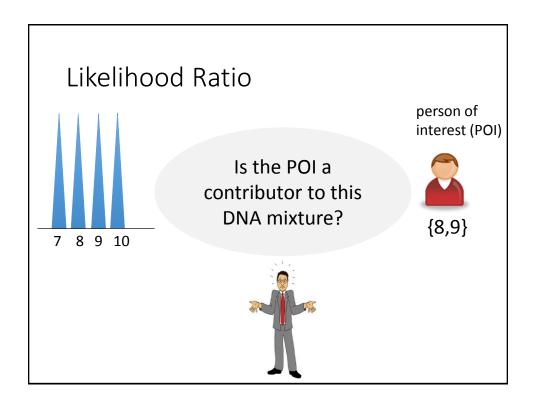
Likelihood Ratios for Mixtures: Binary Approach

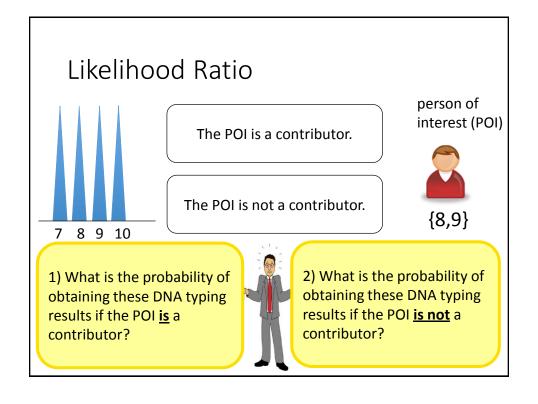
Acknowledgement

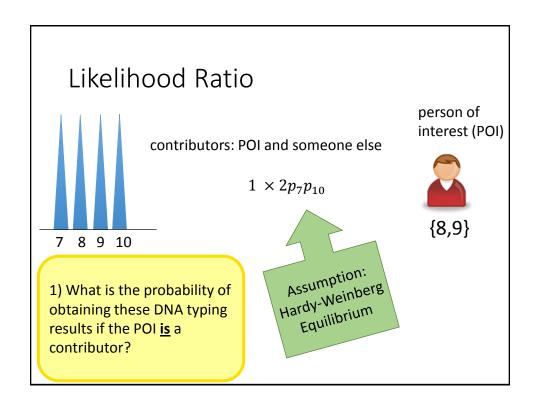
I thank Simone Gittelson, Bruce Weir and John Buckleton for their helpful discussions.

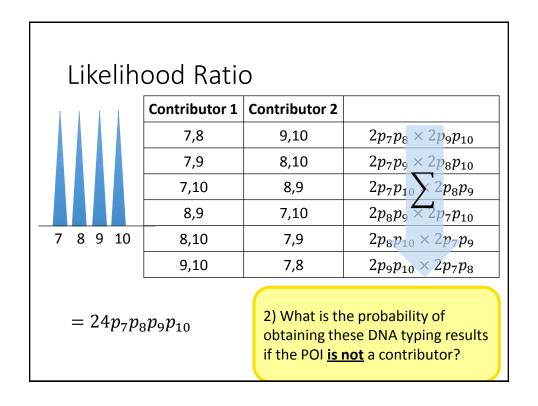
Disclaimer

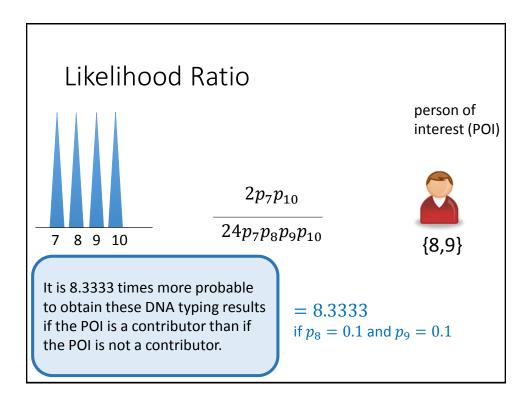
<u>Points of view in this presentation are mine</u> and do not necessarily represent the official position or policies of the National Institute of Standards and Technology.

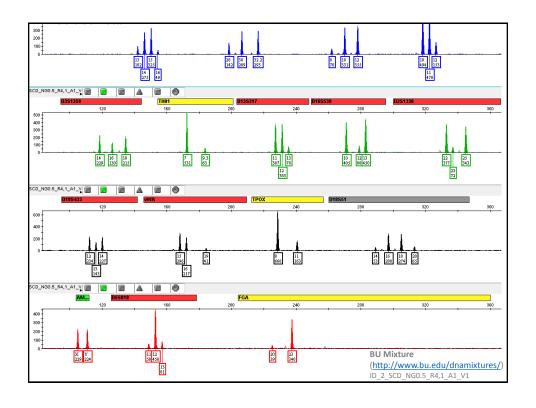




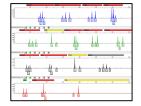








Standpoints of the prosecution and the defense

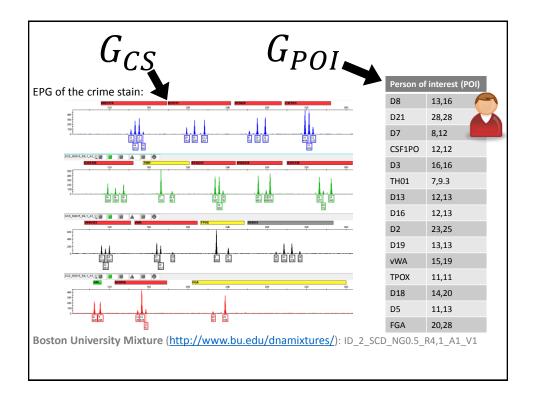


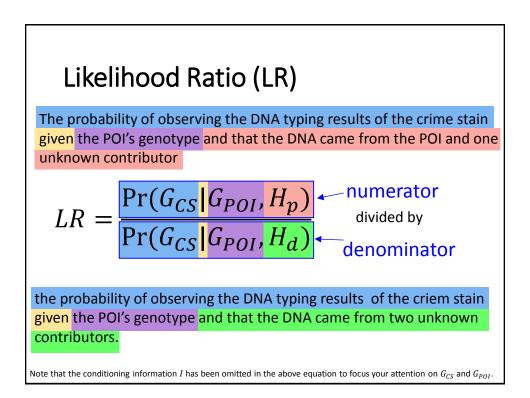


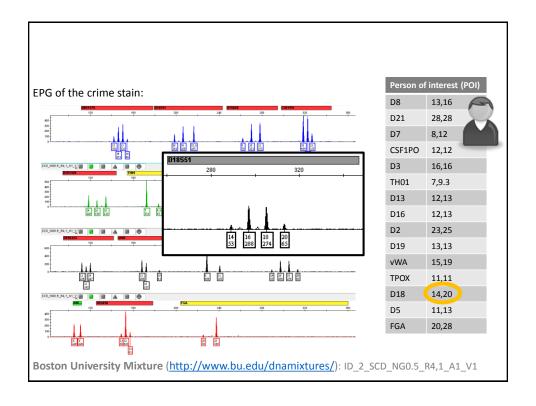
EPG of the crime stain

 H_p : The DNA came from the POI and an unknown contributor.

 H_d : The DNA came from two unknown contributors.





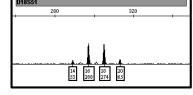


D18S51

 $p_{14} = 0.134$

 $p_{16} = 0.147$ $p_{18} = 0.078$

 $p_{20} = 0.018$





Numerator:

What is the probability of obtaining these DNA typing results for the crime stain if the POI is a contributor and the POI has genotype {14,20}?

Major	Minor
16,18	14,20

$$Pr(16,18) \times Pr(14,20)$$

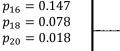
$$=2p_{16}p_{18}\times 1$$

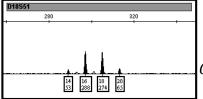
$$=2p_{16}p_{18}$$

Likelihood Ratio (LR)

D18S51

 $p_{14} = 0.134$







 $G_{POI} = \{14,20\}$

Denominator:

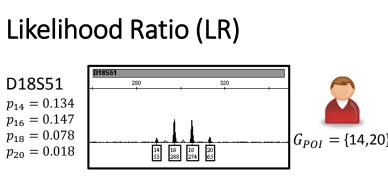
What is the probability of obtaining these DNA typing results for the crime stain if the POI is not a contributor?

Major	Minor
16,18	14,20

$$Pr(16,18) \times Pr(14,20)$$

$$=2p_{16}p_{18}\times 2p_{14}p_{20}$$

$$=4p_{14}p_{16}p_{18}p_{20}$$

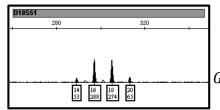


$$LR = \frac{2p_{16}p_{18}}{4p_{14}p_{16}p_{18}p_{20}}$$

$$= \frac{1}{2p_{14}p_{20}}$$

$$= 207.30$$

D18S51 $p_{14} = 0.134$ $p_{16} = 0.147$ $p_{18} = 0.078$ $p_{20} = 0.018$





The DNA typing results are 207 times more probable if the DNA came from the person of interest and an unknown contributor than if the DNA came from two unknown contributors.

Unrestricted LR - Peak Heights are ignored

Likelihood Ratio (LR)

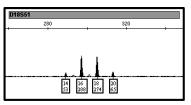
D18S51

 $p_{14} = 0.134$

 $p_{16} = 0.147$

 $p_{18} = 0.078$

 $p_{20} = 0.018$





 $G_{POI} = \{14,20\}$

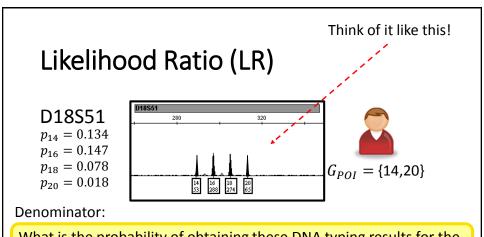
Numerator:

What is the probability of obtaining these DNA typing results for the crime stain if the POI is a contributor and the POI has genotype {14,20}?

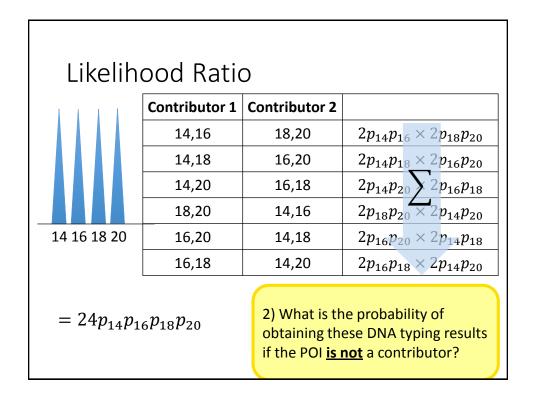
$$Pr(16,18) \times Pr(14,20)$$

 $=2p_{16}p_{18}\times 1$

 $=2p_{16}p_{18}$



What is the probability of obtaining these DNA typing results for the crime stain if the POI <u>is not</u> a contributor?



$$\frac{2p_{16}p_{18}}{24p_{14}p_{16}p_{18}p_{20}} = \frac{1}{12p_{14}p_{20}} = \frac{1}{0.0289}$$

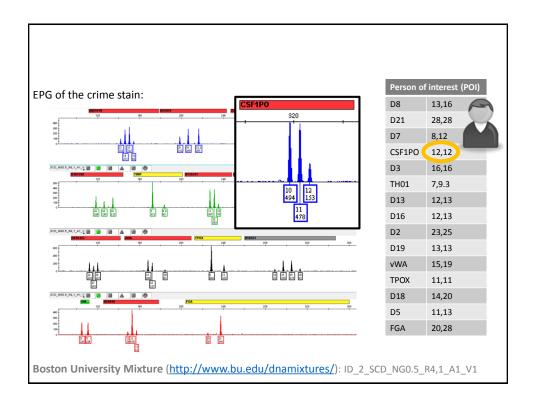
$$= 34.55$$
Restricted LR
$$LR = \frac{2p_{16}p_{18}}{4p_{14}p_{16}p_{18}p_{20}}$$

$$= \frac{1}{12p_{14}p_{20}}$$

$$= \frac{1}{2p_{14}p_{20}}$$

$$= 207.30$$

$$= 34.55$$

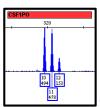






 $p_{10} = 0.220$

 $\begin{array}{l} p_{11} = 0.309 \\ p_{12} = 0.360 \end{array}$





 $G_{POI} = \{12,12\}$

allele at 12 is <u>above</u> the

Numerator: stochastic threshold

What is the probability of obtaining these DNA typing results for the crime stain if the POI is a contributor and the POI has genotype

{12,12}?

Major	Minor
10,11	12,12
10,11	10,12
10,11	11,12

$$Pr(10,11) \times Pr(12,12)$$

$$=2p_{10}p_{11}\times 1$$

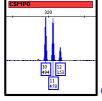
$$=2p_{10}p_{11}$$

Likelihood Ratio (LR)

CSF1PO

 $p_{10} = 0.220$

 $p_{11} = 0.309$ $p_{12} = 0.360$





 $G_{POI} = \{12,12\}$

allele at 12 is <u>above</u> the stochastic threshold

Denominator:

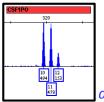
What is the probability of obtaining these DNA typing results for the crime stain if the POI <u>is not</u> a contributor?

Major	Minor
10,11	12,12
10,11	10,12
10,11	11,12





 $p_{10} = 0.220$ $p_{11} = 0.309$ $p_{12} = 0.360$





 $G_{POI} = \{12,12\}$

allele at 12 is <u>above</u> the stochastic threshold

Denominator:

$$Pr(10,11) \times Pr(12,12) + Pr(10,11) \times Pr(10,12) + Pr(10,11) \times Pr(11,12)$$

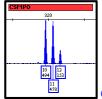
Major	Minor
10,11	12,12
10,11	10,12
10,11	11,12

$$= 2p_{10}p_{11} \times p_{12}^2 + 2p_{10}p_{11} \times 2p_{10}p_{12}$$
$$+2p_{10}p_{11} \times 2p_{11}p_{12}$$
$$= 2p_{10}p_{11}p_{12}(p_{12} + 2p_{10} + 2p_{11})$$

Likelihood Ratio (LR)

CSF1PO

 $p_{10} = 0.220$ $p_{11} = 0.309$ $p_{12} = 0.360$





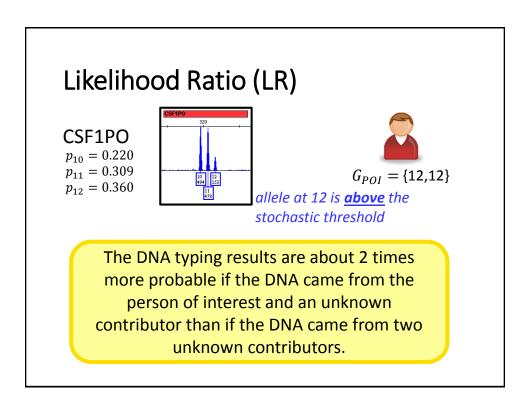
 $G_{POI} = \{12,12\}$

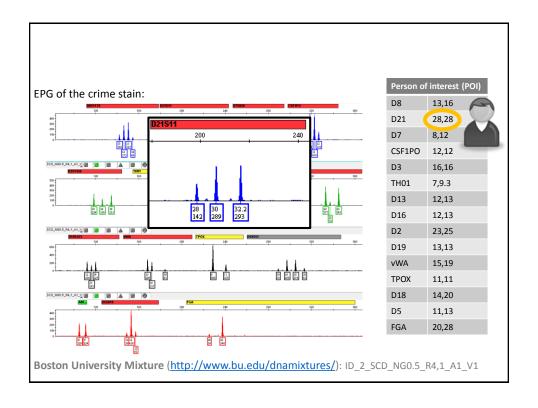
allele at 12 is <u>above</u> the stochastic threshold

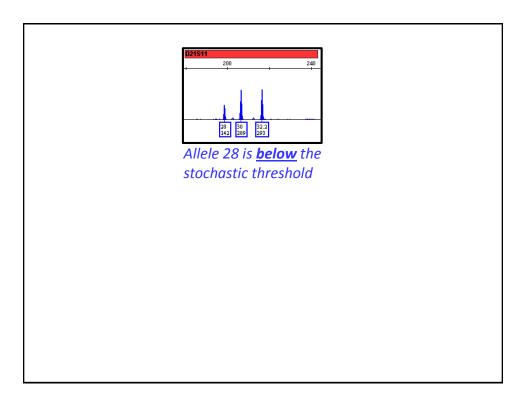
$$LR = \frac{2p_{10}p_{11}}{2p_{10}p_{11}p_{12}(p_{12} + 2p_{10} + 2p_{11})}$$

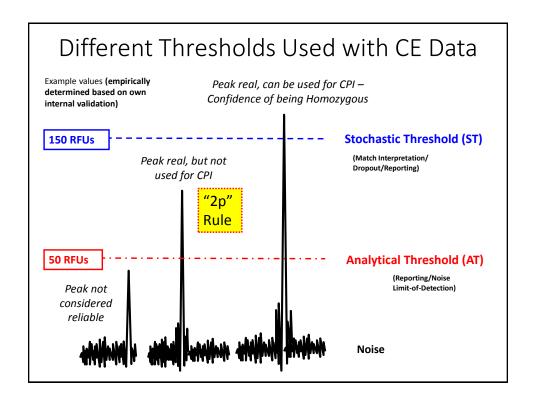
$$= \frac{1}{p_{12}(p_{12} + 2p_{10} + 2p_{11})}$$

$$= 1.96$$





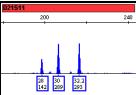






D21S11

 $p_{28} = 0.159$ $p_{30} = 0.283$ $p_{32,2} = 0.090$



stochastic threshold



Allele 28 is **below** the

Numerator:

What is the probability of obtaining these DNA typing results for the crime stain if the POI <u>is</u> a contributor and the POI has genotype {28,28}?

Major	Minor
30,32.2	28,F

$$Pr(30,32.2) \times Pr(28,F)$$

$$=2p_{30}p_{32.2}\times 1$$

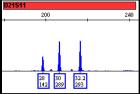
$$=2p_{30}p_{32.2}$$

Likelihood Ratio (LR)

D21S11

 $p_{28} = 0.159$ $p_{30} = 0.283$

 $p_{32,2} = 0.090$



stochastic threshold

Allele 28 is **below** the

"2p" or $p^2 + 2p(1-p)$

 $G_{POI} = \{28,28\}$

Denominator:

What is the probability of obtaining these DNA typing results for the crime stain if the POI is not a contributor?

Major Minor 30,32.2 28,F

$$Pr(30,32.2) \times Pr(28,F)$$

$$=2p_{30}p_{32.2}\times[2p_{28}(1-p_{28})+p_{28}^2]$$

$$=2p_{30}p_{32.2}(2p_{28}-p_{28}^2)$$

"2p" or $p^2 + 2p(1-p)$

Derivation of the 2p Rule

• Two ways to think of it... (easy)

$$2pq \longrightarrow 2pq \longrightarrow 2p$$

"2p" or $p^2 + 2p(1-p)$

Derivation of the 2p Rule

• Two ways to think of it... (mathematical)

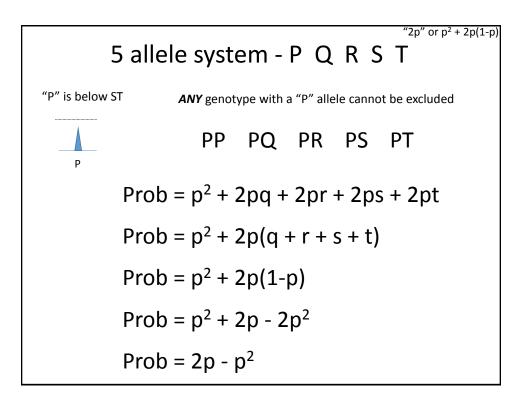
5 allele system - P Q R S T

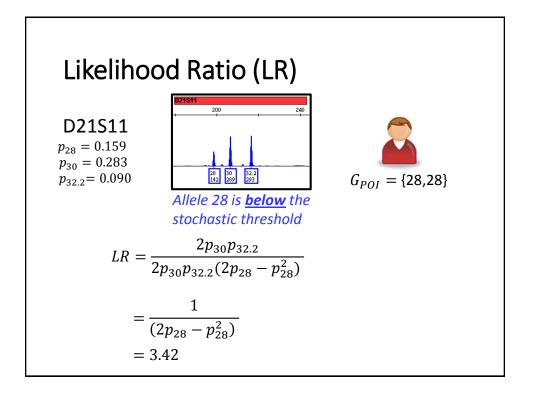
"P" is below ST

ANY genotype with a "P" allele cannot be excluded



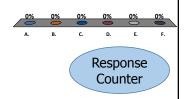
PP PQ PR PS PT

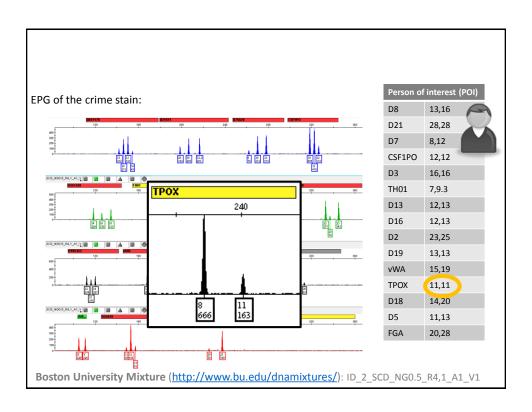




What does a $LR \approx 3$ mean?

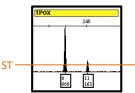
- A. The person of interest committed the crime.
- B. A total of 3 peaks were observed at this locus.
- C. It is about 3 times more probable that the DNA came from the person of interest and an unknown contributor than that the DNA came from two unknown contributors.
- There are 3 contributors to this DNA mixture.
- E. The DNA typing results are about 3 times more probable if the DNA came from the person of interest and an unknown contributor than if the DNA came from two unknown contributors.

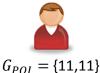




TPOX

 $p_8 = 0.525$ $p_{11} = 0.252$





The peak at 11 is above

Numerator:

the stochastic threshold.

What is the probability of obtaining these DNA typing results for the crime stain if the POI is a contributor and the POI has genotype {11,11}?

Major	Minor
8,8	11,11
8,8	8,11

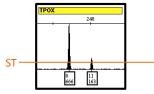
$$Pr(8,8) \times Pr(11,11)$$

 $= \cdots$

Likelihood Ratio (LR)

TPOX

 $p_8 = 0.525$ $p_{11} = 0.252$





 $G_{POI} = \{11,11\}$

The peak at 11 is **above** Denominator:

the stochastic threshold.

What is the probability of obtaining these DNA typing results for the crime stain if the POI is not a contributor?

Major	Minor
8,8	11,11
8,8	8,11

$$Pr(8,8) \times Pr(11,11) + Pr(8,8) \times Pr(8,11)$$

= …

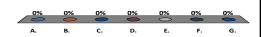


A.
$$\frac{p_8^2}{p_8^2(p_{11}^2 + 2p_8p_{11})} = \frac{1}{p_{11}(p_{11} + 2p_8)}$$

The peak at 11 is <u>above</u> the stochastic threshold.

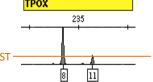
- $B. \frac{1}{p_{11}+2p_8}$
- *C.* 1
- $D. \frac{1}{2p_8p_{11}}$
- E. $\frac{1}{p_{11}^2}$
- F. infinity
- G. ???

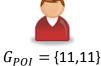
Response Counter



Likelihood Ratio (LR)

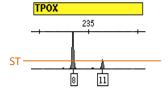
TPOX $p_8 = 0.525$ $p_{11} = 0.252$

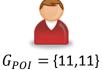




= 3.05







The DNA typing results are about 3 times more probable if the DNA came from the person of interest and an unknown contributor than if the DNA came from two unknown contributors.

Likelihood Ratio (LR) for all loci

 H_p : The DNA came from the POI and an unknown contributor.

 H_d : The DNA came from two unknown contributors.

If H_p is true, is the POI the major contributor or the minor contributor?

If H_p is true, the POI could be either the major contributor or the minor contributor. Let us consider these possibilities to be equally probable. So if H_p is true, there is a probability of $\frac{1}{2}$ that the POI is the major contributor and a probability of $\frac{1}{2}$ that the POI is the minor contributor.

We can only observe these DNA typing results if the POI is the minor contributor.



D18S51:

Major	Minor
16,18	14,20

$$G_{POI} = \{14,20\}$$

CSF1PO:

Major	Minor
10,11	12,12
10,11	10,12
10,11	11,12

$$G_{POI} = \{12,12\}$$

D21S11:

Major	Minor
30,32.2	28,F

$$G_{POI} = \{28,28\}$$

TPOX:

Major	Minor
8,8	11,11
8,8	8,11

$$G_{POI} = \{11,11\}$$

Likelihood Ratio (LR) for all loci

 H_p : The DNA came from the POI and an unknown contributor.

 H_d : The DNA came from two unknown contributors.

Numerator:

Because these DNA typing results are only possible when the POI is the minor contributor, and the POI is the minor contributor with a probability of $\frac{1}{2}$, we multiply the numerator of the likelihood ratio for the entire profile by $\frac{1}{2}$.

Locus	Likelihood Ratio
D8S1179	3.66
D21S11	3.42
D7S820	3.74
CSF1PO	1.96
D3S1358	2.39
TH01	1.75
D13S317	4.58
D16S539	1.89
D2S1338	5.03
D19S433	1.29
vWA	1
TPOX	3.05
D18S51	207.30
D5S818	3.77
FGA	1

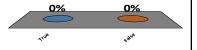
All Loci: $LR = 2.5 \times 10^7$

True or false?

A likelihood ratio of 2.5×10^7 means that it is 2.5×10^7 times more probable that the DNA came from the person of interest and an unknown contributor than that the DNA came from two unknown contributors.

- A. True
- B. False

Response Counter

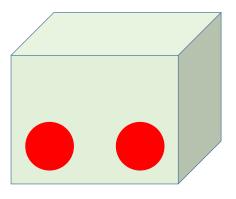


$$LR = 2.5 \times 10^7 = 25 \text{ million}$$

The DNA typing results are about 25 million times more probable if the DNA came from the person of interest and an unknown contributor than if the DNA came from two unknown contributors.

Factor of 2





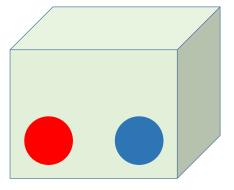
What is the probability you will randomly pick a red marble?

Pr(picking left) x Pr(left is red) + Pr(picking right) x Pr(right is red)

$$0.5 \times 1 + 0.5 \times 1$$

= 1

Suppose...



What is the probability you will randomly pick a red marble?

Pr(picking left) x Pr(left is red) + Pr(picking right) x Pr(right is red)

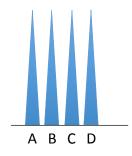
$$0.5 \times 1 + 0.5 \times 0$$

= 0.5



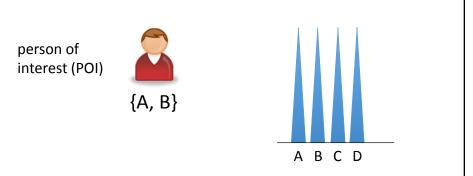
person of interest (POI)





Hp: POI and someone else

Hd: DNA came from 2 unknowns



	Contributor 1	Contributor 2	probability of peaks in crime stain EPG if H_p is true	probability of peaks in crime stain EPG if H_d is true
1	AB	CD	$1 \times 2p_{\mathcal{C}}p_{\mathcal{D}}$	$2p_A p_B \times 2p_C p_D$
2	AC	BD	0	$2p_A p_C \times 2p_B p_D$
3	AD	BC	0	$2p_A p_D \times 2p_B p_C$
4	BC	AD	0	$2p_B p_C \times 2p_A p_D$
5	BD	AC	0	$2p_B p_D \times 2p_A p_C$
6	CD	AB	$2p_{c}p_{D}\times 1$	$2p_C p_D \times 2p_A p_B$

Only 2 genotype combinations are possible for Hp

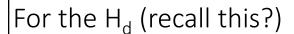
	Contributor 1	Contributor 2	probability of peaks in	probability of peaks in crime
			crime stain EPG if $H_{m p}$ is true	stain EPG if H_d is true
1	AB	CD	$1 \times 2p_{C}p_{D}$	$2p_A p_B \times 2p_C p_D$
2	AC	BD	0	$2p_Ap_C \times 2p_Bp_D$
3	AD	BC	0	$2p_A p_D \times 2p_B p_C$
4	BC	AD	0	$2p_Bp_C \times 2p_Ap_D$
5	BD	AC	0	$2p_B p_D \times 2p_A p_C$
6	CD	AB	$2p_{C}p_{D}\times 1$	$2p_C p_D \times 2p_A p_B$

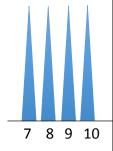
Нр

Row 1 - 0.5 x 1 x
$$2p_c p_d$$

Row 6 - 0.5 x
$$2p_c p_d$$
 x 1

= 0.5 x 1 x
$$2p_c p_d + 0.5$$
 x $2p_c p_d$ x 1
= $2p_c p_d$





Contributor 1	Contributor 2	
7,8	9,10	$2p_7p_8\times 2p_9p_{10}$
7,9	8,10	$2p_7p_9 \times 2p_8p_{10}$
7,10	8,9	$2p_7p_{10}$ $2p_8p_9$
8,9	7,10	$2p_8p_9 \times 2p_7p_{10}$
8,10	7,9	$2p_8p_{10}\times 2p_7p_9$
9,10	7,8	$2p_9p_{10}\times 2p_7p_8$

 $= 24p_7p_8p_9p_{10}$

2) What is the probability of obtaining these DNA typing results if the POI **is not** a contributor?

Likelihood Ratio

$$\frac{2p_c p_d}{24p_a p_b p_c p_d} = \frac{1}{12p_a p_b}$$

This is like the box with 2 red marbles: 0.5 + 0.5 = 1 in the numerator

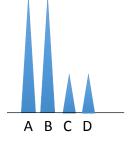
Anytime you have a "mirror image" for the numerator, there is no factor of 0.5 (or factor of 2 in the denominator)

Now Suppose...

person of interest (POI)



{A, B}



Hp: POI and someone else

Hd: DNA came from 2 unknowns

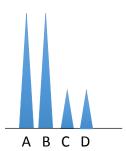
person of interest (POI)



{A, B}

Hp: POI and someone else

Hd: DNA came from 2 unknowns



	Contributor 1 (MAJOR)	Contributor 2 (MINOR)	probability of peaks in crime stain EPG if H_p is true	probability of peaks in crime stain EPG if H_d is true
1	AB	CD	$1 \times 2p_{C}p_{D}$	$2p_A p_B \times 2p_C p_D$
2	AC	BD	0	0
3	AD	BC	0	0
4	BC	AD	0	0
5	BD	AC	0	0
6	CD	AB	0	0

	Contributor 1	Contributor 2	probability of peaks in	probability of peaks in crime
	(MAJOR)	(MINOR)	crime stain EPG if H_p is true	stain EPG if H_d is true
1	AB	CD	$1 \times 2p_{\mathcal{C}}p_{\mathcal{D}}$	$2p_A p_B \times 2p_C p_D$
2	AC	BD	0	0
3	AD	BC	0	0
4	BC	AD	0	0
5	BD	AC	0	0
6	CD	AB	0	0

$$Hp = 0.5 \times 1 \times 2p_c p_d = 0.5 \times 2p_c p_d$$

$$Hd = 2p_a p_b \times 2p_c p_d$$

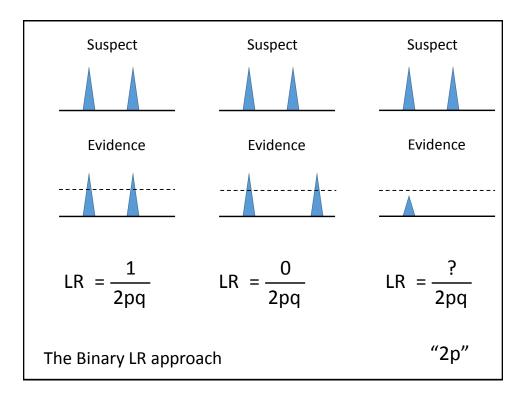
Likelihood Ratio

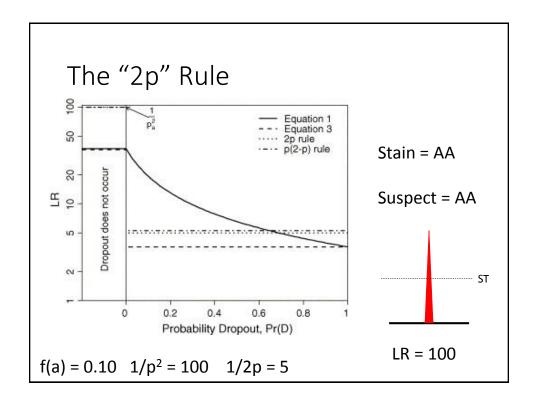
$$\frac{0.5 \times 2p_c p_d}{4p_a p_b p_c p_d} = \frac{0.5}{2p_a p_b} = \frac{1}{4p_a p_b}$$

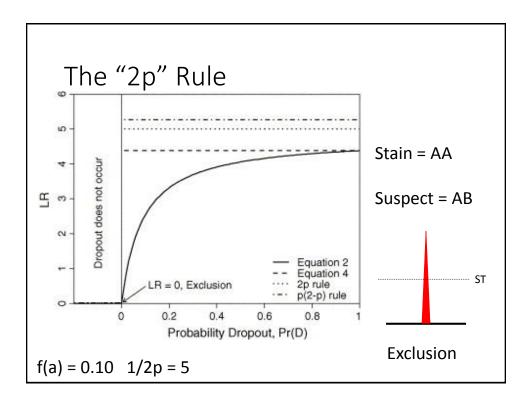
This is like the box with 1 red, 1 blue marble: 0.5 (left)+ 0.5 (right)

Therefore, a factor of 0.5 appears in the numerator (or factor of 2 in the denominator)

So - why do we even need probabilistic genotyping?







Whatever way uncertainty is approached, probability is the *only* sound way to think about it.



-Dennis Lindley