外籍学者讲学材料之十六

植物遗传学

美国堪萨斯大学　梁学礼教授

(1981.5.9——7.1)

上　册

农业部科技局
北京农业大学
1982年1月

前　　言

1981年5月9日至7月1日，中央农业部科技局委托北京农业大学主办《全国植物遗传讲习班》，由美国堪萨斯大学梁学礼教授主讲。全国29个省、市的高等农业院校及农业科学院65名讲师助研以上的人员参加了学习。另有部分人员列席旁听。

梁学礼教授共讲授了十三章，计有概率，细胞和染色体，细胞分裂和繁殖，连锁和重组，染色体结构变异，染色体数量变异和非正倍体，远缘杂交和植物改良，组织培养和谷物改良，抗病遗传，细胞质遗传，多倍体遗传，群体遗传，数量性状遗传。另外两章未讲。讲授时数共计约130学时，外加辅导答疑座谈讨论等各种活动。

本教材分别由林学建、周国玉、夏奇梅、蒋志谦、龚学明、许复华、王国昌，付焕延、陈昌颐、吴一民、王焕丕、肖增宽、薛启汉、戚家华、陈伟程、李正焯、谢德泉、罗耀武、刘克治、陈德鑫、冯启涣、何基娜、陆漱韵整理。除概率一章经梁学礼教授审阅修改外，其余各章分别由付焕延、肖增宽、程经有、陆漱韵、陈德鑫、陈伟程、许启凤、赵世绪、沈克全、王焕丕、芦宗海审阅修改。由李惟基、祖国洪绘制插图。最后由程经有、沈克全、陆漱韵分别翻译图表说明和习题，并作了复审和统编等工作。由于我们水平有限，错误缺点在所难免，希望读者发现不当之处，及时指出，不胜感谢，另外拟将整编教材寄请梁学礼教授修改并补充第十章和第十二章，以便今后出版。

本教材内容丰富，适于农业高等院校教师、研究生、农业科技人员及其它生物学科的教学和科研人员参考。

北京农业大学植物遗传教研组

1982年1月

目　　录

第一章　概率(Probability)

Ⅰ、概率概念（Concept of probability）

概率又叫或然率或可能率，指的是某一事件所发生的机会、可能性等等。概率的公式如下：

$$P_{(A)}=\lim_{n\to\infty}\frac{n_A}{n}$$

$P_{(A)}$ = A事件发生的概率。

n = 群体中之个体数或测验次数。

n_A = A事件在群体中出现的次数。

上述公式是概率在统计学上的定义。$n\to\infty$表示n趋问于无穷大，就是个体数极大或测验次数极多的意思。当n不是无穷大的时候，P值只是一个估计值，在样品与样品之间会有差别，和数学上的定义不完全相同。

例如，扑克牌共52张，其中红26、黑26。抽红的机会或抽黑的机会各是$\frac{26}{52}=\frac{1}{2}$。这在数学上是很清楚的。如果从统计学上看，每次抽一张然后放回，洗匀，再抽一张，再放回，连续抽10次，那么不一定5次是红，5次是黑。也可能4次黑、6次红，或3次黑、7次红等等。当抽牌的次数增加时，红与黑的比例接近1∶1。即50%∶50%。相当于数学上概率的定义，亦即n越大，概率准确性越高。

为什么要用概率？

为什么在遗传、育种上要用概率，因为我们做植物遗传育种试验，面临的最大问题，就是怎样确定我们试验结果的可靠性。因为实际结果常常和理论预期的结果有偏差，需要加以正确分析和判断。例如，孟德尔的豌豆杂交试验。一对性状测交后代在理论上应得到1∶1分离群体。

Rr × rr

↓

$\frac{1}{2}$Rr（园粒）∶$\frac{1}{2}$rr（皱粒）

但实际调查这个群体的结果，可能两种类型的比率与理论数存在差异。这就需要用概率的概念，分析、判断它的真实性，找出能否被我们所接受的概率。此外，作物品种比较试验用A、B两个品种，在不同地点、不同年度间试验结果可能很不一样。我们必需对差异的真实性加以分析和判断，因为造成差异的原因有很多，某些造成差异的因素是无法完全控制的，例如：

一、土壤变异。肥力，水分，酸碱度，微生物的种类及数目等等，都可能在一块试验田内有不同的含量，无法完全控制到均匀的地步。

二、取样误差。群体中经常只能取样品作试验，希望样品的结果能够应用在群体上。最好的样品是随机样品，因避免或尽可能减小人为选择的干预，有较大的代表性。

三、气候条件。气候变化也无法控制，田间试验会由于风、雨量、温度和湿度等等的影响，也会发生偏差。

仍以上述豌豆杂交试验为例，试验结果是1:1或3:1，那么实际数和理论数的差异究竟要大到什么程度或小到什么程度才能接受1:1或3:1的假说，我们常用的卡方测验就是以概率为理论根据的。又如，品种产量的卡方测验，F测验，t测验等等也是根据概率作出判断的。如果显著水准达95%或99%，就是说100次中95次或99次结果可靠，只有5次或1次的偏差。

有了概率概念，我们通常对试验结果不会认定100%正确，只能说观察结果与期待值达到基本符合。

计算概率时，发生机会的表达：

a = 发生，b = 不发生，如 a = b = 1，

则 $$p_a=\frac{a}{a+b}=\frac{1}{1+1}=\frac{1}{2}=P,$$

$$p_b=\frac{b}{a+b}=\frac{1}{1+1}=\frac{1}{2}q。$$

概率的二个定则：

(一) $p+q=1 \qquad p=1-q \qquad q=1-p$

(二) $1\geqslant p\geqslant 0$

例如：

$$\begin{gathered}Rr\\ \downarrow\otimes\\ \underbrace{\frac{1}{4}RR:\frac{2}{4}Rr}:\frac{1}{4}rr\\ \frac{3}{4}R:\frac{1}{4}r\\ 3\quad:\quad 1\end{gathered}$$

$$P_{(R)}=\frac{3}{3+1}=\frac{3}{4}=p\text{(R发生概率)}$$

$$P_{(r)}=\frac{1}{3+1}=\frac{1}{4}=q\text{(R不发生的概率)}$$

$$p+q=\frac{3}{4}+\frac{1}{4}=1$$

∴ 各种概率之和等于1，$1\geqslant P\geqslant 0$或$1\geqslant q\geqslant 0$

Ⅱ、概率规则：（Probability rules）

一、独立事件：

假设A和B都是独立事件。

如　$p_{(A)} = A$（A事件发生的概率）

　　$p_{(B)} = B$（B事件发生的概率）

则　$p_{(A \cdot B)} = p_{(A)} \cdot p_{(B)}$（A、B同时发生的概率）

两个独立事件同时发生的概率等于每个事件各自发生时概率的乘积。

例如，抛掷一个铜币，正、反面出现的机会各占$\frac{1}{2}$，即：$p_{(正)} = \frac{1}{2}$，$p_{(反)} = \frac{1}{2}$ 如果同时一次抛掷二个铜币（二个独立事件），二个都是正面的概率是 $P = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$；同理，二个都是反面的概率也是 $p = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$，因为这二个铜币正面与反面发生的机会是独立的。

以作物为例，如豌豆 GgPpTt 三对基因为独立遗传，则$n = 3$　$2^n = 2^3 = 8$，可出现8种类型配子，每一种配子出现的概率都是$\frac{1}{8}$。其组合如下：

$$GpT = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{8}$$

$$Gpt = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{8}$$

$$\vdots \qquad\qquad \vdots$$

$$gpt = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{8}$$

这是由于每一对基因的一个等位因子（allele）分向二极的机会都是$\frac{1}{2}$，而且三对基因彼此是独立的。同时要附带假设：

（一）　每一种配子或合子成活率一样。

（二）　没有所谓“倾向选择”（Preferntial selection）

再如，同源多倍体形成配子的概率也是如此。例如：

$S_1S_2s_3s_4$（同源四倍体如马铃薯）

↓

$S_1S_2, S_1s_3, S_1s_4, S_2s_3, S_2s_4, s_3s_4$　（形成二倍体的配子，有6种基因型组合，各占$\frac{1}{6}$）

这种配子类型数也可用组合公式加以计算。

$$C_y^n = \binom{n}{r} = \frac{n!}{r!(n-r)!}$$

$$C_2^4 = \frac{n!}{r!(n-r)!} = \frac{4\cdot3\cdot2\cdot1}{2\cdot1\cdot2\cdot1} = 6$$

n ＝总个体数或染色体数

r ＝每个配子所携带的染色体数，

如果，同源六倍体（甘薯）$S_1S_2s_3s_4s_5s_6$形成三倍体的配子类型数是：

$C_3^6 = \frac{6!}{3!(6-3)!} = 20$〔注：$0! = 1$，$n^\circ = 1$（n是有理数）〕。同理，如果同源四倍体是二对基因，如：$S_1S_2S_3s_4T_1T_2T_3t_4$则　形成配子类型的概率应该是：

$C_2^4 = 6$ $C_2^4 = 6$　　　　配子类型　　　　概率

$S_1S_2 = \frac{1}{6}$　　$T_1T_2 = \frac{1}{6}$　　$S_1S_2T_1T_2 = \frac{1}{6}\cdot\frac{1}{6} = \frac{1}{36}$

$S_1S_3 = \frac{1}{6}$　　$T_1T_3 = \frac{1}{6}$　　$S_1S_3T_1T_3 = \frac{1}{6}\cdot\frac{1}{6} = \frac{1}{36}$

⋮　⋮　　⋮　⋮　　⋮　　　⋮

$S_3s_4 = \frac{1}{6}$　　$T_3t_4 = \frac{1}{6}$　　$S_3s_4T_3t_4 = \frac{1}{6}\cdot\frac{1}{6} = \frac{1}{36}$

二、互斥事体：（Matually Exclusive Events）

所谓互斥事件，就是其中一个事件发生，另一个事件就不能发生的现象；即两个不可能同时发生的事件叫互斥事件。如一个铜币，出现正面，就不出现反面；如出现反面，就不出现正面。这就是互斥事件。互斥事件的任何一种单一发生的概率等于它们各自发生概率的和。

如　$p_{(A)} = AP_{(B)} = B$

则　$p_{(A\cdot B)} = p_{(A)} + p_{(B)}$

如一个完整骰子有六面每面机会都是$\frac{1}{6}$，如果我们问不是1点就是2点的概率是多少？因为$p_{(A=1点)} = \frac{1}{6}$，$p_{(B=2点} = \frac{1}{6}$，所以，$p_{(A\cdot B)} = \frac{1}{6} + \frac{1}{6} = \frac{2}{6} = \frac{1}{3}$（互斥事件的和）。

又如，GgPp形成配子类型是$\frac{1}{4}$GP ：$\frac{1}{4}$Gp：$\frac{1}{4}$gP：$\frac{1}{4}$gp，假如问发生1个显性，1个隐性基因组合的这种配子概率是多少？或者问不是Gp就是gP配于的概率是多少？即$p_{(A\cdot B)} = \frac{1}{4} + \frac{1}{4} = \frac{2}{4} = \frac{1}{2}$

三、组合事件：（Combining Probability Rule）

独立事件和互斥事件二者规则都需要用到的事件构成组合事件。

例：RrGg

$\downarrow\otimes$

如果我们问后代个体基因型是3个显性，1个隐性的概率是多少？那么独立事件、互斥事件的概率规则都要用到。

（一）表明独立事件，同时出现的概率是相乘，如：1.♀、♂不同配子的概率是：

$$RG=\frac{1}{2}\cdot\frac{1}{2}=\frac{1}{4}$$

$$Rg=\frac{1}{2}\cdot\frac{1}{2}=\frac{1}{4}$$

$$rG=\frac{1}{2}\cdot\frac{1}{2}=\frac{1}{4}$$

$$rg=\frac{1}{2}\cdot\frac{1}{2}=\frac{1}{4}$$

2.♀、♂配子形成合子时如：

$$\frac{1}{4}RG\times\frac{1}{4}RG=\frac{1}{16}RRGG$$

$$\frac{1}{4}RG\times\frac{1}{4}Rg=\frac{1}{16}RRGg$$

$$\vdots \qquad\qquad \vdots$$

$$\frac{1}{4}rg\times\frac{1}{4}rg=\frac{1}{16}rrgg$$

表1—1　F_1代两对基因杂合体RrGg通过自交，在子代中每个基因组合的可能的基因型排列的概率

表配	♀	♂	概率	♀	♂	概率	♀	♂	概率	♀	♂	概率	♀	♂	概率
子代基因型的排列	RT	RG	$\frac{1}{16}$	RG	Rg	$\frac{1}{16}$	RG	rg	$\frac{1}{16}$	Rg	rg	$\frac{1}{16}$	rg	rg	$\frac{1}{16}$
				RG	rG	$\frac{1}{16}$	rg	RG	$\frac{1}{16}$	rG	rg	$\frac{1}{16}$			
				Rg	RG	$\frac{1}{16}$	Rg	Rg	$\frac{1}{16}$	rg	Rg	$\frac{1}{16}$			
				rG	RG	$\frac{1}{16}$	rG	rG	$\frac{1}{16}$	rg	rG	$\frac{1}{16}$			
							rG	Rg	$\frac{1}{16}$						
							Rg	rG	$\frac{1}{16}$						
组合	4显性基因		$\frac{1}{16}$	3显性基因 1隐性基因		$\frac{4}{16}$	2显性基因 2隐性基因		$\frac{6}{16}$	1显性基因 3隐性基因		$\frac{4}{16}$	4隐性基因		$\frac{1}{16}$

（二）表明互斥事件出现的概率是相加。

如求出现3个显性，1个隐性的这种个体的概率，即

$$\frac{1}{16}RRGg+\frac{1}{16}RrGG+\frac{1}{16}RRGg+\frac{1}{16}RrGG=\frac{4}{16}=\frac{1}{4}$$

四、条件概率：（Conditional probability）

如果从男性和女性中调查色盲的情况，而我们只要求女性中的色盲概率，即有条件的概率。

n ＝群体中的个体数

n_H＝女性的人数

n_A＝有色盲的人数

$p_{(H)}=\frac{n_H}{n}$（女性的概率）

$$p_{(A)}=\frac{n_A}{n}$$ （色盲人的概率）

那么，既是色盲又是女性的概率则是：

$$p_{(A \cdot H)}=\frac{n_{A \cdot H}}{n}$$

换句话说，如果只调查女性中的色盲者，把男性排除，就要用条件概率。

公式：$p_{(A/H)}=\lim\limits_{n\to\infty}\frac{n_{A \cdot H}}{n_H}$

$$=\frac{p_{(A \cdot H)}}{p_{(H)}}\left(p_{A \cdot H)}=\frac{n_{AH}}{n}\,p_{(H)}=\frac{n_H}{n}\right)$$

又如，马铃薯四倍体种和二倍体种间杂交，在后代中获得多单倍体的概率的计算。如果只调查绿株又是多倍单倍体（polyhaploid）的概率，即条件概率。

♀ $\dfrac{\text{S. tuberosum}}{2n=4x=48}$ × $\dfrac{\text{S. phureja}}{2n=2x=24}$ ♂

绿色株 ↓ 红色株

$(n=2x=24)$ F_1 $(n=x=12)$

红色株 $2n=3x=36$（经过受精）

绿色株（可能是母本大孢子形成二倍体配子伪性结实 $n=24$—多倍单倍体；或是母本自交或别株花粉混杂的四倍体后代也表现绿色）

设：A = 多倍单倍体的绿株

H = 绿株

$n=F_1$群体的个体数

$$p_{(A)}=\frac{n_A}{n} \qquad p_{(H)}=\frac{n_H}{n}$$

$$p_{(A/H)}=\lim_{n\to\infty}\frac{n_{AH}}{n_H}$$

所以条件概率的定义：$p_{(A/H)}=\dfrac{p_{(AH)}}{p_H}$

Ⅲ、二项分布（Binomial Expension）

下面讨论二项分布，应用二项展开式，比较容易求得某项的概率。

首先我们可以把群体分成二项群体；

n = 群体的个体数，p = 某一类型的频率 q = 另一类型的频率。

p与q的二项展开式如下：

$$(p+q)^n=p^n+np^{n-1}q+\frac{n(n-1)}{2!}p^{n-2}q^2+\frac{n(n-1)(n-2)}{3!}p^{n-3}q^3+\cdots+q^n$$

例如，在抗病和不抗病（染病）杂交后代群体中，估计抗病及染病个体的不同组合概率。

$$\begin{array}{ccc} & Ss \times ss & \\ (\text{抗病}) & \downarrow & (\text{染病}) \\ & 1Ss : 1ss & \\ & (\text{抗}) \quad (\text{染}) & \end{array}$$

设在群体中随机取样6株。

$n=6, p=q=\frac{1}{2}$（注意在不同群体中p、q也可以等于其他比例），那么，在此样品中，估计各种抗、染病的个体不同组合概率则是：

$$\left(\frac{1}{2}+\frac{1}{2}\right)^6=\left(\frac{1}{2}\right)^6+6\left(\frac{1}{2}\right)^5\cdot\left(\frac{1}{2}\right)+\frac{6.5}{2!}\left(\frac{1}{2}\right)^4\cdot\left(\frac{1}{2}\right)^2+\frac{6\cdot5\cdot4}{3!}\left(\frac{1}{2}\right)^3\cdot\left(\frac{1}{2}\right)^3$$

$$+\frac{6.5.4.3}{4!}\left(\frac{1}{2}\right)^2\cdot\left(\frac{1}{2}\right)^4+\frac{6.5.4.3.2.1}{5!}\left(\frac{1}{2}\right)\cdot\left(\frac{1}{2}\right)^5$$

$$+\frac{6.5.4.3.2.1}{6!}\left(\frac{1}{2}\right)^6$$

$$=\frac{1}{64}+\frac{6}{64}+\frac{15}{64}+\frac{20}{64}+\frac{15}{64}+\frac{6}{64}+\frac{1}{64}$$

6个个体都抗病的概率　5个个体抗病 1个个体染病的概率……………6个个体都染病的概率

〔注：n＝群体数（样品）　n＋1＝项数〕

如果n很大，二项展开式太长。那么，求任一项单独的系数，也可用如下公式：

$$\frac{n!}{r!(n-r)!}$$

设：$n=6$　$r=3$（即求3个个体抗病，3个染病的组合系数）

$$\frac{6!}{3!(6-3)!}=\frac{6.5.4.3.2}{3.2.3.2}=20$$

如果：$n=100$　$r=50$（即求50个抗病，50个染病的组合概率）

即：$\frac{100!}{50!50!}\cdot\left(\frac{1}{2}\right)^{50}\left(\frac{1}{2}\right)^{50}$〔注：数字虽大但可用电子计算器简便计算〕

上述$p=q=\frac{1}{2}$，显示二项式展开各项系数分布是对称的。如果，$p \neq q$也同样可以作出二项分布群体。如：

$$\begin{array}{c} Rr \\ \downarrow\otimes \\ \underbrace{1/4RR : 2/4Rr}_{3/4R} : 1/4rr \\ \end{array}$$

3/4R　　1/4r

（园）　（皱）

则：$p_{(R)}=\frac{3}{4}$，$q_{(r)}=\frac{1}{4}$

∴二项式是 $\left(\frac{3}{4}+\frac{1}{4}\right)^n$,这种二项式展开后，则表现不对称分布，但应用和估算方法同上。

Ⅳ、假说之测定与二项群体之取样
(Testing Hypothesis and sampling from Binomial Popultions)

二项分布群体：

我们把研究的生物群体分成二组归类，如植株高、矮；早熟、晚熟；水果的酸、甜等等分为二组归类进行研究。那么这种群体就叫二项群体。

$$(p+q)^n \qquad p=q=\frac{1}{2}$$

设：n＝50（样品数），二组数目正好是 25：25，即 $P=q=\frac{1}{2}$ $(P+q)^{50}$

$$C_r^n=C_{25}^{50}=\frac{n!}{r!(n-r)!}=\frac{50!}{25!25!}=0.1123$$

也就是说在上述群体中，刚好取得 25:25 的机会，亦即偏差＝0，这种机会是11.23%。如果允许偏差±1，即：24—26，26—24，相差±1，那么机会约10%。如果允许偏差±2，即23—27，27—23，相差±2，机会约7⁺%。据此可以作出如下图：

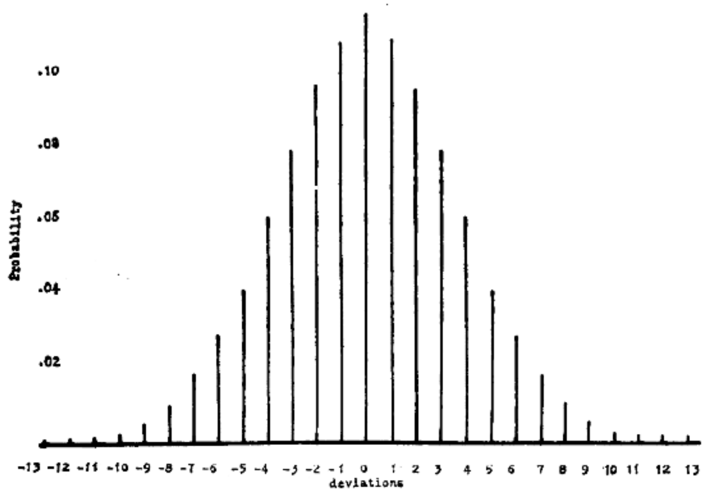


图1—1　二项分布群体

这是：1:1群体的二项分布，1条线的比数占11.23%，叫点估计没有偏差；允许自＋3到－3七条线之和的比数占50%以上，叫段估计，就是允许一定偏差，代表面更大，取得机

会更多些。

理论与实际有时不一样，产生偏差的原因可能是：

一、群体p和q比率原来就不是1∶1或1∶3的理论数字：

假如群体组成p和q比率正确，但实测有偏差则有下列原因，

二、个体生活力不等；如aa之活力低于AA及Aa；

三、取样误差；样品不是随机样品，没有代表性；

四、实际存在的分类困难；如可孕个体与不孕个体之区别。

纠正的办法，应多取样品，多次重复试验。例如：喷药试验，假设某药浓度一定，24小时后，调查死虫和活虫数目，来估计该药剂的有效度。

假设调查数据是100个虫，其中死虫占42，活虫58，那么结论是什么呢？

我们可以说，药的有效度是42%，因为没有考虑存在偏差，只有一个数字，叫点估计。

如果考虑偏差，从低到高范围有两个数字，叫段估计。

利用二项群体分布表（查附表A－1或A－2·CI·95%或99%，在A－1表中（95%可信程度），n＝100（样品），f＝42（死虫数），可得药有效度是32—52%，我们可下这种判断，在100次试验中有95次所得结果将在32—52%之间，只有5次可能超出此范围。

另外，如果f数字大于50，可通过查表，间接计算。比如：n＝100，f＝70

100—70＝30，查A－1表，得0.21—0.40，则药的有效度为：1－〔0.21－0.40〕＝0.60－0.79＝60—79%。

如果样品数在表上不能找到怎么办？也还可以有另一种查表间接计算方法；

比如：n＝55　f＝10

查附表A－1　用n＝50(55前之样品值），得0.10—0.34。

$$\text{间接计算公式}\left\{\begin{aligned}\frac{0.10\times 50}{55}&=0.09\\ \frac{0.34\times 50}{55}&=0.31\end{aligned}\right\}\text{因此，}$$

（n＝55，f＝10）药的有效度为9—31%。

V、多项式分布（multinomial Distribution)

生物中有时群体不仅分为二组，而可分为三组或更多组。我们可以用多项式分布市答案。

例：

$$\begin{array}{c}Rr\\ \downarrow\otimes\\ \frac{1}{4}RR : \frac{2}{4}Rr : \frac{1}{4}rr\end{array}$$

如果研究此群体，同时对三种基因型都考虑，即三项分布。三项式的公式是：

$$(p+q+r)^n$$

求某项系数的公式是：

$$\frac{n!}{r!u!v!}p^{r}\cdot q^{u}\cdot m^{v}$$

r = RR实际观察数（红花）

u = Rr实际观察数（粉红花）

v = rr实际观察数（白花）

$P = 红花频率 = \frac{1}{4}$

$q = 粉红花频率 = \frac{2}{4}$

$m = 白花频率 = \frac{1}{4}$

设： Rr

↓⊗

问在此后代群体中取样4株，刚好1红：2粉红：1白的机会（概率）是什么？

$$n=4 \qquad \frac{4!}{1!2!1!}\left(\frac{1}{4}\right)^{1}\cdot\left(\frac{1}{2}\right)^{2}\cdot\left(\frac{1}{4}\right)^{1}$$

$$=12\times\frac{1}{64}=\frac{3}{16}$$

结论：取16个样品，每个样品包括4株，其中有3个样品刚好是1红：2粉红：1白的比例。

Ⅵ、常态分布（Normal Probability Distribution）

生物学中常用的二种分布是：

二项分布——用于质量性状（上面已讨论），及常态分布——用于数量性状。

还有一种植物上不常用的分布，称为Poisson Distribution，这在n很大，p很小时有此分布，如微生物研究，某细菌突变频率分布等。

质量性状和数量性状之区别：

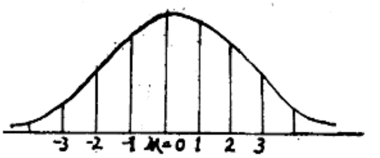
质量性状	数量性状
一、少数基因控制，	一、有关之基因数目多，
二、不连续分布，	二、连续分布，
三、每个基因功用显明（major genes），	三、每个基因功用微小而不显明（minor genes），
四、不受环境影响或影响很小，	四、受环境影响甚大，

常态分布由二个常数（Constants）决定：

平均数：μ，μ 在样品中由m或$\bar{x}$来估计。标准差：σ，σ 在样品中由S来估计。μ 决定分布的位置，σ 决定分散程度。

常态曲线特性：表现左右对称平滑的曲线，曲线下的面积相当于1。

可以把各种不同的常态曲线分布成标准曲线分布(**Standard Normal Distribution**)标准化常态曲线是以$\mu=0$，$\sigma=1$而作出的。怎样把任何曲线变为常态曲线呢？可用如下公式求出标准曲线：



$$Z=\frac{X_i-\mu}{\sigma}$$ $X_i(1\cdots i\cdots n)$ = 代表群体各个变数

Z值求得后，即可自附表A－3中求其相应的概率。

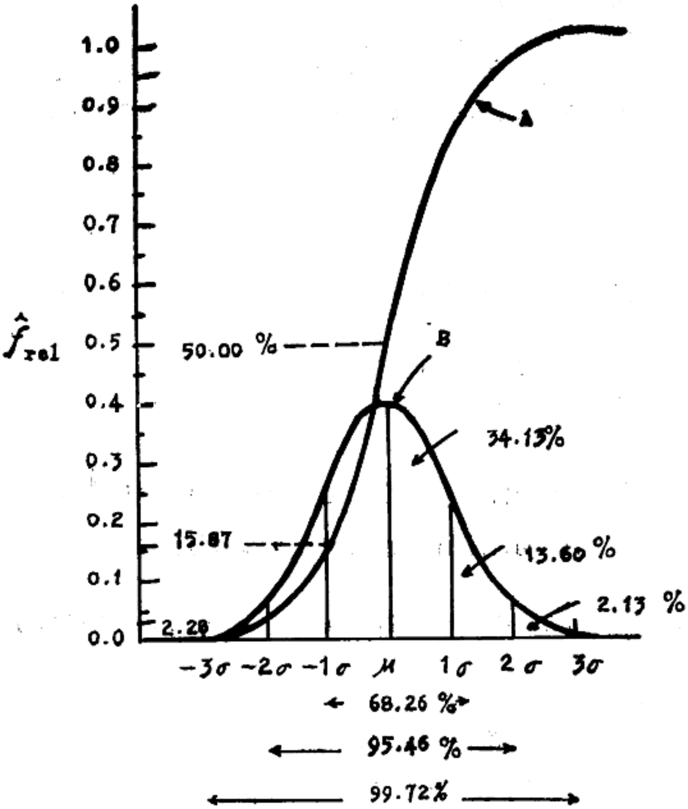


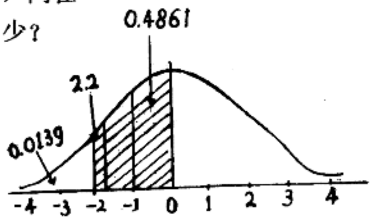
图1—2　在正态概率密度函数和累积正态分布函数下的面积

A：累积正态分布函数，B：正态概率密度函数

例1，假设某高粱群体株高之$\mu=182$cm，$\sigma=10$cm，问在这种群体中选择株高160cm及低于160cm植株的概率是多少？

应用上述公式：

$$\frac{160-182}{10}=2.2$$



用—2.2查附表A—3得0.4861(从原点→2.2的面积)，再用0.5—0.4861＝0.0139（注：0.5表示常态曲线

半边的面积）。

结论：在此群体中逢机选得株高160cm及160cm以下的植株的机会为1.39%。

例2.大豆A，B二个品种杂交后，将F_2群体分为二组

A × B

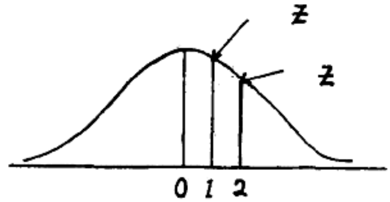
↓

F_1

↓ ⊗

F_2

Ⅰ：旱田　Ⅱ：水田



设：旱田群体每株平均产量$\mu_1 = 70g$，$\sigma_1 = 10g$

水田群体每株平均产量$\mu_2 = 560g$，$\sigma_2 = 408g$

如果我们在旱田中发现一株产量90g，在有灌溉水田中选到一株产量600g，问比较这二株，那一株好些？（不能直接比较，因为环境不一样必须换算标准群体后才能比较）

$$Z_{旱} = \frac{90-70}{10} = 2.0$$

$$Z_{水} = \frac{600-560}{40} = 1.0$$

标准化后比较旱田植株在原点右边两个阶差；灌溉水田植株在原点右边一个阶差，因此，Z旱＞Z水。

结论：旱田选择的这株比灌溉水田的那株好。在水田上发现的这株高产量，是因为受灌溉条件影响造成的，并不是遗传组成优于那株自旱田选得的植株。

Ⅶ、最小样品数之决定（Determiration of Minimum Family Size）

我们做试验时，应利用最小样品，因为它最经济。如何求得最小样品数，而仍可以达到试验的目的？

例1、燕麦（0ats）

有芒(W)是无芒(w)的显性，

黑壳（S）是白壳（s）的显性。

WwSs × wwss

↓

$\underbrace{\text{1WwSs:1Wwss:1wwSs}}$:1wwss

有芒黑壳$\frac{3}{4}$　：无芒白壳$\frac{1}{4}$。

问在这样的群体中最小样品数应多大，才能选到至少1株是无芒白壳的个体？

具体计算如下：利用二项分布决定，

$P = \frac{3}{4}$（不需要的部分）　$P = \frac{1}{4}$（需要的部分）

（C.I.95%或99%根据需要决定）