an AR(2) model:

$Y_t = 1.4\times Y_{t-1} - 0.45 Y_{t-2} + \varepsilon_t$

① Lag polynomial: (令 $Y_t = L Y_{t-1}$, $Y_t = L^2 Y_{t-2}$)

$1-\phi_1 L - \phi_2 L_2 - \cdots - \phi_p L^p$

⇒ $\phi(L) = 1-1.4L+0.45L^2$

② Let $z = 1/L$

⇒ $\phi(z) = z^2 - 1.4z + 0.45$

③ $\phi(z)=0$ ⇒ $(z-0.9)(z-0.5)=0$

$z_1=0.9$ / $z_2=0.5$

② MA model: moving average model

- the current value of the observed series is expressed as a function of current and lagged unobservable shocks.

- MA(1) model

$Y_t = \mu + e_t + \underbrace{\theta}_{\text{weight}} \underbrace{\varepsilon_{t-1}}_{x} \quad \varepsilon_t \sim WN(0,\sigma^2)$

$Y_t = \theta\varepsilon_{t-1} + \varepsilon_t$

$Y_{t-1} = \theta\varepsilon_{t-2} + \varepsilon_{t-1}$

$Y_{t-2} = \theta\varepsilon_{t-3} + \varepsilon_{t-2}$

有关

(无关)

→ if

$\theta > 0$ persistent (两个连续的数值之间有正向关系)

$\theta < 0$ aggressively mean reverts

※ AR(1): $Y_t = \varphi Y_{t-1} + \varepsilon_t + \mu$

MA(1): $Y_t = \mu + \theta\varepsilon_{t-1} + \varepsilon_t$

For MA(1) model:

1) $E(Y_t)=\mu$

2) $Var(Y_t) = (1+\theta^2)\sigma^2$ (ε_t的 variance)

③ Difference between MA(1) and AR(1)

- MA(1) have a short memory
- AR(1) has a much long memory

④ ARMA model:

Autoregressive moving average (ARMA)

- ARMA(1,1)

$Y_t = \delta + \phi_1 Y_{t-1} + \theta\varepsilon_{t-1} + \varepsilon_t$

mean: $E[Y_t]=\mu=\dfrac{\delta}{1-\phi_1}$

Variance:

$V[Y_t]=\gamma_0=\dfrac{\sigma^2(1+2\phi\theta+\theta^2)}{1-\phi_1^2}$

⑤ Application of AR, MA and ARMA.

→ ACF 与 PCAF → define the characteristic of the time series.

ACF	PACF	Model
gradually decay	Sharp cutoff	AR
Sharp cutoff	gradually decay	MA
gradually decay	gradually decay	ARMA

⑥ Hypothesis test for autocorrelation:

- if ARMA, ACF decay to 0.
- Box-pierce statistic & Ljung-Box

→ χ^2 table, Chi-square

→ $H_0: \rho_1=\rho_2=\cdots=\rho_n=0$

Summary:

time series $\xrightarrow{①}$ 去 trend $\xrightarrow{②}$ 去 seasonality

$\xrightarrow{③}$ 检验是否存在 random walk

if $\xrightarrow{④}$ 无 noise

$\xrightarrow{⑤}$ 有 去掉 random walk $\longrightarrow$ noise

$\xrightarrow{⑥}$ covariance stationary $\begin{cases} \text{mean} \\ \text{variance} \\ \rho(h) \end{cases}$ $\xrightarrow{⑦}$

选择3个模型 $\xrightarrow{⑧}$ $\rho(\tau)$ ACF

$\rho(\tau)$ PACF

表格

	ACF	PACF
AR		
MA		
ARMA		

$$Y_t = \mu + \underline{b_1 t} + \underline{b_2 L_1 + b_3 L_2} + \theta Y_{t-1} + \varphi \varepsilon_{t-1} + \varepsilon_t$$

⑥ Simple return vs. Continuously compounded returns

1) Simple return

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

2) continuous compounded return:

$$r_t = \ln P_t - \ln P_{t-1}$$

3) Relationship: $1 + R_t = e^{r_t}$.

⑦ Measure volatility and risk.

- $\sigma_{annual} = \sqrt{252}\, \sigma_{daily}$
- $\hat{\sigma}^2 = \frac{1}{T} \sum_{i=1}^{T} (r_t - \hat{\mu})^2$

⑧ Distribution of Financial Returns

- many return series are skewed and fat-tailed.

⇒ test normality of a distribution.

- $\begin{cases} \text{the Jarque-Bera test} \\ \text{power laws} \end{cases}$

⑨ The Jarque-Bera Test:

→ H_0: skewness $= 0$ & kurtosis $= 3$

→ $JB = (T-1)\left[\frac{\hat{S}^2}{6} + \frac{(\hat{K}-3)^2}{24}\right]$

→ Chi-square 卡方分布

⑩ Power law:

→ compare tail behaviour is to examine the natural log of the tail probability

→ $\ln P(X > x)$

⑪ Dependence:

→ joint density: $f_{X,Y}(x,y) = f_X(x) \cdot f_Y(y)$

→ linear dependence:

$$\rho_{X,Y} = \frac{Cov(X,Y)}{\sigma_x \cdot \sigma_y}$$

→ 排序的相关系数: linear, nonlinear.

- Rank correlation (Spearman's correlation)
- kendall's τ

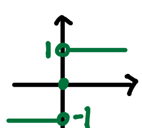
⑫ Rank Correlation — Spearman's correlation:

→ $\rho_s = 1 - \frac{6 \sum_{i=1}^{n} d_i^2}{n(n^2-1)}$

→ from $(-1, 1)$

⑬ Rank Correlation — kendall's τ

sign function:



→ concordant pair (协同组)

$sgn(X_2 - X_1) = sgn(Y_2 - Y_1)$

→ discordant pair

$sgn(X_2 - X_1) = -sgn(Y_2 - Y_1)$

→ $\tau = \dfrac{n_c - n_d}{n(n-1)/2}$

n_c: number of concordant

n_d: number of discordant pairs

→ ordinal correlation are less sensitive to outliers.

Simulation.

1. Monte Carlo simulation.

1) increase the repeating times. N

$\downarrow S_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}\uparrow}$

2) Antithetic variables: negative correlation.

3) control variable.

<use the same draws on a related problem whose solution is known>

2. 抽样：(数据不足)

① Bootstrapping:

1) iid bootstrap

2) circular block bootstrap (CBB)