

COVARIANZA GENÉTICA ENTRE INDIVIDUOS

Conocimientos previos requeridos

álgebra elemental
leyes de Mendel

Conceptos previos requeridos

probabilidad
varianza
covarianza
parentesco
efecto medio de un alelo
modelo genético infinitesimal

Aplicaciones

AC – mayormente de interés académico

EP – estaciones de prueba, pruebas diseñadas, pruebas con datos de campo

GE – cría y mejoramiento de poblaciones generales

PL – cría y mejoramiento de núcleos o plantales

Interpretación de significado y sesgo de métodos de estimación de parámetros genéticos [**AC**]

COVARIANZA GENÉTICA ENTRE INDIVIDUOS

Similaridad, semejanza o parecido son conceptos intuitivos. Es común observar que individuos genéticamente relacionados se parecen. Sin embargo, esa evidencia es insuficiente para establecer relaciones que permitan predecir tendencias genéticas o fenotípicas, o planificar apareamientos. Para tales fines se requiere una caracterización *cuantitativa* del parecido entre parientes.

Similaridad, semejanza, o parecido pueden evaluarse cuantitativamente a través del cálculo de una covarianza. Matemáticamente, una covarianza mide el grado al cual dos variables varían en forma conjunta (i.e. co-varían: una aumenta o disminuye cuando la otra también aumenta o disminuye). Si tener copias de los mismos genes hace que individuos relacionados se parezcan entre sí, deberían existir covarianzas de signo positivo entre individuos para diversas variables cuantitativas de interés (e.g. peso, altura, % de proteína en la leche, etc.)

Diversos efectos génicos (e.g. aditividad, dominancia, interacciones de diversos tipos) contribuyen a esa covarianza, dependiendo del tipo de parentesco que se trate. En una relación *progenitor-descendiente*, los únicos efectos génicos que podrían generar similaridad son los *aditivos* y las interacciones *aditivo x aditivo*. La razón de ello es simple: los cromosomas homólogos se separan al formarse los gametos. *Los progenitores sólo transmiten alelos, no pares alélicos*. En contraste, en una relación de *hermanos completos* existe una probabilidad distinta de cero que parte de los pares alélicos sean comunes a ambos individuos, i.e. que en ambos hermanos se hayan producido las mismas combinaciones de alelos heredadas de sus padres. Los efectos de *dominancia* e interacciones del tipo *dominante x dominante* o *aditivo x dominante* contribuyen al parecido entre hermanos completos. El parentesco afectará entonces la magnitud de la covarianzas genéticas entre individuos.

En síntesis, *los efectos génicos, ponderados de alguna forma por el grado de parentesco serían los determinantes del parecido entre parientes*. Para una derivación cuantitativa de esas relaciones es conveniente comenzar por el caso más simple, la covarianza entre parientes cuando sólo se consideran dos alelos en un locus.

Modelo de un locus

Considérense dos individuos X e Y con genotipos y efectos génicos:

$$\begin{aligned} X : A_{i_x} A_{j_x} &= \alpha_{i_x} + \alpha_{j_x} + \alpha\alpha_{i_x j_x} \\ Y : A_{i_y} A_{j_y} &= \alpha_{i_y} + \alpha_{j_y} + \alpha\alpha_{i_y j_y} \end{aligned}$$

donde los α_i y α_j representan efectos medios de alelos y los $\alpha\alpha_{ij}$ efectos de pares alélicos.

La covarianza entre los genotipos de X e Y será (Apéndice I):

$$\begin{aligned} \text{cov}(A_{i_x} A_{j_x}, A_{i_y} A_{j_y}) &= E\alpha_{i_x} \alpha_{i_y} + E\alpha_{i_x} \alpha_{j_y} + E\alpha_{j_x} \alpha_{i_y} + E\alpha_{j_x} \alpha_{j_y} + E\alpha_{i_x} \alpha\alpha_{i_y j_y} + \\ &+ E\alpha_{j_x} \alpha\alpha_{i_y j_y} + E\alpha_{i_y} \alpha\alpha_{i_x j_x} + E\alpha_{j_y} \alpha\alpha_{i_x j_x} + E\alpha\alpha_{i_x j_x} \alpha\alpha_{i_y j_y} \end{aligned}$$

Para cada uno de esos términos se requieren expresiones analíticas que sean funciones de los efectos génicos y las relaciones de parentesco. Los primeros 4 términos son productos cruzados entre efectos medios. Su valor es cero, excepto en el caso que se trate de alelos idénticos por descendencia. Por lo tanto, para el primer componente, por ejemplo:

$$E\alpha_{i_x} \alpha_{i_y} = p(A_{i_x} \equiv A_{i_y}) E\alpha_i^2 = \frac{a_{XY}}{2} E\alpha_i^2$$

La suma de los 4 términos similares será:

$$\frac{a_{XY}}{2} E\alpha_i^2 + \frac{a_{XY}}{2} E\alpha_i^2 + \frac{a_{XY}}{2} E\alpha_i^2 + \frac{a_{XY}}{2} E\alpha_i^2 = a_{XY} 2E\alpha_i^2 = a_{XY} \sigma_{01}^2$$

Para el término correspondiente al par alélico vale el mismo razonamiento:

$$E\alpha\alpha_{i_x j_x} \alpha\alpha_{i_y j_y} = p(A_{i_x} A_{j_x} \equiv A_{i_y} A_{j_y}) E\alpha\alpha_{ij}^2 = d_{XY} E\alpha\alpha_i^2 = d_{XY} \sigma_{01}^2$$

El resto de los términos son iguales a 0 debido a la definición de efectos medios y efectos de pares alélicos y entonces, cuando se considera un solo locus, la covarianza entre los genotipos de X e Y será:

$$\text{cov}(A_{i_x} A_{j_x}, A_{i_y} A_{j_y}) = a_{xy} \sigma_{10}^2 + d_{xy} \sigma_{01}^2$$

Modelo de dos loci

Sean los individuos X e Y con genotipos:

$$\begin{aligned} X : A_{i_x} A_{j_x} B_{k_x} B_{l_x} &= \alpha_{i_x} + \alpha_{j_x} + \beta_{k_x} + \beta_{l_x} + \alpha\alpha_{i_x j_x} + \beta\beta_{k_x l_x} + \alpha_{i_x} \beta_{k_x} + \alpha_{i_x} \beta_{l_x} + \alpha_{j_x} \beta_{k_x} + \\ &+ \alpha_{j_x} \beta_{l_x} + \alpha_{i_x} \beta\beta_{k_x l_x} + \alpha_{j_x} \beta\beta_{k_x l_x} + \beta_{k_x} \alpha\alpha_{i_x j_x} + \beta_{l_x} \alpha\alpha_{i_x j_x} + \alpha\alpha\beta\beta_{i_x j_x k_x l_x} \\ Y : A_{i_y} A_{j_y} B_{k_y} B_{l_y} &= \alpha_{i_y} + \alpha_{j_y} + \beta_{k_y} + \beta_{l_y} + \alpha\alpha_{i_y j_y} + \beta\beta_{k_y l_y} + \alpha_{i_y} \beta_{k_y} + \alpha_{i_y} \beta_{l_y} + \alpha_{j_y} \beta_{k_y} + \\ &+ \alpha_{j_y} \beta_{l_y} + \alpha_{i_y} \beta\beta_{k_y l_y} + \alpha_{j_y} \beta\beta_{k_y l_y} + \beta_{k_y} \alpha\alpha_{i_y j_y} + \beta_{l_y} \alpha\alpha_{i_y j_y} + \alpha\alpha\beta\beta_{i_y j_y k_y l_y} \end{aligned}$$

La covarianza entre los genotipos de X e Y (Apéndice I) contendrá 15^2 términos que pueden agruparse en tipos de diferente constitución. Interesan en particular los 5 tipos de términos cuyas covarianzas pueden ser diferentes de cero.

De los 16 términos que representan efectos de alelos tomados de a uno (combinaciones de $\alpha_i, \alpha_j, \beta_k, \beta_l$ en X con $\alpha_i, \alpha_j, \beta_k, \beta_l$ en Y) sólo los que involucran alelos de un mismo locus podrán ser diferentes de cero. Habrá 4 términos del tipo:

$$E\alpha_{i_x} \alpha_{i_y} = p(A_{i_x} \equiv A_{i_y}) E\alpha_i^2 = \frac{a_{xy}}{2} E\alpha_i^2$$

y otros 4 del tipo:

$$E\beta_{k_x} \beta_{k_y} = p(B_{k_x} \equiv B_{k_y}) E\beta_k^2 = \frac{a_{xy}}{2} E\beta_k^2$$

por lo que, en total, la contribución de estos términos será:

$$\frac{4a_{xy}}{2} E\alpha_i^2 + \frac{4a_{xy}}{2} E\beta_i^2 = a_{xy} (2E\alpha_i^2 + 2E\beta_i^2) = a_{xy} \sigma_{10}^2$$

Dos de los 4 términos que involucran pares alélicos corresponden a interacciones intra-locus y pueden contribuir componentes distintos de cero:

$$E\alpha\alpha_{ij_x} \alpha\alpha_{ij_y} = d_{xy} E(\alpha\alpha)_i^2 \quad \text{y} \quad E\beta\beta_{kl_x} \beta\beta_{kl_y} = d_{xy} E(\beta\beta)_k^2$$

Estos dos términos contribuyen a la varianza de dominancia:

$$d_{xy} E(\alpha\alpha)_i^2 + d_{xy} E(\beta\beta)_k^2 = d_{xy} [E(\alpha\alpha)_i^2 + E(\beta\beta)_k^2] = d_{xy} \sigma_{01}^2$$

Los 16 términos correspondientes a alelos individuales tomados de a pares (combinaciones de $\alpha_i\beta_k$, $\alpha_i\beta_l$, $\alpha_j\beta_k$, $\alpha_j\beta_l$ en X con $\alpha_i\beta_k$, $\alpha_i\beta_l$, $\alpha_j\beta_k$, $\alpha_j\beta_l$ en Y), son diferentes de cero cuando esos alelos individuales tomados de a pares son idénticos por descendencia. La probabilidad de que lo sean es $(a_{xy}/2)^2$ y por lo tanto:

$$\frac{16 a_{xy}^2 E(\alpha\beta)_{ik}^2}{4} = 4a_{xy}^2 E(\alpha\beta)_{ik}^2 = a_{xy}^2 \sigma_{20}^2$$

Los 8 términos del tipo combinaciones de $\alpha\alpha\beta$ en X con $\alpha\alpha\beta$ en Y o $\alpha\beta\beta$ en X con $\alpha\beta\beta$ en Y son diferentes de cero cuando tanto el alelo individual como el par alélico son idénticos por descendencia. Ello ocurre con una probabilidad $a_{xy}d_{xy}/2$ y entonces:

$$4E(\alpha\alpha\beta)^2 + 4E(\alpha\beta\beta)^2 = \frac{4a_{xy}d_{xy}}{2} [E(\alpha\alpha\beta)^2 + E(\alpha\beta\beta)^2] = a_{xy}d_{xy} \sigma_{11}^2$$

El término de interacción entre dobles pares alélicos contribuirá a la variabilidad cuando estos pares sean idénticos por descendencia, un evento con probabilidad d_{xy}^2 :

$$d_{xy}^2 E(\alpha\alpha\beta\beta)^2 = d_{xy}^2 \sigma_{02}^2$$

La contribución a la covarianza genética entre individuos de todos los términos detallados para un modelo de dos loci será:

$$\text{cov}(A_{i_x} A_{j_x} B_{k_x} B_{l_x}, A_{i_y} A_{j_y} B_{k_y} B_{l_y}) = a_{xy} \sigma_{10}^2 + d_{xy} \sigma_{01}^2 + a_{xy} d_{xy} \sigma_{11}^2 + a_{xy}^2 \sigma_{20}^2 + d_{xy}^2 \sigma_{02}^2$$

El patrón del modelo de 2 loci puede extenderse por inducción a un modelo de n loci.

Modelo de n loci

Si dos individuos X e Y provienen de una población no endocriada que se está apareando al azar y puede asumirse que los efectos de la interacción genotipo-ambiente y el ligamiento son despreciables, entonces la covarianza genética entre ellos puede expresarse:

$$\sigma_{G_{XY}} = \sum_{i=0}^n \sum_{j=0}^n a_{XY}^i d_{XY}^j \sigma_{ij}^2 \quad \text{sujeto a : } 1 \leq i+j \leq n$$

Si los supuestos de apareamiento aleatorio y ausencia de endocria no se cumplen, los coeficientes de relación aditiva y dominancia estarán aumentados y no será válido el supuesto de aditividad de los componentes de varianza. Mather (1949) y Kempthorne (1957) tratan esas situaciones.

Progenitor-progenie

La relación aditiva entre un padre y un hijo es 1/2 y la relación de dominancia es cero. Por lo tanto, la covarianza genética entre ellos será:

$$\sigma_{G_{XY}} = a_{XY} \sigma_{10}^2 + a_{XY}^2 \sigma_{20}^2 + a_{XY}^3 \sigma_{30}^2 + a_{XY}^4 \sigma_{40}^2 + \dots + a_{XY}^n \sigma_{n0}^2 = \frac{\sigma_{10}^2}{2} + \frac{\sigma_{20}^2}{4} + \frac{\sigma_{30}^2}{8} + \dots + \frac{\sigma_{n0}^2}{2^n}$$

Note que a medida que aumenta el orden de los términos de interacción (i.e., 20, 30, 40, ..., n0), las contribuciones son ponderadas por coeficientes que se reducen exponencialmente. Es esperable que sólo los términos de muy bajo orden contribuyan significativamente a la covarianza genética.

Hermanos completos

Las relaciones aditiva y de dominancia son 1/2 y 1/4, respectivamente, de modo que:

$$\begin{aligned} \sigma_{G_{XY}} &= a_{XY} \sigma_{10}^2 + d_{XY} \sigma_{01}^2 + a_{XY} d_{XY} \sigma_{11}^2 + a_{XY}^2 \sigma_{20}^2 + d_{XY}^2 \sigma_{02}^2 + a_{XY}^2 d_{XY}^2 \sigma_{22}^2 + \dots \\ &= \frac{\sigma_{10}^2}{2} + \frac{\sigma_{01}^2}{4} + \frac{\sigma_{11}^2}{8} + \frac{\sigma_{20}^2}{4} + \frac{\sigma_{02}^2}{16} + \frac{\sigma_{22}^2}{64} + \dots \end{aligned}$$

En la Tabla 1 se indican coeficientes de componentes de varianza para los tipos de parentesco más comunes. Como se deduce de la tabla, un progenitor transmite alelos individuales y puede transmitir asimismo parte de las interacciones entre alelos individuales. Las contribuciones de varianza aditiva son las más importantes y en ellas se basa el mejoramiento genético por selección.

Tabla 1. Coeficientes de componentes de varianza genética para diferentes tipos de parientes

Parentesco	a_{XY}	d_{XY}	A	D	A x A		DxD		AxD		
			σ^2_{10}	σ^2_{01}	σ^2_{20}	σ^2_{30}	σ^2_{02}	σ^2_{03}	σ^2_{11}	σ^2_{12}	σ^2_{21}
Gemelos	1	1	1	1	1	1	1	1	1	1	1
Ancestro- descendiente ¹	$(1/2)^n$	0	$(1/2)^n$	0	$(1/2)^{2n}$	$(1/2)^{3n}$	0	0	0	0	0
Hermanos completos	1/2	1/4	1/2	1/4	1/4	1/8	1/16	1/64	1/8	1/32	1/16
Medio hermanos	1/4	0	1/4	0	1/16	1/64	0	0	0	0	0
Primos hermanos	1/8	0	1/8	0	1/64	1/512	0	0	0	0	0
Tío-sobrino	1/4	0	1/4	0	1/16	1/64	0	0	0	0	0
Dobles primos hermanos	1/4	1/16	1/4	1/16	1/16	1/64	1/256	1/4096	1/64	1/4096	1/256

¹Separados n generaciones

Bibliografía

Kempthorne, O. 1957. Introduction to genetic statistics. Wiley, N.Y.
Mather. 1949. Biometrical genetics. Dover. N.Y.

Apéndice I

Para el cálculo de una covarianza entre sumas de variables aplican las siguientes relaciones. Sean:

$$a = w + x$$

$$b = y + z$$

La covarianza entre a y b puede escribirse:

$$\text{cov}(a,b) = [\Sigma ab - (\Sigma a \Sigma b / N)] / N$$

Si a y b están expresadas como desvíos, Σa y Σb valen cero y, por lo tanto:

$$\text{cov}(a,b) = \Sigma ab / N = E(ab) = E(w + x)(y + z) = Ewy + Ewz + Exy + Exz$$

La covarianza tendrá un número de términos igual al cuadrado del número de elementos que contienen las variables compuestas por sumas de variables.