

多因素方差分析的基本检验

信171 李金哲

目录

CONTENT

1

原理阐释

- 1)基本思想
- 2)方差分析的步骤
- 3)多重比较方法

2

实验步骤

- 1)总体流程
- 2)正态检验
- 3)方差检验

3

结果分析

- 1)结果表格
- 2)正态检验结果分析
- 3)方差检验



01

原理阐释

基本思想

方差分析（ANOVA）是由英国统计学家 R.A.Fisher 首创，为纪念 Fisher，以 F 命名，故方差分析又称 F 检验（F test）。用于推断多个总体均数有无差异。是一种典型的还原论思想。

起源

主要功能

方差分析可以同时检验两个或多个平均数之间的差异以及几个因素水平之间的交互作用。主要功能是分析因变量的总变异中不同来源的变异。

与t检验的差别

t检验只适宜检验两个平均数之间是否存在差异。对于一个复杂的问题，t检验只能进行多组平均数两两之间的差异检验。



方差分析的步骤

1、求平方和

总平方和(SS_T)

组间平方和(SS_B)

组内平方和(SS_W)

$$SS_T = SS_W + SS_B$$

2、计算自由度

组间自由度: $df_B = k - 1$ (k 为组数)

组内自由度: $df_W = k(n - 1)$ (n 为每组人数)

总自由度: $df_T = nk - 1$ 或者 $df_T = df_B + df_W$

3、计算均方

组间均方: $MS_B = SS_B / df_B$

组内均方: $MS_W = SS_W / df_W$

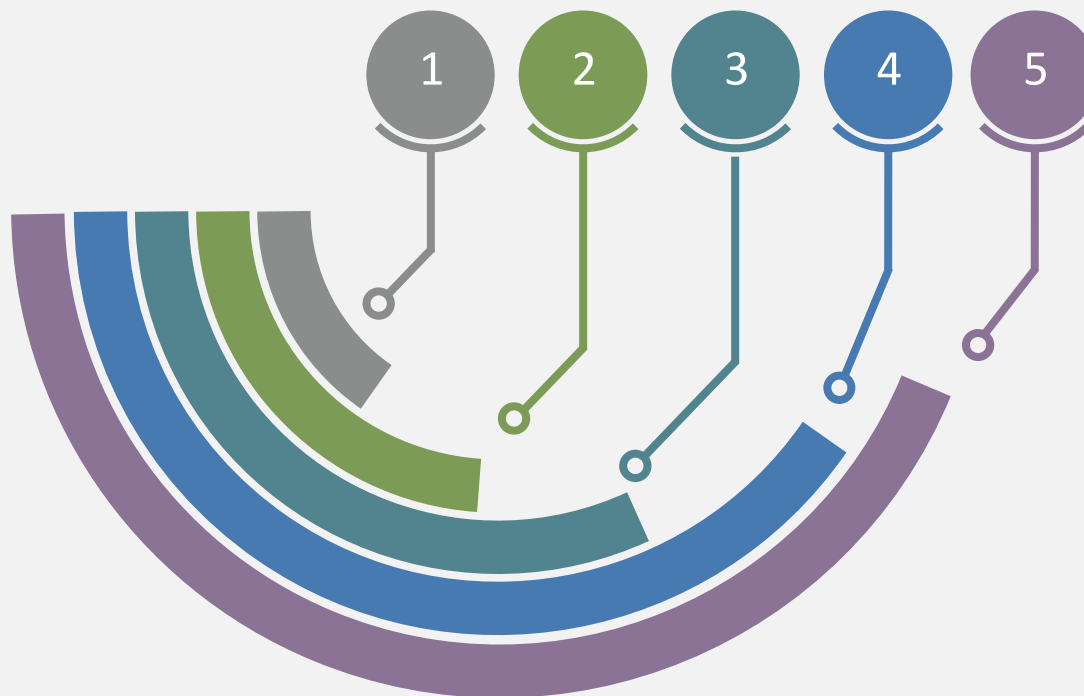
4、计算F值

$$F = MS_B / MS_W \text{ (组间均方/组内均方)}$$

5、进行F检验并作出判断

如果拒绝虚无假设的 p 值定为 $p = 0.05$, 计算的 F 值远小于所确定的显著性水平的临界值, 就可拒绝虚无假设, 说明不同组的平均数之间至少有一对差异显著。

如果 F 值大于所确定的临界值, 就不能拒绝虚无假设。



多重比较方法

LSD最小显著性差异法

利用全部样本数据;在一定程度上克服了放大犯一类错误的问题

LSD



先验对比

多重比较检验中, 若某些水平和另外一些水平的均值有一定的差异, 则可以进一步检验两组的均值是否有显著差异。

先验对比



趋势分析

控制变量为定序型变量, 趋势检验能够分析随着控制变量水平的变换, 观测变量值变化的总体趋势是怎样的 (线性、二次等)

趋势分析





02

实验步骤

总体流程

1.提出原假设

H0:广告形式和地区对销售额无影响
H1:广告形式和地区对销售额有影响



2.正态检验

判断数据是否服从正太分布



3.方差检验

判断检验结果是否具有方差齐性



4.结果阐释

分析表中数据，给出结论



正态检验

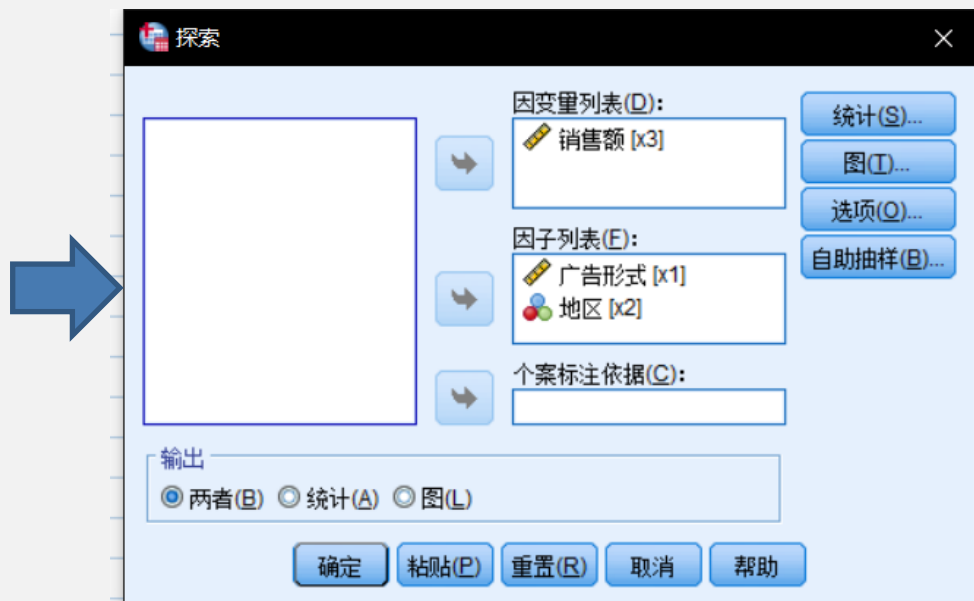
1.选择

分析→描述统计→探索



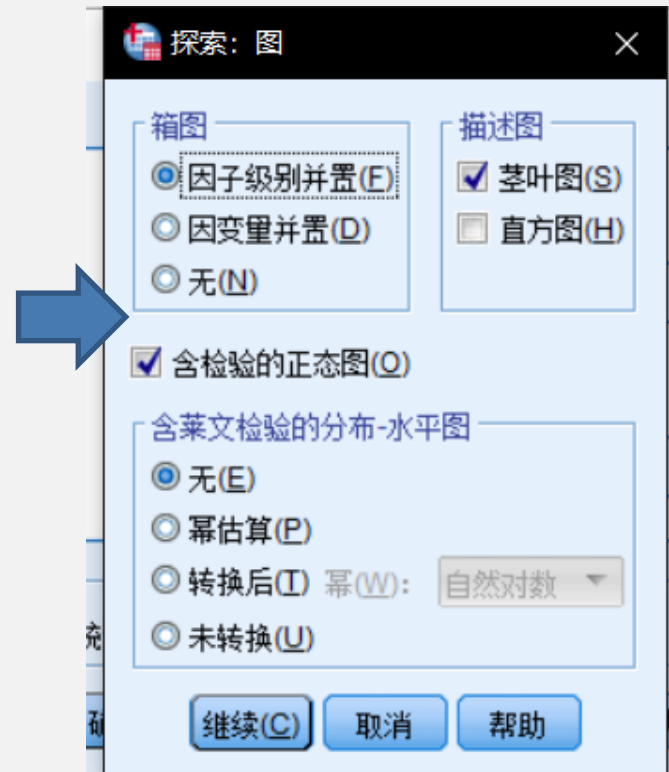
2.变量设置

因变量列表选择销售额，因子列表选择广告形式、地区



3.选择绘制

勾选直方图、带验证的正态图→继续

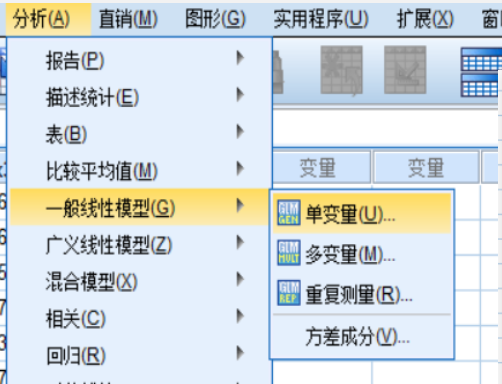


方差检验

1.选择



分析→一般线性模型→单变量



2.变量设置



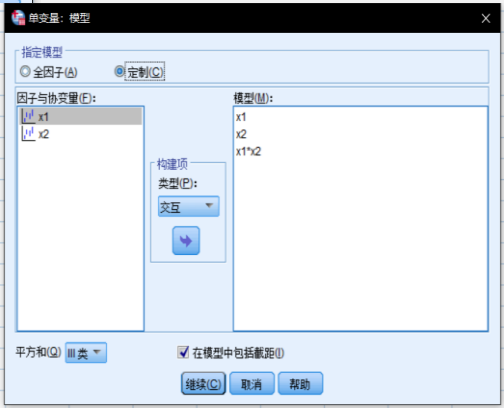
因变量选择销售
额，固定因子选
择广告形式、地
区



3.选择



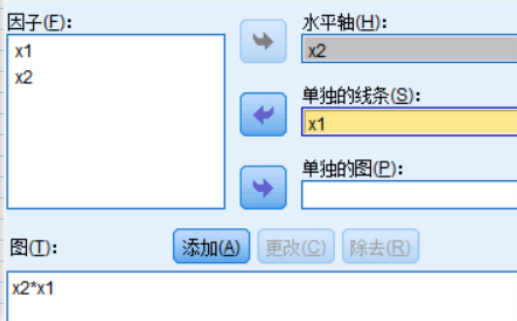
模型→设定→交
互→继续



4.选择



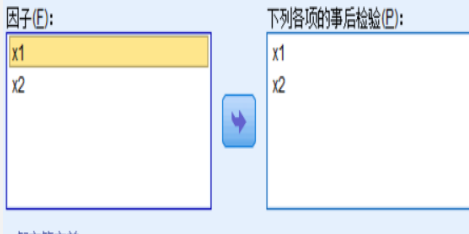
绘制→水平轴选
择地区，单图选
择广告形式→添
加→继续



5.两两比较



两两比较→两两
比较检验广告形
式、地区→继续





03

结果分析

结果表格

假设重述

H0:广告形式和地区对销售额无影响

H1:广告形式和地区对销售额有影响



1)正态性检验

正态性检验							
		柯尔莫戈洛夫-斯米诺夫 ^a			夏皮洛-威尔克		
广告形式	统计	自由度	显著性	统计	自由度	显著性	
销售额	报纸	.128	36	.144	.982	36	.809
	广播	.103	36	.200 [*]	.964	36	.276
	宣传品	.101	36	.200 [*]	.977	36	.629
	体验	.117	36	.200 [*]	.948	36	.092

*. 这是真显著性的下限。

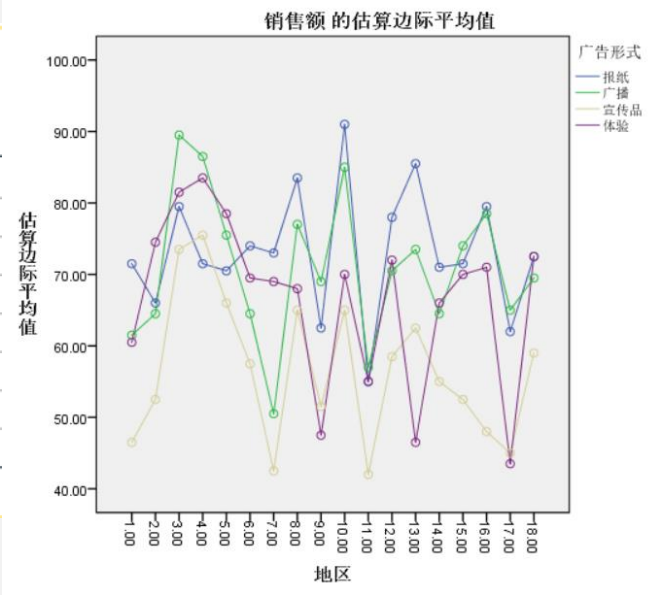
a. 里利氏显著性修正



2)方差检验

主体间效应检验					
因变量: 销售额					
源	III 类平方和	自由度	均方	F	显著性
修正模型	20094.306 ^a	71	283.018	3.354	.000
截距	642936.694	1	642936.694	7619.990	.000
x1	5866.083	3	1955.361	23.175	.000
x2	9265.306	17	545.018	6.459	.000
x1 * x2	4962.917	51	97.312	1.153	.286
误差	6075.000	72	84.375		
总计	669106.000	144			
修正后总计	26169.306	143			

a. R 方 = .768 (调整后 R 方 = .539)



正态检验结果分析

正态性检验							
		柯尔莫戈洛夫-斯米诺夫 ^a			夏皮洛-威尔克		
广告形式		统计	自由度	显著性	统计	自由度	显著性
销售额	报纸	.128	36	.144	.982	36	.809
	广播	.103	36	.200 [*]	.964	36	.276
	宣传品	.101	36	.200 [*]	.977	36	.629
	体验	.117	36	.200 [*]	.948	36	.092

*. 这是真显著性

a. 里利氏显著性

正态性检验							
		柯尔莫戈洛夫-斯米诺夫 ^a			夏皮洛-威尔克		
地区		统计	自由度	显著性	统计	自由度	显著性
销售额	1.00	.142	8	.200 [*]	.975	8	.933
	2.00	.159	8	.200 [*]	.961	8	.823
	3.00	.199	8	.200 [*]	.943	8	.641
	4.00	.162	8	.200 [*]	.973	8	.918
	5.00	.188	8	.200 [*]	.954	8	.747
	6.00	.152	8	.200 [*]	.951	8	.721
	7.00	.329	8	.011	.798	8	.027
	8.00	.196	8	.200 [*]	.928	8	.502
	9.00	.209	8	.200 [*]	.923	8	.458
	10.00	.203	8	.200 [*]	.917	8	.404
	11.00	.186	8	.200 [*]	.932	8	.532
	12.00	.229	8	.200 [*]	.948	8	.688
	13.00	.150	8	.200 [*]	.958	8	.794
	14.00	.154	8	.200 [*]	.991	8	.996
	15.00	.196	8	.200 [*]	.949	8	.699
	16.00	.306	8	.026	.806	8	.033
	17.00	.150	8	.200 [*]	.966	8	.866
	18.00	.152	8	.200 [*]	.949	8	.701



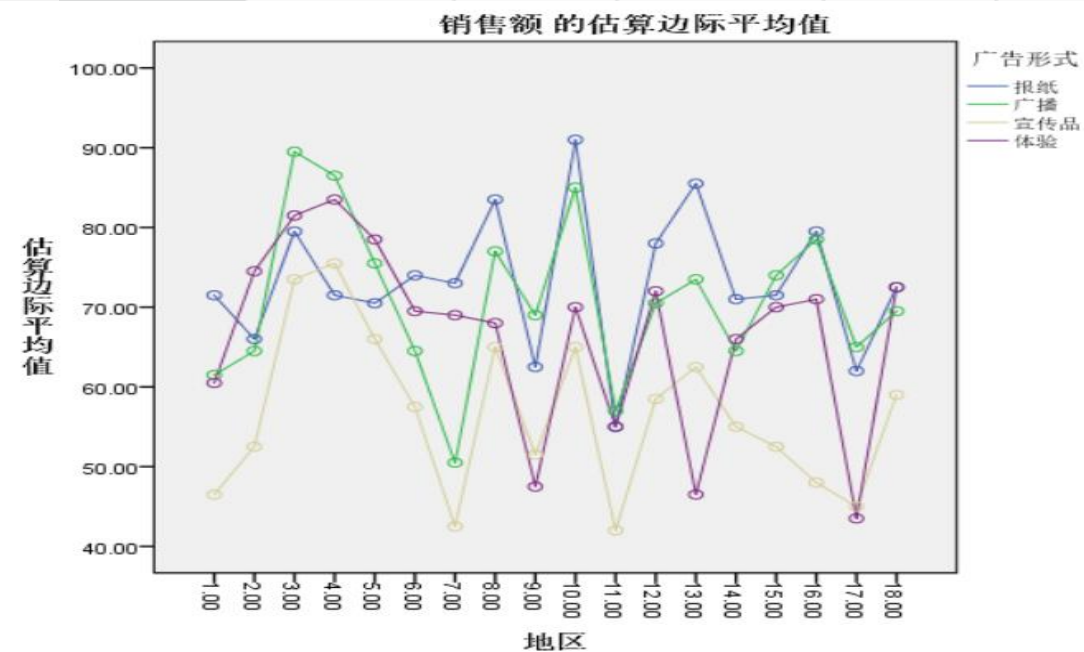
正态检验

结果分析

两个正态检验表格中的p值均大于0.05(红框)，即公告形式和销售额，以及地区和销售额p值均大于0.05。
故接受原假设，因变量与正态分布无明显差异，因此数据服从正态分布

方差检验

主体间效应检验					
因变量: 销售额					
源	III 类平方和	自由度	均方	F	显著性
修正模型	20094.306 ^a	71	283.018	3.354	.000
截距	642936.694	1	642936.694	7619.990	.000
x1	5866.083	3	1955.361	23.175	.000
x2	9265.306	17	545.018	6.459	.000
x1 * x2	4962.917	51	97.312	1.153	.286
误差	6075.000	72	84.375		
总计	669106.000	144			
修正后总计	26169.306	143			



方差检验

在方差齐次检验中:

X1(广告形式)的统计量的值 $F=23.175$, P 值为0.000;

X2(地区) 的统计量的值 $F=6.459$, P 值为0.000;

X1和X2(广告形式和地区)交互作用的统计量的值 $F=1.153$, P 值为0.286



结论

- 1、该数据中的方差齐性检验结果是具有方差齐性的
- 2、
 - X1(广告形式) $\text{sig}=0.000 < 0.05$, 广告形式对销售额有影响;
 - X2(地区) $\text{sig}=0.000 < 0.05$, 地区对销售额有影响;
 - X1和X2 $\text{sig}=0.286 > 0.05$, 广告形式和地区的交互作用对销售额无影响。
- 3、由图像可知, 当地区不同时, 广告形式对于销售额的影响也是不同的。