

5.5元 计量经济学 第一套

一、单项选择题

- 1、双对数模型 $\ln Y = \ln \beta_0 + \beta_1 \ln X + \mu$ 中，参数 β_1 的含义是 (C)
- A. Y 关于 X 的增长率 B. Y 关于 X 的发展速度
C. Y 关于 X 的弹性 D. Y 关于 X 的边际变化
- 2、设 k 为回归模型中的参数个数，n 为样本容量。则对多元线性回归方程进行显著性检验时，所用的 F 统计量可表示为 (B)
- A. $\frac{ESS/(n-k)}{RSS/(k-1)}$ B. $\frac{R^2/(k-1)}{(1-R^2)/(n-k)}$
C. $\frac{R^2/(n-k)}{(1-R^2)/(k-1)}$ D. $\frac{ESS/(k-1)}{TSS/(n-k)}$
- 3、回归分析中使用的距离是点到直线的垂直坐标距离。最小二乘准则是指 (D)
- A. 使 $\left| \sum_{t=1}^n (Y_t - \hat{Y}_t) \right|$ 达到最小值 B. 使 $\min |Y_i - \hat{Y}_i|$ 达到最小值
C. 使 $\max |Y_t - \hat{Y}_t|$ 达到最小值 D. 使 $\sum_{t=1}^n (Y_t - \hat{Y}_t)^2$ 达到最小值
- ~~4、对于一个含有截距项的计量经济模型，若某定性因素有 m 个互斥的类型，为将其引入模型中，则需要引入虚拟变量个数为 (B)~~
- A. m B. m-1 C. m+1 D. m-k
- 5、回归模型中具有异方差性时，仍用 OLS 估计模型，则以下说法正确的是 (A)
- A. 参数估计值是无偏非有效的 B. 参数估计量仍具有最小方差性
C. 常用 F 检验失效 D. 参数估计量是有偏的
- 6、在一元线性回归模型中，样本回归方程可表示为 (C)
- A. $Y_t = \beta_0 + \beta_1 X_t + u_t$ B. $Y_t = E(Y_t / X) + \mu_i$
C. $\hat{Y}_t = \hat{\beta}_0 + \hat{\beta}_1 X_t$ D. $E(Y_t / X_t) = \beta_0 + \beta_1 X_t$
- ~~7、在经济发展发生转折时期，可以通过引入虚拟变量方法来表示这种变化。例如，研究中国城镇居民消费函数时。1991 年前后，城镇居民商品性实际支出 Y 对实际可支配收入 X 的回归关系明显不同。现以 1991 年为转折时期，设虚拟变量 $D_t = \begin{cases} 1, & 1991 \text{年以后} \\ 0, & 1991 \text{年以前} \end{cases}$ ，数据散点图显示消费函数发生了结构性变化：基本~~

消费部分下降了，边际消费倾向变大了。则城镇居民线性消费函数的理论方程可以写作（ D ）

- A. $Y_t = \beta_0 + \beta_1 X_t + u_t$ B. $Y_t = \beta_0 + \beta_1 X_t + \beta_2 D_t X_t + u_t$
 C. $Y_t = \beta_0 + \beta_1 X_t + \beta_2 D_t + u_t$ D. $Y_t = \beta_0 + \beta_1 X_t + \beta_2 D_t + \beta_3 D_t X_t + u_t$

8、对于有限分布滞后模型

$$Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \cdots + \beta_k X_{t-k} + u_t$$

在一定条件下，参数 β_i 可近似用一个关于 i 的阿尔蒙多项式表示（ $i = 0, 1, 2, \dots, m$ ），

其中多项式的阶数 m 必须满足（ A ）

- A. $m < k$ B. $m = k$ C. $m > k$ D. $m \geq k$

9、在自适应预期模型和库伊克模型中，假定原始模型的随机扰动项 u_t 满足古典线性回归模型的所有假设，则对于这两个模型中的滞后解释变量 Y_{t-1} 和误差项 u_t^* ，下列说法正确的有（ D ）

- A. $Cov(Y_{t-1}, u_t^*) = 0, Cov(u_t^*, u_{t-1}^*) = 0$
 B. $Cov(Y_{t-1}, u_t^*) = 0, Cov(u_t^*, u_{t-1}^*) \neq 0$
 C. $Cov(Y_{t-1}, u_t^*) \neq 0, Cov(u_t^*, u_{t-1}^*) = 0$
 D. $Cov(Y_{t-1}, u_t^*) \neq 0, Cov(u_t^*, u_{t-1}^*) \neq 0$

10、设 u_t 为随机误差项，则一阶线性自相关是指（ B ）

- A. $Cov(u_t, u_s) \neq 0 (t \neq s)$ B. $u_t = \rho u_{t-1} + \varepsilon_t$
 C. $u_t = \rho_1 u_{t-1} + \rho_2 u_{t-2} + \varepsilon_t$ D. $u_t = \rho^2 u_{t-1} + \varepsilon_t$

11、利用德宾 h 检验自回归模型扰动项的自相关性时，下列命题正确的是（ B ）

- A. 德宾 h 检验只适用一阶自回归模型
 B. 德宾 h 检验适用任意阶的自回归模型
 C. 德宾 h 统计量渐进服从 t 分布
 D. 德宾 h 检验可以用于小样本问题

12、关于联立方程组模型，下列说法中错误的是（ B ）

- A. 结构式模型中解释变量可以是内生变量，也可以是前定变量
 B. 简化式模型中解释变量可以是内生变量，
 C. 简化式模型中解释变量是前定变量
 D. 结构式模型中解释变量可以是内生变量

13、以下选项中，正确地表达了序列相关的是（ A ）

A. $Cov(\mu_i, \mu_j) \neq 0, i \neq j$ B. $Cov(\mu_i, \mu_j) = 0, i \neq j$

C. $Cov(X_i, X_j) = 0, i \neq j$ D. $Cov(X_i, \mu_j) \neq 0, i \neq j$

14、一元线性回归分析中的回归平方和 ESS 的自由度是（ D ）

A. n B. $n-1$ C. $n-k$ D. 1

15、边际成本函数为 $MC = \alpha + \beta_1 Q + \beta_2 Q^2 + \mu$ （MC 表示边际成本；Q 表示产量），则下列说法正确的有（ A ）

A. 模型中可能存在多重共线性 B. 模型中不应包括 Q^2 作为解释变量

C. 模型为非线性模型 D. 模型为线性模型

~~16、如果某个结构方程是恰好识别的，估计其参数可用（ D ）~~

A. 最小二乘法

B. 极大似然法

C. 广义差分法

D. 间接最小二乘法

17、已知样本回归模型残差的一阶自相关系数接近于 1，则 DW 统计量近似等于（ A ）

A. 0

B. 1

C. 2

D. 4

18、更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

19、设 M 为货币需求量， Y 为收入水平， r 为利率，流动性偏好函数为 $M = \beta_0 + \beta_1 Y + \beta_2 r + \mu$ ，又设 $\hat{\beta}_1$ 、 $\hat{\beta}_2$ 分别是 β_1 、 β_2 的估计值，则根据经济理论，一般来说（ A ）

A. $\hat{\beta}_1$ 应为正值， $\hat{\beta}_2$ 应为负值

B. $\hat{\beta}_1$ 应为正值， $\hat{\beta}_2$ 应为正值

C. $\hat{\beta}_1$ 应为负值， $\hat{\beta}_2$ 应为负值

D. $\hat{\beta}_1$ 应为负值， $\hat{\beta}_2$ 应为正值

20、对于有限分布滞后模型，解释变量的滞后长度每增加一期，可利用的样本数据就会（ B ）

A. 增加 1 个

B. 减少 1 个

C. 增加 2 个

D. 减少 2 个

~~二、多项选择题~~

1、对联立方程模型参数的单一方程估计法包括（ A B D F ）

A. 工具变量法

B. 间接最小二乘法

C. 完全信息极大似然估计法

D. 二阶段最小二乘法

E. 三阶段最小二乘法

F. 有限信息极大似然估计法

2、下列哪些变量一定属于前定变量（ C D ）

A. 内生变量

B. 随机变量

C. 滞后变量

D. 外生内生变量

E. 工具变量

- 3、古典线性回归模型的普通最小二乘估计量的特性有（ **A B C** ）
- A. 无偏性 B. 线性性 C. 最小方差性
- D. 不一致性 E. 有偏性
- 4、利用普通最小二乘法求得的样本回归直线 $\hat{Y}_i = \hat{\beta}_1 + \hat{\beta}_2 X_i$ 的特点（ **A C D** ）
- A. 必然通过点 (\bar{X}, \bar{Y}) B. 可能通过点 (\bar{X}, \bar{Y})
- C. 残差 e_i 的均值为常数 D. \hat{Y}_i 的平均值与 Y_i 的平均值相等
- E. 残差 e_i 与解释变量 X_i 之间有一定的相关性
- 5、关于联立方程模型识别问题，以下说法不正确的有（ **A B** ）
- A. 满足阶条件的方程则可识别
- B. 如果一个方程包含了模型中的全部变量，则这个方程恰好识别
- C. 如果一个方程包含了模型中的全部变量，则这个方程不可识别
- D. 如果两个方程包含相同的变量，则这两个方程均不可识别
- E. 联立方程组中的每一个方程都是可识别的，则联立方程组才可识别
- F. 联立方程组中有一个方程不可识别，则联立方程组不可识别

三、判断题（判断下列命题正误，并说明理由）

- 1、简单线性回归模型与多元线性回归模型的基本假定是相同的。

错

在多元线性回归模型里除了对随机误差项提出假定外，还对解释变量之间提出无多重共线性的假定。

- 2、在模型中引入解释变量的多个滞后项容易产生多重共线性。

对

在分布滞后模型里多引进解释变量的滞后项，由于变量的经济意义一样，只是时间不一致，所以很容易引起多重共线性。

- 3、DW 检验中的 d 值在 0 到 4 之间，数值越小说明模型随机误差项的自相关度越小，数值越大说明模型随机误差项的自相关度越大。

错

DW 值在 0 到 4 之间，当 DW 落在最左边 $(0 < d < d_L)$ 、最右边 $(4 - d_L < d < 4)$

时，分别为正自相关、负自相关；

中间($d_U < d < 4 - d_U$)为不存在自相关区域；

其次为两个不能判定区域。

4、在计量经济模型中，随机扰动项与残差项无区别。

错

它们均为随机项，但随机误差项表示总体模型的误差，残差表示样本模型的误差；另外，残差=随机误差项+参数估计误差。

5、在经济计量分析中，模型参数一旦被估计出来，就可将估计模型直接运用于实际的计量经济分析。

错

参数一经估计，建立了样本回归模型，还需要对模型进行检验，包括经济意义检验、统计检验、计量经济专门检验等。

四、计算题

1、根据某城市 1978——1998 年人均储蓄(y)与人均收入(x)的数据资料建立了如下回归模型

$$\hat{y} = -2187.521 + 1.6843x$$

$$se = (340.0103) (0.0622)$$

$$R^2 = 0.9748, S.E. = 1065.425, DW = 0.2934, F = 733.6066$$

试求解以下问题

(1) 取时间段 1978——1985 和 1991——1998，分别建立两个模型。

模型 1: $\hat{y} = -145.4415 + 0.3971x$

模型 2: $\hat{y} = -4602.365 + 1.9525x$

$$t = (-8.7302) (25.4269)$$

$$t = (-5.0660) (18.4094)$$

$$R^2 = 0.9908, \sum e_1^2 = 1372.202$$

$$R^2 = 0.9826, \sum e_2^2 = 5811189$$

计算 F 统计量，即 $F = \sum e_2^2 / \sum e_1^2 = 5811189 / 1372.202 = 4334.9370$ ，对给定的 $\alpha = 0.05$ ，查 F 分布表，得临界值 $F_{0.05}(6,6) = 4.28$ 。请你继续完成上述工作，并回答所做的是一项什么工作，其结论是什么？

解：该检验为 Goldfeld-Quandt 检验
因为 $F=4334.937>4.28$ ，所以模型存在异方差

(2) 根据表 1 所给资料，对给定的显著性水平 $\alpha = 0.05$ ，查 χ^2 分布表，得临界值 $\chi_{0.05}(3) = 7.81$ ，其中 $p=3$ 为自由度。请你继续完成上述工作，并回答所做的是一项什么工作，其结论是什么？

表 1

ARCH Test:

F-statistic	6.033649	Probability	0.007410
Obs*R-squared	10.14976	Probability	0.017335

Test Equation:

Dependent Variable: RESID^2

Method: Least Squares

Date: 06/04/05 Time: 17:02

Sample(adjusted): 1981 1998

Included observations: 18 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	244797.2	373821.3	0.654851	0.5232
RESID^2(-1)	1.226048	0.330479	3.709908	0.0023
RESID^2(-2)	-1.405351	0.379187	-3.706222	0.0023
RESID^2(-3)	1.015853	0.328076	3.096397	0.0079
R-squared	0.563876	Mean dependent var		971801.3
Adjusted R-squared	0.470421	S.D. dependent var		1129283.
S.E. of regression	821804.5	Akaike info criterion		30.26952
Sum squared resid	9.46E+12	Schwarz criterion		30.46738
Log likelihood	-268.4257	F-statistic		6.033649
Durbin-Watson stat	2.124575	Prob(F-statistic)		0.007410

解：该检验为 ARCH 检验

- (1) 由 Obs*R-squared=10.1498>7.81，表明模型存在异方差；
- (2) 由各系数的 t-值可知，残差各阶滞后系数均大于 2，表明各阶滞后对 RESID

均有影响，揭示存在异方差。

- 2、根据某行业 1955——1974 年的库存量（y）和销售量（x）的资料（见表 2），运用 EViews 软件得如下报告资料，试根据所给资料和图形完成下列问题：
- (1) 完成表 2 的空白处，由报告资料写出估计模型的表达式（用标准书写格式）；
- (2) 根据写出的模型表达式求销售量对库存量影响的短期乘数、动态乘数和长期乘数，同时给出经济解释；
- (3) 根据所给资料对估计模型进行评价（包括经济意义、拟合效果、显著性检验等）。

表 2

Dependent Variable: Y
Method: Least Squares
Date: 06/04/05 Time: 17:42
Sample(adjusted): 1958 1974
Included observations: 17 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-6.419601	2.130157		-
PDL01	1.156862	0.195928		-
PDL02	0.065752	0.176055		-
PDL03	-0.460829	0.181199		-
R-squared	0.996230	Mean dependent var		81.97653
Adjusted R-squared		S.D. dependent var		27.85539
S.E. of regression	1.897384	Akaike info criterion		4.321154
Sum squared resid	46.80087	Schwarz criterion		4.517204
Log likelihood	-32.72981	F-statistic		
Durbin-Watson stat	1.513212	Prob(F-statistic)		0.000000
Lag Distribution of X	i	Coefficient	Std. Error	T-Statistic
. *	0	0.63028	0.17916	
. *	1	1.15686	0.19593	
. *	2	0.76178	0.17820	
* .	3	-0.55495	0.25562	
Sum of Lags		1.99398	0.06785	

$t_{(17)}(0.025) = 2.110, t_{(13)}(0.025) = 2.160, t_{(12)}(0.025) = 2.176,$
 $t_{(17)}(0.05) = 1.740, t_{(13)}(0.05) = 1.771, t_{(12)}(0.05) = 1.782$
 $F_{(4,12)}(0.05) = 3.26, F_{(5,13)}(0.05) = 3.03, F_{(5,17)}(0.05) = 2.81$

解：（1）第一栏的 t 统计量值：

T-Statistic
-3.013675
5.904516
0.373472
-2.513216

第二栏的 t 统计量值：

T-Statistic
3.51797
5.90452
4.27495
-2.17104

Adjusted R-squared	0.99536
F-statistic	1145.20

$$\hat{y}_t = -6.4196 + 0.6303x_t + 1.1569x_{t-1} + 0.7618x_{t-2} - 0.5550x_{t-3}$$
$$t = (-3.0137)(3.5180) \quad (5.9045) \quad (4.2750) \quad (-2.1710)$$
$$\bar{R}^2 = 0.9954, \quad DW = 1.5132, \quad F = 1145.16$$

（2）短期乘数为 0.6303，动态乘数分别为 1.1569，0.7618，-0.5550。长期乘数为 1.994。

（3）模型整体的拟合效果较好，可决系数达到 0.9963，F 统计量为 1145.16，除 x_{t-3} 的系数的 t 统计量外，其余均大于在显著性水平为 0.05，自由度为 12 下的临界值 2.176，说明模型中销售额在滞后第三期对库存量影响较小外，其它各均影响显著。

3、根据某地区居民对农产品的消费 y 和居民收入 x 的样本资料，应用最小二乘法估计模型，估计结果如下，拟合效果见图。由所给资料完成以下问题：

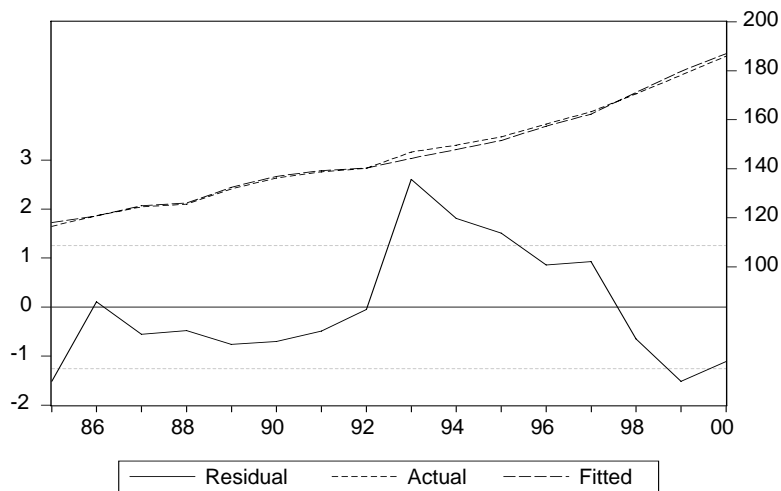
(1) 在 $n=16$, $\alpha=0.05$ 的条件下, 查 D-W 表得临界值分别为 $d_L=1.106, d_U=1.371$, 试判断模型中是否存在自相关;

(2) 如果模型存在自相关, 求出相关系数 $\hat{\rho}$, 并利用广义差分变换写出无自相关的广义差分模型。

$$\hat{y} = 27.9123 + 0.3524x$$

$$se = (1.8690) (0.0055)$$

$$R^2 = 0.9966, \sum_{i=1}^{16} e_i^2 = 22.0506, DW = 0.6800, F = 4122.531$$



解: (1) 因为 $DW=0.68 < 1.106$, 所以模型中的随机误差存在正的自相关。

(2) 由 $DW=0.68$, 计算得 $\hat{\rho}=0.66$, 所以广义差分表达式为

$$y_t - 0.66y_{t-1} = 0.34\beta_1 + \beta_2(x_t - 0.66x_{t-1}) + u_t - 0.66u_{t-1}$$

第二套

一、单项选择题

1、把反映某一总体特征的同一指标的数据，按一定的时间顺序和时间间隔排列起来，这样的数据称为（ **B** ）

- A. 横截面数据 B. 时间序列数据
C. 修匀数据 D. 原始数据

2、多元线性回归分析中，调整后的可决系数 \bar{R}^2 与可决系数 R^2 之间的关系（ **A** ）

- A. $\bar{R}^2 = 1 - (1 - R^2) \frac{n-1}{n-k}$ B. $\bar{R}^2 \geq R^2$
C. $\bar{R}^2 > 0$ D. $\bar{R}^2 = 1 - (1 - R^2) \frac{n-k}{n-1}$

~~3、半对数模型 $Y_i = \beta_1 + \beta_2 \ln X_i + u_i$ 中，参数 β_2 的含义是（ **D** ）~~

- A. Y 关于 X 的弹性
B. X 的绝对量变动，引起 Y 的绝对量变动
C. Y 关于 X 的边际变动
D. X 的相对变动，引起 Y 的期望值绝对量变动

4、已知五元标准线性回归模型估计的残差平方和为 $\sum e_i^2 = 800$ ，样本容量为 46，则随机误差项 u_i 的方差估计量 $\hat{\sigma}^2$ 为（ **D** ）

- A. 33.33 B. 40 C. 38.09 D. 20

5、现用 OLS 法得到的样本回归直线为 $Y_i = \hat{\beta}_1 + \hat{\beta}_2 X_i + e_i$ ，以下说法不正确的是（ **B** ）

- A. $\sum e_i = 0$ B. $Cov(X_i, e_i) \neq 0$
C. $\bar{\hat{Y}} = \bar{Y}$ D. (\bar{X}, \bar{Y}) 在回归直线上

6、Goldfeld-Quandt 检验法可用于检验（ **A** ）

- A. 异方差性 B. 多重共线性 C. 序列相关 D. 设定误差

7、用于检验序列相关的 DW 统计量的取值范围是（ **D** ）

- A. $0 \leq DW \leq 1$ B. $-1 \leq DW \leq 1$
C. $-2 \leq DW \leq 2$ D. $0 \leq DW \leq 4$

~~8、对联立方程组模型估计的方法主要有两类，即（ **A** ）~~

- A. 单一方程估计法和系统估计法
B. 间接最小二乘法和系统估计法

C. 单一方程估计法和二阶段最小二乘法

D. 工具变量法和间接最小二乘法

9、在模型 $Y_t = \beta_1 + \beta_2 X_{2t} + \beta_3 X_{3t} + u_t$ 的回归分析结果报告中，有

$F = 263489.23$ ， F 的 p 值=0.000000，则表明（ **C** ）

A. 解释变量 X_{2t} 对 Y_t 的影响是显著的

B. 解释变量 X_{3t} 对 Y_t 的影响是显著的

C. 解释变量 X_{2t} 和 X_{3t} 对 Y_t 的联合影响是显著的.

D. 解释变量 X_{2t} 和 X_{3t} 对 Y_t 的影响是均不显著

10、如果回归模型中解释变量之间存在完全的多重共线性，则最小二乘估计量的值为（ **A** ）

A. 不确定，方差无限大 B. 确定，方差无限大

C. 不确定，方差最小 D. 确定，方差最小

在序列自相关的情况下，参数估计值仍是无偏的，其原因是（ **C** ）

A. 无多重共线性假定成立 B. 同方差假定成立

C. 零均值假定成立 D. 解释变量与随机误差项不相关假定成立

11、应用 DW 检验方法时应满足该方法的假定条件，下列不是其假定条件的为（ **B** ）

A. 解释变量为非随机的

B. 被解释变量为非随机的

C. 线性回归模型中不能含有滞后内生变量

D. 随机误差项服从一阶自回归

12、在具体运用加权最小二乘法时，如果变换的结果是

$$\frac{Y_i}{X_i} = \beta_1 \frac{1}{X_i} + \beta_2 \frac{X_i}{X_i} + \frac{u_i}{X_i}$$

则 $\text{Var}(u_i)$ 是下列形式中的哪一种? (**B**)

A. $\sigma^2 x$ B. $\sigma^2 x^2$ C. $\sigma^2 \sqrt{x}$ D. $\sigma^2 \log x$

13、经济变量的时间序列数据大多存在序列相关性，在分布滞后模型中，这种序列相关性就转化为（ **B** ）

A. 异方差问题

B. 多重共线性问题

C. 序列相关性问题

D. 设定误差问题

14、关于自适应预期模型和局部调整模型，下列说法错误的有（ **D** ）

A. 它们都是由某种期望模型演变形成的

- B. 它们最终都是一阶自回归模型
C. 它们的经济背景不同
D. 都满足古典线性回归模型的所有假设，故可直接用 OLS 方法进行估计

15、设某地区消费函数中，消费支出不仅与收入 X 有关，而且与消费者的年龄构成有关，若将年龄构成分为小孩、青年人、成年人和老年人 4 个层次。假设边际消费倾向不变，考虑上述年龄构成因素的影响时，该消费函数引入虚拟变量的个数为 (C)

- A. 1 个 B. 2 个 C. 3 个 D. 4 个

16、个人保健支出的计量经济模型为： $Y_i = \alpha_1 + \alpha_2 D_{2i} + \beta X_i + u_i$ ，其中 Y_i 为保健年度支出； X_i 为个人年度收入；虚拟变量 $D_{2i} = \begin{cases} 1 & \text{大学及以上} \\ 0 & \text{大学以下} \end{cases}$ ； u_i 满足古典假定。则大学以上群体的平均年度保健支出为 (B)

- A. $E(Y_i | X_i, D_{2i} = 0) = \alpha_1 + \beta X_i$ B. $E(Y_i | X_i, D_{2i} = 1) = \alpha_1 + \alpha_2 + \beta X_i$
C. $\alpha_1 + \alpha_2$ D. α_1

17、在联立方程结构模型中，对模型中的每一个随机方程单独使用普通最小二乘法得到的估计参数是 (B)

- A. 有偏且一致的 B. 有偏不一致的
C. 无偏但一致的 D. 无偏且不一致的

18、下列宏观经济计量模型中投资 (I) 函数所在方程的类型为 (D)

$$\begin{cases} Y_t = C_t + I_t + G_t \\ C_t = \alpha_0 + \alpha_1 Y_t + u_{1t} \\ I_t = \beta_0 + \beta_1 Y_{t-1} + \beta_2 \gamma_t + u_{2t} \end{cases}$$

- A. 技术方程式 B. 制度方程式
C. 恒等式 D. 行为方程式

19、在有 M 个方程的完备联立方程组中，若用 H 表示联立方程组中全部的内生变量与全部的前定变量之和的总数，用 N_i 表示第 i 个方程中内生变量与前定变量之和的总数时，第 i 个方程过度识别时，则有公式(**A**)成立。

~~A. $H - N_i > M - 1$~~

~~B. $H - N_i = M - 1$~~

~~C. $H - N_i = 0$~~

~~D. $H - N_i < M - 1$~~

20、对自回归模型进行估计时，假定原始模型的随机扰动项满足古典线性回归模型的所有假设，则估计量是一致估计量的模型有(**B**)

~~A. 库伊克模型~~

~~B. 局部调整模型~~

~~C. 自适应预期模型~~

~~D. 自适应预期和局部调整混合模型~~

二、多项选择题

1、设一阶自回归模型是库伊克模型或自适应预期模型，估计模型时可用工具变量替代滞后内生变量，该工具变量应该满足的条件有(**AE**)

A. 与该滞后内生变量高度相关

B. 与其它解释变量高度相关

C. 与随机误差项高度相关

D. 与该滞后内生变量不相关

E. 与随机误差项不相关

2、计量经济模型的检验一般包括内容有(**ABCD**)

A、经济意义的检验

B、统计推断的检验

C、计量经济学的检验

D、预测检验

E、对比检验

3、以下变量中可以作为解释变量的有(**ABCDE**)

A. 外生变量

B. 滞后内生变量

C. 虚拟变量

D. 前定变量

E. 内生变量

4、广义最小二乘法的特殊情况是(**BD**)

A. 对模型进行对数变换

B. 加权最小二乘法

C. 数据的结合

D. 广义差分法

E. 增加样本容量

5、对美国储蓄与收入关系的计量经济模型分成两个时期分别建模，重建时期是 1946—1954;重建后时期是 1955—1963，模型如下：

重建时期: $Y_t = \lambda_1 + \lambda_2 X_t + \mu_{1t}$

重建后时期: $Y_t = \lambda_3 + \lambda_4 X_t + \mu_{2t}$

关于上述模型, 下列说法正确的是 (**A B C D**)

A. $\lambda_1 = \lambda_3, \lambda_2 = \lambda_4$ 时则称为重合回归 B. $\lambda_1 \neq \lambda_3, \lambda_2 = \lambda_4$ 时称为平行回归

C. $\lambda_1 = \lambda_3, \lambda_2 \neq \lambda_4$ 时称为共点回归 D. $\lambda_1 \neq \lambda_3, \lambda_2 \neq \lambda_4$ 时称为相异回归

E. $\lambda_1 \neq \lambda_3, \lambda_2 = \lambda_4$ 时, 表明两个模型在统计意义上无差异

三、判断题 (判断下列命题正误, 并说明理由)

1、线性回归模型意味着因变量是自变量的线性函数。

错

线性回归模型本质上指的是参数线性, 而不是变量线性。同时, 模型与函数不是同一回事。

2、多重共线性问题是随机扰动项违背古典假定引起的。

错

应该是解释变量之间高度相关引起的。

~~3、通过虚拟变量将属性因素引入计量经济模型, 引入虚拟变量的个数与样本容量大小有关。~~

错

~~引入虚拟变量的个数样本容量大小无关, 与变量属性, 模型有无截距项有关。~~

4、双变量模型中, 对样本回归函数整体的显著性检验与斜率系数的显著性检验是一致的。

正确

要求最好能够写出一元线性回归中, F 统计量与 t 统计量的关系, 即 $F = t^2$ 的来历; 或者说明一元线性回归仅有一个解释变量, 因此对斜率系数的 t 检验等价于对方程的整体性检验。

~~5、如果联立方程模型中某个结构方程包含了所有的变量, 则这个方程不可识别。~~

正确

没有唯一的统计形式

四、计算题

1、家庭消费支出（Y）、可支配收入（ X_1 ）、个人个财富（ X_2 ）设定模型如

下: $Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \mu_i$

回归分析结果为:

LS // Dependent Variable is Y

Date: 18/4/05 Time: 15:18

Sample: 1 10

Included observations: 10

Variable	Coefficient	Std. Error	T-Statistic	Prob.
C	24.4070	6.9973		0.0101
X_2	-0.3401	0.4785		0.5002
X_3	0.0823	0.0458		0.1152
R-squared		Mean dependent var		111.1256
Adjusted R-squared		S.D. dependent var		31.4289
S.E. of regression		Akaike info criterion		4.1338
Sum squared resid	342.5486	Schwartz criterion		4.2246
Loglikelihood	-31.8585	F-statistic		
Durbin-Watson stat	2.4382	Prob(F-statistic)		0.0001

回答下列问题

- (1) 请根据上表中已有的数据, 填写表中画线处缺失结果 (注意给出计算步骤);
- (2) 模型是否存在多重共线性? 为什么?
- (3) 模型中是否存在自相关? 为什么?

在0.05显著性水平下, d_l 和 d_u 的显著性点

	$k'=1$		$k'=2$	
n	d_l	d_u	d_l	d_u
9	0.824	1.32	0.629	1.699
10	0.879	1.32	0.697	1.641
11	0.927	1.324	0.658	1.604

答: (1)

Variable	Coefficient	Std. Error	T-Statistic	Prob.
C	24.4070	6.9973	<u>3.4881</u>	0.0101
X_2	-0.3401	0.4785	<u>-0.7108</u>	0.5002
X_2	0.0823	0.0458	<u>1.7969</u>	0.1152

R-squared	<u>0.9615</u>	Mean dependent var	111.1256
Adjusted R-squared	<u>0.9505</u>	S.D. dependent var	31.4289
S.E. of regression	<u> </u>	Akaike info criterion	4.1338
Sum squared resid	342.5486	Schwartz criterion	4.2246
Log likelihood	- 31.8585	F-statistic	<u>87.3336</u>
Durbin-Watson stat	2.4382	Prob(F-statistic)	0.0001

(2) 存在多重共线性； F 统计量和 R^2 显示模型很显著，但变量的 t 检验值都偏小。

(3) $n=10, k'=2$ ，查表 $d_L=0.697$ ； $d_U=1.641$ ； $4-d_L=3.303$ ； $4-d_U=2.359$ 。
DW=2.4382>2.359，因此不能确定模型是否存在自相关。

2、根据某城市 1978—1998 年人均储蓄与人均收入的数据资料建立了如下回归模型：

$$\hat{y} = -2187.521 + 1.6843x$$

$$se = (340.0103) (0.0622)$$

$$R^2 = 0.9748, \quad s.e. = 1065.425, \quad DW = 0.2934, \quad F = 733.6066$$

试求解以下问题：

(1) 取时间段 1978—1985 和 1991—1998，分别建立两个模型。

模型 1: $\hat{y} = -145.4415 + 0.3971x$

$$t = (-8.7302) (25.4269)$$

$$R^2 = 0.9908, \quad \sum e_1^2 = 1372.202$$

模型 2: $\hat{y} = -4602.365 + 1.9525x$

$$t = (-5.0660) (18.4094)$$

$$R^2 = 0.9826, \quad \sum e_2^2 = 5811189$$

计算 F 统计量，即 $F = \sum e_2^2 / \sum e_1^2 = 5811189 / 1372.202 = 4334.9370$ ，给定 $\alpha = 0.05$ ，查 F 分布表，得临界值 $F_{0.05}(6,6) = 4.28$ 。请你继续完成上述工作，并回答所做的是一项什么工作，其结论是什么？

(2) 利用 y 对 x 回归所得的残差平方构造一个辅助回归函数：

$$\hat{\sigma}_t^2 = 242407.2 + 1.2299\hat{\sigma}_{t-1}^2 - 1.4090\hat{\sigma}_{t-2}^2 + 1.0188\hat{\sigma}_{t-3}^2$$

$$R^2 = 0.5659, \text{ 计算 } (n-p)R^2 = 18 * 0.5659 = 10.1862$$

给定显著性水平 $\alpha = 0.05$ ，查 χ^2 分布表，得临界值 $\chi_{0.05}^2(3) = 7.81$ ，其中，自由度 $p=3$ 。请你继续完成上述工作，并回答所做的是一项什么工作，其结论是什么？

(3) 试比较 (1) 和 (2) 两种方法，给出简要评价。

答：(1) 这是异方差检验，使用的是样本分段拟和 (Goldfeld-Quant)， $F = 4334.937 > 4.28$ ，因此拒绝原假设，表明模型中存在异方差。

(2) 这是异方差 ARCH 检验， $(n-p)R^2 = 18 * 0.5659 = 10.1862 > 7.81$ ，所以拒绝原假设，表明模型中存在异方差。

(3) 这两种方法都是用于检验异方差。但二者适用条件不同：

A、Goldfeld-Quant 要求大样本；扰动项正态分布；可用于截面数据和时间序列数据。

B、ARCH 检验仅适宜于时间序列数据，且其渐进分布为 χ^2 -分布。

~~3、Sen 和 Srivastava (1971) 在研究贫富国之间期望寿命的差异时，利用 101 个国家的数据，建立了如下的回归模型：~~

~~$$\bar{Y}_i = -2.40 + 9.39 \ln X_i - 3.36(D_i(\ln X_i - 7))$$~~

~~$$(4.37) \quad (0.857) \quad (2.42)$$~~

~~$$R^2 = 0.752$$~~

~~其中：X 是以美元计的人均收入；~~

~~Y 是以年计的期望寿命；~~

~~Sen 和 Srivastava 认为人均收入的临界值为 1097 美元 ($\ln 1097 = 7$)，若人均收入超过 1097 美元，则被认定为富国；若人均收入低于 1097 美元，被认定为贫穷国。~~

~~(括号内的数值为对应参数估计值的 t-值)。~~

~~(1) 解释这些计算结果。~~

~~(2) 回归方程中引入 $D_i(\ln X_i - 7)$ 的原因是什么？如何解释这个回归解释变量？~~

~~(3) 如何对贫穷国进行回归？又如何对富国进行回归？~~

解：(1) 由 $\ln X = 1 \Rightarrow X = 2.7183$ ，也就是说，人均收入每增加 1.7183 倍，平均意义上各国的期望寿命会增加 9.39 岁。若当为富国时， $D_i = 1$ ，则平均意义

上，富国的人均收入每增加 1.7183 倍，其期望寿命就会减少 3.36 岁，但其截距项的水平会增加 23.52，达到 21.12 的水平。但从统计检验结果看，对数人均收入 $\ln X$ 对期望寿命 Y 的影响并不显著。方程的拟合情况良好，可进一步进行多重共线性等其他计量经济学的检验。

(2) 若 $D_i = 1$ 代表富国，则引入 $D_i(\ln X_i - 7)$ 的原因是想从截距和斜率两个方面考证富国的影响，其中，富国的截距为 $(-2.40 + 3.36 \times 7 = 21.12)$ ，斜率为 $(9.39 - 3.36 = 6.03)$ ，因此，当富国的人均收入每增加 1.7183 倍，其期望寿命会增加 6.03 岁。

(3) 对于贫穷国，设定 $D_i = \begin{cases} 1 & \text{若为贫穷国} \\ 0 & \text{若为富国} \end{cases}$ ，则引入的虚拟解释变量的形式

为 $(D_i(7 - \ln X_i))$ ；对于富国，回归模型形式不变。

第三套

一、单项选择题

1、对样本的相关系数 γ ，以下结论错误的是（ **A** ）

A. $|\gamma|$ 越接近 0， X 与 Y 之间线性相关程度高

B. $|\gamma|$ 越接近 1， X 与 Y 之间线性相关程度高

C. $-1 \leq \gamma \leq 1$

D. $\gamma = 0$ ，则在一定条件下 X 与 Y 相互独立

2、同一时间，不同单位相同指标组成的观测数据称为（ **B** ）

A. 原始数据

B. 截面数据

C. 时间序列数据

D. 修匀数据

3、为了分析随着解释变量变动一个单位，因变量的增长率变化情况，模型应该设定为（ **C** ）

~~A. $\ln Y = \beta_1 + \beta_2 \ln X + u$~~

~~B. $Y = \beta_0 + \beta_1 \ln X + u$~~

C. $\ln Y = \alpha_0 + \alpha_1 X + u$

D. $Y_i = \beta_1 + \beta_2 X_i + u_i$

4、多元线性回归模型中，发现各参数估计量的 t 值都不显著，但模型的 R^2 或

\bar{R}^2 却很大， F 值也很显著，这说明模型存在（ **A** ）

A. 多重共线性

B. 异方差

C. 自相关

D. 设定偏误

5、在异方差性情况下，常用的估计方法是（ **D** ）

A. 一阶差分法

B. 普通最小二乘法

C. 工具变量法

D. 广义差分法

6、DW 检验中要求有假定条件，在下列条件中不正确的是（ **D** ）

A. 解释变量为非随机的

B. 随机误差项为一阶自回归形式

C. 线性回归模型中不应含有滞后内生变量为解释变量

D. 线性回归模型只能为一元回归形式

7、广义差分法是（ **B** ）的一个特例

A. 加权最小二乘法

B. 广义最小二乘法

C. 普通最小二乘法

D. 两阶段最小二乘法

8、在下列引起序列自相关的原因中，不正确的是（ **D** ）

A. 经济变量具有惯性作用

B. 经济行为的滞后性

C. 设定偏误

D. 解释变量之间的共线性

9、假设估计出的库伊克（Koyck）模型如下：

$$\hat{Y}_t = -6.9 + 0.35X_t + 0.76Y_{t-1}$$

$$t = (-2.6521) \quad (4.70) \quad (11.91)$$

$$R^2 = 0.897 \quad F = 143 \quad DW = 1.916$$

则 (**C**)

A. 分布滞后系数的衰减率为 0.34

B. 在显著性水平 $\alpha = 0.05$ 下, DW 检验临界值为 $d_L = 1.3$, 由于 $DW = d = 1.916$

$> d_L = 1.3$, 据此可以推断模型扰动项存在自相关

C. 即期消费倾向为 0.35, 表明收入每增加 1 元, 当期的消费将增加 0.35 元

D. 收入对消费的长期影响乘数为 Y_{t-1} 的估计系数 0.76

~~10. 虚拟变量 (**A**)~~

A. 主要来代表质的因素, 但在有些情况下可以用来代表数量因素

B. 只代表质的因素

C. 只代表数量因素

D. 只代表季节影响因素

~~11. 若想考察某两个地区的平均消费水平是否存在显著差异, 则下列那个模型比较适合 (Y 代表消费支出, X 代表可支配收入, D_2, D_3 表示虚拟变量) (**D**)~~

A. $Y_i = \alpha + \beta X_i + u_i$

B. $Y_i = \alpha_1 + \beta_1 X_i + \beta_2 (D_{2i} X_i) + \mu_i$

C. $Y_i = \alpha_1 + \alpha_2 D_{2i} + \alpha_3 D_{3i} + \beta X_i + \mu_i$

D. $Y_i = \alpha_1 + \alpha_2 D_{2i} + \beta X_i + \mu_i$

12. 逐步回归法既检验又修正了 (**D**)

A. 异方差性

B. 自相关性

C. 随机解释变量

D. 多重共线性

13. 已知模型的形式为 $Y_i = \beta_1 + \beta_2 X_i + u_i$, 在用实际数据对模型的参数进行估计的时候, 测得 DW 统计量为 0.6453, 则广义差分变量是 (**B**)

A. $Y_t - 0.6453Y_{t-1}, X_t - 0.6453X_{t-1}$

B. $Y_t - 0.6774Y_{t-1}, X_t - 0.6774X_{t-1}$

C. $Y_t - Y_{t-1}, X_t - X_{t-1}$

D. $Y_t - 0.05Y_{t-1}, X_t - 0.05X_{t-1}$

14. 目前所学的回归分析中, 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

~~15、在有 M 个方程的完备联立方程组中，当识别的阶条件为 $H - N_i > M - 1$ 时（ H 为联立方程组中内生变量和前定变量的总数， N_i 为第 i 个方程中内生变量和前定变量的总数），则表示(**A**)~~

- A. 第 i 个方程恰好识别 B. 第 i 个方程不可识别
C. 第 i 个方程过度识别 D. 第 i 个方程的识别状态不能确定

16、多元线性回归分析中，调整后的可决系数 \bar{R}^2 与可决系数 R^2 之间的关系 (**B**)

- A. $\bar{R}^2 = 1 - (1 - R^2) \frac{n-k}{n-1}$ B. $\bar{R}^2 = 1 - (1 - R^2) \frac{n-1}{n-k}$
C. $\bar{R}^2 > 0$ D. $\bar{R}^2 \geq R^2$

17、在异方差的情况下，参数估计值的方差不能正确估计的原因是 (**A**)

- A. $E(u_i^2) \neq \sigma^2$ B. $E(u_i u_j) \neq 0 (i \neq j)$
C. $E(x_i u_i) \neq 0$ D. $E(u_i) \neq 0$

18、检验自回归模型扰动项的自相关性，常用德宾 h 检验，下列命题正确的是 (**B**)

- A. 德宾 h 检验只适用一阶自回归模型
B. 德宾 h 检验适用任意阶的自回归模型
C. 德宾 h 统计量服从 t 分布
D. 德宾 h 检验可以用于小样本问题

19、设 $y_i = \beta_1 + \beta_2 x_i + u_i, \text{Var}(u_i) = \sigma_i^2 = \sigma^2 f(x_i)$ ，则对原模型变换的正确形式为(**B**)

- A. $y_i = \beta_1 + \beta_2 x_i + u_i$
B. $\frac{y_i}{\sqrt{f(x_i)}} = \frac{\beta_1}{\sqrt{f(x_i)}} + \beta_2 \frac{x_i}{\sqrt{f(x_i)}} + \frac{u_i}{\sqrt{f(x_i)}}$
C. $\frac{y_i}{f^2(x_i)} = \frac{\beta_1}{f^2(x_i)} + \beta_2 \frac{x_i}{f^2(x_i)} + \frac{u_i}{f^2(x_i)}$
D. $y_i f(x_i) = \beta_1 f(x_i) + \beta_2 x_i f(x_i) + u_i f(x_i)$

~~20、在修正序列自相关的方法中，能修正高阶自相关的方法是 (**C**)~~

- A. 利用 DW 统计量值求出 $\hat{\rho}$ B. Cochrane-Orcutt 法
C. Durbin 两步法 D. 移动平均法

二、多项选择题

1、希斯特（Shisko）研究了什么因素影响兼职工作者的兼职收入，模型及其估计结果为：

$$\hat{w}_m = 37.07 + 0.403 w_0 - 90.06 \text{race} + 113.64 \text{reg} + 2.26 \text{age}$$
$$(0.062) \quad (24.47) \quad (27.62) \quad (0.94)$$
$$R^2 = 0.74 \quad df = 311$$

其中： w_m 为兼职工薪（美元/小时）； w_0 为主业工薪（美元/小时）； race 为虚拟变量，若是白人取值为 0，非白人取值为 1； reg 为虚拟变量，当被访者是非西部人时， reg 取值为 0，当被访者是西部地区人时， reg 取值为 1； age 为年龄；括号中的数据位系数估计值的标准误。关于这个估计结果，下列说法正确的有（**A D E**）

- A. 在其他因素保持不变条件下，非白人的兼职工薪每小时比白人约低 90 美元
- B. 在其他因素保持不变条件下，白人的兼职工薪每小时比白人约低 90 美元
- C. 在其他因素保持不变条件下，非西部人的兼职工薪每小时比西部人约高出 113.64 美元
- D. 在其他因素保持不变条件下，非西部人的兼职工薪每小时比西部人约低出 113.64 美元
- E. 四个变量在 5%显著性水平下统计上是显著的

2、对于二元样本回归模型 $Y_i = \hat{\beta}_1 + \hat{\beta}_{21} X_{2i} + \hat{\beta}_3 X_{3i} + e_i$ ，下列各式成立的有（**A B C**）

- A. $\sum e_i = 0$
- B. $\sum e_i X_{2i} = 0$
- C. $\sum e_i X_{3i} = 0$
- D. $\sum e_i Y_i = 0$
- E. $\sum X_{3i} X_{2i} = 0$

3、能够检验多重共线性的方法有（**A C E**）

- A. 简单相关系数矩阵法
- B. DW 检验法
- C. t 检验与 F 检验综合判断法
- D. ARCH 检验法
- E. 辅助回归法(又称待定系数法)

4、对联立方程模型参数的单方程估计法包括（**A B D**）

- A. 工具变量法
- B. 间接最小二乘法
- C. 完全信息极大似然估计法
- D. 二阶段最小二乘法
- E. 三阶段最小二乘法

- 5、如果模型中存在自相关现象，则会引起如下后果（ **BCDE** ）
- A.参数估计值有偏 B.参数估计值的方差不能正确确定
- C.变量的显著性检验失效 D.预测精度降低
- E.参数估计值仍是无偏的

三、判断题（判断下列命题正误，并说明理由）

1、在实际中，一元回归几乎没什么用，因为因变量的行为不可能仅由一个解释变量来解释。

错

在实际中，在一定条件下一元回归是很多经济现象的近似，能够较好地反映回归分析的基本思想，在某些情况下还是有用的。

2、多重共线性问题是随机扰动项违背古典假定引起的；

错

应该是解释变量之间高度相关引起的。

3、在异方差性的情况下，若采用 Eviews 软件中常用的 OLS 法，必定高估了估计量的标准误。

错

有可能高估也有可能低估。

如：考虑一个非常简单的具有异方差性的线性回归模型：

$$Y_i = \beta X_i + u_i; \text{ 设: } \text{Var}(u_i) = \sigma_i^2 = Z_i^2 \sigma^2$$

$$\text{则: } \text{Var}(\hat{\beta}) = \text{Var}\left(\frac{\sum X_i u_i}{\sum X_i^2}\right) = \frac{\sum X_i^2 \text{Var}(u_i)}{(\sum X_i^2)^2} = \sum \frac{X_i^2}{(\sum X_i^2)^2} \text{Var}(u_i)$$

等方差情形下： $\text{Var}(\hat{\beta}) = \sum \frac{X_i^2}{(\sum X_i^2)^2} \text{Var}(u_i) = \sigma^2 \sum \frac{X_i^2}{(\sum X_i^2)^2}$ ，这也是 Eviews 常用的估计结果；

$$\text{异方差情形下: } \text{Var}(\hat{\beta}) = \sum \frac{X_i^2}{(\sum X_i^2)^2} \text{Var}(u_i) = \sum \frac{X_i^2 Z_i^2}{(\sum X_i^2)^2} \sigma^2 = \sigma^2 \sum \frac{X_i^2 Z_i^2}{(\sum X_i^2)^2};$$

对上述两种情形进行比较：

$$\sigma^2 \sum \frac{X_i^2 Z_i^2}{(\sum X_i^2)^2} \bigg/ \sigma^2 \sum \frac{X_i^2}{(\sum X_i^2)^2} = \sum X_i^2 Z_i^2 / \sum X_i^2 = \begin{cases} >1 & Z_i^2 > 1 \\ <1 & Z_i^2 < 1 \end{cases}, \text{ 故常用的 OLS}$$

估计，有可能低估异方差性条件下的系数标准误（ $Z_i^2 > 1$ ），也有可能高估异方差性条件下的系数标准误（ $Z_i^2 < 1$ ）（请与教材中的情形进行比较）。

~~4、虚拟变量只能作为解释变量。~~

错

虚拟变量还能作被解释变量。

~~5、设估计模型为~~

$$\hat{PCE}_i = -171.4412 + 0.9672PDI_i$$

$$t = (-7.4809) \quad (119.8711)$$

$$R^2 = 0.9940 \quad \text{DW} = 0.5316$$

由于 ~~$R^2 = 0.9940$~~ ，表明模型有很好的拟合优度，则模型不存在伪（虚假）回归。

错

可能存在伪（虚假）回归，因为可决系数较高，而 DW 值过低。

四、计算题

1、某公司在为建造一个新的百货店选址的决策过程中，对已有的 30 个百货店的销售额作为其所处地理位置特征的函数进行回归分析，并且用该回归方程作为新百货店的不同位置的可能销售额，估计得出（括号内为估计的标准差）

$$\hat{Y}_i = 30 + 0.1 \times X_{1i} + 0.01 \times X_{2i} + 10.0 \times X_{3i} + 3.0 \times X_{4i}$$

(0.02) (0.01) (1.0) (1.0)

其中： Y_i = 第*i*个百货店的日均销售额（百美元）；

X_{1i} = 第*i*个百货店前每小时通过的汽车数量（10 辆）；

X_{2i} = 第*i*个百货店所处区域内的人均收入（美元）；

X_{3i} = 第*i*个百货店内所有的桌子数量；

X_{4i} = 第*i*个百货店所处地区竞争店面的数量；

请回答以下问题：

- （1）说出本方程中系数 0.1 和 0.01 的经济含义。
- （2）各个变量前参数估计的符号是否与期望的符号一致？
- （3）在 $\alpha = 0.05$ 的显著性水平下检验变量 X_{1i} 的显著性。

（临界值 $t_{0.025}(25) = 2.06$ ， $t_{0.025}(26) = 2.056$ ， $t_{0.05}(25) = 1.708$ ， $t_{0.05}(26) = 1.706$ ）

解：平均意义上，（1）每小时通过该百货店的汽车增加 10 辆，该店的每日收入就会平均增加 10 美元。该区域居民人均收入每增加 1 美元，该店每日收入就

会平均增加 1 美元。

(2) 最后一个系数与期望的符号不一致，应该为负数，即该区竞争的店面越多，该店收入越低。其余系数的符号符合期望。

(3) 用 t 检验。 $t=0.1/0.02=5$ ，有 $t=5 > t_{0.025}(25)=2.06$ 可知，该变量显著。

2、~~一国的对外贸易分为出口和进口，净出口被定义为出口与进口的差额。影响净出口的因素很多，在宏观经济学中，汇率和国内收入水平被认为是两个最重要的因素，我们根据这一理论对影响中国的净出口水平的因素进行实证分析。~~

~~设 NX 表示我国净出口水平（亿元）；GDP 为我国国内生产总值（亿元），反映我国的国内收入水平；D(GDP)表示 GDP 的一阶差分；E 表示每 100 美元对人民币的平均汇率（元/百美元），反映汇率水平。利用 1985——2001 年我国的统计数据（摘自《2002 中国统计年鉴》），估计的结果见下表。~~

(1) 选择解释我国净出口水平最适合的计量经济模型，写出该模型并说明选择的原因，其它模型可能存在什么问题；

(2) 解释选择的计量经济模型的经济意义。

相关系数矩阵

	NX	GDP	E
NX	1.000000	0.853314	0.786534
GDP	0.853314	1.000000	0.916241
E	0.786534	0.916241	1.000000

Dependent Variable: NX

Method: Least Squares

Date: 03/21/05 Time: 11:02

Sample: 1985 2001

Included observations: 17

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-2135.887	645.9685	-3.306488	0.0048
E	4.851832	0.983587	4.932794	0.0002
R-squared	0.618636	Mean dependent var	879.9059	
Adjusted R-squared	0.593211	S.D. dependent var	1348.206	
S.E. of regression	859.8857	Akaike info criterion	16.46161	
Sum squared resid	11091052	Schwarz criterion	16.55963	
Log likelihood	-137.9237	F-statistic	24.33245	
Durbin-Watson stat	0.890230	Prob(F-statistic)	0.000180	

Dependent Variable: NX

Method: Least Squares

Date: 03/21/05 Time: 11:04

Sample: 1985 2001

Included observations: 17

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-761.6691	313.1743	-2.432093	0.0280
GDP	0.036827	0.005810	6.338492	0.0000
R-squared	0.728145	Mean dependent var		879.9059
Adjusted R-squared	0.710021	S.D. dependent var		1348.206
S.E. of regression	726.0044	Akaike info criterion		16.12312
Sum squared resid	7906237.	Schwarz criterion		16.22115
Log likelihood	-135.0465	F-statistic		40.17648
Durbin-Watson stat	1.289206	Prob(F-statistic)		0.000013

Dependent Variable: NX

Method: Least Squares

Date: 03/21/05 Time: 11:06

Sample: 1985 2001

Included observations: 17

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-822.2318	789.9381	-1.040881	0.3156
E	0.180334	2.145081	0.084069	0.9342
GDP	0.035671	0.015008	2.376855	0.0323
R-squared	0.728282	Mean dependent var		879.9059
Adjusted R-squared	0.689465	S.D. dependent var		1348.206
S.E. of regression	751.2964	Akaike info criterion		16.24026
Sum squared resid	7902248.	Schwarz criterion		16.38730
Log likelihood	-135.0422	F-statistic		18.76202
Durbin-Watson stat	1.279954	Prob(F-statistic)		0.000109

Dependent Variable: NX

Method: Least Squares

Date: 03/21/05 Time: 11:09

Sample(adjusted): 1986 2001

Included observations: 16 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-3036.617	444.7869	-6.827128	0.0000
E	8.781248	0.929788	9.444358	0.0000
D(GDP)	-0.301465	0.054757	-5.505550	0.0001
R-squared	0.878586	Mean dependent var		962.9563
Adjusted R-squared	0.859907	S.D. dependent var		1346.761

S.E. of regression	504.0793	Akaike info criterion	15.45070
Sum squared resid	3303247.	Schwarz criterion	15.59557
Log likelihood	-120.6056	F-statistic	47.03583
Durbin-Watson stat	2.214778	Prob(F-statistic)	0.000001

解：（1）模型选择可依据两个方面：经济学意义和计量经济学的模型选择准则。根据回归结果，从 **Akaike info criterion(=15.45)**和 **Schwarz criterion(= 15.60)**看，可认为最后一个回归模型（第四个）最佳，进一步从模型中各个变量 **t** 检验显著性、模型的 **F** 检验显著性，拟合优度、自相关性等综合考虑，从计量经济学的角度，可认为第四个回归模型是设定较好的模型，即将 **NX**（净出口）关于汇率、**DGDP**（**GDP** 的一阶差分）进行回归的模型较好；在此基础上，还应当进行经济意义或经济理论背景的检验，以进一步确定本问题中模型得设定。

比较而言，从计量经济学的观点看，各自在不同程度和不同方面，存在着这样或那样的不完善的地方：第一个模型，**NX** 关于 **E** 的回归，**Akaike info criterion** 和 **Schwarz criterion** 的值大于第 4 个模型，且拟合优度也不太好，自相关现象存在；第二个模型，与第 2 个模型的情形类似；而第三个模型是将 **NX** 对 **E**、**GDP** 的回归，结果提示，这样的回归模型多重共线性现象严重，且不能正常判断其自相关性。

（2）所选模型的经济意义是：影响净出口的主要因素是汇率和 **GDP** 的增长量。汇率每提高一个单位，净出口就会增加 **8.781248** 个单位（亿元），**DGDP** 每增加一个单位（亿元），则净出口增加 **0.03682** 亿元。

3、下面结果是利用某地财政收入对该地第一、二、三产业增加值的回归结果，根据这一结果试判断该模型是否存在多重共线性，说明你的理由。

Dependent Variable: REV
Method: Least Squares
Sample: 1 10
Included observations: 10

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	17414.63	14135.10	1.232013	0.2640
GDP1	-0.277510	0.146541	-1.893743	0.1071
GDP2	0.084857	0.093532	0.907252	0.3992
GDP3	0.190517	0.151680	1.256048	0.2558
R-squared	0.993798	Mean dependent var	63244.00	
Adjusted R-squared	0.990697	S.D. dependent var	54281.99	
S.E. of regression	5235.544	Akaike info criterion	20.25350	
Sum squared resid	1.64E+08	Schwarz criterion	20.37454	
Log likelihood	-97.26752	F-statistic	320.4848	
Durbin-Watson stat	1.208127	Prob(F-statistic)	0.000001	

答：多重共线性现象较为严重。因为方程整体非常显著，表明三次产业 **GDP** 对财政收入的解释能力非常强，但是每个个别解释变量均不显著，且存在负系数，与理论矛盾，原因是存在严重共线性。

第四套

一、单项选择题

1、设 OLS 法得到的样本回归直线为 $Y_i = \hat{\beta}_1 + \hat{\beta}_2 X_i + e_i$ ，则点 (\bar{X}, \bar{Y})

(**B**)

- A. 一定不在回归直线上 B. 一定在回归直线上
C. 不一定在回归直线上 D. 在回归直线上方

2、在下列各种数据中，以下不应作为经济计量分析所用数据的是 (**C**)

- A. 时间序列数据 B. 横截面数据
C. 计算机随机生成的数据 D. 虚拟变量数据

~~3、在简单线性回归模型中，认为具有一定概率分布的变量是(**A**)~~

- ~~A. 内生变量 B. 外生变量 C. 虚拟变量 D. 前定变量~~

4、根据样本资料估计得出人均消费支出 Y 对人均收入 X 的回归模型为

$\ln Y_i = 2.00 + 0.75 \ln X_i$ ，表明人均收入每增加 1%，人均消费支出将增加(**B**)

- A. 0.2% B. 0.75% C. 2% D. 7.5%

5、多元线性回归分析中的 RSS 反映了 (**C**)

- A. 应变量观测值总变差的大小
B. 应变量回归估计值总变差的大小
C. 应变量观测值与估计值之间的总变差
D. Y 关于 X 的边际变化

~~6、在经济发展发生转折时期，可以通过引入虚拟变量方法来表示这种变化。例如，研究中国城镇居民消费函数时。1991 年前后，城镇居民商品性实际支出 Y 对实际可支配收入 X 的回归关系明显不同。现以 1991 年为转折时期，设虚拟变量 $D_t = \begin{cases} 1 & 1991 \text{ 年以前} \\ 0 & 1991 \text{ 年以后} \end{cases}$ ，数据散点图显示消费函数发生了结构性变化：基本消费部分下降了，边际消费倾向变大了。则城镇居民线性消费函数的理论方程可以写作 (**D**)~~

~~A. $Y_t = \beta_1 + \beta_2 X_t + u_t$~~

~~B. $Y_t = \beta_1 + \beta_2 X_t + \beta_4 D_t X_t + u_t$~~

~~C. $Y_t = \beta_1 + \beta_2 X_t + \beta_3 D_t + u_t$~~

~~D. $Y_t = \beta_1 + \beta_2 X_t + \beta_3 D_t + \beta_4 D_t X_t + u_t$~~

7、已知模型的形式为 $Y_t = \beta_1 + \beta_2 X_t + u_t$ ，在用实际数据对模型的参数进行估

计的时候，测得 DW 统计量为 0.52，则广义差分变量是(**D**)

A. $y_t - 0.48y_{t-1}, x_t - 0.48x_{t-1}$

B. $y_t - 0.7453y_{t-1}, x_t - 0.7453x_{t-1}$

C. $y_t - 0.52y_{t-1}, x_t - 0.52x_{t-1}$

D. $y_t - 0.74y_{t-1}, x_t - 0.74x_{t-1}$

8、在有 M 个方程的完备联立方程组中，若用 H 表示联立方程组中全部的内生变量与全部的前定变量之和的总数，用 N_i 表示第 i 个方程中内生变量与前定变量之和的总数时，第 i 个方程不可识别时，则有公式(**D**)成立。

A. $H - N_i > M - 1$

B. $H - N_i = M - 1$

C. $H - N_i = 0$

D. $H - N_i < M - 1$

9、如果模型中的解释变量存在完全的多重共线性，参数的最小二乘估计量是 (**C**)

A. 无偏的

B. 有偏的

C. 不确定

D. 确定的

10、关于联立方程组模型，下列说法中错误的是 (**B**)

~~A. 结构式模型中解释变量可以是内生变量~~

~~B. 简化式模型中解释变量可以是内生变量~~

~~C. 简化式模型中解释变量是前定变量~~

~~D. 结构式模型中解释变量可以是前定变量~~

11、在序列自相关的情况下，参数估计值仍是无偏的，其原因是 (**A**)

A. 零均值假定成立

B. 同方差假定成立

C. 无多重共线性假定成立

D. 解释变量与随机误差项不相关假定成立

12、在 DW 检验中，当 d 统计量为 2 时，表明 (**C**)

A. 存在完全的正自相关

B. 存在完全的负自相关

C. 不存在自相关

D. 不能判定

13、在下列多重共线性产生的原因中，不正确的是 (**D**)

A. 经济本变量大多存在共同变化趋势

B. 模型中大量采用滞后变量

C. 由于认识上的局限使得选择变量不当

D. 解释变量与随机误差项相关

14、下列说法不正确的是 (**C**)

A. 异方差是一种随机误差现象

- B. 异方差产生的原因有设定误差
- C. 检验异方差的方法有加权最小二乘法
- D. 修正异方差的方法有加权最小二乘法

15、设 k 为回归模型中的参数个数， n 为样本容量。则对多元线性回归方程进行显著性检验时，所用的 F 统计量可表示为 (**A**)

- A. $\frac{R^2/(k-1)}{(1-R^2)/(n-k)}$
- B. $\frac{ESS/(n-k)}{RSS/(k-1)}$
- C. $\frac{R^2/(n-k)}{(1-R^2)/(k-1)}$
- D. $\frac{ESS/(k-1)}{TSS/(n-k)}$

~~16、对联立方程组模型中过度识别方程的估计方法有 (**D**)~~

- ~~A. 间接最小二乘法~~
- ~~B. 普通最小二乘法~~
- ~~C. 间接最小二乘法和二阶段最小二乘法~~
- ~~D. 二阶段最小二乘法~~

17、对模型进行对数变换,其原因是 (**B**)

- A. 能使误差转变为绝对误差
- B. 能使误差转变为相对误差
- C. 更加符合经济意义
- D. 大多数经济现象可用对数模型表示

18、局部调整模型不具有如下的特点 (**D**)

- A. 对应的原始模型中被解释变量为期望变量，它不可观测
- B. 模型是一阶自回归模型
- C. 模型中含有一个滞后被解释变量 Y_{t-1} ，但它与随机扰动项不相关
- D. 模型的随机扰动项存在自相关

19、假设根据某地区 1970—1999 年的消费总额 Y (亿元) 和货币收入总额 X (亿元) 的年度资料，估计出库伊克(Koyck)模型如下：

$$\hat{Y}_t = -6.9057 + 0.2518X_t + 0.8136Y_{t-1}$$

$$t = (-1.6521) \quad (5.7717) \quad (12.9166)$$

$$R^2 = 0.997 \quad F = 4323 \quad DW = 1.216$$

则 (**C**)

- A. 分布滞后系数的衰减率为 0.1864

- B. 在显著性水平 $\alpha = 0.05$ 下，DW 检验临界值为 $d_L = 1.3$ ，而

$d = 1.26 < d_L = 1.3$ ，据此可以推断模型扰动项存在自相关

- C. 即期消费倾向为 0.2518，表明收入每增加 1 元，当期的消费将增加 0.2518 元

- D. 收入对消费的长期影响乘数为 Y_{t-1} 的估计系数 0.8136

20、在模型有异方差的情况下,常用的补救措施是(**D**)

A.广义差分法 B.工具变量法 C.逐步回归法 D.加权最小二乘法

二、多项选择题

1、调整后的判定系数 \bar{R}^2 与判定系数 R^2 之间的关系叙述正确的有 (**CE**)

A. \bar{R}^2 与 R^2 均非负

B. \bar{R}^2 有可能大于 R^2

C. 判断多元回归模型拟合优度时, 使用 \bar{R}^2

D. 模型中包含的解释变量个数越多, \bar{R}^2 与 R^2 就相差越大

E. 只要模型中包括截距项在内的参数的个数大于 1, 则 $\bar{R}^2 < R^2$

2、如果模型中存在序列自相关现象, 则有如下后果 (**BCDE**)

A. 参数估计值有偏 B. 参数估计值的方差不能正确确定

C. 变量的显著性检验失效 D. 预测精度降低

E. 参数估计值仍是无偏的

3、下列说法不正确的有 (**BCF**)

A. 加权最小二乘法是广义最小二乘法的特殊情况

B. 广义最小二乘法是加权最小二乘法的特殊情况

C. 广义最小二乘法是广义差分法的特殊情况

D. 广义差分法是广义最小二乘法的特殊情况

E. 普通最小二乘法是加权最小二乘法的特殊情况

F. 加权最小二乘法是普通最小二乘法的特殊情况

4、关于联立方程模型识别问题, 以下说法正确的有 (**ACDE**)

A. 可识别的方程则满足阶条件

B. 如果一个方程包含了模型中的全部变量, 则这个方程恰好识别

C. 如果一个方程包含了模型中的全部变量, 则这个方程不可识别

D. 如果两个方程包含相同的变量, 则这两个方程均不可识别

E. 联立方程组中的每一个方程都是可识别的, 则联立方程组才可识别

F. 满足阶条件和秩条件的方程一定是过度识别

5、在 DW 检验中, 存在不能判定的区域是 (**CD**)

A. $0 < d < d_L$

B. $d_U < d < 4 - d_U$

C. $d_L < d < d_U$

D. $4 - d_U < d < 4 - d_L$

E. $4 - d_L < d < 4$

三、判断题（判断下列命题正误，并说明理由）

1、随机扰动项的方差与随机扰动项方差的无偏估计没有区别。

错

随机扰动项的方差反映总体的波动情况，对一个特定的总体而言，是一个确定的值。

在最小二乘估计中，由于总体方差在大多数情况下并不知道，所以用样本数据去估计 σ^2 ： $\hat{\sigma}^2 = \sum e_i^2 / (n - k)$ 。其中 n 为样本数， k 为待估参数的个数。 $\hat{\sigma}^2$ 是 σ^2 线性无偏估计，为一个随机变量。

2、经典线性回归模型（CLRM）中的干扰项不服从正态分布的，OLS 估计量将有偏的。

错

即使经典线性回归模型（CLRM）中的干扰项不服从正态分布的，OLS 估计量仍然是无偏的。因为 $E(\hat{\beta}_2) = E(\beta_2 + \sum K_i \mu_i) = \beta_2$ ，该表达式成立与否与正态性无关。

~~3、虚拟变量的取值原则上只能取 0 或 1。~~

对

虚拟变量的取值是人为设定的，主要表征某种属性或特征或其它的存在与否，0 或 1 正好描述了这种特性。当然，依据研究问题的特殊性，有时也可以取其它值。

4、拟合优度检验和 F 检验是没有区别的。

错

（1）F 检验中使用的统计量有精确的分布，而拟合优度检验没有；

（2）对是否通过检验，可决系数（修正可决系数）只能给出一个模糊的推测；而 F 检验可以在给定显著水平下，给出统计上的严格结论。

~~5、联立方程组模型根本不能直接用 OLS 方法估计参数。~~

错

递归方程可以用 OLS 方法估计参数，而其它的联立方程组模型不能直接用 OLS 方法估计参数。

四、计算题

1、通过建模发现，某企业的某种产品价格 P 和可变成本 V 之间满足如下关系： $\ln P = 34.5 + 0.56 \ln V$ 。目前可变成本占产品价格的 20%。现在，企业可以改进该产品，但是改进要增加 10% 可变成本（其他费用保持不变）。问，企业是否该选择改进？

解：（1）由模型可知，价格和可变成本之间的弹性为 0.56。假设改进产品，则可变成本增加 10%，价格的变化率为 $0.56 \times 10\% = 5.6\%$ ，可见价格增加的幅度不如可变成本增加的幅度。

（2）利润增量为 $5.6\% \times P - 10\% \times V$ ，只要利润增量大于 0，就应该选择改进。

（3）易得，只要当 $P/V > (10/5.6)$ ，就有利润大于 0。而目前成本只占价格的 20%，远小于 5.6/10，所以应该选择改进。

2、某公司想决定在何处建造一个新的百货店，对已有的 30 个百货店的销售额作为其所处地理位置特征的函数进行回归分析，并且用该回归方程作为新百货店的不同位置的可能销售额，估计得出（括号内为估计的标准差）

$$\hat{Y}_i = 30 + 0.1 \times X_{1i} + 0.01 \times X_{2i} + 10.0 \times X_{3i} + 3.0 \times X_{4i}$$

(0.02) (0.01) (1.0) (1.0)

其中： Y_i = 第 i 个百货店的日均销售额（百美元）；

X_{1i} = 第 i 个百货店前每小时通过的汽车数量（10 辆）；

X_{2i} = 第 i 个百货店所处区域内的人均收入（美元）；

X_{3i} = 第 i 个百货店内所有的桌子数量；

X_{4i} = 第 i 个百货店所处地区竞争店面的数量；

请回答以下问题：

- （1）说出本方程中系数 0.1 和 0.01 的经济含义。
- （2）各个变量前参数估计的符号是否与期望的符号一致？
- （3）在 $\alpha = 0.05$ 的显著性水平下检验变量 X_{1i} 的显著性。

（临界值 $t_{0.025}(25) = 2.06$ ， $t_{0.025}(26) = 2.056$ ， $t_{0.05}(25) = 1.708$ ， $t_{0.05}(26) = 1.706$ ）

答：（1）每小时通过该百货店的汽车增加 10 辆，该店的每日收入就会平均增加 10 美元。该区域居民人均收入每增加 1 美元，该店每日收入就会平均增加 1 美元。

（2）最后一个系数与期望的符号不一致，应该为负数，即该区竞争的店面越多，该店收入越低。其余符号符合期望。

（3）用 t 检验。 $t = 0.1/0.02 = 5$ ，有 $t > t_{0.025}(25) = 2.06$ 知道，该变量显著。

3、以广东省东莞市的财政支出作为被解释变量、财政收入作为解释变量做计量经济模型，即 $Y = \alpha + \beta X + \mu$ ，方程估计、残差散点图及 ARCH 检验输出结果分别如下：

方程估计结果：

Dependent Variable: Y

Method: Least Squares

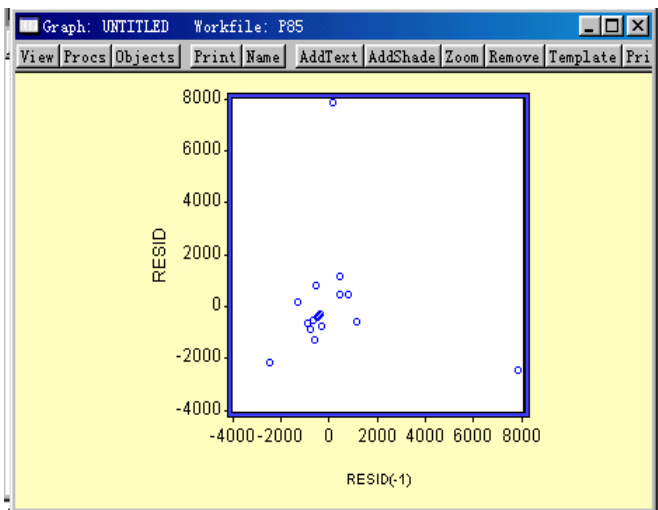
Date: 05/31/05 Time: 12:42

Sample: 1980 1997

Included observations: 18

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-2457.310	680.5738	-3.610644	0.0023
X	0.719308	0.011153	64.49707	0.0000
R-squared	0.996168	Mean dependent var	25335.11	
Adjusted R-squared	0.995929	S.D. dependent var	35027.97	
S.E. of regression	2234.939	Akaike info criterion	18.36626	
Sum squared resid	79919268	Schwarz criterion	18.46519	
Log likelihood	-163.2963	F-statistic	4159.872	
Durbin-Watson stat	2.181183	Prob(F-statistic)	0.000000	

残差与残差滞后 1 期的散点图：



ARCH 检验输出结果：

ARCH Test:

F-statistic	2.886465	Probability	0.085992
Obs*R-squared	7.867378	Probability	0.096559

Test Equation:

Dependent Variable: RESID^2

Method: Least Squares

Date: 06/10/05 Time: 00:33

Sample(adjusted): 1984 1997

Included observations: 14 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-9299857.	7646794.	-1.216177	0.2549
RESID^2(-1)	0.033582	0.308377	0.108900	0.9157
RESID^2(-2)	-0.743273	0.320424	-2.319650	0.0455
RESID^2(-3)	-0.854852	11.02966	-0.077505	0.9399
RESID^2(-4)	37.04345	10.91380	3.394182	0.0079

R-squared	0.561956	Mean dependent var	5662887.
Adjusted R-squared	0.367269	S.D. dependent var	16323082
S.E. of regression	12984094	Akaike info criterion	35.86880
Sum squared resid	1.52E+15	Schwarz criterion	36.09704
Log likelihood	-246.0816	F-statistic	2.886465
Durbin-Watson stat	1.605808	Prob(F-statistic)	0.085992

根据以上输出结果回答下列问题：

- （1）该模型中是否违背无自相关假定？为什么？（ $\alpha = 0.05$ ， $d_L = 1.158$, $d_U = 1.391$ ）
- （2）该模型中是否存在异方差？说明理由（显著性水平为 10%， $\chi^2_{0.1}(4) = 7.7794$ ）。
- （3）如果原模型存在异方差，你认为应如何修正？（只说明修正思路，无需计算）

解：（1）没有违背无自相关假定；第一、残差与残差滞后一期没有明显的相关性；第二、根据 D-W 值应该接受原假设；（写出详细步骤）

（2）存在异方差（注意显著性水平是 0.1）；（写出详细步骤）

（3）说出一种修正思路即可。

第五套

一、单项选择题

1、下列说法正确的有（ C ）

- A. 时序数据和横截面数据没有差异
- B. 对总体回归模型的显著性检验没有必要
- C. 总体回归方程与样本回归方程是有区别的
- D. 判定系数 R^2 不可以用于衡量拟合优度

2、所谓异方差是指（ A ）

- A. $Var(u_i) \neq \sigma^2$
- B. $Var(x_i) \neq \sigma^2$
- C. $Var(u_i) = \sigma^2$
- D. $Var(x_i) = \sigma^2$

3、在给定的显著性水平之下，若DW统计量的下和上临界值分别为 d_L 和 d_U ，则当 $d_L < d < d_U$ 时，可认为随机误差项（ D ）

- A. 存在一阶正自相关
- B. 存在一阶负相关
- C. 不存在序列相关
- D. 存在序列相关与否不能断定

~~4、在利用月度数据构建计量经济模型时，如果一年里的1、3、5、9四个月表现出季节模式，则应该引入虚拟变量个数为（ A ）~~

- A. 4
- B. 3
- C. 11
- D. 6

~~5、假如联立方程模型中，若第 i 个方程包含了模型中的全部变量（即全部的内生变量和全部的前定变量），则第 i 个方程是（ D ）~~

- A. 可识别的
- B. 恰好识别
- C. 过度识别
- D. 不可识别

~~6、在序列自相关的情况下，参数估计值的方差不能正确估计的原因是（ B ）~~

- A. $E(u_i^2) \neq \sigma^2$
- B. $E(u_i u_j) \neq 0 (i \neq j)$
- C. $E(x_i u_i) \neq 0$
- D. $E(u_i) \neq 0$

7、在模型 $Y_t = \beta_1 + \beta_2 X_{2t} + \beta_3 X_{3t} + u_t$ 的回归分析结果报告中，有

$F = 263489.23$ ， F 的 p 值 = 0.000000，则表明（ C ）

- A. 解释变量 X_{2t} 对 Y_t 的影响是显著的

B.解释变量 X_{3t} 对 Y_t 的影响是显著的

C.解释变量 X_{2t} 和 X_{3t} 对 Y_t 的联合影响是显著的

D.解释变量 X_{2t} 和 X_{3t} 对 Y_t 的联合影响均不显著

8、用模型描述现实经济系统的原则是(**B**)

A、模型规模大小要适度，结构尽可能复杂

B、以理论分析作先导，模型规模大小要适度

C、模型规模越大越好；这样更切合实际情况

D、以理论分析作先导，解释变量应包括所有解释变量

9、如果回归模型违背了同方差假定，最小二乘估计是(**A**)

A. 无偏的，非有效的

B. 有偏的，非有效的

C. 无偏的，有效的

D. 有偏的，有效的

10、设线性回归模型为 $y_i = \beta_1 + \beta_2 x_{2i} + \beta_3 x_{3i} + u_i$ ，下列表明变量之间具有完

全多重共线性的是(**A**)

A. $0 \cdot x_1 + 2 \cdot x_2 + 0.2 \cdot x_3 = 0$

B. $0 \cdot x_1 + 2 \cdot x_2 + 0.2 \cdot x_3 + v = 0$

C. $0 \cdot x_1 + 0 \cdot x_2 + 0 \cdot x_3 = 0$

D. $0 \cdot x_1 + 0 \cdot x_2 + 0 \cdot x_3 + v = 0$

其中 v 为随机误差项

11、对于有限分布滞后模型，解释变量的滞后长度每增加一期，可利用的样本数据就会(**B**)

A. 增加 1 个

B. 减少 1 个

C. 增加 2 个

D. 减少 2 个

12、关于自适应预期模型和局部调整模型，下列说法错误的有(**D**)

A.它们都是由某种期望模型演变形成的

B.它们最终都是一阶自回归模型

C.它们的经济背景不同

D. 都满足古典线性回归模型的所有假设，故可直接用 OLS 方法进行估

计

13、在检验异方差的方法中，不正确的是(**D**)

A. Goldfeld-Quandt 方法

B. ARCH 检验法

C. White 检验法

D. DW 检验法

~~14、边际成本函数为 $C = \alpha_0 + \alpha_1 Q + \alpha_2 Q^2 + \mu$ (C 表示边际成本， Q 表示产量)，则下列说法正确的有(**C**)~~

~~A. 模型为非线性模型~~

~~B. 模型为线性模型~~

~~C. 模型中可能存在多重共线性~~

~~D. 模型中不应包括 Q^2 作为解释变量~~

- 15、对自回归模型进行估计时，假定原始模型的随机扰动项 u_i 满足古典线性回归模型的所有假设，则估计量是一致估计量的模型有（ **B** ）
- A. 库伊克模型 B. 局部调整模型
C. 自适应预期模型 D. 自适应预期和局部调整混合模型
- 16、在 DW 检验中，当 d 统计量为 0 时，表明（ **A** ）
- A. 存在完全的正自相关 B. 存在完全的负自相关
C. 不存在自相关 D. 不能判定
- 17、在下列产生序列自相关的原因中，不正确的是（ **D** ）
- A. 经济变量的惯性作用 B. 经济行为的滞后作用
C. 设定偏误 D. 解释变量的共线性
- 18、简化式模型就是把结构式模型中的内生变量表示为（ **B** ）
- A. 外生变量和内生变量的函数关系
B. 前定变量和随机误差项的函数所构成的模型
C. 滞后变量和随机误差项的函数所构成的模型
D. 外生变量和随机误差项的函数模型所构成的模型
- 19、加权最小二乘法是（ **C** ）的一个特例
- A. 广义差分法 B. 普通最小二乘法
C. 广义最小二乘法 D. 两阶段最小二乘法
- 20、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

二、多项选择题

1、关于衣着消费支出模型为： $Y_i = \alpha_1 + \alpha_2 D_{2i} + \alpha_3 D_{3i} + \alpha_4 (D_{2i} D_{3i}) + \beta X_i + u_i$ ，其中 Y_i 为衣着方面的年度支出； X_i 为收入，

$$D_{2i} = \begin{cases} 1 & \text{女性} \\ 0 & \text{男性} \end{cases}, \quad D_{3i} = \begin{cases} 1 & \text{大学毕业及以上} \\ 0 & \text{其他} \end{cases}。$$

则关于模型中的参数下列说法正确的是（ **ABCE** ）

A. α_2 表示在保持其他条件不变时, 女性比男性在衣着消费支出方面多支出 (或少支出) 差额

B. 表示在保持其他条件不变时, 大学文凭及以上比其他学历者在衣着消费支出方面多支出 (或少支出) 差额

C. α_4 表示在保持其他条件不变时, 女性大学及以上文凭者比男性大学以下文凭者在衣着消费支出方面多支出 (或少支出) 差额

D. α_4 表示在保持其他条件不变时, 女性比男性大学以下文凭者在衣着消费支出方面多支出 (或少支出) 差额

E. α_4 表示性别和学历两种属性变量对衣着消费支出的交互影响

2、如果模型中解释变量之间存在共线性, 则会引起如下后果 (**BCD**)

A. 参数估计值确定

B. 参数估计值不确定

C. 参数估计值的方差趋于无限大

D. 参数的经济意义不正确

E. DW 统计量落在了不能判定的区域

3、应用 DW 检验方法时应满足该方法的假定条件, 下列是其假定条件的有 (**ABCDE**)

A. 解释变量为非随机的

B. 截距离项不为零

C. 随机误差项服从一阶自回归

D. 数据无缺失项

E. 线性回归模型中不能含有滞后内生变量

4、利用普通最小二乘法求得的样本回归直线 $\hat{Y}_i = \hat{\beta}_1 + \hat{\beta}_2 X_i$ 具有如下性质 (**ACD**)

A. 样本回归线必然通过点 (\bar{X}, \bar{Y})

B. 可能通过点 (\bar{X}, \bar{Y})

C. 残差 e_i 的均值为常数

D. \hat{Y}_i 的平均值与 Y_i 的平均值相等

E. 残差 e_i 与解释变量 X_i 之间有一定的相关性

5、当结构方程为恰好识别时, 可选择的估计方法是 (**CDE**)

A. 最小二乘法

B. 广义差分法

C. 间接最小二乘法

D. 二阶段最小二乘法

E. 有限信息最大似然估计法

三、判断题（判断下列命题正误，并说明理由）

1、双变量模型中，对样本回归函数整体的显著性检验与斜率系数的显著性检验是一致的；

正确

最好能够写出一元线性回归模型；F 统计量与 T 统计量的关系，即 $F = t^2$ 的来历；或者说明一元线性回归仅有一个解释变量，因此对斜率系数的 t 检验等价于对方程的整体性检验。

2、多重共线性问题是随机扰动项违背古典假定引起的。

错误

应该是解释变量之间高度相关引起的。

3、在模型 $Y_t = \beta_1 + \beta_2 X_{2t} + \beta_3 X_{3t} + u_t$ 的回归分析结果报告中，有

$F = 263489.23$ ， F 的 p 值 $= 0.000000$ ，则表明解释变量 X_{2t} 对 Y_t 的影响是显著的。

错误

解释变量 X_{2t} 和 X_{3t} 对 Y_t 的联合影响是显著的

~~4、结构型模型中的每一个方程都称为结构式方程，结构方程中，解释变量只可以是前定变量。~~

错误

结构方程中，解释变量可以是前定变量，也可以是内生变量。

~~5、通过虚拟变量将属性因素引入计量经济模型，引入虚拟变量的个数与模型有无截距项无关。~~

错误

模型有截距项时，如果被考察的定性因素有 m 个相互排斥属性，则模型中引入 $m-1$ 个虚拟变量，否则会陷入“虚拟变量陷阱”；

模型无截距项时，若被考察的定性因素有 m 个相互排斥属性，可以引入 m 个虚拟变量，这时不会出现多重共线性。

四、计算题

1、已知某公司的广告费用(X)与销售额(Y)的统计数据如下表所示：

X(万元)	40	25	20	30	40	40	25	20	50	20	50	20
Y(万元)	490	395	420	475	385	525	480	400	560	365	510	540

- (1) 估计销售额关于广告费用的一元线性回归模型
(2) 说明参数的经济意义
(3) 在的显著水平下对参数的显著性进行 t 检验。

解：(1)利用 OLS 法估计样本回归直线为： $\hat{Y}_t = 319.086 + 4.185X_t$

(2)参数的经济意义：当广告费用每增加 1 万元，公司的销售额平均增加 4.185 万元。

(3) $t = \frac{\hat{\beta}_1}{\sqrt{Var(\hat{\beta}_1)}} = 3.79 > t_{0.025}(10)$ ，广告费用对销售额的影响是显著的。

2、设某商品的需求模型为 $Y_t = \beta_0 + \beta_1 X_{t+1}^* + u_t$ ，式中， Y 是商品的需求量， X_{t+1}^* 是人们对未来价格水平的预期，在自适应预期假设 $X_{t+1}^* - X_t^* = r(X_t - X_t^*)$ 下，通过适当变换，使模型中变量 X_{t+1}^* 成为可观测的变量。

解：将自适应预期假设写成 $X_{t+1}^* - (1-r)X_t^* = rX_t$

原模型 $Y_t = \beta_0 + \beta_1 X_{t+1}^* + u_t$ ①

将①滞后一期并乘以 $(1-r)$ ，有

$(1-r)Y_{t-1} = \beta_0(1-r) + \beta_1(1-r)X_t^* + (1-r)u_{t-1}$ ②

①式减去②式，整理后得到

$Y_t = r\beta_0 + r\beta_1 X_t + (1-r)Y_{t-1} + v_t$

式中： $v_t = u_t - (1-r)u_{t-1}$

3、根据某城市 1978——1998 年人均储蓄与人均收入的数据资料建立了如下回归模型：

$\hat{y} = -2187.521 + 1.6843x$

$se = (340.0103) (0.0622)$

$R^2 = 0.9748, S.E. = 1065.425, DW = 0.2934, F = 733.6066$

试求解以下问题：

(1) 取时间段 1978——1985 和 1991——1998，分别建立两个模型。

$$\text{模型 1: } \hat{y} = -145.4415 + 0.3971x$$

$$t = (-8.7302) \quad (25.4269)$$

$$R^2 = 0.9908, \sum e_1^2 = 1372.202$$

$$\text{模型 2: } \hat{y} = -4602.365 + 1.9525x$$

$$t = (-5.0660) \quad (18.4094)$$

$$R^2 = 0.9826, \sum e_2^2 = 5811189$$

计算 F 统计量，即 $F = \sum e_2^2 / \sum e_1^2 = 5811189 / 1372.202 = 4334.9370$ ，给定

$\alpha = 0.05$ ，查 F 分布表，得临界值 $F_{0.05}(6,6) = 4.28$ 。请你继续完成上述工作，并

回答所做的是一项什么工作，其结论是什么？

(2) 利用 y 对 x 回归所得的残差平方构造一个辅助回归函数：

$$\hat{\sigma}_t^2 = 242407.2 + 1.2299\hat{\sigma}_{t-1}^2 - 1.4090\hat{\sigma}_{t-2}^2 + 1.0188\hat{\sigma}_{t-3}^2$$

$$R^2 = 0.5659, \text{计算 } (n-p)R^2 = 18 * 0.5659 = 10.1862$$

给定显著性水平 $\alpha = 0.05$ ，查 χ^2 分布表，得临界值 $\chi_{0.05}^2(3) = 7.81$ ，其中 $p=3$ ，

自由度。请你继续完成上述工作，并回答所做的是一项什么工作，其结论是什么？

(3) 试比较 (1) 和 (2) 两种方法，给出简要评价。

解：(1) 这是异方差检验，使用的是样本分段拟和 (Goldfeld-Quant)， $F = 4334.937 > 4.28$ ，因此拒绝原假设，表明模型中存在异方差。

(2) 这是异方差 ARCH 检验， $(n-p)R^2 = 18 * 0.5659 = 10.1862 > 7.81$ ，所以拒绝原假设，表明模型中存在异方差。

(3) 这两种方法都是用于检验异方差。但二者适用条件不同：A、Goldfeld-Quandt 要求大样本；扰动项正态分布；可用于截面数据和时间序列数据。B、ARCH 检验仅适宜于时间序列数据，检验统计量的极限分布为 χ^2 -分布。

第六套

一、单项选择题

- 1、计量经济学的研究方法一般分为以下四个步骤（ **B** ）
 - A. 确定科学的理论依据、模型设定、模型修定、模型应用
 - B. 模型设定、估计参数、模型检验、模型应用
 - C. 搜集数据、模型设定、估计参数、预测检验
 - D. 模型设定、模型修定、结构分析、模型应用
- 2、简单相关系数矩阵方法主要用于检验（ **D** ）
 - A. 异方差性
 - B. 自相关性
 - C. 随机解释变量
 - D. 多重共线性
- 3、在某个结构方程恰好识别的条件下，不适用的估计方法是（ **D** ）
 - ~~A. 间接最小二乘法~~
 - ~~B. 工具变量法~~
 - ~~C. 二阶段最小二乘法~~
 - ~~D. 普通最小二乘法~~
- 4、在利用月度数据构建计量经济模型时，如果一年里的 12 个月全部表现出季节模式，则应该引入虚拟变量个数为（ **C** ）
 - A. 4
 - B. 12
 - C. 11
 - D. 6
- 5、White 检验可用于检验（ **B** ）
 - A. 自相关性
 - B. 异方差性
 - C. 解释变量随机性
 - D. 多重共线性
- 6、如果回归模型违背了无自相关假定，最小二乘估计量是（ **C** ）
 - A. 无偏的，有效的
 - B. 有偏的，非有效的
 - C. 无偏的，非有效的
 - D. 有偏的，有效的
- 7、假如联立方程模型中，第 i 个方程排除的变量中没有有一个在第 j 个方程中出现，则第 i 个方程是（ **D** ）
 - ~~A. 可识别的~~
 - ~~B. 恰好识别~~
 - ~~C. 过度识别~~
 - ~~D. 不可识别~~
- 8、在简单线性回归模型中，认为具有一定概率分布的随机变量是（ **A** ）
 - A. 内生变量
 - B. 外生变量
 - C. 虚拟变量
 - D. 前定变量
- 9、应用 DW 检验方法时应满足该方法的假定条件，下列不是其假定条件的为（ **B** ）
 - A. 解释变量为非随机的
 - B. 被解释变量为非随机的
 - C. 线性回归模型中不能含有滞后内生变量

D.随机误差项服从一阶自回归

10、二元回归模型中，经计算有相关系数 $R_{X_2X_3} = 0.9985$ ，则表明 (**B**)

A. X_2 和 X_3 间存在完全共线性

B. X_2 和 X_3 间存在不完全共线性

C. X_2 对 X_3 的拟合优度等于 0.9985

D.不能说明 X_2 和 X_3 间存在多重共线性

11、在 DW 检验中，存在正自相关的区域是 (**B**)

A. $4 - d_L < d < 4$

B. $0 < d < d_L$

C. $d_U < d < 4 - d_U$

D. $d_L < d < d_U, 4 - d_U < d < 4 - d_L$

12、库伊克模型不具有如下特点 (**D**)

A. 原始模型为无限分布滞后模型，且滞后系数按某一固定比例递减

B. 以一个滞后被解释变量 Y_{t-1} 代替了大量的滞后解释变量 X_{t-1}, X_{t-2}, \dots ，

从而最大限度的保证了自由度

C. 滞后一期的被解释变量 Y_{t-1} 与 X_t 的线性相关程度肯定小于 X_{t-1}, X_{t-2}, \dots

的相关程度，从而缓解了多重共线性的问题

D. 由于 $Cov(Y_{t-1}, u_t^*) = 0$, $Cov(u_t^*, u_{t-1}^*) = 0$ ，因此可使用 OLS 方法估计参数，参数估计量是一致估计量

13、在具体运用加权最小二乘法时，如果变换的结果是 $\frac{Y_t}{X_t} = \beta_1 \frac{1}{X_t} + \beta_2 + \frac{u_t}{X_t}$ ，

则 $Var(u_t)$ 是下列形式中的哪一种?(**B**)

A. $\sigma^2 X_t$

B. $\sigma^2 X_t^2$

C. $\sigma^2 \sqrt{X_t}$

D. $\sigma^2 \log(X_t)$

~~14、下列是简化的三部门宏观经济计量模型，则模型中前定变量的个数为 (**A**)~~

$$\begin{cases} \cancel{Y_t = C_t + I_t + G} \\ \cancel{C_t = \alpha_0 + \alpha_1 Y_t + u_{1t}} \\ \cancel{I_t = \beta_0 + \beta_1 Y_{t-1} + \beta_2 Y_t + u_{2t}} \end{cases}$$

A. 3

B. 4

C. 2

D. 6

~~15、在异方差的情况下，参数估计值仍是无偏的，其原因是 (**D**)~~

- A.零均值假定不成立 B.序列无自相关假定成立
- C.无多重共线性假定成立 D.解释变量与随机误差项不相关假定成立

16、已知 DW 统计量的值接近于 2，则样本回归模型残差的一阶自相关系数 $\hat{\rho}$ 近似等于(**A**)

- A. 0 B. -1 C. 1 D. 4

~~17、对美国储蓄与收入关系的计量经济模型分成两个时期分别建模，重建时期是 1946—1954;重建后时期是 1955—1963，模型如下：~~

~~重建时期： $Y_t = \lambda_1 + \lambda_2 X_t + \mu_{1t}$~~

~~重建后时期： $Y_t = \lambda_3 + \lambda_4 X_t + \mu_{2t}$~~

关于上述模型，下列说法不正确的是 (**D**)

- A. $\lambda_1 = \lambda_3$ $\lambda_2 = \lambda_4$ 时则称为重合回归 B. $\lambda_1 \neq \lambda_3$ $\lambda_2 = \lambda_4$ 时称为平行回归
- C. $\lambda_1 \neq \lambda_3$ $\lambda_2 \neq \lambda_4$ 时称为相异回归 D. $\lambda_1 \neq \lambda_3$ $\lambda_2 = \lambda_4$ 两个模型没有差异

18、对样本的相关系数 γ ，以下结论错误的是 (**A**)

- A. $|\gamma|$ 越接近 0， X 与 Y 之间线性相关程度高
- B. $|\gamma|$ 越接近 1， X 与 Y 之间线性相关程度高
- C. $-1 \leq \gamma \leq 1$
- D. $\gamma = 0$ ，在正态分布下，则 X 与 Y 相互独立

19、对于二元样本回归模型 $Y_i = \hat{\beta}_1 + \hat{\beta}_2 X_{2i} + \hat{\beta}_3 X_{3i} + e_i$ ，下列不成立的有 (**D**)

- A. $\sum e_i = 0$ B. $\sum e_i X_{2i} = 0$
- C. $\sum e_i X_{3i} = 0$ D. $\sum e_i Y_i = 0$

~~20、当联立方程模型中第 i 个结构方程是不可识别的，则该模型是(**B**)~~

A.可识别的 B.不可识别的 C.过度识别的 D.恰好识别的

二、多项选择题

1、关于自适应预期模型和局部调整模型，下列说法不正确的有（ **C E** ）

- A. 它们都是由某种期望模型演变形成的
- B. 它们最终都是一阶自回归模型
- C. 它们都是库伊克模型的特例
- D. 它们的经济背景不同
- E. 都满足古典线性回归模型的所有假设，从而可直接用 OLS 进行估计

2、能够检验多重共线性的方法有（ **A B** ）

- A. 简单相关系数矩阵法
- B. t 检验与 F 检验综合判断法
- C. DW 检验法
- D. $ARCH$ 检验法
- E. $White$ 检验

3、有关调整后的判定系数 \bar{R}^2 与判定系数 R^2 之间的关系叙述正确的有（**B C**）

- A. \bar{R}^2 与 R^2 均非负
- B. 模型中包含的解释个数越多， \bar{R}^2 与 R^2 就相差越大.
- C. 只要模型中包括截距项在内的参数的个数大于 1，则 $\bar{R}^2 < R^2$.
- D. \bar{R}^2 有可能大于 R^2
- E. R^2 有可能小于 0，但 \bar{R}^2 却始终是非负

4、检验序列自相关的方法是（ **C E** ）

- A. F 检验法
- B. $White$ 检验法
- C. 图形法
- D. $ARCH$ 检验法
- E. DW 检验法
- F. $Goldfeld-Quandt$ 检验法

5、对多元线性回归方程的显著性检验，所用的 F 统计量可表示为（ **B E** ）

- A. $\frac{ESS/(n-k)}{RSS/(k-1)}$
- B. $\frac{ESS/(k-1)}{RSS/(n-k)}$
- C. $\frac{R^2/(n-k)}{(1-R^2)/(k-1)}$
- D. $\frac{ESS}{RSS/(n-k)}$
- E. $\frac{R^2/(k-1)}{(1-R^2)/(n-k)}$

三、判断题（判断下列命题正误，并说明理由）

1、在对参数进行最小二乘估计之前，没有必要对模型提出古典假定。

错误

在古典假定条件下，OLS 估计得到的参数估计量是该参数的最佳线性无偏估计（具有线性、无偏性、有效性）。总之，提出古典假定是为了使所作出的估计量具有较好的统计性质和方便地进行统计推断。

2、当异方差出现时，常用的 t 和 F 检验失效；

正确

由于异方差类，似于 t 比值的统计量所遵从的分布未知；即使遵从 t -分布，由于方差不具有最小性。这时往往会夸大 t -检验，使得 t 检验失效；由于 F -分布为两个独立的 χ^2 变量之比，故依然存在类似于 t -分布中的问题。

3、解释变量与随机误差项相关，是产生多重共线性的主要原因。

错误

产生多重共线性的主要原因是：经济本变量大多存在共同变化趋势；模型中大量采用滞后变量；认识上的局限使得选择变量不当；……。

4、如果经典线性回归模型（CLRM）中的干扰项不服从正态分布，OLS 估计量将有偏的。

错误

即使经典线性回归模型（CLRM）中的干扰项不服从正态分布，OLS 估计量仍然是无偏的。因为 $E(\hat{\beta}_2) = E(\beta_2 + \sum K_i \mu_i) = \beta_2$ ，该表达式成立与否与正态性无关。

~~5、由间接最小二乘法与两阶段最小二乘法得到的估计量都是无偏估计。~~

错误

间接最小二乘法适用于恰好识别方程的估计，其估计量为无偏估计；

而两阶段最小二乘法不仅适用于恰好识别方程，也适用于过度识别方程。两阶段最小二乘法得到的估计量为有偏、一致估计。

~~四、计算题~~

~~1、为了研究深圳市地方预算内财政收入与国内生产总值的关系，得到以下数据：~~

年 份	地方预算内财政收入 Y (亿元)	国内生产总值(GDP)X (亿元)
1990	21.7037	171.6665
1991	27.3291	236.6630
1992	42.9599	317.3194
1993	67.2507	449.2889
1994	74.3992	615.1933
1995	88.0174	795.6950
1996	131.7490	950.0446
1997	144.7709	1130.0133
1998	164.9067	1289.0190
1999	184.7908	1436.0267
2000	225.0212	1665.4652
2001	265.6532	1954.6539

资料来源：《深圳统计年鉴 2002》，中国统计出版社
 利用 EViews 估计其参数结果为

Dependent Variable: Y
 Method: Least Squares
 Date: 07/01/05 Time: 20:15
 Sample: 1990 2001
 Included observations: 12

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-3.611151	4.161790	-0.867692	0.4059
GDP	0.134582	0.003867	34.80013	0.0000
R-squared	0.991810	Mean dependent var	119.8793	
Adjusted R-squared	0.990991	S.D. dependent var	79.36124	
S.E. of regression	7.532484	Akaike info criterion	7.027338	
Sum squared resid	567.3831	Schwarz criterion	7.108156	
Log likelihood	-40.16403	F-statistic	1211.049	
Durbin-Watson stat	2.051640	Prob(F-statistic)	0.000000	

- (1)建立深圳地方预算内财政收入对 GDP 的回归模型；
- (2)估计所建立模型的参数，解释斜率系数的经济意义；
- (3)对回归结果进行检验；
- (4)若是 2005 年年的国内生产总值为 3600 亿元，确定 2005 年财政收入的预测值和预测区间($\alpha = 0.05$)。

解：地方预算内财政收入（Y）和 GDP 的关系近似直线关系，可建立线性回归模型：

$$Y_t = \beta_1 + \beta_2 GDP_t + u_t$$

即

$$\hat{Y}_t = -3.611151 + 0.134582 GDP_t$$

(4.161790) (0.003867)

$$t=(-0.867692) \quad (34.80013)$$

$$R^2=0.99181 \quad F=1211.049$$

$R^2=0.99181$ ，说明GDP解释了地方财政收入变动的 99%，模型拟合程度较好。

模型说明当 GDP 每增长 1 亿元，平均说来地方财政收入将增长 0.134582 亿元。

当 2005 年 GDP 为 3600 亿元时，地方财政收入的点预测值为：

$$\hat{Y}_{2005} = -3.611151 + 0.134582 \times 3600 = 480.884 \text{ (亿元)}$$

区间预测：

$$\sum x_i^2 = \sigma_x^2(n-1) = 587.2686^2 \times (12-1) = 3793728.494$$

$$(X_{f1} - \bar{X})^2 = (3600 - 917.5874)^2 = 7195337.357$$

取 $\alpha = 0.05$, Y_f 平均值置信度 95% 的预测区间为：

$$\hat{Y}_f \mp t_{\alpha/2} \hat{\sigma} \sqrt{\frac{1}{n} + \frac{(X_f - \bar{X})^2}{\sum x_i^2}}$$

$GDP_{2005} = 3600$ 时

$$480.884 \mp 2.228 \times 7.5325 \times \sqrt{\frac{1}{12} + \frac{7195337.357}{3293728.494}} = 480.884 \mp 25.2735 \text{ (亿元)}$$

Y_f 个别值置信度 95% 的预测区间为：

$$\hat{Y}_f \mp t_{\alpha/2} \hat{\sigma} \sqrt{1 + \frac{1}{n} + \frac{(X_f - \bar{X})^2}{\sum x_i^2}}$$

$$\text{即} \quad = 480.884 \mp 2.228 \times 7.5325 \times \sqrt{1 + \frac{1}{12} + \frac{7195337.357}{3293728.494}}$$

$$= 480.884 \mp 30.3381 \text{ (亿元)}$$

2、运用美国 1988 研究与开发 (R&D) 支出费用 (Y) 与不同部门产品销售量 (X) 的数据建立了一个回归模型，并运用 Glejser 方法和 White 方法检验异方差，由此决定异方差的表现形式并选用适当方法加以修正。结果如下：

$$\hat{Y} = 192.9944 + 0.0319X$$

$$(0.1948) \quad (3.83)$$

$$R^2 = 0.4783, s.e. = 2759.15, F = 14.6692$$

White Heteroskedasticity Test:

F-statistic	3.057161	Probability	0.076976
Obs*R-squared	5.212471	Probability	0.073812

Test Equation:
Dependent Variable: RESID^2
Method: Least Squares
Date: 08/08/05 Time: 15:38
Sample: 1 18
Included observations: 18

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-6219633.	6459811.	-0.962820	0.3509
X	229.3496	126.2197	1.817066	0.0892
X^2	-0.000537	0.000449	-1.194942	0.2507
R-squared	0.289582	Mean dependent var	6767029.	
Adjusted R-squared	0.194859	S.D. dependent var	14706003	
S.E. of regression	13195642	Akaike info criterion	35.77968	
Sum squared resid	2.61E+15	Schwarz criterion	35.92808	
Log likelihood	-319.0171	F-statistic	3.057161	
Durbin-Watson stat	1.694572	Prob(F-statistic)	0.076976	

$$|\hat{e}| = 6.4435\sqrt{X}$$

(4.5658)

$$R^2 = 0.2482$$

请问：（1）White 检验判断模型是否存在异方差。
（2）Glejser 检验判断模型是否存在异方差。
（3）该怎样修正。

解：（1）给定 $\alpha = 0.05$ 和自由度为 2 下，查卡方分布表，得临界值 $\chi^2 = 5.9915$ ，而 White 统计量 $nR^2 = 5.2125$ ，有 $nR^2 < \chi^2_{0.05}(2)$ ，则不能拒绝原假设，说明模型中不存在异方差。

（2）因为对如下函数形式

$$|e| = \beta_2\sqrt{X} + \varpi$$

得样本估计式

$$|\hat{e}| = 6.4435\sqrt{X}$$

(4.5658)

$$R^2 = 0.2482$$

由此，可以看出模型中随机误差项有可能存在异方差。

(3) 对异方差的修正。可取权数为 $w=1/X$ 。

3、Sen 和 Srivastava (1971) 在研究贫富国之间期望寿命的差异时，利用 101 个国家的数据，建立了如下的回归模型：

$$\hat{Y}_i = -2.40 + 9.39 \ln X_i - 3.36(D_i(\ln X_i - 7))$$

$$(4.37) \quad (0.857) \quad (2.42)$$

$$R^2=0.752$$

其中：X 是以美元计的人均收入；

Y 是以年计的期望寿命；

Sen 和 Srivastava 认为人均收入的临界值为 1097 美元 ($\ln 1097 = 7$)，若人均收入超过 1097 美元，则被认定为富国；若人均收入低于 1097 美元，被认定为贫穷国。

(括号内的数值为对应参数估计值的 t-值)。

(1) 解释这些计算结果。

(2) 回归方程中引入 $D_i(\ln X_i - 7)$ 的原因是什么？如何解释这个回归解释变量？

(3) 如何对贫穷国进行回归？又如何对富国进行回归？

解：(1) 由 $\ln X = 1 \Rightarrow X = 2.7183$ ，也就是说，人均收入每增加 1.7183 倍，平均意义上各国的期望寿命会增加 9.39 岁。若当为富国时， $D_i = 1$ ，则平均意义上，富国的人均收入每增加 1.7183 倍，其期望寿命就会减少 3.36 岁，但其截距项的水平会增加 23.52，达到 21.12 的水平。但从统计检验结果看，对数人均收入 $\ln X$ 对期望寿命 Y 的影响并不显著。方程的拟合情况良好，可进一步进行多重共线性等其他计量经济学的检验。

(2) 若 $D_i = 1$ 代表富国，则引入 $D_i(\ln X_i - 7)$ 的原因是想从截距和斜率两个方面考证富国的影响，其中，富国的截距为 $(-2.40 + 3.36 \times 7 = 21.12)$ ，斜率为 $(9.39 - 3.36 = 6.03)$ ，因此，当富国的人均收入每增加 1.7183 倍，其期望寿命会增加 6.03 岁。

(3) 对于贫穷国，设定 $D_i = \begin{cases} 1 & \text{若为贫穷国} \\ 0 & \text{若为富国} \end{cases}$ ，则引入的虚拟解释变量的形式

为 $(D_i(7 - \ln X_i))$ ；对于富国，回归模型形式不变。