Applied Mathematical Sciences, Vol. 8, 2014, no. 140, 6963 - 6967
HIKARI Ltd, www.m-hikari.com
http://dx.doi.org/10.12988/ams.2014.48617

Bayesian Networks Use in Simple Maternity Problems

Marina A. P. Andrade and M. A. M. Ferreira

Instituto Universitário de Lisboa (ISCTE – IUL)
BRU – IUL, Lisboa, Portugal

Abstract

The use of DNA profiles in forensic identification problems are common procedure nowadays. Here it is intended to exemplify how to use the analysis of DNA profiles to solve the problem of simple maternity search. For this aim it is necessary to make use of a probabilistic expert system (PES), in the case an object-oriented Bayesian network (OOBN).

Keywords: Maternity search, Bayesian networks, DNA profiles

1 Introduction

The use of networks that "carry" probabilities arose with the geneticist Sewall Wright in 1921. Their used was widespread in several forms in various areas as social sciences and economics. The models are, in general, linear. Examples are the Path Diagrams or Structural Equation Models (SEM). The non-linear models, Bayesian networks or Probabilistic Expert Systems (PES), are usually used in artificial intelligence.

Here the aim is the approach of simple maternity problems: to identify if a woman is the mother of a child – commonly a dead child. To solve the real problems an analysis is used through object-oriented Bayesian networks with Hugin[1] software.

---

[1] www.hugin.com – OOBN a resource available in the Hugin 6.4 software.

2 Bayes' law

In each case there the evidence available is unique. But the use of the known Bayes' law allows quantifying the evidence weight in favour of the accused. Generally there are two competing hypotheses, the prosecution hypothesis and the defence hypothesis that must be evaluated. Thus, it is possible to determine the conditional probability of the prosecution hypothesis given the evidence versus the defence hypothesis given the evidence. The ratio of the two hypotheses -. *posterior odds* - measures the weight of the evidence in favour of the prosecution hypothesis given the evidence.

The *posterior odds* may be decomposed as the product of the likelihood ratio and the *prior odds*. The last one should be updated by the judge or jury, or in many occasions is considered equal to one since it is considered equal prior probabilities for the two hypotheses in dispute. Consequently the likelihood ratio is of great importance to measure the weight of the evidence given the evidence observed, in each case.

The set of markers used in forensic identification are in different chromosomes therefore they are independent. To compute the overall likelihood after having the likelihood for each marker it is only necessary to multiply the results for the markers studied in each case.

3 Simple maternity search problems

According to Francisco Corte-Real, INML[2], last year were carried out 5709 kinship biological research exams, concerning 1217 judicial processes. In the previous year (2010) had been carried out 5595 exams in the context of 1379 examinations processes.

The specialist stated that through these tests, paternities and maternities are looked for, being confirmed that the first are the most frequent. In the case of maternity examinations, these are requested in cases like crimes of infanticide when a newborn is found dead and it is necessary to identify the mother, but, according to Francisco Corte-Real, are much less frequent.

According to this charge, approximately 90% of examinations are referring to doubts raised by parents about the paternity of the children and the remaining 10% are made in their private capacity.

In the INML took place, only last year, a total of 4405 forensic examinations for biological processes involving criminal cases as 1165 rapes and murders. In 2010, had been carried out 4795 concerning 1105 examinations processes.

[2] INML – Instituto Nacional de Medicina Legal (National Legal Medicine Institute)

Under 60 processes were conducted, for individual examinations of genetic identification, 203 (142 in 2010, to 88 processes), usually requested by prosecutors and for identification of corpses or parts. *End of citation*[3].

This exhibits how often, and in what context, the DNA profiles analysis is a practice for forensic proposes in Portugal. Here it will be considered the situation of maternity examinations cases, much less usual than the paternity ones, particularly the requested in cases like crimes of infanticide, when a newborn is found dead and it is imperative to identify the mother. It is assumed that are available the genotypes from the putative mother (*pmgt*) and from the child (*chgt*). This is generally called the simple maternity search. The hypotheses in dispute are:

H_P: The putative mother is the true mother of the child

vs.

H_D: The true mother of the child is another individual chosen randomly from the population, not related with the putative mother.

The evidence is $\boldsymbol{E} = (\boldsymbol{chgt}, \boldsymbol{pmgt})$ – the child genotype and the putative mother genotype. The *posterior odds* is, as in (Ferreira and Andrade, 2009),

$$\frac{\boldsymbol{P(H_P|E)}}{\boldsymbol{P(H_D|E)}} = \frac{\boldsymbol{P(E|H_P)}}{\boldsymbol{P(E|H_D)}} * \frac{\boldsymbol{P(H_P)}}{\boldsymbol{P(H_D)}} \qquad (\mathbf{1})$$

and assuming $\boldsymbol{P(H_P) = P(H_D)}$ as usual,

$$\frac{\boldsymbol{P(H_P|E)}}{\boldsymbol{P(H_D|E)}} = \frac{\boldsymbol{P(E|H_P)}}{\boldsymbol{P(E|H_D)}} \qquad (\mathbf{2}).$$

To compute the likelihood ratio it may be used a Bayesian network, Fig. 1, where the nodes pg and mg are of class *founder* (single node network which states are the observed alleles with the observed population frequencies). Nodes pmgt and chgt are of class *genotype* (representing the individuals genotypes). Nodes tpg and tmg specify whether the corresponding allele belongs to the putative mother. If ch_match_pm? is true the child allele is identical to the one of the putative mother, if not it is chosen randomly in the population. Node chmg defines the Mendel inheritance being the allele of the individual chose randomly after the ancestral

---

[3] http://www.publico.pt/Sociedade/mais-de-5700-testes-de-paternidade-realizados-em-2011-1529758, 19.01.212 translated.

alleles. Node chpg is the other element of the child genotype pair chosen randomly from the population.

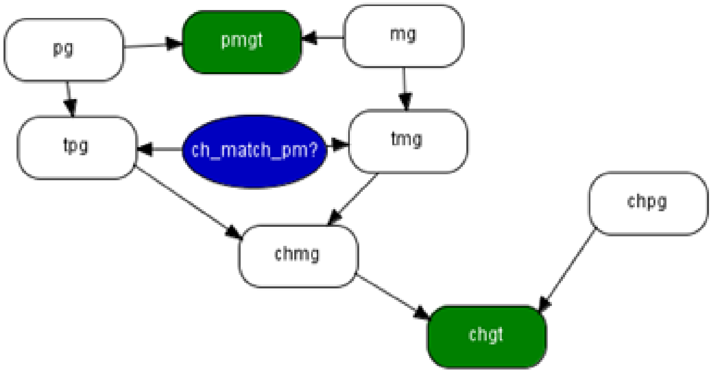

Figure 1: Simple maternity network.

3 Example results

To exemplify the application of the described tools, three markers (FGA, D21S11 and PENTA D) were chosen. In Table 1 are presented the genotypes and the allelic frequencies in the population.

Marker	Allele Frequencies				*chgt, pmgt*
FGA	p_{20}	p_{21}	p_{24}	p_{25}	(20, 24); (24,25)
	0.1421	0.1768	0.1325	0.0718	
D21S11	p_{27}	p_{29}	p_{30}	$p_{31.2}$	(27,31.2); (29,31.2)
	0.0246	0.2136	0.2437	0.1138	
PENTA D	p_9	p_{11}	p_{13}	p_{15}	(11, 13); (9,11)
	0.1984	0.1777	0.2066	0.0250	

Table 1: Genetic profiles and population frequencies for the chosen markers

The obtained results are presented in Table 2. It must be highlighted that when there is a share of an allele with low frequencies in the population the probability of the yes hypothesis is very high, as it is expected.

Marker		FGA	D21S11	PENTA D	Rescaled
ch_match_pm	yes	0.7848	0.6872	0.5845	0.9185
	no	0.2152	0.3128	0.4155	0.0815

Table 2: Analysis results

In the last column, named rescaled, it is presented the product for the three markers and it is rescaled so that the two possible values may add to 1. With the value obtained the Judge or the Jury can multiply the prior to determine the posterior probabilities for each hypothesis and its ratio give the posterior odds for the case.

References

[1] D. Abrantes, M. L. Pontes, M. F. Pinheiro, M. Andrade and M. A. M. Ferreira, Towards a systematic probabilistic evaluation of parentage casework in forensic genetics: A modest attempt to define a general standardized approach to simple and complex cases, *Forensic Science International*: Genetics Supplement **Series 1**, (2008), 635-637.

[2] M. Andrade, A Note on Foundations of Probability, *Journal of Mathematics and Technology*, **vol.. 1 (1)**, (2010), 96-98.

[3] M. Andrade and M. A. M. Ferreira, Evaluation of Paternities with less usual Data using Bayesian Networks", BMEI 2010 – 2010 3rd International Conference on Biomedical Engineering and Informatics. Proceedings, vol.6, pp. 2475-2477, 2010.

[4] M. Andrade, M. A. M. Ferreira, J. A. Filipe and M. Coelho, Paternity dispute: is it important to be conservative?, *Aplimat – Journal of Applied Mathematics*, **vol. 1 (2)**, (2008).

[5] M. A. M. Ferreira and M. Andrade, A note on Dawnie Wolfe Steadman, Bradley J. Adams, and Lyle W. Konigsberg, Statistical Basis for Positive Identification in Forensic Anthropology. American Journal of Physical Anthropology 131: 15-26 (2006), *International Journal of Academic Research*, **vol. 1 (2)**, (2009), 23-26.

[6] M. Andrade and M. A. M. Ferreira, Simple Maternity Search with Bayesian Networks, 12th Conference on Applied Mathematics APLIMAT 2013 Proceedings, (2013), 18-22.

Received: August 11, 2014