Forensic DNA Inference

ICFIS 2008
Lausanne, Switzerland

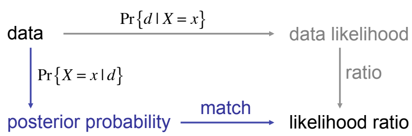
Mark W Perlin, PhD, MD, PhD
Joseph B Kadane, PhD
Robin W Cotton, PhD

Likelihood Ratio

quantitative rarity statistic in forensic science

data likelihood ratio (DLR)

match likelihood ratio (MLR)



Match

item	questioned evidence	suspect
data	d_Q	d_S
type	Q	S

$$M(Q,S):\ for\ some\ value\ x \in X$$
$$Q = x\ \&\ S = x$$

Match Probability

$$\text{for uncertain type } Q, \Pr\{Q = x\} = q(x)$$
$$\text{for uncertain type } S, \Pr\{S = x\} = s(x)$$

$$\Pr\{M(Q,S)\} = \sum_{x \in X} \Pr\{Q = x \;\&\; S = x\}$$ (disjoint values)

$$= \sum_{x \in X} \Pr\{Q = x\} \cdot \Pr\{S = x\}$$ (assume types independent)

$$= \sum_{x \in X} q(x) \cdot s(x)$$

Inner Product

$$\Pr\{M(Q,S)\} = \sum_{x \in X} q(x) \cdot s(x)$$

- standard pattern comparison method
- widely used, powerful math properties

$$\Pr\{M(Q,S)\} = \int_X q(x) \cdot s(x)\,dx$$

$$\Pr\{M(Q,S)\} = \int_X q(x) \cdot s(x) \cdot \mu(x)\,dx$$

Match Rarity

$$\text{for random type } R, \Pr\{R = x\} = r(x)$$

Define the Match Likelihood Ratio (MLR) statistic

$$MLR \equiv \frac{\Pr\{M(Q,S)\}}{\Pr\{M(Q,R)\}}$$

How to compute the MLR

$$MLR = \frac{\Pr\{M(Q,S)\}}{\Pr\{M(Q,R)\}} = \frac{\sum_{x \in X} q(x) \cdot s(x)}{\sum_{x \in X} q(x) \cdot r(x)}$$

MLR is a likelihood ratio

Hypothesis C: suspect S contributed to evidence item Q

$$\Pr\{M \mid C\} = \Pr\{M(Q,S)\}$$

$$\Pr\{M \mid \bar{C}\} = \Pr\{M(Q,R)\}$$

$$MLR = \frac{\Pr\{M(Q,S)\}}{\Pr\{M(Q,R)\}} = \frac{\Pr\{M \mid C\}}{\Pr\{M \mid \bar{C}\}}$$

MLR assesses match observation M under alternative contributor hypotheses C and C'

How to report a MLR

$$MLR = \frac{\Pr\{M(Q,S)\}}{\Pr\{M(Q,R)\}}$$

The probability of a match between
evidence type Q and suspect type S
is (some number) times more likely than
the probability of a match between
evidence type Q and a random type R.

Type Probability

posterior probability

$$q(x) = \Pr\{Q = x \mid d_Q\}$$
$$\propto \Pr\{d_Q \mid Q = x\} \cdot \Pr\{Q = x\}$$

likelihood function × prior probability

Single Source DNA

likelihood $\quad \Pr\{d_Q \mid Q = x\} = \begin{cases} 1 & x = x_0 \\ 0 & otherwise \end{cases}$

prior $\quad \Pr\{Q = x\} \propto 1$

posterior $\quad q(x) \propto \Pr\{d_Q \mid Q = x\} \cdot \Pr\{Q = x\}$

$$MLR = \frac{\sum_{x \in X} q(x) \cdot s(x)}{\sum_{x \in X} q(x) \cdot r(x)} = \frac{1}{r(x_0)}$$

DNA Mixture: Inclusion

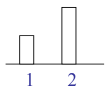
$$\Pr\{d_Q \mid Q = x\} = \begin{cases} 1 & alleles(x) \subset alleles(d_Q) \\ 0 & otherwise \end{cases}$$

$$\Pr\{Q = x\} \propto 1$$

$$q(x) \propto \Pr\{d_Q \mid Q = x\} \cdot \Pr\{Q = x\}$$

$$MLR = \frac{\sum_{x \in X} q(x) \cdot s(x)}{\sum_{x \in X} q(x) \cdot r(x)} = \frac{1}{\sum_{x \in I} r(x)}$$

DNA Inclusion Example



Perpetrator [1 1], [1 2], [2 2]
Suspect [2 2]

$$MLR = \frac{\sum_{x \in X} q(x) \cdot s(x)}{\sum_{x \in X} q(x) \cdot r(x)}$$

$$= \frac{\sum_{x \in I} \frac{1}{3} \cdot s(x)}{\sum_{x \in I} \frac{1}{3} \cdot r(x)}$$

$$= \frac{\frac{1}{3} \cdot 1}{\frac{1}{3} \cdot \left(p_1^2 + 2p_1p_2 + p_2^2\right)}$$

$$= \frac{1}{p_1^2 + 2p_1p_2 + p_2^2}$$

DNA Mixture: Equality

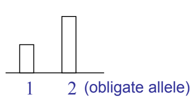
$$\Pr\left\{d_Q \mid Q = x, V = y\right\} = \begin{cases} 1 & alleles(x) \cup alleles(y) = alleles\left(d_Q\right) \\ 0 & otherwise \end{cases}$$

$$\Pr\{Q = x\} \propto 1$$

$$q(x) \propto \Pr\left\{d_Q \mid Q = x, V = y\right\} \cdot \Pr\{Q = x\}$$

$$MLR = \frac{\sum_{x \in X} q(x) \cdot s(x)}{\sum_{x \in X} q(x) \cdot r(x)} = \frac{1}{\sum_{x \in J} r(x)}$$

DNA Equality Example



Victim [1 1]
Perpetrator [1 2], [2 2]
Suspect [2 2]

$$MLR = \frac{\sum_{x \in X} q(x) \cdot s(x)}{\sum_{x \in X} q(x) \cdot r(x)}$$

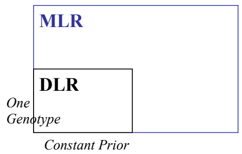
$$= \frac{\sum_{x \in I} \frac{1}{2} \cdot s(x)}{\sum_{x \in I} \frac{1}{2} \cdot r(x)}$$

$$= \frac{\frac{1}{2} \cdot 1}{\frac{1}{2} \cdot \left(2p_1p_2 + p_2^2\right)}$$

$$= \frac{1}{2p_1p_2 + p_2^2}$$

DLR is a special case of MLR

Suspect S genotype probability



Evidence Q genotype prior probability

Mixture: Evidence Q *0/1 likelihood function*

DNA Mixture: Quantitative

$$\Pr\left\{d_Q \mid Q = x, V = y\right\} = N\left(\left[w\mathbf{x} + (1-w)\mathbf{y}\right] \cdot u, \Sigma\right)$$

$$\Pr\left\{Q = x\right\} = r(x)$$

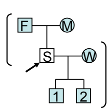
$$q(x) \propto \Pr\left\{d_Q \mid Q = x, V = y\right\} \cdot \Pr\left\{Q = x\right\}$$

$$MLR = \frac{\sum_{x \in X} q(x) \cdot s(x)}{\sum_{x \in X} q(x) \cdot r(x)}$$

1, 2 & 3 unknown
low copy number
damaged DNA
case-to-case
method validation

DNA Kinship

$$\Pr\left\{S = x \mid F = y, M = z\right\}$$



$$\Pr\left\{d_S \mid S = x, W = w\right\}$$

$$s(x) \propto \Pr\left\{d_S \mid S = x, W = w\right\} \cdot \Pr\left\{S = x \mid F = y, M = z\right\}$$

$$MLR = \frac{\sum_{x \in X} q(x) \cdot s(x)}{\sum_{x \in X} q(x) \cdot r(x)}$$

missing persons
paternity
familial search
mass disaster

Population Substructure

Genotypes are not independent: shared ancestry

Coancestry coefficient θ: allele IBD probability

Induces a measure μ on joint genotypes

$$MLR = \frac{\sum_{x \in X} q(x) \cdot s(x) \cdot \mu(\theta, x)}{\sum_{x \in X} q(x) \cdot r(x) \cdot \mu(\theta, x)}$$

Conclusions

- introduced MLR: "match likelihood ratio"
- an inner product ratio statistic
- agrees with "data likelihood" DNA LR
- extends LR to other DNA applications
- statistical framework for non-DNA forensics