

# 常染色体遗传模型

生命是数字的游戏，随着近代生物学的高速发展，数学在生命科学的作用愈发突出，无论是微观方向的发展，还是宏观方向的研究，都必须有精密的数学计算作为推动其前进的不懈动力。用数学方法探索遗传规律是非常有意义的。

## 问题背景

---

- 随着人类的进化，人们为了揭示生命的奥妙，越来越注重遗传学的研究，特别是遗传特征的逐代传播，引起人们更多的注意。无论是人，还是动、植物都会将本身的特征遗传给下一代，这主要是因为后代继承了双亲的基因，形成自己的基因对，基因对确定了后代所表现的特征。

## 问题背景

---

- 在常染色体遗传中，后代是从每个亲体的基因对中各继承一个基因，形成自己的基因对，基因对也称为基因型。如果我们所考虑的遗传特征是由两个基因A和a控制的，那么就有三种基因对，记为AA，Aa，aa。

## 问题背景

- 例如，金鱼草是由两个遗传基因决定它的花的颜色，基因型是 $AA$ 的金鱼草开红花， $Aa$ 型的开粉红色花，而 $aa$ 型的开白花。又如人类眼睛的颜色也是通过常染色体遗传控制的。基因型是 $AA$ 或 $Aa$ 的人，眼睛为棕色，基因型是 $aa$ 的人，眼睛为蓝色。

## 问题背景

- 这里因为Aa和AA都表示了同一外部特征，我们认为基因A支配基因a，也可认为基因a对于A来说是隐性的。当一个亲体的基因型为Aa，而另一个亲体的基因型是aa，那么后代可以从aa型中得到基因a，从Aa型中得到基因A，或得到基因a。这样，后代基因型为Aa或aa的可能性相等

## 问题背景

- 下面给出双亲体基因型的所有可能的结合，使其后代形成每种基因型的概率：

		父体——母体的基因型					
		AA—AA	AA—Aa	AA—aa	Aa—Aa	Aa—aa	aa—aa
后代基因型	AA	1	1/2	0	1/4	0	0
	Aa	0	1/2	1	1/2	1/2	0
	aa	0	0	0	1/4	1/2	1

## 问题描述

- 农场的植物园中某种植物的基因型为 $AA$ ， $Aa$ 和 $aa$ . 农场计划采用 $AA$ 型的植物与每种基因型植物相结合的方案培育植物后代. 那么经过若干年后，这种植物的任一代的三种基因型分布如何？



# 模型假设

- 假设：(i) 设  $a_n$ ,  $b_n$  和  $c_n$  分别表示第  $n$  代植物中，基因型为  $AA$ ,  $Aa$  和  $aa$  的植物占植物总数的百分率。令  $x^{(n)}$  为第  $n$  代植物的基因型分布：

$$x^{(n)} = \begin{pmatrix} a_n \\ b_n \\ c_n \end{pmatrix}$$

显然有  $a_n + b_n + c_n = 1$

$x^{(0)}$  表示植物基因型的初始分布(即培育开始时的分布，可以通过统计数据获得)，



## 模型假设

- (ii)第 $n$ 代的分布只与第 $n-1$ 代的分布有关，关系是通过上面的表确定的，摘录如下：

		父体——母体的基因型		
		AA—AA	AA—Aa	AA—aa
后代基因型	AA	1	1/2	0
	Aa	0	1/2	1
	aa	0	0	0

## 模型建立和求解

- 建模： 根据假设(ii)，先考虑第 $n$ 代中的AA型。由于第 $n-1$ 代的AA型与AA型结合，后代全部是AA型；第 $n-1$ 代的Aa型与AA型结合，后代是AA型的可能性为 $1/2$ ，而第 $n-1$ 代的aa型与AA型结合，后代不可能是AA型。故 $a_n$ 与 $a_{n-1}$ ， $b_{n-1}$ ， $c_{n-1}$ 的关系式由全概率公式可得

$$a_n = 1a_{n-1} + 1/2b_{n-1} + 0c_{n-1}$$

## 模型建立和求解

- 类似可得 $b_n$ ， $c_n$ 与 $a_{n-1}$ ， $b_{n-1}$ ， $c_{n-1}$ 的关系式，综合得下式

$$a_n = 1a_{n-1} + 1/2b_{n-1} + 0c_{n-1}$$

$$b_n = 0a_{n-1} + 1/2b_{n-1} + 1c_{n-1}$$

$$c_n = 0a_{n-1} + 0b_{n-1} + 0c_{n-1}$$

$$\begin{cases} a_n = a_{n-1} + \frac{1}{2}b_{n-1} \\ b_n = \frac{1}{2}b_{n-1} + c_{n-1} \\ c_n = 0 \end{cases}$$

# 模型建立和求解

$$x^{(n)} = \begin{pmatrix} a_n \\ b_n \\ c_n \end{pmatrix} = \begin{pmatrix} 1 & 1/2 & 0 \\ 0 & 1/2 & 1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} a_{n-1} \\ b_{n-1} \\ c_{n-1} \end{pmatrix} = Mx^{(n-1)}$$

$$x^{(n)} = Mx^{(n-1)} \Rightarrow x^{(n)} = M^n x^{(0)}$$

$$M = \begin{pmatrix} 1 & 1/2 & 0 \\ 0 & 1/2 & 1 \\ 0 & 0 & 0 \end{pmatrix} \Rightarrow M^n = ?$$

$$x^{(n)} = M^n x^{(0)}, \quad M = \begin{pmatrix} 1 & 1/2 & 0 \\ 0 & 1/2 & 1 \\ 0 & 0 & 0 \end{pmatrix} \Rightarrow M^n = ?$$

---

$$\lambda_1 = 1, \lambda_2 = 1/2, \lambda_3 = 0$$

$$p_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, p_2 = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, p_3 = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$$

$$P = (p_1, p_2, p_3) = \begin{pmatrix} 1 & -1 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix} \Rightarrow P^{-1} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}$$

# 模型建立和求解

$$M^n = P \Lambda^n P^{-1} = \begin{pmatrix} 1 & -1 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & & \\ & \frac{1}{2} & \\ & & 0 \end{pmatrix}^n \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}$$

$$M^n = \begin{pmatrix} 1 & 1 - \frac{1}{2^n} & 1 - \frac{1}{2^{n-1}} \\ 0 & \frac{1}{2^n} & \frac{1}{2^{n-1}} \\ 0 & 0 & 0 \end{pmatrix} \quad \begin{cases} a_n = 1 - \frac{1}{2^n} b_0 - \frac{1}{2^{n-1}} c_0 \\ b_n = \frac{1}{2^n} b_0 + \frac{1}{2^{n-1}} c_0 \\ c_n = 0 \end{cases}$$

# 模型讨论

若干年后的分布趋势

$$\begin{cases} a_n = 1 - \frac{1}{2^n} b_0 - \frac{1}{2^{n-1}} c_0 \rightarrow 1 \\ b_n = \frac{1}{2^n} b_0 + \frac{1}{2^{n-1}} c_0 \rightarrow 0 \\ c_n = 0 \rightarrow 0 \end{cases}$$

$$\begin{cases} a_n = a_{n-1} - \frac{1}{2} b_{n-1} \\ b_n = \frac{1}{2} b_{n-1} + c_{n-1} \\ c_n = 0 \end{cases}$$



## 模型讨论

- 若在上述问题中，不选用基因型AA植物与每一植物结合，而是将具有相同基因型植物相结合，那么后代所具有三种基因型的概率由下表给出：

		父体——母体的基因型		
		AA—AA	Aa—Aa	aa—aa
后 代 基 因 型	AA	1	1/4	0
	Aa	0	1/2	0
	aa	0	1/4	1

# 模型推广

- 现在世界上已经发现的遗传病将近**4000**种。在一般情况下，遗传疾病是与特殊的种族、部落及群体有关。例如，遗传病库利氏贫血症的患者以居住在地中海沿岸为多，镰状网性贫血症一般流行在黑人中，家族黑蒙性白痴症则流行在东欧犹太人中间。患者经常未到成年就痛苦地死去，而他们的父母则是疾病的病源。

## 模型推广

- 假若我们能识别这些疾病的隐性患者，并且规定隐性患者不能结合(因为两个隐性病患者结合，他们的后代就可能成为显性患者)，那么未来的儿童，虽然有可能是隐性患者，但绝不会出现显性特征，不会受到疾病的折磨。现在，我们考虑在控制结合的情况下，如何确定后代中隐性患者的概率。

# 模型推广

- 假设
- (i) 常染色体遗传的正常基因记为A，不正常基因记为a，并以AA，Aa，aa分别表示正常人，隐性患者，显性患者的基因型
- (ii) 设 $a_n$ ， $b_n$ 分别表示第n代中基因型为AA，Aa 占总人数的百分数，记

$$x^{(n)} = \begin{pmatrix} a_n \\ b_n \end{pmatrix}$$

## 模型推广

- (iii)为使每个儿童至少有一个正常的父亲或母亲，因此正常人或隐性患者必须与正常人结合，其后代的基因型概率由表给出：

		父体——母体的基因型	
		AA—AA	AA—Aa
后代基因型	AA	1	1/4
	Aa	0	1/2
	aa	0	1/4

# 模型推广

---

- 建模：由假设(iii)，从第 $n-1$ 代到第 $n$ 代基因型分布的变化取决于方程.....

# 模型推广

---

- 模型讨论：研究在随机结合的情况下，隐性患者的变化是很有意思的，但随机结合导致了非线性化问题，讨论将会非常复杂，这里就不多述了。