

农作物病虫害预测预报参考用书

昆虫生态学的 常用数学分析方法

鄢祥光 編著

农业出版社



农作物病虫害预测预报参考用书

昆虫生态学的常用数学分析方法

鄔祥光 編著

馬世駿 校閱

农业出版社

內 容 提 要

本书以实例形式介绍了有关昆虫生态学的常用数学分析方法。内容分为四大部分：首先从实态调查的基本知识入手，论述了一群量数本身的变化规律及其应用，然后从昆虫生态学的角度，介绍了昆虫种群时间、空间概念的常用数学分析方法，论述了两群或两群以上量数之间的变化规律及其应用，最后针对田间试验，介绍了进一步掌握昆虫种群特征的数学分析方法。因此本书系一部比较全面的昆虫生态和数学相结合的实用书籍，其中所采用的例子，都考虑了病虫害资料的数学特征，大多数是生产实践中所常遇到的，因此容易明了，可供昆虫科学工作者、植物保护工作者、病虫害预测预报站，以及高等院校生物、植保等专业的参考。

农作物病虫害预测预报参考用书 昆虫生态学的常用数学分析方法

郎 祥 光 編 著
馬 世 駿 校 閱

农 业 出 版 社 出 版

北京老钱局一号

(北京市书刊出版业营业许可證出字第 106 号)

新华书店北京发行所发行 各地新华书店经售

农业出版社印刷厂印刷装订

统一书号 13144.145

1963 年 10 月北京制型

开本 850×1168 毫米

1963 年 12 月初版

三十二分之一

1963 年 12 月北京第一次印刷

字数 383 千字

印数 1—3,000 册

印张 十五又二分之一 封面二

定价 (10) 二元六角

序

数据分析是生物科学研究工作中的一个重要步骤，通过調查或試驗取得了大量数据，如何从这些数据中发现问题，并进而推导出結論来，首先要經過一个数据整理和分析过程，因此，要求生物学工作者掌握一些常用的数学分析方法、根据研究对象和数据的特性，去选择应该采用的分析方法。有时一項研究結果，需要运用几种方法进行比较，从不同的分析結果中，可能发现不同的問題。

在社会主义农业生产建設的带动下，近十年来昆虫生态学及害虫预测预报工作已經有了显著的发展，各地农业科学研究机构及病虫预测预报站都积累了許多有关害虫数量消长的資料，这些資料是我們今后提高害虫预测预报及防治技术措施的基础，如何把这些資料有效地利用起来，則涉及两方面的問題：一、提高昆虫生物学的理論基础；二、熟悉若干分析方法（其中数学分析方法占着重要的地位）。鄔祥光同志是一位昆虫生态学工作者，在水稻螟虫的预测预报中积累有不少的經驗，这次又广泛搜集国内有关資料，結合他自己的体会，編写成这本常用数学分析方法，应当說，这是各地从事预测预报的同志及一部分生态学工作者当前所急需的一本工具书。相信在今后的预测预报的提高以及昆虫生态学研究中将發揮一定的作用。

这本书的内容是相当丰富的，与一般生物統計书籍比較，有特点，例如它是以昆虫生态学的问题为对象，比較系統地介紹算式的来源，以及由簡單到复杂的計算方法。当然，书中某些尚須簡化，另外还待补充若干昆虫发生预测的实例，这些都

各地有关方面的工作同志，不断予以指正和补充，作为一个昆虫生态学工作者，我对这本书的出版，表示欢迎、并寄予良好的愿望。

馬 世 駿

1962 年除夕于北京

作 者 序

在祖国当前社会主义飞跃建设的形势要求下，全国昆虫生态学的研究和农作物病虫害预测预报工作均不断获得新的进展，为了紧跟上形势，对生态资料的数学处理分析的知识，要求就比较迫切。我写作这本书，作为抛砖引玉，为感谢党的多年培养，就个人学习所得的点滴知识和体会，试以个人意见，汇集各地昆虫生态研究和预测预报实践资料，草成此稿，并编纂一些较常用的经验公式和方法，在编纂中，为了使叙述系统化以及安排上的便利，曾摘录了前人的一些个别的篇幅，此外，还附带介绍了日本气象预测学家寺田一彦博士所创造的各种差异显著性测定用的计算尺，以供参考。

本稿得蒙中国科学院动物研究所马世骏老师在百忙之中费神校阅，对本书整个编排、取材、以至具体词句，均作了指导，使本稿得以完成，同时并蒙农业部植物保护局束炎南技师等提出许多意见，谨此致以衷心感谢。

郎 祥 光

于中国科学院中南昆虫研究所

1962 年冬，广州

目 录

序

作者序

緒論	1
第一篇 昆虫种群实态調查統計数字的基本性格	5
一、昆虫調查統計数字的主要特点	6
1. 可量性状和可数性状	6
2. 有效数字	9
3. 統計代換	10
4. 相对单位和绝对单位	25
二、集中性常数及其应用	25
1. 平均数	26
2. 众数	38
3. 中位数	47
4. 三种集中性常数的关系	50
三、分散性常数及其应用	50
1. 平均误差	51
2. 标准差	53
3. 平均数标准差	61
4. 或然差	65
5. 变异系数	67
第二篇 昆虫种群空間結構研究的常用数学分析	70
一、昆虫野外集团的概率分布	71
1. 概率的基本概念	71
2. 概率分布表和概率分布曲綫	73
3. 常見的几种昆虫种群空間結構的概率分布型式	76
二、正态概率分布	84

1. 正态曲线的来源	84
2. 正态曲线的性质	90
3. 正态曲线的横坐标及曲线下的面积	92
4. 正态曲线的拐点	108
5. 正态曲线的纵坐标和基本正态曲线	116
6. 正态曲线的计算方法	117
7. 偏态系数和峰态系数	137
三、概率分布适合性测验	140
1. χ^2 的意义	141
2. 正二项式分布适合性测验	146
3. Poisson 随机分布适合性测验	147
4. Neyman 核心分布适合性测验	149
5. Polya-Eggenberger 核心分布适合性测验	151
6. 嵌纹分布适合性测验	153
四、大数定律与取样技术	157
1. 大数定律原理	159
2. 取样误差	162
3. 在预定精确度下决定取样本数	166
4. 几种常用取样方法	171
五、样本间的比较	177
1. 差数的平均及其标准差	177
2. 根据平均数来测验样本的可靠性和样本间差异的显著性	179
3. 根据变量来测验样本的可靠性和样本间差异的显著性	197
4. 根据变异系数来测验样本的可靠性和样本间差异的显著性	206
第三篇 昆虫种群时间数列研究的常用数学分析	208
一、一次函数规律	209
1. 直线相关及相关系数	210
2. 直回归	220
3. 直回归线在昆虫生态学上的应用	230
4. 回归系数及回归线的估计标准误差	238
二、二次函数规律	240
1. 曲线回归	241
2. 曲线相关及相关率	245

3. 阶差法	247
4. 分割和法	256
5. 指数函数曲线	258
6. 抛物线及双曲线	267
7. 表现为二曲线之和的曲线	284
三、高次函数规律	294
1. 偏相关及多相关	295
2. 多元回归	300
3. 多次回归	343
4. 正态积分函数曲线	345
5. 偏峰态曲线和致死中量	362
6. 低峰态曲线	365
7. Logistic-曲线	368
8. 周期性的波纹曲线	376
四、修匀经验曲线	394
1. 直线的移动平均法	394
2. 抛物线的移动平均法	402
五、长期趋势及周期节奏	412
1. 长期趋势	413
2. 短周期节奏	413
3. 长周期节奏	415
第四篇 昆虫生态田间试验的常用数学分析	420
一、随机排列试验的变量分析	421
1. 变量分析的基本原理	421
2. 随机排列设计	427
3. 拉丁方排列设计	431
二、顺次排列试验多个处理的比较分析	436
1. 对比排列设计	436
2. 多次重复排列设计	440
三、互变量分析	446
1. 互变量分析的性质及其应用	446
2. 互变量分析方法	448
四、复互变量分析	455

1. 复互变量分析的性质	455
2. 复互变量分析方法	455
附：常用公式索引	460
附：常用工具图表索引	466
附图：统计用计算尺及其使用方法	470
附表：自然对数检索表	471

緒 論

昆虫生态学应用的数理分析是应用数学中的一个特殊部分；又是数理統計学和昆虫生态学的一个交叉部分，它是用来研究昆虫数量变化的規律，或者是探索病虫发生发展在具体历史条件下表现在量方面的規律。因此它并不同于研究数为对象的理論数学(Mathematics)；又不同于研究社会中量的現象的数理統計学(Statistics)；也不同于研究昆虫发生发展具体事实的昆虫生态学，而是以理論数学、数理統計学和昆虫生态学三者为基础，在昆虫生态科学技术实践上加以发展的应用数学。因此可以把它列为数理生态学(Mathematical ecology)的一个内容。

本来昆虫生态学的中心問題，就是昆虫种群的数量变动，这就包含着丰富的“数量”的意义。馬世駿*(1962)曾在中国昆虫学会1961年及1962年学术討論会中对昆虫生态学中的数量、時間及空間的概念加以概括性的总结，认为任何一个种群(某一時間在单位空間内所有的个体的組合)都占据一定的空間，任何一个种也都有在一定空間内通过時間积累而組成的生活历史，因此衡量一个种群的昌盛与否，通常根据数量和空間两个指标，而数量的多少和空間的大小又与時間直接相联系，一方面時間长短表达了种群增长的快慢，同时也表达出時間因素在数量积累上的作用。空間和時間

* 馬世駿 (1962): 昆虫生态学的若干新进展，中国昆虫学会 1961 年学术討論会会刊 15—24。

* 馬世駿 (1962): 昆虫种群的時間、空間结构及其在数量变动上的意义，中国昆虫学会 1962 年学术討論会会刊。

之所以被看成是数量生态学上的因素 (Preston 1960, Alan 1960, 等), 主要由于在许多重要生态学问题上都蕴藏或直接关联这两个因素的作用。例如任何一个昆虫种群, 都需要占有繁殖场所、食物供应范围及活动与分布的领域。空间因素所显示的重要性, 随着因时间所增加的数量而变化, 同时数量与时间因素在种群动态中所显示的作用, 又随空间的变化而转移。

空间和时间因素如果作为一个科学数据表达出来, 它就含有广泛的“数”的意义, 而数量变动则含有广泛的“量”的意义, 数和量的关系在昆虫生态学中是相互联系的、又是相互制约的, 因而是辩证的统一。昆虫及其他生物共同生活在一定的自然地理环境中, 环境结构本来就是一系列时间的反映, 每个特定阶段的环境, 都表现有它的时间外貌, 种群的时间结构及环境变化的时间数列, 反映了物质能量转换的序列关系, 构成了动物体内的代谢作用与动物和环境间的代谢作用所组成的生态系统 (Ecosystem), 即是空间结构与时间机制的生态系统。生态系统中的活有机体成员, 常形成了一定的群落 (Community), 而群落是一个大集体 (在数学上的意义是一个大集团), 它是由许多集团 (Group) 即种群所组成的。昆虫生态学的根本任务正是要揭露这些大小集团的运动变化规律, 在这里面必然地产生两个问题: 一个是各种环境因素怎样作用于昆虫的发生和发展? 它们之间存在哪些规律? 另一个问题是昆虫在受到各种环境因素作用后, 本身在发生和发展上产生了哪些反应? 遗憾的是: 目前我们对各种生态因素的作用, 及其相互联系和相互制约的规律, 还未能完全掌握! 我们能够初步掌握的, 仅仅是其中一些较重要的主导因素, 但主导因素是会转化的, 其中一些我们认为不很重要的因素, 在某些情况下也会转化为重要的因素或主导因素, 对昆虫的发生和发展予以较大的影响, 或者有些很重要的因素我们还未找到; 此外在昆虫本身发生和发展的反应问题上, 目前和今后我们都不能把昆虫整个集团的所

有个体一一去数，还是只凭取样方法去調查計算，而取样方法所得的結果总有一定的出入(随着取样方法而有不同)，它只能代表一定情况下的一定准确程度，并不是絕對的真正数字。基于这些原因，可見有必要研究一些較科学的数学方法，使我們能較精确地找出各种环境因素与昆虫发生发展的規律，和較精确地衡量昆虫在数量上所产生的的一系列反应的規律。

由于集团中的所有个体是无法一一去数的东西，因此要用一些集中性常数和分散性常数去代表它和衡量它，而集中性常数和分散性常数有賴于取样方法而求得。取样方法是針對概率分布而不同的，其准确程度又有賴于測驗其差异性与适合性的显著性。研究这些問題，将分別在本书第一、二篇中列述。

昆虫生态学是研究昆虫种群(Population)、群落(Community)或其他生态系统(Eco-system)、以及构成这些系統的有机体(Organism)与环境間相互关系的科学，因此在进行数量分析时，首先要研究各种环境因素与昆虫发生发展的相关規律，在充分了解相关規律之后，根据所求得的規律来研究其規律的具体分布軌跡。研究这些問題，将分別为一次函数、二次函数、高次函数等，在本书第三篇中列述。

本书以实例形式，介紹了有关昆虫数量生态常用的一些数学分析方法，其中所采用的例子，大多数是昆虫生态研究中、特别是昆虫发生預測中所常遇到的，但希望讀者不要机械搬用，而且对具体問題要作具体分析。数学只能是作为一种工具来帮助我們更深入地了解生物学的規律，因此在任何一个数学处理及任何一个数学分析过程中，必須充分考虑其生物学意义。因为生物的規律絕不是机械的循环，如果企图以几个数学公式来概括变化多端的昆虫生态現象，或者以数学的規律来代替生物学規律，这显然是不正确的。因此把多年資料埋头計算，完全不管計算資料的本质上的表現，結果只能說是完成了一些数学形式，这并不是正确使用数理

分析的方法,例如,在病虫预测预报时,完全不管哪一个时期、哪一种生态因素与昆虫发生的生态学关系,只是大量地把各月各旬的气温、降水量、初雷日、终雷日、上一世代期距……等与三化螟蛾的始见期和盛发期的相关系数和迴归式都计算一遍,企图借此找出较大的相关关系而用于发生时期的预测,这样运用数学的结果,只能是使预测预报工作人员长期坐在办公室里摇计算机,而对现实毫无帮助!因为即使在那么多的计算工作量中,偶然地找出一两个较大的相关系数(即使是达到0.9以上的相关系数),但可以认为这只是偶然的,例如在气温问题上,因为昆虫的发生时期的迟早和久暂,并不是刻板地每年都是同某月或某旬的气温有关,而是同一定的积温有关,有些年份,設3月上旬才达到 16°C (假设这是影响昆虫解除滞育的温度),但在另一些年份,2月下旬或中旬已经达到 16°C ,而3月上旬已经超过 16°C 了。显然如果我们去计算3月上旬的气温同某种昆虫发生期的相关关系,倒不如去研究或计算达到 16°C 的那一句的气温与该种昆虫发生时期的相关,会更加有意义些;会有更显著的相关。

有时我们通过辛勤劳动,取得了较多的调查试验结果,但往往只是罗列现象,没有能得出明确的结论,也找不出通过调查试验后应得的规律,有时还会把必然性的联系看成是偶然性的现象;另一方面有时又根据一些数据不足的資料,得出一些不够精确的结论,把一些偶然性的现象,看成是必然性的联系。在这些情况下,就必须借助于进一步的数学处理,加以检查,然后才可以找出事物的规律,获得比较正确的结论。

因此,本书的写作,只是希望在介绍一些数量生态常用数学分析方法的基础上,对資料的数学处理方法及分析问题上有所帮助,以供作实践上的参考。

第 一 篇

昆虫种群实态調查統計数字的基本性格

种群既然是許多个体的組合，因此在一定的時間和一定空間的基础上，对种群实态的描述、对种群实态的調查研究的結果，首先就表現在数量上湧現一群量数，如果这群量数的数值是可以无限多的，就叫做变数（每次对该变数測量的数值叫做該变数的变員数）。例如粘虫的虫口密度，每平方米的昆虫个体数可以由个位数增到十位数、百位数、甚至許多位数；稻螟为害的枯心率也可以由很小增到很大，甚至 100%。所以变数可以用来描述种群实态的数量形态。

这样，变数的調查方法和数字的处理方法就成为种群实态描述的前提，数字的調查方法，即概率分布和取样技术等問題，容俟本书第二篇中列述，但在进行調查試驗之后，所获得的往往是許多杂乱无章的原始数字，这些数字必須善于收集、整理，才能不走样地、确切地表达出昆虫种群实态的客观眞面目，这是一件不可忽視的基础工作。

一群量数有两个基本性格：即其集中性和分散性。集中性和分散性的辯証統一就是該群量数变异的規律，如果我們把这个問題处理得恰当，就能充分地把調查对象的数量性状描画出来。

一、昆虫調查統計數字的主要特点

1. 可量性状和可数性状

上面提到在进行昆虫实态的調查研究时,必然得到一批变数,变数实质上就是一种“性状”,这些性状可以分为可量性状和可数性状两类,亦即是变数的連續性变数和不連續性变数两类:

(1)可量性状 即是連續性的变数(Continuous variation),一批变数,如果把它由小至大排好队以后,在每两个变数中間仍存在无限多的連續变数,因此这类变数是可以精确地測量出来的,叫做可量性状。例如气温的上升和下降是連續性的,因此我們把每一次測定的气温,由小至大排好队,就能清楚地看出它的連續性。設某日內的气温由 20°C 連續上升至 30°C ,我們測量多次的結果,設若得到 20、21、22、24、25…… 30°C 等,但在这多次測量中間,气温总是逐步上升的,因此我們可以在 20°C 与 21°C 两个变数的中間,找出許多个变数,如 20.1、20.2、20.3……等,而 20.1 与 20.2 两个变数中間,仍然可以找出許多个变数,如 20.11、20.12……等等,可見这些連續性变数是无限多的,所以气温的測定值是属于可量性状。

(2)可数性状 即不連續性的变数(Discontinuous variation),在由小至大排好队的一批变数中,两个相邻的变数中間是无法存在另一个变数的,因此这类变数只可逐个去数,而不可以去量,如昆虫数 20 与 21 头之間,絕对不可能有 20.5 头,这是因为昆虫个体数是不可能非整数的,所以昆虫个体数的調查是属于可数性状。

昆虫調查統計數字的主要特点之一,就是昆虫的數字資料以可数性状占居优势,属于可量性状的甚少,所以在进行昆虫調查統計时,應該充分重視这个数量性状特点:

資料数字的連續性和不連續性的性狀，在實踐上有时用另一形式表現出來，例如研究比較几个水稻品种之間的抗病性，就是一种不連續性数字，因为甲品种和乙品种的发病率是毫无关系的，各自独立的，不是从甲的发病率漸变为乙的发病率，它們之間沒有从低至高或从高至低的联系，有人把这样的資料当作是連續性的資料来对待，繪成图 1 的漸变曲綫图，就不够明确。本来繪制該图的要求，只是比較各个水稻品种之間的发病輕重，因此应该用图 2 的直方图的方法来表示这种不連續性資料，才能达到容易比較的目的，因为不同水稻品种的发病率不同于每天的溫度、湿度等呈漸变曲綫規律的事物，如果像图 1 那样，用連續性的漸变曲綫去表示不連續的資料，是容易使人发生錯觉的。

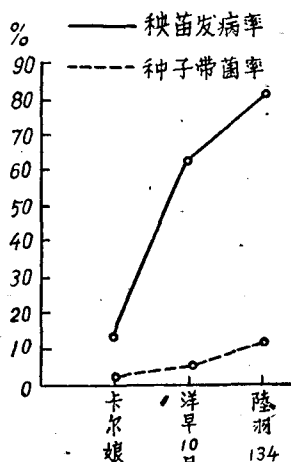


图 1 不同稻种稻瘟病种子带菌率与秧苗发病率的关系
(据广东合浦)

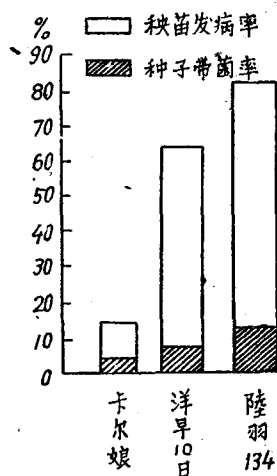


图 2 不同稻种稻瘟病种子带菌率与秧苗发病率的直方图
(作者根据图 1 改制)

同样可以理解，每天的降水量也是不連續性資料，因为从降水量最多的那一天，到最少的那一天中間，并不是逐漸上升或下降的，昨天可以降几百毫米的雨，但今天可以突然不下雨，暴雨之前