

# Observacions sobre el càlcul de la mitjana fenotípica condicional

J. Ignacio Lucas Lledó

3/10/2022

Lynch and Walsh (1998), pàgina 89, ofereix un exemple de com calcular les mitjanes condicionals dels valors genotípics d'un caràcter al·lomètric de dacsca (*Zea mays*) en una població experimental amb només dos gens contribuint a la variància genètica. La mitjana condicional es defineix com "the conditional mean phenotype of individuals with allele\* i at the first locus without regard to the other allele at the locus or to the other genotype at the second locus". Entenc que la mitjana condicional,  $G_{i...}$ , és el valor fenotípic promig dels individus amb (almenys) un al·lel  $i$  al primer locus.

Sense assumir equilibri de Hardy-Weinberg, jo interprete aquesta definició tal que així:

$$G_{X...} = \frac{1}{\sum_{jkl} P(Xjkl)} \sum_{jkl} P(Xjkl) G_{Xjkl} \quad (1)$$

On  $P(Xjkl)$  és la freqüència del genotip  $(Xj, kl)$ . De la mateixa manera que en el segon locus, el genotip  $kl$  pot prendre tres valors (perquè no diferenciem les dues classes d'heterozigot possibles d'acord amb el progenitor del qual han heretat cada al·lel), els genotips  $Xj$  inclouen qualsevol dels dos tipus d'heterozigot ( $Xj$  i  $jX$ ) i l'homozigot  $XX$ .

Així doncs, si assumim equilibri de Hardy-Weinberg, i segregació independent, i si les freqüències al·lelèiques fóren  $p_1$  i  $q_1$  al primer locus i  $p_2$  i  $q_2$  al segon, i si la freqüència  $p_1$  és la de l'al·lel  $X$ , etc., aleshores:

$$G_{X...} = \frac{p_1^2 p_2^2 G_{XXZZ} + p_1^2 2p_2 q_2 G_{XXZW} + p_1^2 q_2^2 G_{XXWW} + 2p_1 q_1 p_2^2 G_{XYZZ} + 2p_1 q_1 2p_2 q_2 G_{XYZW} + 2p_1 q_1 q_2^2 G_{XYWW}}{p_1^2 + 2p_1 q_1} \quad (2)$$

On el denominador és la freqüència en la població d'individus amb almenys un al·lel  $X$ .

Però Lynch and Walsh (1998) adopta una aproximació diferent, que dóna resultats diferents. A la pàgina 88 utilitza probabilitats condicionals per definir:

$$G_{X...} = \sum_{jkl} P(ijkl|i = X) G_{Xjkl} \quad (3)$$

(He canviat  $B_T$  per  $X$ , per estalviar-me algun subíndex), i ve a dir, a la pàgina següent, que  $P(ijkl|i = X) = P(j) \cdot P(kl)$ . És a dir, la probabilitat condicional del genotip  $(ij, kl)$  sabent que el cromosoma  $i$  porta l'al·lel  $X$  seria el producte de la probabilitat de què l'altre cromosoma porte l'al·lel  $j$  (igual a la freqüència al·lelèica) i la probabilitat del genotip  $kl$  al segon locus, d'acord amb l'equilibri de Hardy-Weinberg. Aquesta interpretació és consistent, però diferent de la meua, en tant que jo no diferenciaria entre els dos cromosomes.

La diferència d'interpretacions es pot comprendre amb un únic locus, amb dos al·lells  $A$  i  $a$ . Jo diria que les probabilitats condicionals de ser un individu homozigot  $AA$ , (o heterozigot) sabent que té almenys un al·lel  $A$  són:

$$P(AA|A-) = \frac{p^2}{p^2 + 2pq} = \frac{p}{1 + q}$$

$$P(Aa|A-) = \frac{2pq}{p^2 + 2pq} = \frac{2q}{1 + q}$$

Perquè no faig distinció entre els dos cromosomes, i considere que la informació que se'ns dona és que “almenys hi ha un al · lel  $A$ ”, no en quin dels dos cromosomes està l'al · lel  $A$ . En canvi, si distingim els cromosomes i considerem diferents els heterozigots  $Aa$  dels  $aA$ , aleshores:

$$P(AA|A-) = P(AA| - A) = p$$

$$P(Aa|A-) = P(aA| - A) = q$$

$$P(aA|A-) = P(Aa| - A) = 0$$

Aquesta és la interpretació que fan Lynch and Walsh (1998), i que concorda amb la notació utilitzada, ja que les variables  $i$  i  $j$  no identifiquen els al · lels per l'estat, sinó pel cromosoma en què es troben. Una conseqüència necessària de la interpretació de Lynch i Walsh és que  $G_{X_{jkl}} = G_{iX_{kl}}$ , és a dir: es podrien definir i calcular per separat els efectes additius d'un al · lel heretat del pare i del mateix al · lel heretat de la mare. La conseqüència és que a l'hora de ponderar els valors fenotípics de cada genotip en el càlcul de la mitjana condicional, els heterozigots en el locus focal només compten al 50%. De fet, el meu càlcul es correspon amb el de Lynch and Walsh (1998) si dividisc les freqüències dels heterozigots al locus focal entre 2.

Però, aleshores, si Lynch and Walsh (1998) tenen raó, definir verbalment la mitjana condicional és un maldecap. I d'explicar-la, ja no vull dir-te res! Resulta molt poc intuïtiu que els heterozigots hagen de comptar només la meitat en la mitjana ponderada. Com que no sé d'on ha eixit la igualtat  $\alpha_i = G_{i...} - \mu_G$ , no puc jutjar si per casualitat Lynch and Walsh (1998) poden haver-se equivocat. Però supose que no, i que ha d'haver un bon motiu per calcular la mitjana condicional amb només la meitat d'heterozigots.

Lynch, M., and B. Walsh. 1998. *Genetics and Analysis of Quantitative Traits*. Sunderland, MA, U.S.A.: Sinauer Associates, Inc.