

OPTIMISATION DU PLAN D'ACCOUPEMENT DANS LA SÉLECTION COMBINÉE

B. POUJARDIEU et R. ROUVIER

*Laboratoire de Méthodologie génétique,
Centre de Recherches de Toulouse, I.N.R.A.,
31 - Castanet-Tolosan*

RÉSUMÉ

Nous avons recherché l'optimum du plan d'accouplement de façon à maximiser l'espérance du progrès génétique à la suite d'une sélection combinée effectuée en choisissant les individus d'après leurs valeurs phénotypiques individuelles, moyenne de famille de mère et moyenne de famille de père. Nous avons alors cherché l'optimum du nombre d de mères accouplées à chaque père pour donner un certain nombre n de descendants mesurés, lorsque le produit nd de ces deux nombres, c'est-à-dire la taille de famille de père, est fixé ; cela pour différentes valeurs de ce produit pouvant correspondre sur le plan zootechnique à la sélection des lapins et porcins ou de la poule, pour des valeurs de l'héritabilité variant de 0,100 à 0,500 par pas de 0,100, du rapport p de la composante mère à la composante père de la variance, variant de 1 à 4 par pas de 1. Même pour les faibles valeurs du produit nd et donc de la taille de famille de père, il faut noter l'efficacité nettement plus grande de la sélection combinée par rapport à la sélection individuelle pour les faibles valeurs de l'héritabilité. Pour les faibles valeurs de l'héritabilité ($h^2 = 0,100$ à $0,200$), il convient d'augmenter n autant que possible, même si p est supérieur à 1. Pour des valeurs plus élevées de l'héritabilité, lorsque p est supérieur à 1, il est préférable d'augmenter la valeur de d .

Pour les valeurs élevées de nd (100 à 300, sélection de la volaille), il apparaît qu'il y a peu d'avantages à augmenter la taille de famille de père nd au-delà de 150. Dans la plupart des situations, il y a un optimum de n et de d qui correspond à une valeur d'autant plus grande de d que le rapport p est plus grand pour une héritabilité h^2 donnée, ou que h^2 est plus grande pour p donné. La valeur $d = 10$ souvent utilisée en pratique, répond à des conditions zootechniques et semble devoir être, en moyenne, pour plusieurs caractères de déterminisme génétique différents, une des plus proches de l'optimum.

INTRODUCTION

La définition de plans d'accouplement qui permettent de maximiser l'espérance du progrès génétique, conséquence d'une sélection à la suite d'un jugement sur descendance ou sur collatéraux a préoccupé un certain nombre de chercheurs. La plupart d'entre eux, tant pour les végétaux (COCHRAN, 1951 ; FINNEY, 1957) que pour

les animaux (DICKERSON et HAZEL, 1944, NORDSKÖG, 1959 ; RENDEL, 1959 ; ROBERTSON, 1957 et 1960 ; SMITH, 1960) raisonnent le problème en supposant limitée *a priori* la taille de la population soumise à sélection. POUTOUS et VISSAC (1962) fondent la recherche de l'optimum sur le coût des opérations de contrôle comparé à la rentabilité escomptée. Quel que soit le mode d'approche, l'optimum, s'il existe, procède d'un antagonisme entre la précision du jugement et l'intensité de la sélection en pratique réalisable. Il semble que seul LUSH (1947) ait étudié sans ces contraintes les variations d'efficacité de la sélection combinée en fonction de la taille de famille et de la corrélation intra-classe, pour un seul type de liens de parenté. DEMPSTER et LERNER (1947) dans leur recherche d'un plan d'accouplement chez la poule sélectionnée pour l'intensité de ponte étudient la variation d'efficacité en fonction de la durée de l'intervalle entre générations et de l'augmentation corrélatrice de la précision de l'estimation des valeurs génétiques. Ces auteurs utilisent, outre les valeurs individuelles, l'information apportée par le phénotype des seules pleines sœurs.

Chez les espèces polytoques, on peut choisir les individus d'après leur valeur pour un index de sélection combinant les valeurs phénotypiques individuelles, si les caractères sont mesurables sur l'animal vivant, et les valeurs moyennes des familles de père et de mère. Quel que soit le nombre de pères, il importe de définir le nombre de mères accouplées à chaque père et l'effectif des descendants à contrôler par mère de telle façon que l'efficacité d'une sélection par index soit maximum. Il nous a paru utile d'étudier ce problème et d'envisager les applications possibles à la sélection des oiseaux domestiques et des espèces de mammifères polytoques.

I. — CRITÈRE D'EFFICACITÉ DE LA SÉLECTION COMBINÉE POUR UN CARACTÈRE

I. I. — *Position du problème*

Dans un certain nombre de cas de sélection, il est possible d'estimer la valeur génétique additive d'un caractère chez un individu, à partir des valeurs phénotypiques de ce caractère chez cet individu chez des pleins frères ou sœurs, chez des demi-frères ou sœurs de père. Chez la poule, où les plans de sélection des souches habituellement mis en place permettent d'utiliser cette technique, ces plans sont tels que les productions de tous les animaux d'une même famille, et de toutes les familles, soient des productions contemporaines d'animaux ayant approximativement le même âge. Les générations sont donc séparées. Cela présente l'avantage de permettre la comparaison des productions des animaux dans les mêmes conditions de milieu et de rendre l'estimation des valeurs génétiques plus précise. L'utilisation de tels plans de sélection pourrait être également envisagée dans le cas des lapins et des porcins. Certes, lorsque les productions des individus à comparer ne sont pas contemporaines, on peut essayer d'obtenir des facteurs de correction pour les différents effets du milieu. Mais l'existence d'interactions éventuelles entre ces effets, ou entre génotype et milieu, rend parfois hasardeuse la correction des données. Nous considérerons ici le cas d'une sélection combinée pour des animaux dont les productions sont contemporaines et mesurées dans les mêmes conditions de milieu.

Le plan d'accouplement hiérarchisé est celui suivant lequel s pères ($s > 1$) ont

sailli chacun d mères ($d > 1$) et chaque mère a donné n descendants mesurés (ici, $n \geq 1$). La sélection s'effectue, à chaque génération, parmi les $N = sdn$ individus mesurés de l'échantillon total. Ce plan d'accouplement est utilisé pour la volaille, les lapins, les porcins. Pour des valeurs données des paramètres génétiques et statistiques, la précision de la sélection dépendra de la taille (dn) de famille de père, des nombres de mères (d) par père et de descendants (n) par mère. Nous étudierons la variation de cette précision (efficacité) en fonction de ces paramètres et, pour différentes valeurs des paramètres génétiques et phénotypiques. Nous étudierons la sélection pour un seul caractère et aborderons le cas de la sélection sur plusieurs caractères. En pratique, s'il est possible de maintenir le nombre d de mère par père à peu près constant, il n'en est pas de même pour n qui peut varier d'une famille à l'autre. Pour simplifier, nous considérerons cependant plusieurs séries de valeurs constantes de d et n . Nous n'étudierons pas l'incidence de l'intervalle entre générations et supposerons que les générations sont séparées. Nous envisagerons le cas d'une seule génération de sélection.

1. 2. — Index et efficacité de la sélection pour un caractère

Il convient d'estimer la valeur génétique additive G du caractère considéré. Cette estimation \hat{G} est l'index de sélection. Les individus de l'échantillon sont classés d'après les valeurs de \hat{G} et la sélection consiste à retenir pour la reproduction q p. 100 des individus ayant les valeurs les plus élevées de \hat{G} . Nous nous plaçons dans le cas simple où les reproducteurs, pères et mères, choisis au hasard dans une population très grande, sont sans parenté et non consanguins. Soit P_{jlm} la variable aléatoire phénotypique représentant la valeur susceptible d'être prise par le caractère chez le m^e descendant de la l^e mère accouplée au j^e père. Nous écrivons d'une part le modèle génétique classique

$$P_{jlm} = \mu + G_{jlm} + E_{jlm}$$

μ est une constante,

G_{jlm} est la variable aléatoire valeur génétique additive,

E_{jlm} est la variable aléatoire représentant les effets génétiques de dominance et d'épistasie, et du milieu. Les variables P , G et E sont supposées suivre conjointement une loi normale, G et E étant supposées indépendantes.

Par ailleurs, nous pouvons écrire le modèle statistique également classique (KEMP-THORNE, 1957).

$$P_{jlm} = \mu + s_j + d_{jl} + e_{jlm} \quad (I)$$

où s_j est l'effet du père j .

d_{jl} est l'effet de la mère jl .

e_{jlm} est un écart intra-famille de mère.

Les trois variables aléatoire introduites seront, par hypothèse, distribuées normalement $N\{0, \sigma_s^2\}$, $N\{0, \sigma_d^2\}$, $N\{0, \sigma_e^2\}$ et indépendantes. Cela entraîne notamment que les reproducteurs, pères et mères, de la population, sont non consanguins et non apparentés. L'estimateur \hat{G}_{jlm} de G_{jlm} est la régression de G sur les dn valeurs phénotypiques des individus de la famille de père numéro j . Suivant les hypothèses simples que nous avons prises sur les liens de parenté cela revient au même de considérer la régres-

sion de G sur les valeurs phénotypiques individuelles moyennes de famille de mère et de père, que l'on peut exprimer sous la forme

$$\hat{G}_{jlm} = \beta_1(P_{jlm} - P_{jl.}) + \beta_2(P_{jl.} - P_{j..}) + \beta_3(P_{j..} - \mu) \quad (2)$$

$\beta_1, \beta_2, \beta_3$, étant les coefficients de régression, $P_{jl.}$ et $P_{j..}$, les moyennes phénotypiques des familles de mère et père. Les valeurs de $\beta_1, \beta_2, \beta_3$ s'obtiennent d'après la théorie de la régression linéaire multiple, en résolvant le système d'équations représenté sous forme matricielle par :

$$\begin{bmatrix} \text{Var}(P_{jlm} - P_{jl.}) & \text{Cov}(P_{jlm} - P_{jl.})(P_{jl.} - P_{j..}) \\ & \text{Cov}(P_{jlm} - P_{jl.})(P_{j..} - \mu) \\ \text{Cov}(P_{jl.} - P_{j..})(P_{jlm} - P_{jl.}) & \text{Var}(P_{jl.} - P_{j..}) \\ & \text{Cov}(P_{jl.} - P_{j..})(P_{j..} - \mu) \\ \text{Cov}(P_{j..} - \mu)(P_{jlm} - P_{jl.}) & \text{Cov}(P_{j..} - \mu)(P_{j..} - \mu) \\ & \text{Var}(P_{j..} - \mu) \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} = \begin{bmatrix} \text{Cov } G_{jlm}(P_{jlm} - P_{jl.}) \\ \text{Cov } G_{jlm}(P_{jl.} - P_{j..}) \\ \text{Cov } G_{jlm}(P_{j..} - \mu) \end{bmatrix}$$

Ce système d'équation s'écrit aussi, compte tenu du modèle de l'équation (1) et des règles de la génétique des populations (voir par exemple MALÉCOT, 1966) permettant d'obtenir les éléments du vecteur second membre :

$$\begin{bmatrix} \frac{n-1}{n} \sigma_e^2 & 0 & 0 \\ c & \frac{d-1}{d} \sigma_d^2 + \frac{\sigma_e^2}{n} & 0 \\ 0 & 0 & \sigma_s^2 + \frac{1}{d} \sigma_d^2 + \frac{1}{nd} \sigma_e^2 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} = \begin{bmatrix} \frac{n-1}{2n} \sigma_g^2 \\ \frac{(d-1)(n+2)}{4nd} \sigma_g^2 \\ \frac{2+n(d+1)}{4nd} \sigma_g^2 \end{bmatrix}$$

où σ_g^2 est la variance génétique additive. On obtient alors les valeurs des coefficients de régression qui s'expriment en fonction des variances théoriques dans la population dont est issu l'échantillon, de d et n ;

$$\beta_1 = \frac{1}{2} \frac{\sigma_g^2}{\sigma_e^2} \quad \beta_2 = \frac{(n+2)}{4n} \frac{\sigma_g^2}{\sigma_d^2 + \frac{\sigma_e^2}{n}} \quad \beta_3 = \frac{2+n(d+1)}{4nd} \frac{\sigma_g^2}{\sigma_s^2 + \frac{1}{d} \sigma_d^2 + \frac{1}{nd} \sigma_e^2}$$

La méthode que nous venons d'utiliser pour déterminer les coefficients de régression présente l'intérêt d'être aisément généralisable à des cas plus complexes de sélection (ROUVIER, 1969 *a* et *b*) et de permettre une déduction simple des propriétés statistiques de l'index.

On sait que l'espérance du progrès génétique, lorsque l'on sélectionne q p. 100 des animaux ayant les valeurs les plus élevées de G , s'écrit

$$E(\Delta_G) = i \cdot R_{G\hat{G}} \cdot \sigma_G$$

avec $i = \frac{\Delta \hat{G}}{\sigma_{\hat{G}}}$; $\Delta \hat{G}$ est l'écart entre la moyenne des valeurs de \hat{G} dans le groupe sélectionné et la moyenne générale du toupeau avant la sélection.

Dans le cas de distribution normale, i est déterminé simplement à partir du pour-

centage d'individus sélectionnés. Ce pourcentage est conditionné par la nécessité de renouveler le cheptel, compte tenu des caractéristiques biologiques de la reproduction de l'espèce considérée et de contraintes zootechniques. On peut, sans nuire à la généralité du problème, chercher à maximiser l'espérance du progrès génétique pour i donné. Cela revient à maximiser $R_{G\hat{G}}$ ou $R^2_{G\hat{G}}$:

$$R^2_{GG} = \frac{\sigma_{\hat{G}}^2}{\sigma_G^2} = \frac{1}{\sigma_G^2} \left[\beta_1 \frac{n-1}{2n} \sigma_G^2 + \beta_2 \frac{(d-1)(n+2)}{4nd} \sigma_G^2 + \beta_3 \frac{2+n(d+1)}{4nd} \sigma_G^2 \right]$$

La précision et l'efficacité de l'index étant caractérisée par $R^2_{G\hat{G}}$, nous étudierons sa variation en fonction de nd , ainsi que de n et d pour nd fixé.

Les expressions de β_1 , β_2 , β_3 et $R^2_{G\hat{G}}$ ont été obtenues par OSBORNE (1957) par une méthode différente. Cet auteur a étudié les efficacités relatives de diverses méthodes de sélection de la poule : individuelle, familiale (familles de pleines sœurs, ou de pleines et demi-sœurs) et combinées. Cependant cet auteur se place dans le cas particulier où $\sigma_s^2 = \sigma_d^2 = \frac{1}{4} \sigma_G^2$. Cela exclut la possibilité de trouver un optimum pour d

et n , pour dn fixé. Pour de nombreux caractères la composante mère σ_d^2 de la variance peut être supérieure à la composante père et nous considérerons cette possibilité.

Nous avons supposé la normalité des distributions, ce qui implique la linéarité des liaisons entre valeurs génétique additive et phénotypique du caractère et donne la plus grande généralité à la formule (3), $R^2_{G\hat{G}}$ exprimant complètement la liaison entre G et \hat{G} . Si la distribution de \hat{G} est normale dans la génération initiale, en théorie elle ne le sera plus après la première génération de sélection. Cependant, dans les premières générations de sélection la distribution de \hat{G} restera voisine de la normale, sauf si le pourcentage d'individus sélectionnés est très faible, ou si l'héritabilité du caractère est très élevée. Notons cependant que la propriété la plus importante conditionnant la validité de l'index de sélection est celle de linéarité de la régression de G sur \hat{G} . \hat{G} est toujours un bon estimateur de G pourvu que les liaisons entre toutes les variables considérées soient linéaires.

II. — VARIATION DE L'EFFICACITÉ DE LA SÉLECTION COMBINÉE

La variation de $R^2_{G\hat{G}}$ a été étudiée en fonction de l'héritabilité du caractère, de la taille nd de famille de père, de la valeur du rapport $\frac{\sigma_d^2}{\sigma_s^2}$, de n et d lorsque les trois paramètres précédents sont fixés. Nous avons considéré une variance phénotypique unitaire,

$$\sigma_s^2 + \sigma_d^2 + \sigma_e^2 = 1 \quad (4)$$

L'héritabilité h^2 du caractère a été prise égale à

$$h^2 = \frac{4\sigma_s^2}{\sigma_s^2 + \sigma_d^2 + \sigma_e^2} \quad (5)$$

Il s'agit d'une définition de l'héritabilité dans un sens large qui tient compte d'une variance génétique d'épistasie éventuelle, due à l'interaction entre effets génétiques additif de gènes non allèles.

Nous avons pris $\sigma_g^2 = 4\sigma_s^2$, pour ne pas augmenter le nombre de paramètres. Lorsque cette égalité est vérifiée, il n'y a pas d'épistasie du type ci-dessus. En présence de cette épistasie, on aurait $\sigma_g^2 > 4\sigma_s^2$. Cette hypothèse ne devrait pas nuire à la généralité des résultats. On voit, d'après l'équation (2) et les expressions des coefficients de régression β , que σ_g^2 intervient comme un coefficient de proportionnalité dans l'expression du numérateur de R^2_{GG} . Pour des valeurs $\sigma_g^2 > 4\sigma_s^2$, les valeurs de R^2_{GG} seraient donc systématiquement plus faibles que celles qui vont être données. σ_d^2 peut être égal ou supérieur à σ_s^2 . Il est supérieur à σ_s^2 , notamment s'il y a une variation génétique de dominance et d'épistasie, une variance des effets maternels ou des effets milieu commun aux individus d'une même famille de mère. Nous avons considéré 4 valeurs p du rapport $\frac{\sigma_d^2}{\sigma_s^2} : p = 1, 2, 3, 4$. Ces valeurs de p doivent être compatibles avec les équations (3) et (4), sachant que σ_d^2 est au moins égal à $\frac{1}{4}\sigma_g^2$ et que σ_s^2 est au moins égal à $\frac{1}{2}\sigma_g^2$. Les valeurs de l'héritabilité considérées varient entre 0,100 et 0,500 par pas de 0,100. Nous avons considéré deux séries de valeurs du produit nd pouvant correspondre, sur le plan zootechnique, à la sélection combinée du lapin ou des porcins d'une part, de la volaille d'autre part.

2. 1. — Sélection des lapins ou des porcins

Le tableau 1 donne les valeurs des R^2_{GG} pour différents niveaux de $h^2, nd, p = \frac{\sigma_d^2}{\sigma_s^2}$,

Pour des valeurs de $h^2, \frac{\sigma_d^2}{\sigma_s^2}$ fixés, R^2_{GG} augmente avec nd . La corrélation R_{GG} augmente, en valeur relative, plus vite pour les faibles valeurs de h^2 que pour les fortes valeurs. Pour nd donné, R^2_{GG} augmente toujours avec n , nombre de pleins frères ou sœurs, dans le cas des faibles valeurs de l'héritabilité ($h^2 = 0,100$). Pour des valeurs plus élevées de l'héritabilité, il peut apparaître un optimum des valeurs de n et d , dans les limites de variation considérées des paramètres, pour certaines valeurs de $p > 1$. Le plus souvent, à partir de valeurs de p supérieures à 1 et d'autant plus faibles que l'héritabilité est plus grande, R^2_{GG} diminue lorsque d diminue. Pour des caractères d'héritabilité moyenne à élevée, et lorsque la composante mère de la variance est très supérieure à la composante père, il convient donc de ne pas trop diminuer le nombre de mères par père.

Pour les valeurs de $h^2 = 0,100$ à 0,400, l'efficacité de la sélection combinée, pour nd, n et d fixés, diminue lorsque le rapport $p = \frac{\sigma_d^2}{\sigma_s^2}$ augmente jusqu'à 4.

Pour $h^2 = 0,500$, cette efficacité diminue lorsque p augmente jusqu'à 3 et augmente lorsque p passe de 3 à 4. Cette constatation peut s'expliquer compte tenu de l'expression de R^2_{GG} . Elle est à rapprocher des résultats de LUSH (1947) qui a comparé les efficacités relatives de la sélection combinée et de la sélection individuelle en fonction de la corrélation intra-classe, dans le cas d'un seul type de famille.

Considérons la sélection des lapines sur leurs aptitudes maternelles dans un trou-

peau fermé d'effectif limité, la seule immigration pouvant être due à l'éventuelle introduction d'un petit nombre de reproducteurs mâles à chaque génération. Nous envisagerons le cas où les générations sont séparées et nous étudierons la répartition en famille qu'il semble souhaitable d'obtenir à chaque génération. Les aptitudes maternelles (prolificité, production laitière estimée par le poids de la portée à l'âge de 21 jours des lapereaux) sont des caractères à faible héritabilité ainsi que l'indiquent les résultats de ROLLINS *et al.*, (1963) pour la taille de portée. Les valeurs numériques du tableau 1 indiquent que dans le cas des caractères à faible héritabilité, il convient d'augmenter autant que possible les tailles de famille de père (valeur de nd). Pour une valeur donnée de nd , il faut chercher à augmenter n , même si le rapport de la composante mère de la variance à la composante père est élevé. Cependant, dès que h^2 atteint la valeur 0,200 et p la valeur 2, l'augmentation de $R^2_{G\hat{G}}$ est nettement plus grande lorsque n passe de 1 à 5 que lorsqu'il passe de 5 à 10 ou de 10 à 20. Pratiquement la taille des familles de pleines sœurs élevées et mesurées à chaque génération pourra être de $n = 5$ à 8. Si nous considérons, par exemple, la valeur $n = 5$, cela conduit à sélectionner 20 p. 100 des lapines ayant les valeurs les plus élevées des index de sélection et à élever pour la reproduction lors de la génération suivante, 5 filles de chacune de ces lapines sélectionnées. Ces 5 pleines sœurs seraient obtenues en deux portées successives, en moyenne, de la même lapine. Prendre une valeur plus élevée de n nécessite de retenir des filles dans plus de deux portées successives de la même femelle.

En ce qui concerne la sélection des femelles, pour un caractère relatif aux aptitudes maternelles, pour les faibles héritabilités ($h^2 = 0,100$) il convient de souligner l'importance de l'avantage de la sélection combinée par rapport à la sélection massale ou individuelle. De plus la répartition des femelles en familles permet une sélection des mâles pour les caractères qu'ils n'expriment pas, par un jugement sur descendance ou sur collatéraux, ce qui augmente encore le progrès génétique attendu.

Pour les caractères à héritabilité moyenne à élevée, en présence d'une variance génétique autre qu'additive, ou d'effets maternels, ou d'effets du milieu commun aux membres d'une même famille, il peut exister un optimum pour les valeurs de n et d . Ce cas peut correspondre à la sélection sur la vitesse de croissance et l'indice de consommation des jeunes animaux jusqu'à l'âge commercial d'abattage. La taille de famille de père peut facilement être de l'ordre de $nd = 30$, puisqu'il s'agit de caractères mesurés sur le jeune animal. L'examen du tableau 1 indique que les effectifs de $d = 5$ à 10 mères par père et $n = 6$ à 3 descendants par mère (ce qui correspond à la mesure des descendants d'une portée de chaque lapine) pourraient être un optimum dans un certain nombre de cas. En effet, pour les héritabilités moyennes à élevées, à partir de $p = 2$ ou 3, pour $nd = 30$, l'efficacité de l'index diminue lorsque d devient inférieur à 5.

Le raisonnement ci-dessus pourrait s'appliquer à la sélection d'un troupeau fermé de truies.

2. 2. — Sélection de la volaille

Le tableau 2 donne les valeurs de $R^2_{G\hat{G}}$ pour les mêmes niveaux de h^2 et valeurs de p que précédemment, mais pour des valeurs de nd , de n et d pouvant correspondre à la sélection de la volaille et notamment des poules pondeuses. Certaines valeurs de n et d prises en compte ne correspondent pas à une réalité zootechnique dans la sélection de la volaille, mais sont considérées seulement à titre indicatif pour mieux étu-

dier les variations de R^2_{GG} . On observe les mêmes tendances dans la variation de R^2_{GG} que dans le cas où la valeur de nd était plus faible. L'augmentation d'efficacité lorsqu'on passe de la sélection individuelle à la sélection combinée est beaucoup plus importante pour les faibles que pour les fortes valeurs de l'héritabilité. Pour h^2 donné, cette efficacité augmente avec la taille de famille nd , mais cette augmentation est faible lorsque nd passe de 150 à 200. Il y a donc peu d'avantages à augmenter la taille de famille de père au-delà de 150. Pour nd fixé, R^2_{GG} n'augmente constamment avec n , que pour $p = 1$ et toutes les valeurs de h^2 , et pour $p = 2$ et $h^2 = 0,100$. Dans la plupart des situations il y a un optimum de d et n qui correspond à une valeur d'autant plus grande de d que p est plus grand pour h^2 donné ou que h^2 est plus grand pour p fixé. Dans certains cas, R^2_{GG} est constamment décroissant lorsque n varie de 1 à 100. Comme précédemment on observe que dans certains cas, pour nd , n et d donnés, R^2_{GG} peut diminuer jusqu'à une certaine valeur et augmenter ensuite, lorsque p passe de 1 à 4. On peut considérer que pour la plupart des caractères zootechniques habituellement considérés et pour lesquels la composante mère est supérieure à la composante père, il y aura un optimum de d et n . La valeur de $d = 10$ utilisée souvent en pratique, répond à des conditions zootechniques, et semble devoir être, en moyenne, pour plusieurs caractères de déterminisme génétique différents, une des plus proches de l'optimum.

III. — SÉLECTION SUR PLUSIEURS CARACTÈRES

Nous considérons le cas de la sélection par index sur plusieurs caractères. Dans cette sélection, on cherche à augmenter la valeur d'un génotype global H défini comme

une combinaison linéaire $H = \sum_{i=1}^P a_i g_i$, de coefficients a_i donnés, des valeurs génétiques additives g_i des p caractères sur lesquels porte la sélection. HENDERSON (1963)

a montré que l'index de sélection sur plusieurs caractères peut s'écrire $I = \sum_{i=1}^P a_i \hat{g}_i$

l'estimateur \hat{g}_i étant tel que sa corrélation avec g_i soit maximum. La sélection s'effectue en retenant pour la reproduction q p. 100 des individus ayant les valeurs les plus élevées de I . L'espérance du progrès génétique s'écrit (HAZEL, 1943) :

$$E(\Delta_H) = i R_{HI} \sigma_H$$

avec $i = \frac{\Delta_I}{\sigma_I}$. Δ_I est l'écart entre la moyenne de I dans le groupe sélectionné et sa moyenne générale, σ_I est l'écart type de I , R_{HI} est la corrélation entre H et I , σ_H l'écart type de H .

Pour nd , n et d fixés, l'index I est tel que R_{HI} soit maximum. Pour nd donné on peut aussi rechercher la combinaison de n et d qui maximise R_{HI} . Lorsque on estime H par régression linéaire multiple sur les variables prédictrices, on peut écrire,

$R^2_{HI} = \frac{\sigma_I^2}{\sigma_H^2}$. Pour maximiser R^2_{HI} , par optimisation du plan d'accouplement, il faut donc maximiser σ_I^2 . On a :

$$\sigma_1^2 = \sum_{i=1}^P a_i^2 \sigma_{\hat{g}_i}^2 + 2 \sum_{\substack{i, i' \\ i < i'}} a_i a_{i'} \text{cov}(\hat{g}_i \hat{g}_{i'}) \quad (6)$$

Pour nd fixé, on peut penser qu'un optimum de n et d existe. On ne peut cependant pas facilement donner de règle générale. La variation de $\sum_{i=1}^P a_i^2 \sigma_{\hat{g}_i}^2$ peut se déduire aisément de celles de $R^2_{G_1 \hat{G}_1}$ en utilisant des tableaux établis suivant le même principe que les tableaux 1 et 2. Mais le meilleur estimateur \hat{G}_1 de G_1 tiendra compte des phénotypes de tous les caractères mesurés sur l'individu et ses apparentés et non plus d'un seul. La variation de la somme des doubles produits dans l'expression (6) dépend des coefficients a_i et des covariances $\text{cov}(\hat{g}_i \hat{g}_{i'})$. La recherche du plan d'accouplement optimum, dans la sélection sur plusieurs caractères, nécessite donc d'obtenir notamment la matrice des variances et covariances des estimateurs des valeurs génétiques additives pour chaque valeur de nd , n et d , et de calculer dans chaque cas la valeur numérique de la forme quadratique de l'équation (6).

CONCLUSION

Montrer l'intérêt de la recherche du nombre de femelles accouplées à chaque père et du nombre de descendants à mesurer par mère pour maximiser l'efficacité d'une sélection combinée pour des caractères mesurables sur l'animal vivant, est le but de cette étude. Nous l'avons abordé par le calcul numérique en nous limitant, par souci de simplification, à l'étude approfondie de la sélection pour un seul caractère. De plus, puisque nous avons utilisé les propriétés de la loi normale, nous nous sommes limités à la première génération de sélection. Une simulation sur ordinateur permettrait de s'affranchir de l'hypothèse de normalité et d'étendre l'étude à plusieurs générations successives de sélection.

Reçu pour publication en juillet 1971

SUMMARY

BREEDING PLAN OPTIMIZATION IN COMBINED SELECTION

We have tried to define the optimum of a breeding plan in order to encourage genetic progress to a maximum after combined selection, choosing individuals according to their individual phenotypic values such as maternal family average and paternal family average. We then tried to find the optimum number d of mothers bred with each father to give a certain number n of progeny measured, when the product nd of these two numbers (size of paternal family) is fixed. This was done on an animal husbandry level for different values of this product in the selection of rabbits and pigs or hens for heritability values (ranging from 0.100 to 0.500 by thresholds of 0.100) of the ratio p of the maternal component to the paternal variance component which varied from 1 to 4 by thresholds of 1. Even for low values of the product nd (paternal family size), combined election is clearly more effective than individual selection for low heritability values. It is advisable to increase n as much as possible for low heritability value ($h^2 = 0.100-0.200$), even if p is higher than 1. When p is higher than 1, it is preferable to increase the value of d for higher heritability values.

For high nd values (100-300, poultry selection), it does not seem advantageous to increase paternal family size nd over 150. In most situations, there is an optimum for n and d corresponding to a d value which is greater when the ratio p is greater for a given h^2 heritability, or when h^2 is greater for a given p . The value $d = 10$ often used in practise, satisfies animal husbandry conditions and in general seems to be one of the closest to the optimum for several different genetic determinism characters.

RÉFÉRENCES BIBLIOGRAPHIQUES

- COCHRAN W.G., 1951. Improvement by means of selection in *Proceedings of the second Berkeley symposium on mathematical statistics and probability*. Berkeley University.
- DEMPSTER E. R., LERNER I. M., 1947. The optimum structure of breeding flocks. I. Rate of genetic improvement under different breeding plans. *Genetics*, **32**, 555-579.
- DICKERSON G. E., HAZEL L. N., 1944. Effectiveness of selection on progeny testing performance as supplement to earlier culling in livestock. *J. Agric. Res.*, **69**, 459-476.
- FINNEY D. J., 1957. Statistical problems of plant selection. *Proc. 30th Session of international Statistical Institute*.
- HAZEL L. N., 1943. The genetic basis for constructing selection indices. *Genetics*, **28**, 476-490.
- HENDERSON C. R., 1963. Selection index and expected genetic advance. In *Statistical Genetics and Plant Breeding*. Edited by W. D. HANSON and H. F. ROBINSON. Natl. Acad. Sci. Nat. Res. Council. Publ., **982**, 141-163.
- KEMPTHORNE O., 1957. *An introduction to genetic statistics.*, p. 235-246, p. 510-514. John Wiley and Sons, New York.
- LUSH J. L., 1947. Family merit and individual merit as bases for selection. *Amer. Naturalist*, **81**, 241-261, 362-379.
- MALÉCOT G., 1966. *Probabilités et hérédité*. 1 vol. 356 p. Institut national d'Études démographiques et P. U. F. Paris.
- NORDSKOG A. W., 1959. Note on optimum group size for progeny tests. *Biometrics*, **15**, 513-517.
- OSBORNE R., 1957. The use of sire and dam family averages in increasing the efficiency of selective breeding under hierarcial mating system. *Heredity*, **11**, 93-116.
- POUTOUS M., VISSAC B., 1962. Recherche théorique des conditions de rentabilité maximum de l'épreuve de descendance des taureaux d'insémination artificielle. *Ann. Zootech.*, **11**, 233-256.
- RENDEL J. M., 1959. Optimum group size in half sib family selection. *Biometrics*, **15**, 376-81.
- ROBERTSON A., 1957. Optimum group size in progeny testing and family selection. *Biometrics*, **13**, 442-450.
- ROBERTSON A., 1960. On optimum family size in Selection Programmes. *Biometrics*, **16**, 296-298.
- ROLLINS W. C., CASADY R. B., SITTMANN K., D. B., 1963. Genetic variance component analysis of litter size and weaning weight of *New Zealand White* rabbits. *J. Anim. Sci.*, **22**, 654-657.
- ROUVIER R., 1969 a. *Contribution à l'étude des index de sélection sur plusieurs caractères*. Thèse doct. 3^e cycle. Fac. Sci. Paris. 92 p.
- ROUVIER R., 1969 b. Calcul des index de sélection dans le cas d'individus consanguins et apparentés quelconques. *Ann. Génét. Sel. anim.*, **1**, 67-70.
- SMITH C., 1960. Efficiency of animal testing schemes. *Biometrics*, **16**, 408-415.