

# 计量经济学 第一次作业

2020 年秋

注意事项:

1. 请于 **2020 年 10 月 23 日下午 5 点**前提交电子版至网络学堂，不接受以其他形式提交的作业。对作业提交有疑问请发邮件联系助教唐爱迪 (tangad.17@sem.tsinghua.edu.cn)。
2. 上机作业的数据集请从网络学堂 → 课程文件 → 习题集下载，文件名为 “EduIncome.dta”。
3. 请将所有作业的答案生成一份 PDF 文档（手写扫描或者打印皆可）提交。另外请再提交一份上机作业的 Stata 程序 (do file)。
4. 欢迎在提交作业前参加助教答疑，也可以和同学讨论作业思路，但请提交你**自己**写的答案。

[一元方程参数的方差] 对一元线性回归方程  $y = \beta_0 + \beta_1 x + u$ ，在假设一到假设五成立的情况下，请推导 OLS 估计量  $\beta_0$  和  $\beta_1$  的方差 (以  $x$  为条件) (20 分)

$$\begin{aligned} \text{Var}(\hat{\beta}_0|x) &= \frac{\sigma^2 \sum_{i=1}^n x_i^2}{n \sum_{i=1}^n (x_i - \bar{x})^2} \\ \text{Var}(\hat{\beta}_1|x) &= \frac{\sigma^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \end{aligned}$$

[误差和残差] 在证明一元回归方程  $\hat{\beta}_1$  的无偏性时，我们用到了如下公式：

$$\hat{\beta}_1 = \beta_1 + \frac{\sum (x_i - \bar{x})u_i}{\sum (x_i - \bar{x})^2}.$$

小清提出，根据解  $\hat{\beta}_1$  的一阶条件， $\sum u_i = 0$ ， $\sum x_i u_i = 0$ ，因此  $\hat{\beta}_1 = \beta_1$ 。从而证明了无偏性。小明的说法对吗？为什么？(10 分)

[上机题：收入的决定因素] 为了研究劳动力市场上决定收入的因素，我们收集了一个数据集，每一个观测代表了一个人。包括以下变量：

- id 个人编号
- gender 性别.1= 女性, 2= 男性.
- birthyear 出生年份
- marriage 婚姻状况: 1= 处于婚姻状态, 2= 非结婚状态 (包括未婚、离异、丧偶等)
- empjob\_twage 年收入 (元)
- schooling\_yr 受教育年数

上机作业的数据集请从网络学堂 → 课程文件 → 习题集下载, 文件名为 “EduIncome.dta”。用 Stata 回答以下问题：

- (1) 用 Stata 计算除了 id 之外所有变量的均值、标准差、最小值和最大值。(5 分)
- (2) 创建一个新变量 female, 1 代表女性, 0 代表男性。数据中有多少比例是女性？(5 分)
- (3) 估计下面这个回归方程并给出回归结果 (附上回归结果的截图)。(5 分)

$$\text{empjob\_twage} = \beta_0 + \beta_1 \text{schooling\_yr} + \beta_2 \text{female} + u.$$

在回答以下问题时，假设零条件均值假设成立。

- (4) 根据你的回归结果，在受教育年限一样的情况下，平均来说，男性和女性的收入差别为多少？(5 分)
- (5) 根据你的回归结果，教育年限每增加一年，收入平均增加多少？(5 分)
- (6) 计算  $\text{empjob\_twage}$ ,  $\hat{\text{empjob\_twage}}$  和  $\hat{u}$  的均值。(5 分)
- (7) 在这个方程中，有多少比例的收入波动被教育年限和性别解释？(5 分)

**[完全共线性，矩法与最小二乘法]** 在一次抽样调查中，政府调查了家庭用于食品的月消费额，并收集了  $n$  个样本，其中第 1 到第  $k$  个记录来自甲省，第  $k+1$  到  $n$  个记录来自乙省。我们用虚拟变量  $x_1$  和  $x_2$  分别代表记录来自甲省和乙省。(注意: 虚拟变量表示该变量只能取两个值。在这个例子里，如果记录来自甲省，则  $x_1 = 1$ ，否则  $x_1 = 0$ 。类似的，如果记录来自乙省，则  $x_2 = 1$ ，否则  $x_2 = 0$ 。)

考虑下面这个回归方程：

$$y = \beta_1 x_1 + \beta_2 x_2 + u.$$

我们假设零条件均值成立，即  $E(u|x_1, x_2) = 0$ 。

(1) 小华提出，因为记录不是来自甲省就是来自乙省，因此以上方程违反了不存在完全共线性的要求，因此无法进行估计。小华的说法对吗？为什么？(5 分)

下面我们分别用矩法和最小二乘法来推导  $\hat{\beta}_1$  和  $\hat{\beta}_2$  的表达式。

(2)[矩法] 根据  $E(u|x_1, x_2) = 0$ ，推导  $E(ux_1) = 0$  和  $E(ux_2) = 0$ 。(2 分)

(3)[矩法] 写出样本矩的表达式 (表示为包含  $y_i$  的形式)。(3 分)

(4)[最小二乘法] 写出最小二乘法的目标函数 (表示为包含  $y_i$  的方程)。(2 分)

(5)[最小二乘法] 写出最小二乘法的一阶条件。(3 分)

(6) 请问用矩法和最小二乘法的得到的参数估计一样吗？为什么？(5 分)

(7) 计算  $\hat{\beta}_1$  和  $\hat{\beta}_2$  的表达式，请清楚写出过程，并化简到最简单的形式 (10 分)。根据你的表达式和以下数据，分别计算  $\hat{\beta}_1$  和  $\hat{\beta}_2$ 。(5 分)

y	数据来源省份
5	甲
9	甲
7	甲
3	甲
1	甲
6	乙
4	乙
8	乙
2	乙
10	乙

表 1: 数据集