心理统计
第十讲：方差分析(ANOVA)

严超赣

Chao-Gan Yan, Ph.D.

yancg@psych.ac.cn
http://rfmri.org/yan

Institute of Psychology, Chinese Academy of Sciences

1

1

方差分析

2

2

ANOVA的研究情境

- 一位研究者感兴趣影响儿童阅读能力的因素.研究者认为儿童的年龄和每次阅读时间可能是重要的影响因素。研究者设计了以下实验：选取三个年龄组的儿童: 3 岁, 8 岁, 和 14 岁.将每个年龄组的儿童随机分配到三个阅读条件. 组1阅读时间为5分钟; 组2为15分钟; 组3为30分钟.两个星期之后测试了这些儿童的阅读能力。

3

3 X 3 因素设计

		阅读时间		
		5 分钟	15 分钟	30 分钟
年龄	3 岁			
	8 岁			
	14 岁			

- 如何分析数据？ t-检验和 z-检验不能用于多于 2 组的数据. 处理这类数据需要用 一种新的推论统计程序: 方差分析 (Analysis of Variance, ANOVA).

4

ANOVA能够处理数据的类型

- 在上例中有两个自变量 (称为因素): 年龄和阅读时间. 两个都是组间 (独立样本) 变量. ANOVA 亦可用于分析包含组内 (重复测量) 因素的研究设计，同时包含组间和组内因素的混合设计(e.g. 假设上例中我们用同一些儿童作纵向研究。年龄是组内变量，阅读时间是组间变量).

5

单因素设计和因素设计

- 在方差分析中,因素就是自变量.
- 包含一个组间自变量的研究称为单因素设计（single-factor design）.
- 具有多于一个自变量研究称为因素设计（factorial design）.
- 构成因素的个别处理条件称为因素的水平.
- 上述研究称为因素设计, 两个组间因素,每一个因素有 3 个水平 (称为 3 X 3 组间设计).

6

1. ANOVA的逻辑

- 与假设检验的逻辑是同样的, 只是具体内容有变化
- step 1: 陈述 H_0 (和H_1 ??)，确定标准: a = ?
- step 2: ANOVA 检验总是单尾
- step 3: 指出检验的df (有两个df)
- step 4: 查表找出临界 F统计量
- step 5: 对于样本，计算F统计量
- step 6: 比较 F统计量和临界 F统计量
- step 7: 对于H_0 作出结论

7

单因素, 独立测量研究设计的例子

- 检验三个不同的学习方法的效应。将学生随机分配到3个处理组
- 方法 A：让学生只读课本, 不去上课.
- 方法 B：上课,记笔记，不读课本.
- 方法 C：不读课本，不去上课, 只看别人的笔记

8

Step 1: 陈述假设和设定标准 (选择 α)

- H_0: $\mu_1 = \mu_2 = \mu_3$
- H_1: 其中一个组与另一个（或更多）的组均值不同。备择假设可能的形式很多：
- μ_1不等于 $\mu_2 = \mu_3$
  $\mu_1 = \mu_3$ 不等于 μ_2
  $\mu_1 = \mu_2$ 不等于 μ_3
  μ_1 不等于 μ_2 不等于 μ_3
- 因此，只需给出虚无假设就够了

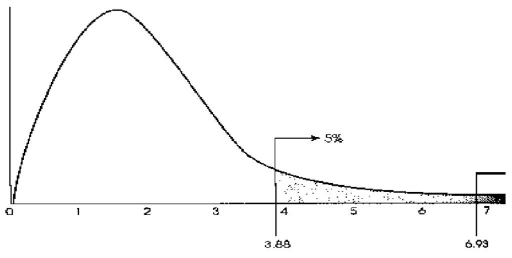
9

Step 2-3

- step 2: ANOVA 检验总是单尾. 因为不存在负的方差. F分布表也只有单侧的Alpha.（F分布图）
- step 3: 找出检验的 df. 注意要考虑2个 df

10

F分布



11

step 4: 从表找出临界 F统计量

表 12.2　F 临界值表样例

分母的 df	α	分子的 df					
		1	2	3	4	5	…
1	0.05	161	200	216	225	230	…
	0.01	4052	4999	5403	5625	5764	
2	0.05	18.51	19.00	19.16	19.25	19.30	…
	0.01	98.49	99.00	99.17	99.25	99.30	
3	0.05	10.13	9.55	9.28	9.12	9.01	…
	0.01	34.12	30.92	29.46	28.71	28.24	
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

- 与 t分布表类似, F分布表也是描述一族 F分布.
- 需要用到两个df,用一个找出正确的行另一个找出正确的列.上面一行对应于 α = 0.05, 下面一行对应于 α = .01.

12

step 5: 计算样本的F统计量观测值

- 两个独立样本的t检验

  $t_{obs} = \dfrac{\text{得到的样本均值间差异}}{\text{期望的机会差异}}$

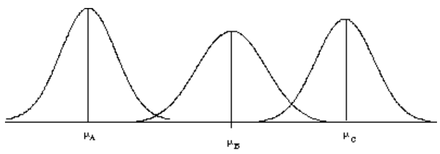
- 对于 ANOVA检验统计量（称为 F比率）类似

  $F = \dfrac{\text{样本均值间方差（差异）}}{\text{期望的误差方差}}$

13

为什么用方差?

- 因为有多于两个组.



- 如何计算一个分数来描述差异间分布? 差异不能够分割, 但是方差能够分割。这就是ANOVA -方差分析名字的由来

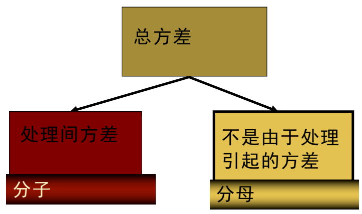
14

方差的来源

- 什么造成样本之间的变异？
  - 处理/组效应 - 处理造成的差异
  - 个体差异效应 - 个体差异变异
  - 随机误差
- 每一个处理内部的变异 (处理内变异)
  - 个体差异效应
  - 随机误差

15

方差的分解



16

方差的分解

方差(Variance)

$$s^2 = \frac{\sum(X_i - \bar{X})^2}{df}$$

s²: 方差、均方差，也可表示为MS

$$\longrightarrow \quad s^2 = \frac{SS}{df}$$

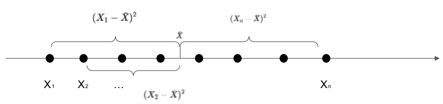
$$SS = \sum(X_i - \bar{X})^2$$

SS: 离差平方和

17

方差的分解

$$SS = \sum(X_i - \bar{X})^2$$



18

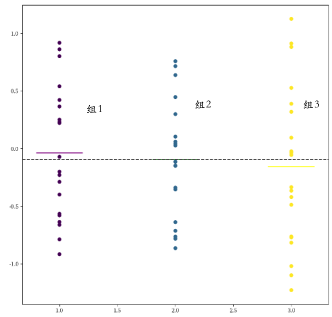
方差的分解

方差的可分解性

方差（或变异）的可分解性是指总的离差平方和可以分解为几个不同来源的平方和。

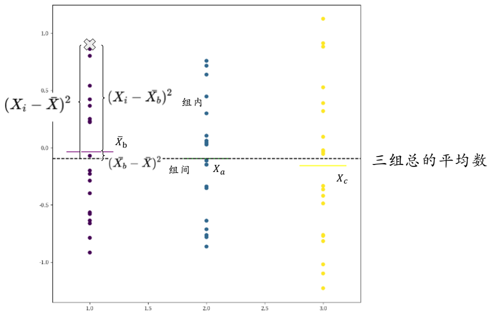
19

方差的分解



20

方差的分解



21

方差的分解

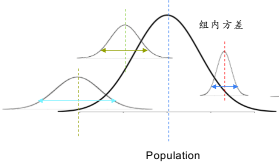
方差的可分解性

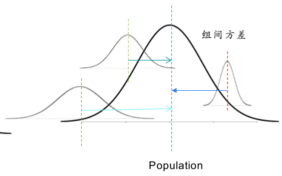
方差（或变异）的可分解性是指总的离差平方和可以分解为几个不同来源的平方和。

总平方和可以分解为组内平方和和组间平方和。

22

方差的分解





23

方差的分解

$$\sum(X_i-\bar{X})^2=\sum(X_i-\bar{X}_b)^2+\sum(\bar{X}_b-\bar{X})^2$$

（$\bar{X}$ 总平均值，$\bar{X}_b$ 组平均值）

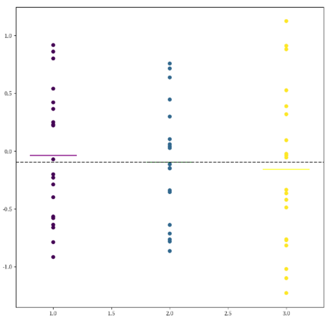
↓

$$SS_T = SS_W + SS_B$$

(SS_T 总平方和，SS_B 组间平方和，SS_W 组内平方和)

24

方差的分解



若想要对三个地区人群的主观幸福感进行比较，我们可以怎么样进行处理呢?

H0: 三个地区人群的主观幸福感没有差异

H1: 三个地区人群的主观幸福感存在差异

25

方差的分解

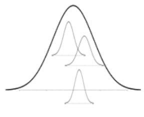
假定H0为真，则各组之间不存在差异 $\bar{X}_b \longrightarrow \bar{X}$

$$\sum(X_i-\bar{X})^2=\sum(X_i-\bar{X}_b)^2+\sum(\bar{X}_b-\bar{X})^2$$

$\downarrow$ $\sum(X_i-\bar{X})^2$ $\quad\downarrow$ 0

$$SS_T = SS_W + SS_B$$

$\downarrow$ SS_T $\quad\downarrow$ 0



26

方差的分解

直接使用离差平方和来进行比较会受到自由度大小的影响，故取方差(均方)来进行比较。

假定H0为真，则各组之间不存在差异 $\bar{X}_b \longrightarrow \bar{X}$

$$\sum(X_i-\bar{X})^2=\sum(X_i-\bar{X}_b)^2+\sum(\bar{X}_b-\bar{X})^2$$

$\downarrow$ $\sum(X_i-\bar{X})^2$ $\quad\downarrow$ 0

$$SS_T = SS_W + SS_B$$

$\downarrow$ SS_T $\quad\downarrow$ 0

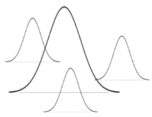
$$s^2=\frac{SS}{df}$$

$$\longrightarrow \frac{MS_B}{MS_W} \longrightarrow 0$$

27

方差的分解

假定H1为真，则各组之间存在差异



$$SS_B \gg SS_W$$

$$\downarrow$$

$$\frac{MS_B}{MS_W} \longrightarrow +\infty$$

28

方差的分解

假定H1为真，则各组之间存在差异

$$SS_B \gg SS_W$$

$$\downarrow$$

$$\boxed{\frac{MS_B}{MS_W}} \longrightarrow +\infty$$

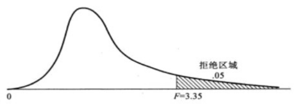
29

F比率

*F*值

$$F=\frac{MS_B}{MS_W}$$

当H0为真时，F值服从一个广为人知的概率分布。为了纪念R.A.Fisher，这个分布称为F分布。

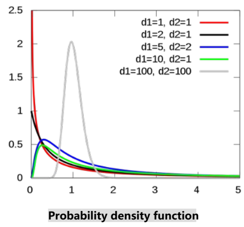


(F分布，自由度为(2,27))

30

F比率

F分布



前提条件：

1.总体正态

观测值来自正态分布的总体。

2.变异的同质性

$\sigma_1^2 = \sigma_2^2 = \ldots = \sigma_k^2$

3.独立性

31

step 6: 比较 F统计量的观测值与临界 F统计量

- 如果 F统计量的观测值 (F_{obs}) 在统计上显著，则拒绝 H_0

32

2. ANOVA的专用符号

- K = 处理条件(或组)的数目
- n = 每一个组的数目(如果它们相等)
- n_i = 第i组的数目(如果 它们不等)
- N = Sn_i = 总的样本容量
- $T_i = SX_{ij}$
- G = ST_i =总的和
- G-bar = G / N = 总的均值
- SS_i = 每一个组的和方 = $S(X_{ij} - {}_i)^2$

33

例 12.2 为了检验三种不同学习方法的效应，将学生随机分配到 3 个处理组。方法 A：让学生只读课本，不去上课；方法 B：学生上课、记笔记，但不读课本；方法 C：学生不读课本、不去上课，只看别人的笔记。经过一段时间后，对学习效果进行测量，得到结果如表 12.5。请问各方法之间是否有差异？（用 $\alpha=0.05$ 的显著性水平）

表 12.5 不同学习方法的效果

研究方法		
方法 A 只读课本	方法 B 只记笔记	方法 C 只看别人的笔记
0	4	1
1	3	2
3	6	2
1	3	0
0	4	0
$T_1=5$	$T_2=20$	$T_3=5$
$SS_1=6$	$SS_2=6$	$SS_3=4$
$n_1=5$	$n_2=5$	$n_3=5$
$\bar{X}_1=1$	$\bar{X}_2=4$	$\bar{X}_3=1$

34

原始数据整理 （1）

例 12.2 为了检验三种不同学习方法的效应，将学生随机分配到 3 个处理组。方法 A：让学生只读课本，不去上课；方法 B：学生上课、记笔记，但不读课本；方法 C：学生不读课本、不去上课，只看别人的笔记。经过一段时间后，对学习效果进行测量，得到结果如表 12.5。请问各方法之间是否有差异？（用 $\alpha=0.05$ 的显著性水平）

表 12.5 不同学习方法的效果

研究方法		
方法 A 只读课本	方法 B 只记笔记	方法 C 只看别人的笔记
0	4	1
1	3	2
3	6	2
1	3	0
0	4	0
$T_1=5$	$T_2=20$	$T_3=5$
$SS_1=6$	$SS_2=6$	$SS_3=4$
$n_1=5$	$n_2=5$	$n_3=5$
$\bar{X}_1=1$	$\bar{X}_2=4$	$\bar{X}_3=1$

35

原始数据整理 （2）

- $\Sigma X^2 = 106$
  G = 30 = 总的和
  N = 15 = 总的样本容量
  G-bar = 30/15 = 2 = 总的均值
  K = 3 =处理条件 (或组) 的数目

36

方差的分解

- F比率 = $\frac{\text{处理间方差}}{\text{处理内方差}}$
  - 需要找出两个方差.
  - 最基本公式s^2 = SS/df.
  - $SS_{和} = \Sigma X^2 - (G^2/N)$
  - $SS_{和} = 106 - (30^2/15) = 106 - 60 = 46$
  - 需要将其分解为组间变异和组内变异.
  - $SS_{和} = SS_{组间} + SS_{组内}$

$$SS = \sum (X - X)^2$$

$$SS = \sum X^2 - \frac{(\sum X)^2}{n}$$

37

组内方差和组间方差的计算

- 如何得到$SS_{组内}$?
  - 将每一个组SS相加
  - $SS_{within} = \Sigma SS_{每一个处理内部} = \Sigma SS_i$
  - = 6 + 6 + 4 = 16
- 如何得到$SS_{组间}$?
  - 快捷的方法是: $SS_{和}$- $SS_{组内}$
- 若数据足够，不推荐用这种方法,因为:
  - 无法检查计算错误
  - 未涉及$SS_{组间}$ 是如何组成.

38

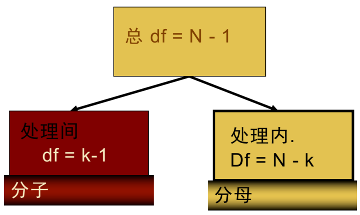
计算 $SS_{组间}$的两个公式

- 定义公式和计算公式

定义公式	计算公式
SS [illegible] = $\Sigma[n_i(\bar{X}$ - G-bar$)^2]$	SS [illegible] = $\Sigma(T^2/n_i) - G^2/N$
$= 5(1-2)^2 + 5(4-2)^2 + 5(1-2)^2$ = 5 + 20 + 5 = 30	$= 5^2/5 + 20^2/5 + 5^2/5 - 30^2/15$ = 5 + 80 + 5 − 60 = 30

39

自由度的分割



40

自由度

- 已计算出SS, 找出 df:
  - 共有两个 (或三个) 自由度, 一个组间方差df,一个组内方差df (以及一个总的 df).
  - $df_{和}$ = N - 1
  - $df_{组间}$= K - 1
  - $df_{组内}$ = (N-1) - (K-1) = N - K
  - $df_{和} = df_{组内} + df_{组间}$
- 在例子中:
  - $df_{组内}$ = 15 - 3 = 12
  - $df_{组间}$= 3 - 1 = 2
  - $df_{和}$ = 15 - 1 = 14, = 12 + 2

41

计算均方和F比率

- 方差 = 均方 = MS = SS/df
  - $MS_{组间} = SS_{组间}/df_{组间}$
  - --> 上例中 = 30/2 = 15
- $MS_{组内} = MS_{误差}$ =误差的均方 = $SS_{组内}/df_{组内}$
  - --> 上例中 = 16/12 = 1.33
- F比率 = $\frac{\text{处理间方差}}{\text{处理内方差}} = \frac{MS_{组间}}{MS_{组内}}$
  - 上例中的F比率是: 15/1.33 = 11.28

42

方差分析表

- 将结果总结到方差分析表中

表 12.6 例 12.2 的方差分析表

来源	SS	df	MS	F	η_p^2
组间	30	2	15		
组内	16	12	1.333	11.25	0.65
总和	46	14			

43

对假设作出结论

- 查 F表 确定 F_{crit} 对假设作出结论
  - $df_{组间}$ = 分子的df
  - $df_{组内}$ = 分母的df (误差)

  --> 上例中:
  - $df_{组内} = 12$; $df_{组间} = 2$
  - 如果选择 $\alpha = .05$, $F_{crit} = 3.88$
  - 如果选择 $\alpha = .01$, $F_{crit} = 6.93$
- F比率的观测值11.28大于 F_{crit}., 所以拒绝 H_0 ($\mu_1 = \mu_2 = \mu_3$).

44

报告结果

- $F(df_{组间}, df_{组内})$ = Fobs, p < ?
  - "单因素方差分析发现学习方法有显著的效应, $F(2,12) = 11.28$, $p < 0.01$.

45

Alpha的预设

- 电脑的结果输出会给出实际 p值. 假设检验的逻辑是必须事前预设Alpha水平. 如果选择了.01, 就必须将其用于所有的检验. 所以,如果有两个实验, 电脑程序得到实验1的 p值 = .001,实验2的p值 = .01. 它们都在统计上显著. 假设检验是 yes/no 决策. 上例中,结论都是 YES. 实验1的结果并不比实验 2"更显著".

46

4. 事后检验（Post hoc tests）

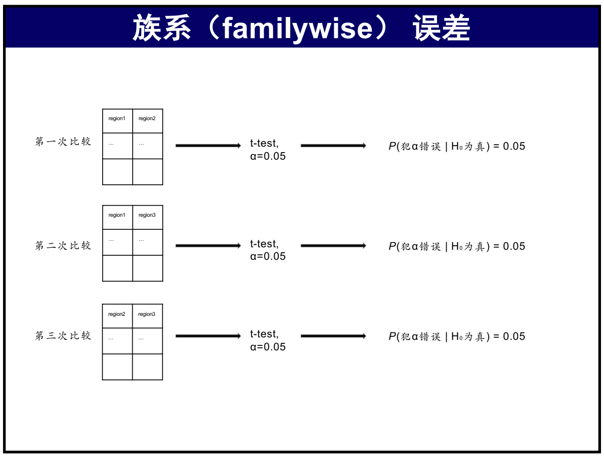
- ANOVA 的结果是检验H_0: $\mu_1 = \mu_2 = \mu_3$,这是一个两点 (拒绝/不拒绝) 决策. 并未提供哪个备择假设得到支持. 也就是说, 只知道一些组与其它组不同, 但并知道差别在哪些组之间.
- 所以从ANOVA得到显著差异的结果 (拒绝H_0)后,一定要做作一些事后检验. 事后检验使我们能够比较各组, 发现差异产生在什么地方.
- 事后检验就是比较每一个处理组与另一个处理组, 一次比较两个. 这称为成对比较.

47

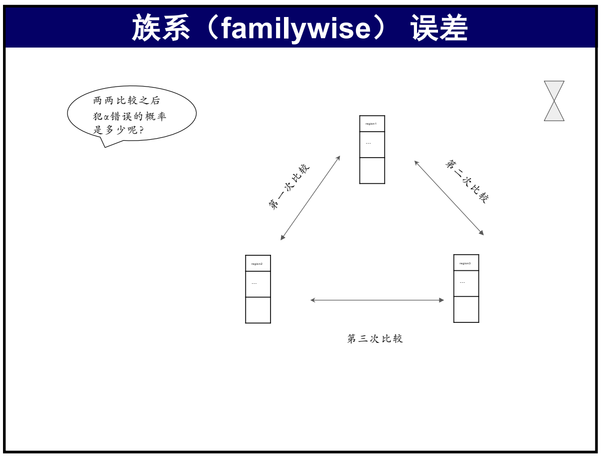
族系（familywise） 误差

- 在上例中, 可以比较 m_1 与 m_2, m_1与 m_3, 以及 m_2与 m_3. 这样的做法有没有问题？
- 每一个比较 都是一个单独的假设检验, 每一个都有犯I类错误的风险. 所以,比较对数越多, 作结论的风险越大。即容易发现实际不存在的差异。 这称为实验导致的（experimentwise）alpha 水平 或族系（familywise） 误差
- $\alpha_{EW} = 1 - (1 - \alpha)^c$ c = 比较对数
- 对于上述例子, 如果选择 α = 0.05 作3 对比较
- $\alpha_{EW} = 1 - (1 - \alpha)^c = 1 - (.95)^3 = 1 - .857 = .143$
- I类错误的机会增加到14.3%而不再是5%，多数事后检验设计中都控制了实验导致误差.

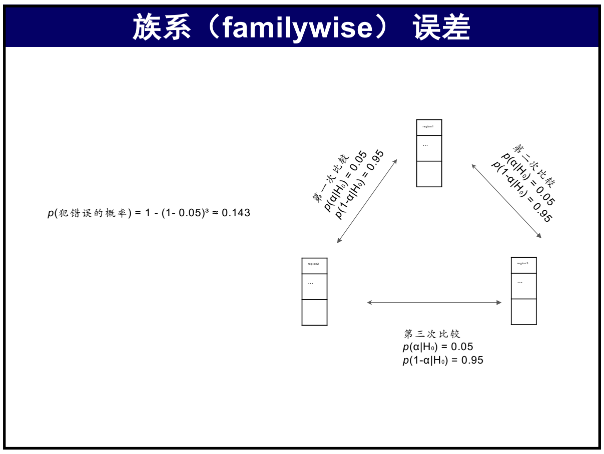
48



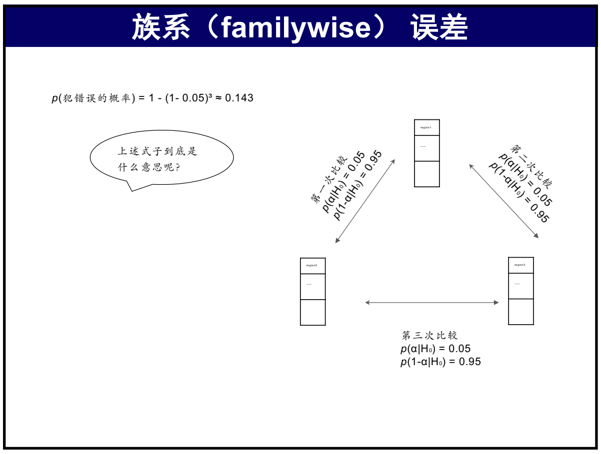
49



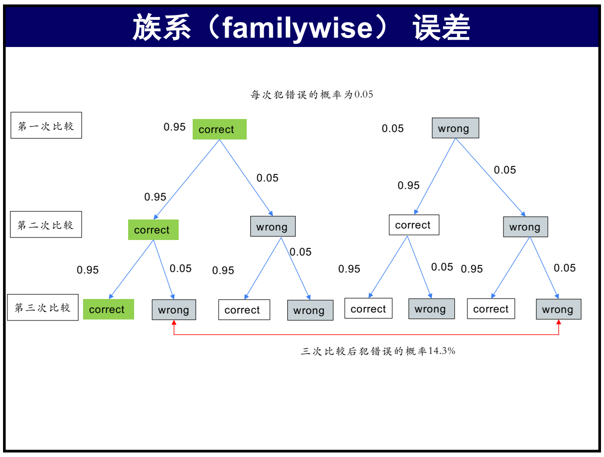
50



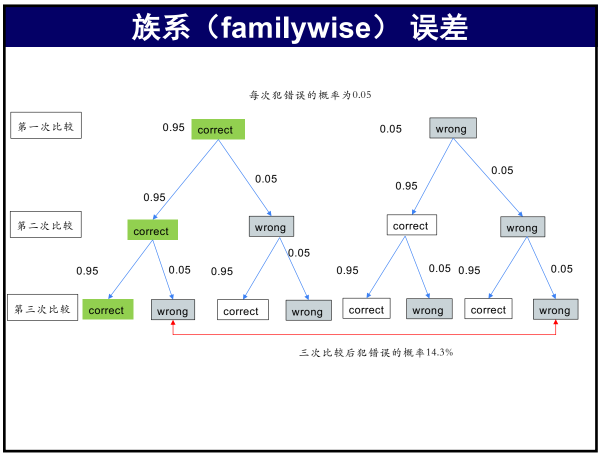
51



52

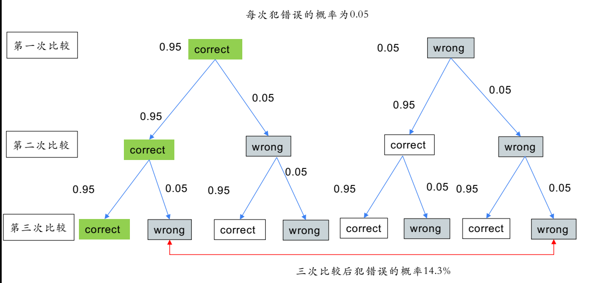


53



54

族系（familywise） 误差



55

Tukey's HSD 检验

- 此检验要求各组有相等的样本容量.
- 可以计算出单一的值确定处理均值间的最小差异，考查此差异在统计上是否显著.
- HSD = q * sqrt（$MS_{组内}$/n）
- q 值可以从表中查出(附表5). 需要用到K和 $df_{组内}$, 以及α_{EW}

56

Tukey's HSD 检验

在上例中 (用α_{EW} = .05):

HSD = q * sqrt（$MS_{组内}$/n）=(3.77) sqrt（1.33/5） = (3.77)(.516) = 1.94

比较 1: H_0: $m_1 = m_2$

2 -1 = 4.0 - 1.0 = 3.0

HSD = 1.94 < 3.0，拒绝 H_0

比较 2: H_0: $m_1 = m_3$

3 -1 = 1.0 - 1.0 = 0.0

HSD = 1.94 > 0.0，不能 拒绝 H_0

比较 3: H_0: $m_2 = m_3$

2 -3 = 4.0 - 1.0 = 3.0

HSD = 1.94 < 3.0， 拒绝 H_0

所以 B 与 A 和 C不同，而A 与 C 没有差异

57

Scheffe检验

- 特别适用于n 不等的情况
- 这是保守的检验(降低 I类错误的风险, 但增加II类错误的风险).
- 用F比率检验差异. 重新计算两项的 $MS_{组间}$, 每次只检验一个比较.注意:用整体的 $df_{组间}$ 和整体的$MS_{组内}$.

58

Scheffe检验

- $T_1 = 5$，$T_2 = 20$，$T_3 = 5$
  $SS_1 = 6$，$SS_2 = 6$，$SS_3 = 4$
  $n_1 = 5$，$n_2 = 5$，$n_3 = 5$
- 比较 1: H_0: $\mu_1 = \mu_2$
  - $SS_{组间} = 5^2/5+20^2/5-25^2/10 = 22.5$
  - $MS_{组间} = 22.5/2 = 11.25$
  - $MS_{组内} = 16/12 = 1.33$ （不变）
  - F比率 = $MS_{组间}$ /$MS_{组内}$ = 11.25/1.33 = 8.46
  - 查 F表. α = .05, $F_{crit}(2,12) = 3.88$
  - 8.46 > 3.88, 拒绝 H_0

59

Scheffe检验

- 比较 2: H_0: $m_1 = m_3$
  - $SS_{组间} = 5^2/5 +5^2/5 - 10^2/10 = 0$
  - $MS_{组间}$ = 0/2 = 0
  - $MS_{组内}$ = 16/12 = 1.33
  - F比率 = $MS_{组间}$ /$MS_{组内}$ = 0/1.33 = 0
  - 0 < 3.88, 不能拒绝 H_0
- 比较 3: H_0: $m_2 = m_3$
  - $SS_{组间} = 5^2/5+20^2/5-25^2/10 = 22.5$
  - $MS_{组间}$ = 22.5/2 = 11.25
  - $MS_{组内}$ = 16/12 = 1.33
  - F比率 = $MS_{组间}$ /$MS_{组内}$ = 11.25/1.33 = 8.46
  - 查 F表. a = .05, $F_{crit}(2,12) = 3.88$
  - 8.46 > 3.88, 拒绝 H_0

60

方差分析与t-检验的关系

- 差异间独立样本 t-检验与两个水平的单因素组间 ANOVA有何区别?
- 没有. F比率 = t^2
- 差异间t-检验和 ANOVA, t-检验是考察两个均值间的差异, ANOVA 是考察方差. 如果只有两个组, t 统计量的平方就是F统计量

61

例2:

- 一位研究者研究三种键盘设计。记录了三组被试的错误次数:
  键盘A: 0 4 0 1 0
  键盘B: 6 8 5 4 2
  键盘C: 6 5 9 4 6
  键盘类型对打字错误有无显著的影响?

62

Step 1. 陈述 H_0 和 H_1; 确定显著性标准: $\alpha = .05$

Step 2. 查表求临界 F 值

Df 组间=k-1=3-1=2; Df 组内=n-k=15-3=12

F crit (2, 12) = 3.88

Step 3:作方差分析,计算 F 值

1)计算每组的 T 和 SS; G 和 Σx^2

2)作方差分析表

来源	SS	df	MS	
处理间	70	2	35	
处理内	46	12	3.83	F=9.14
总和	116	14		

Step 4:

对 H_0 作出结论

因为观察到的 F 统计量9.14大于临界 F 3.88, 所以我们有理由拒绝 H_0. 键盘类型对打字错误有显著的影响

63

- Step 5 事后检验
  - HSD = q * sqrt（$MS_{组内}$/n）=3.77*sqrt (3.83/5) = 3.30
  - 比较 1:A-B
  - 5-1=4>3.3 所以A与B差异显著
  - 比较 1:A-C
  - 6-1=5>3.3 所以A与C差异显著
  - 比较 1:B-C
  - 6-5=1<3.3 所以B与C差异不显著

64

Effect Size

$$\eta^2 = \frac{SS_{组间}}{SS_{组间} + SS_{组内}}$$

$\eta^2 = r^2$

r^2 的大小	效应的评估
$0.01 < r^2 < 0.09$	小的效应
$0.09 < r^2 < 0.25$	中等效应
$r^2 > 0.25$	大的效应

65

作业

1. 下列方差分析表总结了四种条件下的实验结果。在每种实验条件中包含n=10的样本，填充表的值。

来源	SS	df	MS	
处理间			15	F=
处理内	108			
总和				

66

66

作业

2. 以下实验数据是2岁、6岁和10岁儿童瞬时记忆的结果。用α=.05的标准来检验年龄组间平均数的差异。

2岁	n=20	X=2.1	s=1.3
6岁	n=20	X=4.3	s=1.5
10岁	n=20	X=6.9	s=1.8

67

67

作业

3. 一位研究者考察身体吸引力对人们对人知觉和判断的影响。将应聘者分为3组：高吸引力组、中等吸引力组和低吸引力组，每组n=12。以下分数是人事经理对三组被试的整体评价。

a) 计算三组的均值并画图表示

b) 在α=.05的水平上确定3组间有无差异。

高吸引力			中等吸引力			低吸引力		
5	4	4	6	5	3	4	3	1
3	5	6	6	6	7	3	1	2
4	3	8	5	4	6	2	4	3
3	5	4	8	7	8	2	1	2

68

68