MARIO DI BACCO, PIERO CAVINA*,
ELSA PACCIANI**

Age at death diagnosis by cranial suture obliteration: a bayesian approach.

Serie Ricerche 1995, n.1

*Laureato frequentatore del Dipartimento di Scienze Statistiche "Paolo Fortunati"
** Sopraintendenza Archeologica per la Toscana via Della Pergola 65, Firenze

Dipartimento di Scienze Statistiche "Paolo Fortunati"
Università degli studi di Bologna

Finito di stampare nel mese di Maggio 1995
presso le Officine Grafiche TECNOPRINT S.N.C.
Via del Legatore 3, Bologna.

1. Purpose of the paper.

The problem of determining the age at death by cranial suture obliteration has been a topic of discussion since the past century; nevertheless, the scientific debate still remains open and includes on one hand positions that are totally critical, and on the other hand more favorable contributions, however shoving much difficulty in providing the definition of satisfactory indicators.

The low diagnostic power that has been revealed, up to now, by the various works on different populations can be partly explained by the lack on an appropriate use of the statistical reasoning when searching for predicting rules.

Thus, an experimental control has been conducted on a sample of 221 skulls of adult individuals - from the age of 20 - (127 males and 94 females) originating from various regions of Italy,who died approximately at the end of the last century. The sample is housed in the "Museo Nazionale di Antropologia ed Etnologia" in Florence. The age and sex have been recorded in the catalogue of the Museum and were obtained from the Registry Office.

The reference methodology of collecting data is the one proposed by Acsádi and Nemeskéri (1970) and later used by other authors; it is based on the subdivision of the main vault sutures (coronal, sagittal and lambdoid) into segments (sixteen segments on the ectocranial side, on the basis of detectable morphological differences, according to Vallois, 1937; sixteen segments of approximately equal length, on the endocranial side) and on further evaluating the stage of obliteration of each segment by the Ribbe's scale (1885), that we modified by unifying the first two degrees (0 and 1). In this way a scale of four degrees was obtained (1, 2, 3, 4). The reason for this modification is the difficulty encountered in distinguishing between the two stages on the skull, and it is our belief that it may not have a biological meaning but may be a post-mortem artifact.

2. Tools for the Bayesian Analysis.

The aim of our work is to inquire how cranial suture obliteration can be used to diagnose the age at death and whether it is a "good" way to carry out such a diagnosis. In order to answer these questions we will be referring to statistical tools of the classical Bayesian Decision Theory (Berger, 1985).

In this section, our statistical approach will be described and, so as to make our expounding easier, the matter will be organized into two subsections. Following, in subsection 2.1 we will give a rather informal presentation of the Bayesian approach to the problem. Whilst subsection 2.2 will be devoted to a thorough deepening of the matter.

2.1. A sketch of the procedure.

Let us suppose that on the skull of the human H we observed - in segment i of the ectocranial [endocranial] side - the "obliteration degree" (OD) x_i, $1 \le x_i \le 4$ [y_i , $1 \le x_i \le 4$], $i = 1,..., 16$, so that

$(x_1,..., x_{16}, y_1,..., y_{16})$ = (shortly) = (x,y)

is the global information about the OD of H's skull, that is H's Global Obliteration (GO).
Then two questions will arise:

a) by (also) using the information (x,y), what is the number to be chosen for H's age at death? And

b) is (x,y) a "good" piece of information so as to diagnose the age at death?

Our aim is just to answer these questions. Then, let us call *One* the person who carries out the diagnosis, and if H died during his A^{th} year of life - i.e. after his A^{th} and before his $(A+1)^{th}$ birthday - we define A the H's age at death,

$A = 20, 21, ..., 84, 85.$

Then our procedure is arranged in the following steps:

A) Choosing among numbers 20, ..., 85 the "optimal" diagnosis, when *One* is in the initial (or prior) state of information, i.e. before learning the H's GO.

To do this, he assigns the initial (or prior) probabilities P(A) thereby expressing his belief that "H died at age A"

$$(P(A) > 0, \quad \sum_{A=20}^{85} P(A) = 1).$$

Moreover, *One* must also numerically evaluate the consequences to be faced should his diagnosis $\hat{A}$ be incorrect, since he comes to the knowledge that the real age of H at death is A. Then *One* weighs the possible consequences of each diagnosis with probabilities P(A) and his optimal diagnosis will be, of course, the one which shows the most favorable balance. Let us denote $\hat{A}_0$ as this "optimal" diagnosis and R_0 the numerical appraisal of the risk in choosing $\hat{A}_0$.

In subsection 2.2 we shall discuss how *One* - in the present circumstances - could numerically evaluate the consequences of each diagnosis and how he could assign the initial probabilities P(A).

B) Choosing - as in A) - the "optimal diagnosis, when H's GO is known to be (x,y).

Then, first of all, *One* should update his initial probabilities P(A) with final (or "a posteriori") probabilities $P(A \mid x,y)$. In order to make such an updating, *One* assigns the probability to "H's GO is (x,y)" conditional to each hypothesis "H died at age A". These probabilities (also called likelihoods) will be denoted $P(x,y \mid A)$.

In subsection 2.3 a way of assigning them will be suggested . After this assignment, final probabilities $P(A \mid x,y)$ result by compounding initial probabilities P(A) and likelihoods $P(x,y \mid A)$ according to Bayes' theorem.

At this point *One* could choose the "optimal" diagnosis by repeating step A but, obviously, using probabilities $P(A \mid x,y)$ instead of probabilities P(A). Let $\hat{A}_{x,y}$ be the "optimal" diagnosis and - analogously with step A - R(x,y) the risk *One* faces when choosing it.

C) It should be noticed that in step B *One*'s information about H's GO is assumed to be only hypothetical - this is the reason why we have often used the conditional verb. Indeed, inquiring on GO's effectiveness for diagnosis of H's age at death assumes a concrete significance only if *One* has not yet decided whether H's skull is going to be surveyed or not; so that, the evaluation of its effectiveness can be done only simulating that the GO's result (x,y) is known. Therefore step B should be ideally reiterated for all possible GO results.

Then, let us call {X,Y} the set of possible results: by reiterating step B we obtain the corresponding set $\{\hat{A}_{x,y}\}$ of the possible "optimal" diagnoses and the set {R(x,y)} of the possible risks. Clearly, the probability that *One* faces the risk R(x,y) is just the probability *One* assigns to "the GO's result (x,y) will be observed unconditional to any hypothesis on H's age at death": we denote P(x,y) such probability, (x,y) varying in the set {X,Y}.

Then, so to appraise prospectively the effectiveness of GO in age at death diagnosis *One* has - first of all - to weigh the possible consequences of using GO - i.e. the risks in the set {R(x,y)} - with probabilities P(x,y) of the correlative set {X,Y}. Now, let us denote R_1 the numerical result of this balance among the risks, as in step A we have called R_0 the initial risk. It is natural to think that the judgment on the effectiveness of GO rises from the comparison between R_0 and R_1. The simplest way to make this is to calculate

$$\Delta = \frac{R_0 - R_1}{R_0} = 1 - \frac{R_1}{R_0}$$

that we call Expected Relative Improvement (ERI). It is easily perceived that ERI varies between 0 and 1. In subsection 2.2 we will describe ERI a little more formally and we will outline the way we propose to estimate it empirically but reliably, will be outlined.

D) Let us briefly return to step B, where it was mentioned that, in order to update probabilities P(A) with probabilities $P(A \mid x,y)$, *One* needs to assign - or to elicitate - likelihoods $P(x,y \mid A)$. To make such assignments - or elicitations - our sample of skulls will be used; although the procedure followed is the standard Bayesian one, its description involves some formal complications that will take place in subsection 2.2, step D. On the other hand, these likelihood elicitations make the search of the optimal diagnosis $\hat{A}_{x,y}$ rather tedious. In other words if *One* observes H's GO (x,y) and - following step B - wants to reach $\hat{A}_{x,y}$, he will encounter some practical difficulties in the calculations. So as to avoid them, we propose that *One*, after learning GO (x,y), diagnoses H's age at death by using an approximate tool which has been empirically proved to work well: here we roughly explain such an approximate tool.

After using - as it was earlier mentioned - the sample to elicitate likelihoods, we have diagnosed the age at death of each individual following step B. Let us call $\hat{A}(x_j,y_j)$ the diagnosis for the j^{th} individual, being (x_j,y_j) his GO. We will see in subsection 2.3 that vectors x_j and y_j can be reduced to vectors $(n_{1j},...,n_{4j}) = n_j$ and, respectively, $(m_{1j},...,m_{4j}) = m_j$, where n_{kj} $[m_{kj}]$ is the number of the ectocranial [endocranial] segments whose OD is k, k=1,...,4, so that

$$\sum_{k=1}^{4} n_{kj} = \sum_{k=1}^{4} m_{kj} = 16.$$

Having said this, we have ascertained that the regression of $\hat{A}(x_j,y_j)$ on (n_j,m_j) is most satisfactory. Therefore, for the reasons that will be explained in subsection 2.2 step E, *One*, by being aware that on the ectocranial (endocranial) side of the j^{th} skull n_{kj} (m_{kj}) segments have OD k (k=1,...,4) is able to quickly diagnose H's age at death using a linear form in $(n_{1j},\dots,n_{4j})$ and $(m_{1j},\dots,m_{4j})$ whose coefficients are given in section 3.

2.2. More Details

Here we state formally the exact diagnostic procedure - and how to appreciate its effectiveness - described roughly in the previous subsection 2.1. Even now the matter is arranged in steps and in each of them we deepen the topics sketched in the corresponding ones of 2.1.

A) As it was noticed in step A of 2.1., if *One* chooses diagnosis $\hat{A}$ and then learns that the true H's age at death is A, he faces some disagreeable consequences; such consequences are more and more disagreeable as $|A-\hat{A}|$ increases. It is well known that there is a suitable way to express numerically the comparison among degrees of disagreeableness, that is to say the quadratic loss:

$$k\,(A-\hat{A})^2 \tag{2.1.1}$$

being k a constant (which is unessential, as we shall see).

On the other hand , suppose *One* has assigned the initial (or prior) probability P(A) to "H died at age A". Then, if he chooses diagnosis $\hat{A}$, he faces the random consequence $k(A-\hat{A})^2$ whose probability distribution is determined in an obvious way by probabilities P(A), A = 20,...,85. According to von Neumann-Morgenstern theory the consistent tool to synthesize random consequences is to calculate the Expectation or Risk, i.e. in our case:

$$R(\hat{A}) = \sum_{A=20}^{85} k\,(A-\hat{A})^2\,P(A) \tag{2.2.2}$$

and $R(\hat{A})$ is *One*'s risk when he chooses diagnosis $\hat{A}$. Of course the least disagreable consequence is the one minimizing the risk (2.2.2). It is an easy matter to realize that such an "optimal" diagnosis is

$\hat{A}_0$ = the integer nearest to expectation E_0 (2.2.3)

where expectation E_0 is

$$E_0 = \sum_{A=20}^{85} A\,P(A) \tag{2.2.4}$$

Then the *least risk in the initial state of information* - i.e. before learning H's GO obliteration - is

$$R_0 = R(\hat{A}_0) = \sum_{A=20}^{85} k\,(A-\hat{A}_0)^2\,P(A) \tag{2.2.5}$$

Now let us proceed to the evaluation of prior probabilities P(A), A = 20,...,85. As we have already said P(A) is the probability *One* assigns to "H's age at death is A" in his initial state of information, i.e. before knowing that H's GO is (x,y). This means that *One*'s opinions on the age at death of H have been formed on other information. Therefore the evaluation of P(A) appears to be a rather elusive problem since it is obviously unrealistic to suppose that all who have to diagnose the H's age at death know exactly the same things. However this question becomes clear if we put it in the context of our search for the effectiveness of GO for diagnosing age at death. In fact our answer is based - see step C in 2.1 - on the Expected Relative Improvement (ERI), i.e. on the comparison between initial risk R_0 in Eq. (2.2.5) and final risk we will exhaustively describe in the next step. Then suppose that *One*'s initial state of information is very poor. It is easy to realize that in such a situation the information about H's GO has the greatest efficiency for the age at death diagnosis, so that it is possible to evaluate GO's effectiveness when it works in the more favorable conditions. On the other hand it is realistic to assume that in a "very poor" state of information *One* at least knows - approximately, perhaps - when and where H died. It is hard to believe that in a real situation he knows less of this! Indeed, let us accept this view; remember, also, that GO's effectiveness will be tested by means of a sample of skulls of Italian adults dead towards the end of the last century (see section 1). Then we assume the prior probabilities to be:

$$P(A) = \frac{f(A)}{\sum_{A=20}^{85} f(A)} \qquad A = 20,...,85 \qquad (2.2.6)$$

where f(A) is the number of Italian who died at age A during the period 1880-1890. (Source: Istituto Centrale di Statistica - ISTAT). On the other hand, it is instructive to point out that the shape of the resulting frequency polygon is typically lexian; therefore values (2.2.6) are a "good" elicitation of the initial probabilities for any "normal" or "standard" population.

B) According to step B in 2.1, suppose *One* learns that H's GO is (x,y). Then he updates prior probability P(A) with posterior (or final) probability P(A | x,y) compounding P(A) and likelihood P(x,y | A) - see (2.1.1) - via Bayes' theorem. Indeed, he puts:

$$P(A \mid x,y) = \frac{P(A)\, P(x,y \mid A)}{\sum_{A=20}^{85} P(A)\, P(x,y \mid A)} \qquad A = 20,...,85 \qquad (2.2.7)$$

being:

$$P(x,y) = \sum_{A=20}^{85} P(A)\, P(x,y \mid A) \qquad (2.2.8)$$

the (unconditional) probability that *One* must assign to the event "on H's skull GO (x,y) will be observed".

Of course, now *One* shall revise his diagnosis; to do this he replaces P(A) with P(A | x,y) in Eq. (2.2.2); therefore his "new" risk choosing $\hat{A}$ is:

$$R(\hat{A};x,y) = \sum_{A=20}^{85} k\,(A-\hat{A})^2\, P(A \mid x,y)$$

and, consequently his "new" diagnosis is

$$\hat{A}_{x,y} = \text{the integer nearest to } E_{x,y} \qquad (2.2.9)$$

where

$$E_{x,y} = \sum_{A=20}^{85} A\, P(A \mid x,y)$$

Obviously, One's risk for diagnosis $\hat{A}_{x,y}$ becomes:

$$R(\hat{A}_{x,y};x,y) = \sum_{A=20}^{85} k\,(A-\hat{A}_{x,y})^2\, P(A \mid x,y) \qquad (2.2.10)$$

Notice that we need to evaluate P(A | x,y) to reach Eq. (2.2.10). For doing this the likelihoods P(x,y | A), A = 20,...,85 are essential; therefore it is a very important matter to express them using sample data. All this will be discussed in step D.

C) As it was already pointed out (C step in 2.1), testing GO effectiveness implies that we have to imagine step B is reiterated for all of the possible GO results, forming the set {X,Y}.

Then their correlative possible diagnoses and possible risks form respectively the sets {$\hat{A}$(X,Y)} and {R(X,Y)}. That is, *One* - if he decides to survey H's GO - faces the random risk R(X,Y) and the probability that his/her actual risk is R(x,y) obviously coincides with probability P(x,y) in Eq. (2.2.8). In other words R(X,Y) is a random consequence, governed by the probability distribution {P(x,y); (x,y)∈(X,Y)}. Again, according to von Neumann-Morgenster theory, the Expected Risk:

$$R_1 = \sum_{(x,y)\in(X,Y)} R(\hat{A}_{x,y};\, x,y)\, P(x,y) \qquad (2.2.11)$$

synthesizes perspectively what *One* should face diagnosing H's age at death, having to survey H's GO, but before knowing its result.

Then it is obvious that ERI in (2.1.3) can be written:

$$\Delta = \frac{R_0 - R_1}{R_0} = 1 - \sum_{(x,y)\in(X,Y)} \frac{R(\hat{A}_{x,y})}{R_0} P(x,y) \qquad (2.2.12)$$

Direct computation of (2.2.12) involves a lot of calculations, because the set $\{(X,Y)\}$ is a big one indeed. But we have the sample of skulls described in section 1, which is large enough and with a wide assortment of ages at death. Therefore it is possible to estimate ERI in a satisfactory way and with relatively little effort.

Let us denote $_AT$ ($_AT \geq 0$) the number of skulls that in our sample are humans died at age A (A = 20,...,85) and (x_j,y_j) the GO of the j^{th} of them ($j = 1,2,...,{}_AT$).

Then , if $\hat{A}_{x_j,y_j}$ is the age at death that *One* - according to (2.2.9) - should diagnose, put

$$Q(A) = \frac{\sum_{j=1}^{_AT} (A - \hat{A}_{x_j,y_j})^2}{_AT} \qquad \text{when } {}_AT > 0$$

$$Q(A) = 0 \qquad \text{when } {}_AT = 0$$

It is easy to realize that

$$\sum_{A=20}^{85} Q(A)\, P(A)$$

is a "good" estimate of risk R_1, as defined in (2.2.11). Then, agreeing again on the evaluations of P(A) in Eq. (2.2.6), we put

$$Q = \frac{\sum_{A=20}^{85} Q(A)\, f(A)}{\sum_{A=20}^{85} f(A)} \qquad (2.2.13)$$

so that :

$$D = 1 - \frac{Q}{R_0} \qquad (2.2.14)$$

is a satisfactory estimate of ERI.

D) As it was pointed out (D step in 2.1) elicitation of $P(x,y \mid A)$ is necessary for updating P(A) with $P(A \mid x,y)$, A = 20,...,85. Now we take up such matter.

Suppose that it is known that H died at age A. Then let us describe the survey of H's GO with a very simple model. Namely:

$$x = (x_1,...,x_i,...,x_{16})$$

(i.e. the OD on the sixteen ectocranial sections) is obtained as if sixteen drawings have been made with replacement from urn $EC\bigcup_A$ in which the ratio of balls labelled l - where l is obliteration degree l - is $_Ap_l$ (l=1,...,4).

Likewise, the outcomes

$$y = (y_1,...,y_i,...,y_{16})$$

(i.e. the OD on the sixteen endocranial sections) are described with a wholly analogous model, but urn $EC\bigcup_A$ is replaced with urn $EN\bigcup_A$ in which the ratio of balls labelled l is $_Aq_l$ (l = 1,...,4), possibly different from $_Ap_l$. Put:

$$_Ap = ({}_Ap_1,...,{}_Ap_4),\ {}_Aq = ({}_Aq_1,...,{}_Aq_4)$$

and denote

$$n_l, m_l \qquad (l = 1,\dots,4) \qquad (2.2.15)$$

the number of the ectocranial and, respectively, endocranial sections on which OD is l, so that

$$\sum_{l=1}^{4} n_l = \sum_{l=1}^{4} m_l = 16$$

Then the hypothesis that ${}_Ap$, ${}_Aq$ and age at death A are known, our models give probability $P(x,y \mid A, {}_Ap, {}_Aq)$ of "H's GO is (x,y)" and it results:

$$P(x,y \mid A, {}_Ap, {}_Aq) = \frac{16!}{n_1! \dots n_4!} \prod_{l=1}^{4} {}_Ap_l^{\,n_l} \frac{16!}{m_1! \dots m_4!} \prod_{l=1}^{4} {}_Ap_l^{\,m_l} \qquad (2.2.16)$$

(i.e. the product of two multinomial distributions).

But there is no doubt that *One* is uncertain on both ${}_Ap$ and ${}_Aq$; so they are random vectors which are supposed to be continuos. Then denote $\varphi({}_Ap,{}_Aq)$ the personal probability density function (p.d.f.) *One* gives them. On the other hand, the sample of T skulls we have already described in section 1 is available. Then there are the following information:

$${}_AT$$

$$({}_Ax,{}_Ay) = ({}_Ax_1,{}_Ay_1\,; \dots ;\, {}_Ax_j,{}_Ay_j\,; \dots ;\, {}_Ax_{{}_AT},{}_Ay_{{}_AT}) \qquad (2.2.17)$$

i.e. the number of skulls belonging to humans who died at age A and respectively the GO of each of them. It will be thanks to this information that *One* is able to update the initial opinion on ${}_Ap$, ${}_Aq$ in the light of new information. Indeed, since the knowledge of the relation between obliteration and age is very vague, it seems very realistic in the present situation to suppose that the sample information $({}_Ax,{}_Ay)$ is able to wield the greatest possible influence on

the final p.d.f. $\varphi({}_Ap,{}_Aq \mid A;\, {}_Ax,{}_Ay)$. Then to render this task precise, assume that ${}_Ap$ and ${}_Aq$ are stochastically independent, so

$$j({}_Ap,{}_Aq) = j1({}_Ap)\, j2({}_Aq) \qquad (2.2.18)$$

and both φ_1 and φ_2 are uniform.

Using the "vague" prior (2.2.18) information, $({}_Ax,{}_Ay)$ are determining factors of the p.d.f. $\varphi({}_Ap,{}_Aq \mid A;\, {}_Ax,{}_Ay)$. On the other hand let us accept for sample information $({}_Ax,{}_Ay)$ in (2.2.17) our urn-models. Then, after denoting - analogously to (2.2.15) -

$${}_An_{jl},\ {}_Am_{jl} \qquad j = 1,\dots,{}_AT \qquad l = 1,\dots,4$$

the number of ectocranial and, respectively, endocranial sections of j^{th} skull among the ${}_AT$ ones whose OD is l, we have:

$$P({}_Ax,{}_Ay \mid A, {}_Ap, {}_Aq) =$$

$$= \prod_{j=1}^{{}_AT} \frac{16!}{{}_An_{j1}! \dots {}_An_{j4}!} \frac{16!}{{}_Am_{j1}! \dots {}_Am_{j4}!} \prod_{l=1}^{4} {}_Ap_l^{\,{}_An_{jl}}\ {}_Aq_l^{\,{}_Am_{jl}} \qquad (2.2.19)$$

or, putting:

$$\sum_{j=1}^{{}_AT} {}_An_{jl} = {}_AN_l \qquad \sum_{j=1}^{{}_AT} {}_Am_{jl} = {}_AM_l \qquad (2.2.20)$$

$$P({}_Ax,{}_Ay \mid A, {}_Ap, {}_Aq) =$$

$$= \frac{16!}{{}_AN_1! \dots {}_AN_4!} \frac{16!}{{}_AM_1! \dots {}_AM_4!} \prod_{l=1}^{4} {}_Ap_l^{\,{}_AN_l}\ {}_Aq_l^{\,{}_AM_l} \qquad (2.2.21)$$

Then, the non informative p.d.f. (2.2.18) and likelihood (2.2.21) give -via Bayes theorem - the updated p.d.f. for $({}_Ap, {}_Aq)$:

$$\varphi({}_Ap,{}_Aq \mid A; {}_Ax,{}_Ay) = h \prod_{l=1}^{4} {}_Ap_l^{\,{}_AN_l}\ {}_Aq_l^{\,{}_AM_l} \tag{2.2.22}$$

being h a normalizing constant.

Now it only remains to make the last step to have our wanted probability $P(x,y \mid A)$, averaging the conditional probability (2.2.16) by means of the (updated) p.d.f. (2.2.22). Standard calculations give:

$$P(x,y \mid A) = \frac{B(n+{}_AN)}{B({}_AN)}\frac{B(m+{}_AM)}{B({}_AM)} \tag{2.2.23}$$

where:

$$n = (n_1,\ldots,n_4) \qquad\qquad m = (m_1,\ldots,m_4) \tag{2.2.24}$$

being n_l $[m_l]$ the number of ectocranial [endocranial] sections of H's skull whose OD is l, l = 1,...,4 and

$${}_AN = ({}_AN_1,\ldots{}_AN_4) \qquad\qquad {}_AM = ({}_AM_1,\ldots{}_AM_4)$$

with ${}_AN_l$ and ${}_AM_l$ (l = 1,..4) defined in (2.2.19) and at last

$$B(z) = \frac{\prod_{l=1}^{4} z_l!}{\left(\sum_{l=1}^{4} z_l\right)!} \tag{2.2.25}$$

for $z = n+{}_AN$ or $m+{}_AM$ or ${}_AN$ or ${}_AM$.

E) Therefore to make optimal diagnosis of H's age at death after learning her/his GO (x,y), firstly (x,y) is reduced to $n = (n_1,\ldots,n_4)$, $m = (m_1,\ldots,m_4)$ according to (2.2.24). Then, using (2.2.23) likelihoods $P(x,y \mid A)$ are obtained for A = 20,...,85. Inserting such values and initial probabilities (2.2.6) in (2.2.7) *One* obtains updated probabilities $P(A \mid x,y)$, A=20,...,85; then Eq. (2.2.9) gives $P(A \mid x,y)$, A=20,...,85. Lastly, (2.2.9) gives optimal diagnosis $\hat{A}_{x,y}$.

Such procedure needs tedious calculations because of the evaluation just of the likelihoods; so there is some good reason to simplify it.

For this aim, consider again the reduction (n,m) in (2.2.24) of H's GO (x,y). Then define the "linear" diagnosis:

$$A_{x,y}(a_0,\alpha,\beta) = a_0 + \sum_{l=1}^{4} \alpha_l n_l + \sum_{l=1}^{4} \beta_l m_l \tag{2.2.26}$$

where a_0, $\alpha = (\alpha_1,\ldots,\alpha_4)$, $\beta = (\beta_1,\ldots,\beta_4)$ will be chosen so that when replacing the optimal diagnosis $\hat{A}_{x,y}$ with $A_{x,y}(a_0,\alpha,\beta)$ the Expected Relative Improvement (ERI) (2.2.12) decreases as little as possible. Formally, let us denote

$$\{n_1(x),\ldots,n_4(x); m_1(y),\ldots,m_4(y); (x,y) \in (X,Y)\} =$$

$$= \{n(x), m(y); (x,y) \in (X,Y)\} \tag{2.2.27}$$

the set of the "reductions" of all possible GO's results $(x,y) \in (X,Y)$. Then the ERI using $A_{x,y}(a_0,\alpha,\beta)$ is:

$$\Delta(a_0,\alpha,\beta) = 1 - \sum_{(x,y) \in (X,Y)} \frac{R[A_{x,y}(a_0,\alpha,\beta)]}{R_0} P(x,y) \tag{2.2.28}$$

Then the problem is to choose a_0,α,β so that on comparing (2.2.12) and (2.2.28) it results, for a_0^*,α^*,β^*:

$$0 \le \Delta - \Delta(a_0^*,\alpha^*,\beta^*) = \text{minimum} \tag{2.2.29}$$

Since quadratic loss is used, it is easy to realize that Eq. (2.2.28) is equivalent to:

$$\sum_{(x,y) \in (X,Y)} \{\hat{A}_{x,y} - a_0^* - \sum_{l=1}^{4} a_l^* n_l(x) - \sum_{l=1}^{4} \beta_l^* m_l(y)\}^2 P(x,y) = \tag{2.2.30}$$

= minimum with respect to any choice of (a_0,α,β).

Now denote ρ^2 the squared multiple correlation index computed between "dependent" variable $\hat{A}_{x,y}$ and "independent" or "explaining" variables n(x), m(y) with $(x,y) \in (X,Y)$. then it could be easily shown that the minimum (2.2.29) is equal to $1-\rho^2$. On the other hand, it is possible to obtain a good evaluation of a_0^*, α^*, β^* using the sample of skulls we have already described, i.e. minimizing with respect to a_0,α,β the quantity:

$$\sum_{i=1}^{T} \{\hat{A}(x_i,y_i) - a_0 - \sum_{l=1}^{4} a_l n_{il} - \sum_{l=1}^{4} \beta_l m_{il}\}^2$$

where $\hat{A}(x_i,y_i)$ is the "optimal" diagnosis for the i^{th} skull whose GO is (x_i,y_i) and (n_i,m_i) the corresponding "reduction" (i=1,...,T). Likewise, the same sample data can be used in an obvious way to evaluate ρ^2.

2.3. Results.

Following the procedure outlined in 2.1 and specifically stated in 2.2, the age at death of each member of the sample was diagnosed. Therefore each diagnosis was made using the summarized data n_{ij}, m_{ij} i = 1,...,4, where n_{ij} [m_{ij}] is the number of the ectocranial [endocranial] segments of the j^{th} skull whose OD is i.

All diagnosed ages are recorded in the third column of table 1 and each one of them can be immediately compared with the corresponding real age that is recorded in the second column. Such comparisons should suggest that GO is only acceptable to diagnose the youngest ages and the senile ones, while it appears to be disappointing when considering the central ages.

This conclusion is confirmed by the ERI evaluation: according to Eq. (2.2.5), (2.2.13), (2.2.14) we have:

Least risk in the initial state of information

$R_0 = 380.9$

Estimated Expected Risk

$Q = 177.3$

Estimated Expected Relative Improvement

$D = 0.53$

that is, in a "standard" population - i.e. in a population whose death distribution is Lexian - learning about the skull GO reduces the risk of misdiagnosis to a half: it is easy to realize that this unpleasant matter is due to the heavy influence of the GO's diagnostic unreliability just concerning the central ages.

In 2.2, step E, it was said that an high value of the squared multiple correlation coefficient ρ^2 between (n_i,m_i i= 1,...,4) variables and the diagnosed ages allows to replace the exact but painful diagnosis with a linear approximation (see, in particular, 2.2.31).

We actually find:

$\rho^2 = 0.915$

so that, if on the H's skull (x,y) GO was observed, then it seems justified to diagnose the H's age at death using the "reduced" information $\{n_1,...,n_4; m_1,...,m_4\}$ in the linear form:

$$A(a_0,\alpha,\beta) = a_0+\alpha_1 n_1+\alpha_2 n_2+\alpha_3 n_3+\alpha_4 n_4+\beta_1 m_1+\beta_2 m_2+\beta_3 m_3+\beta_4 m_4 \quad (3.1)$$

where the values of the coefficients have been evaluated by means of the sample, and result:

$a_0 = 74.35$

$\alpha_1 = -1.64 \quad \beta_1 = -1.68$

$\alpha_2 = -0.80 \quad \beta_2 = -0.47$

$\alpha_3 = -0.55 \quad \beta_3 = -0.22$

$\alpha_4 = -0.20 \quad \beta_4 = 0.82$

In literature the multiple regression has been already used to solve diagnosis problems using GO. We point out that our linear approximation rests on the sound foundations of the Bayesian Statistical Decision Theory.

Table 1: age-at-death diagnosis for each individual

n *individual No.;*
a *observed age at death;*
b *optimal diagnosis;*
c *estimated "linear" diagnosis;*
d,..,g *number of ectocranial sections whose OD is, respectively, 1,..,4;*
h,..,k *number of endocranial sections whose OD is, respectively, 1,..,4.*

n	a	b	c	d	e	f	g	h	i	j	k
1	20	27	31.2	13	3	0	0	10	5	1	0
2	20	27	34.1	15	1	0	0	6	10	0	0
3	20	28	21.9	15	1	0	0	16	0	0	0
4	20	27	22.3	16	0	0	0	15	1	0	0
5	20	24	25.9	16	0	0	0	12	4	0	0
6	20	25	25.6	15	1	0	0	13	3	0	0
7	20	25	24.7	16	0	0	0	13	3	0	0
8	21	43	42.8	6	7	3	0	9	4	3	0
9	21	26	23.5	16	0	0	0	14	2	0	0
10	21	28	33.7	14	2	0	0	7	9	0	0
11	22	29	21.1	16	0	0	0	16	0	0	0
12	22	25	27.3	13	3	0	0	13	3	0	0
13	22	31	36.4	14	2	0	0	5	10	1	0
14	22	43	45.4	11	4	0	1	1	10	5	0
15	22	29	21.1	16	0	0	0	16	0	0	0
16	23	25	29.6	16	0	0	0	9	7	0	0
17	23	34	40.4	10	3	1	2	7	6	3	0
18	23	36	38.3	9	6	1	0	8	7	1	0
19	23	33	37.8	15	1	0	0	4	7	5	0
20	23	60	41.8	13	3	0	0	1	15	0	0
21	23	53	45.4	8	8	0	0	2	12	2	0
22	23	25	27.3	13	3	0	0	13	3	0	0
23	23	40	45.2	8	8	0	0	4	5	6	1
24	23	56	57.7	6	10	0	0	0	2	9	5
25	23	26	31.3	14	2	0	0	9	7	0	0
26	23	26	24.4	15	1	0	0	14	2	0	0
27	23	25	30.8	16	0	0	0	8	8	0	0
28	23	24	28.0	15	1	0	0	11	5	0	0
29	23	28	22.8	14	2	0	0	16	0	0	0
30	24	55	41.5	12	4	0	0	2	14	0	0
31	24	25	30.4	15	1	0	0	9	7	0	0
32	24	34	37.5	11	5	0	0	6	10	0	0
33	24	27	32.1	13	3	0	0	9	7	0	0
34	24	25	27.3	13	3	0	0	13	3	0	0

35	24	24	27.2	16	0	0	0	11	5	0	0
36	24	25	29.2	15	1	0	0	10	6	0	0
37	24	25	30.8	16	0	0	0	8	8	0	0
38	24	26	23.5	16	0	0	0	14	2	0	0
39	24	35	39.7	16	0	0	0	3	10	1	2
40	24	26	28.7	13	3	0	0	12	3	1	0
41	25	51	55.7	5	10	1	0	0	6	8	2
42	25	52	53.7	7	9	0	0	0	6	7	3
43	25	45	53.1	9	1	0	6	4	5	1	6
44	25	48	57.3	4	10	1	1	1	4	8	3
45	25	36	36.7	14	2	0	0	6	3	7	0
46	26	50	40.0	14	2	0	0	2	13	1	0
47	26	30	31.2	13	3	0	0	11	3	1	1
48	26	26	33.2	16	0	0	0	6	10	0	0
49	26	26	26.1	13	3	0	0	14	2	0	0
50	27	27	31.2	13	3	0	0	10	5	1	0
51	27	33	36.3	14	1	0	1	7	5	3	1
52	27	38	45.7	13	3	0	0	1	8	5	2
53	27	53	56.3	3	13	0	0	0	5	9	2
54	27	45	39.4	15	1	0	0	2	12	2	0
55	27	27	24.0	14	2	0	0	15	1	0	0
56	27	25	28.9	14	2	0	0	11	5	0	0
57	28	28	21.9	15	1	0	0	16	0	0	0
58	28	52	44.5	9	7	0	0	2	12	2	0
59	28	39	44.2	10	6	0	0	3	8	4	1
60	28	39	41.6	8	7	1	0	6	9	1	0
61	28	40	41.2	11	5	0	0	4	11	0	1
62	28	25	26.4	14	2	0	0	13	3	0	0
63	28	39	39.0	12	4	0	0	4	12	0	0
64	29	29	21.1	16	0	0	0	16	0	0	0
65	30	26	24.4	15	1	0	0	14	2	0	0
66	30	28	34.2	8	6	2	0	13	3	0	0
67	30	34	37.1	12	4	0	0	6	8	2	0
68	30	27	22.3	16	0	0	0	15	1	0	0
69	30	25	29.1	14	2	0	0	11	4	1	0
70	31	62	59.7	0	16	0	0	0	6	7	3
71	31	33	37.5	13	3	0	0	5	9	2	0
72	31	59	64.4	1	8	5	2	0	7	8	1
73	32	51	47.1	10	6	0	0	0	10	6	0
74	32	32	36.3	13	3	0	0	6	8	2	0
75	32	66	53.1	6	4	1	5	6	0	4	4
76	33	28	33.3	13	3	0	0	8	8	0	0
77	33	59	60.5	3	12	0	1	0	3	8	5
78	33	56	53.3	4	11	1	0	0	11	5	0
79	33	42	42.1	7	9	0	0	6	5	5	0
80	34	38	39.3	7	8	0	1	8	7	1	0
81	34	39	38.4	12	4	0	0	6	3	7	0

82	34	52	53.3	14	2	0	0	0	5	3	8
83	34	58	58.6	4	12	0	0	0	1	11	4
84	35	65	65.1	3	10	1	2	0	1	8	7
85	35	59	56.7	6	10	0	0	0	6	5	5
86	35	64	64.5	0	13	2	1	0	4	8	4
87	35	54	46.0	8	6	2	0	6	1	8	1
88	35	50	52.2	4	12	0	0	1	8	6	1
89	35	56	58.8	2	10	2	2	0	10	5	1
90	35	26	24.4	15	1	0	0	14	2	0	0
91	35	65	61.2	0	13	0	3	0	3	11	2
92	35	76	68.7	0	5	2	9	0	2	11	3
93	35	62	63.7	2	8	5	1	0	4	11	1
94	36	54	55.5	3	13	0	0	0	0	16	0
95	36	29	21.1	16	0	0	0	16	0	0	0
96	36	67	58.9	7	5	0	4	2	4	2	8
97	36	29	28.9	13	2	1	0	14	0	1	1
98	36	36	46.0	11	5	0	0	3	6	4	3
99	36	57	50.5	6	10	0	0	0	14	1	1
100	37	33	33.5	9	6	1	0	12	3	1	0
101	37	46	40.3	12	4	0	0	3	13	0	0
102	37	59	60.3	2	11	1	2	0	7	6	3
103	37	25	30.1	14	2	0	0	10	6	0	0
104	38	27	31.5	14	2	0	0	9	6	1	0
105	38	25	26.4	15	0	0	0	13	3	0	0
106	38	44	46.2	9	7	0	0	4	2	8	2
107	38	61	42.7	14	2	0	0	0	14	2	0
108	38	40	41.6	7	9	0	0	6	7	3	0
109	38	45	46.6	7	8	1	0	4	4	8	0
110	39	50	51.2	6	10	0	0	0	7	9	0
111	39	25	25.6	15	1	0	0	13	3	0	0
112	39	44	45.5	8	8	0	0	5	2	7	2
113	40	42	45.2	10	6	0	0	2	6	8	0
114	40	54	58.1	5	9	2	0	0	6	7	3
115	40	47	49.8	9	6	1	0	7	0	1	8
116	40	27	23.2	15	1	0	0	15	1	0	0
117	40	28	33.7	14	2	0	0	7	9	0	0
118	40	33	35.9	12	4	0	0	7	7	2	0
119	41	58	57.8	0	16	0	0	0	5	10	1
120	41	57	58.6	1	14	1	0	0	4	11	1
121	42	42	46.9	6	10	0	0	4	5	6	1
122	43	61	48.0	9	6	1	0	0	15	1	0
123	43	57	60.7	2	12	1	1	0	3	10	3
124	43	75	73.5	0	8	6	2	0	0	10	6
125	43	67	66.5	0	12	3	1	0	2	9	4
126	44	47	49.7	7	7	2	0	3	7	5	1
127	44	62	60.8	0	14	1	1	0	5	9	2
128	44	53	53.7	4	12	0	0	0	4	12	0

129	44	29	24.0	14	0	0	2	16	0	0	0
130	44	33	29.1	8	8	0	0	15	1	0	0
131	45	46	51.7	9	4	2	1	6	2	1	7
132	46	56	57.3	2	12	1	1	0	4	12	0
133	46	58	57.2	6	9	1	0	0	1	12	3
134	47	60	63.7	0	9	5	2	0	9	7	0
135	47	46	48.6	7	8	0	1	3	3	9	1
136	48	69	69.3	2	10	3	1	0	0	8	8
137	48	36	40.7	13	3	0	0	3	8	5	0
138	48	64	61.8	4	10	1	1	0	0	11	5
139	48	60	60.4	0	13	0	1	0	6	6	1
140	49	57	55.8	2	12	1	1	0	11	4	0
141	49	47	43.5	7	9	0	0	4	11	1	0
142	50	52	56.0	4	11	1	0	0	4	11	1
143	50	61	60.7	1	13	1	1	0	2	12	2
144	50	32	36.1	13	3	0	0	7	7	1	1
145	50	34	36.2	12	4	0	0	7	6	3	0
146	51	54	53.6	3	13	0	0	1	10	3	2
147	53	57	44.9	8	8	0	0	2	14	0	0
148	53	64	62.5	0	14	1	1	0	4	7	3
149	55	47	44.2	11	5	0	0	2	10	0	0
150	55	55	56.0	7	7	0	2	0	6	6	4
151	55	60	58.6	3	13	0	0	0	0	13	3
152	56	71	67.1	2	8	3	3	0	1	10	5
153	56	46	53.0	5	11	0	0	2	4	7	3
154	56	48	45.8	13	2	1	0	6	0	4	6
155	57	66	63.7	1	11	1	3	0	3	9	4
156	57	79	78.7	1	7	3	5	0	0	2	14
157	58	67	63.9	0	14	0	2	0	2	9	5
158	58	69	63.4	6	6	1	3	0	0	9	7
159	58	72	65.9	1	9	3	3	0	1	12	3
160	58	66	65.9	3	9	3	1	0	2	8	6
161	60	70	68.0	0	10	5	1	0	2	11	3
162	60	52	49.8	7	8	1	0	5	2	5	4
163	60	80	81.5	2	6	5	3	0	0	0	16
164	60	79	78.1	0	8	1	7	0	0	2	14
165	60	44	53.8	6	6	1	3	2	6	6	2
166	60	74	72.0	1	14	1	0	0	0	3	13
167	60	68	66.2	0	12	0	4	0	2	8	6
168	60	74	75.4	2	8	4	2	0	0	4	12
169	60	29	21.1	16	0	0	0	16	0	0	0
170	61	66	62.8	0	12	1	3	0	2	12	2
171	61	53	55.1	4	10	2	0	0	9	7	0
172	62	67	67.2	1	12	3	0	0	1	9	6
173	62	65	65.2	0	14	2	0	0	3	8	5
174	63	57	45.7	9	7	0	0	1	13	2	0
175	64	65	50.2	8	4	3	1	6	0	7	3

176	64	68	59.9	2	8	0	6	0	0	16	0
177	64	53	55.6	3	13	0	0	0	5	5	0
178	65	68	65.3	0	12	1	3	0	0	12	4
179	65	61	63.5	0	11	3	2	0	5	8	1
180	65	59	59.6	0	15	1	0	0	7	4	1
181	66	43	40.7	10	6	0	0	4	12	0	0
182	66	58	53.5	5	9	2	0	0	12	4	0
183	66	71	71.9	0	9	6	1	0	4	6	6
184	67	60	63.4	3	8	5	0	0	4	6	1
185	69	67	63.4	7	6	1	2	0	4	1	9
186	69	53	54.9	3	13	0	0	0	5	8	0
187	70	66	57.9	5	7	1	3	5	0	3	8
188	70	66	59.1	6	6	4	0	0	1	14	1
189	71	72	73.2	2	10	4	0	0	4	0	12
190	71	70	69.6	0	10	6	0	0	0	0	0
191	72	63	66.1	0	10	6	0	0	6	7	1
192	72	61	56.7	6	6	1	3	6	0	1	9
193	72	65	62.0	5	6	3	2	4	2	1	9
194	72	69	72.2	0	5	0	11	0	3	5	8
195	73	65	63.6	0	13	0	1	0	3	9	4
196	74	56	54.6	3	12	1	0	0	10	5	0
197	75	72	70.9	2	9	5	0	0	2	6	8
198	76	77	73.4	0	4	7	5	0	0	10	2
199	76	60	60.5	4	10	0	2	0	2	9	5
200	79	69	70.4	2	9	3	2	0	2	5	9
201	79	78	81.8	0	8	6	2	0	0	2	14
202	80	55	53.5	3	13	0	0	0	8	8	0
203	80	62	61.6	0	16	0	0	0	0	0	0
204	80	65	75.8	0	2	0	14	0	0	7	9
205	80	56	56.9	0	16	0	0	0	10	0	0
206	80	73	73.3	1	4	8	3	0	0	0	0
207	80	48	60.3	6	0	1	9	6	0	1	9
208	81	69	67.0	2	9	3	2	0	1	8	5
209	82	59	58.2	6	7	1	2	4	3	0	9
210	82	76	73.7	4	7	2	3	0	0	2	14
211	82	72	71.4	0	9	4	3	0	0	10	6
212	82	72	71.4	0	9	4	3	0	0	10	6
213	82	84	83.0	0	2	0	14	0	0	0	16
214	82	45	45.9	14	0	2	0	7	0	2	7
215	82	46	53.7	4	11	1	0	3	4	6	3
216	83	64	65.1	0	10	3	2	0	4	10	2
217	83	60	52.4	13	1	0	2	0	4	7	5
218	84	79	76.0	3	9	2	2	0	0	0	16
219	84	84	81.8	0	4	0	12	0	0	0	16
220	84	69	70.6	2	8	4	2	1	1	5	9
221	85	66	64.7	2	13	1	0	0	1	8	7

3. References

G. Acsádi and J. Nemeskéri (1980), *History of human life span and mortality*, Akadémiai Kiado, Budapest.

J.O. Berger (1985), *Statistical decision theory and bayesian analysis (second edition)*, Springer-Verlag, New York.

J.E. Buikstra, L.W. Konigsberg (1985), *American Antropologist*, n. 87 p. 316-333.

C. Masset (1971), *Bulletins et mémoires de la Société d'Anthropologie de Paris*, n. 7 p. 85-105.

C. Masset (1989), in *Iscan M.Y.: Age markers in the human skeleton*, Thomas, Springfield.

W.R.K. Perizonius (1984), *Journal of Human Evolution* n.13 p. 201-216.

R. Pressat (1961), *L'analyse démographique*, Presses universitaires de France, Paris.

F.C. Ribble (1885), *Etude sur l'ordre d'obliteration des sutures du crâne dans les races humaines*. Thèse de Médicine, Paris.

H.V. Vallois (1937), *L'Anthropologie* n.19 p. 499-532.