

## 2 实验设计 (Experimental design)

多元分析方法不再仅仅局限于对数据的探索和新假设的生成。特别对控制实验的数据分析，约束排序(constrained ordination)是一个很强大的工具。在这一节，我们将重温一下实验设计的基本类型，特别是野外控制实验的设计。通常，我们希望实验的目的是比较研究对象（比如生物群落）在不同的处理（treatments）下变化，当然其中有一种处理通常是“对照处理(control treatment)”

如果响应变量是一元变量（比如群落内物种的种数、总生物量等），最常用分析手段是方差分析（ANOVA）、一般线形模型（其实包括方差分析，线形回归及它们的组合），或是广义线形模型（GLM），GLM 是一般线形模型的延伸，在响应变量不符合正态分布的条件下，广义线形模型是比较适合的分析方法。

### 2.1 完全随机化设计 (Completely randomized design)

最简单的实验设计是完全随机化设计（图 2-1），我们先在一个样区里面选择若干实验地（plots），将不同的处理水平随机挑选实验地。这种设计是合理的，但有时并不是最好的，因为并没有考虑到环境异质性的影响。环境异质性作为一种不可预知的变量总是在影响试验结果。如果环境异质性的变化很大，这种随机设计可能会降低实验的可靠性和效果。简单讲，并不能保证样区内的环境是同质的。

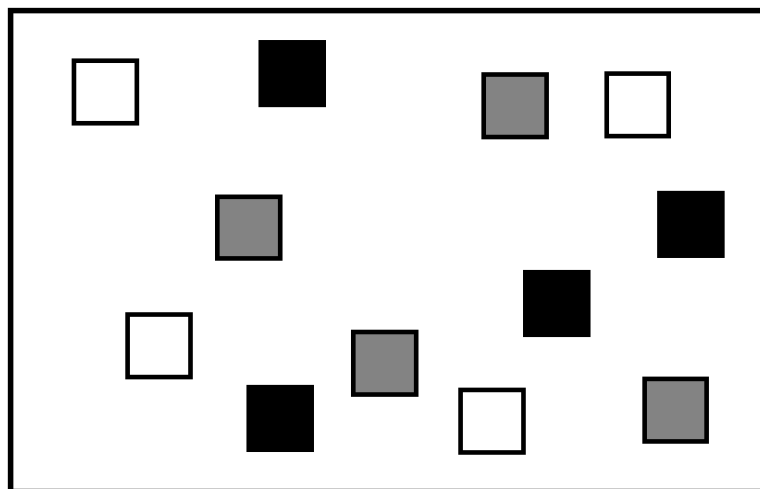


Figure 2-1. Completely randomized design, with three treatment levels and four replicates (or replications).

### 2.2 随机完全区组设计 (Randomized complete blocks)

有几种方法可以控制环境异质性对于实验的影响。在生态学实验设计中最常用的是随机完全区组设计。首先，我们先选择多个区组（blocks），每个区组内是尽量环境同质（比如图 2-2 中长方形区组环境同质，而不同的区组有环境梯度存在）。区组的数量等于重复的数量。每个区组仅仅包含每种处理水平的一个重复，在同一区组内的各个水平的位置是随机的。

如果区组之间的环境条件不同，这种随机完全区组设计显然要比完全随机化的实验设计检验效果更好。不过，如果区组之间的环境差异不大，这种设计方案要比完全随机

设计的检验效果要差，因为随机完全区组设计显然限制了重复的分布，减少了实验设计的自由度。如果处理水平和重复的量比较少的话，随机完全区组设计可能更危险。

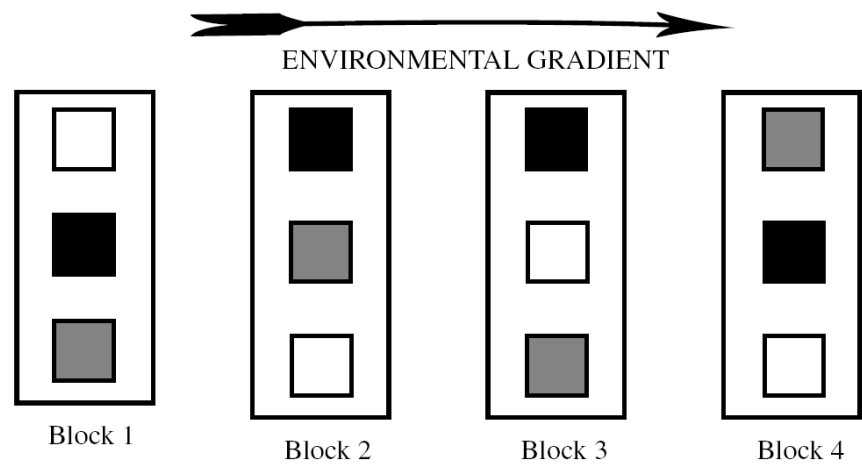


Figure 2-2. The randomized complete blocks design.

2.3 拉丁方设计（Latin square design）

拉丁方设计（图 2-3）假设一个正方形实验地的横向和纵向都有梯度差异。在正方形实验地内，实验设计每个行和每个列只含有一个处理水平的一个重复。所以，拉丁方设计要求重复的数量也等于处理水平的数量。这可能是个不太合理的限制条件。然而拉丁方设计可能在农业研究方面比生态学研究用得更普遍。跟随机完全区组设计一样，拉丁方内的环境异质性得到了很好的控制。当然，如果环境异质性变化很小，对于实验影响不大的情况下，拉丁方的设计比起完全随机设计的效果差得更远，因为拉丁方的自由度更少。

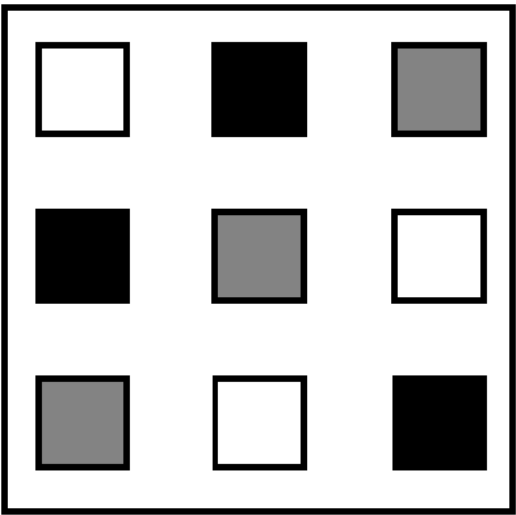


Figure 2-3. Latin square design.

## 2.4 最常犯的错误：假重复（pseudoreplications）

假重复（图 2-4）是生态学研究中最常犯的错误（Hurlbert 1984）。实验设计如图 2-4，每个实验地（plot）代表一种处理，每种处理的重复仅仅在同一个实验地内出现。（译者注：原文讲得比较专业，举个简单的例子说明这个问题更容易明白），如图 2-4，假设我们研究施肥的多少对于植物生长的影响，实验是这么设计的，plot1 是对照，有三个重复，同样，plot2 是施肥 10%，也有三个重复，plot3 是施肥 20%，也有三个重复。我们如果用方差分析得三个 plot 的实验结果有显著差异，但能说明施肥的不同对植物生长有显著影响吗？显然不行，因为，这种显著差异或许是由 plot 之间的差异引起的，并不是施肥的多少引起的。所以，在同一个 plot 里面做重复，是对于验证 treatment 效果，是不可取的，也是不可靠的。

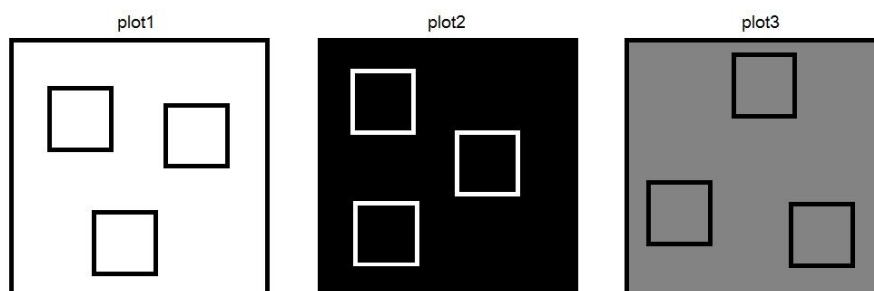


图 2-4 关于假重复的实验

## 2.5 多因子实验设计（Combining more than one factor）

前面介绍的都是只有一种处理（因子）实验设计。但实际上，我们通常在一次实验内同时完成数个因子的差异检验。例如，在生态学研究中，我们经常要去同时检验施肥(fertilization)和刈割(mowing)的效应。

**因子设计（Factorial designs）：**检测双因子组合实验最常用的方法叫做因子设计。这种方法要求每个实验样本含有一种因子的水平跟和另外一种因子的水平。比如，我们考虑施肥（两个水平，施肥与不施肥）和刈割（刈割与不刈割）的影响，我们可以得到四种不同的组合。这四个组合在空间上可以完全随机布置，也可以使用随机完全区组设计和拉丁方设计。

**分层设计（Hierarchical designs）：**在分层设计中，每个主要样方(plot)包含几个小样方（subplots）。比如，我们要研究施肥对土壤生物的影响。考虑到实际条件的限制(比如边缘效应)，样方（plot）至少需要 5m×5m 的面积。这限制了各个水平重复的数量，以保证试验区空间是足够的。尽管如此，土壤生物经常通过直径 3cm 的土钻来取样。常识告诉我们，一个基本样方（plot）需要多钻几次，所获得的样本才具有代表性。这样做是对的。然而，每一个土钻取的样并不是独立观测值。这里的变量又多出来一个水平，即小样方（在方差分析的术语中，这里的样方是一个随机因子）。

这些样方处于处理水平之内（有时我们也把这种设计叫做嵌套设计而不是分层设计）。在这种设计中，处理效应一般来说是通过对最短距层次水平的变异程度的检验得到的。在上面提到的例子中，施肥的效应是对样方（在一个施肥处理水平内部）之间的变异进行检验得到的，而不是对土钻之间的变异进行检验。仅仅通过在一个样方内多打

几个土钻，我们是无法增加误差的自由度的；但使用这种方法，我们却能够降低一个处理水平内部各样方之间的变异程度，并且增加检验结果的解释力度。有时，这种设计也叫做裂区设计。裂区设计更为经常地运用在另一种分层设计中，在这种设计中，两个（或两个以上）实验因素是像图 2-5 一样结合在一起的。在这种情况下，除了主样方（每个水平设置三个重复）之外，还存在着裂区，主样方在这里就该称为整区。

图 2-5，岩床的效应是通过对比岩床类型内各主样方之间变异程度的检验得到的。应当注意到，这种实验设计和岩床-施肥双因素实验设计是不同的，因此也应当使用不同的分析方法--它还多出来一个变异水平，也就是样方（即存在于岩床种类这一主要因素之内的次要因素）。

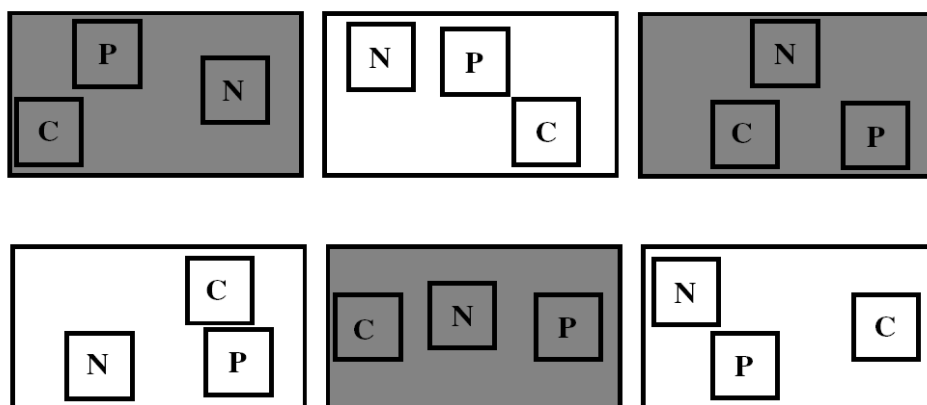


Figure 2-5. The split-plot design. In this example, the effect of fertilization was studied on six plots, three of them on limestone (shaded) and three of them on granite (empty). The following treatments were established in each plot: control (C), fertilized by nitrogen (N), and fertilized by phosphorus (P).

## 2.6 重复观测（repeated observations）

生物或生态的主体（objects）有一基本的特点是经常会随时间发生变化。我们通常感兴趣的不仅是静态的数据，也对动态的数据感兴趣。在一般的情况下，如果本底的数据（baseline data, 未通过实验处理之前的数据）可用的话，那推论具有更强的说服力。在这种情况下，我们需要使用 BACI（before after control impact 前后对照实验影响）的设计（Green 1979）。

比如要研究工厂对河流的影响(图 2-6)，你可以在建工厂前在下游取一次水样，建成后依然在下游同一地点取水样，比较建厂前后水质的变化；还可以建厂后，同时在上游和下游的取样，比较两个水样的差异，分析河水流过工厂后是否有变化。但这样都有点问题，第一种做法，不能保证时间前后取样是上游的水质是一致的（原始状态不一致）。第二种做法，不能保证水质的变化全部是由工厂引起的，也许是有上游他的因素引起的。正确的做法应该是 BACI 设计，在建工厂前，应该同时在上游和下游取样，然后分析得到差别，然后建完工厂后，也一样同时在上游和下游取样，两者的差别减去原来的差别，便是真正的工厂引起的影响。

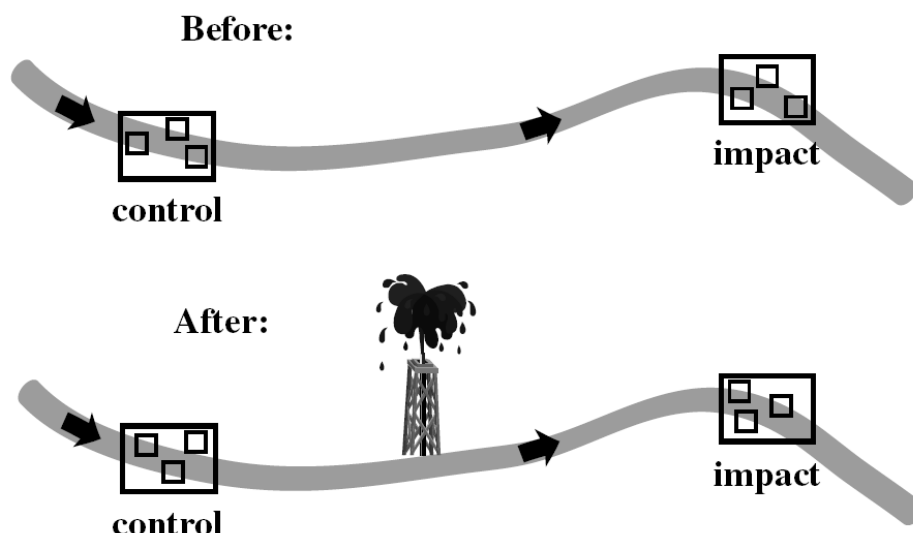


Figure 2-6. The non-replicated BACI (before after control impact) design.

## 2.7 控制实验数据和观测数据 (Experimental and observational data)

控制实验毕竟在时间和空间上受到很大的限制。在大的空间和时间尺度，只能依赖观察数据。生态学研究的好多东西是无法通过控制实验来模拟的，只能通过观测得到，控制实验有是做不到，因此，控制实验和观测结果相结合是非常必要的。通常来讲，观测的数据对于假说的产生更有用，而实验数据更适合假设的验证。然而，情况不一定总是这样的。比如在群落水平的控制实验，我们经常考察的是一系列相关的影响变量，比如本书案例研究第三个例子（第 13 章，关于施肥试验的例子），施肥不仅仅是影响作物的盖度，也直接影响杂草的多度。在群落水平的控制实验，其实那里有很多不可控制的因素（比如是否在同一年做的实验，实验地是否有差异等等），而且这些因素也影响结果。当相应变量是物种的组成的时候，我们经常知道有些物种的生物学特征（这些特征是我们无法控制的）。比如，当响应变量是群落物种组成，我们经常了解一些物种的生物性状（其中我们无法操纵）。我们首要测试的假设处理是否改变物种的组成，但结果往往会引出物种的物种性状关系的新假设。

此外，在非控制实验研究中，我们经常收集数据去验证一个特定的假设。类似于控制实验的数据收集方式是很有用的。Diamond (1986) 甚至提出一个新词叫“自然实验 (natural experiment)”，通过自然的观察当作是控制实验的结果，同样用控制实验的分析方法来分析数据（例如，在是否有蜥蜴的岛上比较蜘蛛的丰富度，此时蜥蜴的有无就可以当作是控制实验因子）。

生态学研究，控制实验数据和观测数据都是必须的。