135. Probabilities on Inheritance in Consanguineous Families. XIV

By Yûsaku KOMATU and Han NISHIMIYA

Department of Mathematics, Tokyo Institute of Technology

(Comm. by T. FURUHATA, M.J.A., Oct. 12, 1955)

XI. General consanguineous lineages

1. Preliminaries

We have considered in the preceding chapter[1] several probabilities on lineages in which consanguineous marriages consist merely of a serial sequence of single ones. In order to generalize them, we first classify all the possible lineages according to a number of consanguineous marriages involved. Lineages of every class are further divided into subclasses according to a manner of intervention of consanguineous marriages.

In general, we designate by a *branching* or *converging position* a position where collateral lines arise or a consanguineous marriage occurs, and by a *critical position* either of them. A lineal part of a lineage will be called a *maximal chain* if it connects two consecutive critical positions. In the following, we suppose that at every branching position there arises a separation into just *two* collateral lines, and that a *panmixia* takes place at any stage of generations. Members at non-critical position as well as spouses coming from other lineages even at branching positions are indifferent to the classification. However, if a number of non-critical members involved in a maximal chain reduces to unity or degenerates to empty, it will be required to make considerably troublesome modifications corresponding to those performed in chapter VI et seq. For the sake of simplicity, we shall restrict ourselves, in principle, to a generic case by omitting these extreme cases.

We first consider a mother-descendants combination $(\mathbf{0}; \mathbf{1}, \mathbf{2})\ldots$, in which $\mathbf{0}$ is a branching member. In a lineage involving t consanguineous marriages, there exist $t+1$ branching positions, t converging positions, and hence $2t+1$ critical positions. Distinguishing all such lineages by means of a manner on consanguineous marriages,

1) Previous parts of the paper have been published in these Proceedings: Y. Komatu and H. Nishimiya, Probabilities on inheritance in consanguineous families. I–XIII. Proc. Japan Acad. **30** (1954), 42–45; 46–48; 49–52; 148–151; 152–155; 236–240; 241–244; 245–247; 636–640; 641–649; 650–654; **31** (1955), 186–189; 190–194. Cf. also a supplementary paper: A remark concerning probabilities on inheritance in consanguineous families. Proc. Japan Acad. **31** (1955), 380–381.

we designate by $M(t)$ the number of different subclasses so obtained. We have, for instance, $M(0)=1$, $M(1)=3$, $M(2)=21$.

Once the problem on mother-descendants combination having been fully dealt with, related problems, e. g. on mother-descendant or descendants combination will be readily derived by means of routine methods.

2. Mother-descendants combination

The reduced probability of *mother-descendants combination* of the Jth subclass, $\kappa\ldots(\mathbf{0};\mathbf{1},\mathbf{2})$, among $M(t)$ subclasses will be designated by

$$K_t^J(\mathbf{0};\mathbf{1},\mathbf{2}) \qquad (J=I, II,\ldots, M(t)).$$

Our object is then to derive an explicit expression for this probability. However, for economy reason of the space, we shall state here only the final result. Detailed discussions, especially the proof of the formula, will be performed fully in another paper.[2]

Now, *our main formula is expressed in the form*

$$K_t^J(\mathbf{0};\mathbf{1},\mathbf{2})=\bar{A}_1\bar{A}_2+[\boldsymbol{P}_2^{(1)}]_t^J\bar{A}_2Q(\mathbf{0};\mathbf{1})+[\boldsymbol{P}_1^{(2)}]_t^J\bar{A}_1Q(\mathbf{0};\mathbf{2})$$
$$+[\boldsymbol{T}]_t^J T(\mathbf{0};\mathbf{1},\mathbf{2})+[\boldsymbol{S}]_t^J S(\mathbf{0};\mathbf{1},\mathbf{2})+[\boldsymbol{O}]_t^J\{\bar{A}_1Q(\mathbf{1};\mathbf{2})+\bar{A}_2Q(\mathbf{2};\mathbf{1})\}$$
$$(J=I, II,\ldots, M(t);\ t=0,1,2,\ldots),$$

$\bar{A}_1$ and $\bar{A}_2$ designating the relative frequencies of the genotypes of $\mathbf{1}$ and $\mathbf{2}$, respectively.

In order to write down explicit expressions for several coefficients contained in the formula, we consider all the possible paths steadily descending from $\mathbf{0}$ to $\mathbf{1}$ and to $\mathbf{2}$, which will be designated by

$$\mathbf{1}_u\ (u=1,\ldots, U_J) \quad \text{and} \quad \mathbf{2}_v\ (v=1,\ldots, V_J),$$

respectively. The length of a path, i. e. the total sum of generation-numbers involved in a path, exclusive of initial member but inclusive of terminal member, will be designated by the same notation as the path itself. Further, for a branching position Ω lying on a path $\mathbf{1}_u$ or $\mathbf{2}_v$, all the possible subpaths which are obtained by rejecting the parts antecedent to Ω will be designated, together with their lengths, by

$$\mathbf{1}_u(\Omega)\ (u=1,\ldots, U_J) \quad \text{or} \quad \mathbf{2}_v(\Omega)\ (v=1,\ldots, V_J),$$

respectively, some among them being eventually vacuous or coincident. Every path and every subpath is supposed to be closed at its both ends, though the length is to be counted by excluding its starting member.

2) To be published in Bull. Tokyo Inst. Tech. Detailed discussions for the previous parts are found also in this Bulletin: Y. Komatu and H. Nishimiya, Probabilistic investigations on inheritance in consanguineous families. Bull. Tokyo Inst. Tech. Nos. 1, 2, 3 (1954), 1–66, 67–152, 153–222; No. 2 (1955), in press.

By making use of these notations, *the coefficients contained in the main formula are expressed in the form*:

$$[\boldsymbol{P}_2^{(1)}]_t^J = 2\sum_{u=1}^{U_J} 2^{-\mathbf{1}_u}, \qquad [\boldsymbol{P}_1^{(2)}]_t^J = 2\sum_{v=1}^{V_J} 2^{-\mathbf{2}_v},$$

$$[\boldsymbol{T}]_t^J = 2 \sum_{\mathbf{1}_u \cap \mathbf{2}_v = \mathbf{0}} 2^{-\mathbf{1}_u - \mathbf{2}_v},$$

$$[\boldsymbol{S}]_t^J = 4\sum_{\Omega} \sum_{\substack{\mathbf{1}_u(\Omega)\cap \mathbf{2}_v(\Omega)=\Omega \\ \mathbf{1}_u - \mathbf{1}_u(\Omega) = \mathbf{2}_v - \mathbf{2}_v(\Omega) = \Omega(\mathbf{0})}} 2^{-\Omega(\mathbf{0}) - \mathbf{1}_u(\Omega) - \mathbf{2}_v(\Omega)},$$

$$[\boldsymbol{O}]_t^J = 2\sum_{\Omega} \sum_{\substack{\mathbf{1}_u(\Omega)\cap \mathbf{2}_v(\Omega)=\Omega \\ \mathbf{1}_u - \mathbf{1}_u(\Omega) = \mathbf{2}_v - \mathbf{2}_v(\Omega)}} 2^{-\mathbf{1}_u(\Omega) - \mathbf{2}_v(\Omega)}.$$

Here the summation affixed by Ω extends over all the branching positions except $\mathbf{0}$, and the summation affixed by $\mathbf{1}_u \cap \mathbf{2}_v = \mathbf{0}$ over all the possible pairs of paths $\mathbf{1}$ and $\mathbf{2}$ disjoint after $\mathbf{0}$, while the summation affixed by $\mathbf{1}_u(\Omega)\cap\mathbf{2}_v(\Omega)=\Omega$ and $\mathbf{1}_u - \mathbf{1}_u(\Omega) = \mathbf{2}_v - \mathbf{2}_v(\Omega) = \Omega(\mathbf{0})$ extends over all the possible pairs of paths $\mathbf{1}_u$ and $\mathbf{2}_v$ common before Ω and disjoint after Ω, $\Omega(\mathbf{0})$ being the length of the common part.

3. Mother-descendant combination

A mother-descendants combination yields, by convoluting ε_n from below, a corresponding *mother-descendant combination*, of which the reduced probability will be designated by

$$K_{t;n}^J(\mathbf{0};\mathbf{1}) = \sum_{p,q} K_t^J(\mathbf{0};p,q)\,\varepsilon_n(p,q;\mathbf{1}) \qquad (J=I, II, \ldots, M(t)),$$

the number of interjacent consanguineous marriages being equal to $t+1$.

We suppose here also that the generation-numbers concerned, except n, are greater than unity. It is then shown that *the final result may be written in the form*

$$K_{t;1}^J(\mathbf{0};\mathbf{1}) = \bar{A}_1 + \tfrac{1}{2}\{[\boldsymbol{P}_2^{(1)}]_t^J + [\boldsymbol{P}_1^{(2)}]_t^J\}\,Q(\mathbf{0};\mathbf{1}) + [\boldsymbol{T}]_t^J T(\mathbf{0};\mathbf{1}) + \tfrac{1}{2}[\boldsymbol{S}]_t^J S(\mathbf{0};\mathbf{1}) + [\boldsymbol{O}]_t^J R(\mathbf{1}),$$

$$K_{t;n}^J(\mathbf{0};\mathbf{1}) = \bar{A}_1 + 2^{-n}\{[\boldsymbol{P}_2^{(1)}]_t^J + [\boldsymbol{P}_1^{(2)}]_t^J\}\,Q(\mathbf{0};\mathbf{1}) \qquad (n>1).$$

The coefficients are constructed in a manner defined in the last section with respect to $\mathbf{0}$ and the last converging members. In particular, the coefficient of $Q(\mathbf{0};\mathbf{1})$ is equal to the sum of the terms $2\times 2^{-\mathbf{1}_w}$ where $\mathbf{1}_w$ extends over the sum of generation-numbers in every possible path numbered with $w=1,\ldots, W_t$, which connects descendingly the initial member $\mathbf{0}$ with the terminal member $\mathbf{1}$, exclusive of the former but inclusive of the latter. This fact represents a generalization of a result noticed in chapter V, §4 for a particular case $t=0$. However, it can be shown that the range of its validity is far extensive, as will be mentioned later.

4. Related combinations

Mother-descendants as well as mother-descendant combinations having been dealt with, results on several related combinations can be obtained by simple operations. First, by eliminating mother's type from a mother-descendants combination, we get the corresponding *descendants combination*, of which *the reduced probability is shown to be expressible in the form*

$$\begin{aligned}\textstyle\sum_t^J(\mathbf{1}, \mathbf{2}) &\equiv \sum_e \bar{A}_e K_t^J(e; \mathbf{1}, \mathbf{2}) \\ &= \bar{A}_1\bar{A}_2 + \{[\boldsymbol{T}]_t^J + [\boldsymbol{O}]_t^J\}\{\bar{A}_1 Q(\mathbf{1}; \mathbf{2}) + \bar{A}_2 Q(\mathbf{2}; \mathbf{1})\};\end{aligned}$$

the coefficient of the second term in the last expression is, in its fully explicit form, given by

$$[\boldsymbol{T}]_t^J + [\boldsymbol{O}]_t^J = 2\sum_Z \sum_{\substack{\mathbf{1}_u(Z)\cap \mathbf{2}_v(Z) = Z \\ \mathbf{1}_u - \mathbf{1}_u(Z) = \mathbf{2}_v - \mathbf{2}_v(Z)}} 2^{-\mathbf{1}_u(Z) - \mathbf{2}_v(Z)},$$

where the summation affixed by Z extends over all the branching positions, inclusive of $\mathbf{0}$.

Next, by eliminating mother's type from a mother-descendant combination, we get a *distribution of genotypes* in the corresponding descendant's generation. Thus, *the probabilistic frequency of genotype A_1 in the generation designated by $(\mathbf{1})_{t;n}^J$ is given by*

$$\begin{aligned}\bar{A}_{t;n}^J(\mathbf{1}) &\equiv \sum_e \bar{A}_e K_{t;n}^J(e; \mathbf{1}) \\ &= \begin{cases}\bar{A}_1 + \{[\boldsymbol{T}]_t^J + [\boldsymbol{O}]_t^J\} R(\mathbf{1}) & \text{for } n=1, \\ \bar{A}_1 & \text{for } n>1,\end{cases}\end{aligned}$$

the coefficient being constructed with respect to the member $\mathbf{0}$ and the last converging members. There appears a deviation from equilibrium in case $n=1$.

On the other hand, it is also easy to derive probabilities of several combinations in which a collateral separation takes place at an intermediate, not original, generation. In fact, we get by means of readily comprehensible notations, the following results:

$$\begin{aligned}K_{l|t}^J(\mathbf{0}; \mathbf{1}, \mathbf{2}) &\equiv \sum_e \kappa_l(\mathbf{0}; e) K_t^J(e; \mathbf{1}, \mathbf{2}) \\ = \bar{A}_1\bar{A}_2 &+ 2^{-l}\{[\boldsymbol{P}_2^{(1)}]_t^J \bar{A}_2 Q(\mathbf{0}; \mathbf{1}) + [\boldsymbol{P}_1^{(2)}]_t^J \bar{A}_1 Q(\mathbf{0}; \mathbf{2})\} \\ &+ \{[\boldsymbol{T}]_t^J + [\boldsymbol{O}]_t^J\}\{\bar{A}_1 Q(\mathbf{1}; \mathbf{2}) + \bar{A}_2 Q(\mathbf{2}; \mathbf{1})\} + 2^{-l}\{[\boldsymbol{T}]_t^J + [\boldsymbol{S}]_t^J\} S(\mathbf{0}; \mathbf{1}, \mathbf{2});\end{aligned}$$

$$\begin{aligned}K_{l|t;1}^J(\mathbf{0}; \mathbf{1}) = \bar{A}_1 &+ 2^{-l-1}\{[\boldsymbol{P}_2^{(1)}]_t^J + [\boldsymbol{P}_1^{(2)}]_t^J\} Q(\mathbf{0}; \mathbf{1}) \\ &+ \{[\boldsymbol{T}]_t^J + [\boldsymbol{O}]_t^J\} R(\mathbf{1}) + 2^{-l}\{[\boldsymbol{T}]_t^J + [\boldsymbol{S}]_t^J\} S(\mathbf{0}; \mathbf{1}),\end{aligned}$$

$$K_{l|t;n}^J(\mathbf{0}; \mathbf{1}) = \bar{A}_1 + 2^{-l-n}\{[\boldsymbol{P}_2^{(1)}]_t^J + [\boldsymbol{P}_1^{(2)}]_t^J\} Q(\mathbf{0}; \mathbf{1}) \qquad (n>1);$$

$$\textstyle\sum_{l|t}^J(\mathbf{1}, \mathbf{2}) = \bar{A}_1\bar{A}_2 + \{[\boldsymbol{T}]_t^J + [\boldsymbol{O}]_t^J\}\{\bar{A}_1 Q(\mathbf{1}; \mathbf{2}) + \bar{A}_2 Q(\mathbf{2}; \mathbf{1})\};$$

$$\bar{\boldsymbol{A}}_{l|t;n}^J(\mathbf{1}) = \begin{cases}\bar{A}_1 + \{[\boldsymbol{T}]_t^J + [\boldsymbol{O}]_t^J\} R(\mathbf{1}) & (n=1), \\ \bar{A}_1 & (n>1).\end{cases}$$

Here, it is supposed that all the generation-numbers concerned, except l and n, are all greater than unity. Further, in all these formulas, every coefficient is to be constructed in a manner defined in §2 with respect to the first branching member and the last converging members. It is noted that the last two quantities $\sum_{l|t}^{\prime}(\mathbf{1},\mathbf{2})$ and $\bar{\mathbf{A}}_{l|t;n}^{\prime}(\mathbf{1})$ are independent of the generation-number l.