第八章 方差分析

第一节｜方差分析的基本思想（ANOVA）

> **一句话抓住本质**
> **方差分析不是在“算方差”，而是在问：**
> **👉「组与组之间的差异，是否大到超过随机波动本身？」**

一、方差分析要解决什么问题？（Purpose）

> **当比较对象 ≥ 2 组时：**

- 不能再一组一组做 t 检验（Ⅰ类错误会爆炸）
- 需要一个**统一的、整体的差异判断**

📌 **ANOVA 的问题表述是：**

> **多个样本均数看起来不一样，这种“不一样”，**
> **是真实处理效应，还是随机波动造成的？**

二、核心直觉：一切差异都来自“波动”（Intuition）

Fisher 的洞察非常简单，也非常狠：

> **只要有数据，就一定有波动（variation）。**
> **问题不在“有没有波动”，而在“波动从哪来”。**

因此，ANOVA 的第一步不是算均数，而是：

👉 **把“总波动”拆开**

三、变异的两大来源（Key Concept）

① 组间变异（Between-group variation）

SS 组间（Systematic variation）

- 不同处理组 **均数之间** 的差异
- 直觉理解：

> **不同处理 / 不同条件，**
> **是否真的“把均值拉开了”？**

🧠 例如：

- 不同降压药 → 平均降压效果不同
- 不同教学方法 → 平均成绩不同

📌 **这部分是：**

- **我们“希望看到”的差异**
- 包含：
    - 系统因素（处理效应）
        -
            - 不可避免的随机因素

---

② 组内变异（Within-group variation）

SS 组内（Random variation）

- **同一组内部** 个体之间的差异
- 直觉理解：

> **就算给的是同一种处理，**
> **个体也不可能完全一样**

🧠 例如：

- 同一药物下，不同人降压程度不同
- 同一课堂下，学生成绩仍有高低

📌 **这部分是：**

- **纯随机噪声**
- 是判断“差异是否值得相信”的**背景噪音**

四、总变异的分解（The Big Picture）

ANOVA 的灵魂公式（不背形式，背逻辑）：

$$\text{总变异} = \text{组间变异} + \text{组内变异}$$

- **SS 总 = SS 组间 + SS 组内**
- 所有统计推断都建立在这个分解上

📌 **你可以把它理解为：**

> 把世界拆成
> 「信号（signal）」和「噪声（noise）」

五、F 值在干什么？（What is F really testing）

F 的直觉定义：

$$F = \frac{\text{组间变异的平均大小}}{\text{组内变异的平均大小}}$$

> 👉 **如果“组与组之间的差异”，**
> **明显大于“组内本来就有的随机波动”，**
> **那我们才相信：均数真的不一样。**

三种直觉情形（非常重要）

1️⃣ F ≈ 1

- 组间差异 ≈ 随机波动
- 👉 看起来不同，其实是噪声

2️⃣ F > 1（但不大）

- 有点差异，但不够稳
- 👉 不能下结论

3️⃣ F ≫ 1

- 组间差异远大于组内噪声
- 👉 **拒绝 H_0：总体均数不全相等**

六、一句话总结（Exam-ready）

> **将总变异分解为组间变异和组内变异，**
> **通过比较二者的相对大小（F 值），**
> **判断多个总体均数是否存在显著差异。**

第二节｜完全随机设计方差分析

> **一句话总览**
> 完全随机设计的方差分析，本质是：
> **把“所有观测值的波动”拆成两块，然后看哪一块更大。**

表 8-1　完全随机设计方差分析的数据结构

	处理因素						合计
	水平 1	水平 2	…	水平 i	…	水平 k	
	X_{11}	X_{21}	…	X_{i1}	…	X_{k1}	
	X_{12}	X_{22}	…	X_{i2}	…	X_{k2}	
	…	…	…	…	…	…	
	X_{1j}	X_{2j}	…	X_{ij}	…	X_{kj}	

(李康,贺佳主编, 2024, p. 78) - Please click Zotero - Refresh in Word/LibreOffice to update all fields.

续表

	处理因素						合计
	水平 1	水平 2	…	水平 i	…	水平 k	
	…	…	…	…	…	…	
	X_{1n}	X_{2n}	…	X_{in}	…	X_{kn}	
n_i	n_1	n_2	…	n_i	…	n_k	n
$\bar{X}_i$	$\bar{X}_1$	$\bar{X}_2$	…	$\bar{X}_i$	…	$\bar{X}_k$	$\bar{X}$
S_i^2	S_1^2	S_2^2	…	S_i^2	…	S_k^2	S^2

(李康,贺佳主编, 2024, p. 79) - Please click Zotero - Refresh in Word/LibreOffice to update all fields.

一、完全随机设计在干什么？（Purpose）

设计层面的问题：

- 只有 **一个处理因素**（single factor）
- 这个因素有 **k 个水平**
- 实验对象 **完全随机** 分配到各处理组

- 比较目标：
  👉 **不同处理水平的总体均值是否相同**

📌 对应的统计假设是：

- $H_0: \mu_1 = \mu_2 = \cdots = \mu_k$
- H_1: 至少有一个总体均值不同

二、最重要的直觉：为什么要“拆方差”？

你这整段推导，其实只在做一件事：

> **解释“为什么数据不一样”。**

任何不一样，只有两种来源：

1️⃣ **组与组之间不一样**（处理因素可能有用）
2️⃣ **同一组内部本来就不一样**（个体随机性）

于是：

> **总变异 = 组间变异 + 组内变异**

这不是技巧，是**世界观。**

三、三个平方和到底在说什么？（Key Concepts）

① 总变异$SS_{总}$

$$SS_{总} = \sum_{i=1}^{k} \sum_{j=1}^{n_i} (X_{ij} - \bar{X})^2$$

直觉解释：

所有观测值，
离“总体均值”有多散？

这是**世界的全部波动。**

---

② 组间变异 $SS_{组间}$

$$SS_{组间} = \sum_{i=1}^{k} n_i(\bar{X}_i - \bar{X})^2$$

直觉解释：

各组“均值”
离总体均值有多远？

- 衡量的是：**处理因素把均值拉开了多少**
- 包含：
    - 真正的处理效应（signal）
        -
        - 抽样带来的随机误差

---

③ 组内变异$SS_{组内}$

$$SS_{组内} = \sum_{i=1}^{k} \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)^2$$

直觉解释：

在同一处理条件下，
个体之间本来就有多乱？

- 这是**纯随机噪声**
- 是判断“差异是否可信”的背景水平

四、那一大坨代数推导在干嘛？

你给出的平方展开：

$$(X_{ij} - \bar{X}) = (X_{ij} - \bar{X}_i) + (\bar{X}_i - \bar{X})$$

一句话翻译：

> 一个观测值离总体均值的偏差
>
> ＝
>
> 它离“本组均值”的偏差
> \+
> 本组均值离“总体均值”的偏差

平方、求和、交叉项为 0
👉 **自然得到**

$$SS_{总} = SS_{组间} + SS_{组内}$$

📌 这不是巧合，是结构必然。

五、自由度为什么这样分？（Very important）

总自由度

$$v_{总} = n - 1$$

👉 一个总体均值被估计掉

组间自由度

$$v_{组间} = k - 1$$

👉 k 个组均值，受 1 个总体均值约束

组内自由度

$$v_{\text{组内}} = n - k$$

👉 每组估计 1 个均值，共用掉 k 个自由度

📌 **自由度也满足同样的分解逻辑：**

$$v_{\text{总}} = v_{\text{组间}} + v_{\text{组内}}$$

六、均方 MS 的真正含义（别只当公式）

组间均方

$$MS_{\text{组间}} = \frac{SS_{\text{组间}}}{k-1}$$

👉 **单位自由度上的“处理 + 随机”变异**

组内均方

$$MS_{\text{组内}} = \frac{SS_{\text{组内}}}{n-k}$$

👉 **单位自由度上的纯随机噪声**

📌 在 ANOVA 里：

$$MS_{\text{组内}}$$

= **随机误差方差的估计量**

七、F 值到底在比什么？ (The soul)

$$F = \frac{MS_{组间}}{MS_{组内}}$$

一句话直觉：

> **“处理造成的差异”，**
> **是否明显大于“世界本来就有的乱”？**

三种结论直觉

- **F ≈ 1**
  👉 组间 ≈ 随机噪声
  👉 没证据认为处理有效
- **F > 1 且显著**
  👉 处理效应超过随机波动
  👉 拒绝 H_0

表 8-2 完全随机设计的方差分析表

变异来源	平方和（SS）	自由度（ν）	均方（MS）	F 值
总变异	$SS_{总} = \sum_{i=1}^{k}\sum_{j=1}^{n_i}(X_{ij} - \bar{X})^2 = (n-1)S^2$	$\nu_{总} = n-1$		
处理组间	$SS_{组间} = \sum_{i=1}^{k} n_i(\bar{X}_i - \bar{X})^2$	$\nu_{组间} = k-1$	$MS_{组间} = \frac{SS_{组间}}{\nu_{组间}}$	$F = \frac{MS_{组间}}{MS_{组内}}$
组内（误差）	$SS_{组内} = SS_{总} - SS_{组间}$	$\nu_{组内} = \nu_{总} - \nu_{组间}$	$MS_{组内} = \frac{SS_{组内}}{\nu_{组内}}$	

注：方差分析表中自由度 ν 常以 DF 表示。

(李康,贺佳主编, 2024, p. 80) - Please click Zotero - Refresh in Word/LibreOffice to update all fields.

【例 8-1】在评价某药物耐受性及安全性的I期临床试验中，将符合纳入标准的 30 名健康志愿者随机分为 3 组，每组 10 名，各组注射剂量分别为 0.5U、1U 和 2U，观察 48 小时部分凝血活酶时间（s），结果见表 8-3。试问不同剂量组的部分凝血活酶时间有无不同？

表 8-3　三种不同剂量组 48 小时部分凝血活酶时间　　时间：s

	0.5U	1U	2U	合计
	36.8	40.0	32.9	
	34.4	35.5	37.9	
	34.3	36.7	30.5	
	35.7	39.3	31.1	
	33.2	40.1	34.7	
	31.1	36.8	37.6	
	34.3	33.4	40.2	
	29.8	38.3	38.1	
	35.4	38.4	32.4	
	31.2	39.8	35.6	
n_i	10	10	10	30（n）
$\bar{X}_i$	33.62	37.83	35.10	35.516 7（$\bar{X}$）
S_i	2.263 6	2.207 1	3.313 3	3.107 2（S）

(李康,贺佳主编, 2024, p. 80) - Please click Zotero - Refresh in Word/LibreOffice to update all fields.

第三节 随机区组设计的方差分析

方差分析具体步骤：

（1）提出检验假设，确定检验水准

H_0: $\mu_1=\mu_2=\mu_3$，即三个组部分凝血活酶时间的总体均数相同

H_1: μ_1, μ_2, μ_3 不全相同，即三个组部分凝血活酶时间的总体均数不全相同

$\alpha=0.05$

（2）计算检验统计量 F 值

$$SS_{总}=(n-1)S^2=(30-1)\times 3.107\,2^2=279.986\,1$$

$$\nu_{总}=30-1=29$$

$$SS_{组间}=\sum_{i=1}^{k} n_i(\bar{X}_i-\bar{X})^2$$

$$=10\times(33.62-35.516\,7)^2+10\times(37.83-35.516\,7)^2+10\times(35.10-35.516\,7)^2$$

$$=91.224\,7$$

$$\nu_{组间}=3-1=2$$

(李康,贺佳主编, 2024, p. 80) - Please click Zotero - Refresh in Word/LibreOffice to update all fields.

$$MS_{组间}=\frac{SS_{组间}}{\nu_{组间}}=\frac{91.224\,7}{2}=45.612\,4$$

$$SS_{组内}=SS_{总}-SS_{组间}=279.986\,1-91.224\,7=188.761\,4$$

$$\nu_{组内}=\nu_{总}-\nu_{组间}=29-2=27$$

$$MS_{组内}=\frac{SS_{组内}}{\nu_{组内}}=\frac{188.761\,4}{27}=6.991\,2$$

$$F=\frac{MS_{组间}}{MS_{组内}}=\frac{45.612\,4}{6.991\,2}=6.52$$

将上述计算结果列于表 8-4 的方差分析表中。

(李康,贺佳主编, 2024, p. 81) - Please click Zotero - Refresh in Word/LibreOffice to update all fields.

表 8-4 完全随机设计的方差分析表

变异来源	*SS*	*DF*	*MS*	*F*	*P*
总变异	279.986 1	29			
处理组间	91.224 7	2	45.612 4	6.52	<0.05
组内（误差）	188.761 4	27	6.991 2		

（3）确定 P 值，做出推断结论

分子自由度 $\nu_{组间}=2$，分母自由度 $\nu_{组内}=27$，查附表 4F 界值表（方差分析用），$F_{0.05(2,27)}=3.35$。由于 $F=6.52$，$F>F_{0.05(2,27)}$，从而 $P<0.05$，按照 $\alpha=0.05$ 的检验水准拒绝 H_0，可以认为三种不同剂量 48 小时部分凝血活酶时间不全相同。

(李康,贺佳主编, 2024, p. 81) - Please click Zotero - Refresh in Word/LibreOffice to update all fields.

第三节｜随机区组设计的方差分析

一句话抓住本质
随机区组设计不是多加一个因素，
而是：**把“已知会造成差异的背景噪声，先扣掉”。**

表 8-5 随机区组设计方差分析的数据结构

区组（B）	处理因素（A）			
	水平 1	水平 2	…	水平 k
区组 1	X_{11}	X_{21}	…	X_{k1}
区组 2	X_{12}	X_{22}	…	X_{k2}
…	…	…	…	…
区组 m	X_{1m}	X_{2m}	…	X_{km}

(李康,贺佳主编, 2024, p. 81) - Please click Zotero - Refresh in Word/LibreOffice to update all fields.

一、为什么要用随机区组设计？（Purpose）

在完全随机设计（CRD）里，我们默认：

> 所有实验对象 **本质相同**，
> 差异主要来自处理因素 + 随机波动。

但现实里经常不是这样。

典型问题场景

- 人与人差异很大（性别、体重、基础值）
- 批次不同（天、地块、动物窝）
- 时间不同（前后顺序）

📌 **这些差异：**

- 不是你关心的处理因素
- 但会显著影响结果
- 如果不控制，会被“算进误差里”

👉 **随机区组设计的目的：**

> 把这些“已知的大差异”，
> 从随机误差中**单独剥离出来。**

二、随机区组设计在结构上做了什么？（Design intuition）

设计步骤（非常重要）

1️⃣ 先按“相同 / 相近条件”把对象分成 **m 个区组**
2️⃣ 每个区组内部，再**随机**分配到 k 个处理水平
3️⃣ 每个区组 × 每个处理 → **只有 1 个观测值**

📌 数据结构关键词：

- **两因素**：处理因素 + 区组因素
- **无重复**：每个格子 1 个值

- 统计上：
  👉 **two-way ANOVA without replication**

三、随机区组设计的核心直觉（非常关键）

CRD 的逻辑：

$$总变异 = 处理 + 随机$$

RBD 的逻辑：

$$总变异 = 处理 + 区组 + 随机$$

📌 **这一步的意义是：**

把“本来要进误差项的区组差异”，
单独拿出来，不让它污染误差。

四、三种平方和分别在干什么？（Key Concepts）

① 处理平方和$SS_{处理}$

回答的问题：

不同处理水平之间的均值差异有多大？

👉 **这是你真正关心的效应（signal）**

---

② 区组平方和$SS_{区组}$

$$SS_{区组} = \sum_{j=1}^{m} k(\bar{X}_j - \bar{X})^2$$

直觉解释：

> 不同区组整体水平差得有多远？

- 衡量的是：
  - 个体基础差异
  - 批次 / 时间 / 地块效应
- **不是研究重点**$SS_{误差}$
- 但**非常真实、非常大**

📌 把它单独拆出来，是 RBD 的灵魂。

---

③ 误差平方和

定义方式：

$$SS_{误差} = SS_{总} - SS_{处理} - SS_{区组}$$

直觉解释：

> 在扣除了“处理效应”和“区组差异”之后，
> 剩下的、解释不了的纯随机波动。

👉 **这是最“干净”的噪声估计。**

五、自由度结构（只看逻辑）

- 总自由度：

$$v_{总} = mk - 1$$

- 处理自由度：

$$v_{处理} = k - 1$$

- 区组自由度：

$$v_{区组} = m - 1$$

- 误差自由度：

$$v_{误差} = (k - 1)(m - 1)$$

📌 和你已经熟的原则完全一致：

> **每多估计一个“均值”，就少一个自由度。**

六、F 检验到底检验什么？（Very important）

处理效应的检验统计量：

$$F_{处理} = \frac{MS_{处理}}{MS_{误差}}$$

一句话翻译：

> 在“区组差异已经被扣掉”的前提下，
> 处理之间的差异，
> 是否仍然显著大于随机噪声？

📌 **这就是 RBD 比 CRD 更有力的原因。**

七、为什么随机区组设计能“提高效率”？（核心优势）

对比一眼看懂

- **CRD：**
  - 区组差异 👉 被算进误差
  - 误差大 → F 小 → 不显著

- **RBD：**
  - 区组差异 👉 单独剥离
  - 误差小 → F 变大 → 更容易显著

👉 **不是数据变好了，是噪声被清理了。**

表 8-6　随机区组设计的方差分析表

变异来源	平方和（SS）	自由度（ν）	均方（MS）	F 值
总变异	$SS_{总}=\sum_{i=1}^{k}\sum_{j=1}^{m}(X_{ij}-\bar{X})^2$	$\nu_{总}=n-1$		
处理间	$SS_{处理}=\sum_{i=1}^{k}m(\bar{X}_i-\bar{X})^2$	$\nu_{处理}=k-1$	$MS_{处理}=\frac{SS_{处理}}{\nu_{处理}}$	$F_{处理}=\frac{MS_{处理}}{MS_{误差}}$
区组间	$SS_{区组}=\sum_{j=1}^{m}k(\bar{X}_j-\bar{X})^2$	$\nu_{区组}=m-1$	$MS_{区组}=\frac{SS_{区组}}{\nu_{区组}}$	$F_{处理}=\frac{MS_{区组}}{MS_{误差}}$
误差	$SS_{误差}=SS_{总}-SS_{处理}-SS_{区组}$	$\nu_{误差}=\nu_{总}-\nu_{处理}-\nu_{区组}$ $=(k-1)(m-1)$	$MS_{误差}=\frac{SS_{误差}}{\nu_{误差}}$	

(李康,贺佳主编, 2024, p. 82) - Please click Zotero - Refresh in Word/LibreOffice to update all fields.

【例 8-2】为探讨不同营养素(A,B,C)喂养对增加小鼠体重的效果,某研究者将 24 只小鼠按性别、年龄、窝别及体重相同或相近原则分成 8 个区组。每区组中 3 只小鼠,用随机的方式分配到 A、B、C 三种不同的处理组,喂养四周后测量各小鼠体重,并计算其增加量(g),数据如表 8-7 所示。请问不同营养素喂养对体重增加的效果有无差别?

表 8-7　三组小鼠体重增加量　　单位:g

区组	A	B	C	$\bar{X}_j$
1	55.0	63.8	79.0	65.93
2	54.0	65.6	76.5	65.37
3	61.1	67.5	79.5	69.37
4	74.5	61.1	86.6	74.07
5	86.7	91.8	94.7	91.07
6	42.0	51.8	43.2	45.67
7	71.9	69.2	61.1	67.40
8	41.5	48.6	64.4	51.50
$\bar{X}_i$	60.84	64.93	73.13	66.30($\bar{X}$)

(李康,贺佳主编, 2024, p. 82) - Please click Zotero - Refresh in Word/LibreOffice to update all fields.

方差分析步骤如下:

(1)建立检验假设,确定检验水准

对于处理组

H_0:$\mu_1=\mu_2=\mu_3$,即三种营养素对体重增加量的总体均值相同

H_1:μ_1,μ_2,μ_3 不全相等,即三种营养素对体重增加量的总体均值不全相同

对于区组

H_0:$\mu_1=\mu_2=\cdots=\mu_8$,即八个区组体重增加量的总体均值相同

H_1:$\mu_1,\mu_2,\cdots,\mu_8$ 不全相等,即八个区组体重增加量的总体均值不全相同

$\alpha=0.05$

(李康,贺佳主编, 2024, p. 82) - Please click Zotero - Refresh in Word/LibreOffice to update all fields.

（2）计算检验统计量 F 值

$$SS_{总}=\sum_{i=1}^{k}\sum_{j=1}^{m}(X_{ij}-\bar{X})^2$$
$$=(55.0-66.30)^2+(54.0-66.30)^2+\cdots+(64.4-66.30)^2$$
$$=5\ 466.210\ 0$$

$$SS_{处理}=\sum_{i=1}^{k}m(\bar{X}_i-\bar{X})^2=8\times(60.84-66.30)^2+8\times(64.93-66.30)^2+8\times(73.13-66.30)^2$$
$$=626.699\ 2$$

$$SS_{区组}=\sum_{j=1}^{m}k(\bar{X}_j-\bar{X})^2$$
$$=3\times(65.93-66.30)^2+3\times(65.37-66.30)^2+\cdots+3\times(51.50-66.30)^2$$
$$=3\ 990.598\ 2$$

$$SS_{误差}=SS_{总}-SS_{处理}-SS_{区组}=5\ 466.210\ 0-626.699\ 2-3\ 990.598\ 2$$
$$=848.912\ 6$$

将上述计算结果列于表 8-8 的方差分析表中。

(李康,贺佳主编, 2024, p. 83) - Please click Zotero - Refresh in Word/LibreOffice to update all fields.

表 8-8　随机区组设计的方差分析表

变异来源	SS	ν	MS	F 值
总变异	5 466.210 0	23		
处理组间	626.699 2	2	313.349 6	5.17
区组间	3 990.598 2	7	570.085 5	9.40
误差	848.912 6	14	60.636 6	

（3）确定 P 值，做出推断结论

对于处理因素，根据分子的自由度和分母的自由度，按照 $\alpha=0.05$ 的检验水准，查附表 4F 界值表（方差分析用），$F_{0.05(2,14)}=3.74$，由于 $F=5.17$，$F>F_{0.05(2,14)}$，故 $P<0.05$，差别具有统计学意义（拒绝 H_0）。结论：可认为三种营养素对于小鼠体重增加量的总体均值不全相同。

对于区组因素，查附表 4F 界值表，$F_{0.05(7,14)}=2.76$，由于 $F=9.40$，$F>F_{0.05(7,14)}$，故 $P<0.05$，差别具有统计学意义，则可认为不同区组小鼠的体重增加量不全相同。

(李康,贺佳主编, 2024, p. 83) - Please click Zotero - Refresh in Word/LibreOffice to update all fields.

第四节 多个样本均数的两两比较

> **一句话先行**
> **ANOVA 只告诉你“至少有不同”，**
> **多重比较才告诉你“到底谁和谁不同”。**

一、为什么 ANOVA 之后还不够？（Purpose）

当你拒绝了 ANOVA 的 H_0：

$$H_0: \mu_1 = \mu_2 = \cdots = \mu_k$$

你真正知道的只有一件事：

> **“这 k 个均数里，至少有一对不一样。”**

但你**不知道**：

- 哪几组不同
- 差异发生在哪
- 是不是只和对照组不同

👉 **这就是多重比较存在的唯一理由。**

二、为什么不能直接用很多次 t 检验？（核心风险）

表面上看很合理：

- k 组 → 两两比较
- 一共 $\binom{k}{2}$ 次 t 检验

但问题是：I 类错误会累积

如果做 m 次独立比较：

$$\alpha' = 1 - (1-\alpha)^m$$

你给的例子非常关键

- $k = 5 \to m = 10$
- 单次 $\alpha = 0.05$

$$\alpha' = 1 - (0.95)^{10} \approx 0.40$$

👉 **40% 的概率至少错一次**

📌 这不是“稍微保守一点”，
这是**统计灾难。**

三、多重比较的本质冲突（一定要想清楚）

所有方法，本质都在**权衡两件事**:

① 少冤枉好人（控制 I 类错误）

② 别放过坏人（保持检验效能）

> **越严格控制错误 → 越不容易显著**
> **越放松 → 越容易“显著但不可信”**

所以不存在“最好”的方法，
只有**“最合适当前问题的”方法。**

四、两种错误控制思路（核心分类）

① CER：比较误差率（comparison-wise error rate）

> **每一次比较，都按 α 控制**

- 单次看没问题

- 比较次数一多 → 总体错误率暴涨

📌 代表方法：

- **LSD-t 检验**

👉 **特点**：

- 非常灵敏
- 非常危险
- **只适合探索性、少数比较**

---

② EER：实验误差率（experiment-wise error rate）

> **把“整个实验”犯错的概率控制在 α**

- 每一次比较都更严格
- 整体更安全

📌 代表方法：

- SNK-q
- Tukey
- Scheffé
- Bonferroni
- Sidak

👉 **这是正规分析的主流思路**

五、常见方法一眼判断（不背表格版）

🟦 Dunnett–t 检验

“只和对照组比”

- 多个处理 vs 一个对照
- 药物、干预研究首选
- **问题明确、功效高**

🟧 LSD–t 检验

“像 t 检验一样随便比”

- 不控制整体 I 类错误
- **只能用于极少数、有明确假设的比较**
- ❗ 考试里要写：

> 理论上只适合两组比较

🟩 SNK–q 检验（重点）

“逐步比较，先比差得最远的”

- 任意两组都可比
- 样本量可不等
- 控制 EER，但**比 Tukey 稍宽松**

👉 **常用、好用、教材爱考**

🟦 Tukey 法

“最标准的任意两两比较”

- 任意两组
- 要求样本量相等
- 控制 EER 较严格

🟪 Scheffé 法

“最保守，但最自由”

- 不仅能比均数
- 还能比**线性组合**
- 样本量不等也可

👉 **几乎不误报，但很难显著**

🟥 Bonferroni / Sidak

“简单粗暴地调 α”

- Bonferroni：$\alpha' = \alpha/m$
- Sidak：$\alpha' = 1 - \sqrt[m]{1-\alpha}$

📌 Sidak 稍微不那么保守
📌 Bonferroni 最好理解、最好解释

六、SNK-q 法在干什么？（你给的公式直觉）

$$q = \frac{\bar{X}_A - \bar{X}_B}{\sqrt{\frac{MS_{误差}}{2}\left(\frac{1}{n_A} + \frac{1}{n_B}\right)}}$$

这其实就是一句话：

> **“两组均值差，
> 是误差波动的多少倍？”**

- 分子：实际差异
- 分母：ANOVA 估计的随机噪声

📌 和 t 检验极像，
但：

- 使用的是 **ANOVA 的误差均方**
- 比较顺序是 **逐步的、分层的**

七、什么时候该用哪种？（真正有用）

我给你一个"决策树版直觉"

- **只关心 vs 对照组？**
  👉 Dunnett
- **明确只比某一两对？探索性？**
  👉 LSD（谨慎）
- **任意两两都想比，教材 / 常规？**
  👉 SNK 或 Tukey
- **想非常保守、还要比线性组合？**
  👉 Scheffé
- **审稿 / 伦理 / 多终点研究？**
  👉 Bonferroni / Sidak

表 8-9　多个样本均数两两比较方法选择策略

选择方法	适用范围
Dunnett-t 检验	实验前确定的多个试验组与一个对照组均数差别的比较
LSD-t 检验	多个组中，根据专业，仅对某一对或某几对在专业上有特殊探索价值的均数间进行的近似比较。理论上只适合两组比较
SNK-q 检验	实验后对任意两两组间均数均进行比较，各比较组样本含量可不相等
Tukey 法	任意两两组间均数均进行比较，要求各比较组样本含量相同
Schéffe 法	既可以进行因素水平的平均效应的比较，还可比较因素水平平均效应的线性组合，多用于对比组样本含量不等的资料
Sidak t 检验	两两比较时检验水准调整为 α'（$\alpha'=1-\sqrt[m]{1-\alpha}$，$\alpha$ 为方差分析原检验水准，m 为两两比较次数），以使多次比较犯 I 类错误的概率控制在 α 以内
Bonferroni t 检验	将两两比较时检验水准调整为 α'（$\alpha'=\alpha/m$），以使多次比较犯 I 类错误的概率控制在 α 以内，是 Sidak t 检验的近似

(李康,贺佳主编, 2024, p. 84) - Please click Zotero - Refresh in Word/LibreOffice to update all fields.

【例 8-3】对例 8-1 中不同注射剂量组受试者部分凝血活酶时间的均数作两两比较。

（1）提出检验假设，确定检验水准

$H_0: \mu_A = \mu_B$，即任意两组的部分凝血活酶时间的总体均数相等

$H_1: \mu_A \neq \mu_B$，即任意两组的部分凝血活酶时间的总体均数不等

$\alpha = 0.05$

（2）计算检验统计量 q 值

首先将三个样本均数由大到小排序，并编组（表 8-10）：

表 8-10　三组样本的部分凝血活酶时间均数排序编组

注射剂量	均数	组次
1U	37.83	1
2U	35.10	2
0.5U	33.62	3

三组均数共需做 $\binom{2}{3} = \dfrac{3!}{2!(3-2)!} = 3$ 次两两比较。

组次 1 与组次 3 比较：

$$MS_{误差} = 6.991\,2,\quad \bar{X}_1 = 37.83,\quad \bar{X}_3 = 33.62,\quad n_1 = 10,\quad n_3 = 10$$

$$q_{1,3} = \frac{\bar{X}_1 - \bar{X}_3}{\sqrt{\dfrac{MS_{误差}}{2}\left(\dfrac{1}{n_1} + \dfrac{1}{n_3}\right)}} = \frac{37.83 - 33.62}{\sqrt{\dfrac{6.991\,2}{2}\left(\dfrac{1}{10} + \dfrac{1}{10}\right)}} = 5.04$$

(李康,贺佳主编, 2024, p. 84) - Please click Zotero - Refresh in Word/LibreOffice to update all fields.

其余类推，可以得到组次 1 与组次 2、组次 2 与组次 3 比较的 q 值，检验的 P 值可以通过查附表 5 得到，将所有计算结果与 $q_{0.05(a,27)}$ 界值列于表 8-11，其中参数 a 为两对比组包含的组数。需要注意：如果组次 1 与组次 3 比较结果 $P>\alpha$，后面就不需要进行检验，直接判定为 $P>\alpha$。

表 8-11 例 8-3 的 SNK 法两两比较计算表

对比组 A 与 B	$\bar{X}_A-\bar{X}_B$	q 值	组数 a	$q_{0.05(a,27)}$ 界值	P 值
组次 1 与 3	4.21	5.04	3	3.52	<0.05
组次 1 与 2	2.73	3.27	2	2.91	<0.05
组次 2 与 3	1.48	1.77	2	2.91	>0.05

注：本例 $v_{误差}=27$，q 界值表中无此自由度下临界值，故采用内插值法计算相应的 q 界值。

（3）确定 P 值，做出推断结论

以误差（组内）自由度和对比组包含的组数 a 查 q 界值表，查表若 q 值大于或等于 q 界值，则可以推断比较的两组间差别具有统计学意义，否则差别无统计学意义。由表 8-10 可知，1U 与 2U 组、1U 与 0.5U 组比较时，$P<0.05$，拒绝 H_0，差别有统计学意义，而 2U 组与 0.5U 组之间的差别无统计学意义。

(李康,贺佳主编, 2024, p. 85) - Please click Zotero - Refresh in Word/LibreOffice to update all fields.

第五节 方差齐性检验｜Bartlett 检验

> **一句话先给结论**
> **Bartlett 检验在问的只有一件事：**
> **各组的“波动大小”，是不是在同一个量级上？**

一、为什么“方差齐性”这么重要？（Why it matters）

ANOVA 的核心比较是：

$$F=\frac{MS_{组间}}{MS_{误差}}$$

而这里面有一个**隐含前提**：

> **$MS_{误差}$ 能代表“所有组共同的随机波动水平”**

如果某些组：

- 波动特别大
- 某些组又非常稳定

👉 那一个“合并误差方差”就**失真了**
👉 F 值的参考基准就被污染
👉 **显著性判断不再可靠**

📌 所以在 ANOVA 之前，
检查方差是否齐同，比检查正态性更关键。

二、什么时候用 Bartlett？什么时候不用？（Usage）

Bartlett 检验的适用前提

- 各组 **相互独立**
- 数据 **近似正态分布**
- 检验目标：

$$H_0: \sigma_1^2 = \sigma_2^2 = \cdots = \sigma_k^2$$

📌 **重要提醒（一定要记）**

> **Bartlett 对“非正态”非常敏感**
> 非正态 → 容易误判“方差不齐”

👉 如果正态性存疑：
优先 Levene 检验

三、Bartlett 检验的核心思想（千万别被公式吓到）

核心直觉只有一句话：

> **如果各组方差真的一样，**
> **那每一组方差都应该“接近合并方差”。**

Bartlett 做的事就是：

1️⃣ 先算一个**合并方差** S_c^2
2️⃣ 看每一组方差 S_i^2
3️⃣ 量化它们**偏离合并方差的程度**

四、检验统计量在干嘛？（拆开看）

1️⃣ 为什么用对数？

$$\ln\left(S_c^2/S_i^2\right)$$

因为：

- 方差是正数
- 方差比例比差值更有意义
- 对数能让分布更接近 χ^2

👉 **这是“稳定方差”的经典技巧**

2️⃣ Q_1：真实差异强度

$$Q_1 = \sum_{i=1}^{k} \quad (n_i - 1)\ln\left(S_c^2/S_i^2\right)$$

直觉解释：

- 每一组：
  - 样本量越大 → 话语权越大
  - 方差偏离越多 → 惩罚越大

👉 **各组方差“不一致程度”的总量**

3️⃣ Q_2：小样本校正项

$$Q_2 = 1 + \frac{1}{3(k-1)}(\sum_{i=1}^{k} \frac{1}{n_i - 1} - \frac{1}{n-k})$$

你只要知道一件事：

> 这是为了让
> χ^2 的近似在小样本下不至于太乐观

📌 **考试不要求你解释它的来源**
📌 只要知道：**它是修正因子**

4️⃣ 最终统计量

$$\chi^2 = \frac{Q_1}{Q_2}, v = k - 1$$

👉 **"偏离强度 / 校正"**

五、检验逻辑一眼就懂（Decision rule）

在原假设成立时：

$$\chi^2 \sim \chi^2(k-1)$$

判断规则

- 若

$$\chi^2 \geq \chi^2_{\alpha, k-1}$$

  👉 **拒绝 H_0**
  👉 方差不齐

- 若

$$\chi^2 < \chi^2_{\alpha,\,k-1}$$

👉 **不拒绝 H_0**
👉 尚不能认为方差不齐

📌 注意措辞：

不是"证明齐性"，
而是"未发现不齐"。

六、和 Levene 的一句话对比（非常好用）

Bartlett：

- 假设正态
- 对非正态敏感
- 在正态条件下最有力

Levene：

- 不要求正态
- 更稳健
- 略保守

👉 **现实数据：Levene 更安全**
👉 **教材 / 理论推导：Bartlett 更经典**

七、考试级标准总结

Bartlett 方差齐性检验用于检验多个独立正态总体方差是否相等，其检验统计量在原假设成立时近似服从自由度为 $k-1$ 的 χ^2 分布。当统计量超过临界值时，拒绝原假设，认为各组方差不齐。

给你一个最后的“防混淆锚点”

> **ANOVA 是“信号 / 噪声”**
> **Bartlett 是在问：**
> **“这些噪声，是不是同一种噪声？”**

只要这句话在，你就永远不会把“方差齐性”当成形式条件。

【例 8-4】对例 8-1 资料作方差齐性检验。

（1）提出检验假设，确定检验水准

H_0: $\sigma_1^2 = \sigma_2^2 = \sigma_3^2$，即三个总体方差相等

H_1: σ_1^2，σ_2^2，σ_3^2 不全相等

$\alpha = 0.10$

（2）计算检验统计量 χ^2 值

本例：

$$k=3,\quad n_1=n_2=n_3=10,\quad n=30$$

$$S_1=2.263\,6,\quad S_2=2.207\,1,\quad S_3=3.313\,3$$

$$MS_{误差}=6.991\,2$$

(李康,贺佳主编, 2024, p. 85) - Please click Zotero - Refresh in Word/LibreOffice to update all fields.

$$Q_1=\sum_{i=1}^{k}(n_i-1)\ln(S_c^2/S_i^2)$$

$$=(10-1)\ln\frac{6.991\,2}{2.263\,6^2}+(10-1)\ln\frac{6.991\,2}{2.207\,1^2}+(10-1)\ln\frac{6.991\,2}{3.313\,3^2}$$

$$=1.987\,2$$

$$Q_2=1+\frac{1}{3(k-1)}\left(\sum_{i=1}^{k}\frac{1}{n_i-1}-\frac{1}{n-k}\right)$$

$$=1+\frac{1}{3\times(3-1)}\times\left(\frac{1}{10-1}+\frac{1}{10-1}+\frac{1}{10-1}-\frac{1}{30-3}\right)$$

$$=1.049\,4$$

$$\chi^2=\frac{Q_1}{Q_2}=\frac{1.987\,2}{1.049\,4}=1.89,\quad \nu=3-1=2$$

（3）确定 P 值,做出推断结论

自由度 $\nu=2$,查 χ^2 界值表，$\chi^2_{0.10,2}=4.61$。由于 $\chi^2=1.89$，$\chi^2<\chi^2_{0.10,2}$，故 $P>0.10$,按照 $\alpha=0.10$ 的检验水准,不拒绝 H_0,尚不能认为三个总体方差不齐。

(李康,贺佳主编, 2024, p. 86) - Please click Zotero - Refresh in Word/LibreOffice to update all fields.

[BIBLIOGRAPHY] Please click Zotero - Refresh in Word/LibreOffice to update all fields