目　录

第一章 绪论

练习题

一、单项选择题

1．经济计量学一词的提出者为（ ）

A．弗里德曼　　B．丁伯根

C．费瑞希　　D．萨缪尔森

2．下列说法中正确的是（ ）

A．经济计量学是经济学、统计学和数学合流而构成的一门交叉学科。

B．经济计量学是经济学、数理统计学和政治经济学合流而构成的一门交叉学科。

C．经济计量学是数理经济学和政治经济学合流而构成的一门交叉学科。

D．经济计量学就是数理经济学。

3．理论经济计量学的主要目的为（ ）

A．研究经济变量之间的依存关系

B．研究经济规律

C．测度由经济计量学模型设定的经济关系式

D．进行经济预测

4．下列说法中不是应用经济计量学的研究目的为（ ）

A．测度经济系统的发展水平

B．经济系统结构分析

C．经济指标预测

D．经济政策评价

5．经济计量学的建模依据为（ ）

A．统计理论　　B．预测理论

C．经济理论　　D．数学理论

6．随机方程式构造依据为（ ）

A．经济恒等式　　B．政策法规

C．变量间的技术关系　　D．经济行为

7．经济计量学模型的被解释变量一定是（ ）

A．控制变量　　B．政策变量

C．内生变量　　D．外生变量

8．在同一时点或时期上，不同统计单位的相同统计指标组成的数据是（ ）

A．时期数据　　B．时点数据

C．时序数据　　D．截面数据

二、多项选择题

1．在一个经济计量模型中，可作为解释变量的有（ ）

A．内生变量　　B．外生变量

C．控制变量　　　　　　　D．政策变量
E．滞后变量

2．对经济计量模型验证的准则有（　　）
A．最小二乘准则　　　　　B．经济理论准则
C．统计准则　　　　　　　D．数学准则
E．经济计量准则

3．经济计量模型的应用在于（　　）
A．设定模型　　　　　　　B．检验模型
C．结构分析　　　　　　　D．经济预测
E．规划政策

三、名词解释

1．经济计量学
2．理论经济计量学
3．应用经济计量学
4．内生变量
5．外生变量
6．随机方程
7．非随机方程
8．时序数据
9．截面数据

四、简答题

1．简述经济计量分析的研究步骤。
2．简述经济计量模型检验的三大原则。
3．简述经济计量模型的用途。

参考答案

一、单项选择题

1．C　　2．A　　3．C　　4．A　　5．C　　6．D　　7．C　　8．D

二、多项选择题

1．ABCDE　　　2．BCE　　　3．CDE

三．名词解释

1．经济计量学：是经济学、统计学和数学合流而构成的一门交叉学科。

2．理论经济计量学：是寻找适当的方法，去测度由经济计量模型设定的经济关系式。

3．应用经济计量学：以经济理论和事实为出发点，应用计量方法，解决经济系统运行过程中的理论问题或实践问题。

4．内生变量：具有一定概率分布的随机变量，由模型自身决定，其数值是求解模型的结果。

5．外生变量：是非随机变量，在模型体系之外决定，即在模型求解之前已经得到了数值。

6．随机方程：根据经济行为构造的函数关系式。

7．非随机方程：根据经济学理论或政策、法规而构造的经济变量恒等式。

8．时序数据：指某一经济变量在各个时期的数值按时间先后顺序排列所形成的数列。

9．截面数据：指在同一时点或时期上，不同统计单位的相同统计指标组成的数据。

四、简答题

1．简述经济计量分析的研究步骤。

用经济计量方法研究社会经济问题是以经济计量模型的建立和应用为基础的，其过程可分为四个连续的步骤：建立模型、估计参数、验证模型和使用模型。

建立模型是根据经济理论和某些假设条件，区分各种不同的经济变量，建立单一方程式或方程体系，来表明经济变量之间的相互依存关系。

模型建立后，必须对模型的参数进行估计；就是获得模型参数的具体数值。

模型估计之后，必须验证模型参数估计值在经济上是否有意义，在统计上是否令人满意。

对经济现象的计量研究是为了使用经济计量模型。经济计量模型的使用主要是用于进行经济结构分析、预测未来和制定或评价经济政策。

2．简述经济计量模型检验的三大原则。

第一，经济理论准则；第二，统计准则；第三，经济计量准则。

（1）经济理论准则

经济理论准则即根据经济理论所阐明的基本原理，以此对模型参数的符号和取值范围进行检验；就是据经济理论对经济计量模型中参数的符号和取值范围施加约束。

（2）统计准则

统计准则是由统计理论决定的，统计准则的目的在于考察所求参数估计值的统计可靠性。由于所求参数的估计值是根据经济计量模型中所含经济变量的样本观测值求得的，便可以根据数理统计学的抽样理论中的几种检验，来确定参数估计值的精确度。

（3）经济计量准则

经济计量准则是由理论经济计量学决定的，其目的在于研究任何特定情况下，所采用的经济计量方法是否违背了经济计量模型的假定。经济计量准则作为二级检验，可视为统计准则的再检验。

3．简述经济计量模型的用途。

对经济现象的计量研究是为了使用经济计量模型。经济计量模型的使用主要是用于进行经济结构分析、预测未来和制定或评价经济政策。

(1) 结构分析。就是利用已估计出参数值的模型，对所研究的经济系统变量之间的相互关系进行分析，目的在于了解和解释有关经济变量的结构构成和结构变动的原因。

(2) 预测未来。就是根据已估计出参数值的经济计量模型来推测内生变量在未来时期的数值，这是经济计量分析的主要目的之一。

(3) 规划政策。这是经济计量模型的最重要用途，也是它的最终目的。规划政策是由决策者从一系列可供选择的政策方案中，挑选出一个最优政策方案予以执行。一般的操作步骤是先据模型运算一个基本方案，然后改变外生变量（政策变量）的取值，得到其它方案，对不同的政策方案的可能后果进行评价对比，从而做出选择，因此又称政策评价或政策模拟。

第二章　　一元线性回归模型

一、单项选择题

1．回归分析的目的为（　）

A．研究解释变量对被解释变量的依赖关系

B．研究解释变量和被解释变量的相关关系

C．研究被解释变量对解释变量的依赖关系

D．以上说法都不对

2．在回归分析中，有关被解释变量 Y 和解释变量 X 的说法正确的为（　）

A．Y 为随机变量，X 为非随机变量

B．Y 为非随机变量，X 为随机变量

C．X、Y 均为随机变量

D．X、Y 均为非随机变量

3．在 X 与 Y 的相关分析中（　）

A．X 是随机变量，Y 是非随机变量

B．Y 是随机变量，X 是非随机变量

C．X 和 Y 都是随机变量

D．X 和 Y 均为非随机变量

4．总体回归线是指（　）

A．解释变量 X 取给定值时，被解释变量 Y 的样本均值的轨迹。

B．样本观测值拟合的最好的曲线。

C．使残差平方和最小的曲线。

D．解释变量 X 取给定值时，被解释变量 Y 的条件均值或期望值的轨迹。

5．随机误差项是指（　）

A．个别的 Y_i 围绕它的期望值的离差

B．Y_i 的测量误差

C．预测值 $\hat{Y}_i$ 与实际值 Y_i 的偏差

D．个别的 X_i 围绕它的期望值的离差

6．最小二乘准则是指（　）

A．随机误差项 u_i 的平方和最小

B．Y_i 与它的期望值 $\bar{Y}$ 的离差平方和最小

C．X_i 与它的均值 $\bar{X}$ 的离差平方和最小

D．残差 e_i 的平方和最小

7．按照经典假设，线性回归模型中的解释变量应为非随机变量，且（　）

A．与被解释变量Y_i不相关

B．与随机误差项u_i不相关

C．与回归值$\hat{Y}_i$不相关

D．以上说法均不对

8．有效估计量是指（　）

A．在所有线性无偏估计量中方差最大

B．在所有线性无偏估计量中变异系数最小

C．在所有线性无偏估计量中方差最小

D．在所有线性无偏估计量变异系数最大

9．在一元线性回归模型中，σ^2的无偏估计量$\hat{\sigma}^2$为（　）

A．$\dfrac{\sum e_i^2}{n}$　　B．$\dfrac{\sum e_i^2}{n-1}$

C．$\dfrac{\sum e_i^2}{n-2}$　　D．$\dfrac{\sum e_i^2}{n-3}$

10．判定系数R^2的取值范围为（　）

A．0¡ÜR^2¡Ü2　　B．0¡ÜR^2¡Ü1

C．0¡ÜR^2¡Ü4　　D．1¡ÜR^2¡Ü4

11．回归系数 $_2$通过了 t 检验，表示（　）

A．$_2$¡Ù0　　B．$\hat{}_2$¡Ù0

C．$_2$¡Ù0，$\hat{}_2$=0　　D．$_2$=0，$\hat{}_2$¡Ù0

12．个值区间预测就是给出（　）

A．预测值$\hat{Y}_0$的一个置值区间

B．实际值Y_0的一个置值区间

C．实际值Y_0的期望值的一个置值区间

D．实际值X_0的一个置值区间

13．一元线性回归模型中，$\hat{}_1$的估计是（　　）

A．$\hat{}_1=\bar{Y}+\hat{}_2\bar{X}$　　B．$\hat{}_1=\bar{Y}+\hat{}_2X$

C．$\hat{}_1=\bar{Y}-\hat{}_2\bar{X}$　　D．$\hat{}_1=Y+\hat{}_2\bar{X}$

二、多项选择题

1．对于经典线性回归模型，回归系数的普通最小二乘估计量具有的优良特性有（　）

A．无偏性　　B．线性性

C．有效性　　D．确定性

E．误差最小性

2．判定系数 R^2 可表示为（　）

A．$R^2 = \frac{RSS}{TSS}$　　B．$R^2 = \frac{ESS}{TSS}$

C．$R^2 = 1 - \frac{RSS}{TSS}$　　D．$R^2 = 1 - \frac{ESS}{TSS}$

E．$R^2 = \frac{ESS}{ESS + RSS}$

3．在经典线性回归模型中，影响 $\hat{\ }_2$ 的估计精度的因素有（　）

A．Y_i 的期望值 $E(Y_i)$　　B．Y_i 的估计值 $\hat{Y}_i$

C．Y_i 的总变异 $\sum(Y_i - \bar{Y})^2$　　D．随机误差项的方差 σ^2

E．X_i 的总变异 $\sum(X_i - \bar{X})^2$

4．对于截距项 $_1$，即使是不显著的，也可不理会，除非（　）

A．模型用于结构分析　　B．模型用于经济预测

C．模型用于政策评价　　D．$_1$ 有理论上的特别意义

E．以上说法都对

5．评价回归模型的特性，主要从如下几个方面入手（　）

A．经济理论评价　　B．统计上的显著性

C．回归模型的拟合优度　　D．回归模型是否满足经典假定

E．模型的预测精度

三、名词解释

1．回归分析

2．相关分析

3．总体回归函数

4．随机误差项

5．有效估计量

6．判定系数

四、简答题

1．简述回归分析与相关分析的关系。

2．简述随机误差项 u 的意义。

3．试述最小二乘估计原理。

4．试述经典线性回归模型的经典假定。

5．叙述高斯一马尔可夫定理，并简要说明之。

6．试述一元线性回归模型$\hat{Y}=\hat{\beta}_1+\hat{\beta}_2X$ 中影响$\hat{\beta}_2$的估计精度[$\hat{\beta}_2$的方差 $Var(\hat{\beta}_2)$]的因素。

7．简述 t 检验的决策规则。

8．如何评价回归分析模型。

五、计算题

1．以 1978~1997 年中国某地区进口总额Y（亿元）为被解释变量，以地区生产总值 X（亿元）为解释变量进行回归，得到回归结果如下：

$$\hat{Y}_t=-261.09+0.2453X_t$$

Se=(31.327)　　(　　)

t=(　　　　)　　(16.616)

R^2=0.9388　　　　　　n=20

要求：(1) 将括号内缺失的数据填入；

(2) 如何解释系数 0.2453 和系数－261.09；

(3) 检验斜率系数的显著性。(计算结果保留三位小数)

2．据 10 年的样本数据得到消费模型为

$$\hat{Y}=-231.80+0.7194X$$

Se=(0.9453)　　(0.0217)

R^2=0.9909

取显著性水平 α= 5％，查 t 分布表可知

$t_{0.025}(8)$=2.306　　$t_{0.05}$ (8) =1.860

$t_{0.025}(10)$=2.228　　$t_{0.05}$ (10) =1.812

要求：(1) 检验回系数的显著性。

(2) 给出斜率系数的 95%置信区间。(计算结果保留三位小数)

3．用 10 年的 GDP 与货币存量的数据进行回归，使用不同度量的货币存量得到如下两个模型：

模型 1：GDP_t ＝ －787.4723+8.0863M_{1t}

Se = (77.9664)　(0.2197)

模型 2：GDP_t ＝ －44.0626+1.5875M_{2t}

Se = (61.0134)　(0.0448)

已知 GDP 的样本方差为 100，模型 1 的残差平方和$\sum_{i=1}^{10}e_{1i}^2$ =100，模型 2 的残差平方和$\sum_{i=1}^{10}e_{2i}^2$ =70，请比较两回归模型，并选择一个合适的模型。(计算结果保留二位小数)

4．用 12 对观测值估计出的消费函数为Y=10.0+0.9X，且已知$\hat{\sigma}^2$=100，$\bar{X}$=200，$\sum(X-\bar{X})^2$ =4000。试预测当X_0=250 时，Y的均值Y_0的值，并求Y_0的 95%置信区间[$t_{0.025}(10)$=2.228，计算结果保留二位小数]。

参考答案

一、单项选择题

1．C　2．A　3．C　4．D　5．A　6．D　7．B

8．C　9．C　10．B　11．A　12．B　13．C

二、多项选择题

1．ABC　2．BCE　3．DE

4．BD　5．ABCD

三、名词解释

1．回归分析：就是研究被解释变量对解释变量的依赖关系，其目的就是通过解释变量的已知或设定值，去估计或预测被解释变量的总体均值。

2．相关分析：测度两个变量之间的线性关联度的分析方法。

3．总体回归函数：$E(Y/X_i)$是 X_i 的一个线性函数，就是总体回归函数，简称总体回归。它表明在给定 X_i 下 Y 的分布的总体均值与 X_i 有函数关系，就是说它给出了 Y 的均值是怎样随 X 值的变化而变化的。

4．随机误差项：为随机或非系统性成份，代表所有可能影响 Y，但又未能包括到回归模型中来的被忽略变量的代理变量。

5．有效估计量：在所有线性无偏估计量中具有最小方差的无偏估计量叫做有效估计量。

6．判定系数：$R^2=\dfrac{\sum(\hat{Y}_i-\bar{Y})^2}{\sum(Y_i-\bar{Y})^2}=\dfrac{ESS}{TSS}$，是对回归线拟合优度的度量。$R^2$ 测度了在 Y 的总变异中由回归模型解释的那个部分所占的比例或百分比。

四、简答题

1．简述回归分析与相关分析的关系。

答：相关分析主要测度两个变量之间的线性关联度，相关系数就是用来测度两个变量之间的线性关联程度的。而在回归分析中，我们的主要目的在于根据其它变量的给定值来估计或预测某一变量的平均值。例如，我们想知道能否从一个学生的数学成绩去预测他的统计学平均成绩。

在回归分析中，被解释变量 Y 被当作是随机变量，而解释变量 X 则被看作非随机变量。而在相关分析中，我们把两个变量都看作是随机变量。

2．简述随机误差项 u 的意义。

答：随机误差项 u 是代表所有对 Y 有影响但未能包括在回归模型中的那些变量的替代变量。因为受理论和实践条件的限制而必须省略一些变量，其理由如下：

（1）理论的欠缺。虽然有决定 Y 的行为的理论，但常常是不能完全确定的，理论常常有一定的含糊性。

（2）数据的欠缺。即使能确定某些变量对 Y 有显著影响，但由于不能得到这些变量的数据信息而不能引入该变量。

（3）核心变量与非核心变量。例如，在居民消费模型中，除了收入 X_1 外，家庭的人

口数 X_2、户主宗教信仰 X_3、户主受教育水平 X_4 也影响家庭消费支出。但很可能 X_2、X_3、X_4 合起来的影响也是很微弱的，是一种非系统的或随机的影响。从效果与成本角度来看，引入它们是不合算的。所以，人们把它们的联合效用当作一个随机变量来看待。

（4）人类行为的内在随机性。即使我们成功地把所有有关的变量都引进到模型中来，在个别的 Y 中仍不免有一些　内在　的随机性，无论我们花了多少力气都解释不了的。随机误差项 u_i 能很好地反映这种随机性。

（5）节省原则，我们想保持一个尽可能简单的回归模型。如果我们能用两个或三个变量就基本上解释了 Y 的行为，就没有必要引进更多的变量。让 u_i 代表所有其它变量是一种很好的选择。

3．试述最小二乘估计原理。

答：样本回归模型为：$Y_i = \hat{\ }_1 + \hat{\ }_2 X_i + e_i = \hat{Y}_i + e_i$，$e_i = Y_i - \hat{Y}_i = Y_i - \hat{\ }_1 - \hat{\ }_2 X_i$，残差 e_i 是实际值 Y_i 与其估计值 $\hat{Y}_i$ 之差。对于给定的 Y 和 X 的 n 对观测值，我们希望样本回归模型的估计值 $\hat{Y}_i$ 尽可能地靠近观测值 Y_i。为了达到此目的，我们就必须使用最小二乘准则，使：

$$\sum e_i^2 = \sum (Y_i - \hat{Y}_i)^2 = \sum (Y_i - \hat{\ }_1 - \hat{\ }_2 X_i)^2$$

尽可能地小。$\sum e_i^2 = f(\hat{\ }_1, \hat{\ }_2)$，残差平方和是估计量 $\hat{\ }_j$ 的函数，对任意给定的一组数据（样本），最小二乘估计就是选择 $\hat{\ }_1$ 和 $\hat{\ }_2$ 值，使 $\sum e_i^2$ 最小。如此求得的 $\hat{\ }_1$ 和 $\hat{\ }_2$ 就是回归模型中回归系数的最小二乘估计，这种方法就称为最小二乘法。

4．试述经典线性回归模型的经典假定。

答：对于总体线性回归模型，其经典假定如下。

假定 1：误差项 u_i 的均值为零。

假定 2：同方差性或 u_i 的方差相等。对所有给定的 X_i，u_i 的方差都是相同的。

假定 3：各个误差项之间无自相关，u_i 和 u_j（$i \neq$j）之间的相关为零。

假定 4：u_i 和 X_i 的协方差为零或 E（u_iX_i）=0

该假定表示误差项 u 和解释变量 X 是不相关的。

假定 5：正确地设定了回归模型，即在经验分析中所用的模型没有设定偏误。

假定 6：对于多元线性回归模型，没有完全的多重共线性。就是说解释变量之间没有完全的线性关系。

5．叙述高斯一马尔可夫定理，并简要说明之。

答：在给定经典线性回归模型的假定下，最小二乘估计量是最佳线性无偏估计量。

该定理说明最小二乘估计量 $\hat{}_j$ 是 $_j$ 的最佳线性无偏估计量。即

第一，它是线性的，即它是回归模型中的被解释变量 Y 的线性函数。

第二，它是无偏的，即它的均值或期望值 $E(\hat{}_j)$ 等于其真值 $_j$，即 $E(\hat{}_j)=\ _j$。

第三，它在所有这样的线性无偏估计量中具有最小方差。具有最小方差的无偏估计量叫做有效估计量。

五、计算题

1．解：

（1）Se=0.015　　　　　t=-8.342

（2）斜率参数 0.2453 表示地区生产总值增加 1 亿元，进口需求增加 0.2453 亿元。截距系数-261.09 无实际意义。

（3）斜率系数的 t 统计量为 16.616，远大于临界水平，据 t 检验应拒绝真实斜率系数为零的假设。

2．解：

（1）t 统计量分别为

$$t_{\hat{}_1}=\frac{\hat{}_1}{Se(\hat{}_1)}=\frac{-231.80}{0.9453}=-245.213$$

$$t_{\hat{}_2}=\frac{\hat{}_2}{Se(\hat{}_2)}=\frac{0.7194}{0.0217}=33.152$$

$$\left|t_{\hat{}_0}\right|=245.213>t_{0.025}(8)=2.306$$

$$\left|t_{\hat{}_1}\right|=33.152>t_{0.025}(8)=2.306$$

所以 $\hat{}_1$，$\hat{}_2$ 均为显著的。

（2）¦Â2 的置信区间为

$$\hat{}_2-t_{\alpha/2}\cdot Se(\hat{}_2)\le\ _2\le\hat{}_2+t_{\alpha/2}\cdot Se(\hat{}_2)$$

0.7194－2.306×0.0217¡Ü¦Â2¡Ü0.7194+2.306×0.0217

0.669¡Ü¦Â2¡Ü0.980

3．解：

模型 1 判定系数为

$$R_1^2=1-\frac{\sum e_{1i}^2}{\sum(Y-\bar{Y})^2}=1-\frac{100}{10\times100}=0.90$$

模型 2 的判定系数为

$$R_2^2 = 1 - \frac{\sum e_{2i}^2}{\sum (Y - \bar{Y})^2} = 1 - \frac{70}{10 \times 100} = 0.93$$

模型 1 的 t 统计量分别为 $t_{\hat{}_1} = -10.10$ $t_{\hat{}_2} = 36.81$

模型 2 的 t 统计量分别为 $t_{\hat{}_1} = -0.72$ $t_{\hat{}_2} = 35.44$

两模型的斜率系数均通过了 t 检验，说明 M_{1t} 与 M_{2t} 均与 GDP_t 有线性关系，但模型 1 的判定系数 R^2 大于模型 2 的判定系数 R^2，具有较好的拟合优度，因此应选择模型 1。

4．解：Y_0 的预测值为 Y_0

$\hat{Y}_0$=10.0+0.9×250=235.0

Y_0 的 95%的置信区间为

$$\hat{Y}_0 - t_{0.025} \cdot Se(\hat{Y}_0) \le Y_0 \le \hat{Y}_0 + t_{0.025} \cdot Se(\hat{Y}_0)$$

$$Se(\hat{Y}_0) = \hat{\sigma}\sqrt{\frac{1}{n} + \frac{(X_0 - \bar{X})^2}{\sum (X - \bar{X})^2}} = 10\sqrt{\frac{1}{12} + \frac{(250 - 100)^2}{4000}} = 8.42$$

235.0－2.228×8.42 ≤ Y_0 ≤ 235.0+2.228×8.42

216.24 ≤ Y_0 ≤ 253.76

第三章　多元线性回归模型

练习题

一、单项选择题

1.在线性回归模型 Y_i=¦Â$_1$+¦Â$_2X_{2i}$+¦Â$_3X_{3i}$+u_i 中，¦Â$_2$表示（　　　）

A．X$_{3i}$，u_i 保持不变条件下，X$_2$ 每变化一单位时，Y 的均值的变化。

B．任意情况下，X$_2$ 每变化一单位时，Y 的均值的变化。

C．X$_{3i}$ 保持不变条件下，X$_2$ 每变化一单位时，Y 的均值的变化。

D．u_i 保持不变条件下，X$_2$ 每变化一单位时，Y 的均值的变化。

2．在线性回归模型 Y_i=¦Â$_1$+¦Â$_2X_{2i}$+¦Â$_3X_{3i}$+u_i 中，¦Â$_1$的含义为（　　）

A．指所有未包含到模型中来的变量对 Y 的平均影响。

B．Y$_i$ 的平均水平。

C．X$_{2i}$，X$_{3i}$ 不变的条件下，Y$_i$ 的平均水平。

D．X$_{2i}$=0，X$_{3i}$=0 时，Y$_i$ 的真实水平。

3．在多元线性回归模型中，调整后的判定系数 $\bar{R}^2$ 与判定系数 R^2 的关系为（　　）

A．$R^2 < \bar{R}^2$　　　　　　　　B．$\bar{R}^2 < R^2$

C．R^2¡Ü$\bar{R}^2$　　　　　　　　D．$\bar{R}^2$¡ÜR^2

4．回归模型中不可使用的模型为（　　）

A．$\bar{R}^2$ 较高，回归系数高度显著；

B．$\bar{R}^2$ 较低，回归系数高度显著；

C．$\bar{R}^2$ 较高，回归系数不显著；

D．$\bar{R}^2$ 较低，回归系数显著。

5．在回归模型 Y =¦Â$_1$+¦Â$_2$X$_2$+¦Â$_3$X$_3$+¦Â$_4$ X$_4$+u 中，X$_3$ 与 X$_4$ 高度相关，X$_2$ 与 X$_3$，X$_4$ 无关，则因为 X$_3$ 与 X$_4$ 的高度相关会使 $\hat{}_2$ 的方差（　）

A．变大　　　　　　　　B．变小

C．不确定　　　　　　　D．不受影响

6．在回归模型 Y =¦Â$_1$+¦Â$_2$X$_2$+¦Â$_3$X$_3$+¦Â$_4$ X$_4$+u 中，如果原假设 H_0：¦Â$_2$= 0 成立，则意味着（　）

A．估计值 $\hat{}_2$=0　　　　　　B．X$_2$ 与 Y 无任何关系

C．回归模型不成立　　　　　　D．X$_2$ 与 Y 无线性关系

7．在对数线性模型 $LnY_i = \alpha + \quad X_i + u$ 中，　度量了（　　　）

A．X 变动 1%时，Y 变动的百分比。

B．Y 变动 1%时，X 变动的百分比。

C．X 变动一个单位时，Y 变动的数量。
D．Y 变动一个单位时，X 变动的数量。

8．在线性到对数模型，$LnY_t = {}_1 + {}_2 t + u_t$ 中，Y_t 代表国内生产总值，t 代表时间变量，则斜率系数 ¦$\hat{A}_2$ 代表（　　）

A．经济发展速度　　B．平均增长量
C．总增长量　　D．经济增长率

9．在对数到线性模型 $Y_t = {}_1 + {}_2 LnX_t + u_t$ 中，斜率系数 ¦$\hat{A}_2$ 的含义为（　　）

A．X 变动 1%时，Y 变动的数量。
B．X 变动一个单位时，Y 变动的数量。
C．X 变动 1%时，Y 变动的百分比。
D．X 变动一个单位时，Y 变动的百分比。

10．在回归模型 $Y_i = \alpha_1 + \alpha_2 X_{2i} + \alpha_3 X_{3i} + u_i$ 中，解释变量 X_3 为无关解释变量，则因为 X_3 的引入，会使 α_2 的最小二乘估计 $\hat{\alpha}_2$（　）

A．无偏、方差变大　　B．无偏、方差不变
C．有偏、方差变大　　D．有偏、方差不变

11.真实的回归模型为 $Y_i = \alpha_1 + \alpha_2 X_{2i} + \alpha_3 X_{3i} + u_i$，但是在回归分析时使用的模型为 $Y_i = \alpha_1 + \alpha_2 X_{2i} + v_i$，漏掉了重要解释变量 X_3，则会使 α_2 的最小二乘估计 $\hat{\alpha}_2$（　　）

A．X_3 与 X_2 相关时有偏　　B．X_3 与 X_2 相关时无偏
C．无偏　　D．有偏

12．对于倒数模型 $Y_t =$ ¦$\hat{A}$+¦$\hat{A}\frac{1}{X_t} + u_t$，当 ¦$\hat{A}$>0，¦$\hat{A}_2$>0 时，可用来描述（ ）

A．增长曲线　　B．菲利普斯曲线
C．恩格尔支出曲线　　D．平均总成本曲线

13．根据判定系数 R^2 与 F 统计量的关系可知，当 R^2=1 时，有（ ）

A．F=1　　B．F=-1
C．F=0　　D．$F=\infty$

14．根据样本资料估计得到人均消费支出 Y 对人均收入 X 的回归模型为 $Ln\hat{Y}_i = 1.00 + 0.75LnX_i$，这表明人均收入每增加 1%，人均消费支出将增加（　　）

A．2%　　B．0.2%
C．0.75%　　D．7.5%

15．对回归系数进行显著性检验时的 t 统计量为（　　）

A．$\dfrac{{}_j}{Se(\hat{}_j)}$　　B．$\dfrac{\hat{}_j}{Var(\hat{}_j)}$

C. $\frac{{}_{j}}{Var(\hat{\ }_{j})}$　　　　　　　　D. $\frac{\hat{\ }_{j}}{Se(\hat{\ }_{j})}$

二、多项选择题

1．多元回归模型 Y_i =¦$\hat{A}_1$+¦$\hat{A}_2$ X_{2i}+¦$\hat{A}_3$ X_{3i}+u_i 通过了整体显著性 F 检验，则可能的情况为（　　）

A．¦$\hat{A}_2$= 0，¦$\hat{A}_3$= 0　　　　B．¦$\hat{A}_2$¡Ù，¦$\hat{A}_3$¡Ù

C．¦$\hat{A}_2$= 0，¦$\hat{A}_3$¡Ù　　　　D．¦$\hat{A}_2$¡Ù，¦$\hat{A}_3$= 0

E．¦$\hat{A}_2$=¦$\hat{A}_3$¡Ù

2．对回归模型进行显著性检验时所用的 F 统计量可表示为（　　）

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

C. $\frac{R^2/(k-1)}{(1-R^2)/(n-k)}$　　　　D. $\frac{(1-R^2)/(n-k)}{R^2(k-1)}$

E. $\frac{R^2/(n-k)}{(1-R^2)/(k-1)}$

3．有关对变量取对数的经验法则下列说法正确的为（　　）

A．对于大于 0 的数量变量，通常均可取对数；

B．以年度量的变量，如年龄等以其原有形式出现；

C．比例或百分比数，可使用原形式也可使用对数形式；

D．使用对数时，变量不能取负值；

E．数值较大时取对数形式。

4．真实模型为 Y_i =¦$\hat{A}_1$+¦$\hat{A}_2$ X_{2i}+¦$\hat{A}_3$ X_{3i}+u_i 时，如果使用模型 Yi = $\alpha_1+\alpha_2 X_{2i}+u_i$ 中，则遗漏了重要解释变量 X_3，此时对参数的最小二乘估计有较大影响，下列说法正确的为（　　）

A．如果 X_3 与 X_2 相关，则 $\hat{\alpha}_1$ 与 $\hat{\alpha}_2$ 是有偏、非一致的；

B．如果 X_3 与 X_2 不相关，则 $\hat{\alpha}_1$ 与 $\hat{\alpha}_2$ 是有偏、非一致的；

C．如果 X_3 与 X_2 不相关，则 $\hat{\alpha}_2$ 是无偏的；

D．如果 X_3 与 X_2 相关，则 $\hat{\alpha}_2$ 是有偏、一致的。

E．如果 X_3 与 X_2 不相关，则 $\hat{\alpha}_2$ 是有偏、一致的。

三、名词解释

1．多元线性回归模型

2．调整的判定系数

3．对数线性模型

四、简答题

1．多元回归分析中为何要使用调整的判定系数。

2．多元经典回归模型中，影响偏回归系数 ¦Â$_j$的最小二乘估计量 $\hat{}_j$ 方差的因素有哪些？

3．简述高斯—马尔可夫定理及其意义；

４．简述多元回归模型的整体显著性检验决策规则。

5．对于多元线性回归模型，为什么在进行了总体显著性 F 检验之后，还要对每个偏回归系数进行是否为 0 的 t 检验。

６．对数线性模型的优点有哪些？

７．什么是回归模型的设定偏误？简要说明其后果。

五、计算题

1．使用 30 年的年度数据样本，得到某地区生产函数模型回归结果如下：

$$LnY = 1.655 + 0.358\mathrm{LnL} + 0.745\mathrm{LnK}$$

(0.185)　(0.125)　　(0.095)

$R^2 = 0.955$

其中，Y=地区生产总值（亿元），L=劳动投入（亿元），K=资本存量（亿元）。(计算结果保留三位小数)。

要求：(1) 检验各回归系数的显著性；

(2) 检验回归模型的整体显著性；[$\alpha = 0.05$，$F_{0.05}(2,27)=3.42$，$F_{0.05}(3,30)=2.92$]

(3) 利用回归结果分析该地区的投入产出状况。

2．对二个解释变量的回归模型 Y$_t$=¦Â$_1$+¦Â$_2$X$_{2t}$+¦Â$_3$X$_{3t}$+u_t，使用 20 年的年度样本数据进行回归，解释平方和 ESS=64.50，总平方和 TSS=66.30。(计算结果保留二位小数)

要求：(1) 求出该回归模型的判定系数 R^2 和 $\bar{R}^2$；

(2) 对该回归模型进行整体显著性检验。

[$\alpha = 0.05$，$F_{0.05}(2,17)=3.59$，$F_{0.05}(3,20)=3.10$]

3．据 1950—1969 年的年度数据得到某国的劳动力薪金模型

$\hat{W}_t$=8.582+0.364$(PF)_t$+0.004$(PF)_{t-1}$－2.560U$_t$

(1.129) (0.080)　(0.072)　　(0.658)

R^2=0.873

其中，W=劳动力平均薪金，PF=生产成本，U=失业率（%）。(计算结果保留三位小数)

要求：(1) 对模型回归系数进行显著性检验。[$\alpha = 0.05$，$t_{0.025}(16)=2.12$]

(2) 引进变量$(PF)_{t-1}$合理吗？

(3) 如要估计劳动力薪金对失业率的弹性应如何处理？

六、分析题

1．设定某商品的需求模型为

Y=¦Â$_0$+¦Â$_1$X$_1$+¦Â$_2$X$_2$+¦Â$_3$X$_3$+¦Â$_4$X$_4$+ u

其中，Y=商品销售量，X$_1$=居民可支配收入，X$_2$=该商品的价格指数，X$_3$=该商品的社会拥有量，X$_4$=其它商品价格指数。搜集到 10 个年份的年度数据，得到如下两个样本回归模型：

模型 1：$\hat{Y} = -12.76 + 0.104X_1 - 0.188X_2 + 0.319X_4$

(6.52)　(0.01)　　(0.07)　　(0.12)

R^2=0.997

模型 2：$\hat{Y} = -13.53 + 0.097X_1 - 0.199X_2 + 0.015X_3 + 0.34X_4$

(7.5) (0.03) (0.09) (0.05) (0.15)

R^2=0.998

模型式下括号中的数字为相应回归系数估计量的标准误。

试对所给出的两个模型进行检验，并选择出一个合适的模型。[$\alpha = 0.05$，$t_{0.025}(5)=2.57$，$t_{0.025}(6)=2.45$；$F_{0.05}(3,6)=4.76$，$F_{0.05}(4,5)=5.19$]（计算结果保留二位小数）

2．据 20 年的年度样本资料，得到如下的劳动力薪金模型

$$\hat{W}_t = 1.073 + 5.288V_t - 0.116X_t + 0.54M_t + 0.046M_{t-1}$$

(0.797) (0.812) (0.111) (0.022) (0.019)

R^2=0.934

其中，W_t=劳动力人均薪金，V=岗位空缺率，X=就业人员人均国内生产总值，M_t=出口额，M_{t-1}=上年出口额（括号内的数字为标准误，计算结果保留三位小数）。

要求：(1) 引进变量 X 的原理为何？理论上，X 的系数符号应为正还是负。

(2) 哪些变量可从模型中删去。[$t_{0.025}(15)=2.131$]

(3) 检验回归模型的总显著性，[$F_{0.05}(4,15)=3.06$]

3．经济学家提出假设，能源价格上升导致资本产出率下降。据 30 年的季度数据，得到如下回归模型：

$$Ln(Y/K) = 1.5492 + 0.7135Ln(L/K) - 0.1081LnP + 0.0045t$$

(16.35) (21.69) (-6.42) (15.86)

R^2=0.98

其中，Y=产出，K=资本流量，L=劳动投入，P_t=能源价格，t=时间。括号内的数字为 t 统计量。(计算结果保留二位小数)

问：(1) 回归分析的结果是否支持经济学家的假设；

(2) 如果在样本期内价格 P 增加 60%，据回归结果，资本产出率下降了多少？

(3) 除了（L/K）和 P 的影响，样本期内的资本产出率趋势增长率如何？

(4) 如何解释系数 0.7135？

参考答案

一、单项选择题

1．C 2．A 3．B 4．C 5．D 6．B 7．A 8．D

9．A 10．A 11．A 12．D 13．D 14．C 15．D

二、多项选择题

1．BCD 2．BC 3．ABCD 4．AC

三、名词解释

1．多元线性回归模型：在模型中包含二个以上的解释变量的线性回归模型。

2．调整的判定系数：$\bar{R}^2 = 1 - \dfrac{\sum e_i^2/(n-k)}{\sum(Y_i - \bar{Y})^2/(n-1)}$，所谓调整，就是指 $\bar{R}^2$ 的计算式中的 $\sum e_i^2$ 和 $\sum(Y_i - \bar{Y})^2$ 都用它们的自由度（$n-k$）和（$n-1$）去除。

3．对数线性模型：$LnY_i = \alpha + \beta LnX_i + u_i$，该模型中 LnY_i 对 α，β 是线性关系，LnY_i 对 LnX_i 也是线性关系。该模型可称为对数—对数线性模型，简称为对数线性模型。

四、简答题

1．多元回归分析中为何要使用调整的判定系数。

答：判定系数 R^2 的一个重要性质是：在回归模型中增加一个解释变量后，它不会减少，而且通常会增大。即 R^2 是回归模型中解释变量个数的非减函数。所以，使用 R^2 来判断具有相同被解释变量 Y 和不同个数解释变量 X 的回归模型的优劣时就很不适当。此时，R^2 不能用于比较两个回归方程的拟合优度。

为了消除解释变量个数对判定系数 R^2 的影响，需使用调整后的判定系数：

$\bar{R}^2 = 1 - \dfrac{\sum e_i^2/(n-k)}{\sum (Y_i - \bar{Y})^2/(n-1)}$，所谓调整，就是指 $\bar{R}^2$ 的计算式中的 $\sum e_i^2$ 和 $\sum (Y_i - \bar{Y})^2$ 都用它们的自由度（$n-k$）和（$n-1$）去除。

2．多元经典回归模型中，影响偏回归系数 β_j 的最小二乘估计量 $\hat{\beta}_j$ 方差的因素有哪些？

答：$\hat{\beta}_j$ 的方差取决于如下三个因素：$\sigma^2, SST_j 和 R_j^2$。

（1）$Var(\hat{\beta}_j)$ 与 σ^2 成正比；σ^2 越大，$\hat{\beta}_j$ 的方差 $Var(\hat{\beta}_j)$ 越大。回归模型的干扰项 u 是对回归结果的干扰，干扰(σ^2)越大，使得估计任何一个解释变量对 Y 的局部影响就越困难。

（2）$Var(\hat{\beta}_j)$ 与 X_j 的总样本变异 SST_j 成反比；总样本变异 SST_j 越大，$\hat{\beta}_j$ 的方差 $Var(\hat{\beta}_j)$ 越小。

（3）$Var(\hat{\beta}_j)$ 与解释变量之间的线性关联程度 R_j^2 正相关；R_j^2 越大，$\hat{\beta}_j$ 的方差 $Var(\hat{\beta}_j)$ 越大。

3．简述高斯—马尔可夫定理及其意义。

答：在多元线性回归模型的经典假定下，普通最小二乘估计量 $\hat{\beta}_1$，$\hat{\beta}_2$，…，$\hat{\beta}_k$ 分别是 β_1，β_2，…，β_k 的最佳线性无偏估计量。就是说，普通最小二乘估计量 $\hat{\beta}_1$，$\hat{\beta}_2$，…，$\hat{\beta}_k$ 是所有线性无偏估计量中方差最小的。

高斯－马尔可夫定理的意义在于：当经典假定成立时，我们不需要再去寻找其它无偏估计量，没有一个会优于普通最小二乘估计量。也就是说，如果存在一个好的线性无偏估计

量，这个估计量的方差最多与普通最小二乘估计量的方差一样小，不会小于普通最小二乘估计量的方差。

4．简述多元回归模型的整体显著性检验决策规则。

答：(1) 设定假设

原假设 H_0： ${}_2 = {}_3 = \mathrm{L} = {}_k = 0$

备择假设 H_1： ${}_j$ 不全为 0，j=2, 3, ，k

(2) 计算 F 统计量

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

(3) 在给定显著性水平 α 的条件下，查 F 分布表得临界值 $F_\alpha(k-1, n-k)$。

(4) 判断

如果 $F > F_\alpha(k-1, n-k)$，则拒绝 H_0，接受备择假设 H_1。

如果 $F \le F_\alpha(k-1, n-k)$，则不拒绝 H_0。

5．对于多元线性回归模型，为什么在进行了总体显著性 F 检验之后，还要对每个偏回归系数进行是否为 0 的 t 检验。

答：多元回归模型的总体显著性就是对原假设 H_0： ${}_2 = {}_3 = \mathrm{L} = {}_k = 0$ 进行检验。检验的目的就是判断被解释变量 Y 是否与 $X_2, X_3,$ ，X_k 在整体上有线性关系。若原假设 H_0： ${}_2 = {}_3 = \mathrm{L} = {}_k = 0$ 被拒绝，即通过了 F 检验，则表明 Y 与 $X_2, X_3,$ ，X_k 在整体上有线性关系。但这并不表明每一个 X 都对 Y 有显著的线性影响，还需要通过 t 检验判断每一个回归系数的显著性。

6．对数线性模型的优点有哪些？

答：对数线性模型的优点为

(1) 对数线性模型中斜率系数度量了一个变量（Y）对另一个变量（X）的弹性。

(2) 斜率系数与变量 X，Y 的测量单位无关，其结果值与 X，Y 的测量单位无关。

(3) 当 $Y > 0$ 时，使用对数形式 LnY 比使用水平值 Y 作为被解释变量的模型更接近经典线性模型。大于零的变量，其条件分布常常是有异方差性或偏态性；取对数后，虽然不能消除这两方面的问题，但可大大弱化这两方面的问题。

(4) 取对数后会缩小变量的取值范围。使得估计值对被解释变量或解释变量的异常值不会很敏感。

7．什么是回归模型的设定偏误？简要说明其后果。

答：多元回归模型的设定偏误主要包括以下三种：

（1）回归模型中包含了无关解释变量

（2）回归模型中遗漏了重要解释变量

（3）回归模型中的函数形式设定偏误

后果为：（1）回归模型中包含了无关解释变量：回归系数的最小二乘估计量的方差非最小。

（2）回归模型中遗漏了重要解释变量：如果遗漏的变量与包含的变量相关，则回归系数的最小二乘估计量是有偏误的，且非一致。

（3）回归模型中的函数形式设定偏误：不能得到有效估计和正确的经济解释。

五、计算题

1．解：（1）t 统计量分别为

$t_{\hat{1}}=8.946$ $t_{\hat{2}}=2.864$ $t_{\hat{3}}=7.842$

t 统计量的绝对值均大于 2，样本量为 30，据简便准则，各回归系数均通过了检验。

（2） $F=\dfrac{\mathrm{R}^2/(k-1)}{(1-R^2)/(n-k)}=\dfrac{0.955^2/(3-1)}{(1-0.955^2)/(30-3)}$

=139.953

$F>F_{0.05}(2,27)=3.42$

回归模型整体显著

（3）该地区劳动力投入每增加 1%，产出将增加 0.358%，资本存量每增加 1%，产出将增加 0.745%

2．解：（1） $R^2=\dfrac{ESS}{TSS}=\dfrac{64.50}{66.30}=0.973$

（2） $F=\dfrac{ESS/(k-1)}{(TSS-ESS)/(n-k)}$

$=\dfrac{64.50/(3-1)}{(66.30-64.50)/(20-3)}$

=304.583

$F=304.583>F_{0.05}(2,17)=3.59$

所以回归模型整体显著

3．解：（1） $t_{\hat{1}}=7.601$ $t_{\hat{2}}=4.55$ $t_{\hat{3}}=0.056$ $t_{\hat{4}}=-3.891$

$(PF)_{t-1}$ 的回归系数不显著，其它回归系数的｜t｜均大于 2.12 ，其它回归系数均显著。

（2）理论上上期成本对本期工资水平有影响，但回归结果表明偏回归系数不显著，

引入该变量不合理。

（3）应将模型设定为对数一对数线性模型。

六、分析题

1．解：（1） $F=\dfrac{R^2/(k-1)}{(1-R^2)/(n-k)}$

模型 1：$F_1=\dfrac{0.997^2/(4-1)}{(1-0.997^2)/(10-4)}=331.834$

模型 2：$F_2=\dfrac{0.998^2/(4-1)}{(1-0.998^2)/(10-4)}=498.501$

F_1=331.834>$F_{0.05}$(3,6)=4.76

F_2=498.501>$F_{0.05}$(4,5)=5.19，所以两模型整体上都显著。

（2）各回归系数检验

$$t_{\hat{\beta}_j}=\frac{\hat{\beta}_j}{Se(\hat{\beta}_j)}$$

模型 1：$t_{\hat{\beta}_0}=\dfrac{-12.76}{6.52}=-1.96$　　$t_{\hat{\beta}_1}=\dfrac{0.104}{0.01}=10.40$

$t_{\hat{\beta}_2}=\dfrac{-0.188}{0.07}=-2.69$　　$t_{\hat{\beta}_4}=\dfrac{0.319}{0.12}=2.66$

$|t_{\hat{\beta}_0}|$=1.96<$t_{0.025}$(6)=2.45，常数项未通过 t 检验，但这并不影响回归模型的使用，其它偏回归系数均显著。

模型 2：$t_{\hat{\beta}_0}=\dfrac{-13.53}{7.5}=1.80$　　$t_{\hat{\beta}_1}=\dfrac{0.097}{0.03}=3.23$

$t_{\hat{\beta}_2}=\dfrac{-0.199}{0.09}=-2.21$　　$t_{\hat{\beta}_3}=\dfrac{0.015}{0.05}=0.30$

$t_{\hat{\beta}_4}=\dfrac{0.34}{0.15}=2.27$

临界值为 $t_{0.05}$(5)=2.57，常数项，X_2, X_3, X_4 的系数均不显著。如降低置信水平如α =10%，X_2, X_4 的系数可能显著，但 X_3 的系数则无法显著，即 X_3 与 Y 无线性关系，并且 X_3 的系数符号也不合理，因此，X_3 不应包含在模型中，应选用模型 1。

2．解：（1）理论上分析人均产值越高，人均薪金就越高，因此 X 是影响W的因素，且回归系数的符号为正。

（2）V, X, M, M_{t-1} 的 t 统计量分别为：6.512，-1.045，24.545，2.421，除 X 的系数外，t 统计量绝对值均大于临界值 $t_{0.025}$(15)=2.131。因为 X 的系数不显著，W 与 X 无线性关系，X 应从模型中删去。

（3） $F=\dfrac{R^2/(k-1)}{(1-R^2)/(n-k)}=\dfrac{0.934^2/(5-1)}{(1-0./974^2)/(20-5)}=25.629$

F>$F_{0.05}$(4,15)=3.06，回归模型整体显著。

3．解：(1）该回归模型支持了假设，因为价格 P 的回归系数符号为负，说明价格每提高 1%，资本产出率将下降 0.1081%。

(2）资本产出率的下降幅度为

$0.1081\%\times60=6.486\%$

(3）时间变量 t 的回归系数代表增长率，资本产出率趋势增长率为 0.45%。

(4）系数 0.7135 表示每单位资本的劳动力投入增加 1%，资本产出率增加 0.7135%。

第四章　违背经典假定的回归模型

练习题

一、单项选择题

1．下列哪种情况说明存在异方差（　）

A．$E(u_i)=0$　　B．$E(u_i u_j)=0 \quad i\neq j$

C．$E(u_i^2)=\sigma^2$（常数）　　D．$E(u_i^2)=\sigma_i^2$

2．当存在异方差时，使用普通最小二乘法得到的估计量是（　）

A．有偏估计量　B．有效估计量　C．无偏估计量　D．渐近有效估计量

３．下列哪种方法不是检验异方差的方法（　）

A．残差图分析法　B．等级相关系数法　C．样本分段比检验　D．DW 检验法

4．如果 $|e_i|$ 与 $\sqrt{X_i}$ 之间存在线性关系，则认为异方差形式为（　）

A．$\sigma_i^2=\sigma^2X_i$　B．$\sigma_i^2=\sigma^2X_i^2$　C．$\sigma_i^2=\sigma^2$　D．$\sigma_i^2=\sigma^2\sqrt{X_i}$

5．如果 $|e_i|$ 与 X_i 之间存在线性关系，则认为异方差的形式为（　）

A．$\sigma_i^2=\sigma^2X_i$　B．$\sigma_i^2=\sigma^2X_i^2$　C．$\sigma_i^2=\sigma^2$　D．$\sigma_i^2=\sigma^2\sqrt{X_i}$

6．如果普通最小二乘法估计残差 e_i 与 X_i 有显著的形式为 $|e_i|=0.2875X_i+v_i$ 的关系，则加权最小二乘法估计模型参数时，权数应为（　）

A．X_i　B．$\dfrac{1}{X_i^2}$　C．$\dfrac{1}{X_i}$　D．$\dfrac{1}{\sqrt{X_i}}$

7．戈德菲尔德—匡特检验适用于检验（　）

A．序列相关　B．异方差　C．多重共线性　D．设定误差

8．异方差情形下，常用的估计方法是（　　）

A．一阶差分法　　B．广义差分法　　C．工具变量法　　D．加权最小二乘法

9．下列哪种情况属于存在序列相关（　　）

A．$Cov(u_i, u_j) = 0$，$i \neq j$　　B．$Cov(u_i, u_j) \neq 0$，$i \neq j$

C．$Cov(u_i, u_j) = \sigma^2$，$i = j$　　D．$Cov(u_i, u_j) = \sigma_i^2$，$i = j$

10．当一个线性回归模型的随机误差项存在序列相关时，直接用普通最小二乘法估计参数，则参数估计量为（　　）

A．有偏估计量　　B．有效估计量　　C．无效估计量　　D．渐近有效估计量

11．下列哪种方法不是检验序列相关的方法（　　）

A．残差图分析法　　B．自相关系数法　　C．方差扩大因子法　　D．DW 检验法

12．DW 检验适用于检验（　　）

A．异方差　　B．序列相关　　C．多重共线性　　D．设定误差

13．若计算的 DW 统计量为 2，则表明该模型（　　）

A．不存在一阶序列相关　　B．存在一阶正序列相关

C．存在一阶负序列相关　　D．存在高阶序列相关

14．如果模型 $Y_t = \ {}_0 + \ {}_1 X_t + u_t$ 存在序列相关则（　　）

A．$Cov(X_t, u_t) = 0$　　B．$Cov(u_t, u_s) = 0$　　$(t \neq s)$

C．$Cov(X_t, u_t) \neq 0$　　D．$Cov(u_t, u_s) \neq 0$　　$(t \neq s)$

15．DW 检验的原假设为（　　）

A．DW=0　　B．$\ = 0$　　C．DW=1　　D．$\ =1$

16．DW 统计量的取值范围是（　　）

A．$-1 \leq DW \leq 0$　　B．$-1 \leq DW \leq 1$

C．$-2 \leq DW \leq 2$　　D．$0 \leq DW \leq 4$

17．根据 20 个观测值估计的一元线性回归模型的 DW=2.3，在样本容量 n=20，解释变量个数 k=1，显著性水平 $\alpha = 0.05$ 时，查得 $d_L = 1.201$，$d_U = 1.411$，则可以判断该模型（　　）

A．不存在一阶自相关　　B．有正的一阶自相关

C．有负的一阶自相关　　D．无法确定

18．当模型存在一阶自相关情况下，常用的估计方法是（　　）

A．加权最小二乘法　　B．广义差分法

C．工具变量法　　D．普通最小二乘法

19．采用一阶差分法估计一阶自相关模型，适合于（　）

A．　≈ 0　　B．　≈ 1　　C．$-1<\quad<0$　　D．$0<\quad<1$

20．在多元线性回归模型中，若某个解释变量对其余解释变量的判定系数接近 1，则表明模型中存在（　）

A．异方差　　B．自相关　　C．多重共线性　　D．设定误差

21．在线性回归模型中，若解释变量 X_1 和 X_2 的观测值成比例，即有 $X_{1i}=kX_{2i}$，其中 k 为非零常数，则表明模型中存在（　）

A．异方差　　B．多重共线性　　C．序列相关　　D．设定误差

22．经验认为，某个解释变量与其他解释变量间多重共线性很严重的判别标准是这个解释变量的方差扩大因子 VIF（　）

A．大于 1　　B．小于 1　　C．大于 10　　D．小于 5

23．若查表得到 d_L 和 d_U，则不存在序列相关的区间为（　）

A．$0\le DW\le d_U$　　B．$d_U\le DW\le 4-d_U$

C．$4-d_U\le DW\le 4-d_L$　　D．$4-d_U\le DW\le 4$

二、多项选择题

1．常用的检验异方差的方法有（　）

A．残差图分析法　　B．等级相关系数法

C．戈德菲尔德—匡特检验　　D．戈里瑟检验

E．怀特检验

2．存在异方差条件下普通最小二乘法具有如下性质（　）

A．线性性　　B．无偏性　　C．最小方差性

D．有偏性　　E．无效性

3．异方差情况下将导致（　）

A．参数估计量是无偏的，但不是最小方差无偏估计　　B．参数显著性检验失效

C．模型预测失效　　D、参数估计量是有偏的，且方差不是最小的

E．模型预测有效

4．当模型存在异方差时，加权最小二乘估计量具有（　）

A．线性性　　B．无偏性　　C．有效性

D．一致性　　E．不是最小方差无偏估计量

5．下列经济计量分析回归模型中哪些可能存在异方差问题（　）

A．用横截面数据建立的家庭消费支出对家庭收入水平的回归模型

B．用横截面数据建立的产出对劳动和资本的回归模型

C．以 20 年资料建立的某种商品的市场供需模型

D．以 20 年资料建立的总支出对总收入的回归模型

E．按照 差错 学习 模式建立的打错数对打字小时数的回归模型

6．以下关于 DW 检验的说法，不正确的有（　　）

A．要求样本容量较大　　B．$-1 \le DW \le 1$

C．可用于检验高阶序列相关　　D．能够判定所有情况

E．只适合一阶线性序列相关

7．序列相关情况下，常用的参数估计方法有（　　）

A．一阶差分法　　B．广义差分法　　C．工具变量法

D．加权最小二乘法　　E．广义最小二乘法

8．若查表得到 DW 上、下限 d_L 和 d_U，则 DW 检验的不确定区间为（　　）

A．$d_U \le DW \le 4-d_U$　　B．$4-d_U \le DW \le 4-d_L$

C．$d_L \le DW \le d_U$　　D．$4-d_L \le DW \le 4$

E．$0 \le DW \le d_L$

9．DW 检验不适用于下列情况下的序列相关检验（　　）

A．随机误差项具有高阶序列相关　　B．样本容量太小

C．含有滞后被解释变量的模型　　D．正的一阶线性自相关形式

E．负的一阶线性自相关形式

10．自相关情况下将导致（　　）

A．参数估计量不再是最小方差线性无偏估计量

B．均方差 MSE 可能严重低估误差项的方差

C．常用的 F 检验和 t 检验失效

D．参数估计量是无偏的

E．利用回归模型进行预测的结果会存在较大的误差

三、名词解释

1．异方差　　2．序列相关

3．多重共线性　　4．方差扩大因子

5．加权最小二乘法　　6．广义差分法

四、简答题

1．举例说明异方差的概念。

2．简述存在异方差时普通最小二乘估计存在的问题。

3．简述样本分段比检验法的应用步骤。

4．简述等级相关系数法的检验步骤。

5．产生序列相关的原因有哪些？

6．序列相关性带来哪些后果？

7．简述 DW 检验及其决策规则。

8．简述 DW 检验的局限性。

9．存在严重共线性时，估计参数产生的后果有哪些？

10．多重共线性直观判定法包括哪些主要方法？

11．多重共线性补救方法有哪几种？

五、计算题

1．现有 X 和 Y 的样本观测值如下：

X	1	2	4	5	10
Y	2	4	2	10	5

假设 Y 对 X 的回归方程为 $Y_i = {}_0 + {}_1X_i + u_i$，且 $Var(u_i) = \sigma^2 X_i^2$，试用适当方法估计此回归方程。(计算结果保留两位小数)

2．10 家自行车厂的自行车定价与评定质量的名次比较如下：

工厂	1	2	3	4	5	6	7	8	9	10
名次	6	9	2	8	5	1	7	4	3	10
价格	480	395	575	550	510	545	400	465	420	370

试计算质量名次与价格的等级相关系数。

3．设 $Y_t = {}_1 + {}_2X_{2t} + \mathrm{L} + {}_kX_{kt} + u_t$

如果 $u_t = {}_1u_{t-1} + {}_2u_{t-2} + \mathrm{L} + {}_pu_{t-p} + v_t$

v_t 满足经典假设，试写出广义差分模型。

4．有一个参数个数 k=4 的线性回归模型，用一个容量为 n=20 个时序数据样本进行普通最小二乘估计，得到如下资料：$\sum_{t=1}^{n} e_t^2 = 40$，$\sum_{t=2}^{n} e_t^2 = 39$，$\sum_{t=2}^{n} e_{t-1}^2 = 36$,

$$\sum_{t=2}^{n} e_t e_{t-1} = 20$$

试根据这些资料计算 DW 统计量。

六、分析题

1．在研究生产函数时，我们得到如下两个模型

模型 I ：$Ln\hat{\theta} = -5.040 + 0.887LnK + 0.893LnL$

$$\begin{array}{llll} Se & (1.400) & (0.087) & (0.137) \end{array}$$

$R^2 = 0.878 \quad n = 21$

模型 II ：$Ln\hat{\theta} = -8.570 + 0.027t + 0.460LnK + 1.285LnL$

$$\begin{array}{lllll} Se & (2.990) & (0.021) & (0.333) & (0.324) \end{array}$$

$R^2 = 0.889 \quad n = 21$

其中，θ =产量，K=资本，L=劳动时数，t=时间，n=样本容量，模型下括号内为参数估计的标准误。[$\alpha = 0.05$， $t_{\alpha/2}(18) = 2.101$， $t_{\alpha/2}(17) = 2.110$]

请回答以下问题：

(1)说明模型 I 中所有系数在统计上都是显著的。

(2)说明模型 II 哪些参数的系数在统计上是不显著的。

(3)若 t和 LnK之间相关系数为 0.97, 你将从中得出什么结论？

(4)模型 I 的规模报酬为多少？

2. 设有柯布 -道格拉斯生产函数，其对数线性形式为

$$LnY = LnA + \alpha LnL + \quad LnK + u$$

其中, Y=国内生产总值, L=劳动力投入，K=资本投入。

时间序列数据中劳动投入 L 和资本投入 K 有很高的相关性，存在较严重多重共线性。如果有已知信息判断该经济系统为规模报酬不变，如何修改上述模型来消除多重共线性。

参考答案

一、单项选择题

1．D　2．C　3．D　4． A　5．B　6．C　7．B　8．D　9． B　10．C

11．C　12．B　13．A　14．D　15．B　16．D　17．A　18．B　19．B　20．C

21． B　22． C　23． B

二、多项选择题

1． ABCDE　2． ABE　3． ABC　4． ABCD　5． ABE　6． BCD　7． ABE

8． BC　9． ABC　10． ABCDE

三、名词解释

1．异方差 ：在回归模型中，随机误差项 u_1， u_2， ， u_n 不具有相同的方差，即

$Var(u_i) \neq Var(u_j)$，当$i \neq j$时 ，则称随机误差的方差为异方差 。

2.序列相关 ：在进行回归分析时，如果不同观测点的误差项之间相关，即$Cov(u_i, u_j) \neq 0$，$i \neq j$，则称随机误差项之间存在着序列相关现象，也称为自相关。

3．多重共线性 ：在多元线性回归模型中，解释变量$X_1, X_2, \cdots, X_k$之间存在完全或近似的线性关系，称解释变量$X_1, X_2, \cdots, X_k$之间存在完全或近似多重共性线。也称为复共线性。

4. 方差扩大因子 ：$\dfrac{1}{1-R_j^2}$度量了由于X_j与其它解释变量之间的线性关联程度对估计量$\hat{\ }_j$的方差的影响，称其为方差扩大因子，定义为 $VIF_j = \dfrac{1}{1-R_j^2}$。

5．加权最小二乘法 ：为了克服方差非齐性，所采用的方法即加权最小二乘法。就是通过对原来的模型进行加权变换，使经过变换的模型具有同方差的随机误差项，然后再应用普通最小二乘法进行参数估计。

6．广义差分法：广义差分法可以克服所有类型的序列相关带来的问题。如果

$$Y_t = {}_1 + {}_2 X_{2t} + \cdots + {}_k X_{kt} + u_t$$

$$u_t = {}_1 u_{t-1} + {}_2 u_{t-2} + \cdots + {}_p u_{t-p} + v_t$$

v_t为经典误差项，则可以将模型变换为

$$\begin{aligned}
& Y_t - {}_1 Y_{t-1} - {}_2 Y_{t-2} - \cdots - {}_p Y_{t-p} \\
= & {}_1(1 - {}_1 - {}_2 - \cdots - {}_p) \\
+ & {}_2(X_{2t} - {}_1 X_{2,t-1} - {}_2 X_{2,t-2} - \cdots - {}_p X_{2,t-p}) + \cdots \\
+ & {}_k(X_{kt} - {}_1 X_{k,t-1} - {}_2 X_{k,t-2} - \cdots - {}_p X_{k,t-p}) + v_t
\end{aligned}$$

此模型即为广义差分模型，该模型不存在序列相关问题。采用普通最小二乘法估计该模型得到的参数估计量，即为原模型参数的无偏、有效的估计量。

四、简答题

1．举例说明异方差的概念。

答：在回归模型中，随机误差项u_1，u_2， ，u_n不具有相同的方差，即

$$Var(u_i) \neq Var(u_j)\text{，当} i \neq j \text{时}$$

在线性模型的基本假定中，u_i关于方差不变的假定不成立，则称存在异方差性。

例如：在研究城镇居民收入与消费的关系时，居民收入与消费水平有着密切的关系。用 X_i 表示第 i 户的收入，Y_i 表示第 i 户的消费额，那么反映收入与消费之间的模型为

$$Y_i = \quad_1 + \quad_2 X_i + u_i\,,\quad i = 1,2,\cdots,n$$

模型中，因为各户的收入不同，消费观念和习惯的差异，导致消费的差异非常大，模型中存在明显的异方差性。

2．简述存在异方差时普通最小二乘估计存在的问题。

答：模型中存在异方差时，如果采用普通最小二乘法估计，存在以下问题：

(1)参数估计量虽是无偏的，但不是最小方差线性无偏估计。

(2)参数的显著性检验失效。

3．简述样本分段比检验法的应用步骤。

答：样本分段比检验也叫戈德菲尔德－匡特检验，步骤是：

(1)将样本按某个解释变量的大小顺序排列，并将样本从中间截成两段。

(2)各段分别用普通最小二乘法拟合回归模型。令第一段为高方差段，第二段为低方差段，并记两段的样本容量分别为 n_1 和 n_2，模型参数个数为 k，两段样本回归残差分别为 e_{1i} 和 e_{2i}，则两段的残差平方和分别为 $RSS_1 = \sum_{i=1}^{n_1} e_{1i}^2$ 和 $RSS_2 = \sum_{i=1}^{n_2} e_{2i}^2$，从而可计算出各段模型的随机误差项的方差估计量分别为 $\hat{\sigma}_1^2 = \dfrac{RSS_1}{n_1 - k}$ 和 $\hat{\sigma}_2^2 = \dfrac{RSS_2}{n_2 - k}$，由此可构造出检验统计量为

$$F = \frac{\hat{\sigma}_1^2}{\hat{\sigma}_2^2} = \frac{RSS_1/(n_1 - k)}{RSS_2/(n_2 - k)}$$

该统计量服从自由度为 $(n_1 - k)$ 和 $(n_2 - k)$ 的 F 分布。在给定的显著性水平 α 之下，若此统计量 F 的值大于临界值 $F_\alpha\left(n_1 - k, n_2 - k\right)$，则可认为有异方差的存在。

4．简述等级相关系数法的检验步骤。

答：等级相关系数法又称斯皮尔曼(Spearman)检验，是一种应用较广泛的方法。这种检验方法既适用于大样本，也适用于小样本。按下式计算出等级相关系数

$$r_s = 1 - \frac{6}{n(n^2 - 1)} \sum_{i=1}^{n} d_i^2$$

其中，n 为样本容量，d_i 为对应于 X_i 和 $|e_i|$ 的等级的差数。通过 t 检验判断是否存在异方差。

在多元回归的情况下，需对每一个解释变量做等级相关系数检验。只有当每个解释变量检验都不存在异方差时模型中才不存在异方差。否则，模型中存在异方差。

5．产生序列相关的原因有哪些？

答：(1) 遗漏了重要的解释变量。

⑵ 经济变量的滞后性。

⑶ 回归模型函数形式的设定错误也可能引起序列相关。

⑷ 实际问题研究中出现的蛛网现象(Cobweb Phenomenon)。

⑸ 对原始数据加工整理

6．序列相关性带来哪些后果？

答：⑴ 参数的估计量是无偏的，但不是有效的。

⑵可能严重低估误差项的方差。

⑶ 常用的 F 检验和 t 检验失效。

⑷回归参数的置信区间和利用回归模型进行预测的结果存在较大的误差。

7．简述 DW 检验及其决策规则。

答：DW 检验只能用于检验随机误差项具有一阶自回归形式的序列相关问题。随机误差项的一阶自回归形式为 $u_t = \quad u_{t-1} + v_t$，构造的原假设是 $H_0: \quad = 0$。为了检验该假设，构造 DW 统计量首先要求出回归估计式的残差 e_t，$DW = \dfrac{\sum_{t=2}^{n}(e_t - e_{t-1})^2}{\sum_{t=1}^{n} e_t^2}$，$n$ 较大时，$DW \approx 2(1 - \hat{\ })$，根据样本容量 n 和解释变量的数目 k'(不包括常数项)，查 DW 分布表，得临界值 d_L 和 d_U，然后依下列准则考察计算得到的 DW 值，以决定模型的自相关状态。

DW 检验决策规则

$0 \le DW \le d_L$	误差项 $u_1, u_2, \cdots, u_n$ 间存在正相关
$d_L < DW \le d_U$	不能判定是否有自相关
$d_U < DW < 4 - d_U$	误差项 $u_1, u_2, \cdots, u_n$ 间无自相关
$4 - d_U \le DW < 4 - d_L$	不能判定是否有自相关
$4 - d_L \le DW \le 4$	误差项 $u_1, u_2, \cdots, u_n$ 间存在负相关

8．简述 DW 检验的局限性。

答：DW 检验的局限性

⑴DW 检验有两个不能确定的区域；

⑵DW 统计量的上、下界表要求 $n \ge 15$；

⑶检验不适应随机误差项具有高阶序列相关的情况；

⑷只适用于有常数项的回归模型并且解释变量中不能含滞后的被解释变量。

9．存在严重共线性时，估计参数产生的后果有哪些？

答：

(1) 多重共线性不改变参数估计量的无偏性。

(2)多重共线性使参数的最小二乘估计的方差很大，即估计值的精度很低。

(3) t 检验和 F 检验失效。

10．多重共线性直观判定法包括哪些主要方法？

答：

(1) R^2 较高，而显著 t 统计量较少时，可能存在多重共线性问题。

(2)当增加或剔除一个解释变量，或者改变一个观测值时，回归系数的估计值发生较大变化，就认为回归方程存在严重的多重共线性。

(3)一些重要的解释变量在回归方程中没有通过显著性检验时，可初步判断存在着严重的多重共线性。

(4)有些解释变量的回归系数所带符号与定性分析结果违背时，可能存在多重共线性问题。

(5)解释变量间的相关系数较大时，可能会出现多重共线性问题。

11．多重共线性补救方法有哪几种？

答：处理多重共线性问题的方法很多，常用的有下面几种。

(1)使用非样本先验信息

如果据先前的经济计量分析或经济理论分析已知模型中的共线性解释变量的参数间具有某种线性关系，则可利用此条件消除解释变量间的多重共线性。

(2)横截面与时间序列数据并用

就是先利用横截面数据估计某一参数，将结果代入原方程后，再利用时间序列数据估计另一参数。

(3)剔除一些不重要的共线性解释变量

(4)增大样本容量

(5)使用有偏估计

五、计算题

1．解：$Y_i = {}_0 + {}_1 X_i + u_i$

两边同除以 X_i 得：$\frac{Y_i}{X_i} = {}_0 \frac{1}{X_i} + {}_1 + \frac{u_i}{X_i}$

$Var(\frac{u_i}{X_i}) = \frac{1}{X_i^2} Var(u_i) = \sigma^2$ 消除了异方差，可用普通最小二乘法估计参数：

$$E(\frac{1}{X}) = \frac{\sum_{i=1}^{5} \frac{1}{X_i}}{5} = 0.41, \qquad E(\frac{Y}{X}) = \frac{\sum_{i=1}^{5} \frac{Y_i}{X_i}}{5} = 1.40$$

$$\hat{}_0 = \frac{\sum_{i=1}^{5} (\frac{1}{X_i} - 0.41)(\frac{Y_i}{X_i} - 1.40)}{\sum_{i=1}^{5} (\frac{1}{X_i} - 0.41)^2} = \frac{0.71}{0.52} = 1.37$$

$$\hat{}_1 = 1.4 - 1.37 \times 0.41 = 0.83$$

回归方程为 $\hat{Y} = 1.37 + 0.83X$

2．解：计算过程如下：

工厂	质量排序	价格排序	d	d^2
1	6	5	1	1
2	9	9	0	0
3	2	1	1	1
4	8	2	6	36
5	5	4	1	1
6	1	3	-2	4
7	7	8	-1	1
8	4	6	-2	4
9	3	7	-4	16
10	10	10	0	0
合计				64

$$r_s = 1 - \frac{6\Sigma d^2}{n(n^2-1)} = 1 - \frac{6\times 64}{10\times(10^2-1)} = 0.61$$

3．解：广义差分模型为

$$\begin{aligned}
& Y_t - {}_1Y_{t-1} - {}_2Y_{t-2} - \mathrm{L} - {}_pY_{t-p} \\
= & {}_1(1 - {}_1 - {}_2 - \mathrm{L} - {}_p) \\
+ & {}_2(X_{2t} - {}_1X_{2,t-1} - {}_2X_{2,t-2} - \mathrm{L} - {}_pX_{2,t-p}) + \mathrm{L} \\
+ & {}_k(X_{kt} - {}_1X_{k,t-1} - {}_2X_{k,t-2} - \mathrm{L} - {}_pX_{k,t-p}) + v_t
\end{aligned}$$

4．解：

$$DW = \frac{\sum_{t=2}^{n}(e_t - e_{t-1})^2}{\sum_{t=1}^{n} e_t^2}$$

$$= \frac{\sum_{t=2}^{n} e_t^2 + \sum_{t=2}^{n} e_{t-1}^2 - 2\sum_{t=1}^{n} e_t e_{t-1}}{\sum_{t=1}^{n} e_t^2}$$

$$= \frac{39+36-2\times 20}{40} = \frac{35}{40} = 0.875$$

六、分析题

1．解：(1)计算各参数的 t 统计量分别为 -3.60,10.20,6.52 绝对值全部大于临界值 2.101，所以各偏回归系数均显著。

(2)计算各参数的 t统计量分别为 -2.87,1.29,1.38,3.97,t和 L nK 的参数所

对应的 t统计量绝对值小于临界值 2.111，所以不显著。

(3)模型 II存在多重共线性问题。

(4)$\hat{\alpha}+\hat{\ }=0.887+0.893=1.78$ ，规模报酬递增。

2. 解：因为是规模报酬不变，所以 $\alpha+\ =1$，将 $\alpha=1-\ $ 代入模型中，得到

$$Ln\frac{Y}{L}=LnA+\ \ Ln\frac{K}{L}+u$$

由二个解释变量的对数模型转变为一个解释变量的对数线性模型，消除了多重共线性。使用普通最小二乘法估计出资本弹性 $\hat{\ }$，则劳动力弹性为 $\hat{\alpha}=1-\hat{\ }$。从而得到柯布道格拉斯生产函数。

第五章　分布滞后模型

练习题

一、单项选择题

1．下列属于有限分布滞后模型的是（　　）

A．$Y_t=\alpha+\ {}_0X_t+\ {}_1Y_{t-1}+u_t$

B．$Y_t=\alpha+\ {}_0X_t+\ {}_1Y_{t-1}+\ {}_2Y_{t-2}+\mathrm{L}+\ {}_kY_{t-k}+u_t$

C．$Y_t=\alpha+\ {}_0X_t+\ {}_1X_{t-1}+\mathrm{L}+u_t$

D．$Y_t=\alpha+\ {}_0X_t+\ {}_1X_{t-1}+\mathrm{L}+\ {}_kX_{t-k}+u_t$

2．设分布滞后模型为 $Y_t=\alpha+\ {}_oX_t+\ {}_1X_{t-1}+\ {}_2X_{t-2}+\ {}_3X_{t-3}+u_t$，为了使模型的自由度达到 30，必须拥有多少年的观测资料（　　）

A．32　　B．33　　C．35　　D．38

3．在分布滞后模型 $Y_t=\alpha+\ {}_oX_t+\ {}_1X_{t-1}+\ {}_2X_{t-2}+\ {}_3X_{t-3}+u_t$ 中，延期过渡性影响乘数是指（　　）

A． $_o$　　B． $_1$、 $_2$、 $_3$　　C． $_1+\ _2+\ _3$　D． $_o+\ _1+\ _2+\ _3$

4．在分布滞后模型的估计中，使用时间序列资料可能存在的序列相关问题就表现为（　　）

A．异方差问题　　　　　　　　B．自相关问题

C．多重共线性问题　　　　　　D．随机解释变量问题

5．下列哪一个不是几何分布滞后模型的变换模型（　　）

A．库伊克变换模型　　　　　　B．自适应预期模型

C．部分调整模型　　　　　　　D．有限多项式滞后模式

6．下列哪个模型的一阶线性自相关问题可用 DW 检验（　　）

A．有限多项式分布滞后模型　　　　B．自适应预期模型

C．库伊克变换模型　　　　　　　　D．部分调整模型

7．对有限分布滞后模型 $Y_t=\alpha+\ _oX_t+\ _1X_{t-1}+\mathrm{L}+\ _kX_{t-k}+u_t$ 进行多项式变换时，多项式的阶数 m 与最大滞后长度 k 的关系是（　　）

A．$m<k$　　　B．$m=k$　　　C．$m>k$　　　D．不确定

8．设无限分布滞后模型为 $Y_t=\alpha+\ _oX_t+\ _1X_{t-1}+\mathrm{L}+u_t$ 该模型满足库伊克（koyck）提出的两个假设，则长期影响乘数为（　　）

A．$_o\lambda^k$　　B．$\dfrac{o}{1-\lambda}$　　C．$\dfrac{(1-\lambda^k)\ \ _0}{1-\lambda}$　　D．不确定

9．使用普通最小二乘法在对自回归模型进行估计时，若随机误差项满足经典线性回归模型的所有假定，则估计量是一致估计量的模型是（　　）

A．Koyck 变换模型　　　　　B．部分调整模型

C．自适应预期模型　　　　　D．自适应预期和部分调整混合模型

10．对于自适应预期模型，采用什么方法估计参数比较合适。（　　）

A．普通最小二乘法　　　　　B．加权最小二乘法

C．工具变量法　　　　　　　D．广义差分法

11．若假定 t 期解释变量观测值与同期被解释变量希望达到水平之间存在线性关系，则这种理论假定属于（　　）

A．Koyck 变换模型　　　　　B．几何分布滞后模型

C．自适应预期模型　　　　　D．部分调整模型

12．若假定影响被解释变量 Y_t 的因素不是 X_t 而是 X_{t+1} 的预期 X^*_{t+1} ，即

$Y_t = {}_0 + {}_1X^*_{t+1} + u_t$，建立在这种经济理论基础上的模型属于（　　）

A．Koyck 变换模型　　　　B．几何分布滞后模型

C．自适应预期模型　　　　D．部分调整模型

13．对于自回归模型，检验是否存在序列相关的方法是（　　）

A．DW 检验　　B．方差比检验　　C．自相关系数检验　　D．h 检验法

14．对于无限分布滞后模型，库伊克（Koyck）提出的假定是（　　）

A．参数符号相同且按几何数列衰减

B．参数符号相同且按几何数列递增

C．参数符号不同但按几何数列衰减

D．参数符号不同但按几何数列递增

二、多项选择题

1．对于有限分布滞后模型，对它应用最小二乘法估计时存在以下困难（　　）

A．产生多重线性　　　　B．产生异方差　　　　C．产生自相关

D．损失自由度　　　　E．最大滞后期 k 较难确定

2．在模型 $Y_t = \alpha + {}_oX_t + {}_1X_{t-1} + {}_2X_{t-2} + {}_3X_{t-3} + u_t$ 中延期乘数是指

A．${}_0$　　　　B．${}_1$　　　　C．${}_2$

D．${}_3$　　　　E．${}_0 + {}_1 + {}_2 + {}_3$

3．对自适应预期模型 $Y_t = {}_0 + {}_1X_t + {}_2Y_{t-1} + v_t$，$Y_{t-1}$ 是随机解释变量，若采用工具变量法估计参数，则工具变量 Z_t 必须满足的条件是（　　）

A．$Cov(Z_t, X_t) = 0$　　　　B．$Cov(Z_t, X_t) \neq 0$

C．$Cov(Z_t, Y_{t-1}) \neq 0$　　　　D．$Cov(Z_t, Y_{t-1}) = 0$

E．$Cov(Z_t, v_t) = 0$

4．对于无限分布滞后模型，库伊克（Koyck）提出的假定是（　　）

A．参数符号要相同　　　　B．参数符号不同

C．参数按几何数列衰减　　　　D．参数按几何数列递增

E．参数变化没有规律性

5．下列哪些模型属于无限分布滞后模型（　　）

A．阿尔蒙多项式滞后模型　　B．部分调整模型

C．自适应预期模型　　D．几何分布滞后模型

E．库伊克（Koyck）变换模型

6．对于有限分布滞后模型，通常采用什么方法估计参数（　　）

A．经验权数法　　B．阿尔蒙多项式变换法

C．普通最小二乘法　　D．工具变量法

E．广义最小二乘法

7．产生滞后的原因包括（　　）

A．心理因素　　B．技术因素　　C．制度因素

D．模型设计原因　　E．估计参数原因

8．阿尔蒙估计法的优点是（　　）

A．克服了自由度不足的问题　　B．阿尔蒙变换具有充分的柔顺性

C．解决了滞后阶数问题　　D．多项式的阶数是固定的

E．可以克服多重共线性问题

三、名词解释

1．分布滞后模型　　2．短期影响乘数

3．延期影响乘数　　4．长期影响乘数

5．几何分布滞后模型

四、简答题

1．产生滞后的原因有哪些？

2．用最小二乘法对分布滞后模型进行参数估计时存在什么困难？

3．经验权数法估计步骤是什么？

4．阿尔蒙多项式滞后模型的原理及优缺点。

5．什么是无限分布滞后模型？简述库伊克（Koyck）所提出两个假设的内容。

6．自适应预期模型的经济理论基础。

7．部分调整模型的经济理论假定。

8．能否直接用 DW 检验自回归模型的自相关问题？为什么？应采用什么方法检验？

五、计算题

1．设有限分布滞后模型为$Y_t = \alpha + {}_o X_t + {}_1 X_{t-1} + {}_2 X_{t-2} + {}_3 X_{t-3} + {}_4 X_{t-4} + u_t$，

假设用 2 阶多项式变换该模型为阿尔蒙多项式模型，使用普通最小二乘法估计这个模型，

得到：$\hat{\alpha}_0 = 0.5$　　$\hat{\alpha}_1 = 0.45$　　$\hat{\alpha}_2 = -0.1$。

要求：(1)求 $_0$， $_1$， $_2$， $_3$， $_4$的估计值。

(2) 求 X 对 Y 的短期影响乘数、长期影响乘数和延期影响乘数。

2．设$Y_t = \alpha + \ _o X_t + \ _1 X_{t-1} + \ _2 X_{t-2} + \ _3 X_{t-3} + u_t$

要求：用 2 阶有限多项式对模型进行阿尔蒙变换。

3．由经济理论得知，现在的消费水平受到现在和过去收入水平影响，假定 Y=消费额 X=收入，试用库伊克（Koyck）变换推导出消费函数适当模型。

六、分析题

1．已知某公司 1998 年至 2003 年库存商品额 Y 与销售额 X 的季度数据资料，假定最大滞后长度 k=3，多项式的阶数 m=2。

要求：(1) 试建立库存商品额 Y 与销售额 X 的分布滞后模型。并利用阿尔蒙多项式进行变换。

(2) 假定用最小二乘法得到阿尔蒙多项式变换模型的估计式为

$$\hat{Y}_t = -120.6278 + 0.5314 Z_{0t} + 0.8026 Z_{1t} - 0.3327 Z_{2t}$$

写出分布滞后模型的估计式。

2．设自回归模型为$Y_t = \alpha X_t + \lambda Y_{t-1} + v_t$ 式中$v_t = u_t - \lambda u_{t-1}$为满足经典假设随机误差项，应该采用什么方法估计参数？写出估计该模型参数的正规方程组。

参考答案

一、单项选择题

1．D　2．D　3．B　4．C　5．D　6．A　7．A　8．B　9．B　10．C　11．D　12．C　13．D　14．A

二、多项选择题

1．ACDE　2．BCD　3．CE　4．AC　5．BCDE　6．AB　7．ABC　8．ABE

三、名词解释

1．分布滞后模型 ：如果一个回归模型不仅包含解释变量的现期值，而且还包含解释变量的滞后值，则这个回归模型就是分布滞后模型。它的一般形式为：

$$Y_t = \alpha + \ _0 X_t + \ _1 X_{t-1} + \ + \ _k X_{t-k} + u_t$$

或

$$Y_t = \alpha + \ _0 X_t + \ _1 X_{t-1} + \ + u_t$$

2．短期影响乘数：在分布滞后模型 $Y_t=\alpha+{}_0X_t+{}_1X_{t-1}+\ +{}_kX_{t-k}+u_t$ 中，${}_0$ 称为短期影响乘数，它表示解释变量 X变化一个单位对同期被解释变量 Y产生的影响。

3．延期影响乘数：在分布滞后模型：$Y_t=\alpha+{}_0X_t+{}_1X_{t-1}+\ +{}_kX_{t-k}+u_t$ 中，

${}_1$，${}_2$，，${}_k$ 称为延期过渡性影响乘数，它们度量解释变量 X的各个前期值变动一个单位对被解释变量 Y的滞后影响。

4．长期影响乘数：在分布滞后模型：$Y_t=\alpha+{}_0X_t+{}_1X_{t-1}+\ +{}_kX_{t-k}+u_t$ 中，所有乘数的和 $\sum {}_i = {}_0+{}_1+{}_2+\mathrm{L}= \ <\infty$ 称为长期影响乘数。

5．几何分布滞后模型：对于无限分布滞后模型 $Y_t=\alpha+{}_0X_t+{}_1X_{t-1}+\mathrm{L}+u_t$ 库伊克(koyck)提出了两个假设：①模型中所有参数的符号都是相同的。②模型中的参数几何数列衰减的，即 ${}_j={}_0\lambda^j$，j=0，1，2，。式中，0＜¦Ë＜1，¦Ë称为分布滞后的衰减率，¦Ë越小，衰减速度就越快，X 滞后的远期值对当期 Y 值的影响就越小。

$$Y_t=\alpha+{}_0X_t+{}_0\lambda X_{t-1}+{}_0\lambda^2X_{t-2}+\mathrm{L}+{}_0\lambda^jX_{t-j}+\mathrm{L}+u_t$$

就称为几何分布滞后模型。

四、简答题

1．产生滞后的原因有哪些？

答：(1)心理上的原因 。

因为人们要改变习惯以适应新的情况往往需要一段时间 ,这种心理因素会造成出现滞后效应。

(2)技术上的原因 。

产品的生产周期有长有短，但都需要一定的周期，造成出现滞后效应。

(3)制度上的原因。

某些规章制度的约束使人们对某些外部变化不能立即做出反应，从而出现滞后现象。

2．用最小二乘法对分布滞后模型进行参数估计时存在什么困难？

答：首先对于无限分布滞后模型，因为其包含无限多个参数，无法用最小二乘法直接对其估计，其次对于有限分布滞后模型，即使假设它满足经典假定条件，对它应用最小二乘估计也存在以下困难。

(1)产生多重共线问题

对于时间序列 X_t 的各期变量之间往往是高度相关的，因而分布滞后模型常常产生多重共线性问题。

(2)损失自由度问题

由于样本容量有限，当滞后变量数目增加时，必然使得自由度减少。由于经济数据的收集常常受到各种条件的限制，估计这类模型时经常会遇到数据不足的困难。

(3)对于有限分布滞后模型，最大滞后期 k较难确定。

(4)分布滞后模型中的随机误差项往往是严重自相关的。

３．经验权数法估计步骤是什么？

答：经验权数法又称为经验法，它是指根据观察及经验为滞后变量的系数指定权数，即根据经验赋予各滞后变量的系数 $_0$，$_1$， ，$_k$ 相应的权数，使滞后变量按权数的线性组合构成新的变量W，然后用最小二乘法估计参数。

４．阿尔蒙多项式滞后模型的原理及优缺点。

答：阿尔蒙多项式滞后模型的基本思想是：如果有限分布滞后模型中的参数 $_i(i=0,1,2,\cdots,k)$的分布可以近似用一个关于 i 的低阶多项式表示，就可以利用多项式减少模型中的参数。

阿尔蒙估计法的优点

（1）克服了自由度不足的问题。

（2）阿尔蒙变换具有充分的柔顺性。

（3）可以克服多重共线性问题。

阿尔蒙估计法的缺点

（1）仍没有能够解决原模型滞后阶数 k应该取什么值为最好的问题。

（2）多项式阶数 m必须事先确定，而 m的实际确定往往带有很大的主观性。

（3）虽然阿尔蒙估计法可能将回归式中的多重共线性程度降低了很多，变量 Z 之间的多重共线性就可能弱于诸 X之间的多重共线性 ,但它并没能完全消除多重共线性问题对回归模型的影响。

５．什么是无限分布滞后模型？简述库伊克（Koyck）所提出两个假设的内容。

答：型如 $Y_t=\alpha+{}_0X_t+{}_1X_{t-1}+\cdots+u_t$ 的模型称为无限分布滞后模型，库伊克(koyck)对模型提出了两个假设：

(1)模型中所有参数的符号都是相同的。

(2)模型中的参数是按几何数列衰减的，即

$$_j={}_0\lambda^j \qquad j=0，1，2，$$

式中，０＜¦Ë＜１，¦Ë称为分布滞后的衰减率，¦Ë越小，衰减速度就越快，X 滞后的远期值对当期 Y 值的影响就越小。

６．自适应预期模型的经济理论基础。

答：自适应预期模型建立在如下的经济理论基础上：影响被解释变量 Y_t 的因素不是 X_t 而是 X_{t+1} 的预期 X^*_{t+1}，即 $Y_t={}_0+{}_1X^*_{t+1}+u_t$

自适应预期假定，就是预期的形成过程如下式所表达的：

$$X^*_{t+1}-X^*_t=\gamma(X_t-X^*_t)$$

式中，γ 称为预期调整系数，且 $0\le\gamma\le1$，$(X_t-X^*_t)$是实际值与预期值的偏差，称为预期误差。

自适应预期模型的自回归形式为

$$Y_t=\gamma{}_0+{}_1\gamma X_t+(1-\gamma)Y_{t-1}+v_t$$

7．部分调整模型的经济理论假定。

答：部分调整模型所根据的行为假定是模型所表达的不应是 t 期解释变量观测值与同期被解释变量观测值之间的关系，而应是 t 期解释变量观测值与同期被解释变量希望达到的水平之间的关系。即

$$Y_t^* = \ {}_0 + \ {}_1 X_t + u_t$$

式中，Y_t^* —被解释变量的希望值（或最佳值），X_t —解释变量在 t 期的真实值，u_t —随机误差项。由于种种原因，被解释变量的实际变动值 $Y_t - Y_{t-1}$ 往往只能达到希望水平与实际水平变动 $Y_t^* - Y_{t-1}$ 的一部分。设只达到了δ 比例的一部分，则部分调整假设可表示为

$$Y_t - Y_{t-1} = \delta(Y_t^* - Y_{t-1})$$

式中，δ 为部分调整系数，且有 $0 \le \delta \le 1$。

部分调整模型的自回归形式为

$$Y_t = \delta \ {}_0 + \delta \ {}_1 X_t + (1-\delta) Y_{t-1} + \delta u_t$$

8．能否直接用 DW 检验自回归模型的自相关问题？为什么？应采用什么方法检验？

答：(1) 在自回归模型中，如果含有滞后被解释变量 Y_{t-1} 作为解释变量，这时需要检查模型中随机误差项是否存在序列相关性，DW 检验就不再适用了。

(2) 因为应用 DW 检验的一个前提条件就是解释变量为非随机变量，否则就会得到错误的结论。

(3) 此时需要用 h 统计量检验，设自回归模型为

$$Y_t = \ {}_0 + \ {}_1 X_t + \ {}_2 Y_{t-1} + u_t$$

定义的 h 统计量为

$$h = \hat{\ } \sqrt{\frac{n}{1 - nVar(\hat{\ }_2)}}$$

其中，$\hat{\ }_2$ 是模型中 Y_{t-1} 的系数 ${}_2$ 的估计量，$Var(\hat{\ }_2)$ 是 $\hat{\ }_2$ 的方差的样本估计值，n 为样本容量，$\hat{\ }$ 是随机误差项一阶自相关系数的估计值，在应用时，可取 $\hat{\ } = 1 - \frac{1}{2}d$，$d$ 是通常意义下 DW 统计量的取值。

h 统计量的原假设为 H_0： $= 0$，备择假设为 H_1： $\neq 0$。

在大样本情形下，Durbin 证明了在原假设 H_0： $= 0$ 成立的条件下，统计量 h 渐进地遵循零均值和单位方差的正态分布。

五、计算题

1．解：(1) $\hat{\ }_{0}=0.5$

$$\hat{\ }_{1}=0.5+0.45-0.1=0.85$$

$$\hat{\ }_{2}=0.5+2\times0.45-4\times0.1=1.00$$

$$\hat{\ }_{3}=0.5+3\times0.45-9\times0.1=0.95$$

$$\hat{\ }_{4}=0.5+4\times0.45-16\times0.1=0.7$$

(2)短期乘数为 0.5

长期乘数为 4.0

延期乘数分别为：0.85， 1.00， 0.95， 0.7。

2.解：系数多项式表达式为

$$_{i}=\alpha_{0}+\alpha_{1}i+\alpha_{2}i^{2} \qquad (i=0,1,2\ ,3)$$

其中，$\alpha_0,\alpha_1,\alpha_2$是待估计的参数。

$$Y_t=\alpha+\alpha_0(X_t+X_{t-1}+X_{t-2}+X_{t-3})+\alpha_1(X_{t-1}+2X_{t-2}+3X_{t-3})+\alpha_2(X_{t-1}+4X_{t-2}+9X_{t-3})+u_t$$

另记：

$$Z_{0t}=X_t+X_{t-1}+X_{t-2}+X_{t-3}$$
$$Z_{1t}=X_{t-1}+2X_{t-2}+3X_{t-3}$$
$$Z_{2t}=X_{t-1}+4X_{t-2}+9X_{t-3}$$

则可变换为

$$Y_t=\alpha+\alpha_0Z_{0t}+\alpha_1Z_{1t}+\alpha_2Z_{2t}+u_t$$

利用样本数据进行最小二乘估计，可得到各个参数的估计值，分别记为$\hat{\alpha},\hat{\alpha}_0,\hat{\alpha}_1,\hat{\alpha}_2$，可得原模型参数的估计值为

$$\hat{\ }_{0}=\hat{\alpha}_0$$
$$\hat{\ }_{1}=\hat{\alpha}_0+\hat{\alpha}_1+\hat{\alpha}_2$$
$$\hat{\ }_{2}=\hat{\alpha}_0+2\hat{\alpha}_1+4\hat{\alpha}_2$$
$$\hat{\ }_{3}=\hat{\alpha}_0+3\hat{\alpha}_1+9\hat{\alpha}_2$$

3.解：据题意，消费函数模型为$Y_t=\alpha+\ _0X_t+\ _1X_{t-1}+\mathrm{L}+u_t$，库伊克(koyck)提出了两个假设：

（1）模型中所有参数的符号都是相同的。

（2）模型中的参数是按几何数列衰减的，即

$$_{\mathrm{j}}=\ _{0}\lambda^{\mathrm{j}} \qquad j=0，1，2，$$

式中，0＜¦Ë＜1，¦Ë称为分布滞后的衰减率，¦Ë越小，衰减速度就越快，X 滞后的远期值对当期 Y 值的影响就越小。

将 $_j$ 代入到模型中，得

进行库伊克(koyck)变换，可得消费函数模型为 $Y_t = \alpha(1-\lambda) + {}_0 X_t + \lambda Y_{t-1} + v_t$ ，

式中，$v_t = u_t - \lambda u_{t-1}$。

六、分析题

1. 解：(1) 假设 $Y_t = \alpha + {}_0 X_t + {}_1 X_{t-1} + {}_2 X_{t-2} + {}_3 X_{t-3} + u_t$

系数多项式表达式为

$$_i = \alpha_0 + \alpha_1 i + \alpha_2 i^2 \qquad (i=0,1,2,3)$$

其中，$\alpha_0, \alpha_1, \alpha_2$ 是待估计的参数。

$$Y_t = \alpha + \alpha_0(X_t + X_{t-1} + X_{t-2} + X_{t-3}) + \alpha_1(X_{t-1} + 2X_{t-2} + 3X_{t-3}) + \alpha_2(X_{t-1} + 4X_{t-2} + 9X_{t-3}) + u_t$$

另记：

$$Z_{0t} = X_t + X_{t-1} + X_{t-2} + X_{t-3}$$
$$Z_{1t} = X_{t-1} + 2X_{t-2} + 3X_{t-3}$$
$$Z_{2t} = X_{t-1} + 4X_{t-2} + 9X_{t-3}$$

则模型为

$$Y_t = \alpha + \alpha_0 Z_{0t} + \alpha_1 Z_{1t} + \alpha_2 Z_{2t} + u_t$$

(2) $\hat{\alpha}_0 = 0.5314$ ， $\hat{\alpha}_1 = 0.8026$ ， $\hat{\alpha}_2 = -0.3327$， $\hat{}_0 = 0.5314$ ， $\hat{}_1 = 1.0013$ ， $\hat{}_2 = 0.8058$ ， $\hat{}_3 = -0.0551$ 。

分布滞后模型的估计式为

$$Y_t = -120.6278 + 0.5314X_t + 1.0013X_{t-1} + 0.8058X_{t-2} - 0.0551X_{t-3}$$

2. 解：应采用工具变量法。采用 X_{t-1} 为工具变量，得到正规方程组为

$$\sum X_t Y_t = \hat{\alpha}\sum X_t^2 + \hat{\lambda}\sum X_t Y_{t-1}$$

$$\sum X_{t-1} Y_t = \hat{\alpha}\sum X_{t-1} X_t + \hat{\lambda}\sum X_{t-1} Y_{t-1}$$

求解该方程组可得到参数的一致估计量。

第六章　虚拟解释变量模型

练习题

一、　单项选择题

1．设$Y_i = {}_0 + {}_1X_i + u_i$，$Y_i$ =居民消费支出，X_i =居民收入，D=1 代表城镇居民，D=0 代表农村居民，则截距变动模型为（　）

A．$Y_i = {}_0 + {}_1X_i + {}_2D + u_i$　　B．$Y_i = ({}_0 + {}_2) + {}_1X_i + u_i$

C．$Y_i = ({}_0 + {}_1) + {}_1X_i + u_i$　　D．$Y_i = {}_0 + {}_1X_i + {}_2DX_i + u_i$

2．设虚拟变量 D 影响线性回归模型中 X_i 的斜率，如何引进虚拟变量，使模型成为斜率变动模型（　）

A．直接引进 D　　B．按新变量 DX_i 引进

C．按新变量 $D+X_i$ 引进　　D．无法引进

3．设截距系统变参数模型为：$Y_i = {}_{0i} + {}_1X_i + u_i$，${}_{0i} = a + bZ_i$，当 Z_i 为什么变量时，该模型实质上是截距变动模型（　）

A．内生变量　　B．外生变量　　C．虚拟变量　　D．滞后变量

4．虚拟变量陷阱是指（　）

A．引进虚拟变量后造成多重共线性问题　　B．引进虚拟变量后造成异方差问题

C．引进虚拟变量造成序列相关问题　　D．引进虚拟变量后造成设定误差问题

5．在建立经济计量模型时，虚拟变量（　）

A．只能做解释变量　　B．可以做解释变量，也可以做被解释变量

C．只能做被解释变量　　D．既不能做解释变量，也不能做被解释变量

6．虚拟变量的赋值原则是（　）

A．给定某一质量变量的某属性出现为 1，未出现为 0。

B．给定某一质量变量的某属性出现为 1，另一属性出现为 0。

C．不用赋值

D．按照某一质量变量属性种类编号赋值

7．有关虚拟变量的表述正确的为（　）

A．用来代表质的因素，有时候也可以代表数量因素　　B．只能用来代表质的因素

C．只能用来代表数量因素　　　　　　　　　　　D．以上都不正确

8．如果一个回归模型不包含截距项，对一个具有 m 个特征的质的因素需要引入的虚拟变量个数为（　　）

A．m　　　B．$(m-1)$　　　C．$(m-2)$　　　D．$(m+1)$

9．如果一个回归模型包含截距项，对一个具有 m 个特征的质的因素需要引入的虚拟变量个数为（　　）

A．m　　　B．$(m-1)$　　　C．$(m-2)$　　　D．$(m+1)$

10．设个人消费函数 $Y_i = {}_1 + {}_2 X_i + u_i$ 中，消费支出 Y 不仅与收入 X 有关，而且与消费者的性别、年龄构成有关，年龄构成可以分为老、中、青三个层次，假定边际消费倾向不变，该消费函数引入虚拟变量的个数为（　　）

A．1 个　　　B．2 个　　　C．3 个　　　D．4 个

11．在一个包含截距项的回归模型 $Y_i = {}_0 + {}_1 D + {}_2 X_i + u_i$ 中，如果一个具有 m 个特征的质的因素引入 m 个虚拟变量，则会产生的问题为（　　）

A．异方差　　　　　　　　　B．序列相关

C．不完全多重线性相关　　　D．完全多重线性相关

12．设消费函数为 $Y_i = {}_0 + {}_1 D + {}_2 X_i + u_i$，式中 Y_i=第 i 个居民的消费水平，X_i=第 i 个居民的收入水平，D 为虚拟变量，D=1 表示正常年份，D=0 表示非正常年份，则（　　）

A．该模型为截距、斜率同时变动模型　　　B．该模型为截距变动模型

C．该模型为分布滞后模型　　　　　　　　D．该模型为时间序列模型

13．设截距和斜率同时变动模型为 $Y_i = {}_0 + {}_1 D + {}_2 X_i + {}_3 (DX_i) + u_i$

下面那种情况成立，该模型为截距变动模型（　　）

A．${}_1 \neq 0, \quad {}_3 \neq 0$　　　B．${}_1 \neq 0, \quad {}_3 = 0$

C．${}_1 = 0, \quad {}_3 = 0$　　　D．${}_1 = 0, \quad {}_3 \neq 0$

14．根据样本资料建立的消费函数如下：$\hat{C}_t = 110.5 + 65D_t + 0.5X_t$

其中，C 为消费，X 为收入，虚拟变量 D=1 城镇家庭，D=0 农村家庭，所有参数均检验显著，则城镇家庭的消费函数为（　　）

A．$\hat{C}_t = 110.5 + 0.5X_t$　　　B．$\hat{C} = 175.5 + 0.5X_t$

C．$\hat{C}_t = 110.5 + 65.5X_t$　　　D．$\hat{C} = 111 + 65.5X_t$

15．假定某需求函数 $Y_t = {}_0 + {}_1 X_t + u_t$，且需求量与季节有关，季节分为春、夏、

秋、冬四季，引入 4 个虚拟变量得到虚拟变量模型，则模型参数估计量为（　　）

A．有效估计量　　B．有偏估计量　　C．非一致估计量　　D．无法估计

16．设消费函数为$Y_i = {}_0 + {}_1X_i + a_1D_1 + a_2D_2 + a_3D_3 + u_i$，其中，$Y$为消费，$X$为收入，$D_1 = \begin{cases} 1 & 第一季度 \\ 0 & 其它 \end{cases}$，$D_2 = \begin{cases} 1 & 第二季度 \\ 0 & 其它 \end{cases}$，$D_3 = \begin{cases} 1 & 第三季度 \\ 0 & 其它 \end{cases}$。

该模型包含了几个质因素（　　）

A．1　　B．2　　C．3　　D．4

17．设消费函数为$\hat{C}_t = {}_0 + {}_1X_t + {}_2D + u_t$，$C$ 为消费，X 为收入，$D = \begin{cases} 1 & 城镇居民 \\ 0 & 农村居民 \end{cases}$，如果统计检验${}_2 = 0$成立，则城镇居民消费函数和农村居民消费函数是（　　）

A．相互平行的　　B．相互垂直的　　C．相互交叉的　　D．相互重叠的

18．设消费函数为$\hat{C}_t = {}_0 + {}_1X_t + {}_2D + u_t$，$C$ 为消费，X 为收入，$D = \begin{cases} 1 & 城镇居民 \\ 0 & 农村居民 \end{cases}$，如果统计检验${}_2 \neq 0$成立，则城镇居民消费函数和农村居民消费函数是（　　）

A．相互平行的　　B．相互垂直的　　C．相互交叉的　　D．相互重叠的

19．假定月收入在 2000 元以内和月收入在 2000 元以上的居民边际消费倾向明显不同，$D = \begin{cases} 1 & X \geq 2000元 \\ 0 & X < 2000元 \end{cases}$，$X_t^* = 2000元$，则描述消费函数变动线性关系应采用（　　）

A．$Y_t = {}_0 + {}_1X_t + {}_2D + u_t$

B．$Y_t = {}_0 + {}_1X_t + {}_3DX_t + u_t$

C．$Y_t = {}_0 + {}_1X_t + {}_2(X_t - X_t^*) + u_t$

D．$Y_t = {}_0 + {}_1X_t + {}_2(X_t - X_t^*)D + u_t$

20．当质的因素引进到经济计量模型时，需要使用（　　）

A．外生变量　　B．前定变量　　C．内生变量　　D．虚拟变量

二、多项选择题

1．设$Y_t = \alpha_0 + \alpha_1X_{1t} + \alpha_2X_{2t} + u_t$，$Y_t$=羊毛衫销售量，$X_{1t}$=羊毛衫价格，$X_{2t}$=居民收入。$D$为虚拟变量，$D$=1 代表春秋季，$D$=0 代表冬夏季，则（　　）

A．截距变动模型为：$Y_t = \alpha_0 + \alpha_1 X_{1t} + \alpha_2 X_{2t} + {}_{0}D + u_t$

B．截距变动模型为：$Y_t = \alpha_0 + \alpha_1 DX_{1t} + \alpha_2 X_{2t} + {}_{0}D + u_t$

C．截距斜率同时变动模型为：$Y_t = \alpha_0 + \alpha_1 X_{1t} + \alpha_2 X_{2t} + {}_{0}D + {}_{1}DX_{1t} + u_t$

D．截距斜率同时变动模型为：$Y_t = \alpha_0 + \alpha_1 X_{1t} + \alpha_2 X_{2t} + {}_{0}D + {}_{2}DX_{2t} + u_t$

E．截距斜率同时变动模型为：$Y_t = \alpha_0 + \alpha_1 X_{1t} + \alpha_2 X_{2t} + {}_{0}D + {}_{1}DX_{1t} + {}_{2}DX_{2t} + u_t$

2．虚拟变量取值为 0 和 1 分别代表某种属性是否存在，其中（　）

A．0 表示存在这种属性　　　B．0 表示不存在这种属性

C．1 表示存在这种属性　　　D．1 表示不存在这种属性

E．0 和 1 所代表的内容可以随意设定

3．设消费函数为$Y_i = {}_{0} + {}_{1}D + {}_{2}X_i + u_i$，$Y_i$=第 i 个居民的消费水平，X_i=第 i 个居民的收入水平，D 为虚拟变量，该模型为（　）

A．截距、斜率同时变动模型　B．系统变参数模型的特殊情况。

C．截距变动模型　　　D．斜率变动模型　　　E．分段回归

4．设$Y_i = {}_{0} + {}_{1}D + {}_{2}X_i + {}_{3}(DX_i) + u_t$，式中，$Y_i$＝第 i 个家庭的消费水平，X_i=第 i 个家庭的收入水平，$D = \begin{cases} 1 & \text{城镇居民家庭} \\ 0 & \text{农村居民家庭} \end{cases}$，通过 t 检验进行判断（　）

A．若 ${}_{1} \neq 0,\ {}_{3} \neq 0$，则为截距和斜率同时变动模型。

B．若 ${}_{1} \neq 0,\ {}_{3} = 0$，则为截距变动模型。

C．若 ${}_{1} = 0,\ {}_{3} = 0$，表明有城市和农村居民有着完全相同的消费模式。

D．若 ${}_{1} = 0,\ {}_{3} \neq 0$，则为斜率变动模型。

E．若 ${}_{1} \neq 0,\ {}_{0} = 0,\ {}_{3} \neq 0$，称为无截距项模型

5．在截距变动模型$Y_i = a_0 + a_1 D + \ X_i + u_i$中，有关模型系数叙述正确的为（　）

A．a_o 是基础模型截距项　　B．a_1 是基础模型截距项

C．a_o 为公共截距项　　D．a_1 是公共截距系数

E．a_1 是差别截距系数

三、名词解释

1．虚拟变量　　　　　　　　2．截距变动模型

3．截距斜率同时变动模型　　　4．分段线性回归

四、简答题

1．回归模型引入虚拟变量的一般规则是什么？

2．举例说明截距和斜率同时变动模型的应用。

3．根据某种商品销售量和个人收入季度数据建立如下模型：

$Y_t = {}_0 + {}_1 D_{1t} + {}_2 D_{2t} + {}_3 D_3 + {}_4 D_{4t} + {}_5 X_t + u_t$，其中，虚拟变量 D_{it} 为第 i 季度时为 1，其余为 0，这时会发生什么问题，参数是否能够用最小二乘法估计？

4．举例说明如何建立斜率变动模型。

5．举例说明如何建立分段线性回归模型。

五、计算题

1． 设 $Y_t = {}_0 + {}_1 D_1 + {}_2 D_2 + {}_4 X_t + u_t$，其中，$Y$=大学教师收入，$X$=教学年份，$D_1 = \begin{cases} 1 & 男性 \\ 0 & 女性 \end{cases}$，$D_2 = \begin{cases} 1 & 白人 \\ 0 & 其它人种 \end{cases}$。

请回答以下问题：

(1) ${}_{1,}\ {}_2$ 的含义是什么？

(2)求 $E(Y_t | D_1 = 1, D_2 = 1, X_t)$，并解释其意义。

2．设消费函数为 $Y_t = {}_0 + {}_1 X_t + u_t$，若月收入 X_t 在 1000 元以内和 1000 元以上的边际消费倾向存在显著差异，如何修改原来的模型？

六、分析题

1．设某地区职工工资的收入模型为 $Y_i = {}_1 + {}_2 X_i + u_i$，式中，$Y$=职工工资收入，$X$=工龄，考虑职工收入受受教育程度（高中以下、高中、大专以上）的影响，请引入合适的虚拟变量，并对上述模型加以改进。

2．已知下述模型：$Y_t = \alpha_{1t} + \alpha_{2t} X_t + u_t$，如果参数 α_{1t}, α_{2t} 都是随时间变化而线性变化的，应如何对以上模型进行变换？

3．根据相关数据得到了如下的咖啡需求函数方程：

$$\begin{aligned} L\hat{nY_t} &= 1.2789 - 0.1647 LnX_1 + 0.5115 LnX_2 + 0.1483 LnX_3 - 0.0089T \\ &\quad - 0.0961 D_{1t} - 0.157 D_{2t} - 0.0097 D_{3t} \\ R^2 &= 0.80 \end{aligned}$$

其中，X_1, X_2, X_3，T，D_{1t}, D_{2t}, D_{3t} 的 t 统计量依次为（?2.14)，(1.23)，(0.55)，(?8.36)，(?3.74)，(?6.03)，(?0.37)。Y_t = 人均咖啡消费量，X_1 = 咖啡价格，X_2 = 人均可支配收入，X_3 = 茶的价格， T = 时间变量，D_{it} 为虚拟变量，第 i 季时取值为 1，其余为零。

要求：(1) 模型中 X_1, X_2, X_3 系数的经济含义是什么？

(2) 哪一个虚拟变量在统计上是显著的？

(3) 咖啡的需求是否存在季节效应？

4．家庭消费支出 C 除了依赖于家庭收入 X 之外，还同下列因素有关：

(1)家庭所属民族：有汉、蒙、满、回；

(2)家庭所在地域：有南方，北方；

(3)户主的文化程度：有大专以下、本科、研究生。

试根据以上资料分析确定家庭消费支出的线性回归模型。

参考答案

一、单项选择题

1．A　2．B　3．C　4．A　5．B　6．A　7．A　8．A
9．B　10．C　11．D　12．B　13．B　14．B　15．D　16．A
17．D　18．A　19．D　20．D

二、多项选择题

1．ACDE　2．BCE　3．BC　4．ABCD　5．ACE

三、名词解释

1．虚拟变量：在经济计量分析中，经常会碰到所建模型的被解释变量受到诸如战争、自然灾害、国际环境、季节变动以及政府经济政策变动等质量变量的影响。给定某一质量变量某属性的出现为 1，未出现为 0，称这样的变量为虚拟变量。

2．截距变动模型：在模型 $Y_i = a_0 + a_1 D + \quad X_i + u_i$ 中，D 表示虚拟变量，D=0 和 D=1 表示两种不同的模型，他们的截距不同，则称其为截距变动模型。

3．截距斜率同时变动模型：例如消费函数不但在斜率上有差异，在截距上也是有可能不一致，将两个问题同时考虑进来，我们可以得到回归方程

$$Y_i = \ _0 + \ _1 D + \ _2 X_i + \ _3 (DX_i) + u_i$$

若 $_1 \neq 0$， $_3 \neq 0$，则为截距和斜率同时变动模型

4．分段线性回归：当解释变量 X 的值达到某水平 X^* 之前，与被解释变量 Y 之间存在

某种线性关系；当解释变量 X 的值达到或超过 X^* 以后，与被解释变量的关系就会发生变化。如果已知 X 的转折点 X^*，可以用虚拟变量模型分别估计每一段的斜率。这就是分段线性回归。

四、简答题

1．回归模型引入虚拟变量的一般规则是什么？

答：回归模型引入虚拟变量的一般规则是：

(1) 如果模型中包含截距项，则一个质变量有 m个特征，只需引入（m-1）个虚拟变量。

(2) 如果模型中不包含截距项，则一个质变量有 m个特征，需引入 m个虚拟变量。

2．举例说明截距和斜率同时变动模型的应用。

答：例如城镇居民和农村居民的消费函数不但在斜率上有差异，在截距上也是有可能不一致的，将两个问题同时考虑进来，我们可以得到回归方程

$$Y_i = {}_0 + {}_1 D + {}_2 X_i + {}_3 (DX_i) + u_i$$

式中，Y_i ＝第 i 个家庭的消费水平，X_i ＝第 i 个家庭的收入水平，

$$D = \begin{cases} 1 & \text{城镇居民家庭} \\ 0 & \text{农村居民家庭} \end{cases}$$

方程可以表示为

$$D=1 \qquad Y_i = {}_0 + {}_1 + ({}_2 + {}_3) X_i + u_i$$

$$D=0 \qquad Y_i = {}_0 + {}_2 X_i + u_i$$

${}_1$ 和 ${}_3$ 分别表示城镇居民家庭和农村居民家庭的消费函数在截距和斜率上的差异。我们一般通过 t 检验来判定它们之间是否有差异。

①若 ${}_1 \neq 0$， ${}_3 \neq 0$，则为截距和斜率同时变动模型。

②若 ${}_1 \neq 0$， ${}_3 = 0$， 则为截距变动模型。

③若 ${}_1 = 0$， ${}_3 = 0$ 则表示城镇居民家庭和农村居民家庭有着完全相同的消费模式。

④若 ${}_1 = 0$， ${}_3 \neq 0$ 则为斜率变动模型，这种情况在现实中出现得不是很多。

3．答：会产生共线性，此时不能用最小二乘法估计，因为他违背了引入虚拟变量的一般规则。回归模型引入虚拟变量的一般规则是：① 如果模型中包含截距项，则一个质变量有 m个特征，只需引入（m-1）个虚拟变量。②如果模型中不包含截距项，则一个质变量有 m个特征，需引入 m个虚拟变量。

4．举例说明如何建立斜率变动模型。

答：斜率变动指的是改变了变动的速率或弹性。例如城镇居民家庭与农村居民家庭的消费函数，在边际消费倾向（斜率）上可能会有所不同。回归方程为 $Y_i = {}_1 + {}_2 X_i + {}_3 (DX_i) + u_i$，式中，D 为虚拟变量，则方程可以表示为

$$D=1 \qquad Y_i = {}_1 + ({}_2 + {}_3) X_i + u_i$$

$$D=0 \qquad Y_i = {}_1 + {}_2 X_i + u_i$$

${}_3$ 表示函数在斜率上的差异。我们一般通过 t 检验来判定它们之间是否有差异。若 ${}_3 \neq 0$，则称其为斜率变动模型。

5．举例说明如何建立分段线性回归模型。

答：在经济关系中常有这样的情况：当解释变量 X 的值达到某水平 X^* 之前，与被解释变量 Y 之间存在某种线性关系；当解释变量 X 的值达到或超过 X^* 以后，与被解释变量的关系就会发生变化。

例如：在 1979 年以前，我国居民的消费支出呈缓慢上升的趋势，从 1979 年开始，居民消费支出为快速上升趋势。显然，1979 年是一个转折点，即 $X^* = 1979$。于是，我们可以用以下模型描述我国居民在 1955 年至 2004 年期间消费支出的变动趋势

$$Y_t = {}_0 + {}_1 t + {}_2 (t - X^*) D + u_t$$

Y_t 为某年的消费支出，t 为年份（$t = 1955$，1956，L，2004），D 为虚拟变量，则

$$D = \begin{cases} 1 & t \geq X^* \\ 0 & t < X^* \end{cases}$$

于是，两个不同时期的消费趋势为：

1979 年以前模型为： $Y_t = {}_0 + {}_1 t + u_t$

1979 年以后模型为 ： $Y_t = {}_0 - {}_2 X^* + ({}_1 + {}_2) t + u_t$

五、计算题

1．（1）${}_1$ 代表性别对收入的影响， ${}_2$ 代表肤色对收入的影响

（2）$E(Y_t / D_1 = 1, \mathrm{D}_2 = 1, \mathrm{X}_t) = ({}_0 + {}_1 + {}_2 + {}_3) + {}_4 X_t$，该模型代表男性白种人的收入模型。

2. 设 $X^* = 1000$ ， $\mathrm{D} = \begin{cases} 1 & \mathrm{X}_t \geq 1000 \\ 0 & \mathrm{X}_t < 1000 \end{cases}$，模型可修改为

$$Y_t = {}_0 + {}_1X_t + {}_2(X_t - X^*)D + u_t$$

1000 元以内组的回归模型为　$Y_t = {}_0 + {}_1X_t + u_t$

1000 元以上组的回归模型为　$Y_t = {}_0 - {}_2X^* + ({}_1 + {}_2)X_t + u_t$。

六．分析题

1．解：将职工受教育程度分为高中以下、高中、大专以上，引入 2 个虚拟变量。

设 $D_1 = \begin{cases} 1 & \text{高中} \\ 0 & \text{其它} \end{cases}$，　$D_2 = \begin{cases} 1 & \text{大专以上} \\ 0 & \text{其它} \end{cases}$，

则模型可改进为

$$Y_i = {}_1 + {}_2X_i + \alpha_1 D_{1i} + \alpha_2 D_{2i} + u_i$$

2．解：设 $\alpha_{1t} = a + bt \quad \alpha_{2t} = c + dt$

则 $Y_t = a + bt + (c + dt)X_t + u_t$

即：$Y_t = +a + bt + cX_t + dtX_t + u_t$

3．解：(1) X_1 的系数表示咖啡价格弹性，X_2 的系数表示人均可支配收入的弹性，X_3 的系数表示茶的价格弹性。

(2) D_{1t} 和 D_{2t}，即第一季度和第二季度影响显著。

(3) 存在。

4．解：设 $D_{11} = \begin{cases} 1 & \text{汉族} \\ 0 & \text{其它} \end{cases}$，　$D_{12} = \begin{cases} 1 & \text{蒙族} \\ 0 & \text{其它} \end{cases}$，　$D_{13} = \begin{cases} 1 & \text{满族} \\ 0 & \text{其它} \end{cases}$，

$D_{21} = \begin{cases} 1 & \text{南方} \\ 0 & \text{其它} \end{cases}$，　$D_{31} = \begin{cases} 1 & \text{大专以下} \\ 0 & \text{其　它} \end{cases}$，　$D_{32} = \begin{cases} 1 & \text{本科} \\ 0 & \text{其它} \end{cases}$。

则家庭消费支出模型为

$$C = {}_0 + {}_1D_{11} + {}_2D_{12} + {}_3D_{13} + {}_4D_{21} + {}_5D_{31} + {}_6D_{32} + {}_7X + u\text{。}$$

第七章　联立方程模型

练习题

一、单项选择题

1. 如果联立方程模型中两个结构方程的统计形式完全相同，则下列结论成立的是（　）

A．二者之一可以识别　　B．二者均可识别

C．二者均不可识别　　D．不确定

2．简化式模型中的简化式参数表示（　）

A．内生解释变量对被解释变量的总影响

B．内生解释变量对被解释变量的直接影响

C．前定变量对被解释变量的直接影响

D．前定变量对被解释变量的总影响

3．结构式方程恰好识别是指（　）

A．结构式参数有唯一数值　　B．简化式参数具有唯一数值

C．结构式参数具有多个数值　D．简化式参数具有多个数值

4．结构式方程过度识别是指（　）

A．结构式参数有唯一数值　　B．简化式参数具有唯一数值

C．结构式参数具有多个数值　D．简化式参数具有多个数值

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（　）

A．恰好识别　B．不可识别　C．过度识别　D．不确定

6．下面关于内生变量的表述，错误的是（　）

A．内生变量都是随机变量。

B. 内生变量受模型中其它内生变量和前定变量的影响，同时又影响其它内生变量。

C．在结构式方程中，解释变量可以是前定变量，也可以是内生变量。

D．滞后内生变量与内生变量具有相同性质。

7．下面关于简化式模型的概念，不正确的是（　）

A．简化式方程的解释变量都是前定变量。

B．在同一个简化式模型中，所有简化式方程的解释变量都完全一样。

C．如果一个结构式方程包含一个内生变量和模型系统中的全部前定变量，这个结构式方程就等同于简化式方程。

D．简化式参数是结构式参数的线性函数。

8．下列宏观经济计量模型中投资函数所在方程的类型为（　）

$$Y_t = C_t + I_t + G_t \qquad \text{(定义方程)}$$

$$C_t = \alpha_0 + \alpha_1 Y_t + u_t \qquad \text{(消费函数)}$$

$$I_t = \ _0 + \ _1 Y_{t-1} + \ _2 R_t + v_t \qquad \text{(投资函数)}$$

A．技术方程　B．制度方程　C．恒等式　D．行为方程

9．在联立方程模型中，下列关于工具变量的表述，错误的是（　）

A．工具变量必须与将要替代的内生解释变量高度相关。

B．工具变量必须是模型中的前定变量，与结构方程中的随机误差项不相关。

C．工具变量与所要估计的结构方程中的前定变量之间的相关性必须很弱，以避免多重共线性。

D．若引入多个工具变量，即使工具变量之间存在多重共线性，也不影响估计结果。

10．使用间接最小二乘法估计参数，结构式参数估计量的性质为（　　）

A．无偏、一致　　　　B．有偏、一致

C．无偏、非一致　　　　D．有偏、非一致

11．将内生变量的前期值作解释变量，这样的变量称为（　　）

A．虚拟变量　　B．控制变量　　C．政策变量　　D．滞后变量

二、多项选择题

1．使用普通最小二乘法直接估计联立方程中有内生解释变量方程的结构参数时，会使该估计量（　　）

A．无效的；　B．无偏　C．有偏　D．一致　E．非一致

2．使用间接最小二乘法估计结构式方程参数时必须满足的条件有（　　）

A．结构方程为恰好识别

B．结构方程为过度识别

C．简化式方程的扰动项满足经典假定

D．前定变量之间无完全的多重共线性

E．结构方程中解释变量间无严重多重共线性

3．在下列宏观经济模型中，前定变量包括（　　）

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

$$I_t = \ _0 + \ _1 Y_t + \ _2 Y_{t-1} + \ _3 R_t + v_t$$

$$Y_t = C_t + I_t$$

其中，C=总消费，I=总投资，Y=总收入，R=利率

A．C_t　B．C_{t-1}　C．Y_t　D．R_t　E．Y_{t-1}

4．当结构方程为恰好识别时，可选择的估计方法为（　　）

A．普通最小二乘法　　　　B．广义差分法

C．间接最小二乘法　　　　D．二阶段最小二乘法

E．加权最小二乘法

5．对于二阶段最小二乘法，下列说法正确的是（　　）

A．二阶段最小二乘法对样本容量没有严格要求

B．二阶段最小二乘法仅适合于过度识别方程

C．二队段最小二乘法要求模型的结构式随机干扰项满足经典假定

D．二阶段最小二乘法要求模型中所有前定变量之间无高度多重共线性

E．二阶段最小二乘法估计量不具有一致性

6．关于联立方程模型，下列说法不正确的有（　　）

A．联立方程偏倚实质是内生变量与前定变量的高度相关；

B．只有当模型中所有方程均可识别时，模型才可识别；

C．结构式方程中解释变量必须是外生变量或滞后内生变量；

D．简化式模型中简化式参数反映了解释变量对被解释变量的间接影响；

E．满足经典假定时，简化式参数的最小二乘估计量具有无偏、一致性。

三、名词解释

1．联立方程模型　　　　　2．联立方程偏倚
3．前定变量　　　　　　　4．行为方程
5．结构式模型　　　　　　6．简化式模型
7．结构式参数　　　　　　8．恰好识别
9．过度识别

四、简答题

1．简述联立方程模型的形式。
2．简述识别的定义及识别的分类。
3．简述识别的阶条件与秩条件。
4．简述间接最小二乘法的假设条件及操作步骤。
5．简述工具变量法的应用规则及局限性。
6．简述二阶段最小二乘法的假设条件及操作步骤。

五、计算题

1．考察下面的模型

$$C_t = \alpha_0 + \alpha_1 Y_t + u_t \qquad \text{（消费函数）}$$

$$Y_t = C_t + S_t \qquad \text{（收入恒等式）}$$

其中，C_t, Ｙ$_t$为内生变量，S_t为外生变量。

根据样本观测值得到以下结果：

$$\sum (C-\bar{C})(S-\bar{S}) = 400, \quad \sum (S-\bar{S})^2 = 100, \quad \bar{C}=500, \quad \bar{S}=100$$

要求：用间接最小二乘法估计消费函数。（计算结果保留二位小数）

2．下面模型中

$$C_t = \ {}_1 + \ {}_1 Y_t + u_t$$

$$Y_t = C_t + S_t$$

C=消费，Ｙ=收入，S=储蓄，C, Ｙ为内生变量，S 为外生变量。已知边际消费倾向β_1=0.75。
要求：计算 S 增加一个单位对Ｙ和Ｃ产生的总影响（计算结果保留二位小数）。

3．考虑下述模型

$$C_t = \alpha_1 + \alpha_2 D_t + u_t \qquad \text{（消费方程）}$$

$$I_t = \ {}_1 + \ {}_2 D_{t-1} + v_t \qquad \text{（投资方程）}$$

$$D_t = C_t + I_t + Z_t$$

其中，C=消费支出，D=收入，I=投资，Ｚ=自发支出；C、I 和 D 为内生变量。
要求：（1）写出消费方程的简化式方程；
（2）用阶条件研究各方程的识别问题。

六、分析题

1．设有货币需求和供给模型

货币需求　　$M_t^d = \ {}_0 + \ {}_1 Y_t + \ {}_2 R_t + \ {}_3 P_t + u_{1t}$

货币供给 $M_t^s = \alpha_0 + \alpha_1 Y_t + u_{2t}$

其中，M =货币存量，Y =收入，R =利率，P =价格，R 和 P 是前定变量，$M_t^s = M_t^d = M$ 。

问：(1) 需求函数、供给函数是可识别的吗？

(2) 用什么方法估计可识别方程的参数，为什么？

(3) 如果给供给函数添加解释变量 Y_{t-1} 和 M_{t-1}，模型两方程识别性会发生什么变化？用什么方法估计参数？

2．考虑以下扩展的凯恩斯收入决定模型

$C_t = {}_0 + {}_1Y_t - {}_2T_t + u_{1t}$ （消费函数）

$I_t = \alpha_0 + \alpha_1 Y_{t-1} + u_{2t}$ （投资函数）

$T_t = \gamma_0 + \gamma_1 Y_t + u_{3t}$ （税收函数）

$Y_t = C_t + I_t + G_t$ （收入恒等式）

其中，C=消费支出，Y =收入，I=投资，T=税收。模型中的内生变量是 C, I, T 和 Y，而前定变量是 C, G（政府支出）和 Y_{t-1}。

要求：(1) 用阶条件检查方程组的可识别性；

(2) 如果将外生变量利率 R_t 列入投资函数的解释变量中，整个模型的识别状况如何？

参考答案

一、单项选择题

1．C　2．D　3．A　4．C　5．B　6．D　7．D　8．D

9．D　10．B　11．D

二、多项选择题

1．CE　2．ACD　3．BE　4．CD　5．CD　6．ACD

三、名词解释

1．联立方程模型：联立方程模型是根据经济理论和某些假设条件，区分各种不同的经济变量，建立一组方程式来描述经济变量间的联立关系。

2．联立方程偏倚：在联立方程模型中，一些变量可能在某一方程中作为解释变量，而在另一方程中又作为被解释变量。这就会导致解释变量与随机干扰项之间存在相关关系，从而违背了最小二乘估计理论的一个重要假定。如果直接使用最小二乘法，就会产生所估计的参数是有偏的、非一致的等问题，称为联立性偏误。

3．前定变量：外生变量和滞后内生变量合称为前定变量。前定变量影响现期模型中的其它变量，但不受它们的影响，因此只能在现期的方程中作解释变量，且与其中的随机干扰项

互不相关。

4．行为方程：解释居民、企业和政府的经济行为，描述它们对外部影响是怎样做出反应的方程称为行为方程。

5．结构式模型：每一个方程都把内生变量表示为其他内生变量、前定变量和随机干扰项的函数，描述经济变量关系结构的联立方程组称为结构式模型。

6．简化式模型：把模型中每个内生变量表示为前定变量和随机干扰项的函数，得到的模型称为简化式模型。

7．结构式参数：结构式模型中的参数称为结构式参数，它表示每个解释变量对被解释变量的直接影响，其正负号表示影响的方向，绝对值表示影响的程度。

8．恰好识别：在可识别的模型中，结构式参数具有唯一数值的方程称为恰好识别。

9．过度识别：在可识别的模型中，结构式参数具有多个数值的方程称为过度识别。

四、简答题

1．简述联立方程模型的形式。

答：联立方程模型按方程的形式可分为结构式模型和简化式模型。

每一个方程都把内生变量表示为其他内生变量、前定变量和随机干扰项的函数，描述经济变量关系结构的联立方程组称为结构式模型。

结构式模型中的参数称为结构式参数，它表示每个解释变量对被解释变量的直接影响，其正负号表示影响的方向，绝对值表示影响的程度。

把模型中每个内生变量表示为前定变量和随机干扰项的函数，得到的模型称为简化式模型。

简化式模型中的参数称为简化式参数，简化式参数表达前定变量对内生变量的直接影响和间接影响的总度量。

2．简述识别的定义及识别的分类。

答：定义：若某一结构方程具有唯一的统计形式，则称此方程是可以识别的；否则，就称此结构方程是不可识别的。若线性联立方程中的每个结构方程都是可以识别的，则称此模型是可以识别的；否则，就称此模型是不可识别的。

模型的识别分为可识别和不可识别两类。可识别的模型又分为恰好识别和过度识别两种情况。在可识别的模型中，结构式参数具有唯一数值的方程称为恰好识别；结构式参数具有多个数值的方程称为过度识别。

3．简述识别的阶条件与秩条件。

答：设结构式模型所含方程的总数（或内生变量总数）为 M，模型包含的变量总数（包括前定变量和内生变量）为 H，待识别的方程包含的变量总数（包括内生变量和前定变量）为 G。

阶条件：若某一个结构式方程是可以识别的，则模型中方程数减一小于或等于此方程排

斥的变量总数，即 M－1≤H－G，若 $M-1>H-G$，则不可识别；$M-1=H-G$，则为恰好识别；$M-1<H-G$，则为过度识别。阶条件是必要条件，不是充分条件。

秩条件：在具有 M 个方程的结构式模型中，任何一个方程可以识别的充分必要条件是：不包括在该方程中的变量（包括内生变量和前定变量）的参数所组成的矩阵（记为 A）的秩为 $M-1$，即 $r(A)=M-1$。秩条件是充分必要条件，也就是说：如果秩条件成立，则方程是可识别的；如果方程是可识别的，则秩条件成立，或者秩条件不成立，则方程是不可识别的。

4．简述间接最小二乘法的假设条件及操作步骤。

答：简化式方程的解释变量均为前定变量，无联立性偏误问题，可以使用普通最小二乘法估计简化式参数，从而导出结构式参数，这就是间接最小二乘法。

间接最小二乘法有以下三条假设条件：

（1）被估计的结构方程必须是恰好识别的。

（2）简化式方程的随机干扰项必须满足最小二乘法的假定。

（3）前定变量之间不存在完全的多重共线性。

间接最小二乘法包括以下三个步骤：

第一步，将结构式模型化为简化式模型。也就是把每一个内生变量表示为前定变量和随机干扰项的函数。

第二步，对简化式模型的各方程用最小二乘法估计参数，从而得到简化式参数估计值。

第三步，把简化式参数的估计值代入结构式参数与简化式参数的关系式，求得结构式参数的估计值。

5．简述工具变量法的应用规则及局限性。

答：工具变量法的就是用合适的预定变量作为工具变量代替结构方程中的内生变量，从而降低解释变量与随机误差项之间的相关程度，再利用最小二乘法进行参数估计。

应用规则：如果内生解释变量 Y_t 与 u_t 相关，我们就选择一个工具变量 Z_t 来代替 Y_t。Z_t 要满足两个条件：一是 Z_t 与 u_t 不相关，即 $Cov(Z_t, u_t)=0$；二是 Z_t 与 Y_t 高度相关，即 $Cov(Z_t, Y_t)\neq 0$。在联立方程模型中，工具变量一般从外生变量中选取。

工具变量法的局限性

（1）在实践中，找到一个既有经济意义，又满足两个条件的工具变量非常困难。

（2）若满足两个条件的工具变量有多个时，在选择方面具有任意性。

（3）检验工具变量与随机误差项不相关有很大困难。

6．简述二阶段最小二乘法的假设条件及操作步骤。

答：二阶段最小二乘法的就是将内生解释变量对联立方程模型中所有外生变量回归，得到内生解释变量的估计值（拟合值），将这个估计值（拟合值）作为工具变量，对结构方

程使用普通最小二乘法。

二阶段最小二乘法必须满足以下假设条件：

（1）被估计的结构式方程必须是可识别的，特别地，二阶段最小二乘法适合于过度识别方程。

（2）结构式模型中的各随机干扰项必须满足最小二乘法经典假定，即零期望值、同方差、无自相关且与全部前定变量无关。

（3）所有前定变量之间不存在高度多重共线性。

（4）解释变量之间不是完全共线性的。

（5）样本容量足够大。

二阶段最小二乘法的步骤

第一阶段：将待估计方程中的内生解释变量Y_t对联立方程模型中的全部前定变量回归，即估计简化式方程，计算内生解释变量Y_t的估计值$\hat{Y}_t$。

第二阶段：用第一阶段得到的内生解释变量的估计值$\hat{Y}_t$代替内生解释变量Y_t，对该结构方程使用普通最小二乘法估计结构式参数。

五、计算题

1．解：消费模型的简化式方程为

$$C_t = \pi_0 + \pi_1 S_t + v_{1t}$$

其中，$\pi_0 = \dfrac{\alpha_0}{1-\alpha_1}, \pi_1 = \dfrac{\alpha_1}{1-\alpha_1}$，解得：$\alpha_0 = \dfrac{\pi_0}{1+\pi_1}, \alpha_1 = \dfrac{\pi_1}{1+\pi_1}$

对消费方程简化式使用普通最小二乘法估计参数

$$\hat{\pi}_1 = \frac{\sum (C-\bar{C})(S-\bar{S})}{\sum (S-\bar{S})^2} = \frac{400}{100} = 4$$

$$\hat{\pi}_0 = \bar{C} - \hat{\pi}_1 \bar{S} = 500 - 4\times 100 = 100$$

则 $$\hat{\alpha}_0 = \frac{\pi_0}{1+\pi_1} = \frac{100}{1+4} = 20$$

$$\hat{\alpha}_1 = \frac{\pi_1}{1+\pi_1} = \frac{4}{1+4} = 0.8$$

因此，消费函数为$\hat{C}_t = 20 + 0.8Y_t$

2．解：联立方程的简化式为

$$C_t = \pi_0 + \pi_1 S_t + v_{1t}$$
$$Y_t = \pi_2 + \pi_3 S_t + v_{2t}$$

其中，$\pi_0 = \frac{{}_0}{1-\ \ {}_1}$£¬$\pi_1 = \frac{{}_1}{1-\ \ {}_1}$，$\pi_2 = \frac{{}_0}{1-\ \ {}_1}$£¬$\pi_3 = \frac{1}{1-\ \ {}_1}$

π_1，π_3 表示外生变量 S 对消费 C 和收入 Y 的总影响。

$$\pi_1 = \frac{0.75}{1-0.75} = 3\text{£¬}\pi_3 = \frac{1}{1-0.75} = 4$$

S 增加一个单对 Y 的总影响为 4 个单位，对 C 的总影响为 3 个单位。

3．解：(1）消费方程的简化式方程为

$$C_t = \pi_0 + \pi_1 D_{t-1} + \pi_2 Z_t + v_{1t}$$

其中，$\pi_0 = \frac{\alpha_{1\ \ } + \alpha_{2\ \ 1}}{1-\alpha_2}$£¬$\pi_1 = \frac{\alpha_{2\ \ 2}}{1-\alpha_2}$£¬$\pi_2 = \frac{\alpha_2}{1-\alpha_2}$£∀${}_{1t} = \frac{\alpha_2 v_t + u_t}{1-\alpha_2}$

（2）对于消费方程

阶条件：排除的变量数=（5-2）=3＞（方程个数-1）=2，方程为过度识别。

对于投资方程

排除的变量数=（5-2）=3＞（方程个数-1）=2，该方程为过度识别。

六、分析题

1．解：(1）货币需求方程

排除的变量数=0＜（方程个数-1）=1，所以，该方程不可识别。

货币供给方程

阶条件：排除的变量个数=2＞（方程个数-1）=1，该方程为过度识别。

秩条件：排除的系数矩阵的秩为 1=（方程个数-1)，该方程可识别。

（2）货币供给方程可识别，因为是过度识别，应使用二阶段最小二乘法估计参数。

（3）供给方程添加变量 Y_{t-1} 和 M_{t-1} 后，识别状态为

货币需求方程

阶条件：排除的变量数=2＞（方程个数-1）

秩条件：排除的变量系数矩阵的秩为 1=（方程个数-1）

货币需求方程可识别且为过度识别。

货币供给方程

阶条件：排除的变量个数=2＞（方程个数-1）=1

秩条件：排除的变量系数矩阵的秩为 1=（方程个数-1)，货币供给方程可识别且为过度识别。

2．解：(1）阶条件：方程数 M=4，变量总数 H=6，消费函数变量数 H_1=3，投资函数变量数 H_2=2，税收函数变量数 H_3= 2。

$H-H_1$=3=M-1=3　　消费函数恰好识别

$H-H_2$=4＞M-1=3　　投资函数过度识别

$H-H_3$=4＞M-1=3　　税收函数过度识别

恒等式不存在识别问题，联立方程模型可识别。

（2）M=4　H=7　H_1=3　H_2=3　H_3=2

H-H_2=4＞M-1=3　　　　　　消费函数过度识别
H-H_2=4＞M-1=3　　　　　　投资函数过度识别
H-H_3=5＞M-1=3　　　　　　税收函数过度识别
恒等式不存在识别问题，联立方程模型可识别。

《经济计量分析》模拟试题一

题号	一	二	三	四	五	六	总分
分值	20	15	20	15	20	10	100
得分							

得分	评阅人

一、单项选择题（共 20 题，每题 1 分）

1．经济计量模型的被解释变量一定是（　）
　A．控制变量　　　　　　　B．政策变量
　C．内生变量　　　　　　　D．外生变量

2．在同一时点或时期上，不同统计单位的相同统计指标组成的数据是（　）
　A．时期数据　　　　　　　B．时点数据
　C．时序数据　　　　　　　D．截面数据

3．设截距和斜率同时变动模型为 $Y_i = {}_0 + {}_1 D + {}_2 X + {}_3 (DX) + u_i$

下面那种情况成立，该模型为截距变动模型（　）

　A． ${}_1 \neq 0,\ {}_3 \neq 0$　　　　B． ${}_1 \neq 0,\ {}_3 = 0$

　C． ${}_1 = 0,\ {}_3 = 0$　　　　D． ${}_1 = 0,\ {}_3 \neq 0$

4．设个人消费函数 $Y_i = {}_1 + {}_2 X_i + u_i$ 中，消费支出 Y 不仅与收入 X 有关，而且与消费者的性别、年龄构成有关，年龄构成可以分为老、中、青三个层次，假定边际消费倾向不变，该消费函数引入虚拟变量的个数为（　　）

　A．1 个　　　B．2 个　　　C．3 个　　　D．4 个

5．下列哪一个不是几何分布滞后模型的变换模型（　）
　A．库伊克变换模型　　　　B．自适应预期模型
　C．部分调整模型　　　　　D．有限多项式滞后模式

6．如果 $|e_i|$ 与 X_i 之间存在线性关系，则认为异方差的形式为（　）

A．$\sigma_i^2 = \sigma^2 X_i$　　　　B．$\sigma_i^2 = \sigma^2 X_i^2$

C．$\sigma_i^2 = \sigma^2$　　　　D．$\sigma_i^2 = \sigma^2 \sqrt{X_i}$　（σ^2为常数）

7．戈德菲尔德—匡特检验适用于检验（　　）

A．序列相关　　　　B．异方差

C．多重共线性　　　　D．设定误差

8．DW 统计量的取值范围是（　）

A．$-1 \le DW \le 0$　　　　B．$-1 \le DW \le 1$

C．$-2 \le DW \le 2$　　　　D．$0 \le DW \le 4$

9．在线性回归模型中，若解释变量 X_1 和 X_2 的观测值成比例，即有 $X_{1i} = kX_{2i}$，其中 k 为非零常数，则表明模型中存在（　）

A．异方差　　B．多重共线性　　C．序列相关　　D．设定误差

10．设分布滞后模型 $Y_t = \alpha + {}_o X_t + {}_1 X_{t-1} + {}_2 X_{t-2} + {}_3 X_{t-3} + u_t$ 中，为了使模型的自由度达到 30，必须拥有多少年的观测资料（ ）

A．32　　　B．33　　　C．35　　　D．38

11．设消费函数为 $Y_i = {}_0 + {}_1 X_i + \alpha_1 D_1 + \alpha_2 D_2 + a_3 D_3 + u_i$，其中 Y 为消费，X 为收入，$D_1 = \begin{cases} 1 & \text{第一季度} \\ 0 & \text{其它} \end{cases}$，$D_2 = \begin{cases} 1 & \text{第二季度} \\ 0 & \text{其它} \end{cases}$，$D_3 = \begin{cases} 1 & \text{第三季度} \\ 0 & \text{其它} \end{cases}$，该模型包含了几个质因素(　　)

A．1　　　B．2　　　C．3　　　D．4

12．在回归分析中，有关被解释变量 Y 和解释变量 X 的说法正确的为（　　）

A．Y 为随机变量，X 为非随机变量；　　B．Y 为非随机变量，X 为随机变量；

C．X、Y 均为随机变量；　　D．X、Y 均为非随机变量。

13．最小二乘准则是指（　　）

A．随机误差项 u_i 的平方和最小；

B．Y_i 与它的期望值 $\bar{Y}$ 的离差平方和最小；

C．X_i 与它的均值 $\bar{X}$ 的离差平方和最小；

D．残差 e_i 的平方和最小。

14．在一元线性回归模型中，σ^2 的无偏估计量 $\hat{\sigma}^2$ 为（　　）

A．$\frac{\sum e_i^2}{n}$　　　　B．$\frac{\sum e_i^2}{n-1}$

C．$\frac{\sum e_i^2}{n-2}$　　　　D．$\frac{\sum e_i^2}{n-3}$

15．在多元线性回归模型中，调整后的判定系数 $\bar{R}^2$ 与判定系数 R^2 的关系为（　　）

A．$R^2 < \bar{R}^2$　　B．$\bar{R}^2 < R^2$　　C．$R^2 \le \bar{R}^2$　　D．$\bar{R}^2 \le R^2$

16．在回归模型 $Y_i=\beta_1+\beta_2X_2+\beta_3X_3+\beta_4X_4+u_i$ 中，X_3 与 X_4 高度相关，X_2 与 X_3、X_4 无关，则因为 X_3 与 X_4 的高度相关会使 $\hat{}_2$ 的方差（　　）

A．变大　　　　B．变小

C．不确定　　　　D．不受影响

17．根据判定系数 R^2 与 F 统计量的关系可知，当 R^2=1 时，有（ ）

A．F=1　　　　B．F=　1

C．F=0　　　　D．$F=\infty$

18．如果联立方程模型中两个结构方程的统计形式完全相同，则下列结论成立的是（　）

A．二者之一可以识别　　　　B．二者均可识别

C．二者均不可识别　　　　D．不确定

19．简化式模型中的简化式参数表示（ ）

A．内生解释变量对被解释变量的总影响。

B．内生解释变量对被解释变量的直接影响。

C．前定变量对被解释变量的直接影响。

D．前定变量对被解释变量的总影响。

20．结构式方程过度识别是指（　）

A．结构式参数具有唯一数值　　　　B．简化式参数具有唯一数值

C．结构式参数具有多个数值　　　　D．简化式参数具有多个数值

得分	评阅人

二、多项选择题（共 5 题，每题 3 分）

1．经济计量模型的应用在于（　）

A．设定模型　　　　B．检验模型

C．结构分析　　　　D．经济预测

E．规划政策

2．对于有限分布滞后模型，对它应用最小二乘法估计时存在以下困难（　　）

A．产生多重线性　　B．产生异方差　　C．产生自相关

D．损失自由度　　E．最大滞后期 k 较难确定

3．存在异方差条件下普通最小二乘法具有如下性质（　　）

A．线性性　　　　B．无偏性

C．最小方差性　　　　D．有偏性　　　　E．无效性

4．对回归模型进行显著性检验时所用的 F 统计量可表示为（　　）

A．$\dfrac{ESS/(n-k)}{RSS/(k-1)}$　　B．$\dfrac{ESS/(k-1)}{RSS/(n-k)}$

C．$\dfrac{R^2/(k-1)}{(1-R^2)/(n-k)}$　　D．$\dfrac{(1-R^2)/(n-k)}{R^2(k-1)}$　　E．$\dfrac{R^2/(n-k)}{(1-R^2)/(k-1)}$

5．判定系数 R^2 可表示为（　　）

A．$R^2=\dfrac{RSS}{TSS}$　　B．$R^2=\dfrac{ESS}{TSS}$

C．$R^2=1-\dfrac{RSS}{TSS}$　　D．$R^2=1-\dfrac{ESS}{TSS}$

E．$R^2=\dfrac{ESS}{ESS+RSS}$

得分	评阅人

三、名词解释（共 5 题，每题 4 分）

1．多重共线性

2．分布滞后模型

3．虚拟变量

4．外生变量

5．判定系数

得分	评阅人

四、简答题（共 3 题，每题 5 分）

1．简述经济计量分析的研究步骤。

2．试写出 DW 检验的判断区间。

3．简述识别的阶条件与秩条件。

得分	评阅人

五、计算题（共 2 题，每题 10 分）

1．现有 X 和 Y 的样本观测值如下：

X	1	2	3	5	10
Y	2	4	2	10	5

假设 Y 对 X 的回归方程为 $Y_i = {}_0 + {}_1X_i + u_i$，且 $Var(u_i)=\sigma^2X_i^2$，试用适当方法估计此回归方程。(保留两位小数)

2．有一参数个数 k=4 的线性回归模型，用一个容量为 n=20 的时序数据样本进行普通最小二乘估计，得到如下资料：$\sum_{t=1}^{n}e_t^2=40$，$\sum_{t=2}^{n}e_t^2=39$，$\sum_{t=2}^{n}e_{t-1}^2=36$,

$\sum_{t=2}^{n} e_t e_{t-1} = 20$。

试根据这些资料计算 DW 统计量。

得分	评阅人

六、分析题（共 1 题，每题 10 分）

在研究生产函数时，我们得到如下两个模型

模型 I ：
$$Ln\hat{\theta} = -5.040 + 0.887LnK + 0.893LnL$$
$$Se \quad (1.400) \quad (0.087) \qquad (0.137)$$

$$R^2 = 0.878 \qquad n = 21$$

模型 II ：
$$Ln\hat{\theta} = -8.570 + 0.027t + 0.460LnK + 1.285LnL$$
$$Se \quad (2.99) \quad (0.021) \quad (0.333) \qquad (0.324)$$

$$R^2 = 0.889 \qquad n = 21$$

其中，θ =产量，K=资本，L=劳动时数，t=时间，n=样容量，模型下括号内为参数估计的标准误。[$\alpha = 0.05$ $\quad t_{\alpha/2}(18) = 2.101$ $\quad t_{\alpha/2}(17) = 2.110$]

请回答以下问题：

(1)说明模型 I 中所有系数在统计上都是显著的。

(2)说明模型 II 哪些参数的系数在统计上是不显著的。

(3)若 t和 LnK之间相关系数为 0.97，你将从中得出什么结论？

(4)模型 I 的规模报酬为多少？

《经济计量分析》模拟试题一答案

一、单项选择题

1-5．CDBCD　6-10．BBDBD　11-15．AADCB　16-20．DDCDC

二、多项选择题

1．CDE　2．ADE　3． ABE　4． BC　5． BCE

三、名词解释

1．多重共线性：在多元线性回归模型中，解释变量 $X_1, X_2, \cdots, X_k$ 之间存在完全或近似的线性关系。称解释变量 $X_1, X_2, \cdots, X_k$ 之间存在完全或近似多重共性线。也称为复共线性。

2．分布滞后模型：分布滞后模型一般定义为 如果一个回归模型不仅包含解释变量现期值，而且还包含解释变量的滞后值，则这个回归模型就是分布滞后模型。它的一般形式为：

$$Y_t = \alpha + {}_0X_t + {}_1X_{t-1} + \cdots + {}_kX_{t-k} + u_t$$

或 $$Y_t = \alpha + {}_0X_t + {}_1X_{t-1} + \cdots + u_t$$

3．虚拟变量：在经济计量分析中，经常会碰到所建模型的被解释变量受到诸如战争、自然灾害、国际环境、季节变动以及政府经济政策变动等质量变量的影响。给定某一质量变量某属性的出现为 1，未出现为 0，称这样的变量为虚拟变量。

4．外生变量：是非随机变量，在模型体系之外决定，即在模型求解之前已经得到了数值。

5．判定系数：$R^2 = \dfrac{\sum(\hat{Y}_i - \bar{Y})^2}{\sum(Y_i - \bar{Y})^2} = \dfrac{ESS}{TSS}$，是对回归线拟合优度的度量。$R^2$ 测度了在 Y 的总变异中由回归模型解释的那个部分所占的比例或百分比。

四、简答题

1．答：用经济计量方法研究社会经济问题是以经济计量模型的建立和应用为基础的，其过程可分为四个连续的步骤：建立模型、估计参数、验证模型和使用模型。

建立模型是根据经济理论和某些假设条件，区分各种不同的经济变量，建立单一方程式或方程体系，来表明经济变量之间的相互依存关系。

模型建立后，必须对模型的参数进行估计；就是获得模型参数的具体数值。

模型估计之后，必须验证模型参数估计值在经济上是否有意义，在统计上是否令人满意。

对经济现象的计量研究是为了使用经济计量模型。经济计量模型的使用主要是用于进行经济结构分析、预测未来和制定或评价经济政策。

2．答：DW 检验的决策区间

$0 \le DW \le d_L$	误差项 $u_1, u_2, \cdots, u_n$ 间存在正相关
$d_L < DW \le d_U$	不能判定是否有自相关
$d_U < DW < 4 - d_U$	误差项 $u_1, u_2, \cdots, u_n$ 间无自相关
$4 - d_U \le DW < 4 - d_L$	不能判定是否有自相关
$4 - d_L \le DW \le 4$	误差项 $u_1, u_2, \cdots, u_n$ 间存在负相关

3．答：答：设结构式模型所含方程的总数（或内生变量总数）为 M，模型包含的变量总数（包括前定变量和内生变量）为 H，待识别的方程包含的变量总数（包括内生变量和前定变量）为 G。

阶条件：若某一个结构式方程是可以识别的，则模型中方程数减一小于或等于此方程排斥的变量总数，即 M－1≤H－G，若 $M - 1 > H - G$，则不可识别；$M - 1 = H - G$，则为恰好识别；$M - 1 < H - G$，则为过度识别。阶条件是必要条件，不是充分条件。

秩条件：在具有 M 个方程的结构式模型中，任何一个方程可以识别的充分必要条件是：不包括在该方程中的变量（包括内生变量和前定变量）的参数所组成的矩阵（记为 A）的秩

为 $M-1$，即 r（A）$=M-1$。秩条件是充分必要条件，也就是说：如果秩条件成立，则方程是可识别；如果方程是可识别的，则秩条件成立，或者秩条件不成立，则方程是不可识别的。

五、计算题

1．解：$Y_i = {}_0 + {}_1 X_i + u_i$

两边同除以 X_i 得：$\frac{Y_i}{X_i} = {}_0 \frac{1}{X_i} + {}_1 + \frac{u_i}{X_i}$

$Var(\frac{u_i}{X_i}) = \frac{1}{X_i^2} Var(u_i) = \sigma^2$ 消除了异方差，可用普通最小二乘法估计参数。

$$E(\frac{1}{X}) = \frac{\sum_{i=1}^{5} \frac{1}{X_i}}{5} = 0.41, \qquad E(\frac{Y}{X}) = \frac{\sum_{i=1}^{5} \frac{Y_i}{X_i}}{5} = 1.40$$

$$\hat{}_0 = \frac{\sum_{i=1}^{5} (\frac{1}{X_i} - 0.41)(\frac{Y_i}{X_i} - 1.40)}{\sum_{i=1}^{5} (\frac{1}{X_i} - 0.41)^2} = \frac{0.71}{0.52} = 1.37$$

$$\hat{}_1 = 1.4 - 1.37 \times 0.41 = 0.83$$

回归方程为 $\hat{Y} = 1.37 + 0.83X$

2．解：

$$DW = \frac{\sum_{t=2}^{n} (e_t - e_{t-1})^2}{\sum_{t=1}^{n} e_t^2}$$

$$= \frac{\sum_{t=2}^{n} e_t^2 + \sum_{t=2}^{n} e_{t-1}^2 - 2\sum_{t=1}^{n} e_t e_{t-1}}{\sum_{t=1}^{n} e_t^2}$$

$$= \frac{39 + 36 - 2 \times 20}{40} = \frac{35}{40} = 0.875$$

六、分析题

解：(1)计算各参数的 t 统计量分别为 -3.60, 10.20, 6.52 绝对值全部大于临界值 2.101，所以各偏回归系数均显著。

(2)计算各参数的 t 统计量分别为 -2.87, 1.29, 1.38, 3.97，t 和 L nK 的参数所对应的 t 统计量绝对值小于临界值 2.111，所以不显著。

(3)模型 II 存在多重共线性问题。

$(4)\hat{\alpha}+\hat{\ }=0.887+0.893=1.78$，规模报酬递增。

经济计量分析》模拟试题二

题号	一	二	三	四	五	六	总分
分值	20	15	20	15	20	10	100
得分							

得分	评阅人

一、单项选择题（共 20 题，每题 1 分）

1．设无限分布滞后模型为 $Y_t=\alpha+\ _oX_t+\ _1X_{t-1}+\cdots+u_t$，该模型满足库伊克（koyck）提出的两个假设，则长期影响乘数为（　）

A． $_o\lambda^k$　　B． $\frac{_o}{1-\lambda}$　　C． $\frac{(1-\lambda^k)\ _0}{1-\lambda}$　　D．不确定

2．在一个包含截距项的回归模型 $Y_i=\ _1+\ _2X_i+u_i$ 中，如果一个具有 m 个特征的质的因素引入 m 个虚拟变量，则会产生的问题为（　　）

A．异方差　　B．序列相关

C．不完全多重线性相关　　D．完全多重线性相关

3．有效估计量是指（　）

A．在所有线性无偏估计量中方差最大

B．在所有线性无偏估计量中变异系数最小

C．在所有线性无偏估计量中方差最小

D．在所有线性无偏估计量变异系数最大

4．回归系数 $_2$ 通过了 t 检验，表示（　）

A． $_2\neq 0$　　B． $\hat{\ }_2\neq 0$

C． $_2\neq 0$，$\hat{\ }_2=0$　　D． $_2=0$，$\hat{\ }_2\neq 0$

5．在回归模型 $Y=\beta_1+\beta_2X_2+\beta_3X_3+\beta_4X_4+u$ 中，X_3 与 X_4 高度相关，X_2 与 X_3、X_4 无关，则因为 X_3 与 X_4 的高度相关会使 $\hat{\ }_2$ 的方差（　）

A．变大　　B．变小　　C．不确定　　D．不受影响

6．下列哪种情况属于存在序列相关（　）

A．$Cov(u_i,u_j)=0$，$i \neq j$　　　B．$Cov(u_i,u_j)\neq 0$，$i \neq j$

C．$Cov(u_i,u_j)=\sigma^2$，$i = j$　　　D．$Cov(u_i,u_j)=\sigma_i^2$，$i = j$

7．DW 的取值范围是（　）

A．$-1\leq DW \leq 0$　　　B．$-1\leq DW \leq 1$

C．$-2\leq DW \leq 2$　　　D．$0\leq DW \leq 4$

8．采用一阶差分法估计一阶自相关模型，适合于（　）

A．≈ 0　　B．≈ 1　　C．$-1< \quad <0$　　D．$0< \quad <1$

9．设截距和斜率同时变动模型为$Y_i = {}_0 + {}_1D + {}_2X + {}_3(DX)+u_i$

下面那种情况成立时，该模型为截距变动模型（　）

A．${}_1 \neq 0,\quad {}_3 \neq 0$　　　B．${}_1 \neq 0,\quad {}_3 = 0$

C．${}_1 = 0,\quad {}_3 = 0$　　　D．${}_1 = 0,\quad {}_3 \neq 0$

10．设消费函数为 $\hat{C}_t = {}_0 + {}_1X + {}_2D + u_t$，$C$ 为消费，X 为收入，$D=\begin{cases}1 & \text{城镇居民}\\ 0 & \text{农村居民}\end{cases}$，如果统计检验 ${}_2 \neq 0$ 成立，则城镇居民消费函数和农村居民消费函数是（　）

A．相互平行的　　B．相互垂直的　　C．相互交叉的　　D．相互重叠的

11．根据判定系数 R^2 与 F 统计量的关系可知，当 R^2=1 时，有（　）

A．F=1　　　B．F=?‘1

C．F=0　　　D．$F=\infty$

12．对回归系数进行显著性检验时的 t 统计量为（　）

A．$\frac{{}_j}{Se(\hat{}_j)}$　　　B．$\frac{\hat{}_j}{Var(\hat{}_j)}$

C．$\frac{{}_j}{Var(\hat{}_j)}$　　　D．$\frac{\hat{}_j}{Se(\hat{}_j)}$

13．结构式方程恰好识别是指（　）

A．结构式参数有唯一数值　　　B．简化式参数具有唯一数值

C．结构式参数具有多个数值　　　D．简化式参数具有多个数值

14．下列宏观经济计量模型中投资函数所在方程的类型为（　）

$Y_t = C_t + I_t + G_t$　　　（定义方程）

$C_t = \alpha_0 + \alpha_1 Y_t + u_t$　　　（消费函数）

$I_t = \ _0 + \ _1 Y_{t-1} + \ _2 R_t + v_t$　　　　　（投资函数）

A．技术方程　　　B．制度方程　　　C．恒等式　　　D．行为方程

15．使用间接最小二乘法估计参数，结构式参数估计量的性质为（　）

A．无偏、一致　　　　　　　B．有偏、一致

C．无偏、非一致　　　　　　D．有偏、非一致

16．在同一时点或时期上，不同统计单位的相同统计指标组成的数据是（　）

A．时期数据　　B．时点数据　　C．时序数据　　D．截面数据

17．经济计量模型的被解释变量一定是（　）

A．控制变量　　B．政策变量　　C．内生变量　　　D．外生变量

18．设个人消费函数 $Y_i = \ _1 + \ _2 X_i + u_i$ 中，消费支出 Y 不仅与收入 X 有关，而且与消费者的性别、年龄构成有关，年龄构成可以分为老、中、青三个层次，假定边际消费倾向不变，该消费函数引入虚拟变量的个数为（　　）

A．1 个　　　B．2 个　　　C．3 个　　　　D．4 个

19．下列哪一个不是几何分布滞后模型的变换模型（　）

A．库伊克变换模型　　　　　　B．自适应预期模型

C．部分调整模型　　　　　　　D．有限多项式滞后模式

20．戈德菲尔德—匡特检验适用于检验（　　）

A．序列相关　　B．异方差　　C．多重共线性　　D．设定误差

得分	评阅人

二、多项选择题（共 5 题，每题 3 分）

1．对于有限分布滞后模型，对它应用最小二乘法估计时存在以下困难（　）

A．产生多重线性　　　B．产生异方差　　　　C．产生随机解释变量

D．损失自由度　　　　E．最大滞后期 k 较难确定

2．存在异方差条件下普通最小二乘法具有如下性质（　　）

A．线性性　　　　　　B．无偏性

C．最小方差性　　　　D．有偏性　　　　　E．无效性

3．使用间接最小二乘法估计结构式方程参数时必须满足的条件有（　）

A．结构方程为恰好识别

B．结构方程为过度识别

C．简化式方程的扰动项满足经典假定

D．前定变量之间无完全的多重共线性

E．结构方程中解释变量间无严重多重共线性

4．对回归模型进行显著性检验时所用的 F 统计量可表示为（　　）

A．$\dfrac{ESS/(n-k)}{RSS/(k-1)}$　　　B．$\dfrac{ESS/(k-1)}{RSS/(n-k)}$

C. $\dfrac{R^2/(k-1)}{(1-R^2)/(n-k)}$　　D. $\dfrac{(1-R^2)/(n-k)}{R^2(k-1)}$　　E. $\dfrac{R^2/(n-k)}{(1-R^2)/(k-1)}$

5．当结构方程为恰好识别时，可选择的估计方法为（　）

A．普通最小二乘法　　B．广义差分法

C．间接最小二乘法　　D．二阶段最小二乘法

E．加权最小二乘法

得分	评阅人

三、名词解释（共 5 题，每题 4 分）

1．序列相关　　2．分布滞后模型

3．虚拟变量　　4．内生变量　　5．多元线性回归模型

得分	评阅人

四、简答题（共 3 题，每题 5 分）

1．试述经典线性回归模型的经典假定。

2．试写出 DW 检验的判断区间。

3．简述识别的阶条件与秩条件。

得分	评阅人

五、计算题 （共 2 题，每题 10 分）

1．设 $Y_t = \alpha + {}_o X_t + {}_1 X_{t-1} + {}_2 X_{t-2} + {}_3 X_{t-3} + {}_4 X_{t-4} + u_t$,假设用 2 阶有限多项式变换估计这个模型，得到：

$\hat{\alpha}_0 = 0.5$ ， $\hat{\alpha}_1 = 0.45$ ， $\hat{\alpha}_2 = -0.1$。

(1) 求 $_0$， $_1$， $_2$， $_3$， $_4$的估计值

(2) 求 X 对 Y 的短期影响乘数、长期影响乘数和延期影响乘数

2．以 1978~1997 年中国某地区进口总额Y（亿元）为被解释变量，以地区生产总值 X（亿元）为解释变量进行回归，得到回归结果如下：

$$\hat{Y}_t = -261.09 + 0.2453 X_t$$

Se=(31.327)　　(　　)

t=(　　　)　　(16.616)

R^2=0.9388　　　n=20

要求：(1) 将括号内缺失的数据填入；

(2) 如何解释系数 0.2453 和系数－261.09。

得分	评阅人

六、分析题（共 1 题，每题 10 分）

家庭消费支出 C 除了依赖家庭收入 X 之外，还同下列因素有关：

(1)家庭所属民族：有汉、蒙、满、回；

(2)家庭所在地域：有南方，北方；

(3)户主的文化程度：有大专以下、本科、研究生；

试根据以上资料分析确定家庭消费支出的线性回归模型。

《经济计量分析》模拟试题二答案

一、单项选择题

1-5．BDCAD　　6-10．BDBBA　　11-15．DDADB　　16-20．　DCCDB

二、多项选择题

1．ADE　　2．ABE　　3．ACD　　4．BC　　5．　CD

三、名词解释

1．序列相关：如果一个回归模型不同时点的误差项之间相关，即 $Cov(u_i, u_j) \neq 0, i \neq j$，则称随机误差项之间存在着序列相关现象，也称为自相关。

2．分布滞后模型：如果一个回归模型不仅包含解释变量的现期值，而且还包含解释变量的滞后值，则这个回归模型就是分布滞后模型。它的一般形式为：

$$Y_t = \alpha + {}_0X_t + {}_1X_{t-1} + \mathrm{L} + {}_kX_{t-k} + u_t$$

或 $$Y_t = \alpha + {}_0X_t + {}_1X_{t-1} + \mathrm{L} + u_t$$

3．虚拟变量：在经济计量分析中，经常会碰到所建模型的被解释变量受到诸如战争、自然灾害、国际环境、季节变动以及政府经济政策变动等质量变量的影响。给定某一质量变量某属性的出现为 1，未出现为 0，称这样的变量为虚拟变量。

4．内生变量：具有一定概率分布的随机变量，由模型自身决定，其数值是求解模型的结果。

5．多元线性回归模型：在回归模型中包含二个以上解释变量的线性回归模型。

四、简答题

1．答：对于总体线性回归模型，其经典假定如下。

假定 1：误差项 u_i 的均值为零。

假定 2：同方差性或 u_i 的方差相等。对所有给定的 X_i，u_i 的方差都是相同的。

假定 3：各个误差项之间无自相关，u_i 和 u_j（i≠j）之间的相关为零。

假定 4：u_i 和 X_i 的协方差为零。

该假定表示误差项 u 和解释变量 X 是不相关的。

假定 5：正确地设定了回归模型，即在经验分析中所用的模型没有设定偏误。

假定 6：对于多元线性回归模型，没有完全的多重共线性。就是说解释变量之间没有完全的线性关系。

2．答：DW 检验的决策区间为

$0 \le DW \le d_L$	误差项 $u_1, u_2, \mathrm{L}, u_n$ 间存在正相关
$d_L < DW \le d_U$	不能判定是否有自相关
$d_U < DW < 4 - d_U$	误差项 $u_1, u_2, \mathrm{L}, u_n$ 间无自相关
$4 - d_U \le DW < 4 - d_L$	不能判定是否有自相关
$4 - d_L \le DW \le 4$	误差项 $u_1, u_2, \mathrm{L}, u_n$ 间存在负相关

3．答：设结构式模型所含方程的总数（或内生变量总数）为 M，模型包含的变量总数（包括前定变量和内生变量）为 H，待识别的方程包含的变量总数（包括内生变量和前定变量）为 G。

阶条件：若某一个结构式方程是可以识别的，则模型中方程数减一小于或等于此方程排斥的变量总数，即 $M-1 \le H-G$，若 $M-1 > H-G$，则不可识别；$M-1 = H-G$，则为恰好识别；$M-1 < H-G$，则为过度识别。阶条件是必要条件，不是充分条件。

秩条件：在具有 M 个方程的结构式模型中，任何一个方程可以识别的充分必要条件是：不包括在该方程中的变量（包括内生变量和前定变量）的参数所组成的矩阵（记为 A）的秩为 $M-1$，即 r（A）$= M-1$。秩条件是充分必要条件，也就是说：如果秩条件成立，则方程是可识别；如果方程是可识别的，则秩条件成立，或者秩条件不成立，则方程是不可识别的。

五、计算题

1．解：(1) $\hat{\ }_0 = 0.50$

$$\hat{\ }_1 = 0.5 + 0.45 - 0.1 = 0.85$$

$$\hat{\ }_2 = 0.5 + 2 \times 0.45 - 4 \times 0.1 = 1.00$$

$$\hat{\ }_3 = 0.5 + 3 \times 0.45 - 9 \times 0.1 = 0.95$$

$$\hat{\ }_4 = 0.5 + 4 \times 0.45 - 16 \times 0.1 = 0.7$$

(2)短期乘数为 0.50

长期乘数为 4.00

延期乘数分别为：0.85　1.00　0.95　0.70。

2．解：

（1）Se=0.015　　　　　t=-8.342

（2）斜率系数 0.2453 表示地区生产总值增加 1 亿元进口需求增加 0.2453 亿元。截距系数-261.09 无实际意义。

（3）斜率系数的 t 统计量为 16.616，远大于临界水平，据 t 检验应拒绝真实斜率系数为零的假设。

六、分析题

解：设 $D_{11}=\begin{cases}1 & \text{汉族}\\ 0 & \text{其它}\end{cases}$，$D_{12}=\begin{cases}1 & \text{蒙族}\\ 0 & \text{其它}\end{cases}$，$D_{13}=\begin{cases}1 & \text{满族}\\ 0 & \text{其它}\end{cases}$，

$D_{21}=\begin{cases}1 & \text{南方}\\ 0 & \text{其它}\end{cases}$，$D_{31}=\begin{cases}1 & \text{大专以下}\\ 0 & \text{其　　它}\end{cases}$，$D_{32}=\begin{cases}1 & \text{本科}\\ 0 & \text{其它}\end{cases}$。

则家庭消费支出模型为

$C = {}_0 + {}_1D_{11} + {}_2D_{12} + {}_3D_{13} + {}_4D_{21} + {}_5D_{31} + {}_6D_{32} + {}_7X + u$。