

实验13-1 软件开发人员的薪金——基本模型（验证，参见教材P332-335）

在下面给出的数据中：

第1列 编号

第2列 薪金 y

第3列 资历 x_1 （从事专业工作的年数）

第4列 管理 x_2 （1表示管理人员，0表示非管理人员）

第5列 教育 x_3, x_4 （1表示中学程度 $x_3x_4=10$ ，2为大学 $x_3x_4=01$ ，3为更高程度 $x_3x_4=00$ ）

01	13876	1	1	1
02	11608	1	0	3
03	18701	1	1	3
04	11283	1	0	2
05	11767	1	0	3
06	20872	2	1	2
07	11772	2	0	2
08	10535	2	0	1
09	12195	2	0	3
10	12313	3	0	2
11	14975	3	1	1
12	21371	3	1	2
13	19800	3	1	3
14	11417	4	0	1
15	20263	4	1	3
16	13231	4	0	3
17	12884	4	0	2
18	13245	5	0	2
19	13677	5	0	3
20	15965	5	1	1
21	12366	6	0	1
22	21352	6	1	3
23	13839	6	0	2
24	22884	6	1	2
25	16978	7	1	1
26	14803	8	0	2
27	17404	8	1	1
28	22184	8	1	3
29	13548	8	0	1
30	14467	10	0	1
31	15942	10	0	2
32	23174	10	1	3
33	23780	10	1	2
34	25410	11	1	2
35	14861	11	0	1
36	16882	12	0	2
37	24170	12	1	3
38	15990	13	0	1

39	26330	13	1	2
40	17949	14	0	2
41	25685	15	1	3
42	27837	16	1	2
43	18838	16	0	2
44	17483	16	0	1
45	19207	17	0	2
46	19346	20	0	1

假设满足多元线性回归模型

$$y = a_0 + a_1 x_1 + a_2 x_2 + a_3 x_3 + a_4 x_4 + \varepsilon$$

实验要求:

在工作区建立 mydata.m;

求回归系数及其置信区间（置信水平 $\alpha=0.05$ ）、检验统计量 R_2 、 F 、 p ，有关散点图的 MATLAB 程序如下:

```
% 软件开发人员的薪金——基本模型
%模型: y=a0+a1*x1+a2*x2+a3*x3+a4*x4+ε
clear;clc;format compact;format short g;
M=dlmread('mydata.m');%读取 ASCII 码文件
y=M(:,2); x1=M(:,3); x2=M(:,4);
for k=1:size(M,1) % M 的行数
    if M(k,5)==1 % 中学程度
        x3(k,1)=1; x4(k,1)=0;
    elseif M(k,5)==2 % 大学程度
        x3(k,1)=0; x4(k,1)=1;
    elseif M(k,5)==3 %更高程度
        x3(k,1)=0; x4(k,1)=0;
    end
end
[a,aint,r,rint,stats]=regress(y,[ones(size(M,1),1) x1 x2 x3 x4]);
disp('a0,a1,a2,a3,a4 估计值 置信区间');
[a,aint]
R2=stats(1)
F=stats(2)
p=stats(3)

subplot(121);
plot(x1,r,'+'); %模型的残差ε 与资历 x1 的关系

subplot(122);
for k=1:size(M,1)
    if M(k,4)==0 & M(k,5)==1
        x234(k)=1;
    elseif M(k,4)==1 & M(k,5)==1
        x234(k)=2;
    elseif M(k,4)==0 & M(k,5)==2
```

```

        x234(k)=3;
elseif M(k,4)==1 & M(k,5)==2
        x234(k)=4;
elseif M(k,4)==0 & M(k,5)==3
        x234(k)=5;
elseif M(k,4)==1 & M(k,5)==3
        x234(k)=6;
end
end
plot(x234,r,'+'); %模型的残差ε 与管理-教育 x2-x3,x4 组合的关系

```

实验报告提交:

程序的运行结果（数据和图形）。

实验13-2 软件开发人员的薪金——更好的模型（编程）

在实验 13-1 的模型中增加 x_2 与 x_3 , x_4 的交互项后，多元线性回归模型为

$$y=a_0+a_1x_1+a_2x_2+a_3x_3+a_4x_4+a_5x_2x_3+a_6x_2x_4+\varepsilon$$

实验要求:

同实验 13-1，区别在于模型不同，所以要根据新模型修改实验 13-1 的程序，运行程序。

实验报告提交:

程序和运行结果（程序、数据和图形）。