

## 计量经济学试题一及答案

### 一、判断题（15分）

1. 线性回归模型中，解释变量是原因，被解释变量是结果。（F）
2. 多元回归模型统计显著是指模型中每个变量都是统计显著的。（F）
3. 在存在异方差情况下，常用的 OLS 法总是高估了估计量的标准差。（F）
4. 总体回归线是当解释变量取给定值时因变量的条件均值的轨迹。（Y）
5. 线性回归是指解释变量和被解释变量之间呈现线性关系。（F）
6. 判定系数  $R^2$  的大小不受回归模型中所包含的解释变量个数的影响。（F）
7. 多重共线性是一种随机误差现象。（F）
8. 当存在自相关时，OLS 估计量是有偏的并且也是无效的。（F）
9. 在异方差的情况下，OLS 估计量误差放大的原因是从属回归的  $R^2$  变大。（F）
10. 任何两个计量经济模型的  $R^2$  都是可以比较的。（F）

### 二、简答题（15分）

1. 计量经济模型分析经济问题的基本步骤。（4分）

答：建立与应用计量经济学模型的主要步骤如下：（1）设定理论模型，包括选择模型包含的变量，确定变量之间的数学关系和拟定模型中的待估参数的数值范围；（2）收集样本数据，要考虑样本数据的完整性、准确性、可比性和一致性；（3）估计模型参数；（4）检验模型，包括经济意义检验、统计检验、计量检验和预测检验（5）计量经济模型的应用。

2. 举例说明如何引进加法模式和乘法模式建立虚拟变量模型。（6分）

答案：设 Y 为个人消费支出；X 表示可支配收入，定义

$$D_{2t} = \begin{cases} 1 & \text{2季度} \\ 0 & \text{其他} \end{cases} \quad D_{3t} = \begin{cases} 1 & \text{3季度} \\ 0 & \text{其他} \end{cases} \quad D_{4t} = \begin{cases} 1 & \text{4季度} \\ 0 & \text{其他} \end{cases}$$

如果设定模型为

$$Y_t = B_1 + B_2 D_{2t} + B_3 D_{3t} + B_4 D_{4t} + B_5 X_t + u_t$$

此时模型仅影响截距项，差异表现为截距项的和，因此也称为加法模型。

如果设定模型为

$$Y_t = B_1 + B_2 D_{2t} + B_3 D_{3t} + B_4 D_{4t} + B_5 X_t \\ + B_6 (D_{2t} X_t) + B_7 (D_{3t} X_t) + B_8 (D_{4t} X_t) + u_t$$

此时模型不仅影响截距项，而且还影响斜率项。差异表现为截距和斜率的双重变化，因此也称为乘法模型。

三．下面是我国 1990-2003 年 GDP 对 M1 之间回归的结果。（5 分）

$$\begin{array}{lcl} \ln(GDP) = 1.37 & + & 0.76 \ln(M1) \\ \text{se} & (0.15) & (0.033) \\ t & (9.13) & (23) \end{array}$$

$P(|t| > 1.782) = 0.05$ , 自由度 = 12;

1. 求出空白处的数值，填在括号内。（2 分）

2. 系数是否显著，给出理由。（3 分）

答：根据 t 统计量，9.13 和 23 都大于 5% 的临界值，因此系数都是统计显著的。

四．试述异方差的后果及其补救措施。（10 分）

答案：

后果：OLS 估计量是线性无偏的，不是有效的，估计量方差的估计有偏。建立在 t 分布和 F 分布之上的置信区间和假设检验是不可靠的。

补救措施：加权最小二乘法（WLS）

1. 假设  $\sigma_i^2$  已知，则对模型进行如下变换：

$$\frac{Y_i}{\sigma_i} = \frac{B_1}{\sigma_i} + B_2 \frac{X_i}{\sigma_i} + \frac{u_i}{\sigma_i}$$

2. 如果  $\sigma_i^2$  未知

(1) 误差与  $X_i$  成比例：平方根变换。

$$\frac{Y_i}{\sqrt{X_i}} = \frac{B_1}{\sqrt{X_i}} + B_2 \frac{X_i}{\sqrt{X_i}} + \frac{u_i}{\sqrt{X_i}}$$

可见，此时模型同方差，从而可以利用 OLS 估计和假设检验。

(2) 误差方差和  $X_i^2$  成比例。即  $E(u_i^2) = \sigma^2 X_i^2$

$$\frac{Y_i}{X_i} = \frac{B_1}{X_i} + B_2 \frac{X_i}{X_i} + \frac{u_i}{X_i}$$

3. 重新设定模型：

五．多重共线性的后果及修正措施。（10 分）

1) 对于完全多重共线性，后果是无法估计。

对于高度多重共线性，理论上不影响 OLS 估计量的最优线性无偏性。但对于个别样本的估计量的方差放大，从而影响了假设检验。

实际后果：联合检验显著，但个别系数不显著。估计量的方差放大，置信区间变宽，t 统计量变小。对于样本内观测值得微小变化极敏感。某些系数符号可能不对。难以解释自变量对应变量的贡献程度。

2) 补救措施：剔出不重要变量；增加样本数量；改变模型形式；改变变量形式；利用先验信息。

六. 试述 D-W 检验的适用条件及其检验步骤？（10 分）

答案：使用条件：

- 1) 回归模型包含一个截距项。
- 2) 变量 X 是非随机变量。
- 3) 扰动项的产生机制： $u_t = \rho u_{t-1} + v_t \quad -1 \leq \rho \leq 1$ 。
- 4) 因变量的滞后值不能作为解释变量出现在回归方程中。

检验步骤

- 1) 进行 OLS 回归，并获得残差。
- 2) 计算 D 值。
- 3) 已知样本容量和解释变量个数，得到临界值。
- 4) 根据下列规则进行判断：

零假设	决策	条件
无正的自相关	拒绝	$0 < d < d_L$
无正的自相关	无法确定	$d_L < d < d_U$
无负的自相关	拒绝	$4 - d_L < d < 4$
无负的自相关	无法决定	$4 - d_U < d < 4 - d_L$
无正的或者负的自相关	接受	$d_U < d < 4 - d_U$

七. （15 分）下面是宏观经济模型

$$\begin{aligned} M_t &= C(1) * P_t + C(2) * Y_t + C(3) * I_t + C(4) * M_{t-1} + u_t^D \\ I_t &= C(5) * M_t + C(6) * Y_t + u_t^C \\ Y_t &= C(7) * I_t + u_t^A \end{aligned}$$

变量分别为货币供给 M、投资 I、价格指数 P 和产出 Y。

1. 指出模型中哪些是内生变量，哪些是外生变量。（5分）

答：内生变量为货币供给  $M_t$ 、投资  $I_t$  和产出  $Y_t$ 。

外生变量为滞后一期的货币供给  $M_{t-1}$  以及价格指数  $P_t$

2. 对模型进行识别。（4分）

答：根据模型识别的阶条件

方程（1）： $R(B_0 \Gamma_0) = k = 0 < m - 1 = 2$ ，不可识别。

方程（2）： $R(B_0 \Gamma_0) = k = 2 = m - 1$ ，恰好识别。

方程（3）： $R(B_0 \Gamma_0) = k = 2 = m - 1$ ，恰好识别。

3. 指出恰好识别方程和过度识别方程的估计方法。（6分）

答：对于恰好识别方程，采用间接最小二乘法。首先建立简化方程，之后对简化方程进行最小二乘估计。

对于过度识别方程，采用两阶段最小二乘法。首先求替代变量（工具变量），再把这个工具变量作为自变量进行回归。

#### 八、（20分）应用题

为了研究我国经济增长和国债之间的关系，建立回归模型。得到的结果如下：

Dependent Variable: LOG(GDP)

Method: Least Squares

Date: 06/04/05 Time: 18:58

Sample: 1985 2003

Included observations: 19

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	6.03	0.14	43.2	0
LOG(DEBT)	0.65	0.02	32.8	0
R-squared	0.981	Mean dependent var		10.53
Adjusted R-squared	0.983	S.D. dependent var		0.86
S.E. of regression	0.11	Akaike info criterion		-1.46
Sum squared resid	0.21	Schwarz criterion		-1.36
Log likelihood	15.8	F-statistic		1075.5

---

若  $k=1, n=19, d_L=1.074, d_U=1.536$ , 显著性水平  $\alpha=0.05$

其中, GDP 表示国内生产总值, DEBT 表示国债发行量。

(1) 写出回归方程。(2 分)

答:  $\text{Log}(\text{GDP}) = 6.03 + 0.65 \text{ LOG}(\text{DEBT})$

(2) 解释系数的经济学含义?(4 分)

答: 截距项表示自变量为零时, 因变量的平均期望。不具有实际的经济学含义。

斜率系数表示 GDP 对 DEBT 的不变弹性为 0.65。或者表示增发 1% 国债, 国民经济增长 0.65%。

(3) 模型可能存在什么问题? 如何检验?(7 分)

答: 可能存在序列相关问题。

因为  $d.w = 0.81$  小于  $d_L = 1.074$ , 因此落入正的自相关区域。由此可以判定存在序列相关。

(4) 如何就模型中所存在的问题, 对模型进行改进?(7 分)

答: 利用广义最小二乘法。根据  $d.w = 0.81$ , 计算得到  $\rho = 0.6$ , 因此回归方程滞后一期后, 两边同时乘以 0.6, 得

$$0.6 \log(\text{GDP}_{t-1}) = 0.4B_1 + 0.6B_2 \log(\text{DEBT}_{t-1}) + 0.6u_{t-1}$$

方程

$$\log(\text{GDP}_t) = B_1 + B_2 \log(\text{DEBT}_t) + u_t$$

减去上面的方程, 得到

$$\log(\text{GDP}_t) - 0.6 \log(\text{GDP}_{t-1}) = 0.6B_1 + B_2 (\log(\text{DEBT}_t) - 0.6 \log(\text{DEBT}_{t-1})) + v_t$$

利用最小二乘估计, 得到系数。

计量经济学试题二及答案

一、判断正误（20 分）

1. 随机误差项  $u_i$  和残差项  $e_i$  是一回事。（ F ）
2. 给定显著性水平  $\alpha$  及自由度，若计算得到的  $|t|$  值超过临界的  $t$  值，我们将接受零假设（ F ）
3. 利用 OLS 法求得的样本回归直线  $\hat{Y}_i = b_1 + b_2 X_i$  通过样本均值点  $(\bar{X}, \bar{Y})$ 。（ T ）
4. 判定系数  $R^2 = TSS/ESS$ 。（ F ）
5. 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。（ F ）
6. 双对数模型的  $R^2$  值可以与对数线性模型的相比较，但不能与线性对数模型相比较。（ T ）
7. 为了避免陷入虚拟变量陷阱，如果一个定性变量有  $m$  类，则要引入  $m$  个虚拟变量。（ F ）
8. 在存在异方差情况下，常用的 OLS 法总是高估了估计量的标准差。（ T ）
9. 识别的阶条件仅仅是判别模型是否可识别的必要条件而不是充分条件。（ T ）
10. 如果零假设  $H_0: B_2=0$ ，在显著性水平 5% 下不被拒绝，则认为  $B_2$  一定是 0。（ F ）

二、以一元回归为例叙述普通最小二乘回归的基本原理。（10 分）

解：依据题意有如下的一元样本回归模型：

$$Y_i = b_1 + b_2 X_i + e_i \quad (1)$$

普通最小二乘原理是使得残差平方和最小，即

$$\min Q = \min \sum e_i^2 = \min \sum (Y_i - b_1 - b_2 X_i)^2 \quad (2)$$

根据微积分求极值的原理，可得

$$\frac{\partial Q}{\partial b_1} = 0 \Leftrightarrow \frac{\partial Q}{\partial b_1} = -2 \sum (Y_i - b_1 - b_2 X_i) = 0 \quad (3)$$

$$\frac{\partial Q}{\partial b_2} = 0 \Leftrightarrow \frac{\partial Q}{\partial b_2} = -2 \sum (Y_i - b_1 - b_2 X_i) X_i = 0 \quad (4)$$

将(3)和(4)式称为正规方程，求解这两个方程，我们可得到：

$$\begin{aligned}\sum Y_i &= nb_1 + b_2 \sum X_i \\ \sum Y_i X_i &= b_1 \sum X_i + b_2 \sum X_i^2\end{aligned}\quad (5)$$

解得：

$$\begin{aligned}b_1 &= \bar{Y} - b_2 \bar{X} \\ b_2 &= \frac{\sum x_i y_i}{\sum x_i^2}\end{aligned}$$

其中  $x_i = X_i - \bar{X}$ ,  $y_i = Y_i - \bar{Y}$ , 表示变量与其均值的离差。

三、下面是利用 1970-1980 年美国数据得到的回归结果。其中 Y 表示美国咖啡消费（杯/日·人），X 表示平均零售价格（美元/磅）。（15 分） 注： $t_{\alpha/2}(9) = 2.262$ ,  $t_{\alpha/2}(10) = 2.228$

$$\begin{aligned}\hat{Y}_t &= 2.6911 - 0.4795 X_t \\ se &= (0.1216) \quad (a) \\ t\text{值} &= (b) \quad 42.06 \quad R^2 = 0.6628\end{aligned}$$

1. 写空白处的数值啊  $a$ ,  $b$ 。（0.0114, 22.066）
2. 对模型中的参数进行显著性检验。
3. 解释斜率系数  $B_2$  的含义，并给出其 95% 的置信区间。

解：1. （0.0114, 22.066）

2.  $B_1$  的显著性检验： $t = 22.066 > t_{\alpha/2}(9) = 2.262$ , 所以  $B_1$  是显著的。

$B_2$  的显著性检验： $t = 42.06 > t_{\alpha/2}(9) = 2.262$ , 所以  $B_2$  是显著的。

3.  $B_2$  表示每磅咖啡的平均零售价格每上升 1 美元，每人每天的咖啡消费量减少 0.479 杯。

$$\begin{aligned}P(-2.262 \leq t \leq 2.262) &= 0.95 \\ P\left(-2.262 \leq \frac{b_2 - B_2}{se(b_2)} \leq 2.262\right) &= 0.95 \\ P(b_2 - 2.262 se(b_2) \leq B_2 \leq b_2 + 2.262 se(b_2)) &= 0.95\end{aligned}$$

$B_2$  的 95% 的置信区间为：

$$\begin{aligned}[-0.479 - 0.026, -0.479 + 0.026] \\ [-0.505454, -0.454]\end{aligned}$$

四、若在模型： $Y_t = B_1 + B_2 X_t + u_t$  中存在下列形式的异方差： $\text{var}(u_t) = \sigma^2 X_t^3$ ，你如何估计参数  $B_1, B_2$ （10 分）

解：对于模型

$$Y_t = B_1 + B_2 X_t + u_t \quad (1)$$

存在下列形式的异方差： $\text{var}(u_t) = \sigma^2 X_t^3$ ，我们可以在(1)式左右两端同时除以  $\sqrt{X_t^3}$ ，可得

$$\begin{aligned} \frac{Y_t}{\sqrt{X_t^3}} &= B_1 \frac{1}{\sqrt{X_t^3}} + B_2 \frac{X_t}{\sqrt{X_t^3}} + \frac{u_t}{\sqrt{X_t^3}} \\ &= B_1 \frac{1}{\sqrt{X_t^3}} + B_2 \frac{X_t}{\sqrt{X_t^3}} + v_t \end{aligned} \quad (2)$$

其中

$$v_t = \frac{u_t}{\sqrt{X_t^3}}$$

代表误差修正项，可以证明

$$\text{var}(v_t) = \text{var}\left(\frac{u_t}{\sqrt{X_t^3}}\right) = \frac{1}{X_t^3} \text{var}(u_t) = \frac{1}{X_t^3} \sigma^2 X_t^3 = \sigma^2$$

即  $v_t$  满足同方差的假定，对(2)式使用 OLS，即可得到相应的估计量。

五、考虑下面的模型： $Y_t = B_0 + B_1 X_t + B_2 D_{2t} + B_3 D_{3t} + B_4 D_{4t} + u_t$  其中，Y 表示大学教师的年薪收入，X 表示工龄。为了研究大学教师的年薪是否受到性别（男、女）、学历（本科、硕士、博士）的影响。按照下面的方式引入虚拟变量：（15 分）

$$D_2 = \begin{cases} 1, & \text{男教师} \\ 0, & \text{女教师} \end{cases} \quad D_3 = \begin{cases} 1, & \text{硕士} \\ 0, & \text{其他} \end{cases} \quad D_4 = \begin{cases} 1, & \text{博士} \\ 0, & \text{其他} \end{cases}$$

1. 基准类是什么？
2. 解释各系数所代表的含义，并预期各系数的符号。
3. 若  $B_4 > B_3$ ，你得出什么结论？

解：1. 基准类为本科女教师。

2.  $B_1$  表示工龄对年薪的影响，即工龄每增加 1 单位，平均而言，年薪将增加  $B_1$  个单位。预期符号为正，因为随着年龄的增加，工资应该增加。



$B_2$  体现了性别差异。

$B_3$  和  $B_4$  体现了学历差异，预期符号为正。

3.  $B_4 > B_3$  说明，博士教师的年薪高于硕士教师的年薪。

六、什么是自相关？杜宾—瓦尔森检验的前提条件和步骤是什么？（15 分）

解：自相关，在时间（如时间序列数据）或者空间（如在截面数据中）上按顺序排列的序列的各成员之间存在着相关关系。在计量经济学中指回归模型中随机扰动项之间存在相关关系。用符号表示：

$$\text{cov}(u_i, u_j) = E(u_i u_j) \neq 0 \quad i \neq j$$

杜宾—瓦尔森检验的前提条件为：

- (1) 回归模型包括截距项。
- (2) 变量  $X$  是非随机变量。
- (3) 扰动项  $u_t$  的产生机制是

$$u_t = \rho u_{t-1} + v_t \quad (-1 \leq \rho \leq 1, \text{表示自相关系数})$$

上述这个描述机制我们称为一阶自回归模型，通常记为 AR(1)。

(4) 在回归方程的解释变量中，不包括把因变量的滞后变量。即检验对于自回归模型是不使用的。

杜宾—瓦尔森检验的步骤为：

- (1) 进行 OLS 的回归并获得  $e_t$ 。
- (2) 计算  $d$  值。
- (3) 给定样本容量  $n$  和解释变量  $k$  的个数，从临界值表中查得  $d_L$  和  $d_U$ 。
- (4) 根据相应的规则进行判断。

七、考虑下面的联立方程模型：
$$\begin{cases} Q_t = A_1 + A_2 P_t + A_3 X_t + u_{1t} \\ Q_t = B_1 + B_2 P_t + u_{2t} \end{cases}$$
 其中， $P$ ， $Q$  是内生变量， $X$  是外生变量， $u$  是随机误差项（15 分）

- 1、求简化形式回归方程？
- 2、判定哪个方程是可识别的（恰好或过度）？

3、对可识别方程，你将用哪种方法进行估计，并简述基本过程？

解 1.

$$P_t = \Pi_1 + \Pi_2 X_t + v_{1t}$$

其中：

$$\Pi_1 = \frac{B_1 - A_1}{A_2 - B_2}, \quad \Pi_2 = -\frac{A_3}{A_2 - B_2}, \quad v_{1t} = \frac{u_{2t} - u_{1t}}{A_2 - B_2} \quad (1)$$

$$Q_t = \Pi_3 + \Pi_4 X_t + v_{2t}$$

其中：

$$\Pi_3 = \frac{A_2 B_1 - A_1 B_2}{A_2 - B_2}, \quad \Pi_4 = -\frac{A_3 B_2}{A_2 - B_2}, \quad v_{1t} = \frac{A_2 u_{2t} - B_2 u_{1t}}{A_2 - B_2} \quad (2)$$

2. 根据阶判断条件， $m = 2$ ，

对于第一个方程， $k=0$ ， $k < m-1$ ，所以第一个方程不可识别。

对于第二个方程， $k=1$ ， $k = m-1$ ，所以第二个方程恰好识别。

3. 对于恰好识别的方程，可以采用二阶段最小二乘法，也可以使用间接最小二乘法。下面将简单介绍间接最小二乘法的基本过程：

步骤 1：从结构方程导出简化方程；

步骤 2：对简化方程的每个方程用 OLS 方法回归；

步骤 3：利用简化方程系数的估计值求结构方程系数的估计值。

### 计量经济学试题三答案

#### 一、判断正误（20 分）

1. 回归分析用来处理一个因变量与另一个或多个自变量之间的因果关系。（ F ）
2. 拟合优度  $R^2$  的值越大，说明样本回归模型对总体回归模型的代表性越强。（ T ）
3. 线性回归是指解释变量和被解释变量之间呈现线性关系。（ F ）
4. 引入虚拟变量后，用普通最小二乘法得到的估计量仍是无偏的。（ T ）
5. 多重共线性是总体的特征。（ F ）
6. 任何两个计量经济模型的  $R^2$  都是可以比较的。（ F ）
7. 异方差会使 OLS 估计量的标准误差高估，而自相关会使其低估。（ F ）
8. 杜宾-瓦尔德森检验能够检验出任何形式的自相关。（ F ）
9. 异方差问题总是存在于横截面数据中，而自相关则总是存在于时间序列数据中。（ F ）
10. 内生变量的滞后值仍然是内生变量。（ F ）

#### 二、选择题（20 分）

1. 在同一时间不同统计单位的相同统计指标组成的数据组合，是（ D ）  
A. 原始数据      B. Pool 数据      C. 时间序列数据      D. 截面数据
2. 下列模型中属于非线性回归模型的是（ C ）  
A.  $Y = \beta_0 + \beta_1 \ln X + u$       B.  $Y = \beta_0 + \beta_1 X + \beta_2 Z + u$   
C.  $Y = \beta_0 + X^{\beta_1} + u$       D.  $Y = \beta_0 + \beta_1 / X + u$
3. 半对数模型  $Y = \beta_0 + \beta_1 \ln X + u$  中，参数  $\beta_1$  的含义是（ C ）  
A. X 的绝对量变化，引起 Y 的绝对量变化  
B. Y 关于 X 的边际变化  
C. X 的相对变化，引起 Y 的期望值绝对量变化  
D. Y 关于 X 的弹性
4. 模型中其数值由模型本身决定的变量是（ B ）  
A. 外生变量      B. 内生变量      C. 前定变量      D. 滞后变量
5. 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，F 统计量的  $p$  值 = 0.0000，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  $Y_t$  的联合影响不显著
6. 根据样本资料估计人均消费支出  $Y$  对人均收入  $X$  的回归模型为  $\ln \hat{Y}_i = 2.00 + 0.75 \ln X_i$ ，这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%
7. 如果回归模型违背了同方差假定，最小二乘估计量是 ( 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
8. 在回归模型满足 DW 检验的前提条件下，当  $d$  统计量等于 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
9. 将一年四个季度对被解释变量的影响引入到包含截距项的回归模型当中，则需要引入虚拟变量的个数为 ( C )
- A. 5      B. 4      C. 3      D. 2
10. 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. 有偏但一致的      B. 有偏且不一致的      C. 无偏且一致的      D. 无偏但不一致的

三、下表给出了三变量模型的回归的结果：（10 分）

方差来源	平方和	自由度 (d. f)	平方和的均值 (MSS)
来自回归 (ESS)	106.58	2	53.29
来自残差 (RSS)	1.8	17	0.106
总离差 (TSS)	108.38	19	_____

注：保留 3 位小数，可以使用计算器。在 5% 的显著性水平下，本题的  $F_\alpha = 4.45$ 。

- 完成上表中空白处内容。
- 求  $R^2$  与  $\bar{R}^2$ 。
- 利用 F 统计量检验  $X_2$  和  $X_3$  对  $Y$  的联合影响，写出简要步骤。

答案： 1. 见题

2. 
$$R^2 = \frac{ESS}{TSS} = \frac{106.58}{108.38} = 0.982$$

$$\overline{R^2} = 1 - (1 - R^2) \frac{n-1}{n-k} = 1 - (1 - 0.982) \frac{19}{17} = 0.980$$

3. 可以利用  $F$  统计量检验  $X_2$  和  $X_3$  对  $Y$  的联合影响。

$$F = \frac{ESS/2}{RSS/17} = \frac{53.29}{0.106} = 502.736 \quad (\text{或} \quad F = \frac{R^2/(k-1)}{(1-R^2)/(n-k)})$$

因为  $F > F_\alpha = 4.45$ ,  $X_2$  和  $X_3$  对  $Y$  的联合影响是显著的。

四、考虑下面的模型： $Y_t = B_0 + B_1X_t + B_2D_{2t} + B_3D_{3t} + B_4D_{4t} + u_t$  其中， $Y$  表示大学教师的年薪收入， $X$  表示工龄。为了研究大学教师的年薪是否受到性别、学历的影响。按照下面的方式引入虚拟变量：（10 分）

$$D_2 = \begin{cases} 1, & \text{男} \\ 0, & \text{女} \end{cases} \quad D_3 = \begin{cases} 1, & \text{硕士} \\ 0, & \text{本科} \end{cases} \quad D_4 = \begin{cases} 1, & \text{博士} \\ 0, & \text{本科} \end{cases}$$

1. 基准类是什么？
2. 解释各系数所代表的含义，并预期各系数的符号。
3. 若  $B_4 > B_3$ ，你得出什么结论？

答案：1. 基准类是本科学历的女教师。

2.  $B_0$  表示刚参加工作的本科学历女教师的收入，所以  $B_0$  的符号为正。

$B_1$  表示在其他条件不变时，工龄变化一个单位所引起的收入的变化，所以  $B_1$  的符号为正。

$B_2$  表示男教师与女教师的工资差异，所以  $B_2$  的符号为正。

$B_3$  表示硕士学历与本科学学历对工资收入的影响，所以  $B_3$  的符号为正。

$B_4$  表示博士学历与本科学学历对工资收入的影响，所以  $B_4$  的符号为正。

3. 若  $B_4 > B_3$ ，说明博士学历的大学教师比硕士学历的大学教师收入要高。

五、若在模型： $Y_t = B_1 + B_2X_t + u_t$  中存在下列形式的异方差： $Var(u_t) = \sigma^2 X_t^3$ ，你如何估计参数  $B_1, B_2$ （10 分）

答案：使用加权最小二乘法估计模型中的参数  $B_1, B_2$ 。

在模型  $Y_t = B_1 + B_2 X_t + u_t$  的两边同时除以  $\sqrt{X_t^3}$ ，我们有：

$$\frac{Y_t}{\sqrt{X_t^3}} = B_1 \frac{1}{\sqrt{X_t^3}} + B_2 \frac{1}{\sqrt{X_t}} + \frac{u_t}{\sqrt{X_t^3}}$$

令  $Y_t^* = \frac{Y_t}{\sqrt{X_t^3}}$ ， $v_t = \frac{u_t}{\sqrt{X_t^3}}$ ，则上面的模型可以表示为：

$$Y_t^* = B_1 \frac{1}{\sqrt{X_t^3}} + B_2 \frac{1}{\sqrt{X_t}} + v_t \quad (1)$$

$$Var(v_t) = Var\left(\frac{u_t}{\sqrt{X_t^3}}\right) = \frac{Var(u_t)}{X_t^3} = \frac{\sigma^2 X_t^3}{X_t^3} = \sigma^2$$

，即变换后的模型（1）的随机误差项满足同方差假定，可以使用 OLS 估计出  $B_1$ ， $B_2$ 。上述方法称为加权最小二乘法。

六、简述自相关后果。对于线性回归模型  $Y_t = B_1 + B_2 X_{1t} + B_3 X_{2t} + u_t$ ，如果存在  $u_t = \rho u_{t-1} + v_t$  形式的自相关，应该采取哪些补救措施？（15 分）

答案：自相关就是指回归模型中随机误差项之间存在相关。用符号表示：

$$Cov(u_i, u_j) = Eu_i u_j = 0 \quad i \neq j$$

对于线性回归模型  $Y_t = B_1 + B_2 X_{1t} + B_3 X_{2t} + u_t$ ，若在模型中存在  $u_t = \rho u_{t-1} + v_t$  形式的自相关问题，我们使用广义差分变换，使得变换后的模型不存在自相关问题。

对于模型： $Y_t = B_1 + B_2 X_{1t} + B_3 X_{2t} + u_t$   
(1)

取模型的一阶滞后： $Y_{t-1} = B_1 + B_2 X_{1t-1} + B_3 X_{2t-1} + u_{t-1}$   
(2)

在（2）式的两边同时乘以相关系数  $\rho$ ，则有：

$$\rho Y_{t-1} = \rho B_1 + \rho B_2 X_{1t-1} + \rho B_3 X_{2t-1} + \rho u_{t-1} \quad (3)$$

用（1）式减（3）式并整理得：

$$Y_t - \rho Y_{t-1} = B_1(1 - \rho) + B_2(X_{1t} - \rho X_{1t-1}) + B_3(X_{2t} - \rho X_{2t-1}) + u_t - \rho u_{t-1}$$

令  $Y_t^* = Y_t - \rho Y_{t-1}$ ， $B_1^* = B_1(1 - \rho)$ ， $X_{1t}^* = X_{1t} - \rho X_{1t-1}$ ， $X_{2t}^* = X_{2t} - \rho X_{2t-1}$   
则有：

$$Y_t^* = B_1^* + B_2 X_{1t}^* + B_3 X_{2t}^* + v_t$$

(4)

在 (4) 中  $v_t$  满足古典假定, 我们可以使用普通最小二乘法估计 (4) 式, 得到  $B_1^*$ ,  $B_2$ ,  $B_3$  的估计量, 再利用  $B_1$  和  $B_1^*$  的对应关系得到  $B_1$  的估计值。

### 七、应用题 (共 30 分)

利用美国 1980-1995 年间人均消费支出 (PCE) 和人均可支配收入 (PDPI) 的数据, 得到了如下回归分析结果:

Dependent Variable: LOG(PCE)

Method: Least Squares

Date: 06/09/05 Time: 23:43

Sample: 1980 1995

Included observations: 16

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LOG(PDPI)	1.205281	0.028891	41.71870	0.0000
C	-2.092664	0.281286	-7.439640	0.0000
R-squared	0.992020	Mean dependent var	9.641839	
Adjusted R-squared	0.991450	S.D. dependent var	0.096436	
S.E. of regression	0.008917	Akaike info criterion	6.485274	-
Sum squared resid	0.001113	Schwarz criterion	6.388701	-
Log likelihood	53.88219	F-statistic	1740.450	
Durbin-Watson stat	2.322736	Prob(F-statistic)	0.000000	

(1) 根据以上结果, 写出回归分析结果报告。(10 分)

答:  $\log(\hat{PCE}) = -2.09 + 1.21\log(PDPI)$

se=(0.28) (0.029)

t=(-7.44) (41.72)

$$p=(0.0000) \quad (0.0000) \quad R^2=0.992$$

(2) 如何解释解释变量的系数和综合判定系数？（10 分）

答：由于该模型是双对数模型，因此，解释变量的系数为因变量对自变量的弹性，在本例中为消费收入弹性，表示收入每增加 1%，消费将平均增加 1.2%。

(3) 对模型中解释变量系数  $B_2$  进行显著性检验。（10 分）

答： 1、  $H_0: B_2=0$        $H_1: B_2 \neq 0$

$$T = \frac{b_2 - B_2}{se(b_2)} \sim t(14)$$

2、构造统计量

$$T = \frac{b_2 - B_2}{se(b_2)} = \frac{1.21 - 0}{0.029} = 41.72$$

3、计算相应的 T 值，

4、查显著性水平为  $\alpha$  的临界值  $t_{\alpha/2}(14) = 2.145$

$$T = \frac{b_2 - B_2}{se(b_2)} = \frac{1.21 - 0}{0.029} = 41.72 > t_{\alpha/2}(15) = 2.131$$

由于

所以拒绝原假设，认为解释变量系数  $B_2$  是统计显著的。



## 复习资料

### 一、判断正误（10 分）

1. 随机变量的条件均值与非条件均值是一回事。（错）
2. 线性回归模型意味着变量是线性的。（错）
3.  $ESS = TSS + RSS$ 。（错）
4. 对于多元回归模型，如果联合检验结果是统计显著的则意味着模型中任何一个单独的变量均是统计显著的。（错）
5. 双对数模型中的斜率表示因变量对自变量的弹性。（对）
6. 为了避免陷入虚拟变量陷阱，如果一个定性变量有  $m$  类，则要引入  $m$  个虚拟变量。（错）
7. 如果回归模型违背了同方差假定，最小二乘估计量是有偏无效的。（错）
8. 在存在接近多重共线性的情况下，回归系数的标准差会趋于变小，相应的  $t$  值会趋于变大。（错）
9. 在任何情况下 OLS 估计量都是待估参数的最优线性无偏估计。（错）
10. 一个联立方程模型中的外生变量在另一个联立方程模型中可能是内生变量。（对）
11. 间接最小二乘法适用于过度识别方程。（错）
12. 假设模型存在一阶自相关，其他条件都满足，则仍用 OLS 法估计参数，得到的估计量仍是无偏的，不再是有效的，显著性检验失效，预测失效。（对）
13. 用一阶差分法消除自相关时，我们假定自相关系数等于-1。（错）
14. 当异方差出现时，最小二乘估计是有偏的和不具有最小方差特性。（错）
15. 在模型  $Y_t = B_0 + B_1X_t + B_2D_t + u_t$  中，令虚拟变量  $D$  取值为  $(0, 2)$  而不是  $(0, 1)$ ，那么参数  $B_2$  的估计值也将减半， $t$  值也将减半。（错）

### 二、选择题（20 分）

- 1、下列不属于线性回归模型经典假设的条件是（ A ）
  - A. 被解释变量确定性变量，不是随机变量。
  - B. 随机扰动项服从均值为 0，方差恒定，且协方差为 0。
  - C. 随机扰动项服从正态分布。
  - D. 解释变量之间不存在多重共线性。

2、参数  $\beta$  的估计量  $\hat{\beta}$  具备有效性是指 ( B )

- A.  $Var(\hat{\beta})=0$     B.  $Var(\hat{\beta})$  为最小    C.  $E(\hat{\beta}-\beta)=0$     D.  $E(\hat{\beta}-\beta)$  为最小

3、利用 OLS 估计模型  $Y_i = \alpha_0 + \alpha_1 X_i + \mu_i$  求得的样本回归线，下列哪些结论是不正确的 ( D )

- A. 样本回归线通过  $(\bar{X}, \bar{Y})$  点    B.  $\sum \hat{\mu}_i = 0$   
C.  $\bar{Y} = \bar{\hat{Y}}$     D.  $Y_i = \hat{\alpha}_0 + \hat{\alpha}_1 X_i$

4、如果联立方程模型中某结构方程包含了模型系统中所有的变量，则这个方程是 ( B )

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

5、对于某样本回归模型，已求得 DW 统计量的值为 1，则模型残差的自相关系数  $\rho$  近似等于 ( B )

- A. 0    B. 0.5    C. -0.5    D. 1    注：dw 近似等于  $2(1-\rho)$

6、一元线性回归中，相关系数  $r=( C )$

- A.  $\frac{(\sum (X_i - \bar{X})(Y_i - \bar{Y}))^2}{\sum (X_i - \bar{X})^2 \sum (Y_i - \bar{Y})^2}$     B.  $\frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sum (X_i - \bar{X})^2 \sum (Y_i - \bar{Y})^2}$   
C.  $\frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum (X_i - \bar{X})^2 \sum (Y_i - \bar{Y})^2}}$     D.  $\frac{\sum (Y_i - \bar{Y})^2}{\sqrt{\sum (X_i - \bar{X})^2 \sum (Y_i - \bar{Y})^2}}$

7、对联立方程模型进行参数估计的方法可以分两类，即：( B )

- A. 间接最小二乘法和系统估计法    B. 单方程估计法和系统估计法  
C. 单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

8、结构式模型中的每一个方程都称为结构式方程，在结构方程中，解释变量可以是滞后变量，也可以是 ( D )

- A. 外生变量    B. 滞后变量    C. 内生变量    D. 外生变量和内生变量

9、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.  $0 \leq DW \leq 4$

10、下面说法正确的是( D )

A. 内生变量是非随机变量

B. 滞后变量是随机变量

C. 外生变量是随机变量

D. 外生变量是非随机变量

11、在同一时间不同统计单位的相同统计指标组成的数据组合，是( D )。

A. 原始数据

B. 时点数据

C. 时间序列数据

D. 截面数据

12、单一方程计量经济模型必然包括( A )。

A. 行为方程

B. 技术方程

C. 制度方程

D. 定义方程

13、在同一时间不同统计单位的相同统计指标组成的数据组合，是( D )。

A. 原始数据

B. 时点数据

C. 时间序列数据

D. 截面数据

14、计量经济模型的被解释变量一定是( C )。

A. 控制变量

B. 政策变量

C. 内生变量

D. 外生变量

15、同一统计指标按时间顺序记录的数据称为( B )。

A. 横截面数据

B. 时间序列数据

C. 修匀数据

D. 原始数据

16、半对数模型  $Y = \beta_0 + \beta_1 \ln X + \mu$  中，参数  $\beta_1$  的含义是( C )。

A. X 的绝对量变化，引起 Y 的绝对量变化

B. Y 关于 X 的边际变化

C. X 的相对变化，引起 Y 的期望值绝对量变化

D. Y 关于 X 的弹性

17、计量经济学是下列哪门学科的分支学科( C )。

A. 统计学

B. 数学

C. 经济学

D. 数理统计学

18、计量经济学成为一门独立学科的标志是( B )。

A. 1930 年世界计量经济学会成立 B. 1933 年《计量经济学》会刊出版

C. 1969 年诺贝尔经济学奖设立

D. 1926 年计量经济学 (Economics) 一词构造出来

19、外生变量和滞后变量统称为( D )。

A. 控制变量

B. 解释变量

C. 被解释变量

D. 先决变量

20、一元线性回归分析中的回归平方和 TSS 的自由度是( B )。

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

21、在总体回归直线  $E(\hat{Y}) = \beta_0 + \beta_1 X$  中,  $\beta_1$  表示 ( B )。

- A. 当  $X$  增加一个单位时,  $Y$  增加  $\beta_1$  个单位  
B. 当  $X$  增加一个单位时,  $Y$  平均增加  $\beta_1$  个单位  
C. 当  $Y$  增加一个单位时,  $X$  增加  $\beta_1$  个单位  
D. 当  $Y$  增加一个单位时,  $X$  平均增加  $\beta_1$  个单位

22、调整的判定系数  $\bar{R}^2$  与多重判定系数  $R^2$  之间有如下关系 ( D )

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

23、假设回归模型为  $y_i = \alpha + \beta x_i + \mu_i$ , 其中  $x_i$  为随机变量,  $x_i$  与  $\mu_i$  相关则  $\beta$  的普通最小二乘估计量 ( D )

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

24、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  $\mu_i$  的方差估计量  $\hat{\sigma}^2$  为 ( D )

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

### 三、名词解释

#### 1. 拟合优度

答案: 样本回归直线与样本观测值之间的拟合程度, 用判定系数表示, 目的是了解解释变量对被解释变量的解释程度。

#### 2. 虚拟变量

答案: 用 0, 1 表示质的因素对回归模型的影响, 主要是事物品质或属性的量化, 反映在截距和斜率变化上。

#### 3. 序列相关

答案：线性回归模型中各个误差项之间的协方差不为零，又叫自相关

#### 4、多重共线性

答案：线性回归模型中的若干或全部解释变量的样本观测值之间具有线性关系

#### 5、工具变量法

答案：找一个变量，该变量与模型中的随机解释变量高度相关，与随机误差项无关，此变量和模型其他变量构造出模型相应回归系数的一致估计量，这种方法叫工具变量法

1. 计量经济学模型：揭示经济现象中客观存在的因果关系，主要采用回归分析方法的经济数学模型。

2. 参数估计的无偏性：它的均值或期望值是否等于总体的真实值。

3. 参数估计量的有效性：它是否在所有线性无偏估计量中具有最小方差。估计量的期望方差越大说明用其估计值代表相应真值的有效性越差；否则越好，越有效。不同的估计量具有不同的方差，方差最小说明最有效。

6. 内生变量：具有某种概率分布的随机变量，它的参数是联立方程系统估计的元素，内生变量是由模型系统决定的，同时也对模型系统产生影响。内生变量一般都是经济变量。

7. 异方差：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，则认为出现了异方差性。

8. 回归分析：研究一个变量关于另一个（些）变量的依赖关系的计算方法和理论。其目的在于通过后者的已知或设定值，去估计和预测前者的（总体）均值。前一变量称为被解释变量或应变量，后一变量称为解释变量或自变量。

### 三、简答题（每小题 6 分，共 24 分）

#### 1、何谓虚拟变量？

答：(1) 在建立模型时，有一些影响经济变量的因素无法定量描述，如职业、性别对收入的影响，教育程度，季节因素等往往需要用定性变量度量。为了在模型中反映这类因素的影响，并提高模型的精度，需要将这类变量“量化”。根据这类变量的属性类型，构造仅取“0”或“1”的人工变量，通常称这类变量为“虚拟变量”。

#### 2、对计量经济模型的检验应从几个方面入手？

对计量经济模型的检验应从几个方面入手？

答：①经济意义检验；②统计准则检验；③计量经济学准则检验；④模型预测检验。

#### 3、在计量经济模型中，为什么会存在随机误差项？

答：随机误差项是计量经济模型中不可缺少的一部分。产生随机误差项的原因有以下几个方面：①模型中被忽略掉的影响因素造成的误差；②模型关系认定不准确造成的误差；③变量的测量误差；④随机因素。

4、古典线性回归模型的基本假定是什么？

答：①零均值假定。即在给定  $x_t$  的条件下，随机误差项的数学期望（均值）为 0，即  $E(u_t)=0$ 。②同方差假定。误差项  $u_t$  的方差与  $t$  无关，为一个常数。③无自相关假定。即不同的误差项相互独立。④解释变量与随机误差项不相关假定。⑤正态性假定，即假定误差项  $u_t$  服从均值为 0，方差为  $\sigma^2$  的正态分布。

5、在满足古典假定条件下，一元线性回归模型的普通最小二乘估计量有哪些统计性质？

答：①线性，是指参数估计量  $\hat{b}_0$  和  $\hat{b}_1$  分别为观测值  $y_t$  和随机误差项  $u_t$  的线性函数或线性组合。②无偏性，指参数估计量  $\hat{b}_0$  和  $\hat{b}_1$  的均值（期望值）分别等于总体参数  $b_0$  和  $b_1$ 。③有效性（最小方差性或最优性），指在所有的线性无偏估计量中，最小二乘估计量  $\hat{b}_0$  和  $\hat{b}_1$  的方差最小。

6、序列相关性的后果。

答：（1）模型参数估计值不具有最优性；（2）随机误差项的方差一般会低估；（3）模型的统计检验失效；（4）区间估计和预测区间的精度降低。

7、在建立计量经济学模型时，什么时候，为什么要引入虚拟变量？

答案：在现实生活中，影响经济问题的因素除具有数量特征的变量外，还有一类变量，这类变量所反映的并不是数量而是现象的某些属性或特征，即它们反映的是现象的质的特征。这些因素还很可能是重要的影响因素，这时就需要在模型中引入这类变量。引入的方式就是以虚拟变量的形式引入。

8、工具变量选择必须满足的条件是什么？

答案：选择工具变量必须满足以下两个条件：（1）工具变量与模型中的随机解释变量高度相关；（2）工具变量与模型的随机误差项不相关。

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。

6、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。

7、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。

8、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。

9、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。

$F = t^2$

10、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

11、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。

12、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。

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. 663
1992	42. 9599	317. 3194
1993	67. 2507	449. 2889
1994	74. 3992	615. 1933
1995	88. 0174	795. 695
1996	131. 749	950. 0446
1997	144. 7709	1130. 013
1998	164. 9067	1289. 019
1999	184. 7908	1436. 027
2000	225. 0212	1665. 465



2001	265.6532	1954.654
------	----------	----------

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：（1）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：

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

（2）即

$$Y_t = -3.611151 + 0.134582 GDP_t$$

(4.161790)      (0.003867)

(t) (-0.867692)      (34.80013)

**R<sup>2</sup> = 0.99181      F = 1211.049**

R<sup>2</sup>=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。

(3) 从斜率的t检验值看，在5%显著性水平下自由度为12-2=10的临界值  $t_{0.025}(10) = 2.228$

，有  $|t| = 34.8 > t_{0.025}(10) = 2.228$ ，拒绝零假设，接受  $\beta_2$  显著不为零，且该值介于0和1之间，符合经济理论。

(4) 当2005年GDP为3600亿元时，地方财政收入的点预测值为：

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

区间预测：

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

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

取  $\alpha = 0.05$ ，平均值置信度95%的预测区间为：

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

$$GDP_{2005} = 3600 \text{ 时, } 480.884 \mp 2.228 * 7.5325 * \sqrt{\frac{1}{12} + \frac{7195337.357}{3293728.494}} = 480.884 \mp 25.2735$$

(亿元)

$Y_f$  个别值置信度95%的预测区间为：

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

$$\text{即, } 480.884 \mp 2.228 * 7.5325 * \sqrt{1 + \frac{1}{12} + \frac{7195337.357}{3293728.494}} = 480.884 \mp 30.3381 \text{ (亿元)}$$