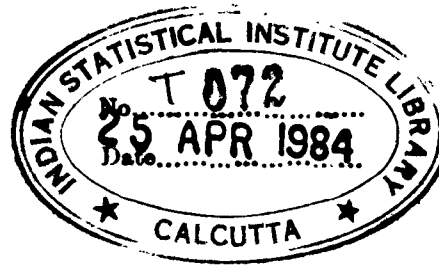
A Statistical Study of Tongue Pigmentation in Man



DABEERU CHANDRASEKHARA RAO

CALCUTTA
1970

Correction:

For page numbers, add
10 to each printed
number after
49

In memory of

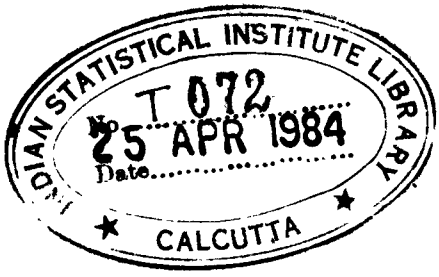
my brother

D. Srinivasa Rao

A STATISTICAL STUDY OF TONGUE PIGMENTATION IN MAN

By

DABEERU CHANDRASEKHARA RAO



A thesis submitted to the Indian Statistical Institute in partial fulfilment of the requirements for the award of the degree of

Doctor of Philosophy

Calcutta

1970

A C K N O W L E D G E M E N T S

It gives me great pleasure to express my deep sense of gratitude to Professor C.R.Rao,F.R.S., under whose valuable supervision this thesis work was carried out. I also express my special regards to him as the Director of the Research and Training School who provided me all the necessary facilities for carrying out the research work.

I am very grateful to Professor R.L.Kirk, C.C.Li, Curt Stern, J.V.Neel, W.J.Schull and James F.Crow, who went through parts of this thesis work and offered many useful comments. I am also very grateful to Professor T.A.Davis, who is mainly responsible for the interest I have developed in tongue pigmentation studies.

I am very much indebted to Mr. R. Chakraborty for his many constructive suggestions and comments pertaining to my research work, and also for his untiring help during the preparation of this thesis. I am very grateful to many faculty members of the Research and Training School for the encouragement and inspiration I have been receiving from them.

Mr. R.Chakraborty, Mr.J.K.Gorai and Miss Monami Bose have kindly permitted me to include our joint results in this thesis, for which I am very thankful to them.

Finally, I thank Mr. Arun Das and Mr. P.Nandi for their excellent and fast typing of the thesis.

D.C.RAO

C O N T E N T S

CHAPTER I

INTRODUCTION

1.1 Tongue pigmentation in man :

The 'particular' type of dark spots and patches on the surface of the tongue, as investigated in the present thesis, was first noticed in human populations by Davis [5], who called it 'tongue pigmentation'. It is to be emphasized that not all colour pigments come under this 'particular' type. There are some pathological conditions giving rise to transient discolouration of the tongue. This aspect will be discussed a little later. Such tongue pigmentation does not seem to have been reported for any population, human or otherwise.

Tongue pigmentation is usually found on the upper surface or borders (rim) of the tongue. In most of the cases, the spots were observed on the upper surface including the tip of the tongue, and the patches on the borders, right and left. It may be recorded here that out of all the tongue pigmented people seen so far, nearly 80% showed only spots, 15% showed only patches and the rest showed both spots and patches. Very few individuals had this pigmentation on the interior surface and/or the borders of the tongue. Thus, one may have either exclusively spots or

patches, otherwise both. The trait exhibits variation in the number as well as the distribution of the spots and patches, and also in the size and shape of the latter. There may be one spot to many spots, isolated or clustered, and patches from the size of about 2 sq. mm. to as large a size as covering nearly half of the tongue surface. Out of several thousand examined so far only one was observed showing just one spot covering an area of nearly 0.5 sq. mm. on the deep interior surface.

Some observations indicate that the trait is not subject to any changes over time. For example, I became aware of a patch of about 1 sq. cm. on my tongue in 1964, but there has not been any change of any sort till now, either in respect of its size or intensity of the colour of this pigmentation. Also, at the time of the tongue examination, most of the literate and pigmented subjects were asked as to whether they remember to have had it since birth. In some cases, their parents were interrogated about it. Though in most of the cases the reply was "I am not sure", quite a few of them confirmed to have it unchanged ever since they first noticed it on their tongues. This provides an encouraging evidence in favour of the above contention.

Method of examination :

The subject was asked to wide open the mouth, extending the tongue outward frontally as far as possible. Most of the tongues were examined in the diffused day-light, though a few had to be examined under torchlight. Generally, the South Indian women were rather shy to show their tongues especially in the presence of others. But most of them could be persuaded to show the tongue, though examination had to be done under torchlight inside their houses. Also, most of the new born babies who were studied in Calcutta hospitals for the trait, were examined under torchlight inside nurseries. Though it was not very difficult to examine the grown-ups under torchlight, considerable difficulties were faced in the examination of the babies at hospitals under torchlight. It is not, however, advisable to examine the tongues under artificial light, or direct sun light.

The habit of chewing betel (may be with tobacco and/or catechu) creats a deep colouration on the upper surface of the tongue, which in most chewers can mask the pigmentation patches completely, if present. It can also produce coloured artifacts on the tongue resembling (to the less experienced investigators) the pigmentation patches.

To avoid this difficulty, it was decided to discard all such cases unless they agreed to be examined in diffused day-light. Inspite of this, if the subject could not be classified one way or the other with complete accuracy, it was decided to discard such cases. It may be noted that a few subjects were discarded this way, all of whom were under the regular habit of chewing too much of betel. Thus, only those subjects who were most favourable for tongue examination were included in our studies on tongue pigmentation.

The spots and patches, coming under 'tongue pigmentation', assume slightly different colours. The spots are dark brown whereas the patches are usually of dark ash-brown colour. The colour of this pigmentation itself makes it quite unique, and thus distinguishing this pigmentation from other pseudo-pigments of the tongue arising out of certain pathological conditions, or from artifacts due to habits of chewing betel or tobacco (many rural Indians chew tobacco leaves regularly). Betel chewing produces discontinuous and large patches, the colour being red (in different shades). And tobacco leaf colour is developed in cases of tobacco-chewers. In any case, if there arises any doubt due to the presence of artificial colour coatings on the tongue, a little scratching of such areas with some sharp edge should clarify it. This method was found to be very useful

in some regular tobacco-chewers. The possibility of confusing between tongue pigmentation and other types of discolouration of the tongue arising out of some pathological conditions is much reduced by an understanding of the latter conditions. Some such pathological conditions are considered below:

(i) White spots and patches : These are usually associated with vitamin deficiency. The 'scratching technique' helps to isolate these spots without any difficulty. This type creates no problem while examining for tongue pigmentation.

(ii) Whitening of the tongue surface : This usually happens due to leukaemia, a condition arising out of a great excess of white corpuscles in the blood, associated with changes in the lymphatic system. This condition does not produce as such any patches, but the normal colour of the tongue changes all over since the relative density of white corpuscles gets elevated. This colour also does not stand on the way of search for tongue pigmentation.

(iii) Light-brown to red spots : This is most likely to be confused with tongue pigmentation. This type of spots appear on the tongue almost surely due to malnutrition. At times, even discontinuous patches appear under this condition. Though it is possible to scratch these spots and isolate them without paining the subject,

it is not known whether this scratching method actually sufficed to identify all these cases with complete accuracy in our studies, because, not all the subjects identified to have this type of pigmentation allowed their tongues to be scratched for this purpose. Thus, there remains an element of doubt in respect of distinguishing tongue pigmentation from this type of pigmentation. Since most of the cases of malnutrition arise out of poverty, data on the families coming from economically very backward colonies are not subjected to any analysis and they are left out of this thesis. This is done just to safeguard the data from possible doubtful observation. In fact, only at the initial stages of data collection, a few families from economically very backward colonies (mostly scheduled caste) were examined and the data recorded. Later on, such families were not even examined. However, some population data were collected from such communities also, which are reported in the next Chapter.

(iv) Black patches : These patches are suspected to be cancerous in origin (Sanghvi [32]). Active research is being carried out at the Cancer Research Institute of the Tata Memorial Centre (Bombay) on this aspect. In any case, this colour raises no problems about identifying tongue pigmentation.

(v) Dark-blue patches : This is caused by carcinoma of the

tongue. Though not very much is known about it, this pigmentation can be accurately distinguished from tongue pigmentation by a thorough examination of the tongue.

These are the few pathological conditions that are known to produce some colour pigments on the tongue. Probably there may be a few more pathological conditions giving rise to such pigments. If there are, presumaly they are diagnosed as cases of tongue pigmentation in this study. But, this possibility does not appear to have effected this present study, for, a panel of medical experts could not identify any pathological condition giving rise to pigments on the tongue among the supposedly tongue pigmented people who were brought to their attention at the Andhra Medical College (Waltair, Andhra Pradesh). In fact, a small control study was carried out at the Andhra Medical College as follows. A team of medical experts, including physicians, pathologists and skin specialists, were taken around the in-patient wards of the Medical College Hospital, and all the in-patients were examined first by me for tongue pigmentation. Each of the many tongue pigmented people thus identified were subsequently examined by each member of the medical team separately for possible pathological conditions showing pigments on the tongue. At the end, the entire medical team sat together,

discussed their findings and came to the conclusion that none except one of the tongue pigmented people had any other pigment due to a pathological condition. The single case, however, showed one patch coming under tongue pigmentation, and a discoloured area on the tongue, which was attributed to leukaemia, with which the patient was affected. They also thoroughly examined the tongue pigmented people in order to arrive at a possible interpretation of this pigmentation. But for suggesting that this could be the melanin pigment, they could not contribute anything more on this aspect. It was reported by most of the members of the medical team that in the past also they had noticed such pigmentation in many patients, but not knowing its relevance to any disease, they were just ignoring it. All this makes the method of examination and identification of tongue pigmentation unambiguous.

Figure 1.1 presents three black and white photographs depicting one case with only spots and the other two with only a patch each. It may be noted that the true colour of tongue pigmentation, which is not possible to be brought in a black and white photograph, could not be caught even through colour photography. Figure 1.1 may be found on the inner side of the back cover.

Material studied :

Altogether 10 pedigrees each one covering 3 to 4 generations, were collected for this trait from Kerala and Andhra Pradesh. Data were also obtained on 406 families (a family consists of two living parents and atleast one living child), covering West Bengal, Andhra Pradesh, Madras, Kerala and Maharashtra, five states in India. Also, data were collected on 18 endogamous caste groups, 4 tribes and 17 more groups which are mixtures of caste groups from Orissa and the above-mentioned 5 states of India. Lastly, we have examined 845 new born babies at three hospitals of Calcutta. All this constitutes the entire material studied for the trait under study, which was collected during 1968-1970.

Concerning the consistency in the identification of tongue pigmentation, a small sample-check was carried out. Out of a total of 132 families studied in West Bengal, which forms a part of the total of 406 families studied for the trait, 27 families were re-examined with a time-lag of nearly two months between the two visits to each of the 27 families. The diagnosis at the second visit was consistent with that of the first, but for the case of an infant who was not available at the second visit.

All the materials studied for tongue pigmentation are analysed

part by part in the next three chapters in order to investigate whether or not this trait has any genetic basis.

Notation : Without any significance attached to the symbols right now, $\bar{A}$ and $\bar{a}$ are used to simply denote 'normal' and 'tongue pigmented' cases. Now on, unless stated otherwise, 'pigmented', 'affected' and 'trait' will be used to mean 'tongue pigmented', in the next three chapters.

It may be recorded here that after tongue pigmentation was first reported by Davis, several people were lead to think (personal communication to Davis) that this pigmentation might appear due to regular chewing of betel leaves. That this explanation is at best of very limited value will become evident once the next section of this chapter and the second section of third chapter are looked into. Next section reports the rates of incidence of this trait in some Indian populations. The distribution of this trait in the different Indian populations studied does not appear to be consistent with their chewing habits. For example, West Bengal shows a significantly smaller rate of incidence than that for the tribes of Andhra Pradesh studied, whereas, West Bengal has relatively more chewers of betel than the tribes of Andhra Pradesh. Also, many tongue pigmented

babies are reported in the first section of fourth chapter, babies having been examined a few hours after their birth. Therefore, explanations like chewing betel, food habits etc. can not account for the incidence of tongue pigmentation.

1.2 Genetics of hypertrichosis of the ear rims :

Hypertrichosis of the ear rims, corresponding to thick, long and dark hairs on the ear rims of men, is taken as a human genetic trait. These hairs may be present on one or both the ears, exhibiting great variation in the number of hairs. By intensity of this trait we shall mean the number of hairs. If both the ears are affected, their intensities are usually the same. This condition is inherited through the male line, from father to son. However, not all the sons of an affected father may exhibit the trait. Manifestation usually starts after 17 years of age.

It is not clearly known whether an affected man with a certain intensity exhibited the same intensity at the time of manifestation (age of onset), or that the intensity grew over age. A few available cases, however, go to the credit of the latter possibility. In my own case, I remember to have noticed just one hair on each of the

two ears sometime during 1968, and no hairs were noticed before 1967. Now there are three hairs on the left ear and two on the right. Similar observations were also reported by Dronamraju [8] and Gates et al [12]. More longitudinal studies are necessary to have any definite idea about the pattern of growth.

The total variation in the intensity of the trait may be artificially broken up into several grades, each grade corresponding to a certain range of variation for the intensity. Thus, depending on the intensity of the trait at the time of examination, an affected man is classified into one of the grades. It is felt that the intensity is also inherited, and therefore grading the intensity is meaningful.

Based on studies of huge pedigrees and family data, this trait was suggested by earlier workers as a Y-linked dominant trait, though some of them doubted it. This means that all the sons of an affe father should carry the dominant gene, not necessarily manifesting the trait. Incomplete penetrance of the dominant gene could very well suppress manifestation.

Method of examination :

Both the ears of each subject were personally examined for

the trait. Ear examination was done from a distance of not more than six inches. In most of the cases where the intensity of the trait is low, the head hairs were separated by hand and the ear rims were closely examined by putting hand on them. In any case, after a thorough examination, each subject was asked if he had the habit of plucking ear-hairs. Judgement about the intensity was based on this information also. The presence of atleast one coarse hair on atleast one ear, anywhere on the rim except the lobe, classifies a man as affected and otherwise normal.

The present study reports data on 168 families (fathers and adult sons) from West Bengal, which were collected during 1969. These data are analysed to study the genetics of the trait. With a brief review of the earlier findings, results of the present study are included in the fifth chapter.

1.3 Estimation of MNS-chromosome frequencies :

The genetics of MNS blood group system is so well known that it is not necessary to give an introduction about it. If the blood samples are tested using the three antisera anti-M, anti-N and anti-S, we have six phenotypes, which are popularly denoted by M, MS, MN, MNS, N and NS. The M-gene frequency is denoted by m, and

that of N by $n = 1 - m$. Also, m_s, m_S, n_s and n_S are used to represent frequencies of the chromosomes Ms, MS, Ns and NS respectively ($m_s + m_S = m$ and $n_s + n_S = n$). Thus, there are altogether 3 independent parameters in this model, say, m, m_s and n_s. Boyd [1] presented the maximum likelihood estimation of these parameters, under random mating. Two other versions of the same estimation problem are considered in the sixth chapter, after giving a brief account of Boyd's results.

CHAPTER II

TONGUE PIGMENTATION IN MAN

PRELIMINARY STUDIES

2.1 Distribution of the trait in some Indian populations:

The failure to identify tongue pigmentation as arising out of some visualizable factors necessitates further studies in new dimensions. It was felt that its distribution in different populations will throw some light on this area. Davis [5] has reported that the rate of incidence of this trait increases as one proceeds from north to south, at least in Indian populations, with a few exceptions. This, he explains, may correspond to some positive correlation of the trait with skin colour, though the implications are not clear. Therefore, several Indian populations were studied for the trait, in order to investigate the distribution pattern afresh.

Though many of the 'populations' studied here are meaningfully 'sub-divided populations' or 'endogamous groups', quite a few of them represent heterogeneous groups with mixed caste affiliations. The latter type of populations are studied here just for the sake of recording the distribution of the trait over different geographical regions. Altogether, 22 endogamous groups and 17 'other' populations are studied. The latter 17 populations are studied separately from the former 22 because

either the endogamous identity is unknown for such populations, or, due to smallness of the sample sizes several endogamous groups from the same geographical region are grouped together to constitute one 'population'. The 22 endogamous groups represent 18 caste groups and 4 tribal groups. Table 2.1 presents the necessary information in two parts, (a) dealing with the 22 endogamous caste groups, and (b) dealing with the remaining 17 groups.

Observe that Table 2.1 (a and b) assigns one identification number (first column) to each of the population groups. Hereafter, the populations will be referred to by their corresponding identification numbers. Numbers 1 to 18 correspond to the endogamous caste groups, and 19 to 22 represent the 4 tribes. Number 32 represents a mixture of two tribes, both from the Srikakulam district of Andhra Pradesh. The samples corresponding to numbers 6,7 and 19 to 22 are deliberately pooled over both males and females because, sex does not create differential rates of incidence for those populations. This aspect is discussed in details in the fourth section, wherein, these 6 populations are also dealt with. With this general introduction to the material studied, we shall now briefly describe the sources of these data.

All these 39 samples were personally studied during 1968-1970. The 6 populations studied from Maharashtra were represented in our samples by school-going students at Poona, whose ages ranged from 5 to 17 years. In West Bengal, 132 families (52 Brahmin and 80 Kayastha) were studied for this trait (vide section 4). 264 parents of these families, together with 4 more individuals, all above 20 years of age, constitute the samples for 6 and 7. Again, school-going students at Jeypore constitute the samples for 8,9,10,24 and 25, all between 7 to 19 years of age (mostly around 12 years). 35 families, representing many caste groups, were studied at Waltair (Andhra Pradesh), and the parents of these families form the samples for numbers 31 and 34. Their ages ranged from 19 years to 58 years. The rest of the samples from Andhra Pradesh caste groups (11 to 18, 26 to 30 and 33) were obtained from the students of Andhra University (Waltair), whose ages ranged from 17 years to 25 years. Samples for 35 and 36 consist of the parents of 79 families examined at Madras, which had mixed caste affiliations and whose ages ranged between 21 to 67 years. Samples from Kerala (37 to 39) were, again, represented by school-going students at Ernakulam, their ages ranging from 6 years to 16 years (mostly around 11 years). The tribes of Orissa (19 and 20) were studied in and around Jeypore, under Koraput district. They were all above 20 years of age. Finally, the tribes

of Andhra Pradesh (21 and 22) were studied in the agency areas of Srikakulam district (around G.L.Puram), who were all adults. It may be noted that the tribes studied here (19 to 22 and 32) are mostly confined to the districts of Srikakulam and Koraput and that these two districts are adjacent to one another.

In Table 2.1, all the populations coming from economically very backward classes are pointed out by asterisk (*) marks alongwith their identity numbers. This is done in accordance with the first chapter. The importance of the rates of incidence for such populations is left to one's own judgement.

The rates of incidence for two populations may be compared by using the statistic

$$(r_1 - r_2)^2 \;/\; (s_1^2 + s_2^2)$$

which follows a chi-square distribution with 1 d.f. A glance at Table 2.1 reveals a great deal of variation in the incidence rates, thus suggesting it as an Anthropological marker in human populations.

We shall now investigate whether or not the incidence rate increases as one proceeds from north to south, as was observed by Davis [5]. In Figure 2.1, we represent each of the 22 endogamous populations [presented in Table 2.1(a)] by a point on the vertical

line corresponding to the rate of incidence. These points are identified by the corresponding identity numbers (as given in Table 2.1), together with the geographical regions from where they were studied. W, E, SE and NE stand for West, East, South-eastern and North-eastern respectively. A study of the sequence of these symbols, from top to bottom in Figure 2,1, fails to support Davis' observation.

A close observation of the figure discovers several clusters or homogeneous groups of populations. For example, the rates of incidence for populations 1, 2, 6, 9 and 18 tend to cluster around a single point.

Before passing on to the next section, we shall attempt here to investigate whether or not there is any association between the rate of incidence and skin colour, a problem that was raised by Davis [5]. Let us consider the Parsis, West Bengal Brahmins, Andhra Pradesh Brahmins (pooled over Vaidiki and Niyogi) and all the four tribes (vide Table 2.1a). These 4 groups are ordered above in the decreasing intensity of fairness of the skin colour. The 4 x 2 contingency table (4 population groups and 2 categories $\bar{A}$ and $\bar{a}$ for the trait) yields $\chi^2 = 65.35$ (3 d.f., P is nearly zero). This provides a strong evidence in favour of an association between the rate of incidence and skin colour, thereby supporting

Davis' observation. It is clear that the rate of incidence is higher for dark skinned populations than that for fair skinned. Since the skin colour is controlled by the amount of melanian pigment, it appears that tongue pigmentation and melanin are closely associated, if not the same.

2.2 Pedigree study:

This section presents a study of altogether ten pedigrees collected from two states of India, Kerala and Andhra Pradesh. The purpose of this study is to detect the probable genetic mechanism, if any, controlling the trait under study. It may be noted that though there is nothing unusual with the unequal distribution of the trait in some Indian populations, as was seen in the previous section, our failure to ascribe any simple reason for its occurrence seems to necessitatefurther studies on it.

The ten pedigrees are presented in Figures 2.2 to 2.11. The symbols used in the pedigrees are explained in Table 2.2. Figure 2.11 runs over five living generations, Figures 2.3 and 2.4 run over 3 generations, the remaining running over 4 generations. The propositi are shown by arrow marks. The populations from where these pedigrees were collected are indicated in Table 2.3. As a first step, it is felt useful to compare the rates of incidence in the pedigrees with those obtained through population

studies (Table 2.1) for the same populations. This analysis is clearly set out in Table 2.3, where, n_1 stands for the number of individuals who were personally examined for the trait; the rate of incidence in a pedigree (r_1) is computed by treating the doubtful cases (marked by ? in the pedigrees) and the unexamined but pigmented cases (indirect information, marked by partial dark shade in the pedigrees) as normal ones. Thus, r_1 represents an underestimate. The other symbols in Table 2.3 carry the usual meaning. Out of the ten pedigrees, five show significantly higher rates of incidence than those at the population level. This provides strong evidence to suspect some genetic basis for the trait. However, since we identify only two phenotypes, there is every reason to think in terms of a single gene hypothesis for the trait. We shall consider autosomal (dominant and recessive), autosomal and sex-limited, Y-linked (dominant and recessive) and X-linked (dominant and recessive) hypotheses as the several possibilities under the single gene theory. It is abundantly clear from the pedigrees that Y-linkage and sex-limited hypotheses can not be supported as possible genetic models for the inheritance of tongue pigmentation. Let us now consider X-linked recessivity as a possibility. Let 'A' & 'a' represent the dominant and recessive alleles respectively. It is clear that,

under such a genetic hypothesis, all the sons should be affected (X_aY) and all the daughters should be normal ($X_A X_a$) from each mating of the type "affected female x normal male" (X_aX_a x X_AY). But, all the ten pedigrees studied provide evidence against such an expectation, thereby totally rejecting X-linked recessive hypothesis. Similarly, under X-linked dominant hypothesis, we expect all the sons to be normal (X_aY) and all the daughters to be affected ($X_A X_a$) out of every mating of the type "normal female x affected male" (X_aX_a x X_AY). Neither this expectation is realised in the pedigrees. Thus, X-linkage is totally ruled out as a possible genetic hypothesis for the inheritance of the trait. Now, let us look at the autosomal dominant hypothesis. Under this, all the children out of "normal x normal" matings (aa x aa) are expected to be normal (aa). This is not supported by the pedigrees, though there are some individual families in the pedigrees that go to the credit of such a hypothesis. Thus at least by enumeration process we arrive at the autosomal receiving as the only possible simple genetic hypothesis for the inheritance of the trait. We shall often refer to this single locus diallelic autosomal recessive hypothesis simply as recessive hypothesis.

Now, under the recessive hypothesis, the two phenotypes are $\bar{A}$ (AA and Aa, normal) and $\bar{a}$ (aa, pigmented). All the pedigrees

provide very good evidence in favour of this hypothesis. Nevertheless, there are several $\bar{a} \times \bar{a}$ matings giving rise to at least one $\bar{A}$ child, which is in contrary to the recessive hypothesis. From now on, we shall refer to such cases as 'exceptional' cases. Notice that these exceptional cases are not many in number, and that they may be disposed off on some grounds like illegitimacy or some variation in the genic action. In all, we have 8 individual families ($\bar{a} \times \bar{a}$) in the ten pedigrees, which produced some exceptional cases. Altogether 25 children in these 8 families were examined, out of whom 9 turned out to be normal. This high proportion of exceptional cases within $\bar{a} \times \bar{a}$ families (nearly 1/3) certainly can not be explained under illegitimacy alone. It may, however, be recorded here that the last seven pedigrees (Figures 2.5 to 2.11) were collected from very much underdeveloped villages, where illegitimacy is known to exist to a considerable extent. However, in none of these cases serological tests could be performed. On the other hand, the pedigree in Figure 2.3 comes from a business community (Hindu) known to have a peculiar marriage custom prevalent among them, particularly in earlier days, by which an unmarried younger brother could be the real father of any of the children born to the wives of his elder brothers. I have gathered this information from the local people, but was unable to investigate this matter personally.

The propositus (arrow marked) had two unmarried younger brothers either of whom could also account for the 13 years old normal daughter born to the wife of the propositus. The author's discussions with the propositus and his brothers were quite revealing and supported the above contention. The question, however, remained undecided as to whether the said girl showed a violation of the proposed genetic hypothesis or not.

In any case, barring the exceptional cases, the pedigrees support the recessive hypothesis for the inheritance of the trait.

2.3 Family data and testing for non-genetic chance hypothesis:

In this section we shall present data on 406 Indian families (parents and children) and investigate whether a non-genetic chance hypothesis can account for the incidence of the trait. The chance hypothesis may be stated as "the occurrence of tongue pigmentation in a population is determined purely by chance without any influence of heredity".

Data on 406 families are presented in Table 2.4, which presents $\bar{a} \times \bar{a}$ families each with at least one $\bar{A}$ child separately. This includes 134 families from West Bengal (represented by two caste groups, Kayastha and Brahmin), 137 from Maharashtra

(represented by two caste groups, Maratha and Brahmin), 56 from Andhra Pradesh and 79 from Madras. The last two groups represent two heterogeneous samples having mixed caste affiliations.

In the family data of Table 2.4, under the chance hypothesis, the ratio between affected and normal children should come out all the same irrespective of the mating type. A simple test for this is provided by a study of the association between the mating type and the offspring phenotype. Considering the data on the offspring alone of Table 2.4, one obtains from the corresponding 3 x 2 contingency table $\chi^2 = 194.29$ ($P \approx 1$). This rules out the chance hypothesis beyond any doubt. It may be noted that, in the above contingency table, all the $\bar{a} \times \bar{a}$ families were pooled together.

We shall, however, go a step further to examine the family data in more details, in order to come to a decisive conclusion about the relevance of the chance hypothesis for the trait. The probability of an individual's tongue being normal ($\bar{A}$) is estimated from the parental data of Table 2.4 as

$$t = \frac{644}{812} = 0.793$$

Let us consider a family of any mating type having c children

(c stands for the number of children actually examined for the trait). Under the chance hypothesis, the probability that such a family has at least one pigmented child is simply given by

$$\pi_c = 1 - t^c$$

Let,

N_c = Observed number of families (of all mating types) of size c.

R_c = Observed number of families (of all mating types) of size c, each family containing at least one pigmented child.

Then, for each value of c, provided that the numbers involved are large enough, the statistic for testing the validity of the chance hypothesis is given by

$$(R_c - N_c\pi_c)^2/N_c\pi_c(1 - \pi_c)$$

which follows a chi-square distribution with 1 d.f. The details of the analysis are incorporated in Table 2.5. It may be remarked here that the analysis for c = 1 under the chance hypothesis coincides exactly with the one under the hypothesis of recessive inheritance for the trait. Therefore, the single sib families have been left out of Table 2.5. All this analysis, thus, provides very strong evidence against the non-genetic chance hypothesis.

Therefore, the recessive hypothesis remains as the only possibility. However, under the recessive hypothesis, the five normal children out of $\bar{a} \times \bar{a}$ matings (Table 2.4) again turn out to be exceptional cases. These cases may be explained on the same grounds as we did in the second section. Our inclination to treat these cases as arising out of assignable causes and not real contradictions to the recessive inheritance, stems from the fact that no other hypothesis, including the non-genetic one, can account for the incidence of the trait. Clearly enough, we are concentrating only on single gene models. More attention will be paid to the problem of exceptional cases in the fourth chapter.

Note that, depending on the context, the symbols $\bar{A}$ and $\bar{a}$ are used to mean either normal and pigmented, or dominant and recessive respectively.

2.4 Variations due to sex and age:

As the next step, we shall now investigate as to whether or not the two factors sex and age create any variations in the rates of incidence of the trait, before proceeding to the detailed genetic studies on the trait. The data for this study are taken from Tables 2.1 and 2.4.

2.4.1 Variations due to sex:

This study is based on the parental data (Table 2.4) from West Bengal, Maharashtra and Madras, and also on the population data on the four tribal groups studied in the first section (numbers 19 to 22 in Table 2.1.a). Note that, out of a total of 134 families from West Bengal, 52 came from Brahmin community and 82 from the Kayastha. The parental data on these two caste groups are separately studied in this section. Similarly, the 137 families from Maharashtra correspond to 34 Brahmin families and 103 Maratha families, which are separately studied here. The 79 families from Madras represented many castes and they are all pooled as such. The analysis of all these 5 caste groups and the 4 tribal groups is presented in Table 2.6. But for the mixed group of Madras which yields $\chi^2_1 = 5.92 (P < 0.02)$, no other sample provides evidence in favour of any association between sex and the trait. The genetic heterogeneity of the Madras sample could be responsible for the outcome. Therefore, it is concluded that the trait occures equally frequently among both males and females, which is also a necessary condition for any autosomal inheritance model.

2.4.2 Variations due to age : The four proper caste groups of West Bengal and Maharashtra, excluding the mixed sample of Madras, and the 146 school-going students of Parsi community (number 1 of Table 2.1.a) are studied for this problem. It may be noted that the mixed Madras sample is not analysed here since anyway the heterogeneous group does not serve much of a purpose. Also, the 4 tribal groups are not included in this study because the ranges of age variations for these samples are rather narrow. Thus, only five samples are analysed here, the details of which are given in Table 2.7. All this suggests that neither age is associated with the incidence rate of tongue pigmentation. However, it does not rule out the possibility of an association between the two at younger age groups, since our study is limited mostly to adults. Kirk [13] detected strong positive association between the two factors, based on a study of 470 individuals from an Australian aboriginal community. Most of them had ages between 6 to 15 years. Also, Davis [6] found similar association among Kindergarten students. Nevertheless, for the time being, we shall confine our attention only to what is suggested by the present studies. This aspect will again be considered in the fourth chapter.

Though the chances of misclassification are remote, in view of the possible confusion between tongue pigmentation and pseudo-pigments due to malnutrition, a well-planned medical study of both normal and pigmented individuals will be of considerable importance.

The pattern of the distribution of the trait in different populations might suggest the origin of the pigmentation gene and its migration path. However, more populations should be studied before attempting to tackle this problem, especially the Mongoloid and the Negroid races.

The problem of association between age and the trait can be tackled better by studying all age groups in an endogamous population.

TABLE 2.1

(a) Distribution of tongue pigmentation in some Indian endogamous groups.

No.	State	Population	Sex	sample size n	Incidence rate r(%)	Variance s^2
1	Maharashtra	Parsi	M	146	8.22	5.15
2	"	C.K.P.	M	108	9.26	7.78
3	"	Brahmin	M	571	13.13	1.99
4	"	Maratha	M	832	17.07	1.69
5	"	Mali	M	35	20.00	45.70
6	West Bengal	Brahmin	M+F	104	9.62	8.36
7	"	Kayastha	M+F	164	18.29	9.11
8	Orissa	Brahmin(Holua)	F	89	23.60	20.26
9	"	Brahmin(Damua)	F	60	10.00	15.00
10	"	Karan	F	42	16.67	33.07
11	Andhra Pradesh	Brahmin(Niyogi)	M	68	17.65	21.37
12	"	Brahmin(Vaidiki)	M	88	17.04	16.06
13	"	Vysya	M	37	13.51	31.58
14	"	Kamma	M	73	23.29	24.47
15	"	Velama	M	13	23.08	136.56
16	"	Naidu	M	25	20.00	64.00
17	"	Kapu	M	44	29.54	47.30
18	"	Reddy	M	30	10.00	30.00
19*	Orissa	Gadaba	M+F	60	43.33	40.92
20*	"	Parja	M+F	69	30.43	30.68
21*	Andhra Pradesh	Jatapu	M+F	160	37.50	14.65
22*	"	Savara	M+F	107	31.78	20.26

* Economically backward communities (see page 6)

Symbols used : M and F for Male and Female.

TABLE 2.1(Contd.)

(b) Distribution of tongue pigmentation in some Indian populations (mixed).

No.	State	Population	Sex	n	r(%)	s^2
23	Maharashtra	Others	M	260	15.77	5.11
24	Orissa	Sarua & Kottigiya Brahmin	F	18	5.56	29.17
25	"	All	M	73	8.22	10.33
26	Andhra Pradesh	Other Brahmin	M	34	17.65	42.75
27	"	K + Y + S	M	31	22.58	56.39
28	"	Christian	M	36	30.56	58.95
29†	"	Scheduled Caste	M	28	32.14	77.89
30	"	Unknown	F	61	19.67	25.90
31	"	Unknown	M	35	20.00	45.70
32†	"	KD + JD	M	25	24.00	72.96
33*	"	Harijan & Girijan	M	18	22.22	96.02
34	"	Unknown	F	35	28.57	58.37
35	Madras	Unknown	M	79	11.39	12.78
36	"	Unknown	F	79	26.58	24.70
37	Kerala	Christian	M	268	19.78	5.90
38	"	Hindu	M	231	24.24	7.95
39	"	Others	M	32	12.50	34.22

† Many are from economically backward families

Symbols used : K for Kshatriya; Y for Yadava; S for Sali; KD for Konda Dora; JD for Jatapu Dora

TABLE 2.2

Explanation of the symbols used in Figures 2.2 to 2.11

□	:	Male
○	:	Female
D	:	Dead
N.S.	:	Not seen
N.A	:	Not available
Numbers below the individuals	:	Corresponding ages
Dark shade	:	Presence of tongue pigmentation, confirmed through personal examination
Partial dark shade	:	Presence of tongue pigmentation, confirmed only through hear-say information
Question mark(?)	:	Examined, but doubtful
Arrow mark (→)	:	Propositus

TABLE 2.3

Comparison of incidence rates obtained through pedigree and population studies

Population studied	Pedigree study				Population study			χ^2_1
	Figure no.	n_1	r_1	s^2_1	n_2	r_2	s^2_2	
Hindu (Kerala)	2.2	78	38.46	30.34	231	24.24	7.95	5.28
Hindu (Kerala)	2.3	49	67.35	44.88	231	24.24	7.95	35.18
Brahmin (Madras)	2.4	15	26.67	130.38	158*	18.99	9.74	0.42
Sali (Andhra Pradesh)	2.5	54	16.67	25.72	31**	22.58	56.39	0.42
Kapu (Andhra Pradesh)	2.6	46	32.61	47.77	44	29.54	47.30	0.10
Jalari (Andhra Pradesh)	2.7	122	27.05	16.17	-	-	-	-
Vysya (Andhra Pradesh)	2.8	80	25.00	23.44	37	13.51	31.58	2.40
Yadava (Andhra Pradesh)	2.9	26	50.00	96.15	31**	22.58	56.39	4.93
Kapu (Andhra Pradesh)	2.10	51	62.74	45.84	44	29.54	47.30	11.83
Reddy (Andhra Pradesh)	2.11	51	33.33	43.57	30	10.00	30.00	7.40

* Corresponds to a mixed sample from Madras, as given in Table 2.1(b).

** Corresponds to a mixed sample from Andhra Pradesh given as K+Y+S in Table 2.1(b).

TABLE 2.4

Data on tongue pigmentation in 406 families from India

Parents		Children	
Mating type	No.	$\bar{A}$	$\bar{a}$
$\bar{A} \times \bar{A}$	257	645	90
$\bar{A} \times \bar{a}$ (including $\bar{a} \times \bar{A}$)	130	261	73
$\bar{a} \times \bar{a}$	15	0	41
$\bar{a} \times \bar{a}$	1	1	1
$\bar{a} \times \bar{a}$	1	2	1
$\bar{a} \times \bar{a}$	1	1	1
$\bar{a} \times \bar{a}$	1	1	2

TABLE 2.5

Analysis of 406 families under non-genetic chance hypothesis

Sibship size c	π_c	N_c	$N_c\pi_c$	R_o	$N_c\pi_c(1-\pi_c)$	χ^2	d.f.	P
2	0.371	136	50.456	34	31.737	8.53	1	< 0.004
3	0.501	99	49.599	39	24.750	4.54	1	< 0.04
4	0.604	62	37.448	25	14.829	10.45	1	< 0.002
5	0.686	30	20.580	12	6.462			
6	0.751	10	7.510	6	1.870	11.84	1	< 0.0007
7	0.803	4	3.212	3	0.633			
Totals	-	341	168.805	119	80.281	35.36	4	< 0.00001

For the totals : $\chi^2 = \frac{(119-168.805)^2}{80.281} = 30.90$ (1 d.f., P $\approx$ 1)

TABLE 2.6

Variation due to sex in the incidence of tongue pigmentation

Population	Pigmented		Normal		χ^2_1	P
	Male	Female	Male	Female		
Kayastha	11	19	71	63	2.61	> .09
Brahmin (West Bengal)	5	5	47	47	0	1
Maratha	24	31	79	72	1.22	> .25
Brahmin (Maharashtra)	6	8	28	26	0.36	> .53
Mixed group (Madras)	9	21	70	58	5.92	< .02
Jatapu (Andhra Pradesh)	48	12	81	19	0.02	> .88
Savara (Andhra Pradesh)	18	16	44	29	0.51	> .44
Gadaba (Orissa)	17	9	25	9	0.46	> .48
Parja (Orissa)	15	6	43	5	3.59	> .05

TABLE 2.7

Variation due to age in the incidence of tongue pigmentation

Population	Age group	Pigmented	Normal	χ^2	d.f.	P
Kayastha	20-30	9	38	0.90	2	> .63
	30-40	12	44			
	≥ 40	9	52			
Brahmin (West Bengal)	< 40	5	44	0.04	1	> .84
	≥ 40	5	50			
Maratha	20-30	8	23	6.25	3	> .09
	30-40	21	56			
	40-50	19	32			
	≥ 50	7	40			
Brahmin (Maharashtra)	< 40	7	21	0.57	1	> .43
	≥ 40	7	33			
Parsi	5-10	4	77	2.60	1	> .09
	10-18	8	57			

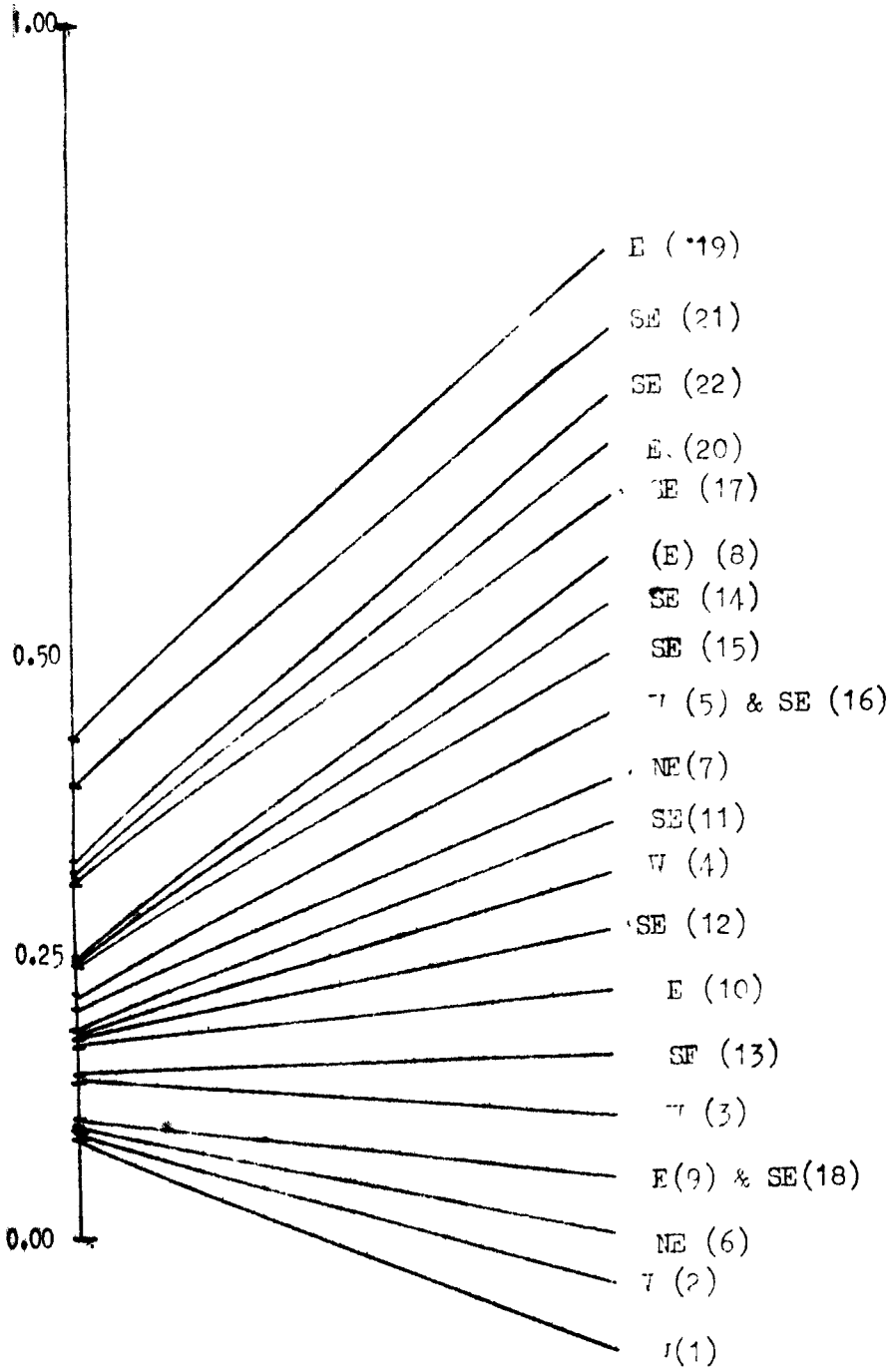

1.00
0.50
0.25
0.00
E (19)
SE (21)
SE (22)
E (20)
E (17)
(E) (8)
SE (14)
SE (15)
(5) & SE (16)
NE(7)
SE(11)
W (4)
SE (12)
E (10)
SE (13)
(3)
E(9) & SE(18)
NE (6)
(2)
(1)

FIGURE 2.1

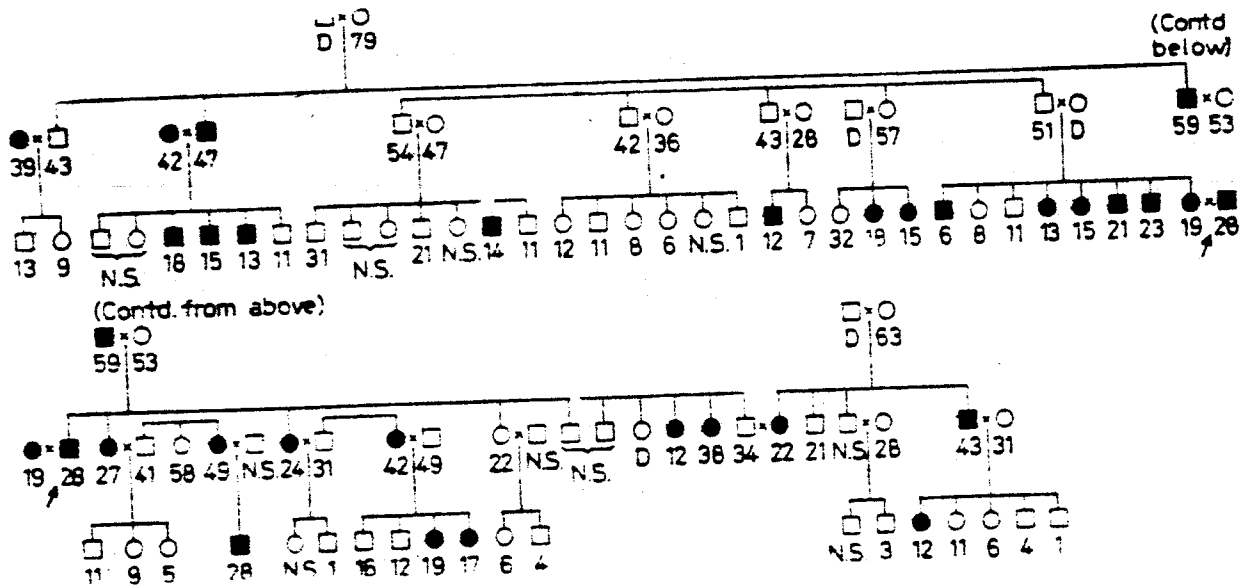

D 79
(Contd below)
39 43
42 47
54 47
42 36
43 28
D 57
51 D
59 53
13 9
N.S.
18 15 13 11 31
N.S.
21 N.S. 14 11 12 11 8 6 N.S. 1 12 7 32 19 15 6 8 11 13 15 21 23 19 28
(Contd. from above)
59 53
D 63
19 28 27 41 58 49 N.S. 24 31 42 49 22 N.S. N.S. D 12 38 34 22 21 N.S. 28 43 31
11 9 5 28 N.S. 1 16 12 19 17 6 4
N.S. 3 12 11 6 4 1

FIGURE 2.2

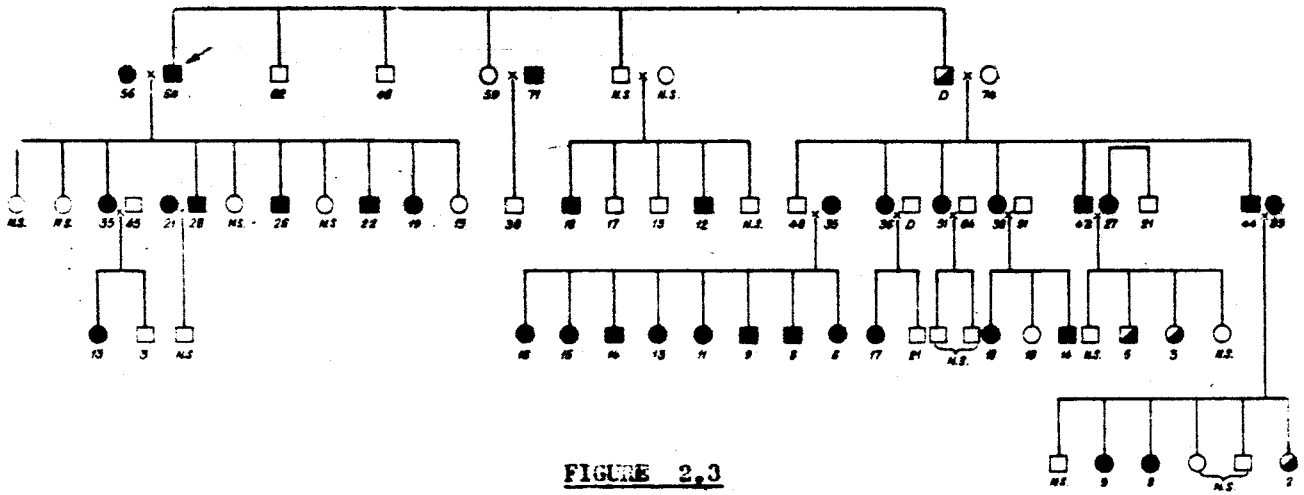


FIGURE 2.3

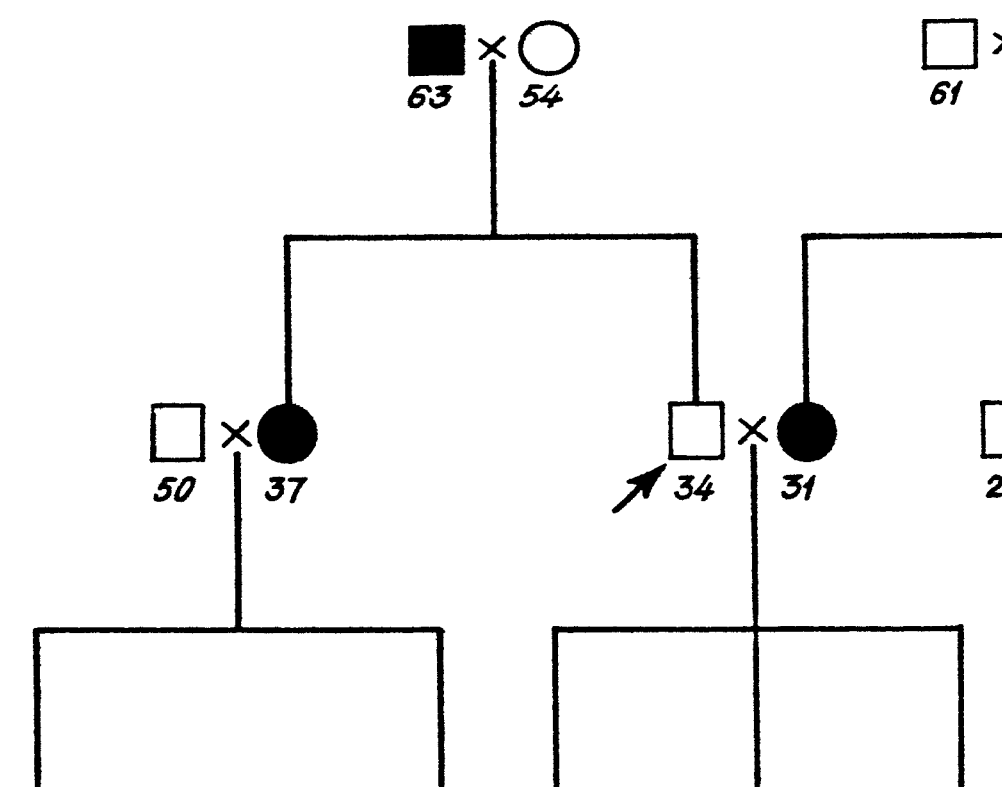


FIGURE 2.4

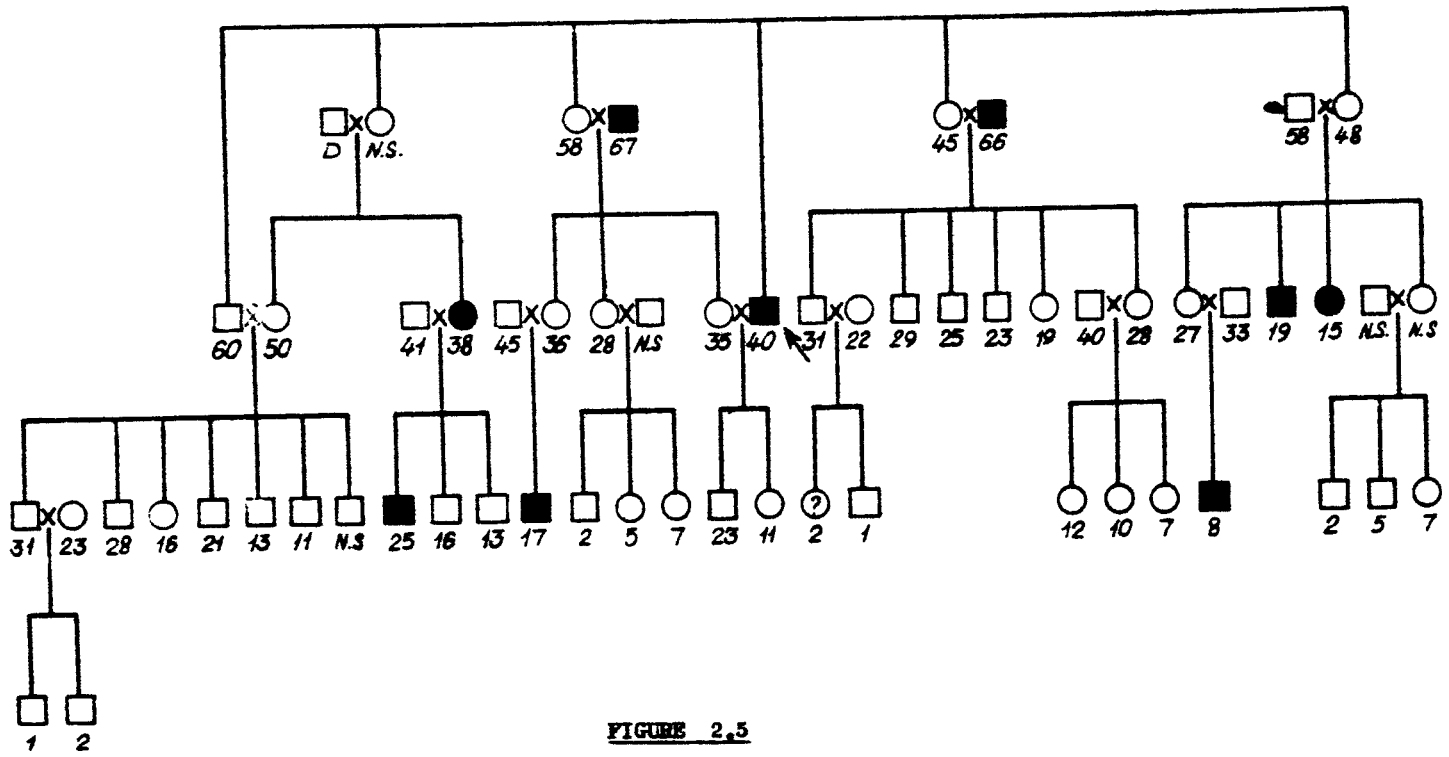

D N.S.
58 67
45 66
58 48
60 50
41 38
45 36
28 N.S
35 40
31 22
29 25 23 19 40 28
27 33 19 15 N.S. N.S
31 23 28 16 21 13 11 N.S
25 16 13 17
2 5 7
23 11
2 1
12 10 7 8
2 5 7
1 2

FIGURE 2.5

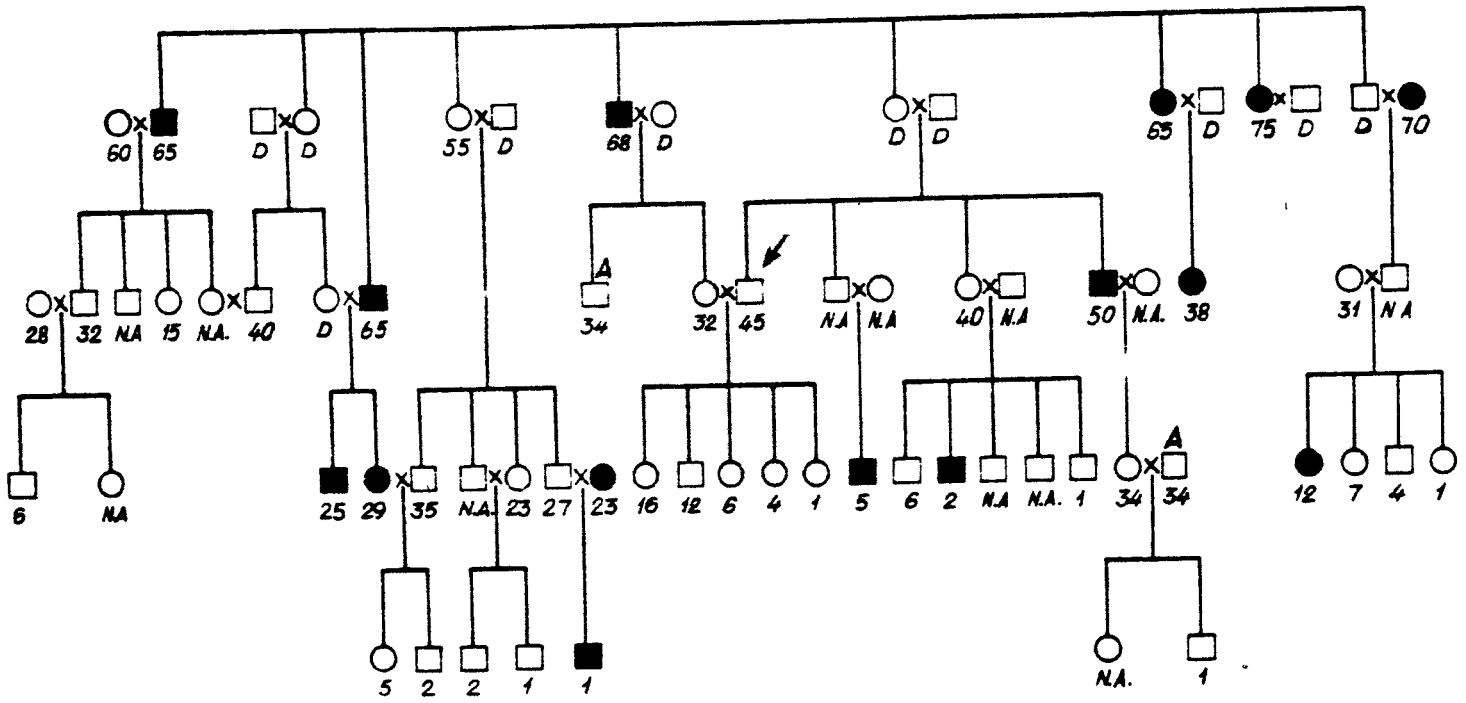

60 65
D D
55 D
68 D
D D
65 D
75 D
D 70
28 32 N.A 15 N.A. 40
D 65
A
34
32 45
N.A N.A
40 N.A
50 N.A. 38
31 N.A
6 N.A
25 29 35 N.A. 23 27 23
16 12 6 4 1
5 6 2 N.A N.A. 1
34 34
12 7 4 1
5 2 2 1 1
N.A. 1

FIGURE 2.6

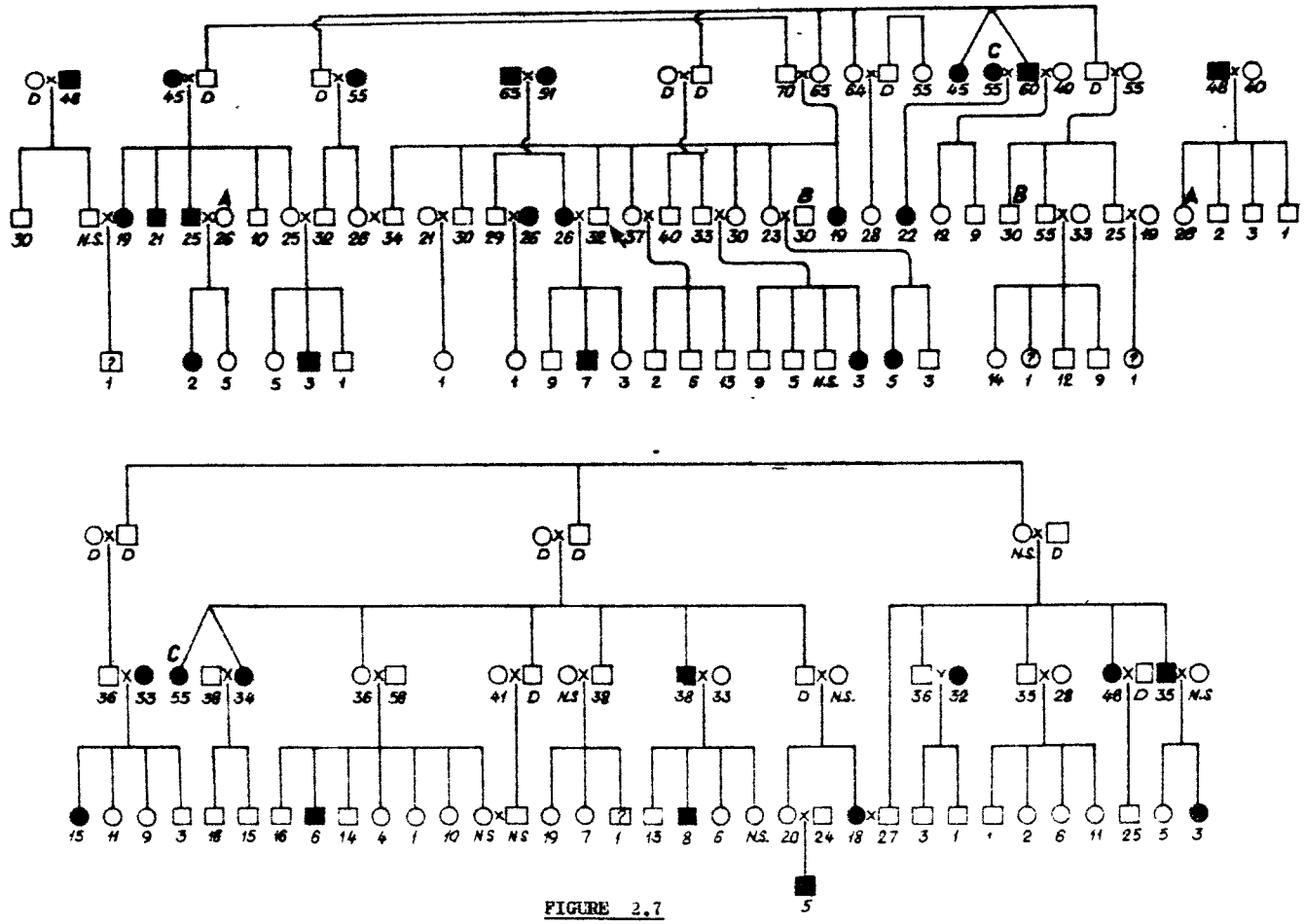


FIGURE 2.7

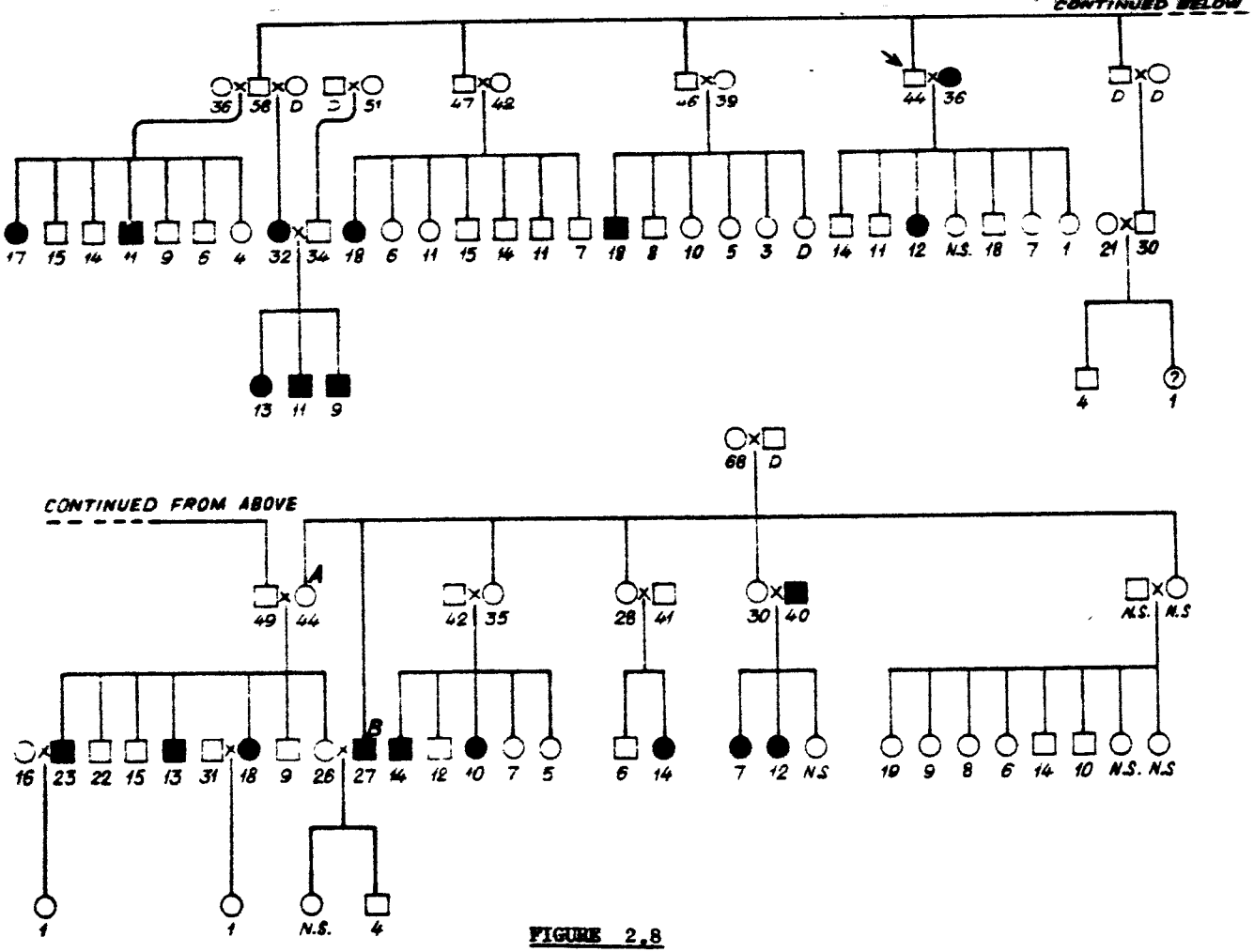

CONTINUED BELOW
36 58 D 51
47 42
46 39
44 36
D D
17 15 14 11 9 6 4 32 34 18 6 11 15 14 11 7 18 8 10 5 3 D 14 11 12 N.S. 18 7 1 21 30
13 11 9
4 1
68 D
CONTINUED FROM ABOVE
49 44
42 35
28 41
30 40
N.S. N.S
16 23 22 15 13 31 18 9 26 27 44 12 10 7 5 6 14 7 12 N.S 19 9 8 6 14 10 N.S. N.S
A
B
1 1 N.S. 4

FIGURE 2.8

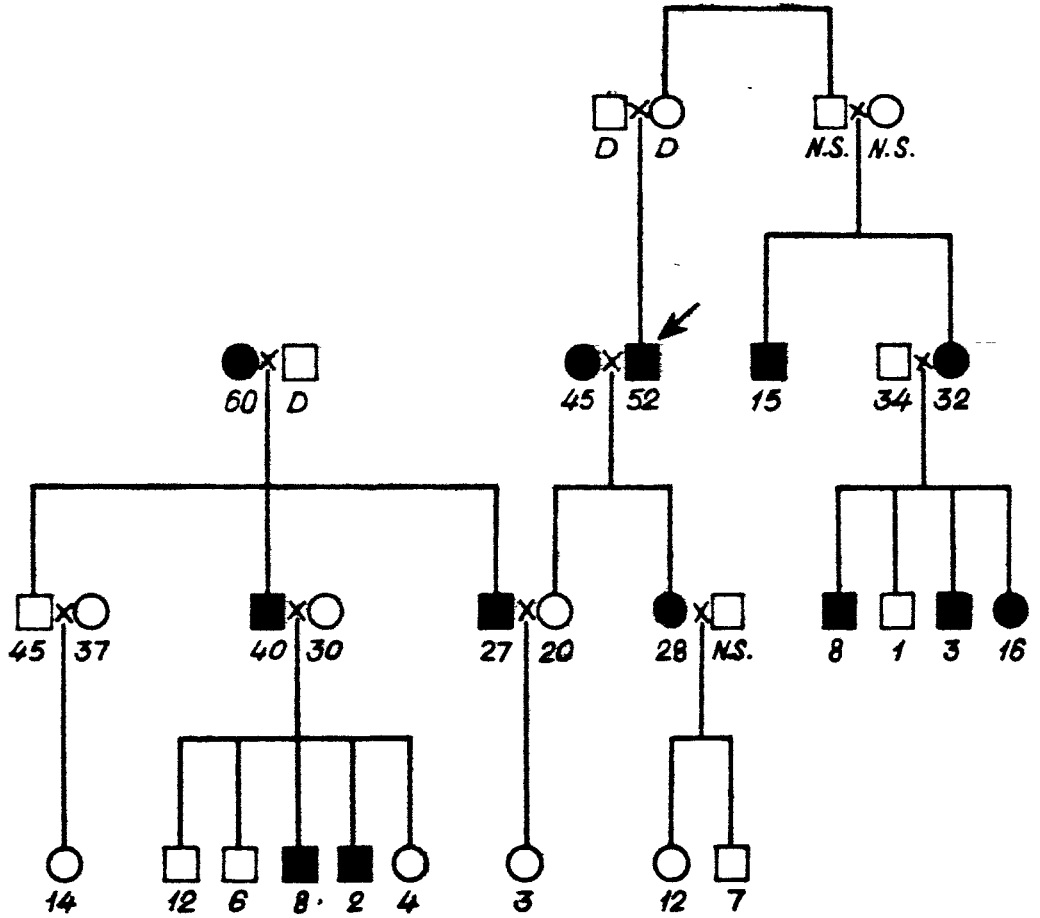

D
D
N.S.
N.S.
60
D
45
52
15
34
32
45
37
40
30
27
20
28
N.S.
8
1
3
16
14
12
6
8
2
4
3
12
7

FIGURE 2.9

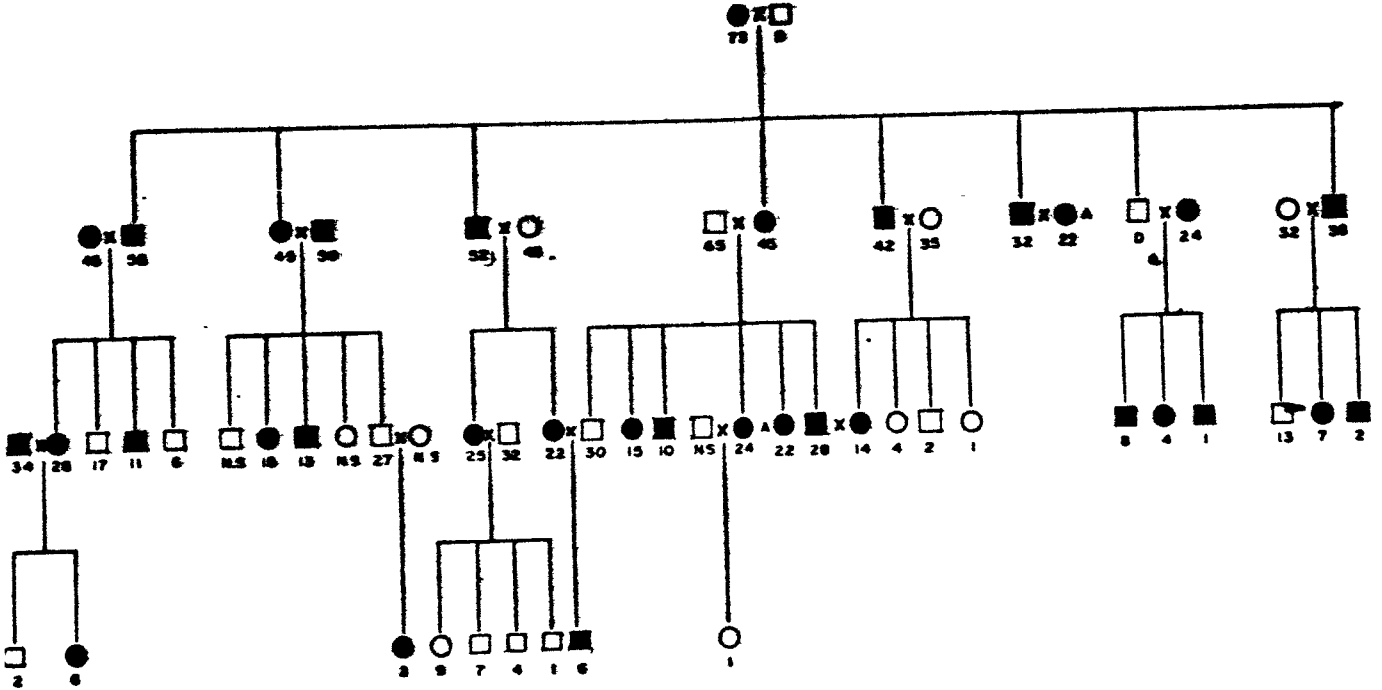


FIGURE 2.10

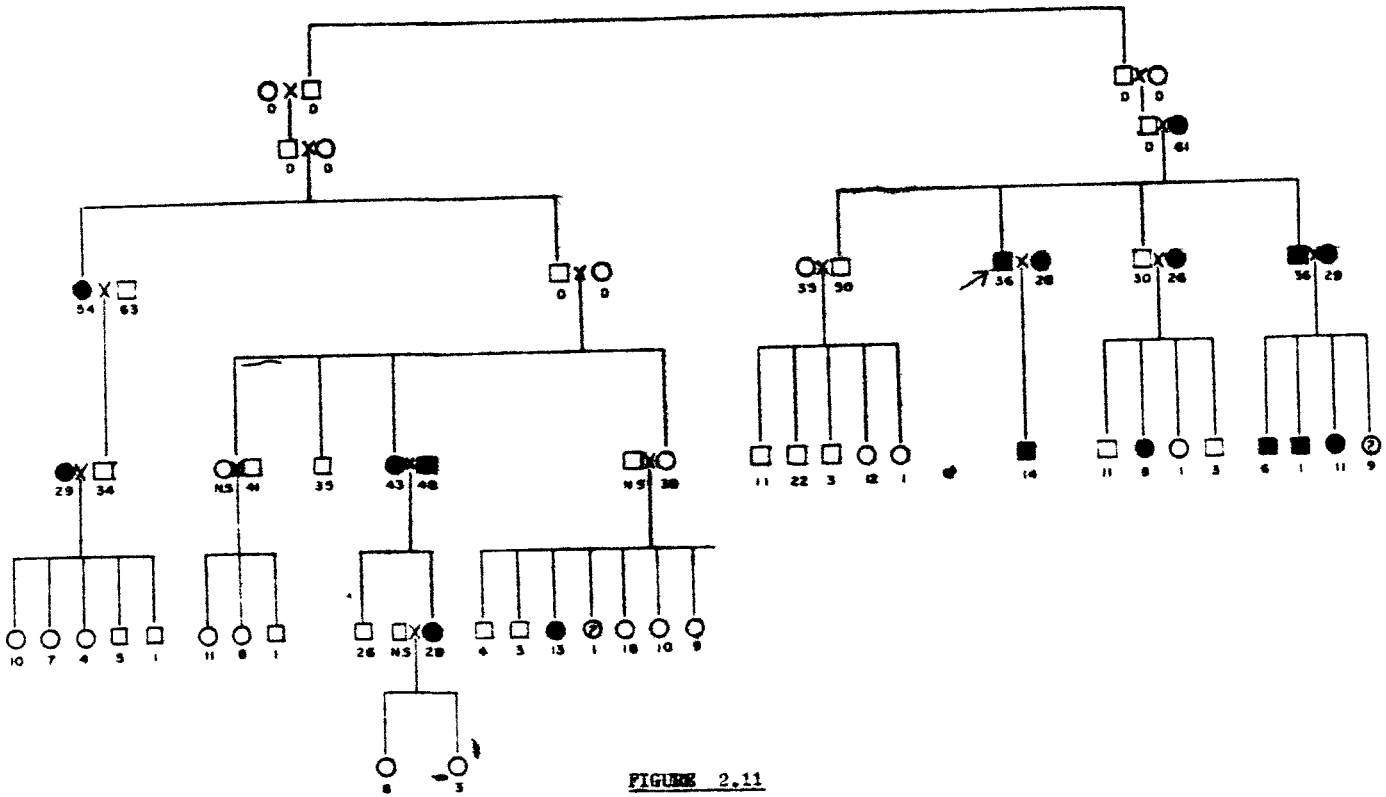


FIGURE 2.11

CHAPTER III

TONGUE PIGMENTATION IN MAN

STATISTICAL ANALYSIS OF FAMILY DATA

Having rejected the chance hypothesis for the occurrence of tongue pigmentation in human populations, it is clear that there are some systematic forces that control the incidence of the trait. Of the several possible simple genetic models considered in the previous Chapter, all the ten pedigrees provided very good evidence in favour of the diallelic autosomal recessive hypothesis for this trait. Nevertheless, one can not afford to come to such a rash conclusion on the basis of a study of just ten pedigrees. Inspite of the good evidence these pedigrees provide in favour of the recessive hypothesis, it could turn out to have a different mode of inheritance, should it at all be an inherited condition. The chances for this alternative possibility are rather considerable. On the other hand, these pedigrees provide good motivation for further search on the genetics of this trait. Also, the accumulative evidence contained in the pedigrees is strongly suggestive of only the simple recessive hypothesis, and that no other simple genetic theory seems to hold good. Therefore, it is likely that this trait is inherited either as a simple recessive condition or has some

complicated genetic basis. But, the fact that there are only two phenotypes (pigmented and normal) makes it rather unbelievable that it could have a complicated genetic basis. This, in any case, does not rule out such a possibility, but merely makes it less probable. Should more phenotypes be identified for this trait (say, based on factors like size and shape of the pigmented areas), a complex genetic theory for it would sound more meaningful. Therefore, it was decided to make detailed analysis of the family data presented in the previous Chapter in order to verify the suggested autosomal recessive hypothesis.

In what follows, expected frequencies for different possible events are obtained under the recessive hypothesis and they are compared with the corresponding observed frequencies through 'tests of significance'. Thus, this problem is tackled as a 'testing' problem in the next two sections, and the same is translated into an 'estimation' problem in the third section.

3.1 A priori expectation of the number of recessives in sibships of different sizes :

Assuming tongue pigmentation to be an autosomal recessive condition, the two phenotypes are denoted by $\bar{A}$ (AA, Aa) and

$\bar{a}$ (aa), where a is the recessive or the 'pigmentation' allele. Throughout the analysis that follows, we shall restrict to only those families each of which have atleast one pigmented child. Thus, by a family in this section we mean a family with the above condition, unless specified otherwise. The analysis is carried out separately for families of the two mating types $\bar{A} \times \bar{A}$ and $\bar{A} \times \bar{a}$ (including $\bar{a} \times \bar{A}$) based on a priori expectation of the number of recessives in sibships of different sizes. A good account of the method of analysis is found in Smith [35].

3.1.1. Analysis of data on $\bar{A} \times \bar{a}$ families.

Observe that all the families of the mating type $\bar{A} \times \bar{a}$ are in fact Aa x aa, since each such family has atleast one $\bar{a}$ child. The probability of a child from such a family being pigmented is 1/2. Therefore, the probability of getting atleast one pigmented child in a sibship of size c is given by $1 - (1/2)^c$. Also, the probability of getting exactly r pigmented sibs in such a sibship is given by

$$\binom{c}{r} (1/2)^c \; / \; [1 - (1/2)^c]$$

the mean a_c and the variance b_c of this truncated binomial

distribution are seen to be

$$a_c = \frac{c}{2[1 - (1/2)^c]}$$

and $$b_c = a_c \left[\frac{1 + c}{2} - a_c\right]$$

Numerical values of a_c and b_c for different values of c were tabulated by Smith [35].

Let

m_c = Observed number of families each having c children

R_c = Observed number of pigmented children in all families, each with c children

$E_c = a_c\, m_c$ = Expected number of pigmented children out of all families, each with c children.

$V_c = b_c\, m_c$ = the corresponding variance.

Observed and expected frequencies are now compared for each value of c, to test for their agreement. The test statistic is

$$(R_c - E_c)^2 / V_c$$

which follows a χ^2 distribution with 1 d.f. Also, the totals

are tested for the overall fit using a statistic similar to the above one. There are altogether 49 families of the mating type Aa x aa. Analysis of this data, carried out as described above, is presented in Table 3.1. Thus, this analysis provides very good evidence in favour of the autosomal recessive hypothesis. It is of concern, however, that the chi-square value corresponding to sibships of size 2 is significant (4.79, $P < 0.03$). This is indicative of some type of heterogeneity in the data. However, the overall fit is quite satisfactory.

3.1.2 Analysis of data on $\bar{A} \times \bar{A}$ families.

Note that all these families are of the type Aa x Aa, since each of these families contains atleast one recessive child. Such a family produces a pigmented child with 1/4 probability and a normal child with 3/4 probability. Therefore, the probability of getting atleast one pigmented child in a sibship of size c is $1 - (3/4)^c$, and the probability of getting r pigmented sibs in a sibship of size c is

$$\binom{c}{r} (1/4)^r (3/4)^{c-r} / \left[1 - (3/4)^c\right]$$

from where one can get the mean A_c and the variance B_c, which

are given by

$$A_c = \frac{c}{4\left[1 - (3/4)^c\right]}$$

and $$B_c = A_c\left[\frac{3 + c}{4} - A_c\right]$$

Smith [35] has tabulated the values of A_c and B_c for different values of c.

Let

M_c = Observed number of families, each having c children

R_c = Observed number of pigmented children in those M_c families

$E_c = A_c M_c$ and $V_c = B_c M_c$.

Agreement between the corresponding observed and expected frequencies are now tested using the same statistic as given in 3.1.1. In all, we have data on 63 $\bar{A} \times \bar{A}$ families, each having atleast one pigmented child. Analysis of the data on the above lines is set out in Table 3.2. One finds convincing evidence from this analysis in favour of the recessive hypothesis.

Though so far things went on rather smoothly inspite of the discrepency detected in $\bar{A} \times \bar{a}$ families with two children each,

the observations on $\bar{a} \times \bar{a}$ families appear to challenge the inference concerning the appropriateness of the recessive hypothesis. Though illegitimacy can account for the 5 normal children born to pigmented parents (see Table 2.4), it can not be taken for granted since no serological tests were performed on those families. For the time being we explain these 'exceptional' cases as arising out of either extra-marital influence or some variation in the genic action. The latter possibility will be explored in the next Chapter.

3.2 Further analysis of the family data :

The previous section presented analysis of the family data which was aimed at studying only the segregation ratios through a purely genetic approach. As was seen, the analysis was free from gene frequency concept. Before having some idea about the genetics of any trait, it would be quite meaningless to talk about gene frequency. Also, the analysis was carried out essentially on the entire family data, which was collected from as many as five states of India covering a large geographical area and also which represented a mixture of many endogamous groups. The present section deals with some further analysis of the family data for extracting

whatever evidence they show in favour of the recessive hypothesis. Since this analysis involves gene frequencies, each endogamous group will be studied here separately. It is quite relevant now to talk about the tongue pigmentation gene frequency since the previous section afforded a strong evidence in favour of the recessive hypothesis.

In all only two endogamous groups are studied here, for, the other groups are represented in the pooled sample only by small numbers of families. The two groups are Brahmin (52 families) and Kayastha (80 families), both from West Bengal. Most of these families were studied at Calcutta, though a few came from nearby places. These 132 families form a part of the total 406 families reported in the previous Chapter.

The two endogamous groups are studied here for the following aspects, whose details will be found respectively in 3.2.1, 3.2.2 and 3.2.3.

(1) Testing the Hardy-Weinberg law (random mating)

(2) Sib-pair analysis to test the recessive hypothesis and the chance hypothesis (as stated in the second chapter), and

(3) Fisher's method of analysis.

3.2.1 Testing the Hardy-Weinberg Law :

The different mating types, their expected proportions under random mating and the corresponding observed frequencies in both the endogamous groups are shown in Table 3.3. The maximum likelihood estimate of q, the recessive or pigmentation gene frequency, and its variance are given by

$$\hat{q} = \left[(2n_4 + n_3 + n_2)/2n\right]^{1/2}$$

and $$\hat{V}(\hat{q}) = (1 - \hat{q}^2)/8n$$

where, n_4, n_3 and n_2 stand respectively for the observed frequencies of families of the types $\bar{a} \times \bar{a}$, $\bar{a} \times \bar{A}$ and $\bar{A} \times \bar{a}$, and n denotes the number of families studied in a group. These estimates, their standard errors and values of the goodness of fit statistic (χ^2 with 2 d.f.) are also included in Table 3.3. The low chi-square values for both the groups (5.87 and 5.49 for Brahmin and Kayastha) exhibit a good fit of the random mating model. Therefore, we shall assume that these two populations are in equilibrium under random mating. The gene frequency estimates given in Table 3.3 will be be used in further analysis.

3.2.2 Sib-pair analysis :

All possible sib-pairs were extracted from the family data on the two populations. For example, a sibship of size c gives rise to $\binom{c}{2}$ sib-pairs. This way, altogether 200 pairs were obtained from the 52 Brahmin families and 260 from the 80 Kayastha families. These sib-pairs are classified into four types :

(i) $\overline{AA}$, both the sibs are normal

(ii) $\overline{Aa}$, elder sib is normal and the younger is pigmented

(iii) $\overline{aA}$, elder sib is pigmented and the younger is normal, and

(iv) $\overline{aa}$, both sibs are pigmented.

The expected frequencies of these four types are derived under the recessive hypothesis as follows. Table 3.4 presents the different genotypic mating types, their frequencies under random mating and the corresponding segregation probabilities. The probability of observing a type (i) sib-pair is obtained from Table 3.4 as the sum of products of column (2) and squares of column (3); i.e.,

$$P(\overline{AA}) = p^4 + 4p^3q + 4p^2q^2(3/4)^2 + 2p^2q^2 + 4pq^3(1/2)^2$$

$$= (4 + 4q - 3q^2 - q^3)p/4$$

where, $p = 1 - q$. Similarly, the probability of observing a

type (ii) sib-pair is obtained from Table 3.4 as the sum of products of columns (2), (3) and (4), which is given by

$$P(\bar{A}\,\bar{a}) = pq^2(3+q)/4 = P(\bar{a}\,\bar{A})$$

In the same way one gets

$$P(\bar{a}\,\bar{a}) = q^2(1+q)^2/4 = 1-P(\bar{A}\,\bar{A}) - 2P(\bar{A}\,\bar{a}).$$

Multiplying these probabilities by the total number of sib-pairs one gets the corresponding expected frequencies. The relevant analysis is set out in Table 3.5, where, hypotheses I and II correspond to the recessive and the chance hypotheses respectively. The probabilities of the four sib-pair types under the chance hypothesis are derived as follows. Let t denote the proportion of pigmented individuals in a random sample study. This t is obtained here as $\hat{q}^2$. Now, if the chance hypothesis were to be true, the probability that an individual is pigmented is estimated by t. Therefore, the probability of $\bar{A}\,\bar{A}$ sib-pair is simply $(1-t)^2$. The four probabilities thus obtained are

$$P(\bar{A}\,\bar{A}) = (1-t)^2$$
$$P(\bar{A}\,\bar{a}) = (1-t)t = P(\bar{a}\,\bar{A}), \text{ and}$$
$$P(\bar{a}\,\bar{a}) = t^2$$

The analysis of sib-pair data under the chance hypothesis

is also included in Table 3.5. This analysis also provides very good evidence against the chance hypothesis, and in favour of the genetic hypothesis.

3.2.3 Fisher's method of analysis :

The advantage of this method of analysis is that we can analyse data on all the families of the mating types $\bar{A} \times \bar{A}$ and $\bar{A} \times \bar{a}$. We notice that earlier, in the first section, we could analyse data on only such families each of which had atleast one pigmented child. This condition is completely relaxed in this type of analysis. The analysis of section one was concerned with numbers of pigmented children, whereas this analysis, known as Fisher's method of analysis, is based on numbers of sibships of a given composition.

Analysis of data on $\bar{A} \times \bar{A}$ families : Table 3.6 presents the necessary material for this analysis. Consider a family of $\bar{A} \times \bar{A}$ mating type with c children, and not necessarily having any pigmented children. The probability that all the c sibs are normal ($\bar{A}$) is given by the ratio of the sum of products of columns (3) and (6) of Table 3.6 relevant to $\bar{A} \times \bar{A}$, and the frequency of the $\bar{A} \times \bar{A}$ mating type.

$$\text{i.e., } \left[p^4 + 4p^3q + 4p^2q^2(3/4)^c\right] / (1 - q^2)^2$$

$$= 1 - \frac{4q^2}{(1 + q)^2}\left[1 - (3/4)^c\right] = P_c, \text{ say}$$

And therefore, the probability that a sibship of size c resulting from a $\bar{A}x\bar{A}$ mating has atleast one pigmented child is given by

$$- \frac{4q^2}{(1 + q)^2}\left[1 - (3/4)^c\right] = 1 - P_c$$

Therefore, each family of the mating type $\bar{A} \times \bar{A}$ can be classified, on expectation basis, into the first type (all c sibs normal) with probability P_c and the second type (atleast one pigmented sib) with probability $1 - P_c$. Variance of the number of families of size c having atleast one pigmented child is simply $n_c P_c(1 - P_c)$, where n_c is the number of families of the mating type $\bar{A} \times \bar{A}$, each having c children. $n_c P_c$ and $n_c(1 - P_c)$ give the expected numbers of families of the two types. The necessary computations for Brahmin and Kayastha are presented in Tables 3.7 and 3.8.

For $\bar{A} \times \bar{A}$ Brahmin families we get from Table 3.7,

$$\chi^2 \text{ (1 d.f.)} = \frac{(38.738 - 37)^2}{38.738} + \frac{(7 - 5.262)^2}{5.262}$$

$$= 0.65 \quad (P > 0.403).$$

which shows excellent fit of the recessive model. Also, following

Smith [35] we get

$$\chi^2 (1 \text{ d.f.}) = \frac{(7 - 5.262)^2}{4.563} = 0.66 \quad (P > 0.403)$$

which again affords the same inference. Fisher's method of analysis for Kayastha sample is summarised in Table 3.10. Following Smith [35], from Table 3.8 we get

$$\chi^2 (1 \text{ d.f.}) = \frac{(12 - 10.422)^2}{8.168} = 0.30 \quad (P > 0.580)$$

which supports the recessive hypothesis.

Analysis of data on $\bar{A} \times \bar{a}$ families. The probability that a family of $\bar{A} \times \bar{a}$ mating type with c children has no pigmented child is again obtained from Table 3.6, which is given by

$$Q_c = \frac{1 - q\left[1 - (1/2)^{c-1}\right]}{1 + q}$$

and the rest of the procedure is just the same as the one presented above for $\bar{A} \times \bar{A}$ families, with Q_c replacing P_c. This analysis is carried out only for Kayastha families, since Brahmin families of this mating type are very few. Table 3.9 presents the relevant computations, whose results under this method are summarised in Table 3.10. Table 3.9 gives, following Smith,

$$\chi^2 \ (1 \text{ d.f.}) = \frac{(7 - 10.186)^2}{5.266} = 1.93 \ (P > 0.157)$$

It is very clear from all that was done in 3.2.3 that the diallelic autosomal recessive hypothesis for the inheritance of tongue pigmentation stands supported.

3.3 Estimation of segregation ratios :

In this section we shall estimate the two well-known Mendelian ratios which are expected to be 1/2 and 1/4, for the two mating types Aa x aa and Aa x Aa respectively. It may be noted that in sections 1 and 2 we have tested whether these ratios are realised in the family data. Instead, the two ratios, defined as the probabilities of getting a pigmented child from the two matings, will be estimated here from the entire family data. The necessary statistical methodology for this can be found in Patil [17], which will be briefly outlined here.

Let us first of all consider families of the mating type $\bar{A}$ x $\bar{a}$. If a family of this mating type produced atleast one $\bar{a}$ child, the genotypic mating type is necessarily Aa x aa, under the recessive hypothesis. And, from all such families (Aa x aa), half of the children are expected to be of the genotype aa. In

the previous two sections we had assumed this segregation ratio (1/2) and tested for the goodness of fit. Instead, we shall now attempt to estimate this ratio and see how close it turns out to 1/2. Let us consider all families (out of the 406 families presented in Table 2.4) of the mating type $\bar{A} \times \bar{a}$, each one of them having atleast one $\bar{a}$ child. Thus, all these families are genotypically Aa x aa. Let π denote the probability for a child from the above mating type to be $\bar{a}$. Clearly, π is expected to be 1/2 under the recessive hypothesis. Now, consider a family of the above type having n_j children (note that, for the sake of ready reference, we shall follow the notations of Patil [17] in this section). The probability that x out of these n_j children are $\bar{a}$ is given by the truncated binomial model

$$\binom{n_j}{x} \pi_j^x (1-\pi_j)^{n_j-x} \Big/ \left[1 - (1-\pi_j)^{n_j}\right]$$

$$x = 1, 2, \ldots, n_j$$

$$n_j = 2, 3, \ldots, K$$

where, K = maximum size (number of children) of any family in our study

Let, N_j = number of families of size n_j

and, T_j = number of $\bar{a}$ children in the N_j families

Then, the likelihood equation for the estimation of π_j is given by

$$\bar{X}_j = \frac{n_j \pi_j}{1 - (1 - \pi_j)^{n_j}}$$

where, $\bar{X}_j = T_j/N_j$.

An iterative solution of this equation yields the maximum likelihood estimate $\hat{\pi}_j$, for each family size ($\hat{\pi}_j$ for families of size n_j). The asymptotic variance of $\hat{\pi}_j$ is given by

$$\text{Var.}(\hat{\pi}_j) = \frac{\pi_j(1 - \pi_j)}{N_j} \Big/ \frac{d\mu^*_j}{d\pi_j}$$

$$= \frac{\pi_j(1 - \pi_j)}{N_j} \cdot \frac{1 - (1 - \pi_j)^{n_j}}{n_j} \Big/ \left[1 + \frac{n_j \pi_j (1 - \pi_j)^{n_j - 1}}{1 - (1 - \pi_j)^{n_j}} \right]$$

where,

$$\mu^*_j = n_j \pi_j / \left[1 - (1 - \pi_j)^{n_j} \right]$$

For different values of π_j spaced at intervals of 0.01, the corresponding values of (μ^*_j / n_j) where tabulated by Patil [17]. This table can be profitably used for the estimation of π_j. In the present study, the estimates are obtained from Patil's table,

using linear interpolation wherever necessary. Let us denote the estimates thus obtained by $\hat{\pi}_{j0}$ (initial estimates). The improved estimates, say $\hat{\pi}_{j1}$, can then be obtained from

$$\hat{\pi}_{j1} = \hat{\pi}_{j0} + \left[(\bar{X}_j - \pi^*_{j0}) \Big/ \left(\frac{d\mu^*_j}{d\pi_j} \right)_{\pi_j = \hat{\pi}_{j0}} \right]$$

where, addition of 'zero' to the subscript implies that $\hat{\pi}_{j0}$ should be substituted for π_j in the concerned expressions. A repeated application of this process will yield very accurate estimates.

For the present study, however, we follow the following procedure. Since there are families of different sizes, we shall first of all obtain the initial estimates of π_j $(\hat{\pi}_{j0})$ as described above. Then we shall test for the homogeneity of the different distributions in respect of the parameter π. Or in other words, we shall test whether the tongue pigmentation gene segregates alike in families of different sizes (i.e. whether all the π_j's are equal). If this test suggests homogeneity, we shall obtain the combined estimate of π very accurately (pooled over all families of all sizes). Otherwise, we shall go back to the initial estimates $\hat{\pi}_{j0}$ and improve them repeatedly till we

arrive at the desired accuracy. It may be noted that testing for homogeneity and combined estimation of π are carried out simultaneously.

Under the hypothesis of homogeneity, let us denote the common value by π. A starting or initial estimate $\hat{\pi}_0$ of π may be chosen as that value around which the $\hat{\pi}_{j0}$'s cluster. The next approximation to the maximum likelihood estimates is then given by

$$\hat{\pi}_1 = \hat{\pi}_0 + \frac{\bar{X} - \bar{\mu}}{\bar{\partial}}$$

where,

$$\bar{X} = \sum_{j=1}^{k} N_j \bar{X}_j / N$$

$$\bar{\mu} = \sum_{j=1}^{k} N_j \mu^*_{j0} / N$$

$$\bar{\partial} = \sum_{j=1}^{k} N_j \partial_{j0} / N$$

$$\partial_{j0} = \frac{d\mu^*_j}{d\pi_j} \Bigg|_{\pi_j = \hat{\pi}_{j0}}$$

and, $$N = \sum_{j=1}^{k} N_j$$

Repeated application of these cyclic operations will ultimately yield the maximum likelihood estimate, to be denoted by π. A test for the homogeneity hypothesis is then given by the statistic

$$\chi^2 = \frac{1}{\hat{\pi}(1-\hat{\pi})} \sum_{j=1}^{k} \frac{N_j [\bar{X}_j - \mu_j(\hat{\pi})]^2}{\left(\frac{d\mu_j(\pi)}{d\pi}\right)_{\pi=\hat{\pi}}}$$

where,

$$\mu_j(\pi) = \frac{\pi n_j}{1 - (1-\pi)^{n_j}}$$

and $$\frac{d\mu_j(\pi)}{d\pi} = \frac{\mu_j}{\pi(1-\pi)} \left[1 + (n_j - 1)\pi - \mu_j\right]$$

The above test statistic, χ^2, follows a χ^2 distribution with $(K - 1)$ degrees of freedom.

Let us consider the data given in Table 3.1 on $\bar{A} \times \bar{a}$ families, each having atleast one $\bar{a}$ child. The detailed analysis of the data is neatly incorporated in Table 3.11. Compuations in columns (1) to (6) are already explained. Observe that the $\hat{\pi}_{j0}$'s are clustered around 0.4, which we take as $\hat{\pi}_0$. Corresponding to this value of $\hat{\pi}_0$, making use of Patil's table we obtain $\mu_j^*(\hat{\pi}_0)$

for each value of n_j, which are presented in column (7). Also, column (8) is computed directly from the formula given already, with $\hat{\pi}_0 = 0.4$. Next, we obtain $\bar{X}$, $\bar{\mu}$, and $\bar{\delta}$ in the last row of Table 3.11, from where one gets the improved estimate

$$\hat{\pi}_1 = 0.4 + \frac{1.545 - 1.556}{1.813} = 0.394$$

and an approximate variance estimate

$$\widehat{\text{Var.}}\,(\hat{\pi}_1) = \frac{\hat{\pi}_1(1 - \hat{\pi}_1)}{N\bar{\delta}} = 0.002993$$

$$\text{S.e.}\,(\hat{\pi}_1) = 0.055.$$

Using $\hat{\pi}_1$, we now compute $\mu_j(\hat{\pi}_1)$ directly from the formula already given, the values of which are given in column (9). Similarly, column (10) is also computed directly from the formula. It is next seen that one more cycle of operations brings no more improvement in $\hat{\pi}_1$ upto the third decimal place. Therefore, we take this value as $\hat{\pi}$ and perform the necessary computations in column (11), from where we get

$$\chi^2_3 = \frac{\text{Total of column (11)}}{\hat{\pi}\,(1 - \hat{\pi})}$$

$$= 3.70$$

which exhibits a very good fit ($P > 0.28$) of the homogeneity hypothesis to our data. Thus we have, for the entire data,

$$\hat{\pi} = 0.394$$

$$\text{and } \mathcal{E}\, e(\hat{\pi}) = 0.055$$

It may be noted that this estimate does not differ significantly from 0.5, at the 5% level. Thus, this finding is quite consistent with the suggested recessive inheritance hypothesis for tongue pigmentation.

Let us now consider all $\bar{A} \times \bar{A}$ families given in Table 3.2, each having atleast one $\bar{a}$ child. Therefore, the value of π is expected to be 0.25 for these families. The procedure being the same as above, the detailed analysis is set out in Table 3.12. We have taken here $\hat{\pi}_0 = 0.25$ and got $\hat{\pi}_1 = 0.265$, with a standard error 0.040. While performing the next cycle of operations, we got $\bar{X} = \bar{\mu}(\hat{\pi}_1)$, and therefore $\hat{\pi}_1$ is taken as $\hat{\pi}$, the maximum likelihood estimate. With this $\hat{\pi}$, we have $\chi^2_5 = 8.28$ $(P > 0.13)$ which supports the homogeneity hypothesis. Hence, $\hat{\pi}$ represents the overall segregation ratio (estimate) for all the 59 $\bar{A} \times \bar{A}$ families. It is clear that $\hat{\pi}$ does not differ significantly from the expected value (0.25) under the recessive

hypothesis (at the 5% level).

The analysis carried out in sections 3.1 and 3.3 provided sufficient evidence to suggest the recessive hypothesis for the inheritance of the trait. Through gene frequency estimates, the family data analysis of section 3.2 also supported the same conclusion. Therefore, we accept this hypothesis as the mode of inheritance and investigate the penetrance aspect in the next Chapter.

TABLE 3.1

Number of $\bar{a}$ children in $\bar{A} \times \bar{a}$ (including $\bar{a} \times \bar{A}$) families each containing at least one $\bar{a}$ child

Sibship size c	No. of sibships m_c	Observed $\bar{a}$ children	Expected $\bar{a}$ children	Variance	χ^2	d.f.	P
1	5	5	5	0	-	-	-
2	15	16	19.995	3.330	4.79	1	< 0.03
3	17	26	29.138	8.330	1.18	1	> 0.27
4	7	13	14.931	5.474	0.68	1	> 0.40
5	5	13	12.905	5.410	0.00	1	1
Totals	49	73	81.969	22.544	6.65	4	> 0.14

For totals : $\chi^2 = \frac{(73 - 81.969)^2}{22.544} = 3.60$ (1 d.f., $P > 0.05$)

TABLE 3.2

Number of $\bar{a}$ children in $\bar{A} \times \bar{A}$ families each containint at least one $\bar{a}$ child

Sibship size c	No. of sibships M_c	Observed $\bar{a}$ children	Expected $\bar{a}$ children	Variance	χ^2	d.f.	P
1	4	4	4	0	-	-	-
2	12	12	13.716	1.464	2.01	1	> 0.15
3	16	20	20.752	4.208	0.13	1	> 0.06
4	17	26	24.871	7.140	0.18	1	> 0.65
5	6	12	9.834	3.552	1.32	1	> 0.24
6	6	9	10.950	4.656	0.16	1	> 0.65
7	2	7	4.040	1.940			
Totals	63	90	88.163	22.960	3.80	5	> 0.57

For totals : $\chi^2 = \frac{(90 - 88.163)^2}{22.960} = 0.15$ (1 d.f., P > 0.65)

TABLE 3.3

Test for Hardy-Weinberg Law

Mating type Male x Female	Mating frequency	Observed frequency	
		Brahmin	Kayastha
$\bar{A} \times \bar{A}$	$(1 - q^2)^2$	44	53
$\bar{A} \times \bar{a}$	$q^2(1 - q^2)$	3	16
$\bar{a} \times \bar{A}$	$q^2(1 - q^2)$	3	6
$\bar{a} \times \bar{a}$	q^4	2	5
Totals (n)	1	52	80
$\hat{q}$		0.310087	0.447214
S.e.		0.046612	0.035355
χ^2(2 d.f.)		5.87	5.49

TABLE 3.4

Genotypic mating types, their frequencies under random mating and the corresponding segregation probabilities

Genotypic mating		Probability of a child being	
Type	Frequency	$\bar{A}$	$\bar{a}$
(1)	(2)	(3)	(4)
AA x AA	p^4	1	0
— x Aa	$4p^3q$	1	0
Aa x Aa	$4p^2q^2$	3/4	1/4
AA x aa	$2p^2q^2$	1	0
Aa x —	$4pq^3$	1/2	1/2
aa x —	q^4	0	1

TABLE 3.5

Sib-pair analysis for testing genetic and non-genetic chance hypotheses

Sib-pair type	Observed frequency		Expected frequency under			
			Hypothesis - I		Hypothesis - II	
	Brahmin	Kayastha	Brahmin	Kayastha	Brahmin	Kayastha
$\bar{A}\,\bar{A}$	161	172	169.790	183.227	163.388	166.400
$\bar{A}\,\bar{a}$	12	22	10.980	24.772	17.382	41.600
$\bar{a}\,\bar{A}$	17	36	10.980	24.772	17.382	41.600
$\bar{a}\,\bar{a}$	10	30	8.250	27.229	1.848	10.400
χ^2 (with 3 d.f.)			4.22	6.37	37.64	47.12
P			$> .222$	$> .093$	$< .00001$	$< .00001$

TABLE 3.6

Phenotypic and genotypic mating types, their frequencies, corresponding segregation probabilities and probabilities thereof for observing all-normal sibship of size c

Mating type		Mating frequency	Probability of a sib being		Probability of all the c sibs being
Phenotypic	Genotypic		$\bar{A}$	$\bar{a}$	$\bar{A}$
(1)	(2)	(3)	(4)	(5)	(6)
$\bar{A} \times \bar{A}$	AA x AA	p^4	1	0	1
	— x Aa	$4p^3q$	1	0	1
	Aa x Aa	$4p^2q^2$	3/4	1/4	$(3/4)^c$
Total		$(1 - q^2)^2$	-	-	-
$\bar{A} \times \bar{a}$	AA x aa	$2p^2q^2$	1	0	1
	Aa x —	$4pq^3$	1/2	1/2	$(1/2)^c$
Total		$2q^2(1-q^2)$	-	-	-
$\bar{a} \times \bar{a}$	aa x aa	q^4	0	1	0

TABLE 3.7

Fisher's method of analysis for $\underline{A} \times \overline{A}$ Brahmin families

Sibship size	Number of sibships	Number of sibships with all $\underline{A}$ sibs		Number of sibships, each with atleast one $\underline{a}$ sib		
		Observed	Expected	Observed	Expected	Variance
1	8	8	7.552	0	0.448	0.423
2	12	12	10.824	0	1.176	1.061
3	7	5	6.097	2	0.903	0.786
4	11	7	9.317	4	1.683	1.426
5	4	4	3.316	0	0.684	0.567
6	2	1	1.632	1	0.368	0.300
Totals	44	37	38.738	7	5.262	4.563

TABLE 3.8

Fisher's method of analysis for $\overline{A} \times \overline{A}$ Kayastha families

1	8	6	7.232	2	0.768	0.694
2	19	15	15.827	4	3.173	2.643
3	14	12	10.906	2	3.094	2.410
4	7	6	5.173	1	1.827	1.350
5	2	1	1.416	1	0.584	0.413
6	1	0	0.686	1	0.314	0.215
7	2	1	1.338	1	0.662	0.443
Totals	53	41	42.578	12	10.422	8.168

TABLE 3.9

Fisher's method of analysis for $\bar{A} \times \bar{a}$ Kayastha families

Sibship size	Number of sibships	Number of sibships with all $\bar{A}$ sibs		Number of sibships, each with atleast one $\bar{a}$ sib		
		Observed	Expected	Observed	Expected	Variance
1	5	5	3.455	0	1.545	1.068
2	10	7	5.360	3	4.640	2.487
3	2	1	0.918	1	1.082	0.497
4	4	2	1.680	2	2.320	0.974
5	1	0	0.401	1	0.599	0.240
Totals	22	15	11.814	7	10.186	5.266

TABLE 3.10

Summary of Fisher's method of analysis of 75 Kayastha families

Mating type	Sibship composition	Number of sibships		χ^2	d.f.	P
		Observed	Expected			
$\bar{A} \times \bar{A}$	All sibs $\bar{A}$	41	42.578	0.30	1	> .580
	Atleast one $\bar{a}$ sib	12	10.422			
$\bar{A} \times \bar{a}$ (including $\bar{a} \times \bar{A}$)	All sibs $\bar{A}$	15	11.814	1.86	1	> .168
	Atleast one $\bar{a}$ sib	7	10.186			
	Totals	-	-	2.16	2	> .332

TABLE 3.11

Computations for the estimation of π from $\bar{A} \times \bar{a}$ family data on tongue pigmentation

n_j (1)	N_j (2)	T_j (3)	$\bar{X}_j$ (4)	$\bar{X}_j/n_j$ (5)	$\hat{\pi}_j$ (6)	$\mu_j^*(\hat{\pi}_o)$ (7)	∂_j (8)	$\mu_j(\hat{\pi}_1)$ (9)	$(\frac{d\mu_j}{d\pi})_{\hat{\pi}_1}$ (10)	χ^2 component (11)
2	15	16	1.067	0.534	0.125	1.250	0.781	1.2457	0.777	0.612
3	17	26	1.529	0.510	0.400	1.531	1.716	1.520	1.706	0.001
4	7	13	1.857	0.464	0.406	1.838	2.772	1.822	2.747	0.003
5	5	13	2.600	0.520	0.504	2.169	3.895	2.145	3.872	0.267
Totals	44	68	-	-	-	-	-	-	-	0.883
Weighted average			1.545	-	-	1.556	1.813	-	-	-

Column (8) : $\partial_j = (\frac{d\mu_j^*}{d\pi})_{\hat{\pi}_o}$

Column (11): χ^2 component $= \frac{N_j[\bar{X}_j - \mu_j(\hat{\pi})]^2}{(d\mu_j/d\pi)_{\hat{\pi}_1}}$

TABLE 3.12

Computations similar to Table 3.11, corresponding to $\bar{A} \times \bar{A}$ Families

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
2	12	12	1.000	0.500	0.000	1.143	0.652	1.152	0.668	0.415
3	16	20	1.250	0.417	0.216	1.297	1.404	1.318	1.434	0.052
4	17	26	1.529	0.382	0.278	1.463	2.239	1.497	2.290	0.008
5	6	12	2.000	0.400	0.356	1.639	3.156	1.686	3.237	0.183
6	6	9	1.500	0.250	0.166	1.825	4.137	1.888	4.236	0.213
7	2	7	3.500	0.500	0.496	2.020	5.171	2.098	5.299	0.742
Totals	59	86	-	-	-	-	-	-	-	1.613
Weighted average			1.458	-	-	1.426	2.075	1.458	-	-

CHAPTER IV

TONGUE PIGMENTATION IN MAN

PENETRANCE AND OTHER STUDIES

It was clear from the previous chapter that this trait has a simple genetic basis. In all probability, the mode of inheritance is autosomal recessivity. The only cases threatening this rigid mode of inheritance (which implicitely assumes full penetrance of the alleles involved) are provided by the few normal children born to pigmented parents. Should these cases represent genuine exceptions, not arising out of extra-marital influence, consideration of incomplete penetrance of the pigmentation allele seems to justify such cases. For example, under the recessive inheritance model, a child born to recessive parents is expected to be of the genotype aa and not necessarily pigmented. He could very well be normal (unpigmented). Failure of penetrance of the pigmentation allele in the child could account for this. After all, it is the genotype that is inherited and not the phenotype.

It should be recalled here that we failed to detect any variation in the rate of incidence due to age factor. But, our study was restricted only to adults. Davis [6], on the other hand, observed some significant variations due to age among

kindergarten students. The same is true for the Malag tribe of an Australian aboriginal community (Kirk [13]). It appears, therefore, that there is significant difference between the rates of incidence,, one corresponding to very young ages and another corresponding to adults or even older, for the same population. To study this aspect more in details it became necessary to look for this pigmentation in the infants. Accordingly, a study of 845 infants is presented in the next section.

4.1 A study of the new born babies :

The purpose of this study is to analyse the rates of incidence of the trait among the new born, and to compare the same with those for the adults as reported in the second Chapter.

A total of 845 new born babies were studied for tongue pigmentation. This study was carried out during March to August 1970 at the maternity wards of Nilratan Sirkar Medical College, Calcutta Medical College and R. G. Kar Medical College hospitals of Calcutta. Nearly one-fourth of the babies were examined in the day light and the rest under torch light. It may be stressed at this point that tongue examination of the new born is extremely difficult. At the first place, observations on 71 babies were

totally discarded since it was not possible to force open the mouth for thorough examination of the tongue. Secondly, even barring these 71 cases, observations on 13 babies out of 845 were 'doubtful' in the sense that it could not be confirmed as to whether they had tongue pigmentation. Two plausible explanations for this could be the difficulty with examination under torch light and the dilute intensities at which pigments are generally formed in the infants.

$\bar{A}$, $\bar{a}$ and d are used to mean 'normal', 'affected' (pigmented) and 'doubtful' respectively. Depending on the context they also stand for the corresponding frequencies. Table 4.1 presents data on 845 babies classified according to age (in days), sex, caste and the trait.

For human genetic problems it is desirable that a trait must be identified with great accuracy leaving almost no doubt. It may be observed that, in the present study, 1.54% of the observations on the trait are doubtful (13 out of 845). Though this figure sounds high, it is, perhaps, justifiable since the study refers to infants only. Concerning the doubtful cases, it is next attempted to see whether at all sex plays a roll in the accuracy of identification. Table 4.2 presents the relevant analysis, which rules out any such

association. This analysis is done separately for two age-groups, since it is felt that the age-group could very well influence the nature of such associations. It should be remembered, however, that some cell frequencies in Table 4.2 are very small. The χ^2 values were computed after applying Yate's correction.

Table 4.3 incorporates results of sex-trait association study. This is also done separately for two age-groups for the same reason as given above. From the low χ^2 values one infers that the two attributes, sex and tongue pigmentation, are independent. Thus, it appears that sex does not play any role in the incidence of the trait, even at birth.

Finally, Table 4.4 presents the rates of incidence of the trait among the new born. They are computed as :

$$\text{Minimum} : \frac{\bar{a}}{\bar{A} + \bar{a} + d} \times 100$$

$$\text{Maximum} : \frac{\bar{a} + d}{\bar{A} + \bar{a} + d} \times 100$$

It should be noted that the true rates of incidence lie between the corresponding minimum and maximum rates. The overall rate for the infants of West Bengal is estimated at 3.90% (maximum),

whereas the corresponding figure for the adults of West Bengal is 15.93% (pooled over the entire data). This figure for the adults was obtained from the data on parents alone, of the families that were studied in West Bengal for the trait, together with data on a few more individuals who were examined just recently. This consisted of 104 Brahmins (10 pigmented) and 166 Kayastha (33 pigmented). The rates for infants and adults are seen to differ significantly.

Another observation is right in order at this stage. Notice from Table 4.1 that the overall maximum rate for Kayastha is significantly higher than that for Brahmins. Interestingly, the same is true for the adults of Bengal (vide Chapter II).

In view of the very low incidence rate of the trait among the new born, the following two explanations are offered, either of which might be accepted at this stage, before further investigation is carried out on this aspect :

(i) Some of the supposedly normal babies are actually pigmented, but the intensity of the colour is so dilute that it becomes too difficult to identify such cases at birth.

(ii) Penetrance of the pigmentation allele increases over age, atleast during early infancy.

Perhaps the truth lies in a combination of the two. It is felt that, in any case, the penetrance increases very rapidly from the time of birth to very early infancy. The fact that penetrance at birth is very low is clear from the following argument. Assuming that the first explanation given above is incorrect, one obtains an upper bound for the penetrance at birth as (3.90/15.93) x 100% = 24.48%. However, should the first explanation be also true, atleast partly, one gets higher penetrance estimates than given above. In any case, incomplete penetrance seems to be an essential phenomenon.

The overall results obtained in this section are alarming enough to create motivation for a thorough search on the penetrance of the pigmentation allele. The first point that bothers us the most is the exceptionally low rate of incidence among the new born. Under the first explanation given to account for this, though some of the phenotypically normal babies might really be tongue pigmented, it sounds rather fantastic that mere dilute intensity of the pigmentatio in the babies should give rise to such a low rate. Of course, it is nothing very unusual since most of the babies were examined under torch light, an artificial light that could very well prevent the detection of tongue pigmentation at dilute intensities of the colour.

On the other hand, the assumption of incomplete penetrance alone does not appear to be sound, for similar reasons. We shall, however, assume that in some cases the pigmentation allele fails to penetrate, thereby classifying some of the recessive genotypes (aa) as dominant phenotypes ($\bar{A}$), and investigate further in the next section.

4.2 Penetrance of the tongue pigmentation allele :

Chapter II reported family data on tongue pigmentation in man, which contained some 'controversial' observations relating to the families with pigmented parents ($\bar{a} \times \bar{a}$) giving some normal ($\bar{A}$) children, thereby opposing the proposed genetic hypothesis for the inheritance of the trait under study. It was then observed that these cases might have been due to extra-marital influence or some variation in the genic action. However, the first possibility could not be verified due to lack of co-operation from the particular families concerned. An attempt is made here to explore the latter possibility, assuming this to be the only possibility. Thus, it should be noted at this stage that the family data should be made free of illegitimate cases before employing the method presented here.

In this section, by introducing the concept of penetrance

probability of a gene, denoted by β, the joint distribution of all the sample families is derived under the assumption of random mating with respect to the locus under study, and the parameters involved are estimated by the method of maximum likelihood.

It may be noted here that Elandt-Johnson [10] has recently considered a similar methodological problem. However, she estimates the parameters for each parental phenotypic mating separately, whereas the present method deals with the estimation of parameters from the joint distribution of parents and offspring (of all mating types taken at a time). The present method will be better especially if the penetrance of the gene is known to be high.

Model

Define

β = Prob. (phenotype is $\bar{a}$ given that the genotype is aa)

q = frequency of the pigmentation allele (a)

p = frequency of the normal allele (A)

Table 4.5 presents the different genotypic mating types, their frequencies under random mating, the probabilities with which the genotypic matings are classified under the possible phenotypic

mating types, and the segregation probabilities corresponding to each genotypic mating type. The joint distribution of parents and offspring is presented in Table 4.6, which is derived from Table 4.5 as follows :

Consider $\bar{A} \times \bar{A}$ mating frequency (λ_1). It is obtained as the sum of products of columns (1) and (2) of Table 4.5, i.e.,

$$\lambda_1 = p^4 + 4p^3q + 4p^2q^2 + 2p^2q^2(1 - \beta) + 4pq^3(1 - \beta) + q^4(1 - \beta)^2$$

$$= (1 - \beta q^2)^2.$$

Similarly, the sum of products of columns (1) and (3) will yield λ_2 and so on.

Now consider the probability of $\bar{a}$ child being a from $\bar{A} \times \bar{A}$ mating type (R_{12}/λ_1). R_{12} is obtained as the sum of products of columns (1), (2) and (6) of Table 4.5, i.e.

$$R_{12} = p^2q^2\beta + 2pq^3\beta(1 - \beta) + q^4(1 - \beta)^2\beta$$

$$= \beta q^2(1 - \beta q)^2$$

Similarly the other entries of Table 4.6 are obtained.

Estimation of the parameters

Consider a random sample of N families with a total of n

offspring. The distribution of these N families according to the parental mating type and the offspring phenotype is shown in Table 4.6 within parentheses.

The joint likelihood is given by (Adhikari, Sarma and Chakraborty, 1970, unpublished, and cited in Chakraborty [2]):

$$L = K \left(\prod_{i=1}^{3} \lambda_i^{N_i} \right) \left\{ \prod_{i=1}^{3} \prod_{j=1}^{2} (R_{ij}/\lambda_i)^{n_{ij}} \right\} f(\theta)$$

where, K is a constant, and $f(\theta)$ refers to the family size distribution, which is not necessary for the estimation problem.

The two parameters q and β are now estimated by the maximum likelihood method. To start with one obtains some initial estimates of q and β, and then improve these estimates by the well known iterative procedure. The expressions for the efficient scores

$$L_q = (\partial \log L / \partial q)$$

$$L_\beta = (\partial \log L / \partial \beta)$$

and for the variance-covariance matrix

V_{qq} = variance of q

$V_{q\beta}$ = covariance of q and β

$V_{\beta\beta}$ = variance of β

are given below. It may be noted that we get the information matrices for the distribution of parents, and the conditional distributions of offspring corresponding to each of the three phenotypic mating types separately and then pool these four matrices giving the following overall information matrix.

$$L_q = \sum_{i=1}^{3} \left(\frac{N_i - n_i}{\lambda_i}\right)\left(\frac{\partial \lambda_i}{\partial q}\right) + \sum_{i=1}^{3}\sum_{j=1}^{2} \frac{n_{ij}}{R_{ij}} \left(\frac{\partial R_{ij}}{\partial q}\right)$$

$$L_\beta = \sum_{i=1}^{3} \left(\frac{N_i - n_i}{\lambda_i}\right)\left(\frac{\partial \lambda_i}{\partial \beta}\right) + \sum_{i=1}^{3}\sum_{j=1}^{2} \frac{n_{ij}}{R_{ij}} \left(\frac{\partial R_{ij}}{\partial \beta}\right)$$

$$I_{qq} = 8 N \beta \lambda + \frac{n_1 f_1^2}{\lambda_1^2 R_{11} R_{12}} + \frac{n_2 f_2^2}{\lambda_2^2 R_{21} R_{22}}$$

$$I_{q\beta} = 4 N q \lambda + \frac{n_1 f_1 f_3}{\lambda_1^2 R_{11} R_{12}} + \frac{n_2 f_2 f_4}{\lambda_2^2 R_{21} R_{22}}$$

$$I_{\beta\beta} = \frac{2 N q^2 \lambda}{\beta} + \frac{n_1 f_3^2}{\lambda_1^2 R_{11} R_{12}} + \frac{n_2 f_4^2}{\lambda_2^2 R_{21} R_{22}} + \frac{n_3}{\beta (1 - \beta)}$$

where,

$$\lambda = (1 + \lambda_2) / \sqrt{\lambda_1} - 2\sqrt{\lambda_3}$$

$$f_1 = \lambda_1 \left(\frac{\partial R_{12}}{\partial q}\right) - R_{12}\left(\frac{\partial \lambda_1}{\partial q}\right)$$

$$f_2 = \lambda_2 \left(\frac{\partial R_{22}}{\partial q}\right) - R_{22}\left(\frac{\partial \lambda_2}{\partial q}\right)$$

$$f_3 = \lambda_1 \left(\frac{\partial R_{12}}{\partial \beta}\right) - R_{12}\left(\frac{\partial \lambda_1}{\partial \beta}\right)$$

$$f_4 = \lambda_2 \left(\frac{\partial R_{22}}{\partial \beta}\right) - R_{22}\left(\frac{\partial \lambda_2}{\partial \beta}\right)$$

where, ∂ stands for the differentiation sign. The pooled variance - covariance matrix is obtained by inverting the information matrix I.

The computational method is briefly as follows. First choose some trial values or initial estimates of q and β, say q_0 and β_0. For example, one can obtain q_0 from data on the parental mating types alone, under random mating. β_0 can be obtained as the proportion of recessive children ($\bar{a}$) among all matings of the type $\bar{a} \times \bar{a}$. With these initial values of q and β, evaluate the efficient scores and the dispersion matrix, from where one obtains the correction factors for q_0 and β_0 by

$$\delta q_0 = V_{q_0 q_0} L_{q_0} + V_{q_0 \beta_0} L_{\beta_0}$$

and

$$\delta \beta_0 = V_{q_0 \beta_0} L_{q_0} + V_{\beta_0 \beta_0} L_{\beta_0}$$

The improved estimates are thus obtained from

$$q_1 = q_0 + \delta q_0 \quad \text{and} \quad \beta_1 = \beta_0 + \delta \beta_0$$

All this consists of one cycle of operations. Unless the correction factors are as small as we wish, the entire computational procedure should be repeated time and again taking the improved estimates in one cycle as the initial estimates for the next one, till the desired accuracy is attained. These estimates, however, converge to a stable value with probability one (Rao [19]).

Table 4.7 presents the family data on tongue pigmentation in man (see Table 2.4). Starting with 0.4 and 0.95 as the initial estimates of q and β respectively, it took altogether 11 cycles or iterations to arrive at the final estimates with accuracy upto the sixth decimal place. Some parts of these computations are incorporated in Table 4.8. Using the final estimates of q and β the expected frequencies have been computed, which are shown in Table 4.7 within parentheses.

From Table 4.8 one obtains the final estimates as

$$\hat{q} = 0.482634$$

and $$\hat{\beta} = 0.861736$$

with respective standard errors 0.018015 and 0.045445.

Thus, under suitable assumptions, it follows that the tongue pigmentation allele (a) has high penetrance. However, it should be reminded that all this is valid only if the 'controversial' observations are not due to extra-marital influence.

The above analysis is carried out on the entire family data, representing several endogamous populations. It was felt that the penetrance is very high and therefore, in order to estimate penetrance values close to 100%, huge body of data are necessary. For this reason the family data where pooled together. However, separate estimation has been done from Brahmin and Kayastha family data. This was done just to see whether the penetrance values differ very much. From $\bar{a} \times \bar{a}$ Brahmin families we have 6 children, all of whom are $\bar{a}$. On the other hand, $\bar{a} \times \bar{a}$ Kayastha families produced 1 $\bar{A}$ and 9 $\bar{a}$ children in all. The $\bar{A}$ Kayastha child is suggestive of incomplete penetrance, but the $\bar{a} \times \bar{a}$ Brahmin families do not suggest such a phenomenon. Since the penetrance is very high, in the absence of $\bar{A}$ children out of $\bar{a} \times \bar{a}$ matings, we may fail to distinguish a penetrance value close to 100% from 100%. This is because, the $\bar{a} \times \bar{a}$ families are most informative about β (Rao [27]). The analysis for Brahmin and Kayastha families, separately and together, is summarised in Table 4.9. Observe that, as remarked

above, the penetrance estimate from Brahmin families is no different from unity. On the whole, the pigmentation allele is seen to have high penetrance.

4.3 Other studies :

Having satisfied ourselves about the genetics of tongue pigmentation, it is of considerable interest to study some other aspects involving this trait. Some such problems are :

(i) Histopathological studies of the pigmented tongue tissue and comparing these findings with those for normal tongue tissue;

(ii) Possible biochemical studies aimed at exploring the nature of this pigmentation; and

(iii) Relation or association between tongue pigmentation and other traits of general interest. This problem includes the linkage studies also.

4.3.1 Pilot studies on Histopathological and Biochemical aspects :

Concerning the first one of the above problems, a preliminary study was carried out on two dead people who were pigmented. They were examined for the trait at the N.R.S.M.C. Hospital of Calcutta

nearly ten days after their death. In fact, altogether eleven dead bodies were examined with the result that two of them were identified as tongue pigmented. Of the many living pigmented people examined at the Andhra Medical College (Waltair, Andhra Pradesh), only two have agreed to extend their co-operation for this study. Accordingly, the pigmented tongue tissues were collected from the two dead bodies, and only biopsy specimens were obtained from the two living individuals. One slide of each of these four specimens were kindly examined by Professor H. Gruneberg, F.R.S., University College London and by Professor I. Doniach, London Hospital Medical School. They both have agreed in their findings, a summary of which is given below (Gruneberg, personal communication).

"In one of the dead tissues (male, 40 years) there are extensive extravasations of blood under the mucosa which would have resulted in a livid discolouration and could thus have given the impression of pigmentation. In the other dead tissue (male, 25 years) there is a marked dilation of venules under the mucosa, but no extravasations. The dilation of the blood vessels under the mucosa could have produced a livid discolouration as in the first case. In both the biopsy specimens (male, 50 years and female 45 years) a moderate amount of brown pigment is present in some places in

the connective tissue."

Thus it is clear that this study should not be carried out on ten days old thoroughly decomposed dead bodies, which could lead to a very misleading results. It was, therefore, decided not to pay any attention to the results obtained from dead tissues.

The results of the two living specimens, on the other hand, are quite interesting. A project to study many more living cases is being carried out by the Indian Statistical Institute. A part of this program includes the study of these tissues with the help of the iron test (Prussian blue reaction) to see whether or not this pigmentation is haemosiderin, and hence derived from haemoglobin. Should this test give negative results, a Dopa test will be carried out to see whether the brown pigment is melanin. It is hoped that this project will solve the problem one way or the other.

Concerning the possible Biochemical studies, analysis of the serum copper oxidase activity is considered to be one of the means of finding out whether this tongue pigmentation is melanin or not. Since it is associated with melanin production through tyrosinase activity, it is hoped that this study will throw much light on this problem. Study of the serum copper oxidase activity among both

tongue pigmented and normal individuals is of current interest to me.

4.3.2 Relation with other traits :

About the possible association this trait may have with other attributes, it will be interesting to carry out an investigation involving as many well defined genetic markers as possible. Such a study can be suggestive of possible linkages also.

Here, we shall report the findings of a pilot study to investigate the relation between tongue pigmentation and mental ability. In some South Indian communities there is a strong belief that "tongue pigmented" people are more intelligent than others. Assuming that this "tongue pigmentation" is the same as the one we have been talking about, this study was carried out to verify the old belief. Also, another motivation for this study was provided by some assertions in the literature sugresting positive association between the two attributes (Vogel, F. : personal communication).

A sample of 342 students, 242 males and 100 females, aged 11 to 19 years, from M. M. P. School at Srikakulam (Andhra Pradesh, India) was examined in the day light for tongue pigmentation.

Although not all of the applicants are admitted to this school, selection is not based on intelligence tests. The sample may, therefore, be considered as a representative of the normal population of children in Andhra Pradesh. In all, 62 male and 24 female students were affected ($\bar{a}$).

The students were subsequently given special written and oral tests, worked out the approval of Psychometricians, and awarded scores (maximum score = 10). The method followed for these tests was as follows :

The students were divided into seven groups on the basis of their educational levels. In order to assess their mental abilities, ten questions were prepared for each group on the basis of the subjects they were taught, and the questions were not of the pure achievement type, but were expected to measure reasoning ability and speed. The candidates were to answer these questions verbally and each question carried 1/2 marks, and hence the maximum for these ten questions were 5 marks. The answers to these questions were evaluated on the spot.

Next, ten more questions were prepared and the candidates were to answer them at their homes and had to submit their answers

the next day. No background knowledge was necessary to answer these questions, and the candidates had to use their reasoning power to arrive at the correct solutions. As before, the answers were evaluated and the maximum marks for these ten questions were 5. The total score thus secured by a student, out of a maximum of 10 marks, is taken as an index of his mental ability. Table 4.10 presents the frequency distribution of scores for the 86 affected (pigmented) and 256 unaffected (normal) students, classified according to the sex. Thus we have four frequency distributions, one each for the normal male, pigmented male, nomal female and pigmented female students. The mean scores and the variances for all these four distributions are also presented in Table 4.10. The hypothesis to be tested may be stated as "mean score of the pigmented group is equal to that of the normal". We shall test this hypothesis separately for males and females. Though this sample also fails to show any variantions in the incidence rate due to sex, we are not pooling males and females together in this problem because of the fact that the denominator of the test statistic used below is likely to be elevated by such a pooling.

Assuming that the scores follow a normal distribution, the following statistic is computed based on data for male students

(Table 4.10):

$$\frac{5.121 - 4.694}{\sqrt{\frac{2.030}{180} + \frac{3.292}{62}}} = 1.688$$

which follows standard normal distribution. Similarly, for the data on female students, the value of the statistic turns out to be 0.823. Thus we see that, both values being low, the hypothesis under examination stands supported. Hence, the two attributes tongue pigmentation and mental ability (as determined here) are independent, thereby disposing off the assertions contrary to it.

TABLE 4.1

Tongue pigmentation data on new born babies

Age (in days)	Sex	Normal (Ā)			Pigmented (ā)			Doubtful (d)		
		Brahmin	Kayastha	Others	Brahmin	Kayastha	Others	Brahmin	Kayastha	Others
<1	Male	25	54	114	1	5	3	1	5	0
	Female	22	68	93	0	4	1	0	1	2
≥1	Male	33	74	95	1	1	1	0	1	1
	Female	28	63	122	0	0	1	0	1	1
Unknown	Male	0	2	8	0	1	0	0	0	0
	Female	0	0	11	0	0	1	0	0	0

TABLE 4.2

Investigation of possible association between sex and 'doubtful observations'

Age (in days)	Male		Female		χ^2	d.f.	P
	$\bar{A} + \bar{a}$	d	$\bar{A} + \bar{a}$	d			
< 1	202	6	188	3	0.30	1	> .583
Rest*	216	2	226	2	0.21	1	> .600
Totals	418	8	414	5	0.51	2	> .741
FOR TOTALS :					0.65	1	> .403

* Includes cases of unknown age-group

TABLE 4.3

Sex-variation in the incidence of tongue pigmentation in new born babies

Age (in days)	Male		Female		χ^2	d.f.	P
	$\bar{A}$	$\bar{a}$	$\bar{A}$	$\bar{a}$			
< 1	193	9	183	5	0.91	1	> .317
Rest*	212	4	224	2	0.22	1	> .584
Totals	405	13	407	7	1.13	2	> .549
FOR TOTALS					1.79	1	> .179

* Includes cases of unknown age-group

TABLE 4.4

Incidence rates of tongue pigmentation in new born babies

Population	Frequency of			Incidence rates (in percentages)	
	$\bar{A}$	$\bar{a}$	d	Minimum	Maximum
Brahmin	108	2	1	1.80	2.70
Kayastha	261	11	8	3.93	6.78
Others	443	7	4	1.54	2.42
Totals	812	20	13	2.37	3.90

TABLE 4.5

Genotypic mating	Mating frequency	Probability with which the corresponding genotypic mating type is classified under			Probability of an offspring being of the phenotype	
		$\bar{A}x\bar{A}$	$\bar{A}x\bar{a}$	$\bar{a}x\bar{a}$	$\bar{A}$	$\bar{a}$
	(1)	(2)	(3)	(4)	(5)	(6)
AA x AA	p^4	1	0	0	1	0
AA x Aa	$4p^3q$	1	0	0	1	0
Aa x Aa	$4p^2q^2$	1	0	0	$1-(\beta/4)$	$\beta/4$
AA x aa	$2p^2q^2$	$1-\beta$	β	0	1	0
Aa x aa	$4p\,q^3$	$1-\beta$	β	0	$1-(\beta/2)$	$\beta/2$
aa x aa	q^4	$(1-\beta)^2$	$2\beta(1-\beta)$	β^2	$1-\beta$	β

Table 4.6

Joint distribution of parents and offspring

Phenotypic mating Type	Frequency	Probability of an offspring being of the phenotype $\bar{A}$	$\bar{a}$	Totals	
$\bar{A} \times \bar{A}$	$\lambda_1 (N_1)$	$R_{11}/\lambda_1 (n_{11})$	$R_{12}/\lambda_1 (n_{12})$	1	(n_1)
$\bar{A} \times \bar{a}$	$\lambda_2 (N_2)$	$R_{21}/\lambda_2 (n_{21})$	$R_{22}/\lambda_2 (n_{22})$	1	(n_2)
$\bar{a} \times \bar{a}$	$\lambda_3 (N_3)$	$R_{31}/\lambda_3 (n_{31})$	$R_{32}/\lambda_3 (n_{32})$	1	(n_3)
TOTALS	1 (N)	...	...	...	(n)

Where

$$\lambda_1 = (1 - \beta q^2)^2$$

$$\lambda_2 = 2\beta q^2 (1 - \beta q^2)$$

$$\lambda_3 = 1 - \lambda_1 - \lambda_2 = \beta^2 q^4$$

$$R_{12} = \beta q^2 (1 - \beta q)^2$$

$$R_{11} = \lambda_1 - R_{12}$$

$$R_{22} = 2\beta^2 q^3 (1 - \beta q)$$

$$R_{21} = \lambda_2 - R_{22}$$

$$R_{32} = \beta^3 q^4 \qquad \text{and}$$

$$R_{31} = \lambda_3 - R_{32}$$

TABLE 4.7

Observed and expected frequencies of mating types and the corresponding offspring phenotypes under the incomplete penetrance model

Mating		Offspring phenotype	
Type	Frequency	$\bar{A}$	$\bar{a}$
$\bar{A} \times \bar{A}$	257(259.37)	645(656.21)	90(78.79)
$\bar{A} \times \bar{a}$	130(130.27)	261(232.48)	73(101.52)
$\bar{a} \times \bar{a}$	19(16.36)	5(7.05)	46(43.95)

TABLE 4.8

Estimation of q and β from family data on tongue pigmentation

Iteration No.	$\hat{q}$	$\hat{\beta}$	$\partial \hat{q}$	$\partial \hat{\beta}$	$L_{\hat{q}}$	$L_{\hat{\beta}}$
1	0.4	0.95	0.0792667	-0.0645190	334.600659	-7.984148
2	0.479267	0.885481	0.0018072	-0.0171108	-10.237299	-11.959057
3	0.481074	0.868370	0.0011396	-0.0048120	0.014650	-2.411703
6	0.482603	0.861873	0.0000228	-0.0000996	-0.004904	-0.049441
10	0.482634	0.861737	0.0000001	-0.0000006	-0.000030	-0.000281
11	0.482634	0.861736	0.0000000	-0.0000002	-0.000010	-0.000078

TABLE 4.9

Estimates of β and q, and their standard errors for Brahmin and Kayastha families

Population	Number of families	Number of iterations	$\hat{\beta}$	S.e.($\hat{\beta}$)	$\hat{q}$	S.e.($\hat{q}$)
Brahmin	52	14	0.999999	0.000340	0.314682	0.038352
Kayastha	80	14	0.818127	0.112766	0.484127	0.042930
Brahmin and Kayastha	132	13	0.891012	0.075305	0.417263	0.029208

TABLE 4.10

Frequency distributions of scores obtained by 342 students, classified according to sex and and tongue pigmentation

Score (out of 10)	Male		Female	
	Normal	Pigmented	Normal	Pigmented
2.0	11	3	9	2
2.5	4	0	0	0
3.0	22	2	3	3
3.5	3	11	1	0
4.0	25	3	21	2
4.5	23	14	8	3
5.0	35	7	19	6
5.5	13	7	3	1
6.0	26	0	4	5
6.5	7	1	4	0
7.0	3	2	1	0
7.5	2	2	1	2
8.0	4	6	1	0
8.5	2	3	1	0
9.5	0	1	0	0
Total	180	62	76	24
Mean	4.694	5.121	4.513	4.792
Variance	2.030	3.292	1.941	2.147

CHAPTER V

GENETICS OF HYPERTRICHOSIS OF THE EAR RIMS

5.1 Review of the earlier findings :

The problem of genetics of hypertrichosis of the ear rims received general attention ever since the first pedigree was published by Tommasi [38,39], and more so after Cockayne [4] postulated a dominant gene on the Y-chromosome to be responsible for the inheritance of this trait. This mode of inheritance seems to have been accepted since then. It was Stern [36] who raised doubt for the first time about the Y-linkage for the trait, and this marked the beginning of serious investigations pertaining to the genetics of the trait. Major contributors to this problem are Gates and Bhaduri [11], Dronamraju [8,9], Sarkar et al [33], Gates et al [12], Slatis and Apelbaum [34], Stern et al [37] and Chakravartti [3].

Gates and Bhaduri have reported a few infants showing hairy ears, though they did not personally examine them. They studied 21 pedigrees, including a huge one. These pedigrees were based, to a great extent, on indirect information (hear-say). They contained several normal adult sons of affected fathers, indicating failures of penetrance of the dominant gene, under Y-linkage. Thus the overall evidence contained in the pedigrees about Y-linked dominant inheritance is quite satisfactory, though, reliability of the data is questionable.

It may be noted that, when an affected son of a normal father, but whose maternal grand father is affected, was observed, the authors tried to explain them as cases of crossing over from Y → X in the maternal grand fathers. However, this argument sounds artificial, since such cases can also be explained as due to incomplete penetrance of the dominant gene in the normal fathers having affected sons. Apparently, the latter possibility did not receive much of their attention.

Subsequently, 3 pedigrees were added by Dronamraju, including one of his own. Y-linkage theory was supported by the pedigrees, barring the young normal sons of affected fathers. An attraction of this article is that two alternative modes of inheritance were also considered at length. However, though the pedigrees are compatible with one or the two alternative theories (autosomal dominant and male sex-limited), the probability of observing such compositions in the 3 pedigrees simultaneously under this theory was shown to be negligibly small.

Sarkar et al have contributed 7 more pedigrees, which are consistent with the Y-linkage theory, accepting occasional non-penetrance of the gene. Incidentally, they were the first to introduce the concept of grades, by artificially breaking u

almost continuous intensity of the trait. They introduced a system of 5 grades for the affected men.

6 more pedigrees were added by Gates et al, all of which provided good evidence in favour of the Y-linkage theory with incomplete penetrance. They concluded saying "it only remains to say that all pedigrees of hairy ears yet published are compatible with a gene in the Y-chromosome with occasional failure of penetrance, and most of them are incompatible with any other form of inheritance". The authors seem to have jumped to conclusions just on the basis of data, whose reliability is very much limited.

A leading contribution to this problem came from Slatis and Apelbaum, who analysed the data in a highly original manner. For the first time, they gave a physical shape to the penetrance of the gene responsible for the incidence of the trait. From a random sample study, the proportion of affected men was obtained for each of several age groups, and these proportions were divided by the maximum proportion occurring in any of the age groups (usually, the oldest age group). The thus revised proportions were called the 'apparent penetrances' for the respective age groups. And, the 'adjusted penetrances' were obtained through the best line of fit between age and apparent penetrance. Using these adjusted penetrances, they have

analysed their family data under the Y-linkage hypothesis, and found good evidence in favour of it. They have also analysed the same data under conditions most favourable to the autosomal dominance and recessive hypotheses, and failed to support either by them. Thus, this study provided good evidence in favour of the Y-linkage theory. It may be noted that they classified a man as affected if he had at least several coarse hairs on the top or side of the ear.

Another leading article on this topic is that of Stern et al. Their method of examination,is, perhaps, the best possible one that minimises the chances of misclassification. A man was classified as affected if he had at least on e coarse hair on at least one pinna. It was observed by them that sometimes the informant classified affected as unaffected and unaffected as affected. This tends to lower the importance of the several assertions made by previous workers in favour of the Y-linkage theory, since observations on many individuals of the pedigrees they studied were obtained through the informants. Next, analysing their family data in an ingenious manner, Stern et al have demonstrated that the trait is of hereditary importance. However, they failed to discriminate between the three possible hypotheses considered by Slatis and Apelbaum. This study, thus, turned out to be a genuine challenge to the appropriateness of

Y-linage for the trait.

Finally, Chakravartti made an attempt to investigate the acceptability of Y-linkage theory. In this study also a man was classified as affected if he had at least several coarse hairs on the top or side of the ear. The family data collected by him supported the Y-linkage inheritance model. This seems to be the last publication pertaining to the genetics of the trait, excepting the present study covered in the next section.

Thus, we see that some studies failed to support the Y-linkage hypothesis, though many of them have supported it. Nevertheless, most of the studies that turned out in favour of Y-linkage, suffered from accuracy of observation, and as such do not carry much of importance. Therefore, it was decided to collect family data and study the problem afresh. The results of this study are presented in the next section.

5.2 Present study :

Data on 168 families from West Bengal constitute the material for this study. Usually, a family was selected for this study only if both the generations were represented by some living adult males. It may be noted that in as many cases as possible, adult female members of the families also were examined, but none was seen to have

expressions. The starting or the initial estimates of m_s and n_s are given by

$$m_s = \frac{m\sqrt{n_1(n_1 + n_3 + n_5)}}{a\sqrt{n_1 + n_2}}$$

and

$$n_s = \frac{n\sqrt{n_5(n_1 + n_3 + n_5)}}{a\sqrt{n_5 + n_6}}$$

where,

$$a = m\sqrt{n_1/(n_1 + n_2)} + n\sqrt{n_5/(n_5 + n_6)}$$

The information matrix $I = ((I_{ij}))$ is computed from

$$I_{11} = 4G\left[1 + \frac{E_1}{E_2} + \frac{E_5}{E_3} + \frac{E_5}{E_4}\right]$$

$$I_{22} = 4G\left[1 + \frac{E_5}{E_6} + \frac{E_1}{E_3} + \frac{E_1}{E_4}\right]$$

$$I_{12} = 2G\left[1 + \frac{E_3}{E_4}\right] = I_{21}$$

and, the variance - covariance matrix $V = ((V_{ij}))$ of the estimates of m_s and n_s is obtained by inverting I. The efficient scores (Rao [19]) for m_s and n_s are given by

$$L_1 = \frac{2n_1 + n_3}{m_s} - \frac{2n_2\, m_s}{E_2} - \frac{n_4\, n_s}{E_4}$$

of affected men can now be computed for each of several age groups, once we classify the n subjects into several age groups. These proportions are analogous to the 'apparent penetrances' of Slatis and Apelbaum [34]. Next, the best line of fit between ages and proportions (obtained above) is obtained. This represents the penetrance curve, from where one reads off the penetrance value for any individual age. These penetrance estimates are used in the analysis that follows. Though this is an ideal method, a slightly different method has been followed here. Since the fathers of all the 168 families are affected, one son from each one of them was drawn at random. Thus, the n = 168 sons selected this way constitute our sub-sample, and the rest carries through as above. This method yields 0.258, 0.386, 0.750 and 1.000 as the proportions of affected sub-sample subjects for the age groups 18-30, 30-40, 40-50 and 50-60 years respectively.

It should be clear that these estimates are obtained under the assumption of Y-linkage of the trait, though the two autosomal theories will also be tested by using the same estimates. It may be noted that, following Slatis and Apelbaum [34], Stern et al [37] reported the apparent penetrances for west Bengal as 0.30, 0.35 and 0.59 for the age groups 20-29, 30-39 and 40-49 respectively.

A part of the differences between their estimates and our estimates can be attributed to some casual factors, and as such there do not seem to be any major differences, which, if present, may arise due to the difference between the two methods. Thus, our estimates may be used for testing the autosomal theories also.

Test for Y-linkage

The analysis is done separately for two groups of families, one consisting of families of grade-1 fathers, and the other group consisting of families of higher grade fathers. Under Y-linkage, all adult sons of all families in each of the two groups are expected to carry the gene. However, some of them may not manifest the trait, due to failures of penetrance. Therefore, the expected number of affected adult sons within each group is obtained by

$$E_A = \sum_i x_i$$

where, x_i is the penetrance value corresponding to the age of one adult son, the summation being taken over all adult sons in a group. Denoting the corresponding expected number of normal adult sons (obtained by subtraction) by E_N, the statistic for testing the Y-linkage hypothesis is given by

$$(E_A + E_N)\ (O_A - E_A)^2 / E_A E_N$$

m_s, n_s and F, denoted by $A = ((a_{ij}))$, is given by

$a_{11} = 2m_s(1-F) + F$ $\quad a_{12} = 0$ $\quad a_{13} = m_s(1-m_s)$

$a_{21} = -a_{11}$ $\quad a_{22} = 0$ $\quad a_{23} = m_S(1-m-m_s)$

$a_{31} = 2(1-F)n_s$ $\quad a_{32} = 2(1-F)m_s$ $\quad a_{33} = -2m_s n_s$

$a_{41} = -a_{31}$ $\quad a_{42} = -a_{32}$ $\quad a_{43} = -2(mn-m_s n_s)$

$a_{51} = 0$ $\quad a_{52} = 2n_s(1-F)+F$ $\quad a_{53} = n_s(1-n_s)$

$a_{61} = 0$ $\quad a_{62} = -a_{52}$ $\quad a_{63} = n_S(1-n-n_s)$.

Using the estimates obtained in Step-1, we can compute the matrix A, from where we proceed to compute the information matrix (of order 3 x 3) $I = ((I_{ij}))$, where,

$$I_{ij} = G\left[\sum_{k=1}^{6} \frac{a_{ki}\, a_{kj}}{E_k}\right] \quad \text{for } i,j = 1,2,3.$$

Denoting the variance - covariance matrix of m_s, n_s and F by $V = ((V_{ij}))$, we proceed to

Step-3: Compute the efficient scores for m_s, n_s and F by

$$L_1 = \frac{n_1+n_3}{m_s} + \frac{n_2(1-F)}{m_S(1-F)+F+2m_s(1-F)} + \frac{n_1(1-F)}{m_s(1-F)+F} - \frac{n_2}{m_s} - \frac{n_4\, n_s}{mn-m_s n_s}$$

PE_A. Similarly the other frequencies are obtained, and the analysis is incorporated in Table 5.2, which provides evidence against this mode of inheritance.

However, if the above test yields insignificant differences between the corresponding observed and expected frequencies, the autosomal theory can neither be accepted. The reason is that we are overestimating the dominant gene frequency by the above method. In such situations, one could estimate the true proportion of affected men over all ages by

$$\frac{1}{N}\sum_i (n_i/x_i)$$

where, N is the total sample size (for a population study), n_i is the number of affected males in the i-th age group, and x_i is the penetrance value for the same age group. The summation is taken over all age groups. Now the dominant gene frequency can be estimated from the above proportion, and the rest of the analysis remains the same.

Test for autosomal recessivity.

Denoting the recessive gene frequency by q, the probability that an adult son of an affected father is also affected is given by $Q = q$. With the same maximum proportion, 0.6, we get

$Q = q = 0.779$, the rest of the analysis being presented in Table 5.3. It is clear that the present study fails to support the autosomal recessive hypothesis.

Taking all the available evidences into account, we conclude that the Y-linked dominance hypothesis, with occasional failures of penetrance, is the most probable mode of inheritance for hypertrichosis of the ear rims.

TABLE 5.1

Test for Y-linkage

Sons above 19 years		Fathers of grade		Totals
		1	2,3 & 4	
Affected	Observed	39	29	68
	Expected	37.142	23.127	60.269
Normal	Observed	108	57	165
	Expected	109.858	62.873	172.731
	χ^2	0.12	2.04	2.16
	d.f.	1	1	2

For totals : $\chi^2 = 1.34$(1 d.f.)

TABLE 5.2

Test for Autosomal Dominance

Sons above 19 years		Fathers of grade 1	Fathers of grade 2,3 & 4	Totals
Affected	Observed	39	29	68
	Expected	28.228	17.576	45.804
Normal	Observed	108	57	165
	Expected	118.772	68.424	187.196
	χ^2	5.09	9.35	14.44
	d.f.	1	1	2

For totals : $\chi^2 = 13.39$(1 d.f.)

TABLE 5.3

Test for Autosomal Recessiveness

Sons above 19 years		Fathers of grade 1	Fathers of grade 2,3 & 4	Totals
Affected	Observed	39	29	68
	Expected	28.971	18.040	47.011
Normal	Observed	108	57	165
	Expected	118.029	67.960	185.989
	χ^2	4.32	8.43	12.75
	d.f.	1	1	2

For totals : χ^2 = 11.74 (1 d.f.)

CHAPTER VI

ESTIMATION OF MNS-CHROMOSOME FREQUENCIES

6.1 Maximum likelihood estimation of MNS-chromosome frequencies under random mating :

The problem of estimating the chromosome frequencies for this blood group system was considered at length by Boyd [1]. Since these estimates will be used in the next section, a brief outline of Boyd's estimation procedure is given here. The three independent parameters to be estimated are m, m_s and n_s, as stated in section 1.3.

For the six phenotypes M,MS,MN,MNS,N,NS let the observed frequencies be n_1, n_2, n_3, n_4, n_5 and n_6 respectively in a sample of size G. The expected phenotypic proportions are given by

$$M : E_1 = m_s^2 \qquad MNS : E_4 = 2(mn - m_s n_s)$$

$$MS : E_2 = m_s^2 + 2m_s m_S \qquad N : E_5 = n_s^2$$

$$MN : E_3 = 2m_s n_s \qquad NS : E_6 = n_S^2 + 2n_s n_S$$

m and n are estimated by simple gene counting,

i.e. $$m = [n_1 + n_2 + (n_3 + n_4)/2] / G$$

and, $$n = [n_5 + n_6 + (n_3 + n_4)/2] / G = 1 - m$$

Now, it remains to estimate m_s and n_s. This is done by the well-known iterative method. We shall merely give the relevant

expressions. The starting or the initial estimates of m_s and n_s are given by

$$m_s = \frac{m\sqrt{n_1(n_1 + n_3 + n_5)}}{a\sqrt{n_1 + n_2}}$$

and $$n_s = \frac{n\sqrt{n_5(n_1 + n_3 + n_5)}}{a\sqrt{n_5 + n_6}}$$

where,

$$a = m\sqrt{n_1/(n_1 + n_2)} + n\sqrt{n_5/(n_5 + n_6)}$$

The information matrix $I = ((I_{ij}))$ is computed from

$$I_{11} = 4G\left[1 + \frac{E_1}{E_2} + \frac{E_5}{E_3} + \frac{E_5}{E_4}\right]$$

$$I_{22} = 4G\left[1 + \frac{E_5}{E_6} + \frac{E_1}{E_3} + \frac{E_1}{E_4}\right]$$

$$I_{12} = 2G\left[1 + \frac{E_3}{E_4}\right] = I_{21}$$

and, the variance - covariance matrix $V = ((V_{ij}))$ of the estimates of m_s and n_s is obtained by inverting I. The efficient scores (Rao [19]) for m_s and n_s are given by

$$L_1 = \frac{2n_1 + n_3}{m_s} - \frac{2n_2\, m_s}{E_2} - \frac{n_4\, n_s}{E_4}$$

and, $$L_2 = \frac{2n_5 + n_3}{n_s} - \frac{2n_6\, n_s}{E_6} - \frac{n_4\, m_s}{E_4}$$

which lead to the correction factors for the initial estimates of m_s and n_s, given by,

$$\Delta m_s = L_1 V_{11} + L_2 V_{12}$$

$$\Delta n_s = L_1 V_{12} + L_2 V_{22}$$

thereby giving the improved estimates

$$m_s^* = m_s + \Delta m_s \qquad n_s^* = n_s + \Delta n_s$$

where, m_s and n_s are the initial estimates. If both Δm_s and Δn_s are negligibly small, our estimates are already obtained above. Otherwise, taking m_s^* and n_s^* as the initial estimates, we need to repeat the entire procedure time and again till Δm_s and Δn_s become as small as we wish. Denoting the final estimates by $\bar{m}_s$ and $\bar{n}_s$, their standard errors (s.e.) are given by

$$\text{s.e.}(\bar{m}_s) = \sqrt{V_{11}} \qquad \text{s.e.}(\bar{n}_s) = \sqrt{V_{22}}$$

$$\text{s.e.}(\bar{m}_s) = \sqrt{V(\bar{m}_s)} \quad \text{and} \quad \text{s.e.}(\bar{n}_s) = \sqrt{V(\bar{n}_s)}$$

where,

$$V(\bar{m}_s) = V_{11} + \frac{(m-2\bar{m}_s)(1-m)}{2G} \quad \text{and} \quad V(\bar{n}_s) = V_{22} + \frac{(n-2\bar{n}_s)(1-n)}{2G}$$

This brings us to the end of this section.

6.2 Maximum likelihood estimation of chromosome frequencies and the parameter F of Wright's model from MNS blood group data :

We shall now estimate the parameters m_s, n_s and F, under Wright's model. The additional parameter F is the so-called inbreeding coefficient. It may be noted that this parameter measures the deviation from random mating. Inbreeding coefficient is a component of F. Estimation of these parameters is split up into 5 steps which are described below. The method is essentially the same as 6.1.

Step-1: From the random sample of G individuals, we estimate m, m_s and n_s following section 6.2. For notational convenience we shall denote the estimates $\bar{m}_s$ and $\bar{n}_s$, thus obtained, simply by m_s and n_s. Choose a trial value of F on a priori grounds.

Step-2: The phenotypes, observed fequencies and expected proportions are denoted by the same symbols as in the preceding section. Here,

$E_1 = m_s^2(1-F) + Fm_s$ $\qquad$ $E_4 = 2(1-F)(mn - m_s n_s)$

$E_2 = m_S^2(1-F) + Fm_S + 2(1-F)m_s m_S$ $\qquad$ $E_5 = n_s^2(1-F) + Fn_s$

$E_3 = 2(1-F)\, m_s n_s$ $\qquad$ $E_6 = n_S^2(1-F)+Fn_S+2(1-F)n_s n_S$.

The matrix of partial derivatives of E's with respect to

m_s, n_s and F, denoted by $A = ((a_{ij}))$, is given by

$a_{11} = 2m_s(1-F) + F$ $\quad a_{12} = 0$ $\quad a_{13} = m_s(1-m_s)$

$a_{21} = -a_{11}$ $\quad a_{22} = 0$ $\quad a_{23} = m_s(1-m-m_s)$

$a_{31} = 2(1-F)n_s$ $\quad a_{32} = 2(1-F)m_s$ $\quad a_{33} = -2m_s n_s$

$a_{41} = -a_{31}$ $\quad a_{42} = -a_{32}$ $\quad a_{43} = -2(mn-m_s n_s)$

$a_{51} = 0$ $\quad a_{52} = 2n_s(1-F)+F$ $\quad a_{53} = n_s(1-n_s)$

$a_{61} = 0$ $\quad a_{62} = -a_{52}$ $\quad a_{63} = n_s(1-n-n_s)$.

Using the estimates obtained in Step-1, we can compute the matrix A, from where we proceed to compute the information matrix (of order 3 x 3) $I = ((I_{ij}))$, where,

$$I_{ij} = G\left[\sum_{k=1}^{6} \frac{a_{ki}\, a_{kj}}{E_k}\right] \quad \text{for} \quad i,j = 1,2,3.$$

Denoting the variance - covariance matrix of m_s, n_s and F by $V = ((V_{ij}))$, we proceed to

Step-3: Compute the efficient scores for m_s, n_s and F by

$$L_1 = \frac{n_1+n_3}{m_s} + \frac{n_2(1-F)}{m_s(1-F)+F+2m_s(1-F)} + \frac{n_1(1-F)}{m_s(1-F)+F} - \frac{n_2}{m_s} - \frac{n_4\, n_s}{mn-m_s n_s}$$

$$L_2 = \frac{n_3+n_5}{n_s} + \frac{n_6(1-F)}{n_S(1-F)+F+2n_s(1-F)} + \frac{n_5(1-F)}{n_s(1-F)+F} - \frac{n_6}{n_S} - \frac{n_4 m_s}{mn - m_s n_s}$$

$$L_3 = \frac{n_2(1-m-m_s)}{m_S(1-F)+F+2m_s(1-F)} + \frac{n_6(1-n-n_s)}{n_S(1-F)+F+2n_s(1-F)} + \frac{n_1(1-m_s)}{m_s(1-F)+F}$$

$$+ \frac{n_5(1-n_s)}{n_s(1-F)+F} - \frac{n_3+n_4}{1-F}$$

Step-4: The corrections Δm_s, Δn_s and ΔF for the initial estimates of the parameters and the improved estimates are now computed following 6.1.

If all the three correction factors are negligibly small then the above estimates are the final ones. Otherwise the cycle should be repeated from step-2 onwards till we get stable estimates. The standard erros can be computed analogous to section 6.1.

Example : The various formulae required for computations having already been given clearly in several steps, only the results of an example will be presented here. The following are the phenotypic frequencies in a sample of 805 individuals, as reported by Lester [14].

M : 145	MNS : 144
MS : 80	N : 135
MN : 240	NS : 61

Simple gene - counting yields

$\hat{m} = 0.518012$ and $\hat{n} = 1 - \hat{m} = 0.481988$

Final estimates of the chromosome frequencies and the so-called inbreeding coefficient, F, together with their standard errors are given in Table 6.1. It may be noted that it took altogether four cycles or iterations to obtain the final estimates with accuracy upto the sixth decimal place. The trial value of F was taken to be 0.2. Notice that the standard error of the estimate of F is quite large. The value of the χ^2 goodness of fit with 1 d.f. turns out to be 0.608, which demonstrates a very good fit of the model.

The rather large standard error for the estimate of F suggests that in order to obtain reliable estimate of this parameter one should base the estimation on even larger body of data. The estimated value of F should not be confused with inbreeding coeficient. F accounts for both inbreeding and several types of heterogeneities in the data which, at least at present, can not be separated one fom the other. A good discussion on this aspect is recently offered by Chakraborty [2].

6.3 Maximum likelihood estimation of chromosome frequencies from family data under restricted random mating :

The Hardy-Weinberg law (HWL) is so powerful a tool that many

population geneticists have exploited it thoroughly in many useful applications. A great application of HWL is indeed in the estimation of gene frequencies for Mendelian characters. Though the assumption of HWL greatly simplifies the labour involved in estimating the frequencies, it seems reasonable, however, that one should relax this assumption to the extent possible since the HWL does not strictly hold good in any natural population. An attempt to this effect is recently made by Adhikari, Sarma and Chakraborty (unpublished). In their work, they considered the problem of estimating gene frequencies from family data on ABO blood groups by slightly relaxing the assumption of HWL. Details of this method are given by Chakraborty [2]. They call such a mating system, the RESTRICTED RANDOM MATING (RRM). The essential purpose of this section is to present the analogus estimation of the MNS-chromosome frequencies under RRM. It may be noted here that these estimates are probably superior to those obtained under HWL. The exact extent of gain due to RRM involves the comparison of HWL and RRM which will not be attempted here. The estimation procedure is also illustrated here.

Other parameters being as in the first section, the following two functions are defined :

$$g = m_s/m \qquad \text{and} \qquad d = n_s/n$$

The 21 phenotypic mating types, their frequencies (λ_i's) and the conditional probabilities of the offspring phenotypes are shown in Table 6.2. The table also includes data from a random sample of G families (parents and their offspring). The conditional probabilities in Table 6.2 are constructed as follows :

Consider the mating type MS x MS. This phenotypic mating is split up into the three corresponding genotypic matings in the proportions as shown below.

genotypic mating type given the phenotypic mating MS x MS	probability
MS/MS x MS/MS	$(1-g)^2$
MS/MS x MS/Ms	$2g(1-g)$
MS/Ms x MS/Ms	g^2

Observe that the first two genotypic matings give offspring of the phenotype MS with probability 1. The last mating gives two types of offspring MS and M with probabilities 3/4 and 1/4 respectively. Therefore, we get

P(an offspring is of phenotype MS/the phenotypic mating is MS x MS)

$$= (1-g)^2 + 2g(1-g) + (3/4)g^2$$
$$= 1-g^2/4 .$$

Similarly the other probabilities are computed.

Observe that the parameters to be estimated are λ_i's ($i = 1,\ldots,21$), m, g and d, from where we can obtain the chromosome frequencies.

The log-likelihood of the sample (parents and offspring) is easily seen to be

$$\text{Log } L = \text{Constant} + \text{Log } L_1 + \text{Log } L_2$$

where,

$$\text{Log } L_1 = \sum_{i=1}^{21} G_1 \log \lambda_i$$

$$\begin{aligned}\text{Log } L_2 = {} & A \log g + B \log d + C \log(1-g) + n_{67} \log(1+g) \\ & + D \log(2-g) + n_1 \log(2+g) + E \log(1-d) \\ & + n_{71} \log(1+d) + F \log(2-d) + n_{113} \log(2+d) \\ & + n_{113} \log(2-g^2) + (n_{15} + n_{81}) \log(2-gd) \\ & + n_{27} \log(4-gd) + n_{69} \log(1-gd) + n_{83} \log(2-d)^2 \\ & + n_{76} \log(g+d) + n_{75} \log (2-g-d)\end{aligned}$$

where,

$$A = 2n_2+n_8+2n_{14}+n_{16}+n_{20}+n_{22}+n_{28}+n_{34}+n_{44}+2n_{68}+n_{70}+n_{74}+n_{82}+n_{88}$$

$$\begin{aligned}B = {} & n_{16}+n_{28}+n_{46}+n_{58}+n_{70}+2n_{72}+n_{78}+n_{82}+2n_{84}+n_{90}+n_{100} \\ & +n_{102}+2n_{114}+n_{120}\end{aligned}$$

$$C = n_{43}+n_{67}+n_{73}+n_{87}$$

$$D = n_{1}+n_{7}+n_{19}+n_{21}+n_{33}$$

$$E = n_{45}+n_{71}+n_{77}+n_{89} \qquad \text{and}$$

$$F = n_{57}+n_{99}+n_{101}+n_{113}+n_{119}$$

It may be noted that the above partition of L into L_1 and L_2 simplifies the estimation procedure since λ_i's are confined only in L_1 and, g and d are confined only in L_2. Thus, one obtains the maximum likelihood estimates of λ_i's as

$$\hat{\lambda}_i = G_i/G$$

and $$\hat{V}(\hat{\lambda}_i) = G_i(G - G_i)/G^3$$

$$\hat{Cov}(\hat{\lambda}_i, \hat{\lambda}_j) = -G_iG_j/G^3 \quad \text{for } i \neq j$$

$$i, j = 1,\ldots,21.$$

Now turning to the estimation of m, by gene-counting method one obtains

$$\hat{m} = (M) + (MS) + 1/2.[(MN) + (MNS)]$$

where, (M), (MS) etc. are the expected phenotypic proportions among the offspring, which are obtained from Table 6.2. One can also show that the above one is the maximum likelihood estimate of m. After substituting the expressions for (M), (MS) etc., we get

$$\hat{m} = \frac{1}{4} \cdot (4\hat{\alpha}_1 + 3\hat{\alpha}_2 + 2\hat{\alpha}_3 + \hat{\alpha}_4)$$

where,

$$\alpha_1 = \lambda_1 + \lambda_2 + \lambda_7$$

$$\alpha_2 = \lambda_3 + \lambda_4 + \lambda_8 + \lambda_9$$

$$\alpha_3 = \lambda_5 + \lambda_6 + \lambda_{10} + \lambda_{11} + \lambda_{12} + \lambda_{13} + \lambda_{16}$$

and, $\alpha_4 = \lambda_{14} + \lambda_{15} + \lambda_{17} + \lambda_{18}$

$\hat{V}(\hat{m})$ can be obtained fom the variance - covariance matrix of $\hat{\lambda}_i$'s. Note tha $\hat{V}(\hat{n}) = \hat{V}(1-\hat{m})$ is the same as $\hat{V}(\hat{m})$.

It remains now to estimate g and d for which it is enough to consider L_2. The estimation is carried out through the wellknown iterative method.

In what follows we first start with the initial estimates of g and d, say g_o and d_o, given by (De Groot [7])

$$g_o = \sqrt{\bar{c}_2/(\bar{c}_1 + \bar{c}_2)}$$

and

$$d_o = \sqrt{\bar{c}_6 / (\bar{c}_5 + \bar{c}_6)}$$

where, $\bar{c}_i$'s are the observed phenotypic proportions of the offspring in the sample of G families (Table 6.2).

In order to obtain the correction factors to the initial estimates, one obtains the efficient scores by

$$L_g = \frac{\partial \log L_2}{\partial g} = \frac{A}{g} - \frac{C}{1-g} + \frac{n_{67}}{1+g} - \frac{D}{2-g} + \frac{n_1}{2+g} - \frac{2g\, n_{13}}{2-g^2} - \frac{d(n_{15}+n_{81})}{2-gd}$$

$$- \frac{d\, n_{27}}{4-gd} - \frac{d\, n_{69}}{1-gd} + \frac{n_{76}}{g+d} - \frac{n_{75}}{2-g-d}$$

$$L_d = \frac{\partial \log L_2}{\partial d} = \frac{B}{d} - \frac{E}{1-d} + \frac{n_{71}}{1+d} - \frac{F}{2-d} + \frac{n_{113}}{2+d} - \frac{2d\, n_{83}}{2-d^2} - \frac{g(n_{15}+n_{81})}{2-gd}$$

$$- \frac{g\, n_{27}}{4-gd} - \frac{g\, n_{69}}{1-gd} + \frac{n_{76}}{g+d} - \frac{n_{75}}{2-g-d}$$

where, A,B,C,D, E and F are defined earlier; and the information matrix $I = ((I_{ij}))$, $i,j = 1,2$, is given by

$$I_{11} = \frac{R_2+R_4+R_6}{g(2-g)} + \frac{2R_8+2R_{15}+R_{13}}{4g(1-g)} + \frac{4R_1}{4-g^2} + \frac{2R_3}{2-g^2} + \frac{R_{12}}{1-g^2} \quad \frac{df}{g} + e$$

$$I_{12} = f + e$$

and, $$I_{22} = \frac{R_{10}+R_{17}+R_{20}}{d(2-d)} + \frac{2R_8+2R_{15}+R_{13}}{4d(1-d)} + \frac{R_{12}}{1-d^2}$$

$$+ \frac{2R_{14}}{2-d^2} + \frac{4R_{19}}{4-d^2} + \frac{gf}{d} + e$$

where,

$$f = \frac{R_3+R_{14}}{2(2-gd)} + \frac{R_5}{4-gd} + \frac{R_{12}}{2(1-gd)}$$

and, $$e = \frac{R_{13}}{4(2-g-d)} + \frac{R_{13}}{4(g+d)}$$

and, R_i's are the totals of offspring for each type of phenotypic mating (Table 6.2).

By substituting the initial values g_o and d_o one can evaluate the efficient scores L_g and L_d and the information matrix, and proceed as in sections 6.1 and 6.2.

Having estimated all the parameters, we now give the maximum likelihood estimates of the chromosome frequencies and their variances (De Groot [7]).

$$\hat{m}_s = \hat{g}\,\hat{m}$$

$$\hat{n}_s = \hat{d}\,\hat{n}$$

$$\hat{m}_S = (1-\hat{g})\,\hat{m}$$

$$\hat{n}_S = (1-\hat{d})\,\hat{n}$$

$$\hat{V}(\hat{m}_s) = \hat{g}^2\,\hat{V}(\hat{m}) + \hat{m}^2\,\hat{V}(\hat{g})$$

$$\hat{V}(\hat{n}_s) = \hat{d}^2\,\hat{V}(\hat{n}) + \hat{n}^2\,\hat{V}(\hat{d})$$

$$\hat{V}(\hat{m}_S) = (1-\hat{g})^2\,\hat{V}(\hat{m}) + \hat{m}^2\,\hat{V}(\hat{g})$$

$$\hat{V}(\hat{n}_S) = (1-\hat{d})^2\,\hat{V}(\hat{n}) + \hat{n}^2\,\hat{V}(\hat{d})$$

Example :

We shall illustrate here the methods developed in this section with the data presented in Table 6.2. The data were extracted from a series of two earlier publications (Sanger et al.[31] and Race et.al.[18]. However, the estimates of λ_i's and their variance-covariance matrix are not presented here. We have

$$\hat{m} = 0.554878 \qquad \hat{n} = 1 - \hat{m} = 0.445122$$

and, $$\hat{V}(\hat{m}) = \hat{V}(\hat{n}) = 0.000412 .$$

The initial estimates of g and d are given by

$$g_o = 0.403786 \quad \text{and} \quad d_o = 0.808122.$$

By using these estimates one evaluates

$$L_g = -57.391564 \quad \text{and} \quad L_d = -31.700293$$

and hence, $\Delta g_o = -0.139217$ and $\Delta d_o = -0.067030$.

Since the correction factors are quite large, we get the following improved estimates after the completion of one cycle

$$g_1 = 0.264469 \quad \text{and} \quad d_1 = 0.741092 .$$

Repeating these computations for five cycles altogether we obtain the final improved estimates, which are presented in Table

Table 6.3 along with their standard errors. Table 6.3 also presents the estimates and standard errors of several other parameters of interest.

An interesting problem is the comparison of HWL and RRM as introduced in the third section. One way of tackling this could be by obtaining the variance-covariance matrix for the estimates of g and d from family data under HWL, and comparing this with the corresponding matrix obtained under RRM. The concept of generalised variance can be profitably used here.

TABLE 6.1

Maximum likelihood estimates of the chromosome frequencies and inbreeding coefficient together with their standard errors from MNS blood group data

parameter	estimate	s.e.
m_S	0.111441	0.002156
m_s	0.406571	0.009644
n_S	0.090170	0.003749
n_s	0.391818	0.009110
F	0.052808	0.033645

TABLE 6.2

Phenotypic mating types, their frequencies and the conditional probabilities of the offspring phenotypes (sample frequencies are shown within parentheses)

Mating Type	Frequency	Offspring Conditional probabilities & observed frequencies MS	M	MNS	MN	Ns	N	Total frequency
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
MS x MS	λ_1	$(4-g^2)/4$	$g^2/4$	0	0	0	0	
	$G_1(6)$	$n_1(15)$	$n_2(0)$	$n_3(0)$	$n_4(0)$	$n_5(0)$	$n_6(0)$	$R_1(15)$
MS x M	λ_2	$(2-g)/2$	$g/2$	0	0	0	0	
	$G_2(3)$	$n_7(7)$	$n_8(1)$	$n_9(0)$	$n_{10}(0)$	$n_{11}(0)$	$n_{12}(0)$	$R_2(8)$
MS x MNS	λ_3	$(2-g^2)/4$	$g^2/4$	$(2-gd)/4$	$gd/4$	0	0	
	$G_3(17)$	$n_{13}(21)$	$n_{14}(1)$	$n_{15}(9)$	$n_{16}(3)$	$n_{17}(0)$	$n_{18}(0)$	$R_3(34)$
MS x MN	λ_4	$(2-g)/4$	$g/4$	$(2-g)/4$	$g/4$	0	0	
	$G_4(10)$	$n_{19}(11)$	$n_{20}(3)$	$n_{21}(10)$	$n_{22}(1)$	$n_{23}(0)$	$n_{24}(0)$	$R_4(25)$
MS x NS	λ_5	0	0	$(4-gd)/4$	$gd/4$	0	0	
	$G_5(6)$	$n_{25}(0)$	$n_{26}(0)$	$n_{27}(12)$	$n_{28}(2)$	$n_{29}(0)$	$n_{30}(0)$	$R_5(14)$
MS x N	λ_6	0	0	$(2-g)/2$	$g/2$	0	0	
	$G_6(6)$	$n_{31}(0)$	$n_{32}(0)$	$n_{33}(6)$	$n_{34}(6)$	$n_{35}(0)$	$n_{36}(0)$	$R_6(12)$

TABLE 6.2 (contd.)

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
M x M	λ_7	0	1	0	0	0	0	
	$G_7(0)$	$n_{37}(0)$	$n_{38}(0)$	$n_{39}(0)$	$n_{40}(0)$	$n_{41}(0)$	$n_{42}(0)$	$R_7(0)$
M x MNS	λ_8	$(1-g)/2$	$g/2$	$(1-d)/2$	$d/2$	0	0	
	$G_8(4)$	$n_{43}(10)$	$n_{44}(0)$	$n_{45}(0)$	$n_{46}(9)$	$n_{47}(0)$	$n_{48}(0)$	$R_8(19)$
M x MN	λ_9	0	1/2	0	1/2	0	0	
	$G_9(7)$	$n_{49}(0)$	$n_{50}(7)$	$n_{51}(0)$	$n_{52}(9)$	$n_{53}(0)$	$n_{54}(0)$	$R_9(16)$
M x NS	λ_{10}	0	0	$(2-d)/2$	$d/2$	0	0	
	$G_{10}(1)$	$n_{55}(0)$	$n_{56}(0)$	$n_{57}(1)$	$n_{58}(0)$	$n_{59}(0)$	$n_{60}(0)$	$R_{10}(1)$
M x N	λ_{11}	0	0	0	1	0	0	
	$G_{11}(5)$	$n_{61}(0)$	$n_{62}(0)$	$n_{63}(0)$	$n_{64}(13)$	$n_{65}(0)$	$n_{66}(0)$	$R_{11}(13)$
MNS x MNS	λ_{12}	$(1-g^2)/4$	$g^2/4$	$(1-gd)/2$	$gd/2$	$(1-d^2)/4$	$d^2/4$	
	$G_{12}(12)$	$n_{67}(8)$	$n_{68}(0)$	$n_{69}(16)$	$n_{70}(2)$	$n_{71}(4)$	$n_{72}(3)$	$R_{12}(33)$
MNS x MN	λ_{13}	$(1-g)/4$	$g/4$	$(2-g-d)/4$	$(g+d)/4$	$(1-d)/4$	$d/4$	
	$G_{13}(15)$	$n_{73}(5)$	$n_{74}(1)$	$n_{75}(15)$	$n_{76}(6)$	$n_{77}(1)$	$n_{78}(7)$	$R_{13}(35)$
MNS x NS	λ_{14}	0	0	$(2-gd)/4$	$gd/4$	$(2-d^2)/4$	$d^2/4$	
	$G_{14}(3)$	$n_{79}(0)$	$n_{80}(0)$	$n_{81}(5)$	$n_{82}(0)$	$n_{83}(2)$	$n_{84}(1)$	$R_{14}(8)$

TABLE 6.2 (contd.)

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
MNS x N	λ_{15}	0	0	(1-g)/2	g/2	(1-d)/2	d/2	
	$G_{15}(10)$	$n_{85}(0)$	$n_{86}(0)$	$n_{87}(6)$	$n_{88}(4)$	$n_{89}(5)$	$n_{90}(7)$	$R_{15}(22)$
MN x MN	λ_{16}	0	1/4	0	1/2	0	1/4	
	$G_{16}(7)$	$n_{91}(0)$	$n_{92}(2)$	$n_{93}(0)$	$n_{94}(11)$	$n_{95}(0)$	$n_{96}(3)$	$R_{16}(16)$
MN x NS	λ_{17}	0	0	(2-d)/4	d/4	(2-d)/4	d/4	
	$G_{17}(4)$	$n_{97}(0)$	$n_{98}(0)$	$n_{99}(2)$	$n_{100}(2)$	$n_{101}(1)$	$n_{102}(1)$	$R_{17}(6)$
MN x N	λ_{18}	0	0	0	1/2	0	1/2	
	$G_{18}(2)$	$n_{103}(0)$	$n_{104}(0)$	$n_{105}(0)$	$n_{106}(2)$	$n_{107}(0)$	$n_{108}(3)$	$R_{18}(5)$
NS x NS	λ_{19}	0	0	0	0	$(4-d^2)/4$	$d^2/4$	
	$G_{19}(0)$	$n_{109}(0)$	$n_{110}(0)$	$n_{111}(0)$	$n_{112}(0)$	$n_{113}(0)$	$n_{114}(0)$	$R_{19}(0)$
NS x N	λ_{20}	0	0	0	0	(2-d)/2	d/2	
	$G_{20}(4)$	$n_{115}(0)$	$n_{116}(0)$	$n_{117}(0)$	$n_{118}(0)$	$n_{119}(4)$	$n_{120}(4)$	$R_{20}(8)$
N x N	λ_{21}	0	0	0	0	0	1	
	$G_{21}(1)$	$n_{121}(0)$	$n_{122}(0)$	$n_{123}(0)$	$n_{124}(0)$	$n_{125}(0)$	$n_{126}(3)$	$R_{21}(3)$
TOTALS	1							
	$G(123)$	$C_1(77)$	$C_2(15)$	$C_3(82)$	$C_4(70)$	$C_5(17)$	$C_6(32)$	$T(293)$

TABLE 6.3

Maximum likelihood estimates and their standard errors

parameter	estimate	standard error
m	0.554878	0.020293
n	0.445122	0.020293
g	0.269478	0.047864
d	0.735158	0.060473
m_s	0.149527	0.027111
m_S	0.405351	0.030414
n_s	0.327235	0.030773
n_S	0.117917	0.027441

REFERENCES

1. Boyd, W.C.(1954). Maximum likelihood method of estimation of gene frequencies from MNS data. Amer.J.Hum.Genet.,6,1.

2. Chakraborty,R.(1970). Some considerations on population structure, genetic correlation and human multiple births. (Unpublished Ph.d.thesis).

3. Chakravartti, M.R.(1968). Hairy pinnae in Indian populations. Acta. Genet., Basel, 18, 511.

4. Cockayne,E.A. (1933). Inherited abnormalities of the skin and its appendages, London, Oxford University Press, 394.

5. Davis, T.A.(1968). 'Biology in the tropics' in Haldane and Modern Biology. ed. Dronamraju,K.R. The Johns Hopkins Press, Baltimore, 327.

6. Davis, T.A.(1970). Personal Communication.

7. De Groot, M.H.(1956). The covariance structure of maximum likelihood gene frequency estimates for the MNS system. Amer. J. Hum. Genet.,8, 229.

8. Dronamraju, K.R.(1960-61). Hypertrichosis of the pinna of the human ear, Y-linked pedigrees. J. Genet., 57,230.

9. Dronamraju, K.R.(1964). Genetic studies of Andhra Pradesh population. (Unpublished Ph.d. thesis).

10. Elandt-Johnson, R.C.(1970). Segregation analysis for complex modes of inheritance. Amer. J. Hum. Genet., 22, 129.

11. Gates, R.R. and Bhaduri.,P.N. (1961). 'The inheritance of hairy ear rims' in Mankind Monographs I.

12. Gates,R.R; Chakravartti,M.R. and Mukherjee,D.P.(1962). Final pedigrees of Y-chromosome inheritance. Amer. J. Hum. Genet., 14, 363.

13. Kirk, R.L. (1970). Personal communication.

14. Lester, F.I. (1961). Population dynamics of the sickle-cell trait in the black caribs of British Hondurus, Central America. Amer. J. Hum. Genet.,13, 233.

15. Li,C.C. and Sacks,L. (1954). The derivation of joint distribution and correlation between relatives by the use of stochastic matrices. Biometrics, 10, 347.

16. Li,C.C.(1955). Population Genetics, University of Chicago Press.

17. Patil, G.P. (1959). Contributions to estimation in a class of discrete distributions. (Unpublished Ph.d. thesis).

18. Race, R.R.; Sanger,R.; Lawler S.D. and Bertinshaw,D.(1949). The inheritance of the MNS blood groups : A second series of families. Heredity, 3, 205.

19. Rao, C.R. (1965). Linear statistical inference and its applications. John Wiley and Sons, New York.

20. Rao,D.C. (1970). Tongue pigmentation in man. Human Heredity, 20, 8.

21. Rao,D.C. (1970). Tongue pigmentation in man : A new genetic trait. Current Science, 39, 161.

22. Rao, D.C.(1970). Genetics of tongue pigmentation in man. Human Heredity (in press).

23. Rao,D.C. (1970). The relation between tongue pigmentation and mental ability. Human Heredity (in press).

24. Rao,D.C.(1970). Further analysis of family data on tongue pigmentation in man. Jap. J. Hum. Genet. (in press).

25. Rao,D.C. and Bose,M. (1970). Tongue pigmentation in new born. Jap. J. Hum. Genet. (in press).

26. Rao,D.C. and Gorai,J.K.(1970). Penetrance of tongue pigmentation allele. Jap. J. Hum. Genet.(in press).

27. Rao,D.C.(1970). 'Complex segregation analysis' : a letter to the editor. Amer. J. Hum. Genet. (in press).

28. Rao,D.C.(1970). A contribution to genetics of hypertrichosis of the ear rims. Human Heredity (in press).

29. Rao,D.C.(1970). 'Statistical methods in blood group' in Proceedings of second Matscience Conference,Madras,India

30. Rao,D.C. and Chakraborty,R. (1970). Maximum likelihood estimation of chromosome frequencies from family data on MNS blood groups. (with Sankhyā).

31. Sanger,R.; Race,R.R.; Walsh,R.J. and Montgomery,C.(1948). An antibody which subdivides the human MN blood groups. Heredity, 2, 131.

32. Sanghvi,L.D. (1969). Personal Communication.

33. Sarkar,S.S.;Banerjee,A.R.;Bhattacharjee,P. and Stern,C.(1961). A contribution to the genetics of hypertrichosis of the ear rims. Amer. J. Hum. Grenet.,13, 214.

34. Slatis,M.M. and Appelbaum,A.(1963). Hairy pinnae of the ear in Israeli populations. Amer.J.Hum.Genet.,15,74.

35. Smith,C.A.B. (1956). A test for segregation ratios in family data. Ann. Hum. Genet.,20, 257.

36. Stern,C.(1957). The problem of complete Y-linkage in man. Amer.J.Hum.Genet.,9, 147.

37. Stern, C.; Centerwall, W.R. and Sarkar, S.S. (1964). New data on the problem of Y-linkage of hairy pinnae. Amer.J. Hum. Genet., 16, 455.

38. Tommasi, C. (1907). Ipertricosi auricolare famigliare. Arch. Psichiatr. Neuropat. Antropol. Crim. Med. Legale., 28, 60.

39. Tommasi, C. (1907). Ipertricosi auricolare famigliare, Giorn. Psych. Clin. Tech. Manic., 35, 1.