

## 计量经济学习题

### 一、名词解释

- 1、普通最小二乘法：为使被解释变量的估计值及观测值在总体上最为接近使  $Q =$  最小，从而求出参数估计量的方法，即之。
- 2、总平方和、回归平方和、残差平方和的定义：TSS 度量  $Y$  自身的差异程度，称为总平方和。TSS 除以自由度  $n-1$  = 因变量的方差，度量因变量自身的变化；RSS 度量因变量  $Y$  的拟合值自身的差异程度，称为回归平方和，RSS 除以自由度（自变量个数-1）= 回归方差，度量由自变量的变化引起的因变量变化局部；ESS 度量实际值及拟合值之间的差异程度，称为残差平方和。RSS 除以自由度（ $n$ -自变量个数-1）= 残差（误差）方差，度量由非自变量的变化引起的因变量变化局部。
- 3、计量经济学：计量经济学是以经济理论为指导，以事实为依据，以数学和统计学为方法，以电脑技术为工具，从事经济关系及经济活动数量规律的研究，并以建立和应用经济计量模型为核心的一门经济学科。而且必须指出，这些经济计量模型是具有随机性特征的。
- 4、最小样本容量：即从最小二乘原理和最大似然原理出发，欲得到参数估计量，不管其质量如何，所要求的样本容量的下限；即样本容量必须不少于模型中解释变量的数目（包括常数项），即之。
- 5、序列相关性：模型的随机误差项违背了相互独立的根本假设的情况。
- 6、多重共线性：在线性回归模型中，如果某两个或多个解释变量之间出现了相关性，那么称为多重共线性。
- 7、工具变量法：在模型估计过程中被作为工具使用，以替代模型中及随机误差项相关的随机解释变量。这种估计方法称为工具变量法。
- 8、时间序列数据：按照时间先后排列的统计数据。
- 9、截面数据：发生在同一时间截面上的调查数据。
- 10、相关系数：指两个以上的变量的样本观测值序列之间表现出来的随机数学关系。
- 11、异方差：对于线性回归模型提出了假设干根本假设，其中包括随机误差项具有同方差；如果对于不同样本点，随机误差项的方差不再是常数，而互不一样，那么认为出现了异方差性。
- 12、外生变量：外生变量是模型以外决定的变量，作为自变量影响内生变量，外生变量决定内生变量，其参数不是模型系统的元素。因此，外生变量本身不能在模型体系内得到说明。外生变量一般是确定性变量，或者是具有临界概率分布的随机变量。外生变量影响系统，但本身并不受系统的影响。外生变量一般是经济变量、条件变量、政策变量、虚变量。一般情况下，外生变量及随机项不相关。





## 二、填空题

- 1、计量经济学中，经济学 提供理论根底，统计学 提供资料依据，数学 提供研究方法。
- 2、研究经济问题时，一般要处理三种类型的数据：(1) 截面 数据；(2) 时间序列 数据；和 (3) 虚拟变量 数据。
- 3、OLS 参数估计量具有如下统计性质，即 线性、无偏性、有效性。
- 4、时间序列数据及横截面数据的最大区别在于 数据的顺序性。
- 5、在模型中引入多个虚拟变量时，虚拟变量的个数应按以下原则确定：如果有 M 个互斥的属性类型，那么在模型中引入 M-1 个虚拟变量。
- 6、在现实经济活动中往往存在一个被解释变量受到多个解释变量的影响的现象，表现为在线性回归模型中有多个解释变量，这样的模型被称为 多元线性回归模型。
- 7、在多元线性回归模型中，参数的最小二乘估计量具 线性性、无偏性、最小方差性，同时多元线性回归模型满足经典假定，所以此时的最小二乘估计量是最优的线性无偏估计量，又称 BLUE 估计量。
- 8、计量经济学的核心内容是建立和应用 计量经济模型。
- 9、 $R^2$  是一个回归直线及样本观测值拟合优度的数量指标，其值越大，拟合优度越好，其值越小，拟合优度就越差。
- 10、自相关 就是指总体回归方程的误差项  $u_i$  之间存在着相关，即：按时间或空间排序的观察值序列的个成员之间存在的相关。

## 三、单项选择题

1. 经济计量模型是指( C )  
A. 投入产出模型  
B. 数学规划模型  
C. 包含随机方程的经济数学模型  
D. 模糊数学模型
2. 回归分析中定义的( B )  
A. 解释变量和被解释变量都是随机变量  
B. 解释变量为非随机变量，被解释变量为随机变量  
C. 解释变量和被解释变量都为非随机变量  
D. 解释变量为随机变量，被解释变量为非随机变量
3. 设 k 为回归模型中的参数个数, n 为样本容量。那么对总体回归模型进展显著性检验 (F 检验) 时构造的 F 统计量为( A )

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

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

$$C. F = \frac{ESS}{RSS}$$

$$D. F = \frac{RSS}{ESS}$$





4. D-W 检验, 即杜宾-瓦尔森检验, 用于检验时间序列回归模型的误差项中的一阶序列相关的统计量, DW 统计量以 OLS 残差为根底:

$$D. W = \frac{\sum_{t=2}^n (\tilde{e}_t - \tilde{e}_{t-1})^2}{\sum_{t=1}^n \tilde{e}_t^2}, \text{ 那么 ( C )}$$

- A. 那么说明存在着正的自相关      B. 那么说明存在着负的自相关  
C. 那么说明无自相关      D. 无法说明任何意义
5. 容易产生异方差的数据为 ( C )  
A. 时序数据      B. 修匀数据  
C. 横截面数据      D. 年度数据
6. 计量经济模型分为单方程模型和 ( C )。  
A. 随机方程模型      B. 行为方程模型
7. 同一统计指标按时间顺序记录的数据列称为 ( B )  
A. 横截面数据      B. 时间序列数据
8. 样本数据的质量问题, 可以概括为完整性、准确性、可比性和 ( B )。  
A. 时效性      B. 一致性
9. 有人采用全国大中型煤炭企业的截面数据, 估计生产函数模型, 然后用该模型预测未来煤炭行业的产出量, 这是违反了数据的 ( A ) 原那么。  
A. 一致性      B. 准确性
10. 对以下模型进展经济意义检验, 哪一个模型通常被认为没有实际价值的 ( B )。  
A.  $C_i$  (消费)  $= 500 + 0.8I_i$  (收入)  
B.  $Q_{di}$  (商品需求)  $= 10 + 0.8I_i$  (收入)  $0.9P_i$  (价格)  
C.  $Q_{si}$  (商品供应)  $= 20 + 0.75P_i$  (价格)  
D.  $Y_i$  (产出量)  $= 0.65K_i^{0.6}$  (资本)  $L_i^{0.4}$  (劳动)

#### 四、多项选择题

1. 不满足 OLS 根本假定的情况, 主要包括: ( ABCD )。  
A. 随机序列项不是同方差, 而是异方差





- B. 随机序列项序列相关，即存在自相关
- C. 解释变量是随机变量，且及随机扰动项相关
- D. 解释变量之间相关，存在多重共线性
- E. 因变量是随机变量，即存在误差

2、随机扰动项产生的原因大致包括如下几个方面，它们是（ ABCD ）。

- A. 客观现象的随机性（人的行为、社会环境及自然影响的随机性）
- B. 模型省略变量（被省略的具有随机性的变量归入随机扰动项）
- C. 测量及归并误差（估计时测量和归并误差都归入随机扰动项）
- D. 数学模型函数的形式的误定
- E. 从根本上看是由于经济活动是人类参与的活动

3、内生变量（ ABDE ）。

A. 在联立方程模型中，内生变量由系统内方程决定，同时又对模型系统产生影响；既作为被解释变量，又可以在不同的方程中作为解释变量。

- B. 一般情况下，内生变量及随机项相关。
- C. 内生变量决定外生变量
- D. 内生变量一般都是经济变量
- E. 内生变量  $Y$  一般满足：

$$\text{Cov}(Y_i, \mu_i) \neq 0, \text{ 即 } E(Y_i \mu_i) \neq 0.$$

4、影响预测精度的因素包括（ ACD ）。

- A. 样本容量愈大，预测的方差愈小，预测的精度愈大
- B. 样本中解释变量的离均差的和愈大，预测的方差愈小，预测的精度愈大
- C. 内插预测的精度比拟有把握，外推预测的能力显著下降，预测精度难以把握
- D. 当其样本容量  $n$  相当大，而预测点的取值  $X_0$  接近于  $X$  的平均值时，预测的方差最小，预测的精度最大
- E. 残差标准差的估计值愈小，回归预测的精度愈准确，所以常常把残差标准差的估计值作为预测精度的标志

5. 以下哪些变量属于前定变量（ CD ）。

- A. 内生变量
- B. 随机变量
- C. 滞后变量
- D. 外生变量
- E. 工具变量

## 五、判断题

- 1、通常把由方程组内决定的变量称为内生变量，而不能由方程组内直接决定的变量为前定变量，又称为先决变量。（√）
- 2、前定（先决）变量既能作为解释变量，也能作为被解释变量。（×）



3、D-W 检验，即杜宾-瓦尔森检验， $D.W = \frac{\sum_{i=2}^n (\tilde{e}_i - \tilde{e}_{i-1})^2}{\sum_{i=1}^n \tilde{e}_i^2}$ ，其最大优点为简单易行。如果

D.W 值接近于零，那么说明越倾向于无自相关。(×)

4、截面数据是一批发生在同一时间截面上的调查数据。例如，在给定的某个时点上对个人、家户、企业、城市、地区、国家或一系列其它单位采集的样本所构成的数据集。(√)

5、内生变量是理论或模型所要解释的变量，即因变量，它是为理论或模型以外的因素所影响的变量，是具有某种概率分布的随机变量。(√)

6、违背根本假设的计量经济学模型是不可估计的。(×)

7、只有满足根本假设的计量经济学模型的普通最小二乘参数估计量才具有无偏性和有效性。(√)

8、要使得计量经济学模型拟合得好，就必须增加解释变量。(×)

9、在拟合优度检验中，拟合优度高，那么解释变量对被解释变量的解释程度就高，可以推测模型总体线性关系成立；反之亦然。(×)

10、样本容量 N 越小，残差平方和 RSS 就越小，模型拟合优度越好。(×)

11、当计量经济学模型出现异方差性，其普通最小二乘法参数估计量仍具有无偏性，但不具有有效性。(√)

12、实际问题中的多重共线性不是自变量之间存在理论上或实际上的线性关系造成的，而是由于所收集的数据之间存在近似的线性关系所致。(√)

13、模型的拟合优度不是判断模型质量的唯一标准，为了追求模型的经济意义，可以牺牲一点拟合优度。(√)

14、如果给定解释变量值，根据模型就可以得到被解释变量的预测值。(×)

15、异方差问题中，随机误差项的方差及解释变量观测值之间都是有规律可循的。(×)

16、计量经济学模型解释经济活动中各因素之间的理论关系，用确定性的数学方程加以描述。(×)

17、计量经济学根据研究对象和内容侧重面不同，可以分为广义计量经济学和狭义计量经济学。(√)

18、计量经济学是一门经济学科，而不是数学或其他。(√)

19、样本数据的收集是计量经济学的核心内容。(×)

20、方法，主要包括模型方法和计算方法，是计量经济学研究的根底。(×)

21、具有因果关系的变量之间一定有数学上的相关关系，具有相关关系的变量之间一定具有因果关系。(×)

22、乘数是变量的变化率之比。(×)



- 23、单方程计量经济学模型是以多个经济现象为研究对象，是应用最为普遍的计量经济学模型。(×)
- 24、对于最小二乘法最合理的参数估计量应该使得从模型中抽取  $n$  组样本观测值的概率最大。(×)
- 25、总体平方和由残差平方和和回归平方和组成。(√)
- 26、校正的判定系数和非校正的判定系数仅当非校正判定系数为 1 时才相等。(√)
- 27、判定所有解释变量是否对应变量有显著影响的方法是看是否每个解释变量都是显著的  $t$  统计量；如果不是，那么解释变量整体是统计不显著的。(×)
- 28、当  $R^2=1$ ,  $F=0$  ;当  $R^2=0$ ,  $F=\infty$ 。(×)
- 29、在模型  $Y_i=B_1+B_2X_{2i}+B_3X_{3i}+u_i$  中，如果  $X_2$  和  $X_3$  负相关且  $B_3>0$ , 那么从模型中略去解释变量  $X_3$  将使  $b_{12}$  的值减小[也即,  $E(b_{12})<(B_2)$ ]。其中  $b_{12}$  是  $Y$  仅对  $X_2$  的回归方程中的斜率系数。(√)
- 30、当我们说估计的回归系数在统计上是显著的，意思是说它显著不为 1。(×)
- 31、要计算  $t$  临界值，仅仅需知道自由度。(×)
- 32、整个多元回归模型在统计上是显著的意味着模型中任何一个单独的变量均是统计显著的。(×)
- 33、就估计和假设检验而言，单方程回归及多元回归没有什么区别。(√)
- 34、无论模型中包括多少个解释变量，总离差平方和的自由度总为  $(n-1)$ 。(√)
- 35、双对数模型的斜率和弹性系数一样。(√)
- 36、对于变量之间是线性的模型而言，斜率系数是一个常数，弹性系数是一个变量。但双对数模型的弹性系数是一个常数，而斜率是一个变量。(√)
- 37、双对数模型的  $R^2$  值可以及对数-线性模型的相比拟，但不能及线性-对数模型的相比拟。(√)
- 38、线性-对数模型的  $R^2$  值可以及线性模型相比拟，但不能及双对数模型或对数线性模型的相比拟。(√)
- 39、模型 A:  $+$ ;  $r^2$ ; 模型 B:  $;$   $r^2$  模型 A 更好一些，因为它的  $r^2$  大。(×)
- 40、在存在异方差情况下，普通最小二乘估计是有偏的和无效的。(×)
- 41、如果存在异方差，通常使用的  $t$  检验和  $F$  检验是无效的。(√)
- 42、在存在异方差情况下，常用的 OLS 估计总是高估了估计量的标准差。(×)
- 43、当存在序列相关时，OLS 估计量是有偏的并且也是无效的。(×)
- 44、消除序列相关的广义差分变换假定自相关系数必须等于 1。(√)
- 45、两个模型，一个是一阶差分形式，一个是水平形式，这两个模型的  $R^2$  是不可以直接比拟的。(√)
- 46、存在多重共线性时，模型参数无法估计。(×)
- 47、尽管存在着完全多重共线性，普通最小二乘估计量仍然是最优线性无偏估计量。



(×)

48、在存在高度多重共线性的情况下，无法估计一个或多个偏回归系数的显著性。(√)

49、一旦模型中的解释变量是随机变量，那么违背了根本假设，使得模型的 OLS 估计量有偏且不一致。(×)

## 六、简答

### 1、随机扰动项产生的原因

答：(1) 客观现象的随机性。引入  $e$  的根本原因，乃是经济活动是人类参与的，因此不可能像科学实验那样准确。

(2) 此外还有社会环境和自然环境的随机性。

(3) 模型省略了变量。被省略的变量包含在随机扰动项  $e$  中。

(4) 测量及归并误差。测量误差致使观察值不等于实际值，汇总也存在误差。

(5) 数学模型形式设定造成的误差。由于认识缺乏或者简化，将非线性设定成线性模型。

经济计量模型的随机性，正是为什么要采用数理统计方法的原因。

### 2、采用普通最小二乘法，已经保证了模型最好地拟合样本观测值，为何还要进展拟合优度检验？

答：普通最小二乘法所保证的最好拟合，是同一个问题内部的比拟，拟合优度检验结果所表示的优劣是不同问题之间的比拟。两个同样满足最小二乘原那么的模型，对样本观测值的拟合程度不一定一样。

### 3、针对普通最小二乘法，线性回归模型的根本假设

答：(1) 解释变量是确定性变量，而且解释变量之间不相关。

(2) 随机误差项具有 0 均值且同方差。

(3) 随机误差项在不同样本点之间独立，不存在序列相关。

(4) 随机误差项及解释变量之间不相关。

(5) 随机误差项服从 0 均值且同方差的正态分布。

## 七、综合题

1、某人试图建立我国煤炭行业生产方程，以煤炭产量为被解释变量，经过理论和经历分析，确定以固定资产原值、职工人数和电力消耗量变量作为解释变量，变量的选择是正确的。于是建立了如下形式的理论模型：

$$\text{煤炭产量} = \alpha_0 + \alpha_1 \text{固定资产原值} + \alpha_2 \text{职工人数} + \alpha_3 \text{电力消耗量} + \mu$$

选择 2000 年全国 60 个大型国有煤炭企业的数据为样本观测值；固定资产原值用资产形成当年当年价计算的价值量，其它采用实物量单位；采用 OLS 方法估计参数。指出该计量经济学问题中可能存在的主要错误，并简单说明理由。



- 答：(1) 模型关系错误。直接线性模型表示投入要素之间完全可以替代，及实际生产活动不符。
- (2) 估计方法错误。该问题存在明显的序列相关性，不能采用 OLS 方法估计。
- (3) 样本选择违反一致性。行业生产方程不能选择企业作为样本。
- (4) 样本数据违反可比性。固定资产原值用资产形成年当年价计算的价值量，不具备可比性。

2、材料：为证明刻卜勒行星运行第三定律，把地球及太阳的距离定为 1 个单位。地球绕太阳公转一周的时间为 1 个单位（年）。那么太阳系 9 个行星及太阳的距离（D）和绕太阳各公转一周所需时间（T）的数据如下：

obs	水星	金星	地球	火星	木星	土星	天王星	海王星	冥王星
DISTANCE			1						
Time			1				84	165	248
D <sup>3</sup>			1				7078	27271	61630
T <sup>2</sup>			1				7056	27225	61504

用上述数据建立计量模型并使用 EViews 计算输出结果如下

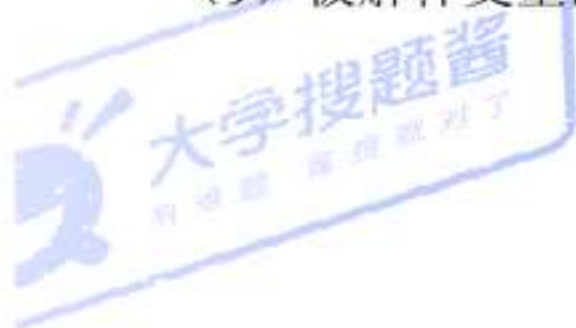
Dependent Variable: LOG(TIME)  
Method: Least Squares  
Date: 02/24/04 Time: 23:30  
Sample: 1 9  
Included observations: 9

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LOG(DISTANCE)	1.500033	0.000334	4492.202	0.0000

R-squared 0.999999    Mean dependent var 2.181016  
Adjusted R-squared 0.999999    S.D. dependent var 2.587182  
S.E. of regression 0.002185    Akaike info criterion -9.309794  
Sum squared resid 3.82E-05    Schwarz criterion -9.287880  
Log likelihood 42.89407    Durbin-Watson stat 0.952167

问题：根据 EViews 计算输出结果答复以下问题

- (1) EViews 计算选用的解释变量是\_\_\_\_\_
- (2) EViews 计算选用的被解释变量是\_\_\_\_\_
- (3) 建立的回归模型方程是\_\_\_\_\_
- (4) 回归模型的拟合优度为\_\_\_\_\_
- (5) 回归函数的标准差为\_\_\_\_\_
- (6) 回归参数估计值的样本标准差为\_\_\_\_\_
- (7) 回归参数估计值的 t 统计量值为\_\_\_\_\_
- (8) 残差平方和为\_\_\_\_\_
- (9) 被解释变量的平均数为\_\_\_\_\_





(10) 被解释变量的标准差为\_\_\_\_\_

答案如下：

- (1) Log(distance)
- (2) Log(time)
- (3) L Log(time)+u
- (4)
- (5)
- (6)
- (7)
- (8)
- (9)
- (10)

3、

(中国) 国内生产总值及投资及货物和效劳净出口

单位：亿元

年份	国内生产总值 (Y)	资本形成额 (X1)	货物和效劳净出口 (X2)
1991			
1992			
1993			
1994			
1995			
1996			
1997			
1998			
1999			
2000			
2001			
2002			
2003			

用上述数据建立计量模型并使用 EViews 计算输出结果如下

Dependent Variable: Y

Method: Least Squares

Date: 10/19/09 Time: 21:40

Sample: 1991 2003

Included observations: 13

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C				
X1				
X2				

R-squared

Mean dependent var

Adjusted R-squared

S.D. dependent var

S.E. of regression

Akaike info criterion



Sum squared resid	1.00E+08	Schwarz criterion
Log likelihood		F-statistic
Durbin-Watson stat		Prob(F-statistic)

(1) 建立投资及净出口及国民生产总值的二元线性回归方程并进展估计，并解释斜率系数的经济意义。

解：建立  $Y$  及  $X_1$ 、 $X_2$  之间的线性回归模型：

$$Y = \hat{\beta}_0 + \hat{\beta}_1 X_1 + \hat{\beta}_2 X_2 + e_i$$

根据普通最小二乘法参数估计有

$$\hat{B} = (X'X)^{-1}X'Y = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{pmatrix} = \begin{pmatrix} 3871.805 \\ 2.177916 \\ 4.051980 \end{pmatrix}$$

故所求回归方程为

$$Y = \quad + X_1 + \quad$$

$X_1$  的系数  $\beta_1=2.177916$  说明，如果其他变量保持不变，为使国民生产总值增加一亿元投资需增加 2.18 亿元，净出口增加 4.05 亿元也能使国民生产总值增加一亿元。

(2) 对偏回归系数及所建立的回归模型进展检验，显著性水平  $\alpha=0.05$ 。  $t_{0.025}(10) = 2.2281$

解：假设  $H_0: \beta_i = 0$ ，  $H_1: \beta_i \neq 0$ 。在  $H_0$  成立的条件下

$$\text{检验统计量 } t_1 = \frac{\hat{\beta}_1 - \beta_1}{S(\hat{\beta}_1)} = \frac{\hat{\beta}_1}{S(\hat{\beta}_1)} \sim t(n-k) \qquad t_2 = \frac{\hat{\beta}_2 - \beta_2}{S(\hat{\beta}_2)} = \frac{\hat{\beta}_2}{S(\hat{\beta}_2)} \sim t(n-k)$$

$$S(\hat{\beta}_1) = \hat{\sigma} C_{11} = \sqrt{\frac{\sum e_i^2}{n-k}} C_{11} = \qquad S(\hat{\beta}_2) = \hat{\sigma} C_{22} = \sqrt{\frac{\sum e_i^2}{n-k}} C_{22} =$$

其中  $C_{ii}$  是  $(X^T X)^{-1}$  对角线的值。  $\sum e_i^2 = \sum (Y_i - \hat{Y}_i)^2$ ，为残差平方和。

$$\text{所以： } t_1 = \frac{\hat{\beta}_1}{S(\hat{\beta}_1)} = \frac{2.177916}{0.120692} = \qquad t_2 = \frac{\hat{\beta}_2}{S(\hat{\beta}_2)} = \frac{4.051980}{1.282402} =$$

$$\text{给定 } \alpha=0.05. \quad w = \left\{ |t| \geq t_{\frac{\alpha}{2}}(n-k) \right\} = \left\{ |t| \geq t_{0.025}(10) \right\} = \left\{ |t| \geq 2.2281 \right\}。从上面结果看出 t$$

$t_1$ 、 $t_2$  的绝对值均大于 2.2281, 故拒绝  $H_0$ ，认为  $\beta_1$ 、 $\beta_2$  均显著不等于 0， $X_1$ 、 $X_2$  对  $Y$  的影响均显著。

(3) 估计可决系数，以显著性水平  $\alpha = 0.05$  对方程整体显著性进展检验，并估计校正



可决系数，说明其含义。  $F_{0.05}(2,10) = 9.39$

解:  $R^2 = 1 - \frac{RSS}{TSS} = 1 - \frac{e'e}{\sum (Y_i - \bar{Y})^2} =$

假设  $H_0: \beta_1 = \beta_2 = 0$ 。  $H_1: \beta_1, \beta_2$  不全为 0。

检验统计量  $F = \frac{ESS}{k} \bigg/ \frac{RSS}{n-k} = \frac{\sum (\hat{Y}_i - \bar{Y})^2}{k} \bigg/ \frac{\sum (Y_i - \hat{Y})^2}{n-k} =$

给定  $\alpha = 0.05$ ,  $w = \{F \geq F_{\alpha}(k, n-k)\} = \{F \geq F_{0.05}(2,10)\} = \{F \geq 9.39\}$ ,  $F$  远大于  $F(2,10)$ ,

故拒绝  $H_0$ , 认为总体参数  $\beta_1, \beta_2$  不全为等于 0, 资本形成额  $X_1$  和货物和效劳净出口  $X_2$  对国民生产总值  $Y$  的影响显著。

4、假设要求你建立一个计量经济模型来说明在学校跑道上慢跑一英里或一英里以上的人数, 以便决定是否修建第二条跑道以满足所有的锻炼者。你通过整个学年收集数据, 得到两个可能的解释性方程:

方程 A:  $\hat{Y} = 125.0 - 15.0X_1 - 1.0X_2 + 1.5X_3$   $\bar{R}^2 = 0.75$

方程 B:  $\hat{Y} = 123.0 - 14.0X_1 + 5.5X_2 - 3.7X_4$   $\bar{R}^2 = 0.73$

其中:  $Y$ —某天慢跑者的人数;  $X_1$ —该天降雨的英寸数;  $X_2$ —该天日照的小时数;  $X_3$ —该天的最高温度 (按华氏温度);  $X_4$ —第二天需交学期论文的班级数。

请答复以下问题:

(1) 这两个方程你认为哪个更合理些, 为什么?

(2) 为什么用一样的数据去估计一样变量的系数得到不同的符号?

答案:

(1) 方程 B 更合理些。原因是: 方程 B 中的参数估计值的符号及现实更接近些, 如日照的小时数同向变化, 天那么慢跑的人 would 多些; 及第二天需交学期论文的班级数成反向变化, 这一点在学校的跑道模型中是一个合理的解释变量。

(2) 解释变量的系数说明该变量的单位变化在方程中其他解释变量不变的条件下对被解释变量的影响, 在方程 A 和方程 B 中由于选择了不同的解释变量, 如方程 A 选择的是“该天的最高温度”而方程 B 选择的是“第二天需交学期论文的班级数”, 由此造成  $X_2$  及这两个变量之间的关系不同, 所以用一样的数据估计一样的变量得到不同的符号。

5、收集 1978-2001 年的消费额  $XF$  (亿元), 国内生产总值  $GDP$  (亿元) 资料, 建立消费函数, Eviews 结果如下:



Dependent Variable: LOG(XF)				
Method: Least Squares				
Date: 10/21/09 Time: 20:16				
Sample: 1978 2001				
Included observations: 24				
	Coefficient	Std. Error	t-Statistic	Prob.
C			t1=	
LOG(GDP)		84454	t2=	
R-squared		Mean dependent var		
Adjusted R-squared	8480	S.D. dependent var		
S.E. of regression		Akaike info criterion		
Sum squared resid		Schwarz criterion		
Log likelihood		Hannan-Quinn criter.		
F-statistic		Durbin-Watson stat		
Prob(F-statistic)				

- 要求：（1）把表中缺失的数据补上；（5 分）  
（2）把回归分析结果报告出来；（5 分）  
（3）进展经济意义、统计学意义和经济计量学意义检验；（6 分）  
（4）解释系数经济含义。（4 分）
- 6、根据广东省数据，把财政支出（CZ）作为因变量，财政收入（CS）作为解释变量进展一元回归分析后，得到回归残差平方的对数对 log（CS）的回归结果如下：

Dependent Variable: LOG(RESID^2)				
Method: Least Squares				
Date: 5/22/09 Time: 20:24				
Sample: 1978 2003				
Included observations: 26				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
LOG(CS)	1.522024	48743	6.118862	14
C				
R-squared	46731	Mean dependent var		26.24844
Adjusted R-squared	16398	S.D. dependent var		28.03465
S.E. of regression	15.475117	Akaike info criterion		7.824553
Sum squared resid	7310.458	Schwarz criterion		7.593271

- 要求：
- （1）写出异方差表达式  $\sigma_i^2=?$  （10 分）  
（2）进展同方差变换，证实变换后的模型不存在异方差。（10 分）

:  $CZ_t = \beta_0 + \beta_1CS_t + u_t$

其中：  $Var(u_t)=\sigma^2f(CS_t)$  (其中  $\sigma^2$ 为常数)， 其中  $f(CS_t)=(CS_t)^{1.522024}$

模型两边同时除以  $\sqrt{f(CS_t)}$  进展变换，得：（3 分）



$$\frac{CZ_t}{\sqrt{f(CS_t)}} = \frac{\beta_0}{\sqrt{f(CS_t)}} + \beta_1 \frac{CZ_t}{\sqrt{f(CS_t)}} + \frac{u_t}{\sqrt{f(CS_t)}} = \frac{\beta_0}{\sqrt{f(CS_t)}} + \beta_1 \frac{CZ_t}{\sqrt{f(CS_t)}} + v_t$$

其中： $v_t = \frac{u_t}{\sqrt{f(CS_t)}}$ ，可以证明误差项 $v_t$ 是同方差的。

证明如下：（4分）

∵  $v_t = \frac{u_t}{\sqrt{f(CS_t)}}$ ， $v_t^2 = \frac{u_t^2}{f(CS_t)}$ ， $E(v_t^2) = E\left(\frac{u_t^2}{f(CS_t)}\right) = \sigma^2$ （根据条件 $\sigma^2$ 为常数），证得变换后

的误差项是同方差的。

