

计量经济学软件应用

——stata软件实验之一元、多元回归分析

内容概要

- 一、实验目的
- 二、简单回归分析的Stata基本命令
- 三、简单回归分析的Stata软件操作实例
- 四、多元回归分析的Stata基本命令
- 五、多元回归分析的Stata软件操作实例

一、实验目的

掌握运用Stata软件进行简单回归分析以及多元回归分析的操作方法和步骤，并能看懂Stata软件运行结果

一、简单回归命令的Stata基本命令

简单线性回归模型（Simple linear regression model）指只有一个解释变量的回归模型。如

$$y = \beta_0 + \beta_1 x + u$$

其中， y 为被解释变量， x 为解释变量， u 为随机误差项，表示除 x 之外影响 y 的因素。 β_1 称为斜率参数或斜率系数， β_0 称为截距项或截距系数。也称为截距项或常数项。简单线性回归模型的一种特殊情况：

$$y = \beta_1 x + u$$

即假定截距系数 $\beta_0 = 0$ 时，该模型被称为过原点回归；过原点回归在实际中有一定的应用，但除非有非常明确的理论分析表明 $\beta_0 = 0$ ，否则不宜轻易使用过原点回归模型。

一、简单回归命令的Stata基本命令

regress y x

以y为被解释变量，x为解释变量进行普通最小二乘（OLS）回归。regress命令可以简写为横线上方的三个字母reg。

regress y x,noconstant

不包括截距项（constant）的意思

predict z

根据最近的回归生成一个新的变量z，其值等于每一个观测的拟合值（即 \hat{y}_i ）。

predict u,residual

根据最近的回归生成一个新变量u，其值等于每一个观测的残差（即 $e_i = y_i - \hat{y}_i$ ）

三、简单回归命令的Stata软件操作实例

实验1：家庭成分对小时工资的影响

1、打开文件

2、首先对数据进行处理：

只保留部分数据：

```
keep lnhw male age mino marriage edu ew ew2 g_health cparty ra_exp_soc2 emp*  
rank3_al c_id
```

给数据加上标签：

```
label var c_id "citycode"
```

```
label var edu "教育年限"
```

```
label var ew "工作经验"
```

```
label var ew2 "工作经验平方"
```

```
label var age "年龄"
```

三、简单回归命令的Stata软件操作实例

实验1：家庭成分对小时工资的影响

```
label var rank3_al "家庭成分"  
label var male    "是否为男性"  
label var mino    "是否为少数民族"  
label var marrige "是否已婚有配偶"  
label var cparty  "是否为党员"  
label var g_health "自评健康状况"  
label var emp_sec  "行业"  
label var emp_ship "所有制"  
label var emp_contr "合同类型"  
label var emp_occup "职业"  
label var emp_num  "单位规模"  
label var lnhw     "小时工资对数"  
label var ra_exp_soc2 "人情往来支出"
```

三、简单回归命令的Stata软件操作实例

实验1：家庭成分对小时工资的影响

label def rank3_v 1 "New-elite" 2 "Old-elite" 3 "Non-elite" //创建一个名为rank3_v的值标签，它是一组单个数字值及其相应的标签。

label val rank3_al rank3_v //将rank3_al的值和rank3_v联系在一起

保存数据

生成新变量

根据rank3_al生成新的变量虚拟变量rank_1,rank_2,rank_3

同理：

```
foreach i in num sec ship contr occup {  
  tab emp_`i',gen(`i')  
}
```

```
tab c_id,gen(cid)
```


三、简单回归命令的Stata软件操作实例

实验1：家庭成分对小时工资的影响

采用全局字符串

```
global cid cid1-cid76
```

```
global dem age male mino marriage
```

```
global hc edu ew ew2 g_health
```

```
global pc cparty ra_exp_soc2
```

```
global wc num1 num2 num3 contr1 contr2 contr3 contr4 ///
```

```
sec1 sec2 sec3 sec4 sec5 sec6 sec7 ship1 ///
```

```
ship2 ship3 ship4 ship5 occup1 occup2 ///
```

```
occup3 occup4
```

三、简单回归命令的Stata软件操作实例

实验1：家庭成分对小时工资的影响

3、对数据进行描述性统计

```
ssc install estout
estpost tabstat lnhw male age mino marriage edu ew      ///
      g_health cparty ra_exp_soc2 ship1-ship6    ///
      sec1-sec8 occup1-occup5 num1-num4    ///
      contr1-contr5, by(rank3_al)    ///
      statistics(mean sd N) col(statistics)

esttab, main(mean %9.3f)      ///
      aux(sd %9.3f) replace nogap compress unstack    ///
      obslast nolabel order (lnhw male age mino      ///
      marriage edu ew g_health cparty ra_exp_soc2)
```

三、简单回归命令的Stata软件操作实例

实验1：家庭成分对小时工资的影响

3、对数据进行描述性统计

```
ssc install estout
estpost tabstat lnhw male age mino marriage edu ew      ///
      g_health cparty ra_exp_soc2 ship1-ship6    ///
      sec1-sec8 occup1-occup5 num1-num4    ///
      contr1-contr5, by(rank3_al)    ///
      statistics(mean sd N) col(statistics)

esttab, main(mean %9.3f)      ///
      aux(sd %9.3f) replace nogap compress unstack    ///
      obslast nolabel order (lnhw male age mino      ///
      marriage edu ew g_health cparty ra_exp_soc2)
```

三、简单回归命令的Stata软件操作实例

实验1：家庭成分对小时工资的影响

4、回归分析

```
reg lnhw rank1 rank2 ${dem}
```

三、简单回归命令的Stata软件操作实例

实验1：家庭成分对小时工资的影响

(1)表下方区域为基本的回归结果。第一列一次为被解释变量gni，解释变量lntrade，截距项constant；第二列是回归系数的OLS估计值；第三列为回归系数的标准误；第四列是回归系数的t统计值；第五列是P值；最后两列是95%的置信区间

三、简单回归命令的Stata软件操作实例

实验1：简单回归分析：贸易开放与我国城乡收入差距的关系

(2)表左上方区域为方差分析表。第二列从上到下依次为回归方差和（SSE）、残差平方和（SSR）和总离差平方和（SST）；第三列为自由度；第四列为均方和（MSS），由各项平方和除以相应的自由度得到。

(3) 表右上方区域给出了样本数（Number of obs）、判定系数（R-squared）、调整后的判定系数（Adj R-squared）、F统计量的值、回归方程标准误或均方根误（Root MSE, $\hat{\sigma}$,或S.E.）以及其他一些统计量的信息。

三、简单回归命令的Stata软件操作实例

实验1：家庭成分对小时工资的影响

```
reg lnhw rank1 rank2 ${dem},r
```

r是使用Robust Standard Error 控制异方差。我们回归大多数情况下，都要控制异方差。

```
est store reg_1
```

```
reg lnhw rank1 rank2 ${dem} ${cid},r
```

```
est store reg_2
```

```
reg lnhw rank1 rank2 ${dem} ${hc} ${cid},r
```

```
est store reg_3
```

三、简单回归命令的Stata软件操作实例

实验1：家庭成分对小时工资的影响

```
reg lnhw rank1 rank2 ${dem} ${hc} ${pc} ${cid},r  
est store reg_4
```

```
reg lnhw rank1 rank2 ${dem} ${hc} ${pc} ${wc} ${cid},r  
est store reg_5
```

5、将回归在表格中自动呈现

```
esttab reg_1 reg_2 reg_3 reg_4 reg_5
```

```
esttab reg_1 reg_2 reg_3 reg_4 reg_5, (%6.3f) se(%6.3f) nogaps compress star(* 0.1  
** 0.05 *** 0.01) replace r2(%6.3f) obslast keep(rank1 rank2 ${dem} ${hc} ${pc}  
${wc} _cons) order(rank2 rank1 ${dem} ${hc} ${pc} ${wc} _cons)
```