

数学与生态学

Joseph Malkevitch / 文 丁 玫 / 译

导言

数学因其证明普遍适用的漂亮“事实”(定理)及因其众多的应用而很重要。数学中那些显示有用的学术领域的范围已随时间不断增大。有一个时期,就数学与其他各学科的关系而言,人们认为数学与物理是最引人注目的合作伙伴。然而就这一点上来看,没有一门学科中数学不发挥重要作用。

在前几篇专栏文章中,我已经选取了几个这样的跨学科合作伙伴关系:数学与化学及数学与心理学。这回我将谈论数学与生态学。在过去,一般认为数学在生物学中的应用是有限的。这是因为生物体太多样化、太复杂了,对数学分析而言就太难以捉摸了。然而,不仅是随机(概率)数学的应用,而且是确定性的“模型”,在生物学中都已经赢得了关注。

生物学与数学一样,是门具有许多丰富而复杂内容的学科。因此,如果一个人瞧瞧生物学的不同“分支”,就会发现它分成许多部分,其中包括以字母“A”开头的解剖学,以及以字母“Z”开头的动物学。我很惊讶遗传学和基因组学却不在花名册上,但我发现它们位于“生命科学”的不同“分支”图上。

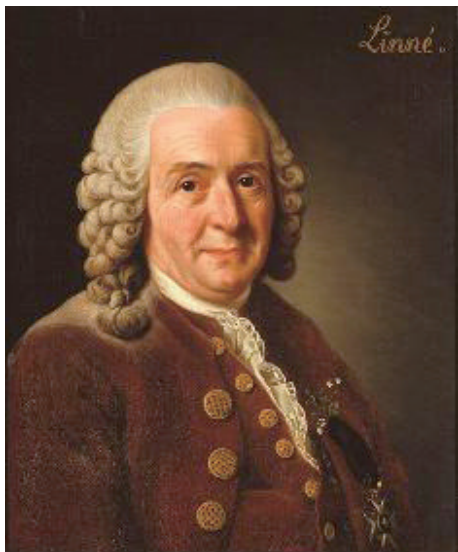
对于思考数学如何用于其它科目,想想分类学往往是有帮助的——结构性地想到一门学科的不同部分怎样组合在一起的方式。例如,没有人会怀疑数学已被证明在遗传学领域中非常有用。然而,时间已经改变了组织及思考该领域结果的方式。从最早的时候起,农民和饲养牲畜的人们就己用“数学思维”以完善结果——更高的作物产量以及更快养肥的牲畜。于是,我们看到下列的传统和新出现术语的混合物:

数学生物学
计算生物学
计算基因组学
生物信息学
生态学

生物学的不同部分已经从数学提供给生物的一系列工具中引进不同的方法。毫不奇怪,通常是一项生物学的发展,导致有关如何用数学来理解生物洞察力的新思路。生物学的突破会导致用数学工具来帮助生物学家的突破。我不试图展现数学帮助生物学家理解现代生物学的浩瀚景观,但可以看看一个相当小的领域,它提供从一个新角度看一些传统课题的方法。这种讨论发生在生物学中称为生态学的那个部分。生态学涉及不同种类的生物彼此之间及和它们的环境相互作用的方式。一旦人们将自己的“显微镜”带进生物学的这个子领域,就会看到生态学虽然是生物学的一小部分,就其自身而言已是具有多变性的一个学科。

分类

自从瑞典学者卡尔·林奈(Carl Linnaeus)引进生物分类后,生物学家们一直关注地球上生命形式的多样性以及怎样理解它们之间的关系。林奈使用以数学术语来思维



卡尔·林奈 (1707-1778)

的语词“串”，来帮助他对所看到的复杂生物世界建模或简化。林奈所做的确实不得不进行现代化升级，以适应他之后的科学家们得到的想法。像所有了解“真实世界”的那些企图一样，那个世界的复杂性往往挑战我们对它描述的尽可能简单的方式！

我们使用分层术语：域、界、门、纲、目、科、属和种。因此，在所在门的范围内有很多纲。然而，这种分类随着时间而改变，并没有被普遍保持。变化的原因之一是“模型”在科学中的使用方式。根据数据，人们试图获得最佳的模型来解释该数据。然后，随着时间推移当新的数据变得可用时，人们试图确保新数据也可被目前的模型所解释；否则的话，模型需要改变，以能更好地理解当前数据的状态。

在克里克-沃森 (Crick-Watson) 试图理解基因的革命前，动物的分类是基于各种特性的，包括视觉外观。狗的尺寸和种类繁多，但大部分动物学家都有共识哪些动物是狗，哪些不是。然而，对某些动物而言，却不清楚哪些“相近”，哪些“远离”。针对不同动物的基因组测序可以帮助建立一个与外貌不同的新“指标”，用来告诉事物是接近的还是相距甚远的。因此，人们可能会倾向于将鸟、昆虫及蝙蝠这样能飞的动物归类为“相近”的，但是并不奇怪的是，从基因的角度看，昆虫和鸟类似乎相距甚远。不同生物的基因组测序是驱使生物分类学近期调整的一部分。过去被视为密切相关的物种 X 和 Y，如果它们的基因组相距甚远，考虑到这一点则有调整其分类系统的动机。

因此，林奈有三个“界”——动物、植物和矿物质。因为矿物质（虽然它们可以在“成长”）不是活着，这第三个界已被丢弃。然而，随着基因组测序的到来，人们发现，理所当然地有更多与植物和动物组织得有所不同的“界”。实际上，已确定在界之上应有一个范畴，它通常被称为域。现在经常使用的三个域是古生菌、细菌和真核生物。古生菌是一种微生物，但它似乎与其它种类的微生物十分不同，值得特别关注。因此，1977年，在卡尔·乌斯 (Carl Woese) 和乔治·E·福克斯 (George E. Fox) 工作的基础上，传统的分类系统被修改，“界”在上述三个区域的范畴内中被重新组织。

什么是生态学？

生态学是生物学的一个分支，它关注的是有机体如何相互关联以及它们居住的环境。如同林奈和他的后继者们所发现的那样，人类与其它许多生命形式共享地球。无论居住何处，不管是纽约城还是亚马逊河，人们只看到地球上物种多样性的一小部分。对这些生命形式的好奇心就是动物园如此受欢迎的原因。这也导致了用科学的理念了解人如何与其它生命形式互动并使用它们（用于食品、医药等）。而且，哪里有科学，哪里就有数学。在生态学中，和其它地方一样，人们使用数字来计数及测量。

当一个人测量东西时，比如说某个地区大象的重量，研究者通常致力于测量所有大象的重量或采集大象的样本，并尝试通过群体的样本信息来外推。当群体数很大时，很难对所有个体进行测量。另一方面，当人们取样本时，往往很难“肯定”样本足以代表群体。美国动物园里的大象是所有大象的“典型”吗？了解一组数据涉及到两个

基本概念。一是中心趋势的量度概念——这是捕获所得数据组价值的单一数字。中心趋势常见的计量是均值（测量值之和除以成员个数）、中值（所有值按递增排序，位于中间的那个），以及众数（最常出现的测量结果）。不令人惊讶的是，由于一组数据可以都和某个数，比如均值相近，而另一组群体则可能与第一组有相同的均值，但数据却“分散开来”，因而单一的数字很难捕捉整个种群的性态。因此，除计算数据集中心趋势量外，自然也要计算“分散性”的程度。分散性测量试图说明数据关于“中心趋向”是怎样分布开来的。一个典型例子是使用均值作为中心趋向的量度及标准差作为关于均值分散性的量度。数据分散性的另一量度是最大和最小测量值之间的相差范围。想到这个范围后，人们可能发明用差值范围除以二的“中间范围”值来度量中心趋向。种群均值和标准差，以及样本均值和标准差是数学家（统计学家）用来比较两个群体以便了解发生了什么的工具。统计学家的一般工具也是生态学家的工具。生态学家已经发明了各种各样的“指数”来衡量并深入了解生命体。

当人试图了解物种多样性时，如果有两套“陷阱”（收集站）并只是从这些陷阱收集所发现物种的数量信息，则可以尝试弄清楚可能有许多物种在第一个陷阱里，而在另一个陷阱较少，但在第一个陷阱里可能每个物种的数量大致相等，而另一个陷阱包括的物种的数目可能相当分散。因此，一个陷阱可能具有5个物种，其中每个物种的数量从1到8变化，而另一个陷阱可能已经发现了11个物种，但每种只有2个或3个个体。为了解生物多样性所形成的诸多观测复杂性也导致了各种各样的定义来捕捉生物多样性的不同方面。复杂性说明，从夏天到冬天，或从一个地理区域到另一个，结果可能会如何不同。另外，大型哺乳动物、鸟类、昆虫和珊瑚的生态学研究对科学家提出了不同的挑战。

稀疏化 (Rarefaction)

数学强大的一个原因是同样的数学工具可以用于完全不同的应用场景。另一方面，对问题建模时，人们可能会使用用于类似情形的相同工具作为第一近似，但是要小心，所得结果在相似情景下应真正有意义。

为了就这一点更加具体化，让我们想象一位生态学家试图了解一个确定的地理区域的物种多样性。抽象地说，有人可能会说需要涉及所测量区域“物种多样性”的一些样本。首先需要考虑的是，计数树种比计数鱼、鸟、蛇、昆虫或藻类来得容易。另一个值得关注的是避免伤及被采样的动物。如果你试图研究一定区域内不同种类的树木，所采用的程序将会与研究老鼠、飞蛾和甲虫有很大的不同。还有，给予不同的地理环境，有趣的相关问题是你想在哪里采取样本。“区域”是湖、江、不规则场地，或森林的矩形截面？虽然在相对明确的样本集情况下思考不同的数学处理颇有迷人之处，我的目标还是多说一点关于这种简单情形如何产生在诸多情况下已大有可为的数学思想。

无论你在何地生活，大城市、郊区或农村地区，你的人生调料之一就是你看到并非人类的其它活体。在大城市，你看到“动物”部门的松鼠、鸽子及“植物”部门的各种花草树木。在乡村你可能会看到鹿或狐狸。然而，各种生命形式的最大水库是没有大量人类居住的地区，或海洋里及“地下”划分出的“野生动植物”。我们往往集中于肉眼能见的生命形式，但也有大量各种各样的细菌存在。

当人类想到生物多样性时，他们往往想象有多少个物种。数学已经以多种方式被用来获得对此问题的一个全面了解。然而，鉴于世界是个充满活力的地方，即便计算物种的个数这一问题都绝非“平凡”。生态学家已经做了很多工作来试图了解地球生态系统的性质以及我们这个地球上生命的复杂局面。大型哺乳动物相对于温度区域的分布是什么？有人可能会直观地认为热带地区两极附近的寒冷地区会有更多的大型哺乳动物。收集数据后人们会得到什么？从赤道向北极或南极移动时树木和开花植物如何

扩展？生命在海洋或陆地上哪个更具多样性？昆虫在不同的温度区域如何分布？

分析物种多样性的最简单方法或许是计数物种。然而，这不是看上去那么简单，因为分辨物种绝非那么容易。可能不难区分老虎和大象，但在给定的时间或位置，如果人们知道正在寻找什么并可分辨实际看到的東西，那么更容易知道一个物种是否存在或不存在！对许多人而言大量不同种类的甲虫看上去很相像。因为许多鸟类一生中并非全生活在一个国家的疆界内，你如何衡量一个国家有多少种不同的鸟类？一些国家有很少的湖泊和河流，因此没有太多的海洋物种。

一个大国的某些地区可能为这些地区不同物种的数目而引以自豪（例如美国的一些州），但是区域的大小、它的纬度（在北半球，更偏北的纬度一般说来有更少的物种）及该地区的人口可能决定了这些物种的数量。撇开上面暗示的许多微妙问题，如何来“测量”存在于特定区域的物种个数？也许最简单的措施是数物种的数量——采取普查。但是，从人口计算我们知道，在纽约州计数人口（不亚于数松鼠）的问题，不是一件容易的事。不令人惊讶地，人们借助于统计这个工具，使用部分信息（样本）以获取所涉及的种群信息。有人可能会尝试通过使用样本采集站来回答有关物种存在或不存在的问题。

数学可能参与的是在什么地方建立样本研究采集工作站。将它们放的位置相互靠近可能不会充分显示全范围的物种多样性，但它们彼此靠近可能有助于“理清”物种在特定区域出现的一致性。也有其它关于位置的问题，取决于计数的本质。要了解一个近似圆形的湖的海洋多样性，应该怎样组合岸边与湖内的采集工作站？因为不同的海洋生物可能出现在不同的深度，水中的采集站应放多深？

因此，为简单起见，让我们想象一下，寻找在湖中的“生物”（鱼、鳗鱼等），在不明显的地方已选择并安装了两个陷阱。为了让你思考一些问题，下面人造的例子显示了两个采集工作站的结果，它们分别位于某个区域的东端与西端，列出的是动物（比如说，啮齿动物）的计数结果及所属物种。

| | 东端 | 西端 |
|-----|----|----|
| 1 | 8 | 1 |
| 2 | 0 | 2 |
| 3 | 2 | 0 |
| 4 | 5 | 0 |
| 5 | 1 | 1 |
| 6 | 1 | 0 |
| 7 | 0 | 3 |
| 8 | 2 | 0 |
| 9 | 0 | 5 |
| 10 | 2 | 0 |
| 11 | 1 | 0 |
| 物种数 | 8 | 5 |
| 动物数 | 22 | 12 |

表 1：两个采集站记录的物种

可以立刻注意到的是，一些物种没有在这两个采集工作站之一出现，当某个物种在两