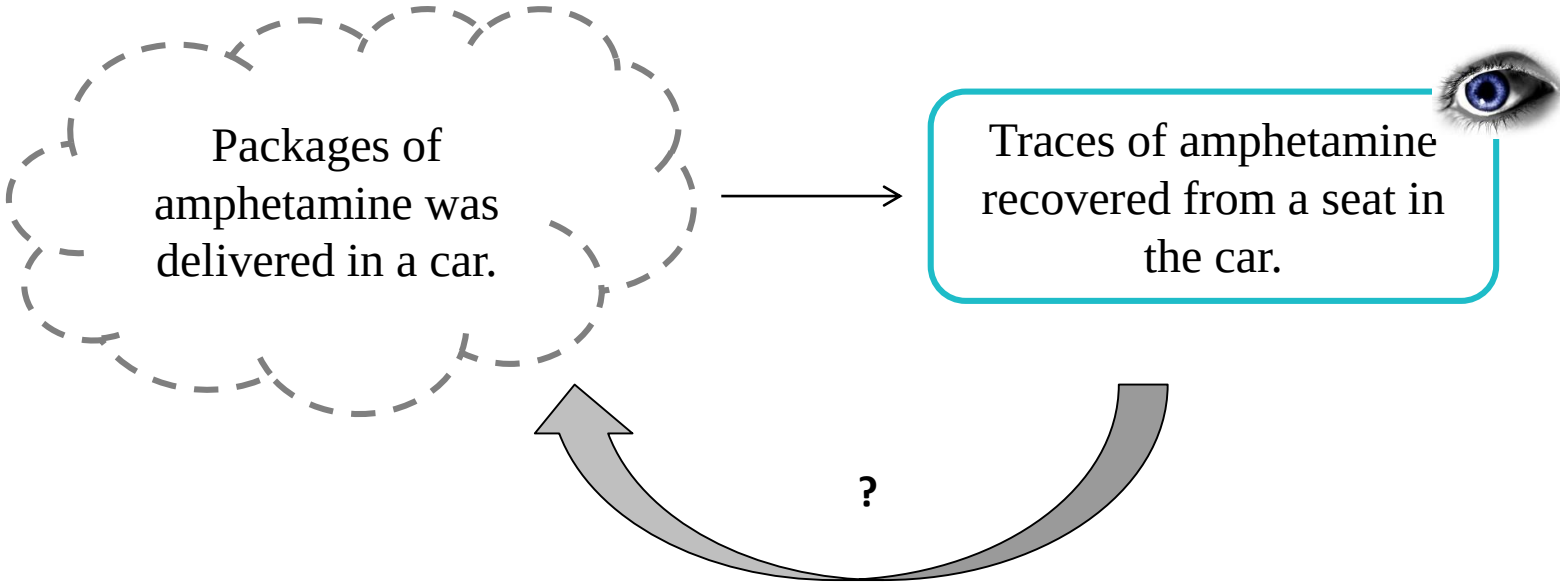
# Meeting 13
# Forensic applications

# Evidence evaluation…

… is an *inductive* inferential process.

”Draw conclusions about what has happened from observed consequences of what happened – what we observe afterwards”



# The two modes of the forensic process

- Investigative mode
- Evaluative mode

Investigative mode:

- Formulation of hypotheses regarding the activities (that have taken place at the crime scene)
- Definition of criteria for subsequent recovery of traces

Evaluative mode:

- Specific questions (formulated hypotheses) about recovered traces and their possible links to suspects or seized goods are treated by use of probabilistic reasoning

…but the two modes come interchangeably in course of the investigation:



# More about the investigative mode

- In the investigative mode *several hypotheses are formulated* about what might have happened at a crime scene, a scene of fire,  a finding-place,…
- Assessment and interpretation of detailed observations lead up to a ranking of the hypotheses formulated
- The assessment is made by deeming how expected detailed observations are under each of the hypotheses formulated. – and falsify such hypotheses under which the observations are considered improbable

This is where the evaluative mode may enter…

- The finally retained hypotheses would form a context that serve as the explanation delivered about what happened at the scene – a kind of giant hypothesis, supported by the investigation, but to be challenged in court

# More about the evaluative mode

- Some questions (hypotheses) cannot be assessed – completely or directly – at the scene (of crime, of fire)
- E.g. hypotheses linking recovered traces to subsequently identified suspects or seized material

  - Was it this shoe that made the recovered footwear mark?
  - Does the blood on the floor originate from the dead person found in the villa?
  - Were the glass fragments recovered from the suspect's jacket transferred from the smashed window at the crime scene?
  - Did somebody burn hazardous waste here?
  - Were the two seized bags of amphetamine parts of the same manufacturing batch?
  - …

Common for these types of questions is that they have the ultimate answers Yes and No.

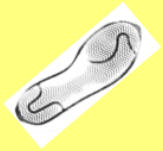
- Conclusions from the evaluative mode may
  - Sort things out at a crime scene/finding-place that helps in deciding along which path the subsequent investigation should continue…

    … should the conclusion be interpreted as a Yes or a No by the CSI
  - Be used as support for an hypothesis about a certain course at the crime scene/finding-place
  - constitute a self-standing piece of forensic evidence that links a recovered material to an individual or another control material, or a specific class/category

Is it then possible to always conclude with a Yes or No?

# Source level attributions

A recovered footwear mark and a pair of shoes seized with a suspect.



Blood recovered from a garment and DNA from swabbing a suspect.

Two seizures of amphetamine – same origin?

# Forensic investigation – and evaluation…

Two seizures of amphetamine – same origin?

Findings:

- Both seizures (materials) have only caffeine as cutting agent
- The dry concentration of amphetamine is about 40 % in both seizures
- The two seizures show similarities in their impurity profiles (presence of small amounts of other substances than amphetamine – bi-products in the manufacturing)

What do these findings signify?

*Same origin with 100 % certainty? 90 % certainty? 50 % certainty?*

Two seizures of amphetamine – same origin?


The question (most probably) put by the commissioner…

> Do the two seizures have a common origin (come from the same batch of manufacturing)?

…can be reformulated into a ***main hypothesis***

> ***$H_m$* :** The two seizures have a common origin

The main hypothesis is a *statement* that constitutes *one* explanation – but not necessarily a good one – to the findings obtained.

- Caffeine as single cutting agent in both
- Similar dry concentrations
- Similarities between impurity profiles

The *weight as evidence* of the statement consists of the belief in the statement – and its relevance for the current alleged activity.

*N.B!* $H_m$ can only be true or false. It is the uncertainty about its truth that is the subject of discussion.

Two seizures of amphetamine – same origin?

# Focusing on the belief (or what would be a proper expression)…

When we cannot categorically state that we are 100% (or 0%) certain about the truth of $H_m$ we must use probability calculus.

Is it then possible to directly estimate the probability that

> ***$H_m$* :** The two seizures have a common origin

is true?

> Answer: No.
>
> This probability is deemed on by ***combining*** the forensic evaluation with other (non-forensic) information from the investigation (supporting or non-supporting $H_m$).

Two seizures of amphetamine – same origin?

Other information

Forensic evaluation

$H_m$ **:** The two seizures have a common origin

Final probability of $H_m$ being true

Other information

**Common platform of evaluation**

Forensic evaluation

***Alternative hypothesis***

e.g. $\boldsymbol{H_a}$ **:** The two seizures have different origins

*Should be chosen to cover all relevant alternatives to the main hypothesis.*

$H_m$ : The two seizures have a common origin

$H_a$ : **The two seizures have different origins**

Forensic evaluation

How expected/probable are…

- Both seizures (materials) have only caffeine as cutting agent
- The dry concentration of amphetamine is about 40 % in both seizures
- The two seizures show similarities in their impurity profiles (presence of small amounts of other substances than amphetamine – bi-products in the manufacturing)

…if ***the main hypothesis*** is true? $\Rightarrow P(\text{Findings}|H_m)$

…if ***the alternative hypothesis*** is true? $\Rightarrow P(\text{Findings}|H_a)$

$$\text{Forensic value of evidence} = V = \frac{P(\text{Findings}|H_m)}{P(\text{Findings}|H_a)}$$

$V > 1 \Rightarrow$ The findings are $V$ times more probable if $H_m$ is true compared to if $H_a$ is true

$V < 1 \Rightarrow$ The findings are $1/V$ times more probable if $H_a$ is true compared to if $H_m$ is true

$H_m$ : The two seizures have a common origin

$H_a$ : **The two seizures have different origins**

Other information

How probable – prior to the forensic investigation – is…

$H_m$ : The two seizures have a common origin ? $\Rightarrow P(H_m)$

…and how probable – prior to the forensic investigation – is …

$H_a$ : The two seizures have different origins ? $\Rightarrow P(H_a)$

$$\text{Prior odds} = O = \frac{P(H_m)}{P(H_a)}$$

$H_m$ : The two seizures have a common origin

$H_a$ : **The two seizures have different origins**

Other information

$$O = \frac{P(H_m)}{P(H_a)}$$

$$V = \frac{P(\text{Findings}|H_m)}{P(\text{Findings}|H_a)}$$

Forensic evaluation

Bayes' theorem :

$$\frac{P(H_m|\text{Findings})}{P(H_a|\text{Findings})} = V \times O$$

*Posterior odds*

$$P(H_m|\text{Findings}) = \frac{V \times O}{V \times O + 1}$$

Final probability of $H_m$ being true – Weight as evidence

$H_m$ : The two seizures have a common origin

**$H_a$ : The two seizures have different origins**

# *Weight as evidence for whom?...*

A source level attribution is generally in Sweden a forensic investigation with the prosecutor as ”destination”.

It is the prosecutor (via the police leader of the preliminary investigation) who (at least in theory) …

- is involved with formulating the alternative hypothesis
- has to deem on the magnitude of the prior odds
- must consider if…

*…the weight as evidence is sufficient to bring this hypothesis as evidence to the indictment*

# Bayes' theorem both in terms of a mathematical formula and as a graphical description

$$\frac{P(H_h)}{P(H_a)} \times \frac{P(E|H_h)}{P(E|H_a)} = \frac{P(H_h|E)}{P(H_a|E)}$$

***Prior odds***
Measures how certain/uncertain the commissioner (prosecutor, police, judge) is about the truth of $H_m$ before considering the outcome of the forensic investigation.

***Forensic value of evidence, V***
States how much more (or less) probable the forensic findings/evidence ***E*** are if $H_m$ is true compared to if $H_a$ is true.

Very often a ***Likelihood ratio***

***Posterior odds***
Measures how certain/uncertain the commissioner is about the truth of $H_m$ upon considering the outcome of the forensic investigation

$\Rightarrow$ ***Weight as evidence***

$$\frac{P(H_h)}{P(H_a)} \times \frac{P(E|H_h)}{P(E|H_a)} = \frac{P(H_h|E)}{P(H_a|E)}$$

Prior odds

Posterior odds

high

1

low

*Forensic value of evidence, V*

high

1

low

Prosecutor

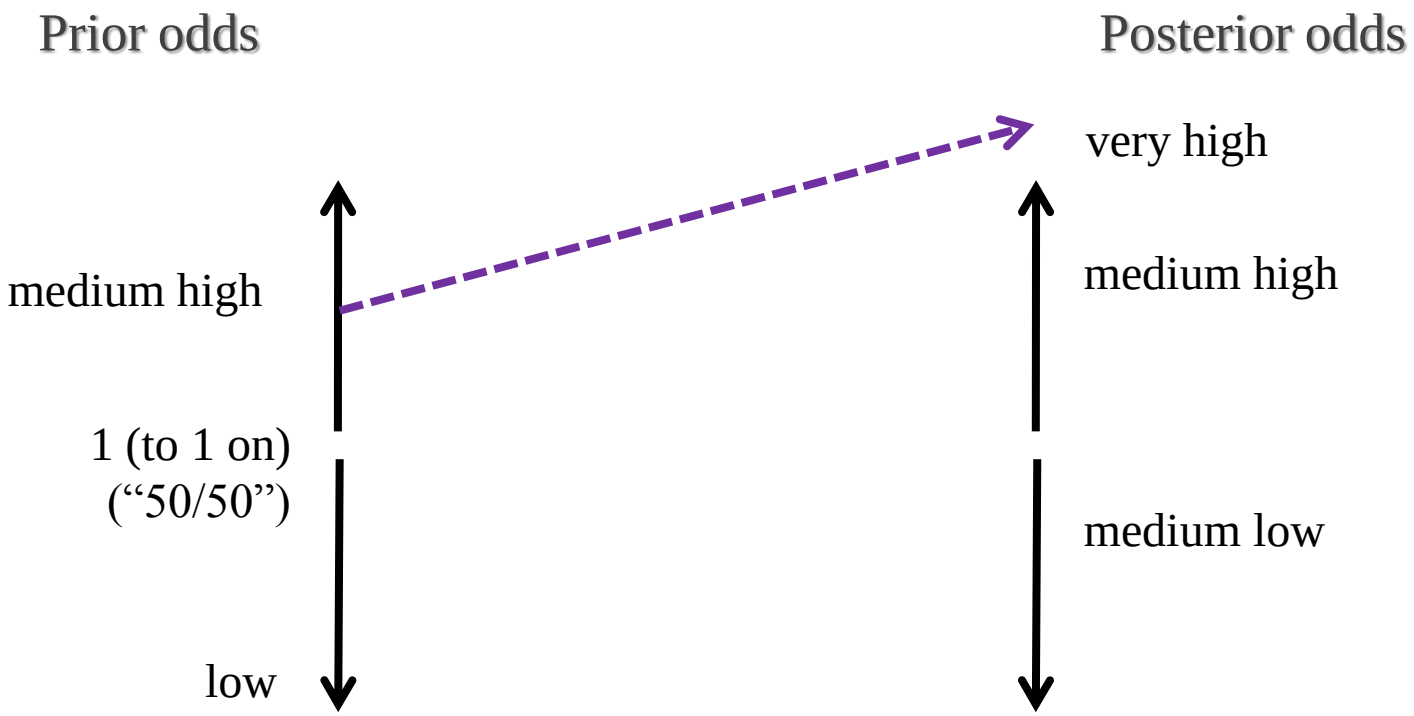
Defence

Prosecutor

Defence

NFC

- With the same pair of main and alternative hypothesis the forensic findings always have the same <u>forensic value of evidence</u> (likelihood ratio) i.e. the arrow has the same angle.
- The <u>weight as evidence</u> (*the posterior odds*) can on the other hand differ depending on the magnitude of the prior odds



# Estimation/calculation – in practice – of the magnitude of the forensic value of evidence (*V*)

In most forensic subject fields today there are no validated mathematical models to support the calculation of the forensic values of evidence.

Lack of background/reference data is the main explanation.

A forensic laboratory should however have uniform standards for reporting their values of evidence.

When models and data are lacking, the components of the value of evidence (i.e. the probabilities in the numerator and denominator of *V*) must be assigned based on (subjective and/or collective) experience and subject knowledge.
$\Rightarrow$ Fairly rough estimates of the magnitudes

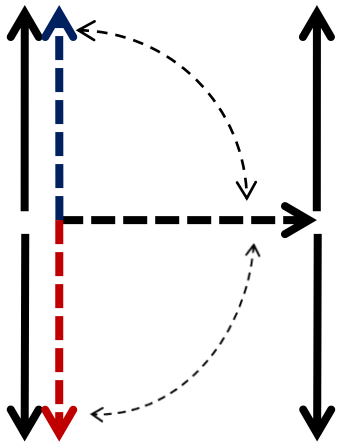
$\Rightarrow$ All reporting of evidence from NFC – with or without using data bases and/or mathematical models – are made using a common ordinal scale of conclusions!

# The scale of conclusion used at NFC:

| Scale level | Magnitude of *V* | "Explanation" |
|---|---|---|
| | | The findings are… |
| +4 | at least one million | …at least one million times more expected… |
| +3 | between 6000 and one million | …at least 6000 times more expected… |
| +2 | between 100 and 6000 | …at least 100 times more expected… |
| +1 | between 6 and 100 | …at least 6 times more expected… |
| 0 | between 1/6 and 6 | … equally expected … |
| | | …if the main hypothesis is true compared to if the alternative hypothesis is true |
| –1 | between 1/100 and 1/6 | …at least 6 times more expected… |
| –2 | between 1/6000 and 1/100 | …at least 100 times more expected… |
| –3 | between 1/(one million) and 1/6000 | …at least 6000 times more expected… |
| –4 | at most 1/(one million) | …at least one million times more expected… |
| | | …if the alternative hypothesis is true compared to if the main hypothesis is true |

# How were the levels of the NFC scale derived?

- The forensic value of evidence is a ratio of two probabilities



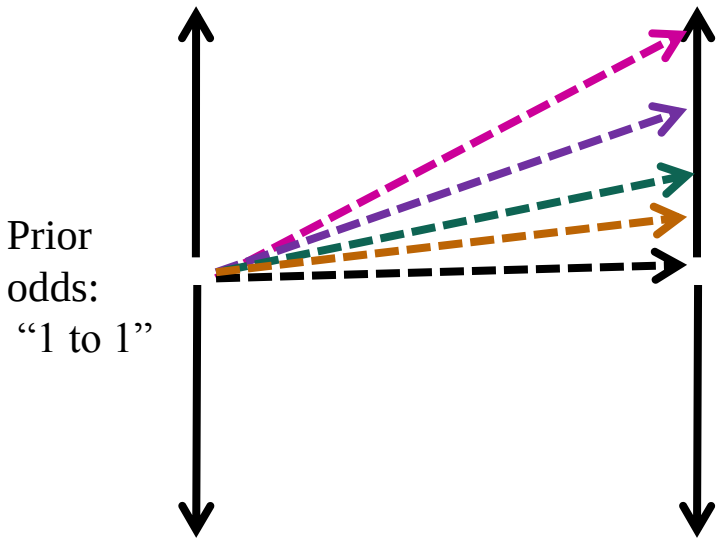
- In theory this value can vary from
  - zero (exclusion of the ***main hypothesis***)

  to
  - infinity (exclusion of the ***alternative hypothesis***)

- …but for a scale to be useful in practice the number of levels must be limited – decomposition of an infinitely long interval into a finite number of intervals

- We chose a set of levels *symmetrically* spread around the value 1, with four supporting levels (positive) and four non-supporting levels (negative), each level corresponding to an interval of values

- The lower limits of the intervals on the supporting side were chosen so that the final weight as evidence (posterior probability) would be acceptably high with respect to each level when the prior odds are equal to one.




| Scale level | Weight-as-evidence $P(H_h \mid E)$ | Lower limit for $V$ |
|---|---|---|
| **+4 :** | **> 0.999999** | **$10^6 \leq V$** |
| **+3:** | **> 0.9998** | **$6000 \leq V$** |
| **+2:** | **> 0.99** | **$100 \leq V$** |
| **+1:** | **> 0.86** | **$6 \leq V$** |
| **0:** | **between 0.14 och 0.86** | **($1/6 < V < 6$)** |

**Level + 2:** Probability 0,99 (= 99%) is generally anticipated among lawyers as sufficiently high to ”confirm” a hypothesis (for bringing a specific evidence to the prosecution)

**Level + 4:** $10^6$ was early chosen as a reasonable limit for the highest level for the evidentiary strength of DNA evidence in Sweden

**Lecels +1 and + 3:** The interval limits have been chosen so that the intervals successively increase in length regularly from a mathematical point of view (close to logarithmic increase). Probabilities 0,9998 and 0,86 became automatic consequences.

Nordgaard A., Ansell R., Drotz W. & Jaeger L.: ”Scale of conclusions for the value of evidence”. *Law, Probability and Risk* 11(1): 1-24.

# Is this something specific for NFC Sweden?

*Paternity index* (Gürtler (1956)) – Early introduction of the *Likelihood Ratio*

Lindley (1977) ”A problem in forensic science”

**Forensic Science Service**, England and Wales: ”Case Assesment and Interpretation”, 1990s

**Ecole des Sciences Criminelles,** Université de Lausanne, CH

**LGC Forensics Ltd, England**

**ESR, New Zealand**

**NFI, The Netherlands**

**Guardia Civil, Spain**

**NFC (SKL), Sweden**

**INCC, Belgium**

**IES, Krakow, Poland**

**EFE, Ireland**

# Eliciting probabilities in the amphetamine seizures case

Two seizures of amphetamine – same origin?

$H_m$ : The two seizures have a common origin

**$H_a$ : The two seizures have different origins**

Conditional probabilities of the findings assuming $H_m$ is true

**$E_1$:** Both seizures (materials) have only caffeine as cutting agent
**$E_2$:** The dry concentration of amphetamine is about 40 % in both seizures
**$E_3$:** The two seizures show similarities in their impurity profiles (presence of small amounts of other substances than amphetamine – bi-products in the manufacturing)

All findings are *consistent* with $H_m$ which means the conditional probability of obtaining them if $H_m$ is true should be close to one.

Reasons for the probability not to equal one may be

- findings are expected to be even more consistent (e.g. exactly the same dry concentration in both seizures)
- more is expected than what has been obtained as findings (e.g.

$H_m$ : The two seizures have a common origin

$H_a$ : **The two seizures have different origins**

Conditional probabilities of the findings assuming $H_a$ is true

If we assume the two seizures consist of amphetamine from different batches of manufacturing generally …

… how probable do we deem …

$E_1$: Both seizures (materials) have only caffeine as cutting agent
$E_2$: The dry concentration of amphetamine is about 40 % in both seizures
$E_3$: The two seizures show similarities in their impurity profiles (presence of small amounts of other substances than amphetamine – bi-products in the manufacturing) ?

---

$E_1$: Both seizures (materials) have only caffeine as cutting agent?
$E_2$: The dry concentration of amphetamine is about 40 % in both seizures?

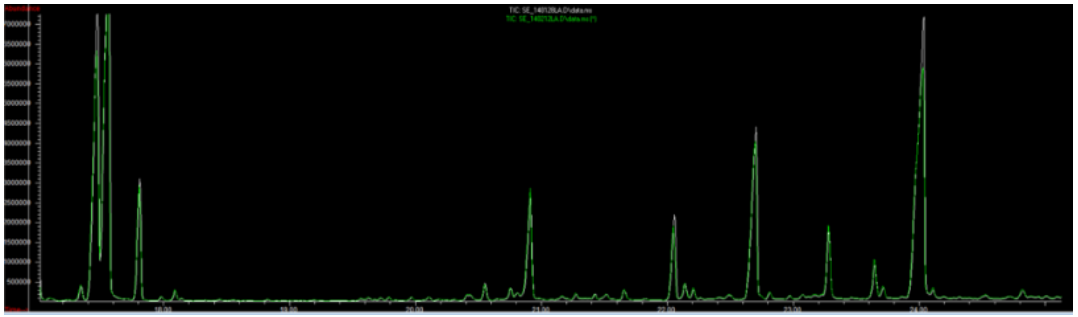
*Use information from historical cases with seizures of amphetamine*

- How often is caffeine the single cutting agent in amphetamine powder material? ⇒ guides the assignment of $P(E_1 \mid H_a)$
- How common is 40% dry concentration? ⇒ guides the assignment of $P(E_2 \mid H_a)$

$H_m$ : The two seizures have a common origin

**$H_a$ : The two seizures have different origins**

***$E_3$*:** The two seizures show similarities in their impurity profiles (presence of small amounts of other substances than amphetamine – bi-products in the manufacturing) ?



*More complex!*

*Experience and knowledge based (subjective) assignment*

- The profiles (set of peak areas) must first be interpreted one at a time
- The peak areas are continuous-valued – a set has a multivariate continuous probability distribution
- It is expected that there are slight discrepancies between the profiles (due to their continuous nature)
- Empirical data must be available from which it should be possible to study the variation among seizures with a common origin (within-variation) [$H_m$ true] and the variation among seizures with different origins (between-variation) [$H_a$ true]

# DNA evidence

- Was “discovered” as a useful tool in forensic investigation of bio-traces in the 1980:s (*Sir Alec Jeffreys*)
- Has undergone an enormous development since then.
  - It started with a few bio-markers that in clear cases could lead to likelihood ratios of magnitude  10 000  (very large at that time)
  - Today, most “kits” in use can in clear cases lead to likelihood ratios of magnitude $10^{20}$
- DNA analysis is used both for crime forensic purposes (comparisons between DNA from a recovered trace with unknown origin and DNA from a suspect/from another recovered trace); and for non-crime purposes (paternity disputes, disaster victim identification (DVI), kinship investigations)
- Today’s technology makes it possible to extract information from very small amounts of DNA (picograms)
  - Challenges for the deeming of whether the DNA recovered is there for innocent purposes or not
  - Many bio-traces would show DNA from several people (mixtures)

# The DNA Double Helix

Consists of so-called nucleobases: *adenine* (A), *thymine* (T), *cytosine* (C) and *guanine* (G) always in the pairs A-T, C-G.

In so-called diploid organisms (like humans) the genetic information comes in 23 *chromosome* pairs, where each chromosome is a double helix. This is referred to as the *genome* of a human being.

Along one chromosome sequences of nucleobase pairs defines so-called *markers* or *loci* (one locus). Such sequences can consist of one up to hundreds of nucleobases. There is a corresponding sequence of nucleobase pairs on the other chromosome but not necessarily of the same length.

The two sequences are called *alleles* and together they form the *genotype* of that locus.

One of the alleles is inherited from the mother and the other from the father, but for most loci it is not possible to know which is what.

Each chromosome pair hosts a great variety of genotypes (or genes). One of the chromosome pairs defines the sex. The others are referred to as autosomes (autosomal DNA).

Most of the genome (more than 90%) is today not verified to have any other function than possibly "assisting" (*non-coding DNA*).

Forensic DNA analysis is concerned with the non-coding part of the DNA.

Several techniques are used to read off the information contained in the genome:

- PCR (investigates so-called short tandem repeats (STR) or single nucleotide polymorphisms (SNP) or other polymorphisms) by multiplying extracted molecules
- Sequencing (todays technique for microorganisms, possibly tomorrow's for humans) strives at projecting the whole genome

# PCR for STR (still a consensus technique among European forensic science institutes (ENFSI) for crime investigation)

A number of so-called *kits* are available. At NFC, Linköping the kit ESX-16 is used:

16 loci on different chromosomes are investigated, 15 autosomal and one sex chromosome locus (used to identify the sex of the human being the source of the DNA).

An example from typing (identifying the genotypes in each autosomal locus):

| Locus | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *Allele* 1 | 15 | 7 | 27 | 14 | 16 | 17.3 | 18 | 11 | 15 | 15 | 12 | 21 | 11 | 17.3 | 15 |
| *Allele* 2 | 15 | 9.3 | 29 | 16 | 16 | 18.3 | 25 | 12 | 16 | 19 | 13 | 22.2 | 11 | 19 | 16 |

The allele codes are  simply number of repeats of a certain sequence.

A complete set of 15 genotypes is referred to as a *DNA profile*.

# DNA comparisons

In a criminal case there is a recovered trace from a crime scene:

- blood stain
- saliva stain
- semen stain
- hairs or body tissues
- vaginal samples
- …

When there is a suspect, ordinary samples can be taken (today buccal swabs are standard, previously blood samples were taken) to recover DNA.

Typed DNA profiles are compared:

➔ match or no match

# How to evaluate a match?

## Population genetics models:

*Hardy-Weinberg equilibrium:*

In a population with random mating the proportion of a certain genotype at a locus can be calculated from the proportions of the alleles defining the genotype:

If the two alleles are the same (homozygote) with allele proportion $p_A$ the genotype proportion is

$$p_{A,A} = p_A^2$$

If the two alleles are different (heterozygote) with allele proportions $p_A$ and $p_B$ the genotype proportion is

$$p_{A,B} = 2 \cdot p_A \cdot p_B$$

Many national populations almost satisfies Hardy-Weinberg (HW) equlibrium (at least such an hypothesis is hard to reject on basis of collected data)

Adjustment (Balding & Nichols, 1995) to take into account so-called *subpopulation effects* (meaning that mating is not random but alleles are structurally inherited along "lines" in the population):

$$p_{A,A} = \frac{\left(2 \cdot F_{ST} + \left(1 - F_{ST}\right) \cdot p_A\right) \cdot \left(3 \cdot F_{ST} + \left(1 - F_{ST}\right) \cdot p_A\right)}{\left(1 + F_{ST}\right) \cdot \left(1 + 2 \cdot F_{ST}\right)}$$

$$p_{A,B} = \frac{2 \cdot \left(F_{ST} + \left(1 - F_{ST}\right) \cdot p_A\right) \cdot \left(F_{ST} + \left(1 - F_{ST}\right) \cdot p_B\right)}{\left(1 + F_{ST}\right) \cdot \left(1 + 2 \cdot F_{ST}\right)}$$

where $F_{ST}$ is the *co-ancestry coefficient* measuring the subpopulation effects (to what extent the mating is non-random).

In Sweden $F_{ST}$ is close to 0.01.

***Example***

A study was made in a population where the coancestry coefficient is estimated to be around 3 % . The following results were obtained for  for locus TH01:

| Allele | Relative frequency |
|---|---|
| 6 | 0.295 |
| 7 | 0.147 |
| 8 | 0.184 |
| 9 | 0.232 |
| 9.3 | 0.026 |
| 10 | 0.116 |

What are the relative frequencies for the genotypes (7,8) and (8,8) respectively assuming

- Hardy-Weinberg equilibrium ?
- subpopulation effects  ?

H-W:

$$p_{7,8} = 2 \cdot 0.147 \cdot 0.184 \approx 0.054$$

$$p_{8,8} = 0.184^2 \approx 0.034$$

| Allele | Relative frequency |
|---|---|
| 6 | 0.295 |
| 7 | 0.147 |
| 8 | 0.184 |
| 9 | 0.232 |
| 9.3 | 0.026 |
| 10 | 0.116 |

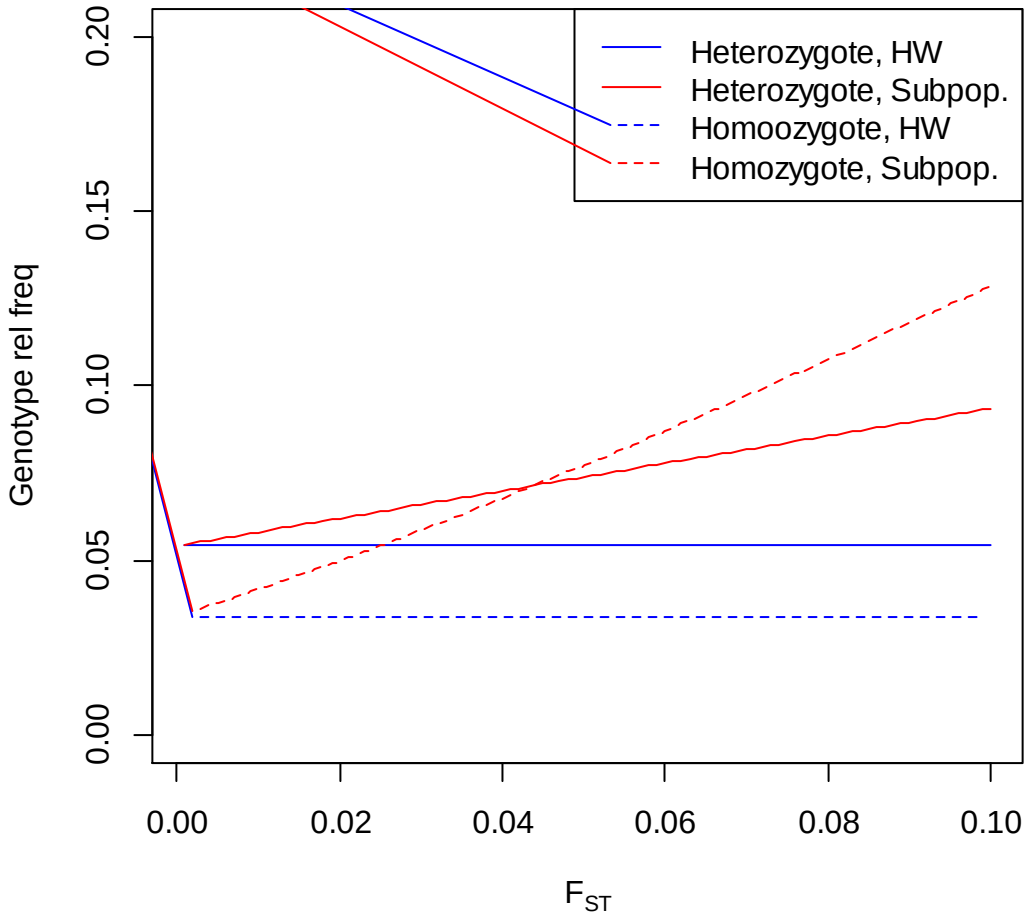
Subpopulation effects:

$$p_{7,8} = \frac{2 \cdot \left(0.03 + \left(1 - 0.03\right) \cdot 0.147\right) \cdot \left(0.03 + \left(1 - 0.03\right) \cdot 0.184\right)}{\left(1 + 0.03\right) \cdot \left(1 + 2 \cdot 0.03\right)} \approx 0.066$$

$$p_{8,8} = \frac{\left(2 \cdot 0.03 + \left(1 - 0.03\right) \cdot 0.184\right) \cdot \left(3 \cdot 0.03 + \left(1 - 0.03\right) \cdot 0.184\right)}{\left(1 + 0.03\right) \cdot \left(1 + 2 \cdot 0.03\right)} \approx 0.059$$

Assuming subpopulation effects always gives higher genotype relative frequencies, but the discrepancy is larger for homozygote genotypes.

With the current genotypes of this locus:




*Linkage equilibrium*:

Genotypes at different loci very often become less statistical dependent with the distance them between in the double helix.

This is in particular true for several loci in the non-coding area that are situated on different chromosomes and for which we may with good approximation assume complete statistical independence.

If two loci are on the same chromosome their genotypes may still be independent if the distance them between is long enough.

The loci chosen in forensic kits for typing short tandem repeats (STR) loci satisfy the assumption of (approximate) independence and are said to be in *linkage equilibrium* (LE).

A small set of single nucleotide polymorphisms (SNPs) may be in LE, but the benefit of using SNPs is that thousands can be analysed in one run. These do not satisfy LE.

With linkage equilibrium the relative frequency of a DNA profile can be calculated from the genotype relative frequencies:

$$p_{\text{profile}} = p_{A_1,B_1} \cdot p_{A_2,B_2} \cdot \ldots \cdot p_{A_L,B_L}$$

where the number of loci in the kit is $L$ and ($A_i$ , $B_i$ ) is the genotype of locus $i$ (with the possibility that $A_i = B_i$ , i.e. homozygote genotype)

Linkage equilibrium implies that a profile relative frequency at a very fast rate goes towards zero when the number of loci used increases.

With a full profile in 15 loci typical relative frequencies are of magnitude less than $10^{-14}$.

*Are these actually to be considered as relative frequencies?*

***Example***

Consider the previously shown profile:

| Locus | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *Allele* 1 | 15 | 7 | 27 | 14 | 16 | 17.3 | 18 | 11 | 15 | 15 | 12 | 21 | 11 | 17.3 | 15 |
| *Allele* 2 | 15 | 9.3 | 29 | 16 | 16 | 18.3 | 25 | 12 | 16 | 19 | 13 | 22.2 | 11 | 19 | 16 |
| $p_{A,B}$ | 0.085 | 0.140 | 0.016 | 0.051 | 0.020 | 0.028 | 0.026 | 0.192 | 0.254 | 0.017 | 0.099 | 0.011 | 0.152 | 0.008 | 0.026 |

The genotype relative frequencies have been calculated using allele relative frequencies obtained from a database from a average modern Swedish population and assuming subpopulation effects with $F_{ST} = 0.01$

The relative frequency of this profile is calculated to $4 \cdot 10^{-21}$

With a population of almost 10 million inhabitants this cannot be a profile belonging to that population if the value is to be taken for a true relative frequency.

# The evaluation model used in a criminal case

Assume there is a stain left at a crime scene and there is a male suspect assumed to have been involved with the criminal activity. DNA is recovered from the stain and from a buccal swab of the suspect.

A full profile is obtained from the stain (sometimes it is not possible to *type* all loci due to degraded DNA) and (as expected) a full profile is obtained from the suspect.

The two profiles match in every locus.

Assume it is the profile previously discussed

| Locus | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *Allele* 1 | 15 | 7 | 27 | 14 | 16 | 17.3 | 18 | 11 | 15 | 15 | 12 | 21 | 11 | 17.3 | 15 |
| *Allele* 2 | 15 | 9.3 | 29 | 16 | 16 | 18.3 | 25 | 12 | 16 | 19 | 13 | 22.2 | 11 | 19 | 16 |

Hypotheses:

$\boldsymbol{H_m}$ : “The suspect is the donor of the stain”

$\boldsymbol{H_a}$ : “Someone else is the donor of the stain”

Evidence:

***E*** : “A match in DNA profile (matches in all 15 autosomal loci of an ESX16-profile and match in the sex-defining locus ) “

Value of evidence (likelihood ratio):

$$V = \frac{P\left(E \middle| H_m\right)}{P\left(E \middle| H_a\right)}$$

How to find (estimates of) the numerator and the denominator?

$P\left(E\middle|H_m\right)$

If the suspect actually left the stain we expect to obtain matches in all loci.

There is no genetic reason for any variation (besides mutations, but such interventions can usually be controlled).

There could be variation due to deficiencies with the equipment or with the operators (reading off the wrong values).

However it is generally non-debatable to set this probability to 1.

$P(E|H_a)$

If someone else left the stain, what is the probability of obtaining the match?

Sometimes things become clearer if we formulate the evidence in terms of the variables

$\boldsymbol{E_c}$ : DNA profile of crime stain

$\boldsymbol{E_s}$ : DNA profile of suspect

The evidence can then be written

$$E = (E_c = \Gamma, E_s = \Gamma)$$

where $\Gamma$ is the profile obtained both with the stain and the suspect.

$$\Rightarrow$$

$$V = \frac{P\left(E \middle| H_m\right)}{P\left(E \middle| H_a\right)} = \frac{P\left(E_c = \Gamma, E_s = \Gamma \middle| H_m\right)}{P\left(E_c = \Gamma, E_s = \Gamma \middle| H_a\right)} = \frac{P\left(E_c = \Gamma \middle| E_s = \Gamma, H_m\right) \cdot P\left(E_s = \Gamma \middle| H_m\right)}{P\left(E_c = \Gamma \middle| E_s = \Gamma, H_a\right) \cdot P\left(E_s = \Gamma \middle| H_a\right)} =$$

$$= \left\langle \begin{matrix} \text{Suspect's profile (isolated) does} \\ \text{not depend on the hypotheses} \end{matrix} \right\rangle = \frac{P\left(E_c = \Gamma \middle| E_s = \Gamma, H_m\right) \cdot P\left(E_s = \Gamma\right)}{P\left(E_c = \Gamma \middle| E_s = \Gamma, H_a\right) \cdot P\left(E_s = \Gamma\right)} =$$

$$\left\langle \begin{matrix} \text{If someone else left the stain, the suspect's profile} \\ \text{cannot have any impact on the profile of the stain} \end{matrix} \right\rangle =$$

$$= \frac{P\left(E_c = \Gamma \middle| E_s = \Gamma, H_m\right)}{P\left(E_c = \Gamma \middle| H_a\right)}$$

Now, the denominator is the probability of obtaining the profile $\Gamma$ of the stain if the stain was left by someone else than the suspect.

This probability should account for the rarity of this profile in the population of potential donors of the stain.

$$P\left(E_c = \Gamma \middle| H_a\right)$$

Is this probability higher for certain groups of the population of potential donors (i.e. is the population stratified with respect to the occurrence of this profile)?

*Note!* Since the stain is from a male (due to the match) the population only consists of males.

What about

- an identical twin of the suspect?
- a full brother of the suspect?
- the suspect's father?
- a son of the suspect?
- a half-brother of the suspect?
- the grand-fathers of the suspect?
- an uncle or a male cousin of the suspect?

If stratification should be taken into account we need to use a so-called full Bayesian approach and compute the value of evidence as the Bayes factor

$$B = \frac{P\left(E_c = \Gamma \middle| E_s = \Gamma, H_m\right)}{\sum_{\substack{\text{all individuals in} \\ \text{the population, except} \\ \text{the suspect}}} P\left(E_c = \Gamma \middle| \text{Individual } i \text{ is the donor}\right) \cdot P\left(\text{Individual } i \text{ is the donor} \middle| H_a\right)}$$

However, this will need knowledge about the prior probabilities

$$P\left(\text{Individual } i \text{ is the donor} \middle| H_a\right)$$

of which the forensic scientist has no opinion (and should not have).

Hence, the evidentiary strength cannot be assessed without prior opinions about which persons could have been involved.

To be able to report measures of evidentiary strength we need to formulate different alternative hypotheses.

First choice: $H_a$ : "Someone else, not closely related to the suspect, left the stain"

$$V = \frac{P\left(E_c = \Gamma \middle| E_s = \Gamma, H_m\right)}{P\left(E_c = \Gamma \middle| H_a\right)}$$

The denominator can now be estimated from a random sample of individuals from the population to which the donor is assumed to belong.

Such a random sample is (today) a kind of panel, i.e. a number of persons from a general population (covering the population of potential donors with negligible effects of over coverage)

➔ DNA population database

Hence, $P\left(E_c = \Gamma \middle| H_a\right)$ is estimated by calculating the relative frequency of this profile using the database.

Less problematic that this relative frequency is not possible to physically obtain in the population, it is used to estimate a probability through a *model* of the population.

For the current profile we previously obtained a calculated relative frequency of $4 \cdot 10^{-21}$.

$$V = \frac{P\left(E_c = \Gamma \middle| E_s = \Gamma, H_m\right)}{P\left(E_c = \Gamma \middle| H_a\right)} = \frac{1}{4 \cdot 10^{-21}} = 2.5 \cdot 10^{20}$$

The match is thus $2.5 \cdot 10^{20}$ times more probable to obtain if the suspect is the donor than if someone else, not closely related to the suspect, is the donor

*Was it him?*

Another alternative hypothesis may be

$H_{a,2}$ : “The stain was left by a full brother of the suspect”

We then need more population genetics to calculate the probability

$$P\left(E_c = \Gamma \middle| H_{a,2}\right)$$

For the current profile an estimate of this probability becomes $1.82 \cdot 10^{-7}$

Hence, the value of evidence is

$$V^{(2)} = \frac{P\left(E_c = \Gamma \middle| E_s = \Gamma, H_m\right)}{P\left(E_c = \Gamma \middle| H_{a,2}\right)} = \frac{1}{1.82 \cdot 10^{-07}} = 5.5 \cdot 10^6$$

The match is thus 5.5 million times more probable to obtain if the suspect is the donor than if a full brother of the suspect is the donor.

Besides identical twins, full siblings of the same sex are the closest related individuals.

Changing the alternative hypothesis to something like

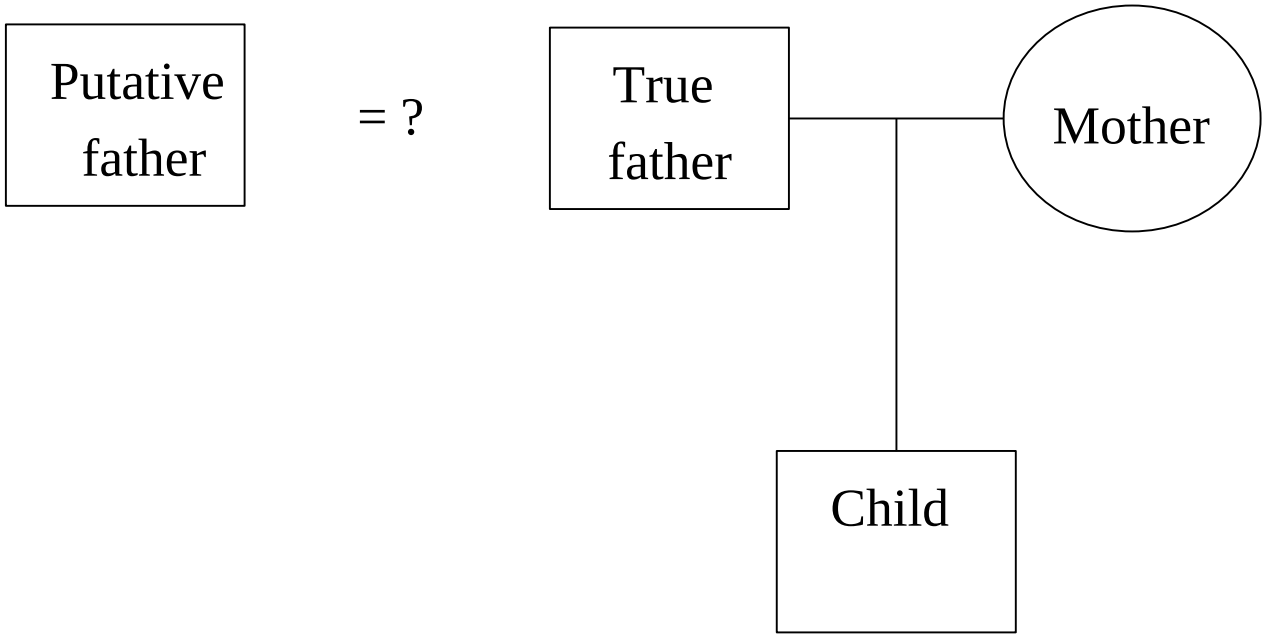
“The donor of the stain is a father or a son of the suspect”

will also render a higher relative frequency (however lower than with a full brother) and as a consequence a lower value of evidence (against the suspect).

It has become more and more common for a suspect to “blame the brother”. The most obvious way to handle this situation is to swab the brother.

- A mismatch directly excludes the brother.
- With a match we’ll have to stick to the lower value of evidence.

# Paternity testing




A simple so-called *pedigree*:

Is the putative (alleged) father the true father of a child?

Let for a specific locus

- *cgt* denote the child's genotype
- *mgt* denote the mother's genotype
- *tfgt* denote the true father's genotype
- *pfgt* denote the putative father's genotype

Note that we assume *cgt*, *mgt* and *pfgt* to be observed, while *tfgt* is unknown.

Hypotheses:

$$H_m : pfgt = tfgt$$

$$H_a : pfgt \neq tfgt$$

The *evidence* consists of the child's genotype (there is no "match" since the child can only have inherited one allele from the mother)

The likelihood ratio is

$$LR = \frac{P(cgt \mid mgt, pfgt, H_m)}{P(cgt \mid mgt, pfgt, H_a)} = \frac{P(cgt \mid mgt, pfgt, H_m)}{P(cgt \mid mgt, H_a)}$$

Now, let

- $A_{mo}$ denote the allele of *cgt* that was inherited from the mother
- $A_{tf}$ denote the allele of *cgt* that was inherited from the (true) father

➔ The numerator of *LR* can be written

$$P(cgt \mid mgt, pfgt, H_m) = P(A_{mo}, A_{tf} \mid mgt, pfgt, H_m) =$$
$$= P(A_{tf} \mid A_{mo}, mgt, pfgt, H_m) \cdot P(A_{mo} \mid mgt, pfgt, H_m) =$$
$$= P(A_{tf} \mid A_{mo}, pfgt, H_m) \cdot P(A_{mo} \mid mgt)$$

Homozygote case, *cgt* = ( *i* , *i* ) :

Numerator is $P\left(A_{mo} = i \mid mgt\right) \cdot P\left(A_{tf} = i \mid pfgt, H_m\right)$

i.e.. $A_{tf}$ is independent of $A_{mo}$

$$P\left(A_{mo} = i \mid mgt = (i,i)\right) = 1$$

$$P\left(A_{mo} = i \mid mgt = (i,j)\right) = \frac{1}{2}$$

$$P\left(A_{tf} = i \mid pfgt = (i,i), H_m\right) = 1$$

$$P\left(A_{tf} = i \mid pfgt = (i,j), H_m\right) = \frac{1}{2}$$

Hence, upon multiplication the numerator is one of the values 1, 0.5 or 0.25.

Heterozygote case, *cgt* = ( *i* , *j* ) :

The general form of the numerator becomes

$$P\left(A_{mo} = i \mid mgt\right) \cdot P\left(A_{tf} = j \mid pfgt, H_m\right) +$$
$$+ P\left(A_{mo} = j \mid mgt\right) \cdot P\left(A_{tf} = i \mid pfgt, H_m\right)$$

since the two alleles must be different.

One of the two probabilities may be = 0 depending on whether the mother or the putative father is heterozygote with “the other allele” being different from *i* and *j* .

The denominator:

$$P(cgt \mid mgt, H_a) = P(A_{mo}, A_{tf} \mid mgt, H_a) =$$
$$= P(A_{tf} \mid A_{mo}, mgt, H_a) \cdot P(A_{mo} \mid mgt, H_a) =$$
$$= P(A_{tf} \mid A_{mo}, H_a) \cdot P(A_{mo} \mid mgt)$$

Homozygote case, *cgt* = ( *i* , *i* ):

$$P(cgt \mid mgt, H_a) = P(A_{tf} = i \mid H_a) \cdot P(A_{mo} = i \mid mgt) =$$
$$= p_i \cdot P(A_{mo} = i \mid mgt)$$

where $p_i$ is the probability (relative frequency) of allele *i* in the population of potential fathers and $P(A_{mo} = i \mid mgt)$ can be developed as for the numerator.

Heterozygote case, $cgt = (\, i \,, j \,)$ :

$$P(cgt \mid mgt, H_a) = P(A_{tf} = i \mid H_a) \cdot P(A_{mo} = j \mid mgt) +$$
$$+ P(A_{tf} = j \mid H_a) \cdot P(A_{mo} = i \mid mgt) =$$
$$= p_i \cdot P(A_{mo} = j \mid mgt) + p_j \cdot P(A_{mo} = i \mid mgt)$$

Again, one of these probabilities may be = 0 if the mother is heterozygote with "the other allele" different from *i* and *j*.

For each locus a likelihood ratio for the hypothesis $H_m$ (vs. $H_a$ ) can be calculated.

If linkage equlibrium can be assumed, the *LR*s for all *L* loci of the kit used can be multiplied to form what is historically referred to as the *paternity index:*

$$PI = LR_1 \cdot LR_2 \cdot \cdots \cdot LR_L$$

but note that an explicit formula would be even more meaningless here than it is for the LR of a match in a crime case.

***Example***

Return to the study where he following results were obtained for  for locus TH01:

| Allele | Relative frequency |
|---|---|
| 6 | 0.295 |
| 7 | 0.147 |
| 8 | 0.184 |
| 9 | 0.232 |
| 9.3 | 0.026 |
| 10 | 0.116 |

Suppose in a disputed paternity case that the child has genotype (8,9), the mother has genotype (8,10) and the putative father has genotype (6,9).

What is factor of this locus in the paternity index?

$$LR_{\mathrm{TH01}} = \frac{P(cgt \mid mgt, pfgt, H_m)}{P(cgt \mid mgt, H_a)}$$

*cgt* = (8,9)
*mgt* = (8,10)
*pfgt* = (6,9)

The child is heterozygote (8,9)

➔ The numerator becomes

$$P(cgt \mid mgt, pfgt, H_m) =$$
$$P(A_{mo} = 8 \mid mgt) \cdot P(A_{tf} = 9 \mid pfgt, H_m) +$$
$$+ P(A_{mo} = 9 \mid mgt) \cdot P(A_{tf} = 8 \mid pfgt, H_m) = \frac{1}{2} \cdot \frac{1}{2} + 0 \cdot 0 = 0.25$$

…and the denominator becomes

$$P(cgt \mid mgt, H_a) = p_i \cdot P(A_{mo} = j \mid mgt) + p_j \cdot P(A_{mo} = i \mid mgt) =$$
$$= 0.184 \cdot 0 + 0.232 \cdot \frac{1}{2} = 0.116$$

$$\Rightarrow LR_{\mathrm{TH01}} = \frac{0.25}{0.116} \approx 2.155$$