数量遗传学
原理

cLiang Yi Chuan Xue Yuan Li

江 苏 科 学 技 术 出 版 社

内 容 简 介

本书是一本数量遗传学教材提纲，包括统计遗传、群体遗传以及生统遗传三部分，内容浅显，叙述简要，介绍了数量遗传学的概貌和基本要点。本书特点是以一系列公式贯穿全书，并由浅入深，循序渐进地加以介绍，了解这些公式，就能了解数量遗传学的基本内容。本书以高中数学为基础，超过此范围的数学内容均有专门解释以便于阅读。

本书可供动、植物遗传育种工作者、医学（遗传病）工作者、及有关专业高等院校师生学习数量遗传学入门的参考。

Principles of Quantitative Genetics
Wharton B. Mather
Burgess Publishing Company, 1964

数量遗传学原理
W.B.马塞 著
盖钧镒 译

江苏科学技术出版社出版
江苏省新华书店发行
江苏新华印刷厂印刷

1979年12月第1版　1979年12月第1次印刷
印数 1—5,000
书号 13196·018　定价 0.58 元

译 者 的 话

数量遗传是遗传学的一个分支学科。广义地说，数量遗传学是用数学，特别是数理统计方法研究生物遗传变异规律的科学。它所研究的内容，迄今主要包括两个方面：一是群体内基因的遗传和变异，一是数量性状的遗传和变异。前者称群体遗传学。后者称生统遗传学，也称数量遗传学（狭义的）。所以数量遗传学是从宏观方面研究生物遗传变异规律的学科之一。

动植物育种实际上就是从大量自然或人工群体中选育具有特殊经济价值的基因与基因型，而这些对人类有特殊经济价值的性状又大量地是数量性状，所以，数量遗传学对于改进动植物育种技术愈益显得重要，它是育种学的重要理论基础。同样，数量遗传，尤其群体遗传对于研究、控制人类遗传病害也是十分重要的。

本书是数量遗传学课程的教材提纲。作者W.B.马塞博士在澳大利亚昆士兰大学动物系工作，他试图将数量遗传各部分的要点综括为一个联结的整体。所以全书分为三篇：第一篇统计遗传学，介绍数量遗传常用的统计方法；第二篇群体遗传学，介绍在系统压力（迁移、突变、选择）及离散压力作用下群体内基因的遗传变异；第三篇生统遗传学，介绍数量性状的遗传变异，包括表型组成、方差的遗传分量、遗传力、选择响应、近交与异交、相关选择等。本书作为一个提纲，写作是

十分简要的，它的循序介绍一系列公式来贯穿全书，了解这些公式就能了解数量遗传学的基本要点。本书以高中数学为基础，超过此范围的数学内容另列为框格，加上注释，以便于阅读。本书还附有习题，通过它，可以帮助理解与掌握。总之，本书是一本数量遗传学入门的提纲和参考书。

鉴于国内关于数量遗传学的系统研究介绍还不多，而动植物遗传育种工作者对这方面的瞭解甚为迫切，1979年10月在成都举行的中国遗传学会作物遗传学报告会上，许多遗传育种工作者提议加强这方面的普及与提高工作。在马育华教授的启示下，将本书译出的目的是为普及数量遗传知识添砖加瓦。原书在排印中有一些错误，翻译时尽力做了注释与更正。限于译者的水平，错误之处一定很多，请批评指正。江苏科学技术出版社在本书符号体系的排印中作了很大努力，使译本更便于阅读。

南京农学院　盖钧镒

1979.10

前　言

在本书中，我试图将统计遗传学、群体遗传学以及生统遗传学的要点综括为一个联结的整体。

我认为在解释数量遗传学各方面，有三本书是卓著的，即：K. Mather, 1951，《生物学的统计分析》(Methuen)；C. C. Li, 1954，《群体遗传学》(University of Chicago Press)；D. S. Falconer, 1961，《数量遗传学》(Oliver and Boyd)。要更全面地讨论我的主要论题，应该参考这些著作。

像 Crow 的 著名的《遗传学摘要》(Burgess) 一样，为免除乱记讲授记录，本书设计于右页空白处供师生随意作附记用。像 Lerner 的《选择的遗传基础》(Wiley)一样，非主题所必需的数学注释和证明另列在框格里。为了学好本书，必须做好习题，答案列于附录。

感谢下列作者和出版社允许以修订形式使用有关图表。

附表 Ⅰ—Ⅳ，F. Yates 和 R. A. Fisher, Oliver and Boyd 出版的《生物、农业及医药研究的统计用表》的原表，此处附表经 K. Mather 及 Methuen 的《生物学的统计分析》更动。附表 Ⅴ，F. Yates 和 R. A. Fisher, Oliver and Boyd 出版的《生物、农业及医药研究的统计用表》。图 3—5, K. Mather, Methuen 出版的《生物学的统计分析》。图9—1、9—2、9—3 和 表 17—1，D. S. Falconer, Oliver and

Boyd 出版的《数量遗传学》。图 12—3， Th. Dobzhansky 及《进化》编辑。图 18—1， I. M. Lerner， Cambridge University Press 出版的《群体遗传学和动物改良》。感谢 Rosalyne Spurway 小姐按照我的草稿绘制了图表。

W. B. 马塞博士

昆士兰大学动物系遗传实验室

1963，12

目　　录

第一篇　统计遗传学

第二篇　群体遗传学

第三篇　生统遗传学

第一篇　统 计 遗 传 学

第一章　基 本 概 念

1·1　引　论

统计学可以看作为试验的数学处理，其目的是双重的：

1．将资料精简为少量简单的常数，及

2．估价这些常数的显著性或重要性。

实验科学的最终目的是由特殊推论到一般，所以实验科学家关心归纳逻辑方法而不是演绎逻辑方法。

1·2　总体和样本

统计学的出发点是由之抽取样本的总体。我们将总体的生统常数定义为参数，而样本的为统计数。统计数由于抽样而受随机误差的影响，但一旦统计数已知，就可以按给定概率指出参数所在的可能范围。习惯上用希腊字母表示参数，用罗马字母表示统计数。

例如，算术平均数为 μ (mu)和 $\bar{x}$，标准差为 σ (sigma)和 s。通篇用 S，而不用 Σ，表示总和。（译者注：为与国内外大多数文献相一致，译文中均用 Σ 代替 S 表示总和）

1·3　概　率

概率的概念是统计学和遗传学的基本概念之一。概率可定义如下：若在 n 次遭遇中有 m 次按特定方式出现，则此特定事件的概率 $P=\frac{m}{n}$。例如，果蝇中 $\frac{e}{e}$ 为黑檀体，$\frac{E}{e}$ 或 $\frac{E}{E}$ 为

野生型。在$\frac{E}{e}\times\frac{e}{e}$杂交中配子为$E$，$e$及$e$，$e$。于是来自$\frac{E}{e}$，配子为$e$的概率是$\frac{1}{2}$；而来自$\frac{e}{e}$，配子为$e$的概率是1。所以果蝇是黑檀体的概率为$\frac{1}{2}$。

注意两个独立事件同时发生的概率为各单独发生时概率的乘积。另一方面，两个独立事件单独发生的概率为其各个别概率之和。

例 1. 在下表中：

AB	aB	Ab	ab
9	3	3	1

aB占$\frac{3}{16}$，Ab占$\frac{3}{16}$，而aB或Ab为$\frac{6}{16}$。

例 2. 由$\frac{E}{e}\times\frac{e}{e}$组合所得一个五果蝇的家系中，全是黑檀体的概率为多少？

$$P=\frac{1}{2}\times\frac{1}{2}\times\frac{1}{2}\times\frac{1}{2}\times\frac{1}{2}=\left(\frac{1}{2}\right)^{5*}$$

同样，全是野生型的概率为多少？

$$P=\left(\frac{1}{2}\right)^5=\frac{1}{32}$$

这两项是二项式$\left(\frac{1}{2}+\frac{1}{2}\right)^5$展开式的"端"项。

$\left(\frac{1}{2}+\frac{1}{2}\right)^5$展开式的其他诸项给出了某些为野生型、某些

译者注：*原文误为$P\times\frac{1}{2}\times\frac{1}{2}\times\frac{1}{2}\times\frac{1}{2}\times\frac{1}{2}=\left(\frac{1}{2}\right)^5$。

为黑檀体的概率。

推而广之，

$$(p+q)^k = p^k + \cdots\cdots + \frac{k!p^r q^{k-r}}{r!(k-r)!}$$

其中 p 是一组的期望值(expectation)，q 是另一组的期望值，k 是家系中的个数，r 是某特定家系中第一组的个数。例如，由 $\frac{E}{e}\times\frac{E}{e}$ 组合所得一个五果蝇的家系中，4 个是黑檀体 1 个是野生型的概率为多少？

这里野生型的概率为 $\frac{3}{4}$，黑檀体为 $\frac{1}{4}$。因为 $\frac{E}{e}\times\frac{E}{e}$ 得 $1\frac{E}{E}:2\frac{E}{e}:1\frac{e}{e}$ 即 3 野生型：1 黑檀体，

$$\therefore \quad P = \frac{5!\left(\frac{1}{4}\right)^4\left(\frac{3}{4}\right)^1}{4!\times 1!} = \frac{15}{1024} = 0.01465$$

1·4 总体的分析

1. 分组或归组

若量测一群有机体的某种数量性状，便得到许多不同数字，处理这些数字的第一步，就是将它们归入相差某一任定数目，例如 5% 的各组。这样就得到量测值的次数分布，每一个组的中点称为组中点值。

2. 次 数 曲 线

若将分组所得的信息(information)图示，以横座标代表组值，纵座标代表各组的个体数，则得到一个次数曲线。由此便可获得关于这个资料更清晰的概念。

3. 正 态 曲 线

这种次数曲线趋近于数学家所称的正态曲线。正态曲线

是 $(p+q)^k$ 式，当 k 无限增大，而 p 或 q 与 $\frac{1}{k}$ 或更小值不同阶时的展开式。我们直接注意到正态分布与概率及二项式展开式的关联。其公式为：

$$m=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

这里，$m=x$ 组的频率，$e=$ 自然对数底数 $=2.71828$，$\pi=\frac{22}{7}$。注意这个公式既不包含 p、q，也不包含 k，而有两个新的参数 μ（mu）和 σ（sigma）。并注意 p 和 q 由假设所定（概率 例题 中 $p=\frac{1}{4}$，$q=\frac{3}{4}$），而这两个参数并非由假设所定，它可由资料估计。μ 称为平均数，σ 称为标准差。

图示如图 1-1

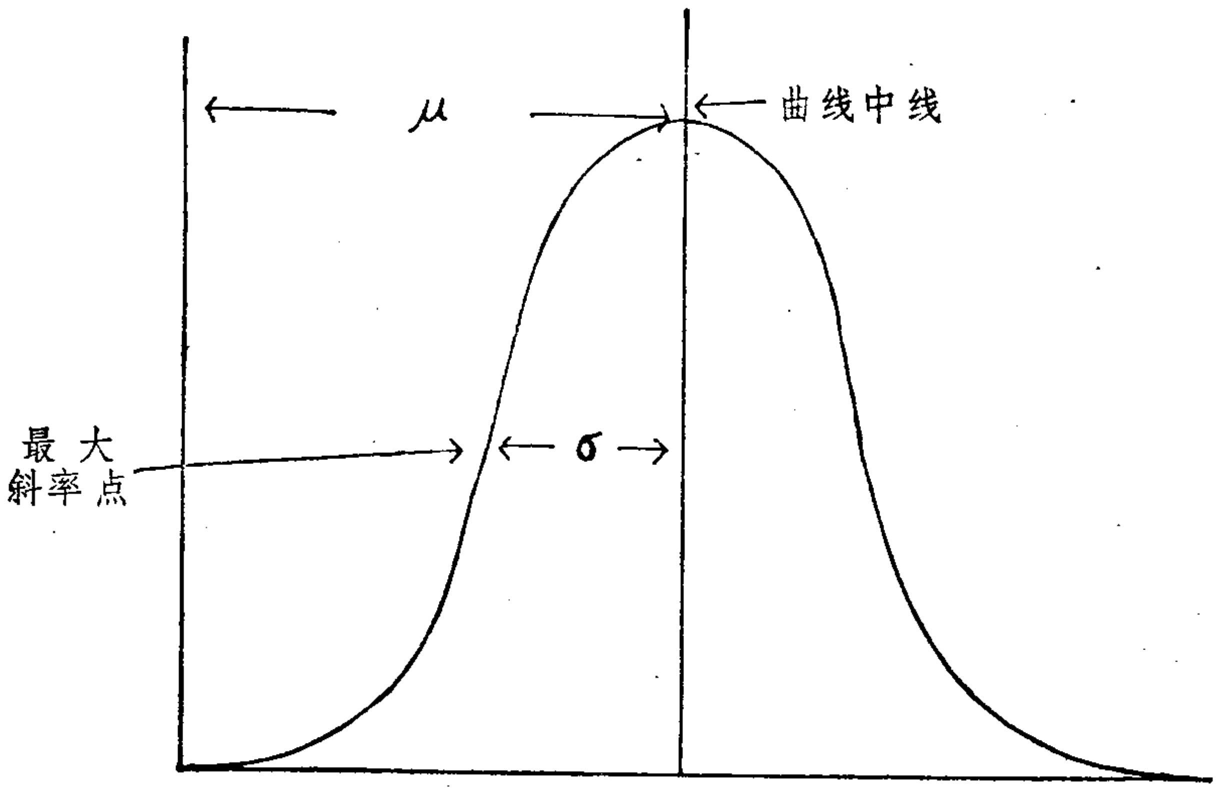

图 1-1

习　　题

1．两棕眼双亲均为具有隐性蓝眼基因的杂合体。(1) 一个五个儿童的家庭中，依次为蓝、棕、蓝、蓝、棕眼的概率为多少？(2) 一个五童之家，三个蓝眼，二个棕眼(不论次序)的概率为多少？

2．某人的祖母和外祖父患隐性纯合所致的一种疾病，此人 (1) 患此病，(2) 带隐性基因的概率各若何？

3．若不管交换，(1) 一个男人继承全部其祖父的常染色体的概率将若何？(2) 一个男人缺乏其外祖父的常染色体的概率将若何？

第二章　生物统计常数

2·1　平均数和标准差

1．公　式

a．平均数 $\bar{x}=\dfrac{\Sigma x}{n}$

b．标准差——离散性测度

最简单的离散性测度或许是观察值离平均数的平均偏差，它等于$\dfrac{\Sigma(x-\bar{x})}{n}$，这里不论 $(x-\bar{x})$ 的符号*。将观察值**在总加前先平方，然后以 n 除，取其平方根，可以克服符号上的困难。因下面“2”所述的理由，分子将为 $n-1$ 而非 n；这样给出标准差的估计值。

译者注：*原文为不论 x 的符号。

**指 $(x-\bar{x})$。

$$s = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n-1}}$$

其中，x 是各个量测值，n 是量测个数。

s^2 = 方差 = V = 离均差平方和(SS)以 $n-1$，即 N(自由度数)除。这就是均方 MS。

在论及成对观察值，而不是单个观察值时，我们遇到的是 x 和 y 的协方差

$$W_{xy} = \frac{\Sigma(x-\bar{x})(y-\bar{y})}{n-1}$$

这里 $\Sigma(x-\bar{x})(y-\bar{y})$ 称为 x 和 y 的乘积和。

注意下式，有时，例如应用计算机时，常较便当：

$$\Sigma(x-\bar{x})^2 = \Sigma x^2 - \frac{1}{n}(\Sigma x)^2$$

Σx^2 称平方和，$(\Sigma x)^2$ 为总和平方。

以下论述可以证实这点。(见框格 2-1)

框格 2-1

证明　$\Sigma(x-\bar{x})^2 = \Sigma x^2 - \frac{1}{n}(\Sigma x)^2$

若引进一个任意的操作平均数 x_m

$$(x-\bar{x}) = (x-x_m) - (\bar{x}-x_m)$$

∴ 两边平方

$$(x-\bar{x})^2 = (x-x_m)^2 - 2(x-x_m)(\bar{x}-x_m) + (\bar{x}-x_m)^2$$

然后总加，

$$\Sigma(x-\bar{x})^2 = \Sigma(x-x_m)^2 - 2\Sigma[(x-x_m)(\bar{x}-x_m)] + \Sigma(\bar{x}-x_m)^2$$

$$= \Sigma(x-x_m)^2 - 2(\bar{x}-x_m)\Sigma(x-x_m) + n(\bar{x}-x_m)^2$$

因为 $\bar{x}$ 和 x_m 是常数。

但是　$n(\bar{x}-x_m) = n\bar{x} - nx_m$，由 $\bar{x}$ 的定义

$$= \Sigma x - n x_m$$
$$= \Sigma (x - x_m)$$

代入 $(\bar{x} - x_m)$ 得

$$\Sigma(x-\bar{x}) = \Sigma(x-x_m)^2 - \frac{2}{n}[\Sigma(x-x_m)]^2 + \frac{1}{n}[\Sigma(x-x_m)]^2$$
$$= \Sigma(x-x_m)^2 - \frac{1}{n}[\Sigma(x-x_m)]^2$$
$$= \Sigma x^2 - \frac{1}{n}(\Sigma x)^2 \quad \text{当 } x_m \text{ 取 0 时}$$

c，二项式分布的平均数和标准差

$\mu = p$ 或 q，依涉及哪一个值的平均数而定。

$\sigma = \sqrt{\frac{pq}{k}}$，$p$ 和 q 是比例，k 是家系 的 大小。（见框格 2-2）

框格 2-2

$\mu = p$ 或 q 及 $\sigma = \sqrt{\frac{pq}{k}}$ 的证实

由某种资料估计 $\bar{x}$ 和 s，可以看出这些公式的真实性，注意 $\bar{x} \to p$ 或 q，$s \to \sqrt{\frac{pq}{k}}$。

以果蝇 $\frac{E}{e}$ 与 $\frac{e}{e}$ 回交为例，每个体为黑檀体的机会是 $\frac{1}{2}$，故在 5 个体的家系中，由上述公式

$$\mu = p = 0.5$$

$$\sigma^2(\text{方差}) = \frac{1}{2} \times \frac{1}{2} \times \frac{1}{5} = 0.05$$

由下列 k 个家系，每家系五个果蝇 的 资 料 可以估计

$\bar{x}$和 s：

野生型比例 (x)	观察次数 (a)	ax	ax^2
0.0	0	0.0	0.0
0.2	2	0.4	0.08
0.4	5	2.0	0.80
0.6	6	3.6	2.16
0.8	3	2.4	1.92
1.0	1	1.0	1.00
	17	9.4	5.96

$$\bar{x}=\frac{\Sigma ax}{n}=\frac{9.4}{17}=0\ 55 \qquad \text{比较：}\ \mu=0.5$$

注意，为要得到实际数目，将诸次数乘以诸比例。

$$\Sigma(x-\bar{x})^2=\Sigma x^2-\frac{1}{n}(\Sigma x)^2$$

$$=5.96-\frac{(9.4)^2}{17}=0.76$$

$$s^2=V=\frac{\Sigma(x-\bar{x})^2}{n-1}=\frac{0.76}{16}=0.05$$

比较 $\sigma^2=0.05$

注意 μ 和 $\bar{x}$，σ^2 和 s^2 的相近显示了以上公式的真实性。

2. 以 N 代替 n 的效用（Bessel 矫正数）

选用 $n-1$ 即 N，自由度数，代替 n 做分母以计算方差，可这样证实：

方差是曲线离散性的测度。

∴ 离散性越大，任何两个 x 值之间的差异也越大。一个差异需由二个值获得。

∴ 得到 $n-1$ 个差异或自由度需要 n 个值。

操作法则是：估计一个参数，丧失一个自由度。例如计算 s 时，平均数是估计得来的。

3. 薛伯氏(Sheppard)分组矫正数

资料分组虽然对简化计算十分有用，（例如，计算平均数时，将组中点乘次数再相加，代替了冗长的加法），但由于稍稍增加了图形的离散度，相应增大标准差而导入了误差。

薛伯氏矫正数是在计算标准差未开方时先减去 $\frac{1}{12}$ 的组距。

2·2 平均数的标准误

平均数标准误是总体中的许多样本平均数间离散的测度。其公式为：

$$s_{\bar{x}} = \frac{s_x}{\sqrt{n}} \quad \text{（推导见框格 2-3）}$$

注意，由此可见 n（样本所含个体数）增大，标准误减小，即增加样本含量减低标准误。

框格 2-3

证明 $s_{\bar{x}} = \frac{s_x}{\sqrt{n}}$

令 x 和 y 为成对的量测值，成对数 n 为样本含量，于是

$$V_x = \frac{\Sigma(x-\bar{x})^2}{n-1} \quad \text{及} \quad V_y = \frac{\Sigma(y-\bar{y})^2}{n-1}$$

以 $x+y$ 代 x