计量经济学期末试卷2020.6（会计投资18）

###### ****一、 单选题（共10题，20分）****

**1、根据样本量n=30的样本估计的结果，一元线性回归的DW值为1.3。已知在5%的显著性水平下，dl=1.352，du=1.489，那么认为原模型**

**A、 存在一阶正自相关**

**B、 存在一阶负自相关**

**C、 不能确定**

**D、**

**不存在自相关性**

**正确答案： A**

**2、在二值因变量模型中，因变量y的预测值为0.6意味着**

**A、 给定解释变量的值，因变量的值为0.6%**

**B、 给定解释变量的值，因变量等于1的概率为60%**

**C、 该模型没有意义，因为因变量只能取0或者1**

**D、 给定解释变量的值，因变量等于1的概率为40%**

**正确答案： B**

**3、以下关于二值因变量模型的说法正确的是**

**A、 OLS方法在所有模型设定下均不再适用**

**B、 应优先考虑使用非线性最小二乘法进行估计**

**C、 当自变量也包含二值变量时，仍可使用线性概率模型**

**D、 可采用正确预测的比例、R方或伪R方衡量模型的拟合优度**

**正确答案： C**

**4、DW检验适用于下列哪种情况的检验**

**A、 AR(2)形式的自相关性**

**B、 正的一阶自回归形式的自相关性**

**C、 解释变量中含有滞后被解释变量的回归模型中的自相关性**

**D、 过原点回归模型中的自相关性**

**正确答案： B**

**5、如果一元线性回归模型中存在自相关性，则OLS估计不具有下列哪个性质**

**A、 线性性**

**B、 无偏性**

**C、 有效性**

**D、 一致性**

**正确答案： C**

**6、虚拟变量陷阱（dummy variable trap)是以下哪个情形**

**A、 不完全多重共线性**

**B、 仅仅是理论所关心的**

**C、 完全多重共线性**

**D、 实际操作中不会发生的**

**正确答案： C**

**7、将因变量的值扩大10，将自变量的值同时扩大100，则：**

**A、 斜率的估计值不变**

**B、 截距的估计值不变**

**C、 回归的R方不变**

**D、 OLS估计量的方差不变**

**正确答案： C**

**8、估计量具有抽样分布的原因是**

**A、 给定X的情况下，误差项的不同实现会导致Y的取值有所不同**

**B、 在现实数据中你往往会重复得到多组样本**

**C、 经济数据是不精确的**

**D、 不同的人可能有不同的估计结果**

**正确答案： A**

**9、对于模型lnY＝β0 + β1ln⁡X+u，其中的系数β1的实际含义是**

**A、 当X成比例增加1%时，Y会成比例变化β1%**

**B、 当X成比例增加1%时，Y会变化0.01β1%**

**C、 X增加一个单位，Y会变化β1**

**D、 X增加一个单位，Y会成比例变化100β1%**

**正确答案： A**

**10、如果一个假设在5%的显著水平下不能被拒绝，则它**

**A、 在10%的显著水平下一定不会被拒绝**

**B、 在10%的显著水平下一定被拒绝**

**C、 在1%的显著水平下可能被拒绝**

**D、 在1%的显著水平下一定不会被拒绝**

**正确答案： D**

###### ****二、 判断题（共10题，20分）****

**1、Probit模型的预测效果好于Logit模型。**

**正确答案： 错误**

**2、自相关性都会造成低估OLS估计量的真实方差。**

**正确答案： 错误**

**3、用DW统计量估计自回归系数ρ 只适用于一阶自相关性。**

**正确答案： 正确**

**4、像横截面数据一样，我们可以假定大多数时间序列观测是独立分布的。**

**正确答案： 错误**

**5、时间序列回归中的OLS估计量在时间序列前三个高斯-马尔科夫假定下是无偏的。**

**正确答案： 正确**

**6、当模型中遗漏了重要变量时，WLS优于OLS。**

**正确答案： 错误**

**7、异方差稳健的标准误估计总是优于同方差的标准误估计。**

**正确答案： 错误**

**8、在多元回归中，一个含有趋势的变量不能作为因变量。**

**正确答案： 错误**

**9、解释变量与残差之间的样本协方差总是为零**

**正确答案： 正确**

**10、存在严重的多重共线性时，参数估计的标准误差变大**

**正确答案： 正确**

###### ****三、 简答题（共2题，20分）****

**1、对于横截面数据，多元线性回归模型的高斯-马尔科夫假定是什么（3分）？无偏性需要哪几个假定（2分）？有效性需要哪几个假定（2分）？简述高斯--马尔科夫定理（3分）。**

**正确答案：**

**假定1线性于参数，假定2随机抽样，假定3不存在完全共线性，假定4零条件均值，假定5同方差。无偏性需要假定1到4，有效性需要假定1到5。高斯-马尔科夫定理是在高斯-马尔科夫假定（或假定1到5）下，OLS估计量是最佳（优）线性无偏估计量（BLUE）。（用公式表达假定也可以）**

**2、利用工资数据文件，此数据是有关935名男性在2010年的月薪，受教育年数，工作经验，婚姻，智商，父亲的受教育年数等的一个横截面数据。为了研究教育回报，小王构建了一个工资对教育和经验的二元线性回归模型，通过分析能力对工资有很大的影响，而且能力与教育有关，但能力不可观测，这样遗漏能力变量会导致教育回报估计有偏不一致。**

**要求（1）请分析教育回报估计偏误的方向。3分**

**（2）为了解决偏误问题，根据目前的数据条件，具体说说可以采用哪两种方法进行估计。 4分**

**（3）简述工具变量满足的两个假定条件。  3分**

**正确答案：**

**（1）因为能力和教育及工资都是正相关，所以偏误是正偏（偏大或高估）。**

**（2）可以用智商作为能力的代理变量进行估计，也可以用父亲的教育作为儿子教育的工具变量进行估计。**

**（3）工具外生性，工具变量和误差项（遗漏变量）不相关；工具相关性，工具变量与内生解释变量相关。（用公式表达也可以得分）。**

###### ****四、 计算题（共3题，40分）****

**1、假设一个关于拘捕率的一个线性概率模型：Y=β0+β1X1+β2X2+β3X3+β4X4+β5X5+u。其中y为一个二值变量，若在2019年间曾被拘捕过则取值1，否则取0。X1表示被捕后定罪的概率，X2表示以前定罪后关进监狱的平均月数，X3表示从18岁到2019年坐牢的总月数，X4表示2019年坐牢的总月数，而X5表示此人在2019年合法就业的季度数。**

**估计结果如下：**

**Y=0.4410-0.1620X1+0.0061X2-0.0023X3-0.022X4-0.043X5**

**（0.017）（0.021）（0.0065）（0.005）（0.005）（0.005）**

**n=2725，R方=0.0474，括号内为标准误。**

**要求：**

**（1）解释截距项0.441的含义。  3分**

**（2）X2的系数符合我们认为的更长判刑会阻止犯罪的直觉吗？  3分**

**（3）在α=0.05的水平下分别检验变量X2和X3系数的显著性。（t=1.96）  3分**

**（4）如果一个人在2019年在监狱里坐牢6个月，与一个其他条件相同但2019没有坐过牢的人相比，其被拘捕的概率有何变化？ 3分**

**（5）讨论就业对拘捕率的影响。一个2019年四个季度都工作的人，与一个其他条件相同但完全不工作的人相比，被拘捕的可能性有何差异？   3分**

**正确答案：**

**（1）0.441表示一个从未定过罪，18岁以后从未坐过牢，2019年也没有进过监狱，而且整个2019年未就业的人，预计被拘捕的概率（可能性）。（或所有自变量都为0时因变量的取值，或其他近似表达即可得分）**

**（2）X2系数为正，表示更长的判刑会增长拘捕（犯罪）率，这不符合（违背）直觉。**

**（3）根据t=β/se，T（X2）=0.938；T（X3）=0.46；均小于1.96，所以不拒绝原假设，系数均不显著。**

**（4）如果一个人在2019年中在监狱里坐牢6个月，与一个其他条件相同但2019没有坐过牢的人相比，其被拘捕的概率减少0.132，因为6\*0.022=0.132。**

**（5）一个2019年四个季度都工作的人，与一个完全不工作的人相比，被拘捕的可能性降低0.172，因为4\*0.043=0.172。**

**2、在一项关于工作技能培训对农民工工资影响的研究中，已知是否参加培训基于自愿原则。研究者通过调查获取500个样本数据并得到以下模型估计结果：**

**log(wage)=0.150+0.061train+0.045train.educ+0.138educ+0.080exper**

**（0.071）（0.018） （0.014）          （0.059）    （0.045）**

**其中，log(wage)表示农民工的小时工资的自然对数，train表示是否参加工作技能培训（1=是），educ表示农民工的受教育年限，exper表示工作时间长短。 （备注：当观测值数量足够大时，1%，5%，10%对应的双边t检验值依次为2.58，1.96，1.65，括号内为标准误，R2=0.42。）问：**

**（1）其他变量保持不变，接受培训的农民工增加1年教育和没有接受培训的农民工增加1年教育对工资的影响相同吗？相差多少？请在1%显著性水平检验其显著性。  3分**

**（2）如果高工资的人群拒绝回答他们的工资水平，这样高工资人群不在样本中，则样本存在什么问题，对估计有什么影响？  3分**

**（3）是否参加培训基于自愿原则或先到先得的原则，这是什么问题？会有什么影响？  3分**

**（4）有研究认为以上模型中应该加入educ2和exper2，在加入这两个平方项后所估计的模型R2=0.56。请问前面所建立的模型是否有可能存在模型设定偏误（遗漏变量）的问题？（5%显著性水平F临界值为3。）  3分**

**正确答案：**

**（1）两者不相同。多1年的教育对培训过的农民工来说工资会比没有培训的工人高4.5%。培训和教育的交互项是显著的（t=0.045/0.014=3.214>2.58 ）**

**（2）样本选择是建立在因变量基础上的，存在内生样本选择问题（或非随机样本问题），带来有偏（不一致的）估计。**

**（3）这是个样本自选择问题，（培训的参与是基于自愿原则实行的，能力低的人更多地参加培训，培训与误差项中的能力相关）会造成有偏（不一致的）估计。**

**（4）F=78>3，所以有证据证明第一个设定的模型存在模型误设（遗漏变量）问题。**

**3、学习通资料中的Eviews数据集hprice.wf1包含了美国波士顿506个社区的房屋中位数价格的横截面数据。考虑如下线性回归模型：**

**其中，lprice为房价的对数，lnox为空气污染程度的对数，ldist表示社区到就业中心距离的对数，rooms表示房屋的平均房间数，stratio表示社区学校的学生-教师比例。(注：要求写出操作命令或简要的菜单操作步骤，步骤中英文太长可以只写前面4个英文字符)**

**（1）试运用统计软件估计上述模型，用方程的形式报告你的回归结果，要求在方程下面注明参数估计值的标准误差、拟合优度R方、F统计量和样本量。4分**

**（2）以5%的显著性，检验“H0：β1=β4” ，写出检验结果的 F统计量及对应P值，结果是拒绝还是不拒绝原假设，说出你的判断标准。 3分**

**（3）以5%的显著性，检验“H0：β3=0.33” ，写出检验结果的F统计量和及对应P值，结果是拒绝还是不拒绝原假设，说出你的判断标准。 3分**

**（4）以5%的显著性，使用带有交互项的怀特检验，写出检验结果的Obs\*R-squared 及对应的P值，根据P值判断是否存在异方差。（要求用菜单操作，写出简单的操作步骤） 3分**

**正确答案：**

**（1）命令：ls  log(price)  c  log(nox)   log(dist)  rooms   stratio 或菜单quick/esti/**

**（0.318）（0.117） （0.043）  （0.019）    （0.006）**

**R2=0.584    F=175.855    n=506**

**(2)命令：wald  c(2)=c(5)或菜单view/coef/wald  输入c(2)=c(5)**

**F统计量为59.434，对应P值为0.000<0.05，因此拒绝原假设“H0：β1=β4”。**

**(3)命令：wald c(4)=0.33 或菜单view/coef/wald  输入c(4)=0.33**

**F统计量为16.589，对应P值为0.0001<0.05，因此拒绝原假设“H0：β3=0.33”。**

**(4)操作路径；view/resi/hete /white**

**Obs\*R-squared =nR2=143.975，对应的P值为0.000<0.05，因此拒绝原假设，表明模型存在异方差。**