The first note of Econometrics

**1.平均因果效应**

在随机分配的情况下，接受干预的个体与未接受干预的个体在结果变量上的平均值差异。公式为：即处理组与对照组结果变量的期望差异。

**2.因变量、自变量**

定义一个简单线性回归模型:.y被称为“因变量”，也就是“被解释变量”；x被称为“自变量”，也就是“解释变**量”**。

**3.高斯·马尔科夫假定**

①假定MLR.1（**线性于参数**）

在总体模型中，因变量y、自变量x和误差项u关系如下：

，其中，β1，β2``````βk 是我们所关心的未知参数（常数），而u则是我们无法观测的随机误差或随机干扰。

②假定MLR.2（**随机抽样**）

我们可以使用总体的一个容量为n的随机样本{（x1，x2,x3,x4``````xk,yi;i=1,2,3···）}，其中i代表观察，j=1,2,3···k代表第j个回归元。

③假定MLR.3（**不存在完全共线性**）

如果方程中一个自变量是其他解释变量的严格线性组合时，我们认为此模型具有完全共线性；在样本中，没有一个自变量是常熟，自变量之间也不存在严格的线性关系。

④假定MLR.4（**零条件均值**）

给定解释变量的任何值，误差的期望为零，即。因此，我们称所有解释是外生的，否则则是内生的。

⑤假定MLR.5（**同方差性**）

给定解释变量的任何值，误差项具有相同的方差，。

**4.一致估计量**

**5.依概率分析**

T检验单个回归系数，F检验是对多个约束的联合检验

**6.无偏估计量**

**7.测量误差不会影响估计的无偏性与一致性**

**8.遗漏变量是对OLS的影响，一般保证不了无偏性与一致性P135**

**9.邹至庄检验**

将实验数据分成两组，比较两组数据的线性模型的残差平方和。假设具有相同未知方差的独立同分布——正态分布。

具体步骤如下：

1. **分别**对两组数据进行回归估计得到残差平方和SSR1和SSR2，两者之和即为无约束模型的残差平方和
2. 将两组数据**混合估计**得到残差平方和SSRP，这就是为受约束模型的残差平方和
3. 构造F统计量，并于给定显著水下F的临界值进行比较，若F值大于临界值，则拒绝假设，即说明两个回归函数存在差异。
4. **虚拟变量**

概念：有关信息通过一个二值变量或者0-1变量来刻画，一般称为虚拟变量（**精准定义：指仅有0-1两个取值的变量，是一种定性变量**。）

**虚拟变量陷进**：在多元线性回归模型中，设置了多个虚拟变量，从而导致了多重共线性

1. **线性概率模型（概念、缺陷）**

**概念：**以虚拟变量为解释变量的多元线性回归模型（）叫做线性概率模型。

**缺陷：**

（1）**随机扰动项uI的非正态性**（在线性概率模型中，ui不再服从正态分布，但是对参数的假设检验和区间估计要求随机扰动项uI服从正态分布）

（2）**随机扰动项uI的异方差性**[可证明Var(ui）=pi(1-pi),OLS的估计值不再具有最小方差性，各参数的估计方差也不是最佳估计，但是估计值仍然具有无偏性]

（3）**不满足0≤E（Yi|Xi）≤1的约束**。线性概率模型估计的概率可能大于1或者是负数。

1. **BP检验原理与步骤**

H0：原模型是同方差的 H1：元模型是异方差的

①对原模型进行回归，得到残差u的预测值，计算其平方

②将残差平方和对所有解释变量做一次项的OLS回归，即，得到拟合优度R2

③利用拟合优度做辅助性的F检验或者LM检验，其中

 

1. **White检验原理与步骤**

H0：原模型是同方差的 H1：元模型是异方差的

①对原模型进行回归，得到残差ui的预测值，计算ui平方、yi的预测值、yi的预测值平方

②将ui平方、yi的预测值、yi的预测值平方做OLS回归，得到拟合优度R2，即,

③利用拟合优度做方程整体显著性的F检验或者LM检验，其中；

**第一章**

**1.横截面数据**

一个或多个变量发生在同一时间截面上的调查数据

**2.时间序列数据**

同一观测个体在不同时点上所观测的数据构成

**3.数学期望具有线性可加性**

**4.；若X与Y独立，相关系数=0，**

**5.中心极限定理的自我表述**

很多个随机变量的和服从正态分布（同分布，相互独立）

**6.小概率事件在一次实验中一般不会发生，若发生可能不是小概率事件，则原假设可能不正确**

**第二章**

**1.简单回归模型**

定义一个简单线性回归模型:.y被称为“因变量”，也就是“被解释变量”；x被称为“自变量”，也就是“解释变**量”**。Β0：常数项、截距项 β1：斜率参数

**2.零条件均值假定**

u的平均值与x无关，且

**3.最小二乘估计**

OLS估计的方法：使残差平方和最小



对β0与β1分别求导**，**

** **（方分之协）

**4.OLS的性质**

①残差和为零

②样本协方差为零

③****

④

**5.名词解释**

：总平方和 ：解释平方和

：残差平方和 （拟合优度）假设拟合优度=60%，则对变量的解释程度为60%

**6.度量单位变化的规律**（P16，表2-3）

**7.无偏性假定的四个条件**

①假定SLR.1（**线性于参数**）

在总体模型中，因变量y、自变量x和误差项u关系如下：

，其中，β1，β2``````βk 是我们所关心的未知参数（常数），而u则是我们无法观测的随机误差或随机干扰。

②假定SLR.2（**随机抽样**）

我们可以使用总体的一个容量为n的随机样本{（x1，x2,x3,x4``````xk,yi;i=1,2,3···）}，其中i代表观察，j=1,2,3···k代表第j个回归元。

③假定SLR.3（**解释变量的样本有变异**）

x的样本结果，即{xi,i=1,2,3······n},数值不完全相同

④假定SLR.4（**零条件均值**）

给定解释变量的任何值，误差的期望为零，即。因此，我们称所有解释是外生的，否则则是内生的。

**8.、OLS的无偏性、OLS估计量的方差(p17)**

**9.残差与误差的区分**

误差无法观测，残差可以根据数据计算得到。

**第三章**

**1.多元线性回归模型**

一般形式：****，其中具有K个自变量、一个截距项、k个斜率参数

**2.外生变量、内生变量**

**外生：**在多元线性模型中，与误差项不相关的解释变量

**内生：**在多元线性模型中，与误差项相关的解释变量

**3.多元线性回归模型中最小二乘法的操作和解释p36**

**4.过度设定不影响OLS的无偏性，因为仍然满足4个假定，故参数估计是无偏的。**

**5.遗漏变量的偏误，影响OLS导致参数估计有偏P**

**偏误公式：**

**6.影响系数误差的因素**

**①误差方差σ2（正向影响）②xj的总样本变异SSTJ（负面影响）③变量间的线性关系RJ2（正向影响）**

**7.的表达式推导**

**第四章**

**1.正态性**

假定总体误差项独立于所有解释变量，且服从均值为零，方差为σ2的正态分布

**2.T检验的步骤（笔记）**

**3.单边检验P98**

**4.双侧检验P102**

**5.置信区间**

**;（区间范围）**

**6.F检验原理及步骤（笔记本）**

**第六章**

**1.度量单位不影响t统计量、F统计量、拟合优度R2，要让斜率参数变大，应使对应自变量的度量单位变大或将Y的单位变小**

**2.Beta系数：也称为标准化系数，将所有的X、Y都标准化后，再做无截距项回归得到系数的估计值。**

**3.调整之后的拟合优度**

**公式：**

**①可能小于0，拟合优度等于0时，调整之后的拟合优度=**

**第七章**

1. **虚拟变量陷进**
2. **邹至庄检验（原理及模型设计）**

**第八章**

**1.异方差性**

**，被解释变量观测值的分散程度是随解释变量的变化而变化的P142**

**2.BP检验步骤**

**3.White检验步骤**

**4.VIF，方差膨胀因子？**

**5.WLS检验（笔记本）**

**6.FGLS检验步骤（笔记本）**

**7.弱平稳的三个条件（协方差平稳）**

**①均值E（xt）=u是与时间t无关的常数**

**②方差Var（xt）=σ2是与时间t无关的常数**

**③协方差Cov（xt,xt+k）是只与时间间隔t相关的函数**