

Distributions of Genotypes after a Panmixia

By Yûsaku KOMATU

(Received Feb. 19, 1954)

1. Introduction.

Consider a population of size $2N$ consisting of N females and N males. We observe a single inherited character which consists of m multiple alleles at one diploid locus denoted by

$$A_i \qquad (i=1, \dots, m)$$

and of which the inheritance is subject to Mendelian law.

There are $m(m+1)/2$ possible genotypes $A_a A_b$ ($a, b=1, \dots, m; a \leq b$) among which m types $A_b A_b$ ($b=1, \dots, m$) are homozygous and $m(m-1)/2$ types $A_a A_b$ ($a, b=1, \dots, m; a < b$) are heterozygous. Let the distributions of these $m(m+1)/2$ genotypes $A_a A_b$ ($a \leq b$) in females and in males be designated by

$$F_{ab} \quad \text{and} \quad M_{ab} \qquad (a, b=1, \dots, m; a \leq b)$$

or, as the aggregates, by

$$\mathfrak{F} = (F_{11}, \dots, F_{mm}, F_{12}, \dots, F_{m-1,m})$$

and

$$\mathfrak{M} = (M_{11}, \dots, M_{mm}, M_{12}, \dots, M_{m-1,m})$$

respectively, so that

$$\sum_{a \leq b} F_{ab} = \sum_{a \leq b} M_{ab} = N.$$

The order of genes in a genotype being immaterial, both genotypes $A_a A_b$ and $A_b A_a$ are regarded as identical each other even when the suffices a and b are distinct. Accordingly, we put $F_{ab} = F_{ba}$ and $M_{ab} = M_{ba}$.

We now introduce a set of stochastic variables

$$\mathfrak{C} = (C_{11}, \dots, C_{mm}, C_{12}, \dots, C_{m-1,m})$$

extending over non-negative integers and satisfying a single equation

$$\sum_{a \leq b} C_{ab} = N.$$

We designate by

$$\psi(\mathbb{C}) \equiv \psi(\mathbb{C} | \mathfrak{F}; \mathfrak{M})$$

the probability that a distribution in the next generation after a panmixia is given by \mathbb{C} , each mating being supposed to produce one child so that, as stated above, $\sum_{a \leq b} C_{ab} = N$.

In order to preserve the size of the whole population in the next generation, it would be rather preferable to suppose that each mating produces two children, one female and one male. In the present purpose, however, one may think to confine oneself to a distribution in either sex in the next generation, by supposing that the fertility of every mating is equal.

The main purpose of the present paper is to establish an explicit expression for a probability-generating function Φ defined by

$$\Phi(\mathfrak{z}) \equiv \Phi(\mathfrak{z} | \mathfrak{F}; \mathfrak{M}) = \sum_{\mathbb{C}} \psi(\mathbb{C} | \mathfrak{F}; \mathfrak{M}) \prod_{a \leq b} z_{ab}^{C_{ab}},$$

\mathfrak{z} designating a set of indeterminate variables:

$$\mathfrak{z} = (z_{11}, \dots, z_{mm}, z_{12}, \dots, z_{m-1,m});$$

the summation extends over the whole range of \mathbb{C} .

By the way, it is noted that a variable involved in \mathbb{C} , for instance $C_{m-1,m}$, similarly a variable involved in \mathfrak{F} and in \mathfrak{M} , and consequently a variable involved in \mathfrak{z} , for instance $z_{m-1,m}$, may be omitted, since we have supposed that total numbers of individuals are fixed in children, in females, and in males. The omission of $z_{m-1,m}$ corresponds, of course, to put $z_{m-1,m} = 1$. The range of summation in the last expression is then replaced by the sets of $m(m+1)/2 - 1$ non-negative integers C_{ab} ($a, b = 1, \dots, m; a \leq b; (a, b) \neq (m-1, m)$) satisfying

$$\sum_{a \leq b} C_{ab} - C_{m-1,m} \leq N.$$

However, we shall retain, for the sake of formal symmetry, the dependent quantities concerning a genotype $A_{m-1}A_m$ in our expressions, since they will give rise to few complications.

We now state a further supplementary remark. Though our original problem concerns m different genes, it is reducible to a problem with a lower m , according to circumstances, for instance, when there concern routine statistics such as means, variances, covariances, etc. on the stochastic variables under consideration. In fact, if we consider a homozygous type A_iA_i , the distinction among the genes except A_i becomes a matter of indifference. Consequently, if we put

$$\begin{aligned} \sum_{b \neq i} F_{ib} &= F_{i\omega}, & \sum_{\substack{a, b \neq i \\ a \leq b}} F_{ab} &= F_{\omega\omega}, \\ \sum_{b \neq i} M_{ib} &= M_{i\omega}, & \sum_{\substack{a, b \neq i \\ a \leq b}} M_{ab} &= M_{\omega\omega}, \\ \sum_{b \neq i} C_{ib} &= C_{i\omega}, & \sum_{\substack{a, b \neq i \\ a \leq b}} C_{ab} &= C_{\omega\omega}, \end{aligned}$$

then the mean and variance of the random variable C_{ii} in the distributions in the next generation coincide really with those which result from considering only two genes, i. e. the gene A_i itself and an imaginary gene A_ω introduced in a manner stated just above. On the other hand, if we consider a heterozygous type A_iA_j , the distinction among the genes except A_i and A_j becomes a matter of indifference. Consequently, if we put

$$\sum_{b \neq i, j} F_{ib} = F_{i\omega}, \quad \sum_{b \neq i, j} F_{jb} = F_{j\omega}, \quad \sum_{\substack{a, b \neq i, j \\ a \leq b}} F_{ab} = F_{\omega\omega}, \quad \text{etc.},$$

the mean and variance of the random variable C_{ij} coincide just with those which result from considering only three genes, i. e. the genes A_i and A_j themselves and an imaginary gene A_ω . Quite similarly, in order to determine the covariance between any two genotypes, it suffices to study a problem in which there concern at most four genes.

In spite of such being the case, and moreover though the arguments for any m can be performed, in principle, quite similarly to those for such lower m , for instance $m=2, 3$ or 4 , we shall treat in the following lines the original problem itself for the sake of completeness.

In a series of previous papers¹⁾, we have dealt with analogous

1) Y. Komatu, Probability-theoretic investigations on inheritance. I-XVI. Proc. Japan Acad. 27-29 (1951-1953); Y. Komatu and H. Nishimiya, Probabilistic investigations on inheritance in consanguineous families. Bull. Tokyo Inst. Tech. (1954), 1-222 etc., of which preliminary announcements are found in Y. Komatu and H. Nishimiya, Probabilities on inheritance in consanguineous families. Proc. Japan Acad. 30 (1954), 42-52, 148-155, 236-247 etc.

problems far extensively but with few preciseness. We have considered there a population of infinite size and studied merely the means of several distributions. The results of the present paper will give generalizations of some of the previous results.

2. Probability-generating function.

We now observe m distinct genes A_i ($i=1, \dots, m$). We first consider a partition of N males into $m(m+1)/2$ classes according to the same number of genotypes $A_a A_b$ ($a, b=1, \dots, m; a \leq b$) of females to be married. Namely, let each of M_{ab} ($a, b=1, \dots, m; a \leq b$) individuals in male-population be divided into $m(m+1)/2$ classes, here empty classes being admissible, in such a manner

$$\begin{aligned} M_{ab} &= \sum_{c \leq d} x_{abcd}, \\ F_{cd} &= \sum_{a \leq b} x_{abcd} \end{aligned} \quad \left(\begin{array}{l} a, b, c, d=1, \dots, m; \\ a \leq b; c \leq d \end{array} \right).$$

Let the matings take place such that, for every pair of c and d with $c, d=1, \dots, m; c \leq d$, x_{abcd} ($a, b=1, \dots, m; a \leq b$) males of genotypes $A_a A_b$ are combined, as a whole, with F_{cd} females of genotype $A_c A_d$.

All the possible permutations of N males consisting of M_{ab} ($a, b=1, \dots, m; a \leq b$) individuals of genotypes $A_a A_b$, respectively, then amount to

$$N! / \prod_{a \leq b} M_{ab}!$$

while the permutations of F_{cd} males to be married with females of genotype $A_c A_d$, these males consisting of x_{abcd} ($a, b=1, \dots, m; a \leq b$) individuals of genotypes $A_a A_b$, respectively, amount to

$$F_{cd}! / \prod_{a \leq b} x_{abcd}!.$$

On the other hand, it is supposed that any mating $A_a A_b \times A_c A_d$ produces each of four possible genotypes $A_a A_c, A_a A_d, A_b A_c$ and $A_b A_d$ *equally likely*, that is, with probability $1/4$. When some of genes are coincident and hence two or four of the genotypes to be produced are identical, the probability is then, of course, interpreted as the corresponding sum, namely, as $1/2$ or 1 , respectively. To state more

precisely, if we denote by i, j, h and k the suffices indicating the genes different each other, then the matings

$$A_i A_i \times A_i A_i, \quad A_i A_i \times A_i A_k, \quad A_i A_i \times A_k A_k, \quad A_i A_i \times A_h A_h,$$

$$A_i A_j \times A_i A_j, \quad A_i A_j \times A_i A_k, \quad A_i A_j \times A_h A_k$$

produce a child of type $A_i A_i; A_j A_j; A_i A_j; A_i A_h; A_i A_k; A_j A_h; A_j A_k$ with probabilities

1,	1/2,	0,	0,	1/4,	1/4,	0;
0,	0,	0,	0,	1/4,	0,	0;
0,	0,	0,	0,	1/2,	1/4,	0;
0,	0,	0,	1/2,	0,	0,	1/4;
0,	1/2,	1,	1/2,	0,	1/4,	1/4;
0,	0,	0,	0,	0,	0,	1/4;
0,	0,	0,	0,	0,	1/4,	1/4,

respectively.

Consequently, the generating function is given by

$$\begin{aligned} \phi(z) = & \sum_{\mathbb{C}} \psi(\mathbb{C}) \prod_{f \leq g} z_{fg}^{\mathbb{C}} = \sum_{\mathbb{C}} \frac{\prod_{a \leq b} M_{ab}!}{N!} \sum_{\mathbb{E}} \prod_{c \leq d} \frac{F_{cd}!}{\prod_{a \leq b} x_{abcd}} \\ & \cdot \sum_{a; c} z_{ac}^{x_{aacc}} \cdot \prod_{a < b; c} \frac{x_{abcc}!}{x_{abcc}^{(ac)}! x_{abcc}^{(bc)}!} \left(\frac{z_{ac}}{2}\right)^{x_{abcc}^{(ac)}} \left(\frac{z_{bc}}{2}\right)^{x_{abcc}^{(bc)}} \\ & \cdot \prod_{a; c < d} \frac{x_{aacd}!}{x_{aacd}^{(ac)}! x_{aacd}^{(ad)}!} \left(\frac{z_{ac}}{2}\right)^{x_{aacd}^{(ac)}} \left(\frac{z_{ad}}{2}\right)^{x_{aacd}^{(ad)}} \\ & \cdot \prod_{a < b} \frac{x_{abab}!}{x_{abab}^{(aa)}! x_{abab}^{(ab)}! x_{abab}^{(bb)}!} \left(\frac{z_{aa}}{4}\right)^{x_{abab}^{(aa)}} \left(\frac{z_{ab}}{2}\right)^{x_{abab}^{(ab)}} \left(\frac{z_{bb}}{4}\right)^{x_{abab}^{(bb)}} \\ & \cdot \prod_{\substack{a < b; c < d \\ (a, b) \neq (c, d)}} \frac{x_{abcd}!}{x_{abcd}^{(ac)}! x_{abcd}^{(ad)}! x_{abcd}^{(bc)}! x_{abcd}^{(bd)}!} \\ & \times \left(\frac{z_{ac}}{4}\right)^{x_{abcd}^{(ac)}} \left(\frac{z_{ad}}{4}\right)^{x_{abcd}^{(ad)}} \left(\frac{z_{bc}}{4}\right)^{x_{abcd}^{(bc)}} \left(\frac{z_{bd}}{4}\right)^{x_{abcd}^{(bd)}}, \end{aligned}$$

where we make an agreement

$$z_{fg} = z_{gf}, \quad x_{abcd}^{(fg)} = x_{abcd}^{(gf)}.$$

In the last expression, we further put

$$\sum_{f \leq g} x_{abcd}^{(fg)} = x_{abcd} \quad (a, b, c, d = 1, \dots, m; a \leq b; c \leq d),$$

in which the summation runs over possible types $A_f A_g$ of a child originated from a mating $A_a A_b \times A_c A_d$. In case where a mating $A_a A_b \times A_c A_d$ can produce only one type of children, namely, when $A_a A_b$ and $A_c A_d$ both are homozygous, i. e. $a = b$ and $c = d$, the quantity $x_{abcd}^{(fg)}$ is interpreted to be equal to x_{abcd} , i. e. $x_{aacc}^{(ac)} = x_{aacc}$. Among the quantities here concerned there exist further several relations expressing their dependency. They may be set out as follows:

$$\sum_{a \leq b; c \leq d} x_{abcd}^{(fg)} = C_{fg} \quad (f, g = 1, \dots, m; f \leq g),$$

$$\sum_{c \leq d} x_{abcd} = M_{ab}, \quad \sum_{a \leq b} x_{abcd} = F_{cd} \quad (a, b, c, d = 1, \dots, m; a \leq b; c \leq d).$$

Taking these relations of dependence into account, the last summation in the above expression for $\Phi(\beta)$ extends over all the possible sets of non-negative integers $\{x_{abcd}^{(fg)}\}$, the second summation over all the possible partitions represented by the sets $\mathfrak{x} = \{x_{abcd}\}$, and the first summation over all the sets of non-negative integers $\mathfrak{C} = \{C_{fg}\}$ with total sum N .

We first perform, by making use of multinomial identities, the summation with respect to suffices fg involved in $\{x_{abcd}^{(fg)}\}$ and $\{C_{fg}\}$, whence follows a relation

$$\begin{aligned} \Phi(\beta) = & \frac{\prod_{a \leq b} M_{ab}!}{N!} \sum_{\mathfrak{x}} \prod_{c \leq d} \frac{F_{cd}!}{\prod_{a \leq b} x_{abcd}!} \\ & \cdot \prod_{a; c} z_{ac}^{x_{aacc}} \cdot \prod_{a < b; c} \left(\frac{z_{ac} + z_{bc}}{2} \right)^{x_{abcc}} \cdot \prod_{a; c < d} \left(\frac{z_{ac} + z_{ad}}{2} \right)^{x_{aacd}} \\ & \cdot \prod_{a < b} \left(\frac{z_{aa} + 2z_{ab} + z_{bb}}{4} \right)^{x_{abab}} \cdot \prod_{\substack{a < b; c < d \\ (a,b) \neq (c,d)}} \left(\frac{z_{ac} + z_{ad} + z_{bc} + z_{bd}}{4} \right)^{x_{abcd}}, \end{aligned}$$

which may be written in a more brief form

$$\Phi(\beta) = \frac{\prod_{a \leq b} M_{ab}!}{N!} \sum_{\mathfrak{x}} \prod_{c \leq d} \frac{F_{cd}!}{\prod_{a \leq b} x_{abcd}!} \prod_{a \leq b; c \leq d} \left(\frac{z_{ac} + z_{ad} + z_{bc} + z_{bd}}{4} \right)^{x_{abcd}}.$$

We now introduce, with parameters

$$t=(t_{ab}) \quad (a, b=1, \dots, m; a \leq b),$$

a function defined by

$$\begin{aligned} \Phi(\mathfrak{z}|t) &\equiv \Phi(\mathfrak{z}|t|\mathfrak{F}; \mathfrak{M}) \\ &= \frac{1}{N!} \prod_{a \leq b} \frac{M_{ab}!}{t_{ab}^{M_{ab}}} \cdot \prod_{c \leq d} \left(\sum_{a \leq b} t_{ab} \frac{z_{ac} + z_{ad} + z_{bc} + z_{bd}}{4} \right)^{F_{cd}}. \end{aligned}$$

Then, by remembering the relations

$$\sum_{a \leq b} x_{abcd} = F_{cd}, \quad \sum_{c \leq d} x_{abcd} = M_{ab},$$

the multinomial theorem implies that *our generating function* $\Phi(\mathfrak{z})$ is given by the constant term, i. e. the coefficient of the term $\prod_{a \leq b} t_{ab}^0$ in the Laurent expansion around the origin of the last quantity $\Phi(\mathfrak{z}|t)$ regarded as a rational function of $m(m+1)/2$ variables $t=(t_{ab})$ ($a, b=1, \dots, m; a \leq b$).

We further introduce, for a later purpose, with parameters

$$\mathfrak{s}=(s_{ab}) \quad (a, b=1, \dots, m; a \leq b),$$

a function defined by

$$\begin{aligned} \Phi(\mathfrak{z}|\mathfrak{s}; t) &\equiv \Phi(\mathfrak{z}|\mathfrak{s}; t|\mathfrak{F}; \mathfrak{M}) \\ &= \frac{1}{N!^2} \prod_{a \leq b} \frac{F_{ab}! M_{ab}!}{s_{ab}^{F_{ab}} t_{ab}^{M_{ab}}} \cdot \left(\sum_{a \leq b, c \leq d} s_{ab} t_{cd} \frac{z_{ac} + z_{ad} + z_{bc} + z_{bd}}{4} \right)^N. \end{aligned}$$

Then, by taking into account the relation

$$\sum_{c \leq d} F_{cd} = N,$$

the multinomial theorem again implies that *our generating function* $\Phi(\mathfrak{z})$ is given also by the constant term in the Laurent expansion around the origin of the last quantity $\Phi(\mathfrak{z}|\mathfrak{s}; t)$ regarded as a rational function of $m(m+1)$ variables $\mathfrak{s}=(s_{ab})$ and $t=(t_{ab})$ ($a, b=1, \dots, m; a \leq b$).

By separating the factors involved in the last two functions according to ones concerning homozygous and heterozygous types, they are brought into the following forms:

$$\begin{aligned} \Phi(\mathfrak{z}|\mathfrak{t}) = & \frac{1}{N!} \prod_i \frac{M_{ii}!}{t_{ii}^{M_{ii}}} \prod_{i < j} \frac{M_{ij}!}{t_{ij}^{M_{ij}}} \\ & \cdot \prod_i \left(t_{ii} z_{ii} + \sum_{b \neq i} \left(t_{bb} z_{ib} + t_{ib} \frac{z_{ii} + z_{ib}}{2} \right) + \sum_{\substack{a, b \neq i \\ a < b}} t_{ab} \frac{z_{ia} + z_{ib}}{2} \right)^{F_{ii}} \\ & \cdot \prod_{i < j} \left(t_{ii} \frac{z_{ii} + z_{ij}}{2} + t_{jj} \frac{z_{jj} + z_{ij}}{2} + t_{ij} \frac{z_{ii} + 2z_{ij} + z_{jj}}{4} \right. \\ & \quad \left. + \sum_{b \neq i, j} \left(t_{bb} \frac{z_{ib} + z_{jb}}{2} + t_{ib} \frac{z_{ii} + z_{ij} + z_{ib} + z_{jb}}{4} \right. \right. \\ & \quad \left. \left. + t_{jb} \frac{z_{jj} + z_{ij} + z_{ib} + z_{jb}}{4} \right) + \sum_{\substack{a, b \neq i, j \\ a < b}} t_{ab} \frac{z_{ia} + z_{ib} + z_{ja} + z_{jb}}{4} \right)^{F_{ij}} \end{aligned}$$

and

$$\begin{aligned} & \Phi(\mathfrak{z}|\mathfrak{s}; \mathfrak{t}) \\ = & \frac{1}{N!^2} \prod_i \frac{F_{ii}! M_{ii}!}{s_{ii}^{F_{ii}} t_{ii}^{M_{ii}}} \prod_{i < j} \frac{F_{ij}! M_{ij}!}{s_{ij}^{F_{ij}} t_{ij}^{M_{ij}}} \cdot \left(\sum_i u_i v_i z_{ii} + \sum_{i < j} (u_i v_j + u_j v_i) z_{ij} \right)^N, \end{aligned}$$

where we put, for the sake of brevity,

$$u_i = s_{ii} + \sum_{b \neq i} \frac{s_{ib}}{2}, \quad v_i = t_{ii} + \sum_{b \neq i} \frac{t_{ib}}{2} \quad (i=1, \dots, m).$$

It would be noticed, in passing, that the generating function $\Phi(\mathfrak{z})$ itself is expressible, for instance, by means of a contour integral in a form

$$\Phi(\mathfrak{z}) = \frac{1}{(2\pi\sqrt{-1})^{m(m+1)}} \int \dots \int \Phi(\mathfrak{z}|\mathfrak{s}; \mathfrak{t}) \prod_{a \leq b} \frac{ds_{ab} dt_{ab}}{s_{ab} t_{ab}},$$

where the multiple integration is taken along the unit circumferences

$$|s_{ab}| = 1, \quad |t_{ab}| = 1 \quad (a, b = 1, \dots, m; a \leq b)$$

in the positive sense on respective complex planes. However, this integral representation will not be availed in the following lines.

3. Means of stochastic variables.

By virtue of its own meaning, there must hold an identity

$$\Phi(e) = 1,$$

where we put

$$e = \left(\frac{1}{\mathbf{1}}, \dots, \frac{m(m+1)/2}{\mathbf{1}} \right).$$

It is a matter of course that our expression derived above conforms to this requirement. In fact, since we have

$$\Phi(e|\mathfrak{s}; t) = \frac{1}{N!^2} \prod_{a \leq b} \frac{F_{ab}! M_{ab}!}{S_{ab}^{F_{ab}} t_{ab}^{M_{ab}}} \cdot \left(\sum_{a \leq b} s_{ab} \sum_{a \leq b} t_{ab} \right)^N,$$

we get, by picking up the constant terms in the Laurent expansion,

$$\Phi(e) = \frac{1}{N!^2} \prod_{a \leq b} F_{ab}! M_{ab}! \cdot \frac{N!}{\prod_{a \leq b} F_{ab}!} \frac{N!}{\prod_{a \leq b} M_{ab}!} = 1.$$

Now, an analytical expression for the generating function having been established in an explicit manner, the mean of every stochastic variable C_{ab} can be readily calculated.

First, for a *homozygous type* $A_i A_i$, we get

$$\frac{\partial \Phi}{\partial z_{ii}}(e|\mathfrak{s}; t) = \frac{1}{N!^2} \prod_{a \leq b} \frac{F_{ab}! M_{ab}!}{S_{ab}^{F_{ab}} t_{ab}^{M_{ab}}} \cdot N \left(\sum_{a \leq b} s_{ab} \sum_{a \leq b} t_{ab} \right)^{N-1} u_i v_i.$$

Consequently, by separating the constant term in the Laurent expansion around the origin of the last quantity regarded as a function of variables s_{ab}, t_{ab} ($a, b = 1, \dots, m; a \leq b$), we obtain for the *mean of* C_{ii} an expression

$$\begin{aligned} \tilde{C}_{ii} &\equiv \tilde{C}_{ii}(\mathfrak{F}; \mathfrak{M}) = \sum_{\mathfrak{C}} C_{ii} \Psi(\mathfrak{C}|\mathfrak{F}; \mathfrak{M}) = \frac{\partial \Phi}{\partial z_{ii}}(e|\mathfrak{F}; \mathfrak{M}) \\ &= \frac{1}{N} \left(F_{ii} + \sum_{b \neq i} \frac{F_{ib}}{2} \right) \left(M_{ii} + \sum_{b \neq i} \frac{M_{ib}}{2} \right). \end{aligned}$$

Next, for a *heterozygous type* $A_i A_j$, we get

$$\frac{\partial \Phi}{\partial z_{ij}}(e|\mathfrak{s}; t) = \frac{1}{N!^2} \prod_{a \leq b} \frac{F_{ab}! M_{ab}!}{S_{ab}^{F_{ab}} t_{ab}^{M_{ab}}} \cdot N \left(\sum_{a \leq b} s_{ab} \sum_{a \leq b} t_{ab} \right)^{N-1} (u_i v_j + u_j v_i),$$

whence follows, similarly as above, for the *mean of* C_{ij} an expression

$$\begin{aligned}\tilde{C}_{ij} &\equiv \tilde{C}_{ij}(\mathfrak{F}; \mathfrak{M}) = \sum_{\mathfrak{C}} C_{ij} \Psi(\mathfrak{C} | \mathfrak{F}; \mathfrak{M}) = \frac{\partial \Phi}{\partial z_{ij}}(e | \mathfrak{F}; \mathfrak{M}) \\ &= \frac{1}{N} \left(\left(F_{ii} + \sum_{b \neq i} \frac{F_{ib}}{2} \right) \left(M_{jj} + \sum_{b \neq j} \frac{M_{jb}}{2} \right) \right. \\ &\quad \left. + \left(F_{jj} + \sum_{b \neq j} \frac{F_{jb}}{2} \right) \left(M_{ii} + \sum_{b \neq i} \frac{M_{ib}}{2} \right) \right).\end{aligned}$$

Thus, by introducing the quantities defined by

$$p_i^{(F)} = \frac{1}{N} \left(F_{ii} + \sum_{b \neq i} \frac{F_{ib}}{2} \right), \quad p_i^{(M)} = \frac{1}{N} \left(M_{ii} + \sum_{b \neq i} \frac{M_{ib}}{2} \right) \quad (i=1, \dots, m),$$

the expressions for the mean just obtained can be brought into brief forms

$$\tilde{C}_{ii} = N p_i^{(F)} p_i^{(M)}, \quad \tilde{C}_{ij} = N (p_i^{(F)} p_j^{(M)} + p_j^{(F)} p_i^{(M)}) \quad (i, j=1, \dots, m; i < j).$$

The last result is quite plausible. In fact, the relative frequencies of the genes A_i ($i=1, \dots, m$) in female-population and in male-population are equal to $p_i^{(F)}$ and $p_i^{(M)}$, respectively. As shown in a previous paper²⁾, the random matings between these two populations produce probabilistically a distribution in the next generation with frequencies of $A_a A_b$ just given by \tilde{C}_{ab} .

4. Variances and covariances.

In order to express the variances, covariances, or, more generally, the quantities concerning the moments of stochastic variables in clearer forms, it will be convenient to introduce an abbreviated notation defined by

$$[A^n]! = \frac{A!}{(A-n)!} = A(A-1) \cdots (A-n+1) \quad (n=0, 1, \dots, A),$$

which represents the number of permutations of n things taken from A different things and is often designated by ${}_A P_n$. If we concern

2) Cf. the first paper cited in¹⁾. Especially, I. Distribution of genes. Proc. Japan Acad. 27 (1951), 371-377; cf. also III. Further discussions on cross breeding. Ibid. 459-483.

merely variances and covariances, it will suffice to introduce the notation only for $n=2$, i. e. $[A^2] = A(A-1)$. However, it will become useful when there concern the moments of higher orders. More generally, for a given polynomial of any number of arguments

$$\mathcal{Q}(A_1, \dots, A_\omega) = \sum_{\Pi} a_{n_1 \dots n_\omega} \prod_{\kappa=1}^{\omega} A_{\kappa}^{n_{\kappa}}$$

we put

$$[\mathcal{Q}(A_1, \dots, A_\omega)]^{\dagger} = \sum_{\Pi} a_{n_1 \dots n_\omega} \prod_{\kappa=1}^{\omega} [A_{\kappa}^{n_{\kappa}}]^{\dagger}.$$

Now, it is readily shown by induction that *there holds for any derivative of arbitrary order a representation*

$$\begin{aligned} & \prod_{i \leq j} \left(\frac{\partial}{\partial z_{ij}} \right)^{n_{ij}} \Phi(e|\xi; t) \\ &= \frac{1}{N!^2} \prod_{a \leq b} \frac{F_{ab}! M_{ab}!}{S_{ab}^{F_{ab}} t_{ab}^{M_{ab}}} \left[N^{\sum_{i \leq j} n_{ij}} \right]^{\dagger} \left(\sum_{a \leq b} S_{ab} \sum_{a \leq b} t_{ab} \right)^{N - \sum_{i \leq j} n_{ij}} \\ & \quad \cdot \prod_i (u_i v_i)^{n_{ii}} \prod_{i < j} (u_i v_j + u_j v_i)^{n_{ij}}, \end{aligned}$$

whence follows

$$\prod_{i \leq j} \left(\frac{\partial}{\partial z_{ij}} \right)^{n_{ij}} \Phi(e) = \left[N^{\sum_{i \leq j} n_{ij}} \right]^{\dagger} \left[\prod_{i \leq j} (N\tilde{C})^{n_{ij}} \right]^{\dagger}.$$

Here the relevant variables in the bracket-notation $[]^{\dagger}$ for the second factor are $m(m+1)$ variables F_{ij} and M_{ij} ($i, j=1, \dots, m; i \leq j$); as shown above, we have

$$\begin{aligned} N\tilde{C}_{ii} &= \left(F_{ii} + \sum_{b \neq i} \frac{F_{ib}}{2} \right) \left(M_{ii} + \sum_{b \neq i} \frac{M_{ib}}{2} \right) & (i=1, \dots, m), \\ N\tilde{C}_{ij} &= \left(F_{ii} + \sum_{b \neq i} \frac{F_{ib}}{2} \right) \left(M_{jj} + \sum_{b \neq j} \frac{M_{jb}}{2} \right) \\ & \quad + \left(F_{jj} + \sum_{b \neq j} \frac{F_{jb}}{2} \right) \left(M_{ii} + \sum_{b \neq i} \frac{M_{ib}}{2} \right) & (i, j=1, \dots, m; i < j). \end{aligned}$$

By means of the last relation, any moments of arbitrary order can be readily computed. As an illustrative example, we shall here deal

with the *variances and covariances*.

First, the variance of C_{ab} is given by

$$\begin{aligned}\text{var}(C_{ab}) &\equiv \text{var}(C_{ab} | \mathfrak{F}; \mathfrak{M}) = \sum_{\mathfrak{C}} (C_{ab} - \tilde{C}_{ab})^2 \Psi(\mathfrak{C}) \\ &= \sum_{\mathfrak{C}} ([C_{ab}^2] + C_{ab}) \Psi(\mathfrak{C}) - \tilde{C}_{ab}^2 \\ &= \frac{\partial^2 \Phi}{\partial z_{ab}^2}(\epsilon) + \tilde{C}_{ab} - \tilde{C}_{ab}^2 \quad (a, b = 1, \dots, m; a \leq b).\end{aligned}$$

Since the value of \tilde{C}_{ab} has been already determined, it is only necessary to substitute the value of a pure derivative of the second order which follows readily from a general formula established above.

Next, the covariance between C_{ab} and C_{cd} is given by

$$\begin{aligned}\text{cov}(C_{ab}, C_{cd}) &\equiv \text{cov}(C_{ab}, C_{cd} | \mathfrak{F}; \mathfrak{M}) = \sum_{\mathfrak{C}} (C_{ab} - \tilde{C}_{ab})(C_{cd} - \tilde{C}_{cd}) \Psi(\mathfrak{C}) \\ &= \sum_{\mathfrak{C}} C_{ab} C_{cd} \Psi(\mathfrak{C}) - \tilde{C}_{ab} \tilde{C}_{cd} = \frac{\partial^2 \Phi}{\partial z_{ab} \partial z_{cd}}(\epsilon) - \tilde{C}_{ab} \tilde{C}_{cd} \\ &\quad (a, b, c, d = 1, \dots, m; a \leq b; c \leq d; A_{ab} \neq A_{cd}).\end{aligned}$$

Since the values of \tilde{C}_{ab} and \tilde{C}_{cd} have been already determined, it is only necessary to substitute the value of a mixed derivative of the second order which also follows readily from a general formula established above.

Now, for the sake of completeness, we shall set out below the values of variance as well as covariance of the stochastic variables, after classified according to homozygous and heterozygous types, in more concrete forms.

In general, if the relevant variables in the bracket-notation are X_{ab} ($a, b = 1, \dots, m, a \leq b; X_{ab} = X_{ba}$), then we get

$$\begin{aligned}\left[\left(X_{ii} + \sum_{b \neq i} \frac{X_{ib}}{2} \right)^2 \right] &= \left(X_{ii} + \sum_{b \neq i} \frac{X_{ib}}{2} \right)^2 - \left(X_{ii} + \sum_{b \neq i} \frac{X_{ib}}{4} \right) \\ \left[\left(X_{ii} + \sum_{b \neq i} \frac{X_{ib}}{2} \right) \left(X_{jj} + \sum_{b \neq j} \frac{X_{jb}}{2} \right) \right] & \\ &= \left(X_{ii} + \sum_{b \neq i} \frac{X_{ib}}{2} \right) \left(X_{jj} + \sum_{b \neq j} \frac{X_{jb}}{2} \right) - \frac{X_{ij}}{4} \quad (i \neq j).\end{aligned}$$

Hence, introducing, as before, notations defined by

$$p_i^{(F)} = \frac{1}{N} \left(F_{ii} + \sum_{b \neq i} \frac{F_{ib}}{2} \right) \quad \text{and} \quad p_i^{(M)} = \frac{1}{N} \left(M_{ii} + \sum_{b \neq i} \frac{M_{ib}}{2} \right),$$

we obtain for the values of variance the following expressions:

$$\begin{aligned} \text{var}(C_{ii}) &= \frac{1}{N(N-1)} \left(N^2 p_i^{(F)2} - \frac{N p_i^{(F)} + F_{ii}}{2} \right) \left(N^2 p_i^{(M)2} - \frac{N p_i^{(M)} + M_{ii}}{2} \right) \\ &\quad + N p_i^{(F)} p_i^{(M)} - N^2 (p_i^{(F)} p_i^{(M)})^2 \\ &= \frac{N^2}{N-1} \left\{ p_i^{(F)} p_i^{(M)} \left(1 - \frac{p_i^{(F)} + p_i^{(M)}}{2} - \frac{1}{N} \right) - \frac{p_i^{(M)2}}{2} \left(\frac{F_{ii}}{N} - p_i^{(F)2} \right) \right. \\ &\quad \left. - \frac{p_i^{(F)2}}{2} \left(\frac{M_{ii}}{N} - p_i^{(M)2} \right) + \frac{1}{4N} \left(\frac{F_{ii}}{N} + p_i^{(F)} \right) \left(\frac{M_{ii}}{N} + p_i^{(M)} \right) \right\}, \\ \text{var}(C_{ij}) &= \frac{1}{N(N-1)} \left\{ \left(N^2 p_i^{(F)2} - \frac{N p_i^{(F)} + F_{ii}}{2} \right) \left(N^2 p_j^{(M)2} - \frac{N p_j^{(M)} + M_{jj}}{2} \right) \right. \\ &\quad + \left(N^2 p_j^{(F)2} - \frac{N p_j^{(F)} + F_{jj}}{2} \right) \left(N^2 p_i^{(M)2} - \frac{N p_i^{(M)} + M_{ii}}{2} \right) \\ &\quad + 2 \left(N^2 p_i^{(F)} p_j^{(F)} - \frac{F_{ij}}{4} \right) \left(N^2 p_i^{(M)} p_j^{(M)} - \frac{M_{ij}}{4} \right) \left. \right\} \\ &\quad + N (p_i^{(F)} p_j^{(M)} + p_j^{(F)} p_i^{(M)}) - N^2 (p_i^{(F)} p_j^{(M)} + p_j^{(F)} p_i^{(M)})^2 \\ &= \frac{N^2}{N-1} \left\{ p_i^{(F)} p_j^{(M)} \left(1 - \frac{p_i^{(F)} + p_j^{(M)}}{2} - \frac{1}{N} \right) \right. \\ &\quad + p_j^{(F)} p_i^{(M)} \left(1 - \frac{p_j^{(F)} + p_i^{(M)}}{2} - \frac{1}{N} \right) - \frac{1}{2} \left(p_j^{(M)2} \left(\frac{F_{ii}}{N} - p_i^{(F)2} \right) \right. \\ &\quad \left. + p_i^{(M)2} \left(\frac{F_{jj}}{N} - p_j^{(F)2} \right) + p_i^{(M)} p_j^{(M)} \left(\frac{F_{ij}}{N} - 2 p_i^{(F)} p_j^{(F)} \right) + p_i^{(M)2} \left(\frac{F_{jj}}{N} - p_j^{(F)2} \right) \right. \\ &\quad \left. - \frac{1}{2} \left(p_j^{(F)2} \left(\frac{M_{ii}}{N} - p_i^{(M)2} \right) + p_i^{(F)} p_j^{(F)} \left(\frac{M_{ij}}{N} - 2 p_i^{(M)} p_j^{(M)} \right) \right. \right. \\ &\quad \left. \left. + p_i^{(F)2} \left(\frac{M_{jj}}{N} - p_j^{(M)2} \right) \right) \right\} \end{aligned}$$

$$+ \frac{1}{4N} \left\{ \left(\frac{F_{ii}}{N} + p_i^{(F)} \right) \left(\frac{M_{jj}}{N} + p_j^{(M)} \right) + \frac{1}{2} \frac{F_{ij}}{N} \frac{M_{ij}}{N} \right. \\ \left. + \left(\frac{F_{jj}}{N} + p_j^{(M)} \right) \left(\frac{M_{ii}}{N} + p_i^{(M)} \right) \right\} \quad (i \neq j).$$

Similarly, we obtain for the values of covariance the following expressions:

$$\text{cov}(C_{ii}, C_{ik}) = \frac{1}{N(N-1)} \left\{ \left(N^2 p_i^{(F)2} - \frac{N p_i^{(F)} + F_{ii}}{2} \right) \left(N^2 p_i^{(M)} p_k^{(M)} - \frac{M_{ik}}{4} \right) \right. \\ \left. + \left(N^2 p_i^{(F)} p_k^{(F)} - \frac{F_{ik}}{4} \right) \left(N^2 p_i^{(M)2} - \frac{N p_i^{(M)} + M_{ii}}{2} \right) \right\} \\ - N^2 p_i^{(F)} p_i^{(M)} (p_i^{(F)} p_k^{(M)} + p_k^{(F)} p_i^{(M)}) \\ = \frac{N^2}{N-1} \left\{ p_i^{(F)} p_i^{(M)} (p_i^{(F)} p_k^{(M)} + p_k^{(F)} p_i^{(M)}) - \frac{p_i^{(M)}}{4} \left(2 p_k^{(M)} \left(\frac{F_{ii}}{N} + p_i^{(F)} \right) \right. \right. \\ \left. \left. + p_i^{(M)} \frac{F_{ik}}{N} \right) - \frac{p_i^{(F)}}{4} \left(2 p_k^{(F)} \left(\frac{M_{ii}}{2} + p_i^{(M)} \right) + p_i^{(F)} \frac{M_{ik}}{N} \right) \right. \\ \left. + \frac{1}{8N} \left\{ \left(\frac{F_{ii}}{N} + p_i^{(F)} \right) \frac{M_{ik}}{N} + \left(\frac{M_{ii}}{N} + p_i^{(M)} \right) \frac{F_{ik}}{N} \right\} \right\} \quad (k \neq i),$$

$$\text{cov}(C_{ii}, C_{kk}) = \frac{1}{N(N-1)} \left(N^2 p_i^{(F)} p_k^{(F)} - \frac{F_{ik}}{4} \right) \left(N^2 p_i^{(M)} p_k^{(M)} - \frac{M_{ik}}{4} \right) \\ - N^2 p_i^{(F)} p_i^{(M)} p_k^{(F)} p_k^{(M)} \\ = \frac{N^2}{N-1} \left\{ p_i^{(F)} p_k^{(F)} p_i^{(M)} p_k^{(M)} - \frac{1}{4} \left(p_i^{(M)} p_k^{(M)} \frac{F_{ik}}{N} + p_i^{(F)} p_k^{(F)} \frac{M_{ik}}{N} \right) \right. \\ \left. + \frac{1}{16N} \frac{F_{ik}}{N} \frac{M_{ik}}{N} \right\} \quad (k \neq i),$$

$$\text{cov}(C_{ii}, C_{hk}) = \frac{1}{N(N-1)} \left\{ N^2 p_i^{(F)} p_h^{(F)} - \frac{F_{ih}}{4} \right\} \left(N^2 p_i^{(M)} p_k^{(M)} - \frac{M_{ik}}{4} \right) \\ + \left(N^2 p_i^{(F)} p_k^{(F)} - \frac{F_{ik}}{4} \right) \left(N^2 p_i^{(M)} p_h^{(M)} - \frac{M_{ih}}{4} \right) \\ - N^2 p_i^{(F)} p_i^{(M)} (p_h^{(F)} p_k^{(M)} + p_k^{(F)} p_h^{(M)}) \\ = \frac{N^2}{N-1} \left\{ p_i^{(F)} p_i^{(M)} (p_h^{(F)} p_k^{(M)} + p_k^{(F)} p_h^{(M)}) - \frac{p_i^{(M)}}{4} \left(p_k^{(M)} \frac{F_{ih}}{N} + p_h^{(M)} \frac{F_{ik}}{N} \right) \right.$$

$$-\frac{p_i^{(F)}}{4} \left(p_k^{(F)} \frac{M_{ih}}{N} + p_h^{(F)} \frac{M_{ik}}{N} \right) + \frac{1}{16N} \left(\frac{F_{ih}}{N} \frac{M_{ik}}{N} + \frac{F_{ik}}{N} \frac{M_{ih}}{N} \right) \Big\} \\ (h, k \neq i; h \neq k),$$

$$\begin{aligned} \text{cov}(C_{ij}, C_{ik}) &= \frac{1}{N(N-1)} \left\{ \left(N^2 p_i^{(F)2} - \frac{N p_i^{(F)} + F_{ii}}{2} \right) \left(N^2 p_j^{(M)} p_k^{(M)} - \frac{M_{jk}}{4} \right) \right. \\ &\quad + \left(N^2 p_j^{(F)} p_k^{(F)} - \frac{F_{jk}}{4} \right) \left(N^2 p_i^{(M)2} - \frac{N p_i^{(M)} + M_{ii}}{2} \right) \\ &\quad + \left(N^2 p_i^{(F)} p_k^{(F)} - \frac{F_{ik}}{4} \right) \left(N^2 p_i^{(M)} p_j^{(M)} - \frac{M_{ij}}{4} \right) \\ &\quad + \left. \left(N^2 p_i^{(F)} p_j^{(F)} - \frac{F_{ij}}{4} \right) \left(N^2 p_i^{(M)} p_k^{(M)} - \frac{M_{ik}}{4} \right) \right\} \\ &\quad - N^2 (p_i^{(F)} p_j^{(M)} + p_j^{(F)} p_i^{(M)}) (p_i^{(F)} p_k^{(M)} + p_k^{(F)} p_i^{(M)}) \\ &= \frac{N^2}{N-1} \left\{ (p_i^{(F)} p_j^{(M)} + p_j^{(F)} p_i^{(M)}) (p_i^{(F)} p_k^{(M)} + p_k^{(F)} p_i^{(M)}) \right. \\ &\quad - \frac{p_j^{(M)}}{4} \left(2 p_k^{(M)} \left(\frac{F_{ii}}{N} + p_i^{(F)} \right) + p_i^{(M)} \frac{F_{ik}}{N} \right) - \frac{p_j^{(F)}}{4} \left(2 p_k^{(F)} \left(\frac{M_{ii}}{N} + p_i^{(M)} \right) \right. \\ &\quad + \left. p_i^{(M)} \frac{M_{ik}}{N} \right) \Big) - \frac{p_i^{(M)}}{4} \left(p_k^{(M)} \frac{F_{ij}}{N} + p_i^{(M)} \frac{F_{jk}}{N} \right) \\ &\quad - \frac{p_i^{(F)}}{4} \left(p_k^{(F)} \frac{M_{ij}}{N} + p_i^{(F)} \frac{M_{jk}}{N} \right) + \frac{1}{16N} \left(2 \left(\frac{F_{ii}}{N} + p_i^{(F)} \right) \frac{M_{jk}}{N} \right. \\ &\quad + \left. 2 \left(\frac{M_{ii}}{N} + p_i^{(M)} \right) \frac{F_{jk}}{N} + \frac{F_{ij}}{N} \frac{M_{ik}}{N} + \frac{F_{ik}}{N} \frac{M_{ij}}{N} \right) \Big\} \quad (i \neq j; k \neq i, j), \end{aligned}$$

$$\begin{aligned} \text{cov}(C_{ij}, C_{hk}) &= \frac{1}{N(N-1)} \left\{ \left(N^2 p_i^{(F)} p_h^{(F)} - \frac{F_{ih}}{4} \right) \left(N^2 p_j^{(M)} p_k^{(M)} - \frac{M_{jk}}{4} \right) \right. \\ &\quad + \left(N^2 p_i^{(F)} p_k^{(F)} - \frac{F_{ik}}{4} \right) \left(N^2 p_j^{(M)} p_h^{(M)} - \frac{M_{jh}}{4} \right) \\ &\quad + \left(N^2 p_j^{(F)} p_h^{(F)} - \frac{F_{jh}}{4} \right) \left(N^2 p_i^{(M)} p_k^{(M)} - \frac{M_{ik}}{4} \right) \\ &\quad + \left. \left(N^2 p_j^{(F)} p_k^{(F)} - \frac{F_{jk}}{4} \right) \left(N^2 p_i^{(M)} p_h^{(M)} - \frac{M_{ih}}{4} \right) \right\} \end{aligned}$$

$$\begin{aligned}
 & -N^2(p_i^{(F)}p_j^{(M)} + p_j^{(F)}p_i^{(M)})(p_h^{(F)}p_k^{(M)} + p_k^{(F)}p_h^{(M)}) \\
 = & \frac{N}{N-1} \left\{ (p_i^{(F)}p_j^{(M)} + p_j^{(F)}p_i^{(M)})(p_h^{(F)}p_k^{(M)} + p_k^{(F)}p_h^{(M)}) \right. \\
 & + p_j^{(M)} \left(p_k^{(M)} \frac{F_{ih}}{N} + p_h^{(M)} \frac{F_{ik}}{N} \right) + p_i^{(M)} \left(p_k^{(M)} \frac{F_{jh}}{N} + p_h^{(M)} \frac{F_{jk}}{N} \right) \\
 & + p_j^{(F)} \left(p_k^{(F)} \frac{M_{ih}}{N} + p_h^{(F)} \frac{M_{ik}}{N} \right) + p_i^{(F)} \left(p_k^{(F)} \frac{M_{jh}}{N} + p_h^{(F)} \frac{M_{jk}}{N} \right) \\
 & \left. + \frac{1}{16N} \left(\frac{F_{ih}}{N} \frac{M_{jk}}{N} + \frac{F_{ik}}{N} \frac{M_{jh}}{N} + \frac{F_{jh}}{N} \frac{M_{ik}}{N} + \frac{F_{jk}}{N} \frac{M_{ih}}{N} \right) \right\} \\
 & (i \neq j; h \neq k; h, k \neq i, j).
 \end{aligned}$$

It would be noted, in passing, that there holds an identical relation of dependence, i. e.

$$0 = \sum_{\mathfrak{C}} \left(\sum_{a \leq b} (C_{ab} - \tilde{C}_{ab}) \right)^2 \Psi(\mathfrak{C}) = \sum_{a \leq b} \text{var}(C_{ab}) + 2 \sum_{a \leq b; c \leq d} \text{cov}(C_{ab}, C_{cd});$$

the last summation extends over all the possible pairs (ab, cd) .

Asymptotic behaviors of variance and covariance as N increases can be readily deduced from the expressions just established. In fact, since there hold always $F_{ab}/N \leq 1$ and $M_{ab}/N \leq 1$, we obtain

$$\begin{aligned}
 \text{var}(C_{ii}) = & N \left\{ p_i^{(F)} p_i^{(M)} \left(1 - \frac{p_i^{(F)} + p_i^{(M)}}{2} \right) \right. \\
 & \left. - \frac{p_i^{(M)2}}{2} \left(\frac{F_{ii}}{N} - p_i^{(F)2} \right) - \frac{p_i^{(F)2}}{2} \left(\frac{M_{ii}}{N} - p_i^{(M)2} \right) \right\} + O(1), \text{ etc.}
 \end{aligned}$$

On the other hand, if the original distributions \mathfrak{F} and \mathfrak{M} show, in particular, the same *equilibrium state*, i. e. when there hold

$$F_{bb} = M_{bb} = Np_b^2, \quad F_{ab} = M_{ab} = N2p_a p_b \quad (a, b = 1, \dots, m; a < b),$$

the expressions may be reduced to fairly simple forms. Namely, we then get

$$\text{var}(C_{ii}) = \frac{N^2}{N-1} p_i^2 (1-p_i) \left(1 - \frac{3+p_i}{3N} \right),$$

$$\text{var}(C_{ij}) = \frac{N^2}{N-1} p_i p_j \left(2 - p_i - p_j - \frac{3 - p_i - p_j - 2p_i p_j}{2N} \right) \quad (i \neq j);$$

$$\text{cov}(C_{ii}, C_{ik}) = -\frac{N^2}{N-1} p_i^2 p_k \left(1 - \frac{1 + p_i}{2N} \right) \quad (k \neq i),$$

$$\text{cov}(C_{ii}, C_{kk}) = \frac{1}{4} \frac{N}{N-1} p_i^2 p_k^2 \quad (k \neq i),$$

$$\text{cov}(C_{ii}, C_{hk}) = \frac{1}{2} \frac{N}{N-1} p_i^2 p_h p_k \quad (h, k \neq i; h \neq k),$$

$$\text{cov}(C_{ij}, C_{ik}) = -\frac{N^2}{N-1} p_i p_j p_k \left(1 - 2p_i - \frac{1 + 2p_i}{2N} \right) \quad (i \neq j; k \neq i, j),$$

$$\text{cov}(C_{ij}, C_{hk}) = \frac{N}{N-1} p_i p_j p_h p_k \quad (i \neq j; h, k \neq i, j; h \neq k).$$

Thus, in this particular case, $\text{cov}(C_{ii}, C_{kk})$, $\text{cov}(C_{ii}, C_{hk})$ and $\text{cov}(C_{ij}, C_{hk})$ are really of order $O(1)$, i. e. bounded regardless of the values of $N > 1$, while the remaining cov's as well as the var's are, in general, unbounded with order $O(N)$.

Department of Mathematics,
Tokyo Institute of Technology.