



全国高等农业院校教材

全国高等农业院校教材指导委员会审定



数量遗传学

● 盛志廉 吴常信 编著

● 畜牧和动物遗传育种专业用

中国农业出版社

全国高等农业院校教材

数量遗传学

盛志廉 吴常信 编著

畜牧和动物遗传育种专业用

中国农业出版社

全国高等农业院校教材

数量遗传学

盛志廉 吴常信 编著

责任编辑 李妍书

出版 中国农业出版社
(北京市朝阳区农展馆北路2号)

发行 新华书店北京发行所

印刷 北京忠信诚胶印厂

* * *

开本 787mm×1092mm16开本

印张 11.75 字数 286千字

版、印次 1995年10月第1版

1999年10月北京第2次印刷

印数 2 701-4 700册 定价 14.80元

书号 ISBN 7-109-03462-3/Q·208

前 言

这是我国第一本适用于畜牧系的数量遗传学教材，可作为动物遗传育种专业的基本教材，也可作为畜牧专业和特种经济动物专业的选修教材。对于广大的高、中级育种工作者，也是一本比较合适的自学教材和参考书。

本教材由东北农学院盛志廉教授和北京农业大学吴常信教授编著。前者编写绪论和第一、二、三、四、五、六、七、九章，后者编写第八、十、十一、十二、十三、十四章。全书承蒙北京农业大学吴仲贤教授主审和华中农业大学彭中镇教授参审，并已根据他们提出的重要意见进行了修改。

全书内容包括群体遗传学基础知识、遗传参数和选择的原理和方法。为了便于初学者学习，还补充了一些必要的统计方法，如通径系数、最小二乘分析和非线性模型；每章都附有习题，书末附汉英名词对照。

根据农业部颁发的《高等学校农科教材编写出版暂行规定》的精神，本教材力求符合大学本科培养规格和课程基本要求，在内容选择上以数量遗传学基本内容为主，尽量吸收一些近年发展的新内容，但一些过于深奥或未经公认的新进展一般不予编入。在表达上努力做到深入浅出、简明扼要，推导也尽量简单。名词术语和符号大多尊重一般习惯，不作大量更换，全书各章也尽量统一，避免代号混乱和重复。为了贯彻少而精的原则，并考虑基本教材的特点，在计算方法的介绍上只选择实用、简明的一两种方法，而不罗列各家提出的诸多方法，这样对初学者可能更为有利。

限于编写人的水平和编写时间匆忙，疏漏和错误都在所难免，欢迎多提意见，不吝赐教。

本教材编写中，“阈性状的参数分析”一章和“最宜选择指数”一节取材于吴仲贤教授编著《统计遗传学》一书；“育种值的矩阵计算方法”和“育种值的BLUP计算方法”两节取材于陈健秋先生所编《现代动物育种》；“非线性模型”一章取材于杨宁先生论文“应用于动物生产中的几种非线性模型”。承蒙允诺，在此深表谢意。

在书稿誉清整理中，徐慧如、白丽华、陈瑶生、张志武、翟国勤等同志给予很大帮助，王宝青同志绘制了大部分的图，一并致谢。

编 者

1994年2月

目 录

前言	
绪论	1
第一章 基因频率与基因型频率	5
§ 1 基本概念	5
1. 群体	5
2. 基因频率	5
3. 基因型频率	6
4. 随机交配	6
§ 2 哈代-温伯格定律	7
1. 定律要点	7
2. 数学证明	7
3. 生物学证明	9
4. 定律的意义	10
§ 3 影响基因频率和基因型频率的因素	10
1. 突变	10
2. 选择	11
3. 遗传漂变	13
4. 迁移	13
5. 同型交配	14
第二章 通径系数	16
§ 1 概念	16
§ 2 相关系数的剖分	17
§ 3 定理	18
1. 定理(一)	18
2. 定理(二)	19
3. 定理(三)	20
4. 定理(四)	21
5. 总定理	22
§ 4 通径链的追溯规则	23
1. 注意通径的方向	23
2. 只能改变一次方向	24
3. 相关线	24
4. 避免重复	24
§ 5 显著性检验	25
第三章 亲属间遗传相关和近交系数	27

§ 1 親子間的通徑關係	27
§ 2 隨機交配下親屬間的遺傳相關	28
§ 3 個體近交系數的計算	29
§ 4 親緣系數	31
§ 5 畜群近交系數的估計	33
1. 親緣協方差表法	33
2. 閉鎖小群體近交系數的估計	35
§ 6 連續規則近交下的近交系數	35
1. 連續全同胞交配	36
2. 連續半同胞交配	37
3. 連續與同一個體回交	37
第四章 近交與雜交	39
§ 1 近交與雜交的遺傳效應	39
1. 近交增加純合子頻率，雜交增加雜合子頻率	39
2. 近交降低群體均值，雜交提高群體均值	41
3. 近交、雜交能改變群體的遺傳方差	43
§ 2 雜種優勢和雜交互補	45
§ 3 雜種優勢的度量與配合力的分析	47
1. 雜種優勢的度量	47
2. 配合力分析	48
第五章 重複率	51
§ 1 概念	51
§ 2 估計方法	52
§ 3 主要用途	53
1. 判斷遺傳估計正確與否	53
2. 確定性狀需要度量的次數	53
3. 估計畜禽終身可能的平均生產力	54
第六章 遺傳力	56
§ 1 概念和原理	56
§ 2 估計方法	59
1. 公畜內女母回歸法	59
2. 半同胞相關法	60
3. 全同胞相關法	61
§ 3 顯著性檢驗	63
§ 4 遺傳力的用途	65
§ 5 減少遺傳力的估計誤差	67
1. 樣本含量預估	67
2. 消除環境誤差	68
3. 混合家系平均親緣相關	70
4. 親本間親緣相關的校正	71
第七章 性狀間遺傳相關	74

§ 1 性状间相关原因的分析	74
§ 2 估计方法	75
1. 亲子协方差法	76
2. 同胞协方差法	77
§ 3 显著性检验	78
§ 4 主要经济性状间的遗传相关	79
第八章 阈性状的参数分析	81
§ 1 重复率	81
1. 回归法	81
2. 列联表法	83
§ 2 遗传力	84
1. 阈性状遗传力的估计	84
2. 阈性状遗传力的显著性测验	86
第九章 最小二乘分析	88
§ 1 单因素最小二乘分析	88
1. 数学模型与最小二乘方程	88
2. 约束条件	89
3. 最小二乘均数	90
§ 2 二因素无互作的最小二乘分析	90
1. 数学模型与正规方程	90
2. 约束条件	91
3. 吸收法	91
4. 最小二乘均数与方差分析	93
§ 3 二因素有互作的最小二乘分析	94
1. 互作效应的显著性检验	94
2. 数学模型和正规方程	95
3. 约束条件	95
4. 最小二乘均数与方差分析	96
§ 4 均数加权二乘分析	97
1. 参数估计	97
2. 方差分析	98
第十章 非线性模型	100
§ 1 非线性模型概述	100
1. 线性模型与非线性模型	100
2. 模型的建立	100
3. 参数的估计	101
§ 2 泌乳曲线	102
1. 数学模型	102
2. 应用	103
§ 3 生长曲线	104
1. 数学模型	104

2. 应用	105
§ 4 产蛋曲线	106
1. 数学模型	106
2. 应用	107
第十一章 选择理论及影响选择的因素	109
§ 1 自然选择与人工选择的比较	109
1. 不同的目的	109
2. 不同的方向	109
3. 不同的效果	109
4. 不同的速度	109
§ 2 改变“遗传的自动调节”	110
1. 保持特殊的环境	110
2. 提高群体的平均数	110
3. 坚持不懈地选择	110
4. 改进选择方法	110
§ 3 选择的创造性及其原因	110
1. 产生了新的突变	111
2. 出现了新的组合	111
3. 基因的固定	111
4. 阈性状的出现	111
§ 4 选择极限	111
§ 5 影响选择效果的因素	112
1. 遗传力与选择差	112
2. 世代间隔	113
3. 性状间的相关	114
4. 不同的选择方法	114
5. 近交与引种	115
6. 环境	116
第十二章 几种基本的选择方法	117
§ 1 选择方法概述	117
1. 外形选择与生产性能选择	117
2. 表型值选择与育种值选择	117
3. 单个性状选择与多个性状选择	118
4. 个体选择与家系选择	118
5. 直接选择与间接选择	118
§ 2 常用的几种选择方法	118
1. 性能测定	118
2. 系谱测定	119
3. 同胞测定	120
4. 后裔测定	121
§ 3 合并选择	123

1. 原理与公式	123
2. 不同选择方法的结果	124
3. 各种选择方法的比较	125
§ 4 间接选择	127
1. 间接选择的意义和作用	127
2. 间接选择的应用举例	128
第十三章 个体育种值估计	130
§ 1 估计育种值的一般方法	130
1. 估计育种值的原理	130
2. 估计育种值的方法	130
§ 2 育种值的矩阵计算方法	140
1. 基本原理	140
2. 公式推导	140
3. 应用举例	141
§ 3 育种值的 BLUP 计算方法	151
1. BLUP 的含义	152
2. BLUP 模型的建立与求解	152
第十四章 多性状选择	161
§ 1 多个性状选择方法概述	161
1. 顺序选择法	161
2. 独立淘汰法	161
3. 选择指数法	162
§ 2 简化选择指数	162
1. 简化选择指数公式的演变	162
2. 制订简化选择指数的步骤	163
§ 3 一般的选择指数公式——无约束选择指数	164
1. 制订过程	164
2. 选择指数的某些参数	166
§ 4 约束选择指数与最宜选择指数	169
1. 约束选择指数	169
2. 最宜选择指数	172
汉英名词对照	174

绪 论

生物的性状可分为两大类，有一类性状的变异可截然区分成若干种相对性状，并可分别以形容词描述，如牛的毛色有黄、黑、红、白花之分，人们称之为质量性状；另一类性状则不然，其变异是连续的，各种变异只能以数量来表述，如猪的平均日增重，从几十克到上千克左右，什么数量都有，称为数量性状。但是这两类性状也不能绝对区分，因为“任何质量都表现为一定数量，没有数量也就没有质量。”表面大致看来都是黑毛色，但细致分析其中的色素含量也不相同；有的以数度量的性状，如豌豆的茎高，可明显的分为高矮两种，因而可视作质量性状；有的性状也可计数，如猪的个体产仔数，但表现都是整数，其变异不完全连续，而其各胎平均产仔数或各群体平均产仔数却表现为连续变异，因而很难断定是质量性状还是数量性状。

家畜的大多数经济性状都是数量性状，如产蛋量、增重速度、产奶量、饲料报酬。毛的细度、密度、胴体的瘦肉率、奶的蛋白质含量等，所以数量性状在畜牧业中显得特别重要。

数量遗传学就是研究数量性状遗传规律的科学。

早在 19 世纪下半叶，英国学者高尔登 (F. Galton) 就开始运用统计方法来研究数量性状的遗传，他在研究人类体高的遗传中正确发现了“回归现象”。他发现在经过男女体高的适当校正之后，子女的平均体高总要比其父母的平均体高更靠近总的群体平均体高，也就是说，体高的遗传中有向群体平均数“回归”的现象。高尔登是生物统计学的创始人，也是研究数量遗传的先驱者，但是由于时代和科学发展的局限，他对“回归现象”没有做出正确的解释和结论，未能成为数量遗传学的创始人。

1900 年孟德尔的“植物杂交的试验”论文被重新发现并引起世界的普遍重视，许多学者用不同的动、植物做了大量的验证工作。英国的实验遗传学家 W. Bateson 对鸡冠的形状遗传进行了大量的试验研究，并与 Saunders 合作发表了关于牛角遗传的论文，第一个证明家畜性状也符合孟德尔遗传规律。以 W. Bateson 和 H. DeVries 为代表的孟德尔学派进而认为孟德尔原理可以普遍用于遗传变异的研究，而连续变异之所以不符合这些定律是因为它是不能遗传的。但以 K. Pearson 和 W. F. R. Weldon 为代表的统计学派则认为只有简单的不连续变异性状才符合孟德尔规律，连续变异的性状其遗传规律只能采用统计学的方法进行研究。这一争论持续约十年之久，争论的焦点在于连续变异是否遗传，也就是说连续变异是受遗传因子制约的，还是由环境效应造成的。到了 1908 年，瑞典遗传学家 Nilsson-Ehle 在小麦种皮颜色的遗传试验中发现，红皮小麦与白皮小麦杂交， F_1 的种皮颜色介于两亲本之间，呈中红色。 F_2 的种皮红色与白色呈 3:1 比例，符合孟德尔一对基因遗传规律。但在 3/4 的红色中，红色的深浅程度不一，有些红似亲本，有些则似 F_1 呈中红色。在另一组试验中发现 F_2 的红色与白色成 15:1 的比例，符合两对因子的分离规律。但按红色深浅

分类时，有 1/16 象红皮亲本，1/16 象白皮亲本，4/16 比 F_1 稍深，4/16 比 F_1 稍浅，6/16 与 F_1 相同。在第三组试验中又发现，种皮颜色红白的比例为 63 : 1；而按颜色深浅程度则形成近乎连续的系列，符合三对因子分离的规律。作者认为，随着涉及遗传因子对数的增加，各种变异逐渐从间断分布趋向连续分布，从而认为连续变异也是遗传的，而且是符合孟德尔遗传规律的，只是涉及的因子对数较多。

经过综合分析，Nilsson-Ehle 提出了多因子假说，其主要论点是：

- (1) 数量性状的遗传是受一系列遗传因子支配的；
- (2) 这种因子单个的效应是微小的，其作用可以累加；
- (3) 相对因子间的显隐性关系通常不存在。

East 在玉米穗长杂交试验中证实了上述观点，并发展了这一假说。

后来许多学者对一些“典型的”数量性状的遗传进行大量研究以后又总结微效基因具有以下特点：

- (1) 微效基因是等效的，作用是累加的；
- (2) 微效基因的等位基因只有增效和减效之别，没有显性抑制或掩盖隐性的现象；
- (3) 由于效应微小，不能予以个别辨认，只能研究性状的总表现，虽然可以对涉及的基因对数作大致的估计。

以后的研究又对多因子假说进行了某些修正。例如 Thoday (1963) 对于果蝇背部刚毛数的研究证明，有时多基因也可以予以个别地辨认，并且表现分离和重组，连锁和交换，即可以确定在染色体上的一定位置。J. W. McArthur 等 (1941) 发现影响番茄果实重的基因的作用不是累加的，而是累积的。更进一步的研究对多基因不存在显性的观点也予以修正，发现在有些情况下，显性与上位效应也是存在的。K. Mather 在 40 年代初还发现多基因也能发生突变。多基因的等效说也为“一效多因，一因多效”说所替代。

Nilsson-Ehle 的多因子假说得到了遗传学界的普遍承认，结束了孟德尔学派与统计学派关于连续变异是否遗传和是否符合孟德尔规律的长期论战，为数量遗传学奠定了重要的理论基础。

但 Nilsson-Ehle 的多因子假说（后改为多基因假说）虽然在理论上似乎统一了数量性状和质量性状的遗传机制，两者都是受基因制约的，但是事实上它并没有为数量遗传研究开辟一条正确的研究途径。受此影响，许多数量遗传研究者钻进了牛角尖，纷纷去追求决定数量性状的基因数目。于是提出了许多有关决定产奶量、产蛋量等经济性状基因数目的推算方法和种种假说。但是一阵热潮过后，很快就发现这些研究实际于事无补，既找不到理论头绪，也指导不了育种实践。因为推算出的基因数目很大，它们相互间的关系又很复杂，这些性状在不同环境下又表现出很大的差别。可以想象，不同的基因对个性状的影响，有的是直接的，有的是间接的，间接程度也不一样，所以不可能每个基因对同一性状都是等效的，因此在等效假定基础上的推算很难得出有价值的结论。

丹麦生物学家 W. Johannsen 通过对菜豆进行一系列研究的总结，于 1909 年提出了纯系学说，对连续变异的原因作了重要补充，对数量遗传学的诞生具有重要作用。

菜豆是自花授粉植物，但在自然界也有混杂的情况。W. Johannsen 在天然混杂的种籽中按种籽重量分别播种，在其后代中选出 19 个株系。种籽百粒重各株系不同，变异从 20g

到 60g 不等。各株系分别纯繁，其后代表现同样的系间差异。每个株系内种籽的重量也有差别，但选择株系内不同重量的种籽分别播种，后代虽然仍有变异，但较重和较轻种籽的后代间的平均百粒重没有显著差异。由此可得到结论：在基因型不同的天然混杂群体中，可分离出许多基因型纯合的株系，因此在混杂群体中选择是有效的；但在由基因型纯合个体自交产生的后代群体——纯系中，个体的基因型都是相同的，它们表型上的差异是由环境造成的，是不能遗传的，因此纯系内继续选择是无效的。纯系学说对遗传学的重要贡献在于区别了遗传的变异和不遗传的变异，只有选择遗传的变异才会产生选择效果；明确了基因型和表现型的概念，同样的基因型可以有不同的表现型，不同的基因型也可以有相同的表现型；指出环境也是造成表现型变异的重要原因，数量性状的变异所以能够完全连续，是多基因与环境共同作用的结果。

数量遗传学的另一大基础是群体遗传学。群体遗传学主要研究孟德尔群体的遗传结构及其变化，群体遗传结构的重要标志是基因频率和基因型频率。在生物的世代传递中，有连续性的只是基因，而基因型则没有。上一代的基因型分离成配子中的基因，而它们的结合又形成下一代新的基因型。

早在 1908 年，英国数学家哈代 (G. H. Hardy) 和德国医生兼遗传学家温伯格 (W. Weinberg) 各自独立地提出后来被定名的哈代-温伯格定律。奠定了群体遗传学的重要基础。以后，英国遗传学家 J. B. S. Haldane (1924) 发表了群体遗传学方面的重要著作《自然和人工选择的数学理论》，对基因频率的短期和长期变化及其重要意义，以及影响因素进行了详细的论述。

英国学者费雪 (R. A. Fisher) 综合了两个学派的意见，于 1918 年发表了一篇关于亲缘相关的论文。美国的 S. Wright 与之紧密配合，他们在前人工作的基础上，创立了群体遗传学，他们的继承者 J. L. Lush 将此应用于家畜育种而创立了数量遗传学。

群体遗传学是孟德尔遗传学由家庭水平向群体水平的进一步发展，而数量遗传学则是统计学和群体遗传学应用于数量性状的产物。在数量性状的研究中可以看到，“基因的累加效应可以按通常习惯的方式用基因频率的变化解释，这是数量遗传学中最重要的一条原理，实际上，它是孟德尔遗传学与统计遗传学的关连环节。因而我们可以完全按孟德尔的方式考虑基因频率的变化在数量性状中所引起的变化，以及某种有利基因在群体中的积累来考虑个体的变化。其结果之一是，个体成为群体的产物，就如群体为其所由组成的个体的产物一样……这可能是现代育种学与经典育种学的最大区别之点，把兴趣由个体转移至群体。”“基因的非累加效应取决于基因型比例的变化，这是数量遗传学的又一条重要原理” (引自吴仲贤著《统计遗传学》)。

数量性状的特征，决定了对其遗传变异的研究必然具有以下特点：

- (1) 性状必须进行度量，而不是简单的区分；
- (2) 必须运用生物统计方法进行分析归纳和比较；
- (3) 必须要以群体作为研究对象，不能限于个体水平或家庭水平。

其实这些特点早已为从事数量遗传研究的先驱者们所掌握，但是为什么他们没有取得决定性的成果呢？总结一下遗传学发展中的正反两方面的经验教训是非常有益的。对于孟德尔的成功关键有许多说法，但是应该看到，孟德尔所以得到成功首先在于他第一个具备

了区分表现型与遗传型的思想，也就是区分现象与实质的思想。千百年来，人们只看到表面的相似与不相似，没有人怀疑相同的现象可能隐藏着不同的实质。孟德尔大胆地设想杂种与其亲本之一虽然表现相同，但其遗传实质不同，一个是杂的，一个是纯的，所以杂种的后代才能分离出两个亲本的性状。这就产生了孟德尔最基本的定律——分离定律。粒子遗传和自由组合都是由此派生出来的。

Lush 等在数量遗传方面的成功也是这样。他同样重视表型值与遗传值的区分，但同时又注意到数量性状与质量性状的差别。如果说造成质量性状表现型与遗传型不一致的主要原因是基因的显性作用的话，那么造成数量性状表型值与遗传值不一致的主要原因则是环境效应。

Lush (1937) 把环境效应值从性状的表型值中剖分出来以后，就找到了性状的遗传值对其表型值的决定程度。这就是他所提出的遗传力概念。

虽然遗传力就等于高尔登早已提出的子女对双亲均值的回归系数，但是 Lush 的解释已完全不同。遗传力所以总是小于 1，即子女均值所以总是要向群体均值“回归”，Lush 认为其原因在于性状表型值中总包含有不能遗传的环境效应值。所以，通过遗传力就可以估计性状的遗传值，就可以掌握数量性状的遗传规律。

Lush 和其学生 L. N. Hazel 提出重复率，遗传力和性状间遗传相关三个遗传参数构成数量遗传学的核心，特别是遗传力这一参数，不但贯穿整门学科，而且具有超出学科范围的意义。D. C. Acker (1987) 甚至将遗传力的概念及其应用列为近百年畜牧科学的五项最大成就之一。

动物数量遗传学一开始就是应动物育种工作的需要而产生的，由于研究对象和研究领域的不同，它采用统计学方法作为研究手段，这也象生化遗传学采用生物化学手段一样是必然的，这不但不排斥采用其他研究手段对数量性状进行研究，而且与其他遗传学分支结合正是数量遗传学发展的重要方向。在发展过程中，它吸收了不少现代数学理论和方法。为了适应育种工作的经济要求，它还引入了不少经济学的内容。数量遗传学是一门朝气蓬勃的现代学科，它的特点就是一方面具有非常明确的为生产服务的目的，另一方面又毫不保守地吸收各种现代科学，兼容各家之长来发展自己，所以在短短半个世纪中，飞快成长，已经成为一门影响很大的成熟学科。近年来，它又与一些新兴学科，如生物化学、分子生物学、遗传工程、电子计算机技术等紧密结合，相互渗透。这样一门充满活力的学科其发展前途肯定是无可估量的。

第一章 基因频率与基因型频率

群体遗传学研究的是孟德尔群体的遗传结构及其变化，而群体遗传结构的主要标志就是基因频率与基因型频率。

§ 1 基本概念

1. 群 体

普通遗传学中所介绍的孟德尔遗传定律是本世纪初前后在“家庭水平”上研究遗传现象所发现的几条规律。所谓“家庭水平”上研究就是限于研究特定双亲与其后裔间的遗传关系。但是在研究生物进化和动物育种中，除了需要在“家庭水平”上，还需要在“群体水平”上了解生物的遗传规律。群体遗传学的任务就是在“群体水平”上研究生物的遗传规律。

这里所说的群体都是指生物的孟德尔群体而言。所谓“孟德尔群体，是一群相互能交配繁殖的个体，它们享有一个共同的基因库”(Th. Dobzhansky, 1955)。大至一个物种，小至一个品系，一个畜群都可以是一个群体。

一个个体的存在是有一定时间限度的，除非发生突变，其遗传性一般是终生不变的，而一个群体的存在一般是长时间的，不受其成员的个体生命的限制。群体可大可小，分布范围可广可窄。群体的遗传性，即遗传结构，可以受多种因素的影响而逐代改变，这一点对育种工作是大为有利的。

2. 基因频率

基因频率是群体遗传组成的基本标志，不同群体的同一基因往往频率不同。

所谓基因频率就是在一个群体中某一基因对其等位基因的相对比率。例如有的牛群(如黑白花牛)大多数牛都有角，只有个别“秃头”；有的牛群(如安格斯牛)则几乎全部都是无角；杂种牛群中有角牛与无角牛的比例更是各不相同。为什么这些牛群在“角”这个性状上表现各不相同呢？是不是它们的基因都不相同呢？不是的，据研究在大多数的牛群，控制“角”这个性状的基因都是相同的，都是由一对等位基因控制的。决定“无角”的基因是显性的，我们用P代表它；决定“有角”的基因是隐性的，用p代表之。只是在不同的牛群，P基因与p基因的比率各不相同而已。譬如在某一牛群(或品种)中，P基因占1%，p基因占99%，则这个牛群中只有约2%的牛是“秃头”，其余都是有角的。另一个牛群(或品种)中，P基因占100%，p基因占0%，这群牛当然都是无角的。上面所说的1% (或写作0.01) 和100% (或写作1) 就是P的基因频率，99% (或写作0.99) 和0% (或写作0) 就是p的基因频率。

各等位基因的基因频率的总和等于 1。如上例前一群体中， $0.01+0.99=1$ ，后一群体中 $1+0=1$ 。例如人的 ABO 血型决定于三个等位基因： I^A 、 I^B 和 i 。据调查 (C. S. Fan, 1944)，中国人(昆明)中 I^A 的基因频率约 0.24， I^B 的基因频率约 0.21， i 的基因频率约 0.55，三者的总和 $0.24+0.21+0.55=1$ 。

由于基因频率是一个相对比率，是以百分率表示的，因此其变动范围在 0~1 之间，一般都写成小数形式，没有负值。

3. 基因型频率

在二倍体生物的体细胞中，基因都是成对存在的。一对或几对基因构成某个性状的基因型，因此一个性状的遗传性不仅决定于基因，更直接地决定于基因型。

一个群体中某一性状的各种基因型间的比率叫做基因型频率。例如牛角的有无，决定于一对等位基因 P 和 p，它们组成的基因型有三种：PP、Pp 和 pp，前两种表现无角，后一种表现为有角，这三种基因型的频率在各群体中各异。譬如在某一牛群中，PP 占 0.01%，Pp 占 1.98%，pp 占 98.01%，也就是说，PP 基因型频率为 0.0001，Pp 基因型频率为 0.0198，pp 基因型频率为 0.9801，三者合计等于 1，即 100%。

基因型不等于表现型，基因型频率并不是表现型的比率。例如上述三种基因型，它们的表现型只有两种，无角与有角，两者的比率是 1.99% (0.01%+1.98%) 和 98.01%。

基因型是由基因组成的，当然两者的频率是密切关联的。为了说明简便起见，不妨以一对基因为例。设 A 与 a 是一对等位基因，它们的频率分别为 p 和 q。这一对基因组成的基因型有三种，即 AA、Aa 与 aa，它们的频率顺序以 D、H 和 R 代表之。在整个群体里，有 D 个 AA 基因型，每个基因型有两个 A 基因，因此有 2D 个 A 基因；另有 H 个 Aa 基因型，其中有 H 个 a 基因和 H 个 A 基因；还有 R 个 aa 基因型，包含有 2R 个 a 基因。这样，A 基因的频率：

$$p = \frac{2D+H}{(2D+H)+(H+2R)} = \frac{2D+H}{2(D+H+R)} = \frac{2D+H}{2} = D + \frac{1}{2}H$$

a 基因的频率：

$$\begin{aligned} q &= \frac{H+2R}{(2D+H)+(H+2R)} = \frac{H+2R}{2(D+H+R)} = \frac{H+2R}{2} \\ &= \frac{1}{2}H+R \end{aligned} \quad (1-1)$$

在伴性基因，我们可以把雌雄看作两个群体，雄异型生物 (XX—♀, XY 或 XO—♂) 的雌性群体中，基因频率与基因型频率的关系与上面所说一般基因一样；在雄性群体中，基因型只有两种 (Ao 和 ao)，所以基因频率就等于基因型频率。雌异型生物 (如家禽) 的情况正相反。

4. 随机交配

随机交配的经典定义是：在一个有性繁殖的生物群体中，任何一个雌性或雄性的个体与任何一个相反性别的个体交配的概率相同，也就是说，任何一对雌雄的结合都是随机的，不受任何选配的影响。

随机交配不是自由交配，自由交配是将公母畜混放在一个群体中，任其自由结合。这样的交配方式实际上还是有选配在其中起作用的。最明显的就是粗野强壮的雄性，其配种的概率就高于其他雄性个体，这就不符合随机交配的定义。

在畜牧业实践中，就畜禽个体而言，完全不加任何选配而随机的交配是不多的，但是就某一性状而言，随机交配的情况还是不少的。就某一性状而说，随机交配指的是在进行个体间的选配时，对该性状的基因型完全未加注意。例如一般在进行猪的配种时，根本不注意交配双方的血型，因此就血型这个性状来说可以说是随机交配的。这里有一个前提，就是猪的血型这个性状与其他被注意选择的性状之间没有相关，否则的话，由于选择了与血型相关的性状而对血型进行了间接选择，就可能不再是随机交配了。当然在近交情况下，对于每个性状都不是随机交配。

§ 2 哈代-温伯格定律

1. 定律要点

英国数学家哈代 (Hardy) 和德国医生兼遗传学家温伯格 (Weinberg)，经过各自独立的研究，于 1908 年同一年分别发表了有关基因频率与基因型频率的主要遗传规律，现在公称为哈代-温伯格定律，或者叫做基因平衡定律。这个定律的要点是：

(1) 在随机交配的大群体中，若没有其他因素的影响，基因频率一代一代下去始终保持不变。

(2) 在一个大群体中，无论其基因型频率如何，只要经过一代随机交配，常染色体基因型频率就达到平衡状态。没有其他因素的影响，以后一代一代随机交配下去，这种平衡状态始终保持不变。

(3) 在平衡状态下，基因型频率与基因频率的关系是：

$$D=p^2, H=2pq, R=q^2 \quad (1-2)$$

不难看出，这些关系式来自以下等式：

$$(p+q)^2=p^2+2pq+q^2$$

由此推理，如有 A、B、C 三个复等位基因，其频率分别为 p、q、r。它们组成的基因型共有 6 种，即 AA、AB、AC、BB、BC、CC。在平衡状态下，各基因型频率与基因频率的关系是：

$$(p+q+r)^2=p^2+2pq+2pr+q^2+2qr+r^2$$

为数更多的复等位基因情况，也依此类推。

多对基因组成的基因型不是一代随机交配就能达到平衡的，涉及的基因位点愈多，达到平衡需要的代数也愈多，有连续的多位点基因型达到平衡的速度更慢。伴性基因也需要若干代才能逐渐达到平衡。

2. 数学证明

这个定律的证明是很简单的。以一对基因为例：设 0 世代的基因频率为 p_0 和 q_0 ，基因型频率为 D_0 、 H_0 和 R_0 ；一世代的基因频率为 p_1 和 q_1 ，基因型频率为 D_1 、 H_1 和 R_1 ；二世

代的基因频率为 p_2 和 q_2 ，基因型频率为 D_2 、 H_2 和 R_2 。

0 世代的个体所产生的配子带有 A 基因或 a 基因的概率各为多少呢？在一个大群体中，任何一个配子带有某一基因的概率就等于该基因在这个群体中的频率。因此 0 世代的个体所产生的配子，带有 A 基因的概率为 p_0 ，带有 a 基因的概率为 q_0 ，也就是说，有 p_0 个配子带有 A 基因， q_0 个配子带有 a 基因。

表 1-1 0 世代基因频率与一世代基因型频率

配子	精子			
	基因	A		a
卵 子	A	频率	p_0	q_0
		p_0	AA	Aa
	a	q_0	Aa	aa
			p_0q_0	q_0^2

在随机交配下，各雌雄配子随机结合，结合成一世代个体的各基因型的频率如表 1-1 所列。

$$D_1 = p_0^2, H_1 = 2p_0q_0, R_1 = q_0^2$$

由此计算一世代基因频率：

$$p_1 = D_1 + \frac{1}{2}H_1 = p_0^2 + p_0q_0 = p_0(p_0 + q_0) = p_0$$

$$q_1 = \frac{1}{2}H_1 + R_1 = p_0q_0 + q_0^2 = q_0(p_0 + q_0) = q_0$$

同样可以证明

$$p_2 = p_0, q_2 = q_0;$$

$$p_n = p_0, q_n = q_0;$$

也就是一代一代下去，基因频率不变。

无论 0 世代的基因型频率如何，其基因频率总是 p_0 和 q_0 。在随机交配下，一世代的基因型频率就是 p_0^2 、 $2p_0q_0$ 和 q_0^2 ，基因频率仍为 p_0 和 q_0 。因而二世代的基因频率仍为 p_0^2 、 $2p_0q_0$ 和 q_0^2 。由于 $p_0 = p_1 = p_2 \cdots = p_n$ ， $q_0 = q_1 = q_2 \cdots = q_n$ ，因此足码可以取消。从一世代开始，每个世代的基因型频率都是 p^2 ， $2pq$ 和 q^2 ，始终保持不变。

举一实例说明之。譬如 0 世代的基因型频率： $D_0 = 0.6$ ， $H_0 = 0.4$ ， $R_0 = 0$ 。基因频率： $p_0 = D_0 + \frac{1}{2}H_0 = 0.6 + 0.2 = 0.8$ ； $q_0 = \frac{1}{2}H_0 + R_0 = 0.2 + 0 = 0.2$

一世代的基因型频率：

$$D_1 = p_0^2 = 0.8^2 = 0.64$$

$$H_1 = 2p_0q_0 = 2 \times 0.8 \times 0.2 = 0.32$$

$$R_1 = q_0^2 = 0.2^2 = 0.04$$

一世代的基因频率：