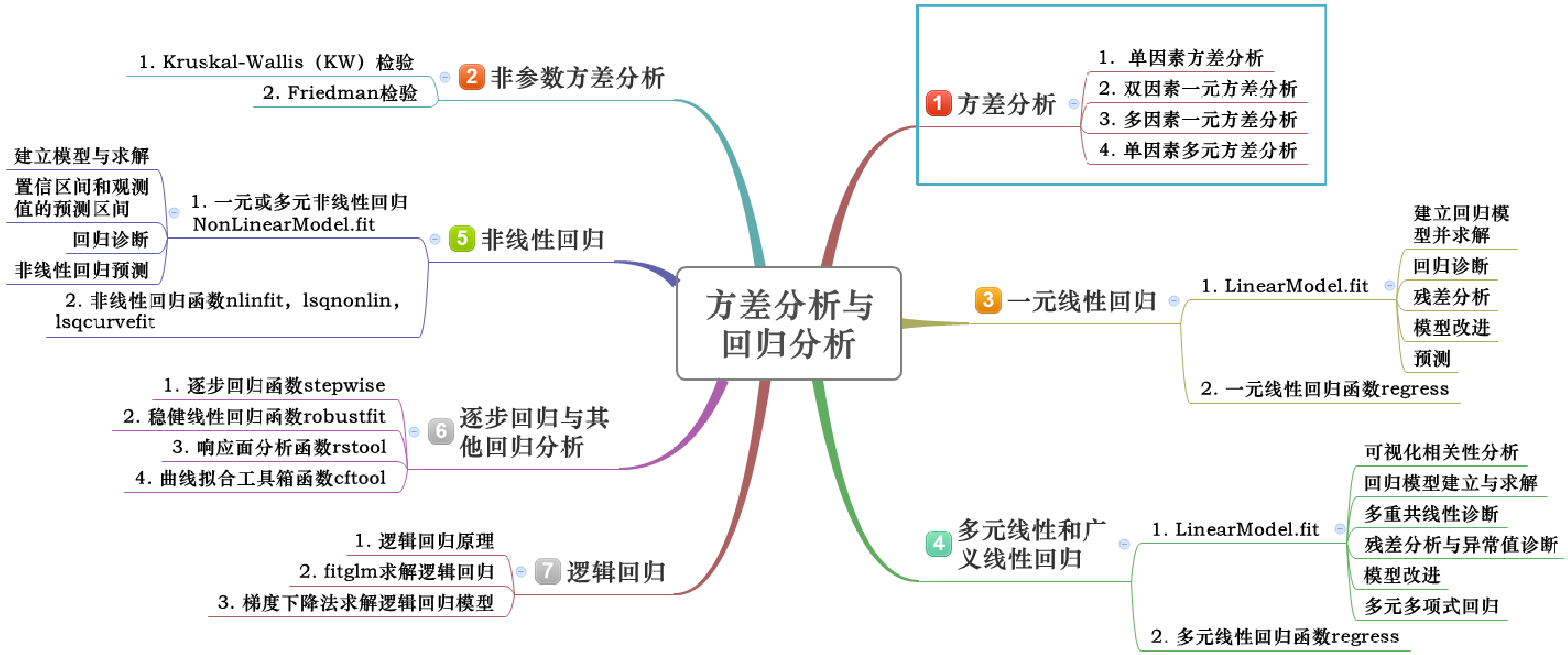
信陽師範學院
数学与统计学院
SCHOOL OF MATHEMATICS AND STATISTICS

第11章 方差分析与回归分析

讲授人：牛言涛

日期：2020年4月14日

第11章 方差分析与回归分析知识点思维导图



方差分析简介

信阳师范学院
数学与统计学院
SCHOOL OF MATHEMATICS AND STATISTICS

- 方差分析(Analysis of Variance，简称ANOVA)，又称“变异数分析”，英国统计学家R.A.Fisher在20世纪20年代提出的一种统计方法，用于两个及两个以上样本均数差别的显著性检验。
- 方差分析的基本原理是认为不同处理组的均数间的差别基本来源有两个：
  - 实验条件，即不同的处理造成的差异，称为组间差异。用变量在各组的均值与总均值之偏差平方和的总和表示，记作SSb，组间自由度dfb。
  - 随机误差，如测量误差造成的差异或个体间的差异，称为组内差异，用变量在各组的均值与该组内变量值之偏差平方和的总和表示， 记作SSw，组内自由度dfw。
- 在生产实践和科学研究中，经常要研究生产条件或实验条件的改变对产品的质量或产量的影响。如在农业生产中，需要考虑品种、施肥量、种植密度等因素对农作物收获量的影响；又如某产品在不同的地区、不同的时期，采用不同的销售方式，其销售量是否有差异。
- 在诸多影响因素中，哪些是主要的，哪些是次要的，以及主要因素处于何种状态时，才能使农作物的产量和产品的销售量达到一个较高的水平，这就是方差分析所要解决的问题。

方差分析简介

- 例：对A1~A4共4种不同灯丝的灯泡进行抽样检测灯泡寿命，根据测试数据，现在要问：灯泡寿命是否与灯丝材料的不同有关?

灯丝＼灯泡寿命	1	2	3	4	5	6	7	8
A1	1600	1610	1650	1680	1700	1700	1780	
A2	1500	1640	1400	1700	1750			
A3	1640	1550	1600	1620	1640	1600	1740	1800
A4	1510	1520	1530	1570	1640	1680		

- 问题特点:
  - 1 项指标（因变量）：寿命
  - 影响指标的因素（因子）：灯丝
  - 因素存在多个不同状态（水平: A1,A2,A3,A4），要求分析因素的不同状态是否对指标有显著影响
- 用数理统计分析试验结果、鉴别各因素对结果影响程度的方法称为方差分析（Analysis Of Variance），记作ANOVA。

方差分析简介

- 方差分析按影响分析指标的因素(自变量)个数的多少，分为单因素方差分析、双因素方差分析、三因素方差分析...
- 方差分析按分析指标（也可简单称为因变量）的个数多少，分为一元方差分析（即ANOVOA）、多元方差分析（即MANOVOA）...
- 多自变量多因变量的方差分析，可以简单称为多元方差分析，当然更精确的称为“X因素Y元方差分析”，如二因素二元方差分析。
- 本节介绍：单因素方差分析、双因素一元方差分析、多因素一元方差分析、单因素多元方差分析。

1. 单因素方差分析

单因素方差分析是比较两组或多组数据的均值，它返回原假设——均值相等的概率，函数anova1

- [p,table] = anova1(X,group,'displayopt')
  - ✓ X的各列为彼此独立的样本观察值，其元素个数相同，p为各列均值相等的概率值，若p值接近于0，则原假设受到怀疑，说明至少有一列均值与其余列均值有明显不同。
  - ✓ group是与X对应的字符或字符串数组，用来声明X每一列中数据的名字或意义，可以省略；
  - ✓ displayopt=on/off表示显示与隐藏方差分析表图和盒图
  - ✓ table为方差分析表

1. 单因素方差分析

anova1函数产生两个图：标准的方差分析表图和盒图。方差分析表中有6列：

第1列(source)显示：X中数据可变性的来源，方差来源有组间、组内和总计3种；

第2列(SS)显示：各方差来源所对应的离均差平方和；

第3列(df)显示：各方差来源所对应的自由度；

第4列(MS)显示：各方差来源所对应的均方，SS/df的比值；

第5列(F)显示：F检验统计量的观测值，它是组间均方与组内均方的比值；

第6列显示：从F累积分布中得到的概率，当F增加时，p值减少。

ANOVA Table

Source	SS	df	MS	F	Prob>F
Columns	0.00053	2	0.00027	41.43	5.91789e-09
Error	0.00017	27	0.00001		
Total	0.0007	29			

$$SSA=\sum_{i=1}^{k}n_i\left(\bar{x}_i-\bar{\bar{x}}\right)^2,\quad SSE=\sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(x_{ij}-\bar{x}_i\right)^2,\quad SST=\sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(x_{ij}-\bar{\bar{x}}\right)^2$$

1. 单因素方差分析

1. 正态性检验：函数lillietest，大样本可用jbtest。
2. 方差齐次性检验：函数vartestn，是对控制变量不同水平下各观测变量总体方差是否相等进行检验。控制变量不同水平下观测变量总体方差无显著差异是方差分析的前提要求。如果没有满足，就不能认为各总体分布相同。因此，有必要对方差是否齐性进行检验。
3. 方差分析：函数anova1
4. 多重比较：函数multcompare，如果控制变量确实对观测变量产生了显著影响，进一步还应确定控制变量的不同水平对观测变量的影响程度如何，其中哪个水平的作用明显区别于其他水平，哪个水平的作用是不显著的等。多重比较交互图，随multcompar输出参数而生成

注意：进行正态性和方差性齐次性检验，要求因素数据比较多；在数据比较少的情况下正态检验的结果是不可靠的，即使不满足方差分析的假定，方差分析的结果通常也是比较稳定的。

1. 单因素方差分析

例1：设有3台机器，用来生产规格相同的铝合金薄板。取样测量薄板的厚度，精确至‰厘米。得结果如下：

机器1：0.251 0.248 0.249 0.250 0.247 0.246 0.250 0.249 0.2451 0.252

机器2：0.257 0.253 0.255 0.254 0.261 0.251 0.253 0.252 0.254 0.256

机器3：0.258 0.264 0.259 0.262 0.262 0.260 0.261 0.255 0.257 0.258

检验各台机器所生产的薄板的厚度有无显著的差异？

```
>> X=[0.251  0.248  0.249  0.250  0.247  0.246  0.250 0.249 0.251  0.252;
  0.257  0.253  0.255  0.254  0.261  0.251  0.253  0.252  0.254  0.256;
  0.258  0.264  0.259  0.262   0.262  0.260  0.261 0.255  0.257  0.258]';
```

说明：可以不进行正态性和方差齐次性检验，因素数据少，但本题为了学习必要，采用检验。

```
>> [h1,p1] = lillietest(X(:,1))
>> [h2,p2] = lillietest(X(:,2))
>> [h3,p3] = lillietest(X(:,3))
% 三台机器的正态性检验均通过
```

1. 单因素方差分析

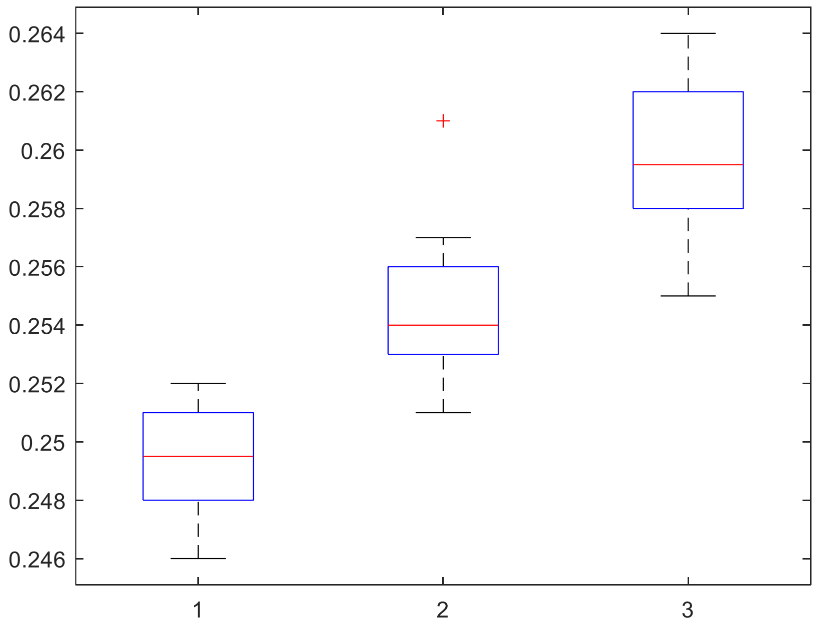
```
>> [p,stats] = vartestn(X)  %方差齐次性检验
p =
    0.4460
stats =
  包含以下字段的 struct:
    chisqstat: 1.6149
           df: 2
```

检验的p值=0.4460 > 0.05，说明在显著性水平0.05下接受原假设，认为3个机器所生产的薄板的厚度服从方差相同的正态分布，满足方差分析的基本假定。

Group Summary Table

Group	Count	Mean	Std Dev
1	10	0.2493	0.00189
2	10	0.2546	0.00288
3	10	0.2596	0.00272
Pooled	30	0.2545	0.00253

Bartlett's statistic	1.6149
Degrees of freedom	2
p-value	0.44599

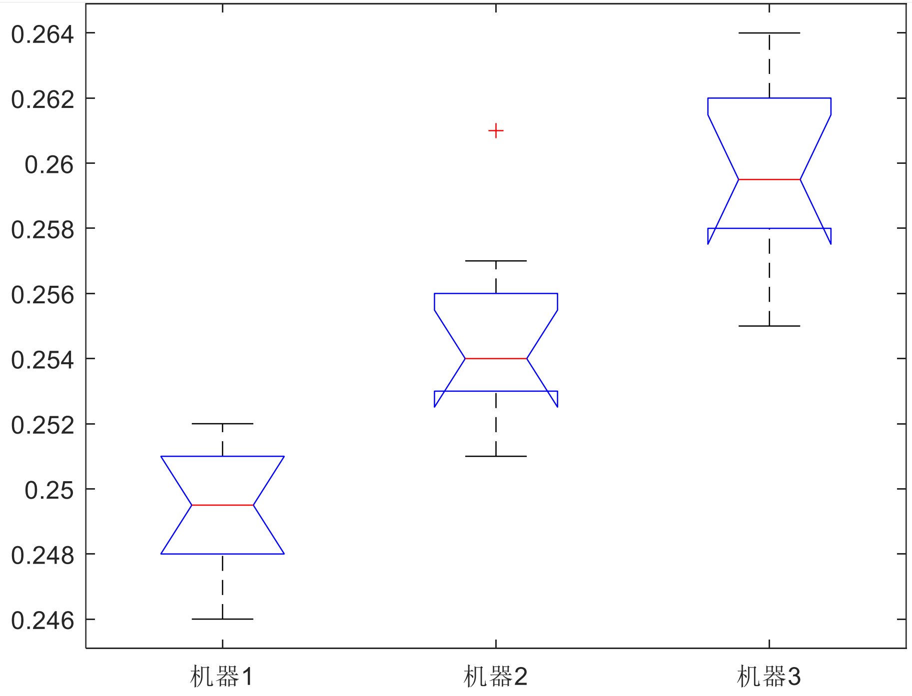


1. 单因素方差分析


```
>> group = {'机器1','机器2','机器3'};
>> [p,table,stats] = anova1(X,group)   %方差分析
p =
   5.9179e-09
table =
stats =
  包含以下字段的 struct:
    gnames: {3×1 cell}
         n: [10 10 10]
    source: 'anova1'
     means: [0.2493 0.2546 0.2596]
        df: 27
         s: 0.0025
```

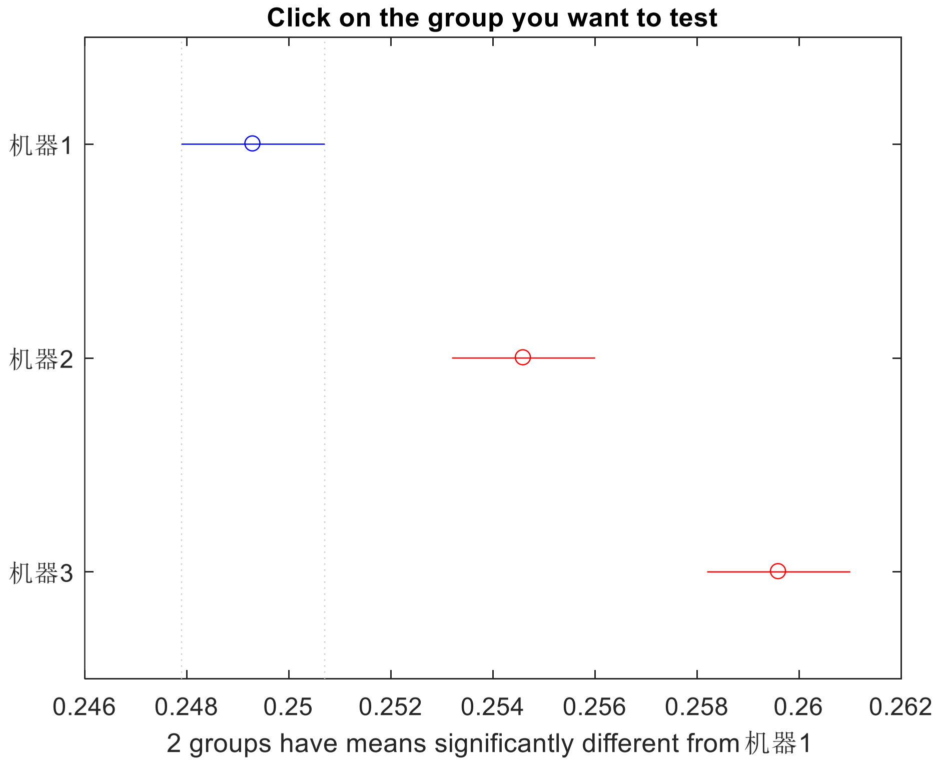
ANOVA Table

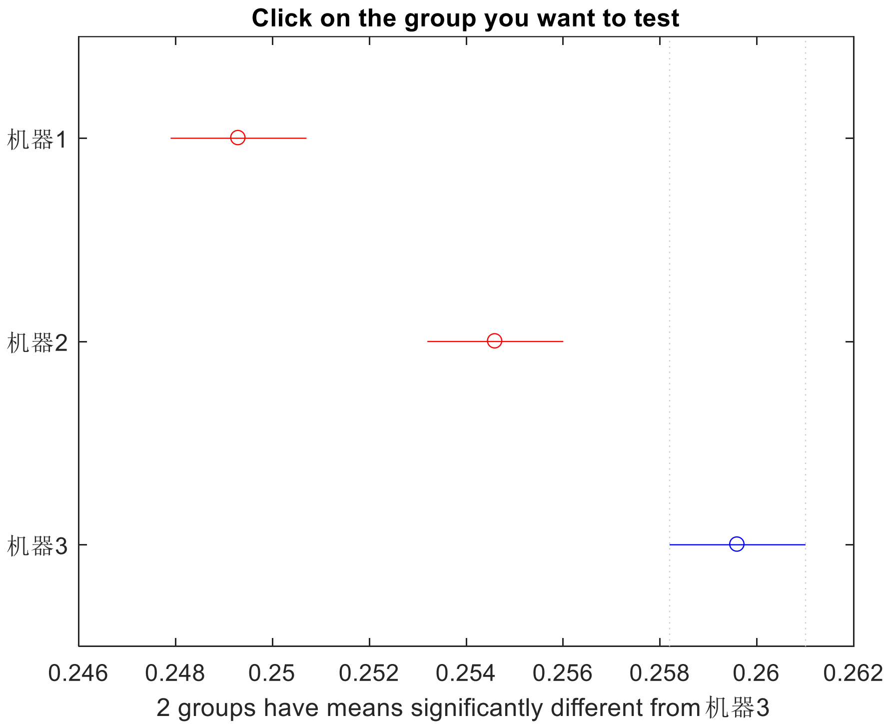
Source	SS	df	MS	F	Prob>F
Columns	0.00053	2	0.00027	41.43	5.91789e-09
Error	0.00017	27	0.00001		
Total	0.0007	29			



1. 单因素方差分析

```
>> [c,m,b,gnames]=multcompare(stats)  %多重分析
c =
    1.0000    2.0000   -0.0081   -0.0053   -0.0025    0.0002
    1.0000    3.0000   -0.0131   -0.0103   -0.0075    0.0000
    2.0000    3.0000   -0.0078   -0.0050   -0.0022    0.0004
m =
    0.2493    0.0008
    0.2546    0.0008
    0.2596    0.0008
```





1. 单因素方差分析

例2：将四种工艺下生产的灯泡进行寿命测试，得到数据表：

试验 \ 工艺	A_1	A_2	A_3	A_4
1	1620	1580	1460	1500
2	1670	1600	1540	1550
3	1700	1640	1620	1610
4	1750	1720		1680
5	1800			

试检验工艺对寿命有无显著影响。($\alpha=0.05$)

```
>>X=[1620,1670,1700,1750,1800,1580,1600,1640,1720,1460,1540,1620,1500,1550,1610,1680];
>> group=[1,1,1,1,1,2,2,2,2,3,3,3,4,4,4,4];
```

1. 单因素方差分析

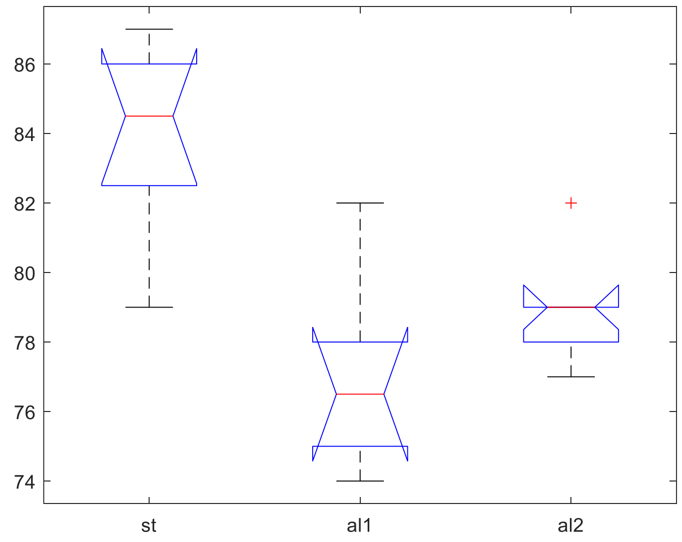
例3：建筑横梁强度的研究：3000磅力量作用在一英寸的横梁上来测量横梁的挠度，钢筋横梁的测试强度是：82 86 79 83 84 85 86 87；其余两种更贵的合金横梁强度测试为合金1：74 82 78 75 76 77；合金2：79 79 77 78 82 79。

检验这些合金强度有无明显差异?

```
>> strength = [82 86 79 83 84 85 86 87 74 82 78 75 76 77 79 79 77 78 82 79];
>> alloy = {'st','st','st','st','st','st','st','st', 'al1','al1','al1','al1','al1','al1','al2','al2','al2','al2','al2','al2'};
>> [p,table,stats] = anova1(strength,alloy)   %方差分析
p = 1.5264e-04
```

ANOVA Table

Source	SS	df	MS	F	Prob>F
Groups	184.8	2	92.4	15.4	0.0002
Error	102	17	6		
Total	286.8	19			



1. 单因素方差分析

例4：现有一学校的计算机科学与技术学院6个专业共718名学生的数学成绩表格，现要分析不同学院的学生的成绩是否有显著性差别。

高等数学期末成绩列表

学号	姓名	开课学期	上课院系	专业名称	课程名称	总成绩	专业编号	考试性质
041740101		2017-2018-1	计算机科学与技术	软件工程	高等数学	95	1	正常考试
041740102		2017-2018-1	计算机科学与技术	软件工程	高等数学	75	1	正常考试
041740103		2017-2018-1	计算机科学与技术	软件工程	高等数学	78	1	正常考试
041740104		2017-2018-1	计算机科学与技术	软件工程	高等数学	66	1	正常考试
041740105		2017-2018-1	计算机科学与技术	软件工程	高等数学	92	1	正常考试
041740106		2017-2018-1	计算机科学与技术	软件工程	高等数学	76	1	正常考试
041740107		2017-2018-1	计算机科学与技术	软件工程	高等数学	80	1	正常考试
041740108		2017-2018-1	计算机科学与技术	软件工程	高等数学	65	1	正常考试
041740109		2017-2018-1	计算机科学与技术	软件工程	高等数学	84	1	正常考试
041740110		2017-2018-1	计算机科学与技术	软件工程	高等数学	79	1	正常考试
041740111		2017-2018-1	计算机科学与技术	软件工程	高等数学	63	1	正常考试
041740112		2017-2018-1	计算机科学与技术	软件工程	高等数学	75	1	正常考试
041740113		2017-2018-1	计算机科学与技术	软件工程	高等数学	70	1	正常考试
041740114		2017-2018-1	计算机科学与技术	软件工程	高等数学	76	1	正常考试
041740115		2017-2018-1	计算机科学与技术	软件工程	高等数学	72	1	正常考试
041740116		2017-2018-1	计算机科学与技术	软件工程	高等数学	73	1	正常考试
041740117		2017-2018-1	计算机科学与技术	软件工程	高等数学	65	1	正常考试
041740118		2017-2018-1	计算机科学与技术	软件工程	高等数学	72	1	正常考试
041740119		2017-2018-1	计算机科学与技术	软件工程	高等数学	83	1	正常考试

第一步：正态性检验

在调用anova1函数做方差分析之前，应先检验样本数据是否满足方差分析的基本假定，即检验正态性和方差齐次性。第一步首先做正态性检验，调用lillietest函数检验6个专业的学生的考试成绩是否服从正态分布，原假设是6个专业的学生的考试成绩服从正态分布，备择假设是不服从正态分布。

```
>> [score,major] = xlsread('computer_math.xlsx',1,'E3:H720');
>> major(:,2) = []; %只保留专业字段
>> major_id = score(:,2); %提取专业编号
%调用lillietest函数分布对6个专业的成绩进行正态性检验
for i=1:6
    scorei = score(major_id == i); %提取第i个专业的成绩数据
    [h,p] = lillietest(scorei(:,1));   %正态性检验
    result(i,:) = p;           %把检验的p值赋给result变量
end
```

```
result =
    0.1812
    0.5000
    0.2928
    0.2417
    0.1107
    0.1250
%均大于0.05，认为服从正态分布。
```

第二步：方差齐次性检验

调用vartestn函数检验6个专业的学生的考试成绩是否服从方差相同的正态分布，原假设是6个专业的学生的成绩服从方差相同的正态分布，备择假设是服从方差不同的正态分布。

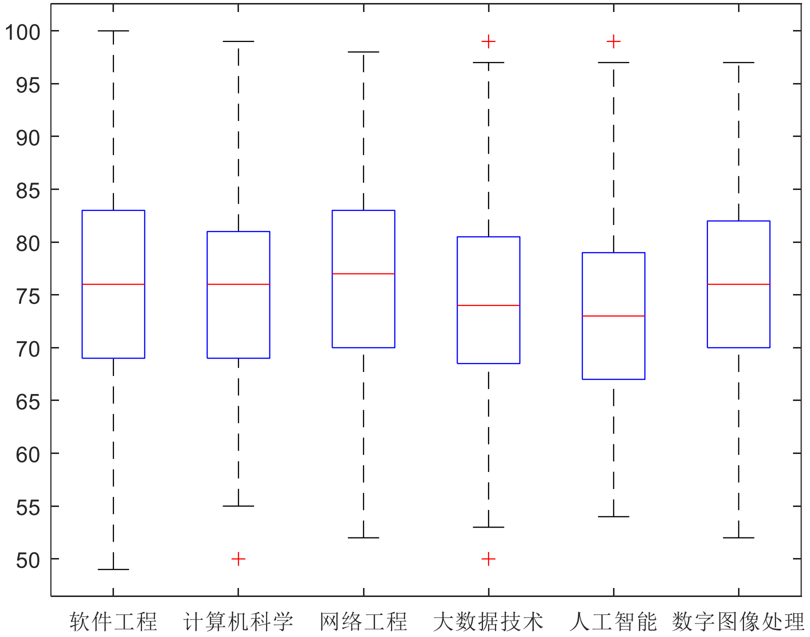
```
>> [p,stats] = vartestn(score(:,1),major)
p = 0.9758
stats = ……
```

检验的p值=0.9758>0.05，说明在显著性水平0.05下接受原假设，认为6个专业的学生的考试成绩服从方差相同的正态分布，满足方差分析的基本假定。

Group Summary Table

Group	Count	Mean	Std Dev
软件工程	118	76.2034	10.0448
计算机科学	120	75.3917	9.4472
网络工程	120	76.5333	9.736
大数据技术	120	74.4583	9.5547
人工智能	120	73.775	9.4633
数字图像处理	120	75.45	9.3484
Pooled	718	75.2994	9.6006

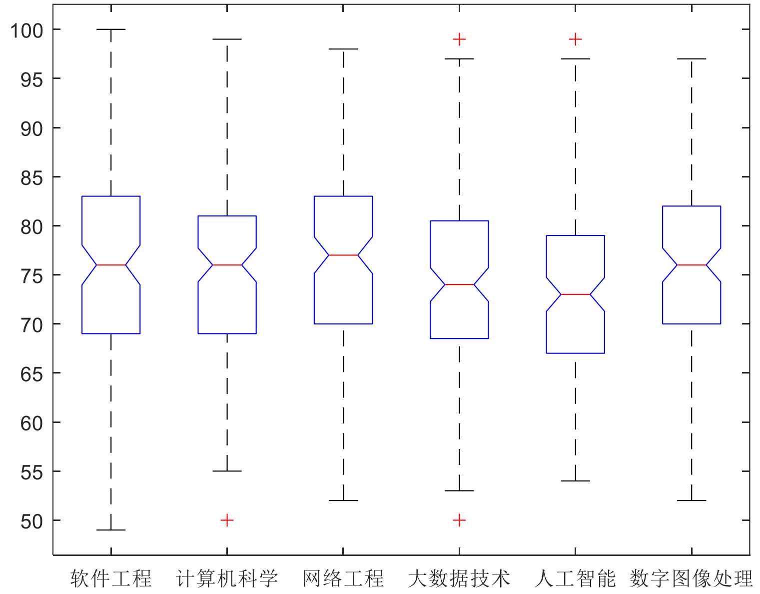
Bartlett's statistic	0.8191
Degrees of freedom	5
p-value	0.9758



第三步：方差分析

- 调用anova1函数进行单因素一元方差分析，检验不同专业的学生的考试成绩有无显著差别，原假设是没有显著差别，备择假设是有显著差别。

```
>> [p,table,stats]=anova1(score(:,1),major)
p = 0.2210
```



ANOVA Table

Source	SS	df	MS	F	Prob>F
Groups	646.6	5	129.325	1.4	0.221
Error	65626	712	92.171		
Total	66272.6	717			

anova1函数返回的p值>0.05，故接受原假设，认为不同专业的学生的考试成绩没有显著的差别，同时还返回的table是一个标准的单因素一元方差分析表，它的各列依次是方差来源、平方和、自由度、均方、F值和p值。还生成了箱线图。

虽然不同专业的学生的成绩没有非常显著的差别，但这并不意味着任意两个专业学生的成绩都没有显著性差别，因此还需要进行两两的比较检验，即多重比较，找出考试成绩存在显著性差别的专业。

第四步：多重比较

数学与统计学院

多重比较函数[c,m,b,gnames]=multcompare(stats)

- 根据结构体变量stats中的信息进行多重比较，返回两两比较的结果矩阵c。c是一个多行6列的矩阵，它的每一行对应一次两两比较的检验，每一行上的元素包括作比较的两个组的组标好，两个组的均值差、均值差的置信区间。
- 返回一个多行2列的矩阵m，第1列为每一组均值的估计值，第2列为相应的标准误差。
- 返回交互式多重比较的图形的句柄值h，可通过h修改图形属性，如图形标题和X轴标签等。
- 组名变量gnames，是一个元胞数组，每一行对应一个组名。

第四步：多重比较

信阳师范学院
数学与统计学院

```
>> [p,table,stats]=anova1(score(:,1),major);
>> [c,m,h,gnames]=multcompare(stats);
```

%设置表头，以元胞数组形式显示矩阵c

```
>> head={'组序号','组序号','置信下限','组均值差','置信上限','置信度上限'};
```

%将矩阵c转为元胞数组，并与head一起显示

```
>> ch = [head;num2cell(c)]
```

%将m转为元胞数组，和gnames一起显示

```
>> mg =[gnames num2cell(m)]
```

```
>> ch = [head;num2cell(c)]
ch =
  16×6 cell 数组
    '组序号'    '组序号'    '置信下限'    '组均值差'    '置信上限'    '置信度上限'
    [     1]    [     2]    [-2.7352]    [ 0.8117]    [ 4.3587]    [    0.9869]
    [     1]    [     3]    [-3.8769]    [-0.3299]    [ 3.2170]    [    0.9998]
    [     1]    [     4]    [-1.8019]    [ 1.7451]    [ 5.2920]    [    0.7259]
    [     1]    [     5]    [-1.1186]    [ 2.4284]    [ 5.9753]    [    0.3709]
    [     1]    [     6]    [-2.7936]    [ 0.7534]    [ 4.3003]    [    0.9907]
    [     2]    [     3]    [-4.6737]    [-1.1417]    [ 2.3903]    [    0.9412]
    [     2]    [     4]    [-2.5987]    [ 0.9333]    [ 4.4653]    [    0.9751]
    [     2]    [     5]    [-1.9153]    [ 1.6167]    [ 5.1487]    [    0.7828]
    [     2]    [     6]    [-3.5903]    [-0.0583]    [ 3.4737]    [    1.0000]
    [     3]    [     4]    [-1.4570]    [ 2.0750]    [ 5.6070]    [    0.5490]
    [     3]    [     5]    [-0.7737]    [ 2.7583]    [ 6.2903]    [    0.2259]
    [     3]    [     6]    [-2.4487]    [ 1.0833]    [ 4.6153]    [    0.9527]
    [     4]    [     5]    [-2.8487]    [ 0.6833]    [ 4.2153]    [    0.9940]
    [     4]    [     6]    [-4.5237]    [-0.9917]    [ 2.5403]    [    0.9676]
    [     5]    [     6]    [-5.2070]    [-1.6750]    [ 1.8570]    [    0.7560]
>> mg =[gnames num2cell(m)]
mg =
  6×3 cell 数组
    '软件工程'        [76.2034]    [0.8838]
    '计算机科学'      [75.3917]    [0.8764]
    '网络工程'        [76.5333]    [0.8764]
    '大数据技术'      [74.4583]    [0.8764]
    '人工智能'        [73.7750]    [0.8764]
    '数字图像处理'    [75.4500]    [0.8764]
```

多重比较交互图

信阳师范学院
数学与统计学院
SCHOOL OF MATHEMATICS AND STATISTICS

- multcompar函数还生成一个交互式图形，可以通过鼠标单击的方式进行两两比较e检验。
- 该交互式图形上用一个符号（圆圈）标出了每一组的组均值，用一条之间段标出了每个组的组均值的置信区间。
- 如果某两条线段不相交，即没有重叠的部分，则说明这两个组的组均值之间的差异是显著的。如果某两条直线段有重叠部分，则说明这两个组的组均值之间的差异是不显著的。
- 用鼠标在图上任意选一个组，选中的组以及与选中的组有显著的其他组均用高亮显示，选中的组用蓝色显示，与选中的组差异显著的其他组用红色显示。

2. 双因素一元方差分析

信阳师范学院
数学与统计学院
SCHOOL OF MATHEMATICS AND STATISTICS

- 双因素方差分析法可以用来分析两个因素的不同水平对结果是否有显著影响，以及两因素之间是否存在交互效应。一般运用双因素方差分析法，先对两个因素的不同水平的组合进行设计试验，要求每个组合下所得到的样本的含量都是相同的。
- 双因素方差分析有两种类型：
  - 一个是无交互作用的双因素方差分析，它假定因素A和因素B的效应之间是相互独立的，不存在相互关系；
  - 另一个是有交互作用的双因素方差分析，它假定因素A和因素B的结合会产生出一种新的效应。

不同品牌的彩电在各地区的销售量数据

品牌(因素A)	销售地区(因素B)				
	B_1	B_2	B_3	B_4	B_5
A_1	365	350	343	340	323
A_2	345	368	363	330	333
A_3	358	323	353	343	308
A_4	288	280	298	260	298

2. 双因素一元方差分析

- 双因素方差分析[p,table,stats] = anova2(X,reps,'displayopt')
  - 说明：执行平衡的双因素试验的方差分析来比较X中两个或多个列（行）的均值，不同列的数据表示因素A的差异，不同行的数据表示另一因素B的差异。如果行列对有多于一个的观察点，**则变量reps指出每一单元观察点的数目**，每一单元包含reps行，如：reps=2；

$$
\begin{array}{cc}
B=1 & B=2
\end{array}
$$
$$
\begin{bmatrix}
y111 & y121 \\
y112 & y122 \\
y211 & y221 \\
y212 & y222 \\
y311 & y321 \\
y312 & y322
\end{bmatrix}
\begin{array}{l}
\left.\begin{array}{l} \\ \\ \end{array}\right\} A=1 \\
\left.\begin{array}{l} \\ \\ \end{array}\right\} A=2 \\
\left.\begin{array}{l} \\ \\ \end{array}\right\} A=3
\end{array}
$$

  - 其余参数与单因素方差分析参数相似。
  - 当 reps=1（默认值）时，anova2将两个p值返回到向量p中。H0A：因素A的所有样本（X中的所有列样本）取自相同的总体；H0B：因素B的所有样本（X中的所有行样本）取自相同的总体。
  - 当reps>1时，anova2还返回第三个p值：H0AB：因素A与因素B没有交互效应。
  - 解释：如果任意一个p值接近于0，则认为相关的零假设不成立。

2. 双因素一元方差分析

例5：有四个品牌的彩电在五个地区销售，为分析彩电的品牌(因素A)和销售地区(因素B)对销售量是否有影响，对每个品牌在各地区的销售量取得以下数据，见下表。试分析品牌和销售地区对彩电的销售量是否有显著影响?

不同品牌的彩电在各地区的销售量数据

品牌(因素A)	销售地区(因素B)				
	B_1	B_2	B_3	B_4	B_5
A_1	365	350	343	340	323
A_2	345	368	363	330	333
A_3	358	323	353	343	308
A_4	288	280	298	260	298

2. 双因素一元方差分析

信阳师范学院
数学与统计学院
SCHOOL OF MATHEMATICS AND STATISTICS

```
>> X = [365 350 343 340 323;345 368 363 330 333;358 323 353 343 308;288 280 298 260 298];
>> [p,table,stats] = anova2(X)
p =
    0.1437    0.0001
```

ANOVA Table

Source	SS	df	MS	F	Prob>F
Columns	2011.7	4	502.93	2.1	0.1437
Rows	13004.5	3	4334.85	18.11	0.0001
Error	2872.7	12	239.39		
Total	17889	19			

销售地区(因素B)无显著性差异 p = 0.1437；

品牌（因素A）有显著性差异 p = 0.0001.

无交互作用。

$$SS_A = b\sum_{i=1}^{a}\left(\overline{X}_{i.} - \overline{X}\right)^2$$ 称为因素A的离差平方和，反映因素 A 对试验指标的影响。

$$SS_B = a\sum_{j=1}^{b}\left(\overline{X}_{.j} - \overline{X}\right)^2$$ 称为因素B的离差平方和，反映因素 B 对试验指标的影响。

$$SS_E = \sum_{i=1}^{a}\sum_{j=1}^{b}\left(X_{ij} - \overline{X}_{i.} - \overline{X}_{.j} + \overline{X}\right)^2$$

称为误差平方和，反映试验误差对试验指标的影响。

$$SS_T = \sum_{i=1}^{a}\sum_{j=1}^{b}\left(X_{ij} - \overline{X}\right)^2$$ 总离差平方和

$$SS_T = SS_A + SS_B + SS_E$$

2. 双因素一元方差分析

方差来源	平方和	自由度	均方和	F 值	F 值临介值
因素A	SS_A	df_A	$MS_A=\dfrac{SS_A}{df_A}$	$F_A=\dfrac{MS_A}{MS_E}$	$F_\alpha((a-1),(a-1)(b-1))$
因素B	SS_B	df_B	$MS_B=\dfrac{SS_B}{df_B}$	$F_B=\dfrac{MS_B}{MS_E}$	$F_\alpha((b-1),(a-1)(b-1))$
误差	SS_E	df_E	$MS_E=\dfrac{SS_E}{df_E}$		
总和	SS_T	df_T			

ANOVA Table

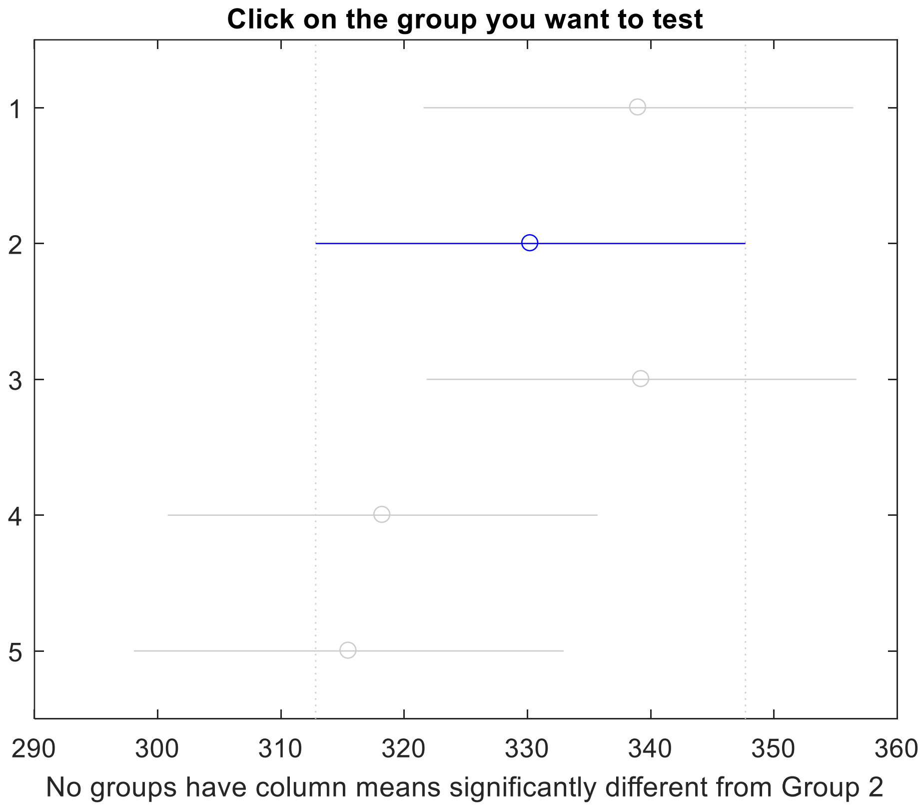
Source	SS	df	MS	F	Prob>F
Columns	2011.7	4	502.93	2.1	0.1437
Rows	13004.5	3	4334.85	18.11	0.0001
Error	2872.7	12	239.39		
Total	17889	19			

2. 双因素一元方差分析


```
>> [c,m,h,gnames] = multcompare(stats,'estimate','column')
```

```
c =
    1.0000    2.0000   -26.1223     8.7500    43.6223    0.9257
    1.0000    3.0000   -35.1223    -0.2500    34.6223    1.0000
    1.0000    4.0000   -14.1223    20.7500    55.6223    0.3691
    1.0000    5.0000   -11.3723    23.5000    58.3723    0.2624
    2.0000    3.0000   -43.8723    -9.0000    25.8723    0.9186
    2.0000    4.0000   -22.8723    12.0000    46.8723    0.8050
    2.0000    5.0000   -20.1223    14.7500    49.6223    0.6688
    3.0000    4.0000   -13.8723    21.0000    55.8723    0.3584
    3.0000    5.0000   -11.1223    23.7500    58.6223    0.2540
    4.0000    5.0000   -32.1223     2.7500    37.6223    0.9990
m =
  339.0000    7.7361
  330.2500    7.7361
  339.2500    7.7361
  318.2500    7.7361
  315.5000    7.7361
```

从多重分析中可以看出，销售地区（因素B）并无显著性差异。



2. 双因素一元方差分析

```
>> [c,m,h,gnames] = multcompare(stats,'estimate','row')
```

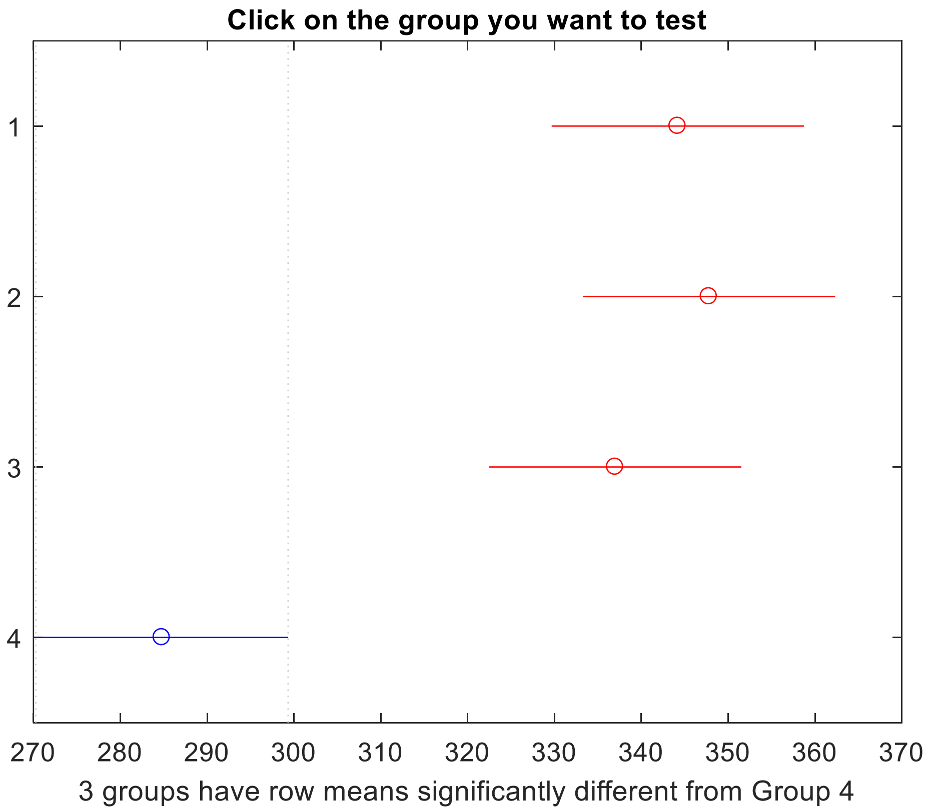
c =

1.0000	2.0000	-32.6523	-3.6000	25.4523	0.9822
1.0000	3.0000	-21.8523	7.2000	36.2523	0.8809
1.0000	4.0000	30.3477	59.4000	88.4523	0.0003
2.0000	3.0000	-18.2523	10.8000	39.8523	0.6941
2.0000	4.0000	33.9477	63.0000	92.0523	0.0002
3.0000	4.0000	23.1477	52.2000	81.2523	0.0009

m =

344.2000	6.9194
347.8000	6.9194
337.0000	6.9194
284.8000	6.9194

从多重分析中可以看出，品牌4与其他三种品牌存在显著性差异，但品牌1、2和3并无显著性差异。



2. 双因素一元方差分析

-1975-
信阳师范学院
数学与统计学院
SCHOOL OF MATHEMATICS AND STATISTICS

例6：有3个工人分别在4台机器上加工某种零件，工作的3天中日产量列表如下：

机器A \ 工人B	B1	B2	B3
A1	15 15 17	19 19 16	16 18 21
A2	17 17 17	15 15 15	19 22 22
A3	15 17 16	18 17 16	18 18 18
A4	18 20 22	15 16 17	17 17 17

试考察工人与机器对日产量的影响是否显著？检验操作工人的技术水平有无显著差异？机器性能有无显著差异？交互作用的影响是否显著？($\alpha=0.05$)

2. 双因素一元方差分析

%列对应因素A机器，行对应因素B工人

```
>> X = [15,15,17,19,19,16,16,18,21;17,17,17,15,15,15,19,22,22;
   15,17,16,18,17,16,18,18,18;18,20,22,15,16,17,17,17,17]';
>> reps=3;
>> [p,tab,stats]=anova2(X,reps,'on')
p =
   0.6645    0.0023    0.0002
stats = 
  包含以下字段的 struct:

      source: 'anova2'
     sigmasq: 1.7222
    colmeans: [17.3333 17.6667 17 17.6667]
        coln: 9
    rowmeans: [17.1667 16.5000 18.5833]
        rown: 12
       inter: 1
        pval: 1.9217e-04
          df: 24
```

ANOVA Table

Source	SS	df	MS	F	Prob>F
Columns	2.75	3	0.9167	0.53	0.6645
Rows	27.167	2	13.5833	7.89	0.0023
Interaction	73.5	6	12.25	7.11	0.0002
Error	41.333	24	1.7222		
Total	144.75	35			

结果解释:

- ✓ 机器性能无显著差异 p = 0.6645
- ✓ 操作工人的技术水平有显著差异 p = 0.0023
- ✓ 交互作用的影响显著 p = 0.0002

此时，交互作用显著，采用多重交互分析主效应可能存在问题，可采用多因素方差分析进行多重比较，分析出哪种机器哪个工人状态下，日产量达到最大。

2. 双因素一元方差分析


```
[c,m,h,gnames] = multcompare(stats,'estimate','column')
```

Note: Your model includes an interaction term that is significant at the level you specified. Testing main effects under these conditions is questionable.

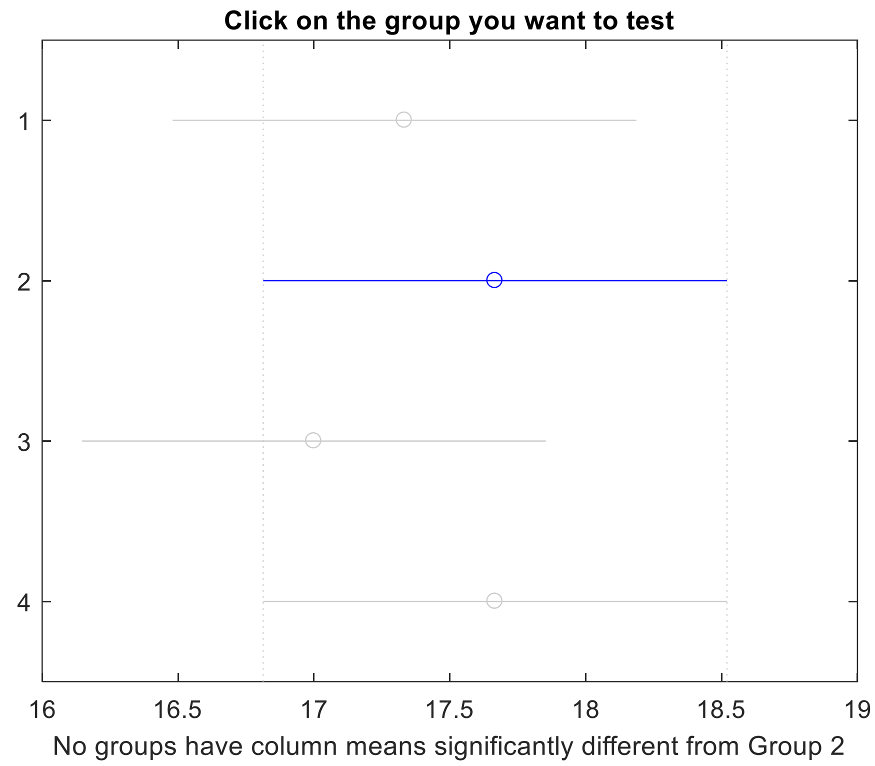
c =

1.0000	2.0000	-2.0399	-0.3333	1.3733	0.9486
1.0000	3.0000	-1.3733	0.3333	2.0399	0.9486
1.0000	4.0000	-2.0399	-0.3333	1.3733	0.9486
2.0000	3.0000	-1.0399	0.6667	2.3733	0.7061
2.0000	4.0000	-1.7066	0	1.7066	1.0000
3.0000	4.0000	-2.3733	-0.6667	1.0399	0.7061

m =

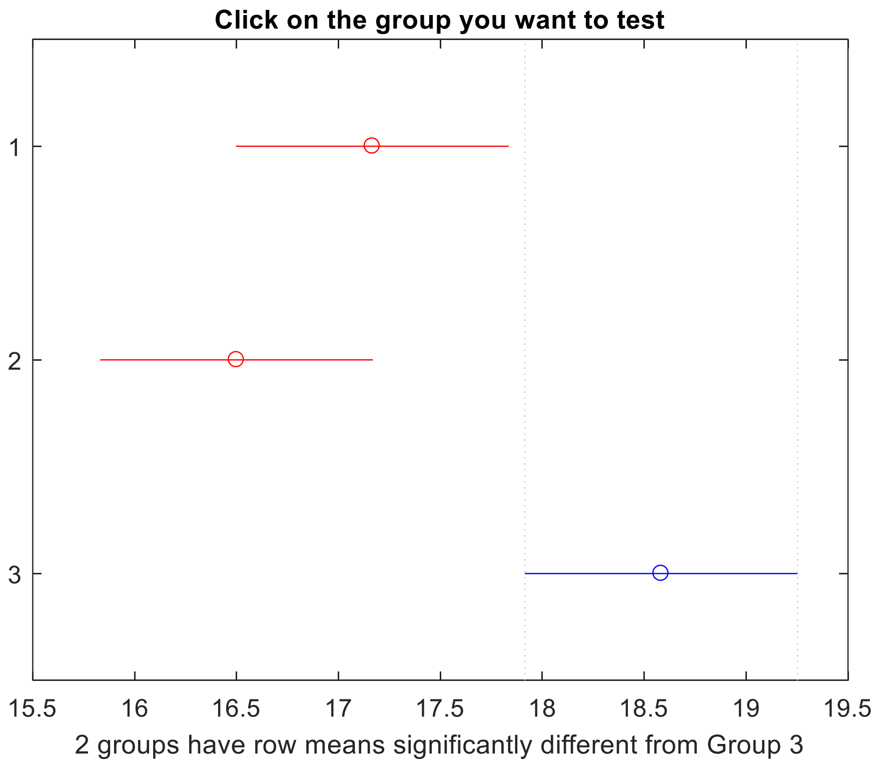
17.3333	0.4374
17.6667	0.4374
17.0000	0.4374
17.6667	0.4374

机器之间无显著差异；可从c中看出置信区间包含了0值，且p值均显著大于0.05，从交互图上看出机器均值存在重叠。



2. 双因素一元方差分析

```
>> [c,m,h,gnames] = multcompare(stats,'estimate','row')
Note: Your model includes an interaction term that is significant at the level
you specified. Testing main effects under these conditions is questionable.
c =
    1.0000    2.0000   -0.6713    0.6667    2.0046    0.4395
    1.0000    3.0000   -2.7546   -1.4167   -0.0787    0.0365
    2.0000    3.0000   -3.4213   -2.0833   -0.7454    0.0019
m =
   17.1667    0.3788
   16.5000    0.3788
   18.5833    0.3788
```



工人之间有显著差异，且B3与B1、B2均有显著性差异。

2. 双因素一元方差分析

例7：为了研究肥料使用量对水稻产量的影响，某研究所做了氮（因素A）、磷（因素B）两种肥料施用量的二因素试验。氮肥用量设低、中、高三个水平，分布使用N1、N2和N3表示；磷肥用量设低、高2个水平，分别用P1、P2表示。供$3 \times 2 = 6$个处理，重复4次，随机区组设计，测得水稻产量如表所示。根据表中数据，不考虑区组因素，分析氮、磷两种肥料的施用量对水稻产量是否有显著性影响，并分析交互作用是否显著，$\alpha = 0.05$。

	区组1	区组2	区组3	区组4
N1P1	38	29	36	40
N1P2	45	42	37	43
N2P1	58	46	52	51
N2P2	67	70	65	71
N3P1	62	64	61	70
N3P2	58	63	71	69

2. 双因素一元方差分析

信阳师范学院
数学与统计学院

- 这里不需要进行正态性和方差性齐次性检验，因素数据少，在数据比较少的情况下正态检验的结果是不可靠的，即使不满足方差分析的假定，方差分析的结果通常也是比较稳定的。
- 对原数据表进行转置，然后提取对应元素，使得每一列对应一个A因素（氮）水平，每一行对应一个B因素（磷）水平；

N1P1	N1P2	N2P1	N2P2	N3P1	N3P2
38	45	58	67	62	58
29	42	46	70	64	63
36	37	52	65	61	71
40	43	51	71	70	69

	N1	N2	N3
P1	38	58	62
P1	29	46	64
P1	36	52	61
P1	40	51	70
P2	45	67	58
P2	42	70	63
P2	37	65	71
P2	43	71	69

2. 双因素一元方差分析

%定义一个矩阵，输入原始数据

```
>> yield=[38 29 36 40; 45 42 37 43; 58 46 52 51
   67 70 65 71; 62 64 61 70; 58 63 71 69];
>> yield=yield'; %矩阵转置
```

%将数据矩阵yield转换成8行3列的矩阵，列对应因素A（氮），行对应因素B（磷）

```
>> yield=[yield(:,[1,3,5]);yield(:,[2,4,6])];
```

%定义元胞数组，以元胞数组形式显示转换后的数据

```
>> top={'因素','N1','N2','N3'};
>> left={'P1';'P1';'P1';'P1';'P2';'P2';'P2';'P2'};
>>[top;left,num2cell(yield)] %显示数据
```

'因素'	'N1'	'N2'	'N3'
'P1'	[38]	[58]	[62]
'P1'	[29]	[46]	[64]
'P1'	[36]	[52]	[61]
'P1'	[40]	[51]	[70]
'P2'	[45]	[67]	[58]
'P2'	[42]	[70]	[63]
'P2'	[37]	[65]	[71]
'P2'	[43]	[71]	[69]

2. 双因素一元方差分析

信阳师范学院
数学与统计学院
SCHOOL OF MATHEMATICS AND STATISTICS

%调用anova2函数作双因素方差分析，返回检验的p值向量，方差分析表，结构体标量stats

```
>>reps = 4;
>>[p,table,stats]=anova2(yield,reps)
p =
    0.0000    0.0004    0.0080
table =   %见图方差分析表数据
stats =
      source: 'anova2'
     sigmasq: 19.5833
    colmeans: [38.7500 60 64.7500]
        coln: 8
    rowmeans: [50.5833 58.4167]
        rown: 12
       inter: 1
        pval: 0.0080
          df: 18
```

ANOVA Table

Source	SS	df	MS	F	Prob>F
Columns	3067	2	1533.5	78.31	0
Rows	368.167	1	368.17	18.8	0.0004
Interaction	250.333	2	125.17	6.39	0.008
Error	352.5	18	19.58		
Total	4038	23			

因素A、因素B以及他们的交互作用对应的检验p值均小于给定的显著性水平0.05，所以可以认为氮、磷两种肥料的施用量对水稻的产量均有显著性影响，并且他们之间的交互作用也是非常显著的。

由于氮、磷两种肥料的用量对水稻的产量均有非常显著的影响，可以作进一步分析，例如进行多重分析，找出因素A、B在哪种水平的组合下水稻的平均产量最高。

2. 双因素一元方差分析

信阳师范学院 数学与统计学院 SCHOOL OF MATHEMATICS AND STATISTICS

%对列（因素A）进行多重比较

```
>>[c_A,m_A]=multcompare(stats,'estimate','column')
c_A =
    1.0000    2.0000  -26.8971  -21.2500  -15.6029    0.0000
    1.0000    3.0000  -31.6471  -26.0000  -20.3529    0.0000
    2.0000    3.0000  -10.3971   -4.7500    0.8971    0.1084
m_A =
   38.7500    1.5646
   60.0000    1.5646
   64.7500    1.5646
```

%对行（因素B）进行多重比较

```
>>[c_B,m_B]=multcompare(stats,'estimate','row')
c_B =
    1.0000    2.0000  -11.6289   -7.8333   -4.0378    0.0004
m_B =
   50.5833    1.2775
   58.4167    1.2775
```

由上面结果可以看出，若单独考虑A因素，它的第1个水平与后两个水平差异显著，它的第2个水平与第3个水平差异不显著，并且当A取第3个水平（N3）时，水稻产量的均值达到最大64.75；

如果单独考虑B因素，它的两个水平差异显著；

在因素A，B的水平组合N3P2下，水稻的平均产量达到最大值，然而这却是错误的。因为A,B之间存在着非常显著的交互作用，在这种情况下对主效应进行检验可能存在问题，这时应该对因素A、B的每种水平组合进行多重比较，找出所要的水平组合。

3. 多因素一元方差分析

anovan函数用来根据样本观测值向量进行均衡或非均衡试验的多因素一元方差分析，检验多个（N个）因素的主效应或交互式效应是否显著。调用格式如下：

- [p,table,stats,terms]=anovan(y, group)：根据样本观测值向量y进行均衡或非均衡试验的多因素一元方差分析，检验多个因素的主效应是否显著。输入参数group是一个元胞数组，它的每一个元胞对应一个因素，是该因素的水平列表，与y等长，用来标记y中的每个观测所对应的因素的水平。anovan函数还生成1个图形，用来显示一个标准的多因素一元方差分析表。
- [p,table,stats,terms]=anovan(y, group)：
  - ✓ 返回元胞数组形式的方差分析表table。返回一个结构体变量stats，用于进行后续的多重比较。当某因素对实验指标的影响显著时，在后续的分析中，可用调用multcompare函数，把stats作为其输入，进行多重比较。返回方差分析计算中的主效应和交互效应矩阵terms。

3. 多因素一元方差分析

p=anovan(y,group,param1,val1,param2,val2,......)：可用的参数名与参数值如下表

参数名	参数值	说明
‘alpha’	（0,1）内的标量	指定置信水平
‘continuous’	下标向量	用来指明哪些分组变量被作为连续变量，而不是离散的分类变量
‘display’	‘on’或‘off’	用来指定是否显示方差分析表
‘model’	‘linear’、‘interaction’、‘full’	用来指定所用模型的类型，‘linear’（默认）只对N个主效应进行检验，不考虑交互效应；‘interaction’对主效应和两个因素的交互效应进行检验；‘full’对N个主效应和全部的交互效应进行检验。也可以通过0和1的矩阵自定义效应项
‘nested’	由0和1构成的矩阵M	指定分组变量之间的嵌套关系
‘random’	下标向量	用来指明哪些分组变量是随机的
‘sstype’	1,2或3	指定平方和的类型，默认值为3
‘varnames’	字符矩阵或字符串元胞数组	指定分组变量的名称。当没有指定时，默认用‘X1’，‘X2’，...‘XN’做为名称

3. 多因素一元方差分析

例8：仍用上节中的数据，研究肥料使用量对水稻产量的影响：

处理	区组			
	1	2	3	4
N1P1	38	29	36	40
N1P2	45	42	37	43
N2P1	58	46	52	51
N2P2	67	70	65	71
N3P1	62	64	61	70
N3P2	58	63	71	69

调用anovan函数进行分析，在不考虑区组因素的情况下，分析氮、磷两种肥料的施用量对水稻的产量是否有显著影响，并分析交互作用是否显著，然后找出在因素A，B的哪种水平组合下水稻的平均产量高，显著性水平为0.05.

3. 多因素一元方差分析

信阳师范学院
数学与统计学院
SCHOOL OF MATHEMATICS AND STATISTICS

```
yield=[38 29 36 40; 45 42 37 43; 58 46 52 51; 67 70 65 71; 62 64 61 70; 58 63 71 69];
yield=yield';  %矩阵转置
yield=yield(:);  %将数据矩阵yield按列拉长成24行1列的向量
%定义因素A(氮)的水平列表向量
A=strcat({'N'},num2str([ones(8,1);2*ones(8,1);3*ones(8,1)]));
%定义因素B(磷)的水平列表向量
B=strcat({'P'},num2str([ones(4,1);2*ones(4,1)]));
B=[B;B;B];
%将因素A、B的水平列表向量与yield向量放在一起构成一个元胞数组
[A,B,num2cell(yield)]
varnames={'A','B'}; %指定因素名称，A表示氮肥施用量，B表示磷肥施用量
```

```
[A,B,num2cell(yield)]
ans =
  24×3 cell 数组
    'N1'    'P1'    [38]
    'N1'    'P1'    [29]
    'N1'    'P1'    [36]
    'N1'    'P1'    [40]
    'N1'    'P2'    [45]
    'N1'    'P2'    [42]
    'N1'    'P2'    [37]
    'N1'    'P2'    [43]
    'N2'    'P1'    [58]
    'N2'    'P1'    [46]
    'N2'    'P1'    [52]
    'N2'    'P1'    [51]
    'N2'    'P2'    [67]
    'N2'    'P2'    [70]
    'N2'    'P2'    [65]
    'N2'    'P2'    [71]
    'N3'    'P1'    [62]
    'N3'    'P1'    [64]
    'N3'    'P1'    [61]
    'N3'    'P1'    [70]
    'N3'    'P2'    [58]
    'N3'    'P2'    [63]
    'N3'    'P2'    [71]
    'N3'    'P2'    [69]
```

3. 多因素一元方差分析

varnames={'A','B'}; %指定因素名称，A表示氮肥施用量，B表示磷肥施用量

[p,table,stats,term]=anovan(yield,{A,B},'model','full','varnames',varnames)

p = 0.0000 0.0004 0.0080

term =

1 0

0 1

1 1

Analysis of Variance

Source	Sum Sq.	d.f.	Mean Sq.	F	Prob>F
A	3067	2	1533.5	78.31	0
B	368.167	1	368.17	18.8	0.0004
A*B	250.333	2	125.17	6.39	0.008
Error	352.5	18	19.58		
Total	4038	23			

Constrained (Type III) sums of squares.

返回的向量p是主效应A，B和交互效应AB所对应的检验的p值，table是元胞数组的方差分析表，stats是一个结构体变量，可用于后续的分析中（如多重比较），矩阵term的3行分布表示了3个效应项：主效应项A，主效应项B和交互式效应项AB，还生成了一个方差分析表。

返回的结果与调用anova2函数得到的结果一样，因素A，因素B以及他们的交互作用所对应的p值均小于给定的显著想水平0.05，所以可以认为氮、磷两种肥料的施用量对水稻的产量均有非常显著的影响，并且他们之间的交互作用也是非常显著的。

3. 多因素一元方差分析

信阳师范学院
数学与统计学院

调用multcompare函数进行多重比较，对因素A、B的每种水平的组合进行多重比较，找出在因素A、B的哪种水平组合下水稻的平均产量最高。

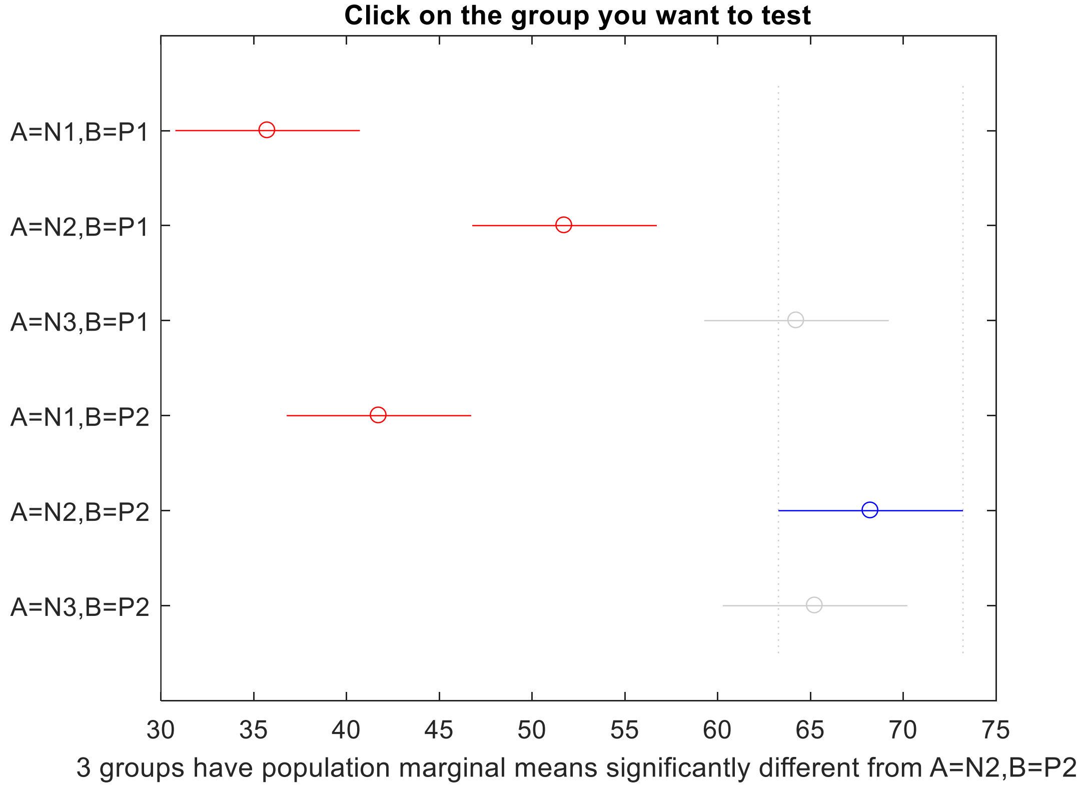
```
[c,m,h,gnames]=multcompare(stats,'dimension',[1 2])
```

c =

1.0000	2.0000	-25.9446	-16.0000	-6.0554	0.0009
1.0000	3.0000	-38.4446	-28.5000	-18.5554	0.0000
1.0000	4.0000	-15.9446	-6.0000	3.9446	0.4236
1.0000	5.0000	-42.4446	-32.5000	-22.5554	0.0000
1.0000	6.0000	-39.4446	-29.5000	-19.5554	0.0000
2.0000	3.0000	-22.4446	-12.5000	-2.5554	0.0093
2.0000	4.0000	0.0554	10.0000	19.9446	0.0483
2.0000	5.0000	-26.4446	-16.5000	-6.5554	0.0006
2.0000	6.0000	-23.4446	-13.5000	-3.5554	0.0047
3.0000	4.0000	12.5554	22.5000	32.4446	0.0000
3.0000	5.0000	-13.9446	-4.0000	5.9446	0.7926
3.0000	6.0000	-10.9446	-1.0000	8.9446	0.9995
4.0000	5.0000	-36.4446	-26.5000	-16.5554	0.0000
4.0000	6.0000	-33.4446	-23.5000	-13.5554	0.0000
5.0000	6.0000	-6.9446	3.0000	12.9446	0.9251

m =

35.7500	2.2127
51.7500	2.2127
64.2500	2.2127
41.7500	2.2127
68.2500	2.2127
65.2500	2.2127



m矩阵列出了6种处理的平均值，很明显第5个处理（即N2P2）的平均值最大，由矩阵c或交互式多重比较的图中可以得到，处理5与处理3,6差异不显著，所以可以认为第3个和第6个处理也是可以的，所以，综上，可以在处理3,5,6中做出选择，即N3P1，N2P2，N3P2。

3. 多因素一元方差分析

例9（续）：有3个工人分别在4台机器上加工某种零件

机器 \ 工人	B1	B2	B3
A1	15 15 17	19 19 16	16 18 21
A2	17 17 17	15 15 15	19 22 22
A3	15 17 16	18 17 16	18 18 18
A4	18 20 22	15 16 17	17 17 17

```
>> X = [15,15,17,19,19,16,16,18,21;17,17,17,15,15,15,19,22,22;
   15,17,16,18,17,16,18,18,18;18,20,22,15,16,17,17,17,17]';
>> XX = X(:);   %转换为一列
>> A = strcat({'A'},num2str([ones(9,1);2*ones(9,1);3*ones(9,1);4*ones(9,1)]));
>> B = strcat({'B'},num2str([ones(3,1);2*ones(3,1);3*ones(3,1)]));  %B交替出现
>> B = [B;B;B;B];
>> [A,B,num2cell(XX)]
```

```
ans =
  36×3 cell 数组
    'A1'   'B1'   [15]
    'A1'   'B1'   [15]
    'A1'   'B1'   [17]
    'A1'   'B2'   [19]
    'A1'   'B2'   [19]
    'A1'   'B2'   [16]
    'A1'   'B3'   [16]
    ……
```

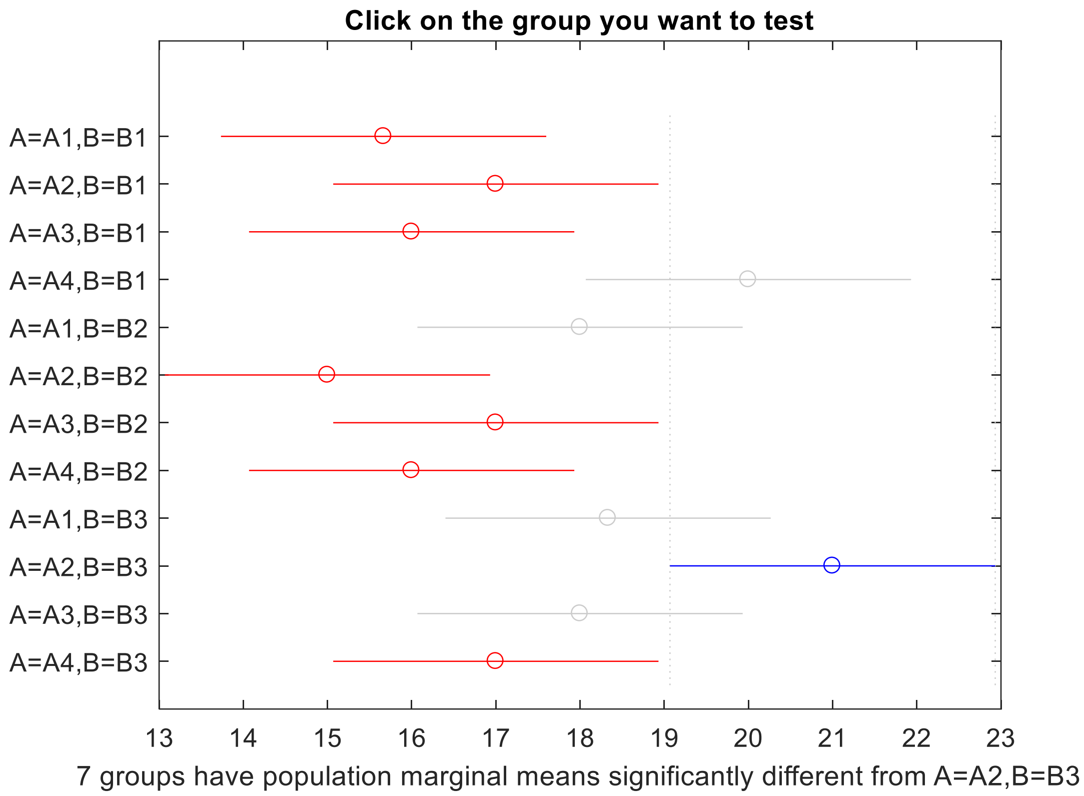
3. 多因素一元方差分析

```
>> varnames={'A','B'};
>> [p,table,stats,term]=anovan(XX,{A,B},'model','full','varnames',varnames)
>> [c,m,h,gnames]=multcompare(stats,'dimension',[1 2])
```

Analysis of Variance

Source	Sum Sq.	d.f.	Mean Sq.	F	Prob>F
A	2.75	3	0.9167	0.53	0.6645
B	27.167	2	13.5833	7.89	0.0023
A*B	73.5	6	12.25	7.11	0.0002
Error	41.333	24	1.7222		
Total	144.75	35			

Constrained (Type III) sums of squares.



方差分析认为机器之间无显著差异，工人之间有显著差异，交互作用显著。

多重分析认为可选择A2B3，均值最大，与其无显著差异的还有A4B1、A1B2、A1B3、A3B3，可在其中作出选择。

4. 单因素多元方差分析

manoval函数用来做单因素多元方差分析，检验多个多元正态总体是否具有相同的均值向量。调用格式如下：[d,p,stats]=manova1(X,group,alpha)

- 样本观测值矩阵X，比较X中的各组观测是否具有相同的均值向量，原假设是各组的组均值是相同的多元向量。
- X是一个$m \times n$的矩阵，它的每一列对应一个变量，每一行对应一个观测，每一个观测都是n元的。
- 输入参数group是一个分组变量，用来标示X中的每个观测所在的组，group可以是一个分类变量、向量、字符串数组或字符串元胞数组，group的长度应该与X的行数相等，group中相同元素对应的X中的观测来自同一个总体（组）的样本。

4. 单因素多元方差分析

信阳师范学院 数学与统计学院 SCHOOL OF MATHEMATICS AND STATISTICS

[d,p,stats]=manova1(X,group,alpha)

- alpha指定检验的显著性水平。返回的d是满足p>alpha的最小维数，其中p是检验的p值，此时检验各组的均值向量是否位于一个d维空间。
- 各组的均值向量生成了一个向量空间，输出参数d是这个空间维数的估计，当d=0时，接受原假设，当d=1时在显著性水平0.05下，拒绝原假设，认为各组的组均值不全相同，但是不能拒绝他们贡献的假设；当d=2时，拒绝原假设，此时各组的组均值可能共面，但是不共线。
- 返回检验的p值向量，它的第i个元素对应的原假设是：各组的均值向量位于一个i-1维空间，若p的第i个元素小于或等于给定的显著性水平，则拒绝这样的原假设。还返回一个结构体变量stats。

4. 单因素多元方差分析

例9：为了研究销售方式对商品的销售额的影响，选择四种商品（甲，乙，丙和丁，分别记为x1，x2，x3，x4）按三种不同的销售方式（1,2,3）进行销售。其数据如下表，根据数据分析不同销售方式对销售额是否有显著影响？$\alpha=0.05$.

编号	销售方式 I				销售方式 II				销售方式 III			
	x1	x2	x3	x4	x1	x2	x3	x4	x1	x2	x3	x4
1	125	60	338	210	66	54	455	310	65	33	480	260
2	119	80	233	330	82	45	403	210	100	34	468	295
3	63	51	260	203	65	65	312	280	65	63	416	265
4	65	51	429	150	40	51	477	280	117	48	468	250
5	130	65	103	205	67	54	481	293	114	63	395	380
6	69	45	350	190	38	50	468	210	55	30	546	235
7	46	60	585	200	42	45	351	190	64	51	507	320
8	146	66	273	250	113	40	390	310	110	90	442	225
9	87	54	585	240	80	55	520	200	60	62	440	248
10	110	77	507	270	76	60	507	189	110	69	377	260
11	107	60	364	200	94	33	260	280	88	78	299	360
12	130	61	391	200	60	51	429	190	73	63	390	320
13	80	45	429	270	55	40	390	295	114	55	494	240
14	60	50	442	190	65	48	481	177	103	54	416	310
15	81	54	260	280	69	48	442	225	100	33	273	312
16	135	87	507	260	125	63	312	270	140	61	312	345
17	57	48	400	285	120	56	416	280	80	36	286	250
18	75	52	520	260	70	45	468	370	135	54	468	345
19	76	65	403	250	62	66	416	224	130	69	325	360
20	55	42	411	170	69	60	377	280	60	57	273	260

4. 单因素多元方差分析

```
>> sale = xlsread('sale.xlsx',1,'B3:M22');
>> X = [sale(:,1:4);sale(:,5:8);sale(:,9:12)];
>> group=[ones(20,1);2*ones(20,1);3*ones(20,1)]; %生成销售方式数据
>> [d,p,stats]=manova1(X,group) %调用manova1函数作多元方差分析
d =
     1
p =
    0.0035
    0.0917
stats =
```

从manova1函数返回回的结果来看，p值分别为0.0035<0.05和0.0917>0.05，说明在显著性水平0.05下拒绝假设：3种销售方式所对应的销量的均值向量不全相同；接受假设：3种销售方式多对应的销售量的均值向量位于一个1维空间（即共线），因此维数的估计值为d=1。

总的来说，在显著性水平0.05下，可认为不同销售方式对销售额有显著影响，但是究竟对四种商品中的哪种商品的销售额影响最显著，还需要对四种商品的销售额分别做一元方差分析。

4. 单因素多元方差分析


```
%调用anova1函数对甲商品的销售额作一元方差分析
[p1,table]=anova1(X(:,1),group);
%调用anova1函数对乙商品的销售额作一元方差分析
[p2,table2]=anova1(X(:,2),group);
%调用anova1函数对丙商品的销售额作一元方差分析
[p3,table3]=anova1(X(:,3),group);
%调用anova1函数对丁商品的销售额作一元方差分析
[p4,table4]=anova1(X(:,4),group);
>> p = [p1,p2,p3,p4]
p =
    0.0411    0.2078    0.6443    0.0009
```

对四种商品的销售额分别做了一元方差分析，得到的检验p值分别为：

$$p1<0.05,\ p2>0.05,\ p3>0.05,\ p4<0.05$$

所以在显著性水平0.05下，可认为不同销售方式对甲商品的销售额有显著影响，对丁商品的销售额有十分显著的影响，对乙和丙商品的销售额没有显著影响。

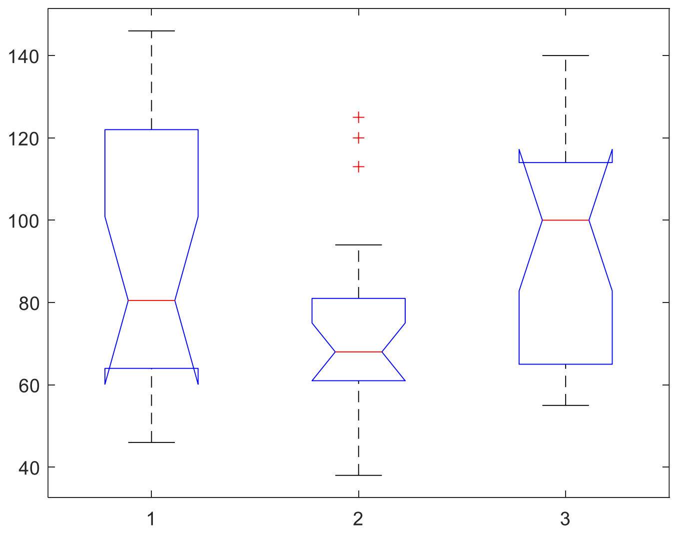
4. 单因素多元方差分析

信阳师范学院
数学与统计学院
SCHOOL OF MATHEMATICS AND STATISTICS

- 甲产品销售三种销售方式

ANOVA Table

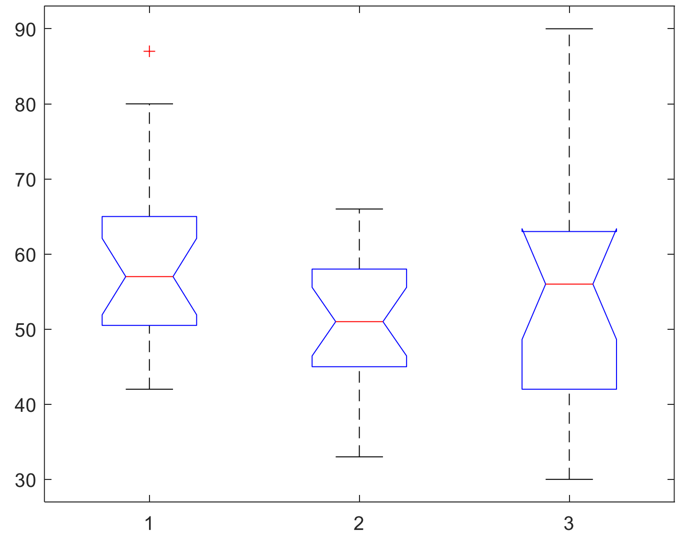
Source	SS	df	MS	F	Prob>F
Groups	5221.3	2	2610.65	3.38	0.0411
Error	44069.6	57	773.15		
Total	49290.9	59			



- 乙产品销售三种销售方式

ANOVA Table

Source	SS	df	MS	F	Prob>F
Groups	518.53	2	259.267	1.62	0.2078
Error	9148.05	57	160.492		
Total	9666.58	59			

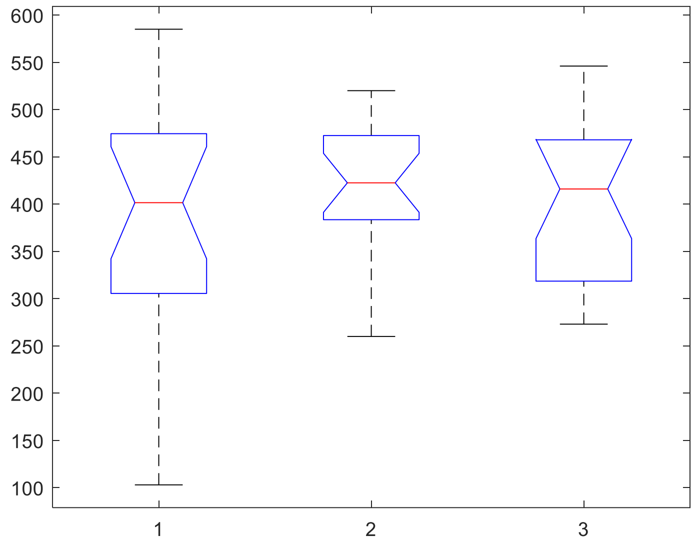


4. 单因素多元方差分析

- 丙产品销售三种销售方式

ANOVA Table

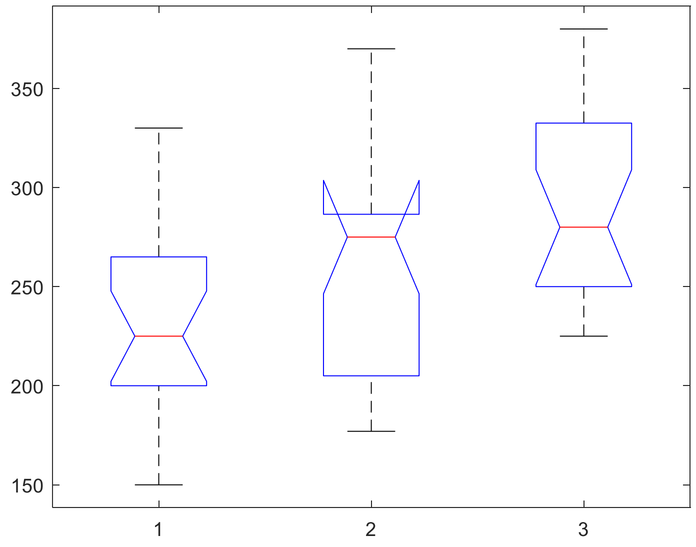
Source	SS	df	MS	F	Prob>F
Groups	7980.8	2	3990.42	0.44	0.6443
Error	513428.5	57	9007.52		
Total	521409.3	59			



- 丁产品销售三种销售方式

ANOVA Table

Source	SS	df	MS	F	Prob>F
Groups	38529.3	2	19264.7	8.01	0.0009
Error	137115.1	57	2405.5		
Total	175644.4	59			




01
IDEA
02
LEARNING
03
CREATIVE
INSPIRATION
EDUCATION
BUSINESS
PHOTO BOOK

感谢聆听