计量经济学

chapter2 经典单方程计量经济学模型：一元线性回归模型

2.1 回归分析概述

2.1.1 回归分析的基本概念

- 变量X与Y的样本相关系数 ρ_{XY}

$$\rho_{XY}=\frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}}$$

- 协方差 $Cov(X,Y)$

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

- X与Y的一组样本相关系数 r_{XY}

$$r_{XY}=\frac{\sum_{i=1}^{n}(X_i-\overline{X})(Y_i-\overline{Y})}{\sqrt{\sum_{i=1}^{n}(X_i-\overline{X})^2}\cdot\sqrt{\sum_{i=1}^{n}(Y_i-\overline{Y})^2}}$$

2.1.2 总体回归函数

> 根据自变量的给定值，考察因变量的总体均值

- 条件分布：$P(Y=Y_j|X=X_i)$

- 条件均值：$E(Y|X=X_i)$

- （双变量）总体回归线：$E(Y|X)=f(X)$

- **线性总体回归函数PRF：**

$$E(Y|X)=\beta_0+\beta_1 X$$

  $\beta_0\beta_1$**是未知参数，称为回归系数**

2.1.3 随机干扰项

> 不可观测的随机变量

- 离差：对于每个个别样本的观察值Y围绕它的期望值$E(Y|X)$的差值

$$\mu=Y-E(Y|X)$$

- 由（6）式：

$$Y=\mu+E(Y|X)$$

- 在线性假设下：

$$Y=\beta_0+\beta_1 X+\mu$$

  > **确定性部分$E(Y|X)$+随机部分μ**

  (6) (7) 被称为总体回归函数的随机设定形式

- **随机干扰性的原因：**

  > 1. 未知的影响因素
  > 2. 残缺的数据
  > 3. 众多细小影响因素
  > 4. 数据观测误差
  > 5. 模型设定误差
  > 6. 变量的内在随机性

2.1.4 样本回归函数

- 样本回归函数SRF

$$\hat{Y}=f(X)=\hat{\beta}_0+\hat{\beta}_1 X$$

将(8)视为(7)的近似替代，则$\hat{Y}$为$E(Y|X)$的估计量，$\hat{\beta}_0$为β_0的估计量，$\hat{\beta}_1$为β_1的估计量，也有如下随机形式：

$$Y=\hat{Y}+\hat{\mu}=\hat{\beta}_0+\hat{\beta}_1 X+e$$

> e被称为样本残差项，代表其他影响Y的随机因素的集合，可以看作μ的估计量$\hat{\mu}$

- 回归分析的主要目的

  根据样本回归函数即(9)估计(6)

$$\begin{gathered}Y=\hat{Y}+e=\hat{\beta}_0+\hat{\beta}_1 X+e\\ \Rightarrow\\ Y=\mu+E(Y|X)=\beta_0+\beta_1 X+\mu\end{gathered}$$

  设计一种方法构造SRF尽可能接近PRF，使$\hat{\beta}_j$尽可能接近β_j

2.2 一元线性回归模型的基本假设

一元线性回归模型的一般形式：

$$Y=\beta_0+\beta_1 X+\mu$$

其中Y为被解释变量，X为解释变量

β_0与β_1为待估参数，μ为随机干扰项

在有n个样本点$\{(X_i,Y_i):i=1,2,\ldots,n\}$的情况下，也可写成：

$$Y_i=\beta_0+\beta_1 X_i+\mu_i$$

为保证参数估计量有良好的性质，通常对模型提出若干基本假设

2.2.1 对模型设定的假设

假设1　回归模型是被正确设定的

1. 模型设定了正确的变量
   - 满足时，称为模型 没有设定偏差
2. 模型选择了正确的函数形式

2.2.2 对解释变量的假设

假设2　解释变量X在所抽取的样本中具有变异性，而且随着样本容量的无限增加，解释变量X的样本方差趋于一个非零的有限常数

$$\frac{\sum_{i=1}^{n}(X_i-\hat{X})^2}{n}\rightarrow Q,\ when\ n\rightarrow\infty$$

2.2.3 对随机干扰项的假设

假设3　给定解释变量X的任何值，随机干扰项的均值为0

$$E(\mu_i|X)=0$$

> μ的期望不依赖于X的任何观测点取值变化，且总为0
>
> $\rightarrow\mu$与X不存在任何形式的相关性
>
> $\rightarrow$X为外生解释变量（严格外生的）,否则为内生解释变量

只有符合该假设，才有(4)=(7)：

$$E(Y|X)=\beta_0+\beta_1 X==Y=\beta_0+\beta_1 X+\mu$$

满足该假设时，根据期望迭代法则：

$$E(\mu_0) = E(E(\mu_i|X)) = E(0) = 0$$

一定可以得到随机干扰项与解释变量间的不相关性

$$Cov(X, \mu_i) = E(X\mu_i) - E(X)E(\mu_i) = E(X\mu_i) = 0$$

> 任何观测点处的X都与μ_i不相关，也包括第i个点处的X_i与μ_i的不相关性
>
> →X是同期外生的
>
> →X与μ同期不相关

假设4　随机干扰项μ具有给定X任何值条件下的同方差性及序列不相关性

$$Var(\mu_i|X) = \sigma^2,\ i = 1, 2, \ldots, n$$
$$Cov(\mu_i, \mu_j|X) = 0,\ i \neq j$$

> 随机干扰项μ的方差不依赖于X的变化，且总为σ^2

在该假设下：

$$Var(\mu_i|X) = E(\mu_i^2|X) - [E(\mu_i|X)]^2 = E(\mu_i^2|X) = \sigma^2$$
$$Var(\mu_i) = E(\mu_i^2) - [E(\mu_i)]^2 = E(\mu_i^2) = \sigma^2$$

> 给定解释变量的任何值时，任意两个观测点的随机干扰项不想管

(18.2)可被解释为：

$$Cov(\mu_i, \mu_j|X) = E[(\mu_i|X)(\mu_j|X)] = 0$$

假设5　随机干扰项服从正态分布

$$\mu_i|X \sim N(0, \sigma^2)$$

经典线性回归模型

> 服从上述5个假设
>
> 高斯-马尔可夫假设：假设1-4

在上述经典假设下：

$$Y|X \sim N(\beta_0 + \beta_1 X, \sigma^2)$$

2.3 一元线性回归模型的参数估计

在一组样本观测值$(X_i, Y_i) : i = 1, 2, \ldots, n$下，通过一定的参数估计方法，估计出样本回归线

2.3.1 参数估计的普通最小二乘法(OLS)

判断标准：

$$Q = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2 = \sum_{i=1}^{n} (Y_i - (\hat{\beta}_0 + \hat{\beta}_1 X))^2$$

> 在给定样本观测值下，选择$\hat{\beta}_0$与$\hat{\beta}_1$使得Y_i与$\hat{Y}_i$之差的平方和最小

当Q对$\hat{\beta}_0$与$\hat{\beta}_1$的一阶偏导数为0时，Q达到最小，即

$$\begin{cases} \frac{\partial Q}{\partial \hat{\beta}_0} = 0 \\ \frac{\partial Q}{\partial \hat{\beta}_1} = 0 \end{cases}$$

可推得正规方程组：

$$\begin{cases} \sum(Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i) = 0 \\ \sum(Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i)X_i = 0 \end{cases}$$
$$\text{或}$$
$$\begin{cases} \sum Y_i = n\hat{\beta}_0 + \hat{\beta}_1 \sum X_i \\ \sum Y_i X_i = \hat{\beta}_0 \sum X_i + \hat{\beta}_1 \sum X_i^2 \end{cases}$$

解得：

$$\begin{cases} \hat{\beta}_0 = \frac{\sum X_i^2 \sum Y_i - \sum X_i \sum Y_i X_i}{n\sum X_i^2 - (\sum X_i)^2} \\ \hat{\beta}_1 = \frac{n\sum Y_i X_i - \sum Y_i \sum X_i}{n\sum X_i^2 - (\sum X_i)^2} \end{cases}$$

记：

$$\begin{cases} \sum x_i^2 = \sum(X_i - X)^2 \\ \qquad = \sum X_i^2 - \frac{1}{n}(\sum X_i)^2 \\ \sum x_i y_i = \sum(X_i - \overline{X})(Y_i - \overline{Y}) \\ \qquad = \sum X_i Y_i - \frac{1}{n}\sum X_i \sum Y_i \end{cases}$$

则(26)可被记为普通最小二乘法估计量的离差形式：

$$\begin{cases} \hat{\beta}_1 = \frac{\sum x_i y_i}{\sum x_i^2} \\ \hat{\beta}_0 = \overline{Y} - \hat{\beta}_1 \overline{X} \end{cases}$$

记$y = \hat{Y}_i - Y$，则有：

$$\begin{aligned} \hat{y}_i &= (\hat{\beta}_0 + \hat{\beta}_1 X_i) - (\hat{\beta}_0 + \hat{\beta}_1 \overline{X} + \overline{e}) \\ &= \hat{\beta}_1 (X_i - \overline{X}) - \frac{1}{n}\sum e_i \\ &= \hat{\beta}_1 x_i \end{aligned}$$

2.3.2 参数估计的最大似然法(ML)

随机抽取容量为n的样本观测值$\{(X_i, Y_i) : i = 1, 2, \ldots, n\}$，$Y_i$服从正态分布：

$$Y_i \sim N(\beta_0 + \beta_1 X_i, \sigma^2)$$

由上式，可得Y_i的概率函数：

$$P(Y_i) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(Y_i - \beta_0 - \beta_1 X_i)^2}, i = 1, 2, \ldots, n$$

极大似然函数

$$\begin{aligned} L(\beta_0, \beta_1, \sigma^2) &= P(Y_1, Y_2, \ldots, Y_n) \\ &= \frac{1}{(2\pi)^{\frac{\pi}{2}} \sigma^n} e^{-\frac{1}{2\sigma^2}\sum(Y_i - \beta_0 - \beta_1 X_i)^2} \end{aligned}$$

极大似然估计量

使极大似然函数最大化，即可求得模型参数

对极大似然函数进行β_0与β_1求导：

$$\begin{cases} \frac{\partial}{\partial\hat{\beta}_0}\sum(Y_i-\hat{\beta}_0-\hat{\beta}_1X_i)^2=0 \\ \frac{\partial}{\partial\hat{\beta}_1}\sum(Y_i-\hat{\beta}_0-\hat{\beta}_1X_i)^2=0 \end{cases}$$

即：

$$\begin{cases} \hat{\beta}_0=\frac{\sum X_i^2\sum Y_i^2-\sum X_i\sum Y_iX_i}{n\sum X_i^2-(\sum X_i)^2} \\ \hat{\beta}_1=\frac{n\sum X_iY_i-\sum X_i\sum Y_i}{n\sum X_i^2-(\sum X_i)^2} \end{cases}$$

> **因此，在满足基本假设的情况下，模型结构参数的极大似然估计值与最小二乘估计值相等**

2.3.3 参数估计的矩估计法(MN)

> 用相应样本矩来估计总体矩

根据随机干扰项的条件零均值假设：

$$\begin{cases} E(\mu_i)=0 \\ Cov(X_i,\mu_i)=0 \end{cases}$$

可被写为

$$\begin{cases} \frac{1}{n}\sum(Y_i-\hat{\beta}_0-\hat{\beta}_1X_i)=0 \\ \frac{1}{n}\sum(Y_i-\hat{\beta}_0-\hat{\beta}_1X_i)X_i=0 \end{cases}$$

与普通最小二乘法中的**正规方程式**（25）相同：

$$\hat{\beta}_1=\frac{\sum x_iy_i}{\sum x_i^2}$$
$$\hat{\beta}_0=\overline{Y}-\hat{\beta}_1\overline{X}$$

2.4 最小二乘估计量的统计性质

最佳线性无偏估计量满足以下三个性质（**有限样本性质/小样本性质**）

> 1. 线性性：是否是另一个随机变量的线性函数
> 2. 无偏性：均值/期望是否是总体的真实值
> 3. 有效性：它是否在所有线性无偏估计量中具有最小的方差

在有限样本情况下很难找到最佳线性无偏估计量，因此需要考察**样本容量无限增大时估计量的渐进性质**

> 4. 渐进无偏性：样本容量趋于无穷大时，它的均值序列是否趋于总体真值
> 5. 一致性：样本容量趋于无穷大时，它是否依概率收敛于总体的真值
> 6. 渐进有效性：样本容量趋于无穷大时，它在所有的一致估计量中是否具有最小的渐近方差

- 小样本更关注无偏性和有效性，线性性不是必须的
- 大样本更关注一致性

2.4.1 线性性

$$\hat{\beta}_1 = \frac{\sum x_i y_i}{\sum x_i^2} = \frac{\sum x_i(Y_i - \overline{Y})}{\sum X_i^2} = \frac{\sum x_i Y_i}{\sum x_i^2} - \frac{\overline{Y}\sum x_i}{\sum x_i^2} = \sum k_i Y_i$$
$$\text{其中} k_i = \frac{x_i}{\sum x_i^2}$$

$$\hat{\beta}_0 = \overline{Y} - \hat{\beta}_1 \overline{X} = \frac{1}{n}\sum Y_i - \overline{X}\sum k_i Y_i = \sum Y_i(\frac{1}{n} - k_i\overline{X}) = \sum \omega_i Y_i$$
$$\text{其中}\omega_i = \frac{1}{n} - k_i\overline{X}$$

2.4.2 无偏性

> $$E(\hat{\beta}_1|X) = \beta_1$$
> $$E(\hat{\beta}_0|X) = \beta_0$$

由线性性得：

$$\begin{aligned}\hat{\beta}_1 &= \sum k_i Y_i = \sum k_i(\beta_0 + \beta_1 X_i + \mu_i) \\ &= \beta_0 \sum k_i + \beta_1 \sum k_i X_i + \sum k_i \mu_i \\ &= \beta_1 + \sum k_i \mu_i\end{aligned}$$

因此：

$$\therefore E(\hat{\beta}_1|X) = E[(\beta_1 + \sum k_i\mu_i)|X] = \beta_1 + \sum k_i E(\mu_i|X) = \beta_1$$

$$E(\hat{\beta}_0|X) = E[(\beta_0 + \sum \omega_i\mu_i)|X] = \beta_0 + \sum \omega_i E(\mu_i|X) = \beta_0$$

2.4.3 有效性（最小方差性）

> 在所有线性无偏估计量中，普通最小二乘估计量具有最小方差

$$\begin{aligned}Var(\hat{\beta}_1|X) &= Var(\sum k_i Y_i|X) = \sum k_i^2 Var(Y_i|X) \\ &= \sum k_i^2 Var(\beta_0 + \beta_1 X_i + \mu_i|X) \\ &= \sum k_i^2 Var(\mu_i|X) = \sum (\frac{x_i}{\sum x_i^2})^2 \sigma^2 = \frac{\sigma^2}{\sum x_i^2}\end{aligned}$$

$$\begin{aligned}Var(\hat{\beta}_0|X) &= Var(\sum \omega_i Y_i|X) = \sum \omega_i^2 Var(\beta_0 + \beta_1 X_i + \mu_i|X) \\ &= \sum \omega_i^2 Var(\mu_i|X) = \sum (\frac{1}{n} - \overline{X}k_i)^2\sigma^2 \\ &= \sum(\frac{1}{n^2} - \frac{2}{n}\overline{X}k_i + \overline{X}^2 k_i^2)\sigma^2 = (\frac{1}{n} - 2\overline{X}\sum k_i + \overline{X}^2\sum k_i^2)\sigma^2 \\ &= (\frac{1}{n} + \frac{\overline{X}^2}{\sum x_i^2})\sigma^2 = (\frac{n\overline{X}^2 + \sum x_i^2}{n\sum x_i^2})\sigma^2 = \frac{\sum X_i^2}{n\sum x_i^2}\sigma^2\end{aligned}$$

> - 高斯——马尔科夫定理
>
> 普通最小二乘估计量具有线性性、无偏性和有效性

2.5 参数估计量的概率分布及随机干扰项方差的估计

2.5.1 参数估计$\hat{\beta}_0$与$\hat{\beta}_1$的概率分布

> $$\hat{\beta}_0 \sim N(\beta_0, \frac{\sum X_i^2}{n\sum x_i^2}\sigma^2)$$
> $$\hat{\beta}_1 \sim N(\beta_1, \frac{\sigma^2}{\sum x_i^2})$$

$\hat{\beta}_0$与$\hat{\beta}_1$的标准差：

$$\sigma_{\hat{\beta}_0}=\sqrt{\frac{\sum X_i^2}{n\sum x_i^2}\sigma^2}$$
$$\sigma_{\hat{\beta}_1}=\sqrt{\frac{\sigma^2}{\sum x_i^2}}$$

标准差可用来衡量估计量接近其真实值的成图，进而判断估计量的可靠性

2.5.2 随机干扰项μ_i与σ^2的估计

> 由于σ^2在实际上是未知的，因此$\hat{\beta}_0$与$\hat{\beta}_1$的方差实际上不可计算，所以需要对其进行估计
>
> 随机干扰项μ_i不可观测，只能从它的估计——**残差**e_i出发，对总体方差σ^2进行估计
>
> 残差：实际观察值与估计值（拟合值）之间的差

关于σ^2的无偏估计量：（最小二乘估计量）

$$\hat{\sigma}^2=\frac{\sum e_i^2}{n-2}$$

$\hat{\beta}_1$的样本方差：

$$S_{\hat{\beta}_1}^2=\frac{\hat{\sigma}^2}{\sum x_i^2}$$

$\hat{\beta}_1$的样本标准差：

$$S_{\hat{\beta}_1}=\frac{\hat{\sigma}}{\sqrt{\sum x_i^2}}$$

$\hat{\beta}_0$的样本方差：

$$S_{\hat{\beta}_0}^2=\frac{\hat{\sigma}^2\sum X_i^2}{n\sum x_i^2}$$

$\hat{\beta}_0$的样本标准差：

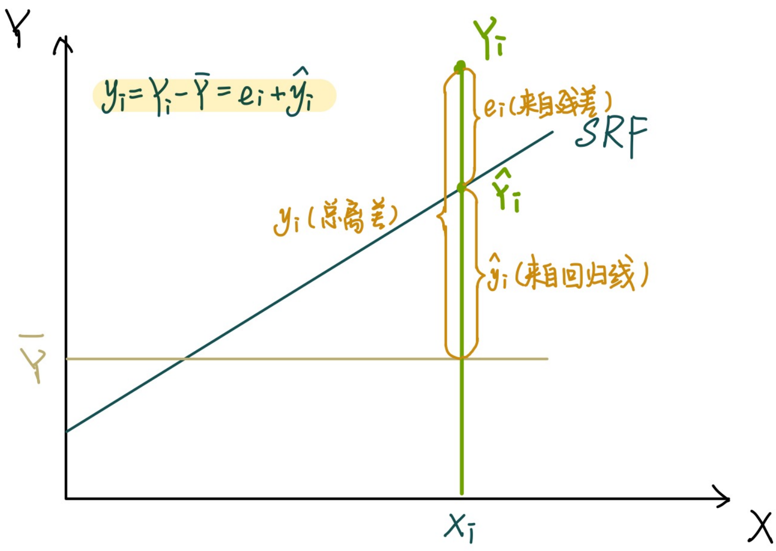
$$S_{\hat{\beta}_0}=\hat{\sigma}\sqrt{\frac{\sum X_i^2}{n\sum x_i^2}}$$

2.4 一元线性回归模型的统计检验

2.4.1 拟合优度检验

> 检验模型对样本观测值的拟合程度

2.4.1.1 总离差平方和的分解



> 如果Y_i落在样本回归线上，则Y的第i个观测值与样本均值间的离差，全部来自样本回归拟合值与样本均值的离差，即完全可由样本回归线解释，表明在该点处实现完全拟合

对于所有样本，考虑**总离差平方和**，反应样本观测值总体离差大小：

$$\sum y_i^2 = \sum(\hat{y_i}^2 + 2\hat{y}_i e_i + e_i^2) = \sum \hat{y_i}^2 + \hat{e_i}^2 = TSS$$

回归平方和反应由模型中解释变量所解释的那部分离差大小：

$$\sum \hat{y_i}^2 = \sum(\hat{Y}_i - \overline{Y})^2 = ESS$$

残差平方和反应样本观测值与估计值偏离的大小，即解释变量未解释的那部分离差大小：

$$\sum e_i^2 = \sum(Y_i - \hat{Y}_i)^2 = RSS$$

2.4.1.2 可决系数R^2统计量

> 可用来自回归线的回归平方和占Y的总离差平方和的比例来判断样本回归线与样本观测值的拟合优度

$$R^2 = \frac{ESS}{TSS} = 1 - \frac{RSS}{TSS}$$

- 如果模型和样本观测值完全拟合，则 $R^2 = 1$
- $0 \leq R^2 \leq 1$
- 随抽样而变动的统计量

$$\because ESS = \sum \hat{y_i}^2 = \sum(\hat{\beta}_1 x_i)^2 = \hat{\beta_1}^2 \sum x_i^2$$
$$\therefore R^2 = \hat{\beta_1}^2 (\frac{\sum x_i^2}{\sum y_i^2})$$

2.4.2 变量的显著性检验

> 对模型中被解释变量与解释变量之间的线性关系是否显著成立的推断
>
> 解释变量是否对被解释变量有显著的先行影响

t检验

已知：

$$\hat{\beta}_1 \sim N(\beta_1, \frac{\sigma^2}{\sum x_i^2})$$

因为真实的σ未知，采用它的无偏估计量$\frac{\sum e_i^2}{n-2}$替代，可构造统计量t：

$$t=\frac{\hat{\beta}_1-\beta_1}{\sqrt{\frac{\hat{\sigma}^2}{\sum x_i^2}}}=\frac{\hat{\beta}_1-\beta_1}{S_{\hat{\beta}_1}}\sim t(n-2)$$

> 如果变量X是显著的，那么参数β_1应该显著的不为零

$$H_0:\beta_1=0,\ H_1:\beta_1\neq 0$$

- 若 $|t|>t_{\frac{\alpha}{2}}(n-2)$，则在 α 显著性水平下拒绝原假设，即 X 是显著的，通过变量显著性测试
- 若 $|t|<t_{\frac{\alpha}{2}}(n-2)$，则在 α 显著性水平下不拒绝原假设，即 X 是不显著的

2.4.3 参数检验的置信区间估计

随机区间：

$$P(\hat{\beta}_j-\delta\leq\beta_j\leq\hat{\beta}_j+\delta)=1-\alpha$$

> 置信区间：$(\hat{\beta}_j-\delta,\hat{\beta}_j+\delta)$
>
> 置信度：$1-\alpha$
>
> 显著性水平：α
>
> 置信限/临界值：$\hat{\beta}_j-\delta,\hat{\beta}_j+\delta$

根据（52）：

$$\begin{gathered}\because P(-t_{\frac{\alpha}{2}}<t<t_{\frac{\alpha}{2}})=1-\alpha\\ \therefore P(-t_{\frac{\alpha}{2}}<\frac{\hat{\beta}_j-\beta_j}{S_{\hat{\beta}_j}}<t_{\frac{\alpha}{2}})=1-\alpha\\ \text{即：}P(\hat{\beta}_j-t_{\frac{\alpha}{2}}\times S_{\hat{\beta}_j}<\beta_j<\hat{\beta}_j+t_{\frac{\alpha}{2}}\times S_{\hat{\beta}_j})=1-\alpha\\ \therefore 1-\alpha\text{的置信度下}\beta_j\text{的置信区间为：}(\hat{\beta}_j-t_{\frac{\alpha}{2}}\times S_{\hat{\beta}_j},\hat{\beta}_j+t_{\frac{\alpha}{2}}\times S_{\hat{\beta}_j})\end{gathered}$$

2.5 一元线性回归分析的应用：预测问题

2.5.1 $\hat{Y_0}$可以作为$E(Y|X=X_0)$与Y_0的预测值

$$\begin{cases}E(\hat{Y}_0)=E(\hat{\beta}_0+\hat{\beta}_1X_0)=E(\hat{\beta}_0)+E(\hat{\beta}_1X_0)=E(\hat{\beta}_0)+X_0E(\hat{\beta}_1)=\beta_0+X_0\beta_1\\ E(Y|X=X_0)=\beta_0+X_0\beta_1\\ E(Y_0)=E(\beta_0+\beta_1X_0+\mu)=\beta_0+\beta_1X_0+E(\mu)=\beta_0+\beta_1X_0\end{cases}$$

2.5.2 总体条件均值与个别值预测值的置信区间

2.5.2.1 总体条件均值预测值的置信区间

$$\begin{gathered}\because \hat{Y}_0=\hat{\beta}_0+\hat{\beta}_1X_0,\text{且}\begin{cases}\hat{\beta}_1\sim N(\beta_1,\frac{\sigma^2}{\sum x_i^2})\\ \hat{\beta}_0\sim N(\beta_0,\frac{\sum X_i^2}{n\sum x_i^2}\sigma^2)\end{cases}\\ \therefore\begin{cases}E(\hat{Y}_0)=\beta_0+\beta_1X_0\\ Var(\hat{Y}_0)=Var(\hat{\beta}_0)+2X_0Cov(\hat{\beta}_0,\hat{\beta}_1)+X_0^2Var(\hat{\beta}_1)\end{cases}\end{gathered}$$

> $$\text{可以证明：}Cov(\hat{\beta}_0,\hat{\beta}_1)=\frac{-\sigma^2\overline{X}}{\sum x_i^2}$$

因此：

$$
\begin{aligned}
\because Var(\hat{Y}_0) &= \frac{\sum X_i^2}{n\sum x_i^2}\sigma^2 - \frac{2\sigma^2\overline{X}X_0}{\sum x_i^2} + \frac{X_0^2\sigma^2}{\sum x_i^2} \\
&= \frac{\sigma^2}{\sum x_i^2}(\frac{\sum X_i^2}{n} - 2X_0\overline{X} + X_0^2) \\
&= \frac{\sigma^2}{\sum x_i^2}(\frac{\sum x_i^2}{n} + (X_0-\overline{X})^2) \\
&= \sigma^2(\frac{1}{n} + \frac{(X_0-\overline{X})^2}{\sum x_i^2}) \\
\therefore \hat{Y}_0 &\sim N[\beta_0+\beta_1X_0, \sigma^2(\frac{1}{n} + \frac{(X_0-\overline{X})^2}{\sum x_i^2})]
\end{aligned}
$$

用σ^2的无偏估计量$\hat{\sigma}^2$替代，构造t统计量：

$$
\begin{aligned}
t &= \frac{\hat{Y}_0 - (\beta_0+\beta_1X_0)}{S_{\hat{Y}_0}} \sim t(n-2) \\
\text{其中，} S_{\hat{Y}_0} &= \sqrt{\hat{\sigma}^2[\frac{1}{n} + \frac{(X_0-\overline{X})^2}{\sum x_i^2}]}
\end{aligned}
$$

在$1-\alpha$的置信度下，总体均值$E(Y|X_0)$的置信区间：

$$\hat{Y}_0 - t_{\frac{\alpha}{2}} \times S_{\hat{Y}_0} < E(Y|X_0) < \hat{Y}_0 + t_{\frac{\alpha}{2}} \times S_{\hat{Y}_0}$$

2.5.2.2 总体个别值预测值的置信区间

> $$
> \begin{aligned}
> \because Y_0 &\sim N(\beta_0+\beta_1X_0, \sigma^2) \\
> \therefore \hat{Y}_0 - Y_0 &\sim N(0, \sigma^2[1 + \frac{1}{n} + \frac{(X_0-\overline{X})^2}{\sum x_i^2}])
> \end{aligned}
> $$

用σ^2的无偏估计量$\hat{\sigma}^2$替代，构造t统计量：

$$
\begin{aligned}
t &= \frac{\hat{Y}_0 - Y_0}{S_{\hat{Y}_0-Y_0}} \sim t(n-2) \\
\text{其中，} S_{\hat{Y}_0-Y_0} &= \sqrt{\hat{\sigma}^2[1 + \frac{1}{n} + \frac{(X_0-\overline{X})^2}{\sum x_i^2}]}
\end{aligned}
$$

在$1-\alpha$的置信度下，个别值 Y_0的置信区间：

$$\hat{Y}_0 - t_{\frac{\alpha}{2}} \times S_{\hat{Y}_0-Y_0} < Y_0 < \hat{Y}_0 + t_{\frac{\alpha}{2}} \times S_{\hat{Y}_0-Y_0}$$

---

chapter3 经典单方程计量经济学模型：多元线性回归模型

3.1 多元线性回归模型

3.1.1 多元线性回归模型的形式

一般形式/总体回归函数的随机表达形式：

$$Y = \beta_0 + \beta_1X_1 + \beta_2X_2 + \cdots + \beta_kX_k + \mu$$

非随机表达式：

$$E(Y|X_1, X_2, \cdots, X_k) = \beta_0 + \beta_1X_1 + \beta_2X_2 + \cdots + \beta_kX_k$$

上式所表示的n个随机方程的举证表达式为$Y = X\beta + \mu$

$$Y=\begin{pmatrix}Y_1\\Y_2\\\vdots\\Y_n\end{pmatrix}_{n\times1},X=\begin{pmatrix}1&X_{11}&X_{12}&\cdots&X_{1k}\\1&X_{11}&X_{22}&\cdots&X_{2k}\\\vdots&\vdots&\vdots&\vdots&\vdots\\1&X_{n1}&X_{n2}&\cdots&X_{nk}\end{pmatrix}_{n\times(k+1)},\beta=\begin{pmatrix}\beta_0\\\beta_1\\\vdots\\\beta_k\end{pmatrix}_{(k+1)\times1},\mu=\begin{pmatrix}\mu_1\\\mu_2\\\vdots\\\mu_n\end{pmatrix}_{n\times1},$$

样本回归函数：

$$\hat{Y}=\hat{\beta}_0+\hat{\beta}_1X_1+\cdots+\hat{\beta}_kX_k=X\hat{\beta}$$

随机表达式：

$$Y=\hat{\beta}_0+\hat{\beta}_1X_1+\cdots+\hat{\beta}_kX_k+e=X\hat{\beta}+e$$

3.1.2 多元线性回归模型的基本假设

- <u>假设1：</u>回归模型是正确设定的
- <u>假设2：</u>解释变量$X_1,X_2,\cdots,X_k$在所抽取的样本中具有变异性，各X_j之间不存在严格线性相关性
  - X 矩阵满秩， $R(X)=k+1$
- <u>假设3：</u>随机干扰项具有条件零均值性

$$E(\mu_i|X_1,X_2,\cdots,X_k)=0$$
$$E(\mu|X)=0$$

> 任何观测点处的随机干扰项μ与任何观测点处各个X都是不相关的，即各X是同期外生的或与μ同期不相关的即
>
> $$E(X_i'\mu_i)=0$$

- <u>假设4：</u>随机干扰项具有条件同方差及序列不相关性

$$\begin{cases}Var(\mu_i|X_1,X_2,\cdots,X_k)=\sigma^2 & i=1,2,\cdots,n\\ Cov(\mu_i,\mu_j|X_1,X_2,\cdots,X_k)=0 & i\neq j,\ i,j=1,2,\cdots,n\end{cases}$$

$$Var(\mu|X)=E(\mu\mu'|X)=E(\begin{matrix}\mu_1^2&\cdots&\mu_1\mu_n\\\vdots&&\vdots\\\mu_n\mu_1&\cdots&\mu_n^2\end{matrix}|X)=\begin{pmatrix}\sigma^2&\cdots&0\\\vdots&&\vdots\\0&\cdots&\sigma^2\end{pmatrix}=\sigma^2I$$

- <u>假设5：</u>随机干扰项满足正态分布

$$\mu_i|X_1,X_2,\cdots,X_k\sim N(0,\sigma^2)$$
$$\mu|X\sim N(0,\sigma^2I)$$

3.2 多元线性回归模型的参数估计

3.2.1 普通最小二乘估计

根据最小二乘原理，参数估计值应使得

$$\begin{aligned}Q&=\Sigma_{i=1}^ne_i^2=\Sigma_{i=1}^n(Y_i-\hat{Y}_i)^2\\&=\Sigma_{i=1}^n(Y_i-(\hat{\beta}_0+\hat{\beta}_1X_{i1}+\hat{\beta}_2X_{i2}+\cdots+\hat{\beta}_kX_{ik}))^2\end{aligned}$$

最小，即可得到待估参数值的正规方程组：

$$\begin{cases}\Sigma(\hat{\beta}_0+\hat{\beta}_1X_{i1}+\hat{\beta}_2X_{i2}+\cdots+\hat{\beta}_kX_{ik})=\Sigma Y_i\\ \Sigma(\hat{\beta}_0+\hat{\beta}_1X_{i1}+\hat{\beta}_2X_{i2}+\cdots+\hat{\beta}_kX_{ik})X_{i1}=\Sigma Y_iX_{i1}\\ \Sigma(\hat{\beta}_0+\hat{\beta}_1X_{i1}+\hat{\beta}_2X_{i2}+\cdots+\hat{\beta}_kX_{ik})X_{i2}=\Sigma Y_iX_{i2}\\ \cdots\cdots\cdots\\ \Sigma(\hat{\beta}_0+\hat{\beta}_1X_{i1}+\hat{\beta}_2X_{i2}+\cdots+\hat{\beta}_kX_{ik})X_{ik}=\Sigma Y_iX_{ik}\end{cases}$$

矩阵形式为：

$$(X'X)\hat{\beta} = X'Y$$

故而

$$\hat{\beta} = (X'X)^{-1}X'Y$$

离差形式的普通最小二乘估计

样本回归模型的离差形式：

$$y = x\beta + e$$

$$y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}_{n\times 1}, x = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1k} \\ x_{11} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nk} \end{pmatrix}_{n\times k}, \hat{\beta} = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_k \end{pmatrix}_{k\times 1}$$

离差形式下参数的最小二乘估计结果：

$$\begin{cases} \hat{\beta} = (x'x)^{-1}x'y \\ \hat{\beta}_0 = \overline{Y} - \hat{\beta}_1\overline{X}_1 - \cdots - \hat{\beta}_k\overline{X}_k \end{cases}$$

随机干扰项μ的方差的普通最小二乘估计

$$\hat{\sigma}^2 = \frac{\Sigma e_i^2}{n-k-1} = \frac{e'e}{n-k-1}$$

3.2.2 最大似然估计

因为

$$Y_i \sim N（X_i\beta, \sigma^2）$$

Y的随机抽取n组样本观测值的联合概率为

$$L(\beta,\sigma^2) = P(Y_1, Y_2, \cdots. Y_n) = \frac{1}{(2\pi)^{\frac{n}{2}}\sigma^n} e^{-\frac{1}{2\sigma^2}(Y-X\beta)'(Y-X\beta)}$$

使概率达到极大值，则可得出一组参数估计量与随机干扰项方差的估计：

$$\begin{cases} \hat{\beta} = & (X'X)^{-1}X'Y \\ \\ \hat{\sigma}^2 = & \frac{\Sigma e_i^2}{n} \end{cases}$$

3.2.3 矩估计*

3.2.4 参数估计量的统计性质

1. 线性性：参数估计量是被解释变量Y的线性组合

$$\hat{\beta} = (X'X)^{-1}X'Y = CY$$

2. 无偏性：

$$\begin{aligned} E(\hat{\beta}|X) =& E((X'X)^{-1}X'Y|X) \\ =& E((X'X)^{-1}X'(X\beta+\mu)|X) \\ =& \beta + (X'X)^{-1}X'E(\mu|X) \\ =& \beta \end{aligned}$$

3. 有效性：

$$Var(\hat{\beta}|X) = \sigma^2(X'X)^{-1}$$

在所有无偏估计量的方差中是最小的

4. 一致性*

3.2.5 样本容量问题

最小样本容量： $n \geq k+1$，样本容量必须不少于模型中解释变量的数目（包括常数项）

满足基本要求的样本容量： 一般经验认为，当$n \geq 30$或者至少$n \geq 3(k+1)$时，才能满足模型估计的基本要求

3.2.6 多元线性回归模型的参数估计实例

3.3 多元线性回归模型的统计检验

3.3.1 拟合优度检验

可决系数

$$R^2 = \frac{ESS}{TSS} = 1 - \frac{RSS}{TSS}$$

其中

$$\begin{array}{ll} \text{总离差平方和} & TSS = \Sigma(Y_i - \overline{Y})^2 \sim (n-1) \\ \text{回归平方和} & ESS = \Sigma(\hat{Y}_i - \overline{Y})^2 \sim (k) \\ \text{残差平方和} & RSS = \Sigma(Y_i - \hat{Y}_i)^2 \sim (n-k-1) \end{array}$$

且有

$$TSS = ESS + RSS$$

> 可决系数越接近1，模型的拟合度越高

调整的可决系数： 剔除变量个数对拟合优度的影响

$$\overline{R}^2 = 1 - \frac{RSS/(n-k-1)}{TSS/(n-1)} = 1 - (1-R^2)\frac{n-1}{n-k-1}$$

赤池信息准则和施瓦茨准则

$$\begin{aligned} AIC &= \ln\frac{e'e}{n} + \frac{2(k+1)}{n} + 1 + \ln(2\pi) \\ SC &= \ln\frac{e'e}{n} + \frac{k+1}{n}\ln n + 1 + \ln(2\pi) \end{aligned}$$

仅当所增加的解释变量能够减少AIC值或者SC值时才在原模型中增加该解释变量

3.3.2 方程总体线性的显著性检验（F检验）

方程显著性的F检验： 检验参数$\beta_0, \cdots, \beta_k$显著不为0

$$H_0: \beta_1 = 0, \beta_2 = 0, \cdots, \beta_k = 0$$

$$H_1: \beta_j\text{不全为}0$$

统计量

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

当$F > F_\alpha(k, n-k-1)$时拒绝原假设H_0

关于拟合优度检验与方程总体现行的显著性检验关系的讨论：

$$\bar{R}^2 = 1 - \frac{n-1}{n-k-1+kF}$$

$$F = \frac{R^2/k}{(1-R^2)/(n-k-1)}$$

> F、R^2同向变化：
>
> $R^2 = 0 \leftrightarrow F = 1$
>
> $R^2 = 1 \leftrightarrow F = +\infty$
>
> $H_0 : \beta_1 = 0, \beta_2 = 0, \cdots, \beta_k = 0 \Leftrightarrow R^2 = 0$

在应用中不必对$\bar{R}^2$过于苛求，重要的是需考察模型的经济关系是否合理

3.3.3 变量的显著性检验（t检验）

t统计量：记c_{jj}为矩阵$(X'X)^{-1}$主对角线上的第j个元素：

$$Var(\hat{\beta}_j) = \sigma^2 c_{jj}$$

通过$\hat{\beta}_j \sim N(\beta_j, \sigma^2 c_{jj})$，可以构造如下t统计量

$$t = \frac{\hat{\beta}_j - \beta_j}{S_{\hat{\beta}_j}} = \frac{\hat{\beta}_j - \beta_j}{\sqrt{c_{jj}\frac{e'e}{n-k-1}}} \sim t(n-k-1)$$

t检验

$$H_0 : \beta_j = 0$$

$$H_1 : \beta_j \neq 0$$

> 在一元线性回归中，t检验与F检验时一致的

$|t| > t_{\frac{\alpha}{2}}(n-k-1)$时拒绝原假设，从而判定对应的解释变量是否应被包含在模型中

3.3.4 参数的置信区间估计

在$1-\alpha$的置信度下β_j的置信区间为

$$(\hat{\beta}_j - t_{\frac{\alpha}{2}} \times S_{\hat{\beta}_j}, \hat{\beta}_j + t_{\frac{\alpha}{2}} \times S_{\hat{\beta}_j})$$

缩小置信区间的方法：

1. 增大样本容量n
2. 提高模型的拟合优度以减少残差平方和 $e'e$
3. 提高样本观测值的分散度以降低 c_{jj}

3.4 多元线性回归模型的预测

3.4.1 E(Y0)的置信区间

$1-\alpha$的置信度下$E(Y_0)$的置信区间：

$$\hat{Y}_0 - t_{\frac{\alpha}{2}} \times \hat{\sigma}\sqrt{X_0(X'X)^{-1}X_0'} < E(Y_0) < \hat{Y}_0 + t_{\frac{\alpha}{2}} \times \hat{\sigma}\sqrt{X_0(X'X)^{-1}X_0'}$$

3.4.2 Y0的置信区间

$1-\alpha$的置信度下Y_0的置信区间：

$$\hat{Y}_0 - t_{\frac{\alpha}{2}} \times \hat{\sigma}\sqrt{1 + X_0(X'X)^{-1}X_0'} < Y_0 < \hat{Y}_0 + t_{\frac{\alpha}{2}} \times \hat{\sigma}\sqrt{1 + X_0(X'X)^{-1}X_0'}$$

3.5 可化为线性的多元非线性回归模型

3.5.1 模型的类型与变换

倒数模型、多项式模型与变量的直接置换法

进行变量替换

幂函数模型、指数函数模型与函数变换法

取对数后进行变量替换

复杂函数模型与级数展开法

3.5.2 可化为线性的非线性回归实例

3.6 含有虚拟变量的多元线性回归模型

3.6.1 含有虚拟变量的模型

一些影响经济变量的因素无法定量度量，构造只取0、1的人工变量，通常被称为虚拟变量，记为D

基础类型和肯定类型取值为1；比较类型和否定类型取值为0

3.6.2 虚拟变量的引入

加法方式：可以考察截距的不同

$$Y_i = \beta_0 + \beta_1 X_i + \beta_2 D_1 + \beta_3 D_2 + \mu_i$$

乘法方式：测度斜率的变化

$$Y_i = \beta_0 + \beta_1 X_i + \beta_2 D_i X_i + \mu_i$$

3.6.3 虚拟变量的设置原则

每一定性变量所需的虚拟变量个数要比类别少1，否则会出现虚拟变量陷阱

3.7 受约束回归

3.7.1 模型参数的线性约束

F检验

在统一数据样本下，记无约束样本回归模型的矩阵式为：

$$Y = X\hat{\beta} + e$$

受约束样本回归模型的矩阵式为：

$$Y = X\hat{\beta}_* + e_*$$

判断约束条件是否为真的检验量：

$$F = \frac{(RSS_R - RSS_U)/(k_U - k_R)}{RSS_U/(n - k_U - 1)} \sim F(k_U - k_R, n - k_U - 1)$$

k为回归模型中解释变量的个数（不包括常数项）

若$F < F_{\frac{\alpha}{2}}(k_U - k_R, n - k_U - 1)$，则不拒绝约束条件

3.7.2 对回归模型增加或减少解释变量

考虑两个回归模型：

$$\begin{aligned} Y &= \beta_0 + \beta_1 X_1 + \cdots + \beta_k X_k + \mu \\ Y &= \beta_0 + \beta_1 X_1 + \cdots + \beta_{k+q} X_{k+q} + \mu \end{aligned}$$

$$H_0: \beta_{k+1} = \cdots = \beta_q = 0$$

相应的F统计量：

$$F=\frac{(RSS_R-RSS_U)/q}{RSS_U/(n-(k+q-1))}\sim F(q,n-(k+q+1))$$

3.7.3 检验不同组之间回归函数的差异

邹氏参数稳定检验

F统计量：

$$F=\frac{[RSS_R-(RSS_1+RSS_2)]/(k+1)}{(RSS_1+RSS_2)/[n_1+n_2-2(k+1)]}\sim F[k+1,n_1+n_2-2(k+1)]$$

3.7.4 非线性约束

最大似然比检验LR

$$LR=-2[\ln L(\tilde{\beta},\tilde{\sigma}^2)-\ln L(\hat{\beta},\hat{\sigma}^2)]\sim\chi^2(h)$$

h为约束条件的个数

沃尔德检验WD：只估计无约束模型

若对模型$Y=\beta_0+\beta_1X_1+\cdots+\beta_kX_k+\mu$检验$\beta_1+\beta_2=1$的约束，可建立沃尔德统计量：

$$W=\frac{(\hat{\beta}_1+\hat{\beta}_2-1)^2}{\tilde{\sigma}^2_{\hat{\beta}_1+\hat{\beta}_2}}\sim\chi^2(1)$$

对于检验非线性约束$\beta_1\beta_2=1$：

$$W=\frac{(\hat{\beta}_1\hat{\beta}_2-1)^2}{\tilde{\sigma}^2_{\hat{\beta}_1\hat{\beta}_2}}\sim\chi^2(1)$$

拉格朗日乘数检验LM：只估计受约束模型

$$LM=nR^2\sim\chi(h)$$

其中h为约束条件个数，R^2为辅助回归的可决系数：

$$e_R=\delta_0+\delta_1X_1+\cdots+\delta_kX_k$$

一般有$LM\le LR\le W$

---

chapter4 经典单方程计量经济学模型：放宽基本假定的模型

4.1 多重共线性

4.1.1 多重共线性的含义

某两个或多个解释变量之间出现了相关性，存在某个解释变量可以由其他解释变量代替

对于模型$Y_i=\beta_0+\beta_1X_{i1}+\cdots+\beta_kX_{ik}+\mu_i$：

完全共线性：存在c_i不全为0，使

$$c_1X_{i1}+c_2X_{i2}+\cdots+c_kX_{ik}=0$$

近似共线性：存在c_i不全为0、随机干扰项v_i，使

$$c_1X_{i1}+c_2X_{i2}+\cdots+c_kX_{ik}+v_i=0$$

4.1.2 实际经济问题中的多重共线性

1. 经济变量相关的共同趋势
2. 模型设定不严谨
3. 样本资料的限制

4.1.3 多重共线性的后果

1. **完全共线性下参数估计量不存在**

   $\hat{\beta} = (X'X)^{-1}X'Y$在出现参数完全共线性时时不存在的

2. **近似共线性下普通最小二乘法参数估计量的方差变大**

   $Var(\hat{\beta}) = \sigma^2(X'X)^{-1}$，此时$(X'X) \approx 0$，使得参数估计量的方差增大

   方差膨胀因子VIF

$$r^2 = \frac{(\Sigma x_{i1}x_{i2})^2}{\Sigma x_{i1}^2 \Sigma x_{i2}^2}$$
$$VIF = \frac{1}{1-r^2}$$

是否共线性	$Var(\hat{\beta}_1)$
完全不共线性	$\frac{\sigma^2}{\Sigma x_{i1}^2}$
近似共线性	$\frac{\sigma^2}{\Sigma x_{i1}^2} \cdot \frac{1}{1-r^2}$
完全共线性	$+\infty$

3. **参数估计量经济意义不合理**
4. **变量的显著性检验和模型的预测功能失去意义**

4.1.4 多重共线性的检验

多重共线性的存在性

1. 对两个解释变量的模型：求出 X_1、X_2 的相关系数 r，如果 $|r| \to 1$ 则说明两变量存在较强的多重共线性
2. 对多个解释变量的模型：若在普通最小二乘法下，模型的 R^2、F 值均较大，但各参数估计的 t 检验值较小，说明各解释变量对Y的联合线性作用显著，但各变量间存在共线性使得他们对于Y的独立作用不能分辨

判明存在多重共线性的范围

1. 判定系数检验

   - $$F_j = \frac{R_j/(k-1)}{(1-R_j^2)/(n-k)} \sim F(k-1, n-k)$$

     R_j为第j个解释变量对其他解释变量的回归方程的决定系数

     若存在较强的共线性，则$R_j \to 1$，$F_j \to +\infty$。原假设设定为X_j与其他解释变量间不存在显著的线性关系

   - 在模型中排除某个解释变量后，拟合优度改变不是很明显，则说明该解释变量与其他解释变量之间存在共线性

2. 逐步回归法

   - 以Y为被解释变量，逐个引入解释变量，构成回归模型
   - 如果拟合优度变化显著，则说明新引入的变量是一个独立解释变量

4.1.5 克服多重共线性的方法

1. 第一类方法：排除引起共线性的变量
2. 第二类方法：减小参数估计量的方差

多重共线性是一种样本现象，同一个模型在一个样本中出现多重共线性，但在另一个样本中可能不存在

只要方程估计的参数标准差较小，t统计值较大，就没有必要过于担心该问题

4.2 异方差性

对于不同的样本点，随机干扰项的方差不再是常数，而是互不相同

4.2.1 异方差的类型

1. 单调递增型：σ_i^2 随 X 的增大而增大
2. 单调递减型：σ_i^2 随 X 的增大而减小
3. 复杂型：σ_i^2 随 X 的增大呈复杂形式

4.2.3 异方差性的后果

1. 参数估计量非有效

   模型出现异方差性时，普通最小二乘法参数估计量依然具有线性性、无偏性，但失去有效性

2. 变量的显著性检验失去意义

   在有偏误的方差基础上构造的t统计量不再服从真实的t分布，t检验失去意义

3. 模型的预测失效

4.2.4 异方差性的检验

检验随机干扰项方差与解释变量观测值之间的相关性

1. **布罗施-帕甘检验Breusch-Pagan(B-P检验)**

   针对辅助回归

$$e_i^2 = \delta_0 + \delta_1 X_{i1} + \cdots + \delta_k X_{ik} + \epsilon_i$$

   检验同方差性即检验

$$H_0 : \delta_0 = \delta_1 = \delta_2 = \cdots = \delta_k = 0$$

   可以利用F检验或LM检验：

$$F = \frac{R_{e^2}^2 / k}{(1 - R_{e^2}^2)/(n-k-1)} \sim F(k, n-k-1)$$

$$LM = n \cdot R_{e^2}^2 \sim \chi(k)$$

2. 怀特检验White

   相比B-P检验引入解释变量的更高次方

4.2.5 异方差的修正

加权最小二乘法WLS

对原模型加权，使之变成一个新的不存在异方差性的模型，然后采用普通最小二乘法估计其参数

对较小的残差平方e_i^2赋予较大权数，对较大的残差平方e_i^2赋予较小权数

> 若$Var(\mu_i) = f(X_{ji})\sigma^2$
>
> 则构建新模型
> $\frac{1}{\sqrt{f(X_{ji})}} Y_i = \beta_0 \frac{1}{\sqrt{f(X_{ji})}} + \beta_1 \frac{1}{\sqrt{f(X_{ji})}} X_{1i} + \cdots + \beta_k \frac{1}{\sqrt{f(X_{ji})}} X_{ki} + \frac{1}{\sqrt{f(X_{ji})}} \mu_i$

加权最小二乘法又称广义最小二乘法GLS

若有

$$Var(\mu_i|X_{i1}, X_{i2}, \cdots, X_{ik}) = f(X_{i1}, X_{i2}, \cdots, X_{ik})\sigma^2$$

则加权最小二乘法中的权即为$\frac{1}{\sqrt{f(X_{i1}, X_{i2}, \cdots, X_{ik})}}$

异方差稳健标准误法

用普通最小二乘法估计的残差平方e_i^2代替σ_i^2

$$Var(\hat{\beta}_1) = \frac{\Sigma x_i^2\sigma_i^2}{(\Sigma x_i^2)^2} \Rightarrow Var(\hat{\beta}_1) = \frac{\Sigma x_i^2 e_i^2}{(\Sigma x_i^2)^2}$$

4.3 内生解释变量问题

4.3.1 内生解释变量问题的提出

1. 内生解释变量与随机干扰项同期无关但异期相关

$$Cov(X_{i2}, \mu_{i-s}) = E(X_{i2}\mu_{i-s}) \neq 0$$

2. 内生解释变量与随机干扰项同期相关

$$Cov(X_{i2}, \mu_i) = E(X_{i2}\mu_i) \neq 0$$

4.3.3 内生解释变量的后果

参数估计量有偏且不具有一致性

4.3.4 工具变量法

1. 工具变量Z的选取
   - 与所替代的随机解释变量高度相关： $Cov(Z, X_j) \neq 0$
   - 与随机干扰项不相关： $Cov(Z, \mu) = 0$
   - 与模型中其他解释变量不高度相关，以避免出现严重的多重共线性
2. 获得工具变量法估计量

$$\tilde{\beta} = (Z'X)^{-1}Z'Y$$

3. 工具变量法是一致有偏估计量
4. 两阶段最小二乘法

   工具变量并没改变原模型，只是在原模型的参数估计过程中用工具变量的信息替代了内生解释变量的信息

   step1：用普通最小二乘法进行X关于工具变量Z的回归：

$$\hat{X}_i = \hat{\alpha}_0 + \hat{\alpha}_1 Z_i$$

   step2：用第一步得到的$\hat{X}_i$为解释变量，进行普通最小二乘回归：

$$Y_i = \beta_0 + \beta_1\hat{X}_i + \mu_i X_i + \mu_i$$

4.3.5 内生性检验与过度识别约束检验

1. 内生性检验

   对于模型$Y_i = \beta_0 + \beta_1 X_i + \beta_2 Z_{i1} + \mu_i$怀疑$X$是同期内生变量

   step1：将X关于外生变量Z_1、Z_2做普通最小二乘估计：

$$X_i = \alpha_0 + \alpha_1 Z_{i1} + \alpha_2 Z_{i2} + v_i$$

   得到残差项$\hat{v}_i$

   step2：将$\hat{v}_i$加入到原模型中，在进行普通最小二乘估计：

$$Y_i = \beta_0 + \beta_1 X_i + \beta_2 Z_{i1} + \delta\hat{v}_i + \epsilon_i$$

   如果$\hat{v}_i$前参数δ显著为0，则X与μ同期无关

2. 过度识别约束检验

对工具变量的外生性进行检验

对有一个内生变量X与一个外生变量Z的二元线性回归模型

$$Y_i = \beta_0 + \beta_1 X_i + \beta_2 Z_i + \mu_i$$

如果内生变量X找到两个工具变量Z_1、Z_2，记两阶段最小二乘回归2SLS的参数估计为$\tilde{\beta}_0$、$\tilde{\beta}_1$、$\tilde{\beta}_2$，残差为$\tilde{\mu}_i$

$$\tilde{\mu}_i = Y_i - (\tilde{\beta}_0 + \tilde{\beta}_1 X_i + \tilde{\beta}_2 Z_i)$$

将$\tilde{\mu}_i$关于所有工具变量Z_1、Z_2及原模型中的外生变量做辅助回归

$$\tilde{\mu}_i = \delta_0 + \delta_1 Z_{i1} + \delta_2 Z_{i2} + \delta_3 Z_i + \epsilon_i$$

记该辅助回归的可决系数为R^2

$$H_0 : Z_1\text{、}Z_2\text{同时为外生变量}$$
$$J = nR^2 \sim \chi^2(2-1)$$

- 当工具变量的个数恰好等于内生变量的个数时，无法检验工具变量的外生性

4.4 模型设定偏误问题

4.4.1 模型设定偏误的类型

1. 相关变量的遗漏
2. 无关变量的误选
3. 错误的函数形式

4.4.2 模型设定偏误的后果

1. 遗漏相关变量偏误
    1. 漏掉的 X_2 与 X_1 相关，结果在小样本下有偏，在大样本下非一致
    2. 漏掉的 X_2 与 X_1 不相关，则常数项系数 α_0 的估计有偏且非一致， α_1 的估计无偏一致
    3. 随机干扰项的方差估计 $\hat{\sigma}^2$ 有偏
    4. α_1 是正确参数 β_1 的有偏估计量
2. 包含无关变量偏误

    普通最小二乘估计量仍是无偏一致的，σ^2 也能被正确估计，但是估计量往往是无效的

    错误模型的普通最小二乘估计量的方差一般较大
3. 错误函数的偏误

    两者具有完全不同的经济意义

4.4.3 模型设定偏误的检验

1. 检验是否含有无关变量：其系数是否显著为0
2. 检验是否有相关变量的遗漏或函数形式设定偏误
    - 残差图示法
        - 残差序列随某解释变量持续上升：遗漏与该解释变量正相关的变量
        - 残差序列随某解释变量持续下降：遗漏与该解释变量负相关的变量
    - 一般性设定偏误检验RESET检验
        - 先估计$\hat{Y} = \hat{\alpha}_0 + \hat{\alpha}_1 X_1$
        - 用$\hat{Y}$的若干次幂来替代变量

        $$Y = \beta_0 + \beta_1 X_1 + \gamma_1 \hat{Y}^2 + \gamma_2 \hat{Y}^3 + \mu$$

        再判断γ的参数是否显著不为零

采用高次幂的形式既可以避免自由度的损失和多重共线性的问题，同时它已包含解释变量的高次幂及交叉项提供的信息

chapter5 时间序列计量经济学模型

5.1 时间序列模型的序列相关性

5.1.1 序列相关性

对于模型

$$Y_t = \beta_0 + \beta_1 X_{t1} + \cdots + \beta_k X_{tk} + \mu_t$$

随机干扰项序列相关则意味着

$$Cov(\mu_i, \mu_j) = E(\mu_i \mu_j) \neq 0 或 Var(\mu) \neq \sigma^2 I$$

一阶序列相关/自相关

$$\mu_t = \rho \mu_{t-1} + \epsilon_t, 1- \leq \rho \leq 1$$

其中ρ为自协方差系数/一阶自相关系数，ϵ_i是满足普通最小二乘法假定的随机干扰项：

$$\begin{aligned} &E(\epsilon_t) = 0 \\ &Var(\epsilon_t) = \sigma^2 \\ &Cov(\epsilon_t, \epsilon_{t-s}) = 0 \end{aligned}$$

5.1.2 实际经济问题中的序列相关性

1. 经济变量固有的惯性
2. 模型设定的偏误
3. 数据的“编造”

5.1.3 序列相关性的后果

1. 参数估计量非有效，依然具有线性无偏性但不再具有渐进有效性
2. 变量的显著性检验失去意义，估计的参数方差 $S_{\hat{\beta}_j}$ 出现偏误
3. 模型的预测失效

5.1.4 序列相关性的检验

采用普通最小二乘法估计模型，求得残差序列e_t

$$\tilde{e}_t = Y_t - (\hat{Y}_t)_{OLS}$$

1. 图示法
2. 回归检验法

   以e_t为被解释变量，各自可能的相关量如e_{t-1}、e_{t-2}、e_t^2等作为解释变量，建立各种方程

   $$e_t = \rho_1 e_{t-1} + \rho_2 e_{t-2} + \epsilon_t$$

   对方程进行估计并进行显著性检验

   优点：一旦确定了模型存在序列相关性，也就同时知道了相关的形式，适用于任何类型的序列相关性问题检验
3. D.W. 检验

   条件：

   - 解释变量X非随机
   - 随机干扰项μ_t为一阶自回归形式

$$\mu_t = \rho\mu_{t-1} + \epsilon_t$$

- 回归模型中不含有滞后应变量作为解释模型，即不出现Y_{t-1}
- 回归模型含有截距项

原假设：

$$H_0 : \rho = 0$$

过程：

1. 根据样本容量T和解释变量数目k查询D.W.分布表，获得临界值d_L、d_U
2. 对比D.W.值：
   - $0 < D.W. < d_L$，存在正自相关
   - $d_L < D.W. < d_U$，不能确定
   - $d_U < D.W. < 4-d_U$，不存在自相关
   - $4-d_U < D.W. < 4-d_L$，不能确定
   - $4-d_L < D.W. < 4$，存在负自相关
3. $$D.W. = \frac{\Sigma_{t=2}^T e_t^2 + \Sigma_{t=2}^T e_{t-1}^2 - 2\Sigma_{t=2}^T e_t e_{t-1}}{\Sigma_{t=1}^T e_t^2} \approx 2(1 - \frac{\Sigma_{t=2}^T e_t e_{t-1}}{\Sigma_{t=2}^T e_t^2}) \approx 2(1-\rho)$$

缺陷：存在一个不能确定的D.W.区域，且只能检验一阶自相关，对存在滞后被解释变量的模型无法检验

4. 拉格朗日乘数检验LM

   克服D.W.缺陷

   怀疑随机干扰项存在p阶序列相关：

$$\mu_t = \rho_1\mu_{t-1} + \rho_2\mu_{t-2} + \cdots + \rho_p\mu_{t-p} + \epsilon_t$$

   可以使用LM检验如下受约束回归方程：

$$Y_t = \beta_0 + \beta_1 X_{t1} + \cdots + \beta_k X_{tk} + (\rho_1\mu_{t-1} + \rho_2\mu_{t-2} + \cdots + \rho_p\mu_{t-p} + \epsilon_t)$$

   约束条件

$$H_0 : \rho_1 = \rho_2 = \cdots = \rho_p = 0$$

   进行辅助回归：

$$\tilde{e}_t = \beta_0 + \beta_1 X_{t1} + \cdots + \beta_k X_{tk} + \rho_1\tilde{e}_{t-1} + \cdots + \rho_p\tilde{e}_{t-p} + \epsilon_t$$

   获得辅助回归的样本容量$T-p$与可决系数R^2:

$$LM = nR^2 \sim \chi^2(p)$$

5.1.5 序列相关的补救

广义最小二乘法

$$\hat{\beta}_* = (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}Y$$

$$\Omega^{-1} = \frac{1}{1-\rho^2}\begin{pmatrix} 1 & -\rho & 0 & \cdots & 0 & 0 & 0 \\ -\rho & 1+\rho^2 & -\rho & \cdots & 0 & 0 & 0 \\ 0 & -\rho & 1+\rho^2 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1+\rho^2 & -\rho & 0 \\ 0 & 0 & 0 & \cdots & -\rho & 1+\rho^2 & -\rho \\ 0 & 0 & 0 & \cdots & 0 & -\rho & 1 \end{pmatrix}$$

其中ρ可从下式中得到

$$\mu_t = \rho\mu_{t-1} + \epsilon_t$$

满足

$$
\begin{aligned}
Var(\mu_t) &= \sigma^2 \\
Cov(\mu_t, \mu_{t-s}) &= \rho^s \sigma^2
\end{aligned}
$$

广义差分法

将原模型变化为满足普通最小二乘法的差分模型，再进行普通最小二乘估计

本质就是广义最小二乘法，但损失部分样本观测值

若原模型存在

$$\mu_t = \rho\mu_{t-1} + \epsilon_t$$

可对模型进行变换为

$$Y_t - \rho Y_{t-1} = \beta_0(1-\rho) + \beta_1(X_{t,1} - \rho X_{t-1,1}) + \epsilon_t$$

普莱斯-温斯特变换

$$
\begin{aligned}
Y_1^* &= \sqrt{1-\rho^2}Y_1 \\
X_{1,j}^* &= \sqrt{1-\rho^2}X_{1,j}
\end{aligned}
$$

进行样本补充后，广义差分法估计结果与广义最小二乘法完全一致

随机干扰项相关系数的估计

科克伦-奥科特迭代法

估计量具有渐进有效性

序列相关稳健标准误法

$$var(\hat{\beta}_1) = \frac{\sigma^2}{\Sigma x_t^2} + \frac{2\sigma^2}{\Sigma x_t^2}[\rho\frac{\Sigma_{t=1}^{T-1}x_t x_{t+1}}{\Sigma x_t^2} + \rho^2\frac{\Sigma_{t=1}^{T-2}x_t x_{t+2}}{\Sigma x_t^2} + \cdots + \rho^{T-1}\frac{x_1 x_T}{\Sigma x_t^2}]$$

5.1.6 虚假序列相关问题

模型中遗漏了重要的解释变量或函数设定有误导致的随机干扰项序列相关。

5.2 时间序列的平稳性及其检验

5.2.1 问题的提出

1. 时间序列的平稳性可以替代随机抽样假定
2. 可以有效减少虚假回归

> **虚假回归**
>
> 杜绝虚假回归最根本的方法，是正确的设定模型

5.2.2 时间序列数据的平稳性

如果X_t满足：

1. 均值 $E(X_t) = \mu$ ，与时间t无关的常数
2. 方差 $Var(X_t) = \sigma^2$ ，与时间t无关的常数
3. 协方差 $Cov(X_t X_{t+k}) = \gamma_k$ ，只与时间间隔k有关，与时间t无关的常数

则说该随机时间序列是平稳的，该随机过程是一个**平稳随机过程**

白噪声：$X_t = \mu_t \sim M(0, \sigma^2)$平稳的

随机游走：$X_t = X_{t-1} + \mu_t$非平稳的

5.2.3 平稳性的图示判断

自相关函数

如果由白噪声产生则近似为正态分布

$$r_k=\frac{\Sigma_{t=1}^{n-k}(X_t-\overline{X})(X_{t+k}-\overline{X})}{\Sigma_{t=1}^{n}(X_t-\overline{X})^2}\sim N(0,\frac{1}{n^2})$$

检验对所有$k>0$，自相关系数ρ_k都为0的联合假设：

$$Q_{LB}=n(n+1)\Sigma_{k+1}^{m}\frac{r_k^2}{n-k}\sim\chi^2(m)$$

m为滞后长度

5.2.4 平稳性的单位根检验

DF检验

将序列记为$X_t=\rho X_{t-1}+\mu_t$

可以表示为

$$\Delta X_t=X_t-X_{t-1}=(\rho-1)X_{t-1}+\mu_t=\delta X_{t-1}+\mu_t$$

一般来说，检验一个时间序列的平稳性，可以通过检验其一阶自回归模型

$$X_t=\alpha+\rho X_{t-1}+\mu_t$$

中单位根$\rho=1$是否存在来判断。

- $\rho\geq 1$： X_t 是非平稳的
- $\rho<1$： X_t 是平稳的

> 在零假设下，即使是在大样本下统计量也是有偏误的（向下偏移），通常的t检验无法使用
>
> 采用DF分布检验
>
> 如果$\tau<DF$，则拒绝原假设

ADF检验

时间序列可能由更高阶的自回归过程生成，或者随机干扰项并非白噪音

$$\begin{aligned}
&\text{模型1:} && \Delta X_t=\delta X_{t-1}+\Sigma_{i=1}^{m}\beta_i\Delta X_{t-i}+\epsilon_t\\
&\text{模型2:} && \Delta X_t=\alpha+\delta X_{t-1}+\Sigma_{i=1}^{m}\beta_i\Delta X_{t-i}+\epsilon_t\\
&\text{模型3:} && \Delta X_t=\alpha+\beta T+\delta X_{t-1}+\Sigma_{i=1}^{m}\beta_i\Delta X_{t-i}+\epsilon_t
\end{aligned}$$

5.2.5 单整时间序列

一次差分过程：

$$X_t=X_{t-1}+\mu_t$$
$$\Delta X_t=\mu_t$$

如果一个时间序列经过d次差分变成平稳序列，则称原序列是d阶单整序列，记为$I(d)$

5.3 协整与误差修正模型

5.3.1 长期均衡关系与协整

假设X、Y的长期均衡关系描述为

$$Y_t=\alpha_0+\alpha_1X_t+\mu_t$$

在$t-1$期末：

1. $Y_{t-1} = \alpha_0 + \alpha_1 X_t$
2. $Y_{t-1} < \alpha_0 + \alpha_1 X_t$
3. $Y_{t-1} > \alpha_0 + \alpha_1 X_t$

非均衡误差

$$\mu_t = Y_t - \alpha_0 - \alpha_1 X_t \sim I(0)$$

如果两个变量都是单整变量，只有当他们的单整阶相同时，才可能协整，如果他们的单整阶不相同，就不可能协整

如果序列X_{1t}、X_{2t}、$\cdots$、X_{kt}都是d阶协整的，存在向量$\alpha = (\alpha_0, \alpha_1, \cdots, \alpha_k)$使得$Z_t = \alpha X_T' \sim I(d-b)$，则认为该解释变量序列是$(d,b)$阶协整的，记为$CI(d,b)$

(d,d)阶协整的经济意义：

两个变量虽然具有各自的长期波动规律，但是之间存在着一个长期稳定的比例关系

5.3.2 协整的检验

两变量的Engle-granger检验——EG检验

检验两个I(1)的单整变量Y_t、X_t是否是协整的

1. 协整回归：使用OLS估计$Y_t = \alpha_0 + \alpha_1 X_t + \mu_t$并计算非均衡误差，得到

$$\hat{Y}_t = \hat{\alpha}_0 + \hat{\alpha}_1 X_t$$
$$e_t = Y_t - \hat{Y}_t$$

2. 检验e_t的单整性：如果$e_t \sim I(0)$则Y_t、X_t为$CI(1,1)$协整，使用模型：

$$\Delta e_t = \delta e_{t-1} + \Sigma_{i=1}^{p} \theta_i \Delta e_{t-i} + \epsilon_t$$

   进行ADF检验并获得适当的检验模型，如果拒绝零假设$H_0: \delta = 0$则说明e_t是平稳序列，从而说明X与Y是协整的

多变量的协整关系检验

对于多变量的协整检验过程，需要检验变量是否具有同阶单整性，以及是否存在稳定的线性组合。后者需要通过设置一个变量为被解释变量，其他变量为解释变量，进行普通最小二乘估计并检验残差序列是否平稳。如果不平稳则更换被解释变量，进行同样的OLS与相应的残差项检验。当所有变量都被作为被解释变量之后，若仍不能得到平稳的残差项序列，则认为这些变量间不存在(d,d)阶协整

5.3.3 关于均衡与协整的再讨论

协整方程不一定是均衡方程

1. 协整方程具有统计意义，均衡方程具有经济意义
2. 均衡方程中应该包含均衡系统中的所有时间序列，而协整方程中可以只包含其中一部分时间序列
3. 协整方程只要求随机项是平稳的，而均衡方程要求随机项时白噪声

5.3.4 误差修正模型ECM

简单差分不一定能解决非平稳时间序列所遇到的全部问题

DHSY模型

假设变量X与Y的长期均衡关系为

$$Y_t = \alpha_0 + \alpha_1 X_t + \mu_t$$

短期/非均衡的关系有如下(1, 1)阶分布滞后形式

$$Y_t = \beta_0 + \beta_1 X_t + \beta_2 X_{t-1} + \delta Y_{t-1} + \mu_t$$

适当变化后可得

$$\Delta Y_t = \beta_1 \Delta X_t - \lambda(Y_{t-1} - \alpha_0 - \alpha_1 X_{t-1}) + \mu_t$$
$$\lambda = 1 - \delta \in (0,1); \ \ \alpha_0 = \frac{\beta_0}{1-\delta}; \ \ \alpha_1 = \frac{\beta_1 + \beta_2}{1-\delta}$$

Y_t又可以写成

$$\Delta Y_t = \beta_1 \Delta X_t - \lambda \cdot ecm_{t-1} + \mu_t$$

ecm为误差修正项

- $Y_{t-1} < \alpha_0 + \alpha_1 X_{t-1}$：ecm为负
- $Y_{t-1} > \alpha_0 + \alpha_1 X_{t-1}$：ecm为正

Y关于X的长期弹性：α_1

Y关于X的短期弹性：β_1

5.4 格兰杰因果关系检验

5.4.1 时间序列自回归模型

自回归模型：$X_t = F(X_{t-1}, X_{t-2}, \cdots, \mu_t)$

- **1阶自回归过程$AR(1)$**：线性方程、1期滞后、白噪声随机扰动项

  $$X_t = \phi X_{t-1} + \epsilon_t$$

- **p阶自回归过程$AR(p)$**：

  $$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \cdots + \phi_p X_{t-p} + \mu_t$$

  纯$AR(p)$过程:随机扰动项是一个白噪声

  $$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \cdots + \phi_p X_{t-p} + \epsilon_t$$

- **q阶移动平均过程$MA(q)$**:随机扰动项不是一个白噪声

  $$\mu_t = \epsilon_t - \theta_1 \epsilon_{t-1} - \cdots - \theta_q \epsilon_{t-q}$$

- **自回归移动平均过程$ARMA(p,q)$**

  $$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \cdots + \phi_p X_{t-p} + \epsilon_t - \theta_1 \epsilon_{t-1} - \cdots - \theta_q \epsilon_{t-q}$$

  一个随机时间序列可以通过一个自回归移动平均过程生成，即序列可以由其自身的过去或滞后值以及随机扰动项来解释

$AR(p)$模型的平稳性条件

引入滞后算子L:

$$LX_t = X_{t-1}, \ \ L^2 X_t = X_{t-2}, \ \ \cdots, L^p X_t = X_{t-p},$$

则$AR(p)$模型可以变化为

$$(1 - \phi_1 L - \phi_2 L^2 - \cdots - \phi_p L^p) X_t = \epsilon_t$$

$AR(p)$的特征方程可以记为

$$\Phi(z) = (1 - \phi_1 z - \phi_2 z^2 - \cdots - \phi_p z^p) = 0$$

如果该特征方程的所有根的模都大于1，则该模型是平稳的

对于一个平稳的随机事件序列，识别其为纯AR过程，所需要使用的是时间序列的自相关函数ACF与偏自相关函数PACF。如果为一纯AR过程，可以采用OLS等方法估计参数。

5.4.2 时间序列向量自回归模型

将上述单个时间序列自回归模型扩展到多个时间序列，即构成向量自回归模型

5.4.3 格兰杰因果关系检验及其应用

当两个变量间在时间上有先导-滞后关系时，考察这种关系是单向的还是双向的

格兰杰因果关系检验的表述

定义：若在包含了变量X、Y的过去信息的条件下，对变量Y的预测效果要优于单独由Y的过去信息对Y进行的预测效果，即变量X有助于解释变量Y的将来变化，则认为变量X是引致变量Y的格兰杰原因。考察X是否影响Y的问题，主要看当期的Y在多大程度上能够被过去的X解释，在Y_t方程中加入X的滞后值是否能使解释程度显著提高。如果X在Y的预测中有帮助，或者X、Y的相关系数在统计上显著时，则说**“X是Y的Granger原因”**。

$$Y_t = \beta_0 + \Sigma_{i=1}^{m}\beta_i Y_{t-i} + \Sigma_{i=1}^{m}\alpha_i X_{t-i} + \mu_t$$

$$X_t = \delta_0 + \Sigma_{i=1}^{m}\delta_i X_{t-i} + \Sigma_{i=1}^{m}\lambda_i Y_{t-i} + v_t$$

1. X对Y有单向影响：Y_t 式中 X 各滞后项前的参数整体不为0，X_t 式中 Y 各滞后项前的参数整体为0
2. Y对X有单向影响：Y_t 式中 X 各滞后项前的参数整体为0，X_t 式中 Y 各滞后项前的参数整体不为0
3. Y与X间有双向影响：Y_t 式中 X 各滞后项前的参数整体不为0，X_t 式中 Y 各滞后项前的参数整体不为0
4. X与Y是独立的：Y_t 式中 X 各滞后项前的参数整体为0，X_t 式中 Y 各滞后项前的参数整体为0

使用F检验判断参数是否整体为0

$$F = \frac{(RSS_R - RSS_U)/m}{RSS/(n-k)} \sim F(m, n-k)$$

应用中的实际问题

1. 滞后期长度的选择问题
2. 时间序列的平稳性问题：对同阶单整非稳定序列也有一定程度的可靠性
3. 样本容量问题：样本容量越大，存在格兰杰因果关系的概率显著增大
4. 格兰杰因果关系检验是必要性条件检验，而不是充分性条件检验：是统计意义上的，不是经济意义上的

---

chapter6 非经典截面数据计量经济学模型

6.1 选择性样本计量经济学模型

受限被解释变量模型：选择性样本模型+持续时间被解释变量模型

6.1.1 经济生活中的选择性样本问题

截断问题：只能从一部分个体中随机抽取被解释变量的样本观测值

归并问题：将解释变量处于某一范围的样本观测值都用一个相同的值代替

6.1.2 “截断”问题的计量经济学模型

通过函数的极大化求得模型的参数估计量

截断分布

$$P(\epsilon > a) = 1 - \Phi(\frac{a-\mu}{\sigma})$$

截断被解释变量数据计量经济学模型的极大似然估计

$$\frac{\partial \ln L}{\partial \begin{pmatrix} \beta \\ \sigma^2 \end{pmatrix}} = \Sigma_{i=1}^{n} \begin{pmatrix} \frac{Y_i - X_i\beta}{\sigma^2} X_i \\ -\frac{1}{2\sigma^2} + \frac{(Y_i - X_i\beta)^2}{2\sigma^4} - \frac{\alpha_i\lambda_i}{2\sigma^2} = \Sigma_{i=1}^{n} g_i = 0 \end{pmatrix}$$

$$\alpha_i = (a - X_i\beta)/\sigma$$
$$\lambda_i = \Phi(\alpha_i)/(1 - \Phi(\alpha_i))$$

如果采用OLS估计会产生**选择性偏误**

1. 忽略了一个非线性项 λ_i
2. 忽略了随机误差项实际上的异方差性

6.1.3 “归并”问题的计量经济学模型

$$\begin{cases} Y = a, & Y^* \le a \\ Y = Y^*, & Y^* > a \end{cases}$$

tobit模型

$$\begin{cases} Y_i^* = X_i\beta + \mu_i, & \mu_i \sim N(0, \sigma^2) \\ Y_i = max(Y_i^*, 0) \end{cases}$$

归并模型的正态分布

$$P(Y = 0) = 1 - \Phi(\frac{\mu}{\sigma})$$

归并模型的极大似然估计

$$\ln L = \Sigma_{Y_i>0} - \frac{1}{2}(\ln(2\pi) + \ln\sigma^2 + \frac{(Y_i - X_i\beta)^2}{\sigma^2}) + \Sigma_{Y_i=0} \ln(1 - \Phi(\frac{X_i\beta}{\sigma}))$$

由两部分组成

1. 没有限制的观测值
2. 受到限制的观测值

6.2 二元离散选择模型

随机效用模型

$$U_i^1 = X_i\beta^1 + \epsilon_i^1$$

记

$$Y_i* = U_i^1 - U_I^0 = X_i(\beta_1 - \beta_0) + (\epsilon_i^1 - \epsilon_i^0) = X_i\beta + \mu_i^*$$

6.2.3 二元Probit离散选择模型及其参数估计

将正态分布作为μ_i^*的概率分布而推导得到的

$$F(t) = \int_{-\infty}^{t} (2\pi)^{-\frac{1}{2}} \exp(-\frac{x^2}{2})dx$$

概率密度函数

$$f(x) = (2\pi)^{-\frac{1}{2}} \exp(-\frac{x^2}{2})$$

1.重复观测值不可以得到的情况下Probit的参数估计

2.重复观测值可以得到的情况下Probit的参数估计

6.2.4 二元Logit离散选择模型及其参数估计

$$F(t) = \frac{1}{1 + e^{-t}}$$

$$f(t)=\frac{e^{-t}}{(1+e^{-t})^2}$$

6.2.5 二元离散选择模型的检验

1. 拟合优度检验

   设L_0为所有解释变量的系数都为0时的似然函数值

$$\ln L_0=n[P\ln P+(1-P)\ln(1-P)]$$

   设L为模型估计得到的似然函数值

$$R^2=1-\frac{\ln L}{\ln L_0}$$

2. 总体显著性检验

$$\begin{gathered}H_0:\beta_1=\beta_2=\cdots=\beta_k=0\\ LR=-2(\ln L_0-\ln L)\sim\chi^2(k)\end{gathered}$$

3. 回代效果检验

6.3 固定效应面板数据计量经济学模型

面板数据计量经济学模型理论是基于样本信息的充分利用而发展的

6.3.1 截面个体变系数模型

$$Y_{it}=\alpha_i+X_{it}\beta_i+\mu_{it}$$

在截面个体之间，存在个体影响，也存在变化的经济结构，因此结构参数在不同横截面个体上是不同的

6.3.2 截面个体变截距模型

$$Y_{it}=\alpha_i+X_{it}\beta+\mu_{it}$$

在横截面个体之间，存在个体影响，但不存在变化的经济结构，因此结构参数在不同横截面个体上是相同的

6.3.3 界面个体截距、系数不变模型

$$Y_{it}=\alpha+X_{it}\beta+\mu_{it}$$

在横截面个体间，不存在个体影响，也不存在变化的经济结构，因此模型的截距与结构参数在不同横截面个体上都相同

6.3.4 截面个体不变截距、变系数模型

$$Y_{it}=\alpha+X_{it}\beta_i+\mu_{it}$$

在横截面个体之间，不存在个体影响，但是存在变化的经济结构，因此模型截距相同

6.3.5 时点变系数模型

$$Y_{it}=\alpha_t+X_{it}\beta_t+\mu_{it}$$

在不同的时点间，存在个体影响，也存在变化的经济结构，因此结构参数在不同时点是不同的

6.3.6 截面个体和时点变截距模型

$$Y_{it}=\alpha_i+\gamma_t+X_{it}\beta+\mu_{it}$$

在横截面个体间，存在个体影响，同时在不同的时点之间存在个体影响，但不存在变化的经济结构

6.3.7 模型设定检验

模型设定的F检验（协变分析检验）

$$
\begin{gathered}
H_0: y_{it}=\alpha_i+X_{it}\beta_i+\mu_{it} \\
H_1: y_{it}=\alpha_i+X_{it}\beta+\mu_{it} \\
H_2: y_{it}=\alpha+X_{it}\beta+\mu_{it}
\end{gathered}
$$

$$
\begin{gathered}
F_2=\frac{(S_3-S_1)/[(n-1)(K+1)]}{S_1/[nT-n(K+1)]}\sim F[(n-1)(K+1),[nT-n(K+1)]] \\
F_1=\frac{(S_2-S_1)/[(n-1)K]}{S_1/[nT-n(K+1)]}\sim F[(n-1)K,n(T-K-1)]
\end{gathered}
$$

6.4 固定效应变截距模型

固定影响变截距模型：个体影响可以用常数项α_i代替

随机影响变截距模型：个体影响可以用不变的常数项α_i与变化的随机项ϵ_i之和代替

LSDV最小二乘虚拟变量模型

$$
\begin{pmatrix}\hat{\alpha}\\ \hat{\beta}\end{pmatrix}=((DX)'(DX))^{-1}（DX)'y
$$

$$
D=\begin{pmatrix}e & 0 & \cdots & 0\\ 0 & e & \cdots & 0\\ \vdots & \vdots & & \vdots\\ 0 & 0 & \cdots & e\end{pmatrix}_{nT\times n}
$$

协方差分析模型的性质

$$
\begin{gathered}
Var(\hat{\beta}_{CV})=\sigma_\mu^2(\Sigma_{i=1}^n X_i'QX_i)^{-1} \\
Var(\hat{\alpha}_i)=\frac{\sigma_\mu^2}{T}+\overline{X}_i Var(\hat{\beta}_{CV})\overline{X}_i' \\
s^2=\Sigma_{i=1}^n\Sigma_{t=1}^T\frac{(y_{it}-\hat{\alpha}_i-x_{it}\hat{\beta}_{CV})^2}{nT-n-K}
\end{gathered}
$$

6.5 固定效应变系数模型

$$
\begin{gathered}
\hat{\beta}_{GLS}=(X'V^{-1}X)^{-1}X'V^{-1}y \\
V=\begin{pmatrix}\Omega_{11} & \Omega_{12} & \cdots & \Omega_{1n}\\ \Omega_{21} & \Omega_{22} & \cdots & \Omega_{2n}\\ \vdots & \vdots & & \vdots\\ \Omega_{n1} & \Omega_{n2} & \cdots & \Omega_{nn}\end{pmatrix},\Omega_{ij}=E(\mu_i\mu_j)
\end{gathered}
$$