

CHAPTER 2: KINSHIP THEORY

2.0 Introduction

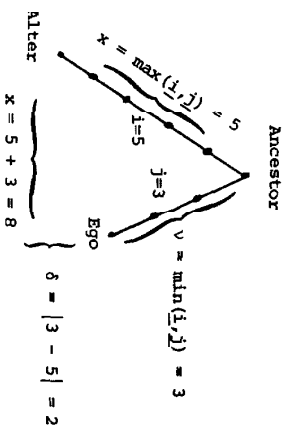
Studies in the algebra of kinship have been increasingly numerous. From a small beginning with the work of the Scottish mathematician Macfarlane (1882), the works are so numerous in the last fifteen years that we may claim the existence of a small "field" which studies the algebraic ways of describing social relationships. In genetics of course "kinship" is a name of a measure not a body of theory. However, as we will note briefly in this chapter, there are obvious relationships between the social and the genetic concepts of kinship. Sections 2.2 and following are specifically devoted to exploring new concepts for genetics based on developments given here.

The purpose of this chapter is not to review the algebra of kinship, but to present sufficient notions of kinship algebra to allow development of subsequent chapters from a common foundation. The most compact source for this purpose turns out to be two papers by Atkins (1974a, 1974b), and what follows draws from both papers.

We attempt to follow Atkins' notation literally for kinship measures, but will in general develop our own notation for other purposes.

2.1 The Fundamental Numbers

The basic problem of kinship theory, whether in social anthropology or genetics, is to study the relationship through genealogical ties of two given individuals. Following anthropological tradition, we call one of these persons "ego" and the other "alter". To define the measures of interest, consider the following diagram taken from Atkins (1974a):



Stated in ordinary language, we have six significant measures to contend with between alter and ego with respect to each common ancestor. Where j is the number of descent links from ego to the ancestor, and i from the ancestor to alter, these six measures are summarized in the following table:

Table 2.1 Summary of Definitions and Algebraic Relationships of Kinship Measures (From Atkins, 1974a, Table 2)

Measure	Definition	Range of Values	Useful Equalities
γ (Generation)	$\gamma = j - i$	$\dots, 2, 1, 0, 1, 2, \dots$	$\gamma = \delta \phi$
σ (Generation sign)	$\sigma = \text{sgn}(j - i)$	$-1, 0, 1$	$\sigma = \text{sgn } \gamma$
δ (Generation removal)	$\delta = i - j $	$0, 1, 2, \dots$	$\delta = \gamma $ $\delta = X - \gamma = K - 2\gamma$
ν (Collateral removal)	$\nu = \min(i, j)$	$0, 1, 2, \dots$	$\nu = X - \delta = K - X$
X (Canon degree)	$X = \max(i, j)$	$0, 1, 2, \dots$	$X = \delta + \nu = K - \nu$
K (Civil degree)	$K = i + j$	$0, 1, 2, \dots$	$K = \nu + X = \delta + 2\nu = 2X - \delta(*)$

* From Ballinoff, 1975.

Since the measures defined in terms of the diagram are quite straightforward, there is little need to explain their "meaning" or the origin of the "useful equalities", most of which are obvious from the definitions. Atkins also gives lengthy explanation of the history of these and several other related measures.

The simplicity of these measures is more apparent than real. In fact, we hardly exploit even the simplest measures at their simplest values. For example, chapter three depends only upon the values δ , X , and K . In its normal development, population genetics depends only upon K . In chapter 3 on the graph theory of marriage systems, and in chapter 4 on structural numbers, we essentially assume $\delta = 0$ and $\nu = X = (1/2)K = i = j$. However, the methods of chapter 4 are extendable to any values, and are defined in their general sense, while in chapter 6 we indicate how to introduce less discrete conditions though retaining the theory derived from $\delta = 0$, at least for demographic purposes.

In Atkins' terminology, these six measures are referred to as "fundamental consanguineal measures". We shall retain this terminology, and also refer to them as "kinship" measures. In chapter 4, we will develop a separate theory of yet other numbers. We will refer to the numbers found in chapter 4 as "structural numbers", since they depend on the types of graphs (structures) found under specified values of kinship measures.

2.2 Genetic Uses of the Fundamental Numbers*

In population genetics, following Malécot (1969), we are concerned with the probability of identity by descent of an allele chosen at random from a locus of ego, with an allele chosen at random from the homologous locus of alter. In the special case of ego and alter being mates, we may study the inbreeding of their offspring, which is the probability that the two homologous alleles at a given loci of the offspring are identical by descent. Clearly, the inbreeding of the offspring is identical to the kinship of the parents.

Denote ego by I and alter by J ; let K be the offspring of I and J ; use " f_a " for the inbreeding of the common ancestor " a " of I and J ; and use f_K for inbreeding of K , θ_{IJ} for the kinship of I, J . Then

$$(1) \quad f_K = \theta_{IJ} = \frac{1}{2} f_a^{K+1} (1 + f_a)$$

where K_a is the civil degree relationship of I, J through ancestor a , and we sum over all common ancestors.

As is evident from reading texts on population genetic theory (e.g., Jacquard 1974 or Crow and Kimura, 1971, a considerable amount of the abstract theory depends upon the population average values of f or of θ . Further, many empirical studies from civil records are essentially studies of distributions of values of f , which is the same as studying distributions of values of K .

In the next sections, we note that K is not the only way to compute f or θ .

2.3 Semi-continuous and Continuous Genealogical Measures

The "normal" methods of genealogical counting of links in pedigrees is to count the numbers of "steps" between offspring and ancestors. Thus, from a person to his grandparent, there are two "steps", or three individuals in the path, and from a person to their first cousin, there are five individuals in the path, or four connecting links.

If k is the number of persons in the path, or g the number of connecting links then we have $k = g + 1$. Genetic relationship ρ is computed as $\rho = \frac{1}{2^g}$ through a path (of a non-inbred ancestor) and kinship as

$$(2) \quad \theta = \left(\frac{1}{2}\right)^k = \left(\frac{1}{2}\right)^{g+1} = \frac{1}{2^p}$$

*Readers uninterested in genetics may omit the remainder of this chapter.

In all of these measures, values of x and g are always integers, so that we may consider this method of computation as "discrete".

Now, if we look at the genealogical path between two individuals I, J through a common ancestor, we note from Ballouff (1975) and section 2.1 that this path breaks into two parts: the path from individual I to the ancestor, and the path from the ancestor to J . Then the number k above is the sum of these two paths, and is equal to k of table 2.1. Also, from table 2.1, the maximum of these two paths is x , and the difference between them, δ , we have from the above references the fact that

$$(1) \quad g = 2x - \delta = k.$$

From Ballouff (1975), we have the following method of computing δ in the absence of information other than x , and the birth dates of I, J . If the difference in years of the birth dates of I, J is d_{IJ} , and the average time span between generations of the population (average age of childbearing for females) is \bar{g} , then we can estimate $\bar{d}_{IJ} = \frac{d_{IJ}}{\bar{g}}$ and use \bar{d}_{IJ} as an estimate of δ for each pair I, J which apparently reasonable results. For example, the mean computed result from the example in Ballouff (1975) is evidently a much better estimate of the "true" δ of a population than are estimates found from isonymy, summarized by Yasuda and Morton (1967).

In this model, although δ is discrete valued, \bar{d} is continuous valued. For this reason, if we compute θ or ρ as in (1), but using (2) to estimate g , and estimating δ with \bar{d} , we get,

$$(1) \quad \theta = \frac{1}{2\rho} = \left(\frac{1}{2}\right) 2x - \bar{d} + 1$$

then we have a semi-continuous measure, dependent on the discrete valued x and the continuous valued \bar{d} .

We may go one step further, and also estimate x continuously by measuring the maximum difference in dates of birth of the ancestor and I , or the ancestor and J .

Calling this difference D_{IJ} , and using \bar{g} as above, we get an estimate $\bar{x}_{IJ} = \frac{D_{IJ}}{\bar{g}}$ of x for each pair I, J . We then estimate g as

$$(1) \quad \hat{g} = 2 \left(\frac{D_{IJ}}{\bar{g}} \right) - \bar{d}_{IJ} = \frac{2}{\bar{g}} (D_{IJ} - \bar{d}_{IJ})$$

Equation (5) gives us a completely continuously valued estimate of g , which has the unique property that we may estimate the genetic relationship of two individuals provided only that we know both the individuals have some common ancestor(s), and, the birthdates of both individuals and of all common ancestors.

Using continuous estimates, the following problem arises: the connecting path from each common ancestor to each of individuals I and J is not always "relatively simple". Here, relatively simple means that for each ancestor A common to I and J , and for any particular path connecting I and J through A , there is no relative born later in time who also has A for an ancestor and who is an ancestor of both I and J . Obviously, if we find such an intermediate ancestor on a given path, then we must first estimate the inbreeding of that ancestor, and then compute the relationship p_{IJ} via the closer ancestor using

$$p_{IJ} = \left(\frac{1}{2}\right)^f (1 + f)$$

where f is the inbreeding of the closer ancestor.

This also implies that we need to know how many paths connect each pair through each ancestor. This problem is not unique to kinship computation by continuous or semi-continuous methods. However, we have learned that there is yet another discrete valued measure embedded in computation of f or ρ , namely the multiplicity of connection via each ancestor. In discrete and semi-continuous computation, since we compute each path separately, the multiplicity of each 1 , so we are not necessarily conscious of the number of paths through each ancestor. However, in continuous computation, we are in danger of losing information by using only age structure information, unless we also discover techniques for finding multiplicity of connection.

With this background, in the following we discuss use of semi-continuous measures on a particular problem, namely modeling the prediction of amount of inbreeding to be found in artificial populations of any salmon, which have an "age structure" which is simply a proportionate mixing of generations, moving over time. The example is useful since it shows that marriage theory and in particular the branch of marriage theory dealing with kinship is in fact much more broadly applicable than simply to human populations.

2.4 Introduction to Use of Semi-Continuous Models

For each individual's parents, we may calculate the coefficient of kinship θ or of relationship $\rho = 2\theta$. The inbreeding of the individual offspring is therefore also $\frac{1}{2}\rho$ of the parents, and the inbreeding of a population is the average of the inbreeding of individuals, hence is computable from θ or ρ .

Now, for a given individual with parents I, J , the number of steps (generations) from I to a particular common ancestor is i , and from J , it is j , so $\rho = \frac{1}{2} (1 + f)$ or $\theta = \frac{1}{2} (1 + f)$. To not worry about the "extra" $+1$ in the exponent, we use ρ instead of θ .

Let $x = \max (i, j)$, $\delta_{ij} = |i - j|$. For each matched pair in the population

we can compute $\rho = \frac{1}{2} (i_a + j_a) (1 + f_a)$ for each ancestor a and for one or more ancestors, with inbreeding. Now, from section 2.3, we know that $\kappa = 2X_{IJ} - \delta_{IJ}$ is

an estimate of $i + j$. Therefore $\rho_a = \frac{1}{2} X_a - \frac{\delta_{IJ}}{2}$ is an estimate of the relationship due to ancestor a of couple I, J . And $\rho_{IJ} = \frac{1}{2} X_{IJ} - \frac{\delta_{IJ}}{2}$ is an

estimate of the total relationship of pair I, J .

Now consider a population divided into m classes each of males and females,

and let $1 \leq i \leq m$ index these classes in the males and $1 \leq j \leq m$ index these classes in the females. Let $p_i^{(t)}$ be the proportion of males in class i at time t and let

$p_j^{(t)}$ be the proportion of females in class j at time t . (Notice that we can generalize for different numbers of classes in males and females by allowing all $i_1 = 0$, or $p_j = 0$ for $1 \geq k, k < m$.)

If we are now interested in the classes of pairs of classes i, j (a male from i , female from j) we can let this proportion at time t be $p_{ij}^{(t)}$.

We may now consider the matrix $P_c = [p_{ij}^{(t)}]$ which shows the proportions of all pairs of classes at time t . Assume that every member of class i or j has the same genetic history as every other member of class i or j . That is, if an individual in class i has a particular individual ancestor at a particular degree of removal, then every other individual in class i has this same ancestor at the same degree of removal, etc.

Therefore, for each class pair (i, j) the expected genetic relationship is $p_{ij}^{(t)}$. And in the whole population, the average genetic relationship is, at time :

$$\bar{\rho}_c = \sum_{i,j} p_{ij}^{(t)}$$

For convenience, define the matrix multiplication $*$ which is the element by element multiplication of square matrices of the same order. If S, M, R are $n \times n$ square matrices, $S = M * R$, if and only if $[S_{ij}] = [M_{ij}R_{ij}]$ for all i, j . Clearly $M * R = R * M$, etc.

Then denoting $P_c = [p_{ij}^{(t)}]$ we have

$$\bar{\rho}_c = \text{perm } P_c * R_c$$

where $\text{perm } M, M$ a matrix, is the sum of all entries in the matrix.

2.5 Estimating Proportions

We must now discuss how to estimate P , how to estimate R , and their effects on $\bar{\rho}^{(t)}$. We have several possible cases for P and R . Looking first at P , we have several cases:

$$(1) P_{ij}^{(t)} = p_i^{(t)} p_j^{(t)}$$

Or, the proportion of i, j pairs in the population is the simple product of their relative proportions in the entire population at time t ;

$$(11) P_{ij}^{(t)} = \frac{p_i^{(t)}}{\sum_w p_w^{(t)}} * \frac{p_j^{(t)}}{\sum_v p_v^{(t)}}$$

where $\frac{p_w}{\sum_w p_w^{(t)}}$ is the total of proportions of a set of w of the classes of males,

$1 \leq w \leq m$, and class i is one of the w classes, with similar interpretation of $\frac{p_v}{\sum_v p_v^{(t)}}$ for females. Expression (11) has the effect of saying the pairs form selectively from the population, but randomly within this selected set. Notice that case (1) is case (11) with $w = v = m$;

$$(111) P_{ij}^{(t)} \text{ and } p_j^{(t)}$$

depend on the productivity of mating types in the previous generation, so $p_{ij}^{(t)}$ is defined as a finite difference equation.

Note that these models essentially only duplicate the developments of chapter 1.

2.6 Estimating Genetic Relationship

Cases for types of ρ computations are:

CASE (I): Parental populations of i, j unrelated for all $i \neq j$. Then we simply have

$$\bar{\rho}_c = \text{perm } P_c * R_c = \text{perm } [R_{ij}] * P_c = \sum_{i,j} p_{ij}^{(t)}$$

since for $i \neq j, p_{ij} = 0$, and E is the identity matrix. In more detail, for each $i = j$, all members of the population have the same ancestry; for $i \neq j$, ancestry is not in common. Then $p_{i_1} = 0$ for $i \neq j$, and

$$\bar{\rho}_c = \sum_{i,j} p_{ij}^{(t)} = \sum_{i,j} \frac{1}{2} X_{ij} - \frac{\delta_{ij}}{2} = \sum_{i,j} \frac{1}{2} X_{ij} - \frac{\delta_{ij}}{2} (1 + f_a)$$

In particular, if the common ancestors of classes i, j are their immediate parents, then $X_{ij} = 1$. Also, the generation difference between i and j is zero, so $\delta_{ij} = 0$,

$$\text{and } \bar{\rho}_e = \frac{1}{2} \left(\frac{1}{2} \right)^2 (1 + \bar{f}_a)$$

If there are no genealogical "complications" other than the common parents, we have

$$\bar{\rho}_e = \frac{1}{4} \frac{1}{2} \frac{1}{2} \frac{1}{2} (1 + \bar{f}_a)(t)$$

where \bar{f}_a is the inbreeding of a parent of pair i, j , and the subscript (t) on $(1 + \bar{f}_a)(t)$ is simply a reminder that we are talking about ancestors of pairs in generation t .

CASE (II): All classes i, j are bred from the same parental stock. Therefore, for all $i, j, \rho_{ij}^{(t)}$ equals a constant h , which is independent of time and of the classes or class-pairs. Then $\bar{\rho}_e = h$ perm $P = h$, so the relationship at any time is independent also of the class composition.

CASE (III): A set of each class $i = j$ of case (ii) are kept as breeders for the classes $i + 1, j + 1$. Therefore, at time t , $X_{ii} = i$ for $i = j$, $X_{ij} = \delta_{ij} + 1$ for $i \neq j$ or $X_{ij} = \delta_{ij} + 1$ for all i, j . We get, for each pair-type i, j

$$\rho_{ij} = \frac{1}{2} \left[\frac{2(\delta_{ij} + 1) - \delta_{ij}}{2} \right] (1 + \bar{f}_a)(t)$$

Concentrating on the exponent, we notice

$$2\delta_{ij} - \frac{\delta_{ij}}{2} + 2 = \delta_{ij} \left(\frac{2\bar{G}_e - 1}{\bar{G}_e} \right) + 2 = \delta_{ij} \left(2 - \frac{1}{\bar{G}_e} \right) + 2$$

which we call $\Delta_{ij}^{(t)}$ for convenience. Note that for $\bar{G}_e = 1$, $\Delta_{ij}^{(t)} = \delta_{ij} + 2$; for $\bar{G}_e = \frac{1}{2}$, $\Delta_{ij}^{(t)} = 2$; for $\bar{G}_e < \frac{1}{2}$, or δ_{ij} "small", $\Delta_{ij}^{(t)} < 2$. Also, recall that in general, δ_{ij} is a small integer such as 0, 1, 2, 3, 4.

Looking at the inbreeding of each pair's parents, we can notice that the inbreeding \bar{f}_a of the ancestors of a particular pair will, in fact, be defined by a recurrence: the parents of $\{(i, j) \mid i = j\}$ are precisely the pair $(i - 1, j - 1)$, or at least the inbreeding is the same, so that \bar{f}_a here becomes

$$\bar{f}_{i-1, j-1} = \frac{1}{2} \bar{f}_{i-2, j-2} = \frac{1}{2} \left(\frac{1}{2} \right)^{(\Delta_{i-2, j-2}^{(t)} + 1)} (1 + \bar{f}_{a-2, j-2})$$

For diagonal entries in R , we have entry

$$(1, 1) = \frac{1}{2} \Delta_{1,1}^{(t)} (1 + \frac{1}{2} \bar{f}_{1-2, 1-2}) (2) = \frac{1}{2} \Delta_{1,1}^{(t)} (1 + \frac{1}{2} \bar{f}_{1-2, 1-2})$$

Now, $\Delta_{11}^{(t)} = \delta_{11}^{(t)} (2 - 1/\bar{G}_e) + 2$, and since $\delta_{11} = 0$ for all t , and $t, \Delta_{11}^{(t)} = 2$.

The diagonal entries look like $\rho_{i,i} = \frac{1}{2} + \frac{1}{4} \rho_{i-2, i-2}$

The arbitrary entry looks like

$$\rho_{i,j}^{(t)} = \frac{1}{2} \Delta_{i,j}^{(t)} (1 + \frac{1}{2} \rho_{\max(i,j), \max(i,j)}^{(t)}) (2) = \frac{1}{2} \Delta_{i,j}^{(t)} - 1 + \frac{1}{2} \Delta_{i,j}^{(t)} \rho_{\max(i,j)-2, \max(i,j)-2}^{(t-2)}$$

Using, for simplicity, $h_e = (2 - 1/\bar{G}_e)$, we have

$$\Delta_{ij}^{(t)} = \delta_{ij}^{(t)} h_e + 2$$

$$\Delta_{ij}^{(t)} - 1 = \delta_{ij}^{(t)} h_e + 1$$

and $\rho_{ij}^{(t)} = \frac{1}{2} \delta_{ij}^{(t)} h_e + 1 + \frac{1}{2} \delta_{ij}^{(t)} h_e + 2 \rho_{\max(i,j)-2, \max(i,j)-2}^{(t-2)}$

The matrix of δ_{ij} is:

$i = 0$	0	1	2	3	4	5	6
$j = 0$	0	1	1	1	1	1	1
	1	1	0	1	2	3	4
	2	2	1	0	1	2	3
	3	3	2	1	0	1	2
	4	4	3	2	1	0	1
	5	5	4	3	2	1	0
	6	6	5	4	3	2	1

which is simply generalization removal (see table 2.1) and the matrix showing $\max(1, j) - 2$, $\max(1, j) - 2$ is:

	j =	0	1	2	3	4	5	6
i = 0		(-2, -2)	(-1, -1)	.	.	.	(3, 2)	(4, 4)
1		(-1, -1)	(-1, -1)
2	
3	
4	
5		(3, 3)	(3, 3)	.
6		(4, 4)	(4, 4)

which is $X = 2$, from table 2.1. So the matrix $R(t)$ looks like:

$$R(t) = \begin{bmatrix} \frac{1}{2} 0^{t+1} + \frac{1}{2} 0^{t+3} (t-2) & . & . & . & \frac{1}{2} 6^{t+1} + \frac{1}{2} 6^{t+2} (t-2) & \rho(4,4) \\ \frac{1}{2} 6^{t+1} + \frac{1}{2} 6^{t+2} (t-2) & \rho(4,4) & . & . & \frac{1}{2} 0^{t+1} + \frac{1}{2} 0^{t+2} (t-2) & \rho(4,4) \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ \frac{1}{2} 6^{t+1} + \frac{1}{2} 6^{t+2} (t-2) & \rho(4,4) & . & . & \frac{1}{2} 0^{t+1} + \frac{1}{2} 0^{t+2} (t-2) & \rho(4,4) \end{bmatrix}$$

or, $R(t) = \begin{bmatrix} \frac{1}{2} \delta_{1j} h^{t+1} \\ \frac{1}{2} \delta_{1j} h^{t+2} \end{bmatrix} * R(t-2)$.

Essentially, everything depends on the value of h , and on the initial β for the initial population(s). If h is positive or 0, then we have a system with "regular" behavior. If h is negative, then the dynamics are unusual, first growing as δ_{1j} grows, then growing (even faster!) as $\delta_{1j} k > 3$, depending on initial R , and dependent on the initial inbreeding of the initial pair, and on the "age structure" P_1, P_2 of breeding pairs after that.

Notice, we do not have recursion in the proportions P_{ij} or dependence of \bar{P}_t on any values of $P_{ij}^{(t)}$ other than for the same time period t , because we assume no relationship (independence) if fertility and matrix type by "age class", providing the class exists at all in a particular time period t .

However, if we wish in this model, we can set $P_{ij}^{(t)} = P_{i-1, j-1}^{(t-1)}$ so that the entire matrix P of frequencies of classes moves over time in a regular way. In this case, the same method allows us to extend the computations as long as the stock is maintained as indicated.

CASE (IV): The proportions of matings between relatives of type $X=1, 2, \dots$ for each pair $\{i, j \mid i=j\}$ is given by a vector X of relative proportions of types $\{V_1, V_2, \dots, V_k\}$ of relationship types $X=1, X=2, \dots, X=k$.

Case (IVa) $\{i, j \mid i \neq j\}$ has the same vector v of proportions of kin types.

then for each class type i, j we must compute the relationship as the mean of the types:

$$\bar{P}_{ij} = \frac{1}{K} \sum_{\text{all ancestors}} \left[\frac{1}{2} (2k - \delta_{ij}) \left(1 + \frac{F}{\text{each ancestor at degree } (\max(1, j) - 1, \max(1, j) - 1)} \right) \right]$$

Let us assume all initial F values are zero for now. Also, assuming there is one common ancestor, we find

$$\bar{P}_{ij} = \frac{1}{K} \sum_{k=1}^K \frac{1}{2} (2k - \delta_{ij})$$

or, two common ancestors (both parents):

$$\bar{P}_{ij} = \frac{1}{K} \sum_{k=1}^K \frac{1}{2} (2k - \delta_{ij} - 1)$$

Case (IVb) the vector v of relative proportions of kin types (X values) differs. For example, in cases II and III above, the relative proportions of possible relatives is strongly related to or determined by the assumptions on mating system. In case II, we assumed all ancestors of type (i, j) were of a specific class.

For other cases, we must make specific computations. In these cases, we compute the matrix equation

$$R(t) = \begin{bmatrix} \frac{1}{2} X_{ij}^{(t)} - \delta_{ij}^{(t)} \\ \frac{1}{2} X_{ij}^{(t)} \end{bmatrix} + \frac{1}{2} \begin{bmatrix} X_{ij}^{(t)} - \delta_{ij}^{(t)} \\ X_{ij}^{(t)} \end{bmatrix} * R(t-2)$$

$$= X(t) + \frac{1}{2} X(t) * R(t-2)$$

Also, the rows below show the number return in numbers of years from time of release the columns show population composition at a given time from t_0

$\tau_0^{(0)}$	$\tau_1^{(0)}$	$\tau_2^{(0)}$	$\tau_3^{(0)}$	$\tau_4^{(0)}$	$\tau_5^{(0)}$	$\tau_6^{(0)}$...
	$\tau_0^{(1)}$	$\tau_1^{(1)}$	$\tau_2^{(1)}$	$\tau_3^{(1)}$	$\tau_4^{(1)}$	$\tau_5^{(1)}$...
		$\tau_0^{(2)}$	$\tau_1^{(2)}$	$\tau_2^{(2)}$	$\tau_3^{(2)}$	$\tau_4^{(2)}$...
			$\tau_0^{(3)}$	$\tau_1^{(3)}$	$\tau_2^{(3)}$	$\tau_3^{(3)}$...
				$\tau_0^{(4)}$	$\tau_1^{(4)}$	$\tau_2^{(4)}$...
					$\tau_0^{(5)}$	$\tau_1^{(5)}$...
						$\tau_0^{(6)}$...

release at time t , and $\tau_j^{(j)}, j > 0$ later returns in periods j units from t .

2.7 Proportions of Wild Types

Suppose that in addition to the classes $1 \leq i \leq m, 1 \leq j \leq m$ there are classes $j=0$, $j=0$ of individuals which are not produced by the same source as the other classes i, j . In other words, a set of "wild" types unrelated to the classes i, j .

Then, for any pair $(i, j), i \neq 0, j \neq 0$, we have the same computations as above. For any pairs $(0, j), (i, 0), i \neq 0, j \neq 0$, the genetic relationship must be zero, since there are no common ancestors. Likewise, all $(0, 0)$ pairs have no history in common with i, j , but there may be an amount of w of genetic relationship ($\frac{1}{2}w = \bar{r}_0^{(t)}$). If the proportions of the classes are

$$p_c = [p_{ij}^{(t)}] \text{ for } i, j \text{ pairs}$$

$$p_{0,j}^{(t)} = [p_{0j}^{(t)}] \text{ for } 0, j \text{ pairs}$$

$$p_{i,0}^{(t)} = [p_{i0}^{(t)}] \text{ for } i, 0 \text{ pairs}$$

$$p_{0,0}^{(t)} = [s^{(t)}] \text{ for } 0, 0 \text{ pairs}$$

We can also treat 0 as the lower index of i and j and look at the new $(m+1) \times (m+1)$ matrices P, R , but we can set a clearer picture by setting $q_{ij} + q_{0j} + q_{i0} + q_{00} = 1$, where q is the proportion of the total represented population in the matrix of the same subscript. We could make these q 's into matrices by use of multiplication to study each element independently. The same analysis as follows would then hold for each element independently and the final average value will be the average for all elements.

We have:

$$p_c = q_{ij} p_{i,j}^{(t)} + R_{i,j}^{(t)} + q_{0,j} p_{0,j}^{(t)} + q_{i,0} p_{i,0}^{(t)} + R_{i,0}^{(t)} + q_{0,0} p_{0,0}^{(t)} + R_{0,0}^{(t)}$$

Notice that since $R_{0,j}^{(t)} = R_{j,0}^{(t)} = [0]$, we can eliminate two terms, but since $q_{i,j} + q_{0,0} \neq 0$ in general, we can say little about the total relationship if we assume $R_{0,0} \neq 0$. If we can compute values for $q_{0,1}, q_{1,0}, q_{0,0}$ then direct computations are possible. Otherwise, if we assume $R_{0,0} = [0]$, then $q_{0,i}^{(t)} + q_{i,0}^{(t)} + q_{0,0} = q^{(t)}$ and $q_{i,j}^{(t)} = 1 - q^{(t)}$. In this case, we may set

$$\beta_c = \bar{p}_c(1 - q^{(t)}) + q^{(t)} \cdot 0 = \bar{p}_c(1 - q^{(t)}).$$

The total relationship is reduced by the product $q^{(t)} \bar{p}_c$. If we assume we have correctly computed \bar{p}_c from (i, j) information we may then estimate the proportion $q^{(t)}$ of matings with wild types and between wild types, hence the proportion $1 - q^{(t)}$ of matings between "factory produced" types:

- (1) providing we can estimate β_c from other evidence (e.g., biochemical studies of the population), or
- (2) if a direct estimate of $q^{(t)}$ can be made from catch proportions and a specific assumption on proportions of mating types.

In case (1): $q^{(t)} = \frac{\bar{p}_c - \beta_c}{\bar{p}_c}$

In case (2): estimate $q_{i,0}, q_{0,j}$ and $q_{0,0}$ by the specific assumptions, then compute $q^{(t)}$ as above.

In the case where we treat each component separately, so that the q 's become matrices $Q_{0,i}$, etc., we have:

$$\beta_c = \text{perm} \begin{bmatrix} Q_{1,j}^{(t)} * P_{1,j}^{(t)} * R_{1,j}^{(t)} \\ Q_{0,j}^{(t)} * P_{0,j}^{(t)} * R_{0,j}^{(t)} \end{bmatrix} + \text{perm} \begin{bmatrix} Q_{i,0}^{(t)} * P_{i,0}^{(t)} * R_{i,0}^{(t)} \\ Q_{0,0}^{(t)} * P_{0,0}^{(t)} * R_{0,0}^{(t)} \end{bmatrix} + \text{perm} \begin{bmatrix} Q_{i,0}^{(t)} * P_{i,0}^{(t)} * R_{i,0}^{(t)} \\ Q_{0,0}^{(t)} * P_{0,0}^{(t)} * R_{0,0}^{(t)} \end{bmatrix}$$

One estimate of $q^{(t)}$ is then

$$q^{(t)} = \text{perm} Q_{i,j}^{(t)} P_{0,j}^{(t)} + \text{perm} Q_{i,0}^{(t)} P_{0,i}^{(t)} + \text{perm} Q_{0,0}^{(t)} P_{0,0}^{(t)}$$

We then estimate \hat{c} from the exterior source, and compute

$$q^{(t)} = \frac{\text{perm } Q_{1,j}^{(t)} R_{1,j}^{(t)} - \beta}{\text{perm } Q_{1,j}^{(t)} R_{1,j}^{(t)}}$$

Note that the fact that we are using p and not θ or f is irrelevant, since as is clear in these last two expressions, the proportionality constants p, f, θ , will all cancel out of the estimation equations for $q^{(t)}$.

These equations therefore allow computations of $q^{(t)}$ from prior knowledge of computations in the earlier sections.