Sampling in Bayesian networks

Roberto Puch-Solis

FORSTAT Workshop
Edinburgh, Scotland, June 13th - 14th 2008

FORENSIC SCIENCE SERVICE

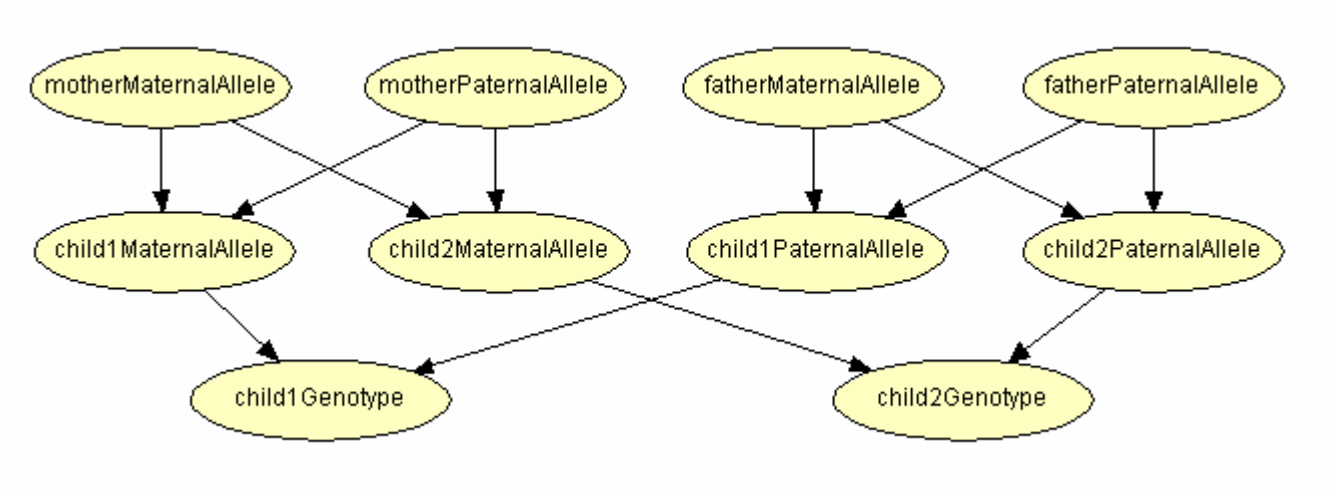
Familial searching query

- I have a DNA profile
- No match against National DNA database
- It is believed that crime runs in families
- How does profiles of siblings look like?
  - How many alleles in common do they have?
- In a locus, sibling genotypes can have none, one or two alleles in common
- Quantification of practitioners experience/expert opinion/intuition

FORENSIC SCIENCE SERVICE

Overview

- Aim is to generate genotypes of siblings at random
- Draw a Bayesian network, displayed below, step by step
- Obtain samples from the network



A Bayesian network for modeling the pedigree of a pair of siblings

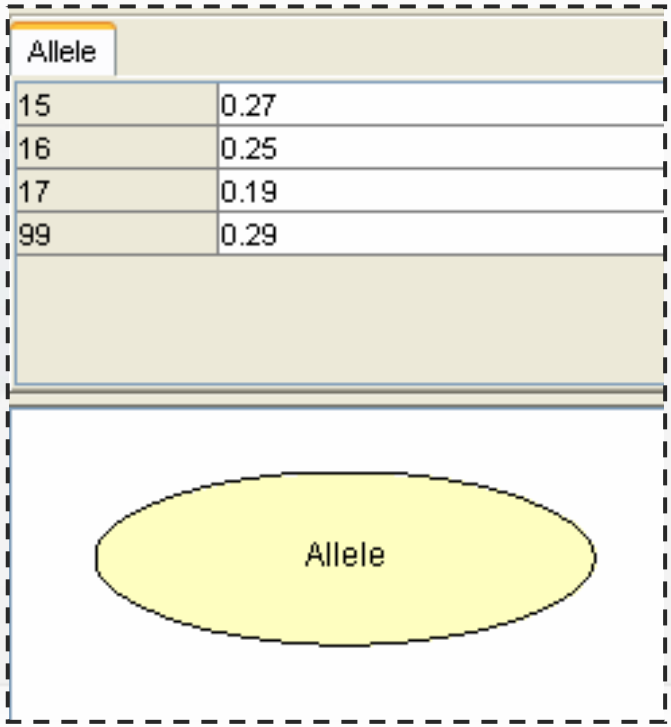
(A.P. Dawid, J. Mortera, V.L. Pascali, D. van Boxel (2002). Probabilistic expert systems for forensic inference from genetic markers, Scan. J. Stat. 29, 577-595.)

FORENSIC SCIENCE SERVICE

Modeling an allele

Allele: the allele that a person will pass to a child

We know nothing about the person, so he/she can pass any allele



FORENSIC SCIENCE SERVICE

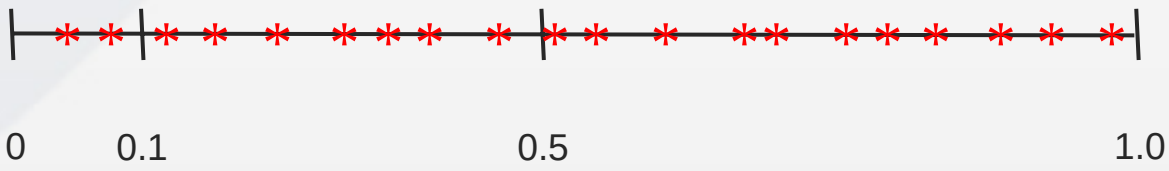
An example of Sampling a discrete distribution

$$X = \begin{cases} 1 & \text{with prob. } 0.1 \\ 2 & \text{with prob. } 0.4 \\ 3 & \text{with prob. } 0.5 \end{cases}$$

Shoot points uniformly and at random

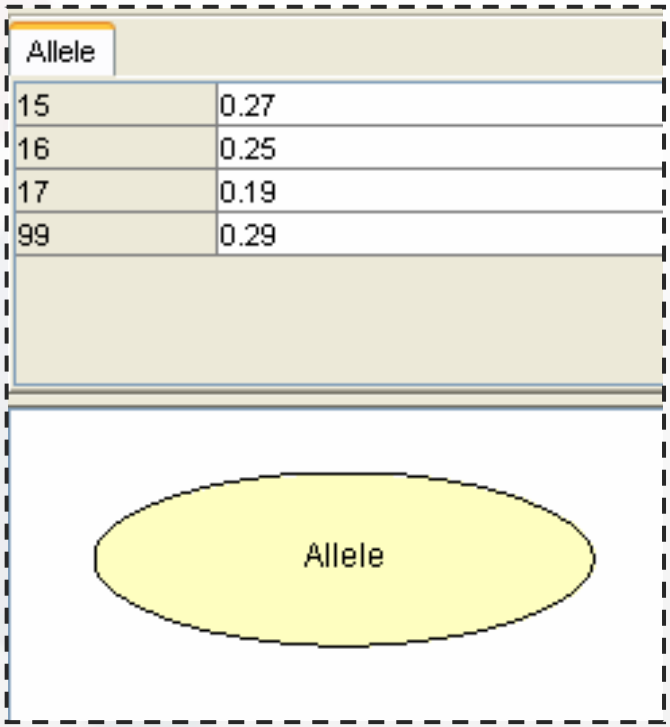
and

count how many fall in each of interval



FORENSIC SCIENCE SERVICE

Allele modeling



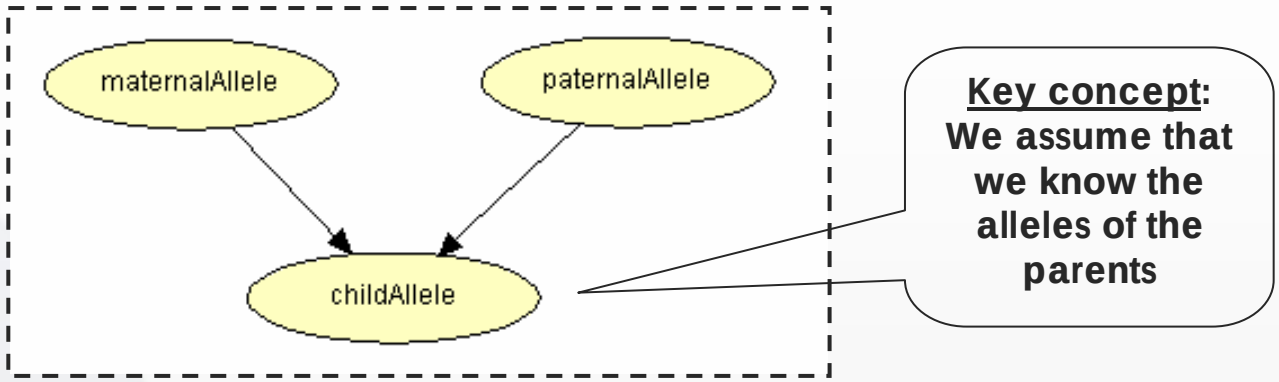
sampling

Allele
15
16
17
99
99
15
15
17
15
15

FORENSIC SCIENCE SERVICE

Allele inheritance

allele: the allele that a person will pass to a child



maternalAllele	paternalAllele	childGenotype	childAllele
15	16	15,16	15 or 16
15	15	15,15	15

Allele inheritance

$$\Pr\left(childAllele = 15 \mid maternalAllele = 15,\ paternalAllele = 16\right)$$

$$= \begin{cases} 15 & \text{with prob. } 0.5 \\ 16 & \text{with prob. } 0.5 \end{cases}$$

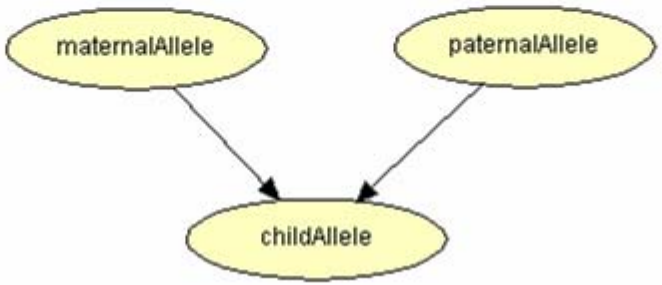
$$\Pr\left(childAllele = 15 \mid maternalAllele = 15,\ paternalAllele = 15\right) = 1$$

FORENSIC SCIENCE SERVICE

Allele inheritance

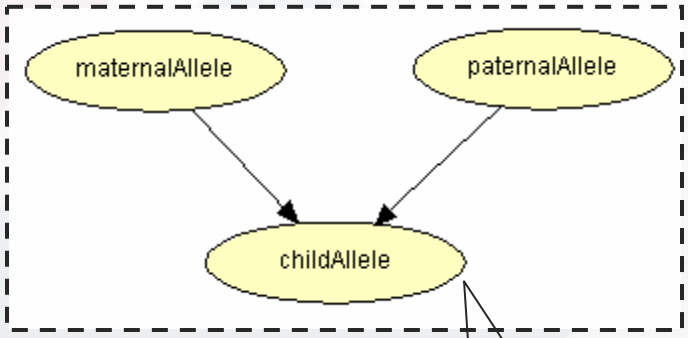
maternalAllele | paternalAllele | childAllele

paternalAllele	15				16				17				99			
maternalAllele	15	16	17	99	15	16	17	99	15	16	17	99	15	16	17	99
15	1	0.5	0.5	0.5	0.5	0	0	0	0.5	0	0	0	1	0.5	0.5	0.5
16	0	0.5	0	0	0.5	1	0.5	0.5	0	0.5	0	0	0	0.5	0	0
17	0	0	0.5	0	0	0	0.5	0	0.5	0.5	1	0.5	0	0	0.5	0
99	0	0	0	0.5	0	0	0	0.5	0	0	0	0.5	0	0	0	0.5



FORENSIC SCIENCE SERVICE

Allele inheritance: sample



sampling

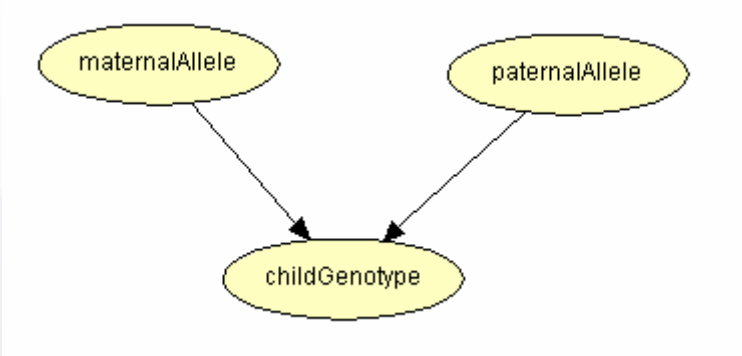
maternalAllele	paternalAllele	childAllele
99	17	17
17	16	16
17	16	17
15	15	15
99	15	99
15	17	17
15	15	15
16	15	15
16	99	16
17	99	17

Key concept:
We assume that we know the alleles of the parents

FORENSIC SCIENCE SERVICE

Genotype inheritance

childGenotype: the allele that a person will pass to a child



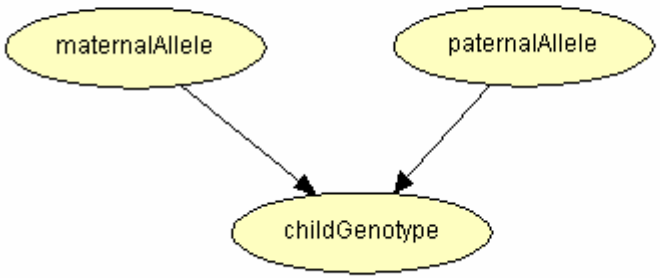
maternalAllele	paternalAllele	childGenotype
15	16	15,16
15	15	15,15

FORENSIC SCIENCE SERVICE

Genotype inheritance

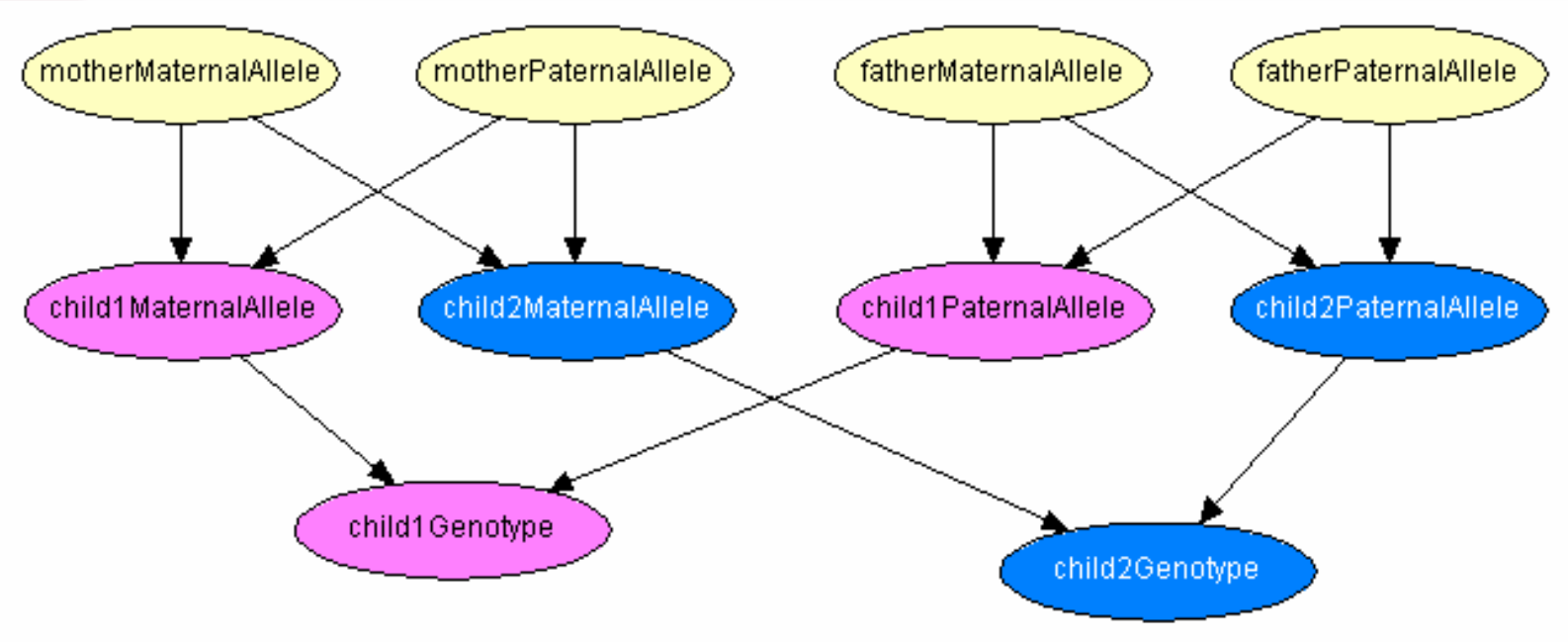
childGenotype

paternalAllele	15				16					
maternalAllele	15	16	17	99	15	16	17	99	15	
15;15	1	0	0	0	0	0	0	0	0	0
15;16	0	1	0	0	1	0	0	0	0	0
15;17	0	0	1	0	0	0	0	0	1	0
15;99	0	0	0	1	0	0	0	0	0	0
16;16	0	0	0	0	0	1	0	0	0	0



FORENSIC SCIENCE SERVICE

A Bayesian network for siblings



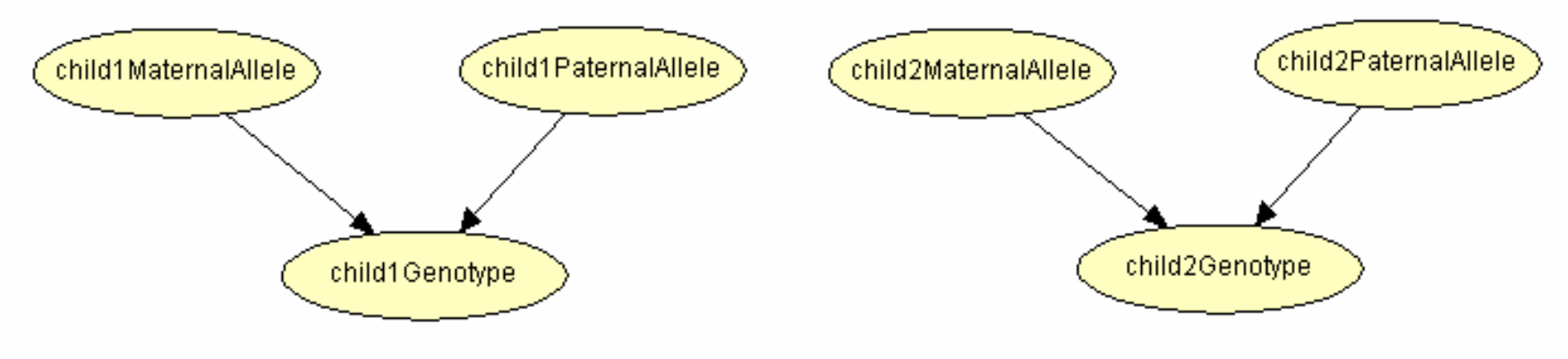
FORENSIC SCIENCE SERVICE

Siblings samples

Mother Maternal Allele	Mother Paternal Allele	Father Maternal Allele	Father Paternal Allele	Child1 Maternal Allele	Child2 Maternal Allele	Child1 Paternal Allele	Child2 Paternal Allele	Child1 Genotype	Child2 Genotype
17	99	17	15	99	17	15	15	15;99	15;17
16	15	17	15	15	15	17	17	15;17	15;17
16	16	99	15	16	16	15	15	15;16	15;16
15	99	17	99	15	15	17	99	15;17	15;99
16	17	16	99	16	17	99	99	16;99	17;99
99	15	15	99	99	99	99	99	99;99	99;99
16	99	15	15	99	16	15	15	15;99	15;16
15	15	15	99	15	15	15	15	99;99	15;15
16	99	17	17	99	99	17	17	17;99	17;99
99	16	15	15	16	16	15	15	15;16	15;16

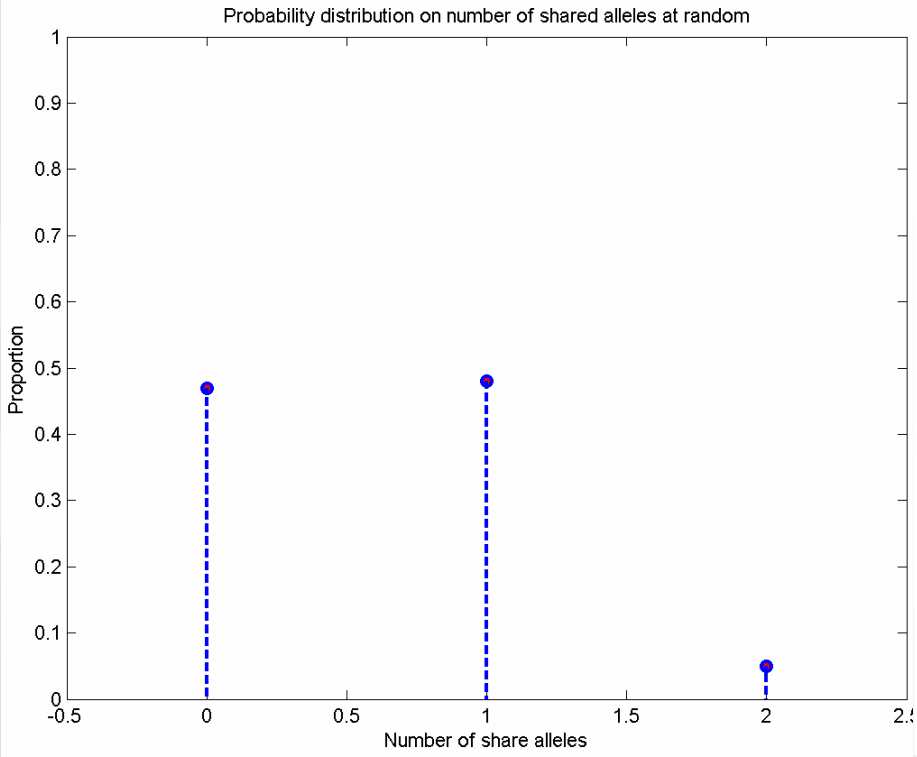
FORENSIC SCIENCE SERVICE

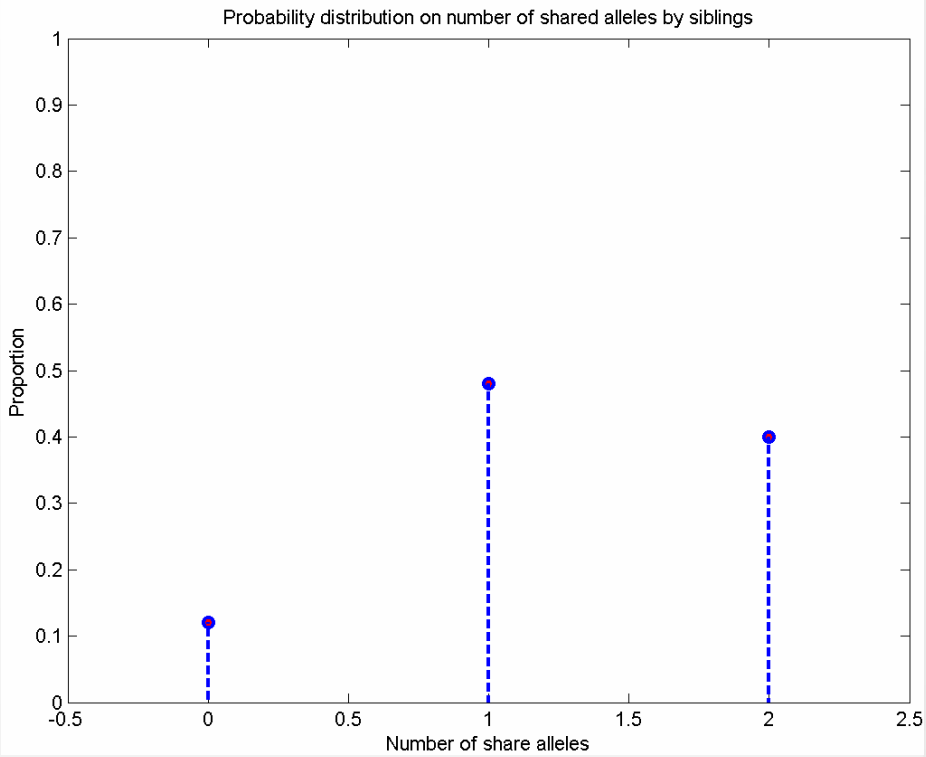
A Bayesian network for unrelated people



FORENSIC SCIENCE SERVICE

Query: proportion on the number of shared alleles





Concluding remarks

- Bayesian networks offer a powerful for modeling pedigrees: **modular**
- Generating samples can be done with existing computer packages
  - Hugin
  - Genie (**free!**)
- The samples so generated aid in the understanding of the behavior of other calculations, such as allele counts or likelihood ratios
- Generate more questions: it is experimentation

FORENSIC SCIENCE SERVICE