名词解释、简答、计算说明各若干道题。

一、名词解释

1、最小二乘法

2、横截面数据

3、时间序列数据

4、面板数据

5、无偏性

6、有效性

7、一致性

8、高斯-马尔科夫定理

9、哑变量

10、离散变量

11、连续变量

12、随机抽样

13、大数定理

14、中心极限定理

15、分布函数

二、简答

1、写出一元线性回归要得到好的估计参数性质的五个假设。

2、请说明使用stata软件（或其他统计软件）的命令：（1）检验两个变量x1和x2的均值是否相同；（2）将因变量y对自变量x进行回归分析；（3）将y和x分别为纵轴和横轴做散点图。

3、写出拟合优度R2的计算公式，说明其经济意义及存在问题，以及调节办法。

4、说明判断参数估计性质的三个标准及其含义

5、当分析变量X对变量Y影响时，如何能够保证排除其他因素影响的干扰？

6、说出你所知道的三个微观数据库？

7、交通方式有步行、公交、自驾、骑车，当此变量进入回归模型做自变量时，应该如何加入模型？

8、一个城市内建设了一个垃圾处理厂，要评估对周围房价产生的影响，请说出你的评估策略。

9、常见分布函数有标准正态分布、卡方分布、t分布和F分布，说明这些分布函数在回归分析时的主要价值何在？

10、分别说明使用横截面数据和使用时间序列数据回归分析时可能存在的问题。

三、计算说明

1、分析收入随年龄的变化关系，得到如下估计方程：

Log(Income) = 10 + 0.5 \* age - 0.01 \*age^2 + e

(1.6) (0.1) (0.003)

括号内为标准误估计量，e为残差。请说明：

1. 年龄和收入间存在怎样的关系？
2. 年龄多大时收入最高？

2、如果我们要分析妇女吸烟支数对生育子女重量影响，可以建立一元线性回归模型如下：

Bw= b0+b1\*cigs+u

其中，Bw代表出生儿体重，单位盎司，cigs代表吸烟支数。问：

（1）建立这样的方程，误差项u里面可能包含哪些因素？

（2）如果b0和b1估计值分别为119和-0.5,他们的含义是什么？

（3）如果吸烟的单位由支数变为包数（每包10只烟），那么cigs的估计系数将变为多少？

3、城市住房租金（rent）受到人口量（pop）和人均收入（aveinc）影响，建立如下回归模型：

Ln(rent)= b0+b1\*ln(pop)+ b2\*ln(aveinc) +u

（1）请预测b1，b2的符号。

（2）如果估计出来ln(aveinc)系数为0.51，标准误0.05，请问这个系数是显著不为0的么？

（3）同（2）问的估计结果，解释这个系数估计量对应的经济意义。

4、使用一元线性回归估计房价与距离垃圾焚化炉远近的影响，得到如下估计方程：

Ln(price)= 9.4 + 0.312 ln(dist)

R2=0.2

1. 你认为ln(dist)的系数符号符合预期吗？

（2）你认为这种简单回顾系数反映了真实的距离对房价的影响吗？

（3）除了距离焚化炉距离，还有哪些影响房价的因素？这些因素可能和距离焚化炉远近相关吗？

5、分析CEO薪酬影响因素，得到如下估计方程：

Ln(salary)= 4.5 + 0.18 ln(sales) + 0.10 ln(mrkvalue) -0.002 profmkt + 0.017 ceoten +u

(0.25) (0.04) (0.04) (0.002) (0.005)

Salary为CEO的薪酬，sales为公司销售额，mrkvalue为公司市值，profmkt为利润占销售额比重，ceoten为CEO担任当前公司CEO年数，问：

（1）评价profmkt对薪酬影响

（2）销售额对薪酬影响显著吗？

（3）解释ceoten系数的经济内涵。