Heredity **60** (1988) 347-353

Received 13 July 1987

A new look at Jinks–Hayman's method for the estimation of genetical components in diallel crosses

Pedro E. Ferreira

Catie, Turrialba 7170, Costa Rica.

A new interpretation of the assumption which states that genes are independently distributed across the parents leads to the construction of new estimates of the genetical components D, H_1, H_2 and F. Simulation experiments are used to compare the new estimates with the classical ones and with an improved (unbiased) version of those estimates. The new estimates show a better behaviour than their competitors in terms of bias and mean square deviation, especially for small diallels.

INTRODUCTION

Jinks and Hayman's method of analysis of diallel crosses has been widely used to investigate the genetical structure which explains a quantitative trait. This methodology, presented in a series of papers (Jinks and Hayman, 1953; Hayman, 1954; and Jinks, 1954), has been reviewed by several authors (Gilbert, 1958; Nassar, 1965; Feyt, 1976; Baker, 1978) who criticised the validity of its assumptions and analysed the consequences of their failure. These assumptions are: (a) diploid segregation, (b) no difference between reciprocal crosses, (c) no non-allelic interaction, (d) no multiple allelism, (e) homozygous parents and (f) genes independently distributed across the parents.

Let genotypes II, ii and Ii, at the ith locus, have genotypic values $c+d_i$, $c-d_i$ and $c+h_i$ ($d_i \geqq 0$), and be represented by a variable θ_i which takes values 1, -1 and 0, respectively. Let us further assume that the diallel cross contains n parents and that k genes control the character of interest. Then it is easy to see that to obtain the formulae given by Hayman (1954), section 2.3, the following equalities

$$\sum_{r=1}^{n} (\theta_{ri} - \theta_{.i})(\theta_{rj} - \theta_{.j}) = 0, \quad (i \neq j), \qquad (1.1)$$

are required, where r is an index for parents and $\theta_{.i}$ and $\theta_{.j}$ are averages across parents. This strong hypothesis, which is related to assumption (f), has been criticised by several authors such as Gilbert (1958) and Feyt (1976). In particular Feyt observed that having n loci, we need at least 2^n parents to obtain a non-trivial theoretical situation in which (1.1) is satisfied.

Sokol and Baker (1977) and Baker (1978) state that (f) will be assured if parents are produced by random mating followed by nonselective inbreeding. This makes the expected value of the left hand side of (1.1) to be zero and therefore, the equality holds for large n, *i.e.*, it is asymptotically true. In practice, however, a small number of parents is used in diallel experiments and asymptotic arguments are not reliable. Lack of validity of (1.1) may induce large biases in the estimates, especially for small diallels, and is responsible for a great deal of criticism which arises from simulation studies (Nassar, 1965; Feyt, 1976).

Working from a population point of view, Hayman (1960 and 1963) derived formulae for the expectations of the classical estimates of D, F, H_1 and H_2, providing unbiased estimates of the corresponding population parameters. An important point to be noticed is that these formulae were obtained assuming that the *expectation* of (1.1) is zero, thus overcoming the difficulties mentioned above.

In this paper, a new theoretical framework for fixed parent diallel crosses which assumes a prior distribution on the parameters that determine the parental genotypes is discussed. The expectations derived by Hayman (1963) are used as a first step

to derive new estimates of the genetical components. A simulation study comparing the new estimates with those of Jinks and Hayman and with Hayman's (1963) estimates is presented. On each simulated diallel the observed frequencies of positive and negative homozygotes at each locus are computed and used to obtain the genetical components. Mean biases and mean square deviations, are computed. The results show that the new estimates improve both the classical estimates and the unbiased ones given by Hayman.

ESTIMATING THE USUAL GENETICAL COMPONENTS

The statistical model

In the following it is assumed that the parents in the diallel constitute a fixed set (and not a random sample from a population), having genotypes represented by vector parameters

$$\theta_s = (\theta_{s1}, \ldots, \theta_{sk}), \qquad s = 1, \ldots, n.$$

Our prior knowledge about θ is expressed by the following assumptions,

(A) $P(\theta_{si} = 1) = U_i$, $\quad P(\theta_{si} = -1) = V_i$, $\quad U_i + V_i = 1$,

(B) θ_{si} are independently distributed.

Further, the genotypic contribution to the phenotype of the rth parent is

$$\eta_{rr} = \sum_i d_i \theta_{ri},$$

and analogously, for the $\theta_r \times \theta_s$ progeny we have

$$\eta_{rs} = \tfrac{1}{2} \sum_i \{d_i(\theta_{ri} + \theta_{si}) + h_i(1 - \theta_{ri}\theta_{si})\}. \quad (2.1)$$

Therefore, adding the environmental contribution e_{rs} to (2.1), the statistical model may be written as

$$y_{rs} = \eta_{rs} + e_{rs}. \quad (2.2)$$

Given the genotypes θ_s, the errors e_{rs} are independently and identically distributed with zero means and equal variances.

Notice that the assumption (1.1) has been changed and is now interpreted as meaning stochastic independence in the prior distribution of the parameters.

Genetical components

Let the genotype of the rth parent be denoted by $\theta_r = (\theta_{r1}, \ldots, \theta_{rk})$ and let the frequency of positive homozygotes at the ith locus be u_i. Then, the arithmetic mean $\theta_{.i}$ will be $w_i = u_i - v_i$ where $v_i = 1 - u_i$.

We are interested in estimating the usual genetical components,

$$D = \sum_{i=1}^{k} d_i^2(1 - w_i^2)$$

$$F = 2 \sum_{i=1}^{k} d_i h_i w_i (1 - w_i^2)$$

$$H_1 = \sum_{i=1}^{k} h_i^2 (1 - w_i^2)$$

$$H_2 = \sum_{i=1}^{k} h_i^2 (1 - w_i^2)^2$$

$$h = \sum_{i=1}^{k} h_i (1 - w_i^2), \quad (2.3)$$

and some functions of them, namely ADD, NGR, FP and RDR, defined as follows,

$ADD = (H_1/D)^{1/2}$, average degree of dominance,

$NGR = h^2/H_2$, number of groups of genes which control the character and exhibit dominance,

$FP = H_2/4H_1$, frequency product $(u_i v_i)$ at loci exhibiting dominance,

$RDR = \dfrac{(4DH_1)^{1/2} + F}{(4DH_1)^{1/2} - F}$, ratio of dominant to recessive alleles in all the parents. (2.4)

We should distinguish these genetical components from their analogue obtained from (2.3) replacing w_i by $W_i = U_i - V_i$. These new parameters, which do not depend on θ, will be denoted D^p, H_1^p, H_2^p, F^p and h^p. They correspond to the *population* genetical components which are considered under a random model approach.

The estimating procedure

The estimation of the parameters in (2.3) and (2.4) starts by considering the equations given by Hayman (1963),

$$\underline{E}(D') = D^p$$

$$\underline{E}(F') = F^p - 2F^p/n$$

$$\underline{E}(H_1') = H_1^p + (h^{p2} - H_2^p)/n$$

$$\underline{E}(H_2') = 4H_1^p \frac{n-1}{n^2} + H_2^p \frac{n^2-6n+6}{n^2} + \frac{1}{n} h^{p2}. \quad (2.5)$$

Notice that a correction was made in the last equation. We also need the equations

$$\underline{E}(\delta^2) = \underline{E}(m_{L1} - m_{L0})^2 = H_1^p \frac{(n-1)^2}{n^3} - H_2^p \frac{(n-1)(2n-3)}{2n^3} + \tfrac{1}{4} h^{p2} \left(\frac{n-1}{n}\right)^2, \quad (2.6)$$

and

$$\underline{E}(\delta) = \underline{E}(m_{L1} - m_{L0}) = \tfrac{1}{2} h^p \frac{n-1}{n}, \quad (2.7)$$

where m_{L0} and m_{L1} are the means of the parents and of the whole diallel, as usual.

In the above expectations the environmental component e_{rs} in (2.2) was not considered. This will be done in the Environmental Variation Section.

Further, taking expectations in (2.3) we obtain

$$\underline{E}(D) = D^p \left(1 - \frac{1}{n}\right)$$

$$\underline{E}(F) = F^p \left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right)$$

$$\underline{E}(H_1) = H_1^p \left(1 - \frac{1}{n}\right)$$

$$\underline{E}(H_2) = H_2^p \left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right)\left(1 - \frac{3}{n}\right) + 4H_1^p \frac{1}{n}\left(1 - \frac{1}{n}\right)^2 \quad (2.8)$$

and

$$\underline{E}(h) = h^p \left(1 - \frac{1}{n}\right). \quad (2.9)$$

From (2.5), (2.6) and (2.8) we obtain the following estimates

$$\hat{D} = V_{0L0} \frac{n-1}{n}$$

$$\hat{F} = (2V_{0L0} - 4W_{0L01}) \frac{n-1}{n}$$

$$\hat{H}_1 = \left(V_{0L0} - 4W_{0L01} + 4V_{1L1} \frac{n-1}{n-2}\right) \frac{n-1}{n} - 4\left(V_{0L1} \frac{n-1}{n} + (m_{L1} - m_{L0})^2\right) \frac{1}{n-2}$$

$$\hat{H}_2 = 4(V_{1L1} - V_{0L1}) \frac{(n-1)^2}{n(n-2)} - 4(m_{L1} - m_{L0})^2 \frac{1}{n-2}. \quad (2.10)$$

Further, from (2.7) and (2.9) we obtain

$$\hat{h} = 2(m_{L1} - m_{L0}). \quad (2.11)$$

Replacement of (2.10) and (2.11) into (2.4) provides the remaining estimates.

SIMULATION STUDIES

Simulating according to the prior distribution

Complete diallels having n parents ($n = 6$, 10 and 15) were simulated using (2.1) and (2.2). The number of simulated loci was 10 and the additive and dominance parameters (d_i, h_i) were taken all equal to 0·5 as in Nassar (1965). A Fortran program was run on a IBM 4331 computer and a random number generator was used to simulate independent -1 and $+1$ variables θ_{ri} ($r = 1, \ldots, n$; $i = 1, \ldots, 10$). Three values were considered for the probabilities U_i of θ_{ri} being $+1$, namely 0·25, 0·50 and 0·75, for the case of $n = 6$. For larger diallels, *i.e.*, $n = 10$ and 15, the probabilities U_i were fixed at 0·50. One thousand diallels were simulated for each of the cases studied. Assumptions (a)-(f), where (f) is given by the expectation of (1.1), are obviously satisfied.

From each simulated diallel, three sets of estimates were computed: (a) the classical ones (Hayman (1954)), henceforth referred to as "JH estimates", (b) Hayman's (1963) unbiased estimates obtained from (2.5) and denoted "HU-estimates", and (c) the new ones given by (2.10) and (2.11), and denoted "N estimates".

The performance of the three sets of estimates was measured through the mean bias and the mean square deviation from the usual parameters.

To obtain these measures, on each simulated diallel, the frequencies u_i of positive homozygotes at the ith locus ($i = 1, \ldots, 10$) were computed on the n simulated parents. Then the usual genetical parameters (2.3) and (2.4) were obtained. The estimates were computed on each diallel and their mean deviation (MD), or bias, and mean square deviation (MSD) from the usual parameters were

obtained as follows:

$$\mathrm{MD} = \frac{1}{N}\sum_{j=1}^{N}(\hat{\nu}_j - \nu_j)$$

$$\mathrm{MSD} = \frac{1}{N}\sum_{j=1}^{N}(\hat{\nu}_j - \nu_j)^2. \qquad (3.1)$$

where ν_j represents a parameter associated to the jth diallel and $\hat{\nu}_j$ its estimate. $N = 1000$ is the number of simulated diallels. The means of the usual parameters simulated were also computed.

The results are given in tables 1, 2 and 3.

The statistics MD and MSD are estimates of $\underline{E}(\hat{D} - D)$ and $\underline{E}(\hat{D} - D)^2$, respectively, *i.e.*, the average bias and the average quadratic loss (or Bayesian risk).

Simulating a random sample of diallels

Since the relative performances of the three sets of estimates observed through the simulation study reported in the previous section depend on the prior distribution, they are of little interest for those who reject Bayesian arguments.

To overcome this difficulty a random sample of one thousand diallels from the population of all possible 6×6 and 10×10 diallels was simulated. The number of loci and the values assumed by (d_i, h_i) were the same as in the previous section.

Mean biases and mean square deviations were computed following (3.1) and are reported in table 4. These figures do not depend on the prior distribution and are useful to compare the estimates from a frequentist or classical point of view.

Comparing the performance of the estimates

The results in tables 2 to 4 show a better performance of the N estimates as compared with JH and HU estimates. In particular, considerable improvement is obtained when estimating D, H_1, H_2, ADD and NGR. Some improvement is also observed in

Table 1 Mean values of usual genetical components according to the number of parents (n) and the probability (U) of positive homozygote at a single locus

n	U	D	H_1	H_2	F	h	ADD	NGR	FP	RDR
	0·25	1·556	1·556	1·254	−1·041	3·113	1·000	7·771	0·200	0·501
6	0·50	2·081	2·081	1·845	−0·005	4·162	1·001	9·419	0·221	1·019
	0·75	1·557	1·557	1·259	1·027	3·115	1·001	7·759	0·201	2·089
10	0·50	2·249	2·249	2·067	0·014	4·497	1·001	9·802	0·230	1·020
15	0·50	2·331	2·331	2·195	0·011	4·663	1·001	9·920	0·235	1·015

Table 2 Mean biases of three sets of estimates from the usual genetical components according to the number of parents (n) and the probability of positive homozygote at a single locus (U)

n	U	Estimate*	D	H_1	H_2	F	ADD	NGR	FP**	RDR
6	0·25	JH	0·38	2·42	2·35	−0·27	0·64	−5·01	25·78	0·14
		HU	0·38	1·19	−0·54	−0·92	0·36	23·08	−136.73	−0·05
		N	0·05	−0·03	0·02	−0·05	0·12	0·79	0·73	0·03
6	0·50	JH	0·49	4·15	4·11	−0·03	0·83	−6·45	17·48	0·00
		HU	0·49	2·02	0·04	−0·04	0·48	1·34	−107·25	0·06
		N	0·06	0·00	0·00	−0·02	0·15	1·49	−0·59	0·04
6	0·75	JH	0·27	2·40	2·34	0·20	0·67	−4·97	25·83	−0·45
		HU	0·27	1·16	−0·52	0·81	0·38	17·49	−134·85	0·87
		N	−0·03	0·00	0·00	0·00	0·13	1·03	0·82	0·02
10	0·50	JH	0·31	2·53	2·51	0·02	0·48	−5·32	9·90	0·00
		HU	0·31	0·81	0·81	0·03	0·19	−2·51	4·81	0·05
		N	0·05	0·03	0·03	0·02	0·08	0·30	3·18	0·02
15	0·50	JH	0·22	1·67	1·66	0·01	0·32	−4·24	5·48	0·00
		HU	0·22	0·41	0·48	0·01	0·09	−1·64	8·83	0·03
		N	0·05	0·01	0·01	0·01	0·04	0·15	−0·12	0·01

* JH = Jinks Hayman, HU = Hayman's unbiased, N = new estimates.
** Mean square deviations $\times 10^3$.

Table 3 Mean square deviations of three sets of estimates from the usual genetical components according to the number of parents (n) and the probability of positive homozygote at a single locus (U)

n	U	Estimate*	D	H_1	H_2	F	ADD	NGR	FP**	RDR
6	0·25	JH	1·55	6·78	6·29	0·93	0·80	26·16	0·83	0·04
		HU	1·55	1·96	0·62	2·83	0·39	360·28†	20·94	0·04
		N	0·98	0·32	0·20	0·59	0·18	10·25	0·30	0·03
6	0·50	JH	2·87	18·19	17·74	0·53	1·23	41·98	0·39	0·04
		HU	2·87	5·00	0·47	1·30	0·61	429·78	12·33	0·16
		N	1·83	0·58	0·51	0·37	0·28	23·16	0·21	0·10
6	0·75	JH	1·32	6·73	6·28	0·80	0·77	25·92	0·85	0·37
		HU	1·32	1·92	0·62	2·41	0·37	361·73†	20·48	7·56
		N	0·86	0·33	0·21	0·53	0·17	12·87	0·34	0·30
10	0·50	JH	1·48	6·66	6·57	0·22	0·39	28·58	0·13	0·02
		HU	1·48	0·95	0·96	0·36	0·15	8·10	0·18	0·05
		N	1·12	0·23	0·21	0·18	0·10	4·77	0·07	0·04
15	0·50	JH	0·86	2·91	2·88	0·10	0·17	18·16	0·04	0·01
		HU	0·86	0·28	0·36	0·14	0·06	3·85	0·13	0·02
		N	0·70	0·09	0·10	0·10	0·05	1·99	0·03	0·02

* JH = Jinks Hayman, HU = Hayman's unbiased, N = new estimates.
** Mean square deviations $\times 10^3$.
† Mean square deviations $\times 10^{-3}$.

Table 4 Mean biases and mean square deviations of three sets of estimates from the usual genetical components computed from a sample of one thousand diallels, according to the number of parents (n)

		Mean Biases						
n	Estimate*	D	H_1	H_2	F	ADD	NGR	FP**	RDR
6	JH	0·35	4·17	4·12	−0·04	0·89	−6·44	17·10	−0·01
	HU	0·35	2·04	0·06	−0·05	0·53	3·33	−0·11	0·05
	N	−0·06	0·02	0·02	−0·03	0·18	1·44	−0·77	0·02
10	JH	0·20	2·51	2·48	−0·03	0·52	−5·31	9·53	0·00
	HU	0·20	0·80	0·78	−0·04	0·21	−2·43	3·37	0·02
	N	−0·05	0·02	0·01	−0·03	0·10	0·39	0·58	0·01
		Mean Square Deviations						
n	Estimate*	D	H_1	H_2	F	ADD	NGR	FP**	RDR
6	JH	2·42	18·40	17·94	0·56	1·38	41·86	0·36	0·04
	HU	2·42	5·12	0·51	1·37	0·69	858·30	12·37	0·16
	N	1·60	0·58	0·51	0·38	0·31	27·06	0·22	0·10
10	JH	1·30	6·55	6·41	0·21	0·44	28·37	0·12	0·02
	HU	1·30	0·90	0·89	0·35	0·16	7·73	0·20	0·05
	N	1·02	0·21	0·20	0·17	0·11	4·95	0·08	0·03

* Jinks Hayman's, HU = Hayman's unbiased, N = new estimates.
** Figures $\times 10^3$.

the estimation of F and FP. The relative advantage of using N estimates diminishes as the size of the diallel increases.

JH's estimates show large positive biases in the estimation of ADD, attaining a maximum of 85 per cent of the parameter value when $n = 6$ and $U = 0{\cdot}5$ (table 2). This overestimation of ADD was also reported by other authors (see Feyt, 1976 and Baker, 1978). In addition, large negative biases were observed for the JH's estimates of NGR, attaining 70 per cent of the parameter values, approximately. Since the parameter $NGR = h^2/H_2$

under-evaluates the number of genes (see table 1), negative biases from *NGR* increase the overall bias from the number of genes.

Hayman's (1963) unbiased estimates (from (2.5)) almost always improve the JH estimates in the case of large diallels, but are not reliable in small diallels. In particular they fail in the estimation of *NGR*, *FP* and *RDR*.

The new estimate proposed for *NGR* overestimates the parameter. A correction factor $(n-1)/n$, where n is the number of parents in the diallel, is recommended for small diallels. In the 6×6 case, a reduction of 60 per cent was observed in the mean square deviation when using this correction.

ENVIRONMENTAL VARIATION

We will now consider the contributions of environment to phenotype as given in (2.2). The variance of e_{rs} is denoted with the letter E, as usual.

In this situation the following expectations may be obtained:

$$\underline{E}(V_{0L0}) = D^p + E$$

$$\underline{E}(W_{0L01}) = \frac{1}{2}D^p - \frac{1}{4}\left(\frac{n-2}{n}\right)P^p + \frac{1}{h}E$$

$$\underline{E}(V_{1L1}) = \frac{1}{4}D^p - \frac{1}{4}\left(\frac{n-2}{n}\right)F^p + \frac{1}{4}H_1^p + \frac{1}{4}\frac{1}{n}(h^{p2} - H_2^p) + \frac{1}{2}\left(\frac{n+1}{n}\right)E$$

$$\underline{E}(V_{0L1}) = \frac{1}{4}D^p - \frac{1}{4}\left(\frac{n-2}{n}\right)F^p + \frac{1}{4}\left(\frac{n-2}{n}\right)^2 H_1^p - \frac{1}{4}\frac{(n-2)(n-3)}{2}H_2^p + \frac{1}{2n}E$$

$$\underline{E}(\delta^2) = \underline{E}(m_{L1} - m_{L0})^2 = \frac{1}{4}\left(\frac{n-1}{n}\right)^2 h^{p2} + \frac{(n-1)^2}{n^3}H_1^p - \frac{1}{2}\frac{(n-1)(2n-3)}{n^3}H_2^p + \left(\frac{n-1}{n^2}\right)E. \tag{4.1}$$

Replacing E in (4.1) by its estimate $\hat{E}$, usually obtained from the mean square error in a randomised block experiment, solving the system of equations, and using the correction factors given by (2.8) we obtain:

$$\hat{D} = (V_{0L0} - \hat{E})\frac{n-1}{n}$$

$$\hat{F} = \left(2V_{0L0} - 4W_{0L01} - 2\hat{E}\frac{n-2}{n}\right)\frac{n-1}{n}$$

$$\hat{H}_1 = \left(V_{0L0} - 4W_{0L01} + 4V_{1L1}\frac{n-1}{n-2}\right)\frac{n-1}{n} - 4\left(V_{0L1}\frac{n-1}{n} + (m_{L1} - m_{L0})^2\right)\frac{1}{n-2} - 3\hat{E}\frac{n-1}{n}$$

$$\hat{H}_2 = 4(V_{1L1} - V_{0L1})\frac{(n-1)^2}{n(n-2)} - 4(m_{L1} - m_{L0})^2\frac{1}{n-2} - 2\hat{E}\frac{n^2-1}{n^2}.$$

Further notice that no environmental correction term is needed in (2.11).

The above treatment of environmental variation, similar to Hayman's (1954), is useful for diallel experiments arranged in balanced completely randomised (CR) or randomised complete blocks (CB) designs. In fact, y_{rs} in (2.1) may be written as

$$\text{or} \quad \begin{aligned} y_{rsl} &= x_{rs} + e_{rsl} && \text{(CR)} \\ y_{rsl} &= x_{rs} + \beta_l + e_{rsl} && \text{(CB)} \end{aligned} \tag{4.2}$$

where $l = l, \ldots, L$ stand for replicates or blocks, and x_{rs} and β_k represent the genotypic and block contribution, respectively. Assuming independence among all terms in (4.2) we may compute the expectations in (4.1) quite easily. For example, for V_{0L0} we have:

$$\underline{E}(V_{0L0}) = \underline{E}\left\{\frac{1}{L}\sum_l \frac{l}{n-1}\sum_i (y_{il} - y_{.l})^2\right\} = D + \underline{E}\left(\frac{1}{L}\sum_l \frac{l}{n-1}\sum_i (e_{il} - e_{.l})^2\right) = D + E,$$

where y_{il} denotes y_{ill} in (4.2), as usual.

DISCUSSION

Most breeders perform diallel cross experiments using a highly selected collection of parents. In this case, the fixed, non-random model of Jinks

and Hayman is appropriate if the assumptions listed in the Introduction are met.

One of these assumptions, namely the one which states that genes are independently distributed across the parents, has been most criticised (Gilbert, 1958; Feyt, 1976; and Baker, 1978). In its simplest formulation, this assumption may be written as in (1.1), but having n loci we need at least 2^n parents to obtain a non-trivial theoretical situation in which (1.1) is satisfied (Feyt, 1976).

This difficulty may be overcome if we adopt a statistical Bayesian approach to the problem in relation to the parameters θ which determine the parental genotype. This approach is perfectly compatible with the fixed parent model and enables us to interpret the independence assumption (f) as a requisite in the prior distribution of θ.

The expectations in Hayman (1963) fit perfectly within this perspective and were used as a first step in the derivation of the new estimates. As a second step, equations (2.9) were used to modify the estimates to attain average-unbiasedness, which means, for example, that

$$\underset{\sim}{E}(\hat{H}_2 - H_2) = 0,$$

a property of interest in the fixed model approach. On the other hand, notice that Hayman's (1960, 1963) unbiased estimates are unbiased for the *population* parameters $(D^p, H_1^p, \ldots)$, a property of interest in the random model situation.

The simulation study evaluates the average biases and the average quadratic losses and shows that the new estimates perform better than their competitors from both a Bayesian and frequentist or classical point of view.

Finally, notice that the random sample of diallels considered in the section "Simulating a random sample of diallels" was taken from the whole population of diallels. This means that assumption (f) or (1.1) was *not* required. Now, if we compare the performance of the estimates in tables 2 and 3 (take $U = 0{\cdot}50$) with the results in table 4, we conclude that the estimates were quite robust against departures from this assumption, a crucial point against those who criticise Jinks-Hayman's methodology.

Acknowledgements The author is grateful to a referee and to Professor I. Bos for their helpful comments. Part of this work was written at EMBRAPA/CNPH, Brasilia, and at the Dept of Statistics, Univ. of Waterloo, on a period of research as Visiting Professor.

REFERENCES

BAKER, R. J. 1978. Issues in diallel analysis. *Crop Science, 18,* 533-536.

FEYT, H. 1976. Étude critique de l'analyse des croisements diallèles au moyen de la simulation. *Ann. Amelior. Plantes, 26,* 173-193.

GILBERT, N. E. G. 1958. Diallel cross in plant breeding. *Heredity, 12,* 477-492.

HAYMAN, B. I. 1954. The theory and analysis of diallel crosses. *Genetics, 39,* 789-809.

HAYMAN, B. I. 1960. The theory and analysis of diallel crosses III. *Genetics, 45,* 155-172.

HAYMAN, B. I. 1963. Notes on diallel-cross theory. In Hanson, W. D. and Robinson, H. F. (eds.) *Statistical Genetics and Plant Breeding,* National Academy of Sciences, Washington, pp. 571-578.

JINKS, J. L. 1954. The analysis of continuous variation in a diallel cross of Nicotiana rustica varieties. *Genetics, 39,* 767-788.

JINKS, J. L. AND HAYMAN, B. I. 1953. The analysis of diallel crosses. *Maize Genetics, Coop. Newsletter, 27,* 48-54.

NASSAR, R. F. 1965. Effect of correlated gene distribution due to sampling on the diallel analysis. *Genetics, 52,* 9-20.

SOKOL, M. J. AND BAKER, R. J. 1977. Evaluation of the assumptions required for the genetic interpretation of diallel experiments in self pollinating crops. *Can. J. Plant Sci., 57,* 1185-1191.