

## ON THE BIOMETRICAL ANALYSIS OF QUANTITATIVE GENETICS

W. Y. TAN

*Washington State University, Pullman, WA. USA*

(Received for publication June 1, 1974)

### Abstract

In this paper some distribution theories of the segregating diploid populations from crosses between two inbred lines are derived, assuming that some information concerning the ancestors are available. These results are then used to obtain the expected values of some statistics, thus extending the results given in Hayman (1960) and Mather and Jinks (1971).

### Introduction

In a recent paper, Tan and Chang (1972) derived some distribution theories of the segregating populations from a cross between two inbred lines. As a continuation of that paper, I derive in this paper the distribution theories of the segregation populations when information about their ancestors are available. As in my previous paper, in this paper I shall also restrict myself to sufficiently large diploid populations and assume that there are no genetic-environmental interactions, no selection, no linkage and no epistatic effects among loci, etc. Moreover, it will be assumed further that the errors are independently and identically distributed and that the effects of genes are equal over different loci. It is of interest to note that the formulas for variances and covariances hold also without the assumption of equality of genotypic effects from different loci.

In section 2 some representation formulas for the distribution theories of  $F_k$ ,  $k \geq 2$ , populations are given under the above set of assumptions. These results are then utilized in sections 3 and 4 to derive probability distributions of various statistics and the expectations of certain quantities. For notations and genetic terminologies, I shall follow in essence Mather and Jinks (1971) and Tan and Chang (1972), unless otherwise stated. Thus,  $V_{1F_3}$  is the variance of  $F_3$  line means,  $V_{2F_3}$  the expected value of the  $F_3$  within line variance(\*);

(\*) Mather and Jinks (1971) called it the mean variance of  $F_3$  families. We rewrite the language so that it is more clearly understood. Similarly, we rewrite the descriptions for the  $F_4$  variances in order to make it more understandable.

$V_{1F_4}$  the variance of means of groups of  $F_4$  families descended from one  $F_2$  ancestor,  $V_{2F_4}$  the expected value of  $F_4$  family means within  $F_2$  ancestor groups, and  $V_{3F_4}$  the expected value of the  $F_4$  within family variance, etc.

*The representation formulas for the distribution theories of  $F_k$  ( $k \geq 2$ )*

Let  $Y_j$  denote the random variable for the phenotypic value of the  $F_j$  population and assume that the effect of the plots in the field layout is the same. Then, using results given in Tan and Chang (1972),

$$Y_j = \mu_j - nd + 2dZ_1^{(j)} + (d+h)Z_2^{(j)} + e_j, \quad j=2, 3, 4, \dots \quad (2.1)$$

where  $\mu_j$  is the grand mean of the  $j$ th generation,  $d$  and  $h$  the additive and dominance effects, respectively,  $(Z_1^{(j)}, Z_2^{(j)})$  the random variables for the genetic segregation of  $F_j$  generation so that

$$(Z_1^{(j)}, Z_2^{(j)}) \sim \text{Mult}\left(n; \frac{1}{2}\left(1 - \frac{1}{2^{j-1}}\right), \frac{1}{2^{j-1}}\right),$$

and  $e_j$  the random disturbance for  $Y_j$ , which is assumed to be independently distributed with mean 0 and variance  $\sigma_e^2$ .

Suppose now information about the ancestors are available. Then it is seen that  $Y_3$  may also be represented by

$$Y_3 = \mu_3 - nd + 2dX_1^{(3)} + (d+h)X_2^{(3)} + 2dX_1^{(2)} + e_3, \quad (2.2)$$

where  $\underline{x}_{(2)}' = (X_1^{(2)}, X_2^{(2)}) \sim \text{Mult}(n; \frac{1}{4}, \frac{1}{2})$  denote the random variables for the genetic segregation of  $F_2$  ancestors, and  $\underline{x}_3' = (X_1^{(3)}, X_2^{(3)})$  the random variables for the genetic segregation of  $F_3$  individuals within  $F_2$  ancestors so that the conditional distribution of  $(X_1^{(3)}, X_2^{(3)})$  given  $\underline{x}_{(2)}$  is

$$(X_1^{(3)}, X_2^{(3)}) | \underline{x}_{(2)} \sim \text{Mult}(X_2^{(2)}; \frac{1}{4}, \frac{1}{2}).$$

Let  $h_3(t)$  be the characteristic function of  $e_3$ . Then, since the  $e_3$ 's are independent of the random variables for genetic segregations, the characteristic function of  $Y_3$  using (2.2) is then, with  $i = \sqrt{-1}$ ,

$$\begin{aligned} \phi_3(t) &= \mathbb{E}e^{iY_3t} = e^{i(\mu_3 - nd)t} h_3(t) \mathbb{E}e^{it[2d(X_1^{(3)} + X_2^{(3)}) + (d+h)X_2^{(3)}]} \\ &= e^{i(\mu_3 - nd)t} h_3(t) \mathbb{E}_{\substack{\underline{x}_2 \\ \underline{x}_3 | \underline{x}_2}} \left\{ e^{it[2dX_1^{(3)} + (d+h)X_2^{(3)}]} \right\} e^{it2dX_1^{(2)}} \\ &= e^{i(\mu_3 - nd)t} h_3(t) \mathbb{E}_{\substack{\underline{x}_2}} \left\{ e^{it2dX_1^{(2)}} \left[ \frac{1}{4}e^{it2d} + \frac{1}{2}e^{it(d+h)} + \frac{1}{4} \right]^{X_2^{(2)}} \right\} \\ &= e^{i(\mu_3 - nd)t} h_3(t) \left[ \frac{1}{4}e^{it2d} + \frac{1}{2} \left( \frac{1}{4}e^{it2d} + \frac{1}{2}e^{it(d+h)} + \frac{1}{4} \right) + \frac{1}{4} \right]^n \\ &= e^{i(\mu_3 - nd)t} h_3(t) \left[ \frac{3}{8}e^{it2d} + \frac{1}{4}e^{it(d+h)} + \frac{3}{8} \right]^n, \end{aligned} \quad (2.3)$$

which is also the characteristic function of  $Y_3$  using (2.1).

Hence, the representations (2.1) and (2.2) are equivalent in the sense that they provide the same probability distribution. If  $e_3$  is normally distributed, then the probability density function of  $Y_3$  using (2.1) and (2.2) is given by

$$\begin{aligned} f_3(y) &= \sum_{s_1=0}^n \sum_{s_2=0}^{n-s_1} \binom{n}{s_1, s_2} \left(\frac{1}{2}\right)^{s_2} \left(\frac{1}{4}\right)^{n-s_2} \\ &\quad \cdot \sum_{t_1=0}^{s_2} \sum_{t_2=0}^{s_2-t_1} \binom{s_2}{t_1, t_2} \left(\frac{1}{2}\right)^{t_2} \left(\frac{1}{4}\right)^{s_2-t_2} \frac{1}{\sqrt{2\pi\sigma_3}} \\ &\quad \cdot \exp\left\{-\frac{1}{2\sigma_3^2}[y-\mu_3+nd-2dt_1-(d+h)t_2-2ds_1]^2\right\} \\ &= \sum_{s=0}^n \sum_{t=0}^{n-s} \binom{n}{s, t} \left(\frac{1}{4}\right)^t \left(\frac{3}{8}\right)^{n-t} \frac{1}{\sqrt{2\pi\sigma_3}} \\ &\quad \cdot \exp\left\{-\frac{1}{2\sigma_3^2}[y-\mu_3-(2s+t-n)d-th]^2\right\}. \end{aligned} \quad (2.4)$$

Similarly,  $Y_4$  may also be represented by

$$Y_4 = \mu_4 - nd + 2dX_1^{(4)} + (d+h)X_2^{(4)} + 2dX_1^{(3)} + 2dX_1^{(2)} + e_4, \quad (2.5)$$

where  $\underline{x}_2' = (X_1^{(2)}, X_2^{(2)}) \sim \text{Mult}(n; \frac{1}{4}, \frac{1}{2})$  denote the random variables for the genetic segregation of  $F_2$  ancestors,  $\underline{x}_3' = (X_1^{(3)}, X_2^{(3)}) | \underline{x}_2' \sim \text{Mult}(X_2^{(2)}; \frac{1}{4}, \frac{1}{2})$  the random variables for the genetic segregation of  $F_3$  ancestors given  $F_2$  and  $\underline{x}_4' = (X_1^{(4)}, X_2^{(4)}) | \underline{x}_3' \sim \text{Mult}(X_2^{(3)}; \frac{1}{4}, \frac{1}{2})$  the random variables for the genetic segregation of  $F_4$  individuals given  $F_3$  ancestors.

Again, it is easy to show that (2.1) and (2.5) give the same characteristic function  $\psi_4(t)$  of  $F_4$  phenotypic values as

$$\psi_4(t) = h_4(t) e^{it(\mu_4 - nd)} \left(\frac{7}{16} e^{2itd} + \frac{1}{8} e^{(d+h)it} + \frac{7}{8}\right)^n, \quad (2.6)$$

where  $h_4(t)$  is the characteristic function of  $e_4$ .

Hence, the representations (2.1) and (2.5) are also equivalent. If  $e_4$  is normally distributed, the probability density function for  $Y_4$  is then

$$\begin{aligned} f_4(y) &= \sum_{i_1=0}^n \sum_{j_1=0}^{n-i_1} \binom{n}{i_1, j_1} \left(\frac{1}{2}\right)^{j_1} \left(\frac{1}{4}\right)^{n-j_1} \\ &\quad \cdot \sum_{i_2=0}^{j_1} \sum_{j_2=0}^{j_1-i_2} \binom{j_1}{i_2, j_2} \left(\frac{1}{2}\right)^{j_2} \left(\frac{1}{4}\right)^{j_1-j_2} \\ &\quad \cdot \sum_{i_3=0}^{j_2} \sum_{j_3=0}^{j_2-i_3} \binom{j_2}{i_3, j_3} \left(\frac{1}{2}\right)^{j_3} \left(\frac{1}{4}\right)^{j_2-j_3} \frac{1}{\sqrt{2\pi\sigma_4}} \\ &\quad \cdot \exp\left\{-\frac{1}{2\sigma_4^2}[y-\mu_4+nd-2(i_1+i_2+i_3)d-j_3(d+h)]^2\right\} \\ &= \sum_{s=0}^n \sum_{t=0}^{n-s} \binom{n}{s, t} \left(\frac{1}{4}\right)^t \left(\frac{7}{16}\right)^{n-t} \frac{1}{\sqrt{2\pi\sigma_4}} \\ &\quad \cdot \exp\left\{-\frac{1}{2\sigma_4^2}[y-\mu_4-(2s+t-n)d-th]^2\right\}. \end{aligned} \quad (2.7)$$

In general, for  $k > 2$ ,  $Y_k$  can also be represented by

$$Y_k = \mu_k - nd + 2d \sum_{j=2}^k X_1^{(j)} + (d+h)X_2^{(k)} + e_k, \quad (2.8)$$

where  $(X_1^{(2)}, X_2^{(2)}) \sim \text{Mult}(n; \frac{1}{4}, \frac{1}{2})$ , and

$$(X_1^{(s)}, X_2^{(s)}) | (X_1^{(s-1)}, X_2^{(s-1)}) \sim \text{Mult}\left(X_2^{(s-1)}; \frac{1}{4}, \frac{1}{2}\right), \quad s=3, 4, \dots, k.$$

Again, it is easily seen that (2.1) and (2.8) give the same characteristic function  $\psi_k(t)$  as

$$\begin{aligned} \psi_k(t) = h_k(t) e^{it(\mu_k - nd)} &\left[ \frac{1}{2} \left(1 - \frac{1}{2^{k-1}}\right) e^{2itd} \right. \\ &\left. + \frac{1}{2^{k-1}} e^{(d+h)it} + \frac{1}{2} \left(1 - \frac{1}{2^{k-1}}\right) \right]^n, \end{aligned} \quad (2.9)$$

where  $h_k(t)$  is the c.f. of  $e_k$ . Hence, (2.1) and (2.8) provide the same probability distribution which can be derived by applying the inversion formula on  $\psi_k(t)$ .

*The distribution theories of the  $F_k$ ,  $k \geq 3$ , phenotypic values when information about their ancestors is available*

Using results given in the previous section, we derive in this section the probability distributions for some quantities and use them to obtain the expected values of some second degree statistics. We notice that the results for the means and variances to be given hold also without the assumption of equality of genotypic values from different loci; in the latter case we need only replace  $nh$ ,  $D$  and  $H$  by  $\sum_{j=1}^n h_j$ ,  $D = \sum_{j=1}^n d_j^2$  and  $H = \sum_{j=1}^n h_j^2$ , where  $d_j$  and  $h_j$  are the additive and dominance effects of the  $j$ th locus.

Let  $Y_{k s_1 s_2 \dots s_{k-1}}$  be the observed value of the  $s_{k-1}$ th  $F_k$  individual from the  $s_{k-2}$ th  $F_{k-1}$  ancestor, the  $s_{k-3}$ th  $F_{k-2}$  ancestor, ... and the  $s_1$ th  $F_2$  ancestor. Then, using (2.8),

$$\begin{aligned} Y_{k s_1 s_2 \dots s_{k-1}} = \mu_k - nd + 2d [X_{1 s_1}^{(2)} + X_{1 s_2}^{(3)} + X_{1 s_3}^{(4)} + & \\ \dots + X_{1 s_{k-1}}^{(k)}] & \\ + (d+h)X_{2 s_{k-1}}^{(k)} + e_{k s_1 s_2 \dots s_{k-1}}, & \end{aligned} \quad (3.1)$$

where  $(X_{1 s_1}^{(2)}, X_{2 s_1}^{(2)}) \sim \text{Mult}(n; \frac{1}{4}, \frac{1}{2})$  independently for all  $s_1 = 1, 2, \dots, m_1$  and conditional on  $(X_{1 s_{r-1}}^{(r)}, X_{2 s_{r-1}}^{(r)}), \dots, (X_{1 s_{r-1}}^{(r)}, X_{2 s_{r-1}}^{(r)})$ ,

$$\begin{aligned} (X_{1 s_r}^{(r+1)}, X_{2 s_r}^{(r+1)}) | (X_{1 s_{r-1}}^{(r)}, X_{2 s_{r-1}}^{(r)}) & \\ \sim \text{Mult}\left(X_{2 s_{r-1}}^{(r)}; \frac{1}{4}, \frac{1}{2}\right), & \end{aligned}$$

independently for all  $s_r = 1, 2, \dots, m_r$ ,  $r = 2, 3, \dots, k-1$ ; furthermore, the

$e_{k s_1 s_2 \dots s_{k-1} s}$  are independently and identically distributed with mean 0 and variance  $\sigma_k^2$ , and also independently distributed of the random variables for genetic segregation (i.e. the X's). From (3.1), the characteristic function (c.f. in short) of  $Y_{k s_1 s_2 \dots s_{k-1}}$  is obtained as

$$\begin{aligned}\psi_1(\omega) &= h_k(\omega) e^{i\omega(\mu_k - nd)} \left[ \frac{1}{2} \left(1 - \frac{1}{2^{k-1}}\right) e^{2di\omega} \right. \\ &\quad \left. + \frac{1}{2^{k-1}} e^{(d+h)i\omega} + \frac{1}{2} \left(1 - \frac{1}{2^{k-1}}\right) \right]^n,\end{aligned}$$

where  $h_k(\omega)$  is the c.f. of  $e_{k s_1 s_2 \dots s_{k-1} s}$ . Using  $\psi_1(\omega)$ , the mean and variance of  $Y_{k s_1 s_2 \dots s_{k-1}}$  are given respectively by

$$\mathbb{E} Y_{k s_1 s_2 \dots s_{k-1}} = \mu_k + \frac{1}{2^{k-1}} nh$$

and

$$\text{Var}(Y_{k s_1 s_2 \dots s_{k-1}}) = \left(1 - \frac{1}{2^{k-1}}\right) D + \frac{1}{2^{k-1}} \left(1 - \frac{1}{2^{k-1}}\right) H + \sigma_k^2,$$

where  $D = nd^2$  and  $H = nh^2$ . Put now

$$\begin{aligned}\bar{Y}_{k s_1 s_2 \dots s_{k-2}} &= \frac{1}{m_{k-1}} \sum_{s_{k-1}=1}^{m_{k-1}} Y_{k s_1 s_2 \dots s_{k-2} s_{k-1}}, \\ \bar{Y}_{k s_1 s_2 \dots s_r \dots} &= \frac{1}{m_{r+1}} \sum_{s_{r+1}=1}^{m_{r+1}} \bar{Y}_{k s_1 s_2 \dots s_{r+1} \dots}, \quad r=1, 2, \dots, k-3,\end{aligned}$$

and

$$\bar{Y}_k \dots = \frac{1}{m_1} \sum_{s_1=1}^{m_1} \bar{Y}_{k s_1 \dots}.$$

Then, for  $r=2, 3, \dots, k-1$ , we have for the representation of  $\bar{Y}_{k s_1 s_2 \dots s_{k-r} \dots}$ :

$$\begin{aligned}\bar{Y}_{k s_1 s_2 \dots s_{k-r} \dots} &= \mu_k - nd + 2d [X_{1 s_1}^{(2)} + X_{1 s_2(s_1)}^{(3)} + \dots + X_{1 s_{k-r}(s_1 s_2 \dots s_{k-r-1})}^{(k-r+1)}] \\ &\quad + 2d [\bar{X}_{1 \cdot (s_1 \dots s_{k-r})}^{(k-r+2)} + \bar{X}_{1 \cdot (s_1 \dots s_{k-1})}^{(k-r+3)} + \dots + \bar{X}_{1 \cdot (s_1 s_2 \dots s_{k-r})}^{(k)}] \\ &\quad + (d+h) \bar{X}_{2 \cdot (s_1 s_2 \dots s_{k-r})}^{(k)} + \bar{e}_{k s_1 s_2 \dots s_{k-r} \dots}, \quad (3.2)\end{aligned}$$

where

$$\begin{aligned}\bar{X}_{j \cdot (s_1 s_2 \dots s_{k-r})}^{(a)} &= \frac{1}{m_{k-r+1} m_{k-r+2} \dots m_{a-1}} \sum_{s_{k-r+1}=1}^{m_{k-r+1}} \sum_{s_{k-r+2}=1}^{m_{k-r+2}} \dots \\ &\quad \sum_{s_{a-1}=1}^{m_{a-1}} X_{j s_{a-1}(s_1 s_2 \dots s_{k-r} \dots s_{a-2})}^{(a)},\end{aligned}$$

for all  $k \geq a > k-r+1$  and  $r=2, 3, \dots, k-1$ , and where the  $\bar{e}_{k s_1 s_2 \dots s_{k-r} \dots}$  are similarly defined as the  $\bar{Y}_{k s_1 s_2 \dots s_{k-r} \dots}$ 's.

Using (3.2), the c.f. of  $\bar{Y}_{k s_1 s_2 \dots s_{k-r} \dots}$  can readily be derived. For this purpose, we put  $\Phi_0(\omega) = e^{(d+h)i\omega}$  and

$$\Phi_r(\omega) = \left[ \frac{1}{4} e^{2di(\omega m_{k-r})} + \frac{1}{2} \Phi_{r-1}\left(\frac{\omega}{m_{k-r}}\right) + \frac{1}{4} \right]^{m_{k-r}}, \\ r=1, 2, \dots, (k-2),$$

and define  $\bar{Y}_{k s_1 s_2 \dots s_{k-1}} = \bar{Y}_{k s_1 s_2 \dots s_{k-1}}$  and

$$h_k(\omega) = \left[ h_k\left(\frac{\omega}{m_{k-r+1} m_{k-r+2} \dots m_{k-1}}\right) \right]^{m_{k-r+1} m_{k-r+2} \dots m_{k-1}}, \\ \text{if } r=1. \quad (3.3)$$

Then, it can readily be shown that, for  $r=1, 2, \dots, (k-1)$ , the c.f. of  $\bar{Y}_{k s_1 s_2 \dots s_{k-r}}$  is

$$\begin{aligned} \psi_r(\omega) &= e^{i\omega(\mu_k - nd)} \left[ h_k\left(\frac{\omega}{m_{k-r+1} m_{k-r+2} \dots m_{k-1}}\right) \right]^{m_{k-r+1} m_{k-r+2} \dots m_{k-1}} \\ &\cdot \left\{ \frac{1}{2} \left( 1 - \frac{1}{2^{k-r}} \right) e^{2di\omega} + \frac{1}{2^{k-r}} \Phi_{r-1}(\omega) + \frac{1}{2} \left( 1 - \frac{1}{2^{k-r}} \right) \right\}^n \end{aligned} \quad (3.4)$$

By the inversion formula, the probability distribution of  $\bar{Y}_{k s_1 s_2 \dots s_{k-r}}$  is then

$$\begin{aligned} f_{k(r)}(y) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-iy\omega} \psi_r(\omega) d\omega \\ &= \sum_{s=0}^n \sum_{t=0}^{n-s} \binom{n}{s, t} \left[ \frac{1}{2} \left( 1 - \frac{1}{2^{k-r}} \right) \right]^{n-t} \left[ \frac{1}{2^{k-r}} \right]^t \\ &\cdot \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\omega(\mu_k - nd) + 2di\omega s - iy\omega} \\ &\cdot \left\{ h_k\left(\frac{\omega}{m_{k-r+1} m_{k-r+2} \dots m_{k-1}}\right) \right\}^{m_{k-r+1} m_{k-r+2} \dots m_{k-1}} \\ &\cdot (\Phi_{r-1}(\omega))^t d\omega. \end{aligned}$$

If the  $e_{k s_1 s_2 \dots s_{k-1}}$ 's are normally distributed, then

$$\begin{aligned} f_{k(1)}(y) &= \sum_{s_1=0}^n \sum_{t_1=0}^{n-s} \binom{n}{s_1, t_1} \left[ \frac{1}{2} \left( 1 - \frac{1}{2^{k-1}} \right) \right]^{n-t} \left[ \frac{1}{2^{k-1}} \right]^t \\ &\cdot g(y; \mu_k - nd + 2ds_1 + (d+h)t_1, \sigma_k^2), \\ f_{k(r)}(y) &= \sum_{s_1=0}^n \sum_{t_1=0}^{n-s_1} \sum_{s_2=0}^{t_1 m_{k-1} - t_1 m_{k-1} - s_2} \sum_{t_2=0}^{t_1 m_{k-1} - s_2} \dots \sum_{s_r=0}^{t_{r-1} m_{k-r+1} - t_{r-1} m_{k-r+1} - s_r} \sum_{t_r=0}^{t_{r-1} m_{k-r+1} - s_r} \binom{n}{s_1, t_1} \\ &\cdot \binom{t_1 m_{k-1}}{s_2, t_2} \dots \binom{t_{r-1} m_{k-r+1}}{s_r, t_r} \left[ \frac{1}{2} \left( 1 - \frac{1}{2^{k-r}} \right) \right]^{n-t_1} \left[ \frac{1}{2^{k-r}} \right]^{t_1} \\ &\cdot \left( \frac{1}{4} \right)^{t_1 m_{k-1} - t_2 + t_2 m_{k-2} - t_3 + \dots + t_{r-1} m_{k-r+1} - t_r} \left( \frac{1}{2} \right)^{t_2 + t_3 + \dots + t_r} \\ &\cdot g(y; \mu_k - nd + 2ds_1 + 2d \frac{s_2}{m_{k-1}} + 2d \frac{s_3}{m_{k-1} m_{k-2}} + \\ &\dots + 2d \frac{s_r}{m_{k-1} m_{k-2} \dots m_{k-r+1}} + (d+h) \frac{t_r}{m_{k-1} m_{k-2} \dots m_{k-r+1}}; \\ &\frac{\sigma_k^2}{m_{k-1} m_{k-2} \dots m_{k-r+1}}), \quad k-1 \geq r \geq 2, \quad (3.5) \end{aligned}$$

where  $g(y; \mu, \sigma^2)$  denotes a normal density with mean  $\mu$  and variance  $\sigma^2$ .

For deriving the moments of  $\bar{Y}_{k s_1 s_2 \dots s_{k-r} \dots}$ , one may use (3.4) directly. We now proceed to derive the expectation and the variance of  $\bar{Y}_{k s_1 s_2 \dots s_{k-r} \dots}$  by using (3.4). For this purpose, we notice the following two lemmas.

**LEMMA 1.**  $((d/d\omega)\Phi_r(\omega/c))_{\omega=0} = 1/c((d/d\omega)\Phi_r(\omega))_{\omega=0}$  and  $((d^2/d\omega^2)\Phi_r(\omega/c))_{\omega=0} = ((1/c^2)(d^2/d\omega^2)\Phi_r(\omega))_{\omega=0}$ , where  $c \neq 0$ .

**LEMMA 2.** Putting  $\Delta_r = (1/i)((d/d\omega)\Phi_r(\omega))_{\omega=0}$  and  $L_{r-1} = (1/i^2)((d^2/d\omega^2)\Phi_{r-1}(\omega))_{\omega=0}$ , then  $\Delta_r = d + (h/2^r)$ ,  $r = 1, 2, \dots, k-1$  and

$$\begin{aligned} L_{r-1} &= \left(d + \frac{h}{2^{r-1}}\right)^2 \\ &+ \sum_{a=1}^{r-1} \frac{1}{2^{a-1} m_{k-r+1} m_{k-r+2} \dots m_{k-r+a}} \left(\frac{1}{2} d^2 + \frac{1}{2^{2(r-a)}} h^2\right) \\ &+ \frac{1}{2^{r-2} m_{k-r+1} m_{k-r+2} \dots m_{k-1}} \left[\frac{1}{2} d^2 + \frac{1}{4} h^2\right], \\ r &= 2 \dots k-1. \end{aligned}$$

The proof of Lemma 1 is trivial. To prove Lemma 2, we notice that, by taking derivatives and applying Lemma 1, we have:

$$\Delta_r = \frac{d}{2} + \frac{1}{2} \Delta_{r-1}$$

and

$$L_{r-1} = \left(d + \frac{h}{2^{r-1}}\right)^2 + \frac{1}{m_{k-r+1}} \left(d^2 - \left(d + \frac{h}{2^{r-1}}\right)^2\right) + \frac{1}{2m_{k-r+1}} L_{r-2}.$$

Solving the above difference equations and noting that  $\Delta_0 = (1/i)[(d/d\omega)\Phi_0(\omega)]_{\omega=0} = (d+h)$  and  $L_0 = (1/i^2)[(d^2/d\omega^2)\Phi_0(\omega)]_{\omega=0} = (d+h)^2$  as  $\Phi_0(\omega) = e^{(d+h)i\omega}$ , Lemma 2 is established. Now,

$$\begin{aligned} E\bar{Y}_{k s_1 s_2 \dots s_{k-r} \dots} &= \frac{1}{i} \left( \frac{d}{d\omega} \log \Psi_r(\omega) \right)_{\omega=0} \\ &= \mu_k - nd + n \left[ \left(1 - \frac{1}{2^{k-r}}\right)d + \frac{1}{2^{k-r}} \Delta_{r-1} \right] \\ &= \mu_k - nd + n \left[ d - \frac{1}{2^{k-r}} d + \frac{1}{2^{k-r}} d + \frac{1}{2^{k-1}} h \right] \\ &= \mu_k + \frac{1}{2^{k-1}} nh \end{aligned} \tag{3.6}$$

by Lemma 2. For  $r = 2, 3, \dots, k-1$ ,

$$\begin{aligned} \text{Var} [\bar{Y}_{k s_1 s_2 \dots s_{k-r} \dots}] &= \frac{1}{i^2} \left( \frac{d^2}{d\omega^2} \log \Psi_r(\omega) \right)_{\omega=0} \\ &= \frac{1}{m_{k-r+1} m_{k-r+2} \dots m_{k-1}} \sigma_k^2 \\ &+ n \left[ 2 \left(1 - \frac{1}{2^{k-r}}\right) d^2 + \frac{1}{2^{k-r}} L_{r-1} - \left(d + \frac{h}{2^{k-1}}\right)^2 \right] \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{m_{k-r+1} m_{k-r+2} \dots m_{k-1}} \sigma_k^2 \\
&\quad + n \left[ 2 \left( 1 - \frac{1}{2^{k-r}} \right) d^2 + \frac{1}{2^{k-r}} \left( d + \frac{h}{2^{r-1}} \right)^2 \right. \\
&\quad \left. + \sum_{a=1}^{r-2} \frac{1}{2^{k-r+a-1} m_{k-r+1} m_{k-r+2} \dots m_{k-r+a}} \left( \frac{1}{2} d^2 + \frac{1}{2^{2(r-a)}} h^2 \right) \right. \\
&\quad \left. + \frac{1}{2^{k-2} m_{k-r+1} m_{k-r+2} \dots m_{k-1}} \left( \frac{1}{2} d^2 + \frac{1}{4} h^2 \right) - \left( d + \frac{h}{2^{k-1}} \right)^2 \right]
\end{aligned}$$

by Lemma 2.

After simplification, we have then:

$$\begin{aligned}
&\text{Var} [\bar{Y}_{k s_1 s_2 \dots s_{k-r}}] \\
&= \frac{1}{m_{k-r+1} m_{k-r+2} \dots m_{k-1}} \left[ \sigma_k^2 + \frac{1}{2^{k-1}} D + \frac{1}{2^k} H \right] \\
&\quad + \sum_{a=1}^{r-2} \frac{1}{2^{k-r+a-1} m_{k-r+1} m_{k-r+2} \dots m_{k-r+a}} \left[ \frac{1}{2} D + \frac{1}{2^{2(r-a)}} H \right] \\
&\quad + \left[ \left( 1 - \frac{1}{2^{k-r}} \right) D + \frac{1}{2^{k-2}} \left( \frac{1}{2^r} - \frac{1}{2^k} \right) H \right], \quad r=2, 3, \dots, (k-1).
\end{aligned} \tag{3.7}$$

Notice that the second term is taken to be zero if  $r \leq 2$ . In (3.7), if  $k=3$  and  $r=2$ , then

$$\text{Var} [\bar{Y}_{3 s_1}] = \frac{1}{m_2} \left( \sigma_3^2 + \frac{1}{4} D + \frac{1}{8} H \right) + \left[ \frac{1}{2} D + \frac{1}{16} H \right] = V_{1 F_3}$$

as defined in Hayman (1960) and Mather and Jinks (1971); if  $k=4$  and  $r=3$ , then

$$\begin{aligned}
\text{Var} [\bar{Y}_{4 s_1}] &= \frac{1}{m_2 m_3} \left[ \sigma_4^2 + \frac{1}{8} D + \frac{1}{16} H \right] \\
&\quad + \frac{1}{m_2} \left( \frac{1}{4} D + \frac{1}{32} H \right) + \left( \frac{1}{2} D + \frac{1}{64} H \right) = V_{1 F_4}
\end{aligned}$$

as defined in Hayman (1960) and Mather and Jinks (1971).

Suppose now  $s_1, s_2, \dots, s_{k-r-1}$  are given; then, using (3.1) and (3.2), we obtain the conditional c.f. of  $\bar{Y}_{k s_1 s_2 \dots s_{k-r}}$  given  $s_1, s_2, \dots, s_{k-r-1}$  as,

$$\begin{aligned}
\psi_{k(1)}^{(*)}(\omega) &= h_k(\omega) e^{i\omega \Gamma \mu_k - nd + 2d(x_{1s_1}^{(2)} + x_{1s_2(s_1)}^{(3)} + \dots + x_{1s_{k-2}(s_1s_2\dots s_{k-3})}^{(k-1)})} \\
&\quad \cdot \left\{ \frac{1}{4} e^{2di\omega} + \frac{1}{2} e^{(d+h)i\omega} + \frac{1}{4} \right\}^{x_{2s_{k-2}(s_1s_2\dots s_{k-3})}^{(k-1)}}
\end{aligned}$$

and

$$\begin{aligned}
\psi_{k(r)}^{(*)}(\omega) &= \left\{ h_k \left( \frac{\omega}{m_{k-r+1} m_{k-r+2} \dots m_{k-1}} \right) \right\}^{m_{k-r+1} m_{k-r+2} \dots m_{k-1}} \\
&\quad \cdot e^{i\omega \Gamma \mu_k - nd + 2d(x_{1s_1}^{(2)} + x_{1s_2(s_1)}^{(3)} + \dots + x_{1s_{k-r-1}(s_1s_2\dots s_{k-r-2})}^{(k-r)})} \\
&\quad \cdot \left\{ \frac{1}{4} e^{2di\omega} + \frac{1}{2} \phi_{r-1}(\omega) + \frac{1}{4} \right\}^{x_{2s_{k-r-1}(s_1s_2\dots s_{k-r-2})}^{(k-r)}}, \\
&\quad r=2, 3, \dots, k-2. \tag{3.8}
\end{aligned}$$

Using (3.8), we obtain the joint c.f. of  $\bar{Y}_{k s_1 s_2 \dots s_{k-r-1}^1}, \bar{Y}_{k s_1 s_2 \dots s_{k-r-1}^2}, \dots, \bar{Y}_{k s_1 s_2 \dots s_{k-r-1}^m}, \dots$  as

$$\begin{aligned} & \psi_{k(r)}^{(\ast\ast)}(\omega_1, \omega_2, \dots, \omega_{m_{k-r}}) \\ &= e^{\sum_{j=1}^{m_{k-r}} \omega_j (\mu_k - nd)} \prod_{j=1}^{m_{k-r}} \left\{ h_k \left( \frac{\omega_j}{m_{k-r+1} m_{k-r+2} \dots m_{k-1}} \right) \right\}^{m_{k-r+1} m_{k-r+2} \dots m_{k-1}} \\ & \quad \cdot \left\{ \frac{1}{2} \left( 1 - \frac{1}{2^{k-r-1}} \right) e^{2di \sum_{j=1}^{m_{k-r}} \omega_j} \right. \\ & \quad \left. + \frac{1}{2^{k-r-1}} \prod_{j=1}^{m_{k-r}} \left[ \frac{1}{4} e^{2di \omega_j} + \frac{1}{2} \phi_{r-1}(\omega_j) + \frac{1}{4} \right] + \frac{1}{2} \left( 1 - \frac{1}{2^{k-r-1}} \right) \right\}^n, \end{aligned} \quad (3.9)$$

Using (3.9), we obtain the covariance of  $\bar{Y}_{k s_1 s_2 \dots s_{k-r-1}^j}, \dots$  and  $\bar{Y}_{k s_1 s_2 \dots s_{k-r-1}^l}, \dots$  as

$$\begin{aligned} & \text{Cov}(\bar{Y}_{k s_1 s_2 \dots s_{k-r-1}^j}, \dots, \bar{Y}_{k s_1 s_2 \dots s_{k-r-1}^l}) \\ &= \frac{1}{i^2} \left[ \frac{\partial^2}{\partial \omega_j \partial \omega_l} \psi_{k(r)}^{(\ast\ast)}(\omega_1, \omega_2, \dots, \omega_{m_{k-r}}) \right]_{\omega_1 = \omega_2 = \dots = \omega_{m_{k-r}} = 0} \\ &= n \left\{ \frac{1}{2} \left( 1 - \frac{1}{2^{k-r-1}} \right) (2d)^2 + \frac{1}{2^{k-r-1}} \left[ \frac{1}{4} (2d) + \frac{1}{2} \Delta_{r-1} \right]^2 \right. \\ & \quad \left. - \left[ \frac{1}{2} \left( 1 - \frac{1}{2^{k-r-1}} \right) (2d) + \frac{1}{2^{k-r-1}} \left( \frac{1}{4} (2d) + \frac{1}{2} \Delta_{r-1} \right) \right]^2 \right\} \\ &= n \left\{ 2 \left( 1 - \frac{1}{2^{k-r-1}} \right) d^2 + \frac{1}{2^{k-r-1}} \left( d + \frac{1}{2^r} h \right)^2 \right. \\ & \quad \left. - \left[ \left( 1 - \frac{1}{2^{k-r-1}} \right) d + \frac{1}{2^{k-r-1}} \left( d + \frac{1}{2^r} h \right) \right]^2 \right\} \\ &= n \left\{ 2 \left( 1 - \frac{1}{2^{k-r-1}} \right) d^2 + \frac{1}{2^{k-r-1}} \left( d + \frac{1}{2^r} h \right)^2 - \left( d + \frac{1}{2^{k-1}} h \right)^2 \right\} \\ &= n \left\{ \left( 1 - \frac{1}{2^{k-r-1}} \right) d^2 + \left( \frac{1}{2^{k-r-1}} - \frac{1}{2^{2k-2}} \right) h^2 \right\} \\ &= \left( 1 - \frac{1}{2^{k-r-1}} \right) D + \frac{1}{2^{k-2}} \left( \frac{1}{2^{r+1}} - \frac{1}{2^k} \right) H \end{aligned} \quad (3.10)$$

Using (3.8) further, we have for the conditional expectation and variance of  $\bar{Y}_{k s_1 s_2 \dots s_{k-r}}, \dots$  given  $s_1, s_2, \dots, s_{k-r-1}$ :

(i) For  $2 \leq r \leq k-2$ ,

$$\begin{aligned} & E(\bar{Y}_{k s_1 s_2 \dots s_{k-r}}, \dots | s_1, s_2, \dots, s_{k-r-1}) \\ &= \mu_k - nd + 2d [X_{1s_1}^{(2)} + X_{1s_2(s_1)}^{(3)} + \dots + X_{1s_{k-r-1}(s_1 s_2 \dots s_{k-r-2})}^{(k-r)}] \\ & \quad + X_{2s_{k-r-1}(s_1 s_2 \dots s_{k-r-2})}^{(k-r)} \left[ \frac{1}{2} d + \frac{1}{2} \Delta_{r-1} \right] \\ &= \mu_k - nd + 2d [X_{1s_1}^{(2)} + X_{1s_2(s_1)}^{(3)} + \dots + X_{1s_{k-r-1}(s_1 s_2 \dots s_{k-r-2})}^{(k-r)}] \\ & \quad + X_{2s_{k-r-1}(s_1 s_2 \dots s_{k-r-2})}^{(k-r)} \left[ d + \frac{1}{2^r} h \right] \end{aligned}$$

by Lemma 2, and

$$\begin{aligned} \text{Var} [\bar{Y}_{k s_1 s_2 \dots s_{k-r}} | s_1, s_2, \dots, s_{k-r-1}] &= \frac{1}{m_{k-r+1} m_{k-r+2} \dots m_{k-1}} \sigma_k^2 \\ &+ X_{2s_{k-r-1}(s_1 s_2 \dots s_{k-r-2})}^{(k-r)} \left[ d^2 + \frac{1}{2} L_{r-1} - \left( d + \frac{h}{2^r} \right)^2 \right], \end{aligned}$$

and

$$\begin{aligned} (\text{ii}) \quad \mathbb{E}(Y_{k s_1 s_2 \dots s_{k-1}} | s_1, s_2, \dots, s_{k-2}) \\ = \mu_k - nd + 2d [X_{1s_1}^{(2)} + X_{1s_2(s_1)}^{(3)} + \dots + X_{1s_{k-2}(s_1 s_2 \dots s_{k-3})}^{(k-1)}] \\ + X_{2s_{k-2}(s_1 s_2 \dots s_{k-3})}^{(k-1)} \left[ d + \frac{1}{2} h \right] \end{aligned}$$

and

$$\begin{aligned} \text{Var} [Y_{k s_1 s_2 \dots s_{k-1}} | s_1, s_2, \dots, s_{k-2}] \\ = \sigma_k^2 + X_{2s_{k-2}(s_1 s_2 \dots s_{k-3})}^{(k-1)} \left[ d^2 + \frac{1}{2} (d+h)^2 - \left( d + \frac{h}{2} \right)^2 \right] \\ = \sigma_k^2 + X_{2s_{k-2}(s_1 s_2 \dots s_{k-3})}^{(k-1)} \left[ \frac{1}{2} d^2 + \frac{1}{4} h^2 \right]. \end{aligned}$$

Since

$$\begin{aligned} &\mathbb{E}\{X_{1s_{k-r-1}(s_1 s_2 \dots s_{k-r-2})}^{(k-r)} | (X_{1s_{k-r-2}(s_1 s_2 \dots s_{k-r-3})}^{(k-r-1)}, X_{2s_{k-r-2}(s_1 s_2 \dots s_{k-r-3})}^{(k-r-1)})\} \\ &= \frac{1}{4} X_{2s_{k-r-2}(s_1 s_2 \dots s_{k-r-3})}^{(k-r-1)}, \\ &\mathbb{E}\{X_{2s_{k-r-1}(s_1 s_2 \dots s_{k-r-2})}^{(k-r)} | (X_{1s_{k-r-2}(s_1 s_2 \dots s_{k-r-3})}^{(k-r-1)}, X_{2s_{k-r-2}(s_1 s_2 \dots s_{k-r-3})}^{(k-r-1)})\} \\ &= \frac{1}{2} X_{2s_{k-r-2}(s_1 s_2 \dots s_{k-r-3})}^{(k-r-1)}, \end{aligned}$$

$\mathbb{E}X_{1s_1}^{(2)} = \frac{1}{4}n$  and  $\mathbb{E}X_{2s_1}^{(3)} = \frac{1}{2}n$ , we have then,

$$\begin{aligned} &\mathbb{E}\{\mathbb{E}(\bar{Y}_{k s_1 s_2 \dots s_{k-r}} | s_1, s_2, \dots, s_{k-r-1})\} \\ &= \mu_k - nd + nd \left[ \left( \frac{1}{2} \right) + \left( \frac{1}{2} \right)^2 + \dots + \left( \frac{1}{2} \right)^{k-r-1} \right] + n \left( \frac{1}{2} \right)^{k-r-1} \left( d + \frac{1}{2^r} h \right) \\ &= \mu_k + \frac{1}{2^{k-1}} nh, \quad r=2, 3, \dots, k-2, \end{aligned} \tag{3.11}$$

$$\begin{aligned} &\mathbb{E}\{\mathbb{E}(Y_{k s_1 s_2 \dots s_{k-1}} | s_1, s_2, \dots, s_{k-2})\} \\ &= \mu_k - nd + nd \left[ \frac{1}{2} + \left( \frac{1}{2} \right)^2 + \dots + \left( \frac{1}{2} \right)^{k-2} \right] + n \left( \frac{1}{2} \right)^{k-2} \left( d + \frac{1}{2} h \right) \\ &= \mu_k + \frac{1}{2^{k-1}} nh, \end{aligned} \tag{3.12}$$

$$\begin{aligned} &\mathbb{E}\{\text{Var}(\bar{Y}_{k s_1 s_2 \dots s_{k-r}} | s_1, s_2, \dots, s_{k-r-1})\} \\ &= \frac{1}{m_{k-r+1} m_{k-r+2} \dots m_{k-1}} \sigma_k^2 \\ &+ \left( \frac{1}{2} \right)^{k-r-1} n \left[ d^2 + \frac{1}{2} L_{r-1} - \left( d + \frac{h}{2^r} \right)^2 \right] \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{m_{k-r+1} m_{k-r+2} \dots m_{k-1}} \sigma_k^2 \\
&\quad + \left(\frac{1}{2}\right)^{k-r-1} n \left\{ d^2 - \left(d + \frac{h}{2^r}\right)^2 + \frac{1}{2} \left(d + \frac{h}{2^{r-1}}\right)^2 \right. \\
&\quad \left. + \sum_{a=1}^{r-2} \frac{1}{2^a m_{k-r+1} m_{k-r+2} \dots m_{k-r+a}} \left( \frac{1}{2} d^2 + \frac{1}{2^{2(r-a)}} h^2 \right) \right. \\
&\quad \left. + \frac{1}{2^{r-1} m_{k-r+1} m_{k-r+2} \dots m_{k-1}} \left( \frac{1}{2} d^2 + \frac{1}{4} h^2 \right) \right\} \\
&= \frac{1}{m_{k-r+1} m_{k-r+2} \dots m_{k-1}} \left( \sigma_k^2 + \frac{1}{2^{k-1}} D + \frac{1}{2^k} H \right) \\
&\quad + \sum_{a=1}^{r-2} \frac{1}{2^{k-r+a-1} m_{k-r+1} m_{k-r+2} \dots m_{k-r+a}} \left[ \frac{1}{2} D + \frac{1}{2^{2(r-a)}} H \right] \\
&\quad + \left( \frac{1}{2^{k-r}} D + \frac{1}{2^{k+r-1}} H \right), \quad r=2, 3, \dots, (k-2), \quad (3.13)
\end{aligned}$$

and

$$\begin{aligned}
&\mathbb{E}\{\text{Var}(Y_{k s_1 s_2 \dots s_{k-1}} | s_1, s_2, \dots, s_{k-2})\} \\
&= \sigma_k^2 + \left(\frac{1}{2}\right)^{k-2} n \left[ \frac{1}{2} d^2 + \frac{1}{4} h^2 \right] = \sigma_k^2 + \frac{1}{2^{k-1}} D + \frac{1}{2^k} H. \quad (3.14)
\end{aligned}$$

In (3.13) and (3.14), if  $k=3$ , then

$$\mathbb{E}\{\text{Var}(Y_{3 s_1 s_2} | s_1)\} = \sigma_3^2 + \frac{1}{4} D + \frac{1}{8} H = V_{2F_3}$$

as defined in Mather and Jinks (1971); if  $k=4$ , then

$$\mathbb{E}\{\text{Var}(\bar{Y}_{4 s_1 s_2} | s_1)\} = \frac{1}{m_3} \left( \sigma_4^2 + \frac{1}{8} D + \frac{1}{16} H \right) + \left( \frac{1}{4} D + \frac{1}{32} H \right) = V_{2F_4}$$

and

$$\mathbb{E}\{\text{Var}(Y_{4 s_1 s_2 s_3} | s_1, s_2)\} = \sigma_4^2 + \frac{1}{8} D + \frac{1}{16} H = V_{3F}$$

as defined in Hayman (1960) and Mather and Jinks (1971).

*The expected values of some second degree statistics*

Define

$$\text{Var}[\bar{Y}_{k s_1 \dots}] = V_{1F_k},$$

and

$$\mathbb{E}\{\text{Var}(\bar{Y}_{k s_1 s_2 \dots s_r} | s_1, s_2, \dots, s_{r-1})\} = V_{rF_k}, \quad r=2, 3, \dots, k-1;$$

$$S_{k(1)} = \frac{1}{m_1 - 1} \sum_{s_1=1}^{m_1} (\bar{Y}_{k s_1 \dots} - \bar{Y}_k \dots)^2,$$

and

$$\begin{aligned}
S_{k(r)} &= \frac{1}{m_1 m_2 \dots m_{r-1} (m_r - 1)} \sum_{s_1=1}^{m_1} \sum_{s_2=1}^{m_2} \\
&\quad \dots \sum_{s_{r-1}=1}^{m_{r-1}} \sum_{s_r=1}^{m_r} (\bar{Y}_{k s_1 s_2 \dots s_r} - \bar{Y}_{k s_1 s_2 \dots s_{r-1}} \dots)^2, \\
&\quad r=2, 3, \dots, k-1,
\end{aligned}$$

with  $\bar{Y}_{k s_1 s_2 \dots s_{k-1}} = Y_{k s_1 s_2 \dots s_{k-1}}$ .

In this section, we proceed to show that

$$\mathbb{E} S_{k(r)} = V_{1 F_k}, \quad r=1, 2, \dots, k-1.$$

Now,

$$\begin{aligned} S_{k(1)} &= \frac{1}{m_1 - 1} \sum_{s_1=1}^{m_1} (\bar{Y}_{k s_1 \dots} - \bar{Y}_{k \dots})^2 \\ &= \frac{1}{m_1 - 1} \left\{ \sum_{s_1=1}^{m_1} \left( \bar{Y}_{k s_1 \dots} - \mu_k - \frac{1}{2^{k-1}} nh \right)^2 - m_1 \left( \bar{Y}_{k \dots} - \mu_k - \frac{1}{2^{k-1}} nh \right)^2 \right\}, \\ S_{k(r)} &= \frac{1}{m_1 m_2 \dots m_{r-1} (m_r - 1)} \sum_{s_1=1}^{m_1} \sum_{s_2=1}^{m_2} \dots \sum_{s_{r-1}=1}^{m_{r-1}} \sum_{s_r=1}^{m_r} (\bar{Y}_{k s_1 s_2 \dots s_r \dots} - \bar{Y}_{k s_1 s_2 \dots s_{r-1} \dots})^2 \\ &= \frac{1}{m_1 m_2 \dots m_{r-1} (m_r - 1)} \left\{ \sum_{s_1=1}^{m_1} \sum_{s_2=1}^{m_2} \dots \sum_{s_r=1}^{m_r} \left( \bar{Y}_{k s_1 s_2 \dots s_r \dots} - \mu_k - \frac{1}{2^{k-1}} nh \right)^2 \right. \\ &\quad \left. - m_r \sum_{s_1=1}^{m_1} \sum_{s_2=1}^{m_2} \dots \sum_{s_{r-1}=1}^{m_{r-1}} \left( \bar{Y}_{k s_1 s_2 \dots s_{r-1} \dots} - \mu_k - \frac{1}{2^{k-1}} nh \right)^2 \right\}, \\ &\quad r=2, 3, \dots, k-1. \end{aligned}$$

Hence,

$$\begin{aligned} \mathbb{E} S_{k(1)} &= \frac{1}{m_1 - 1} \{ m_1 \text{Var}(\bar{Y}_{k s_1 \dots}) - \text{Var}(\bar{Y}_{k s_1 \dots}) \} \\ &= \text{Var}(\bar{Y}_{k s_1 \dots}) = V_{1 F_k}, \end{aligned}$$

and

$$\begin{aligned} \mathbb{E} S_{k(r)} &= \frac{1}{m_1 m_2 \dots m_{r-1} (m_r - 1)} \{ m_1 m_2 \dots m_r \text{Var}(\bar{Y}_{k s_1 s_2 \dots s_r}) \\ &\quad - m_1 m_2 \dots m_{r-1} m_r \text{Var}(\bar{Y}_{k s_1 s_2 \dots s_{r-1} \dots}) \} \\ &= \frac{m_r}{(m_r - 1)} \{ \text{Var}(\bar{Y}_{k s_1 s_2 \dots s_r \dots}) - \text{Var}(\bar{Y}_{k s_1 s_2 \dots s_{r-1} \dots}) \}. \end{aligned}$$

But, from (3.7),

$$\begin{aligned} \text{Var}(\bar{Y}_{k s_1 s_2 \dots s_r \dots}) &= \frac{1}{m_{r+1} m_{r+2} \dots m_{k-1}} \left( \sigma_k^2 + \frac{1}{2^{k-1}} D + \frac{1}{2^k} H \right) \\ &\quad + \sum_{a=1}^{k-r-2} \frac{1}{2^{r+a-1} m_{r+1} m_{r+2} \dots m_{r+a}} \left[ \frac{1}{2} D + \frac{1}{2^{2(k-r-a)}} H \right] \\ &\quad + \left[ \left( 1 - \frac{1}{2^r} \right) D + \frac{1}{2^{k-2}} \left( \frac{1}{2^{k-r}} - \frac{1}{2^k} \right) H \right] \quad (4.1) \end{aligned}$$

and

$$\begin{aligned}
 & \text{Var} [\bar{Y}_{k s_1 s_2 \dots s_{r-1} \dots}] \\
 &= \frac{1}{m_r m_{r+1} \dots m_{k-1}} \left( \sigma_k^2 + \frac{1}{2^{k-1}} D + \frac{1}{2^k} H \right) \\
 &\quad + \sum_{a=1}^{k-r-1} \frac{1}{2^{r+a-2} m_1 m_{r+1} \dots m_{r+a-1}} \left( \frac{1}{2} D + \frac{1}{2^{2(k-r-a+1)}} H \right) \\
 &\quad + \left[ \left( 1 - \frac{1}{2^{r-1}} \right) D + \frac{1}{2^{k-2}} \left( \frac{1}{2^{k-r+1}} - \frac{1}{2^k} \right) H \right] \\
 &= \frac{1}{m_r m_{r+1} \dots m_{k-1}} \left( \sigma_k^2 + \frac{1}{2^{k-1}} D + \frac{1}{2^k} H \right) \\
 &\quad + \sum_{a=0}^{k-r-2} \frac{1}{2^{r+a-1} m_r m_{r+1} \dots m_{r+a}} \left( \frac{1}{2} D + \frac{1}{2^{2(k-r-a)}} H \right) \\
 &\quad + \left[ \left( 1 - \frac{1}{2^{r-1}} \right) D + \frac{1}{2^{k-2}} \left( \frac{1}{2^{k-r+1}} - \frac{1}{2^k} \right) H \right] \\
 &= \frac{1}{m_r} \left[ \frac{1}{m_{r+1} m_{r+2} \dots m_{k-1}} \left( \sigma_k^2 + \frac{1}{2^{k-1}} D + \frac{1}{2^k} H \right) \right. \\
 &\quad \left. + \sum_{a=1}^{k-r-2} \frac{1}{2^{r+a-1} m_{r+1} m_{r+2} \dots m_{r+a}} \left( \frac{1}{2} D + \frac{1}{2^{2(k-r+a)}} H \right) \right] \\
 &\quad + \frac{1}{2^{r-1} m_r} \left( \frac{1}{2} D + \frac{1}{2^{2(k-r)}} H \right) + \left( 1 - \frac{1}{2^{r-1}} \right) D \\
 &\quad + \frac{1}{2^{k-2}} \left( \frac{1}{2^{k-r+1}} - \frac{1}{2^k} \right) H \tag{4.2}
 \end{aligned}$$

From (4.1) and (4.2), it follows that

$$\begin{aligned}
 & \text{Var} (\bar{Y}_{k s_1 s_2 \dots s_r \dots}) - \text{Var} (\bar{Y}_{k s_1 s_2 \dots s_{r-1} \dots}) \\
 &= \left( 1 - \frac{1}{m_r} \right) \left[ \frac{1}{m_{r+1} m_{r+2} \dots m_{k-1}} \left( \sigma_k^2 + \frac{1}{2^{k-1}} D + \frac{1}{2^k} H \right) \right. \\
 &\quad \left. + \sum_{a=1}^{k-r-2} \frac{1}{2^{r+a-1} m_{r+1} m_{r+2} \dots m_{r+a}} \left( \frac{1}{2} D + \frac{1}{2^{2(k-r+a)}} H \right) \right] \\
 &\quad + \left( 1 - \frac{1}{m_r} \right) \left[ \frac{1}{2^r} D + \frac{1}{2^{2k-r-1}} H \right]
 \end{aligned}$$

so that

$$\begin{aligned}
 \mathbb{E} S_{k(r)} &= \frac{1}{m_{r+1} m_{r+2} \dots m_{k-1}} \left( \sigma_k^2 + \frac{1}{2^{k-1}} D + \frac{1}{2^k} H \right) \\
 &\quad + \sum_{a=1}^{k-r-2} \frac{1}{2^{r+a-1} m_{r+1} m_{r+2} \dots m_{r+a}} \left( \frac{1}{2} D + \frac{1}{2^{2(k-r+a)}} H \right) \\
 &\quad + \left( \frac{1}{2^r} D + \frac{1}{2^{2k-r-1}} H \right) = V_{r F_k} \\
 &\quad \text{(see (3.13)), } r=2, 3, \dots, k-1.
 \end{aligned}$$

If  $k=3$ , then

$$\begin{aligned}
 & \mathbb{E} \left\{ \frac{1}{m_1-1} \sum_{s_1=1}^{m_1} (\bar{Y}_{3 s_1} - \bar{Y}_{3 ..})^2 \right\} \\
 &= V_{1 F_3} = \left( \frac{1}{2} D + \frac{1}{16} H \right) + \frac{1}{m_2} \left( \sigma_3^2 + \frac{1}{4} D + \frac{1}{8} H \right),
 \end{aligned}$$

and

$$\mathbb{E} \left\{ \frac{1}{m_1(m_2-1)} \sum_{s_1=1}^{m_1} \sum_{s_2=1}^{m_2} (\bar{Y}_{3s_1s_2} - \bar{Y}_{3s_1})^2 \right\} = V_{2F_3} = \sigma_3^2 + \frac{1}{4}D + \frac{1}{8}H;$$

if  $k=4$ , then

$$\begin{aligned} & \mathbb{E} \left\{ \frac{1}{m_1-1} \sum_{s_1=1}^{m_1} (\bar{Y}_{4s_1\dots} - \bar{Y}_{4\dots})^2 \right\} = V_{1F_4} \\ &= \left( \frac{1}{2}D + \frac{1}{64}H \right) + \frac{1}{m_2} \left( \frac{1}{4}D + \frac{1}{32}H \right) + \frac{1}{m_2 m_3} \left( \sigma_4^2 + \frac{1}{8}D + \frac{1}{16}H \right), \\ & \mathbb{E} \left\{ \frac{1}{m_1(m_2-1)} \sum_{s_1=1}^{m_1} \sum_{s_2=1}^{m_2} (\bar{Y}_{4s_1s_2\dots} - \bar{Y}_{4s_1\dots})^2 \right\} = V_{2F_4} \\ &= \left( \frac{1}{4}D + \frac{1}{32}H \right) + \frac{1}{m_3} \left( \sigma_4^2 + \frac{1}{8}D + \frac{1}{16}H \right), \end{aligned}$$

and

$$\begin{aligned} & \mathbb{E} \left\{ \frac{1}{m_1 m_2 (m_3-1)} \sum_{s_1=1}^{m_1} \sum_{s_2=1}^{m_2} \sum_{s_3=1}^{m_3} (\bar{Y}_{4s_1s_2s_3\dots} - \bar{Y}_{4s_1\dots})^2 \right\} \\ &= V_{3F_4} = \sigma_4^2 + \frac{1}{8}D + \frac{1}{16}H. \end{aligned}$$

#### The covariance between $F_{k_1}$ and $F_{k_2}$ phenotypic values

Using the representation results given in Sections 2 and 3, we may also derive the covariance between the  $F_{k_1}$  and  $F_{k_2}$  ( $k_1 > k_2$ ) phenotypic values. Now, from (3.1) and (3.2), if the  $Y_{k_2 s_1 s_2 \dots s_{k_2-1}}$ 's are offsprings of  $Y_{k_1 s_1 s_2 \dots s_{k_1-1}}$ :

$$\begin{aligned} & Y_{k_1 s_1 s_2 \dots s_{k_1-1}} \\ &= \mu_{k_1} - nd + 2d [X_{1s_1}^{(2)} + X_{1s_2(s_1)}^{(3)} + \dots + X_{1s_{k_1-1}(s_1 s_2 \dots s_{k_1-2})}^{(k_1)}] \\ &+ (d+h) X_{2s_{k_1-1}(s_1 s_2 \dots s_{k_1-2})}^{(k_1)} + e_{k_1 s_1 s_2 \dots s_{k_1-1}} \quad (5.1) \end{aligned}$$

and

$$\begin{aligned} & \bar{Y}_{k_2 s_1 s_2 \dots s_{k_1-1} \dots} \\ &= \mu_{k_2} - nd + 2d [X_{1s_1}^{(2)} + X_{1s_2(s_1)}^{(3)} + \dots + X_{1s_{k_1-1}(s_1 s_2 \dots s_{k_1-2})}^{(k_1)}] \\ &+ 2d [\bar{X}_{1\dots(s_1 s_2 \dots s_{k_1-1})}^{(k_1+1)} + \dots + \bar{X}_{1\dots(s_1 s_2 \dots s_{k_1-1} \dots)}^{(k_2+1)}] \\ &+ (d+h) \bar{X}_{2\dots(s_1 s_2 \dots s_{k_1-1} \dots)}^{(k_2+1)} + \bar{e}_{k_2 s_1 s_2 \dots s_{k_1-1} \dots}, \quad (5.2) \end{aligned}$$

From (5.1) and (5.2), it follows that the joint c.f. of  $Y_{k_1 s_1 s_2 \dots s_{k_1-1}}$  and  $\bar{Y}_{k_2 s_1 s_2 \dots s_{k_1-1} \dots}$  is readily obtained as, if it is assumed that  $e_{k_1 s_1 s_2 \dots s_{k_1-1}}$  and  $\bar{e}_{k_2 s_1 s_2 \dots s_{k_1-1} \dots}$  are independently distributed,

$$\begin{aligned} \psi_c^*(\omega_1, \omega_2) = & e^{i\omega_1(\mu_{k_1}-nd)+i\omega_2(\mu_{k_2}-nd)} \\ & \cdot h_{k_1}(\omega_1) \left\{ h_{k_2} \left( \frac{\omega_2}{m_{k_1} m_{k_1+1} \dots m_{k_2-1}} \right) \right\}^{m_{k_1} m_{k_1+1} \dots m_{k_2-1}} \\ & \cdot \left\{ \frac{1}{2} \left( 1 - \frac{1}{2^{k_1-1}} \right) e^{2di(\omega_1+\omega_2)} + \frac{1}{2^{k_1-1}} \Phi_0(\omega_1) \Phi_{k_2-k_1}(\omega_2) \right. \\ & \left. + \frac{1}{2} \left( 1 - \frac{1}{2^{k_1-1}} \right) \right\}^n, \end{aligned} \quad (5.3)$$

Using (5.3), the covariance between  $Y_{k_1 s_1 s_2 \dots s_{k_1-1}}$  and  $\bar{Y}_{k_2 s_1 s_2 \dots s_{k_1-1} \dots}$  is then obtained as ( $k_2 > k_1$ ),

$$\begin{aligned} \text{Cov } (k_1, k_2) &= \frac{1}{i^2} \left\{ \frac{\partial^2}{\partial \omega_1 \partial \omega_2} \log \psi_c^*(\omega_1, \omega_2) \right\}_{\omega_1=\omega_2=0} \\ &= n \left\{ 2 \left( 1 - \frac{1}{2^{k_1-1}} \right) d^2 + \frac{1}{2^{k_1-1}} (d+h) \Delta_{k_2-k_1} \right. \\ &\quad - \left[ \left( 1 - \frac{1}{2^{k_1-1}} \right) d + \frac{1}{2^{k_1-1}} (d+h) \right] \\ &\quad \times \left[ \left( 1 - \frac{1}{2^{k_1-1}} \right) d + \frac{1}{2^{k_1-1}} \Delta_{k_2-k_1} \right] \} \\ &= n \left\{ 2 \left( 1 - \frac{1}{2^{k_1-1}} \right) d^2 + \frac{1}{2^{k_1-1}} (d+h) \left( d + \frac{h}{2^{k_2-k_1}} \right) \right. \\ &\quad - \left( d + \frac{1}{2^{k_1-1}} h \right) \left[ \left( 1 - \frac{1}{2^{k_1-1}} \right) d + \frac{1}{2^{k_1-1}} \left( d + \frac{1}{2^{k_2-k_1}} h \right) \right] \} \\ &= n \left\{ 2 \left( 1 - \frac{1}{2^{k_1-1}} \right) d^2 + \frac{1}{2^{k_1-1}} (d^2+dh) + \frac{1}{2^{k_2-1}} (dh+h^2) \right. \\ &\quad - \left( d + \frac{1}{2^{k_1-1}} h \right) \left( d + \frac{1}{2^{k_2-1}} h \right) \} \\ &= n \left\{ \left( 1 - \frac{1}{2^{k_1-1}} \right) d^2 + \frac{1}{2^{k_2-1}} \left( 1 - \frac{1}{2^{k_1-1}} \right) h^2 \right\} \\ &= \left( 1 - \frac{1}{2^{k_1-1}} \right) \left( D + \frac{1}{2^{k_2-1}} H \right). \end{aligned} \quad (5.4)$$

In (5.4), if  $k_1=2$  and  $k_2=3$ , then  $\text{Cov } (2, 3) = \frac{1}{2} \left( D + \frac{1}{4} H \right)$ ;

if  $k_1=2$ , then  $\text{Cov } (2, k_2) = \frac{1}{2} \left( D + \frac{1}{2^{k_2-1}} H \right)$ ;

if  $k_1=3$  and  $k_2=4$ , then  $\text{Cov } (3, 4) = \frac{3}{4} \left( D + \frac{1}{8} H \right)$ , etc.

Some simple results such as  $\text{Cov } (2, 3)$  have been given in Mather and Jinks (1971).

#### Literature Cited

- HAYMAN, B. I. 1960. Maximum likelihood estimation of genetic components of variation. *Biometrics* **16**: 369-381.  
 MATHER, K. and J. L. JINKS. 1971. *Biometrical Genetics*. Cornell Univ. Press, Ithaca, New York. 382 p.

TAN, W. Y. and W. C. CHANG. 1972. Convolution approach to the genetic analysis of quantitative characters of self-fertilized population. *Biometrics* 28: 1073-1090.

## 論統計遺傳學之理論

譚外元

美國華盛頓州立大學數學系

1972年本文作者等曾導出自交植物分離集團之機率分布 (*Biometrics* 28:1073-1090)。本文之結果為該文之推廣。

本文求出在已知其祖先之訊識下的分離集團之機率分布，並由此數學模型而導出各種變方及共變方。此等結果頗為一般化，使 Hayman (1960) 及 Mather 與 Jinks (1971) 之結果均成為本法之特例。