

CALCULO DE PROBABILIDADES DE PARÁMETROS FORENSES EN SISTEMAS LIGADOS AL SEXO.

Sánchez García M. Sobrón Fernández M.I.
Facultad de Medicina. Facultad de Matemáticas.
Departamento de Estadística e Investigación Operativa.
Universidad Complutense de Madrid.
28040 Madrid.

Arroyo Pardo E.
Facultad de Medicina.
Departamento de Medicina Legal y Toxicología. (Biología Forense)
Universidad Complutense de Madrid.
28040 Madrid.

ABSTRACT

In this paper, mathematical expressions for calculating the probabilities of several forensic parameters in sex-linked genetic systems are shown, i. e. heterozygosity index, power of discrimination, allelic diversity, power of exclusion (conditioned and unconditioned to mother and child typing) and averaged chance of exclusion.

Also, formulas for calculating confidence intervals for probabilities of allelic frequencies in the general population are given.

KEY WORDS: Forensic Parameters, Heterozygosity Index, Power of Discrimination, Allelic Diversity, Power of Exclusion, Probability.

AMS. CLASSIFICATION 60C05

RESUMEN.

En el presente artículo, se derivan las fórmulas matemáticas para el cálculo de probabilidades de diversos parámetros forenses de un sistema ligado al sexo, como son : el índice de heterocigosidad, el poder discriminante, la diversidad alélica, el poder de exclusión , con y sin tipaje; el poder de exclusión medio. También se dan fórmulas para el cálculo de intervalos de confianza para las probabilidades de los distintos alelos del sistema, relativos a la población de referencia.

PALABRAS CLAVE: Parámetros Forenses, Heterocigosidad, Poder de Discriminación, Diversidad Alélica, Poder de Exclusión, Probabilidades.

1 INTRODUCCION.

En técnicas forenses, los sistemas ligados al sexo son de uso relativamente reciente en pruebas de paternidad y reconocimiento e identificación de individuos. Estos sistemas se caracterizan porque los hombres tienen uno de sus alelos en el cromosoma X y el otro en el cromosoma Y; mientras que las mujeres tienen cada uno de sus alelos en un cromosoma X. Este comportamiento diferente entre hombres y mujeres conlleva fórmulas especiales para el cálculo de parámetros forenses, en estos sistemas genéticos.

El objetivo de este artículo es recoger, de forma sistemática, métodos y fórmulas para el cálculo de parámetros forenses, tales como el índice de heterocigosidad, la diversidad alélica, el poder discriminante, el poder de exclusión, así como explicar métodos para el cálculo de intervalos de confianza para estos parámetros y para las probabilidades de los diferentes alelos de cualquier sistema ligado al sexo. El artículo sistematiza las fórmulas para la estimación de los citados parámetros; y proporciona métodos originales para obtener algunos de los citados cálculos.

El trabajo se estructura en cuatro epígrafes. El primero es ésta introducción. En el segundo se exponen los conceptos fundamentales de los parámetros. En el tercero, se desarrollan los métodos para el cálculo de probabilidades; mientras que en el último se proporciona una metodología para determinar intervalos de confianza asociados con los parámetros.

2 CONCEPTOS FUNDAMENTALES.

Los sistemas genéticos que se consideran son los STR, letras que son las iniciales en inglés de Short Tandem Repeat; según se expone en (1) Bruce S. Weir. En general, sobre cada STR, cada persona tiene dos alelos A_i ; A_j ; estando situado cada alelo sobre uno de los cromosomas del par. Los dos alelos pueden ser iguales ó distintos. Esto sucede en todos los STR situados sobre pares de cromosomas que no están ligados al sexo, tanto en hombres, como en mujeres.

Sin embargo, los varones, en los sistemas ligados al sexo, tienen alelos distintos sobre el cromosoma X, que los situados sobre el cromosoma Y; mientras que las mujeres pueden tener el mismo conjunto de alelos en los dos cromosomas; pues ambos son cromosomas X. Este hecho tiene su relevancia; ya que los hijos varones heredan un alelo del STR situado sobre un cromosoma X de la madre y el correspondiente alelo situado sobre el cromosoma Y del padre; mientras que las hijas heredan un alelo del STR de la madre y el otro alelo pertenece al STR situado sobre el único cromosoma X del padre. Como consecuencia de estos hechos, tienen más complejidad las pruebas de paternidad en las hijas que en los varones.

Teniendo en cuenta lo dicho anteriormente, sobre cada mujer de la población de referencia Ω , se observan dos bandas alélicas $\{A_i : A_j\}$ que forman su sistema STR; mientras que sobre los hombres solo se observa un alelo. Se denota por Ω_M , a la población de mujeres y por Ω_H a la población de hombres. Por $\Omega = \Omega_M \cup \Omega_H$ se denota la población conjunta de hombres y mujeres.

Para cada $\omega \in \Omega_M$ se definen las variables aleatorias ξ_1 y ξ_2 ; de tal forma que $\xi_1(\omega) = A_i$ y $\xi_2(\omega) = A_j$; cuando los alelos de ω son $G(\omega) = \{A_i : A_j\}$. La ley de Hardy-Weinberg dice que ξ_1 y ξ_2 son probabilísticamente independientes; esto es:

$$P\{\omega \mid G(\omega) = \{A_i : A_j\}\} = P\{\omega \mid \xi_1(\omega) = A_i\} \cdot P\{\omega \mid \xi_2(\omega) = A_j\}.$$

Esta ley facilita el cálculo de probabilidades, pues sólo es necesario para realizar dicho cálculo conocer las probabilidades de los alelos, cuyo número siempre es inferior

al de las combinaciones genotípicas de alelos del STR. Con esta terminología pasamos a enunciar los conceptos de los principales parámetros forenses.

Se llama **Índice de Heterocigosidad** a la probabilidad de que un individuo, elegido aleatoriamente de la población de referencia, tenga sus dos alelos distintos; esto es, la probabilidad de que el individuo sea heterocigótico.

La **Diversidad Alélica** es la varianza de las probabilidades de los alelos. Si el sistema tiene k alelos y p_i es la probabilidad del alelo A_i ; la diversidad alélica es:

$$DA_1 = \frac{1}{k} \sum_{i=1}^k \left(p_i - \frac{1}{k} \right)^2$$

Otra forma de medir la diversidad alélica es mediante la métrica taxicab ó L_1 :

$$DA_1 = \frac{1}{k} \sum_{i=1}^k \left| p_i - \frac{1}{k} \right|$$

El **poder de discriminación** es la probabilidad de que se puedan distinguir genéticamente dos individuos elegidos aleatoriamente de la población de referencia. Para ello los dos individuos tienen que tener uno ó los dos alelos distintos.

El **poder de exclusión** es la probabilidad de que un padre sea incompatible con una madre y una hija, compatibles entre si; esto es, la probabilidad de que el padre no tenga en su genotipo el alelo ó alelos de la hija no aportados por la madre.

Cuando se conocen los alelos de la madre y de la hija, por haber sido previamente tipados u observados, el poder de exclusión depende de los alelos observados. El **poder de exclusión medio** es la media de los poderes de exclusión observados para cada una de las posibilidades de compatibilidad entre madre e hija.

3- FÓRMULAS PARA EL CALCULO DE PROBABILIDADES FORENSES.

Se analiza un sólo STR. Se supone que el STR estudiado tiene k alelos, denotados por $SA = \{A_1, A_2, A_3, \dots, A_k\}$; con probabilidades $P = \{p_1, p_2, p_3, \dots, p_k\}$; elementos que se usarán en los siguientes cálculos.

3-1. INTERVALOS DE CONFIANZA PARA LAS PROBABILIDADES ALÉLICAS.

1) MUESTRA DE SÓLO HOMBRES.

Se supone que se elige una muestra aleatoria simple de la población de hombres Ω_H ; de tamaño nh. Sobre cada ω_j de la muestra se definen las variables aleatorias:

$$\xi_i(\omega_j) = \begin{cases} 1 & \text{si el genotipo del hombre } \omega_j \text{ tiene el alelo } A_i \text{ en el cromosoma X.} \\ 0 & \text{en otro caso} \end{cases}$$

Llamando $\hat{p}_i = \frac{1}{nh} \sum_{j=1}^{nh} \xi_i(\omega_j)$, a la probabilidad estimada del alelo A_i ; el

intervalo de confianza para dicha probabilidad p_i , calculado de forma aproximada por el teorema central del límite, es:

$$p_i \in \left[\hat{p}_i - n(\alpha/2) \sqrt{\frac{\hat{p}_i(1-\hat{p}_i)}{nh}} \dots; \dots \hat{p}_i + n(\alpha/2) \sqrt{\frac{\hat{p}_i(1-\hat{p}_i)}{nh}} \right]$$

fórmula que se aplica para $i=1,2,\dots,k$.

2) MUESTRA DE SÓLO MUJERES.

Se supone que se elige una muestra aleatoria simple de la población de mujeres Ω_M ; de tamaño nm . Sobre cada ω_j de la muestra se definen las variables aleatorias:

$\xi_i(\omega_j) = n^\circ$ de alelos A_i , que tiene el genotipo de la mujer ω_j en sus dos cromosomas X.

Supuesto que se cumple la ley de H-W, la media y varianza de $\xi_i(\omega_j)$ son:

$$E(\xi_i(\omega_j)) = 0 \cdot (1-p_i)^2 + 1 \cdot 2 \cdot p_i(1-p_i) + 2 \cdot p_i^2 = 2 \cdot p_i.$$

Para obtener un estimador insesgado de p_i , se debe usar la variable $\zeta_i(\omega_j) = 0.5 \cdot \xi_i(\omega_j)$. La varianza de $\zeta_i(\omega_j)$ es:

$$V(\zeta_i(\omega_j)) = E(\zeta_i^2(\omega_j)) - (E(\zeta_i))^2 = (0.5)^2 \cdot (1-p_i)^2 + (0.5)^2 \cdot 2 \cdot p_i(1-p_i) + (0.5 \cdot 2)^2 \cdot p_i^2 - p_i^2 = 0.5 \cdot p_i(1-p_i).$$

Llamando $\hat{p}_i = \frac{1}{nm} \sum_{j=1}^{nm} \xi_i(\omega_j)$, a la probabilidad estimada del alelo A_i ; el intervalo de confianza para dicha probabilidad p_i , calculado de forma aproximada por el teorema central del límite, es:

$$p_i \in \left[\hat{p}_i - n(\alpha/2) \sqrt{\frac{\hat{p}_i(1-\hat{p}_i)}{2 \cdot nm}} \dots; \dots \hat{p}_i + n(\alpha/2) \sqrt{\frac{\hat{p}_i(1-\hat{p}_i)}{2 \cdot nm}} \right]$$

3 MUESTRA CONJUNTA DE HOMBRES Y MUJERES.

Se supone que se elige una muestra aleatoria simple de la población conjunta de hombres y mujeres Ω ; de tamaño n . Sobre cada ω_j de la muestra se definen las variables aleatorias:

$\xi_i(\omega_j) = n^\circ$ de alelos A_i , que tiene el genotipo del individuo ω_j en los cromosomas X.

Supuesto que se cumple la ley de H-W, la media y varianza de $\xi_i(\omega_j)$ son:

$$E(\xi_i(\omega_j)) = 0 \cdot (0.5p_i + 0.5(1-p_i))^2 + 1 \cdot (0.5 \cdot p_i + 0.5 \cdot 2 \cdot p_i(1-p_i)) + 0.5 \cdot 2 \cdot p_i^2 = 1.5 \cdot p_i.$$

Para obtener un estimador insesgado de p_i , se debe usar la variable $\zeta_i(\omega_j) = (2/3) \cdot \xi_i(\omega_j)$. La varianza de $\zeta_i(\omega_j)$ es:

$$V(\zeta_i(\omega_j)) = E(\zeta_i^2(\omega_j)) - (E(\zeta_i))^2 = (2/3)^2 \cdot 0 \cdot (0.5 \cdot p_i + 0.5 \cdot (1-p_i))^2 + (2/3)^2 \cdot 1^2 \cdot (0.5 \cdot p_i + 0.5 \cdot 2 \cdot p_i(1-p_i)) + (2/3)^2 \cdot 2^2 \cdot 0.5 \cdot p_i^2 - p_i^2 = (5/9) \cdot p_i(1-p_i) + (1/9) \cdot p_i.$$

Llamando $\hat{p}_i = \frac{2}{3 \cdot n} \sum_{j=1}^n \xi_i(\omega_j)$, a la probabilidad estimada del alelo A_i ; el intervalo de confianza para dicha probabilidad p_i , calculado de forma aproximada por el teorema central del límite, es:

$$p_i \in \left[\hat{p}_i - n(\alpha/2) \sqrt{\frac{\frac{5}{9} \hat{p}_i (1 - \hat{p}_i) + \frac{1}{9} \hat{p}_i}{n}}; \dots; \hat{p}_i + n(\alpha/2) \sqrt{\frac{\frac{5}{9} \hat{p}_i (1 - \hat{p}_i) + \frac{1}{9} \hat{p}_i}{n}} \right]$$

En los tres supuestos $n(\alpha/2)$ es tal que $P(\zeta \geq n(\alpha/2)) = \alpha/2$ y ζ es la variable que se distribuye como una normal de media 0 y desviación típica 1.

3-2. PARÁMETROS FORENSES.

1) INDICE DE HETEROZIGOSIDAD(H).

Sólo tiene sentido en la población de mujeres. Teóricamente el índice de heterocigosidad H es la probabilidad de que una mujer, elegida aleatoriamente de la población de referencia Ω_M ; sea heterocigótica. Por tanto H es:

$$H = 1 - \sum_{i=1}^k p_i^2 = 1 - \sum_{i=1}^k p_{ii}; \text{ que se estima por } \hat{H} = \frac{nH}{nm} = 1 - \sum_{i=1}^k \frac{n_{ii}}{nm};$$

siendo nH el número de mujeres heterocigóticas en la muestra de tamaño nm; y n_{ii} es el número de mujeres homocigóticas, en la muestra, con el genotipo $A_i:A_i$.

2) DIVERSIDAD ALÉLICA.

Se suelen usar dos medidas para la diversidad alélica; según se utilice la distancia euclídea L_2 ó la de Chebycheff L_1 , dando lugar a las fórmulas:

$$DA_2 = h_2 = 1 - \frac{1}{k} \sum_{i=1}^k (p_i - \frac{1}{k})^2 \quad \text{y} \quad DA_1 = h_1 = 1 - \frac{1}{k} \sum_{i=1}^k \left| p_i - \frac{1}{k} \right|$$

Que se estiman por:

$$D\hat{A}_2 = \hat{h}_2 = 1 - \frac{1}{k} \sum_{i=1}^k (\hat{p}_i - \frac{1}{k})^2 \quad \text{y} \quad D\hat{A}_1 = \hat{h}_1 = 1 - \frac{1}{k} \sum_{i=1}^k \left| \hat{p}_i - \frac{1}{k} \right|.$$

Entre más diversidad alélica haya, mayor será la heterocigosidad. Cuando $DA_2 = DA_1 = 1$, la heterocigosidad vale $H = 1 - (1/k)$.

3) PODER DE DISCRIMINACIÓN.

El poder de discriminación es la probabilidad de que dos individuos, tomados aleatoriamente de la población de referencia, se puedan distinguir genéticamente. El poder de discriminación se calcula para cada STR, en el supuesto de que se cumpla la ley de H-W. La hipótesis de independencia entre diferentes STR, permite calcular el poder de discriminación para diferentes sistemas genéticos.

CASO 1º. LA POBLACIÓN DE REFERENCIA ES SÓLO DE HOMBRES Ω_H .

Si se toman aleatoriamente dos individuos ω_1 y $\omega_2 \in \Omega_H$, la probabilidad de que ambos tengan el mismo alelo en un STR del cromosoma X; es:

$$PND = \sum_{j=1}^k p_j^2; \text{ que se estima por } P\hat{N}D = \sum_{j=1}^k \hat{p}_j^2;$$

El poder de discriminación es $PD=1-PND=1-\sum_{j=1}^k p_j^2$; que se estima por :

$$P\hat{D} = 1 - P\hat{N}D = 1 - \sum_{j=1}^k \hat{p}_j^2 .$$

CASO 2º. LA POBLACION DE REFERENCIA ES SÓLO DE MUJERES Ω_M .

Cuando se toman aleatoriamente dos mujeres ω_1 y $\omega_2 \in \Omega_M$, los casos posibles en que no se pueden distinguir; ó de falta de discriminación son:

MUJER 1= ω_1 ; MUJER 2= ω_2 ; MUJER 1= ω_1 ; MUJER 2= ω_2 .

GENOTIPOS: $A_i:A_i$ $A_i:A_i$ $A_i:A_j$ $i \neq j$ $A_j:A_i$ ó $A_i:A_j$.

Las probabilidades teóricas se calculan por las siguientes fórmulas:

$$PND = \sum_{j=1}^k p_j^2 + 2 \cdot \sum_{i=1}^k \sum_{j=1, j \neq i}^k p_i^2 \cdot p_j^2 = - \sum_{j=1}^k p_j^2 + 2 \cdot \sum_{i=1}^k \sum_{j=1}^k p_i^2 \cdot p_j^2 = 2 \cdot \sum_{j=1}^k (p_j^2)^2 - \sum_{j=1}^k p_j^4 .$$

Como consecuencia el poder de discriminación es;

$$PD = 1 - NPD = 1 - 2 \cdot \sum_{j=1}^k (p_j^2)^2 + \sum_{j=1}^k p_j^4 .$$

PND y PD se estiman al sustituir en las fórmulas las probabilidades por sus estimadores.

CASO 3º. LA POBLACION ES DE HOMBRES Y MUJERES Ω .

Cuando se toman aleatoriamente dos elementos ω_1 y $\omega_2 \in \Omega_M$, los casos posibles en que no se pueden distinguir los dos individuos; ó de falta de discriminación son:

MUJER 1= ω_1 ; MUJER 2= ω_1 ; MUJER 1= ω_1 ; MUJER 2= ω_2 .

GENOTIPOS: $A_i:A_i$ $A_i:A_i$ $A_i:A_j$ $i \neq j$ $A_j:A_i$ ó $A_i:A_j$.

HOMB1= ω_1 ; HOMB2= ω_2 ; HOMB= ω_1 ; MUJER= ω_2 ; MUJER= ω_1 ; HOMB= ω_2 .

GENOTIPOS: A_j A_j A_j $A_j:A_j$ $A_j:A_j$ A_j .

Por tanto, como consecuencia de las técnicas de medida, un hombre no se distingue genéticamente de una mujer, cuando la mujer es homocigótica y tiene el mismo alelo que el del hombre.

Las probabilidades teóricas de falta de discriminación PND y el poder discriminante PD, suponiendo que la probabilidad de los alelos en el hombre es igual a la probabilidad de los alelos en la mujer, se calculan por las siguientes fórmulas:

$$PND = 0.25 \left(\sum_{j=1}^k p_j^2 + 2 \cdot \sum_{i=1}^k \sum_{j=1, j \neq i}^k p_i^2 \cdot p_j^2 + \sum_{j=1}^k p_j^2 + 2 \cdot \sum_{j=1}^k p_j^3 \right) =$$

$$= 0.25 \left(\sum_{j=1}^k p_j^2 + 2 \cdot \sum_{j=1}^k p_j^3 + 2 \cdot \left(\sum_{j=1}^k p_j^2 \right)^2 - \sum_{j=1}^k p_j^4 \right)$$

Como consecuencia el poder de discriminación es;

$$PD = 1 - NPD = 1 - 0.25 \left(\sum_{j=1}^k p_j^2 + 2 \cdot \sum_{j=1}^k p_j^3 + 2 \cdot \left(\sum_{j=1}^k p_j^2 \right)^2 - \sum_{j=1}^k p_j^4 \right) .$$

PND y PD se estiman al sustituir en las fórmulas las probabilidades por sus estimadores.

3-4 PODER DE EXCLUSIÓN.

El poder de exclusión es la probabilidad de inconsistencia genética entre presuntos padres e hijos. El complemento a 1 de la exclusión, es la consistencia ó incapacidad de exclusión. Se estudia la exclusión bajo diversos supuestos.

CASO 1. EXCLUSIÓN ENTRE PRESUNTO PADRE E HIJA.

Se eligen aleatoriamente de la población de referencia un presunto padre y una hija. Se analiza un sistema STR ligado al sexo. Por ello el padre tiene un solo alelo en el cromosoma X, mientras que la hija tiene dos alelos.

La consistencia ó compatibilidad entre padre e hija, se da en los casos siguientes:

	PADRE	HIJA	PADRE	HIJA.
GENOTIPOS	A_i	$A_i:A_i$	A_i	$A_i:A_j$ ó $A_j:A_i$.

La probabilidad de consistencia, en este supuesto, es:

$$PC = \sum_{j=1}^k p_j^3 + 2 \cdot \sum_{j=1}^k p_j^2 (1 - p_j) = 2 \cdot \sum_{j=1}^k p_j^2 - \sum_{j=1}^k p_j^3.$$

Como consecuencia, el poder de exclusión es :

$$PE = 1 - PC = 1 - 2 \cdot \sum_{j=1}^k p_j^2 + \sum_{j=1}^k p_j^3.$$

Las estimaciones de PE y PD se harían por métodos similares a los expuestos en los párrafos correspondientes al poder de discriminación.

CASO 2°. COMPATIBILIDAD ENTRE UNA PRESUNTA MADRE Y UNA HIJA.

Se eligen aleatoriamente de la población de referencia una presunta madre y una hija. Se analiza un sistema STR ligado al sexo. En él ,tanto la madre, como la hija, tienen dos alelos, uno en cada uno de sus dos cromosomas X.

La consistencia entre madre e hija, se produce en los siguientes casos:

	MADRE	HIJA	MADRE	HIJA.
GENOTIPOS	$A_i : A_i$	$A_i:A_i$ $A_i:A_j$ ó $A_j:A_i$::: $A_i:A_j$ $i \neq j$	$A_i : A_i$	$A_i:A_i$; $A_j:A_j$; $A_i:A_j$ ó $A_j:A_i$.
	MADRE	HIJA		
GENOTIPOS:	$A_i:A_j$ $i \neq j$	$A_i:A_u$; $A_u:A_i$; $A_j:A_u$; $A_u:A_j$;		para $i \neq j \neq u \neq i$.

La probabilidad de consistencia ó compatibilidad, en este caso, se obtiene al evaluar los casos posibles anteriormente descritos, por la fórmula,:

$$\begin{aligned} PC &= \sum_{j=1}^k p_j^4 + 4 \sum_{j=1}^k p_j^3 (1 - p_j) + 2 \cdot \sum_{i=1}^k \sum_{j=1, j \neq i}^k p_i^2 \cdot p_j^2 + 4 \cdot \sum_{i=1}^k \sum_{j=1, j \neq i}^k p_i^2 \cdot p_j (1 - p_i - p_j) = \\ &= \sum_{j=1}^k p_j^4 + 4 \cdot \sum_{j=1}^k p_j^2 (1 - p_j) + 4 \sum_{j=1}^k p_j^3 (1 - p_j) - 4 \sum_{j=1}^k p_j^3 (1 - p_j) - 2 \cdot \sum_{i=1}^k \sum_{j=1, j \neq i}^k p_i^2 \cdot p_j^2 = \\ &= 4 \cdot \sum_{j=1}^k p_j^2 - 4 \cdot \sum_{j=1}^k p_j^3 + 3 \cdot \sum_{j=1}^k p_j^4 - 2 \cdot \sum_{i=1}^k \sum_{j=1}^k p_i^2 \cdot p_j^2. \end{aligned}$$

Por tanto; la probabilidad de inconsistencia entre la presunta madre y la hija es:

$$PE=1-4 \cdot \sum_{j=1}^k p_j^2 + 4 \cdot \sum_{j=1}^k p_j^3 - 3 \cdot \sum_{j=1}^k p_j^4 + 2 \cdot \sum_{i=1}^k \sum_{j=1}^k p_i^2 \cdot p_j^2.$$

Esta probabilidad de exclusión se debe entender respecto a haber elegido aleatoriamente la madre y la hija, y que ambas sean incompatibles, exclusivas ó inconsistentes. Las estimaciones se harían por técnicas similares a las descritas en los epígrafes precedentes.

CASO 2º. COMPATIBILIDAD ENTRE UNA PRESUNTA MADRE Y UNA HIJA.

Se eligen aleatoriamente, de la población de referencia, una presunta madre y una presunta hija, compatible con la madre. Después se elige, aleatoriamente, un presunto padre. Se analiza un sistema STR ligado al sexo. En él, tanto la madre, como la hija, tienen dos alelos, uno en cada uno de sus dos cromosomas X; mientras que el padre sólo tiene uno.

La consistencia entre madre, padre e hija, se produce en los siguientes casos:

	MADRE	HIJA	PADRE	:::	MADRE	HIJA	PADRE
GENOTIPOS	A_i	A_i	A_i	:::	A_i	A_j	A_i
		$A_i:A_i$	A_j		$A_i:A_j$	$A_i:A_i$	A_j
		$A_i:A_j$	A_j		$A_i:A_j$	$A_j:A_j$	A_j
		ó $A_j:A_i$	A_j		$A_i:A_u$	$A_u:A_i$	A_u
	MADRE	HIJA	PADRE	MADRE	HIJA	PADRE	
GENOTIPOS:	$A_i:A_j$	$i \neq j$	$A_i:A_j$	ó	$A_j:A_i$	A_i	ó
	A_j	A_j	A_j	ó	A_j	$A_i:A_j$	ó
	$A_i:A_j$	$A_j:A_u$	A_j	ó	$A_u:A_j$	A_u	A_u

La probabilidad de consistencia ó compatibilidad en este caso, se calcula por las fórmula, que se obtiene al evaluar los casos posibles, anteriormente descritos:

$$PC = \sum_{j=1}^k p_j^5 + 2 \sum_{j=1}^k p_j^4 (1 - p_j) + 6 \cdot \sum_{i=1}^k \sum_{j=1, j \neq i}^k p_i^3 \cdot p_j^2 + 4 \cdot \sum_{i=1}^k \sum_{j=1, j \neq i}^k p_i^2 \cdot p_j (1 - p_i - p_j) =$$

$$= 4 \cdot \sum_{i=1}^k \sum_{j=1}^k p_i^2 \cdot p_j^2 - 2 \cdot \sum_{j=1}^k p_j^4 - 2 \cdot \left(\sum_{i=1}^k p_i^3 \right) \left(\sum_{j=1}^k p_j^2 \right) + \sum_{j=1}^k p_j^5.$$

Por tanto; la probabilidad de inconsistencia entre el presunto padre, la presunta madre y la hija es:

$$PE=1-4 \cdot \sum_{i=1}^k \sum_{j=1}^k p_i^2 \cdot p_j^2 + 2 \cdot \sum_{j=1}^k p_j^4 + 2 \cdot \left(\sum_{i=1}^k p_i^3 \right) \left(\sum_{j=1}^k p_j^2 \right) - \sum_{j=1}^k p_j^5.$$

Si se denomina CMH al suceso que representa la consistencia entre madre e hija, y CMHP al suceso consistencia entre madre, hija y padre; la inconsistencia ó exclusión del padre, con una madre y una hija consistentes se representa por el suceso CMH-CMHP.

Con estas reglas, la probabilidad de este suceso es:

$$P(\text{CMH-CMHP}) = P(\text{CMH}) - P(\text{CMHP}) = 4 \cdot \sum_{j=1}^k p_j^2 - 4 \cdot \sum_{j=1}^k p_j^3 + 3 \cdot \sum_{j=1}^k p_j^4 - 2 \cdot \sum_{i=1}^k \sum_{j=1}^k p_i^2 \cdot p_j^2 -$$

$$- 4 \cdot \sum_{i=1}^k \sum_{j=1}^k p_i^2 \cdot p_j^2 + 2 \cdot \sum_{j=1}^k p_j^4 + 2 \cdot \left(\sum_{i=1}^k p_i^3 \right) \left(\sum_{j=1}^k p_j^2 \right) - \sum_{j=1}^k p_j^5 = 4 \cdot \sum_{j=1}^k p_j^2 - 4 \cdot \sum_{j=1}^k p_j^3 + 5 \cdot \sum_{j=1}^k p_j^4$$

$$- 6 \cdot \sum_{i=1}^k \sum_{j=1}^k p_i^2 \cdot p_j^2 + 2 \cdot \left(\sum_{i=1}^k p_i^3 \right) \left(\sum_{j=1}^k p_j^2 \right) - \sum_{j=1}^k p_j^5.$$

La probabilidad del suceso condicionado, P(CMHP/CMH), se calcula por la fórmula:

$$P(\text{CMHP/CMH}) = 1 - (4 \cdot \sum_{i=1}^k \sum_{j=1}^k p_i^2 \cdot p_j^2 - 2 \cdot \sum_{j=1}^k p_j^4 - 2 \cdot (\sum_{i=1}^k p_i^3)(\sum_{j=1}^k p_j^2) + \sum_{j=1}^k p_j^5) / (4 \cdot \sum_{j=1}^k p_j^2 - 4 \cdot \sum_{j=1}^k p_j^3 + 3 \cdot \sum_{j=1}^k p_j^4 - 2 \cdot \sum_{i=1}^k \sum_{j=1}^k p_i^2 \cdot p_j^2)$$

3-4 EXCLUSIÓN CUANDO SE TIP A LA MADRE Y A LA HIJA.

Cuando se puede observar ó tipar el STR del ADN de la madre y de la hija, el cálculo de la probabilidad de exclusión queda condicionado al resultado del proceso de tipado, dando lugar a los siguientes resultados:

GENOTIPO	MADRE	HIJA	PROBABILIDAD DE EXCLUSIÓN.
	$A_i:A_i$	$A_i:A_i$	$PE=1-p_i$
	$A_i:A_i$	$i \neq j$	$PE=1-p_j$
	$A_i:A_j$	$i \neq j$	$PE=1-p_i \cdot p_j$
	$A_i:A_j$	$i \neq j \neq u \neq i$	$PE=1-p_u$
	$A_i:A_j$	$i \neq j \neq u \neq i$	$PE=1-p_u$
	$A_i:A_j$	$A_j:A_u$	$PE=1-p_u$

En el espacio de probabilidad total, el poder de exclusión medio es:

$$PEM = \sum_{j=1}^k (1-p_j) \cdot \left[p_j^4 + p_j \cdot \sum_{i=1, i \neq j}^k p_i^3 + 2 \cdot p_j \cdot \sum_{i=1, i \neq j}^k p_i^2 \cdot (1-p_i-p_j) \right] + \sum_{i=1}^k \sum_{j=1, j \neq i}^k (1-p_i-p_j) \cdot p_i^2 \cdot p_j^2$$

En el espacio de probabilidad condicionado a la consistencia entre madre e hija, el poder de exclusión medio condicionado es:

$$PEMC = \sum_{i=1}^k p_i^3 \cdot \left[p_i(1-p_i) + 2 \cdot \sum_{j=1, j \neq i}^k p_j(1-p_j) \right] / (p_i(2-p_i) + \sum_{i=1}^k \sum_{j=1, j \neq i}^k p_i \cdot p_j \cdot \left[2 \cdot p_i \cdot p_j(1-p_i-p_j) + 4 \cdot p_i \cdot \sum_{u=1, u \neq i, j}^k p_u \cdot (1-p_u) \right]) / (2 \cdot p_i(1-p_i-p_j) + 2p_j)$$

Las probabilidades de consistencia entre madre e hija son:

$$P(\text{Cons Hija/Madre } A_i:A_i) = P(A_i:A_i; A_i:A_j; A_j:A_i) = p_i^2 + 2 \cdot p_i \cdot \sum_{j=1}^k p_j = p_i(2-p_j)$$

$$P(C H/MA_i:A_j; j \neq i) = P(A_i:A_j; A_j:A_u; A_i:A_u; j \neq i \neq u \neq j) = 2 \cdot (p_i(1-p_i-p_j) + p_j(1-p_j))$$

4 COMENTARIOS SOBRE ESTIMACIONES.

La forma más fácil de estimar los parámetros descritos anteriormente consiste en sustituir las probabilidades teóricas p_i por sus estimadores obtenidos de las muestras; esto es, sustituir p_i por \hat{p}_i .

Si se sigue la técnica de estimación precedente, los intervalos de confianza se forman con los valores mínimo y máximo de las funciones de probabilidad que definen los correspondientes conceptos, definidas sobre los intervalos de confianza asociados a las respectivas probabilidades.

No obstante, también se pueden estimar directamente las probabilidades de los parámetros. Para ello, se toma una muestra aleatoria simple de parejas de la población de referencia:

$$M = \{ (\omega_{11}, \omega_{12}), (\omega_{21}, \omega_{22}), (\omega_{31}, \omega_{32}), \dots, (\omega_{n1}, \omega_{n2}) \}$$

Si en la muestra existen k parejas que tienen los dos alelos iguales; entonces el estimador del poder discriminante es:

$$PD = 1 - k/n = (n - k)/n.$$

Para la consistencia entre padre, madre e hija, se deberían tomar muestras de tres elementos de la población de referencia, y observar el número de triples consistentes $n(MHP)$; así como el número de parejas consistentes entre madre e hija $n(MH)$.

El poder de exclusión condicionado a la consistencia entre madre e hija se mide por la fórmula: $PEC = 1 - n(MHP)/n(MH)$; mientras que el poder de exclusión absoluto se mediría por la fórmula: $PE = 1 - n(MHP)/n(MH)$.

Con este procedimiento, los intervalos de confianza se calcularían aplicando el teorema central del límite a la variable binomial.

BIBLIOGRAFIA.

- (1) Bruce S, Weir (1996). *“Genetic Data Analysis II: Methods for discrete Population Genetic Data”*. Sinauer Associates, Inc. Publishers; Sunderland, Massachusetts.
- (2) Desmarais D.; Zhong Y. y otros. (1998) *“Development of a Highly Polimorphic STR Marker for Identity Testing Purposes at the Human Androgene Receptor Gene”* J. Forensic Sci. Vol 43; N° 5; pp 1046, 1049.