实验08：虚拟变量

印度工人案例

任课教师：胡华平

2024-12-25

学生姓名：\_\_\_\_\_\_\_\_\_；学生学号：\_\_\_\_\_\_\_\_\_\_\_；专业班级：\_\_\_\_\_\_\_\_\_\_

# 1. 作业提交

**实验发布时间**：2024-12-25（周三）24:00:00

**提交截止时间**：2025-01-01（周三）24:00:00

**实验提交材料**：

（1）根据实验要求，完成Office Word电子文档一份（注意不能是wps文档），提交前请将文件命名为下述格式：lab08\_word\_张三\_2019000001.docx。

（2）根据实验要求，完成EViews相关操作，保存并提交1份EViews工作文件.wfl文件，提交前请将文件命名为下述格式：lab08\_eviews\_张三\_2019000001.wfl。

（3）根据实验要求，完成EViews相关操作，保存并提交1份EViews编程代码文件.prg文件，提交前请将文件命名为下述格式：lab08\_code\_张三\_2019000001.prg。

**实验提交方式**：

* 登陆西北农林科技大学[在线教育综合平台](https://eol.nwafu.edu.cn/meol/index.do) ->> 进入课程《计量经济学》（胡华平主讲，课程编号3133101） ->> 进入【课程作业】进行作业资料下载和作业提交。
* 请按上述要求命名各个提交文件。 系统提交页面中，务必让每个上传文件之间换行，保持提交页面布局美观！

# 2. 作业提示

## 2.1 如何在word中编辑数学公式？

（1）如果使用Office 2003版：“插入” $⇒$“对象”$⇒$“microsoft公式3.0”

（2）如使用Office 2007/2010版：“插入”$⇒$“新公式”

（3）使用独立公式软件Mathtype，任何Office版本都可以

a.在Mathtype中编写公式$⇒$ 确定无误后复制公式$⇒$然后粘贴到word中。

b.在Word中修改Mathtype形式的公式：$⇒$ 双击公式则可以打开Mathtype软件，然后按上一步骤操作，修改完成后，点击保存即可。

# 3. 作业内容

**温馨提示**： （1）作业配套数据请在作业发布平台界面中自行下载。 （2）每个同学的数据都不一样（但样本数相同$n=114$）。请下载数据表后，按后面作业要求找到自己的数据，并进行Excel预处理（以便导入到Eviews）。

为了分析一国存货量与销售额的关系，我们考虑如下所示的一些变量（具体定义见 [表 1](#tbl-vars) ）：

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 表 1: 变量定义及说明

| 变量\_代码 | 变量\_中文 | 定义和取值 |
| --- | --- | --- |
| obs | 工人编号 | 序号(observations) |
| wage | 工人工资 | 美元/周($/week) |
| age | 年龄 | 岁(year) |
| edu | 教育水平 | ill=文盲(illiteracy)；pri=初等教育(primary)；mid=中等教育(middle)；hig=高等教育(higher) |
| dpt | 合同类型 | tem=短期合同(temporary)；per=长期合同(permanent) |
| sex | 性别 | f=女(female)；m=男(male) |

 |

下 [图 1](#fig-data-show) 给出了存货与销售额实验案例的学生上机操作数据。

|  |
| --- |
| 图 1: 印度工人工资案例的学生实验数据（n=114） |

# 4. 作业任务

（1）请大家下载本次作业数据文件到本地电脑。

**温馨提示**：a.文件尽量不要放在电脑桌面，而是保存在自己清楚的文件夹路径下（如”D://econometrics//lab08”）；b.注意下载工具的使用，不是直接打开xlsx文件，而是要下载到本地电脑，然后再打开！

答：此问不用作答，完成指定操作即可！

（2）打开EViews软件，创建工作文件（WF），命名为lab08-dummy；以及建立工作页（page），命名为dummy。

答：此问不用作答，完成指定操作即可！

（3）将工作文件项目保存到本地电脑.wfl文件，并命名为“lab08\_eviews\_张三\_2019000001.wfl”的形式。

**要求**：注意记住保存的文件夹路径，这个文件要提交到作业系统的！！！

答：此问不用作答，完成指定操作并确保正确即可！

（4）对xlsx数据文件进行清洗处理，只保留必要数据列：obs、wage、age、edu、dpt、sex）。删除其他个人信息列。

答：此问不用作答，完成指定操作并确保正确即可！

（4）将清洗处理好后的数据导入到刚才建好的Eviews工作文件中。

答：此问不用作答，完成指定操作并确保正确即可！

## 4.1 题目1：虚拟变量编码

根据案例情形，数据集中存在三个定性变量$\left(edu\_{i},dpt\_{i},sex\_{i}\right)$，我们需要依次将它们转换为各自的虚拟变量体系：

list(wage, I(dpt\_per \* sex\_m), I(dpt\_per \* edu\_pri), I(dpt\_per \*
 edu\_mid), I(dpt\_per \* edu\_hig), I(sex\_m \* edu\_pri), I(sex\_m \*
 edu\_mid), I(sex\_m \* edu\_hig))

$$\begin{matrix}&edu\{ill,pri,mid,hig\}⇒\left\{\begin{matrix}edu\\_ill=\left\{\begin{matrix}1,&ill\\0,&not ill\end{matrix}\right.\\edu\\_pri=\left\{\begin{matrix}1,&pri\\0,&not pri\end{matrix}\right.\\edu\\_mid=\left\{\begin{matrix}1,&mid\\0,&not mid\end{matrix}\right.\\edu\\_hig=\left\{\begin{matrix}1,&hig\\0,&not hig\end{matrix}\right.\end{matrix}\right.\end{matrix}$$

$$\begin{matrix}&dpt\{tem,per\}⇒\left\{\begin{matrix}dpt\\_tem=\left\{\begin{matrix}1,&tem\\0,&not tem\end{matrix}\right.\\dpt\\_per=\left\{\begin{matrix}1,&per\\0,&not per\end{matrix}\right.\end{matrix}\right.\end{matrix}$$

$$\begin{matrix}&sex\{f,m\}⇒\left\{\begin{matrix}sex\\_f=\left\{\begin{matrix}1,&f\\0,&not f\end{matrix}\right.\\sex\\_m=\left\{\begin{matrix}1,&m\\0,&not m\end{matrix}\right.\end{matrix}\right.\end{matrix}$$

（1）请你按照上述规则，在EViews中 把定性变量**教育水平**$edu\_{i}$（m=4）设置成虚拟变量体系。然后把$edu\_{i}$序列以及其4个虚拟变量序列，以组对象的形式保存并命名为：group\_edu。最后将这个组对象元素截图到如下空白处。（**要求**：截图10个样本长度即可）。

**温馨提示**：在EViews命令窗口中输入并运行如下代码series edu\_ill=@recode(edu="ill",1,0)

答：截图如下

（2）请你按照上述规则，在EViews中 把定性变量**合同类型**$dpt\_{i}$（m=2）设置成虚拟变量体系。然后把$dpt\_{i}$序列以及其2个虚拟变量序列，以组对象的形式保存并命名为：group\_dpt。最后将这个组对象元素截图到如下空白处。（**要求**：截图10个样本长度即可）。

答：截图如下

（3）请你按照上述规则，在EViews中 把定性变量**性别**$sex\_{i}$（m=2）设置成虚拟变量体系。然后把$sex\_{i}$序列以及其2个虚拟变量序列，以组对象的形式保存并命名为：group\_sex。最后将这个组对象元素截图到如下空白处。（**要求**：截图10个样本长度即可）。

答：截图如下

## 4.2 题目2：定量变量的线性回归模型

如果可以将工资变量（$wage\_{i}$）对年龄变量（$age\_{i}$）进行如下的线性建模（见模型 [式 1](#eq-poly-model) ）：

$$\begin{matrix}\begin{matrix}\hat{wage}\_{i}=&+\hat{β}\_{1}+\hat{β}\_{2}age\_{i}+\hat{β}\_{3}age\_{i}^{2}\end{matrix} (polynomial model) \end{matrix}  \left(1\right)$$

（1）运用Eviews菜单（Quick $⇒$ Estimate Equation），对上述模型（见式 [式 1](#eq-poly-model) ）进行回归分析。在Eviews软件中，以方程对象（Equation）形式保存上述回归结果，并命名为eq\_poly。最后截图到下列空白处。

答：

（2）根据以上模型（见式 [式 1](#eq-poly-model) ）的EViews回归结果，请你计算出**年龄为30岁**工人的工资的估计值。（**要求**：写出主要计算过程，结果保留2位小数。）

答：

## 4.3 题目3：仅含定性变量模型（加法形式）

（1）考虑将工资变量（$wage\_{i}$）对三个定性变量（$edu\_{i},dpt\_{i},sex\_{i}$）进行**加法形式**的虚拟变量线性回归建模。请你写出以**{仅有文盲学历和临时合同的男性}**工人为基础组的虚拟变量回归总体模型。（**要求**：回归系数参数用$α\_{1},α\_{2},\cdots $表示；随机干扰项用$v\_{i}$表示）。

答：

$$\begin{matrix}wage\_{i}&=   &&(quality model)\end{matrix}  \left(2\right)$$

（2）运用Eviews菜单（Quick $⇒$ Estimate Equation），对上述模型（见式 [式 2](#eq-quality-model) ）进行回归分析。在Eviews软件中，以方程对象（Equation）形式保存上述回归结果，并命名为eq\_quality。最后截图到下列空白处。

答：

（2）根据以上模型（见式 [式 1](#eq-poly-model) ）的EViews回归结果，请你计算出**{仅有文盲学历和临时合同的男性}**工人的工资的估计值；以及**{有高学历和临时合同的女性}**工人的工资的估计值。（**要求**：写出主要计算过程，结果保留2位小数。）

答：a）**{仅有文盲学历和临时合同的男性}**工人的工资的估计值

b）**{有高学历和临时合同的女性}**工人的工资的估计值

## 4.4 题目4：仅含定性变量模型（完全乘法形式）

（1）考虑将工资变量（$wage\_{i}$）对三个定性变量（$edu\_{i},dpt\_{i},sex\_{i}$）进行**完全乘法形式**的虚拟变量线性回归建模（也即只出现交叉项，不含独立加法项）。请你写出以**{仅有文盲学历和临时合同的男性}**工人为基础组的虚拟变量回归总体模型。（**要求**：回归系数参数用$γ\_{1},γ\_{2},\cdots $表示；随机干扰项用$v\_{i}$表示）。

答：

$$\begin{matrix}wage\_{i}&=   &&(cross full)\end{matrix}  \left(3\right)$$

（2）运用Eviews菜单（Quick $⇒$ Estimate Equation），对上述模型（见式 [式 2](#eq-quality-model) ）进行回归分析。在Eviews软件中，以方程对象（Equation）形式保存上述回归结果，并命名为eq\_cross\_full。最后截图到下列空白处。

答：

（2）根据以上模型（见式 [式 3](#eq-cross-full) ）的EViews回归结果，请你计算出**{仅有文盲学历和临时合同的男性}**工人的工资的估计值；以及**{有高学历和临时合同的女性}**工人的工资的估计值。（**要求**：写出主要计算过程，结果保留2位小数。）

答：a）**{仅有文盲学历和临时合同的男性}**工人的工资的估计值

b）**{有高学历和临时合同的女性}**工人的工资的估计值

## 4.5 题目5：仅含定性变量模型（部分混合形式）

（1）考虑将工资变量（$wage\_{i}$）对三个定性变量（$edu\_{i},dpt\_{i},sex\_{i}$）进行**部分混合形式**的虚拟变量线性回归建模（也即：定性变量的虚拟变量独立项加法形式进入；但明确要求$sex\_{i}$与$edu\_{i}$的交叉项进入，且其他交叉项不进入模型）。请你写出以**{仅有文盲学历和临时合同的男性}**工人为基础组的虚拟变量回归总体模型。（**要求**：回归系数参数用$γ\_{1},γ\_{2},\cdots $表示；随机干扰项用$v\_{i}$表示）。

答：

$$\begin{matrix}wage\_{i}&=   &&(cross sex)\end{matrix}  \left(4\right)$$

（2）运用Eviews菜单（Quick $⇒$ Estimate Equation），对上述模型（见式 [式 4](#eq-cross-sex) ）进行回归分析。在Eviews软件中，以方程对象（Equation）形式保存上述回归结果，并命名为eq\_cross\_sex。最后截图到下列空白处。

答：

（2）根据以上模型（见式 [式 4](#eq-cross-sex) ）的EViews回归结果，请你计算出**{仅有文盲学历和临时合同的男性}**工人的工资的估计值；以及**{有高学历和临时合同的女性}**工人的工资的估计值。（**要求**：写出主要计算过程，结果保留2位小数。）

答：a）**{仅有文盲学历和临时合同的男性}**工人的工资的估计值

b）**{有高学历和临时合同的女性}**工人的工资的估计值

## 4.6 题目6：同时含定性和定量变量模型（部分混合形式）

（1）考虑将工资变量（$wage\_{i}$）对定量变量（$age\_{i}$,$age\_{i}^{2}$），以及三个定性变量（$edu\_{i},dpt\_{i},sex\_{i}$）进行**部分混合形式**的虚拟变量线性回归建模（也即：定性变量的虚拟变量独立项加法形式进入；但明确要求$sex\_{i}$与$edu\_{i}$的交叉项进入，其他交叉项不进入模型）。请你写出以**{仅有文盲学历和临时合同的男性}**工人为基础组的虚拟变量回归总体模型。（**要求**：回归系数参数用$γ\_{1},γ\_{2},\cdots $表示；随机干扰项用$v\_{i}$表示）。

答：

$$\begin{matrix}wage\_{i}&=   &&(cross hybrid)\end{matrix}  \left(5\right)$$

（2）运用Eviews菜单（Quick $⇒$ Estimate Equation），对上述模型（见式 [式 5](#eq-cross-hybrid) ）进行回归分析。在Eviews软件中，以方程对象（Equation）形式保存上述回归结果，并命名为eq\_cross\_hybrid。最后截图到下列空白处。

答：

（2）根据以上模型（见式 [式 4](#eq-cross-sex) ）的EViews回归结果，请你计算出**{仅有文盲学历和临时合同的男性}**工人的工资的估计值；以及**{有高学历和临时合同的女性}**工人的工资的估计值。（**要求**：写出主要计算过程，结果保留2位小数。）

答：a）**{仅有文盲学历和临时合同的30岁男性}**工人的工资的估计值

b）**{有高学历和临时合同的30岁女性}**工人的工资的估计值

 ## 题目7：模型比较分析

根据上述多个回归模型（见式 [式 1](#eq-poly-model) 、式 [式 2](#eq-quality-model) 、式 [式 3](#eq-cross-full) 、式 [式 4](#eq-cross-sex) 、式 [式 5](#eq-cross-hybrid) ）和EViews分析结果，你最终希望采用哪一个分析模型？请你判断女性工人是否受到了工资歧视？请简要说明你的理由和看法。

答：

## 4.7 题目8：线性约束条件及假设检验

无约束模型如下：

$$\begin{matrix}\begin{matrix}wage\_{i}=&+β\_{1}+β\_{2}age\_{i}+β\_{3}edu\\_pri\_{i}+β\_{4}edu\\_mid\_{i}\\&+β\_{5}edu\\_hig\_{i}+β\_{6}dpt\\_per\_{i}+β\_{7}dpt\\_tem:sex\\_f\_{i}+β\_{8}dpt\\_per:sex\\_f\_{i}+u\_{i}\end{matrix} (无约束) \end{matrix}  \left(6\right)$$

受约束模型如下：

$$\begin{matrix}\begin{matrix}wage\_{i}=&+α\_{1}+α\_{2}age\_{i}+α\_{3}edu\\_pri\_{i}+α\_{4}edu\\_mid\_{i}\\&+α\_{5}edu\\_hig\_{i}+α\_{6}dpt\\_per\_{i}+u\_{i}\end{matrix} (受约束) \end{matrix}  \left(7\right)$$

（1）运用Eviews菜单（Quick $⇒$ Estimate Equation），分别对上述模型（见 [式 6](#eq-prm-full) 和 [式 7](#eq-prm-cs) ）进行最小二乘回归分析。在Eviews软件中，以方程对象（Equation）形式保存上述回归结果，无约束回归命名为eq\_u，受约束回归命名为eq\_r。最后截图到下列空白处。

答：分别截图如下

（2）根据无约束模型最小二乘法的EViews回归结果，请你对模型中两个交互项变量（$dpt\\_tem\*sex\\_f$ 和$dpt\\_per\*sex\\_f$）前面的斜率参数（$β\_{7}$ 和$β\_{8}$）分别进行t检验并得到检验结论。给定显著性水平$α=0.05$，t查表值$t\_{1−α/2}\left(n−8\right)=t\_{0.975}\left(106\right)=1.98$。（**要求**：简单直接给出检验依据和结论即可）。

答：

（3）根据上述题（2）的t检验结论，你能否认为对于女性工人而言，无论是获得一份临时工作（$dpt\\_tem=1$）还是获得一份合同工作（$dpt\\_per=1$），对于女性工人的工资收入都是有显著影响的？

答：

（4）为了进一步验证上述题（3）的观点，我们可以结合无约束模型 （见 [式 6](#eq-prm-full) ）和受约束模型 （见 [式 6](#eq-prm-full) ）进行线性约束条件是否成立的假设检验（F检验）。结合无约束模型和受约束模型的EViews回归结果，请依次回答如下问题：

a）请你写出上述线性约束条件假设检验的原假设$H\_{0}$和备择假设$H\_{1}$。

答：

b）在EViews中，请分别提取无约束模型的判定系数$R\_{u}^{2}$，以及受约束模型的判定系数$R\_{r}^{2}$，保存为标量对象，并分别命名为r2\_u和r2\_r。同时，请分别计算无约束模型的残差平方和$RSS\_{u}$，以及受约束模型的残差平方和$RSS\_{r}$，保存为标量对象，并分别命名为rss\_u和rss\_r。

答：此题无需作答，完成相关操作即可。

c）请利用无约束模型的判定系数$R\_{u}^{2}$，以及受约束模型的判定系数$R\_{r}^{2}$，手动计算合适的F统计量$F\_{1}^{\*}$，保存为标量对象，并命名为F\_str1。

答：计算结果为（保留4位小数）$F\_{1}^{\*}=$

d）请利用无约束模型的残差平方和$RSS\_{u}$，以及受约束模型的残差平方和$RSS\_{r}$，手动计算合适的F统计量$F\_{2}^{\*}$，保存为标量对象，并命名为F\_str2。

答：计算结果为（保留4位小数）$F\_{2}^{\*}=$

e）为了进行上述F检验，请你利用EViews软件正确计算得到相应的F查表值。给定显著性水平$α=0.05$，请你计算F统计量的第一自由度$df\_{1}$和第二自由度$df\_{2}$，分别保存为标量对象，并分别命名为df\_1和df\_2。同时，请正确计算得到相应的F查表值$F\_{1−α/2}\left(df\_{1},df\_{2}\right)$，并保存为标量对象F\_val。

答：F理论查表值结果为（保留4位小数$F\_{1−α/2}\left(df\_{1},df\_{2}\right)=$

f）利用F查表值结果，请对上述题c）和题d）分别进行F检验，并得出检验结论。（**要求**：简要写出检验过程和结论即可。）

答：

g）请你利用EViews菜单操作，进行上述线性约束假设检验，将分析结果冻结Freeze保存为表格对象，并命名为test\_restrain。最后，请你核对是否与前述题c）和题d）的自己手动计算结果一样。（**要求**：将表格对象截图在下面空白处。）

EViews操作提示：（1）双击打开无约束回归对象eq\_u （2）对象视窗下点击菜单View，下拉项选择Coefficients Diagnostics；（3）下拉项选择Wald Test-Coefficient Rstrictions...，进入参数约束引导窗口，按提示填写约束条件。



答：

h）根据上述线性约束条件F检验，你能否认为对于女性工人而言，无论是获得一份临时工作（$dpt\\_tem=1$）还是获得一份合同工作（$dpt\\_per=1$），对于女性工人的工资收入都是有显著影响的？

答：