

生物统计学

第十四章 抽样调查与试验设计

云南大学 生命科学学院



會澤百家 至公天下

获取数据的方式主要有两种：

- 一是在开展研究之前数据已经存在，或者说感兴趣的研究现象已经发生，研究者需要做的是收集分散的数据；
- 二是在开展研究之前数据还不存在，研究者需要首先通过必要的过程产生数据。

- 调查 (survey)
- 试验 (experiment)

- ① 抽样调查概述
- ② 常用抽样调查方法
- ③ 试验设计的基本原理
- ④ 常用试验设计方法

① 抽样调查概述

② 常用抽样调查方法

③ 试验设计的基本原理

④ 常用试验设计方法

14.1 抽样调查概述

调查是对已发生的现象进行数据的汇总、整理、分析，以期了解研究对象内在规律的过程。

- **普查**(census)，是指对研究总体中的每个个体，进行逐一调查，也称**全面调查**(complete survey)。
- **抽样调查**(sampling survey)，一种非全面调查，是根据一定的原则从研究**总体**中抽取一部分个体构成**样本**，对其进行观察、测量或度量，然后利用样本的信息来反映总体特征的方法。

- ① 抽样调查概述
- ② 常用抽样调查方法
- ③ 试验设计的基本原理
- ④ 常用试验设计方法

14.2 常用抽样调查方法

生物学研究中常用的抽样方法有：

- 随机抽样
- 顺序抽样
- 典型抽样

14.2 常用抽样调查方法



① 抽样调查概述

② 常用抽样调查方法

③ 试验设计的基本原理

试验设计的基本要素

试验误差的控制

试验设计的基本原则

④ 常用试验设计方法

① 抽样调查概述

② 常用抽样调查方法

③ 试验设计的基本原理

试验设计的基本要素

试验误差的控制

试验设计的基本原则

④ 常用试验设计方法

14.3 试验设计的基本原理

14.3.1 试验设计的基本要素

- 处理因素
单因素试验(single factor experiment) 和多因素试验(multiple-factor experiment)
处理因素(treatment factor) 和非处理因素(non-treatment factor)
- 试验对象
- 处理效应

① 抽样调查概述

② 常用抽样调查方法

③ 试验设计的基本原理

试验设计的基本要素

试验误差的控制

试验设计的基本原则

④ 常用试验设计方法

14.3 试验设计的基本原理

14.3.2 试验误差的控制

试验误差的来源：

- ① 试验材料本身的差异
- ② 试验条件不一致
- ③ 试验操作不一致
- ④ 偶然因素的影响

14.3 试验设计的基本原理

14.3.2 试验误差的控制

控制误差的途径：

- ① 选择纯和一致的试验材料
- ② 改进和标准化试验的操作流程
- ③ 采用合理的试验设计

① 抽样调查概述

② 常用抽样调查方法

③ 试验设计的基本原理

试验设计的基本要素

试验误差的控制

试验设计的基本原则

④ 常用试验设计方法

14.3 试验设计的基本原理

14.3.3 试验设计的基本原则

- ① 重复 (replication)
- ② 随机 (randomness)
- ③ 局部控制 (local control)

① 抽样调查概述

② 常用抽样调查方法

③ 试验设计的基本原理

④ 常用试验设计方法

完全随机设计

成组设计与配对设计

随机区组设计

平衡不完全区组设计

拉丁方设计

裂区设计

正交设计

① 抽样调查概述

② 常用抽样调查方法

③ 试验设计的基本原理

④ 常用试验设计方法

完全随机设计

成组设计与配对设计

随机区组设计

平衡不完全区组设计

拉丁方设计

裂区设计

正交设计

14.4 常用试验设计方法

14.4.1 完全随机设计

完全随机设计(completely randomized design), 是根据试验处理数 k ($k \geq 2$), 将全部 N 个试验单位随机地分成 k 组, 然后对不同的组实施不同处理的设计方法。

14.4 常用试验设计方法

14.4.1 完全随机设计

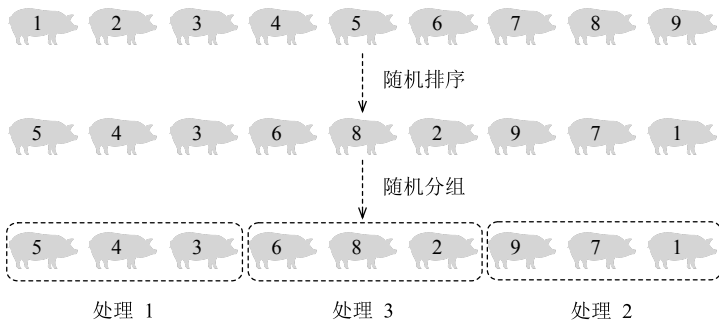


图 14.1 完全随机设计

① 抽样调查概述

② 常用抽样调查方法

③ 试验设计的基本原理

④ 常用试验设计方法

完全随机设计

成组设计与配对设计

随机区组设计

平衡不完全区组设计

拉丁方设计

裂区设计

正交设计

14.4 常用试验设计方法

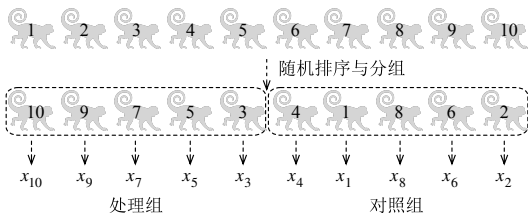
14.4.2 成组设计与配对设计

当处理 $k = 2$ 时，完全随机设计又被称为成组设计(two-group design)。

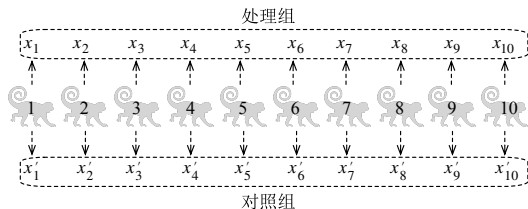
配对设计(paired design) 要求将试验单位俩俩配对，然后将配成对的两个试验单位随机分配到两个处理组中。

14.4 常用试验设计方法

14.4.2 成组设计与配对设计



(a) 成组设计



(b) 配对设计

① 抽样调查概述

② 常用抽样调查方法

③ 试验设计的基本原理

④ 常用试验设计方法

完全随机设计

成组设计与配对设计

随机区组设计

平衡不完全区组设计

拉丁方设计

裂区设计

正交设计

14.4 常用试验设计方法

14.4.3 随机区组设计

随机区组设计(randomized block design) 将试验材料按照性质一致的原则分为几个**区组**(block), 随机化只在区组内进行, 同时要求在每个区组内要包含所有处理。

14.4 常用试验设计方法

14.4.3 随机区组设计

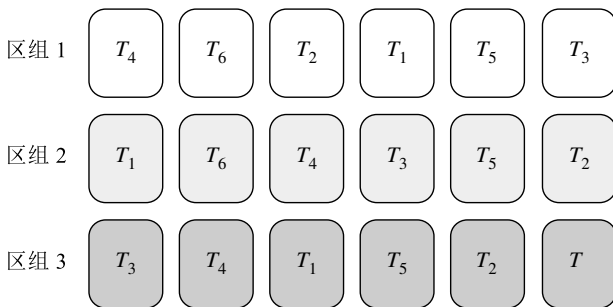


图 14.3 随机区组设计（灰度不同表示非试验因素的影响）

① 抽样调查概述

② 常用抽样调查方法

③ 试验设计的基本原理

④ 常用试验设计方法

完全随机设计

成组设计与配对设计

随机区组设计

平衡不完全区组设计

拉丁方设计

裂区设计

正交设计

14.4 常用试验设计方法

14.4.4 平衡不完全区组设计

随机区组设计中要求每个区组内包含所有的处理，这种区组称为
完全区组(complete block)。

如果一个区组只包含部分处理，则称为
不完全区组(incomplete block)。

14.4 常用试验设计方法

14.4.4 平衡不完全区组设计

平衡不完全区组设计 (balanced incomplete block design)

- 首先，根据试验处理数 k 、重复数 n ，以及每个区组中的小区数 a 计算区组数 A ($Aa = nk$)；以及两个处理在一个区组内相遇的次数 $\lambda = \frac{n(a-1)}{k-1}$ 。
- 然后，根据每一区组的小区数将处理排入区组，保证任意两个处理在一个小区相遇的概率相等。
- 最后，对各区组内的处理随机排列，再对区组进行随机排列。

14.4 常用试验设计方法

14.4.4 平衡不完全区组设计

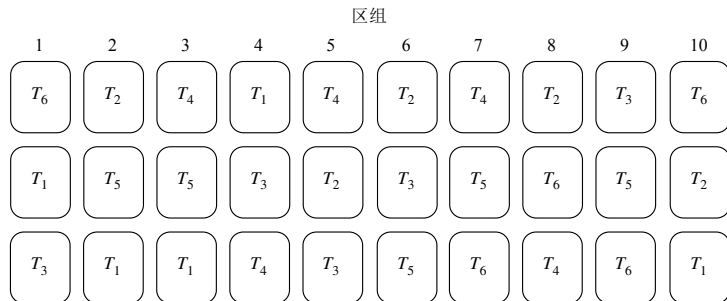


图 14.4 平衡不完全区组设计

① 抽样调查概述

② 常用抽样调查方法

③ 试验设计的基本原理

④ 常用试验设计方法

完全随机设计

成组设计与配对设计

随机区组设计

平衡不完全区组设计

拉丁方设计

裂区设计

正交设计

14.4 常用试验设计方法

14.4.5 拉丁方设计

拉丁方设计 (Latin square design)。

拉丁方设计就是利用拉丁方来安排试验的设计方法，是一种二维设计，用于有三个因素（可以视作一个试验因素和两个非试验因素），且每个因素的水平数都相同的研究。拉丁方设计需要安排的试验数等于 k^2 。

$n \times n$ 拉丁方：将 n 个不同符号排成一个 n 阶方阵，使每一个符号在每一行、每一列都仅出现一次。

14.4 常用试验设计方法

14.4.5 拉丁方设计

例 (14.1)

为研究五种饲料类型对乳牛产乳量的影响，考虑试验用牛和泌乳时间两个非试验因素，试作拉丁方设计。

14.4 常用试验设计方法

14.4.5 拉丁方设计

拉丁方设计流程如下：

① 选择标准拉丁方

- 第一行和第一列的字母按照自然顺序排列的拉丁方称为标准拉丁方(standard square)。
- 标准拉丁方的个数 K 、拉丁方阶数 n 、拉丁方的个数 S 之间的关系为 $S = K \times n! \times (n-1)!$ 。

14.4 常用试验设计方法

14.4.5 拉丁方设计

拉丁方设计流程如下：

② 行、列随机化

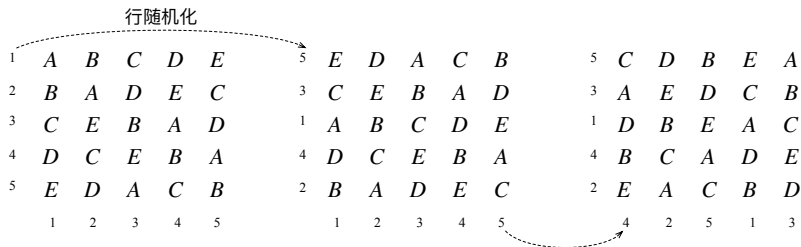


图 14.5 拉丁方设计的行、列随机化

14.4 常用试验设计方法

14.4.5 拉丁方设计

拉丁方设计流程如下：

③ 处理随机化

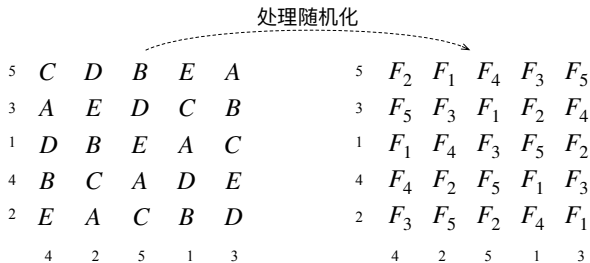


图 14.6 拉丁方设计的处理随机化

14.4 常用试验设计方法

14.4.5 拉丁方设计

拉丁方设计在行和列两个方向上都实施了局部控制，行列两方向皆成完全区组，而且行列互为对方的重复。

处理数、重复数、行数、列数都相等是拉丁方设计的特点。

① 抽样调查概述

② 常用抽样调查方法

③ 试验设计的基本原理

④ 常用试验设计方法

完全随机设计

成组设计与配对设计

随机区组设计

平衡不完全区组设计

拉丁方设计

裂区设计

正交设计

14.4 常用试验设计方法

14.4.6 裂区设计

裂区设计(split plot design) 和拉丁方设计一样也是一种多因素试验设计。

裂区设计中非试验因素有主次之分，适合于安排对不同因素精度要求不同、或因素可控性存在差异的试验。

14.4 常用试验设计方法

14.4.6 裂区设计

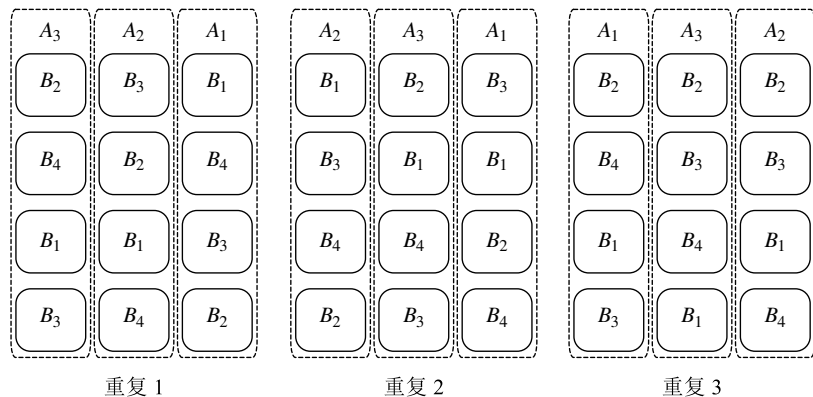


图 14.7 裂区设计（虚线框为主区，实线框为副区）

14.4 常用试验设计方法

14.4.6 裂区设计

裂区设计通常适用于下列情况：

- ① 一个因素的各处理比另一个因素的各处理需要更大的试验区域，或者说一个因素的实施需要更集中的区域。
- ② 试验中某一个因素的主效比另一个因素的主效更重要，而且要求更精确的比较，或者两个因素的交互作用更为重要。
- ③ 根据以往的研究经验，某些因素的效应比其它因素的效应更大，可将表现较大差异的因素作为主处理。
- ④ 对于已开展的试验，临时需要增加因素时，可在原试验设计中的小区中再划分副区。

① 抽样调查概述

② 常用抽样调查方法

③ 试验设计的基本原理

④ 常用试验设计方法

完全随机设计

成组设计与配对设计

随机区组设计

平衡不完全区组设计

拉丁方设计

裂区设计

正交设计

14.4 常用试验设计方法

14.4.7 正交设计

对于多因素试验，且考虑因素之间的交互作用，又要使用完全随机化方案，这样的试验设计称为析因设计(factorial design)。

14.4 常用试验设计方法

14.4.7 正交设计

对于多因素试验，且考虑因素之间的交互作用，又要使用完全随机化方案，这样的试验设计称为析因设计(factorial design)。

正交试验设计(orthogonal design)，特点是在全部试验处理的组合中，通过正交表(orthogonal table) 挑选部分有代表性的处理组合进行试验。

14.4 常用试验设计方法

14.4.7 正交设计

$$L_9(3^4)$$

- L 表示正交表
- 括号内幂的底数 3 表示因素的水平数
- 幂的指数 4 表示最多可安排的因素个数
- L 的下标 9 表示试验次数（处理组合数）

14.4 常用试验设计方法

14.4.7 正交设计

表 14.1 $L_9(3^4)$ 正交表

| 试验号 | 因素 | | | | 处理组合 |
|-----|----------|----------|----------|----------|-------------------|
| | <i>A</i> | <i>B</i> | <i>C</i> | <i>D</i> | |
| 1 | 1 | 1 | 1 | 1 | $A_1 B_1 C_1 D_1$ |
| 2 | 1 | 2 | 2 | 2 | $A_1 B_2 C_2 D_2$ |
| 3 | 1 | 3 | 3 | 3 | $A_1 B_3 C_3 D_3$ |
| 4 | 2 | 1 | 2 | 3 | $A_2 B_1 C_2 D_3$ |
| 5 | 2 | 2 | 3 | 1 | $A_2 B_2 C_3 D_1$ |
| 6 | 2 | 3 | 1 | 2 | $A_2 B_3 C_1 D_2$ |
| 7 | 3 | 1 | 3 | 2 | $A_3 B_1 C_3 D_2$ |
| 8 | 3 | 2 | 1 | 3 | $A_3 B_2 C_1 D_3$ |
| 9 | 3 | 3 | 2 | 1 | $A_3 B_3 C_2 D_1$ |

14.4 常用试验设计方法

14.4.7 正交设计

用正交表进行的试验安排则具有以下特性：

- 均衡分散性

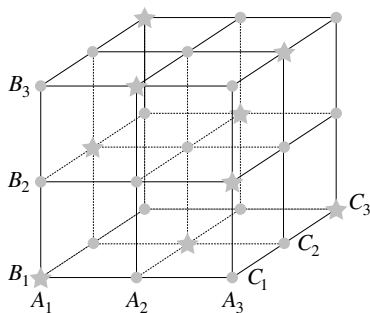


图 14.8 正交表的均衡分散性

14.4 常用试验设计方法

14.4.7 正交设计

用正交表进行的试验安排则具有以下特性：

- 整齐可比性

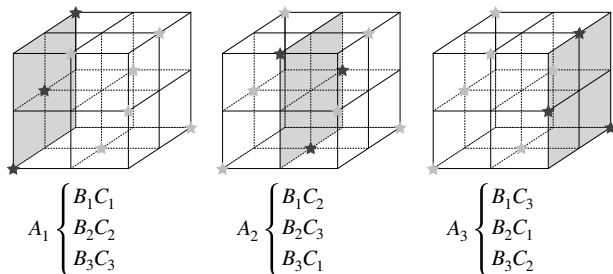


图 14.9 正交表的整齐可比性

14.4 常用试验设计方法

14.4.7 正交设计

以下三种变换可得到同构的正交表(isomorphic orthogonal table):

- ① 任意两列互换，换句话说也就是试验因素可以安排在任一系列上；
- ② 任意两行互换，这就使得处理组合没有顺序，可以自由选择；
- ③ 每一列中不同数字代表的水平可以互换，使得因素的水平可以自由安排。

14.4 常用试验设计方法

14.4.7 正交设计

正交试验设计的基本流程：

- 选择正交表

$$L_{\text{No. of Exp.}}(\text{No. of Levels}^{\text{No. of Factors}})$$

- 不考虑互作时，可选择幂大于或等于因素个数的正交表。
- 考虑互作，幂必须大于因素个数。
- 如果为每个处理组合只安排一个试验单元（无重复），则幂必须大于因素个数与要考虑的互作数之和。

所选的正交表既要能安排下全部试验因素，又要使处理组合数尽可能的少。

14.4 常用试验设计方法

14.4.7 正交设计

正交试验设计的基本流程：

- 选择正交表

某药厂进行了提高某抗生素药物发酵单位的试验，设有 8 个试验因素，各有 3 个水平。如果采用正交试验，并考虑 $A \times B$ 和 $A \times C$ 互作效应，则最少需要试验次数为

$$(3 - 1) \times 8 + 1 + (3 - 1) \times (3 - 1) \times 2 = 25$$

各因素的水平数减 1 求和后加 1，再加上交互作用的自由度，即是所需最少试验次数或处理组合数。

14.4 常用试验设计方法

14.4.7 正交设计

正交试验设计的基本流程：

- 表头设计

表头设计是指将试验所要研究的因素，以及互作分配给正交表中的各列。

- ① 不考虑互作，可以任意确定；
- ② 考虑互作，必须参考正交表的**交互作用表**来确定。

14.4 常用试验设计方法

14.4.7 正交设计

正交试验设计的基本流程：

- 表头设计

表 14.2 $L_8(2^7)$ 正交表

| 试验号 | 列号 | | | | | | |
|-----|----|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 2 | 1 | 1 | 1 | 2 | 2 | 2 | 2 |
| 3 | 1 | 2 | 2 | 1 | 1 | 2 | 2 |
| 4 | 1 | 2 | 2 | 2 | 2 | 1 | 1 |
| 5 | 2 | 1 | 2 | 1 | 2 | 1 | 2 |
| 6 | 2 | 1 | 2 | 2 | 1 | 2 | 1 |
| 7 | 2 | 2 | 1 | 1 | 2 | 2 | 1 |
| 8 | 2 | 2 | 1 | 2 | 1 | 1 | 2 |

14.4 常用试验设计方法

14.4.7 正交设计

正交试验设计的基本流程：

- 表头设计

14.3 $L_8(2^7)$ 二列间的交互作用

| 列号 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|----|-----|-----|-----|-----|-----|-----|-----|
| 1 | (1) | 3 | 2 | 5 | 4 | 7 | 6 |
| 2 | | (2) | 1 | 6 | 7 | 4 | 5 |
| 3 | | | (3) | 7 | 6 | 5 | 4 |
| 4 | | | | (4) | 1 | 2 | 3 |
| 5 | | | | | (5) | 3 | 2 |
| 6 | | | | | | (6) | 1 |
| 7 | | | | | | | (7) |

14.4 常用试验设计方法

14.4.7 正交设计

正交试验设计的基本流程：

- 表头设计

14.4 $L_8(2^7)$ 表头设计

| 因素数 | 列号 | | | | | | |
|-----|-------------------|-------------------|------------------------------|-------------------|------------------------------|-----------------------------------|-------------------|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 3 | A | B | $A \times B$ | C | $A \times C$ | $B \times C$ | |
| 4 | A | B | $A \times B$ $C \times D$ | C | $A \times C$ $B \times D$ | $B \times C$ $A \times D$ | D |
| 4 | A | B $C \times D$ | $A \times B$ | C $B \times D$ | $A \times C$ | D $B \times C$ | $A \times D$ |
| 5 | A $D \times E$ | B $C \times D$ | $A \times B$ $C \times E$ | C $B \times D$ | $A \times C$ $B \times E$ | D $A \times E$ $B \times C$ | E $A \times D$ |

14.4 常用试验设计方法

14.4.7 正交设计

正交试验设计的基本流程：

- 制定方案

根据表头设计的结果，按照 $L_8(2^7)$ 正交表（表 14.2）各因素所在列中的数字进行处理的组合。

各处理的组合根据主效列实施！

本章小结

- ① 抽样调查概述
- ② 常用抽样调查方法
- ③ 试验设计的基本原理
- ④ 常用试验设计方法