

(多组) 计量资料的分析与比较

本模块可以对一组计量资料进行统计分析，也可以对多组计量资料进行比较。涉及到的统计方法包括：

- 正态性检验
- 方差齐性检验
- t 检验
- 方差分析
- 多组均数的两两比较
- 秩和检验
- 非参数方差分析
- 多元回归分析
- 交互作用检验。

例 1：（单纯随机抽样设计）A、B、C 三组大鼠肝脏 ATP 测量结果

A 组	B 组	C 组
7.76	11.14	10.85
7.71	11.6	8.58
8.43	11.42	7.19
8.47	13.85	9.36
10.3	13.53	9.59
6.67	14.16	8.81
11.73	6.94	8.22
5.78	13.01	9.95
6.61	14.18	11.26
6.97	17.72	8.68
7.01	12.5	NA
12.95	NA	NA

NA 表示缺失

首先把数据放到 EXCEL 里，对缺失的格子用 NA 补上。然后从 EXCEL 里拷贝粘贴到下图数据框内（要求数据字段间用制表符分隔，如果是从 EXCEL 复制过来则自动为制表符分隔）。

例 1 数据输入界面

计量资料的比较与分析

分析项目标题: 多组计量数据比较

输出文件名: anovatest

复制粘贴m列数据(制表符分隔)到下框 (第一行为列标题或称变量名, 第二行开始为数据)
如有分层变量, 置前面 (前面 列为分层因素); (后面)每列代表一组(参照组置第一列)。

```
A组      B组      C组
7.76     11.14    10.85
7.71     11.6     8.58
8.43     11.42    7.19
8.47     13.85    9.36
10.3     13.53    9.59
6.67     14.16    8.81
11.73    6.94     8.22
5.78     13.01    9.95
6.61     14.18    11.26
6.97     17.72    8.68
7.01     12.5     NA
12.95    NA       NA
```

- 正态性检验 方差齐性检验 秩和检验及两两比较
 t检验或方差分析 随机区组设计, 双向方差分析 两两(列分组间)均数比较
 两两(行区组间)均数比较 分层因素与列分组交互作用检验

结果存到 (默认为我的文档):

开始分析 >>>

- 选正态性检验、方差齐性检验、秩和检验及两两比较、t 检验或方差分析、两两（列分组）均数比较。
- 该数据没有分层因素（如有放到第一列），所以没有选分层因素与列分组交互作用检验。

输出结果与解释

多组计量数据比较

数据

Row. groups	A 组	B 组	C 组
1	7.76	11.14	10.85
2	7.71	11.6	8.58
3	8.43	11.42	7.19
4	8.47	13.85	9.36
5	10.3	13.53	9.59
6	6.67	14.16	8.81
7	11.73	6.94	8.22
8	5.78	13.01	9.95
9	6.61	14.18	11.26

这里首先列出数据，以便你检查数据读入是否正确。

10	6.97	17.72	8.68
11	7.01	12.5	NA
12	12.95	NA	NA

本例数据符合正态与方差齐性，可以跳过下面的非参数方差分析与非参数的两两比较，而直接看后面的方差分析、回归分析、HSD法两两比较结果。

	统计量(w)	p 值	检验方法	备注
正态性检验	0.937664	0.05676729	Shapiro-Francia	残差分析

	统计量	p 值	检验方法	备注
方差齐性检验	0.885164	0.423149	Brown-Forsythe Levene	Modified robust

	统计量(w)	自由度	p 值	检验方法
非参数方差分析	14.08563	2	0.0008736653	Kruskal-Wallis rank sum test

非参数方差分析后两两比较(multiple test after Kruskal-Wallis)

组-组	obs. dif	critical. dif	difference
1-2	14.8787878787879	9.66280553470837	1
1-3	4.53333333333333	9.91167388442734	0
2-3	10.3454545454545	10.1143875831714	1

这是非参数的两两比较结果：
difference=0，表示无显著差异
difference=1，表示有显著差异

critical. dif: p-value=0.05

基本统计

列分组	N	均数	标准差	最小值	Q1	中位数	Q3	最大值
A组	12	8.3658	2.2016	5.78	6.895	7.735	8.9275	12.95
B组	11	12.7318	2.648	6.94	11.51	13.01	14.005	17.72
C组	10	9.249	1.2243	7.19	8.605	9.085	9.86	11.26

方差分析

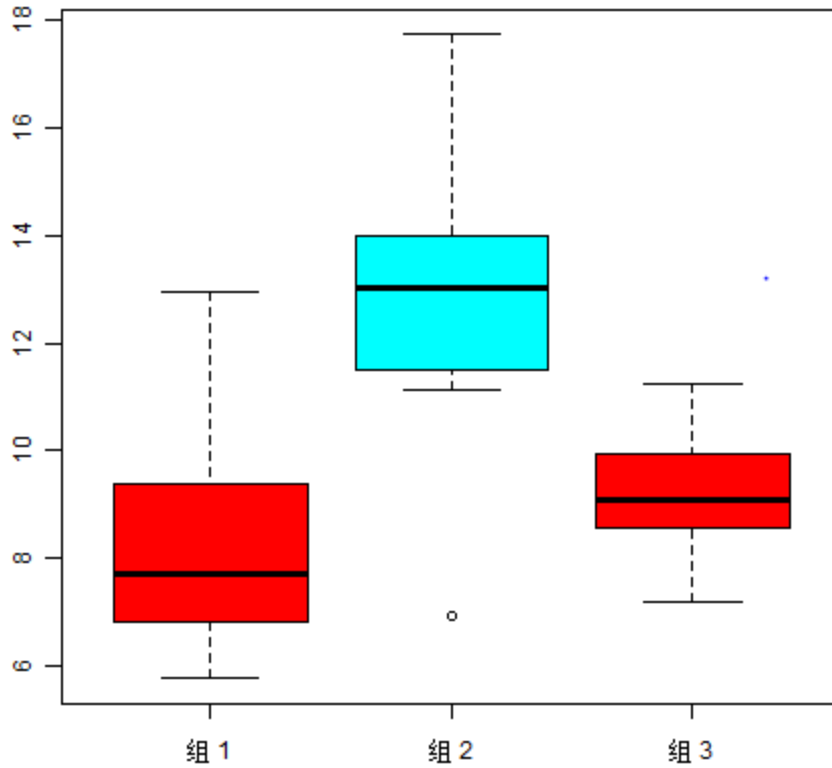
	自由度	离均差平方和	均方	F 值	p 值
factor(列分组)	2	119.5170	59.7585	13.0929	1e-04
Residuals	30	136.9253	4.5642	NA	NA

回归分析

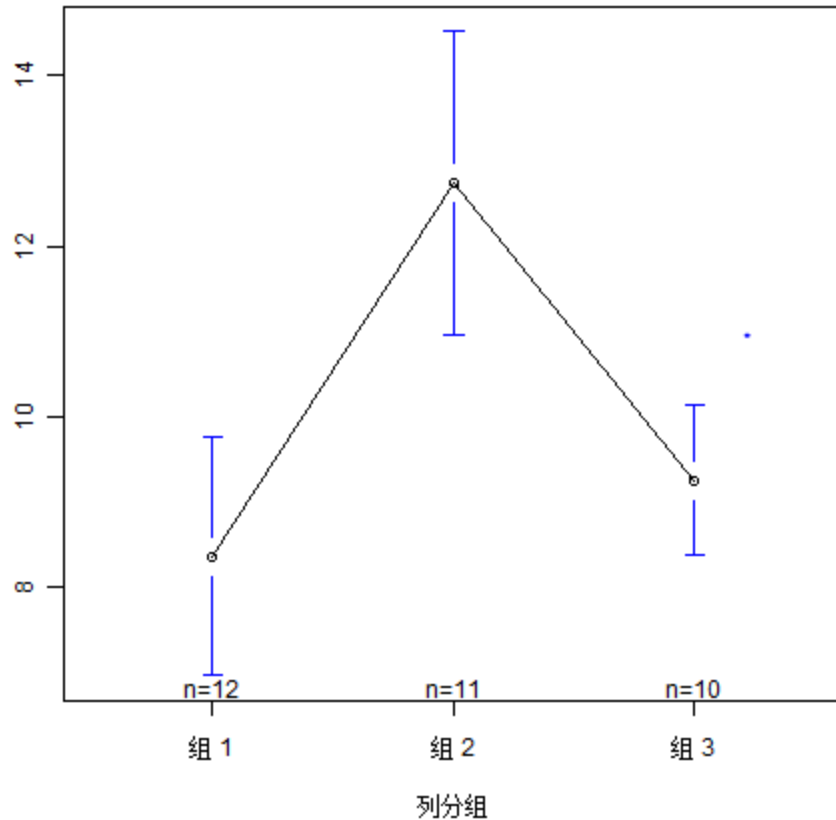
	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	8.36583333333334	0.616723748040649	13.5649605839111	2.4588640650921e-14
factor(列分组)2	4.36598484848485	0.891781066290582	4.89580348083126	3.12621290402499e-05
factor(列分组)3	0.883166666666666	0.914749145435006	0.965474164227482	0.342026303545181

两两(列分组间)比较(Tukey HSD 法)

factor(列分组)	均数差	95%区间下限	95%区间上限	调整的 p 值
2-1	4.366	2.1675	6.5645	1e-04
3-1	0.8832	-1.3719	3.1383	0.6039
3-2	-3.4828	-5.784	-1.1816	0.0022



这里画出 3 组数据分布的箱图



回归分析结果解释：

这里有三组，自动以第一组作为参照，第二组、第三组与之相比较(生成 2 个哑变量分别表示第二、第三组)，得出来的回归方程如下：

$$Y = 8.366 + 4.366 * X2 + 0.833 * X3$$

方程中 X2 与 X3 对应输出结果表中的 factor(列分组)2 与 factor (列分组)3。因为是哑变量，其赋值为：

第一组：X2=0, X3=0

第二组：X2=1, X3=0

第三组：X2=0, X3=1

因此，从方程中可以得出，第一组 Y 的均数是 8.366，第二组比第一组高 4.366， $p=3.126 \times 10^{-5}$ ，第三组比第一组高 0.833， $p=0.342$ 。对照前面输出的各组基本统计指标（均数、标准差等），完全一致。

例 2：（随机区组设计）大白鼠注射不同剂量雌激素后的子宫重量

种类	雌激素剂量		
	0.2	0.4	0.8
A	106	116	145
B	42	68	115
C	70	111	133
D	42	63	87

同例 1，先把数据放到 excel 里然后复制粘贴到数据框内，如下图：

例 2 数据输入界面

分析项目标题: 多组计量数据比较 输出文件名: anovatest

复制粘贴m列数据(制表符分隔)到下框（第一行为列标题或称变量名，第二行开始为数据）
如有分层变量，置前面（前面 列为分层因素）；(后面)每列代表一组(参照组置第一列)。

0.2	0.4	0.8
106	116	145
42	68	115
70	111	133
42	63	87

正态性检验 方差齐性检验 秩和检验及两两比较
 t检验或方差分析 随机区组设计, 双向方差分析 两两(列分组间)均数比较
 两两(行区组间)均数比较 分层因素与列分组交互作用检验

结果存到 (默认为我的文档):

开始分析 >>>

因为这是随机区组设计，所以选“随机区组设计、双向方差分析”，其它参照例 1。

输出结果与解释

多组计量数据比较
数据

Row. groups	0.2	0.4	0.8
1	106	116	145

这里首先列出数据，以便你检查数据读入是否正确。

2	42	68	115
3	70	111	133
4	42	63	87

本例数据符合正态与方差齐性，可以跳过下面的非参数方差分析与非参数的两两比较，而直接看后面的方差分析、回归分析、HSD法两两比较结果。

	统计量(w)	p 值	检验方法	备注
正态性检验	0.9310607	0.3280343	Shapiro-Francia	残差分析

	统计量	p 值	检验方法	备注
方差齐性检验	0.1652459	0.8501996	Brown-Forsythe Levene	Modified robust

	统计量(w)	自由度	p 值	检验方法
非参数方差分析	5.133333	2	0.07679109	Kruskal-Wallis rank sum test

非参数方差分析后两两比较(multiple test after Kruskal-Wallis)

组-组	obs. dif	critical. dif	difference
1-2	2.5	6.10347485721076	0
1-3	5.75	6.10347485721076	0
2-3	3.25	6.10347485721076	0

这是非参数的两两比较结果：
difference=0，表示无显著差异
difference=1，表示有显著差异

critical. dif: p-value=0.05

基本统计

列分组	N	均数	标准差	最小值	Q1	中位数	Q3	最大值
0.2	4	65	30.3535	42	42	56	79	106
0.4	4	89.5	27.8628	63	66.75	89.5	112.25	116
0.8	4	120	25.219	87	108	124	136	145

行区组	N	均数	标准差	最小值	Q1	中位数	Q3	最大值
1	3	122.3333	20.2567	106	111	116	130.5	145
2	3	75	37	42	55	68	91.5	115
3	3	104.6667	31.9739	70	90.5	111	122	133
4	3	64	22.5167	42	52.5	63	75	87

方差分析

	自由度	离均差平方和	均方	F 值	p 值
factor(列分组)	2	6074.0000	3037.0000	33.5374	6e-04
factor(Row.groups)	3	6457.6667	2152.5556	23.7706	1e-03
Residuals	6	543.3333	90.5556	NA	NA

回归分析

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	95.8333	6.7289	14.2421	0
factor(列分组)2	24.5	6.7289	3.641	0.0108
factor(列分组)3	55	6.7289	8.1737	2e-04
factor(Row.groups)2	-47.3333	7.7698	-6.0919	9e-04
factor(Row.groups)3	-17.6667	7.7698	-2.2737	0.0633
factor(Row.groups)4	-58.3333	7.7698	-7.5077	3e-04

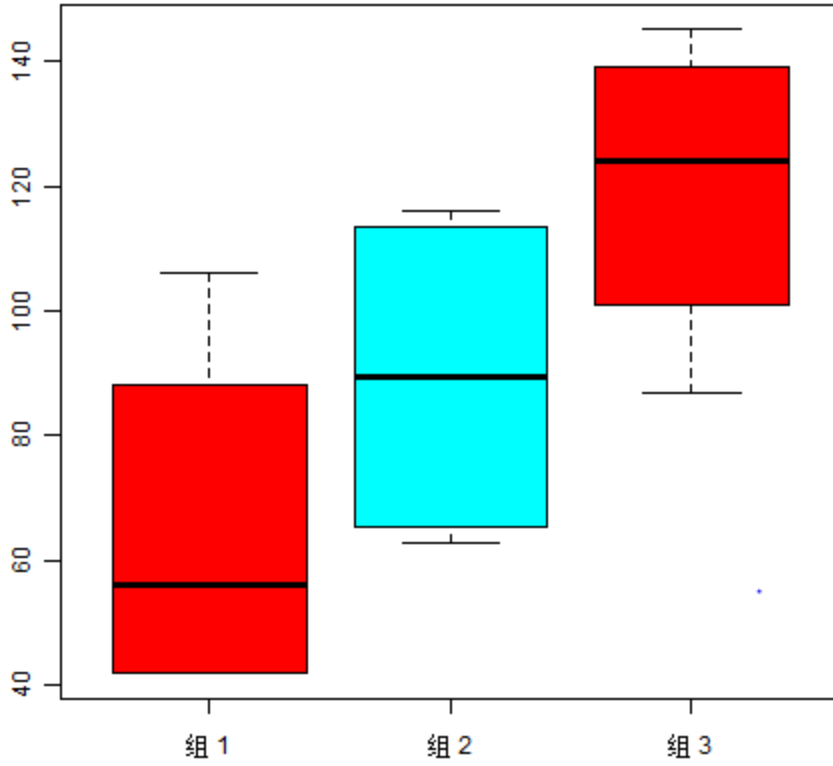
两两(列分组间)比较(Tukey HSD 法)

factor(列分组)	均数差	95% 区间下限	95% 区间上限	调整的 p 值
2-1	24.5	3.854	45.146	0.0252
3-1	55	34.354	75.646	4e-04
3-2	30.5	9.854	51.146	0.0094

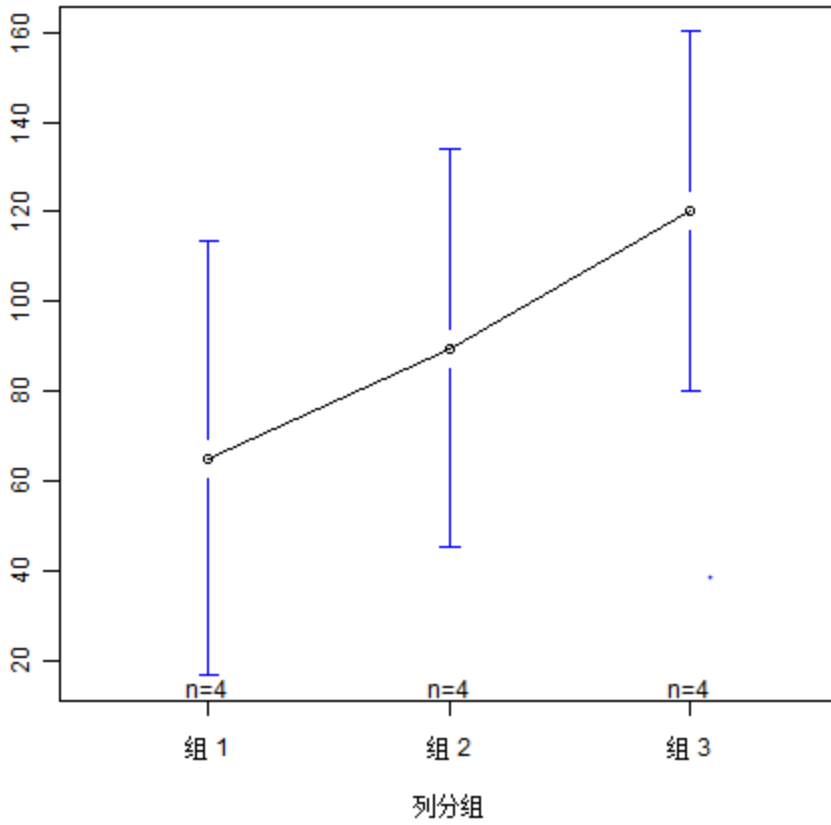
这里回归方程如下：

$$Y_{ij} = \beta_0 + \beta_1 * (i=0.4) + \beta_2 * (i=0.8) + \beta_3 * (j=B) + \beta_4 * (j=C) + \beta_5 * (j=D) + e_{ij}$$

其中 Y_{ij} 表示 i 剂量组第 j 种类大白鼠的观测值， $i=0.2、0.4、0.8$ ， $j=A、B、C、D$ ，分别为哑变量表列分组与行分组，均以第一列、第一行为参照。



这里画出 3 组 (列) 数据分布的箱图



这里画出 3 组 (列) 均数及其 95%可信区间