

2022 年秋季学期《计量经济学》第 一 次课堂测试

测试日期：2022 年 10 月 日 测试时长：45 分钟 成绩：_____

一、判断题（正确的打“√”，错误的打“×”，每小题 4 分，共 5 小题，共 20 分）

- (√) 1. 相对而言，遗漏变量的后果比不相干变量的后果更加严重。
- (×) 2. 方程函数形式的选择主要取决于哪个函数形式拟合的最好。
- (×) 3. 如果回归方程选取 X 和 X^2 为解释变量，由于 X^2 是 X 的函数，方程存在完全多重共线性。
- (×) 4. 回归方程中，被解释变量等于其估计值与随机误差项之和。
- (×) 5. 假设检验通常将希望出现的结果作为原假设，将不希望出现的结果作为备择假设。

二、单项选择题（每小题 6 分，共 5 小题，共 30 分）

- (C) 1. 以下关于模型设定准则的说法**错误**的是_____
- A. 回归方程包含一个变量的最重要的准则是理论，而不是统计上的显著性
- B. 参数 t 检验不显著的变量不能简单地从方程中剔除
- C. 若变量加入方程后，判定系数 R^2 增大，方程应包含该变量
- D. 若变量加入方程，其他变量的系数符号发生改变且不再符合预期，方程不应

包含该变量

- (B) 2. 对多元线性回归方程的显著性检验，所用的 F 统计量可表示为

$$\begin{array}{ll} \text{A. } \frac{ESS/(n-k)}{RSS/(k-1)} & \text{B. } \frac{ESS/(k-1)}{RSS/(n-k)} \\ \text{C. } \frac{R^2/(n-k)}{(1-R^2)/(k-1)} & \text{D. } \frac{ESS}{RSS/(n-k)} \end{array}$$

- (A) 3. 讨论回归结果时不用花费太多时间去分析常数项的估计值，这主要依据的假设是_____

- A. 误差项总体均值为 0 B. 所有解释变量与误差项都不相关
- C. 误差项观测值互不相关 D. 误差项具有同方差

- (D) 4. 在体重 Weight 对身高 Height 的回归方程 $Weight_i = \beta_0 + \beta_1 Height_i + \varepsilon_i$ 中，加入新的变量邮箱号 Mail，重新回归，则以下表述**错误**的是_____

- A. Height 的参数估计值会发生改变 B. 判定系数 R^2 会增大
- C. 调整的判定系数 \bar{R}^2 会减小 D. Mail 的参数估计值一定为 0

- (D) 5. 若回归结果中参数的符号与预期不一致，我们**不应该**采取的做法有_____

- A. 检查数据是否存在录入错误或异常值

- B. 检查方程中是否存在遗漏变量
- C. 检查方程中是否存在严重的多重共线性
- D. 根据结果改变参数的预期符号

三、计算分析题（2 小题，共 50 分）

1. （共 25 分，每小题 5 分）

假如你决定建一个离你学校最近的冷冻酸奶商店的销售量模型。店主很乐意帮助收集数据，因为她相信你们学校的学生是她的主要顾客。经过长时间的数据收集以及没完没了的冷冻酸奶供给之后，你估计得到以下回归方程（括号内的数值为标准误）：

$$\hat{Y}_t = 262.5 + 3.9T_t - 46.94P_t + 134.3A_t - 152.1C_t$$

(0.6) (20.0) (108.0) (138.3)

$N = 29 \quad R^2 = 0.78$

式中， Y_t 代表第 t 个两周内冷冻酸奶的销售总量； T_t 代表 t 期的平均温度（单位：华氏温度）； P_t 代表 t 期该商店冷冻酸奶价格（单位：美元）； A_t 代表反映是否在学校报纸发布广告的虚拟变量（1=店主在学校报纸上做了广告）； C_t 代表反映是否为学校学期时间的虚拟变量（1= t 期是学校学期时间，即 9 月初到 12 月初、1 月初到 5 月底）。

自由度	单侧：10%	5%	2.5%
24	1.318	1.711	2.064
25	1.316	1.708	2.060

A. 为什么要假定“冷冻酸奶供给是没完没了的”？

供给可能影响销售量，且与价格相关。若不假定冷冻酸奶供给是没完没了的，则可能造成随机误差项与解释变量相关，违背基本假设。

B. 说明变量 A 的参数估计值 134.3 的经济含义。

在其他变量保持不变的情况下，如果店主在学校报纸上做了广告，则会使销售总量比不做广告时增加 134.3 个单位。

C. 对变量 T 的参数建立适当的假设，并在 5% 的显著性水平下进行检验。

自由度为 24，单侧 5% 水平下，临界值为 1.711。

变量	预期符号	t 值	临界值	是否拒绝原假设
T	正	6.5	1.711	拒绝

D. 对变量 P 的参数建立适当的假设，并在 5% 的显著性水平下进行检验。

自由度为 24，单侧 5% 水平下，临界值为 1.711。

变量	预期符号	t 值	临界值	是否拒绝原假设
P	负	2.357	1.711	拒绝

E. 你和店主对变量 C 的参数符号都很惊讶。你能解释为什么吗？

回归结果显示，学校学期时间的酸奶销售量要少于非学期时间，不符合预期。可能模型缺少了变量。-152.1 不只是学期期间对销售量的平均影响，而包含了其他未考虑的因素（例如季节因素）。

2. （共 25 分，每小题 5 分）

假设你是美国普外科的一名医生，研究决定吸烟行为的因素，你利用所有 50 个州的数据估计出了以下截面模型（括号内的数值为标准误）：

$$\hat{C}_i = 100 - 9.0 E_i + 1.0 I_i - 0.04 T_i - 3.0 V_i + 1.5 R_i$$

(3.0) (1.0) (0.04) (1.0) (0.5)

$t = -3.0 \quad 1.0 \quad -1.0 \quad -3.0 \quad 3.0$

$\bar{R}^2 = 0.40 \quad N = 50(\text{州})$

式中， C_i 代表第 i 个州每人每天消费香烟的数量； E_i 代表第 i 个州 21 岁以上人群的平均受教育年限； I_i 代表第 i 个州的平均收入（单位：1000 美元）； T_i 代表第 i 个州每包香烟的税负； V_i 代表第 i 个州三大网络播放的电视反吸烟广告的数量； R_i 代表第 i 个州五大网络播放的电台反吸烟广告的数量。

自由度	单侧：5%	2.5%
40	1.684	2.021
60	1.671	2.000

（1）建立关于该方程中变量的参数的恰当假设，并在 5% 的显著性水平上进行检验。自由度为 44，单侧 5% 水平下，临界值为（1.671, 1.684）。

变量	E	I	T	V	R
预期符号	负	正	负	负	负
t 值	-3.0	1.0	-1.0	-3.0	3.0
是否拒绝原假设	拒绝	不拒绝	不拒绝	拒绝	不拒绝

（2）在你看来方程中是否有不相干变量？是否有遗漏变量？请解释原因。

①方程中 T_i 可能是不相干变量。从理论上讲，香烟的税负对每人每天消费香烟的数量的影响较微弱，且税负不会经常变动。而在方程中 T_i 的参数方差很小，且 t 检验并不显著。

②方程中有遗漏变量。首先从理论上，香烟的价格是影响消费香烟数量的重要因素，而这一变量并没有在方程中出现。其次，方程中其余变量的参数出现了未预期的符号。

（3）假设你对 b 中两个问题的回答都是肯定的。这两者中哪个问题更重要，需要首先解决不相干变量还是遗漏变量？为什么？

遗漏变量的问题更重要。需要首先解决遗漏变量，因为遗漏变量会导致回归方程中其余变量的参数估计值有偏，而不相干变量的引入不会引起偏误。

（4）对方程进行回归的目的之一是确定电视与广播电台中反吸烟广告的有效性。你的结论是什么？

方程表明，电视中反吸烟广告对控制香烟消费是有效的，而电台中反吸烟广告是无效的。但由前面的推断得出方程中存在遗漏变量产生了偏误，因此电台中吸烟广告的有效性不能确定。

(5) 普外科医生认为税率与吸烟量无关，要求将这个变量从方程中剔除。假设（剔除该变量后）结果如下，运用四个设定准则判断你是否应赞同她的判断。请说明理由（括号中数值是标准误）。

①理论：因为每人每天消费香烟的数量可能对香烟的税负十分不敏感，故剔除该变量在理论上似乎是合理的。

② t 检验：这个疑似不相干变量（香烟的税负）的 t 统计量值是 1.0，在任何水平上都是不显著的。

③调整的判定系数：当该变量剔除后， R^2 没有变化，方程的整体拟合优度没有得到改善。

④偏误：当 T_i 被剔除方程后，其余变量的参数只是略有变动，意味着除去该变量不会造成偏误。

依据上述分析，可能得出香烟的税负是不相干变量的结论。但是，由于方程遗漏了价格变量，而税率与价格有一定相关性，因此应该加上价格变量之后再分析。