105. Probability-theoretic Investigations on Inheritance. III₄. Further Discussions on Cross-Breeding.

(Further and Ultimate Continuation.)

By Yûsaku KOMATU.

Department of Mathematics, Tokyo Institute of Technology and
Department of Legal Medicine, Tokyo University.

(Comm. by T. FURUHATA, M.J.A., Oct. 12, 1951.)

5. Continuous process.

In preceding sections we have considered a stepwise process of cross-breeding and determined an explicit form of general formulae representing distribution in nth generation. In the present section we shall idealize the process such that it proceeds continuously with the lapse of the time. According to such a manner of treatment, the sums or differences of various kinds in the previous discussion will then be replaced by the corresponding integrals or derivatives. It will also be recognized that several formulae will be considerably clarified by such an idealization.

We have considered, in §3, the general class $X'^{2^n-u}X''^u$ ($0 \leq u \leq 2^n$) appearing in the nth generation, which possesses the frequency of the gene given by (3.10), namely by

$$2^{-n}((2^n-u)\,p_i'+u\,p_i'') = (1-2^{-n}u)\,p_i'+2^{-n}u\,p_i'' \quad (i=1,\ldots,m).$$

When the discrete variable n restricted to integral values runs from 0 to 2^n, this frequency varies from p_i' to p_i''. Correspondingly, we introduce a frequency of the gene A_i given by

$$(5.1) \qquad p_i(x) \equiv (1-x)\,p_i'+x\,p_i'' \qquad (i=1,\ldots,m),$$

depending linearly on a continuous variable x running over the closed interval $0 \leq x \leq 1$, where $\{p_i'\}$ and $\{p_i''\}$ denote two given fixed distributions; x being a quantity which corresponds to $2^{-n}u$ in the previous discrete case. Evidently, it holds always

$$\sum_{i=1}^{m} p_i(x) = (1-x) \sum_{i=1}^{m} p_i'+x \sum_{i=1}^{m} p_i'' = (1-x)+x = 1.$$

The frequency $p_i(x)$ is also contained in the interval between p_i' and p_i'', both ends inclusive.

We further introduce the variable t representing the time, and denote by $\delta(x;t)$ the population-density at the time t. Supposing, for the sake of brevity, that the total population remains constant, it will be expressed by

$$(5.2) \qquad \varDelta = \int_0^1 \delta(x;t)\,dx.$$

While we could consider a more general case where the total population varies with t, our assumption represents that $\varDelta$ in (5.2) is independent t. We now define the relative density by

(5.3) $\rho(x;t) = \delta(x;t)/\varDelta$ $(0 \leq x \leq 1 ;\ 0 \leq t)$.

Because of (5.2), we then get

(5.4) $\int_0^1 \rho(x;t)\,dx = 1$

for any t.

We denote by $\chi(x, y; t)$ the *mating-velocity* between the individuals belonging to the classes with distribution $p_i(x)$ and with $p_i(y)$; namely, let the matings between such individuals, within the time-interval from t to $t+dt$, take place with the density given by $\chi(x, y; t)\,dt$. More precisely, let the number of matings between the individuals belonging to the classes with distributions from $p_i(x)$ to $p_i(x+dx)$ and from $p_i(y)$ to $p_i(y+dy)$ be

$$\chi(x, y;\ t)\ dx\,dy\,dt\,.$$

We further introduce the mating-velocity of population-density expressed by

(5.5) $\varphi(x, y; t) = \chi(x, y; t)/\varDelta$ $(0 \leq x,\ y \leq 1 ;\ 0 \leq t)$.

We put, for convenience, $\chi(x, y; t) = 0$ outside $0 \leq x,\ y \leq 1$, and then correspondingly

(5.6) $\varphi(x, y; t) = 0$ (outside $0 \leq x,\ y \leq 1$).

In view of the definition we remember the symmetry character $\chi(x, y; t) = \chi(y, x; t)$, whence it follows also that the symmetry relation

(5.7) $\varphi(x, y; t) = \varphi(y, x; t)$

is valid for all possible values of x, y and of t.

We first conclude, because of the definition, that the increase of the relative density (5.3) for fixed x is expressed in the form

$$\rho(x; t+dt) - \rho(x; t) = -\int_{y=0}^{y=1} \varphi(x, y; t)\,dt\,dy + 2\int_{y+z=2x} \varphi(y, z; t)\,dt\,dy$$

$$= dt\int_0^1 (-\varphi(x, y; t) + 2\varphi(y, 2x-y; t))\,dy\,.$$

Here the convention (5.6) is used; in particular, it may be understood that

$$\int_0^1 \varphi(y, 2x-y; t)\,dy = \begin{cases} \int_0^{2x} \varphi(y, 2x-y; t)\,dy & (2x \leq 1)\,, \\[2mm] \int_{2x-1}^1 \varphi(y, 2x-y; t)\,dy & (2x > 1)\,. \end{cases}$$

The above equation yields immediately a *fundamental relation*

$$(5.8) \qquad \frac{\partial \rho(x;t)}{\partial t} = \int_0^1 (-\varphi(x,y;t) + 2\varphi(y,2x-y;t))\, dy\,,$$

and, by integration,

$$(5.9) \qquad \rho(x;t) = \rho(x;0) + \int_0^t dt \int_0^1 (-\varphi(x,y;t) + 2\varphi(y,2x-y;t))\, dy\,.$$

Now, the frequency of the gene A_i over the whole population is evidently given by

$$(5.10) \qquad p_i = \int_0^1 p_i(x)\, \rho(x;t)\, dx\,.$$

It will be expected, based on our above mentioned assumption, that this quantity expresses the frequency of A_i in the limit distribution and is consequently *independent of t*. The fact of this independency may indeed also be shown analytically in the following manner. In fact, differentiation of the expression in the right-hand side of (5.10) leads us, by taking (5.8) into account, to

$$\frac{d}{dt}\int_0^1 p_i(x)\, \rho(x;t)\, dx = \int_0^1 p_i(x)\frac{\partial \rho(x;t)}{dt}dx$$

$$= \int_0^1\!\!\int_0^1 p_i(x)(-\varphi(x,y;t)+2\varphi(y,2x-y;t))\, dx\, dy$$

$$= -\int_0^1\!\!\int_0^1 p_i(x)\, \varphi(x,y;t)\, dx\, dy$$

$$+ \int_0^1\!\!\int_{-y}^{2-y} p_i\!\left(\frac{y+z}{2}\right)\varphi(y,z;t)\, dz\, dy\,.$$

Replacing the letter z for integration variable by x everywhere in the last integral, and remembering the convention (5.6) and the symmetry relation (5.7), we further get

$$\frac{d}{dt}\int_0^1 p_i(x)\, \rho(x;t)\, dx = \int_0^1\!\!\int_0^1 \left(-p_i(x) + p_i\!\left(\frac{x+y}{2}\right)\right)\varphi(x,y;t)\, dx\, dy$$

$$= \frac{p_i' - p_i''}{2}\int_0^1\!\!\int_0^1 (x-y)\, \varphi(x,y;t)\, dx\, dy\,.$$

By using now again the symmetry relation (5.7), it will be easily verified that the last integral evidently vanishes, yielding the required independency of the expression in the right-hand side on t. The now proved fact corresponds just to the last statement of §3.

We substitute for $p_i(x)$ its expression (5.1), i.e.,

$$(5.11) \qquad p_i(x) = p_i' - (p_i' - p_i'')\, x\,,$$

in the right-hand member of (5.10). In view of (5.4), it becomes then

$$(5.12) \qquad p_i = p_i' - (p_i' - p_i'')\int_0^1 x\, \rho(x;t)\, dx\,.$$

Since, as just noticed, p_i is independent t, so is also the last integral. Consequently, we get immediately a useful relation :

$$(5.13) \qquad \int_0^1 x(\rho\,(x;\,t) - \rho(x;\,0))\,dx = 0 \,.$$

Denoting, in general, the frequencies of homozygote A_{ii} and heterozygote $A_{ij}(i \neq j)$ at t by $P_{ii}(t)$ and $P_{ij}(t)$ respectively, we then get

$$P_{ii}(t) = \int_0^1 p_i(x)^2\, \rho(x;\,t)\,dx \,,$$

$$(5.14) \qquad\qquad\qquad\qquad (i, j = 1, \ldots, m;\ i < j) \,,$$

$$P_{ij}(t) = 2\int_0^1 p_i(x)\, p_j(x)\, \rho(x;\,t)\,dx$$

the expressions which are again after substituting (5.11) and remembering (5.4), written also in the form

$$P_{ii}(t) = p_i'^2 - 2(p_i' - p_i'')\int_0^1 x\rho(x;\,t)dx + (p_i' - p_i'')^2\int_0^1 x^2\rho(x;\,t)dx \,,$$

$$P_{ij}(t) = 2p_i'\,p_j' - 2(2p_i'\,p_j' - p_i'p_j'' - p_i''p_j')\int_0^1 x\rho(x;\,t)dx$$

$$+ 2(p_i' - p_i'')\,(p_j' - p_j'')\int_0^1 x^2\rho(x;\,t)dx \qquad\qquad (i \neq j) \,.$$

In view of (5.13), we deduce the result

$$P_{ii}(t) - P_{ii}(0) = -(p_i' - p_i'')^2\, \Gamma(t) \,,$$

$$(5.15) \qquad\qquad\qquad\qquad (i, j = 1, \ldots, m;\ i < j) \,,$$

$$P_{ij}(t) - P_{ij}(0) = -2(p_i' - p_i'')(p_j' - p_j'')\, \Gamma(t)$$

by introducing the expression

$$(5.16) \qquad \Gamma(t) = \int_0^1 x^2\, (\rho(x;\,0) - \rho(x;\,t))\,dx \,,$$

where the function $\rho(x;\,t)$ is given by the expression (5.9). It is evidently seen that this result corresponds to the main result in the discrete case; namely, (5.15) corresponds to (2.30) and $\Gamma(t)$ to $\Gamma^{(n)}$ in (2.31).

The above discussion thus shows that, if we consider a continuous process, the main result can be obtained in quite a clear from. However, it is not regarded as an immediate generalization of the previous result on discrete case. But, in order to establish such a generalization, we have only to introduce the concept of *Stieltjes integral* instead of that of ordinary integral. For instance, instead of making use of the relative density $\rho(x;\,t)$ (≥ 0) in (5.3), we introduce a monotone increasing set-function

$$(5.17) \qquad\qquad \sigma(x;\,t) \qquad (0 \leq x \leq 1;\ 0 \leq t)$$

which, in case of non-existence of the density-function $\rho(x; t)$, will be used as a substitute and a generalization of the indefinite integral

$$\int_0^x \rho(x; t)\, dx .$$

This set-function is supposed to satisfy the identity

$$(5.18) \qquad \int_0^1 d\sigma(x; t) \equiv \int_{x=-0}^{x=1+0} d\sigma(x; t) = 1$$

corresponding to (5.4), where $d\sigma(x; t)$ denotes of course the differential with respect to the variable x; the eventual jumps appearing at both ends must be taken into account, as emphasized explicitly in the second member in (5.18).

Under these circumstances, our main result (5.15) and (5.16) will take the following form:

$$P_{ii}(t) - P_{ii}(0) = -(p_i' - p_i'')^2 \Gamma'(t) ,$$
$$(5.19) \qquad\qquad\qquad\qquad (i, j = 1, \ldots, m ; i < j)$$
$$P_{ij}(t) - P_{ij}(0) = -2(p_i' - p_i'')(p_j' - p_j'')\Gamma'(t)$$

and (instead of (5.16))

$$(5.20) \qquad \Gamma(t) = \int_0^1 x^2 d(\sigma(x; 0) - \sigma(x; t)) .$$

Now, the result corresponding to (4.1) is quite evident already from (5.15); namely, we get

$$(5.21) \qquad P_{ij}(t) = P_{ij}(0) \quad (i, j = 1, \ldots, m ; 0 \leqq t) ,$$

provided only $p_i' = p_i''$ for some i. But, if $\Gamma(t)$ has been generalized into the Stieltjes from (5.20), the relation (4.3) in discrete process can also be deduced as a special case of our present result. In fact, we now suppose especially that the initial distribution is represented by a function $\sigma(x; 0)$ such that

$$[\sigma(x; 0)]_{x=-0}^{x=+0} = \lambda' , \qquad [\sigma(x; 0)]_{x=1-0}^{x=1+0} = \lambda'' ,$$

$$\sigma(x; 0) = \lambda' \qquad (0 < x < 1) ;$$

that is, $\sigma(x; 0)$ be a step function possessing its jumps only at the end-points of the interval. It then follows that

$$\int_0^1 x^2 d\sigma(x; 0) = \lambda'' , \qquad \int_0^1 x\, d\sigma(x; 0) = \lambda''$$

and hence, in view of (5.13)

$$\Gamma(t) = \lambda'' - \int_0^1 x^2 d\sigma(x; t) = \int_0^1 (x - x^2) d\sigma(x; t) .$$

Because of monotone increasing character of $\sigma(x; t)$ the inequality $(x-x^2)d\sigma(x; t) \geqq 0$ being valid always in the interval $0 \leqq x \leqq 1$, it surely follows that $\Gamma'(t) \geqq 0$ always. Moreover, it is evidently seen that $\Gamma'(t) > 0$ for $t > 0$ unless $\sigma(x; t) \equiv \sigma(x; 0)$, whence it is concluded that the result

(5.22) $P_{ii}(t) < P_{ii}(0)$ $(0 < t)$

corresponding to (4.3) is also valid.

Various results in the discrete case will also be generalized correspondingly. In particular, our previous main result itself may also be regarded as a special case of the last result (5.19) and (5.20) represetened in Stieltjes form, where the set-function $\sigma(x; 0) - \sigma(x; t)$ is derived from an another suitable function of x, y, t generalizing the mating-velocity which reduces particularly to a step-function with respect to the first two variables.