

一、解释概念：

- 1、多重共线性：是指在多元线性回归模型中，解释变量之间存在的线性关系。
- 2、SRF：就是样本回归函数。即是将样本应变量的条件均值表示为解释变量的某种函数。
- 3、解释变量的边际贡献：在回归模型中新加入一个解释变量所引起的回归平方和或者拟合优度的增加值。
- 4、一阶偏相关系数：反映一个经济变量与某个经济变量的线性相关程度时，剔除另一个变量对它们的影响的真实相关程度的指标。
- 5、最小方差准则：在模型参数估计时，应当选择其抽样分布具有最小方差的估计式，该原则就是最佳性准则，或者称为最小方差准则。
- 6、OLS：普通最小二乘估计。是利用残差平方和为最小来求解回归模型参数的参数估计方法。
- 7、偏相关系数：反映一个经济变量与某个经济变量的线性相关程度时，剔除其它变量（部分或者全部变量）对它们的影响的真实相关程度的指标。
- 8、WLS：加权最小二乘法。是指估计回归方程参数时，按照残差平方加权求和最小的原则进行的估计方法。
- 9、 U_t 自相关：即回归模型中随机误差项逐项值之间的相关。即 $\text{Cov}(U_t, U_s) \neq 0$ $t \neq s$ 。
- 10、二阶偏相关系数：反映一个经济变量与某个经济变量的线性相关程度时，剔除另两个变量对它们的影响的真实相关程度的指标。
- 11、技术方程式：根据生产技术关系建立的计量经济模型。。
- 13、零阶偏相关系数：反映一个经济变量与某个经济变量的线性相关程度时，不剔除任何变量对它们的影响的相关程度的指标。也就是简单相关系数。
- 14、经验加权法：是根据实际经济问题的特点及经验判断，对滞后经济变量赋予一定的权数，利用这些权数构成各滞后变量的线性组合，以形成新的变量，再用最小二乘法进行参数估计的有限分布滞后模型的修正估计方法。

密

15、虚拟变量：在计量经济学中，我们把取值为 0 和 1 的人工变量称为虚拟变量，用字母 D 表示。（或称为属性变量、双值变量、类型变量、定性变量、二元型变量）

16、不完全多重共线性：是指在多元线性回归模型中，解释变量之间存在的近似的线性关系。

17、多重可决系数：用来说明多元线性回归模型对观测值的拟合优度的指标，它是回归平方和 ESS 与总离差平方和 TSS 的比值。

18、边际贡献的 F 检验：是用来检验解释变量的边际贡献是否显著的 F 统计量。它是边际贡献与残差均方差的比值。 $F = \text{边际贡献} / \text{残差均方差}$ 。

19、OLSE 普通最小二乘估计量。即是利用残差平方和最小的原则对模型的参数进行估计得到的参数估计量。

20、PRF 总体回归函数。即是将总体应变量的条件均值表示为解释变量的某种函数。

21、阿尔蒙法：在对有限分布滞后模型的修正估计时，为了消除多重共线性的影响，阿尔蒙提出利用多项式来减少待估计参数的数目的一种修正估计方法。

22、BLUE：在模型的参数估计中，应该遵循的参数估计量应该是最佳线性无偏估计量的准则。

23、复相关系数：是指在多元线性回归模型中，反映多个变量之间存在的线性关系程度的相关系数。

24、滞后效应：应变量受到自身或其它经济变量过去值影响的现象。

25、异方差性：线性回归模型中的随机误差项的方差随某个解释变量的变化而变化的现象。

26、高斯-马尔可夫定理：在古典假定全部满足的条件下用 OLS 估计回归模型的参数，得到的参数估计值即是同时满足线性特性、无偏性和最小方差的参数估计量。

27、可决系数：回归平方和在总变差中所占的比重。

密

二、单项选择题：

- | | | | | | |
|-----|-------|------|-------|-------|-------|
| 1-5 | BDDCB | 6-10 | CAABB | 11-15 | BDBDD |
| 1-5 | CBCDB | 6-10 | ABABA | 11-15 | ABACB |
| 1-5 | ABCD | A | C | C | A |
| 1-5 | CBDBD | 6-10 | BACAC | 11-15 | BBABA |
| 1-5 | CBDAD | 6-10 | BCABB | 11-15 | CCAAB |
| 1-5 | CADBB | 6-10 | ACBAD | 11-15 | BDDCB |
| 1-5 | ABCE | ABCD | D | A | A |
| 1-5 | BACBD | 6-10 | BBBAB | 11-15 | DCCDD |
| 1-5 | BADDD | 6-10 | ADAAC | 11-15 | BBBBC |
| 1-5 | A | C | ABCD | D | B |

三、多项选择题：

- | | | | |
|--------|---------|---------|--------|
| 1、BCDE | 2、AB | 3、CE | 4、BE |
| 1、BC | 2、ABC | 3、ABE | 4、ACDE |
| 1、ADE | 2、ABDF | 3、ABC | 4、CE |
| 1、BCDE | 2、ABDE | 3、BCD | 4、BCE |
| 1、BCD | 2、ABCDE | 3、CDE | 4、BE |
| 1、BCDE | 2、BD | 3、CDE | 4、ABDE |
| 1、BCD | 2、ABCD | 3、ABCDE | 4、BD |

四、判断题（判断命题正误，并说明理由）：

1、在对参数进行最小二乘估计之前，没有必要对模型提出古典假定。
错误。

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

2、当异方差出现时，常用的 t 和 F 检验失效；

密

正确。

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

3、解释变量与随机误差项相关，是产生多重共线性的主要原因。

错误。

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

4、由间接最小二乘法与两阶段最小二乘法得到的估计量都是无偏估计。

错误。

间接最小二乘法适用于恰好识别方程的估计，其估计量为无偏估计；而两阶段最小二乘法不仅适用于恰好识别方程，也适用于过度识别方程。

两阶段最小二乘法得到的估计量为有偏、一致估计。

5、半对数模型 $Y = \beta_0 + \beta_1 \ln X + \mu$ 中，参数 β_1 的含义是 X 的绝对量变化，引起 Y 的绝对量变化。

错误。

半对数模型的参数 β_1 的含义是当 X 的相对变化时，绝对量发生变化，引起因变量 Y 的平均值绝对量的变动。

6、对已经估计出参数的模型不需要进行检验。

错误。

有必要进行检验。首先，因为我们在设定模型时，对所研究的经济现象的规律

性可能认识并不充分，所依据的得经济理论对研究对象也许还不能做出正确的解释和说明。或者虽然经济理论是正确的，但可能我们对问题的认识只是从某些局部出发，或者只是考察了某些特殊的样本，以局部去说明全局的变化规律，必然会导致偏差。其次，我们用以及参数的统计数据或其他信息可能并不十分可靠，或者较多采用了经济突变时期的数据，不能真实代表所研究的经济关系，也可能

密

由于样本太小，所估计的参数只是抽样的某些偶然结果。另外，我们所建立的模型，所用的方法，所用的统计数据，还可能违反计量经济的基本假定，这也是会导致错误的结论。

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

错误。

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

8、随机误差项和残差是有区别的。

正确。

随机误差项 $u_i = Y_i - E(Y/X_i)$ 。当把总体回归函数表示成 $Y_i = \hat{Y}_i + e_i$ 时，其中的 e_i 就是残差。它是用 \hat{Y}_i 估计 Y_i 时带来的误差 $e_i = Y_i - \hat{Y}_i$ ，是对随机误差项 u_i 的估计。

9. T

10.F 异方差性

11.F $Y = Xb + u$

12.T

13.T

14、错误

可决系数是对模型拟合优度的综合度量，其值越大，说明在 Y 的总变差中由模型作出了解释的部分占的比重越大，模型的拟合优度越高，模型总体线性关系的显著性越强。反之亦然。斜率系数的 t 检验是对回归方程中的解释变量的显著性的检验。在简单线性回归中，由于解释变量只有一个，当 t 检验显示解释变量的影响显著时，必然会有该回归模型的可决系数大，拟合优度高。

15、正确。

异方差的出现总是与模型中某个解释变量的变化有关。...

密

自相关性是各回归模型的随机误差项之间具有相关关系。……

16、错误

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

引入 $m-1$ 个虚拟变量，否则会陷入“虚拟变量陷阱”；

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

17、错误

阶条件只是一个必要条件，即满足阶条件的方程也可能是不可识别的。

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

错。

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

19、假定个人服装支出同收入水平和性别有关，由于性别是具有两种属性（男女）的定性因素，因此，用虚拟变量回归方法分析性别对服装支出的影响时，需要引入两个虚拟变量。

错。

是否引入两个虚拟变量，应取决于模型中是否有截距项。如果有截距项则引入一个虚拟变量；如果模型中无截距项，则可引入两个虚拟变量。

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

正确。

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

密

21、随机扰动项的方差与随机扰动项方差的无偏估计没有区别。

错。

随机扰动项的方差反映总体的波动情况，对一个特定的总体而言，是一个确定的值。在最小二乘估计中，由于总体方差在大多数情况下并不知道，所以用样本

数据去估计 σ^2 ： $\hat{\sigma}^2 = \sum e_i^2 / (n-k)$ 。其中 n 为样本数， k 为待估参数的个数。 $\hat{\sigma}^2$ 是 σ^2 线性无偏估计，为一个随机变量。

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

前定变量。

错误。

结构方程中，解释变量可以是前定变量，也可以是内生变量。

24、线性回归模型意味着因变量是自变量的线性函数。

错误。

因为线性回归模型表示的是因变量是自变量的近似的线性关系，除了模型中的自变量以外，还有许多影响因变量的其它因素。

25、G—Q 检验要求的样本容量必须是大样本。

错误。

因为 G—Q 检验要求的样本容量是尽可能是大样本，当然也可以是小样本，但样本数目不能少于模型参数个数的 2 倍以上。

26.F 也可以是小样本，但检验方法有所变化即不删除中间项

27. F 是 $\sum e^2 / (n-2)$

28.T

29.F 在共线性的检验中

30.T

31、在异方差性的情况下，常用的 OLS 法必定高估了估计量的标准误。

错误。有可能高估也有可能低估；如：考虑一个非常简单的具有异方差性的线性回

密

归模型：

$$Y_i = \beta X_i + u_i; \text{Var}(u_i) = \sigma_i^2 = Z_i^2 \sigma^2, \text{ 则: } \text{Var}(\hat{\beta}) = \text{Var}\left(\frac{\sum X_i u_i}{\sum X_i^2}\right) = \frac{\sum X_i^2 \text{Var}(u_i)}{(\sum X_i^2)^2}$$

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

正确。

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

33、多重共线性问题是随机扰动项违背古典假定引起的；

错误。

应该是解释变量之间高度相关引起的。

34、秩条件是充要条件，因此，单独利用秩条件就可以完成联立方程识别状态的确定。

错误。

虽然秩条件是充要条件，但其前提是，只有在通过了阶条件的条件下。在对联立方程进行识别时，还应该结合阶条件判断是过度识别，还是恰好识别。

35、多重共线性是随机误差项的古典假定违背。

错误。

因为多重共线性是回归模型中多个解释变量之间存在的完全的或者近似的共线性，与随机误差项无关。

36、自相关性的后果是：参数估计值仍然是无偏的但不再具有最小方差性。

正确。

37、线性回归模型意味着因变量是自变量的线性函数。

错误。

因为线性回归模型表示的是因变量是自变量的近似的线性关系，除了模型中的自变量以外，还有许多影响因变量的其它因素。

38、D-W 检验要求的样本容量是小样本。

密

错误。

因为 D-W 检验要求的样本容量是尽可能是大样本，当然如果是小样本，就会影响检验的结论的精确性。D-W 检验统计量 $d \approx 2(1-\rho)$ 的结论是在大样本的条件下推导出来的。

39.T

40.F 检验异方差性

41.F Z 检验

42.F 小于下限时必存在，但在上限和下限之间时不能确定

43.F 变大

五、计算分析：

1、解：地方预算内财政收入（Y）和 GDP 的关系近似直线关系，可建立线性

回归模型： $Y_i = \beta_1 + \beta_2 GDP_i + u_i$

即 $\hat{Y}_i = -3.611151 + 0.134582 GDP_i$

(4.161790) (0.003867)

t=(-0.867692) (34.80013)

$R^2=0.99181$ F=1211.049

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

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

当 2005 年 GDP 为 3600 亿元时，地方财政收入的点预测值为：

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

区间预测：因为 $R^2 = \frac{\sum \hat{y}^2}{\sum e^2 + \sum \hat{y}^2}$ 而 $\sum \hat{y}^2 = \hat{\beta}_2^2 \sum x^2$

密

$$\text{所以有: } \sum x^2 = \frac{R^2 \sum e^2}{1-R^2}$$

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

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

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

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

GDP₂₀₀₅ = 3600 时

$$480.884 \pm 2.228 \times 7.5325 \times \sqrt{\frac{1}{12} + \frac{7195337.357}{3793728.494}} = 480.884 \pm 25.2735 \text{ (亿元)}$$

Y_f 个别值置信度 95% 的预测区间为:

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

$$\begin{aligned} \text{即} \quad &= 480.884 \pm 2.228 \times 7.5325 \times \sqrt{1 + \frac{1}{12} + \frac{7195337.357}{3793728.494}} \\ &= 480.884 \pm 30.3381 \text{ (亿元)} \end{aligned}$$

2、解: (1) 给定 $\alpha = 0.05$ 和自由度为 2 下, 查卡方分布表, 得临界值 $\chi^2 = 5.9915$,

而 White 统计量 $nR^2 = 5.2125$, 有 $nR^2 < \chi^2_{0.05}$ (2), 则不能拒绝原假设, 说明模型中不存在异方差。

(2) 因为对如下函数形式

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

得样本估计式

密

$$|\hat{e}| = 6.4435\sqrt{X}$$
$$(4.5658)$$
$$R^2 = 0.2482$$

由此，可以看出模型中随机误差项有可能存在异方差。

(3) 对异方差的修正。可取权数为 $w=1/X$ 。

3、答：（1）由模型估计结果可看出：旅行社职工人数和国际旅游人数均与旅游外汇收入正相关。平均说来，旅行社职工人数增加 1 人，旅游外汇收入将增加 0.1179 百万美元；国际旅游人数增加 1 万人次，旅游外汇收入增加 1.5452 百万美元。

（2）取 $\alpha = 0.05$ ，查表得 $t_{0.025} (31-3) = 2.048$

因为 3 个参数 t 统计量的绝对值均大于 $t_{0.025} (31-3) = 2.048$ ，说明经 t 检验 3 个参数均显著不为 0，即旅行社职工人数和国际旅游人数分别对旅游外汇收入都有显著影响。

取 $\alpha = 0.05$ ，查表得 $F_{0.05} (2, 28) = 3.34$ ，由于 $F = 199.1894 > F_{0.05} (2, 28) = 3.34$ ，说明旅行社职工人数和国际旅游人数联合起来对旅游外汇收入有显著影响，线性回归方程显著成立。

4、令 $x=1/X$ $y=1/Y$

得 $y=b_0+b_1x+U$

令 $x^1=x-E(x)$ $y^1=y-E(y)$

则 $\hat{a}_1=\Sigma x^1 y^1 / \Sigma x^{1^2}$

$\hat{a}_0=E(y) - \hat{a}_1 E(x)$

$\hat{a}_1=397.32$

$\hat{a}_0=-11.32$

$1/Y=-11.32+397.32/X$

密

5、(1) 根据表中数据，通过计算得到： $\sum xy = 40200$

$$\sum x^2 = 74250 \quad \sum y^2 = 22189.6 \quad E(Y) = 142.8 \quad E(X) = 195$$

将上述数据带入下式 $\hat{a}_1 = \sum xy / \sum x^2 = 40200 / 74250 = 0.5414$

$$\hat{a}_0 = E(Y) - \hat{a}_1 E(X) = 142.8 - 0.5414 * 195 = 37.227$$

故样本回归方程为： $\hat{Y} = 37.227 + 0.5414X$

(2) 原假设 $H_0: a_1 = 0$

备择假设 $H_1: a_1 \neq 0$

取 $\alpha = 0.05$ ，检验统计量 $t = \hat{a}_1 / \text{se}(\hat{a}_1) = 0.5414 / 0.0267 = 20.2272 > t_{0.025}(8) = 2.306$

因此拒绝原假设，接受备择假设，说明 X 对 Y 存在显著影响。

同理对 \hat{a}_0 而言有：原假设 $H_0: a_0 = 0$

备择假设 $H_1: a_0 \neq 0$

取 $\alpha = 0.05$ ，检验统计量 $t = \hat{a}_0 / \text{se}(\hat{a}_0) = 37.227 / 5.7008 = 6.5301 > t_{0.025}(8) = 2.306$

因此拒绝原假设，接受备择假设 $H_1: a_0 \neq 0$ 。

结论：回归方程显著成立。

6、(1) $\hat{Y} = 0.8767 + 0.238X \quad R^2 = 0.939$

(2) 第一个子样本的回归分析结果如下：

$$\hat{Y}_1 = 0.6 + 0.276X \quad R^2 = 0.966 \quad \sum e_1^2 = 0.3$$

第二个子样本的回归分析结果如下：

$$\hat{Y}_2 = 1.54 + 0.2X \quad R^2 = 0.854 \quad \sum e_2^2 = 2.024$$

$$F = \sum e_2^2 / \sum e_1^2 = 2.024 / 0.3 = 6.747 \quad \text{而 } F_{0.05}(8, 8) = 3.44$$

$F > F_{0.05}(8, 8) = 3.44$ 显著存在异方差。

(3) 设 $\text{Var}(u_i) = \sigma^2 X_i^2$ 则

$Y/X = b_1 + b_0/X + u/X$ 满足等方差，对 Y/X ， $1/X$ 做 OLS 得

$$Y' = 0.25 + 0.75X' \quad R^2 = 0.76$$

7、解：（1）建立中国 1978 年-1997 年的财政收入 Y 和国内生产总值 X 的线性回归方程

$$Y_i = \beta_1 + \beta_2 X_i + U_i$$

利用 1978 年-1997 年的数据估计其参数，结果为

$$\hat{Y}_i = 857.8375 + 0.100036 X_i$$

$$t = (12.77955) \quad (46.04910)$$

$$R^2 = 0.991583 \quad F = 2120.52$$

GDP 增加 1 亿元，平均说来财政收入将增加 0.1 亿元。

（2） $r^2 = ESS/TSS = 0.991583$ ，模型的拟合程度较高。

$$H_0: \beta_2 = 0 \quad H_1: \beta_2 \neq 0$$

$$t = \frac{\hat{\beta}_2}{SE(\hat{\beta}_2)} \sim t(18)$$

$$t = 46.0491 > t_{0.025}(18), \text{ 拒绝 } H_0$$

说明，国内生产总值对财政收入有显著影响。

（3）若是 1998 年的国内生产总值为 78017.8 亿元，确定 1998 年财政收入的点预测值为

$$\hat{Y} = 857.8375 + 0.100036 \times 78017.8 = 8662.426141 \text{ (亿元)}$$

1998 年财政收入平均值预测区间 ($\alpha = 0.05$) 为：

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

$$8662.426 \pm 2.101 \times 208.5553 \times (1/20 + 9216577098/3112822026)^{1/2}$$

8、解：从模型拟合结果可知，样本观测个数为 27，消费模型的判定系数 $R^2 = 0.95$ ，F 统计量为 107.37，在 0.05 置信水平下查分子自由度为 3，分母自由度为

密

23 的 F 临界值为 3.028，计算的 F 值远大于临界值，表明回归方程是显著的。模型整体拟合程度较高。

依据参数估计量及其标准误，可计算出各回归系数估计量的 t 统计量值：

$$t_0=8.133/8.92=0.91, t_1=1.059/0.17=6.10, t_2=0.452/0.66=0.69, t_3=0.121/1.09=$$

0.11, 除 t_1 外，其余的 t 值都很小。工资收入 X_1 的系数的 t 检验值虽然显著，但该系数的估计值过大，该值为工资收入对消费边际效应，因为它为 1.059，意味着工资收入每增加一美元，消费支出的增长平均将超过一美元，这与经济理论和常识不符。另外，理论上非工资—非农业收入与农业收入也是消费行为的重要解释变量，但两者的 t 检验都没有通过。这些迹象表明，模型中存在严重的多重共线性，不同收入部分之间的相互关系，掩盖了各个部分对解释消费行为的单独影响。

9、解：描述投诉率（Y）依赖航班按时到达正点率（X）的回归方程：

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

$$\hat{Y}_i = 6.017832 - 0.070414 X_i$$

$$(1.052260) \quad (0.014176)$$

$$t = (5.718961) \quad (-4.967254)$$

$$R^2 = 0.778996 \quad F = 24.67361$$

这说明当航班正点到达比率每提高 1 个百分点，平均说来每 10 万名乘客投诉次数将下降 0.07 次。

如果航班按时到达的正点率为 80%，估计每 10 万名乘客投诉的次数为

$$\hat{Y}_i = 6.017832 - 0.070414 \times 80 = 0.384712 \text{ (次)}$$

10、解：（1）因为 $f(X_i) = X_{2i}^2$ ，所以取 $W_{2i} = \frac{1}{X_{2i}}$ ，用 W_i 乘给定模型两端，得：

密

$$\frac{Y_i}{X_{2i}} = \beta_1 \frac{1}{X_{2i}} + \beta_2 + \beta_3 \frac{X_{3i}}{X_{2i}} + \frac{u_i}{X_{2i}}$$

上述模型的随机误差项的方差为一固定常数，即：

$$Var(\frac{u_i}{X_{2i}}) = \frac{1}{X_{2i}^2} Var(u_i) = \sigma^2$$

(2) 根据加权最小二乘法，可得修正异方差后的参数估计式为：

$$\begin{aligned} \hat{\beta}_1 &= \bar{Y}^* - \hat{\beta}_2^* \bar{X}_2^* - \hat{\beta}_3^* \bar{X}_3^* \\ \hat{\beta}_2 &= \frac{(\sum W_{2i} y_i^* x_{2i}^*) (\sum W_{2i} x_{3i}^{*2}) - (\sum W_{2i} y_i^* x_{3i}^*) (\sum W_{2i} x_{2i}^* x_{3i}^*)}{(\sum W_{2i} x_{2i}^{*2}) (\sum W_{2i} x_{3i}^{*2}) - (\sum W_{2i} x_{2i}^* x_{3i}^*)^2} \\ \hat{\beta}_3 &= \frac{(\sum W_{2i} y_i^* x_{3i}^*) (\sum W_{2i} x_{2i}^{*2}) - (\sum W_{2i} y_i^* x_{2i}^*) (\sum W_{2i} x_{2i}^* x_{3i}^*)}{(\sum W_{2i} x_{2i}^{*2}) (\sum W_{2i} x_{3i}^{*2}) - (\sum W_{2i} x_{2i}^* x_{3i}^*)^2} \\ \bar{X}_2^* &= \frac{\sum W_{2i} X_{2i}}{\sum W_{2i}}, \quad \bar{X}_3^* = \frac{\sum W_{2i} X_{3i}}{\sum W_{2i}}, \quad \bar{Y}^* = \frac{\sum W_{2i} Y_i}{\sum W_{2i}} \\ \text{其中} \quad x_{2i}^* &= X_{2i} - \bar{X}_2^* \quad x_{3i}^* = X_{3i} - \bar{X}_3^* \quad y_i^* = Y_i - \bar{Y}^* \end{aligned}$$

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

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

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

12、解：将自适应预期假设写成 $X_{i+1}^* - (1-r)X_i^* = rX_i$

$$\text{原模型 } Y_i = \beta_0 + \beta_1 X_{i+1}^* + u_i \quad (1)$$

将①滞后一期并乘以 (1-r)，有

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

①式减去②式，整理后得到

$$Y_t = r\beta_0 + r\beta_1X_t + (1-r)Y_{t-1} + v_t$$

式中：

$$v_t = u_t - (1-r)u_{t-1}$$

13、(1) 是通过样本容量和解释变量的个数并查 D-W 统计量表得到的。

A 存在。因为 A: $DW=0.8252 < d_l=1.11$

B 不存在。因为 B: $d_u=1.54 < DW=1.82 < 4-d_u$

(2) A 模型的函数形式不正确。

模型 A 是一个关于劳动份额与时间年度之间的一元线性回归模型，而该模型在上述分析中已知存在序列相关，而模型 B 却是一个二次函数模型，且不存在序列自相关，因此模型 A 的函数形式的错误导致了序列自相关的出现。

14、(1) 求回归方程：列计算表得

$$\sum X=36 \quad \sum Y=248946 \quad EX=4.5 \quad EY=31118.25$$

$$\sum X^2=204 \quad \sum Y^2=7792592256 \quad \sum XY=1161095$$

$$\sum x^2 = \sum X^2 - n(EX)^2 = 42$$

$$\sum y^2 = \sum Y^2 - n(EY)^2 = 46828391.5 \quad \sum xy = \sum XY - nEXEY = 40838$$

$$\hat{a}_1 = \sum xy / \sum x^2 = 40838 / 42 = 972.3333$$

$$\hat{a}_0 = EY - \hat{a}_1 EX = 26742.7502$$

$$\text{因此 } \hat{Y} = 26742.7502 + 972.3333X$$

(2) 检验：①经济意义的检验。

②拟合优度的检验

$$r^2 = \hat{a}_1^2 \sum x^2 / \sum y^2 = 0.848$$

③总体参数估计值的可靠性检验

密

$$\sum e^2 = \sum y^2 - \hat{a}_1^2 \sum x^2 = 7120242.833$$

$$Se^2(u_i) = \sum e^2 / (n-2) = 1186707.139$$

$$Se(\hat{a}_1) = (Se^2 / \sum x^2)^{1/2} = 168.09$$

$$Se(\hat{a}_0) = [Se^2 \sum X^2 / (n \sum x^2)] = 848.42$$

④、假设检验（只对 a_1 进行）

$$H_0: a_1 = 0 \quad H_1: a_1 \neq 0$$

$$T = \hat{a}_1 / Se(\hat{a}_1) = 972.3333 / 168.09 = 5.78$$

给定 $\alpha = 0.05$ ，得临界值 $t_{0.025}(6) = 2.45$

显然有 $T > t_{0.025}$ ，接受被择假设 $H_1: b_1 \neq 0$

(3)、预测 1993 年粮食产量

因为 $X_0 = 9$ ，所以 $\hat{Y}_0 = 35493.7499$ （万吨）

给定 $\alpha = 0.05$

$$Se(e_0) = Se[1 + 1/n + (X_0 - EX)^2 / \sum x^2]^{1/2}$$

$$= 1089.36 [1 + 1/8 + (9 - 4.5)^2 / 42]^{1/2} = 1381$$

$$t_{0.025} Se(e_0) = 3383.4$$

$$\hat{Y}_0 - t_{0.025} Se(e_0) < Y_0 < \hat{Y}_0 + t_{0.025} Se(e_0)$$

$$32110.35 < Y_0 < 38877.15$$

15、解：

密

Dependent Variable: Y

Method: Least Squares

Sample: 134

Included observations: 34

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	5.540940	0.968608	5.7205	0.0000
X	0.474514	0.055077	8.6155	0.0000
R-squared	0.6987	Mean dependent var		13.64118
Adjusted R-squared	0.6893	S.D. dependent var		2.436480
S.E. of regression	1.358008	Akaike info criterion		3.506937
Sum squared resid	59.01394	Schwarz criterion		3.596723
Log likelihood	-57.61793	F-statistic		74.227
Durbin-Watson stat	1.796718	Prob(F-statistic)		0.000000

模型结果支持了理论，因为期望回报及其标准差之间存在显著的线性关系。

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

（2）存在异方差（注意显著性水平是 0.1）；（写出详细步骤）

（3）说出一种修正思路即可。

17、解：（1）由模型估计结果可看出：旅行社职工人数和国际旅游人数均与旅游外汇收入正相关。平均说来，旅行社职工人数增加 1 人，旅游外汇收入将增加 0.1179 百万美元；国际旅游人数增加 1 万人次，旅游外汇收入增加 1.5452 百万美元。

（2）取 $\alpha = 0.05$ ，查表得 $t_{0.025} (31-3) = 2.048$

因为 3 个参数 t 统计量的绝对值均大于 $t_{0.025} (31-3) = 2.048$ ，说明经 t 检验 3 个参数均显著不为 0，即旅行社职工人数和国际旅游人数分别对旅游外汇收入都有显著影响。取 $\alpha = 0.05$ ，查表得 $F_{0.05} (2, 28) = 3.34$ ，由于 $F = 199.1894 > F_{0.05} (2, 28) = 3.34$ ，说明旅行社职工人数和国际旅游人数联合起来对旅游外汇收入有显著影响，

密

线性回归方程显著成立。

18、(1) 回归方程: $\hat{\beta}_2 = \Sigma xy / \Sigma x^2 = 40200 / 74250 = 0.5414$

$$\hat{\beta}_1 = \Sigma Y / n - \hat{\beta}_2 \Sigma X / n = 37.227$$

$$Y = 37.227 + 0.5414X$$

(2) $r^2 = \hat{\beta}_2^2 \Sigma x^2 / \Sigma y^2 = 0.5414^2 * 74250 / 22189.6 = 0.98$

(3) $H_0: \beta_2 = 0$, $H_1: \beta_2 \neq 0$

$$t = (\hat{\beta}_2 - \beta_2) / \text{se } \hat{\beta}_2 = 0.5414 / 0.0267 = 20.277$$

因 t 大于 t 临界值, 所以可支配收入显著影响消费支出。

(4) $Y_F = 37.227 + 0.5414 * 370 = 237.5$

$$\begin{aligned} E(Y_F / X_F) &= 237.5 \pm 2.306 * 7.2865 [1/10 + (370 - 195)^2 / 74250]^{1/2} \\ &= 237.5 \pm 12.0289 \end{aligned}$$

即在 95% 的概率水平下, 均值的置信区间为 (225.198, 249.529)

19、(1) 这是异方差的 G-Q 检验, 使用的是样本分段拟合, $F = 4334.937 > 4.28$, 因此拒绝原假设, 表明模型中存在异方差。

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

(3) 这两种方法都是检验异方差的, 但适用条件不同:

A、G-Q 要求尽可能大样本, 扰动项正态分布; 可用于截面数据和时间序列数据。

B、ARCH 检验仅适用于时间序列数据, 且其渐近分布为 χ^2 分布。

六、简述和问答

1、联系: 复决定系数和修正可决系数都是被用来说明拟合优度的而且

$$R^{2*} = 1 - (1 - R^2)(n-1)/(n-k-1)$$

区别: 复决定系数是用来准确反映解释变量对被解释变量的解释程度的。用来反映

拟合优度会给分析人员带来分析错觉，因为只要在回归模型中增加解释变量的个数就会增大复决定系数的值,就是说，要提高模型的拟合优度只需在模型中增加解释变量的个数就可以了，事实并非如此。而修正可决系数是用来准确反映回归线对散布点的拟合优度。（3分）

- 2、
 - (1) 将观测值排序，删除中间约 1/4 项。
 - (2) 分别对前后两个子样本回归，求出各自的残差平方和。
 - (3) 提出假设检验。
 - (4) 构造 F 统计量进行检验。
 - (5) 判断：当 F 统计量大于 F 临界值时，表明存在异方差。
- 3、
 - (1) 经济变量惯性的作用；
 - (2) 经济行为的滞后性；
 - (3) 一些随机偶然因素的干扰；
 - (4) 设定偏误。
- 4、在某种原假设成立的条件下，利用适当的统计量和给定的显著性水平，构造一个小概率事件，如果该事件发生了，就认为原假设不真，接受备择假设。
- 5、
 - (1) 求出回归方程。
 - (2) 求出残差。
 - (3) 将残差前后期数据对应排列成两个残差序列，并绘制散点图。
- 6、
 - (1) 参数估计式仍然是无偏的，但方差会随共线性程度的提高而增大；
 - (2) t 值会变小，其检验失效；
 - (3) 参数的区间估计失去意义。
- 7、
 - (1) 用 F 统计量对边际贡献进行检验。
 - (2) 首先作出原假设和备择假设，选定 F 统计量， $F = \text{边际贡献} / \text{残差均方差}$ ，并计算出 F 的值，在选定的显著性水平下查 F 分布表找出 F 临界值，将其和 F 统计量作比较，如果统计量大于临界值，则说明边际贡献显著；否则不显著。
- 8、
 - (1) 解释变量为非随机的；

- (2) 随机误差项为一阶自回归形式；
- (3) 模型中不应含有滞后内生变量作为解释变量；
- (4) 截距项不为零；
- (5) 数据无缺失项。

当然还应该是在大样本的条件下。

9、 因为复决定系数只是说明了全体解释变量作为一个整体对被解释变量的综合解释程度，随着模型中解释变量的增加，它的值会增大，但这时模型的拟合优度并没有随之提高，所以它不能准确说明拟合优度。而修正可决系数就克服了它的不足，因此能准确反映模型的拟合优度。

10、参数估计值不确定；参数估计值的方差无限大。