



单因素方差分析

信171 李金哲

单因素方差分析原理

- 在观测变量总离差平方和中，如果组间离差平方和所占比例较大，则说明观测变量的变动主要是由控制变量引起的，可以主要由控制变量来解释，控制变量给观测变量带来了显著影响；反之，如果组间离差平方和所占比例小，则说明观测变量的变动不是主要由控制变量引起的，不可以主要由控制变量来解释，控制变量的不同水平没有给观测变量带来显著影响，观测变量值的变动是由随机变量因素引起的。

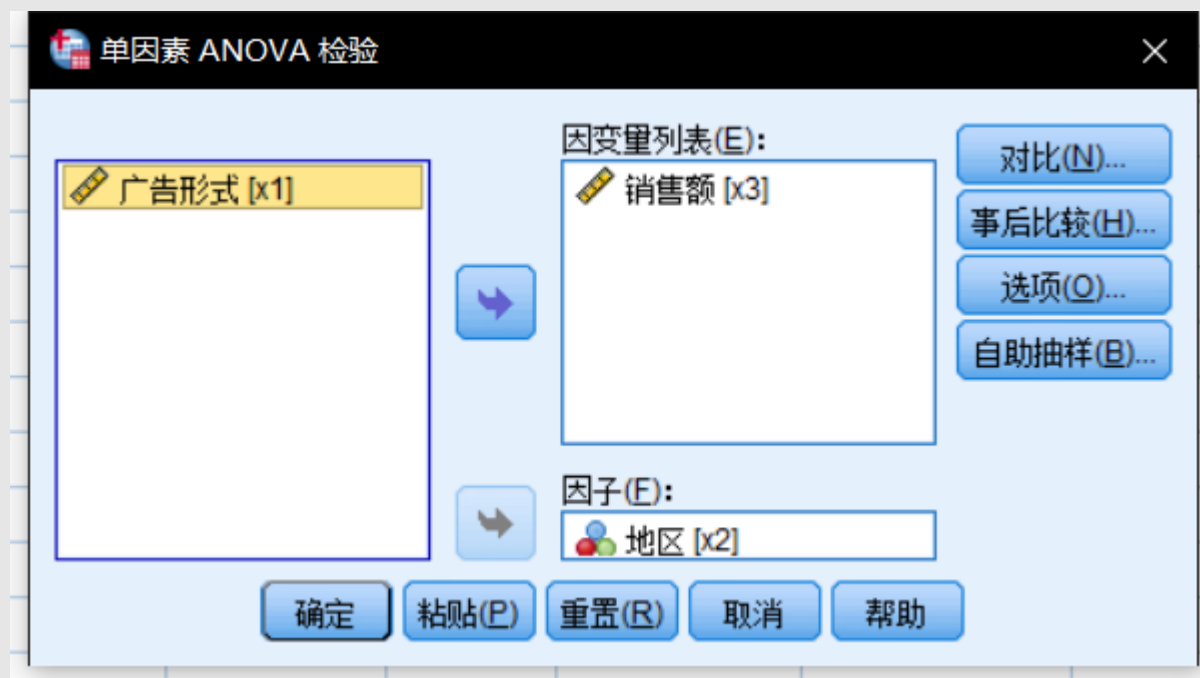
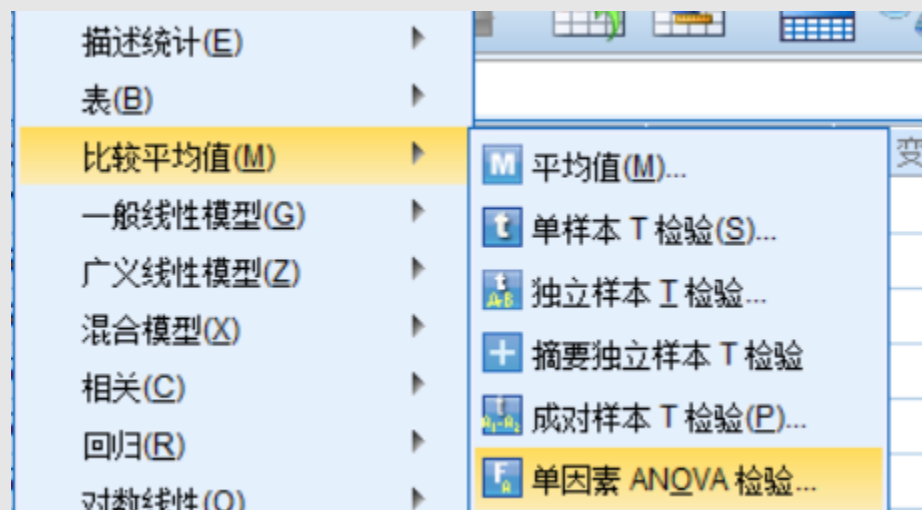
单因素方差分析理解步骤

- 1、明确观测变量和控制变量
- 2、剖析观测变量的方差
- 3、通过比较观测变量总离差平方和各部分所占的比例，推断控制变量是否给观测变量带来了显著影响

单因素方差分析基本步骤

- 1、提出原假设： H_0 ——无差异； H_1 ——有显著差异
- 2、选择检验统计量： 方差分析采用的检验统计量是F统计量，即F值检验
- 3、计算检验统计量的观测值和概率P值： 该步骤的目的就是计算检验统计量的观测值和相应的概率P值。
- 4、通过给定显著性水平，并作出决策

实验过程



先提出原假设：

H0：地区对销售额无相关影响，即 $\mu_1 = \mu_2 = \mu_3 \dots$

H1：地区对销售额存在影响

ANOVA					
销售额	平方和	自由度	均方	F	显著性
组间	9265.306	17	545.018	4.062	.000
组内	16904.000	126	134.159		
总计	26169.306	143			

由方差检验的结果，可以明显得到销售额与地区的相关性概率p值十分小，可以约等于0；而且远小于接受H0的概率 $\alpha = 0.05$ 。所以拒接原假设H0，接受假设H1。即地区对销售额存在显著性影响

由于地区对销售额存在显著性影响，但是无法确定其中每个因素的影响显著性，故而进行事后检验分析

单因素 ANOVA 检验：事后多重比较

假定等方差

☒ LSD ☐ S-N-K ☐ 沃勒-邓肯(W)

☐ 邦弗伦尼(B) ☐ 图基(T) I类/II类误差率: 100

☐ 斯达克(I) ☐ 图基 s-b(K) ☐ 邓尼特(E)

☐ 雪费(C) ☐ 邓肯(D) 控制类别(Y): 最后一个

☐ R-E-G-W F ☐ 霍赫伯格 GT2(H) 检验

☐ R-E-G-W Q ☐ 加布里埃尔(G) ☒ 双侧(2) ☐ < 控制(Q) ☐ > 控制(N)

不假定等方差

☒ 塔姆黑尼 T2(M) ☐ 邓尼特 T3 ☐ 盖姆斯-豪厄尔(A) ☐ 邓尼特 C(U)

显著性水平(E): 0.05

继续(C) 取消 帮助

单因素 ANOVA 检验：...

统计

☒ 描述(D) ☐ 固定和随机效应(F)

☒ 方差齐性检验(H) ☐ 布朗-福塞斯(B)

☐ 韦尔奇(W)

☐ 平均值图(M)

缺失值

☒ 按具体分析排除个案(A) ☐ 成列排除个案(L)

继续(C) 取消 帮助

根据上一问的概率p值约等于0，远小于0.05，得知本次检验使用 塔姆黑尼 的检验数据

塔姆黑尼	1.00	2.00	-4.37500	6.15263	1.000	-33.7112	24.9612
		3.00	-21.00000	5.49025	.248	-46.8557	4.8557
		4.00	-19.25000	4.71226	.194	-42.3107	3.8107
		5.00	-12.62500	4.96033	.976	-36.3412	11.0912
		6.00	-6.37500	4.93869	1.000	-30.0229	17.2729
		7.00	1.25000	7.24261	1.000	-34.7730	37.2730
		8.00	-13.37500	5.04245	.950	-37.3668	10.6168
		9.00	2.37500	5.54185	1.000	-23.7263	28.4763
		10.00	-17.75000	6.43303	.918	-48.7219	13.2219
		11.00	7.75000	5.37105	1.000	-17.5595	33.0595
		12.00	-9.75000	5.25680	1.000	-34.5671	15.0671
		13.00	-7.00000	6.82956	1.000	-40.3955	26.3955
		14.00	-4.12500	4.73753	1.000	-27.2383	18.9883
		15.00	-7.00000	5.55813	1.000	-33.1799	19.1799
		16.00	-9.25000	6.37728	1.000	-39.8908	21.3908
		17.00	6.12500	5.68343	1.000	-20.6760	32.9260
		18.00	-8.37500	4.93869	1.000	-32.0229	15.2729
	2.00	1.00	4.37500	6.15263	1.000	-24.9612	33.7112
		3.00	-16.62500	6.15263	.935	-45.9612	12.7112
		4.00	-14.87500	5.46968	.954	-42.8695	13.1195
		5.00	-8.25000	5.68480	1.000	-36.4141	19.9141
		6.00	-2.00000	5.66592	1.000	-30.1376	26.1376
		7.00	5.62500	7.75677	1.000	-31.5531	42.8031
		8.00	-9.00000	5.75660	1.000	-37.2820	19.2820
		9.00	6.75000	6.19872	1.000	-22.7445	36.2445
		10.00	-13.37500	7.00685	1.000	-46.4237	19.6737
		11.00	12.12500	6.04651	1.000	-16.8716	41.1216
		12.00	5.37500	5.31525	1.000	-24.8224	22.8224

由于检验组别过多，在此不做一一展示，仅阐释分析结果。根据塔姆尼黑中的显著性差异得出，任意两组的显著性水平平均大于0.05，说明每两个组之间均没有差异，即地区之间没有显著性差异。