

**PROBABILIDADES DE ABSORCION BAJO APAREAMIENTO ENTRE HERMANOS ENTEROS
CON SELECCION GENOTIPICA Y ALELICA**

N. BARTOLONI

Recibido: 14-09-88

Aceptado: 29-11-89

RESUMEN

Se considera un sistema unigénico con un par de alelos y valores de selección para los alelos (gametas) y para los genotipos combinando en quince casos a lo largo del tiempo bajo apareamiento entre hermanos enteros. Se ha analizado la variable "número de generaciones necesarias para alcanzar un estado absorbente" y sus parámetros: promedio, variación y asimetría, aplicando la teoría de las cadenas de Markov con barreras absorbentes. Los efectos de la selección gamética y zigótica sobre el sistema a lo largo de las generaciones ha sido evaluado. Las operaciones matemáticas han sido desarrolladas por Bosso et al. (1969).

Palabras clave: selección, cadenas markovianas, apareamiento entre hermanos.

ABSORPTION PROBABILITIES UNDER FULL SIB MATING WITH ALLELIC AND GENOTYPIC SELECTION

SUMMARY

It has been considered a system of one gene with a pair of alleles (gametes) and for the genotypes combining in fifteen cases along the time under full sib mating. It has been analysed the variable "number of generations needed for reach an absorbent state" and its parameters: average, variation, and asymmetry, applicating the theory of Markov chains. The effect of gametic and genotypic selection on the system along the generations have been evaluated. The mathematical manipulations were developed by Bosso et al. (1969).

Key words: selection, Markov chains, full sib-mating.

Cátedra de Genética. Facultad de Agronomía. Universidad de Buenos Aires.
Avda. San Martín 4453 (1417) Buenos Aires - Argentina -

INTRODUCCION

En la conducción de poblaciones de plantas y animales de importancia agronómica surge el interés en prever la constitución de la población en el futuro. Al considerar las conclusiones a las que arribaron Bosso *et al.* (1969) y Tan (1973) se partió de la hipótesis de que los distintos tipos de selección afectan los parámetros de tendencia central, dispersión y asimetría de la variable "número de generaciones hasta absorción". Se ha analizado el caso de apareamiento regular entre hermanos enteros (FS) siguiendo las posibles trayectorias de un alelo de un gene comenzando en un genotipo heterocigota a lo largo del tiempo por medio de la variable mencionada más arriba. Como queda dicho, el propósito de esta ejercitación es analizar la influencia de la selección sobre la variable mencionada utilizando un sistema de apareamiento extensivamente usado en mejoramiento.

METODOS

Se ha considerado el caso de un gene con dos alelos (A_1 y A_2) con aptitud gamética (w_1 para A_1 y w_2 para A_2) y aptitud genotípica (w_1 para A_1A_1 , w_2 para A_1A_2 y w_3 para A_2A_2). Con FS tenemos seis posibles clases de familia:

- (A) $A_1A_1 \times A_1A_1$; (B) $A_2A_2 \times A_2A_2$; (C) $A_1A_1 \times A_2A_2$; (D) $A_1A_1 \times A_1A_2$; (E) $A_1A_2 \times A_2A_2$;
 (F) $A_1A_2 \times A_1A_2$.

Siguiendo la teoría de las cadenas de Markov llamamos a los estados (A) y (B) como estados absorbentes (AS) y a los restantes como estados transitorios (TS) y la matriz de probabilidades de transición (TPM) será:

generación n+1

$$TPM = \begin{bmatrix} I & | & O \\ \hline R & | & Q \end{bmatrix}$$

donde I es una matriz identidad 2x2, O es una matriz nula 2x4, R es la submatriz 4x2 de probabilidades de transición entre TS y AS y Q es la submatriz 4x4 de probabilidades de transición entre TS, y :

$$R = \begin{bmatrix} 0 & 0 \\ (w_1w_1)^2/c_D & 0 \\ 0 & (w_3w_2)^2/c_E \\ (w_1w_1)^2/c_F & (w_3w_2)^2/c_F \end{bmatrix}$$

donde $c_D = (w_1w_1 + w_2w_2)^2$; $c_E = (w_2w_1 + w_3w_2)^2$
 y $c_F = (w_1w_1^2 + 2w_1w_2w_1 + w_3w_2^2)^2$

$$Q = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 2w_1w_2w_1/c_D & 0 & (w_2w_2)^2/c_D \\ 0 & 0 & 2w_2w_3w_1/c_E & (w_2w_1)^2/c_E \\ 2w_1w_3w_1/c_F & 4w_1w_2w_1^3/c_F & 4w_2w_3w_1^3/c_F & (2w_2w_1w_2)^2/c_F \end{bmatrix}$$

Quince casos de selección fueron considerados en este trabajo (Cuadro N° 1). Para todos los casos se ha calculado la TPM usando los valores selectivos para fecundidad y viabilidad (Tan, 1973). Luego, se han obtenido las siguientes matrices y vectores (Bosso *et al.*, 1969).

- a) La matriz promedio **N**: para esta matriz cada elemento N_{ij} indica el número medio de generaciones que un proceso que comienza en el estado S_i pasa a través del S_j antes de alcanzar un AS.
- b) El vector columna **T**: para este vector cada elemento T_i indica el número medio de generaciones que un proceso que comienza en S_i necesita para alcanzar un estado absorbente (incluyendo la posición inicial).
- c) El vector columna T_2 : para este vector cada elemento T_{2i} denota la variancia del número de etapas para alcanzar un AS a partir de S_i .
- d) El vector columna **CV**: con elementos CV_i que indican el coeficiente de variación de la variable "número de etapas para alcanzar un AS".
- e) El vector columna T_3 : cuyos elementos T_{3i} indican el momento de tercer orden de la variable alrededor de la media T_i .
- f) El vector columna **ASI**: cada elemento denota el coeficiente de asimetría de la misma variable de (d).

Las expresiones de los cálculos (Kemeny y Snell, 1960) han sido omitidas aquí por una razón de espacio. Las operaciones con las matrices y vectores siguen las reglas del álgebra elemental (Máltsev, 1978).

RESULTADOS

Los resultados generales en términos de **T**, **CV**, y **ASI** están en el Cuadro N° 2, Cuadro N° 3 y Cuadro N° 4. Sólo se agregan que las expresiones para el **CV** y el **ASI** que son:

$$CV_i = \frac{(T_{2i})^{0,5}}{T_i} (100) \text{ y } ASI_i = \frac{T_{3i}}{(T_{2i})^{1,5}}$$

DISCUSION Y CONCLUSIONES

i) Vector **T**.

1) Entre los **TS**.

Para (C), **T** es siempre mayor que el **T** para los otros **TS** ya que **C** implica un pasaje obligatorio **C→F** antes de llegar a la absorción. Para **D**, **T** es siempre el menor excepto en **XI** puesto que existe ventaja selectiva de A_2 sobre A_1 . **T** es más grande que T_2 en **I** y **II**. Cuando existe alguna ventaja de $A_1 A_2$ (**III**, **IV** y **V**), **E** va hacia **B** en el cual el homocigota $A_2 A_2$ es selectivamente desventajoso. A su vez, **F** va hacia **A** y **B**. Por esta razón el sistema tarda más para alcanzar un **AS** desde **E** que

Cuadro N° 1: Diferentes casos combinando valores selectivos para alelos y genotipos.

	A'	B'	C'	D'	E'
$w_2 = 1,0$	I	II	III	IV	V
$w_2 = 0,8$	VI	VII	VIII	IX	X
$w_2 = 1,2$	XI	XII	XIII	XIV	XV

En todos los casos $w_1 = 1,0$

Referencias:

A: $w_1 = w_2 = w_3 = 1,0$

En B, C, D y E $w_1 = 1,0$ y $w_3 = 0,6$

y difieren de w_2 : $w_2(B) = 0,4$;

$w_2(C) = 0,8$; $w_2(D) = 1,0$ y $w_2(E) = 1,2$

Cuadro N° 2: Valores de T para todos los casos de selección.

Cuadro N° 3: Valores de CV para todos los casos de selección.

Cuadro N° 4: Valores de ASI para todos los casos de selección.

Cuadro N° 2

Cuadro N° 3

Cuadro N° 4

TS	Cuadro N° 2				Cuadro N° 3				Cuadro N° 4			
	(C)	(D)	(E)	(F)	(C)	(D)	(E)	(F)	(C)	(D)	(E)	(F)
I	6,66	4,83	4,83	5,66	71,4	95,6	95,6	84,0	10,2	5,6	5,6	5,3
II	4,38	2,16	2,96	3,38	57,5	85,4	83,8	74,6	17,5	7,6	7,1	9,8
III	6,38	4,06	5,41	5,38	70,2	100,0	85,5	83,2	10,6	5,3	6,6	7,1
IV	7,73	5,37	6,84	6,73	75,4	102,0	87,0	86,6	8,9	5,0	6,4	6,6
V	9,17	6,82	8,32	8,17	79,4	103,0	88,7	89,0	8,2	4,9	5,1	6,2
VI	6,40	3,94	5,40	5,40	69,7	100,0	85,5	82,6	10,8	5,3	6,6	7,2
VII	4,02	1,86	3,22	3,02	56,0	78,9	77,4	74,4	18,9	7,1	8,2	9,1
VIII	5,54	3,24	5,23	4,55	66,0	97,3	75,4	80,5	12,2	5,7	9,6	7,6
IX	6,47	4,11	6,22	5,46	70,4	100,0	76,7	83,3	10,5	5,3	8,4	7,1
X	7,53	5,13	7,31	6,53	74,3	102,0	78,8	85,7	9,3	5,2	7,9	6,7
XI	6,44	5,20	4,22	5,44	70,5	88,0	99,4	83,4	10,5	6,2	4,8	7,1
XII	4,54	2,44	2,68	3,54	58,1	87,8	87,2	74,5	17,1	7,0	6,8	8,5
XIII	6,81	4,79	5,21	5,81	72,1	97,9	93,3	84,5	9,9	5,3	5,3	6,9
XIV	8,32	5,21	6,77	7,32	77,4	99,5	94,8	88,0	8,4	5,0	5,5	6,4
XV	10,10	8,07	8,58	9,10	81,5	100,0	95,8	90,5	7,5	5,0	5,1	6,0

desde F. En el caso en que $w_1 > w_2$ (VI a X) este proceso está enfatizado a medida que la frecuencia de aparición de A_1A_2 decrece. Si $w_2 > w_1$, la tendencia se invierte (XI a XV). Recapitulando: el sistema emplea para entrar en algún AS en promedio, mayor cantidad de generaciones si comienza desde C. Luego, en orden decreciente de número de generaciones, D, E y F.

2) Entre los w.

Si $w_1 > w_2$ implica un incremento en la frecuencia de A_1A_1 en las progenies. Si no hay selección genotípica, $w_1 \neq w_2$ implica que un decrecimiento en los valores de T debido a la relación entre la frecuencia de homocigotas y la frecuencia alélica que es del tipo $2x - 2x + 1$ siendo x la frecuencia alélica (la cual está directamente determinada por w). Esta ecuación tiene un mínimo en $x=0,5$ y los máximos relativos, en $x=0$ y $x=1$. Así que, cualquier cambio en w (en ausencia de W) implicará un incremento en la probabilidad de ingresar en algún AS y, entonces, un decrecimiento en T. Cuando $w_1 > w_3$ los valores de T decrecen aún más. $w_1 > w_2$ implica un incremento en T que se transforma en decrecimiento cuando $w_2 > w_3$. Esto indica una entrada en AS incrementada a través de F, es decir, $E \rightarrow F \rightarrow A$. Cuando $w_2 > w_1$, la ventaja $w_1 > w_3$ se balancea y, aún se invierte la tendencia (compárense los casos II a V con los casos XII a XV). Otra vez, para T_E el incremento está equilibrado por la ventaja de A_1A_2 y, aún la tendencia se invierte (XV).

En general, se observa un incremento en los valores de T cuando los valores de w_2 crecen siendo los casos de no selección (I, VI y XI) intermedios. Esto implica que la presencia de ventaja para el heterocigota produce un incremento en el número de generaciones que tarda el sistema en entrar en algún AS y cuando mayor sea w_2 con respecto a w_1 y w_3 mayor será dicho incremento.

ii) Vector CV.

1) Entre los TS.

CV_C es siempre el más pequeño así que no se esperan entradas tempranas en AS para sistemas que han comenzado desde C. Para D, CV es el mayor, excepto en XI (a causa de la diferencia $w_2 > w_1$). Así que, aunque el sistema entre en un AS más rápidamente cuando parte de D₁, que cuando parte de los otros TS, se esperará un periodo extendido sobre el mayor número de generaciones durante el cual el pasaje hacia AS es continuo. En general, E presenta más variación que F, excepto en algunos casos (V, VIII, IX y X) en los cuales la ventaja A_1A_2 produce un gran aumento en la variación para F.

2) Entre los w.

La variación es más pequeña cuando $w_1 > w_2$ que cuando $w_1 = w_2$ y más grande cuando $w_2 > w_1$. Para F esta tendencia es cierta solo cuando $w_2 > w_3$. De modo que, en general, el periodo de entrada en AS se hace ancho cuando $w_2 > w_1$ y se hace angosto cuando $w_1 > w_2$, comparando con $w_1 = w_2$.

3) Entre los W.

La variación crece con el crecimiento de w_2 siendo los casos de no selección en valores intermedios entre $w_2 > w_3$ y $w_2 > w_1$, excepto para E en el cual siempre existe mayor variación sin selección genotípica que con ella a causa de la diferencia $w_1 > w_3$ lo cual está siendo considerado en todos los casos de selección (la aptitud está positivamente correlacionada con la variación). Así que en general, los mayores valores de w_2 serán acompañados de la mayor tasa de fijación (tanto en generaciones tempranas como en generaciones avanzadas).

iii) Vector ASI.

Los valores de asimetría son positivos en todos los casos considerados, así que las distribuciones de frecuencia tienen una larga cola hacia la derecha.

1) Entre los TS.

En general, los mayores valores de asimetría corresponden a C y, en orden decreciente, según los estados F, E y D. De modo que, para sistemas que comienzan desde C y teniendo en cuenta sus valores de media y CV, se esperaran muy bajas frecuencias de fijación en generaciones tempranas, habiendo una alta concentración en las generaciones iniciales (y cerca de la media) y, aún, habrá entrada en AS (aunque con baja frecuencia) en generaciones muy avanzadas de endogamia; más que para cualquier otro TS. Para D, a su vez se esperará la más alta frecuencia de entradas en AS para generaciones tempranas y con alta concentración en generaciones intermedias (pequeña asimetría, gran variación). Para E y F las distribuciones para el número de generaciones necesarias para alcanzar un AS son muy similares con alta concentración de entradas en AS en las generaciones iniciales anteriores a la medida, aunque no tan iniciales como en C. Sus asimetrías relativas son muy similares, también, más grandes que las de D y mucho más pequeñas que para C.

2) Entre los w .

El caso $w_1 = w_2$ tiene valores intermedios de media, variación y asimetría. Con $w_1 > w_2$ existe una gran concentración de entradas en AS al comienzo del proceso a causa de su gran asimetría y su bajo promedio. Con $w_2 > w_1$ se esperarán entradas en AS aún en muy avanzadas generaciones de FS a causa de su alto promedio y su alta variación.

3) Entre los W.

Los valores de media y variación crecen con w_2 y en la misma forma decrece la asimetría. Eso implica que con $w_2 = 0,4$ tenemos la más alta frecuencia de entradas en AS al comienzo del proceso y, por otro lado, se tiene los más altos valores de heterozigosis con $w_2 > 1$, aún en generaciones muy avanzadas de FS.

BIBLIOGRAFIA

- 1) BOSSO, J.A., O.M. SORPARAIN and F.A. FAVRET. 1969. Application of finite absorbent Markov chains to sib mating populations with selection, *Biometrics*, 22:17-26.
- 2) MALTSEV, A. 1978. Fundamentos de álgebra lineal. Mir (eds), Moscú, pp. 240-252.
- 3) TAN, W. 1973. Application of some finite Markov chain theories to two locus selfing model with selection. *Biometrics*, 29:331-346.