



Valoración estadística de las pruebas en genética forense

Cláudia Vieira da Silva

Instituto Nacional de Medicina Legal e Ciências Forenses

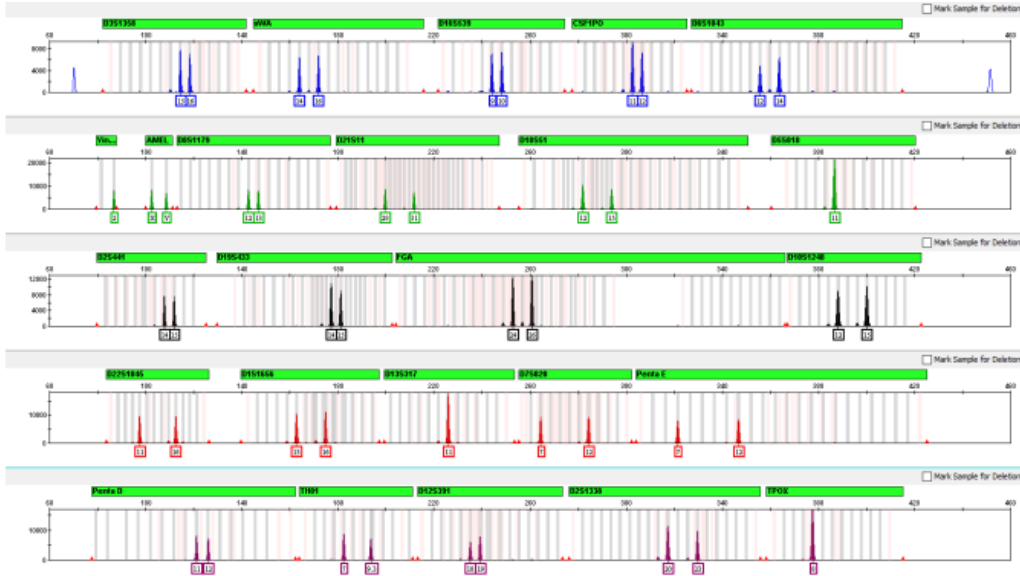
Serviço de Genética e Biologia Forenses

índice

- ▶ Probabilidad condicional
- ▶ *Likelihood Ratio*- coeficiente de verosimilitud
- ▶ Cálculos de parentesco
- ▶ Cálculos con mezclas
- ▶ Desafíos futuros

Fundamentos probabilísticos de las pruebas genéticas

Las pruebas genéticas son claramente probabilísticas

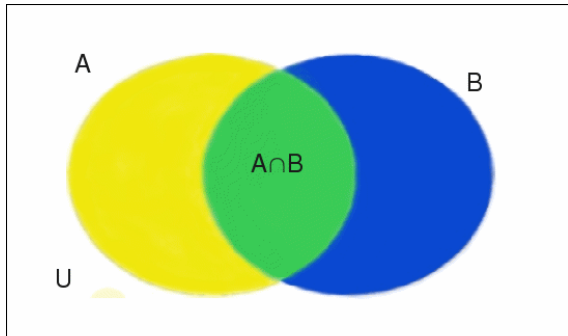


Control 007

- En muchas situaciones, un testigo de un delito no puede decir con absoluta certeza quién cometió un delito concreto.
- El perfil genético se obtiene a partir del estudio de determinadas regiones del ADN. El perfil genético se obtiene a partir del estudio de determinadas regiones del ADN, por lo que, aunque supongamos que cada individuo tiene un perfil genético propio y único, al estudiar sólo determinadas regiones de la cadena de ADN podemos predecir que puede haber otros individuos con el mismo perfil genético.

Las leyes de la probabilidad importantes en genética forense

- ▶ A cada suceso A le corresponde un número real $P(A)$ (*Probabilidad del suceso*);
- ▶ La regla del producto: probabilidad de que dos alelos independientes se presenten juntos $P(A \cap B) = P(A) \times P(B)$;
- ▶ Probabilidad condicionada: probabilidad del suceso A condicionado por B



$$P(A \setminus B) = \frac{P(A \cap B)}{P(B)}$$

$$P(B \setminus A) = \frac{P(B \cap A)}{P(A)}$$

$$P(A \cap B) = P(B) \cdot P(A \setminus B) = P(A) \cdot P(B \setminus A)$$

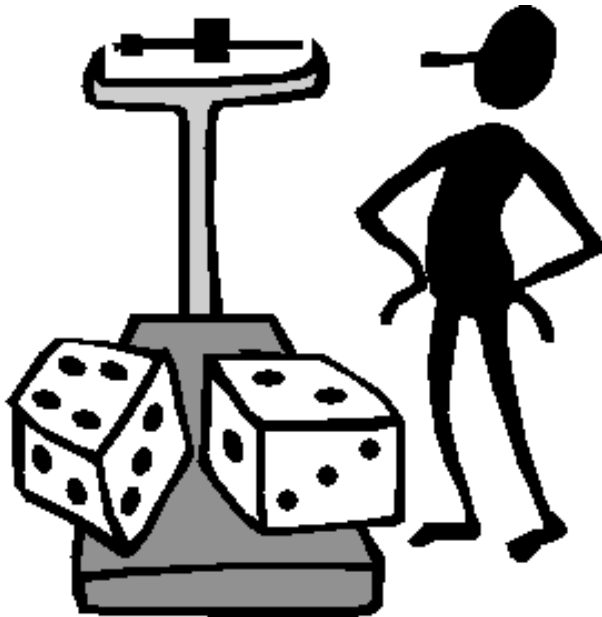
$$P(A \setminus B) = \frac{P(A \cap B)}{P(B)}$$



Teorema de Bayes

Construcción de preposiciones

- ▶ En todos los casos es necesario formalizar las preposiciones adecuadas evaluación de la prueba.
- ▶ Surgen de la información sobre los casos.
- ▶ Resumen ambos puntos de vista en la corte
- ▶ Contrasta la hipótesis de la acusación (H0) con la hipótesis de la defensa (H1):



Hipótesis H0-Proposición a favor de la acusación (en un caso penal) o a favor del parentesco (en un caso de investigación de parentesco)

Ej: H0 - la traza fue producida por el sospechoso Y)
H0 - El individuo X es el padre - demanda civil

Hipótesis H1-Proposición relativa a la defensa de un sospechoso concreto en un caso penal o de un presunto padre (en un caso de investigación de parentesco)

Ej: H1 - la traza fue producida por otro individuo aleatorio de la población)

H1 - El padre biológico es un individuo aleatorio de la población

Aplicación del Teorema de Bayes

Tomemos las siguientes proposiciones:

H0: Hipótesis de acusación

Ejs: El individuo es el padre biológico

El hijo de A y B es el padre biológico

La traza biológica fue producida por la víctima (un perfil genético individual)

La traza biológica fue producida por y el sospechoso (una mezcla)

$$P[H_0 | E, I] = \frac{P[E | H_0, I]P[H_0 | I]}{\underbrace{P[E | H_0, I]P[H_0 | I] + P[E | H_1, I]P[H_1 | I]}_{\approx 1}}$$

H1: Hipótesis de defensa

$$P[H_1 | E, I] = \frac{P[E | H_1, I]P[H_1 | I]}{\underbrace{P[E | H_0, I]P[H_0 | I] + P[E | H_1, I]P[H_1 | I]}_{\approx 1}}$$

E- información de las pruebas genéticas;
I - toda la información relevante para el caso

La razón de verosimilitud

- ▶ La razón de verosimilitud es un estadístico utilizado para cuantificar la bondad relativa del ajuste entre dos modelos estadísticos o hipótesis.
- ▶ dos modelos estadísticos o hipótesis, que se calcula tomando el cociente de sus correspondientes probabilidades condicionadas.

$$LR = \frac{P(H_0 | E, I)}{P(H_1 | E, I)} = \frac{P(E | H_0, I)}{P(E | H_1, I)}.$$

Investigación de parentesco

► Cálculo de la probabilidad de parentesco

Considerando un trío P.Padre (A)/Madre (B)/Hijo (C)

H0: El presunto padre, A, es el padre biológico de C (hipótesis a favor de la paternidad)

H1: Otro individuo aleatorio de la población es el padre biológico de C (hipótesis de defensa).

Considerando: padre biológico E-todas las pruebas genéticas, es decir, GA, GB y GC I-todas las pruebas no genéticas

Por aplicación directa del Teorema de Bayes:

$$P[H_0|E,I] = \frac{P[E|H_0,I]P[H_0|I]}{P[E|H_0,I]P[H_0|I] + P[E|H_1,I]P[H_1|I]}$$

$$P[H_1|E,I] = \frac{P[E|H_1,I]P[H_1|I]}{P[E|H_0,I]P[H_0|I] + P[E|H_1,I]P[H_1|I]}$$

$$LR = \frac{P(H_0|E,I)}{P(H_1|E,I)} = \frac{P(E|H_0,I)}{P(E|H_1,I)} \cdot \frac{P(H_0|I)}{P(H_1|I)}$$

$$W = P(H_0|E,I) = \frac{LR}{LR + 1}$$

Resultados de LR

$$\frac{P(H_0 | E, I)}{P(H_1 | E, I)} = \frac{P(E | H_0, I)}{P(E | H_1, I)} > 1$$

Las pruebas apoyan la hipótesis de que el presunto padre es el padre biológico

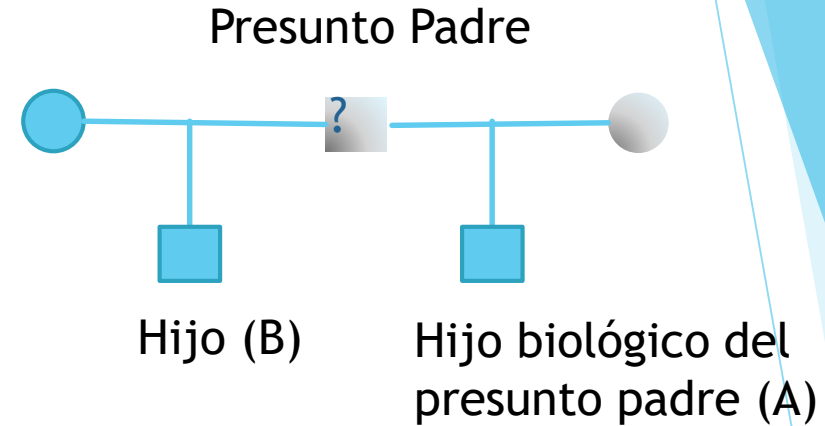
$$\frac{P(H_0 | E, I)}{P(H_1 | E, I)} = \frac{P(E | H_0, I)}{P(E | H_1, I)} < 1$$

Las pruebas no apoyan la hipótesis de que el presunto padre sea el padre biológico. El padre biológico debe ser otro individuo aleatorio de la población.

$$W_{Total} = \frac{LR_{TOTAL}}{LR_{TOTAL} + 1}$$

Casos Prácticos- caso 1

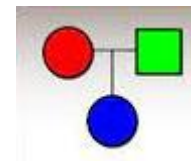
Marcadores Genéticos	Hijo del presunto padre	Madre	Hijo
D8S1179	13, 15	13	13, 14
D21S11	29	30, 32.2	30
D7S820	10, 11	12	12
CSF1PO	12	12	11, 12
D3S1358	16, 18	15, 16	15, 16
TH01	6, 8	9	9
D13S317	11	11, 14	8, 11
D16S539	11	9	9, 12
D2S1338	18, 21	17, 19	17, 25
D19S433	12, 15	14, 15	14, 18.2
VWA	15, 17	16, 17	17, 18
TPOX	8	8, 11	8, 9
D18S51	16	15, 19	17, 19
D5S818	11, 13	11, 12	11, 12
FGA	22, 25	21, 24	24, 26
Penta E	7, 13	5, 15	5, 13
Penta D	11, 15	9, 12	9, 10



H_0 : El padre biológico de B es también el padre biológico de A.

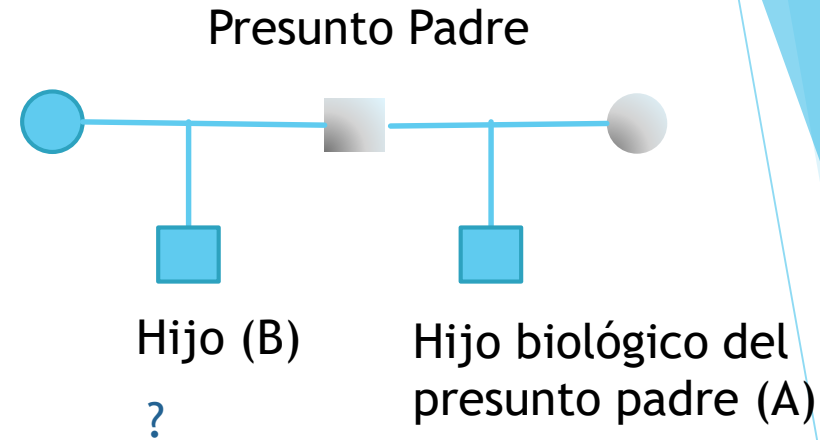
H_1 : El padre biológico de B es otro individuo aleatorio de la población, no emparentado con A.

$$\frac{P(E | H_0, I)}{P(E | H_1, I)} = 0.001$$



Casos Prácticos- caso 1

Locus	Hijo del presunto padre	Hijo
DYS456	12	15
DYS389 I	12	13
DYS390	22	22
DYS389 II	28	28
DYS458	17	-
DYS19	15	16
DYS385 a/b	13, 16	12, 19
DYS393	12	13
DYS391	11	10
DYS439	11	11
DYS635	23	21
DYS392	11	11
GATA H4	11	10
DYS437	16	14
DYS438	9	10
DYS448	21	20



H_0 : El padre biológico de B es también el padre biológico de B.

H_1 : El padre biológico de B es otro individuo aleatorio de la población, no emparentado con A.

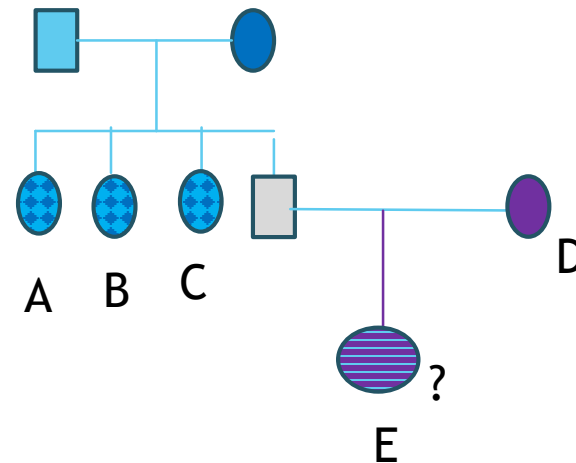


corroboración H_1

Casos Prácticos- caso 2

	A	B	C	D	E
Amelogenina	XX	XX	XX	XX	XX
D3S1358	16, 18	16	14, 16	16	15, 16
vWA	16	16	17, 18	16, 18	16, 17
D16S539	11, 12	11, 12	11, 12	11, 12	11, 12
CSF1PO	12	12	11	12	10, 11
TPOX	8, 12	8, 12	9, 11	8, 12	9, 11
D8S1179	10	10, 13	11, 14	10, 13	11, 14
D21S11	27, 31	27, 32.2	29	31, 32.2	29, 30.2
D18S51	15, 16	13, 16	12, 20	13, 16	18, 20
D2S441	14, 15	14, 15	11.3, 14	14, 15	11, 11.3
D19S433	13, 15	13	12, 14	13	14
TH01	8, 9	8, 9	9, 9.3	8, 9.3	9, 9.3
FGA	22, 25	21, 27	25, 26	21, 26	23, 25
D22S1045	15, 16	11, 16	11, 18	11, 16	11, 18
D5S818	11, 12	11, 12	9, 11	11, 12	9, 11
D13S317	10, 12	10, 12	10, 12	11, 14	11, 12
D7S820	12	10, 12	9, 11	12	8, 9
SE33	20, 27.2	20, 27.2	20, 25.2	18, 27.2	19, 25.2
D10S1248	13, 15	13, 15	14, 16	13, 15	14, 16
D1S1656	12	12	12, 16	12	16

- A - Presunta tía paterna
- B- Presunta tía paterna
- C- Presunta tía paterna
- D- Madre
- E- Hija



H_0 : El padre biológico de E es un hermano de $A/B/C$.

H_1 : El padre biológico de E es otro individuo aleatorio de la población, no emparentado con $A/B/C$.

Casos Prácticos- caso 2

resultados obtenidos con el software Familias 3.0

Presunta Tía Paterna A

Prior probability: 0,5

Posterior probability: 1,0530E-03

Likelihood: 5,1320E-96

Likelihood ratio versus Ped1: 1,0541E-03

Presunta Tía Paterna B

Prior probability: 0,5

Posterior probability: 1,9721E-04

Likelihood: 1,6883E-94

Likelihood ratio versus Ped1: 1,9725E-04

Presunta Tía Paterna C

Prior probability: 0,5

Posterior probability: 1,1857E-03

Likelihood: 1,9758E-94

Likelihood ratio versus Ped1: 1,1857E-03

los resultados obtenidos son claramente favorables a la hipótesis de la defensa.

Casos Prácticos- caso 2

utilizando también información del cromosoma X

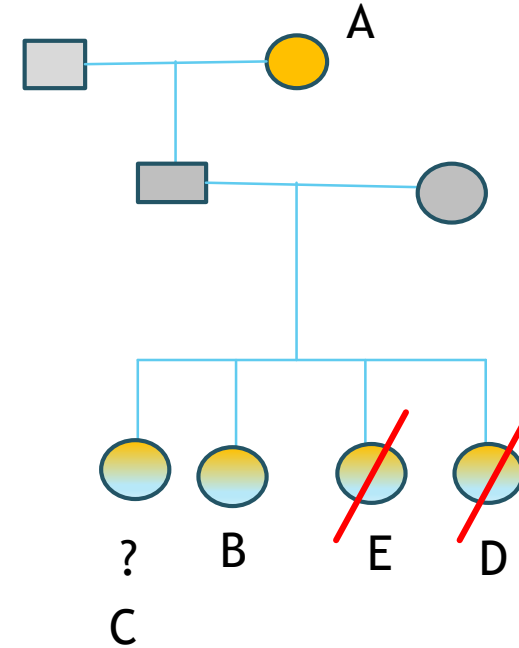
La Presunta tía *A*, la presunta tía *B* y la presunta tía *C*, tienen incompatibilidades con *E*, hija biológica de *D*, en los marcadores del cromosoma X: DXS10074, DXS10101, DXS10135, DXS10146, DXS10148.

	A	B	C	D	E
DXS10103 *	16, 18	18, 19	16, 18	16, 17	17, 18
DXS8378 *	11, 12	11, 12	11, 12	12	10, 12
DXS7132 *	13, 14	13, 14	13, 14	14, 15	14, 15
DXS10134 *	34, 35	34, 35	34, 37	34, 36	35, 36
DXS10074 *	18	18	18	13, 17	8, 13
DXS10101 *	32, 34	30.2, 32	32, 34	28.2, 33	31, 33
DXS10135 *	26, 30	30, 32	26, 30	18.1, 29	18.1, 28
DXS7423 *	15, 16	15, 16	15, 16	14	14, 15
DXS10146 *	30, 31	30, 31	29, 30	28, 29	29, 45.2
DXS10079 *	18, 20	18, 20	16, 18	19, 20	19, 20
HPRTB *	11, 14	11, 13	11, 14	12, 13	12, 13
DXS10148 *	18, 25.1	25.1, 29.1	18, 25.1	24.1	24.1, 27.1

De acuerdo con los resultados obtenidos, el hermano de Presunta tía *A*, presunta tía *B* y presunta tía *C* queda excluido de la paternidad de *E*, hija biológica de *D*.

Casos Prácticos- caso 3

		A	B	C
Amelogenina		X	X	X
D8S1179	(a)	15, 16	14, 15	10, 15
D21S11	(a)	29, 30	29, 30	28, 29
D7S820	(a)	10, 12	9, 11	8, 12
CSF1PO	(a)	10, 12	10, 12	12
D3S1358	(a)	14, 17	17	16, 17
TH01	(a)	8, 9.3	6, 9	9, 9.3
D13S317	(a)	11, 12	11, 13	9, 11
D16S539	(a)	10, 12	10, 12	9, 10
D2S1338*	(a)	19, 20	17, 20	20, 22
D19S433*	(a)	13, 14	15	13, 15
VWA	(a)	17, 20	17, 20	16, 18
TPOX	(a)	9, 11	8, 11	9, 11
D18S51	(a)	12	12, 19	17, 19
D5S818	(a)	11, 12	11	11, 13
FGA	(a)	19, 26	21	20, 26
PentaE	(a)	7, 13	11, 12	7, 10



Casos Prácticos- caso 3 - Estudio inicial

H_0 : El padre biológico de C es un hijo de A e padre biológico de B.

H_1 : El padre biológico de C es otro individuo aleatorio de la población, no emparentado con A y B.

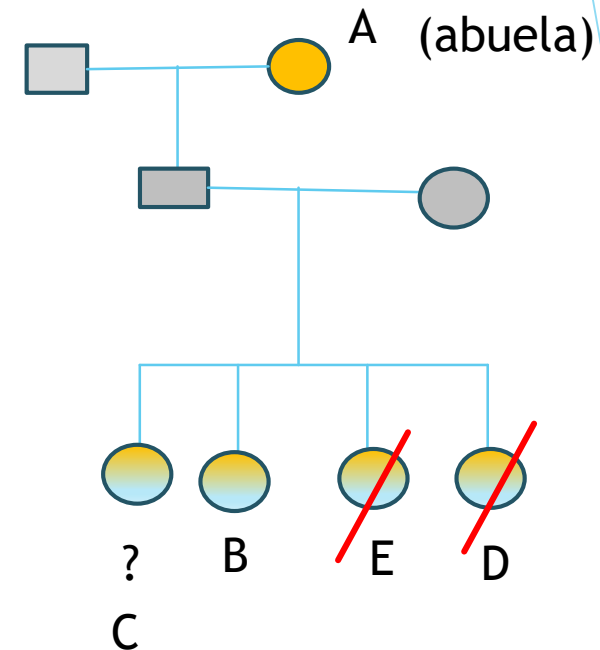
Prior probability: 0,5

Posterior probability: 0,999056708626327

Likelihood: 5,00945878624239E-65

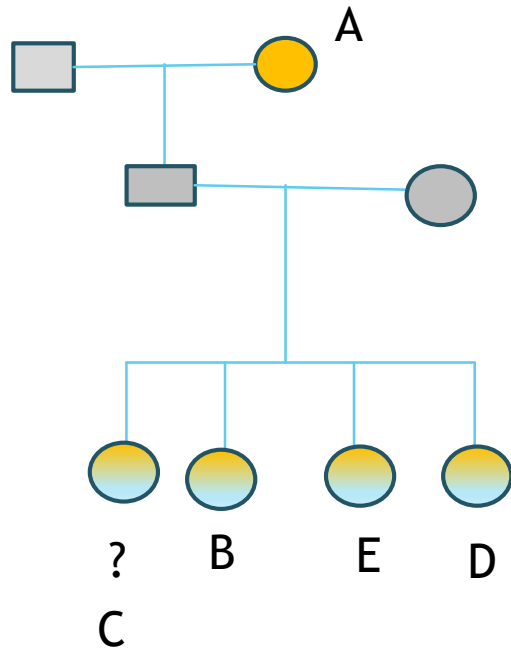
Posterior ratio versus Ped1: 1059,11782

Likelihood ratio versus Ped1: 1059,11782



Casos Prácticos- caso 3- Estudio final

Recibimos muestras de las hermanas E y D



H_0 : El padre biológico de C es un hijo de A e padre biológico de B y E.

El padre biológico de C es un hijo de A e padre biológico de B, E y D.

H_1 : El padre biológico de C es otro individuo aleatorio de la población, no emparentado con A, B y E.

H_1 : El padre biológico de C es otro individuo aleatorio de la población, no emparentado con A, B y E.

LR1=2 117 016 126- El padre biológico de C es un hijo de A e padre biológico de B, E y D.

los resultados obtenidos son claramente favorables a la hipótesis de la acusación.

DESAFÍOS FUTUROS

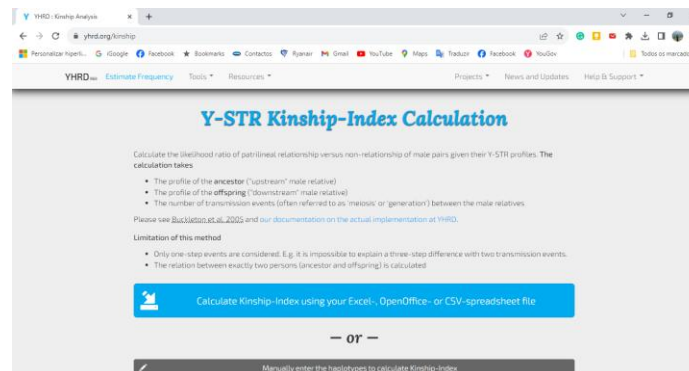
- ❑ introducir nuevos kits en la acreditación
- ❑ utilizar todos los programas informáticos de cálculo para valorar las pruebas.



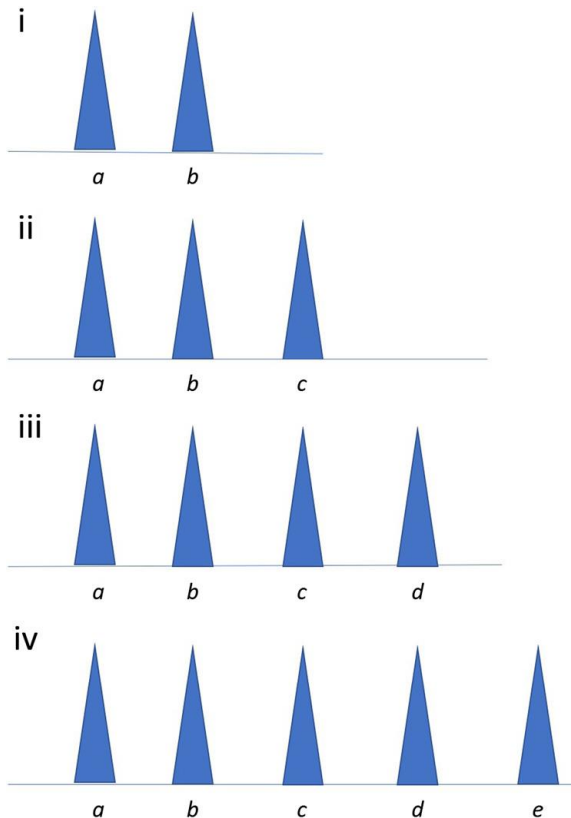
FamLinkX – implementation of a general model for likelihood computations for X-chromosomal marker data

[Daniel Kling](#)^{a, b}, [Barbara Dell'Amico](#)^c, [Andreas O. Tillmar](#)^{c, d}

Show more ▾



Análisis de casos penales



- ▶ Los avances recientes en las tecnologías de tipificación de ADN han llevado al análisis de mezclas de ADN cada vez más complejas de dos, tres o incluso más contribuyentes

Cálculo de LR en casos penales

$$\frac{P(H_0 | E, I)}{P(H_1 | E, I)} = \frac{P(E | H_0, I)}{P(E | H_1, I)} > 1$$

Las pruebas genéticas apoyan la hipótesis de la acusación



$$\frac{P(H_0 | E, I)}{P(H_1 | E, I)} = \frac{P(E | H_0, I)}{P(E | H_1, I)} < 1$$

Las pruebas no apoyan la hipótesis de la acusación. Favorecen la posibilidad de que sea otro individuo al azar de la población

Caso 1- una mezcla normal

Hisopo vaginal recogido à la victima A.

	vestigios
D3S1358	14 17 18
VWA	14 17 18
D16S539	9 13
CSF1PO	9 10 12
TPOX	8 11
D8S1179	10 13 15
D21S11	29 30.2 31 32.2
D18S51	12 16
D2S441	9 11 14
D19S433	13 14 15
TH01	7 9.3
FGA	19 20 25
D22S1045	11 15 16 17
D5S818	9 11 12 13
D13S317	8 11
D7S820	8 10
SE33	20 22.2 23.2 27
D10S1248	13 15 16
D1S1656	11 17.3 18.3
D12S391	19 21 23 24
D2S1338	17 19 20

H0 - la traza fue producida por la victima X y el sospechoso Y;

H1- la traza fue producida por la victima X y por otro individuo;

$$LR \text{ TOTAL} = 9,73479E044$$

los resultados obtenidos son claramente favorables a la hipótesis de la acusación. La víctima y lo sospechosos Y .

Caso 2- mezcla Una Victima y un Sospechoso

Trozos de pañuelos, positivos para sémen. Puede pertenecer a una victima y a un Sospechosos A

	caso-4-am.problema
D3S1358	14 15
VWA	15 16 17 18
D16S539	9 11
CSF1PO	9 10 11
TPOX	8 9 11
D8S1179	10 13 14
D21S11	29 30 32.2 33.2
D18S51	13 14 15 21
D2S441	11 14
D19S433	14 15 15.2
TH01	6 9
FGA	19 20 21 23
D22S1045	11 15 17
D5S818	10 11
D13S317	10 12 13
D7S820	9 10 12
SE33	18 20 28.2
D10S1248	14 16
D1S1656	13 14 15.3
D12S391	15 18 19.3 22
D2S1338	17 23 24 25

H0 - la traza fue producida por la victima X y un desconocido;

H1- la traza fue producida por dos desconocidos

$$\text{LR TOTAL} = 1,56043\text{E}014$$

H0 - la traza fue producida por lo sospechoso A y una desconocida;

H1- la traza fue producida por dos desconocidos

$$\text{LR TOTAL} = 4,91186\text{E}021$$

Caso 2- mezcla Una Victima y un Sospechoso

Trozos de pañuelos, positivos para sémen. Puede pertenecer a una victima y a un Sospechosos A

	caso-4-am.problema
D3S1358	14 15
VWA	15 16 17 18
D16S539	9 11
CSF1PO	9 10 11
TPOX	8 9 11
D8S1179	10 13 14
D21S11	29 30 32.2 33.2
D18S51	13 14 15 21
D2S441	11 14
D19S433	14 15 15.2
TH01	6 9
FGA	19 20 21 23
D22S1045	11 15 17
D5S818	10 11
D13S317	10 12 13
D7S820	9 10 12
SE33	18 20 28.2
D10S1248	14 16
D1S1656	13 14 15.3
D12S391	15 18 19.3 22
D2S1338	17 23 24 25

H0 - la traza fue producida por la victima X y el sospechoso A;
H1- la traza fue producida por dos desconocidos

LR TOTAL= 2,97578E044

los resultados obtenidos son claramente favorables a la hipótesis de la acusación. La víctima y lo sospechosos A .

Caso 3-mezcla con 3 contribuyentes no relacionados

	2022-2214-25.C1- Víctima
D3S1358	15 16 17 18
VWA	14 17 18 19
D16S539	8 10 11 12
CSF1PO	11 12 13
TPOX	8 10 11
D8S1179	13 14 15 17
D21S11	28 29 30 32.2
D18S51	12 14 15 16 17 18
D2S441	11 11.3 12 14 16
D19S433	13 14 15
TH01	7 8 9 9.3
FGA	20 21 22 23 23.2 24 25
D22S1045	15 16 17 18
D5S818	11 12 13
D13S317	8 9 11 12 13
D7S820	9 10 11 12 13
SE33	17 19 20 23.2 29.2 30.2
D10S1248	13 14 15 16 17
D1S1656	12 14 15 16 17.3 18.3
D12S391	15 17 20 22
D2S1338	17 18 19 20 23

H0 - la traza fue producida por la víctima X y dos desconocidos;

H1- la traza fue producida por tres desconocidos

$$\text{LR TOTAL} = 1,67124\text{E}007$$

H0 - la traza fue producida por la víctima X, el sospechoso A y un desconocido;

H1- la traza fue producida por tres desconocidos

$$\text{LR TOTAL} = 3,96819\text{E}024$$

Caso 3-mezcla con 3 contribuyentes no relacionados

	2022-2214-25.C1- IÁçmina
D3S1358	15 16 17 18
VWA	14 17 18 19
D16S539	8 10 11 12
CSF1PO	11 12 13
TPOX	8 10 11
D8S1179	13 14 15 17
D21S11	28 29 30 32.2
D18S51	12 14 15 16 17 18
D2S441	11 11.3 12 14 16
D19S433	13 14 15
TH01	7 8 9 9.3
FGA	20 21 22 23 23.2 24 25
D22S1045	15 16 17 18
D5S818	11 12 13
D13S317	8 9 11 12 13
D7S820	9 10 11 12 13
SE33	17 19 20 23.2 29.2 30.2
D10S1248	13 14 15 16 17
D1S1656	12 14 15 16 17.3 18.3
D12S391	15 17 20 22
D2S1338	17 18 19 20 23

H0 - la traza fue producida por la víctima X, el sospechoso B y un desconocido;

H1- la traza fue producida por tres desconocidos

$$\text{LR TOTAL} = 2,41624\text{E}013$$

H0 - la traza fue producida por la víctima X, el sospechoso A y el sospechoso B ;

H1- la traza fue producida por tres desconocidos

$$\text{LR TOTAL} = 5,39725\text{E}041$$

los resultados obtenidos son claramente favorables a la hipótesis de la acusación. La víctima y los dos sospechosos A e B.

Dasafíos futuros

acreditación de pruebas con muestras problema
validar el software más reciente



Forensic Science International: Genetics
Volume 41, July 2019, Pages 83-92



CaseSolver: An investigative open source expert system based on EuroForMix

[Øyvind Bleka](#)^a, [Lourdes Prieto](#)^b, [Peter Gill](#)^{a,c}

Show more

+ Add to Mendeley Share Cite

The screenshot shows a web browser window with the URL yhrd.org/mixture. The page title is "Y-STR Mixture Calculation". The main heading is "Y-STR Mixture Calculation". Below the heading, there is a paragraph: "Calculate the likelihood ratio of donorship of a given suspect versus non-donorship. The calculation requires a column named 'Role' in your input file and consists of the following rows (specified at the 'Role' column):".

Trace	A trace (all given markers will be used for calculation)
Suspect	The profile of the suspect (or accused or defendant, or the profiles of the suspects if multiple)
Known Contributor	The profiles of additional known contributors (optional)

The only thing to enter after uploading your file is the number of unknown contributors.

Calculate Y-STR Mixture using your Excel-, OpenOffice- or CSV-spreadsheet

— or —

Manually enter the haplotypes to calculate Y-STR Mixture

If you don't know how to export your samples using Applied Biosystems® GeneMapper® ID/ID-X please read our instructions.
If you don't know how to set up an Excel-, OpenOffice- or CSV-spreadsheet for Y-STR mixture calculation, please see our example or read our instructions.

PRÓXIMOS DESAFÍOS



Forensic Science International: Genetics

Volume 41, July 2019, Pages 83-92



CaseSolver: An investigative open source expert system based on EuroForMix

Øyvind Bleka ^a, Lourdes Prieto ^b, Peter Gill ^{a,c}



```
dens <- density(data, n = npts)
dx <- dens$x
dy <- dens$y
if(add == TRUE)
  plot(0., 0., main,
       ylab,
       if(orientat
dx2 <- (dx - min(dx), dx)
x[1.]
dy2 <- (dy - min(dy), dy)
y[1.]
seqbelow <- rep(y[1.], length(dx))
if(Fill == T)
  confshade(dx2, seqbelow, dy2
```

Muchas Gracias



claudia.v.silva@inmlcf.mj.pt