

0100315

橡胶·林业生物统计与田间试验设计

内部刊物 注意保存



广东省农垦厅编

1 9 6 5

前 言

自从毛主席提出三大革命运动的伟大号召之后，我省农垦生产战线和全国各地一样，以阶级斗争为纲、以生产为中心，有计划、有步骤开展科学实验活动。许多农场成立了科学技术推广站、科学实验小组，普遍地开展了样板田工作，群众性的科学实验活动已蓬勃地发展起来。

为了满足科学实验活动发展的需要，我厅举办了两期生物统计与试验设计训练班，培训了一批在职技术干部。现将训练班上采用的讲义补充修订，编印了《橡胶林业生物统计与田间试验设计》一书，以普及这门科学技术知识。这本书是我国以橡胶林业为对象的生物统计与试验设计的专门论著，它广泛地搜集了橡胶生产上的具体实例，详细地阐明了统计分析和试验设计在橡胶生产试验和调查研究的应用。全书二十二章，内容分为生物统计与田间试验设计两大部分。在统计方面，指出生产试验和调查研究上如何做比较，如何找差距，怎样分析数据，怎样判断生产技术措施的效果等问题。在设计方面，指出如何针对橡胶树多年生木本植物的特点结合我们垦区的环境特点去布置田间试验，告诉我们怎样去找寻合理的对比关系，怎样使试验误差减少到最低限度，从而获得准确可靠的试验结果。本书为橡胶林业的试验设计、调查研究、整理数字资料提供了科学的方法，可以作为农、经作物的试验和调查工作的参考，也可作教材用，适合广大技术人员、特别是从事橡胶栽培育种工作人员阅读。我们掌握了这门科学技术，在调查和试验工作上就可少走弯路。

要学好和应用好这门科学技术，首先必须认真学习毛主席著作，特别是要学习好毛主席的四篇哲学论文，以毛泽东思想为统帅，做好科学实验。科学真理来自实践，我们一定要深入实践，坚持理论联系实际，以理论指导科学试验。试验和调查所得的数据，要用唯物辩证的观点来全面分析，抓住间

題本質和主流，反对孤立片面地以数字就事論事，代替了具体問題具体分析；反对罗列数字，搞煩瑣哲学，使人不得要領；反对脫离实际，生搬硬套，机械搬用公式；反对为統計而統計，脫离生产的單純統計观点。要很好运用这門科学技术，还必須与总结群众生产經驗密切結合起来，發揮它更大的作用。

本书是我厅第一屆生物統計与試驗設計訓練班教員邓鳴科、林培基、邱蒙、杨成龙等四位同志，参照生物統計专家王鉴明的生物統計学和試驗設計两书，結合广东垦区生产实际和試驗設計的經驗及橡胶树是多年生林木的特点，参考了有关文献編写的；并經第二屆訓練班教員邱蒙、张少若、刘訓初、丘迁以及全体學員的討論訂正；最后由我厅热带作物处詳細审校出版。編写期間得到王鉴明、杨瑾英和徐广泽同志的指导，封面請徐元文同志設計，在此，一并致以謝忱。由于經驗不足，錯漏之处，敬祈指正。

广东省农垦厅

1965年12月

目 录

生 物 统 计

第一章 緒論	(1)
第一节 生物統計学的定义和基本內容	(1)
第二节 掌握生物統計方法的重要性	(3)
第三节 生物統計主要術語的涵义	(3)
第四节 生物統計的基本符号的运算及符号的定义	(6)
第五节 农艺、土化和昆虫生物統計的特点	(8)
第六节 学习生物統計应有的認識	(9)
第二章 試驗資料的整理	(10)
第一节 試驗資料整理的基本內容	(10)
第二节 簡化运算	(10)
第三节 次数分布表与次数分布图	(11)
第三章 平均数 (集中常数)	(17)
第一节 平均数的意义	(17)
第二节 平均数的种类	(17)
第三节 算术平均数的基本性质	(17)
第四节 算术平均数的求法	(18)
第四章 变异常数 (离中常数)	(22)
第一节 变异常数的意义	(22)
第二节 变异常数的应用范围	(23)
第三节 变异常数的种类	(23)
第四节 标准差的特性及标准差的求法	(25)
第五节 变异系数的求法	(29)
第五章 t 測驗——差异显著性測定	(32)
第一节 統計数值差异显著性測定的应用意义	(32)
第二节 t 測驗的基本原理	(32)
第三节 t 測驗的方法和实例	(39)

第四节	t 测验公式一览表	(69)
第六章	卡方测验 (χ^2 测验)	(70)
第一节	卡方测验的意义	(70)
第二节	适合性测验	(73)
第三节	独立性测验	(76)
第七章	变量分析法	(85)
第一节	变量的意义	(85)
第二节	变量分析法的应用意义和基本内容	(85)
第三节	变量分析法的应用范围	(87)
第四节	应用变量分析法的条件	(87)
第五节	一般的变量分析法的计算步骤和方法	(88)
第六节	多个处理的均衡的比较	(96)
第七节	需要进行统计代换的变量分析法	(104)
第八章	简捷近似统计分析法	(114)
第一节	样本均数可靠性测定	(114)
第二节	两样本均数差异显著性测定	(118)
第三节	简捷近似变量分析法	(122)
第九章	相关	(130)
第一节	简单相关	(130)
第二节	净相关 (部分相关)	(139)
第三节	复相关	(141)
第十章	回归	(144)
第一节	回归的意义与作用	(144)
第二节	直线回归	(145)
第三节	配合抛物线的曲线回归	(152)
第十一章	互变量分析法	(167)
第一节	互变量分析法的意义和作用	(167)
第二节	在什么情况下运用互变量分析法	(168)
第三节	互变量分析法计算的步骤和方法	(168)
第十二章	抽样技术	(177)
第一节	抽样的必要性及其概念	(177)
第二节	样本大小的决定	(177)

第三节	巢式抽样法	(181)
第四节	双重抽样法	(190)
第十三章	生物统计应用于生产调查研究应注意的几个问题	(197)

田间试验设计

第十四章	试验设计的基本概念	(201)
第一节	田间试验设计的意义与作用	(201)
第二节	试验的内容	(201)
第三节	试验设计的三个基本原则	(206)
第四节	试验的主要误差及其控制	(209)
第五节	几个具体问题	(213)
第十五章	空白试验	(216)
第一节	空白试验的意义及其功用	(216)
第二节	利用空白试验确定区组小区的排列方向	(217)
第三节	空白试验应用于决定试区的适当形状和大小	(222)
第四节	空白试验应用于决定试区的适当重复次数	(230)
第五节	空白试验的一些试验结果——几种主要作物的试区形状及大小	(233)
第十六章	随机区组设计	(234)
第一节	单因子随机区组设计	(234)
第二节	复因子随机区组设计	(242)
第十七章	拉丁方设计	(258)
第一节	拉丁方的定义	(258)
第二节	拉丁方设计的优缺点	(259)
第三节	设计方法	(259)
第四节	分析方法	(260)
第十八章	对比法设计与分析	(263)
第一节	对比法设计田间排列的几种形式	(264)
第二节	对比法设计的统计分析方法	(266)
第十九章	裂区试验设计	(282)
第一节	裂区试验设计的特点	(282)
第二节	裂区设计在橡胶试验的应用意义	(283)
第三节	应用裂区试验的原则	(285)

第四节	裂区試驗的設計与分析实例	(286)
第二十章	多品种比較試驗	(300)
第一节	分組共同标准种法	(300)
第二节	二元二等組群(格形設計)拟复因子試驗	(308)
第三节	二元三等組群(三重格形設計)拟复因子試驗	(318)
第四节	平衡不完全随机区組	(322)
第廿一章	混杂試驗設計技术	(328)
第一节	混杂試驗設計的概念	(328)
第二节	2^n 混杂試驗之混杂方法	(330)
第三节	2^3 混杂試驗部分混杂設計的实例	(335)
第四节	2^3 完全混杂之分析示例	(340)
第五节	每因子具有三平准的混杂設計及統計分析方法	(343)
第六节	3^3 混杂試驗之分析实例之一(仅有一重复)	(347)
第七节	3^3 混杂試驗之分析实例之二(有两重复)	(351)
第廿二章	資料遺缺的处理	(356)
第一节	缺株的处理	(356)
第二节	随机区組設計缺区的处理	(357)
第三节	拉丁方設計缺区的处理	(361)
附表: 一、	費雪氏 t 表	(365)
二、	費雪氏卡方 (χ^2) 表	(366)
三、	斯奈迪格氏之 F 与 t 表	(367)
四、	百分数換算为角度值表	(371)
五、	r 与 R5% 与 1% 显著值表	(375)
六、	費雪氏轉 r 为 Z 表	(376)
七、	費雪氏轉 z 为 r 表	(377)
八、	費雪氏均衡多項式表	(377)
复习思考题与练习题		(380)
参考文献		(404)

生物統計

第一章 緒論

第一節 生物統計学的定义和基本內容

我們在認識自然界、生物界的过程中，經常注意到事物的数量的变化，通过反复的生产实践和科学实验，掌握了一定的数据，进行統計分析，从中找出規律——即找出数据中共同的特征和变化发展的規律，从而获得了明确的概念，对事物有了客观的認識。再进而指导实践，使生产和科学技术水平不断提高，不断向前发展。

自然界，生物界的各种現象是千变万化、錯綜复杂的，但是，一切事物或現象都是相互联系、互相制約的，事物的变化和发展是遵循着一定規律而不是杂乱无章的。掌握了科学方法就能揭露自然界生物界客观現象的相互联系相互約制的辯証規律。

我們說某項生产技术措施的好坏，就是通过調查研究和科学实验，掌握一定数据，进行比較分析，从而作出結論的。

要判断生产技术措施的好坏，要分析生物界客观現象的可靠性和真实性，必須掌握正确的統計分析方法。所謂分析，就是怎样做比較的問題。

生物統計学就是：在农业科学技术中，研究如何运用数理統計的原理和方法去寻找合理的对比关系去分析数据，找出差距，判断数据可靠不可靠，效果是真是假的一門科学。更确切地說，生物統計学是应用統計邏輯来研究生物群体內个体間的变异性和生物性状观察过程中的誤差的一門科学。

通过数据分析，再結合专业知識，我們就能对生物界調查研究和試驗結果作出全面的正确的論断，得出可靠的結論。更确切地反映客观規律。从而能地改造客观自然界。

生物統計学的基本內容有四：

一 以样本估計集团

就是通过局部去了解全体的一种方法，科学家要知道麻雀的特征特性，就用解剖少数麻雀的方法去了解全体麻雀的特征特性，这是一种少而精的方法，也就是通过对样本

的研究分析来認識群体的特性特征的一种方法，我們日常所接触到的技术工作中如样本区生长調查、产量調查、植株发病情况調查、栽培技术經驗的調查、或者栽培試驗和品种比較試驗等等，往往都是通过对样本的調查研究来了解群体的特征的。又如我們要进行橡胶树的有性杂交(人工授粉)，当然必須首先了解亲本无性系的开花习性，但是我們不可能也不必要对亲本无性系各齡的全部植株的所有花序都去調查，而只需要选有代表性的几种树齡的少数植株的少数花序进行調查，从而根据这些抽样調查結果来推断該亲本无性系的开花习性。抽样技术一章和統計数值可靠性測定一章就是研究这些問題的。

二 确定某种性状的数量特征与变异情况

例如研究某无性系在不同种植密度的栽培条件下植株树围的增长速度，产量(包括单株产量和单位面积产量)，树皮再生力等。調查分析試驗地的肥力情况和肥力变化趋势；橡胶有性系或无性系的产量水平，和个体間产量变异的情况，无性系的抗风力、抗寒力、抗病力等，平均数和变异常数一章就是研究这个問題的。

三 判断試驗和調查結果的可靠性

一、測驗試驗和調查結果与理論假設的符合程度：例如測驗橡胶有性系的产量性状在无性繁殖中传递能力是否符合亲本实生母树高割綫高产的无性后代也高产的理論假設，对指导选种具有重要意义。

二、測驗試驗和調查結果各处理(或各品种)間的差异程度：例如不同肥料处理对橡胶树围增长和增产的影响效果的比較；在无性系系比区中，比較参試无性系間产量和其他付性状的差异显著程度，从而选拔出較优良的无性系；測驗不同割胶制度的增产效果，或不同刺激剂处理的增产效果等是。两均数差异显著性測驗、变量分析、互变量分析和卡方測驗及第八章簡捷近似統計分析法等章就是研究这些問題的。差异显著性測定的方法是生物統計的核心部分。

四 測定两种或两种以上性状之間的相关关系(包括相关与回归的研究)

例如在預見性研究上，研究橡胶树早期产量鉴定——刺检与試割产量鉴定与正式割胶产量的相关；橡胶幼苗特性特徵与高产性状的相关研究；綜合性的气候因素与病虫害发生的相关——用以指导病虫害預測預报；又如研究种植密度与抗风力、单位面积产量的相关、迴归关系等。相关、迴归就是研究有关这方面問題的。

綜上所述，生物統計学的作用可归结为三点：

- 1.为調查研究和試驗研究結果作出科学的判断，判明差异的真假，可靠的程度，克服經驗主义，使对結果的分析更加精确可靠。
- 2.为試驗本身提供了合理的設計的依据。
- 3.提供了整理数字資料最适当的方法。

第二节 掌握生物統計方法的重要性

一 生物統計学的应用范围日益广泛

統計方法早已广泛应用到工业、农业、經濟、生物、医学、物理、化学和气象等科学技术部門了，就以生物統計学來說，也已广泛应用于耕作与栽培、选育种与遗传、病虫害防除与預測預报、丰凶預測、土壤微生物的研究、毒理与及抽样技术等方面，并且广泛应用于田間試驗和室內試驗，在國內，生物統計学已得到普遍重視。

多年生热带作物如橡胶、油棕等經濟作物也可以广泛地应用生物統計学来研究抽样調查技术；研究某种栽培条件下某品种的生长、产量、抗性特征及其变异情况；比較不同栽培条件对作物生长、产量、抗性的影响；比較不同品种的生产力、抗逆力；比較不同农葯的防治病虫害，除茅的效果；分析两种或两种以上性状之間的相关和迴归关系，用以指导选育种、产量預測、病虫害和气象預測預报与及确定合理的栽培技术措施等等。

二 生物統計学是科研工作者必須掌握的工具

当进行某一項試驗或調查，对所得的資料进行整理分析下結論时，如果没有通过周密的統計分析，而仅凭直觉的观察和判断，这样不但不能作出正确的結論，有时反而作出似是而非的錯誤的結論。科研工作者，必須借重統計学工具，对試驗資料进行統計分析，得出統計結論，进而根据专业理論知識，对統計結論进行辯証的科学分析，最后作出科学結論。

我国在解放初期，在科学技术界中，对生物統計曾經有过不正确的偏見，因此，这門科学的普及和发展甚为緩慢，以致沒有能够在生产及科学研究中起到应有的作用。自从党提倡在学术中必須貫徹“百家爭鳴，百花齐放”的方針以后，尽管不同学派对某些問題尚持有不同的見解，但是对于生物統計学在生产中确能起指导作用，这一点則是无容置疑的。目前在农业生物研究中，“生物統計及試驗設計”已成为不可缺少的工具了，目前大部分农业生物学报，論文，期刊等内容应用生物統計計算方法相当普遍，如果不懂得生物統計学，就不但不能联系、渗透和借用別門科技知識，甚至連論文本身也看不懂了，因此，我們必須急起直追，认真地掌握好这一工具，这样才能使我們更好地接受前人的研究成果来正确地运用到我們的热带作物生产和科研工作中去。

第三节 生物統計主要术语的涵义

1. 变数：数据是自然界、生物界客观現象的反映，它的基本特点是具有变异性，就是說自然界生物界的任何一种現象总是存在着差异的，凡事物都有变异性，这是一个重要的特征，变数（性状）就是生物界数量特征的一种表現，变数是变員数的总和，也可以說由变員数组成的具体性状称为变数，例如橡胶苗木茎粗是一个性状，也是一个变数。

变数可分为两大类：

(1) 連續性变数：又称可量性状。变員数按大小排列，任何两个相邻的变員数之間，不管它們差异多么小，我們总可以再指定另一个变員数插入其中。例如株高、树围、树皮厚度、胶乳和干胶产量、干胶含量、有效 N.P.K. 含量、有机质含量、溫度、湿度、雨量、风速和气压等。可量性状一般为常态分布或偏态分布。

(2) 非連續性变数：又称可数性状。变員数按大小排列，任何两个相邻的变員数之間只相差一个个位整数，而不能再指定另一个变員数插入其間。可数性状可分为正二項式分布和潘松分布两种：

① 正二項式分布的可数性状：亚样本內含有的变員数数目較少，且有一定数目，各亚样本內含有的变員数均相等，出現某一性状的变員数的机率（所謂机率就是某事件中出現与不出現的数与两数和的比率）并不太小，而不出現的机率也并不太大。例如：100 粒花粉的发育花粉粒数，100 粒种子的发芽粒数，100 节甘蔗的螟害节数，100 株苗的芽接成活或者定植成活株数（100 为假設的亚样本內变員数数目）等是。

② 潘松分布的可数性状：亚样本內含有的变員数很多，且无一定数目（变員数数目不相等），其出現某性状的机率非常小，而不出現的机率則非常大，例如 1 两毛蔓豆种子中混杂有其他复盖作物种子粒数，亚样本虽然一律是 1 两，但是由于种子有大有小，有飽滿的和不飽滿的，所以亚样本內含有的变員数——毛蔓豆种子粒数无一定数目，亚样本間的变員数数目并不相等，而且其中混杂别的复盖作物种子很少。又如 100 亩林段內风害断干树株数，亚样本虽然一律是 100 亩，但由于林段內株行距不一定是百分之百准确，加以地形变化的影响，存在缺株弱株补換植苗，所以亚样本內含有的变員数——株数无一定数目，亚样本間的变員数数目并不相等，而且其中出現风害断干株数很少。又如橡胶无性系 100 个花序結实数。亚样本虽然都是 100 个花序，但是由于花序有大有小，花序上着生的雌花数目沒有一定，所以亚样本內含有的变員数——雌花数无一定数目，亚样本間的变員数数目并不相等，而且最后結实数又是很少，这些例子均可反映出潘松分布的可数資料的特点。

可量資料（連續性变数）与可数資料（非連續性变数）是两类性质不同的資料，其統計分析方法各有不同，应分別对待，因此，必須明确这两种变数的概念。一般而論，农艺上的数据大多数是可量資料，昆虫和病害資料大多数是可数資料。

2. 变員数：某一性状或变数的各个具体数值是为該变数的变員数。变員数也相当于在某一特定条件下的性状。例如同一設計的試驗进行多年，每年的試驗可視為一个变員数。又如某品种分布在垦区的各种类型环境中，該品种在某一类型区栽培环境条件中的产量，也可視為一变員数。

3. 变异：变員数間彼此差异的現象称为变异。变异的原因大致有三种：

(1) 个体間的差异；

(2) 同一个体不同观測次数存在差异；

(3) 同一集团不同样本或同一样本不同亚样本的数值間存在差异。例如在同一地段內抽样調查树围，每次随机抽样 30 株，抽样 20 次，如是不同的样本所得的平均树围数值并不相同，各平均值間存在差异。

4. 集团：集团乃指某事物的全体，包括属于该事物的所有个体。組成集团的每个成员称为个体。统计学上的母体、总体、群体或集团均为同义异名之词。

5. 样本：样本是由集团中所选取的若干个体，是某一特定时间（条件）内的观察资料，是该集团的一部分。

我們研究集团的特性是通过样本来研究的。故以样本的特性来估计集团的特性时，样本必须确有代表性。

样本中所含的个体数多少，叫做该样本的大小。样本根据其含有个体数多少而分为大样本与小样本两种。至于个体数多少才算大样本或是小样本，至今学者意见不一致，但一般以样本个体数30作为大小样本的界限者较为普遍， ≥ 30 时为大样本， < 30 时小样本。

6. 抽样：从集团中抽取部分变数来组成一个样本叫抽样。抽样技术是根据集团的性质，抽样可靠性的要求和抽样所能付出的代价（如劳动力、物力与费用等）来决定。

7. 集团统计常数：或称常轨数又称统计常数，用以描述集团而从集团资料测得的某种统计数值指标称为集团统计常数，这种数值可以表现集团的某种特征，如平均数、标准差、变异系数、回归系数、相关系数等均称为统计常数，是集团的真正数值，固定不变。

集团往往无法或不易测得，因而也无法测得集团统计常数，故通常均以样本统计常数来估计之。生物统计所研究的统计常数大都是样本统计常数。

8. 样本统计常数：或称估计常轨数，又称估计常数，用以描述样本而从样本资料测得的某种统计数值指标称为样本统计常数。样本平均数也就是样本统计常数的一种，样本统计常数由于所选样本的不同而有所变化。

9. 误差：广义的误差包括错误、恒误差和试验误差。

(1) 错误：这是由于工作上疏忽所造成的结果。例如，设计上的错误，处理操作上的错误，看错砝码、量尺、仪表、记录错数字，数值计算错误等是。错误可以避免而必须避免发生，一旦产生则误差较大，且难更正。若能复查的应及时复查更正，不能复查的应重做或取销错误数值，以免一误再误。

(2) 恒误差：由于观测人员的观测能力的差异或仪表失修所产生的偏差，这种误差是恒定的，一旦发现则易于改正或改算。

(3) 试验误差（简称误差或机误）：在试验中同一品种（或处理）在不同重复的差异称误差，这种差异是由机会做成的，由机会做成的不一致现象称为试验误差，亦称机误。如土壤、环境的差异，季节的变化，试验材料的变异，病虫害的骚扰，边际影响等等原因所造成的误差，这种误差的发生是由于机会所致，致因甚为复杂，或知其原因或不知其原因，只能尽量减少而无法避免或消灭。田间、工厂试验的试验误差较大，室内试验的试验误差较小；理化试验的试验误差较小，生物试验的试验误差较大。我们可以从技术改良和试验设计两方面着手来减少试验误差。

第四节 生物统计的基本符号的运算及符号的定义

1.
$$\sum_{i=1}^N X_i \equiv X_1 + X_2 + X_3 + \dots + X_N$$

左边是变数 X 的连加， Σ 读做 Sigma 即各变数总和， Σ 为连加号，符号

\equiv 即定义； $\sum_{i=1}^N X_i$ 即变数 X_i 从第一个变数连续相加至第 n 个变数，右边为展

开式，其中的每一个数字即为具体的变数。

例： $\Sigma X = 10 + 8 + 6 + 9 + 12 = 45$

2.
$$\left(\sum_{i=1}^N X_i\right)^2 = (X_1 + X_2 + X_3 + X_4 + \dots + X_N)^2$$

(各变数总和的平方)

例： $(\Sigma X)^2 = (10 + 8 + 6 + 9 + 12)^2 = (45)^2 = 2025$

3.
$$\sum_{i=1}^N X_i^2 = X_1^2 + X_2^2 + X_3^2 + \dots + X_N^2$$

(各变数平方的总和)

例： $\sum_{i=1}^N X_i^2 = 10^2 + 8^2 + 6^2 + 9^2 + 12^2$

$= 100 + 64 + 36 + 81 + 144 = 425$

株号	产量 X	X^2
1	10	100
2	8	64
3	6	36
4	9	81
5	12	144
N=5	$\Sigma X = 45$	$\Sigma X^2 = 425$
	$(\Sigma X)^2 = 2025$	

4.
$$\sum_{i=1}^N \sum_{j=1}^K X_{ij} \equiv (X_{11} + X_{12} + \dots + X_{1K}) + (X_{21} + X_{22} + \dots + X_{2K}) + \dots$$

$+ (X_{N1} + X_{N2} + X_{N3} + \dots + X_{NK})$

(等于各小区产量总和)

例： $\sum_{i=1}^N \sum_{j=1}^K X_{ij} = (3 + 6 + 1 + 10 + 7) + (4 + 9 + 3 + 11 + 6) + (2 + 7 + 4 + 14 + 8) + (5 +$

$8 + 4 + 12 + 9) + (4 + 9 + 3 + 15 + 5) = 27 + 33 + 35 + 38 + 36 = 169$

(双重连加式的展开， i 代表横行， j 代表直行)

处理一直行(j)

	1	2	3	4	5(K)
1	3	6	1	10	7
2	4	9	3	11	6
3	2	7	4	14	8
4	5	8	4	12	9
5	4	9	3	15	5
(N)					

重复一横行(i)

例: X_{13} 表示第一横行, 第三直行的变员数。 $X_{13} = 1$; 又如 $X_{24} = 11$ 是。

$$5. \sum_{i=1}^N \sum_{j=1}^K X_{ij} = \sum_{j=1}^K \sum_{i=1}^N X_{ij}$$

(各小区产量总和, 横行连加的总和等于直行连加的总和)

$$\text{例: } \sum_{j=1}^K \sum_{i=1}^N X_{ij} = (3+4+2+5+4) + (6+9+7+8+9) + (1+3+4+4+3) + (10+11+14+12+15) + (7+6+8+9+5) = 18+39+15+62+35 = 169$$

$$\therefore \sum_{i=1}^N \sum_{j=1}^K X_{ij} = \sum_{j=1}^K \sum_{i=1}^N X_{ij}$$

$$6. \sum_{i=1}^N \sum_{j=1}^K X_{ij}^2 = (X_{11}^2 + X_{12}^2 + \dots + X_{1K}^2) + (X_{21}^2 + X_{22}^2 + \dots + X_{2K}^2) + \dots + (X_{N1}^2 + X_{N2}^2 + \dots + X_{NK}^2)$$

(各小区产量平方和)

$$\text{例: } \sum_{j=1}^K \sum_{i=1}^N X_{ij}^2 = (3^2 + 4^2 + 2^2 + 5^2 + 4^2) + (6^2 + 9^2 + 7^2 + 8^2 + 9^2) + (1^2 + 3^2 + 4^2 + 4^2 + 3^2) + (10^2 + 11^2 + 14^2 + 12^2 + 15^2) + (7^2 + 6^2 + 8^2 + 9^2 + 5^2) = 70 + 311 + 51 + 786 + 255 = 1473$$

$$7. \sum_{i=1}^N X_i Y_i = (X_1 \cdot Y_1) + (X_2 \cdot Y_2) + (X_3 \cdot Y_3) + \dots + \dots + (X_N \cdot Y_N)$$

(两变数的乘积和)

$$\text{例: } \sum_{i=1}^N X_i Y_i = (5.3 \times 6.6) + (4.7 \times 6.0) + (3.9 \times 6.5) + (3.9 \times 7.0) + (4.3 \times 7.7) + (4.9 \times 5.7) + (4.4 \times 7.0) + (8.5 \times 9.1) + (5.8 \times 7.6) + (6.2 \times 6.5) = 369.40$$

田号	新植甘蔗		$X_i \cdot Y_i$
	有效茎数 (千茎/亩) X_i	蔗茎产量 (吨/亩) Y_i	
1	5.3	6.6	34.98
2	4.7	6.0	28.20
3	3.9	6.5	25.35
4	3.9	7.0	27.30
5	4.3	7.7	33.11
6	4.9	5.7	27.93
7	4.4	7.0	30.80
8	8.5	9.1	77.35
9	5.8	7.6	44.08
10	6.2	6.5	40.30
$N=10$	$\sum X = 51.9$	$\sum Y = 69.7$	$\sum XY = 369.40$

$$8. \sum_{i=1}^N (X_i \pm Y_i \pm Z_i) = \sum_{i=1}^N X_i \pm \sum_{i=1}^N Y_i \pm \sum_{i=1}^N Z_i = (X_1 \pm Y_1 \pm Z_1) + (X_2 \pm Y_2 \pm Z_2) + \dots + (X_N \pm Y_N \pm Z_N) = (X_1 + X_2 + \dots + X_N) \pm (Y_1 + Y_2 + \dots + Y_N) \pm (Z_1 + Z_2 + \dots + Z_N) = \sum_{i=1}^N X_i \pm \sum_{i=1}^N Y_i \pm \sum_{i=1}^N Z_i.$$

(多項式的連加总和 = 連加式的多項的总和)

$$9. \sum_{i=1}^N CX_i = C \sum_{i=1}^N X_i = CX_1 + CX_2 + \dots + CX_N = C(X_1 + X_2 + X_3 + \dots + X_N)$$

(常数与变員数的乘积和 = 常数与变員数連加总和的乘积)

$$10. \sum_{i=1}^N C = NC$$

(1—N个常数連加总和 = N个常数的乘积)

第五节 农艺、土化和昆虫生物統計的特点

1. 农艺生物統計的特点:

(1) 資料大多数是可量的, 可数資料占絕少数, 所以大多数資料采用可量資料分析法, 資料很少需要代換为其他单位。

(2) 試驗多为田間的, 室內較少, 所以誤差較大, 而且較难控制。但由于选种、栽培处理效应相对地大, 所以一般不要求过分灵敏的測驗。

(3) 誤差方面一般是試驗誤差較重要于抽样誤差。

(4) 試驗单位(小区)一般为一块土地面积的概念。

(5) 重視相关研究, 如間接鉴定, 性状相关, 丰凶預測。

(6) 多年生的作物, 植株占土地面积較大, 試驗結果随年龄增长与年中气候, 剖面变迁等綜合变化而有异, 往往需要以处理前的資料改正处理后的資料。

(7) 由于栽培技术条件的不善, 往往会造成缺株、缺区現象, 需要估計缺区、缺株产量。

(8) 抽样研究着重在: 調查标准的决定, 生长量与产量抽样調查, 作物估产与土壤抽样等。

2. 土化生物統計的特点:

(1) 室內試驗与田間試驗都一样多。室內試驗宜多注意抽样誤差, 田間試驗則宜多注意試驗誤差。室內試驗中土壤微生物試驗誤差較大, 室內理化試驗誤差較小。

(2) 抽样研究着重在: 土壤微生物工作, 吸肥診斷, 根外施肥等方面。

(3) 相关研究着重在: 土地肥力变化規律与輪作、施肥关系等方面。

(4) 土壤微生物是可数資料, 其他多数是可量資料。

3. 昆虫生物統計的特点:

(1) 資料絕大部分為可數資料，可量資料只占少數。所以大多數資料都採用可數資料的分析方法。

(2) 在某種情況下，可數資料須進行統計代換方可分析。

(3) 重視研究昆蟲分布的特點以便對某種害蟲採取相應的抽樣方法。

(4) 重視相關研究來解決害蟲發生規律和進行蟲害發生的預測預報，同時重視時間數列的研究。

(5) 試驗單位的概念不是一塊土地的概念，而是半葉、一葉、一枝、一株、十株、一行的概念。

第六節 學習生物統計應有的認識

一、不應該認為生物統計是萬能的，更不應認為統計結論是最后的結論。統計方法不能代替全面的科學方法，它只有助於判斷試驗結果的可靠性和現象的規律的真實性，而一點也無助於解釋規律或解釋現象。解釋要靠專業知識來解決。

二、隨着自然科學的發展，儀表愈趨精密，觀察方法愈趨細膩，生物統計愈是需要的，而不是可有可無。

三、生物統計有助於室內與室外試驗研究，個體與群體研究，兩者不可偏廢。

四、統計分析是重要，但正確的試驗設計與有設計地進行調查以及搜集準確的試驗數據更為重要。只有在試驗或調查取得可靠試驗數據的前提下，才能談得上統計分析。試驗設計可以減少試驗誤差，同時也要注意技術改良，兩條腿走路。

五、當統計分析結論無顯著差異時，並非為絕對無差異，不過差異仍在誤差範圍內，不足以推翻假設，可以重復試驗，進一步分析多個試驗結果是否有同一趨勢。然後再作出明確的結論。

六、統計分析方法如因設計複雜而繁復時，不應認為浪費人力，不切合實際，因為試驗過程所花的代價遠較統計分析所花者為高。和整個試驗比較起來，計算分量只占很少比例。

七、不要過分追求公式的來源，因為有些公式的推算需要高深的數學知識，屬於數理統計學的任务，我們的任务是過橋而不是架橋，我們對統計公式的正確態度應該是：明確公式的應用範圍和受那些條件的限制，弄清楚在什麼情況下應用這種或那種公式，從我們本身的應用範圍去思考、弄清楚問題的性質是怎樣的，如何用公式去解決它。否則會變成好高騖遠。當然，我們要反對教條主義，不考慮材料的特點，不根據試驗和調查的目的要求，填鴨式生搬硬套地應用統計公式。要做到靈活應用各種公式，首先必須切實弄清楚各種基本概念（例如隨機區組、拉丁方、 t 分布……等等），其次還要多多聯系本行專業的實際，反復思考，多做練習。

第二章 試驗資料的整理

进行試驗或調查获得了一定資料数据，当資料很少时，纵使不加整理，也可以分析而明其意义，但当資料繁多时，若不进行适当的整理，則資料仅是一堆看不出問題看不出規律的数字，它是没有什么参考价值的，只有把原始資料加以整理，我們才能看出問題，了解資料的分布特征，所以，統計分析的第一步工作就是整理原始的数字資料，求出一系列的統計常数，了解样本的特征，从而作出統計結論。

第一节 試驗資料整理的基本內容

- 一、审核原始資料的可靠性，把不可靠的除去，宁可加以估計(如缺区产量估計)。
- 二、采用簡化方法进行运算，既无損于原始資料的原来消息，又可使計算大大簡化。
- 三、整理成次数分布表或次数分布图，使人对試驗資料获得初步印象，并便于計算各种統計常数。
- 四、利用試驗資料繪制統計图，便于表示資料的分布情况，一目了然。

第二节 簡化运算

統計分析过程中有时要利用簡化运算方法。

一 簡化运算应注意的事項

- 一、在簡化运算过程中以某一数加、减、乘或除全数資数的任何一个数字，則必須以該数加、减、乘或除其余各数。
- 二、平均数受到各种簡化运算的影响，即加、减、乘、除对之均有影响，故最后的结果应进行反簡化运算来校正之。同时要注意在反簡化运算过程中要把正簡化运算过程逆轉来运算。
- 三、标准差只受到乘、除簡化运算影响，而不受加、减簡化运算的影响，故在进行反簡化运算时，只須进行乘或除反簡化运算。
- 四、簡化运算不影响計算的准确性，无損于原始資料的原来面目，但必須进行反簡化运算才能获得真实的統計結論。
- 五、要取舍小数后的位数时，一般运用四舍五入法，比五大的进一位，比五小的舍去之；5之前遇奇数則进一位，遇偶数則舍弃5。

二 簡化运算方法

例 1，应用簡化运算方法求下列数值的平均数及标准差：0.0238、0.0238、0.0241、