

# Desafío Teórico Avanzado ISFG-GHEP 2026: Parentesco

Diseñado por: Magnus Dehli Vigeland

## Introducción

Esta es una prueba de opción múltiple que consta de 20 preguntas. Para cada pregunta, sólo una alternativa es correcta. Puede usar el software que desee, pero tenga en cuenta que algunos programas poseen convenciones integradas (por ejemplo, redondeo) que pueden afectar al resultado. Si su respuesta no coincide exactamente con ninguna de las opciones, elija la más cercana.

El tema del desafío de este año es la endogamia (*inbreeding*): la estimación de coeficientes de endogamia a partir de datos genotípicos y el efecto de la endogamia sobre los LR de parentesco en casos de incesto.

## Supuestos generales

- No hay desviaciones del equilibrio de Hardy-Weinberg.
- Todos los marcadores son independientes (*unlinked*, no ligados) y no presentan desequilibrio de ligamiento (*linkage disequilibrium*).
- No hay *drop-outs*, *drop-ins*, alelos silentes ni mutaciones.
- 

## Archivos necesarios

- `db35.txt`: Base de datos de frecuencias para 35 marcadores STR
- `mother_child.txt`: Genotipos de la madre y del hijo
- `father.txt`: Genotipos del padre

## Algunas definiciones

Los alelos homólogos son idénticos por descendencia (IBD, *identical by descent*) si tienen el mismo origen dentro de un pedigrí determinado. El *coeficiente de endogamia*  $f$  de un miembro del pedigrí es la probabilidad de que, en un *locus* autosómico aleatorio, los dos alelos sean IBD. El coeficiente de endogamia de un hijo es igual al *coeficiente de parentesco*  $\varphi$  entre los padres:  $f_{\text{child}} = \varphi_{\text{parents}}$ .

Los coeficientes  $\kappa$  ( $\kappa_0, \kappa_1, \kappa_2$ ) entre dos individuos no consanguíneos son las probabilidades de que compartan 0, 1 o 2 alelos IBD, respectivamente, en un *locus* autosómico aleatorio. Por ejemplo, un par de hermanos completos tiene  $(\kappa_0, \kappa_1, \kappa_2) = (0.25, 0.5, 0.25)$ . Los coeficientes  $\kappa$  están relacionados con el coeficiente de parentesco mediante la fórmula  $\varphi = \kappa_1/4 + \kappa_2/2$ .

## Lecturas recomendadas

- Vigeland, 2024. *Measures of relatedness*. (Apuntes del curso)
- Ritland, 1996. *Estimators for pairwise relatedness and individual inbreeding coefficients*. *Genetical Research*, 67(2), 175-185.

**BUENA SUERTE!**

## Prólogo

Al detective Vargas ya no le gustaba este caso. Había empezado mal, y probablemente terminaría peor. Sentía lástima por la joven sentada frente a él, con un niño dormido en un cochecito a su lado.

—No hay padre —dijo ella, tirando de la manga para cubrir un hematoma que ya se estaba desvaneciendo—. Mi hijo no necesita padre. Me tiene a mí.

Era evidente que estaba huyendo de alguien. Pero ¿de quién? Cuando Vargas preguntó por su familia, ella se cerró en silencio, con la mirada fija en el cochecito. Protectora, pero también asustada.

El detective Vargas volvió a revisar el informe genético. Ella era la madre del niño, de eso no había ninguna duda. Pero había algo en los datos que lo confundía. Vargas no era especialista, pero el perfil del niño parecía demasiado simple. Demasiados números repetidos. En su mente empezó a tomar forma una idea. Si la mujer no quería identificar al padre, bien. Él ya estaba ahí, escrito en el ADN del niño. La mitad del perfil, sino algo más...?

Necesitaba a alguien capaz de exprimir secretos del ADN. ¿Qué tal aquel experto que lo ayudó el año pasado? Con una sonrisa seca, levantó el teléfono...

## Ejercicios

Considere un marcador STR con  $m$  alelos y frecuencias  $p_1, \dots, p_m$ , tales que  $p_1 + \dots + p_m = 1$ .

1. Sea  $H$  la *homocigosidad esperada* para este marcador, es decir, la probabilidad de que un individuo aleatorio, no consanguíneo, sea homocigoto. Entonces  $H$  es igual a
  - a)  $p_1^2 + \dots + p_m^2$
  - b)  $1 - p_1^2 - \dots - p_m^2$
  - c)  $(1 - p_1)^2 + \dots + (1 - p_m)^2$
  - d)  $1 - (1 - p_1)^2 - \dots - (1 - p_m)^2$
  - e)  $1/m$
2. Supongamos que se genotipa para ese marcador a un niño con un coeficiente de endogamia  $f$ . La probabilidad de que el niño sea homocigoto es
  - a)  $fH$
  - b)  $f + (1 - f)H$
  - c)  $f(1 - H) + (1 - f)H$
  - d)  $(1 - f)H$
  - e)  $1 - fH$
3. Supongamos que se genotipa al niño para  $L$  marcadores independientes, con homocigosidades esperadas  $H_1, \dots, H_L$ , respectivamente. Entonces, el número esperado de genotipos homocigotos es
  - a)  $f(H_1 + \dots + H_L)$
  - b)  $(1 - f)(H_1 + \dots + H_L)$
  - c)  $fL + (1 - f)(H_1 + \dots + H_L)$
  - d)  $fL + (1 - 2f)(H_1 + \dots + H_L)$
  - e)  $\frac{L}{m}(H_1 + \dots + H_L)$

Sea  $H^*$  la suma  $H^* = H_1 + \dots + H_L$ . A partir de las preguntas anteriores se puede demostrar que un estimador del coeficiente de endogamia  $f$  mediante el *método de los momentos* (*method-of-moments*) viene dado por

$$\hat{f}_1 = \frac{N_{\text{hom}} - H^*}{L - H^*},$$

donde  $N_{\text{hom}}$  es el número observado de marcadores homocigotos en el niño.

Para los siguientes ejercicios, use los datos proporcionados en el archivo `mother_child.txt` and `db35.txt`.

4. El estimador  $\hat{f}_1$  da un coeficiente de endogamia estimado para el niño de aproximadamente
- a) 0.10
  - b) 0.14
  - c) 0.17
  - d) 0.21
  - e) 0.28

Otro estimador del coeficiente de endogamia ampliamente utilizado es el estimador de *Ritland* (Ritland, 1996), definido por

$$\hat{f}_2 = \frac{(T_1 + \dots + T_L) - L}{(m_1 + \dots + m_L) - L}.$$

Aquí  $m_j$  es el número de alelos en el marcador  $j$ , y  $T_j$  se define como  $1/p_a$  si el genotipo del marcador  $j$  es  $a/a$ , y como 0 en caso contrario.

5. El estimador de Ritland proporciona un coeficiente de endogamia estimado para el niño de aproximadamente
- a) -0.08
  - b) 0.14
  - c) 0.16
  - d) 0.22
  - e) 0.28
6. El estimador de Ritland es muy sensible a los alelos raros. Supongamos (¡solo para este ejercicio!) que el genotipo del niño para el marcador *PentaD* cambia de 7/9 a 7/7. Entonces, la estimación del coeficiente de endogamia mediante el estimador de Ritland sería aproximadamente
- a) -0.08
  - b) 0.18
  - c) 0.38
  - d) 0.58
  - e) 0.78

Cuando se dispone de datos genotípicos del niño y de uno de los progenitores, es posible aplicar otros estimadores  $f_{\text{child}}$ . Se puede demostrar que, para cualquier marcador dado, tenemos probabilidades iguales

$$P(\text{madre e hijo tengan el mismo genotipo}) = P(\text{hijo sea homocigoto}).$$

(Esta curiosa igualdad es interesante por sí misma, ¡y le animo a que intente demostrarla!). Esto conduce a un tercer estimador de  $f_{\text{child}}$ , reemplazando  $N_{\text{hom}}$  en la ecuación de  $\hat{f}_1$ , por el número  $N_{\text{eq}}$  de marcadores en los que la madre y el niño tienen genotipos idénticos:

$$\hat{f}_3 = \frac{N_{\text{eq}} - H^*}{L - H^*}.$$

7. El estimador  $\hat{f}_3$  da un coeficiente de endogamia estimado para el niño de aproximadamente

- a) 0.14
- b) 0.17
- c) 0.21
- d) 0.24
- e) 0.28

Como  $\hat{f}_3$  se basa en las mismas probabilidades que  $\hat{f}_1$ , generalmente tienen propiedades similares. Sin embargo, se puede obtener un estimador más preciso combinándolos, definido por

$$\hat{f}_4 = \frac{\hat{f}_1 + \hat{f}_3}{2}.$$

8. Usando  $\hat{f}_4$ , el coeficiente de endogamia estimado para el niño es aproximadamente de

- a) 0.15
- b) 0.17
- c) 0.19
- d) 0.21
- e) Ninguno de los anteriores

Las gráficas de la Figura 1 muestran las distribuciones de los cuatro estimadores  $\hat{f}_1, \hat{f}_2, \hat{f}_3, \hat{f}_4$ , obtenidas mediante la simulación de 10,000 pares madre-hijo en dos escenarios: cuando los progenitores no están emparentados, y cuando los progenitores son hermanos completos.

9. Basándose en la Figura 1, ¿cuál de estas afirmaciones **no** es correcta?

- a) En el escenario de ausencia de consanguinidad (*outbred*), el estimador de Ritland  $\hat{f}_2$  tiene la mayor precisión.
- b) En el escenario de hermanos completos, el estimador de Ritland  $\hat{f}_2$  es el menos preciso.
- c) Los estimadores de momentos  $\hat{f}_1$  y  $\hat{f}_3$  son casi idénticos en su distribución, en ambos escenarios.
- d) En el escenario de hermanos completos, el estimador combinado  $\hat{f}_4$  tiene la mayor precisión.

- e) La altura de los picos de las distribuciones determina la precisión de los estimadores.

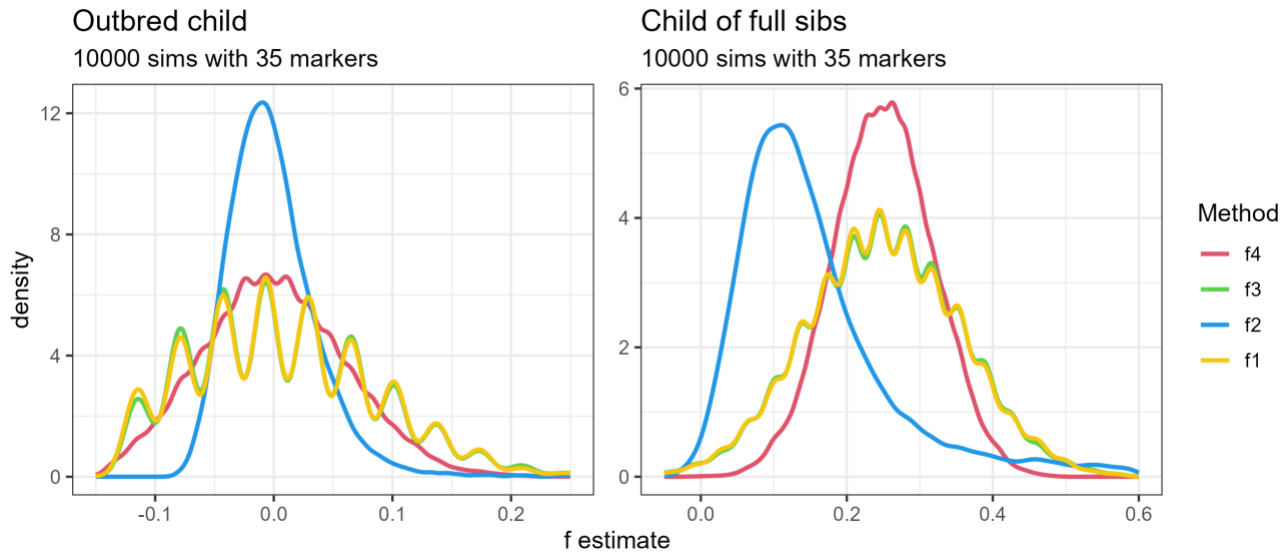


Figura 1: Distribuciones de las estimaciones del coeficiente de endogamia basadas en 10 000 simulaciones para cada escenario.

Considere las siguientes hipótesis para el coeficiente de endogamia del niño:

$$\begin{aligned} H_0: & \quad f = 0 && \text{(ausencia de endogamia)} \\ H_A: & \quad f > 0 && \text{(endogamia).} \end{aligned}$$

10. Sea  $c$  la estima observada del estimador  $\hat{f}_4$ , es decir, el valor obtenido en el Ejercicio 8. En 10 000 simulaciones con progenitores no consanguíneos, 9 dieron como resultado  $\hat{f}_4 \geq c$ . De manera similar, de las 10 000 simulaciones con progenitores hermanos completos, 7706 dieron  $\hat{f}_4 \geq c$ . Con base a esto, un  $p$ -valor aproximado para probar la hipótesis nula  $H_0$  frente a la alternativa  $H_A$ , es

- a) 0.00045
- b) 0.0009
- c) 0.3853
- d) 0.7706
- e) Ninguno de los anteriores

Ahora pasamos a un análisis de los datos basado valores de LR. Considere las siguientes hipótesis:

- $U$ : Los progenitores no están relacionados
- $H$ : Los progenitores son medio hermanos
- $S$ : Los progenitores son hermanos completos
- $P$ : Los progenitores son padre-hija

11. Usando únicamente los datos del niño, el LR que compara la hipótesis  $S$  con la hipótesis  $U$ , es aproximadamente

- a) 1.02
- b) 38
- c) 43
- d) 44
- e) 3076

12. Usando únicamente los datos del niño, los LR de las hipótesis satisfacen

- a)  $1 < LR_{H:U} < LR_{S:U} < LR_{P:U}$
- b)  $1 < LR_{H:U} < LR_{P:U} < LR_{S:U}$
- c)  $1 < LR_{H:U} < LR_{S:U} = LR_{P:U}$
- d)  $1 < LR_{S:U} = LR_{P:U} < LR_{H:U}$
- e)  $1 < LR_{P:U} < LR_{S:U} < LR_{H:U}$

13. Usando los datos de la madre y del niño, el LR que compara  $P$  con  $U$  es aproximadamente

- a) 1.02
- b) 38
- c) 156
- d) 3076
- e) 48 335

14. Usando los datos tanto de la madre como del niño, los LR satisfacen

- a)  $1 < LR_{H:U} < LR_{S:U} < LR_{P:U}$
- b)  $1 < LR_{H:U} < LR_{P:U} < LR_{S:U}$
- c)  $1 < LR_{H:U} < LR_{S:U} = LR_{P:U}$
- d)  $1 < LR_{S:U} = LR_{P:U} < LR_{H:U}$
- e)  $1 < LR_{P:U} < LR_{S:U} < LR_{H:U}$

Algún tiempo después, se arresta a un hombre tras un robo en el edificio donde vive la madre. Ella continuó sin querer hablar, pero se tomó una muestra del hombre y se genotipó. El resultado arrojó una conclusión clara: él era el padre biológico del niño. Pero, ¿cuál era su parentesco con la madre?

Para los siguientes ejercicios, utilice los datos anteriores junto con los genotipos del padre del archivo *father.txt*.

15. Usando todos los datos disponibles, el número de marcadores incompatibles con la hipótesis  $P$  es

- a) 0
- b) 1
- c) 2
- d) 3
- e) Ninguno de los anteriores

16. Usando todos los datos disponibles, los LRs satisfacen

- a)  $1 < LR_{H:U} < LR_{S:U} < LR_{P:U}$
- b)  $1 < LR_{H:U} < LR_{P:U} < LR_{S:U}$
- c)  $LR_{H:U} < LR_{P:U} < 1 < LR_{S:U}$
- d)  $LR_{P:U} < LR_{H:U} < 1 < LR_{S:U}$
- e)  $LR_{P:U} < 1 < LR_{H:U} < LR_{S:U}$

17. El LR comparando la hipótesis más probable con la hipótesis  $U$ , es aproximadamente

- a)  $8.5 \cdot 10^9$
- b)  $3.1 \cdot 10^6$
- c)  $1.9 \cdot 10^6$
- d)  $7.0 \cdot 10^5$
- e)  $2.3 \cdot 10^3$

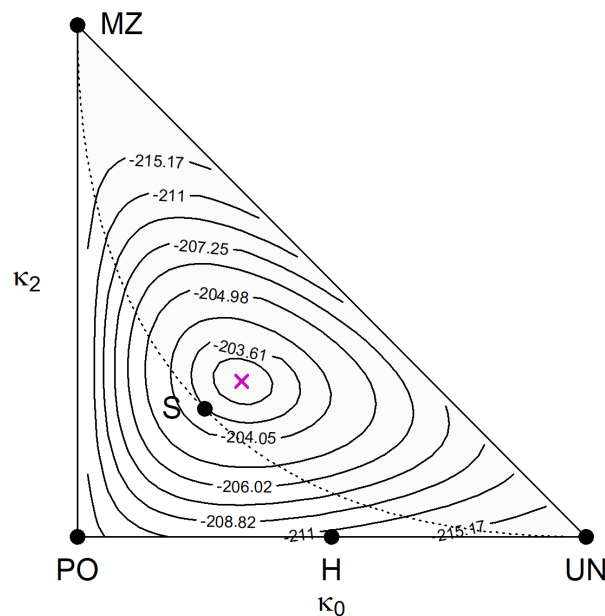


Figura 2: Gráfico de contorno de la función log-verosimilitud para la relación entre los progenitores del niño. Relaciones indicadas: no emparentados (UN), progenitor-hijo (PO), gemelos idénticos (MZ), hermanos completos (S) y medio-hermanos (H). Cada eje varía de 0 a 1.

18. La Figura 2 muestra contornos de la función log-verosimilitud para la relación entre el padre y la madre, en función de sus coeficientes  $\kappa$  ( $\kappa_0, \kappa_1, \kappa_2$ ). Teniendo en cuenta la figura, la estimación de máxima verosimilitud ( $\hat{\kappa}_0, \hat{\kappa}_1, \hat{\kappa}_2$ ) es aproximadamente

- a) (0.32, 0.38, 0.30)
- b) (0.25, 0.39, 0.36)
- c) (0.52, 0.45, 0.03)
- d) (0.29, 0.53, 0.08)
- e) Ninguna de las anteriores

Recuerde que el coeficiente de parentesco  $\varphi$  entre dos individuos está determinado por sus coeficientes  $\kappa$  mediante la fórmula  $\varphi = \frac{1}{4}\kappa_1 + \frac{1}{2}\kappa_2$ . Dado que el coeficiente de endogamia de un individuo es igual al coeficiente de parentesco entre sus progenitores, esto da lugar a otro estimador del coeficiente de endogamia del hijo:

$$\hat{f}_5 = \frac{1}{4}\hat{\kappa}_1 + \frac{1}{2}\hat{\kappa}_2$$

19. Usando  $\hat{f}_5$ , el coeficiente de endogamia estimado del niño es de aproximadamente

- a) 0.13
- b) 0.17
- c) 0.25
- d) 0.27
- e) 0.28

20. Considere un trio formado por un hijo y ambos progenitores, genotipados para un conjunto de marcadores autosómicos. Sea  $\mathcal{R}$  la relación de parentesco entre los progenitores. ¿Cuál de las siguientes afirmaciones **no** es cierta?

- a) Los datos del niño, por si solos, contienen información sobre  $\mathcal{R}$ .
- b) Los datos del niño y de uno de los progenitores contienen más información sobre  $\mathcal{R}$  que sólo los datos del niño.
- c) Los datos del niño y de uno de los progenitores no pueden distinguir si  $\mathcal{R}$  son hermanos completos o padre-hija.
- d) Considerando los datos del niño y uno de los progenitores, el otro progenitor no proporciona información adicional sobre  $\mathcal{R}$ .
- e) Considerando los datos de ambos progenitores, el niño no proporciona información adicional sobre  $\mathcal{R}$ .

**FIN**

Responsable del desarrollo del ejercicio:

Magnus Dehli Vigeland

[magnusdv@gmail.com](mailto:magnusdv@gmail.com)

Traducción:

Lourdes Prieto ([lourditasmt@gmail.com](mailto:lourditasmt@gmail.com))

Comité Ejecutivo del GHEP-ISFG

[info@ghep-isfg.org](mailto:info@ghep-isfg.org)