

育种实用统计学

(2018. 8, 中国农业大学出版社出版)

勘误

亲爱的读者，您好！

感谢您购买《育种实用统计学》（2018. 8, 中国农业大学出版社出版），由于水平有限，书中 p50-52 页第四章遗传率章节内容有错误，请参阅以下内容，抱歉。

李奉令

2024. 03. 04

三、遗传力（Heritability）

（一）遗传力定义

遗传力这个概念，很多教科书都表述其为数量遗传学最重要的概念，可见其重要性。但是植物育种中，一会出现“遗传力”一会又表述为“遗传率”，反映了对这一概念内涵把握的模糊性，目前大有用“遗传率”代替“遗传力”的迹象。实际育种工作中，在单因素品种比较试验，计算考察试验的广义遗传力，本身就反应了对遗传力的理解和育种实际应用中存在问题。BLUP MME 巧妙引入了性状遗传力，作为分析计算的重要参数，对遗传力的准确理解，直接跟 BLUP 分析应用挂钩。因此，准确理解和运用遗传力，对植物育种尤其重要。

广义遗传力指数量性状基因型方差占表现型方差的比率。

$$h_B^2 = \frac{V_G}{V_P} = \frac{V_A + V_D + V_I}{V_P}$$

狭义遗传力指数量性状加性遗传方差占表现型方差的比率。

$$h_N^2 = \frac{V_A}{V_P}$$

注意上述公式是方差之比。以 V_P 来说，是表型的方差而不是表型的观测值。遗传力要放到群体角度去考虑。依据一个群体的均值和群体内每个个体的表型值，可以得到表型方差，其描述的是这个群体的该性状的总变异。但是只有一个群体是获得不了狭义遗传力的。因为无法分离得到 V_A 。

亲属间由于拥有相同的基因而表现相似。《植物数量遗传学》介绍的世代对比法，亲子回归法等，和《家畜育种学》（张沅）介绍的亲子资料估计遗传力，同胞资料估计遗传力，都是利用亲代和子代数据来计算性状的遗传力。遗传力作为重要的遗传参数，其估计方法，通常采用约束最大似然法 REML（Restricted Maximum Likelihood）。狭义遗传力是育种值对表型值的决定系数。从实际估计来看，遗传力是育种值对表型值的回归系数。

数量性状的遗传力具有两个特异性：一是性状特异性，不同的性状有不同的遗传力。二是群体特异性，任何数量性状的遗传力都没有固定的理论值，而是在特定的群体所获得的表型值基础上，通过特定的统计方法估计出的。性状遗传力虽然在不同群体间各有差异，但仍具有相对的恒定性。一般使用其他群体的遗传力估计值，同样可以取得较好的效果。

简单理解，性状遗传力就是性状从亲代向子代传递的能力，因此，遗传力（Heritability）这个叫法没有问题，比遗传率更符合概念内涵，不能仅从计算公式强调其比率含义。遗传力是对性状而言的，因此，不应该考虑某群体的遗传力，某个体的遗传力。遗传力用更通俗一点的话语来描述就是，以玉米为例，只要是玉米，其产量性状的遗传力大概就是多少数值，其衡量的是这个作物的性状，不是某个品种，也不是某次试验。但是可以通过某次试验来估计这个性状的遗传力。

动物育种中一般将低于 0.1 的数量性状，称为低遗传力性状，0.1-0.35 中等遗传力性状，高于 0.35 的性状称为高遗传力性状。这里提到的遗传力都是指狭义遗传力。数量性状的遗传力不高。动物如此，那植物的数量性状遗传力呢，也应该遵循同样的遗传规律。

（二）经验育种（差值选择育种法）中的遗传力

传统的经验育种本质上是差值选择育种，其实使用的就是遗传力，且为狭义遗传力。亲代所表现的优良性状，在后代希望延续，希望提高，或者把两个材料单独具有的优良性状用杂交的方法，组合到后代 1 个个体中去，实现优良性状的聚合叠加，遵循的是狭义遗传力。亲属间由于拥有相同的基因而表现相似。亲代能传递给子代的是基因，不是基因型，等位基因间的显性互作，非等位基因间的上位性互作跟基因型有关。基因的加性效应才能遗传给子代，互作效应，上位性效应都不能，子代跟亲代表型的相似，主要来自基因的加性效应，即育种值，以方差的形式表示就是加性遗传方差 V_A ，因此在简单估计遗传力的时候，会采用亲子回归的计算方法，

$$h_N^2 = 2b_{op}, h_N^2 = \frac{V_A}{V_P}$$

h_N^2 是狭义遗传力， b_{op} 是子代对亲代的表型回归系数。

狭义遗传力指数量性状加性遗传方差 V_A 占表现型方差 V_P 的比例。

表型回归系数通俗的理解，就是回归系数越大，子代和亲代就越像，系数越小就越不像。经验育种者依据亲代的表型在后代群体中寻找目标植株，就是人工的回归搜寻。狭义遗传力大的性状，子代跟亲代更像，就更容易发现目标植株。

亲代和子代间，会有人为选择，选择改变的是群体基因频率，进而改变了后代基因型及表型。一个群体中某个基因的存在，以效应或者方差的形式被计算度量，或者以表型差异的方式呈现在育种者面前被育种者在田间捕获，前者就属于“计算机育种”，后者就属于经验育种，其实本质没有区别，遵循的也是同样的

原理。如何选择因人而异，同时也因育种目标和育种需要而异，有好的材料也不是人人都能利用好，也不是人人都能选出好品种，选择技术手段也是有高下的。

“计算机育种”和经验育种，是目前的两种育种现状。经验育种，就要花费大量的心血和精力时间，待在田里，思考观察性状的细微表现差异和性状间的表型相关关系（遗传相关），积累提炼成个人的育种经验和选择依据，只有这样才有希望育出品种。

经验育种（差值选择法）的差值指的是亲子代间的基因型值差异，并不是育种值当然就不是基因的加性效应差值，这是差值选择法的理论和方法的局限性，这个要引起注意。

（三）BLUP 中的遗传力

使用 BLUP，就是要找出群体中基因加性效应（育种值）最大的个体。基因加性效应的 BLUP（最佳线性无偏预测）分析，**最佳**：是指育种值估计的误差方差（PEV）最小， $\hat{u}$ （估计值）和 a （真值）之间的相关最大，就是用误差方差（PEV）最小的策略，使该回归最接近真值。**无偏**：是指估计值的数学期望等于真值，以育种值为例说明， $\hat{u}$ （估计值）就是 a （真值）。从传统用系谱的 BLUP，到分子育种使用的 GBLUP，再到 ssGBLUP，都能看到遗传力的存在，且在计算分析中，最重要的遗传参数。

下面是 3 种 BLUP 分析的 MME 方程组：

A、传统 BLUP

$$\begin{bmatrix} X'R^{-1}X & X'R^{-1}Z \\ Z'R^{-1}X & Z'R^{-1}Z + A^{-1} \end{bmatrix} \begin{bmatrix} \hat{b} \\ \hat{u} \end{bmatrix} = \begin{bmatrix} X'R^{-1}y \\ Z'R^{-1}y \end{bmatrix}$$

B、分子育种 GBLUP

$$\begin{bmatrix} X'R^{-1}X & X'R^{-1}Z \\ Z'R^{-1}X & Z'R^{-1}Z + G^{-1} \end{bmatrix} \begin{bmatrix} \hat{b} \\ \hat{u} \end{bmatrix} = \begin{bmatrix} X'R^{-1}y \\ Z'R^{-1}y \end{bmatrix}$$

C、分子育种 ssGBLUP（一步法）

$$\begin{bmatrix} X'R^{-1}X & X'R^{-1}Z \\ Z'R^{-1}X & Z'R^{-1}Z + H^{-1} \end{bmatrix} \begin{bmatrix} \hat{b} \\ \hat{u} \end{bmatrix} = \begin{bmatrix} X'R^{-1}y \\ Z'R^{-1}y \end{bmatrix}$$

遗传力在哪里呢？

以传统 BLUP 为例，

$$\begin{bmatrix} \mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{Z} \\ \mathbf{Z}'\mathbf{X} & \mathbf{Z}'\mathbf{Z} + \mathbf{A}^{-1}k \end{bmatrix} \begin{bmatrix} \hat{\mathbf{b}} \\ \hat{\mathbf{u}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{y} \\ \mathbf{Z}'\mathbf{y} \end{bmatrix}, \quad k = \frac{\sigma_e^2}{\sigma_a^2}$$

$$k = \frac{\sigma_e^2}{\sigma_a^2} = \frac{\sigma_y^2 - \sigma_a^2}{\sigma_a^2} = \frac{1 - \sigma_a^2 / \sigma_y^2}{\sigma_a^2 / \sigma_y^2} = \frac{1 - h^2}{h^2}$$

h^2 ，就是狭义遗传力，3 种分析方法都需要输入遗传力参数。

在进行 BLUP 分析时，模型需要用户提供所分析性状的狭义遗传力。

育种的选择是在性状遗传力的限定下进行的，所谓限定，就是一种作物的某个性状的遗传力，是由这个作物本身决定的。经验育种需要逐步甄别这个尺寸，而 BLUP 分析，是在基于合适模型的基础上，分离去掉表型中多种环境效应（比如，年份效应，地点效应），用材料间的亲缘关系（传统 BLUP 靠系谱确定，GBLUP 靠分子测序结果确定更准确的亲缘关系），依据遗传力参数，建立表型数据和育种材料育种值间的回归分析，找到基因加性效应最大的育种材料。

（四）组内相关法估算广义遗传力

《植物数量遗传学》中介绍了这一计算广义遗传力的方法，因此在此做介绍。依据方差分析表，可以直接估算单因素随机完全区组的广义遗传力。

13								
14	变异来源	DF	SS	MS	F	概率	F0.05	F0.01
15	区组间	2	27.561	13.780	8.398*	0.995	3.740	6.510
16	材料间	7	34.080	4.869	2.967*	0.960	2.760	4.280
17	误差	14	22.972	1.641				
18	总变异	23	84.613					
19								

下面是试验的期望均方（EMS）构成。

品种	8-1	SS_r	MS_r	$\sigma_e^2 + 3\sigma_v^2$
误差	(3-1)(8-1)	SS_e	MS_e	σ_e^2

$$V_E = \sigma_e^2 = 1.641 \quad V_G = \sigma_v^2 = (4.869 - 1.641)/3 = 1.076$$

$$V_P = \sigma_e^2 + \sigma_G^2 = 1.641 + 1.076 = 2.717$$

$$\text{广义遗传力: } h_b^2 = V_G / V_P = 1.076 / 2.717 = 0.396$$

单因素随机完全区组试验适合采用组内相关法，但仅能估算广义遗传力。
该方法在植物育种的品种比较试验中，不建议采用。

（五）亲子回归法估算狭义遗传力

以下两个表格中的公式推导，参照马育华先生的《植物育种的量遗传学基础》，在此仅列出结果。亲属间的协方差和遗传方差间存在数量关系。

不同亲属关系的协方差所代表的遗传方差成分

亲属间关系	协方差	近交系数 F=0	近交系数 F=1
亲子代：子代平均与一个亲代	COV_{OP}	$\frac{1}{2}V_A$	$\frac{1}{2}V_A$
亲子代：子代平均与中亲平均	COV_{OP}	$\frac{1}{2}V_A$	$\frac{1}{2}V_A$
半同胞间	COV_1 或 COV_2	$\frac{1}{4}V_A$	$\frac{1}{2}V_A$
全同胞间	COV	$\frac{1}{2}V_A + \frac{1}{4}V_D$	$V_A + V_D$

遗传力 h_N^2 从不同亲属关系所得回归系数与组内相关估计

亲属关系	近交系数 F=0		近交系数 F=1	
	回归 b 或 t	遗传力 h_N^2	回归 b 或 t	遗传力 h_N^2
子代与一个亲本	$b = \frac{1}{2} \frac{V_A}{V_P}$	$h_N^2 = 2b$	$b = \frac{1}{2} \frac{V_A}{V_P}$	$h_N^2 = 2b$
子代与中亲值	$b = \frac{V_A}{V_P}$	$h_N^2 = b$	$b = \frac{V_A}{V_P}$	$h_N^2 = b$
半同胞	$t = \frac{1}{4} \frac{V_A}{V_P}$	$h_N^2 = 4t$	$t = \frac{1}{2} \frac{V_A}{V_P}$	$h_N^2 = 2t$
全同胞	t $= \frac{\frac{1}{2}V_A + \frac{1}{4}V_D}{V_P}$	$h_N^2 < 2t$	$t = \frac{V_A + V_D}{V_P}$	$h_N^2 < t$

自花授粉作物（或者异花授粉作物人工控制自交，如玉米套袋自交，F1 套袋，F2 套袋）可利用 F2 单株和 F3 系统均值的协方差，估算狭义遗传率。假定上位性效应=0，适用于表格中的子代与中亲值回归。

$$h_N^2 = b_{F_2 \overline{F_3}}$$

$$h_N^2 = \frac{cov_{F_2 \overline{F_3}}}{V_{F_2}} = \frac{\frac{1}{2}V_a + \frac{1}{8}V_d}{\frac{1}{2}V_a + \frac{1}{4}V_d + V_e}$$

$$\text{标准误, } \hat{\sigma}(h_N^2) = \sqrt{\frac{SS_{F_3} - b_{F_3/F_2}^2 SS_{F_2}}{(n-2)SS_{F_2}}}$$

由于分子中多了 $\frac{1}{8}V_d$ ，这种计算方法， h_N^2 估计值偏大。

一般来说，F2 个体的方差 (V_{F_2}) 总比 F3 家系平均的方差 ($V_{\overline{F_3}}$) 大，而且 F2 与 F3 家系的方差是在不同年份估计的，因而，受环境的影响很大。由于这些原因，采用回归系数方法所估计的遗传方差往往出现大于表型方差，得到的 h_N^2 大于 1，这是不合理的估计结果。而且因年份的环境条件不同， V_{F_2} ，总方差量变化很大，因此不同年份所估计的遗传力值不能相互比较。可以用以下方法进行校正：

$$h_N^2 = \frac{cov_{F_2 \overline{F_3}}}{V_{F_2}} \times \sqrt{\frac{V_{F_2}}{V_{\overline{F_3}}}} = \frac{cov_{F_2 \overline{F_3}}}{\sqrt{V_{F_2} V_{\overline{F_3}}}} = r_{\overline{F_3} F_2}$$

h_N^2 为 F2 单株和 F3 系统均值间的相关系数，故该法亦称亲子相关法。

遗传力估计的其它方法建议参照有关专业书籍，从动物育种的实践看，不需要花费过多精力，知晓遗传力的原理和用途可能更重要。

（六）实际育种中遗传力育种应用思考

1、单因素随机区组试验有必要计算产量性状的遗传力吗？没必要。一年多点试验同样没有必要。狭义遗传力的计算需要亲代子代数据，这两种试验都不具

备这个数据条件，只能依据方差分析得出环境方差（ V_e ），计算出广义遗传力，没有任何育种价值。

2、不知道某个性状的遗传力，也查不到权威资料，怎么办？方法很简单：比较子代性状跟亲代的相似性，越像遗传力越高。

①狭义遗传力高，表型与基因型的相关程度高，采用混合选择较好。

②狭义遗传力低，表型不能很好的反应上下代间的相似程度，重点应放在系谱法和后裔鉴定上，必要时也可采用间接选择法或综合选择法。

③有性繁殖作物，狭义遗传力决定着亲子间相似程度，决定着由表型预测育种值的可靠程度，但是无性繁殖作物，是由广义遗传力决定亲属间相似程度。

④性状的狭义遗传力高，说明性状的遗传受环境影响小，而且说明亲子代间的相似程度高，由表型判断基因型的命中率高，选择容易见效，反之，选择不易见效。

⑤性状的狭义遗传力高，可在早代选择，且种植规模可较小，低时，宜在晚代选择，且种植规模宜较大。

⑥产量为低遗传力性状。