

# 计量经济学-作业5



## P108 习题2

- 2 假设你受雇于一家新成立的印度尼西亚汽车制造公司 “Indo”。为帮助公司确定其在美国销售汽车的削价水平，需要建立一个美国汽车销售价格的模型。你建立了一个包含 35 种不同款式的 2004 年美国产汽车的价格模型（括号内的数值为标准差）：

模型 A：

$$\hat{P}_i = 3.0 + 0.28W_i + 1.2T_i + 5.8C_i + 0.19L_i$$
$$(0.07) \quad (0.4) \quad (2.9) \quad (0.20)$$
$$\bar{R}^2 = 0.92$$

式中， $\hat{P}_i$  代表第  $i$  款汽车的价格（单位：千美元）； $W_i$  代表第  $i$  款汽车的质量（单位：百磅）； $T_i$  代表第  $i$  款汽车是否可自动驾驶，为虚拟变量，如果是为 1，否则为 0； $C_i$  代表第  $i$  款汽车是不是带有 GPS 定位装置，为虚拟变量，如果带有为 1，否则为 0； $L_i$  代表第  $i$  款汽车发动机的排量大小（单位：升）。

- 你所在公司的价格专家假设模型 A 中所有参数的符号均为正。在 5% 的显著性水平下，对她的预期进行检验。
- 模型 A 中存在什么样的计量经济学问题？特别是， $C$  的参数大小是否引起了你的注意？为什么？问题出在哪里？
- 你决定将  $L$  剔出方程，再对方程进行回归，以此来检验  $L$  是不相干变量的可能性，得到模型 T。你更倾向于哪个模型？为什么？（提示：请确保你使用了四个设定准则。）

模型 T：

$$\hat{P}_i = 18 + 0.29W_i + 1.2T_i + 5.9C_i$$
$$(0.07) \quad (0.30) \quad (2.9)$$
$$\bar{R}^2 = 0.93$$

## 回答

- a. 在显著性水平 5% 的条件下分别对 4 个参数进行假设检验。其中，样本总数为 35，可以求得自由度为  $n - k - 1 = 35 - 4 - 1 = 30$ ，查表得单侧  $t_c = 1.697$

- W: 做出假设  $H_0: \beta \leq 0$   $H_A: \beta > 0$ ，且预测参数为正。  $t_1 = \frac{0.28 - 0}{0.07} = 4$ ，可知  $|4| > 1.697$  即  $|t_1| > t_c$ ，且其预期符号和对立假设相同（均为正数），因此可以拒绝  $H_0$

- **T**:做出假设  $H_0: \beta \leq 0$   $H_A: \beta > 0$ ,且预测参数为正。  $t_2 = \frac{1.2 - 0}{0.4} = 3$ , 可知  $|3| > 1.697$  即  $|t_2| > t_c$ , 且其预期符号和对立假设相同 (均为正数), 因此**可以拒绝**  $H_0$
- **C**:做出假设  $H_0: \beta \leq 0$   $H_A: \beta > 0$ ,且预测参数为正。  $t_3 = \frac{5.8 - 0}{2.9} = 2$ , 可知  $|2| > 1.697$  即  $|t_3| > t_c$ , 且其预期符号和对立假设相同 (均为正数), 因此**可以拒绝**  $H_0$
- **L**:做出假设  $H_0: \beta \leq 0$   $H_A: \beta > 0$ ,且预测参数为正。  $t_4 = \frac{0.19 - 0}{0.20} = 0.95$ , 可知  $|0.95| < 1.697$  即  $|t_1| < t_c$ , 因此**不能拒绝**  $H_0$

b.

- 存在问题: 这个方都可能存在遗漏变量的问题, 另外, 也可能含有不相关变量 (如  $L$ )。
- 关于  $C$  的参数大小:  $C$  的参数**估计值较大**, 很可能表明至少**遗漏了一个变量**。
  - 因为如果遗漏了一个变量, 那么C也会包含其他解释变量, 比如其他奢侈的功能、汽车的总体质量等。因此其估计值就会比较大。

c. 运用四个设定原则的评判过程

- **理论**: 发动机越大的轿车价格越高, 这个符合经济学和生活常理, 所以从理论上讲, 该变量在方程中似乎不是不相干变量。但是, 发动机大的轿车的重量更大, 所以  $W$  和  $L$  在一定程度上衡量的是同一个东西。
- **t检验**: 第一小问可知, 变量的参数估计值在预期的方向上**不显著**, 其t统计量大小仅为  $t_4 = \frac{0.19 - 0}{0.20} = 0.95$
- **F检验**: 当变量被剔除出方程后, 调整的判定系数  $\overline{R^2}$  变高了0.01, 方程总体的拟合优度提高了。
- **偏误**: 当变量被剔除出方程后, 参数估计值改变并不明显。

综上, 我更倾向于模型T。

## P108 习题6

- 6 为了确定性别或者种族是否对一个孩子的经济理解力的测试成绩有显著影响，可以做一个回归。<sup>⑨</sup>将第  $i$  个同学在基础经济学知识测试上的得分 ( $S_i$ ) 看作第  $i$  个学生基本能力测试的合成分数、性别虚拟变量 (如果第  $i$  个学生为女生则为 1，否则为 0)、第  $i$  个学生父母的平均受教育年限和种族虚拟变量 (如果第  $i$  个学生为白人则为 1，否则为 0) 的函数。不幸的是，一场暴风雨冲毁了计算机中心，使计算机输出的各个变量无法分辨，只知道回归估计方程 (括号内的数值为标准差) 为：

$$\hat{S}_i = 5.7 - 0.63X_{1i} - 0.22X_{2i} + 0.16X_{3i} + 1.20X_{4i}$$

$$(0.63) \quad (0.88) \quad (0.08) \quad (0.10)$$

$$N = 24 \quad \bar{R}^2 = 0.54$$

- 确定各个结果分别对应的是哪个变量，请具体说明。
- 解释问题 a 中得出的答案。
- 假设你的答案是正确的，建立合适的假设，并在 5% 的显著性水平下进行假设检验，得出该样本中种族和性别对测试分数产生何种影响。
- 在问题 c 中，使用的检验是单侧检验还是双侧检验？为什么？

## 回答

- $X_1$  : 两个虚拟变量之一；  $X_2$  : 两个虚拟变量之一；  $X_3$  : 父母的平均受教育年限；  $X_4$  : 学生基本能力测试的合成分数
- 根据生活常识，预期父母的平均受教育年限和学生基本能力测试的合成分数为正数，
  - 同时，预计学生基本能力测试的合成分数这一项比父母的平均受教育年限更显著。因此，区分了  $X_3$ 、 $X_4$
  - 那么剩下的两个变量就是虚拟变量，但是并没有常识或者佐证可以进一步区分  $X_1$ 、 $X_2$  具体代表哪个虚拟变量
- 假设检验过程如下
  - 在显著性水平 5% 的条件下先对  $X_1$ 、 $X_2$  进行双侧假设检验 (因为不确定其符号)。其中，样本总数为 24，可以求得自由度为  $n - k - 1 = 24 - 4 - 1 = 19$ ，查表得单侧  $t_c = 2.093$ 
    - $X_1$  : 做出如下假设：  $H_0 : \beta = 0$   $H_A : \beta \neq 0$ ，那么  $t_1 = \frac{-0.63 - 0}{0.63} = -1$ ，可知  $|-1| < 2.093$  即  $|t_1| < t_c$ ，因此不能拒绝  $H_0$
    - $X_2$  : 做出如下假设：  $H_0 : \beta = 0$   $H_A : \beta \neq 0$ ，那么  $t_2 = \frac{-0.22 - 0}{0.88} = -0.25$ ，可知  $|-0.25| < 2.093$  即  $|t_2| < t_c$ ，因此不能拒绝  $H_0$
  - 在显著性水平 5% 的条件下再对  $X_3$ 、 $X_4$  进行单侧假设检验 (因为已经预期其符号)。其中，样本总数为 24，可以求得自由度为  $n - k - 1 = 24 - 4 - 1 = 19$ ，查表得单侧  $t_c = 1.729$ 
    - $X_3$  : 做出假设  $H_0 : \beta \leq 0$   $H_A : \beta > 0$ ，且预测参数为正。  $t_3 = \frac{0.16 - 0}{0.08} = 2$ ，可知  $|2| > 1.729$  即  $|t_3| > t_c$ ，且其预期符号和对立假设相同 (均为正数)，因此可以拒绝  $H_0$

- $X_4$ : 做出假设  $H_0: \beta \leq 0$   $H_A: \beta > 0$ , 且预测参数为正。  $t_4 = \frac{1.2 - 0}{0.1} = 12$ , 可知  $|12| > 1.729$  即  $|t_4| > t_c$ , 且其预期符号和对立假设相同 (均为正数), 因此可以拒绝  $H_0$

**结论：种族和性别对测试分数的影响并不显著。**

- d. 对先前已经确定符号的参数 ( $X_3$ 、 $X_4$ ) 采用单侧检验，而对没有预估过符号的 ( $X_1$ 、 $X_2$ ) 就采取双侧检验。
- 原因：对于已经确定符号的参数，我们只需要检验参数值超过或小于该特定值的概率，而另一侧的概率不重要，因此采用单侧检验；对于没有预估过符号的参数，我们无法确定备择假设所关注的方向。在这种情况下，我们需要考虑参数值两个方向的可能性，因此采用双侧检验