第 7 章 方差分析

- 7.1 方差分析的引论
- 7.2 单因素方差分析
- 7.3 双因素方差分析

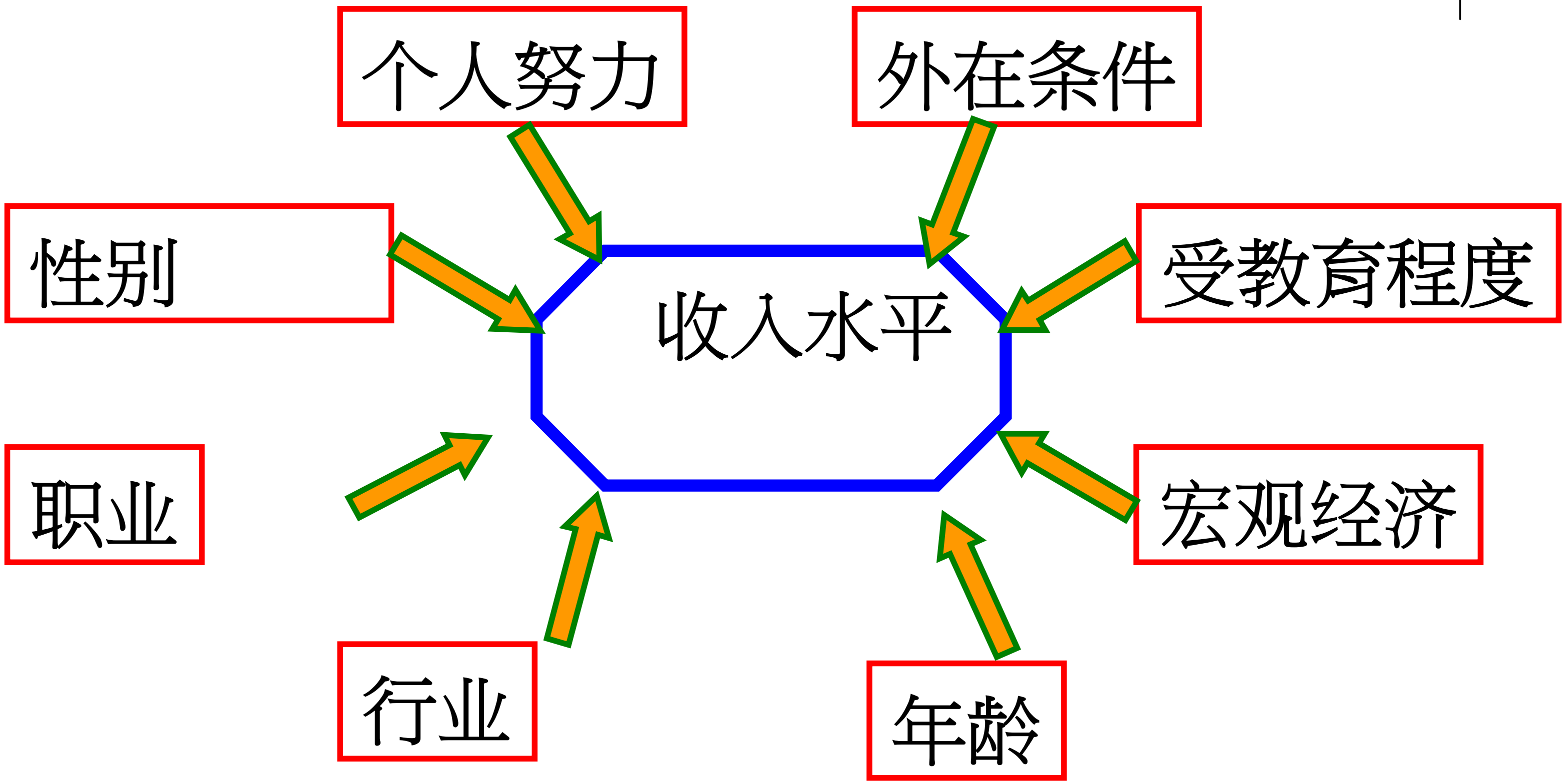
7.1方差分析引论

一、方差分析及其有关术语

二、方差分析的基本思想和原理

三、方差分析的基本假定

四、问题的一般提法


个人努力
外在条件
性别
收入水平
受教育程度
职业
宏观经济
行业
年龄

试验指标——试验中要考察的指标，即收入.

因　　素——影响试验指标的条件，如性别，年龄.

因素 → 可控因素

因素 → 不可控因素

水　　平——因素所处的状态，如性别分男和女.

单因素试验——在一项试验中只有一个因素改变.

多因素试验——在一项试验中有多个因素在改变.

设有三台机器,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千分之一厘米.

铝合金板的厚度

机器 I	机器 II	机器 III
0.236	**0.257**	**0.258**
0.238	**0.253**	**0.264**
0.248	**0.255**	**0.259**
0.245	**0.254**	**0.267**
0.243	**0.261**	**0.262**

试验指标:薄板的厚度

因　　素:机器

水　　平:不同的三台机器是因素的三个不同的水平

试验目的:考察各台机器所生产的薄板的厚度有无显著的差异.即考察机器这一因素对厚度有无显著的影响

单因素试验

一火箭用四种燃料,三种推进器作射程试验.每种燃料与每种推进器的组合各发射两次,得射程如下

火箭的射程　　　　单位：（海里）

推进器(B)		B_1	B_2	B_3
燃料(A)	A_1	58.2 52.6	56.2 41.2	65.3 60.8
	A_2	49.1 42.8	54.1 50.5	51.6 48.4
	A_3	60.1 58.3	70.9 73.2	39.2 40.7
	A_4	75.8 71.5	58.2 51.0	48.7 41.4

试验指标:射程
因　　素:推进器和燃料
水　　平:推进器有3个,燃料有4个
实验目的:推进器和燃料对射程的影响是否有显著的差异
　　　　或二者产生交互作用对射程的影响是否有显著差异
　　　　或二者不产生交互作用对射程的影响是否有显著差异

双因素试验

方差分析(ANOVA)
(analysis of variance)

检验多个总体均值是否相等
- 通过分析数据的误差判断各总体均值是否相等

研究分类型自变量对数值型因变量的影响
- 一个或多个分类尺度的自变量
  - 两个或多个 (k 个) 处理水平或分类
- 一个间隔或比率尺度的因变量

有单因素方差分析和双因素方差分析
- 单因素方差分析：涉及一个分类的自变量
- 双因素方差分析：涉及两个分类的自变量

【例】为了对几个行业的服务质量进行评价，消费者协会在四个行业分别抽取了不同的企业作为样本。最近一年中消费者对总共23家企业投诉的次数如下表

消费者对四个行业的投诉次数				
	行业			
观测值	零售业	旅游业	航空公司	家电制造业
1	57	68	31	44
2	66	39	49	51
3	49	29	21	65
4	40	45	34	77
5	34	56	40	58
6	53	51		
7	44			

试验指标:投诉次数
因　　素:行业
水　　平:行业有4种
实验目的:行业对投诉次数的影响是否有显著差异或四个行业服务质量是否显著不同

方差分析
(例题分析)

分析四个行业之间的服务质量是否有显著差异，也就是要判断“行业”对“投诉次数”是否有显著影响

作出这种判断最终被归结为检验这四个行业被投诉次数的均值是否相等

若它们的均值相等，则意味着“行业”对投诉次数是没有影响的，即它们之间的服务质量没有显著差异；若均值不全相等，则意味着“行业”对投诉次数是有影响的，它们之间的服务质量有显著差异

方差分析的基本思想和原理

方差分析：检验均值是否相等

比较的基础是方差比

如果系统(处理)误差明显地不同于随机误差，则均值不相等；反之，均值相等

误差是由各部分的误差占总误差的比例来测度的

误差包括：随机误差和系统误差

方差分析的基本思想和原理
(两类误差)

随机误差

- 因素的同一水平(总体)下，样本各观察值之间的差异
- 比如，同一行业下不同企业被投诉次数是不同的
- 这种差异可以看成是随机因素的影响，称为*随机误差*

系统误差

- 因素的不同水平(不同总体)下，各观察值之间的差异
- 比如，不同行业之间的被投诉次数之间的差异
- 这种差异*可能*是由于抽样的随机性所造成的，*也可能*是由于行业本身所造成的，后者所形成的误差是由系统性因素造成的，称为*系统误差*

方差分析的基本思想和原理
(两类方差)

数据的误差用平方和(sum of squares)表示，称为方差

组内方差(within groups)

- 因素的同一水平(同一个总体)下样本数据的方差
- 组内方差只包含*随机误差*

组间方差(between groups)

- 因素的不同水平(不同总体)下各样本之间的方差
- 组间方差既包括*随机误差*，也包括*系统误差*

方差分析的基本思想和原理
(方差的比较)

若因素不同水平对试验指标没有影响，则组间误差中只包含随机误差，没有系统误差。这时，组间误差与组内误差经过平均后的数值就应该很接近，它们的比值就会接近1

若因素不同水平对试验指标有影响，在组间误差中除了包含随机误差外，还会包含有系统误差，这时组间误差平均后的数值就会大于组内误差平均后的数值，它们之间的比值就会大于1

当这个比值大到某种程度时，就可以说不同水平之间存在着显著差异，也就是自变量对因变量有影响

方差分析的基本假定

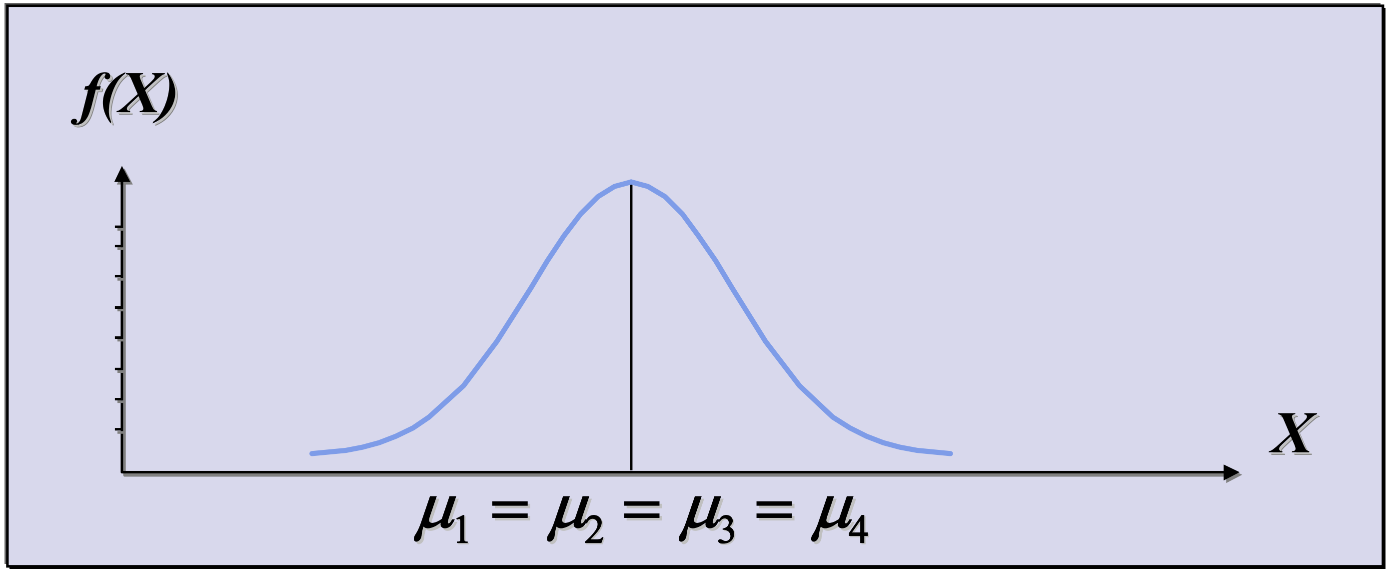
每个总体都应服从正态分布
- 对于因素的每一个水平，其观察值是来自服从正态分布总体的简单随机样本
- 比如，每个行业被投诉的次数必需服从正态分布

各个总体的方差必须相同
- 各组观察数据是从具有相同方差的总体中抽取的
- 比如，四个行业被投诉次数的方差都相等

观察值是独立的
- 比如，每个行业被投诉的次数与其他行业被投诉的次数独立

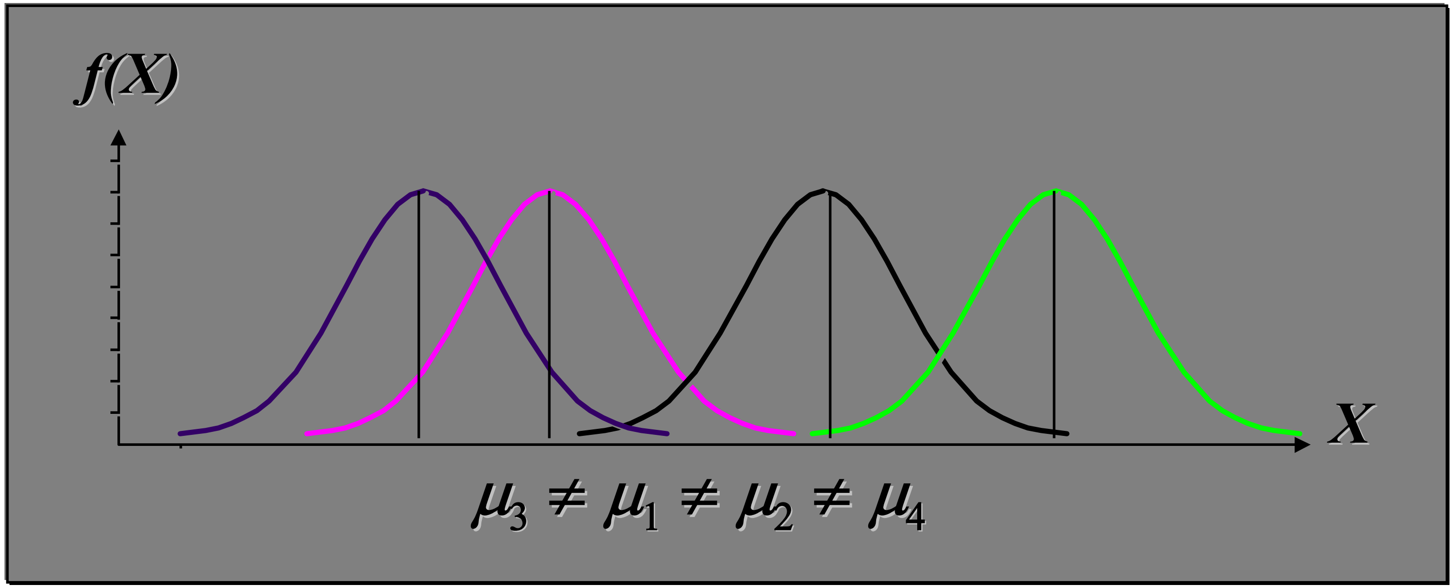
方差分析中基本假定

- 如果原假设成立，即H_0： $\mu_1 = \mu_2 = \mu_3 = \mu_4$
  - 四个行业被投诉次数的均值都相等
  - 意味着每个样本都来自均值为 μ、方差为σ^2的同一正态总体



方差分析中基本假定

- 若备择假设成立，即H_1： μ_i $(i=1,2,3,4)$不全相等
  - 至少有一个总体的均值是不同的
  - 四个样本分别来自均值不同的四个正态总体



7.2 单因素方差分析

一、数据结构

二、分析步骤

三、关系强度的测量

四、用**Excel**进行方差分析

单因素方差分析的数据结构
(one-way analysis of variance)

观察值（j）	因素(A) i			
	水平A_1	水平A_2	…	水平A_k
1	x_{11}	x_{21}	…	x_{k1}
2	x_{12}	x_{22}	…	x_{k2}
:	:	:	:	:
:	:	:	:	:
n	x_{1n}	x_{2n}	…	x_{kn}

提出假设

一般提法

- H_0： $\mu_1=\mu_2=\ldots=\mu_k$
  - 自变量对因变量没有显著影响
- H_1： μ_1， μ_2， … ， μ_k不全相等
  - 自变量对因变量有显著影响

注意：拒绝原假设，只表明至少有两个总体的均值不相等，并不意味着所有的均值都不相等

构造检验的统计量
(计算水平的均值)

假定从第i个总体中抽取一个容量为n_i的简单随机样本，第i个总体的样本均值为该样本的全部观察值总和除以观察值的个数

计算公式为

$$\bar{x}_i = \frac{\sum_{j=1}^{n_i} x_{ij}}{n_i} \qquad (i = 1,2,\cdots,k)$$

式中：n_i为第 i 个总体的样本观察值个数

x_{ij} 为第 i 个总体的第 j 个观察值

构造检验的统计量
(计算全部观察值的总均值)

全部观察值的总和除以观察值的总个数

计算公式为

$$\bar{\bar{x}} = \frac{\sum_{i=1}^{k}\sum_{j=1}^{n_i} x_{ij}}{n} = \frac{\sum_{i=1}^{k} n_i \bar{x}_i}{n}$$

式中：$n = n_1 + n_2 + \cdots + n_k$

构造检验的统计量
(例题分析)

	A	B	C	D	E
1	观测值	行业			
2		零售业	旅游业	航空公司	家电制造业
3	1	57	68	31	44
4	2	66	39	49	51
5	3	49	29	21	65
6	4	40	45	34	77
7	5	34	56	40	58
8	6	53	51		
9	7	44			
10	样本均值	$\bar{x}_1 = 49$	$\bar{x}_2 = 48$	$\bar{x}_3 = 35$	$\bar{x}_4 = 59$
11	样本容量（n）	7	6	5	5
12	总均值	$\bar{\bar{x}} = \frac{57+66+\cdots+77+58}{23} = 47.869565$			

构造检验的统计量
(计算总误差平方和 *SST*)

全部观察值 x_{ij} 与总平均值 $\bar{\bar{x}}$ 的离差平方和

反映全部观察值的离散状况

其计算公式为

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

- 前例的计算结果：

$SST = (57-47.869565)^2+\ldots+(58-47.869565)^2$

$=115.9295$

构造检验的统计量
(计算水平项平方和 *SSA*)

各组平均值 $\bar{x}_i \ (i=1,2,\cdots,k)$ 与总平均值 $\bar{\bar{x}}$ 的离差平方和

反映各总体的样本均值之间的差异程度，又称组间平方和

该平方和既包括随机误差，也包括系统误差

计算公式为

$$SSA=\sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(\bar{x}_i-\bar{\bar{x}}\right)^2=\sum_{i=1}^{k}n_i\left(\bar{x}_i-\bar{\bar{x}}\right)^2$$

- 前例的计算结果：*SSA* = 1456.608696

构造检验的统计量
(计算误差项平方和 *SSE*)

每个水平或组的各样本数据与其组平均值的离差平方和

反映每个样本各观察值的离散状况，又称组内平方和

该平方和反映的是随机误差的大小

计算公式为

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

- 前例的计算结果：*SSE* = 2708

构造检验的统计量
(三个平方和的关系)

- ➡总离差平方和(*SST*)、误差项离差平方和(*SSE*)、水平项离差平方和 (*SSA*) 之间的关系

$$\sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(\bar{x}_{ij}-\bar{\bar{x}}\right)^2=\sum_{i=1}^{k}n_i\left(\bar{x}_i-\bar{\bar{x}}\right)^2+\sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(x_{ij}-\bar{x}\right)^2$$

SST = SSA + SSE

- 前例的计算结果：

  4164.608696=1456.608696+2708

构造检验的统计量
(三个平方和的作用)

SST反映全部数据总的误差程度；SSE反映随机误差的大小；SSA反映随机误差和系统误差的大小

如果原假设成立，则表明没有系统误差，组间平方和SSA除以自由度后的均方与组内平方和SSE和除以自由度后的均方差异就不会太大；如果组间均方显著地大于组内均方，说明各水平(总体)之间的差异不仅有随机误差，还有系统误差

判断因素的水平是否对其观察值有影响，实际上就是比较组间方差与组内方差之间差异的大小

构造检验的统计量
(计算均方*MS*)

各误差平方和的大小与观察值的多少有关，为消除观察值多少对误差平方和大小的影响，需要将其平均，这就是*均方*，也称为方差

计算方法是用误差平方和除以相应的自由度

三个平方和对应的自由度分别是

- ***SST*** 的自由度为n-1，其中n为全部观察值的个数
- ***SSA***的自由度为k-1，其中k为因素水平(总体)的个数
- ***SSE*** 的自由度为n-k

构造检验的统计量
(计算均方 *MS*)

组间方差：*SSA*的均方，记为*MSA*，计算公式为

$$MSA = \frac{SSA}{k-1}$$ 前例计算结果：$MSA = \frac{1456.608696}{4-1} = 485.536232$

2. 组内方差：*SSE*的均方，记为*MSE*，计算公式为

$$MSE = \frac{SSE}{n-k}$$ 前例计算结果：$MSE = \frac{2708}{23-4} = 142.526316$

构造检验的统计量
(计算检验统计量 *F*)

将*MSA*和*MSE*进行对比，即得到所需要的检验统计量*F*

当H_0为真时，二者的比值服从分子自由度为k-1、分母自由度为 n-k 的 *F* 分布，即

$$F=\frac{MSA}{MSE}\sim F(k-1,n-k)$$

前例计算结果：$F=\dfrac{485.536232}{142.526316}=3.406643$

统计决策

- ➡ 将统计量的值F与给定的显著性水平α的临界值F_α进行比较，作出对原假设H_0的决策
  - 根据给定的显著性水平α，在F分布表中查找与第一自由度$df_1 = k$-1、第二自由度$df_2=n-k$ 相应的临界值F_α
  - 若$\boldsymbol{F>F_\alpha}$，则拒绝原假设H_0，表明均值之间的差异是显著的，所检验的因素对观察值有显著影响
  - 若$\boldsymbol{F<F_\alpha}$，则不能拒绝原假设H_0，表明所检验的因素对观察值没有显著影响

输出结果

方差分析：单因素方差分析

SUMMARY

组	观测数	求和	平均	方差
列 1	7	343	49	116.6667
列 2	6	288	48	184.8
列 3	5	175	35	108.5
列 4	5	295	59	162.5

方差分析

差异源	SS	df	MS	F	P-value	F crit
组间	1456.609	3	485.5362	3.406643	0.038765	3.12735
组内	2708	19	142.5263			
总计	4164.609	22				

结果：拒绝原假设表明因素(自变量)与观测值之间有关系

关系强度的测量

只要组间平方和SSA不等于0，就表明两个变量之间有关系(只是是否显著的问题)

当组间平方和比组内平方和(SSE)大，而且大到一定程度时，就意味着两个变量之间的关系显著。反之，就意味着两个变量之间的关系不显著

因此，变量间关系的强度用自变量平方和(SSA)占总平方和(SST)的比例大小来反映

自变量平方和占总平方和的比例记为R2 ,即其平方根R就可以用来测量两个变量之间的关系强度

$$R^2 = \frac{SSA\ (组间平方和)}{SST\ (总平方和)}$$

关系强度的测量
(例题分析)

$$R^2 = \frac{SSA}{SST} = \frac{1456.608696}{4146.608696} = 0.349759 = 34.9759\%$$

- R=0.591404
- 结论：
  - 行业(自变量)对投诉次数(因变量)的影响效应占总效应的34.9759%，而残差效应则占65.0241%。即行业对投诉次数差异解释的比例达到近35%，而其他因素(残差变量)所解释的比例近为65%以上
  - R=0.591404，表明行业与投诉次数之间有中等以上的关系

7.3双因素方差分析
(two-way analysis of variance)

分析两个因素(行因素Row和列因素Column)对试验结果的影响

如果两个因素对试验结果的影响是相互独立的，分别判断行因素和列因素对试验数据的影响，这时的双因素方差分析称为*无交互作用的双因素方差分析*或*无重复双因素方差分析*(Two-factor without replication)

如果除了行因素和列因素对试验数据的单独影响外，两个因素的搭配还会对结果产生一种新的影响，这时的双因素方差分析称为*有交互作用的双因素方差分析*或*可重复双因素方差分析* (Two-factor with replication)

双因素方差分析的基本假定

每个总体都服从正态分布

- 对于因素的每一个水平，其观察值是来自正态分布总体的简单随机样本

各个总体的方差必须相同

- 对于各组观察数据，是从具有相同方差的总体中抽取的

观察值是独立的

无交互作用的双因素方差分析
(例题分析)

【例】有4个品牌的彩电在5个地区销售，为分析彩电的品牌(品牌因素)和销售地区(地区因素)对销售量是否有影响，对每种品牌在各地区的销售量取得以下数据。试分析品牌和销售地区对彩电的销售量是否有显著影响？(α=0.05)

不同品牌的彩电在各地区的销售量数据

品牌因素	地区因素				
	地区1	地区2	地区3	地区4	地区5
品牌1	365	350	343	340	323
品牌2	345	368	363	330	333
品牌3	358	323	353	343	308
品牌4	288	280	298	260	298

数据结构

	A	B	C	D	E	F	G
1			列因素（j）				平均值
2			列1	列2	…	列r	$\bar{x}_{i.}$
3	行因素（i）	行1	x_{11}	x_{12}	…	x_{1k}	$\bar{x}_{1.}$
4		行2	x_{21}	x_{22}	…	x_{2k}	$\bar{x}_{2.}$
5		⋮	⋮	⋮	⋮	⋮	⋮
6		行k	x_{r1}	x_{r2}	…	x_{rk}	$\bar{x}_{r.}$
7–8	平均值 $\bar{x}_{.j}$		$\bar{x}_{.1}$	$\bar{x}_{.2}$	…	$\bar{x}_{.k}$	$\bar{\bar{x}}$

数据结构

- $\bar{x}_{i.}$ 是行因素的第i个水平下各观察值的平均值

$$\bar{x}_{i.}=\frac{\sum_{j=1}^{r}x_{ij}}{r}\qquad(i=1,2,\cdots,k)$$

- $\bar{x}_{.j}$ 是列因素的第j个水平下的各观察值的均值

$$\bar{x}_{.j}=\frac{\sum_{i=1}^{k}x_{ij}}{k}\qquad(j=1,2,\cdots,r)$$

- $\bar{\bar{x}}$ 是全部 kr 个样本数据的总平均值

$$\bar{\bar{x}}=\frac{\sum_{i=1}^{k}\sum_{j=1}^{r}x_{ij}}{kr}$$

分析步骤
(提出假设)

- 提出假设
  - 对行因素提出的假设为
    - H_0: $\mu_1 = \mu_2 = \ldots = \mu_i = \ldots = \mu_k$ (μ_i为第i个水平的均值)
    - H_1: μ_i (i =1,2, … , k) 不全相等
  - 对列因素提出的假设为
    - H_0: $\mu_1 = \mu_2 = \ldots = \mu_j = \ldots = \mu_r$ (μ_j为第j个水平的均值)
    - H_1: μ_j (j =1,2,…,r) 不全相等

分析步骤
(构造检验的统计量)

- 计算平方和(SS)
  - 总误差平方和 $$SST=\sum_{i=1}^{k}\sum_{j=1}^{r}\left(x_{ij}-\bar{\bar{x}}\right)^2$$
  - 行因素误差平方和 $$SSR=\sum_{i=1}^{k}\sum_{j=1}^{r}\left(\bar{x}_{i.}-\bar{\bar{x}}\right)^2$$
  - 列因素误差平方和 $$SSC=\sum_{i=1}^{k}\sum_{j=1}^{r}\left(\bar{x}_{.j}-\bar{\bar{x}}\right)^2$$
  - 随机误差项平方和 $$SSE=\sum_{i=1}^{k}\sum_{j=1}^{r}\left(x_{ij}-\bar{x}_{i.}-\bar{x}_{.j}+\bar{\bar{x}}\right)^2$$

分析步骤
(构造检验的统计量)

- ➡ 总离差平方和(*SST*)、水平项离差平方和(*SSR*和*SSC*) 、误差项离差平方和(*SSE*) 之间的关系

$$\sum_{i=1}^{k}\sum_{j=1}^{r}\left(x_{ij}-\bar{\bar{x}}\right)^2$$

$$=\sum_{i=1}^{k}\sum_{j=1}^{r}\left(\bar{x}_{i.}-\bar{\bar{x}}\right)^2+\sum_{i=1}^{k}\sum_{j=1}^{r}\left(\bar{x}_{.j}-\bar{\bar{x}}\right)^2+\sum_{i=1}^{k}\sum_{j=1}^{r}\left(x_{ij}-\bar{x}_{i.}-\bar{x}_{.j}+\bar{\bar{x}}\right)$$

SST = SSR +SSC+SSE

分析步骤
(构造检验的统计量)

- 计算均方(*MS*)
  - 误差平方和除以相应的自由度
  - 三个平方和的自由度分别是
    - 总离差平方和*SST*的自由度为 $kr-1$
    - 行因素的离差平方和*SSR*的自由度为 $k-1$
    - 列因素的离差平方和*SSC*的自由度为 $r-1$
    - 随机误差平方和*SSE*的自由度为 $(k-1)\times(r-1)$

分析步骤
(构造检验的统计量)

- 计算均方(*MS*)
  - 行因素的均方，记为*MSR*，计算公式为

$$MSR=\frac{SSR}{k-1}$$

  - 列因素的均方，记为*MSC*，计算公式为

$$MSC=\frac{SSC}{r-1}$$

  - 随机误差项的均方，记为*MSE*，计算公式为

$$MSE=\frac{SSE}{(k-1)(r-1)}$$

分析步骤
(构造检验的统计量)

- 计算检验统计量(F)
  - 检验行因素的统计量

$$F_{\mathrm{R}}=\frac{MSR}{MSE}\sim F\big(k-1,(k-1)(r-1)\big)$$

  - 检验列因素的统计量

$$F_{\mathrm{C}}=\frac{MSC}{MSE}\sim F\big(r-1,(k-1)(r-1)\big)$$

- 将统计量的值F与给定的显著性水平α的临界值F_α进行比较，作出对原假设H_0的决策
  - 若$F_{\mathrm{R}}>F_\alpha$，则拒绝原假设H_0，表明均值之间的差异是显著的，即所检验的行因素对观察值有显著影响
  - 若$F_{\mathrm{C}}>F_\alpha$，则拒绝原假设H_0，表明均值之间有显著差异，即所检验的列因素对观察值有显著影响

双因素方差分析表
(基本结构)

	A	B	C	D	E	F	G
1	误差来源	误差平方和 SS	自由度 df	均方 MS	F 值	P 值	F 临界值
2	行因素	SSR	$k-1$	MSR	F_R		
3	列因素	SSC	$r-1$	MSC	F_C		
4	误差	SSE	$(k-1)\times(r-1)$	MSE			
5	总和	SST	$kr-1$				

彩电品牌和地区的双因素方差分析

- 对品牌因素提出的假设为
  - H0： m1=m2=m3=m4 (品牌对销售量没有影响)
  - H1： mi (i =1,2, … , 4) 不全相等 (品牌对销售量有影响)

  对地区因素提出的假设为
  - H0： m1=m2=m3=m4=m5 (地区对销售量没有影响)
  - H1： mj (j =1,2,…,5) 不全相等 (地区对销售量有影响)

差异源	*SS*	df	*MS*	*F*	P-value	F crit
行(品牌)	13004.6	3	4334.85	18.1078	9.46E-05	3.4903
列(地区)	2011.7	4	502.925	2.10085	0.14367	3.2592
误差	2872.7	12	239.392			
总和	17889	19				

- F_R＝18.10777>F_α＝3.4903，拒绝原假设*H*0，说明彩电的品牌对销售量有显著影响
- F_C＝2.100846< F_α＝3.2592，不能拒绝原假设*H*0，说明销售地区对彩电的销售量没有显著影响

双因素方差分析
(关系强度的测量)

行平方和(行*SS*)度量品牌对销售量的影响效应
列平方和(列*SS*)度量地区对销售量的影响效应
两个平方和加在一起度量品牌和地区对销售量的联合效应
联合效应与总平方和的比值定义为R^2

$$R^2=\frac{\text{联合效应}}{\text{总效应}}=\frac{SSR+SSC}{SST}$$

其平方根*R*反映这两个自变量合起来与因变量之间的关系强度

$$R^2=\frac{SSR+SSC}{SST}=\frac{13004.55+2011.70}{17888.95}=0.8394=83.94\%$$

品牌因素和地区因素合起来总共解释了销售量差异的83.94%
R=0.9162，表明品牌和地区两个因素联合与销售量有较强的关系

可重复双因素分析
(例题)

- 【例】城市道路交通管理部门为研究不同的路段和不同的时间段对行车时间的影响，让一名交通警察分别在两个路段和高峰期与非高峰期亲自驾车进行试验，通过试验取得共获得20个行车时间(分钟)的数据，如下表。试分析路段、时段以及路段和时段的交互作用对行车时间的影响

	A	B	C	D
1			路段（列变量）	
2			路段1	路段2
3	时段(行变量)	高峰期	26	19
4			24	20
5			27	23
6			25	22
7			25	21
8		非高峰期	20	18
9			17	17
10			22	13
11			21	16
12			17	12

	A	B	C	D	E	F	G
13							
14	方差分析： 可重复双因素分析						
15	SUMMARY	路段1	路段2	总计			
16	1						
17	观测数	5	5	10			
18	求和	127	105	232			
19	平均	25.4	21	23.2			
20	方差	1.3	2.5	7.066667			
21	6						
22	观测数	5	5	10			
23	求和	97	76	173			
24	平均	19.4	15.2	17.3			
25	方差	5.3	6.7	10.23333			
26	总计						
27	观测数	10	10				
28	求和	224	181				
29	平均	22.4	18.1				
30	方差	12.93333	13.43333				
31	方差分析						
32	差异源	SS	df	MS	F	P-value	F crit
33	样本	174.05	1	174.05	44.06329	5.7E-06	4.493998
34	列	92.45	1	92.45	23.40506	0.000182	4.493998
35	交互	0.05	1	0.05	0.012658	0.911819	4.493998
36	内部	63.2	16	3.95			
37	总计	329.75	19				