

# 清华大学经济管理学院计量经济学期末试题及答案

(2002 年)

(2 小时, 开卷, 满分 100 分)

## 1. (共 40 分, 每小题 4 分) 建立中国居民消费函数模型

$$C_t = \alpha_0 + \alpha_1 I_t + \alpha_2 C_{t-1} + \varepsilon_t \quad \varepsilon_t \sim N(0, \sigma^2) \quad t=1978, 1979, \dots, 2001$$

其中  $C$  表示居民消费总额,  $I$  表示居民收入总额。

- (1) 能否用历年的人均消费额和人均收入数据为样本观测值估计模型? 为什么?
- (2) 人们一般选择用当年价格统计的居民消费总额和居民收入总额作为样本观测值, 为什么? 这样是否违反样本数据可比性原则? 为什么?
- (3) 如果用矩阵方程  $\mathbf{Y} = \mathbf{XB} + \mathbf{E}$  表示该模型, 写出每个矩阵的具体内容, 并标明阶数;
- (4) 如果所有古典假设都满足, 分别从最小二乘原理和矩方法出发, 推导出关于参数估计量的正规方程组;
- (5) 如果  $C_{t-1}$  与  $I_t$  存在共线性, 证明: 当去掉变量  $C_{t-1}$  以消除共线性时,  $\alpha_1$  的估计结果将发生变化;
- (6) 如果模型中  $C_{t-1}$  为随机解释变量且与  $\varepsilon_t$  相关, 证明: 如果用 OLS 估计该消费函数模型, 其参数估计量是有偏的;
- (7) 如果模型中  $C_{t-1}$  为随机解释变量且与  $\varepsilon_t$  相关, 选择政府消费  $G_t$  为  $C_{t-1}$  的工具变量 ( $G_t$  满足工具变量的所有条件), 写出关于参数估计量的正规方程组;
- (8) 如果经检验表明模型存在一阶序列相关, 而需要采用广义差分法估计模型, 指出在常用的软件中是如何实现的?
- (9) 在不受限制的情况下,  $C_t$  的值域为  $(0, \infty)$ , 写出  $C_t$  的对数似然函数;
- (10) 试分析, 以  $t=1978, 1979, \dots, 2001$  数据为样本观测值, 能否说“样本是从母体中随机抽取的”? 那么采用 OLS 估计模型参数, 估计结果是否存在偏误? 为什么?

答:

(1) 不可以。因为历年的人均消费额和人均收入并不是从居民消费总额和居民收入总额的总体中随机抽取的样本, 违背了样本与母体的一致性。

(2) 因为历年的居民消费总额和居民收入总额具有大致相同的“价格”指数, 是否将它们转换为不变价数据并不重要, 不影响数据在样本点之间的可比性。

(3)  $\mathbf{Y} = \mathbf{XB} + \mathbf{E}$  其中

$$\mathbf{Y} = \begin{matrix} C \\ 1978 \\ C_{1979} \\ M \\ C \\ 2001 \end{matrix} \quad \mathbf{X} = \begin{matrix} 1 & I & C \\ 1 & I_{1978} & C_{1978} \\ M & I & C \\ 1 & I & C \\ 2001 & I & C \end{matrix} \quad \mathbf{B} = \begin{matrix} \alpha_0 \\ \alpha_1 \\ \alpha_{2 \cdot 1} \end{matrix} \quad \mathbf{E} = \begin{matrix} \varepsilon \\ \varepsilon_{1978} \\ M \\ \varepsilon \\ 2001 \end{matrix}$$

(4) 从最小二乘原理出发, 推导出关于参数估计量的正规方程组:

$$Q = \sum_{i=1}^n e_i^2 = e'e = (Y - X\hat{\beta})'(Y - X\hat{\beta})$$

$$\frac{\partial}{\partial \beta} (Y'Y - \hat{\beta}'X'Y - Y'X\hat{\beta} + \hat{\beta}'X'X\hat{\beta}) = 0$$

$$\frac{\partial}{\partial \beta} (Y'Y - \hat{\beta}'X'Y - Y'X\hat{\beta} + \hat{\beta}'X'X\hat{\beta}) = 0$$

$$\frac{\partial}{\partial \beta} (Y'Y - 2Y'X\hat{\beta} + \hat{\beta}'X'X\hat{\beta}) = 0$$

$$-X'Y + X'X\hat{\beta} = 0$$

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

从矩方法出发，推导关于参数估计量的正规方程组：

$$X'Y = X'X\beta + X'\mu$$

$$X'(Y - X\beta) = X'\mu$$

$$E(X'(Y - X\beta)) = 0$$

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

(5) 从矩方法出发推导关于参数估计量的正规方程组的第一步可以写成：

$$\begin{aligned} \sum_t C_t &= \sum_t (\alpha_0 + \alpha_1 I_t + \alpha_2 C_{t-1} + \varepsilon_t) \\ \sum_t I_t C_t &= \sum_t I_t (\alpha_0 + \alpha_1 I_t + \alpha_2 C_{t-1} + \varepsilon_t) \\ \sum_t C_{t-1} C_t &= \sum_t C_{t-1} (\alpha_0 + \alpha_1 I_t + \alpha_2 C_{t-1} + \varepsilon_t) \end{aligned}$$

导出的方程组为：

$$\begin{aligned} \sum_t C_t &= \sum_t (\hat{\alpha}_0 + \hat{\alpha}_1 I_t + \hat{\alpha}_2 C_{t-1}) \\ \sum_t I_t C_t &= \sum_t I_t (\hat{\alpha}_0 + \hat{\alpha}_1 I_t + \hat{\alpha}_2 C_{t-1}) \\ \sum_t C_{t-1} C_t &= \sum_t C_{t-1} (\hat{\alpha}_0 + \hat{\alpha}_1 I_t + \hat{\alpha}_2 C_{t-1}) \end{aligned}$$

当去掉变量  $C_{t-1}$ ，构成一个一元模型，其关于参数估计量的正规方程组为

$$\begin{aligned} \sum_t C_t &= \sum_t (\alpha_0 + \alpha_1 I_t + \varepsilon_t) \\ \sum_t I_t C_t &= \sum_t I_t (\alpha_0 + \alpha_1 I_t + \varepsilon_t) \end{aligned}$$

由于  $C_{t-1}$  与  $I_t$  存在共线性，上式第 2 个方程中缺少的  $C_{t-1}$  与  $I_t$  乘积项不为 0，所以去掉该项会影响方程组的解，使得  $\alpha_1$  的估计结果将发生变化。

(6) 如果模型中  $C_{t-1}$  为随机解释变量且与  $\varepsilon_t$  相关，上述方程组中的第 3 个方程非齐次。而用 OLS 估计该消费函数模型，认为正规方程组是齐次方程组，所以其参数估计量是有偏的。

(7) 选择政府消费  $G_t$  为  $C_{t-1}$  的工具变量，得到关于参数估计量的正规方程组为：

$$\begin{aligned}\sum_t C_t &= \sum_t (\alpha_0 + \alpha_1 I_t + \alpha_2 C_{t-1}) \\ \sum_t I_t C_t &= \sum_t I_t (\alpha_0 + \alpha_1 I_t + \alpha_2 C_{t-1}) \\ \sum_t G_t C_t &= \sum_t G_t (\alpha_0 + \alpha_1 I_t + \alpha_2 C_{t-1})\end{aligned}$$

(8) 在解释变量中增加  $AR(1)$ 。

(9)  $C_t$  的对数似然函数为

$$\begin{aligned}L^* &= Ln(L) \\ &= -nLn(\sqrt{2\pi}\sigma) - \frac{1}{2\sigma^2} (\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}})' (\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}})\end{aligned}$$

(10) 严格地，不能说“样本是从母体中随机抽取的”，因为  $C_t$  的值域为  $(0, \infty)$ ，而实际的样本观测值集中于某一区域。那么采用 OLS 估计模型参数，估计结果是存在偏误的，因为样本实际上是选择性的。

## 2. (共 20 分，每小题 5 分) 下列为一完备的联立方程计量经济模型

$$\begin{aligned}C_t &= \alpha_0 + \alpha_1 Y_t + \alpha_2 C_{t-1} + \mu_{1t} \\ I_t &= \beta_0 + \beta_1 Y_t + \mu_{2t} \\ Y_t &= C_t + I_t + G_t\end{aligned}$$

其中  $C$  为居民消费总额、 $I$  为投资总额、 $Y$  为国内生产总值、 $G_t$  为政府消费总额，样本取自 1978—2000 年。

(1) 说明：对于消费方程，用 IV、ILS、2SLS 方法分别估计，参数估计结果是等价的。

(2) 说明：对于投资方程，能否用 IV、ILS 方法估计？为什么？

(3) 对于该联立方程计量经济模型，如果采用 2SLS 估计指出其优缺点。

(4) 如果该模型的每个结构方程的随机项具有同方差性和序列不相关性，而不同结构方程的随机项之间具有同期相关性。写出它们的方差协方差矩阵。

答：

(1) 因为消费方程是恰好识别的结构方程，用 IV、ILS、2SLS 方法分别估计，都可以看成为工具变量方法，而且都以所有先决变量的结合为工具变量，所以参数估计结果是等价的。

(2) 投资方程是过度识别的结构方程，所以不能用 IV、ILS 方法估计。如果用 IV、ILS 方法估计，会得到多组不同的参数估计结果。

(3) 2SLS 估计的优点是：既适用于恰好识别的消费方程，又适用于过度识别的投资方程；由于第一阶段采用所有先决变量作为解释变量，所以在分别估计消费方程和投资方程时，都利用了所有先决变量的信息；克服了每个方程中内生解释变量  $Y_t$  与随机项相关的问题。缺点是没有利用方程之间相关性信息，对于该模型系统，消费方程和投资方程的随机项显然是相关的。

(4)

$$Cov(\mu_1, \mu_2) = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{pmatrix}$$

3. (共 30 分，每小题 5 分) 简单回答以下问题：

(1) 分别指出两要素 C-D 生产函数、两要素一级 CES 生产函数和 VES 生产函数关于要素替代弹性的假设。

(2) 在一篇博士论文中设计的生产函数模型为：

$$Y_t = \delta_0 K_t^{-\alpha\rho} L_t^{-(1-\alpha)\rho} + \sum_{i=1}^k \delta_i G_{it}^{-\rho} \quad -\frac{1}{\rho}$$

其中，Y 为产出量，K、L 为资本和劳动投入量， $G_i$  为第 i 种能源投入量，其它为参数。试

指出该理论模型设计的主要问题，并给出正确的模型设计。

(3) 建立城镇居民食品类需求函数模型如下：

$$\ln(V) = \beta_0 + \beta_1 \ln(Y) + \beta_2 \ln(P_1) + \beta_3 \ln(P_2) + \mu$$

其中 V 为人均购买食品支出额、Y 为人均收入、 $P_1$  为食品类价格、 $P_2$  为其它商品类价格。

拟定每个参数的数值范围，并指出参数之间必须满足的关系。

(4) 指出在实际建立模型时虚变量的主要用途。

(5) 两位研究者分别建立如下的中国居民消费函数模型

$$C_t = \alpha_0 + \alpha_1 I_t + \alpha_2 C_{t-1} + \varepsilon_t \quad \varepsilon_t \sim N(0, \sigma^2)$$

$$\text{和} \quad C_t = \alpha_0 + \alpha_1 I_t + \varepsilon_t \quad \varepsilon_t \sim N(0, \sigma^2)$$

其中 C 表示居民消费总额，I 表示居民收入总额。由相同的样本和相同的估计方法，得到了不同的居民边际消费倾向估计值。如何解释这种现象？由此指出经典计量经济学模型的缺点。

(6) 从经典计量经济学模型设定理论出发，在建立中国宏观计量经济模型时，一般应该如何对第三产业的生产方程进行分解，并指出其理由。

答：

(1) C-D 生产函数的要素替代弹性始终为 1，不随着研究对象、样本区间而变化，当然也不随着样本点而变化；两要素一级 CES 生产函数模型对要素替代弹性的假设为：随着研究对象、样本区间而变化，但是不随着样本点而变化；VES 生产函数的要素替代弹性除了随着研究对象、样本区间而变化外，还随着样本点而变化。

(2) 在该模型中, 将  $K$  和  $L$  首先组合成为一个组合要素:

$$y_{klt} = K_t^\alpha L_t^{1-\alpha}$$

然后, 将该组合要素  $y_{kl}$  与每种能源投入量  $G_i$  一起, 建立多要素一级 CES 生产函数。那么,

其假设是  $y_{kl}$  与  $G_i$  以及  $G_i$  之间具有相同的替代弹性, 这显然是错误的。各种能源之间, 例

如煤炭和石油具有很强的替代性, 而每种能源与  $y_{kl}$  之间的替代性显然要差得多。

应该采用多级 CES 生产函数。例如第一级包含两个函数:

$$y_{klt} = f(K_t, L_t) \quad y_{Gt} = g(G_1, G_2, L)$$

第二级为:

$$Y_t = A(\delta_1 y_{klt}^{-\rho} + \delta_2 y_{Gt}^{-\rho})^{-\frac{\rho}{1-\rho}}$$

(3) 参数  $\beta_1$ 、 $\beta_2$ 、 $\beta_3$  估计量的经济意义分别为人均收入、食品类价格、其它商品类价格

的需求弹性; 由于食品为必需品,  $V$  为人均购买食品支出额, 所以  $\beta_1$  应该在 0 与 1 之间,  $\beta_2$

应该在 0 与 1 之间,  $\beta_3$  在 0 左右, 三者之和为 1 左右。

(4) 在实际建立模型时虚变量主要用于表示定性变量, 例如政策变量、条件变量等。例如建立我国粮食生产模型, 联产承包制度的实施对粮食产量影响很大, 可以作为一个虚变量引入模型, 实行该制度的年份取值为 1, 其它年份取值为 0。

(5) 由于两位研究者依据不同的消费理论, 建立了不同的消费模型。前者依据相对收入假设, 后者依据绝对收入假设。同时, 由于  $I_t$  和  $C_{t-1}$  之间存在一定程度的线性关系, 所以两

个模型得到了不同的居民边际消费倾向  $\alpha_1$  的估计值。这反映了经典计量经济学模型的理论导向所存在的任意性, 不同的人对行为理论理解不同, 就可能建立不同的模型。

(6) 由于第三产业内部包括许多部门, 不同的部门差异很大, 很难发现能够对所有主要部门的产出水平起解释作用的共同变量; 另外, 由于第三产业内部部门之间的结构处于变化之中, 所有建立单一模型缺少结构性功能。所以, 在建立中国宏观计量经济模型时, 一般应该如何对第三产业的生产方程进行分解。

**计量经济学期末试题**  
(2003 年 6 月, 满分 70 分)

1. (12 分) 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, 经过理论和经验分析, 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。

2. (12 分) 以  $Q_t$  表示粮食产量,  $A_t$  表示播种面积,  $C_t$  表示化肥施用量, 经检验, 它们取对数后都是  $I(1)$  变量且互相之间存在  $CI(1,1)$  关系。同时经过检验并剔除不显著的变量 (包括滞后变量), 得到如下粮食生产模型:

$$\ln Q_t = \alpha_0 + \alpha_1 \ln Q_{t-1} + \alpha_2 \ln A_t + \alpha_3 \ln C_t + \alpha_4 \ln C_{t-1} + \mu_t \quad (1)$$

- (1) 写出长期均衡方程的理论形式;
- (2) 写出误差修正项  $ecm$  的理论形式;
- (3) 写出误差修正模型的理论形式;
- (4) 指出误差修正模型中每个待估参数的经济意义。

3. (6 分) 对于上述粮食生产模型(1), 假设所有解释变量与随机误差项都不相关。

- (1) 如果采用普通最小二乘法估计, 用非矩阵形式写出关于参数估计量的正规方程组;
- (2) 从以上正规方程组出发说明, 为什么不能采用分部回归方法分别估计每个参数;

4. (9 分) 投资函数模型

$$I_t = \beta_0 + \beta_1 Y_t + \beta_2 Y_{t-1} + \mu_t$$

为一完备的联立方程计量经济模型中的一个方程, 模型系统包含的内生变量为  $C$  (居民消费总额)、 $I$  (投资总额) 和  $Y$  (国内生产总值), 先决变量为  $G_t$  (政府消费)、 $C_{t-1}$  和  $Y_{t-1}$ 。样本容量为  $n$ 。

- (1) 可否用狭义的工具变量法估计该方程? 为什么?
- (2) 如果采用 2SLS 估计该方程, 分别写出 2SLS 估计量和将它作为一种工具变量方法的估计量的矩阵表达式;
- (3) 如果采用 GMM 方法估计该投资函数模型, 写出一组等于 0 的矩条件。

5. (6 分) 建立城镇居民食品类需求函数模型如下:

$$\ln(V) = 1350. + 0.923 \ln(Y) - 0115. \ln(P_1) + 0.357 \ln(P_2)$$

其中  $V$  为人均购买食品支出额、 $Y$  为人均收入、 $P_1$  为食品类价格、 $P_2$  为其它商品类价格。

- (1) 指出参数估计量的经济意义是否合理，为什么？
- (2) 为什么经常采用交叉估计方法估计需求函数模型？

6. (9 分) 选择两要素一级 CES 生产函数的近似形式建立中国电力行业的生产函数模型：

$$\ln Y = \ln A + \gamma t + m \delta \ln K + m(1 - \delta) \ln L - \frac{1}{2} m \rho \delta (1 - \delta) (\ln K_L)^2 + \mu$$

其中 Y 为发电量，K、L 分别为投入的资本与劳动数量，t 为时间变量。

- (1) 指出参数  $\gamma$ 、 $\rho$ 、 $m$  的经济含义和数值范围；
- (2) 指出模型对要素替代弹性的假设，并指出它与 C-D 生产函数、VES 生产函数在要素替代弹性假设上的区别；
- (3) 指出模型对技术进步的假设，并指出它与下列生产函数模型

$$\ln Y = \ln A + \gamma t + \alpha \ln K + \beta \ln L + \mu$$

在技术进步假设上的区别；

7. (8 分) 试指出在目前建立中国宏观计量经济模型时，下列内生变量应由哪些变量来解释，简单说明理由，并拟定关于每个解释变量的待估参数的正负号。

- |            |               |
|------------|---------------|
| (1) 轻工业增加值 | (2) 衣着类商品价格指数 |
| (3) 货币发行量  | (4) 农业生产资料进口额 |

8. (8 分) 回答：

- (1) 随机时间序列的平稳性条件是什么？证明随机游走序列不是平稳序列。
- (2) 单位根检验为什么从 DF 检验扩展到 ADF 检验？

### 计量经济学期末试题答案

(2003 年 6 月，满分 70 分)

1. (12 分) 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，经过理论和经验分析，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。

答案：（答出 4 条给满分）

- (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，与实际生产活动不符。
- (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。
- (5) 变量间可能不存在长期均衡关系。变量中有流量和存量，可能存在 1 个高阶单整的序列。

应该首先进行单位根检验和协整检验。

2. (12分) 以  $Q_t$  表示粮食产量,  $A_t$  表示播种面积,  $C_t$  表示化肥施用量, 经检验, 它们取

对数后都是  $I(1)$  变量且互相之间存在  $CI(1,1)$  关系。同时经过检验并剔除不显著的变量 (包括滞后变量), 得到如下粮食生产模型:

$$\ln Q_t = \alpha_0 + \alpha_1 \ln Q_{t-1} + \alpha_2 \ln A_t + \alpha_3 \ln C_t + \alpha_4 \ln C_{t-1} + \mu_t \quad (1)$$

- (1) 写出长期均衡方程的理论形式;
- (2) 写出误差修正项  $ecm$  的理论形式;
- (3) 写出误差修正模型的理论形式;
- (4) 指出误差修正模型中每个待估参数的经济意义。

答案:

(1) 长期均衡方程的理论形式为:

$$\ln Q_t = \beta_0 + \beta_1 \ln A_t + \beta_2 \ln C_t + \varepsilon_t$$

(2) 误差修正项  $ecm$  的理论形式为:

$$ecm_t = \ln Q_t - \beta_0 - \beta_1 \ln A_t - \beta_2 \ln C_t$$

(3) 误差修正模型的理论形式为:

$$\ln Q_t = \alpha_2 \ln A_t + \alpha_3 \ln C_t - \lambda ecm_{t-1} + \mu_t$$

(4) 误差修正模型中每个待估参数的经济意义为:

$\alpha_2$ : 播种面积对产量的短期产出弹性;

$\alpha_3$ : 化肥施用量对产量的短期产出弹性;

$\lambda$ : 前个时期对长期均衡的偏离程度对当期短期变化的影响系数。

3. (6分) 对于上述粮食生产模型(1), 假设所有解释变量与随机误差项都不相关。

- (1) 如果采用普通最小二乘法估计, 用非矩阵形式写出关于参数估计量的正规方程组;
- (2) 从以上正规方程组出发说明, 为什么不能采用分部回归方法分别估计每个参数。

答案:

(1) 在所有解释变量与随机误差项都不相关的条件下, 如果采用普通最小二乘法估计, 关于参数估计量的正规方程组为:

$$\sum_t \ln Q_t = \sum_t (\alpha_0 + \alpha_1 \ln Q_{t-1} + \alpha_2 \ln A_t + \alpha_3 \ln C_t + \alpha_4 \ln C_{t-1})$$

$$\sum_t \ln Q_{t-1} \ln Q_t = \sum_t \ln Q_{t-1} (\alpha_0 + \alpha_1 \ln Q_{t-1} + \alpha_2 \ln A_t + \alpha_3 \ln C_t + \alpha_4 \ln C_{t-1})$$

$$\sum_t \ln A_t \ln Q_t = \sum_t \ln A_t (\alpha_0 + \alpha_1 \ln Q_{t-1} + \alpha_2 \ln A_t + \alpha_3 \ln C_t + \alpha_4 \ln C_{t-1})$$

$$\sum_t \ln C_t \ln Q_t = \sum_t \ln C_t (\alpha_0 + \alpha_1 \ln Q_{t-1} + \alpha_2 \ln A_t + \alpha_3 \ln C_t + \alpha_4 \ln C_{t-1})$$



$$\sum_t \ln C_{t-1} \ln Q_t = \sum_t \ln C_{t-1} (\alpha_0 + \alpha_1 \ln Q_{t-1} + \alpha_2 \ln A_t + \alpha_3 \ln C_t + \alpha_4 \ln C_{t-1})$$

(2) 如果采用分部回归方法分别估计每个参数, 例如估计  $\alpha_2$ , 建立一元模型, 其正规方程组为:  $\sum_t \ln A_t \ln Q_t = \sum_t \ln A_t (\alpha_2 \ln A_t)$ , 与上述(1)中第 3 个方程相比较, 则要求方程右边其余各项均为 0。但是, 由于解释变量之间存在一定程度的共线性, 这一要求显然不能满足。所以, 两种情况下的  $\alpha_2$  的估计结果不相同。

#### 4. (9 分) 投资函数模型

$$I_t = \beta_0 + \beta_1 Y_t + \beta_2 Y_{t-1} + \mu_t$$

为一完备的联立方程计量经济模型中的一个方程, 模型系统包含的内生变量为 C (居民消费总额)、I (投资总额) 和 Y (国内生产总值), 先决变量为  $G_t$  (政府消费)、 $C_{t-1}$  和  $Y_{t-1}$ 。样本容量为  $n$ 。

(1) 可否用狭义的工具变量法估计该方程? 为什么?

(2) 如果采用 2SLS 估计该方程, 分别写出 2SLS 估计量和将它作为一种工具变量方法的估计量的矩阵表达式;

(3) 如果采用 GMM 方法估计该投资函数模型, 写出一组等于 0 的矩条件。

答案:

(1) 不能用狭义的工具变量法估计该方程。因为该结构方程是过度识别的。

(2) 如果采用 2SLS 估计该方程, 可以将 2SLS 估计看作为一种工具变量方法。估计量的矩阵表达式分别为:

$$\begin{aligned} \begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix} &= \begin{pmatrix} 1 \\ Y_t \\ Y_{t-1} \end{pmatrix}^{-1} \begin{pmatrix} 1 \\ Y_t \\ Y_{t-1} \end{pmatrix} \begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix} \\ \begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix} &= \begin{pmatrix} 1 \\ Y_t \\ Y_{t-1} \end{pmatrix}^{-1} \begin{pmatrix} 1 \\ Y_t \\ Y_{t-1} \end{pmatrix} \begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix} \end{aligned}$$

前者为 2SLS 估计, 后者为其等价的工具变量估计。

(3) 如果采用 GMM 方法估计该投资函数模型, 用模型系统的所有先决变量作为工具变量。可以写出如下一组等于 0 的矩条件:

$$\begin{aligned} \sum_t I_t &= \sum_t (\hat{\beta}_0 + \hat{\beta}_1 Y_t + \hat{\beta}_2 Y_{t-1}) \\ \sum_t I_t Y_{t-1} &= \sum_t (\hat{\beta}_0 + \hat{\beta}_1 Y_t + \hat{\beta}_2 Y_{t-1}) Y_{t-1} \\ \sum_t I_t G_t &= \sum_t (\hat{\beta}_0 + \hat{\beta}_1 Y_t + \hat{\beta}_2 Y_{t-1}) G_t \end{aligned}$$

$$\sum_{t=1}^T I_t C_{t-1} = \sum_{t=1}^T (\hat{\beta}_0 + \hat{\beta}_1 Y_t + \hat{\beta}_2 Y_t - 1) C_{t-1}$$

5. (6 分) 建立城镇居民食品类需求函数模型如下:

$$\ln(V) = 1.350 + 0.923 \ln(Y) - 0.115 \ln(P_1) + 0.357 \ln(P_2)$$

其中  $V$  为人均购买食品支出额、 $Y$  为人均收入、 $P_1$  为食品类价格、 $P_2$  为其它商品类价格。

- (1) 指出参数估计量的经济意义是否合理, 为什么?
- (2) 为什么经常采用交叉估计方法估计需求函数模型?

答案:

(1) 对于以购买食品支出额位被解释变量的需求函数模型, 即

$$\ln(V) = \alpha_0 + \alpha_1 \ln(Y) + \alpha_2 \ln(P_1) + \alpha_3 \ln(P_2) + \mu$$

参数  $\alpha_1$ 、 $\alpha_2$ 、 $\alpha_3$  估计量的经济意义分别为人均收入、食品类价格、其它商品类价格的需

求弹性; 由于食品为必需品,  $V$  为人均购买食品支出额, 所以  $\alpha_1$  应该在 0 与 1 之间,  $\alpha_2$  应

该在 0 与 1 之间,  $\alpha_3$  在 0 左右, 三者之和为 1 左右。所以, 该模型估计结果中  $\alpha_2$  的估计量缺少合理的经济解释。

(2) 由于该模型中包含长期弹性  $\alpha_1$  和短期弹性  $\alpha_2$  与  $\alpha_3$ , 需要分别采用截面数据和时序数据进行估计, 所以经常采用交叉估计方法估计需求函数模型。

6. (9 分) 选择两要素一级 CES 生产函数的近似形式建立中国电力行业的生产函数模型:

$$\ln Y = \ln A + \gamma t + m \delta \ln K + m(1 - \lambda) \ln L - \frac{1}{2} m \rho \delta (1 - \delta) (\ln K_L)^2 + \mu$$

其中  $Y$  为发电量,  $K$ 、 $L$  分别为投入的资本与劳动数量,  $t$  为时间变量。

- (1) 指出参数  $\gamma$ 、 $\rho$ 、 $m$  的经济含义和数值范围;
- (2) 指出模型对要素替代弹性的假设, 并指出它与 C-D 生产函数、VES 生产函数在要素替代弹性假设上的区别;
- (3) 指出模型对技术进步的假设, 并指出它与下列生产函数模型

$$\ln Y = \ln A + \gamma t + \alpha \ln K + \beta \ln L + \mu$$

在技术进步假设上的区别;

答案:

(1) 参数  $\gamma$  为技术进步速度, 一般为接近 0 的正数;  $\rho$  为替代参数, 在  $(-1, \infty)$  范围内;  $m$  为规模报酬参数, 在 1 附近。

(2) 该模型对要素替代弹性的假设为: 随着研究对象、样本区间而变化, 但是不随着样本点而变化。而 C-D 生产函数的要素替代弹性始终为 1, 不随着研究对象、样本区间而变化, 当然也不随着样本点而变化; VES 生产函数的要素替代弹性除了随着研究对象、样本区间而

变化外，还随着样本点而变化。

(3) 该模型对技术进步的假设为希克斯中性技术进步；而生产函数模型

$$\ln Y = \ln A + \gamma t + \alpha \ln K + \beta \ln L + \mu$$

的技术进步假设为中性技术进步，包括 3 种中性技术进步。

7. (8 分) 试指出在目前建立中国宏观计量经济模型时，下列内生变量应由哪些变量来解释，简单说明理由，并拟定关于每个解释变量的待估参数的正负号。

- |            |               |
|------------|---------------|
| (1) 轻工业增加值 | (2) 衣着类商品价格指数 |
| (3) 货币发行量  | (4) 农业生产资料进口额 |

答案：

(1) 轻工业增加值应该由反映需求的变量解释。包括居民收入（反映居民对轻工业的消费需求，参数符号为正）、国际市场轻工业品交易总额（反映国际市场对轻工业的需求，参数符号为正）等。

(2) 衣着类商品价格指数应该由反映需求和反映成本的两类变量解释。主要包括居民收入（反映居民对衣着类商品的消费需求，参数符号为正）、国际市场衣着类商品交易总额（反映国际市场对衣着类商品的需求，参数符号为正）、棉花的收购价格指数（反映成本对价格的影响，参数符号为正）等。

(3) 货币发行量应该由社会商品零售总额（反映经济总量对货币的需求，参数符号为正）、价格指数（反映价格对货币需求的影响，参数符号为正）等变量解释。

(4) 农业生产资料进口额应该由国内第一产业增加值（反映国内需求，参数符号为正）、国内农业生产资料生产部门增加值（反映国内供给，参数符号为负）、国际市场价格（参数符号为负）、出口额（反映外汇支付能力，参数符号为正）等变量解释。

8. (8 分) 回答：

- (1) 随机时间序列的平稳性条件是什么？证明随机游走序列不是平稳序列。
- (2) 单位根检验为什么从 DF 检验扩展到 ADF 检验？

答案：

(1) 随机时间序列  $\{X_t\}$  ( $t=1, 2, \dots$ ) 的平稳性条件是：1) 均值  $E(X_t) = \mu$ ，是与时间  $t$  无关的常数；2) 方差  $\text{var}(X_t) = \sigma^2$ ，是与时间  $t$  无关的常数；3) 协方差  $\text{cov}(X_t, X_{t+k}) = \gamma_k$ ，只与时期间隔  $k$  有关，与时间  $t$  无关的常数。

对于随机游走序列  $X_t = X_{t-1} + \mu_t$ ，假设  $X_t$  的初值为  $X_0$ ，则易知

$$\begin{aligned} X_1 &= X_0 + \mu_1 \\ X_2 &= X_1 + \mu_2 = X_0 + \mu_1 + \mu_2 \\ &\vdots \\ X_t &= X_0 + \mu_1 + \mu_2 + \dots + \mu_t \end{aligned}$$

由于  $X_0$  为一常数， $\mu_t$  是一个白噪声，因此  $\text{var}(X_t) = t\sigma^2$ ，即  $X_t$  的方差与时间  $t$  有关而非

常数，所以它是一非平稳序列。

(2) 在采用 DF 检验对时间序列进行平稳性检验中，实际上假定了时间序列是由具有白噪声随机误差项的一阶自回归过程（AR(1)）生成的。但在实际检验中，时间序列可能是由更高阶的自回归过程生成的，或者随机误差项并非白噪声，这样用 OLS 法进行估计均会表现出随机误差项出现自相关，导致 DF 检验无效。另外，如果时间序列包含有明显的随时间变化的某种趋势（如上升或下降），则也容易导致 DF 检验中的自相关随机误差项问题。为了保证 DF 检验中随机误差项的白噪声特性，Dickey 和 Fuller 对 DF 检验进行了扩充，形成了 ADF 检验。

## 清华大学本科生考试试题专用纸

考试课程 计量经济学

2006 年 6 月 16 日

(开卷, 2 小时, 满分 70 分)

1. (24 分, 每小题 4 分) 在清华大学一年级学生中按照抽样理论随机抽取 500 个学生作为样本, 建立学生消费函数模型。经分析, 认为第  $i$  学生的消费水平  $C_i$  主要受收入  $I_i$  的影响, 同时由于消费的示范性, 该学生所在班级的平均收入水平  $\bar{I}_i$  也作为解释变量, 不考虑其他因素, 被解释变量  $C_i$  与解释变量之间为线性关系。

(1) 写出该计量经济学问题的总体回归函数、样本回归函数、总体回归模型、样本回归模型。

(2) 证明: 在满足所有基本假设的情况下, OLS 估计量是无偏估计量。

(3) 试分析该模型可能违背哪些基本假设。

(4) 利用正规方程组, 证明: 假定  $\bar{I}_i$  与  $I_i$  以及常数项完全独立 (仅作为假定, 不考虑实际上是否成立), 那么, 在模型中去掉变量  $\bar{I}_i$ , 变量  $I_i$  的系数估计结果不发生变化。

(5) 如果欲检验学生消费是否满足绝对收入假设, 试构造约束回归检验的  $0$  假设和  $F$  统计量; 并回答给定显著性水平  $\alpha = 0.01$  时, 如果计算得到  $F = 5.0$ , 是否接受  $0$  假设。

(6) 如果用 OLS 估计该模型得到残差平方和为 50000, 试计算采用最大似然法估计时的最大对数似然函数值: ( $\ln 2 = 0.693, \ln \pi = 1.145, \ln 10 = 2.303$ )。

2. (10 分, 每小题 5 分) 在一篇研究中国国际收支与经济增长的关系的论文中, 作者分别建立了如下两个模型:

$$GDP_t = 1.465FDI_t + 0.943DI_t$$

$$GDP_t = 4.941FDI_t - 0.338NEX_t$$

试图分别分析外商直接投资  $FDI$ 、国内投资  $DI$  与国内生产总值  $GDP$  之间的关系和外商直接投资  $FDI$ 、净出口  $NEX$  与国内生产总值  $GDP$  之间的关系。

(1) 试指出该论文在模型设定方面的主要问题。

(2) 试设计你认为正确的模型来分析中国国际收支与经济增长之间的关系。

3. (16 分, 每小题 4 分) 投资函数模型

$$I_t = \beta_0 + \beta_1 Y_t + \beta_2 Y_{t-1} + \mu_t$$

为一完备的联立方程计量经济模型中的一个方程, 该方程是可以识别的。模型系统包含的内生变量为  $C$  (消费总额)、 $I$  (投资总额) 和  $Y$  (支出法国内生产总值), 先决变量为  $NEX_t$  (净

出口额)、 $C_{t-1}$  和  $Y_{t-1}$ 。样本容量为  $n$ 。

(1) 根据给定的变量, 试写出模型系统的另外两个方程, 并保证模型系统是可以识别的。

(2) 可否用狭义的工具变量法估计该投资函数模型? 为什么?

(3) 如果采用 2SLS 估计该投资函数模型, 分别写出 2SLS 估计量和将它作为一种工具变量方法的估计量的矩阵表达式:

(4) 如果采用 GMM 方法估计该投资函数模型, 写出一组等于 0 的矩条件。

4. (12 分, 每小题 4 分) 以宏观时间序列数据居民消费总额  $C_t$ 、国内生产总值  $GDP_t$  和居民储蓄余额  $S_t$  为样本观测值, 建立如下居民消费模型:

$$C_t = \alpha + \beta GDP_t + \gamma S_t + \mu_t \quad t = 1980, 1981, \dots, 2005$$

经 ADF 检验, 有:  $C_t \sim I(1)$ ,  $GDP_t \sim I(1)$ ,  $S_t \sim I(2)$ 。

(1) 试回答: 该模型的样本能否被认为是“随机抽取”的? 为什么?

(2) 试分析该模型的随机项的平稳性, 并简单说明理由。

(3) 根据已经知道的检验结果, 写出对  $C_t$  和  $S_t$  分别进行 ADF 检验的最后待检验模型的可能形式。

5. (8 分, 每小题 4 分) 二元 probit 模型  $y_i = \alpha + \beta x_{1i} + \gamma x_{2i} + \varepsilon_i, i = 1, 2, \dots, n$ ,  $y_i$  取 1 或者 0。

估计结果的回归方程输出如下:  $\hat{y}_i = 1 - @CNORM(-(\alpha + \beta x_{1i} + \gamma x_{2i}))$ 。

(1) 该模型区别于经典计量经济学模型的主要之处是什么? 模型中的随机误差项是否满足白噪声假设? 为什么?

(2) 如果对于第  $i$  个体, 计算得到  $\alpha + \beta x_{1i} + \gamma x_{2i} = 0.200$ , 查标准正态分布表有  $P(z > 0.2) = 0.4207$ , 试计算  $\hat{y}_i$  的值, 并指出它的含义。

计量经济学期末试卷标准答案（2006）  
（开卷，2 小时，满分 70 分）

1. (24 分，每小题 4 分)

(1) 总体回归函数为  $E(C_i | I_i, \bar{I}_i) = \beta_0 + \beta_1 I_i + \beta_2 \bar{I}_i$

样本回归函数为  $\hat{C}_i = \hat{\beta}_0 + \hat{\beta}_1 I_i + \hat{\beta}_2 \bar{I}_i$

总体回归模型为  $C_i = \beta_0 + \beta_1 I_i + \beta_2 \bar{I}_i + \mu_i$

样本回归模型为  $\hat{C}_i = \hat{\beta}_0 + \hat{\beta}_1 I_i + \hat{\beta}_2 \bar{I}_i + \hat{\mu}_i$

(2)

$$E\left(\begin{matrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \end{matrix} \right) = E\left(\begin{matrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{matrix} \right) = \begin{pmatrix} 1 & \bar{I} & \bar{I}^2 \end{pmatrix}^{-1} \begin{pmatrix} E(C) \\ E(CI) \\ E(C\bar{I}) \end{pmatrix} = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{pmatrix}$$

$$E\left(\begin{matrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \end{matrix} \right) = \begin{pmatrix} 1 & \bar{I} & \bar{I}^2 \end{pmatrix}^{-1} \begin{pmatrix} E(C) \\ E(CI) \\ E(C\bar{I}) \end{pmatrix} = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{pmatrix}$$

其中利用了所有解释变量与随机项不相关的假设。

(3) 该问题采用截面数据为样本，由于各个个体其它方面的差异，容易造成随机项的异方差，违背同方差假设。

(4) 包括变量  $\bar{I}_i$  的模型的关于待估参数的正规方程组为：

$$\begin{aligned} \sum C_i &= n\beta_0 + \beta_1 \sum I_i + \beta_2 \sum \bar{I}_i \\ \sum C_i I_i &= \beta_0 \sum I_i + \beta_1 \sum I_i^2 + \beta_2 \sum \bar{I}_i I_i \\ \sum C_i \bar{I}_i &= \beta_0 \sum \bar{I}_i + \beta_1 \sum \bar{I}_i I_i + \beta_2 \sum \bar{I}_i^2 \end{aligned}$$

利用假设条件，其前两个方程为：

$$\begin{aligned} \sum C_i &= n\beta_0 + \beta_1 \sum I_i \\ \sum C_i I_i &= \beta_0 \sum I_i + \beta_1 \sum I_i^2 \end{aligned}$$

与不包括变量  $\bar{I}_i$  的模型的关于待估参数的正规方程组完全相同。所以在模型中去掉变量  $\bar{I}_i$ ，

变量  $I_i$  的系数估计结果不发生变化。

(5) 欲检验学生消费是否满足绝对收入假设，即对变量  $I_i$  的参数施加 0 约束。即  $H_0: \beta_2 = 0$ 。

相应的 F 统计量为:  $F = \frac{RSS_r - RSS_u}{RSS_u / 497} \sim F(1, 497)$ 。

给定显著性水平  $\alpha = 0.01$  时, 查得  $F_{0.01}(1, 497) = 6.70$ , 如果计算得到  $F = 5.0$ , 接受  $H_0$  假设。

(6) 计算如下:

$$\sigma_2 = \frac{e'e}{n} = 100$$

$$\begin{aligned} L = \ln(L) &= -n \ln\left(\frac{\sqrt{2\pi\sigma_\mu^2}}{2\pi\sigma_\mu^2}\right) - \frac{1}{2\sigma_\mu^2} e'e \\ &= -500 \cdot \ln\left(\frac{\sqrt{2\pi \cdot 10}}{2\pi \cdot 10}\right) - \frac{1}{2 \cdot 100} \cdot 50000 = -1861 \end{aligned}$$

2. (10 分, 每小题 5 分)

(1) 该论文在模型设定方面的主要问题是违背了从一般到简单的原则, 而是对于同一个被解释变量, 建立了两个互相不能包容的简单的模型。

(2) 正确的分析思路是: 从需求方面考虑, 国内生产总值受消费 (C)、投资 (包括外商直接投资和国内投资) 和净出口的影响, 所以应该建立如下的一般模型:

$$GDP = f(C, DI, FDI, NEX) + \mu$$

例如, 可以采用线性模型形式:

$$GDP_t = \beta_0 + \beta_1 C_t + \beta_2 DI_t + \beta_3 FDI_t + \beta_4 NEX_t + \mu_t$$

根据模型的变量显著性检验和参数估计结果, 判断国际收支 (外商直接投资  $FDI$  和净出口  $NEX$ ) 与经济增长之间的关系。

3. (16 分, 每小题 4 分)

(1) 根据给定的变量写出模型系统的另外两个方程为:

$$C_t = \alpha_0 + \alpha_1 Y_t + \alpha_2 C_{t-1} + \mu_{1t}$$

$$Y_t = C_t + I_t + NEX_t$$

或者 
$$C_t = \alpha_0 + \alpha_1 Y_t + \alpha_2 C_{t-1} + \mu_{1t}$$

$$Y_t = \gamma_0 + \gamma_1 C_t + \gamma_2 I_t + \gamma_3 NEX_t + \mu_{3t}$$

根据方程之间的关系, 可以判断该模型系统是可以识别的。

(2) 因为投资函数模型为过度识别, 所以, 可以用狭义的工具变量法估计该模型, 但是在工具变量的选择上会遇到困难。所以一般不采用。(如果答案为不能用, 也认为是正确的)



(3) 如果采用 2SLS 估计该投资函数模型，其 2SLS 估计量的矩阵表达式为：

$$\begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix} = ((1, Y, Y_{t-1}) (1, Y, Y_{t-1}))^{-1} (1, Y, Y_{t-1})' I$$

其中  $Y_t = \pi_0 + \pi_1 NEX_t + \pi_2 C_{t-1} + \pi_3 Y_{t-1}$

如果将 2SLS 作为一种工具变量方法，其估计量的矩阵表达式为：

$$\begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix} = ((1, Y, Y_{t-1}) (1, Y, Y_{t-1}))^{-1} (1, Y, Y_{t-1})' I$$

(4) 如果采用 GMM 方法估计该投资函数模型，构造的一组等于 0 的矩条件为：

$$\begin{aligned} \sum I_t - \sum (\hat{\beta}_0 + \hat{\beta}_1 Y_t + \hat{\beta}_2 Y_{t-1}) &= 0 \\ \sum I_t NEX_t - \sum (\hat{\beta}_0 + \hat{\beta}_1 Y_t + \hat{\beta}_2 Y_{t-1}) NEX_t &= 0 \\ \sum I_t C_{t-1} - \sum (\hat{\beta}_0 + \hat{\beta}_1 Y_t + \hat{\beta}_2 Y_{t-1}) C_{t-1} &= 0 \\ \sum I_t Y_{t-1} - \sum (\hat{\beta}_0 + \hat{\beta}_1 Y_t + \hat{\beta}_2 Y_{t-1}) Y_{t-1} &= 0 \end{aligned}$$

4. (12 分，每小题 4 分)

(1) 该模型的样本不能被认为是“随机抽取”的。因为样本之间存在序列相关，因此不能随便改变样本次序。

(2) 该模型的随机项不具有平稳性。因为

$$C_t - (\alpha + \beta GDP_t + \gamma S_t) = \mu_t$$

而  $S_t \sim I(2)$ ，所以上式左边不可能为 0 阶单整。

(3) 根据已经知道的检验结果， $C_t \sim I(1)$ ，它的一阶差分为平稳序列，所以进行 ADF 检验

的最后待检验模型的可能形式只能是

$${}^2C_t = \alpha + \beta t + \delta C_{t-1} + \sum_{i=1}^m \gamma_i {}^2C_{t-i} + \varepsilon_t$$

$$\text{或者 } {}^2C_t = \alpha + \delta C_{t-1} + \sum_{i=1}^m \gamma_i {}^2C_{t-i} + \varepsilon_t \text{ 和 } {}^2C_t = \delta C_{t-1} + \sum_{i=1}^m \gamma_i {}^2C_{t-i} + \varepsilon_t$$

同理，根据已经知道的检验结果， $S_t \sim I(2)$ ，它的二阶差分为平稳序列，所以进行 ADF

检验的最后待检验模型的可能形式只能是

$$^3S_t = \alpha + \beta t + \sigma^2 S_{t-1} + \sum_{i=1}^m \gamma_i ^3S_{t-i} + \varepsilon_t$$

$$\text{或者 } ^3S_t = \alpha + \sigma^2 S_{t-1} + \sum_{i=1}^m \gamma_i ^3S_{t-i} + \varepsilon_t \text{ 和 } ^3S_t = \sigma^2 S_{t-1} + \sum_{i=1}^m \gamma_i ^3S_{t-i} + \varepsilon_t$$

5. (8 分，每小题 4 分)

(1) 与经典计量经济学模型的主要区别之处是被解释变量是离散变量。模型中的随机误差项不满足白噪声假设，存在异方差，均值不为 0。

(2) 计算过程为

$$\hat{y}_i = 1 - @CNORM(-(\hat{\alpha} + \hat{\beta} x_{1i} + \hat{\gamma} x_{2i})) = 1 - @CNORM(-0.200) = 1 - (0.4207) = 0.5793$$

其含义是根据模型预测第 i 个体选择 1 的概率为 0.5793。

## 清华大学本科生考试试题专用纸

考试课程 计量经济学      2007 年 7 月 4 日      A 卷  
(满分 70 分)

(一) (32 分) 将中国城镇居民按照人均年收入分成 20 组, 以 2005 年的组平均数为样本观测值, 建立中国城镇居民消费函数模型, 以人均年消费额  $C_i$  为被解释变量, 经过理论分析和经验检验, 选择人均年收入  $I_i$  和人均储蓄余额  $S_i$  作为解释变量, 解释变量和被解释变量之间的关系为直接线性关系。模型形式为:

$$C_i = \beta_0 + \beta_1 I_i + \beta_2 S_i + \varepsilon_i \quad i = 1, 2, \dots, 20$$

- (1) (4 分) 分别写出该问题的总体回归函数、总体回归模型、样本回归函数和样本回归模型。
- (2) (3 分) 分别写出随机误差项具有同方差且无序列相关、具有异方差但无序列相关、具有异方差且具有一阶序列相关时的方差—协方差矩阵。
- (3) (3 分) 当模型满足基本假设时, 写出关于普通最小二乘法参数估计量的正规方程组。
- (4) (4 分) 如果用矩阵方程  $\mathbf{Y} = \mathbf{XB} + \mathbf{E}$  表示该模型, 写出每个矩阵的具体内容, 并标明阶数。
- (5) (3 分) 直观判断该模型是否具有异方差性? 为什么?
- (6) (3 分) 指出“偏回归系数”  $\beta_1$  的实际含义, 并指出解释变量满足什么条件时可以用一元回归模型得到相同的  $\beta_1$  的估计结果?
- (7) (6 分) 写出检验下述命题的原假设: “人均储蓄余额对人均年消费额无影响”; 写出两个可进行该检验的统计量。
- (8) (6 分) 如果在中国城镇居民中按照随机抽样原理抽取 1000 个家庭, 以家庭人均数作为样本观测值, 建立同样形式的中国城镇居民消费函数模型, 试比较两种模型的优劣, 并简单说明理由。

(二) (12 分) 某联立方程模型系统共有 3 个内生变量 ( $y_1, y_2, y_3$ )、3 个先决变量 ( $x_1, x_2, x_3$ ) 和样本观测值始终为 1 的虚变量。

- (1) (6 分) 试写出一个全部由随机方程组成的完备的结构式模型, 每个方程都包含常数项、内生解释变量和先决解释变量, 且保证每个随机方程都是可以识别的。
- (2) (3 分) 指出其中每个随机方程的识别状态。
- (3) (3 分) 指出其中每个随机方程的合适的单方程估计方法。

(三) (8 分) 建立关于失业的 Panel Data 计量经济学模型。选择 31 个省、市的 1995—2004 年数据为样本观测值，以失业率 ( $Y$ ) 为被解释变量，人均 GDP 与当年全国人均 GDP 的比值 ( $X$ ) 和重工业增加值在 GDP 中所占比例 ( $Z$ ) 为解释变量。如果经过检验为一截距和系数都随省、市而变的模型，且假设都只存在固定影响。

(1) (4 分) 写出该模型的一般形式，并指出该模型中各个变量参数的经济含义和符号 (正或者负)。

(2) (4 分) 试从实际经济行为出发分析：可否将模型分成 31 个截面个体方程分别利用各自的时间序列数据为样本进行 OLS 估计？为什么？

(四) (12 分) 试回答：

(1) (3 分) 单整变量的单位根检验为什么从 DF 检验发展到 ADF 检验？

(2) (3 分) 如果  $X_t \sim I(1), Y_t \sim I(1), X_t, Y_t \sim CI(1,1)$ ，而且从经济意义分析表明， $X$  是  $Y$  的唯一原因变量。写出  $X$  和  $Y$  之间长期均衡方程的一般形式。

(3) (3 分) 条件同(2)，写出反映  $X$  和  $Y$  之间短期关系的 ECM 模型的一般形式。

(4) (3 分) 指出 ECM 模型中  $ecm$  项参数的符号 (正或者)，并简述它所反映的经济机制。

(五) (6 分) 分别指出 ELES 需求函数模型系统

$$q_i = r_i + \frac{b_i}{p_i} \left( I - \sum_j p_j r_j \right) \quad i = 1, 2, \dots, n$$

中  $\sum_{j=1}^n b_j$ 、 $\sum_{j=1}^n p_j r_j$ 、 $\frac{b_i I}{p_i q_i}$  的经济含义和数值范围。

2007 年计量经济学期末试题答案 (A 卷, 满分 70 分)

(一) (32 分)

(1) (4 分)

$$\text{总体回归函数: } E(C_i | I_i, S_i) = \beta_0 + \beta_1 I_i + \beta_2 S_i \quad i = 1, 2, \dots, 20$$

$$\text{总体回归模型: } C_i = \beta_0 + \beta_1 I_i + \beta_2 S_i + \varepsilon_i \quad i = 1, 2, \dots, 20$$

$$\text{样本回归函数: } \hat{C}_i = \hat{\beta}_0 + \hat{\beta}_1 I_i + \hat{\beta}_2 S_i \quad i = 1, 2, \dots, 20$$

$$\text{样本回归模型: } \hat{C}_i = \hat{\beta}_0 + \hat{\beta}_1 I_i + \hat{\beta}_2 S_i + \hat{\varepsilon}_i \quad i = 1, 2, \dots, 20$$

(2) (3 分)

随机误差项具有同方差且无序列相关时的方差—协方差矩阵:

$$\text{Cov}(\varepsilon_i, \varepsilon_j) = \begin{matrix} & \sigma^2 & 0 & \text{L} & 0 \\ & 0 & \sigma^2 & \text{L} & 0 \\ \text{M} & & \text{M} & & \text{M} \\ & 0 & 0 & \text{L} & \sigma^2 \end{matrix}$$

随机误差项具有异方差但无序列相关时的方差—协方差矩阵:

$$\text{Cov}(\varepsilon_i, \varepsilon_j) = \begin{matrix} \sigma_1^2 & 0 & \text{L} & 0 \\ & \sigma_2^2 & & \\ 0 & & \text{L} & 0 \\ \text{M} & \text{M} & & \text{M} \\ & 0 & 0 & \text{L} & \sigma_{20}^2 \end{matrix}$$

随机误差项具有异方差且具有一阶序列相关时的方差—协方差矩阵:

$$\text{Cov}(\varepsilon_i, \varepsilon_j) = \begin{matrix} \text{L} & 0 & \sigma \\ \text{L} & \sigma_1^2 & 0 & \sigma_{12} & 0 \\ \sigma_{21} & 0 & \sigma_2^2 & \sigma_{23} & \\ & \sigma & \sigma & \sigma & \sigma^2 \\ 0 & 0 & \sigma_{32} & \sigma_3^2 & \\ \text{M} & \text{M} & \text{M} & & \\ \text{K} & \sigma_{20}^2 & & & \\ & 0 & 0 & \text{L} & \end{matrix}$$

(3) (3 分)

当模型满足基本假设时, 关于普通最小二乘法参数估计量的正规方程组:

$$\begin{aligned} \sum_i C_i - \sum_i \hat{\beta}_0 + \sum_i \hat{\beta}_1 I_i + \sum_i \hat{\beta}_2 S_i &= 0 \\ \sum_i I_i C_i - \sum_i I_i \hat{\beta}_0 + \sum_i \hat{\beta}_1 I_i + \sum_i \hat{\beta}_2 S_i &= 0 \\ \sum_i S_i C_i - \sum_i S_i \hat{\beta}_0 + \sum_i \hat{\beta}_1 I_i + \sum_i \hat{\beta}_2 S_i &= 0 \end{aligned}$$

(4) (4 分)

如果用矩阵方程  $\mathbf{Y} = \mathbf{XB} + \mathbf{E}$  表示该模型，每个矩阵的具体内容：

$$\mathbf{Y} = \begin{bmatrix} C_1 \\ C_2 \\ \vdots \\ C_M \\ \vdots \\ C_{20} \end{bmatrix} \quad \mathbf{X} = \begin{bmatrix} 1 & I_1 & S_1 \\ 1 & I_2 & S_2 \\ \vdots & \vdots & \vdots \\ 1 & I_M & S_M \\ \vdots & \vdots & \vdots \\ 1 & I_{20} & S_{20} \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix} \quad \mathbf{E} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_M \\ \vdots \\ \varepsilon_{20} \end{bmatrix}$$

(5) (3 分)

直观判断，该模型具有异方差性。

采用组平均数据作为样本，因为收入呈现正态分布，每组中包含的个体数目不同，致使得到的不同的组平均数的误差不同。而这些误差的影响进入随机项，使得随机项具有不同的方差。

(6) (3 分)

“偏回归系数”  $\beta_1$  的实际含义是收入对消费的直接贡献，即边际消费倾向。

如果解释变量  $I$  和  $S$  互相独立，可以用一元回归模型得到相同的  $\beta_1$  的估计。

(7) (6 分)

检验命题“人均储蓄余额对人均年消费额无影响”的原假设为： $H_0: \beta_2 = 0$ 。

可以分别采用变量显著性的  $t$  统计量和约束检验的  $F$  统计量进行检验：

$$t = \frac{\hat{\beta}_2}{S_{\hat{\beta}_2}} \sim t(17) \quad F = \frac{(RSS_R - RSS_U) / 1}{RSS_U / 17} \sim F(1, 17)$$

(8) (6 分)

如果在中国城镇居民中按照随机抽样原理抽取 1000 个家庭，以家庭人均数作为样本观测值，建立同样形式的中国城镇居民消费函数模型，应该比本题提供的模型更合理。

一是在中国城镇居民中随机抽样，符合“从总体中随机抽取样本”的要求，用这样的样本估计总体参数，更加可靠。二是样本容量较大。三是样本观测值误差较小。

(二) (12 分)

(1) (6 分) 有多种答案。例如:

$$\begin{aligned} y_1 &= \alpha_0 + \alpha_1 y_2 + \alpha_2 x_1 + \mu_1 \\ y &= \beta_0 + \beta_1 y + \beta_2 x + \beta_3 x + \mu \end{aligned}$$

(2) (3 分)

第 1 个方程为过度识别; 第 2 个方程为恰好识别; 第 3 个方程为过度识别。(3) (3 分)

第 1 个方程适合于采用 2SLS 和 GMM; 第 2 个方程适合于采用 IV、ILS、2SLS; 第 3 个方程适合于采用 2SLS 和 GMM。

(三) (8 分)

(1) (4 分)

该模型的一般形式为:

$$\begin{aligned} y_{it} &= \alpha_i + \beta_{1i} x_{it} + \beta_{2i} z_{it} + \mu_{it} & E(\mu_{it}) &= 0, E(\mu_{it}^2) = \sigma_\mu^2 \\ i &= 1, 2, \dots, 31 & t &= 1, 2, \dots, 10 \end{aligned}$$

该模型中各个参数的经济含义:  $\beta_{1i}$  反映第  $i$  地区经济发展水平对失业率的影响程度, 应该为负数;  $\beta_{2i}$  反映第  $i$  地区产业结构对失业率的影响程度, 应该为正数。

(2) (4 分)

从实际经济行为出发分析, 失业率还受到其它因素的影响, 例如: 教育水平、技术进步、改革力度等, 而这些因素在不同地区之间是相关的, 所以引起模型随机项在不同地区之间是相关的。这样, 不能将模型分成 31 个截面个体方程分别利用各自的时间序列数据为样本进行 OLS 估计, 而应该采用 GLS 估计。

(四) (12 分)

(1) (3 分)

单整变量的单位根检验从 DF 检验发展到 ADF 检验, 简单说是为了保证检验模型中随机项的白噪声性。包括消除时间序列中可能存在的确定性趋势和高阶自回归。

(2) (3 分)

如果  $X_t \sim I(1), Y_t \sim I(1), X_t, Y_t \sim CI(1,1)$ , 而且从经济意义分析表明,  $X$  是  $Y$  的唯一

原因变量。 $X$  和  $Y$  之间长期均衡方程的一般形式为：

$$y_t = \alpha_0 + \alpha_1 x_t + \sum_{i=1}^p \beta_i y_{t-i} + \sum_{i=1}^q \gamma_i x_{t-i} + \mu_t \quad t = 1, 2, \dots, T$$

(3) (3 分)

反映  $X$  和  $Y$  之间短期关系的 ECM 模型的一般形式：

$$y_t = \alpha_0 + \alpha_1 y_{t-1} + \sum_{i=1}^p \beta_i (y_t - y_{t-i}) + \sum_{i=1}^q \gamma_i (x_t - x_{t-i}) + \lambda ecm_{t-1} + \varepsilon_t \quad t = 1, 2, \dots, T$$

(4) (3 分)

ECM 模型中  $ecm$  项参数的符号应该为负。

如果在  $(t-1)$  时刻  $y$  对于均衡状态有一个大于 0 的偏离，即  $ecm_{t-1}$  为正，那么在  $t$  时刻必须对  $y$  的增量施加一个小于 0 的修正，使之在  $t$  时刻回复到均衡状态；反之亦然。

(五) (6 分)

在 ELES 需求函数模型系统  $q_i = r_i + \frac{b_i}{p_i} (I - \sum_j p_j r_j)$   $i = 1, 2, \dots, n$  中，

$\sum_{j=1}^n b_j$  表示所有商品的边际消费倾向之和， $\sum_{j=1}^n b_j \leq 1$ ；

$\sum_{j=1}^n p_j r_j$  表示基本生活线， $0 < \sum_{j=1}^n p_j r_j < I$ ；

$\frac{b_i I}{p_i q_i}$  表示需求的收入弹性， $0 < p_i q_i < 1$ 。

计量经济学期末试题（2007 年秋季）

（开卷，满分 70 分，A 卷）

1. (21 分，每小题 3 分) 多元线性回归模型：

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_k X_{ki} + \mu_i \quad \mu_i \sim N(0, \sigma^2) \quad i = 1, 2, \dots, n$$

其矩阵形式为： $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\mu}$ ，满足所有基本假设。

(1) 分别写出  $\mu_2$  的分布、 $Y_2$  的分布和  $Y_2$  的分布。



(2) 指出“偏回归系数”  $\beta_2$  的含义，并指出解释变量满足什么条件时可以用一元回归模型得到相同的  $\beta_2$  的估计结果？

(3) 如果  $Var(\mu_i) = (x_{1i} + x_{2i})^2 \sigma^2$ ，采用 WLS 估计得到  $\beta = (X'W^{-1}X)^{-1}X'W^{-1}Y$ ，写出其中  $W$  的具体表达式。

(4) 证明： $\sigma^2 = \frac{(Y - X\hat{\beta})'(Y - X\hat{\beta})}{n - k - 1}$  是  $\sigma^2$  的无偏估计。

(5) 如果解释变量  $X_k$  和  $X_{k-1}$  为与  $\mu$  相关的随机变量，仍然采用 OLS 估计得到

$\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_{k-1}, \hat{\beta}_k$ ，指出其中哪些是有偏估计？哪些是无偏估计？简单说明理由。

(6) 如果受到条件限制，被解释变量只能取大于  $a$  的样本观测值，用 OLS 和 ML 分别估计模型，参数估计量是否等价？为什么？

(7) 如果  $X_1$  为只有 2 种类别（A、B）的定性变量， $X_2$  为具有 3 种类别（C、D、E）的定性变量，重新写出该线性回归模型的表达式。

2. (15 分) 指出下列论文中的主要错误之处：

在一篇关于中国石油消费预测研究的论文中，作者选择石油年消费量（OIL，单位：万吨标准煤）为被解释变量，国内生产总值（GDP，按当年价格计算，单位：亿元）为解释变量，1990—2006 年年度数据为样本。首先假定边际消费倾向不变，建立了线性模型：

$$OIL_t = \alpha + \beta GDP_t + \mu_t \quad t = 1990, 1991, \dots, 2006$$

采用 OLS 估计模型，得到

$$\hat{OIL}_t = 13390.30 + 0.183125GDP_t \quad t = 1990, 1991, \dots, 2006$$

然后假定消费弹性不变，建立了对数线性模型：

$$\ln OIL_t = \alpha' + \beta' \ln GDP_t + \mu'_t \quad t = 1990, 1991, \dots, 2006$$

采用 OLS 估计模型，得到

$$\ln \hat{OIL}_t = 5.122385 + 0.458338 \ln GDP_t \quad t = 1990, 1991, \dots, 2006$$

分别将 2020 年国内生产总值预测值（500000 亿元）代入模型，计算得到两种不同假定情况下的 2020 年石油消费预测值分别为 104953 和 68656 万吨标准煤。

3. (12 分，每小题 4 分) 回答下列问题：

(1) 为什么要对联立方程模型中的随机方程进行识别？为什么恒等方程不能被排除在识别问题之外？

(2) 联立方程模型第  $i$  个结构方程识别状态的结构式条件为：如果  $R(B_0 \Gamma_0) < g - 1$ ，则第  $i$

个结构方程不可识别；如果  $R(B_0 \Gamma_0) = g - 1$ ，则第  $i$  个结构方程可以识别，并且如果

$k - k_i = g_i - 1$ ，则第  $i$  个结构方程恰好识别，如果  $k - k_i > g_i - 1$ ，则第  $i$  个结构方程过度识别。为什么在秩条件中没有讨论  $R(\mathbf{B}_0 \Gamma_0) > g - 1$  的情况？为什么在阶条中没有讨论  $k - k_i < g_i - 1$  的情况？

(3) 为什么 2SLS 既适合于恰好识别方程又适合于过度识别方程的估计？为什么它在实际联立方程模型的估计中并不具有实用性？

4. (共 12 分，每小题 4 分) 改进的 CES 生产函数模型的表达式为：

$$Y = A e^{\lambda t} (\delta_0 K^{-\rho} + \delta_1 L^{-\rho})^{-\frac{m}{\rho}}$$

其中， $Y$  为产出量， $K$ 、 $L$  为资本和劳动投入量， $t$  为时间变量，其它为参数。

- (1) 指出模型中每个参数的经济意义和数值范围。
- (2) 写出由该理论模型导出的用于参数估计的线性总体回归模型。
- (3) 设计一个假设检验，检验该生产函数模型可否采用 C-D 生产函数的形式，并写出两种不同的检验统计量。

5. (10 分) 在一篇研究我国工业资本配置效率的论文中，作者利用我国 39 个工业行业（编号  $i=1, 2, \dots, 39$ ）的 9 年（ $t=1991, 1991, \dots, 1999$ ）的 351 组数据为样本，以固定资产存量  $I$  的增长率为被解释变量，以利润  $V$  的增长率为解释变量。分别建立了如下 3 个模型：

$$\ln \frac{I_{it}}{I_{i,t-1}} = \alpha + \beta \ln \frac{V_{it}}{V_{i,t-1}} + \mu_{it} \quad (1)$$

$$\ln \frac{I_{it}}{I_{i,t-1}} = \alpha_i + \beta_i \ln \frac{V_{it}}{V_{i,t-1}} + \mu_{it} \quad (2)$$

$$\ln \frac{I_{it}}{I_{i,t-1}} = \alpha_t + \beta_t \ln \frac{V_{it}}{V_{i,t-1}} + \mu_{it} \quad (3)$$

利用全部样本，采用 OLS 估计模型①，结果表明我国资本配置效率不仅低于发达国家，也低于大多数发展中国家。为了分析我国工业行业的成长性，分别利用每个行业的时间序列数据，对模型②进行 OLS 估计，从结果中发现了最具发展潜力的 5 个行业。为了定量刻画我国每年的资本配置效率，分别用每年的行业数据，采用 OLS 估计模型③，从估计结果可以看出，我国资本配置效率呈逐年下滑趋势。

分别从 Panel Data 的模型设定和估计方法两方面，指出该论文存在的问题，并简单说明理由。

计量经济学期末试题及答案  
(2008 年 6 月，开卷，满分 70 分)

1. (15 分，每小题 3 分) 多元线性回归模型：

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_k X_{ki} + \mu_i \quad \mu_i \sim N(0, \sigma^2) \quad i = 1, 2, \dots, n$$

模型设定是正确的。如果遗漏了显著的变量  $X_k$ ，构成一个新模型

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_{k-1} X_{k-1i} + \varepsilon_i$$

试回答：

- (1) 如果  $X_k$  与其它解释变量完全独立，用 OLS 分别估计原模型和新模型， $\beta_0, \beta_1, \dots, \beta_{k-1}$  的估计结果是否变化？为什么？

不会发生变化，因为由于  $X_k$  与其他解释变量独立，去掉  $X_k$  的正规方程组与新模型一致。

- (2) 如果  $X_k$  与其它解释变量线性相关，用 OLS 分别估计原模型和新模型， $\beta_0, \beta_1, \dots, \beta_{k-1}$  的估计结果是否变化？为什么？

会，因为由于  $X_k$  与其他解释变量不独立，去掉  $X_k$  的正规方程组与新模型不一致，估计量发生变化。

- (3) 如果  $X_k$  是确定性变量，写出新模型中  $\varepsilon_i$  的分布。

$$\varepsilon_i \sim N(\mu_i + \beta_k X_{ki}, \sigma^2)$$

- (4) 如果  $X_k$  是随机变量，且服从正态分布，指出新模型中的  $\varepsilon_i$  是否服从正态分布？为什么？

是， $\mu_i$  服从正太分布，减去一个服从正态分布的随机变量仍然服从正态分布。

- (5) 如果  $X_k$  是随机变量，且服从正态分布，指出新模型是否存在异方差性？为什么？

若原模型存在异方差，则新模型也存在异方差，若原模型不存在异方差，新模型不存在异方差。

2. (12 分，每小题 4 分) 多元线性回归模型：

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_k X_{ki} + \mu_i \quad \mu_i \sim N(0, \sigma^2) \quad i = 1, 2, \dots, n$$

现有  $n$  组样本观测值，其中  $a < Y_i < b$  ( $i = 1, 2, \dots, n$ )，将它们看着是在以下 3 种不同的情况下抽取获得的：①完全随机抽取，②被解释变量被限制在大于  $a$  的范围内随机抽取，③被解释变量被限制在大于  $a$  小于  $b$  的范围内随机抽取。

- (1) 用 OLS 分别估计 3 种情况下的模型，结构参数估计量是否等价？为什么？

等价。OLS 只要样本观测值相同，无论被解释变量是否受到限制，其估计量相同。

- (2) 用 ML 分别估计 3 种情况下的模型，结构参数估计量是否等价？为什么？

不等价，对于 ML，当被解释变量受到限制时，抽取同一个样本的概率发生了变化，因而似然函数发生变化，估计结果也发生变化。

- (3) 用 ML 分别估计 3 种情况下的模型，比较 3 种情况的似然函数值。

情况三的似然函数值>情况二的似然函数值>情况一的似然函数值。

3. (9 分, 每小题 3 分) 下列联立方程模型是一个完备的结构式模型:

$$y_{1t} = \alpha_0 + \alpha_1 y_{2t} + \alpha_2 y_{3t} + \alpha_3 y_{5t} + \mu_{1t}$$

$$y_{2t} = \beta_0 + \beta_1 y_{1t} + \beta_2 y_{4t} + \mu_{2t}$$

$$y_{3t} = \gamma_0 + \gamma_1 y_{1t} + \gamma_2 y_{6t} + \mu_{3t}$$

(1) 指出每个结构方程和该联立方程模型的识别状态。

恰好识别 过度识别 过度识别

(2) 指出分别采用 OLS 和 2SLS 估计第一个方程的优点和缺点。

OLS: 优点: 实际操作性强, 最小均方差性。缺点: 有偏, 方差大, 不利用方程之间的相关信息。

2SLS: 利用方程之间的相关信息。缺点: 不易实际操作。

(3) 如果采用 2SLS 估计第二个方程, 分别写出第 1 阶段和第 2 阶段所估计的模型形式。

第一阶段:  $y_{1t} = \pi_1 y_{4t} + \pi_2 y_{5t} + \pi_3 y_{6t} + \varepsilon$

$$y_{2t} = (1, y_{1t}, y_{4t}) \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix} + \mu$$

第二阶段:  $\beta_2$

4. (10 分, 每题 2.5 分) 以某地区的农村居民人均年消费  $C_t$  为被解释变量, 以人均年家庭经营纯收入  $Y_t^p$  (可视为持久收入) 和人均年其它收入  $Y_t^t$  (可视为瞬时收入) 为解释变量,

建立农村居民消费模型, 以 1978 年至 2007 年数据为样本。经过检验,  $C_t$ 、 $Y_t^p$ 、 $Y_t^t$  都是  $I(1)$  序列。最后建立的 ECM 模型为:

$$C_t = \alpha + \beta_1 Y_t^p + \beta_2 Y_t^t + \gamma_1 C_{t-1} + \gamma_2 Y_{t-1}^p + \gamma_3 Y_{t-1}^t + \lambda ecm_{t-1} + \varepsilon_t$$

(1) 回答: ECM 模型中为什么引入  $C_{t-1}$ 、 $Y_{t-1}^p$ 、 $Y_{t-1}^t$  ?

满足白噪声, 消除序列相关性。

(2) 写出反映  $C_t$ 、 $Y_t^p$ 、 $Y_t^t$  之间协整关系的协整方程的理论形式。

$$\mu_t = C_t - \alpha - \beta_1 Y_t^p - \beta_2 Y_t^t$$

(3) 写出反映  $C_t$ 、 $Y_t^p$ 、 $Y_t^t$  之间长期均衡关系的长期均衡方程的理论形式。

$$C_t = \alpha + \beta_1 Y_t^p + \beta_2 Y_t^t + \sum_{i=1}^m C_{t-i} + \sum_{i=1}^g Y_{t-i}^p + \sum_{i=1}^h Y_{t-i}^t$$

(4) 写出  $ecm_{t-1}$  的表达式。

$$ecm_{t-1} = C_{t-1} - \alpha - \beta Y_{t-1}^p - \gamma Y_{t-1}^l$$

5. (24 分, 每题 6 分) 回答以下问题:

(1) 一位同学在综合练习中根据需求法则建立中国食品需求模型, 以 31 个省会城市 2006 年数据为样本, 以人均年食品消费量为被解释变量, 以食品价格指数为解释变量, 建立一元回归模型, 估计得到食品价格指数的参数为正, 于是发现“需求法则不适用于中国”。试回答: ①该问题的主要错误在哪里? ②试建立一个你认为正确的模型。

遗漏显著变量: 收入。 增加收入解释变量

(2) 在一篇研究农村建设用地流转中的农民收入问题的论文中, 经过分析认为, 影响农民的土地流转收入的因素包括 3 组 (集体组织内部治理、农村集体建设用地市场、国家土地管制政策) 共 13 个解释变量, 为了定量分析它们之间的关系, 作者直接建立了 13 个一元回归模型。试回答: ①论文的建模思路是否正确? 为什么? ②在什么情况下论文模型参数估计结果仍然是可用的? 为什么?

否, 这 13 个解释变量之间具有很强的相关性。

只有当各个解释变量互相独立时, 才可以建立 13 个一元回归模型。

(3) 在一篇研究居民收入差距与经济增长之间关系的论文中, 作者分析了居民收入差距将直接影响固定资产投资和人力资本投资, 而固定资产投资和人力资本投资直接影响经济增长, 于是建立了一个定量分析模型, 以 GDP 增长率为被解释变量, 居民收入差距及其平方项、固定资产投资增长率和人力资本投资增长率同时作为解释变量。试回答: ①该模型的解释变量选择是否正确? 为什么? ②如果认为不正确, 应该如何建立该问题的计量经济学模型? (只需说明思路)

居民收入差距直接影响固定资产投资和人力资本投资, 因此解释变量之间具有共线性。

应建立联立方程模型。

(4) 某人为了研究城市的环境问题, 选择了国内 100 个城市 10 年的数据为样本, 以环境质量综合指数为被解释变量, 以城市规模、人口密度、产业结构、能源结构等连续变量为解释变量, 建立 Panel Data 模型。经过检验, 模型为固定影响变截距模型, 采用 LSDV 方法进行模型估计。然后, 作者将港口虚拟变量 D 引入模型 (D=1 为港口, D=0 为非港口), 希望分析港口因素是否对环境质量产生影响, 采用 LSDV 方法进行模型估计, 显示为 “Near singular matrix”, 无法估计。试回答: ①该问题是否真的出现完全共线性? ②港口虚拟变量的引入是否必要? 为什么?

不是。固定影响变截距模型中本身含有虚拟变量的设置, 因此港口虚拟变量的引入重复, 没有必要。