**Homework-01**

1. What difficulties does the old paradigm of the hypothesis test encounter?

答：

（1）与物理、化学等学科不同，生态学问题的答案不是简单的“是”或“不是”，通过H0显著性检验来验证生态学假说并不严谨。例如，检验群落中竞争的作用，不是通过一个简单实验就可证伪的，因为群落中同时存在竞争、捕食和寄生、干扰等多因素作用。在生态学上构建单一主导因子的H0是很难的。

（3）H0假说是小概率事件，小概率事件在一次实验中不可能发生。但生态过程不断演化，无法预见原假说是小概率事件，即。因此，生态学的所有试验验证逻辑上应倾向于“证实”假说H0成立。

（2）假说-验证范式的另一个困境是操控实验。操控实验包括正交实验设计法与析因法，无论哪种实验，都遵循重复、随机化、区组化的原则。由于研究对象所处的环境时开放的，其中影响因子及相互关系难以被发现和确定。另外，在大尺度环境下，要采取严谨和一致操作才可能排除无关干扰。可是许多专家认为，无论如何精心设计和操控，都“不可能包括所有的多样性和空间”。因此，野外实验存在着难以重复的困难，不能作为假说的判决性实验。

2. What advances does the paradigm of data science used for ecological data?

答：

（1）寻找相关性，而不是因果律。无论是经验科学范式，还是假说-验证范式，都认为科学研究的目的就是寻找现象之间的因果关系。与其他研究范式不同，数据科学范式认为分析变量之间的相关性比探寻因果律更重要，基于相关分析的预测是大数据研究的核心。

（2）数据规律本身是从过去所积累的数据中挖掘出来的，完全可解释过去的现象或问题，另一方面，大数据包含了海量的各种现实数据，通过机器学习过去的经验来推测未来。因此，与因果律相比，基于大数据相关性的预测更准确，而且不易受偏见的影响。

（3）数据科学范式沿袭了经验科学的归纳逻辑，不同的是数据科学采用的是全数据模式，即“样本=整体”的完全归纳法，客服了小样本不完全归纳法的局限性，利于发现异常值。