Error in the likelihood ratio: false match probability

American Academy of Forensic Sciences
February, 2017
New Orleans, LA

Mark W Perlin, PhD, MD, PhD
Cybergenetics, Pittsburgh, PA

Cybergenetics

Forensic evidence

Rape kit

?

?

Suspects

Likelihood ratio

Pr	probability
E	evidence
H	guilt
~H	innocence
X	type variable
x	type instance
p(x)	prior probability
q(x)	posterior probability
log	logarithm

$$LR = \frac{\Pr\{E|H\}}{\Pr\{E|\sim H\}} = \frac{\Pr\{X=x|E\}}{\Pr\{X=x\}} = \frac{q(x)}{p(x)}$$

$$\log(LR) = \log\left(\frac{q}{p}\right)$$

Forensic statistics

log(LR)

Rape kit

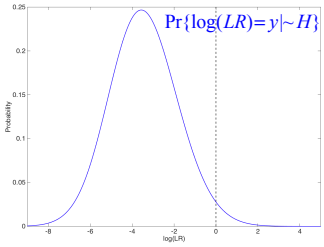
+6

Suspects

-6

Presumption of innocence

The principle that there is a presumption of innocence in favor of the accused is the undoubted law, axiomatic and elementary, and its enforcement lies at the foundation of the administration of our criminal law.
Coffin v. United States (1895)

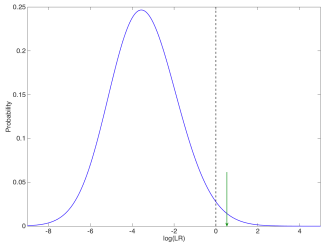


Innocence hypothesis ~H follows prior probability p

Match statistic distribution if **presumed innocent**, **did not contribute** to the evidence

Forensic statistic

Likelihood ratio **summarizes the evidence** against a suspect

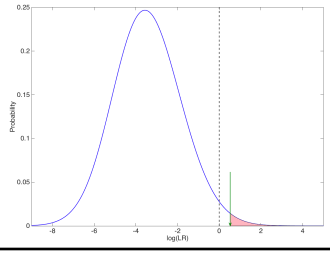


summary
LR = 3
log LR = ½

How much evidence is needed?

Beyond reasonable doubt

False match probability (FMP)
is the **chance of error** that an
innocent person has at least this LR



summary
LR = 3
log LR = ½

error
FMP = 1/100

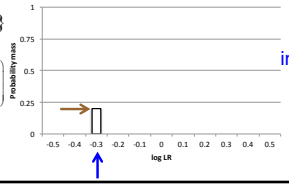
Non-contributor distribution

Type	Prior	Posterior	LR	log LR
x	p(x)	q(x)	q(x)/p(x)	log q/p
1	0.20	0.10	0.5	-0.301
2	0.30	0.15	0.5	-0.301
3	0.25	0.25	1.0	0.000
4	0.25	0.50	2.0	0.301

$$\Pr\{\log(LR)=y|\sim H\}$$

$$\left\{\log\left[q(x)/p(x)\right],\ p(x)\right\}$$

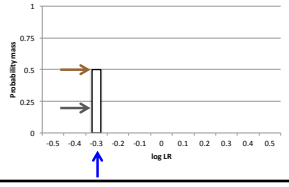
$$for\ every\ x\in X$$



How is log(LR) distributed for innocent people?

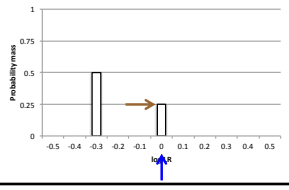
Non-contributor #2

Type	Prior	Posterior	LR	log LR
x	p(x)	q(x)	q(x)/p(x)	log q/p
1	0.20	0.10	0.5	-0.301
2	0.30	0.15	0.5	-0.301
3	0.25	0.25	1.0	0.000
4	0.25	0.50	2.0	0.301



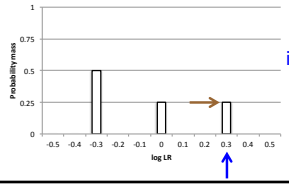
Non-contributor #3

Type	Prior	Posterior	LR	log LR
x	p(x)	q(x)	q(x)/p(x)	log q/p
1	0.20	0.10	0.5	-0.301
2	0.30	0.15	0.5	-0.301
3	**0.25**	**0.25**	**1.0**	**0.000**
4	0.25	0.50	2.0	0.301



Non-contributor #4

Type	Prior	Posterior	LR	log LR
x	p(x)	q(x)	q(x)/p(x)	log q/p
1	0.20	0.10	0.5	-0.301
2	0.30	0.15	0.5	-0.301
3	0.25	0.25	1.0	0.000
4	**0.25**	**0.50**	**2.0**	**0.301**



The log(LR) distribution for innocent people at one locus.

At many loci?

Combine by convolution

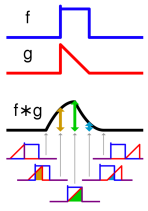
Joint LR is the **product** of independent locus values

$$LR = \prod_{l=1}^{L} LR_l$$

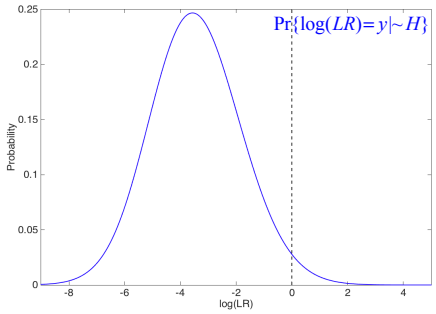
Total **log**(LR) is the **sum** of independent locus values

$$\log LR = \sum_{l=1}^{L} \log LR_l$$

Distribution of a **sum** is the **convolution** of their distributions
(for independent values)



Non-contributor genotype



Southampton rapist

Stranger rape on New Year's Eve

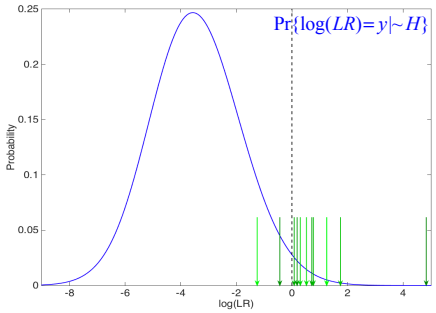
DNA rape kit evidence

A trace amount of semen detected on vaginal swabs was submitted for DNA profiling

The DNA profile contained limited information. A speculative search was done of the UK National DNA Database. The search found 13 profiles, 1 in the Hampshire area.

Due to the low level and incomplete nature of the DNA profile, computer-based statistical evaluation was performed.

Database search



Evidence & error

Item	LR	log(LR)	FMP = Pr(error)	1/FMP
1	1/(17.7)	-1.2485	0.09155110	11
2	1/(2.72)	-0.4339	0.03595410	28
3	1.21	0.0824	0.01818210	55
4	1.54	0.1878	0.01569030	64
5	2.01	0.3025	0.01330630	75
6	3.35	0.5248	0.00958381	104
7	3.35	0.5248	0.00958381	104
8	5.21	0.7166	0.00713871	140
9	5.90	0.7709	0.00655932	152
10	17.8	1.2513	0.00297871	336
11	17.9	1.2535	0.00296855	337
12	55.6	1.7455	0.00123809	808
SB	**67,890**	**4.8318**	**0.00000092**	**1,087,000**

More information

http://www.cybgen.com/information

- Courses
- Newsletters
- Newsroom
- Presentations
- Publications
- Webinars

http://www.youtube.com/user/TrueAllele
TrueAllele YouTube channel

Cybergenetics

YouTube

perlin@cybgen.com