关于加性—显性遗传模型数据处理的计算方法研究

袁 志 发

（西北农学院基础课部）

本文对加性——显性模型的尺度检验及遗传力的估计的计算法进行了探讨，为电子计算机进行数据处理提供了一些普遍而有效的方法。

一、世代均数分量及变异分量

设基因频率为$\frac{1}{2}$，大亲为P_1，小亲为P_2，中亲值为[m]，离中平均效应为[d]F_1代离中亲平均效应为[h]。这些效应为累加性的，且对均数、方差、协方差等统计量的贡献是彼此独立的，这样的模型叫加性——显性模型。经计算各世代均数分量如表1。表1中，B_1为$P_1\times F_1$（回交方式之一），B_2为$P_2\times F_1$（回交方式之二），S_3为F_2中同胞交配的子代，S_4为F_3中同胞交配的子代。

表1、 各世代均数的分量

世代	平均表型			基因型频率		
（群体）	[m]	[d]	[h]	AA	Aa	aa
P_1	1	1	0	1	0	0
P_2	1	-1	0	0	0	1
F_1	1	0	1	0	1	0
F_2	1	0	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$
F_3	1	0	$\frac{1}{4}$	$\frac{3}{8}$	$\frac{1}{4}$	$\frac{3}{8}$
⋮	⋮	⋮	⋮	⋮	⋮	⋮
F_n	1	0	$(\frac{1}{2})^{n-1}$	$\frac{2^{n-1}-1}{2^n}$	$(\frac{1}{2})^{n-1}$	$\frac{2^{n-1}-1}{2^n}$
B_1	1	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	0
B_2	1	$-\frac{1}{2}$	$\frac{1}{2}$	0	$\frac{1}{2}$	$\frac{1}{2}$
S_3	1	0	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$
S_4	1	0	$\frac{3}{8}$	$\frac{5}{16}$	$\frac{3}{8}$	$\frac{5}{16}$

本文写作过程中，宋哲民同志参加讨论，李正德教授给予指导，特此致谢。

对于加性显性模型的表型方差可分解为加性分量D，显性分量H，环境分量σ_e^2及表示显性方向的F分量四部分。经计算，各世代变异分量如表2。

表2、 各世代方差的分量

世代（群体）	D	H	F	σ_e^2
P_1	0	0	0	1
P_2	0	0	0	1
F_1	0	0	0	1
F_2	$\frac{1}{2}$	$\frac{1}{4}$	0	1
F_3	$\frac{3}{4}$	$\frac{3}{16}$	0	1
⋮	⋮	⋮	⋮	⋮
F_n	$1-(\frac{1}{2})^{n-1}$	$(\frac{1}{2})^{n-1}-(\frac{1}{2})^{2n-2}$	0	1
B_1	$\frac{1}{4}$	$\frac{1}{4}$	$-\frac{1}{2}$	1
B_2	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{2}$	1
S_3	$\frac{1}{2}$	$\frac{1}{4}$	0	1
S_4	$\frac{5}{8}$	$\frac{15}{64}$	0	1

二、单尺度检验与均数分量估计的计算法

单尺度检验法是由P_1、P_2及F_1来直接估计[m]、[d]和[h]的，其公式为：

$$\begin{cases}[m]=\frac{1}{2}(\overline{P}_1+\overline{P}_2)\\ [d]=\frac{1}{2}(\overline{P}_1-\overline{P}_2)\\ [h]=\overline{F}_1-\frac{1}{2}(\overline{P}_1+\overline{P}_2)\end{cases}\tag{1}$$

由式（1）进行估计的标准差为：

$$\begin{cases}S_m=\sqrt{V(m)}=\frac{1}{2}\sqrt{V(\overline{P}_1)+V(\overline{P}_2)}\\ S_d=\sqrt{V(d)}=S_m\\ S_h=\sqrt{V(h)}=\sqrt{V(\overline{F}_1)+V(m)}\end{cases}\tag{2}$$

用单尺度来检验模型是通过P_1，P_2，F_1来检验其它各世代而实现的。这是因为从表1可看出有如下期望关系：

$$\begin{cases}A=\overline{B}_1-\frac{1}{2}\overline{P}_1-\frac{1}{2}\overline{F}_1=0,\\ B=\overline{B}_2-\frac{1}{2}\overline{P}_2-\frac{1}{2}\overline{F}_1=0,\\ C=\overline{F}_2-\frac{1}{4}\overline{P}_1-\frac{1}{4}\overline{P}_2-\frac{1}{2}\overline{F}_1=0,\\ D=\overline{F}_3-\frac{3}{8}\overline{P}_1-\frac{3}{8}\overline{P}_2-\frac{1}{4}\overline{F}_1=0,\\ \cdots\cdots\cdots\cdots\end{cases}$$

$$
\begin{aligned}
\sigma_A{}^2 &= \sigma_3{}^2 + \tfrac{1}{4}\sigma_1{}^2 + \tfrac{1}{4}\sigma_5{}^2 \\
\sigma_B{}^2 &= \sigma_4{}^2 + \tfrac{1}{4}\sigma_2{}^2 + \tfrac{1}{4}\sigma_5{}^2 \\
\sigma_C{}^2 &= \sigma_6{}^2 + \tfrac{1}{16}\sigma_1{}^2 + \tfrac{1}{16}\sigma_2{}^2 + \tfrac{1}{4}\sigma_5{}^2 \\
\sigma_D{}^2 &= \sigma_7{}^2 + \frac{9}{64}\sigma_1{}^2 + \frac{9}{64}\sigma_2{}^2 + \frac{1}{16}\sigma_5{}^2 \\
&\cdots\cdots
\end{aligned}
\qquad (3)
$$

其中$\sigma_1{}^2$，$\sigma_2{}^2$，$\sigma_3{}^2$，$\sigma_4{}^2$，$\sigma_5{}^2$，$\sigma_6{}^2$，$\sigma_7{}^2$分别为$\overline{P}_1$，$\overline{P}_2$，$\overline{B}_1$，$\overline{B}_2$，$\overline{F}_1$，$\overline{F}_2$，$\overline{F}_3$的方差。由（3）式知，若所研究的材料是适合加性——显性模型的，则有：

$$
\begin{cases}
F_A = \dfrac{A^2}{\sigma_A{}^2} \sim F\,[\,1,\ n(P_1)+n(B_1)+n(F_1)-3\,] \\
F_B = \dfrac{B^2}{\sigma_B{}^2} \sim F\,[\,1,\ n(P_2)+n(B_2)+n(F_1)-3\,] \\
F_C = \dfrac{C^2}{\sigma_C{}^2} \sim F\,[\,1,\ n(P_1)+n(P_2)+n(F_1)+n(F_2)-4\,] \\
F_D = \dfrac{D^2}{\sigma_D{}^2} \sim F\,[\,1,\ n(P_1)+n(P_2)+n(F_1)+n(F_3)-4\,] \\
\cdots\cdots
\end{cases}
\qquad (4)
$$

单尺度检验一般只进行B_1，B_2，F_2三个世代，更高世当然也是可以的。经过式（4）检验，结果均不显著，表明研究对象是可以用加性——显性模型来进行遗传研究的。如果出现一个或一个以上的尺度与零差异显著，则表明所用模型不当，或有非等位基因的互作，或对基因作用分析不当，未进行尺度转换。

上述数据分析可通过矩阵运算来实现。为此，定义如下矩阵：

单尺度遗传模型矩阵：

$$
G_1 = \begin{pmatrix}
-\tfrac{1}{4} & -\tfrac{1}{4} & -\tfrac{1}{2} & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -\tfrac{1}{2} & -\tfrac{1}{2} & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & -\tfrac{1}{2} & -\tfrac{1}{2} & 1
\end{pmatrix}
$$

单尺度世代方差巨阵：

$$
V_1 = \begin{pmatrix}
V(P_1) & & & & & & & & & \\
& V(P_2) & & & & & & & \bigcirc & \\
& & V(F_1) & & & & & & & \\
& & & V(F_2) & & & & & & \\
& & & & V(P_1) & & & & & \\
& & & & & V(F_1) & & & & \\
& & & & & & V(B_1) & & & \\
& \bigcirc & & & & & & V(P_2) & & \\
& & & & & & & & V(F_1) & \\
& & & & & & & & & V(B_2)
\end{pmatrix}
$$

单尺度世代样本容量矩阵：

$$N_1 = \begin{bmatrix} n(P_1) & & & & & & & & & \\ & n(P_2) & & & & & & & \bigcirc & \\ & & n(F_1) & & & & & & & \\ & & & n(F_2) & & & & & & \\ & & & & n(P_1) & & & & & \\ & & & & & n(F_1) & & & & \\ & & & & & & n(B_1) & & & \\ & & & & & & & n(P_2) & & \\ & \bigcirc & & & & & & & n(F_1) & \\ & & & & & & & & & n(B_2) \end{bmatrix}$$

单尺度世代均数矩阵：

$$\overline{X}_1 = \begin{bmatrix} \overline{P}_1 & \overline{P}_2 & \overline{F}_1 & \overline{F}_2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \overline{P}_1 & \overline{F}_1 & \overline{B}_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \overline{P}_2 & \overline{F}_1 & \overline{B}_2 \end{bmatrix}$$

单尺度系数矩阵：

$$Q = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{2} & -\frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -\frac{1}{2} & -\frac{1}{2} & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

单尺度均数分量矩阵：

$$\theta_1 = \begin{bmatrix} [m] & 0 & 0 \\ [d] & 0 & 0 \\ [h] & 0 & 0 \end{bmatrix}$$

均数分量方差矩阵：

$$\Sigma_{\theta 1} = \begin{bmatrix} V(m) & \times & \times \\ \times & V(d) & \times \\ \times & \times & V(h) \end{bmatrix}$$

尺度矩阵：

$$Z = \begin{bmatrix} C & 0 & 0 \\ 0 & A & 0 \\ 0 & 0 & B \end{bmatrix}$$

尺度方差矩阵：

$$\Sigma_Z = \begin{bmatrix} \sigma_C^2 & 0 & 0 \\ 0 & \sigma_A^2 & 0 \\ 0 & 0 & \sigma_B^2 \end{bmatrix}$$

F检验的F值矩阵：

$$F_0=\begin{pmatrix} F_C & 0 & 0 \\ 0 & F_A & 0 \\ 0 & 0 & F_B \end{pmatrix}$$

由上述矩阵可得：

$$\begin{cases} \text{均数分量估计：} & \theta_1=Q\overline{X}_1' \\ \text{均数分量方差估计：} & \Sigma_{\theta 1}=Q(V_1N^{-1})Q' \\ \text{尺度估计：} & Z=G_1\overline{X}_1' \\ \text{尺度方差估计：} & \Sigma_Z=G_1(V_1N_1^{-1})G_1' \\ \text{F值估计：} & F_0=(G_1\overline{X}_1')(G_1\overline{X}_1')[G_1(V_1N_1^{-1})G_1']^{-1} \end{cases} \qquad (5)$$

如果检验更高世代，只须根据表1在上述有关矩阵上加一行即可。

三、联合尺度检验与均数分量估计的计算法

单尺度测验，虽适用于各种组合的群体，但不能有效地把各个尺度合而为一，对式（3）中各期望关系进行一次综合的检验。为此Cavalli（1952）提出了联合尺度检验法。他运用了最小二乘法原理从各世代的均数中来估计均数各分量及各世代均数的理论值，然后利用χ^2来检验其适合度。均数各分量的估计归结为解如下方程。

（其中：　$I_i=\dfrac{n(i)}{V(i)}$，i为P_1，P_2，F_1，F_2，B_1，B_2）：

$$\begin{pmatrix} I_1+I_2+I_3+I_4+I_5+I_6 & I_1-I_2+\frac{1}{2}I_5-\frac{1}{2}I_6 & I_3+\frac{1}{2}I_4+\frac{1}{2}I_5+\frac{1}{2}I^6 \\ I_1-I_3+\frac{1}{2}I_5-\frac{1}{2}I_6 & I_1+I_2+\frac{1}{4}I_5+\frac{1}{4}I_6 & \frac{1}{4}I_5-\frac{1}{4}I_6 \\ I_3+\frac{1}{2}I_4+\frac{1}{2}I_5+\frac{1}{2}I_6 & \frac{1}{4}I_5-\frac{1}{4}I_6 & I_3+\frac{1}{4}I_4+\frac{1}{4}I_5+\frac{1}{4}I_6 \end{pmatrix}\begin{pmatrix} [m] \\ [d] \\ [h] \end{pmatrix}$$

$$=\begin{pmatrix} I_1\overline{P}+I_2\overline{P}_2+I_3\overline{F}_1+I_4\overline{F}_2+I_5\overline{B}_1+I_6\overline{B}_2 \\ I_1\overline{P}_1-I_2\overline{P}_2+\frac{1}{2}I_5\overline{B}_1-\frac{1}{2}I_6\overline{B}_2 \\ I_3\overline{F}_1+\frac{1}{2}I_4\overline{F}_2+\frac{1}{2}I_5\overline{B}_1+\frac{1}{2}I_6\overline{B}_2 \end{pmatrix} \qquad (6)$$

若令：

世代均数分量遗传模型矩阵为：

$$G_2=\begin{pmatrix}1 & 1 & 0\\ 1 & -1 & 0\\ 1 & 0 & 1\\ 1 & 0 & \frac{1}{2}\\ 1 & \frac{1}{2} & \frac{1}{2}\\ 1 & -\frac{1}{2} & \frac{1}{2}\end{pmatrix}$$

各世代样本容量矩阵为：

$$N_2=\begin{pmatrix}n(P_1) & & & & & \\ & n(P_2) & & & \bigcirc & \\ & & n(F_1) & & & \\ & & & n(F_2) & & \\ & \bigcirc & & & n(B_1) & \\ & & & & & n(B_2)\end{pmatrix}$$

各世代均数矩阵为：$\overline{X}_2=\begin{pmatrix}\overline{P}_1\\ \overline{P}_2\\ \overline{F}_1\\ \overline{F}_2\\ \overline{B}_1\\ \overline{F}_2\end{pmatrix}$

各世代方差矩阵为：

$$V_2=\begin{pmatrix}V(P_1) & & & & & \\ & V(P_2) & & & \bigcirc & \\ & & V(F_1) & & & \\ & & & V(F_2) & & \\ & \bigcirc & & & V(B_1) & \\ & & & & & VB_2)\end{pmatrix}$$

均数分量矩阵为：　$\theta_2=\begin{pmatrix}[m]\\ [d]\\ [h]\end{pmatrix}$

均数分量的协方差矩阵为：

$$\Sigma_{\theta 2}=\begin{pmatrix}D(m) & Cov(m, d) & Cov(m, h)\\ Cov(d, m) & D(d) & Cov(d, h)\\ Cov(h, m) & Cov(h, d) & D(h)\end{pmatrix}$$

则方程（6）可为：

$$[G_2'(N_2V_2^{-1})G_2]\theta_2=G_2'(N_2V_2^{-1})\overline{X}_2 \qquad (7)$$

故有如下各估计：

$$\begin{cases}\text{均数分量估计：} & \theta_2=[G_2'(N_2V_2^{-1})G_2]^{-1}G_2'(N_2V_2^{-1})\overline{X}_2\\ \text{均数分量的协方差估计：} & \Sigma_{\theta_2}=[G_2'(N_2V_2^{-1})G_2]^{-1}\\ \text{各世代均数估计：} & \hat{\overline{X}}_2=G_2\theta_2\\ \chi^2\text{值为：} & \chi^2=(\overline{X}_2-G_2\theta_2)'(N_2V_2^{-1})(\overline{X}_2-G_2\theta_2)\\ & \quad=(\overline{X}_2-\hat{\overline{X}}_2)'(N_2V_2^{-1})(\overline{X}_2-\hat{\overline{X}})\end{cases} \qquad (8)$$

（自由度为世代数减3）

若增加世代，只须分别在G_2，V_2，N_2，$\overline{X}_2$上各加一行即可。

四、方差分量估计的计算法

知道了P_1，P_2，F_1F_2，B_1，B_2等六种世代的表型方差，根据表2里很容易计算出方差各分量的，但是不能估计出各分量的标准误。今令

方差分量矩阵为：$V_\theta=\begin{pmatrix}D\\H\\F\\\sigma_e^2\end{pmatrix}$

各世代表型方差矩阵为：$V_3=\begin{pmatrix}V(P_1)\\V(P_2)\\V(F_1)\\V(F_2)\\V(B_1)\\V(B_2)\end{pmatrix}$

方差分量模型矩阵为：

$$U=\begin{pmatrix}0&0&0&4&-2&-2\\-1&-1&-2&-4&4&4\\0&0&0&0&-1&1\\\frac{1}{4}&\frac{1}{4}&\frac{1}{2}&0&0&0\end{pmatrix}$$

各世代遗传方差模型矩阵为：$G_g=\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{4} & 0 & 0 \\ \frac{1}{4} & \frac{1}{4} & -\frac{1}{2} & 0 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} & 0 \end{pmatrix}$

世代遗传方差矩阵为：$V_g=\begin{pmatrix} V(P_1g) \\ V(P_2g) \\ V(F_1g) \\ V(F_2g) \\ V(B_1g) \\ V(B_2g) \end{pmatrix}$

则有：

$$\begin{cases} \text{方差分量估计为：} V_\theta = U V_3 \\ \text{世代遗传方差估计为：} V_g = G_g(UV_3) \end{cases} \quad (9)$$

由式（9）可算出：

$$\begin{cases} \text{广义遗传力} \quad h_B^2=\dfrac{V(F_2g)}{V(F_2)} \\ \text{狭义遗传力} \quad h_N^2=\dfrac{\frac{1}{2}D}{V(F_2)} \\ \text{显性质} \quad \sqrt{H/D} \quad \text{及} \quad \dfrac{F}{\sqrt{DH}} \end{cases} \quad (10)$$

五、举　　例

D.S.V irk 于1979年所作黄花烟草遗传试验结果及分析如下：（见下面表3）

1. 单尺度检验

$$\theta_1=\begin{pmatrix} 107.3750 & 0 & 0 \\ 8.9250 & 0 & 0 \\ 10.3000 & 0 & 0 \end{pmatrix} \quad \Sigma_{\theta 1}=\begin{pmatrix} 0.6321 & -0.1154 & -0.6321 \\ -0.1154 & 0.6321 & 0.1154 \\ -0.6321 & 0.1154 & 1.6020 \end{pmatrix}$$

$$Z=\begin{pmatrix} -0.7469 & 0 & 0 \\ 0 & -0.9875 & \\ 0 & 0 & 1.0985 \end{pmatrix}$$

$$\Sigma_Z=\begin{pmatrix} 0.8921 & 0 & 0 \\ 0 & 0.9896 & 0 \\ 0 & 0 & 1.2298 \end{pmatrix}$$

$$F_0=\begin{pmatrix} 0.6253 & 0 & 0 \\ 0 & 0.9854 & 0 \\ 0 & 0 & 0.9812 \end{pmatrix}$$

表3　黄花烟草西纯育品种22×73的六个群体资料（株高，单位：cm）

世　代	n	$\bar{X}$	V
P_1（22）	20	116.3000	20.6684
P_2（73）	20	98.4500	29.0500
F_1	60	117.6750	57.4260
F_2	160	111.7781	77.6533
B_1	120	116.0000	59.5288
B_2	120	109.1610	66.1747

计算结果表明，F检验不显著（$F<1$），其资料可用加性——显性模型分析。均数各分量的估计为：$m=197.3750\pm0.7950$，$d=8.9250\pm0.7950$，$h=10.3000\pm1.2657$。分离世代F_2，B_1，B_2的尺度估计为：$C=-0.7469\pm0.9445$，$A=-0.9875\pm0.9948$，$B=1.0985\pm1.1090$。

2．联合尺度检验

$$\theta_2=\begin{pmatrix} 107.3220 \\ 8.1997 \\ 10.0587 \end{pmatrix},$$

$$\Sigma_{\theta_2}=\begin{pmatrix} 0.4567853 & -0.0613140 & -0.7194160 \\ -0.0613140 & 0.4002201 & 0.0799914 \\ -0.7194160 & 0.0799914 & 1.540951 \end{pmatrix},$$

$$\hat{X}_2 = \begin{pmatrix} 115.5217 \\ 99.1223 \\ 117.3798 \\ 112.3514 \\ 116.4512 \\ 108.2515 \end{pmatrix} \qquad \chi^2 = 3.41 \quad (\text{自由度为}3)$$

计算结果表明，χ^2检验不显著，说明资料可用加性——显性模型来分析。均数分量估计为：[m]=107.3220±0.6759，[d]=8.1997±0.6326，[h]=10.0587±1.2412。

3. 方差分量估计

$$V_1 = \begin{pmatrix} 59.2062 \\ 27.6304 \\ 6.6459 \\ 41.1426 \end{pmatrix}, \qquad V_2 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 36.5107 \\ 20.0477 \\ 23.3706 \end{pmatrix},$$

$$h_B{}^2 = \frac{36.5107}{77.6533} = 47\%,$$

$$h_N{}^2 = \frac{\frac{1}{2} \times 59.2062}{77.6533} = 38.12\%, \qquad \sqrt{H/D} = 0.6831,$$

$$\frac{F}{\sqrt{DH}} = 0.1643。$$

参 考 文 献

1. 马育华：1981 数量遗传学及其应用 南京农学院大豆研究室。

2. K·马瑟，J.L.金克斯：生统遗传学导论 1981 农业出版社。