

LE CALCUL DE L'EFFECTIF GÉNÉTIQUE DES POPULATIONS ANIMALES

L. OLLIVIER

*Station de Génétique quantitative et appliquée,
Centre national de Recherches zootechniques, I. N. R. A.,
78350 Jouy en Josas*

RÉSUMÉ

Les formules habituellement utilisées pour le calcul de l'effectif génétique des populations animales aboutissent à surestimer cet effectif. Des formules plus exactes sont données et une application est faite au cas des populations animales où l'insémination artificielle est pratiquée : nombre de mâles limité, nombre de femelles illimité. N_m étant le nombre des mâles, l'effectif génétique (N_g) est donné par $\frac{1}{N_g} = \frac{2}{16N_m} \left(1 + \frac{2N_m}{2N_m - 1} \right)$ au lieu de $\frac{1}{N_g} = \frac{1}{4N_m}$, quand le choix des mâles est aléatoire, et $\frac{1}{N_g} = \frac{2}{16N_m} \left(1 + \frac{N_m}{2N_m - 1} \right)$ au lieu de $\frac{1}{N_g} = \frac{3}{16N_m}$, quand chaque mâle laisse un fils à la génération suivante.

La notion d'effectif génétique, due à WRIGHT (1931), permet de prédire simplement l'accroissement du coefficient de consanguinité F et la dérive aléatoire des fréquences géniques dans les populations d'effectif limité, quel que soit le mode de reproduction.

Dans le cas des populations animales (avec N_m mâles et N_f femelles), les deux formules suivantes de l'effectif génétique (N_g) sont classiquement utilisées (voir par exemple FALCONER, 1960) :

$$\frac{1}{N_g} = \frac{1}{4N_m} + \frac{1}{4N_f} \quad (1)$$

$$\frac{1}{N_g} = \frac{3}{16N_m} + \frac{1}{16N_f} \quad (2)$$

La formule (1), établie par WRIGHT (1931), s'applique lorsque le choix des reproducteurs à chaque génération est *purement aléatoire*. La formule (2), établie par GOWE *et al.* (1959), s'applique lorsque chaque père laisse un fils et N_f/N_m filles à la génération suivante et chaque mère laisse une fille à la génération suivante. Dans ce cas, seul est aléatoire le nombre de fils par mère. Ces deux formules sont obtenues en remarquant que la dérive des fréquences géniques qui accompagne le passage de chaque génération à la suivante à deux composantes : l'une provient de l'échantillonnage des gamètes dans l'individu et l'autre de l'inégale contribution des individus à la génération suivante. La démonstration détaillée est fournie par LATTER (1959) qui montre que la dérive génétique est une fonction :

- des nombre N_m et N_f (supposés constants d'une génération à l'autre)
- des distributions des nombres de fils et de filles d'un père et d'une mère caractérisées par leurs variances σ_1^2 (père \rightarrow fils), σ_2^2 (père \rightarrow fille), σ_3^2 (mère \rightarrow fils) et σ_4^2 (mère \rightarrow fille).

En admettant que ces distributions sont indépendantes, l'effectif génétique peut s'écrire sous la forme suivante (voir HILL, 1972) :

$$\frac{I}{N_g} = \frac{I}{16N_m} \left[2 + \sigma_1^2 + \left(\frac{N_m}{N_f} \right)^2 \sigma_2^2 \right] + \frac{I}{16N_f} \left[2 + \sigma_3^2 + \left(\frac{N_f}{N_m} \right)^2 \sigma_4^2 \right] \quad (3)$$

L'application de la formule (3) à une situation concrète requiert donc le calcul des variances σ_1^2 , σ_2^2 , σ_3^2 et σ_4^2 en fonction du mode de choix des reproducteurs. Lorsque les nombres de descendants de chaque individu sont égaux, ces quatre variances sont nulles. Ces conditions ne peuvent être remplies que si $N_m = N_f$ et que si chaque individu laisse un fils et une fille dans la génération suivante. On a alors :

$$\frac{I}{N_g} = \frac{I}{8N_m} + \frac{I}{8N_f} \quad (4)$$

La dérive est réduite de moitié par rapport au cas de la formule (1) (alternativement on peut dire que l'effectif génétique est doublé).

Quand le choix des reproducteurs est aléatoire, plusieurs distributions sont possibles :

a) le nombre des parents est supérieur à celui des descendants et chaque parent ne donne pas plus d'un descendant. Si p est le rapport nombre de descendants/nombre de parents, un parent quelconque a soit un descendant retenu (avec une probabilité p) soit aucun descendant (avec une probabilité $1 - p$). La distribution du nombre de descendants suit donc *une loi de Bernoulli*, dont la variance est $p(1 - p)$.

b) le choix des descendants est totalement aléatoire, sans limitation du nombre de descendants par individu. Si N et N' sont les nombres de parents et de descendants, le nombre de descendants d'un individu suit *une distribution binomiale* dont les paramètres sont $\frac{I}{N}$ (probabilité de l'événement élémentaire) et N' (nombre d'épreuves).

D'où la variance recherchée : $N' \left(\frac{I}{N} \right) \left(1 - \frac{I}{N} \right)$.

c) lorsque N est suffisamment grand, cette distribution binomiale tend vers une *loi de Poisson* dont la moyenne est N'/N et la variance est égale à la moyenne.

Ainsi, lorsque N_m mâles et N_f femelles sont choisis au hasard parmi les descendants des N_m mâles et N_f femelles de la génération suivante, on a (pour N_m et N_f suffisamment grands) :

$$\begin{aligned}\sigma_1^2 &= 1 && \text{(nombre moyen de fils par père)} \\ \sigma_2^2 &= N_f/N_m && \text{(nombre moyen de filles par père)} \\ \sigma_3^2 &= N_m/N_f && \text{(nombre moyen de fils par mère)} \\ \sigma_4^2 &= 1 && \text{(nombre moyen de filles par mère)}\end{aligned}$$

En reportant ces variances dans la formule (3), on obtient la formule (1). Les conditions d'application de la formule (2), indiquées dans le deuxième paragraphe de cet article, entraînent $\sigma_1^2 = \sigma_2^2 = \sigma_4^2 = 0$. Par contre, la distribution du nombre de fils d'une mère quelconque suit une loi de Bernoulli dont la variance est :

$$\sigma_3^2 = (N_m/N_f) (1 - N_m/N_f)$$

Le report de ces variances dans (3) conduit à la formule (2). Le tableau 1 donne quelques exemples d'effectif génétique pour des situations intermédiaires entre celles de la formule (1) et de la formule (4).

On peut remarquer que si $N_m = N_f = 1$ on se trouve dans le cas F du tableau 1 qui donne $N_g = 4$ au lieu de la valeur exacte $N_g = 2$. Cet apparent paradoxe provient

TABLEAU I

Effectif génétique (N_g) dans diverses situations

(N_m et N_f sont supposés grands)

Mode de choix des reproducteurs	Cas	Distribution et variance du nombre de descendants d'un parent				$\frac{1}{N_g}$
		Voie père-fils (σ_1^2)	Voie père-fille (σ_2^2)	Voie mère-fils (σ_3^2)	Voie mère-fille (σ_4^2)	
Choix total. aléatoire	A	Poisson (1)	Poisson (N_f/N_m)	Poisson (N_m/N_f)	Poisson (1)	$\frac{4}{16N_m} + \frac{4}{16N_f}$
	B	Fixé (0)	Poisson (N_f/N_m)	Poisson (N_m/N_f)	Poisson (1)	$\frac{3}{16N_m} + \frac{4}{16N_f}$
	C	Fixé (0)	Poisson (N_f/N_m)	Bernoulli $\left[\frac{N_m}{N_f} \left(1 - \frac{N_m}{N_f} \right) \right]$	Poisson (1)	$\frac{3}{16N_m} + \frac{3}{16N_f}$
Choix partiellement aléatoire	D	Fixé (0)	Fixé (0)	Bernoulli $\left[\frac{N_m}{N_f} \left(1 - \frac{N_m}{N_f} \right) \right]$	Poisson (1)	$\frac{3}{16N_m} + \frac{2}{16N_f}$
	E	Fixé (0)	Fixé (0)	Bernoulli $\left[\frac{N_m}{N_f} \left(1 - \frac{N_m}{N_f} \right) \right]$	Fixé (0)	$\frac{3}{16N_m} + \frac{1}{16N_f}$
Choix totalement fixé	F	(0)	(0)	(0)	(0)	$\frac{2}{16N_m} + \frac{2}{16N_f}$

de ce que l'équation (3) n'est pas rigoureusement applicable quand N_m et N_f sont faibles. CROW et KIMURA (1970, p. 358) montrent que, dans une population limitée en reproduction panmictique, l'effectif génétique s'exprime, en fonction du nombre N de parents, de la moyenne m et de la variance σ^2 du nombre de descendants par parent, par la formule suivante :

$$N_g = \frac{(2N - 1)m}{2(1 + \sigma^2/m)}$$

Ainsi pour la voie père-fils on a :

$$N_1 = \frac{2N_m - 1}{2(1 + \sigma_1^2)} \quad \text{puisque } m = 1$$

On calcule de même N_2 , N_3 et N_4 pour les trois autres voies. La fréquence génique étant la moyenne des quatre fréquences géniques dans les gamètes empruntant les quatre voies de transmission et l'effectif génétique étant inversement proportionnel à la variance du changement de fréquence génique d'une génération à la suivante, on a :

$$\frac{1}{N_g} = \frac{1}{16} \left(\frac{1}{N_1} + \frac{1}{N_2} + \frac{1}{N_3} + \frac{1}{N_4} \right)$$

On obtient finalement l'expression suivante de l'effectif génétique :

$$\begin{aligned} \frac{1}{N_g} = \frac{2N_m}{2N_m - 1} \left[\frac{1}{16N_m} \left(1 + \sigma_1^2 + \sigma_2^2 \frac{N_m^2}{N_1^2} \right) + \frac{1}{16N_f} \right] \\ + \frac{2N_f}{2N_f - 1} \left[\frac{1}{16N_f} \left(1 + \sigma_4^2 + \sigma_3^2 \frac{N_f^2}{N_m^2} \right) + \frac{1}{16N_m} \right] \end{aligned} \quad (5)$$

A mesure que N_m et N_f augmentent $\frac{2N_m}{2N_m - 1}$ et $\frac{2N_f}{2N_f - 1}$ se rapprochent de 1 et on retrouve la formule (3). Celle-ci aboutit donc, pour N_m et N_f faibles, à une surestimation de N_g , puisque $2N_m/(2N_m - 1)$ est toujours supérieur à 1.

La formule (5) permet d'apporter aux effectifs génétiques du tableau 1 les correctifs voulus lorsque les nombres de reproducteurs sont faibles. A titre d'exemple, dans le cas A on obtient

$$\frac{1}{N_g} = \left(\frac{2}{16N_m} + \frac{2}{16N_f} \right) \left(\frac{2N_m}{2N_m - 1} + \frac{2N_f}{2N_f - 1} \right)$$

expression qui tend vers $\frac{4}{16N_m} + \frac{4}{16N_f}$ lorsque N_m et N_f deviennent grands, et dans le cas E

$$\frac{1}{N_g} = \frac{2}{16N_m} \left(\frac{2N_f}{2N_f - 1} + \frac{N_m}{2N_m - 1} \right) + \frac{1}{16N_f} \left(\frac{2N_m}{2N_m - 1} \right),$$

expression qui tend vers $\frac{3}{16N_m} + \frac{1}{16N_f}$.

L'effectif génétique dans une population où les femelles sont en nombre très élevé et les mâles en nombre limité (cas de l'insémination artificielle) se calcule à l'aide de la formule (1) lorsque le choix des mâles est aléatoire et on a $\frac{1}{N_g} = \frac{1}{4N_m}$ puisque $\frac{1}{N_f}$

est négligeable. Lorsque le choix des mâles est tel que chaque père laisse un fils à la génération suivante, on voit d'après le tableau 1 que le mode d'échantillonnage des voies 2, 3 et 4 est indifférent et on a toujours $\frac{1}{N_g} = \frac{3}{16N_m}$ (cas B, C, D et E du tableau 1). Le cas F, qui suppose $N_m = N_f$ comme nous l'avons vu plus haut, n'est pas à considérer et il n'y a donc, dans le cas d'une population femelle illimitée que deux cas à envisager. Selon que $\sigma_1^2 = 1$ ou 0, on a, en appliquant les formules (1) et (2) soit $N_g = 4N_m$, soit $N_g = 16N_m/3$.

Mais une valeur plus exacte de N_g peut être obtenue à partir de l'expression (5). Ainsi lorsque le choix des mâles est aléatoire — ce qui suppose que leur nombre soit au moins égal à 2 — l'expression suivante s'applique :

$$\frac{1}{N_g} = \frac{2}{16N_m} \left(1 + \frac{2N_m}{2N_m - 1} \right) \quad (6)$$

et lorsque chaque père laisse un fils à la génération suivante — en particulier quand $N_m = 1$ — on a :

$$\frac{1}{N_g} = \frac{2}{16N_m} \left(1 + \frac{N_m}{2N_m - 1} \right) \quad (7)$$

On voit que, à mesure que N_m augmente, $1/N_g$ tend à se rapprocher de $1/4N_m$ et $3/16N_m$ respectivement. Le tableau 2 donne les valeurs de l'effectif génétique calcu-

TABLEAU 2

*Effectif génétique en fonction du nombre de mâles
dans une population où les femelles sont en nombre illimité*

Nombre de mâles (N_m)	Effectif génétique			
	Quand le choix des mâles est aléatoire		Quand chaque mâle laisse un descendant à la génération suivante	
	D'après la formule (6)	$4N_m$	D'après la formule (7)	$16N_m/3$
1	(ces formules ne s'appliquent pas)		4	5,3
2	6,9	8	9,6	10,7
3	10,9	12	15,0	16,0
4	14,9	16	20,4	21,3
5	18,9	20	25,7	26,7
—	—	—	—	—
—	—	—	—	—
10	39,0	40	52,4	53,3
—	—	—	—	—
—	—	—	—	—
20	79,0	80	105,8	106,7
—	—	—	—	—
—	—	—	—	—
100	399,0	400	532,4	533,3

lées à partir de (6) et (7) pour des nombres de mâles variables en comparaison avec les valeurs obtenues à partir des équations (1) et (2). On voit que celles-ci aboutissent à surestimer d'environ une unité l'effectif génétique.

Reçu pour publication en mai 1973.

REMERCIEMENTS

L'auteur remercie le professeur G. MALÉCOT d'avoir bien voulu examiner le manuscrit de cet article.

SUMMARY

THE EFFECTIVE SIZE OF ANIMAL POPULATIONS

The calculation of the effective population number from usual formulae tends to overestimate that number. More exact formulae are given and applied to the case of animal populations with artificial insemination : limited number of males, unlimited number of females. N_m being the number of males, the effective population size (N_g) is given by $\frac{1}{N_g} = \frac{2}{16N_m} \left(1 + \frac{2N_m}{2N_m - 1} \right)$ instead of $\frac{1}{N_g} = \frac{1}{4N_m}$, when the choice of the males is random, and $\frac{1}{N_g} = \frac{2}{16N_m} \left(1 + \frac{N_m}{2N_m - 1} \right)$ instead of $\frac{1}{N_g} = \frac{3}{16N_m}$, when each male has one son in the next generation.

RÉFÉRENCES BIBLIOGRAPHIQUES

- CROW J. F., KIMURA M., 1970. *Introduction to Population Genetics Theory*. Harper and Row, New York.
- FALCONER D. S., 1960. *Introduction to Quantitative Genetics*. Oliver and Boyd. London.
- GOWE R. S., ROBERTSON A., LATTER B. D. H., 1959. Environment and poultry breeding problems. 5. The design of poultry control strains. *Poult. Sci.*, **38**, 462-471.
- HILL W. G., 1972. Estimation of genetic change. I. General theory and design of control populations. *Anim. Breeding Abstr.*, **40**, 1-15.
- LATTER B. D. H., 1959. Genetic sampling in a random mating control population of constant size and sex ratio. *Aust. J. Biol. Sci.*, **12**, 500-505.
- WRIGHT S., 1931. Evolution in Mendelian populations. *Genetics*, **16**, 97-159.