



THÉORIE ET CALCUL DES INDEX DE SÉLECTION

J. P. Boyer

► To cite this version:

J. P. Boyer. THÉORIE ET CALCUL DES INDEX DE SÉLECTION. Annales de zootechnie, 1958, 7 (3), pp.193-242. <hal-00886689>

HAL Id: hal-00886689

<https://hal.science/hal-00886689v1>

Submitted on 11 May 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

THÉORIE ET CALCUL DES INDEX DE SÉLECTION

PAR

J. P. BOYER

Station de Recherches Avicoles, Jouy-en-Josas.

PLAN DU MÉMOIRE ⁽¹⁾

Introduction.

Construction des index.

I. — Sélection pour un seul caractère.

- A. — Sélection massale.
- B. — Sélection familiale.
- C. — Sélection parentale.
- D. — Sélection pédigrée.
- E. — Sélection combinée.
- Conclusions.

II. — Sélection de plusieurs caractères.

- A. — Sélection massale.
- B. — Sélection combinée.
- C. — Sélection à partir des index partiels.
- D. — Généralisation des index.

Conclusions.

Note.

Bibliographie.

INTRODUCTION

Le but que poursuit tout sélectionneur est d'augmenter la productivité de son troupeau par l'amélioration de son patrimoine héréditaire ou génotype moyen. Le choix des reproducteurs, pierre d'achoppement de cette amélioration, se ramène donc à l'estimation du génotype de chaque animal, qui est *a priori* susceptible de participer à la reproduction.

C'est précisément cette estimation — ou du moins une valeur y proportionnelle — que les index de sélection se proposent de résoudre, au sens mathématique du terme, à partir d'un certain nombre d'observations.

On supposera dans ce qui suit des données telles qu'elles permettent

⁽¹⁾ Les dessins et graphiques ont été établis par M^{me} Vincent. La figure 5 est l'œuvre de M. Ferré. Qu'ils en soient ici remerciés.

l'analyse de la variance (condition d'indépendance de la moyenne et de l'écart-type). Dans la majorité des cas, il s'agira de mesures distribuées normalement. Lorsque cette condition n'est pas remplie, il est possible de normaliser les données par un changement de variable approprié : transformation arc sinus pour les pourcentages (FISHER), transformation probit, d'une portée plus générale (BLISS, 1935 ; COCHEZ, 1954).

On supposera par ailleurs que la variance phénotypique est égale à la somme des variances génotypique et due au milieu (condition d'indépendance du génotype et du milieu), soit : $\sigma_p^2 = \sigma_g^2 + \sigma^2$. Cette condition est loin d'être respectée dans tous les cas. Toutefois, elle constitue une hypothèse de travail commode en attendant que des recherches nouvelles permettent de préciser les valeurs des corrélations et des interactions génotype-milieu ; les premières étant d'ailleurs approximativement éliminées par la randomisation des conditions d'élevage.

Enfin, pour raison de simplification, on ne se préoccupera que de la fraction additive de l'action des gènes, la régression du génotype sur les mesures enregistrées étant alors linéaire. Cette hypothèse restreint donc le champ d'application des index, plus précisément leur maximum d'efficacité, aux cas où les écarts (dus à la dominance et à l'épistasie) du génotype vrai à la régression sont relativement faibles. Néanmoins, le fractionnement de l'action génique est au moins en partie un artifice de calcul et il est possible d'augmenter la portion additive de la variance génotypique par un changement de variable adéquat : racine carrée, logarithme, probit, etc... bien qu'on ne puisse donner d'indication systématique (LERNER, 1950 ; COCHEZ, 1957).

Dans ces conditions, l'index de sélection dérivera d'une façon simple de l'équation de régression multiple du génotype sur les données :

$$\boxed{\hat{G} - \bar{G} = k_A(A - \bar{A}) + k_B(B - \bar{B}) + \dots + k_Z(Z - \bar{Z})} \quad (1)$$

\hat{G} étant la valeur probable du génotype estimé et \bar{G} le génotype moyen du troupeau ; A, B, \dots, Z sont des données relatives à l'animal étudié pour divers caractères — ces données pouvant être performances individuelle ou parentale : ascendants, descendants ou collatéraux. $\bar{A}, \bar{B}, \dots, \bar{Z}$, sont les moyennes correspondantes du troupeau — ou de la souche, de la race, de l'espèce. Enfin, k_A, k_B, \dots, k_Z sont les coefficients de régression multiple.

Notons qu'il est facile de simplifier cette équation en lui substituant une valeur équivalente par suppression des termes constants. L'expression de cette nouvelle valeur, à laquelle on réservera le nom d'index, dépend de la méthode de sélection utilisée. Nous la signalerons dans chaque cas particulier, mais raisonnerons toujours à partir de l'équation (1) pour lui conserver son caractère de généralité.

Vue sous cet angle, la théorie des index permet de grouper sous une forme logique les différentes méthodes de sélection usuelles : massale, familiale, combinée, sur la descendance, ou l'ascendance — et leurs divers arrangements, pour un ou plusieurs caractères.

Il nous a paru utile d'effectuer une telle synthèse, sur des bases plus larges que celles utilisées par HAZEL (1943) et ROBERTSON (1955), tout en développant l'aspect analytique de l'élaboration des index qui est en somme une application directe à la régression multiple de la théorie des « path-coefficients » de SEWALL WRIGHT (1923), CHING-CHUN-LI (1948). Par contre, nous avons supposé connue l'estimation des paramètres fondamentaux, notamment héritabilité et corrélation génétique (cf. à ce sujet WINSOR et CLARKE, 1940 ; LUSH, 1940 ; WHATLEY, 1942 ; HAZEL, 1943 ; HAZEL, BAKER et REINMILLER, 1943 ; LUSH, 1947 et 1948 ; ROBERTSON et LERNER, 1949 ; LERNER, 1950 ; KING et HENDERSON, 1954 ; JÉROME et al., 1956) et volontairement limité la discussion à l'efficacité des diverses méthodes de sélection ; ceci afin d'éviter une trop grande confusion dans un domaine non encore exploré totalement, mais de dégager au contraire le raisonnement mathématique servant d'ossature à la construction des index.

CONSTRUCTION DES INDEX

I. — SÉLECTION POUR UN SEUL CARACTÈRE

A. — Sélection massale

On estime le génotype d'un animal à partir de la performance brute — ou phénotype P , de moyenne \bar{P} pour l'ensemble du troupeau. Le diagramme de causalité se réduit alors à sa plus simple expression (fig. 1) et l'équation de régression s'écrit :

$$\hat{G} - \bar{G} = k_P(P - \bar{P}).$$

La figure 1 indique que p , path-coefficient de G par P , est aussi la corrélation r_{GP} entre le génotype et le phénotype, celui-ci étant considéré comme cause « indépendante » de celui-là. Mais, par définition du path-coefficient, on a :

$$p = k_P \frac{\sigma_P}{\sigma_G} \quad \text{d'où} \quad k_P = p \frac{\sigma_G}{\sigma_P} = r_{GP} \frac{\sigma_G}{\sigma_P}.$$

La corrélation r_{GP} peut se calculer directement à partir de la relation additive entre le phénotype P , le génotype G et le milieu E : $P = G + E$, G et E étant indépendants. On a ainsi $r_{GP} = \frac{\sigma_G}{\sigma_P}$. Par suite, $k_P = \frac{\sigma_G^2}{\sigma_P^2}$, ce

qui est la définition même de l'héritabilité h^2 . L'équation de régression pour la sélection massale est donc simplement :

$$\boxed{\hat{G} - \bar{G} = h^2(P - \bar{P})} \quad (2)$$

On voit que cette estimation du génotype est fidèle si l'héritabilité du caractère sélectionné est voisine de l'unité. Dans ce cas, en effet, la corrélation r_{GP} est très forte : le phénotype est l'image presque exacte du génotype.

On voit en outre que $\hat{G} - \bar{G}$ n'est autre que l'écart du génotype estimé au génotype moyen, c'est-à-dire le gain génétique ΔG selon la notation de LERNER (1950) ; $P - \bar{P}$ mesure l'intensité i de la sélection effectuée sur les phénotypes. De sorte que l'on peut écrire symboliquement :

$$\Delta G = h^2 i.$$

Cette égalité, que l'on peut appeler « Loi de la sélection massale » exprime que le gain obtenu en une génération sera d'autant plus grand

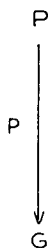


FIG. 1.

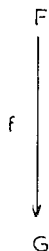


FIG. 2.

que la proportion d'animaux gardés pour la reproduction sera plus faible et que l'héritabilité sera plus forte. Sous réserve de cette dernière remarque, qui rappelle l'efficacité limitée de la sélection massale, on peut substituer à l'équation (2) une valeur proportionnelle par suppression des termes constants. Dans ces conditions, l'index de sélection massale sera purement et simplement la performance observée :

$$\boxed{I = P}.$$

B. — Sélection familiale

Nous définissons la « famille » comme un groupe d'animaux ayant une parenté génétique r constante lorsqu'on les compare deux à deux. L'estimation du génotype individuel s'effectue à partir de la moyenne de famille F , de moyenne \bar{F} pour l'ensemble des familles composant le

troupeau. L'équation de régression du génotype sur la moyenne de famille est donc :

$$\hat{G} - \bar{G} = k_F(F - \bar{F}).$$

D'après la figure 2, on a

$$f = k_F \frac{\sigma_F}{\sigma_G} \quad \text{d'où} \quad k_F = f \frac{\sigma_G}{\sigma_F}.$$

Le path-coefficient f de G par F est aussi la corrélation r_{GF} entre le génotype de l'individu étudié et la moyenne de famille, celle-ci étant

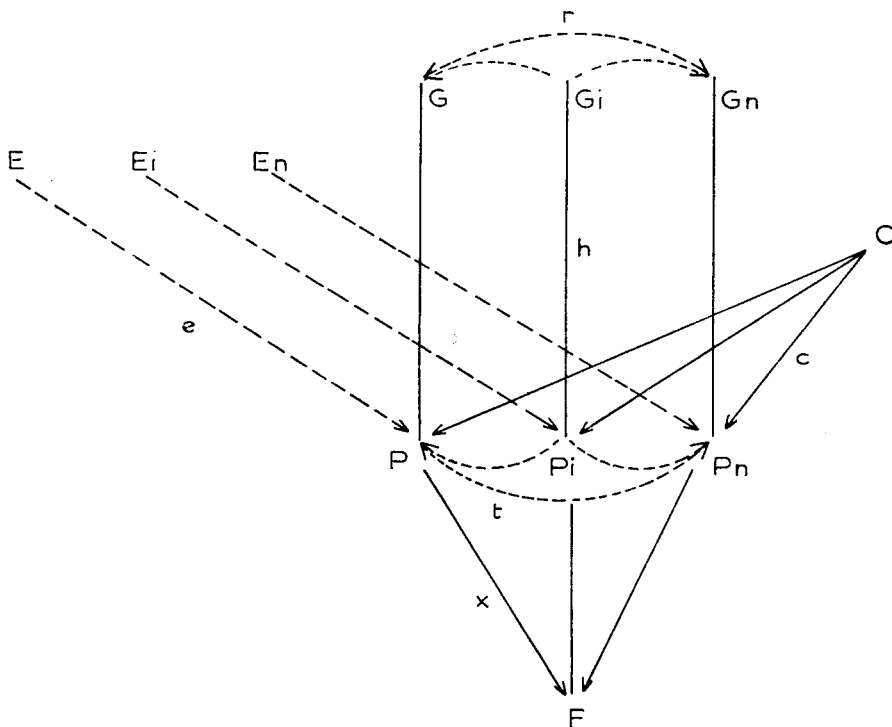


FIG. 3. — Diagramme de causalité de la moyenne de famille. GG_iG_n sont les génotypes de l'animal étudié, du $i^{\text{ème}}$ et du $n^{\text{ème}}$ membres de sa famille ; PP_iP_n sont les phénotypes correspondants, servant au calcul de la moyenne F . r est la parenté génétique, t la corrélation phénotypique intra-famille, h la corrélation génotype-phénotype, x le path-coefficient de F par chacun des phénotypes individuels, c le path-coefficient du phénotype par le milieu C commun aux membres de la famille. On a donc $t = rh^2 + c^2$ (voir note à la fin pour le calcul direct de t). Enfin, e est le path-coefficient du phénotype par le milieu E, E_i, E_n différent pour chaque individu. De sorte que l'équation de détermination du phénotype P est simplement : $h^2 + e^2 + c^2 = 1$.

considérée ici comme cause « indépendante » de celui-là. On retrouverait algébriquement ce résultat en notant que par définition de coefficient de régression

$$k_F = r_{GF} \frac{\sigma_G}{\sigma_F}, \quad \text{d'où} \quad f = r_{GF}.$$

Le diagramme de causalité représenté figure 3 permet de calculer rapidement cette corrélation.

On a en effet

$$r_{GF} = hx + (n-1)rhx = hx[1 + (n-1)r].$$

Mais, d'autre part, l'équation de détermination de F s'écrit :

$$1 = nx^2 + 2 \frac{n!}{2!(n-2)!} tx^2 = nx^2[1 + (n-1)t].$$

D'où l'on tire $x = \frac{1}{\sqrt{n[1 + (n-1)t]}}$ et par suite $r_{GF} = \frac{1 + (n-1)r}{\sqrt{n[1 + (n-1)t]}} h$

Il reste à calculer le rapport σ_G/σ_F . Par définition de la moyenne,

$$F = \frac{1}{n}(P + \dots + Pi + \dots + Pn).$$

Par suite, la variance de la moyenne de famille est égale à :

$$\sigma_F^2 = \frac{1}{n^2}(n\sigma_P^2 + 2 \frac{n!}{2!(n-2)!} t\sigma_P^2) = \frac{1 + (n-1)t}{n} \sigma_P^2.$$

On a par conséquent :

$$\frac{\sigma_P^2}{\sigma_F^2} = \frac{n}{1 + (n-1)t} = \frac{\sigma_G^2}{h^2 \sigma_P^2}$$

par définition de $h^2 = \frac{\sigma_G^2}{\sigma_P^2}$

On peut donc en tirer :

$$\frac{\sigma_G}{\sigma_F} = h \sqrt{\frac{n}{1 + (n-1)t}}.$$

L'équation de régression du génotype sur la moyenne de famille devient alors :

$$\hat{G} - \bar{G} = h \frac{1 + (n-1)r}{\sqrt{n[1 + (n-1)t]}} \cdot h \sqrt{\frac{n}{1 + (n-1)t}} (F - \bar{F}).$$

Ou plus simplement :

$$\boxed{\hat{G} - \bar{G} = h^2 \frac{1 + (n-1)r}{1 + (n-1)t} (F - \bar{F})} \quad (3)$$

Il est facile de voir que le coefficient de régression ainsi trouvé n'est autre que l'héritabilité h_F^2 de la moyenne de famille. En effet, la fraction génétique σ_{GF}^2 de la variance de F est égale à $\frac{h^2 + (n-1)rh^2}{n} \sigma_P^2$ puisqu'on a $t = rh^2 + c^2$ et $h^2 + e^2 + c^2 = 1$.

Par suite,

$$h_F^2 = \frac{\sigma_{GF}^2}{\sigma_F^2} = \frac{1 + (n-1)r}{1 + (n-1)t} h^2 \quad \text{C. Q. F. D.}$$

De sorte que l'on peut écrire la loi de la sélection familiale de façon symbolique, comme précédemment, ΔGF étant le gain génétique obtenu et i_F l'intensité de la sélection :

$$\Delta GF = h_F^2 \cdot i_F.$$

Efficacité de la sélection familiale.

Le coefficient de corrélation r_{GF} mesure la précision avec laquelle la moyenne de famille rend compte du génotype individuel. Ce coefficient est relativement complexe car il dépend de quatre paramètres : n , r , h^2 et c^2 . Il est toutefois intéressant de chiffrer l'efficacité de la sélection familiale par rapport à la sélection massale en comparant les gains génétiques procurés par l'une et l'autre méthodes. D'une manière générale, X étant le critère de sélection utilisé, le gain génétique s'écrit :

$$\Delta_{GX} = r_{GX} \cdot \frac{\sigma_G}{\sigma_X} (X - \bar{X}).$$

A l'intensité de sélection égale, on a évidemment :

$$\frac{X - \bar{X}}{\sigma_X} = \frac{F - \bar{F}}{\sigma_F} = \frac{P - \bar{P}}{\sigma_P} = \frac{i}{\sigma_P}.$$

De sorte que l'on a l'égalité :

$$\Delta_{GX} = r_{GX} \cdot hi.$$

Dans le cas présent, l'efficacité de la sélection familiale est donc mesurée par le rapport

$$\varepsilon = \frac{\Delta GF}{\Delta G} = \frac{r_{GF}}{r_{GP}},$$

soit :

$$\varepsilon = \frac{1 + (n-1)r}{\sqrt{n[1 + (n-1)t]}}.$$

En admettant que l'effet c^2 inclus dans t est négligeable on trouve facilement que ce rapport est égal à l'unité pour $h^2 = [nr^2 - (1-r)^2]/nr$; soit pour $r = 1/2$, $h^2 = (n-1)/2n$ et pour $r = 1/4$, $h^2 = (n-9)/4n$. On peut donc préciser, pour un effectif familial donné n , les valeurs de l'hérabilité au-dessous desquelles il convient de préférer la sélection familiale à la sélection massale. Soit pour des familles de frères/sœurs ($r = 1/2$) ou de demi-frères/sœurs ($r = 1/4$) :

| | | | | | | | | | | | | | | | | |
|-----------|-----------|-------|---|----|----|----|----|----|----|----|----|-----|----|----|----|----------|
| | | $n =$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 15 | 20 | 50 | ∞ |
| $i = 1/2$ | % $h^2 =$ | | 0 | 25 | 33 | 38 | 40 | 42 | 43 | 44 | 44 | 45 | 47 | 48 | 49 | 50 |
| $r = 1/4$ | % $h^2 =$ | | — | — | — | — | — | — | — | — | 0 | 2,5 | 10 | 14 | 21 | 25 |

En somme, la sélection familiale est avantageuse pour les caractères relativement peu héréditaires et des familles assez nombreuses (fig. 4).

L'utilisation de cette méthode avec des familles de demi-frères/sœurs peut être illustrée par la sélection pour la rusticité ($\% h^2 \neq 10$) des familles de pères. Mais on est alors extrêmement limité par la sélection différentielle si le troupeau est issu d'un petit nombre de mâles, comme c'est généralement le cas.

Il est intéressant d'autre part, d'étudier l'efficacité de la sélection familiale en fonction de l'effectif considéré. Aux valeurs extrêmes, pour

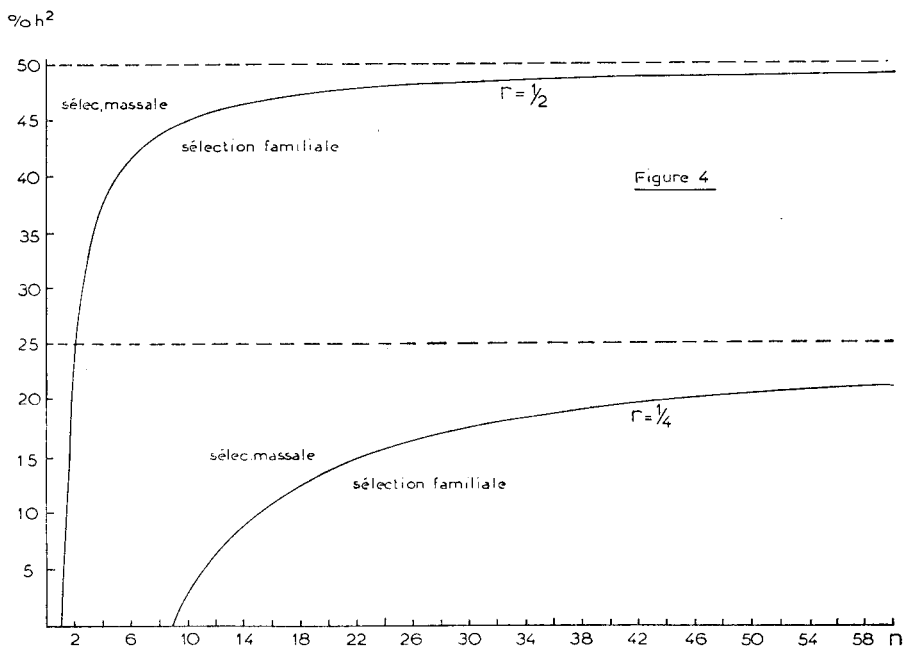


FIG. 4. — Dans tous les cas où $h^2 > 50\%$, la sélection massale est plus efficace que la sélection familiale. Par contre, celle-ci est d'autant plus intéressante que la parenté r est plus étroite et la ressemblance phénotypique t plus faible : c'est-à-dire lorsque h^2 est petit.

$n = 1$ on retombe dans le cas de la sélection massale avec une efficacité $\varepsilon = 1$; pour un effectif infini, ε tend vers la limite r/\sqrt{t} qui n'est pas forcément supérieure à l'unité. Contrairement à ce que l'on croit souvent *a priori*, nous voyons donc que la sélection familiale n'est pas obligatoirement d'autant plus intéressante que l'effectif est plus nombreux. Si l'on admet comme précédemment que l'effet c^2 est négligeable, l'efficacité limite devient \sqrt{r}/h , valeur supérieure à 1 pour $r > h^2$ et qui correspond à la zone située au-dessous des courbes représentées sur la figure 4.

Enfin, il est possible de représenter l'efficacité ε en fonction de deux paramètres, par exemple l'héritabilité h^2 et l'effectif de la famille n , en admettant un effet c^2 nul et une parenté $r = 1/2$ (fig. 5).

Soit $\varepsilon = \frac{n+1}{\sqrt{2n[2+(n-1)h^2]}}$. On trouve, en annulant sa dérivée

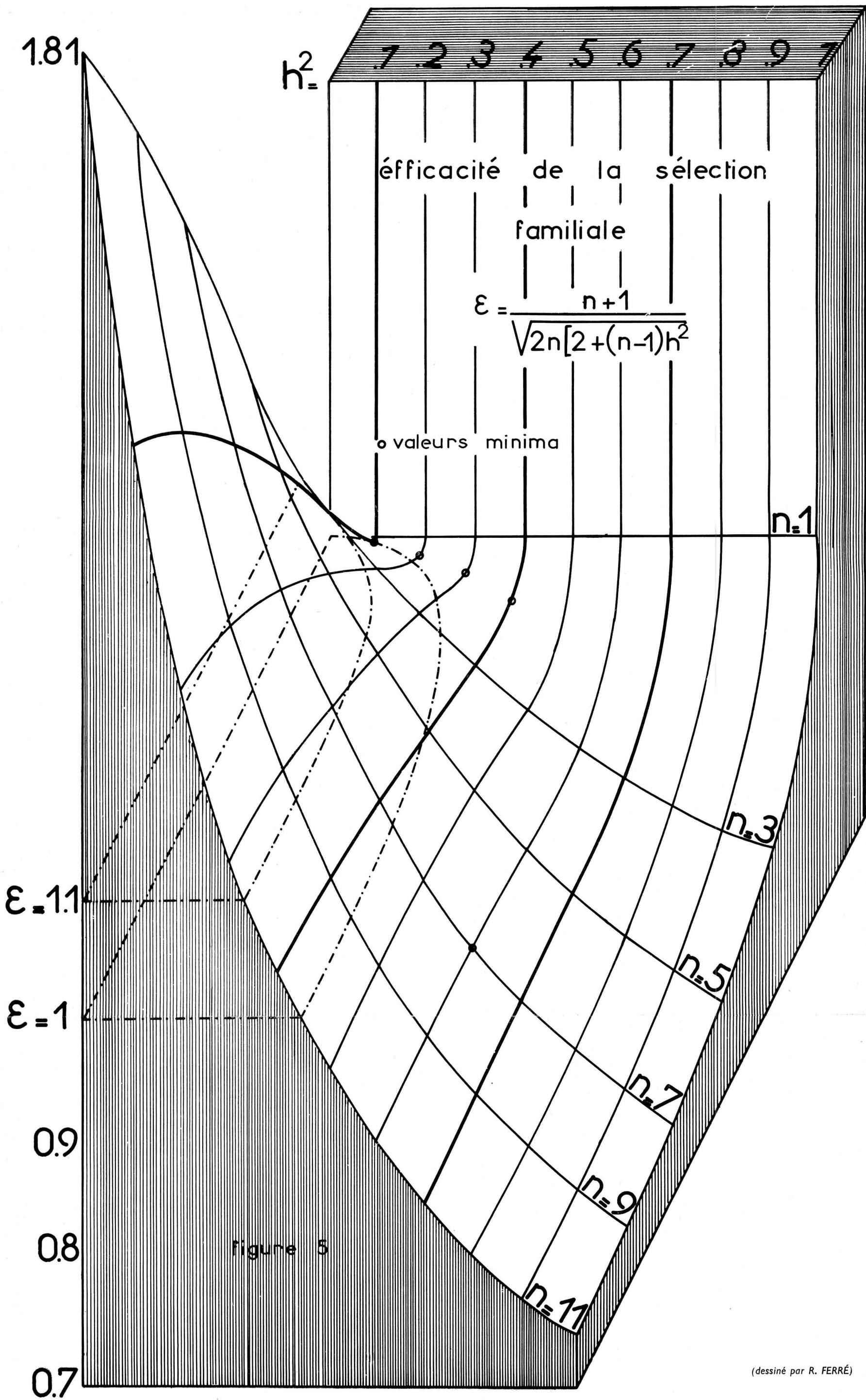


FIG. 5. — Efficacité de la sélection familiale par rapport à la sélection massale pour une parenté $r = 1/2$, en admettant que $t = rh^2$.
 Noter que l'efficacité ε passe par un minimum pour une valeur donnée de h^2 inférieure à $2/3$, minimum atteint pour un effectif $n = (2 - h^2)/(2 - 3h^2)$.

par rapport à n que l'efficacité passe par un minimum pour $n = \frac{2 - h^2}{2 - 3h^2}$ qui nécessite la condition $h^2 < 2/3$ pour que n soit positif. La valeur de ce minimum est alors $\varepsilon = \frac{2\sqrt{1 - h^2}}{2 - h^2}$.

L'efficacité minimum est obtenue pour les valeurs suivantes de n et h^2 :

| | | | | | | | |
|---------|-----|-----|-----|-----|-----|-----|----------|
| $h^2 =$ | 0,1 | 0,2 | 0,3 | 0,4 | 0,5 | 0,6 | 0,66 |
| $n =$ | 1,1 | 1,3 | 1,6 | 2 | 3 | 7 | ∞ |

Si l'on remarque que pour $h^2 > 0,5$, la sélection massale est préférable à la sélection familiale, que d'autre part, il n'est pas possible d'avoir moins de deux animaux pour calculer une moyenne, on peut conclure qu'il vaut mieux n'utiliser la sélection familiale qu'avec des familles d'au moins 4 individus.

Notons enfin que pour $h^2 = 0,4$, l'efficacité est égale à 1,04 pour une dizaine de frères/sœurs, ordre de grandeur pratiquement obtenu chez les sélectionneurs de volailles. Cette efficacité est suffisamment faible pour justifier la sélection massale des caractères ayant une héritabilité supérieure à 40 %. La valeur limite de ε pour un effectif infini est alors $1/h\sqrt{2}$, qui n'est supérieure à l'unité que pour $h^2 < 0,50$.

Index de sélection familiale

Compte tenu de la discussion précédente qui permet de conclure à l'opportunité ou à l'importunité de la sélection familiale, il est facile de substituer à l'équation de régression (2) une valeur simplifiée équivalente qui sera l'index de sélection familiale proprement dit.

Les moyennes de famille étant exprimées en écarts-réduits, on a évidemment $\bar{F} = 0$. D'autre part, l'héritabilité étant donnée et par suite constante pour le caractère étudié, il suffit de connaître le rap-

port $\frac{1 + (n - 1)r}{1 + (n - 1)t}$.

La parenté r peut être choisie à l'avance ($r = 1/2$ pour des familles de frères/sœurs, $r = 1/4$ pour des familles de demi-frères/sœurs). Il est alors simple de construire un jeu de tables donnant directement l'index de sélection familiale :

$$I_F = \frac{1 + (n - 1)r}{1 + (n - 1)t} F.$$

De tels exemples d'index sont donnés dans le tableau I (avec $r = 1/2$ et $t = 0,1$) et le tableau II (avec $r = 1/4$ et $t = 0,5$). Dans le cas parti-

culier où $t = r$ l'index devient simplement $I = F$. De même, dans les familles réduites à un seul individu, I est alors égal à la performance individuelle P et l'on est ramené à l'index de sélection massale, ce qui est évident.

TABLEAU I

| F = | 0,1 | 0,2 | 0,4 | 0,6 | 0,8 | I | 1,2 | 1,4 | 1,6 | 1,8 | 2,0 | 2,2 | 2,4 | 2,6 |
|--------|------|------|------|------|------|------|------|------|------|------|------|------|-------|-------|
| $n =$ | | | | | | | | | | | | | | |
| 1 ... | 0,10 | 0,20 | 0,40 | 0,60 | 0,80 | 1,00 | 1,20 | 1,40 | 1,60 | 1,80 | 2,00 | 2,20 | 2,40 | 2,60 |
| 2 ... | 0,14 | 0,27 | 0,55 | 0,82 | 1,09 | 1,36 | 1,64 | 1,91 | 2,18 | 2,45 | 2,73 | 3,00 | 3,27 | 3,55 |
| 3 ... | 0,17 | 0,33 | 0,67 | 1 | 1,33 | 1,67 | 2,00 | 2,33 | 2,67 | 3,00 | 3,33 | 3,67 | 4,00 | 4,33 |
| 4 ... | 0,19 | 0,39 | 0,77 | 1,16 | 1,54 | 1,92 | 2,31 | 2,69 | 3,08 | 3,46 | 3,85 | 4,23 | 4,62 | 5,00 |
| 5 ... | 0,21 | 0,43 | 0,86 | 1,29 | 1,71 | 2,14 | 2,57 | 3,00 | 3,43 | 3,86 | 4,29 | 4,71 | 5,14 | 5,57 |
| 6 ... | 0,23 | 0,47 | 0,93 | 1,40 | 1,87 | 2,33 | 2,80 | 3,27 | 3,73 | 4,20 | 4,67 | 5,13 | 5,60 | 6,07 |
| 7 ... | 0,25 | 0,50 | 1 | 1,50 | 2,00 | 2,50 | 3,00 | 3,50 | 4,00 | 4,50 | 5,00 | 5,50 | 6,00 | 6,50 |
| 8 ... | 0,26 | 0,53 | 1,06 | 1,59 | 2,12 | 2,65 | 3,18 | 3,71 | 4,24 | 4,77 | 5,29 | 5,82 | 6,35 | 6,88 |
| 9 ... | 0,28 | 0,56 | 1,11 | 1,67 | 2,22 | 2,78 | 3,33 | 3,89 | 4,44 | 5,00 | 5,56 | 6,11 | 6,67 | 7,22 |
| 10 ... | 0,29 | 0,58 | 1,16 | 1,74 | 2,32 | 2,90 | 3,47 | 4,05 | 4,63 | 5,21 | 5,79 | 6,37 | 6,95 | 7,53 |
| 11 ... | 0,30 | 0,60 | 1,20 | 1,80 | 2,40 | 3,00 | 3,60 | 4,20 | 4,80 | 5,40 | 6,00 | 6,60 | 7,20 | 7,80 |
| 12 ... | 0,31 | 0,62 | 1,24 | 1,86 | 2,48 | 3,10 | 3,71 | 4,33 | 4,95 | 5,57 | 6,19 | 6,81 | 7,43 | 8,05 |
| 15 ... | 0,33 | 0,67 | 1,33 | 2,00 | 2,67 | 3,33 | 4,00 | 4,67 | 5,33 | 6,00 | 6,67 | 7,33 | 8,00 | 8,67 |
| 20 ... | 0,36 | 0,72 | 1,45 | 2,17 | 2,90 | 3,62 | 4,35 | 5,07 | 5,79 | 6,52 | 7,24 | 7,97 | 8,69 | 9,41 |
| 25 ... | 0,38 | 0,77 | 1,53 | 2,29 | 3,06 | 3,82 | 4,59 | 5,35 | 6,12 | 6,88 | 7,65 | 8,41 | 9,18 | 9,94 |
| 50 ... | 0,43 | 0,86 | 1,73 | 2,59 | 3,46 | 4,32 | 5,19 | 6,05 | 6,92 | 7,78 | 8,64 | 9,51 | 10,37 | 11,24 |
| | 0,50 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 |

Index de sélection familiale pour $r = 1/2$ et $t = 0,1$ (voir note à la fin du mémoire pour le calcul de t).

TABLEAU II

| F = | 0,1 | 0,2 | 0,4 | 0,6 | 0,8 | I | 1,2 | 1,4 | 1,6 | 1,8 | 2,0 | 2,2 | 2,4 | 2,6 |
|--------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|
| $n =$ | | | | | | | | | | | | | | |
| 1 ... | 0,10 | 0,20 | 0,40 | 0,60 | 0,80 | 1,00 | 1,20 | 1,40 | 1,60 | 1,80 | 2,00 | 2,20 | 2,40 | 2,60 |
| 2 ... | 0,08 | 0,17 | 0,33 | 0,50 | 0,67 | 0,83 | 1,00 | 1,17 | 1,33 | 1,50 | 1,67 | 1,83 | 2,00 | 2,17 |
| 3 ... | 0,08 | 0,15 | 0,30 | 0,45 | 0,60 | 0,75 | 0,90 | 1,05 | 1,20 | 1,35 | 1,50 | 1,65 | 1,80 | 1,95 |
| 4 ... | 0,07 | 0,14 | 0,28 | 0,42 | 0,56 | 0,70 | 0,84 | 0,98 | 1,12 | 1,26 | 1,40 | 1,54 | 1,68 | 1,82 |
| 5 ... | 0,07 | 0,13 | 0,27 | 0,40 | 0,53 | 0,67 | 0,80 | 0,93 | 1,07 | 1,20 | 1,33 | 1,47 | 1,60 | 1,73 |
| 6 ... | 0,06 | 0,13 | 0,26 | 0,39 | 0,51 | 0,64 | 0,77 | 0,90 | 1,03 | 1,16 | 1,29 | 1,41 | 1,54 | 1,67 |
| 7 ... | 0,06 | 0,13 | 0,25 | 0,38 | 0,50 | 0,63 | 0,75 | 0,88 | 1,00 | 1,13 | 1,25 | 1,38 | 1,50 | 1,63 |
| 8 ... | 0,06 | 0,12 | 0,24 | 0,37 | 0,49 | 0,61 | 0,73 | 0,86 | 0,98 | 1,10 | 1,22 | 1,34 | 1,47 | 1,59 |
| 9 ... | 0,06 | 0,12 | 0,24 | 0,36 | 0,48 | 0,60 | 0,72 | 0,84 | 0,96 | 1,08 | 1,20 | 1,32 | 1,44 | 1,56 |
| 10 ... | 0,06 | 0,12 | 0,24 | 0,36 | 0,47 | 0,59 | 0,71 | 0,83 | 0,95 | 1,06 | 1,18 | 1,30 | 1,42 | 1,54 |
| 11 ... | 0,06 | 0,12 | 0,23 | 0,35 | 0,47 | 0,58 | 0,70 | 0,82 | 0,93 | 1,05 | 1,17 | 1,28 | 1,40 | 1,52 |
| 12 ... | 0,06 | 0,12 | 0,23 | 0,35 | 0,46 | 0,58 | 0,69 | 0,81 | 0,92 | 1,04 | 1,15 | 1,27 | 1,39 | 1,50 |
| 15 ... | 0,06 | 0,11 | 0,23 | 0,34 | 0,45 | 0,56 | 0,68 | 0,79 | 0,90 | 1,01 | 1,13 | 1,24 | 1,35 | 1,46 |
| 20 ... | 0,05 | 0,11 | 0,22 | 0,33 | 0,44 | 0,55 | 0,66 | 0,77 | 0,88 | 0,99 | 1,10 | 1,21 | 1,31 | 1,42 |
| 25 ... | 0,05 | 0,11 | 0,22 | 0,32 | 0,43 | 0,54 | 0,65 | 0,75 | 0,86 | 0,97 | 1,08 | 1,19 | 1,29 | 1,40 |
| 50 ... | 0,05 | 0,10 | 0,21 | 0,31 | 0,42 | 0,52 | 0,62 | 0,73 | 0,83 | 0,94 | 1,04 | 1,14 | 1,25 | 1,35 |
| | 0,05 | 0,10 | 0,20 | 0,30 | 0,40 | 0,50 | 0,60 | 0,70 | 0,80 | 0,90 | 1,00 | 1,10 | 1,20 | 1,30 |

Index de sélection familiale pour $r = 1/4$ et $t = 0,5$. Noter que lorsque t est plus grand que r — (c'est-à-dire lorsque l'effet c^2 est important) l'index diminue si l'effectif familial augmente.

C. — Sélection parentale

(« Progeny-test » et « sib-test »).

Le choix des reproducteurs d'après l'épreuve de leur descendance peut être considéré comme une éventualité particulière du cas plus général où la performance propre de l'individu estimé n'est pas incluse dans le calcul de la moyenne de famille.

Une telle situation est fréquente. Elle peut résulter d'une raison de force majeure, l'impossibilité de mesurer directement cette performance : sélection laitière des taureaux, sélection ponte des coqs et plus généralement des caractères exprimés par un seul sexe ; ou encore la sélection des caractères nécessitant le sacrifice, telle la qualité de la carcasse, ou rendant l'animal impropre à la reproduction, telles l'aptitude

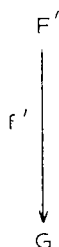


FIG. 6.

à l'engraissement par les œstrogènes ou la réponse à tel ou tel traitement des facteurs extérieurs : carence alimentaire, sensibilité aux maladies. Enfin, l'absence de mesures individuelles peut se justifier par une raison d'économie des moyens mis en œuvre, telle que la possibilité d'échantillonnage dans les familles très nombreuses.

Dans de telles circonstances, le génotype individuel est estimé à partir de la moyenne des performances des descendants (progeny-test) ou des collatéraux (sib-test), ce qui revient à effectuer une sélection familiale d'un genre un peu particulier.

L'équation de régression du génotype sur la moyenne parentale de famille F' est ainsi (fig. 6) :

$$\hat{G} - \bar{G} = k_F (F' - \bar{F}') \quad \text{avec} \quad k_F = f' \frac{\sigma_G}{\sigma_{F'}} = r_{GF}, \quad \frac{\sigma_G}{\sigma_{F'}}.$$

La seule différence entre cette méthode de sélection et la précédente réside donc dans l'expression mathématique de la corrélation r_{GF} que l'on peut calculer facilement à partir du diagramme de causalité représenté figure 7.

On a en effet (fig. 7)

$$r_{GF'} = nrhx = rh\sqrt{\frac{n}{1 + (n-1)t}}.$$

En se reportant au paragraphe précédent pour les autres éléments du calcul, on peut donc écrire l'équation de régression :

$$\hat{G} - \bar{G} = rh\sqrt{\frac{n}{1 + (n-1)t}} \cdot h\sqrt{\frac{n}{1 + (n-1)t}}(F' - \bar{F}')$$

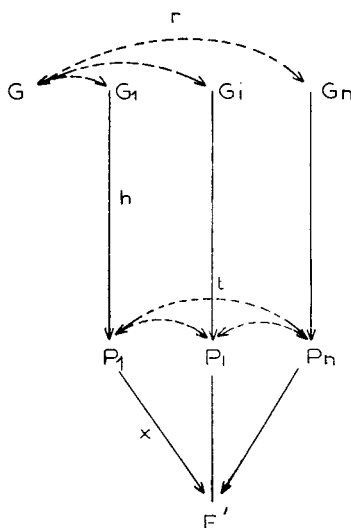


FIG. 7. — Diagramme de causalité de la sélection parentale.

ou plus simplement :

$$\hat{G} - \bar{G} = h^2 \frac{nr}{1 + (n-1)t} (F' - \bar{F}'). \quad (4)$$

Dans le cas du progeny-test, G sera généralement le génotype du père éprouvé. Sa valeur probable sera donnée par l'équation (4) avec la parenté $r = 1/2$ en l'absence de consanguinité ⁽¹⁾.

Dans le cas du sib-test, G sera le génotype d'un collatéral. On prendra donc $r = 1/2$ pour l'estimation des frères/sœurs et $r = 1/4$ pour celle des demi-frères/sœurs ⁽²⁾.

⁽¹⁾ Dans le cas de consanguinité, la parenté devient $r = \frac{1}{2} \cdot \frac{1 + 2f + f'}{\sqrt{(1+f)(1+f')}}$ f et f' étant les coefficients de consanguinité des générations « enfants » et « père » respectivement, selon la notation de WRIGHT et non les path-coefficients de G par F ou F' utilisés ici. D'où la nécessité de ce renvoi.

⁽²⁾ Dans le cas de consanguinité systématique, la parenté entre frères/sœurs devient $r = \frac{1}{2} \cdot (1 + 2f + f')/(1 + f)$

et la parenté entre demi-frères/sœurs devient $r = \frac{1}{4} \cdot (1 + 6f + f')/(1 + f)$.

Notons en passant l'identité foncière de la valeur probable d'un génotype G considéré comme demi-frère/sœur par le père des individus G_1, G_2, \dots, G_n (sib-test) ou comme enfant à venir d'un mâle éprouvé sur un échantillon (progeny-test). En d'autres termes, si l'on admet que l'effet familial c^2 est négligeable — soit $t = rh^2$ — la valeur génotypique probable d'un mâle testé est donnée par l'égalité :

$$\hat{G} - \bar{G} = h^2 \frac{n}{2 + (n-1)h^2} (\bar{F}' - \bar{F}).$$

Tandis que la valeur génotypique probable de ses enfants est donnée par l'égalité :

$$\hat{G} - \bar{G} = h^2 \frac{n}{4 + (n-1)h^2} (\bar{F}' - \bar{F}).$$

Cette dernière manière de s'exprimer représente sans doute mieux à l'esprit les potentialités d'un mâle, que sa valeur génotypique *sensu stricto*.

Efficacité de la sélection parentale.

De même que pour la sélection familiale, il est commode de mesurer l'efficacité de la sélection parentale par rapport à la sélection massale en étudiant la fraction :

$$\frac{f'}{\bar{p}} = \frac{r_{GF}'}{r_{GF}} = r \sqrt{\frac{n}{1 + (n-1)t}}.$$

En admettant que l'effet c^2 inclus dans t est négligeable, on trouve que ce rapport est égal à l'unité pour

$$h^2 = \frac{nr^2 - 1}{(n-1)r}.$$

Soit pour $r = 1/2$, $h^2 = (n-4)/2(n-1)$ et pour $r = 1/4$, $h^2 = (n-16)/4(n-1)$. On peut donc préciser, pour un effectif parental donné n , les valeurs de l'héritabilité au-dessous desquelles il convient de préférer la sélection parentale à la sélection massale. Soit, pour des groupes de frères/sœurs ou de demi-frères/sœurs :

| | | | | | | | | | | | | | | | |
|-------------|-------|---|----|----|----|----|----|----|----|----|----|-----|-----|------|----------|
| | $n =$ | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 12 | 14 | 16 | 18 | 20 | 50 | ∞ |
| $r = 1/2$ % | $h =$ | 0 | 13 | 20 | 25 | 29 | 31 | 33 | 36 | 39 | 40 | 41 | 42 | 47 | 50 |
| $r = 1/4$ % | $h =$ | — | — | — | — | — | — | — | — | — | 0 | 2,9 | 5,3 | 17,3 | 25 |

En somme, la sélection parentale n'offre d'intérêt que pour des caractères peu héritable et des familles nombreuses (fig. 8). Il est clair d'autre part que la sélection d'après les demi-frères/sœurs n'est qu'un pis aller. On peut l'illustrer par la sélection « quantité de lait » des jeunes taureaux, mais pas par leur sélection « taux butyreux » ou « azote total du lait », caractères dont l'héritabilité est nettement supérieure à 25 %.

Il semble dès lors qu'il faille songer soit à utiliser une certaine consanguinité, permettant d'obtenir des parentés supérieures à $1/4$, soit à rechercher des mesures équivalentes ou en forte corrélation génétique avec les caractères parentaux que l'on veut améliorer, en vue d'une première estimation préalable au testage.

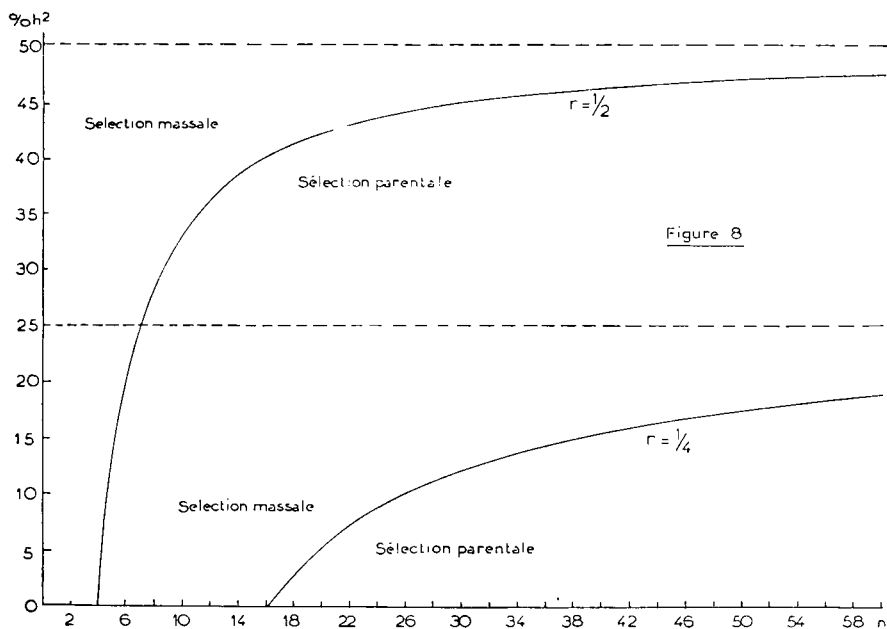


FIG. 8. — La sélection parentale n'offre d'intérêt que pour des familles nombreuses, des caractères peu héritables et des parentés étroites.

Remarques.

Il est d'ailleurs facile de comparer sélection familiale et sélection parentale en étudiant le rapport $\frac{f}{f'} = \frac{r_{GF}}{r_{GF'}} = 1 + \frac{1-r}{nr}$ efficacité toujours supérieure à l'unité, ce qui était évident *a priori*. Ce qui l'est moins, c'est l'importance du terme complémentaire $(1-r)/nr$ qui représente en somme la supériorité de la sélection familiale (ou massale pour $n = 1$) sur la sélection parentale, soit en pourcentages :

| | $n =$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 20 | 50 | ∞ |
|---------------------------------------|-------|-----|-----|-----|----|----|----|----|----|----|----|----|----|----------|
| $r = 1/2$ supériorité de la sélection | | 100 | 50 | 33 | 25 | 20 | 17 | 14 | 13 | 11 | 10 | 5 | 2 | 0 |
| $r = 1/4$ familiale, % | | 300 | 150 | 100 | 75 | 60 | 50 | 43 | 38 | 33 | 30 | 15 | 6 | 0 |

On voit de cette manière que la sélection familiale (c'est-à-dire individu estimé inclus dans le calcul de la moyenne de famille) présente une efficacité supérieure à la sélection parentale (c'est-à-dire individu estimé exclus du calcul de la moyenne de famille), de 15 à 30 % par

exemple dans le cas de 10 à 20 demi-frères/sœurs, chiffres non négligeables chez les grandes espèces. Ceci permet de conclure à l'intérêt que présenterait la recherche de mesures directes équivalentes aux caractères exprimés par un seul sexe.

Il n'est pas moins intéressant de comparer la sélection parentale à une sélection réduite, la moyenne de famille F'' étant calculée sur une fraction φ des n collatéraux de l'animal estimé, soit un échantillon

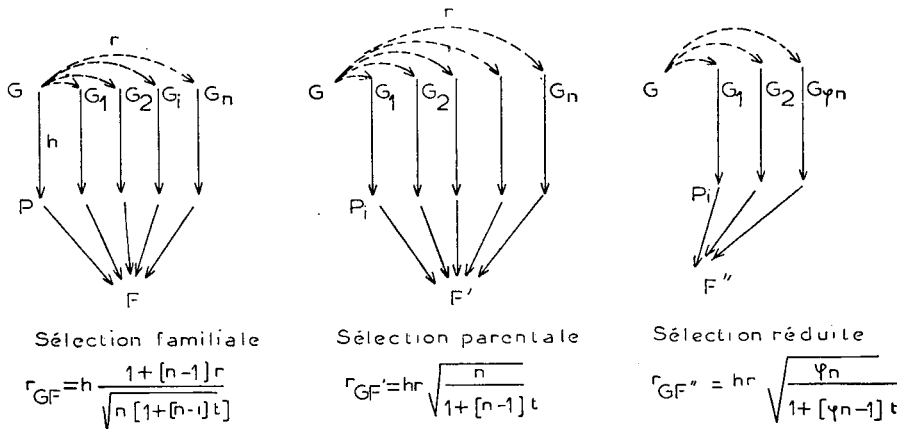


FIG. 9. — Valeurs comparées des corrélations génotype-moyenne de famille selon les diverses méthodes de sélection.

d'effectif φn . La figure 9 indique les diagrammes de causalité des trois manières d'opérer et les corrélations génotype — moyenne de famille correspondantes.

On a de cette façon :

$$\frac{r_{GF'}}{r_{GF''}} = \sqrt{\frac{\frac{1}{\varphi} \cdot \frac{1 + (\varphi n - 1)t}{1 + (n - 1)t}}{1 + \frac{1 - \varphi}{\varphi} \cdot \frac{1 - t}{1 + (n - 1)t}}}$$

rapport toujours supérieur à l'unité, valeur asymptotique lorsque φ ou t tend vers 1 ou lorsque l'échantillon augmente indéfiniment.

Le tableau III indique diverses valeurs de ce rapport pour un échantillon donné, en admettant que $t = rh^2$.

D'après ce tableau, il est clair que la sélection réduite, sur un échantillon suffisamment important de demi-frères/sœurs (par exemple $\varphi = 2/3$) est plus intéressante que sur un échantillon moins nombreux de frères/sœurs (par exemple $\varphi = 1/2$). Cette conclusion est d'autant plus utile à noter qu'il est plus facile d'obtenir un grand nombre de demi-frères-sœurs qu'un nombre relativement plus faible de frères/sœurs, particulièrement chez les grandes espèces. Ainsi, la sélection d'un mâle d'après une vingtaine sur trente demi-sœurs ne souffre qu'une perte de 2 à 13 %

par rapport à la sélection parentale ; alors que d'après une demi-douzaine sur 12 sœurs — cas seulement applicable aux espèces prolifiques — la perte s'élève de 5 à 27 %. L'utilisation pratique de la sélection réduite pour le choix des mâles peut donc permettre une économie de travail d'un 1/3 qui compense largement, à un certain point de vue, la perte de gain y conséquente.

TABLEAU III

| n | $h^2 =$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---------|---------|------|------|------|------|------|------|------|------|------|----|
| 2..... | 1,38 | 1,35 | 1,32 | 1,29 | 1,27 | 1,24 | 1,22 | 1,20 | 1,17 | 1,16 | |
| 4..... | 1,35 | 1,30 | 1,26 | 1,23 | 1,20 | 1,17 | 1,15 | 1,13 | 1,11 | 1,10 | |
| 6..... | 1,33 | 1,27 | 1,22 | 1,18 | 1,15 | 1,13 | 1,11 | 1,10 | 1,08 | 1,07 | |
| 8..... | 1,31 | 1,24 | 1,19 | 1,15 | 1,13 | 1,11 | 1,09 | 1,08 | 1,06 | 1,05 | |
| 10..... | 1,29 | 1,21 | 1,17 | 1,13 | 1,11 | 1,09 | 1,08 | 1,06 | 1,05 | 1,04 | |
| 12..... | 1,27 | 1,20 | 1,15 | 1,12 | 1,10 | 1,08 | 1,07 | 1,06 | 1,05 | — | — |
| 14..... | 1,26 | 1,18 | 1,14 | 1,11 | 1,08 | 1,07 | 1,06 | 1,05 | 1,04 | 1,03 | — |
| 16..... | 1,24 | 1,17 | 1,12 | 1,10 | — | 1,06 | 1,05 | 1,04 | — | — | — |
| 18..... | 1,23 | 1,15 | 1,11 | 1,09 | 1,07 | — | — | — | 1,03 | — | — |
| 20..... | 1,22 | 1,13 | 1,10 | 1,08 | 1,06 | 1,05 | 1,04 | 1,03 | — | 1,02 | — |
| 30..... | 1,18 | 1,11 | 1,08 | 1,06 | 1,04 | 1,04 | 1,03 | 1,02 | 1,02 | — | — |
| 40..... | 1,15 | 1,09 | 1,06 | 1,04 | 1,03 | 1,03 | 1,02 | — | — | 1,01 | — |
| 50..... | 1,13 | 1,07 | 1,05 | — | — | 1,02 | — | — | 1,01 | — | — |
| 60..... | 1,11 | 1,06 | 1,04 | 1,03 | 1,02 | — | — | 1,01 | — | — | — |
| 70..... | 1,10 | 1,06 | — | — | — | — | 1,01 | — | — | — | — |
| 80..... | 1,09 | 1,05 | 1,03 | 1,02 | — | 1,01 | — | — | — | — | — |
| 90..... | 1,08 | 1,04 | — | — | — | — | — | — | — | — | — |
| | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 2..... | 1,22 | 1,21 | 1,20 | 1,19 | 1,18 | 1,17 | 1,16 | 1,15 | 1,15 | 1,14 | |
| 4..... | 1,21 | 1,19 | 1,17 | 1,16 | 1,15 | 1,14 | 1,13 | 1,12 | 1,11 | 1,10 | |
| 6..... | 1,20 | 1,18 | 1,16 | 1,14 | 1,13 | 1,12 | 1,11 | 1,10 | 1,09 | 1,08 | |
| 8..... | 1,19 | 1,16 | 1,14 | 1,12 | 1,11 | 1,10 | 1,09 | 1,08 | 1,07 | 1,07 | |
| 10..... | 1,18 | 1,15 | 1,13 | 1,11 | 1,10 | 1,09 | 1,08 | 1,07 | 1,06 | 1,06 | |
| 12..... | — | 1,14 | 1,12 | 1,10 | 1,09 | 1,08 | 1,07 | 1,06 | 1,05 | 1,05 | |
| 14..... | 1,17 | 1,13 | 1,11 | 1,09 | 1,08 | 1,07 | 1,06 | 1,05 | — | 1,04 | |
| 16..... | 1,16 | — | 1,10 | — | 1,07 | 1,06 | — | — | 1,04 | — | |
| 18..... | — | 1,12 | — | 1,08 | — | — | 1,05 | 1,04 | — | — | |
| 20..... | 1,15 | — | 1,09 | — | 1,06 | 1,05 | — | — | — | 1,03 | |
| 30..... | 1,13 | 1,09 | 1,07 | 1,07 | 1,05 | 1,04 | 1,03 | 1,03 | 1,03 | 1,02 | |
| 40..... | 1,12 | 1,08 | 1,06 | 1,05 | 1,04 | 1,03 | — | 1,02 | 1,02 | — | |
| 50..... | 1,10 | 1,07 | 1,05 | 1,04 | 1,03 | — | 1,02 | — | — | 1,01 | |
| 60..... | 1,09 | 1,06 | 1,04 | 1,03 | — | 1,02 | — | — | 1,01 | — | |
| 70..... | — | 1,05 | — | — | 1,02 | — | — | 1,01 | — | — | |
| 80..... | 1,08 | — | 1,03 | — | — | — | 1,01 | — | — | — | |
| 90..... | 1,07 | 1,04 | — | 1,02 | — | 1,01 | — | — | — | — | |
| | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |

Efficacité de la sélection parentale par rapport à la sélection réduite. En haut valeurs du rapport r_{GF}'/r_{GF}'' pour $r = 1/2$ et $\varphi = 1/2$; en bas valeurs de ce rapport pour $r = 1/4$ et $\varphi = 2/3$.

Il convient néanmoins de rappeler ici le domaine limité de la sélection parentale pour des valeurs de r faibles : elle ne concerne en fait que les caractères peu héréditaires. Mais dans un tel cas il peut être avantageux de lui substituer la sélection réduite. Enfin, le cas général de la comparaison entre la sélection familiale et la sélection réduite (qui tend

vers la sélection parentale lorsque φ tend vers 1) est donné par le rapport :

$$\frac{r_{GF}}{r_{GF}'} = \left(1 + \frac{1-r}{nr}\right) \sqrt{1 + \frac{1-\varphi}{\varphi} \cdot \frac{1-t}{1+(n-1)t}}.$$

Index de sélection parentale.

De même que pour la sélection familiale, il est commode de simplifier l'équation (4) par suppression des termes et facteurs constants. Soit l'index de sélection parentale :

$$I_P = \frac{nr}{1 + (n-1)t} F'.$$

Cet index est susceptible d'être donné directement par des tables en fonction de n et F' pour une parenté r donnée et pour les diverses valeurs de t . L'index de sélection réduite est du même coup construit, n étant remplacé par φn .

D. — Sélection pedigree

L'estimation du génotype individuel s'effectue à partir de la performance d'un ou plusieurs ascendants.

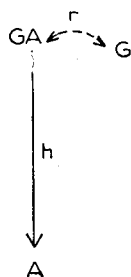


FIG. 10.

Dans le cas le plus simple, un animal est choisi d'après la valeur phénotypique d'un ascendant direct au premier degré — par exemple la sélection des coqs sur la ponte de leur mère.

Soit G le génotype estimé et A le phénotype de l'ascendant considéré, GA son génotype. L'équation de régression de G sur A s'écrit :

$$\hat{G} - \bar{G} = k_A(A - \bar{A}) \quad \text{avec} \quad k_A = r_{GA} \frac{\sigma_G}{\sigma_A}.$$

La corrélation r_{GA} est donnée simplement par la figure 10, soit $r_{GA} = rh$.

D'autre part, si l'ascendant A n'a pas été l'objet d'une sélection antérieure, on a $\sigma_G/\sigma_A = \sigma_G/\sigma_P = h$ et le coefficient de régression se réduit rh^2 .

Si au contraire, l'ascendant a été l'objet d'une sélection, il est nécessaire de calculer la variance σ_A^2 en fonction de la variance originale σ_P^2 .

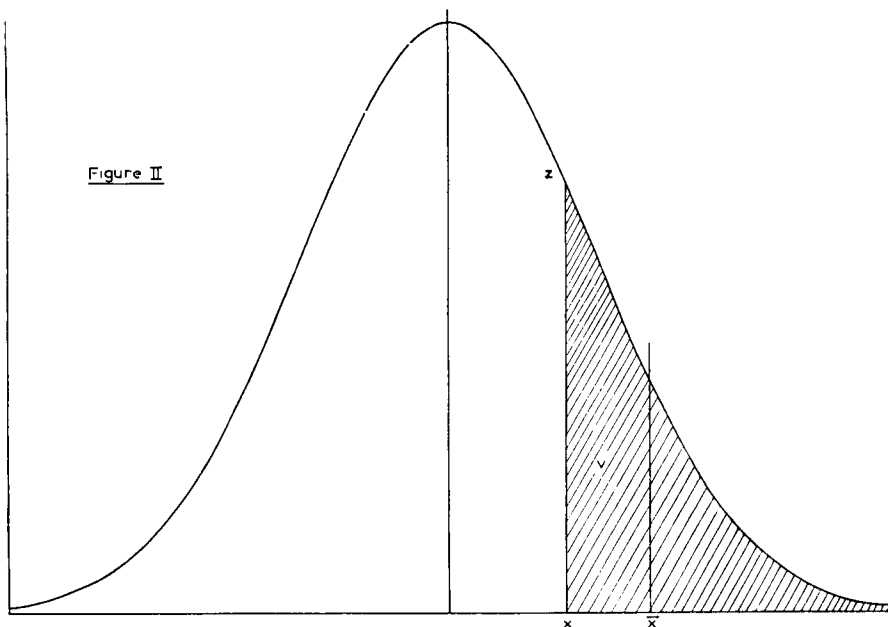


FIG. II. — On a par définition de la moyenne :

$$\bar{x} = \frac{\int_{-\infty}^{+\infty} x z dx}{\int_{-\infty}^{+\infty} z dx} = \sigma_P^2 \frac{z}{v} \quad \text{avec} \quad z = \frac{1}{\sigma_P \sqrt{2\pi}} e^{-\frac{x^2}{2\sigma_P^2}}$$

$$\text{et } v = \frac{1}{2} - \frac{1}{\sigma_P \sqrt{2\pi}}$$

$$\left[x - \frac{1}{2\sigma_P^2} \cdot \frac{x^3}{3} + \frac{1}{2 \cdot 4\sigma_P^4} \cdot \frac{x^5}{5} - \frac{1}{2 \cdot 4 \cdot 6\sigma_P^6} \cdot \frac{x^7}{7} + \frac{1}{2 \cdot 4 \cdot 6 \cdot 8\sigma_P^8} \cdot \frac{x^9}{9} - \dots \pm \frac{1}{2n \cdot n! \sigma_P^{2n}} \cdot \frac{x^{2n+1}}{(2n+1)} \right]$$

On a ensuite, par définition de la variance :

$$\sigma_A^2 = \frac{\int_{-\infty}^{+\infty} (x - \bar{x})^2 z dx}{\int_{-\infty}^{+\infty} z dx} = \frac{x\sigma_P^2 z + \sigma_P^2 v - 2\bar{x}\sigma_P^2 z + \bar{x}^2 v}{v} = \sigma_P^2 + x\bar{x} - \bar{x}^2$$

$$\text{ou mieux, en posant } X = \frac{x}{\sigma_P} \text{ (écart-réduit) : } \sigma_A^2 = (1 + X\bar{X} - \bar{X}^2) \sigma_P^2.$$

Dans le cas théorique d'une sélection tronquée à partir d'une distribution normale (fig. II) on a :

$$\bar{x} = \sigma_P^2 \cdot z/v \quad \text{et} \quad \sigma_A^2 = (1 + X\bar{X} - \bar{X}^2) \sigma_P^2.$$

Par suite,

$$\sigma_A = \sqrt{1 + X\bar{X} - \bar{X}^2} \cdot \sigma_P \quad \text{et} \quad \frac{\sigma_G}{\sigma_A} = \frac{h}{\sqrt{1 + X\bar{X} - \bar{X}^2}}.$$

D'où enfin l'équation de régression :

$$\boxed{\hat{G} - \bar{G} = \frac{rh^2}{\sqrt{1 + X\bar{X} - \bar{X}^2}} (A - \bar{A})} \quad (5)$$

On voit que, par rapport à la sélection massale, l'efficacité de cette méthode est égale à $r_{GA}/r_{GP} = r$, toujours inférieure à l'unité. Mais, il est plus intéressant de comparer la sélection pedigree à la sélection paren-

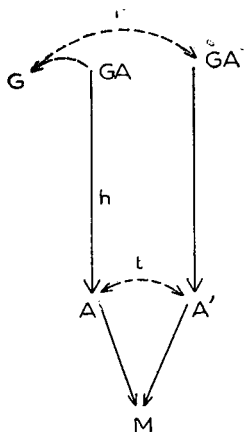


FIG. 12. — Sélection sur le « parent moyen ».

tale, l'une et l'autre pouvant être utilisée pour des caractères exprimés par un seul sexe. L'efficacité est alors mesurée par le rapport

$$\frac{r_{GA}}{r_{GP}} = \sqrt{\frac{1 + (n - 1)t}{n}}$$

qui est, également, toujours inférieur à l'unité, sauf pour $n = 1$, ce qui permet de dire que la sélection pedigree équivaut à une sélection parentale d'après un seul collatéral — pour des valeurs égales de r ; résultat qui devrait être évident à tous les éleveurs.

Il est aussi fréquent d'estimer un animal à partir de la moyenne M de ses parents — alias « parent moyen » de GALTON.

Dans ce cas, l'équation de régression devient :

$$\hat{G} - \bar{G} = k_M (M - \bar{M}) \quad \text{avec} \quad k_M = r_{GM} \frac{\sigma_G}{\sigma_M}.$$

Le diagramme de causalité de la figure 12 donne :

$$r_{GM} = 2rhx \quad \text{avec} \quad 1 = 2x^2(1 + t)$$

soit $r_{GM} = rh \sqrt{\frac{2}{1+t}}$. D'autre part, on a

$$\sigma_M^2 = \frac{1+t}{2} \sigma_A^2 \quad \text{d'où} \quad \sigma_M^2 = \frac{1+t}{2} \cdot \frac{1 + X\bar{X} - \bar{X}^2}{h^2} \sigma_G^2$$

ce qui permet de calculer le coefficient de régression

$$k_M = rh \sqrt{\frac{2}{1+t}} \cdot \sqrt{\frac{2}{1+t} \frac{h}{1 + X\bar{X} - \bar{X}^2}}.$$

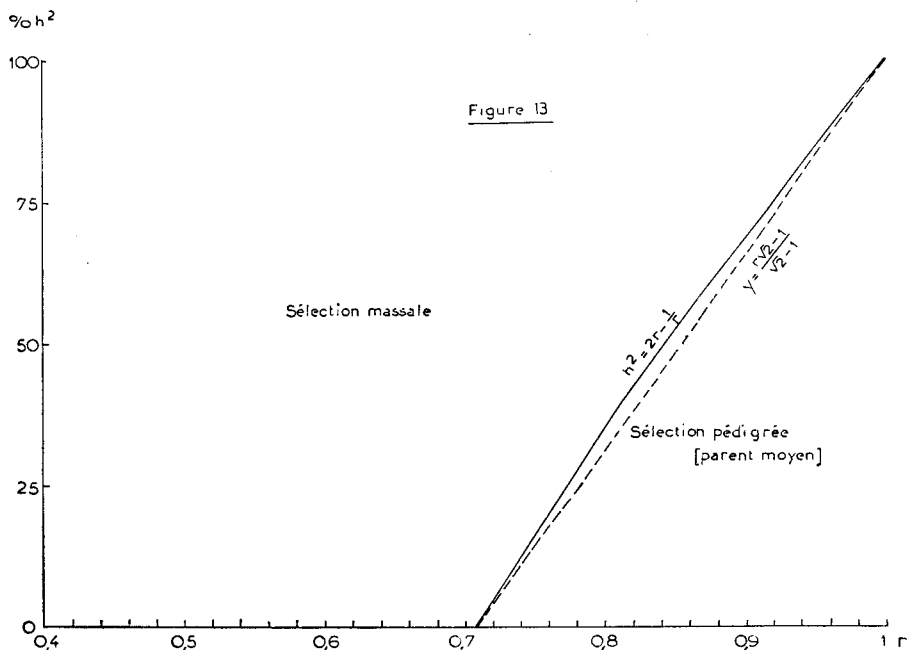


FIG. 13. — La sélection sur le parent moyen n'offre d'intérêt que pour des héritabilités relativement faibles et des parentés très étroites.

D'où enfin l'équation de régression :

$$\boxed{\hat{G} - \bar{G} = \frac{2rh^2}{(1+t)\sqrt{1 + X\bar{X} - \bar{X}^2}} (M - \bar{M})} \quad (6)$$

L'efficacité de cette méthode, rapportée à la sélection massale, est égale à $r_{GM}/r_{GP} = r \sqrt{\frac{2}{1+t}}$. Même en admettant que l'effet c^2 inclus dans t est négligeable, ce rapport ne peut être supérieur à l'unité que pour des héritabilités inférieures à $(2r^2 - 1)/r$. Pour respecter la condition $h^2 > 0$, il faut que l'on ait $r > 1/\sqrt{2}$ pour que cette méthode de sélection soit préférable à la simple sélection massale. Or, une telle parenté correspond

à 5 ou 6 générations de consanguinité frères/sœurs. C'est dire le peu d'intérêt d'une méthode fondée sur le seul critère d'un ascendant ou du « parent moyen ». Cet intérêt se limite aux caractères peu héritables en présence de parentés très étroites (fig. 13).

Très insuffisante en soi, la sélection pedigree peut néanmoins être utile, en tant qu'information complémentaire à d'autres méthodes.

E. — Sélection combinée

Par sélection combinée, on entend l'estimation du génotype d'un animal à partir de la performance brute ou phénotype P , de moyenne \bar{P} , et de la moyenne de famille F , de valeur moyenne \bar{F} , r étant la parenté moyenne des membres de cette famille.

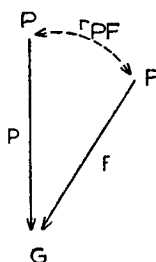


FIG. 14. — Sélection combinée.

Conformément à l'équation (1), la régression multiple du génotype sur les données s'écrit :

$$\hat{G} - \bar{G} = k_P(P - \bar{P}) + k_F(F - \bar{F}).$$

Avec (fig. 14)

$$p = k_P \frac{\sigma_P}{\sigma_G} \quad \text{et} \quad f = k_F \frac{\sigma_F}{\sigma_G}$$

d'où la forme

$$\hat{G} - \bar{G} = p \frac{\sigma_G}{\sigma_P}(P - \bar{P}) + f \frac{\sigma_G}{\sigma_F}(F - \bar{F}).$$

Les rapports σ_G/σ_P et σ_G/σ_F ont été calculés précédemment. Il suffit donc de connaître les path-coefficients p et f . Le diagramme de causalité de la figure 14 donne le système d'équations :

$$r_{GP} = p + fr_{PF} \quad \text{et} \quad r_{GF} = pr_{PF} + f.$$

$$\text{D'où} \quad p = \frac{r_{GP} - r_{GF} \cdot r_{PF}}{1 - r_{PF}^2} \quad \text{et} \quad f = \frac{r_{GF} - r_{GP} \cdot r_{PF}}{1 - r_{PF}^2}.$$

La corrélation r_{GP} est simplement la racine carrée de l'héritabilité (cf. paragraphe sélection massale) soit $r_{GP} = h$. D'autre part, r_{GF} a été calculé plus haut (cf. paragraphe sélection familiale) soit :

$$r_{GF} = h \frac{1 + (n-1)r}{\sqrt{n[1 + (n-1)t]}}.$$

Enfin, le calcul de la corrélation r_{PF} entre le phénotype de l'animal étudié et la moyenne de sa famille s'effectue rapidement à partir du

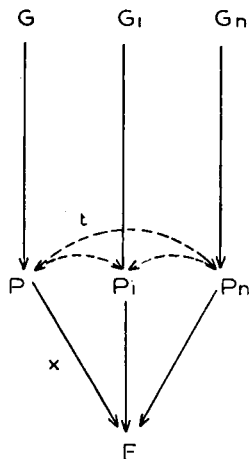


FIG. 15.

diagramme représenté figure 15. On a en effet : $r_{PF} = x + (n-1)tx$, avec pour équation de détermination de F : $1 = nx^2 + n(n-1)tx^2$, qui donne x .

Par suite, $r_{PF} = \sqrt{\frac{1 + (n-1)t}{n}}$. Il suffit de remplacer r_{GP} , r_{GF} et r_{PF} par leur valeur pour obtenir les path-coefficients p et f . Soient :

$$p = h \frac{1-r}{1-t} \quad \text{et} \quad f = h \frac{r-t}{1-t} \sqrt{\frac{n}{1+(n-1)t}}.$$

Enfin, connaissant les rapports

$$\sigma_G/\sigma_P = h \quad \text{et} \quad \sigma_G/\sigma_F = h \sqrt{\frac{n}{1+(n-1)t}},$$

on obtient les coefficients de régression multiple k_P et k_F ; d'où l'équation cherchée :

$$\hat{G} - \bar{G} = h^2 \frac{1-r}{1-t} (P - \bar{P}) + h^2 \frac{r-t}{1-t} \cdot \frac{n}{1+(n-1)t} (F - \bar{F}). \quad (7)$$

Efficacité de la sélection combinée.

Cette efficacité peut être mesurée par référence à la sélection massale en étudiant le rapport $R_{G\hat{G}}/r_{GP}$, $R_{G\hat{G}}$ étant le coefficient de corrélation multiple entre le génotype G d'une part, P et F d'autre part.

On a évidemment : $R_{G\hat{G}} = \frac{\text{cov} \cdot G\hat{G}}{\sigma_G \cdot \sigma_{\hat{G}}}$ avec l'écart-type libre $\sigma_{\hat{G}} = R_{G\hat{G}} \cdot \sigma_G$, G ne différant de sa valeur théorique \hat{G} donnée par la régression, que par un terme aléatoire. On peut donc écrire

$$R_{G\hat{G}}^2 = \frac{\text{cov } G\hat{G}}{\sigma_G^2} = k_P \frac{\text{cov } GP}{\sigma_G^2} + k_F \frac{\text{cov } GF}{\sigma_G^2}$$

ou encore :

$$R_{G\hat{G}}^2 = k_P \frac{r_{GP} \cdot \sigma_G \cdot \sigma_P}{\sigma_G^2} + k_F \frac{r_{GF} \cdot \sigma_G \cdot \sigma_F}{\sigma_G^2} \quad \text{avec} \quad k_P = \rho \frac{\sigma_G}{\sigma_P} \quad \text{et} \quad k_F = f \frac{\sigma_G}{\sigma_F}.$$

Par suite, on peut écrire :

$$R_{G\hat{G}}^2 = \rho r_{GP} + f r_{GF}.$$

Or, tous les termes du second membre de cette égalité ont été calculés précédemment. Il vient, en remplaçant terme par terme :

$$R_{G\hat{G}}^2 = h \frac{1-r}{1-t} \cdot h + h \frac{r-t}{1-t} \sqrt{\frac{n}{1+(n-1)t}} \cdot h \frac{1+(n-1)r}{\sqrt{n[1+(n-1)t]}}.$$

Ou plus simplement :

$$R_{G\hat{G}}^2 = h^2 \frac{1-r}{1-t} + h^2 \frac{r-t}{1-t} \cdot \frac{1+(n-1)r}{1+(n-1)t}.$$

Ainsi l'efficacité de la sélection combinée par rapport à la sélection massale ($r_{GP} = h$) est donnée par le rapport :

$$\frac{R_{G\hat{G}}}{r_{GP}} = \sqrt{\frac{1-r}{1-t} + \frac{r-t}{1-t} \cdot \frac{1+(n-1)r}{1+(n-1)t}} = \sqrt{1 + \frac{(r-t)^2}{1-t} \cdot \frac{n-1}{1+(n-1)t}}.$$

Formule donnée, sans commentaire, par LUSH (1945). Ce rapport étant toujours supérieur à l'unité, on voit que la sélection combinée est toujours préférable à la sélection massale. Toutefois, le terme fractionnaire sous radical s'amenuise lorsque h^2 augmente. En posant $r = 1/2$ et $t = rh^2$ on peut l'écrire :

$$\frac{(1-h^2)^2(n-1)}{(2-h^2)[2+(n-1)h^2]}$$

Il est évidemment nul pour $n = 1$ et $h^2 = 1$.

En se fixant une limite d'efficacité, par exemple 1,01 ou 1,10 on trouve les valeurs limites de h^2 au-dessus desquelles la sélection massale

est pratiquement aussi avantageuse que la sélection combinée. Soit pour divers effectifs n :

| | | | | | |
|-------------------------|------------|------|------|--------|----------|
| $n =$ | 5 | 10 | 20 | 100 | ∞ |
| $\varepsilon = 1,01 \%$ | $h^2 = 82$ | 84 | 85 | $85,8$ | 86 |
| $\varepsilon = 1,10 \%$ | $h^2 = 40$ | 53 | 56 | $58,6$ | 59 |

On voit que pour des héritabilités supérieures à 80 p. 100, la sélection combinée n'offre plus guère d'intérêt. Grosso modo, on pourra se contenter d'une sélection massale au-dessus de $h^2 = 60$ p. 100. En effet, pour cette valeur de h^2 , l'efficacité de la sélection combinée serait :

$$\varepsilon = \sqrt{\frac{n + 1,80}{0,84n + 1,96}}.$$

Soit pour divers effectifs n :

| | | | | |
|-----------------|------|------|------|----------|
| $n =$ | 5 | 10 | 20 | ∞ |
| $\varepsilon =$ | 1,05 | 1,07 | 1,08 | 1,09 |

En utilisant la sélection massale, la perte d'efficacité sera donc de l'ordre de 5 à 10 p. 100. Mais comme il s'agit de caractères très héréditaires, il semble qu'il n'y ait pas d'inconvénient majeur à opérer de cette manière. Pour des héritabilités de 40 à 50 p. 100 la perte s'élève de 10 à 20 p. 100. La valeur proposée $h^2 = 0,6$ paraît donc un seuil raisonnable.

Inversement, pour des héritabilités faibles — par exemple $h^2 = 0,2$, l'efficacité de la sélection combinée est $\sqrt{\frac{n + 2,60}{0,36n + 3,24}}$

Soit pour des effectifs n :

| | | | | |
|-----------------|------|------|------|----------|
| $n =$ | 5 | 10 | 20 | ∞ |
| $\varepsilon =$ | 1,23 | 1,36 | 1,47 | 1,66 |

La valeur combinée permet alors un gain de 23 à 66 p. 100 en sus de celui obtenu par sélection massale pour un caractère peu héréditable.

Index de sélection combinée.

L'équation (7) de régression multiple, peut être remplacée par une valeur équivalente plus simple à calculer, qui sera l'index proprement dit. On a en effet :

$$(7) \quad \hat{G} - \bar{G} = \frac{1-r}{1-t} h^2 (P - \bar{P}) + \frac{r-t}{1-t} \cdot \frac{n}{1+(n-1)t} h^2 (F - \bar{F}).$$

Si les données sont exprimées en écarts-réduits, les moyennes \bar{P} et \bar{F} sont nulles. D'autre part, on substituera à l'équation (7) une valeur proportionnelle en la divisant par le coefficient $\frac{1-r}{1-t} h^2$. En posant (LUSH, 1945) $W = \frac{r-t}{1-r} \cdot \frac{n}{1+(n-1)t}$ l'index de sélection combinée est alors simplement :

$$\boxed{I_c = WF + P}.$$

Notons qu'il est possible de construire des tables donnant la valeur du terme WF, en fonction de F, n et h^2 pour une parenté donnée et en admettant que $t = rh^2$. Un tel exemple est donné tableau IV pour une parenté $r = 1/2$ et une héritabilité $h^2 = 0,25$. Il suffit alors d'ajouter la valeur phénotypique de l'individu estimé à la valeur donnée par la table pour obtenir l'index.

TABLEAU IV

| W = | 0,75 | 1,34 | 1,80 | 2,18 | 2,50 | 2,77 | 3,00 | 3,20 | 3,38 | 3,53 | 3,67 | 3,79 |
|-----|------|------|------|------|------|------|------|------|------|------|------|------|
| n = | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| F = | ± | ± | ± | ± | ± | ± | ± | ± | ± | ± | ± | ± |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0,1 | 0,07 | 0,13 | 0,18 | 0,21 | 0,25 | 0,27 | 0,30 | 0,32 | 0,33 | 0,35 | 0,36 | 0,37 |
| 0,2 | 0,15 | 0,26 | 0,36 | 0,43 | 0,50 | 0,55 | 0,60 | 0,64 | 0,67 | 0,70 | 0,73 | 0,75 |
| 0,3 | 0,22 | 0,40 | 0,54 | 0,65 | 0,75 | 0,83 | 0,90 | 0,96 | 1,01 | 1,05 | 1,10 | 1,13 |
| 0,4 | 0,30 | 0,53 | 0,72 | 0,87 | 1,00 | 1,10 | 1,20 | 1,28 | 1,35 | 1,41 | 1,46 | 1,51 |
| 0,5 | 0,37 | 0,67 | 0,90 | 1,09 | 1,25 | 1,38 | 1,50 | 1,60 | 1,69 | 1,76 | 1,83 | 1,89 |
| 0,6 | 0,45 | 0,80 | 1,08 | 1,30 | 1,50 | 1,66 | 1,80 | 1,92 | 2,02 | 2,11 | 2,20 | 2,27 |
| 0,7 | 0,52 | 0,95 | 1,26 | 1,52 | 1,75 | 1,93 | 2,10 | 2,24 | 2,36 | 2,47 | 2,56 | 2,65 |
| 0,8 | 0,60 | 1,07 | 1,44 | 1,74 | 2,00 | 2,21 | 2,40 | 2,56 | 2,70 | 2,82 | 2,93 | 3,03 |
| 0,9 | 0,67 | 1,20 | 1,62 | 1,96 | 2,25 | 2,49 | 2,70 | 2,88 | 3,04 | 3,17 | 3,30 | 3,41 |
| 1,0 | 0,75 | 1,34 | 1,80 | 2,18 | 2,50 | 2,77 | 3,00 | 3,20 | 3,38 | 3,53 | 3,67 | 3,79 |
| 1,1 | 0,82 | 1,47 | 1,98 | 2,39 | 2,75 | 3,04 | 3,30 | 3,52 | 3,71 | 3,88 | 4,03 | 4,16 |
| 1,2 | 0,90 | 1,60 | 2,16 | 2,61 | 3,00 | 3,32 | 3,60 | 3,84 | 4,05 | 4,23 | 4,40 | 4,54 |
| 1,3 | 0,97 | 1,74 | 2,34 | 2,83 | 3,25 | 3,60 | 3,90 | 4,16 | 4,39 | 4,58 | 4,77 | 4,92 |
| 1,4 | 1,05 | 1,87 | 2,52 | 3,05 | 3,50 | 3,87 | 4,20 | 4,48 | 4,73 | 4,94 | 5,13 | 5,30 |
| 1,5 | 1,12 | 2,01 | 2,70 | 3,27 | 3,75 | 4,15 | 4,50 | 4,80 | 5,07 | 5,29 | 5,50 | 5,68 |
| 1,6 | 1,20 | 2,14 | 2,88 | 3,48 | 4,00 | 4,43 | 4,80 | 5,12 | 5,40 | 5,64 | 5,87 | 6,06 |
| 1,7 | 1,27 | 2,27 | 3,06 | 3,70 | 4,25 | 4,70 | 5,10 | 5,44 | 5,74 | 6,00 | 6,23 | 6,44 |
| 1,8 | 1,35 | 2,41 | 3,24 | 3,92 | 4,50 | 4,98 | 5,40 | 5,76 | 6,08 | 6,35 | 6,60 | 6,82 |
| 1,9 | 1,42 | 2,54 | 3,42 | 4,14 | 4,75 | 5,26 | 5,70 | 6,08 | 6,42 | 6,70 | 6,97 | 7,20 |
| 2,0 | 1,50 | 2,68 | 3,60 | 4,36 | 5,00 | 5,54 | 6,00 | 6,40 | 6,76 | 7,06 | 7,34 | 7,58 |

L'index de sélection combinée $I = WF + P$ se réduit à une simple addition lorsqu'on connaît le terme WF. La table ci-dessus est utilisable pour des familles de frères/sœurs et un caractère d'héritabilité 25 % — soit grossièrement les caractères « ponte ». P est exprimé en écart réduit.

CONCLUSIONS

En prenant pour angle de vue l'efficacité des index, on peut les classer par ordre décroissant selon le critère utilisé, soit d'une manière générale : valeur combinée, valeur familiale, valeur individuelle, valeur pedigree.

Notons toutefois que pour des caractères assez hératables ($h^2 \geq 0,50$ ou même 0,40) la sélection massale devient préférable à la sélection familiale.

Notons aussi le faible intérêt de la sélection pedigree utilisée d'une manière exhaustive.

La principale réserve qu'il convient d'émettre d'autre part est que dans toutes les illustrations de calcul nous avons supposé nul l'effet c^2 du milieu commun aux membres d'une même famille. Cette optique

est assez arbitraire et ne doit être entendue que dans un sens didactique. On peut toutefois se rapprocher d'une telle condition en ne créant pas, dans toute la mesure du possible, de milieu familial particulier, ou en ne comparant entre eux que des animaux élevés dans les mêmes conditions. Une telle occurrence est relativement facile à obtenir avec une petite espèce, lapin ou volaille par exemple, en mélangeant les familles dans les divers bâtiments d'élevage. Il est plus délicat de l'obtenir chez d'autres espèces moins justiciables de « l'organisation » de l'élevage. D'une manière plus aiguë, l'étude d'une grande espèce peut être faussée par l'existence d'un milieu « troupeau » dans la mesure où cette entité s'accompagne d'un certain *line breeding* (cf. corrélations et interactions génotype-milieu). Il est alors nécessaire d'envisager les corrections indispensables ou de normaliser les conditions d'élevage ou encore de se contenter d'un outil plus grossier. Enfin, il est bien évident que la sélection ne porte pratiquement jamais sur un seul caractère, plus précisément sur une seule information — autre que la moyenne du troupeau (cf. ROBERTSON, 1955). Mais nous avons suffisamment indiqué le raisonnement qui aboutit à la construction d'un index. Il convient à présent de l'appliquer au cas de la sélection pour plusieurs caractères.

II. — SÉLECTION SUR PLUSIEURS CARACTÈRES

Il est bien évident que la sélection sur plusieurs caractères est plus complexe que la sélection pour un seul, ne serait-ce que par l'intervention de nouveaux paramètres dans les calculs, tels que les corrélations génétiques et les coefficients d'importance relative des caractères étudiés. Cependant, il n'y a pas de différence de méthode entre ce que nous pouvons appeler la mono-sélection et la poly-sélection ; aussi nous limiterons l'exposé de celle-ci aux procédés le plus pratique (sélection masale) et le plus efficace (sélection combinée).

A. — Sélection masale de plusieurs caractères

L'équation de régression multiple du génotype sur les mesures enregistrées s'écrit :

$$\hat{G} - \bar{G} = k_A (A - \bar{A}) + k_B (B - \bar{B}) + \dots + k_Z (Z - \bar{Z}) \quad (1)$$

AB... Z étant les critères de sélection utilisés — c'est-à-dire des performances individuelles dans le cas présent. Les coefficients de régression multiple sont, d'après la figure 16 :

$$k_A = a \frac{\sigma_G}{\sigma_A} \quad k_B = b \frac{\sigma_G}{\sigma_B} \quad \dots \quad k_Z = z \frac{\sigma_G}{\sigma_Z}$$

L'équation est donc résolue si l'on connaît les path-coefficients $ab...z$ et les rapports σ_G/σ_A , $\sigma_G/\sigma_B, \dots, \sigma_G/\sigma_Z$.

Nous noterons ici G le génotype total de l'animal, correspondant à l'ensemble des caractères $AB...Z$ et $GA, GB...GZ$ les génotypes partiels correspondant à chacun d'eux.

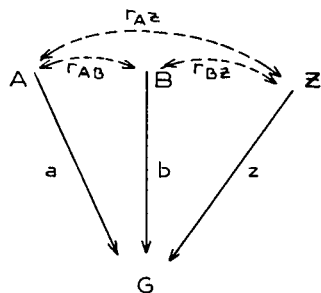


FIG. 16.

Calcul des path-coefficients $ab...z$.

Le diagramme de causalité représenté figure 16 donne le système d'équations :

$$\begin{aligned} r_{G \cdot A} &= a + br_{AB} + \dots + zr_{AZ} \\ r_{G \cdot B} &= ar_{BA} + b + \dots + zr_{BZ} \\ \dots & \dots \dots \dots \dots \dots \dots \\ r_{G \cdot Z} &= ar_{ZA} + br_{ZB} + \dots + z. \end{aligned}$$

Soit un système d'équations linéaires à z inconnues, que l'on peut mettre sous la forme matricielle :

$$\begin{bmatrix} I & r_{AB} & \dots & r_{AZ} \\ r_{BA} & I & \dots & r_{BZ} \\ \dots & \dots & \dots & \dots \\ r_{ZA} & r_{ZB} & \dots & I \end{bmatrix} \begin{Bmatrix} a \\ b \\ \dots \\ z \end{Bmatrix} = \begin{Bmatrix} r_{G \cdot A} \\ r_{G \cdot B} \\ \dots \\ r_{G \cdot Z} \end{Bmatrix}$$

\boxed{M}

En prémultipliant par la matrice inverse

$$\boxed{M}^{-1} = \frac{\boxed{M}^T}{|\boxed{M}|}$$

(soit matrice adjointe/déterminant) on obtient la solution :

$$\begin{Bmatrix} a \\ b \\ \dots \\ z \end{Bmatrix} = \begin{bmatrix} I & r_{AB} & \dots & r_{AZ} \\ r_{BA} & I & \dots & r_{BZ} \\ \dots & \dots & \dots & \dots \\ r_{ZA} & r_{ZB} & \dots & I \end{bmatrix}^{-1} \begin{Bmatrix} r_{G \cdot A} \\ r_{G \cdot B} \\ \dots \\ r_{G \cdot Z} \end{Bmatrix}$$

à cette tradition — est l'équivalence donnée par WINTERS (1940) entre 1 kg de laine et 3,4 de viande chez le mouton. La difficulté majeure naît du fait que l'importance relative des caractères sélectionnés est généralement considérée sous son aspect économique. Il est certes très difficile, sinon d'évaluer, du moins de comparer le rôle économique du centimètre de tour de poitrine de poulet et du pourcentage d'éclosivité — pour donner un exemple. Aussi, LERNER, ASMUNDSON et CRUDEN (1948) ont-ils utilisé les index pour la sélection d'un seul caractère, mais à partir de plusieurs sortes de mesures, ou informations complémentaires, en attribuant la valeur $\alpha = 1$ au caractère choisi et les valeurs $\beta = \zeta = \dots = 0$

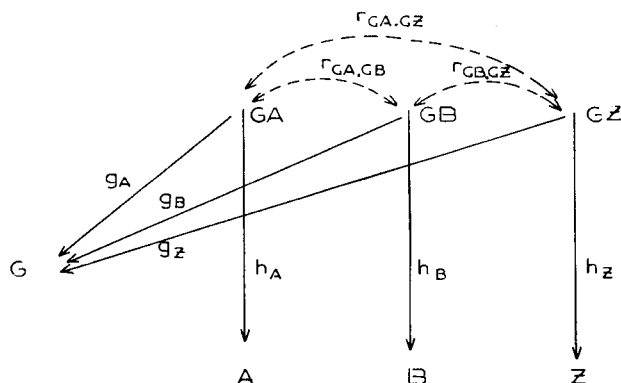


FIG. 17.

à tous les autres. Dans ces conditions le système d'équations extrait de la figure 17 se réduit sensiblement — soit :

$$\begin{aligned} r_{G \cdot A} &= h_A g_A \\ r_{G \cdot B} &= h_B r_{GB \cdot GA} g_A \\ \dots &\dots \dots \\ r_{G \cdot Z} &= h_Z r_{GZ \cdot GA} g_A \end{aligned}$$

On pourra donc prendre des valeurs α, β, \dots arbitraires, faute d'une recherche économique achevée en l'état actuel. Toutefois, l'on pourrait dénoncer le côté illusoire d'une telle étude, l'importance économique des caractères se pliant *ipso facto* aux fluctuations des marchés locaux et du goût des consommateurs. On peut également songer à déterminer l'importance relative à partir de la régression d'un caractère sur un autre plus facile à chiffrer, tel que le prix de l'œuf ou du kilo de viande.

Si cette méthode permet d'éviter les lenteurs d'une étude économique de très grande envergure, elle est néanmoins sujette aux variations saisonnières et annuelles des prix choisis pour référence.

On peut dès lors songer à aborder le problème non plus sous son aspect économique, mais technique, en choisissant des coefficients pro-

portionnels aux intensités de sélection fournissant les gains génétiques que l'on désire obtenir. Ceci nécessite la définition d'un animal idéal, mais n'implique pas que cet idéal ne puisse être modifié d'année en année. Telle est du moins l'approche du problème proposée par COCHEZ (1956), qui a le mérite d'être réaliste.

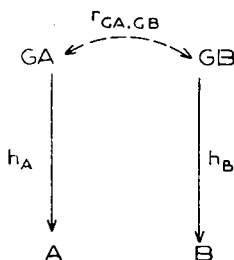


FIG. 18.

Il faut bien reconnaître pourtant que sur le point des coefficients d'importance relative des caractères, la théorie des index présente une faiblesse importante pour son application à la sélection. Elle n'en mérite que mieux d'être davantage étudiée à ce point de vue.

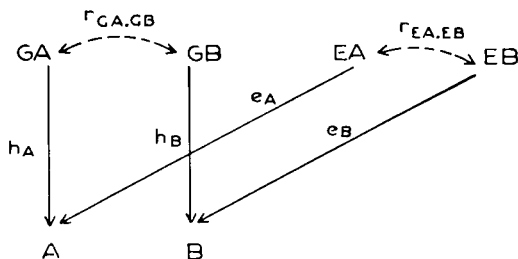


FIG. 19.

Ceci étant, nous supposons déterminées les valeurs de $\alpha, \beta, \dots, \zeta$. Le calcul de $g_A = \alpha \tau_{GA} / \tau_G$ etc... est alors achevé si l'on connaît le rapport τ_{GA} / τ_G .

Par définition de l'héritabilité on a $\tau_{GA} = h_A \tau_A$, calculable à partir des données. Le calcul de τ_G reviendrait à estimer l'héritabilité de l'index que l'on cherche précisément à construire. On ne calcule donc pas τ_G qui s'élimine d'ailleurs en facteur dans la suite des opérations.

b) Calcul des corrélations $r_{GA,GB}, r_{GA,GZ}, \dots$

Ce sont les corrélations génétiques entre les génotypes partiels GA, GB... GZ pris deux à deux.

On pourrait penser à calculer $r_{GA,GB}$ à partir du schéma représenté.

figure 18 : $r_{AB} = h_A h_B r_{GA \cdot GB}$, donnant la corrélation génétique lorsqu'on connaît les autres termes de cette égalité. Malheureusement ce calcul est faussé par l'existence toujours possible d'une corrélation due au milieu, soit $r_{EA \cdot EB}$; on a en fait (fig. 19) $r_{AB} = h_A h_B r_{GA \cdot GB} + e_A e_B r_{EA \cdot EB}$. Le second terme peut apporter une forte erreur d'estimation. On évite cet inconvénient, du moins en partie, en calculant $r_{GA \cdot GB}$ à partir d'animaux parents pour lesquels la corrélation due au milieu a plus de chances d'être nulle. Connaissant la régression j_{AB} de GA/GB et réciproquement, on a bien entendu :

$$r_{GA \cdot GB} = \sqrt{j_{AB} \cdot j_{BA}}$$

D'après la figure 20 on peut écrire :

$$r_{AA'} = rh_A^2 \quad \text{et} \quad r_{BB'} = rh_B^2.$$

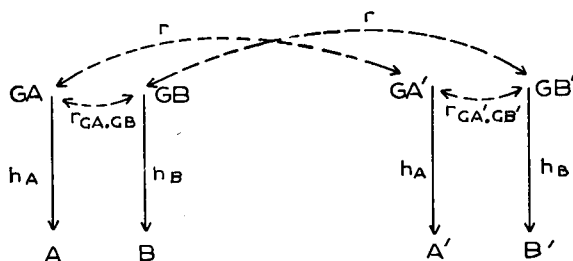


FIG. 20.

De cette figure on tire les schémas de calcul de la figure 21 (LUSH, 1948). On obtient :

$$r_{AB'} = rh_A h_B j_{AB} \quad \text{et} \quad r_{A'B} = rh_A h_B j_{BA}$$

on a donc :

$$\frac{r_{AB'} \cdot r_{A'B}}{r_{AA'} \cdot r_{BB'}} = \frac{rh_A h_B j_{AB} \cdot rh_A h_B j_{BA}}{rh_A^2 \cdot rh_B^2} = j_{AB} \cdot j_{BA} = r_{GA \cdot GB}^2$$

d'où l'on tire finalement :

$$r_{GA \cdot GB} = \sqrt{\frac{\text{cov } AB' \cdot \text{cov } A'B}{\text{cov } AA' \cdot \text{cov } BB'}}$$

calcul qui ne présente pas de difficulté à partir des données. Mais il n'est pas inutile de rappeler que cette estimation suppose l'absence de corrélation due au milieu entre le caractère A d'un animal et le caractère B' d'un animal parent.

Enfin, si l'une des covariances est négative, on prendra la moyenne arithmétique $1/2 (\text{cov } AB' + \text{cov } A'B)$ au lieu de la moyenne géométrique trouvée ici (HAZEL, 1943).

On peut aussi estimer les corrélations génétiques par l'analyse hiérar-

chique de la covariance, méthode analogue à celle utilisée pour l'estimation de l'héritabilité, les variances étant remplacées ici par les covariances. Cette manière d'opérer est sans doute plus rapide lorsque les calculs sont effectués concurremment — leur disposition étant indiquée tableau V.

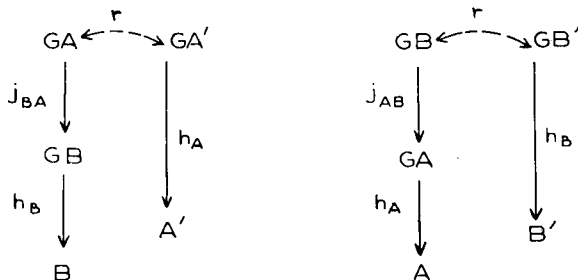


FIG. 21.

Nous n'insisterons pas davantage, ayant donné précédemment quelques références bibliographiques sur ce sujet.

On est ainsi en mesure de calculer les path-coefficients $ab...z$.

Calcul des rapports σ_G/σ_A , $\sigma_G/\sigma_B...$

Les coefficients de régression multiple sont déterminés par les égalités $k_A = a\sigma_G/\sigma_A$ etc... les écarts-types $\sigma_A\sigma_B...$ se déduisent directement des données tandis que σ_G s'élimine des calculs par mise en facteur, comme indiqué précédemment. En effet, dans l'équation de régression (I) σ_G est en facteur au numérateur. Par ailleurs, dans les calculs qui précèdent on voit que, STU étant « des expressions mathématiques excluant σ_G », on a $g_A = S/\sigma_G$ qui reporté dans le système d'équations précédent donne $r_{G \cdot A} = T/\sigma_G$ qui reporté à son tour dans le système d'équations relatif aux path-coefficients donne $a = U/\sigma_G$ qui reporté enfin dans l'équation de régression multiple donne le coefficient :

$$k_A = a\sigma_G/\sigma_A = U \sigma_G/\sigma_G\sigma_A = U/\sigma_A,$$

où ne figure pas σ_G .

Les coefficients de régression multiple étant alors déterminés, on est en possession de tous les éléments nécessaires à la construction de l'index de sélection.

Efficacité de cet index.

Nous l'étudierons pour raison de simplicité dans le cas de deux caractères A et B. La corrélation multiple entre le génotype total G d'une part, et les caractères A et B d'autre part, est telle que :

$$R_{G\hat{G}}^2 = \frac{I}{\sigma_G^2} (k_A \text{ cov GA} + k_B \text{ cov GB}) \text{ avec cov GA} = r_{G \cdot A} \sigma_G \sigma_A \text{ etc.}$$

TABLEAU V. — Calcul simultané de l'héritabilité de 2 caractères et de leurs corrélations génétique, phénotypique et due au milieu (données personnelles — poids de l'œuf et intensité de ponte, exprimés en probits).

| N° mère | Σx^2 | $(\Sigma x)^2/n$ | n | Σy^2 | Σy | $(\Sigma y)^2/n$ | n | Σxy | $\Sigma x \cdot \Sigma y/n$ | n^2 | n' | $\frac{\Sigma n^2}{n'}$ |
|--|---------------------|--------------------|---------|--------------|------------|------------------|-----|-------------|-----------------------------|---------------|----------------|-------------------------|
| 6014 | 28,09 | 5,30 | 1 | 35,28 | 5,94 | 35,28 | 1 | 31,48 | 31,48 | 1 | | |
| 6017 | 75,33 | 15,00 | 3 | 103,47 | 17,49 | 101,97 | 3 | 87,34 | 87,45 | 9 | | |
| 6052 | 204,65 | 41,76 | 9 | 210,10 | 42,45 | 200,22 | 9 | 201,82 | 196,97 | 81 | | |
| 6067 | 177,26 | 34,83 | 7 | 173,39 | 38,33 | 109,88 | 7 | 194,20 | 190,72 | 49 | | |
| 6234 | 149,06 | 34,26 | 8 | 265,13 | 45,78 | 261,08 | 8 | 197,84 | 196,05 | 64 | | |
| 6332 | 132,09 | 22,95 | 4 | 153,82 | 24,30 | 147,62 | 4 | 139,98 | 139,42 | 16 | | |
| 6390 | 192,88 | 33,61 | 6 | 141,34 | 23,70 | 138,14 | 6 | 101,39 | 101,27 | 36 | | |
| 6391 | 65,77 | 13,88 | 3 | 74,63 | 14,69 | 71,93 | 3 | 69,35 | 67,97 | 9 | | |
| Total 7 ♂ A5 | 1025,13 | 201,59 | 41 = n' | 1197,86 | 217,77 | 1167,02 | 41 | 1083,40 | 1071,33 | 265 | 1681 | 6,46 |
| | $(\Sigma x)^2/n' =$ | $(\Sigma y)^2/n =$ | | | | 1156,68 | | | 1070,74 | | | |
| 6683 | 168,79 | 33,78 | 7 | 197,07 | 36,81 | 193,57 | 7 | 179,25 | 177,63 | 49 | | |
| 6748 | 49,69 | 12,05 | 3 | 120,20 | 18,96 | 119,83 | 3 | 75,75 | 76,16 | 9 | | |
| 6752 | 23,01 | 6,30 | 2 | 48,41 | 9,82 | 48,22 | 2 | 31,96 | 31,23 | 4 | | |
| 6907 | 92,00 | 19,04 | 4 | 90,13 | 19,45 | 94,58 | 4 | 92,24 | 92,58 | 16 | | |
| 6907 | 134,8 | 13,48 | 3 | 90,96 | 16,44 | 90,09 | 3 | 73,66 | 73,87 | 9 | | |
| 7022 | 107,56 | 23,05 | 5 | 118,20 | 24,17 | 116,84 | 5 | 111,71 | 111,42 | 25 | | |
| 7075 | 141,7 | 66,93 | 3 | 72,84 | 14,73 | 72,32 | 3 | 69,52 | 69,57 | 9 | | |
| Total 7 ♂ B4 | 570,43 | 121,93 | 27 = n' | 743,81 | 140,38 | 735,45 | 27 | 634,09 | 632,46 | 121 | 729 | 4,48 |
| | $(\Sigma x)^2/n' =$ | $(\Sigma y)^2/n =$ | | | | 729,87 | | | 633,95 | | | |
| Total 107 ♂ A5 | 8440,29 | 1595,94 | 313 = N | 9204,60 | 1675,99 | 9074,83 | 313 | 8580,13 | 8555,53 | 1353/N = 4,32 | 7277/N = 23,25 | 59,83 |
| | $(\Sigma x)^2/n =$ | $(\Sigma y)^2/n =$ | | | | 8993,38 | | | 8562,67 | | | |
| $\begin{aligned} N_{x^2} &= 8137,46 \\ (8440,29 - 8295,64) &= (313 - 107)Hx \\ (8295,64 - 8210,59) &= (313 - 59,83)Mx + (107 - 15)Hx \\ (8210,59 - 8137,46) &= (313 - 23,25)Px + (59,83 - 4,32)Mx + (15 - 1)Hx \\ &= 15Hx \end{aligned}$ | | | | | | | | | | | | |
| $\begin{aligned} d'où \quad Hx &= 0,70218 \\ Mx &= 0,08678 \\ Px &= 0,20300 \\ &\text{composantes de la variance} \end{aligned}$ | | | | | | | | | | | | |
| $\begin{aligned} N_{y^2} &= 8974,26 \\ 120,77 &= 206Hy \\ 81,45 &= 253,17My + 92Hy \\ 19,12 &= 289,75Py + 55,51My + 14Hy \\ &\text{d'où} \quad Hy &= 0,62095 \text{ hasard} \\ &\quad My &= 0,09278 \text{ mère} \\ &\quad Py &= 0,01782 \text{ père} \end{aligned}$ | | | | | | | | | | | | |
| $\begin{aligned} Sy &= 0,74055 \\ h_{Mx}^2 &= 50,1\% \\ h_{Hy}^2 &= 9,6\% \\ h_{Px}^2 &= 29,9\% \end{aligned}$ | | | | | | | | | | | | |
| $\begin{aligned} h_{Mx}^2 &= 50,1\% \\ h_{Hy}^2 &= 9,6\% \\ h_{Px}^2 &= 29,9\% \end{aligned}$ | | | | | | | | | | | | |
| $\begin{aligned} h_{Mx}^2 &= 50,1\% \\ h_{Hy}^2 &= 9,6\% \\ h_{Px}^2 &= 29,9\% \end{aligned}$ | | | | | | | | | | | | |
| $\begin{aligned} h_{Mx}^2 &= 50,1\% \\ h_{Hy}^2 &= 9,6\% \\ h_{Px}^2 &= 29,9\% \end{aligned}$ | | | | | | | | | | | | |
| $\begin{aligned} h_{Mx}^2 &= 50,1\% \\ h_{Hy}^2 &= 9,6\% \\ h_{Px}^2 &= 29,9\% \end{aligned}$ | | | | | | | | | | | | |
| $\begin{aligned} h_{Mx}^2 &= 50,1\% \\ h_{Hy}^2 &= 9,6\% \\ h_{Px}^2 &= 29,9\% \end{aligned}$ | | | | | | | | | | | | |
| $\begin{aligned} h_{Mx}^2 &= 50,1\% \\ h_{Hy}^2 &= 9,6\% \\ h_{Px}^2 &= 29,9\% \end{aligned}$ | | | | | | | | | | | | |
| $\begin{aligned} h_{Mx}^2 &= 50,1\% \\ h_{Hy}^2 &= 9,6\% \\ h_{Px}^2 &= 29,9\% \end{aligned}$ | | | | | | | | | | | | |
| $\begin{aligned} h_{Mx}^2 &= 50,1\% \\ h_{Hy}^2 &= 9,6\% \\ h_{Px}^2 &= 29,9\% \end{aligned}$ | | | | | | | | | | | | |
| $\begin{aligned} h_{Mx}^2 &= 50,1\% \\ h_{Hy}^2 &= 9,6\% \\ h_{Px}^2 &= 29,9\% \end{aligned}$ | | | | | | | | | | | | |
| $\begin{aligned} h_{Mx}^2 &= 50,1\% \\ h_{Hy}^2 &= 9,6\% \\ h_{Px}^2 &= 29,9\% \end{aligned}$ | | | | | | | | | | | | |
| $\begin{aligned} h_{Mx}^2 &= 50,1\% \\ h_{Hy}^2 &= 9,6\% \\ h_{Px}^2 &= 29,9\% \end{aligned}$ | | | | | | | | | | | | |
| $\begin{aligned} h_{Mx}^2 &= 50,1\% \\ h_{Hy}^2 &= 9,6\% \\ h_{Px}^2 &= 29,9\% \end{aligned}$ | | | | | | | | | | | | |
| $\begin{aligned} h_{Mx}^2 &= 50,1\% \\ h_{Hy}^2 &= 9,6\% \\ h_{Px}^2 &= 29,9\% \end{aligned}$ | | | | | | | | | | | | |
| $\begin{aligned} h_{Mx}^2 &= 50,1\% \\ h_{Hy}^2 &= 9,6\% \\ h_{Px}^2 &= 29,9\% \end{aligned}$ | | | | | | | | | | | | |
| $\begin{aligned} h_{Mx}^2 &= 50,1\% \\ h_{Hy}^2 &= 9,6\% \\ h_{Px}^2 &= 29,9\% \end{aligned}$ | | | | | | | | | | | | |
| $\begin{aligned} h_{Mx}^2 &= 50,1\% \\ h_{Hy}^2 &= 9,6\% \\ h_{Px}^2 &= 29,9\% \end{aligned}$ | | | | | | | | | | | | |
| $\begin{aligned} h_{Mx}^2 &= 50,1\% \\ h_{Hy}^2 &= 9,6\% \\ h_{Px}^2 &= 29,9\% \end{aligned}$ | | | | | | | | | | | | |
| $\begin{aligned} h_{Mx}^2 &= 50,1\% \\ h_{Hy}^2 &= 9,6\% \\ h_{Px}^2 &= 29,9\% \end{aligned}$ | | | | | | | | | | | | |
| $\begin{aligned} h_{Mx}^2 &= 50,1\% \\ h_{Hy}^2 &= 9,6\% \\ h_{Px}^2 &= 29,9\% \end{aligned}$ | | | | | | | | | | | | |
| $\begin{aligned} h_{Mx}^2 &= 50,1\% \\ h_{Hy}^2 &= 9,6\% \\ h_{Px}^2 &= 29,9\% \end{aligned}$ | | | | | | | | | | | | |
| $\begin{aligned} h_{Mx}^2 &= 50,1\% \\ h_{Hy}^2 &= 9,6\% \\ h_{Px}^2 &= 29,9\% \end{aligned}$ | | | | | | | | | | | | |
| $\begin{aligned} h_{Mx}^2 &= 50,1\% \\ h_{Hy}^2 &= 9,6\% \\ h_{Px}^2 &= 29,9\% \end{aligned}$ | | | | | | | | | | | | |
| $\begin{aligned} h_{Mx}^2 &= 50,1\% \\ h_{Hy}^2 &= 9,6\% \\ h_{Px}^2 &= 29,9\% \end{aligned}$ | | | | | | | | | | | | |
| $\begin{aligned} h_{Mx}^2 &= 50,1\% \\ h_{Hy}^2 &= 9,6\% \\ h_{Px}^2 &= 29,9\% \end{aligned}$ | | | | | | | | | | | | |
| $\begin{aligned} h_{Mx}^2 &= 50,1\% \\ h_{Hy}^2 &= 9,6\% \\ h_{Px}^2 &= 29,9\% \end{aligned}$ | | | | | | | | | | | | |
| $\begin{aligned} h_{Mx}^2 &= 50,1\% \\ h_{Hy}^2 &= 9,6\% \\ h_{Px}^2 &= 29,9\% \end{aligned}$ | | | | | | | | | | | | |
| $\begin{aligned} h_{Mx}^2 &= 50,1\% \\ h_{Hy}^2 &= 9,6\% \\ h_{Px}^2 &= 29,9\% \end{aligned}$ | | | | | | | | | | | | |
| $\begin{aligned} h_{Mx}^2 &= 50,1\% \\ h_{Hy}^2 &= 9,6\% \\ h_{Px}^2 &= 29,9\% \end{aligned}$ | | | | | | | | | | | | |
| $\begin{aligned} h_{Mx}^2 &= 50,1\% \\ h_{Hy}^2 &= 9,6\% \\ h_{Px}^2 &= 29,9\% \end{aligned}$ | | | | | | | | | | | | |
| $\begin{aligned} h_{Mx}^2 &= 50,1\% \\ h_{Hy}^2 &= 9,6\% \\ h_{Px}^2 &= 29,9\% \end{aligned}$ | | | | | | | | | | | | |
| $\begin{aligned} h_{Mx}^2 &= 50,1\% \\ h_{Hy}^2 &= 9,6\% \\ h_{Px}^2 &= 29,9\% \end{aligned}$ | | | | | | | | | | | | |
| $\begin{aligned} h_{Mx}^2 &= 50,1\% \\ h_{Hy}^2 &= 9,6\% \\ h_{Px}^2 &= 29,9\% \end{aligned}$ | | | | | | | | | | | | |
| $\begin{aligned} h_{Mx}^2 &= 50,1\% \\ h_{Hy}^2 &= 9,6\% \\ h_{Px}^2 &= 29,9\% \end{aligned}$ | | | | | | | | | | | | |
| $\begin{aligned} h_{Mx}^2 &= 50,1\% \\ h_{Hy}^2 &= 9,6\% \\ h_{Px}^2 &= 29,9\% \end{aligned}$ | | | | | | | | | | | | |
| $\begin{aligned} h_{Mx}^2 &= 50,1\% \\ h_{Hy}^2 &= 9,6\% \\ h_{Px}^2 &= 29,9\% \end{aligned}$ | | | | | | | | | | | | |
| $\begin{aligned} h_{Mx}^2 &= 50,1\% \\ h_{Hy}^2 &= 9,6\% \\ h_{Px}^2 &= 29,9\% \end{aligned}$ | | | | | | | | | | | | |
| $\begin{aligned} h_{Mx}^2 &= 50,1\% \\ h_{Hy}^2 &= 9,6\% \\ h_{Px}^2 &= 29,9\% \end{aligned}$ | | | | | | | | | | | | |
| $\begin{aligned} h_{Mx}^2 &= 50,1\% \\ h_{Hy}^2 &= 9,6\% \\ h_{Px}^2 &= 29,9\% \end{aligned}$ | | | | | | | | | | | | |
| $\begin{aligned} h_{Mx}^2 &= 50,1\% \\ h_{Hy}^2 &= 9,6\% \\ h_{Px}^2 &= 29,9\% \end{aligned}$ | | | | | | | | | | | | |
| $\begin{aligned} h_{Mx}^2 &= 50,1\% \\ h_{Hy}^2 &= 9,6\% \\ h_{Px}^2 &= 29,9\% \end{aligned}$ | | | | | | | | | | | | |
| $\begin{aligned} h_{Mx}^2 &= 50,1\% \\ h_{Hy}^2 &= 9,6\% \\ h_{Px}^2 &= 29,9\% \end{aligned}$ | | | | | | | | | | | | |
| $\begin{aligned} h_{Mx}^2 &= 50,1\% \\ h_{Hy}^2 &= 9,6\% \\ h_{Px}^2 &= 29,9\% \end{aligned}$ | | | | | | | | | | | | |
| $\begin{aligned} h_{Mx}^2 &= 50,1\% \\ h_{Hy}^2 &= 9,6\% \\ h_{Px}^2 &= 29,9\% \end{aligned}$ | | | | | | | | | | | | |
| $\begin{aligned} h_{Mx}^2 &= 50,1\% \\ h_{Hy}^2 &= 9,6\% \\ h_{Px}^2 &= 29,9\% \end{aligned}$ | | | | | | | | | | | | |
| $\begin{aligned} h_{Mx}^2 &= 50,1\% \\ h_{Hy}^2 &= 9,6\% \\ h_{Px}^2 &= 29,9\% \end{aligned}$ | | | | | | | | | | | | |
| < | | | | | | | | | | | | |

d'où encore :

$$R_{GG}^2 = k_A \frac{\sigma_A}{\sigma_G} r_{G \cdot A} + k_B \frac{\sigma_B}{\sigma_G} r_{G \cdot B} = ar_{G \cdot A} + br_{G \cdot B},$$

formule analogue à celle trouvée pour la sélection combinée d'un caractère (soit $R^2 = pr_{G \cdot P} + fr_{G \cdot F}$).

Enfin, en explicitant $r_{G \cdot A}$ et $r_{G \cdot B}$ suivant le système d'équations extrait de la figure 17, on écrit :

$$R_{GG}^2 = \frac{1}{\sigma_G^2} [k_A h_A \sigma_A (\alpha \sigma_{GA} + r_{GA \cdot GB} \zeta \sigma_{GB}) + k_B h_B \sigma_B (r_{GA \cdot GB} \alpha \sigma_{GA} + \zeta \sigma_{GB})]$$

formule que nous appliquerons à quelques exemples.

1° Soit $\alpha = 1$ et $\beta = 0$.

Dans ces conditions, le génotype total estimé G se réduit au génotype partiel GA ; mais nous supposons que la corrélation génétique $r_{GA \cdot GB}$ n'est pas nulle. Il vient :

$$R^2 = \frac{1}{\sigma_{GA}^2} (k_A h_A \sigma_A \sigma_{GA} + k_B h_B \sigma_B r_{GA \cdot GB} \sigma_{GA})$$

$$\text{Avec} \quad h_A \sigma_A = \sigma_{GA} \quad \text{et} \quad h_B \sigma_B = \sigma_{GB}$$

$$\text{d'où} \quad R^2 = k_A + k_B \frac{\sigma_{GB}}{\sigma_{GA}} r_{GA \cdot GB} = a \frac{\sigma_{GA}}{\sigma_A} + b \frac{\sigma_{GB}}{\sigma_B} r_{GA \cdot GB}$$

en explicitant k_A et k_B . D'où enfin :

$$R^2 = ah_A + bh_B r_{GA \cdot GB}.$$

Si A est le phénotype P de l'individu et B la moyenne de sa famille F, on peut écrire :

$$a = p ; \quad h_A = h = r_{GP} ; \quad b = f ; \quad h_B r_{GA \cdot GB} = h_F r_{G \cdot GF} = r_{GF}$$

On a donc :

$$R^2 = pr_{GP} + fr_{GF}.$$

En d'autres termes, la sélection combinée d'un seul caractère est un cas particulier de la sélection pour deux caractères, en attribuant à l'un d'eux une importance relative nulle ; sa prise en considération dans l'index n'intervenant que par le truchement de la corrélation génétique $r_{GA \cdot GB}$. Dans le cas de la sélection combinée, on attribue donc l'importance relative 1 et 0 au phénotype individuel et à la moyenne de famille, respectivement.

On voit par cet exemple que les coefficients d'importance relative ne correspondent pas du tout aux coefficients de régression multiple. C'est à tort que dans la sélection combinée (cf. plus haut) l'on dit que l'importance relative de P et F est 1 et W. Ceci permet d'ajouter qu'il serait possible d'utiliser la sélection combinée sous une autre forme, en choisissant $\beta \neq 0$, c'est-à-dire en considérant la moyenne de famille

comme un caractère individuel. Pour théorique que soit cette remarque, elle ne semble pas avoir été suffisamment soulignée.

2° Soit $\alpha = 1$, $\beta = 0$, $r_{GA,GB} = 0$.

Nous sommes ramenés au cas de la sélection massale d'un seul caractère — ou encore au cas de la « sélection successive » de plusieurs caractères (HAZEL et LUSH, 1942). Le coefficient de corrélation entre le génotype total G et sa valeur probable donnée par l'équation de régression est alors tel que : $R^2 = \frac{1}{\sigma_G^2} k_A h_A \sigma_A \sigma_{GA}$; soit $R^2 = r^2 = h_A^2$ si l'on considère que le génotype total G se réduit au génotype partiel GA.

D'autre part, si l'on pose $\alpha = 1$, $\beta = 1$, et $r_{GA,GB} = 0$, nous nous plaçons dans le cas de deux caractères indépendants ayant même importance relative — ou encore dans le cas de la sélection simultanée de deux caractères (ou « mérite total » de HAZEL et LUSH, 1942). Dans ces conditions, il vient :

$$R^2 = \frac{1}{\sigma_G^2} (k_A h_A \sigma_A \sigma_{GA} + k_B h_B \sigma_B \sigma_{GB}).$$

Si, pour la démonstration, nous supposons que $h_A = h_B$ et $\sigma_A = \sigma_B$, nous voyons que $R^2 = 2r^2$. L'efficacité de la sélection simultanée par rapport à la sélection successive est donc $R/r = \sqrt{2}$. Comme le note LERNER (1950) cette efficacité serait égale à \sqrt{n} pour la sélection de n caractères ayant mêmes héritabilités et mêmes variances.

En somme, dans tous les cas où $r_{GA,GB}$ a une valeur positive ou nulle, la sélection simultanée de plusieurs caractères est plus intéressante que la sélection d'un seul. Dans le cas de corrélations génétiques négatives, le gain génétique obtenu sur chaque caractère sera moindre que s'il était amélioré seul. Mais cela est vrai aussi en sens inverse : la détérioration d'un caractère non sélectionné sera moindre. En somme, un index de sélection sur plusieurs caractères permet soit d'obtenir un gain génétique total plus grand que celui obtenu sur un seul caractère, soit d'équilibrer la valeur génétique d'un troupeau au lieu d'en exagérer les qualités et les défauts. Bien plus, les corrélations étant rarement parfaites, il est clair que si leurs valeurs négatives amoindrissent le gain génétique, cela ne signifie pas nécessairement que ce gain soit nul ou négatif : il est des animaux pour lesquels ce que l'on pourrait appeler la « corrélation individuelle » est positive. Le résultat dépendra en fin de compte de la pression de sélection exercée.

B. — Sélection combinée de plusieurs caractères

Dans cette hypothèse, l'estimation du génotype total G est fonction des mesures individuelles AB... Z et des moyennes de familles correspondantes FA, FB... FZ. La méthode est donc analogue à la précédente : elle

Par suite il vient :

$$r_{A \cdot FA} = \sqrt{\frac{I + (n - I)t_A}{n}}$$

3° Calcul des corrélations du type $r_{A \cdot FB}$.

C'est la corrélation entre la valeur individuelle pour un caractère

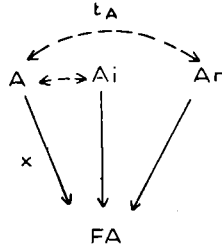


FIG. 23.

et moyenne de famille pour un autre caractère. D'après la figure 24 on peut écrire :

$$r_{A \cdot FB} = r_{AB}y + (n - I)t_{AB}y, \quad \text{avec} \quad I = ny^2 + n(n - I)t_By^2,$$

donnant y . D'où enfin,

$$r_{A \cdot FB} = \frac{r_{AB} + (n - I)t_{AB}}{\sqrt{n[I + (n - I)t_B]}}$$

(Cf. note à la fin du mémoire, pour le calcul direct de t_{AB}).

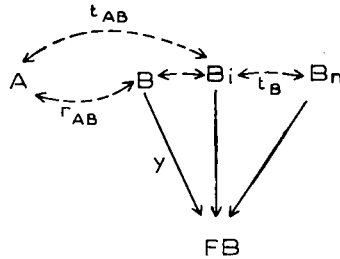


FIG. 24.

4° Calcul des corrélations du type $r_{FA \cdot FB}$.

C'est la corrélation entre moyennes de famille pour deux caractères différents. D'après la figure 25 on a :

$$r_{FA \cdot FB} = xy[nr_{AB} + n(n - I)t_{AB}]$$

les path-coefficients x et y , de la moyenne de famille par le phénotype, ont été calculés précédemment. Il vient :

$$r_{FA \cdot FB} = \frac{n[r_{AB} + (n - I)t_{AB}]}{\sqrt{n[I + (n - I)t_A]}\sqrt{n[I + (n - I)t_B]}}$$

ou plus simplement :

$$r_{FA \cdot FB} = \frac{r_{AB} + (n-1)t_{AB}}{\sqrt{[1 + (n-1)t_A][1 + (n-1)t_B]}}$$

5° Calcul des corrélations du type $r_{G \cdot A}$.

Il s'agit de la corrélation entre génotype total et phénotype indivi-

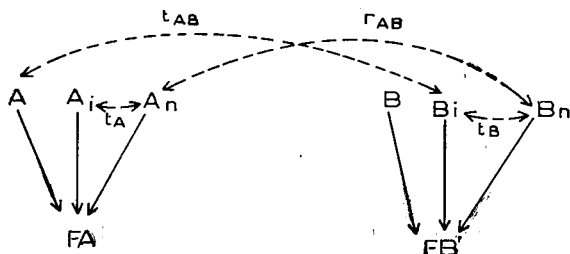


FIG. 25.

duel pour un caractère. Ce calcul a déjà été indiqué à propos de la sélection massale de plusieurs caractères. Soit (fig. 26) :

$$r_{G \cdot A} = h_A(g_A + r_{GA \cdot GB}g_B + \dots + r_{GA \cdot GZ}g_Z) \quad \text{avec} \quad g_A = \alpha h_A \frac{\sigma_A}{\sigma_G}$$

et

$$r_{GA \cdot GH} = \sqrt{\text{cov } AB' \cdot \text{cov } A'B / \text{cov } AA' \cdot \text{cov } BB'}.$$

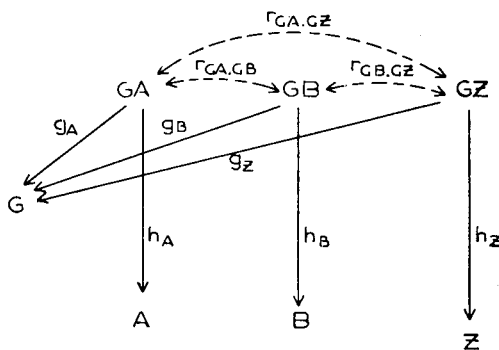


FIG. 26.

6° Calcul des corrélations du type $r_{G \cdot FA}$.

C'est la corrélation entre génotype total et moyenne de famille pour un caractère. D'après le diagramme de causalité représenté figure 27 on peut écrire :

$$r_{G \cdot FA} = \alpha(r_{G \cdot A} + r_{G \cdot A1} + \dots + r_{G \cdot An})$$

avec toujours

$$\alpha = \frac{1}{\sqrt{n[1 + (n-1)t_A]}}$$

Par ailleurs il est nécessaire de calculer les corrélations $r_{G \cdot A}$,

$r_{G \cdot A_i} \dots r_{G \cdot A_n}$.

On a le système suivant :

$$\begin{aligned} r_{G \cdot A} &= h_A(g_A + r_{GA \cdot GB}g_B + \dots + r_{GA \cdot GZ}g_Z) \\ r_{G \cdot A_i} &= h_A(r_{GA} + r_{GA_i \cdot GB}g_B + \dots + r_{GA_i \cdot GZ}g_Z) \\ &\dots \dots \dots \\ r_{G \cdot A_n} &= h_A(r_{GA} + r_{GA_n \cdot GB}g_B + \dots + r_{GA_n \cdot GZ}g_Z). \end{aligned}$$

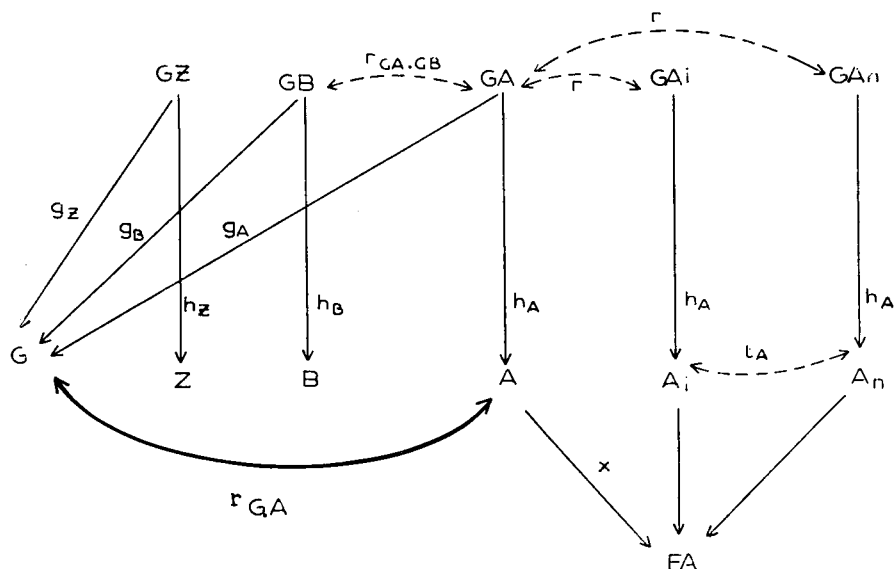


FIG. 27.

On est ainsi amené à calculer $r_{GA_i \cdot GB}$. A partir des régressions de A_i sur B et réciproquement, on peut écrire :

$$r_{GA_i \cdot GB} = \sqrt{j_{A_iB} \cdot j_{BA_i}}.$$

D'après les schémas de calcul représentés figure 28, on a les égalités :

$$\begin{aligned} \text{(I)} \quad r_{A_iB} &= rh_A h_B j_{A_iB} & \text{(II)} \quad r_{AB} &= rh_A h_B j_{BA_i} \\ r_{BBi} &= rh_B^2 & r_{AA_i} &= rh_A^2. \end{aligned}$$

Par suite il vient :

$$r_{GA_i \cdot GB} = \sqrt{j_{A_iB} \cdot j_{BA_i}} = \sqrt{\frac{r_{AB} \cdot r_{A_iB}}{r_{AA_i} \cdot r_{BBi}}}.$$

Mais on a d'autre part :

$$\begin{aligned} \text{(III)} \quad r_{A_iB} &= rh_A h_B j_{A_iB} & \text{(IV)} \quad r_{AB} &= rh_A h_B j_{A_iB} \\ r_{A_iB} &= h_A h_B j_{A_iB} & r_{ABi} &= rh_A h_B j_{A_iB} \end{aligned}$$

d'où

$$r_{A_iB} = r \cdot r_{A_iB} \quad \text{et} \quad r_{AB} = r \cdot r_{ABi}.$$

En remplaçant r_{AiB} et r_{AiBi} par leur valeur il vient :

$$r_{GAi \cdot GB} = r \sqrt{\frac{r_{ABi} \cdot r_{AiB}}{r_{AAi} \cdot r_{BBi}}}.$$

Dans le terme sous-radical, on reconnaît la corrélation génétique $r_{GA \cdot GB}$ précédemment calculée. On a donc enfin :

$$r_{GAi \cdot GB} = r \cdot r_{GA \cdot GB}.$$

Par suite le système d'équations précédent devient :

$$\begin{aligned} r_{G \cdot A} &= h_A(g_A + r_{GA \cdot GB}g_B + \dots + r_{GA \cdot GZ}g_Z) \\ r_{G \cdot Ai} &= h_A(rg_A + r \cdot r_{GA \cdot GB}g_B + \dots + r \cdot r_{GA \cdot GZ}g_Z) = r \cdot r_{G \cdot A} \\ &\dots \dots \dots \\ r_{G \cdot An} &= h_A(rg_A + r \cdot r_{GA \cdot GB}g_B + \dots + r \cdot r_{GA \cdot GZ}g_Z) = r \cdot r_{G \cdot A}. \end{aligned}$$

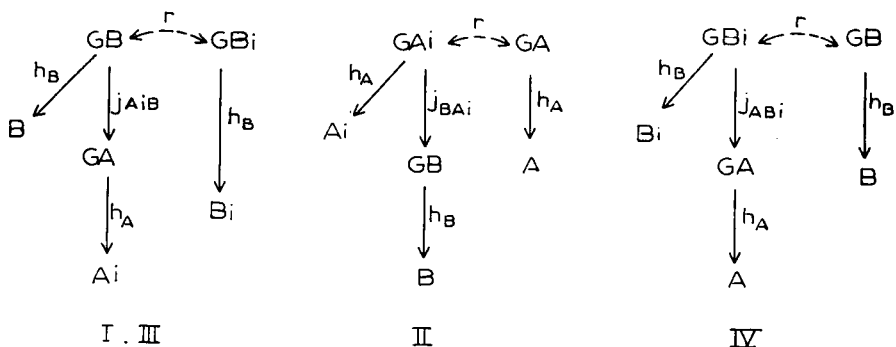


FIG. 28.

On peut donc écrire :

$$r_{G \cdot FA} = x[r_{G \cdot A} + (n - 1)r \cdot r_{G \cdot A}].$$

ou enfin :

$$r_{G \cdot FA} = r_{G \cdot A} \frac{1 + (n - 1)r}{\sqrt{n[1 + (n - 1)t]}}.$$

On est alors en possession de tous les éléments nécessaires au calcul de l'index de sélection combinée pour plusieurs caractères.

C. — Sélection de plusieurs caractères à partir des index partiels

Nous avons vu, dans la première partie de ce mémoire, la construction d'un index en sélection combinée pour un seul caractère. Soit l'équation :

$$\hat{G} - \bar{G} = \frac{1 - r}{1 - t} h^2(P - \bar{P}) + \frac{r - t}{1 - t} \cdot \frac{n}{1 + (n - 1)t} h^2(F - \bar{F})$$

Pour résoudre cette matrice il faut connaître par ailleurs les corrélations entre index partiels $r_{IA, IB}$ etc... dont le calcul n'offre aucune difficulté, et les corrélations $r_{G, IA}$ etc... entre le génotype total G et chaque index partiel.

Calcul des corrélations du type $r_{G, IA}$

Toutes les corrélations n'ont pas été représentées sur la figure 30, de façon à alléger le diagramme de causalité. On voit néanmoins les

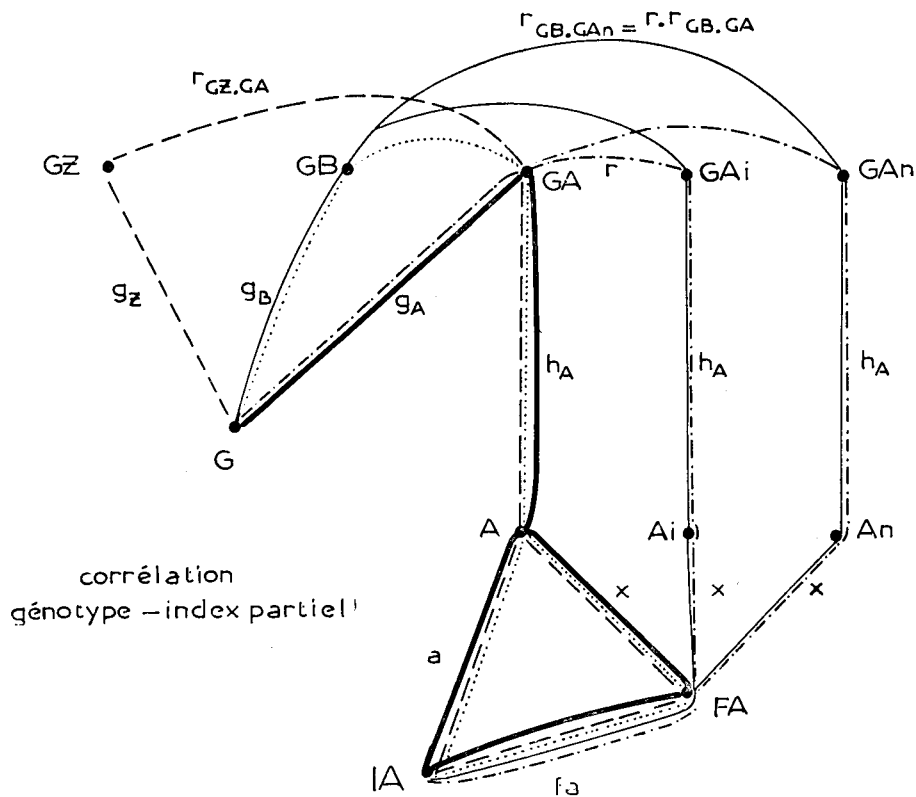


FIG. 30.

différentes sortes de « chemins » impliqués dans le calcul : 1° — les chemins passant par A — soit d'une part, le chemin direct $G \rightarrow GA \rightarrow A$, d'autre part les chemins passant par les génotypes partiels GB... GZ, par raison de corrélation génétique, $G \rightarrow GB \rightarrow GA \rightarrow A$, etc... 2° — les chemins ne passant pas par A mais passant par Ai... An. Il y en a donc $(n - 1)$, intervenant seulement par l'intermédiaire de la moyenne de famille. Soit d'une part les chemins passant par GA : $G \rightarrow GA \rightarrow GAn \rightarrow An \rightarrow FA$, pour lesquels la corrélation

$r_{GB \cdot GAi}$ est simplement la parenté r ; d'autre part, les chemins ne passant pas par GA : $G - GB - GAi - Ai - FA \dots G - GZ - GAn - An - FA$ pour lesquels la corrélation $r_{GB \cdot GAi}$ est égale au produit $r \cdot r_{GA \cdot GB}$ comme nous l'avons démontré précédemment.

On a donc :

$$\begin{array}{l} \text{-----} \quad r_{G \cdot IA} = g_A h_A (a + xfa) \\ \text{.....} \quad + g_B r_{GB \cdot GA} h_A (a + xfa) \\ \text{.....} \quad + g_Z r_{GZ \cdot GA} h_A (a + xfa) \\ \text{-----} \quad + g_A r h_A xfa (n - 1) \\ \text{-----} \quad + g_B (r \cdot r_{GB \cdot GA}) h_A xfa (n - 1) \\ \text{-----} \quad + g_Z (r \cdot r_{GZ \cdot GA}) h_A xfa (n - 1) \end{array} \left. \begin{array}{l} \\ \\ \\ \\ \end{array} \right\} \begin{array}{l} \text{chemins passant par A} \\ \\ \\ \text{chemins passant} \\ \text{par Ai} \dots \text{An.} \end{array}$$

Cette égalité se simplifie par mise en facteurs. Il vient :

$$\begin{aligned} r_{G \cdot IA} &= (a + xfa) h_A (g_A + g_B r_{GB \cdot GA} + \dots + g_Z r_{GZ \cdot GA}) \\ &+ (n - 1) r xfa h_A (g_A + g_B r_{GB \cdot GA} + \dots + g_Z r_{GZ \cdot GA}) \end{aligned}$$

ou plus succinctement :

$$r_{G \cdot IA} = \{ a + fa[1 + (n - 1)r]x \} h_A (g_A + g_B r_{GB \cdot GA} + \dots + g_Z r_{GZ \cdot GA}).$$

On reconnaît dans ce produit l'expression analytique de la corrélation $r_{G \cdot A}$ calculée précédemment. Par ailleurs, on a vu que

$$x = \frac{1}{\sqrt{n[1 + (n - 1)t_A]}}.$$

Par suite, en remplaçant terme à terme, il vient :

$$r_{G \cdot IA} = r_{G \cdot A} \left(a + fa \frac{1 + (n - 1)r}{\sqrt{n[1 + (n - 1)t_A]}} \right).$$

Or, par définition du path-coefficient, on a :

$$a = k_A \frac{\sigma_A}{\sigma_{IA}} \quad \text{avec} \quad k_A = \frac{1 - r}{1 - t_A} h_A^2$$

(cf. par. sélection combinée pour un caractère).

$$\text{Et} \quad fa = k_{FA} \frac{\sigma_{FA}}{\sigma_{IA}} \quad \text{avec} \quad k_{FA} = \frac{r - t_A}{1 - t_A} \cdot \frac{n}{1 + (n - 1)t_A} h_A^2$$

$$\text{et} \quad \sigma_{FA} = \sqrt{\frac{1 + (n - 1)t_A}{n}} \sigma_A.$$

Par suite,

$$a = \frac{1 - r}{1 - t_A} h_A^2 \frac{\sigma_A}{\sigma_{IA}} \quad \text{et} \quad fa = \frac{r - t_A}{1 - t_A} \sqrt{\frac{n}{1 + (n - 1)t_A}} h_A^2 \frac{\sigma_A}{\sigma_{IA}}.$$

Par substitution dans l'équation précédente, on obtient donc :

$$r_{G \cdot IA} = r_{G \cdot A} \frac{\sigma_A}{\sigma_{IA}} \left[\frac{1 - r}{1 - t_A} h_A^2 + \frac{r - t_A}{1 - t_A} \cdot \frac{1 + (n - 1)r}{1 + (n - 1)t_A} h_A^2 \right].$$

Dans le terme entre crochets on reconnaît le carré de la corrélation multiple entre le génotype GA d'une part et ses composantes A et FA (valeur individuelle et moyenne de famille), d'autre part ; soit $R_{GA \cdot IA}^2$ cette expression, l'index partiel I_A étant identique à la valeur probable de GA donnée par l'équation de régression de la sélection combinée.

On peut donc écrire :

$$r_{G \cdot IA} = r_{G \cdot A} R_{GA \cdot IA}^2 \frac{\sigma_A}{\sigma_{IA}}.$$

D'une manière inverse, si l'on considère l'index partiel I_A comme cause « indépendante » du génotype partiel GA (fig. 31), le path-coefficient de GA par I_A est égal à la corrélation $R_{GA \cdot IA}$; soit :

$$R_{GA \cdot IA} = \frac{\sigma_{IA}}{\sigma_{GA}} = \frac{\sigma_{IA}}{h_A \sigma_A}.$$

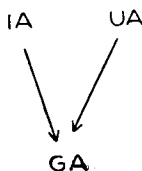


FIG. 31.

Si les données sont exprimées en écarts-réduits, on a évidemment $\sigma_A = 1$. Par suite, il vient :

$$r_{G \cdot IA} = r_{G \cdot A} \frac{\sigma_{IA}}{h_A^2}.$$

On est ainsi en possession de tous les éléments nécessaires au calcul de l'index total de sélection combinée pour plusieurs caractères.

*
* *

Notons que dans ce calcul nous avons attribué la même importance relative à tous les index partiels. Dans le cas général, il faut écrire $R_{GA \cdot IA} = \alpha' \sigma_{IA} / h_A \sigma_A$ etc... α' étant le coefficient d'importance relative de l'index partiel I_A . Nous en avons discuté d'une manière analogue à propos de la sélection massale de plusieurs caractères. Nous n'y reviendrons donc pas puisqu'il suffit de remplacer « valeur individuelle » par « index partiel ».

En résumé, les index partiels I_A, I_B, \dots, I_z pouvant être calculés rapidement à partir d'un jeu de tables établies une fois pour toutes ; les données étant transformées en écarts-réduits ; l'estimation d'un génotype

total G à partir des index partiels se réduit à des calculs moins longs qu'en utilisant la sélection combinée sous sa forme analytique.

Il suffit en effet de calculer les corrélations entre index partiels pris deux à deux : $r_{IA,IB}$ etc... et les corrélations du type $r_{G,IA}$ etc... A leur tour, celles-ci se ramènent à l'estimation de $r_{G,A}$ etc... qui se réduit elle-même au calcul des corrélations génétiques. On économise en fin de compte le calcul des corrélations du type $r_{A,FA}$, $r_{A,FB}$, $r_{FA,FB}$ et $r_{G,FA}$ par rapport à la méthode précédente — gain d'autant plus appréciable que ces corrélations devraient être traduites en chiffres pour chaque famille.

D. — Généralisation de la sélection par index

Nous avons vu la méthode de construction des index à partir de données diverses. Nous ne poursuivrons pas ici davantage. Disons néanmoins qu'il est possible d'établir un index de sélection familiale ou parentale à partir des éléments analysés précédemment. Notons seulement que

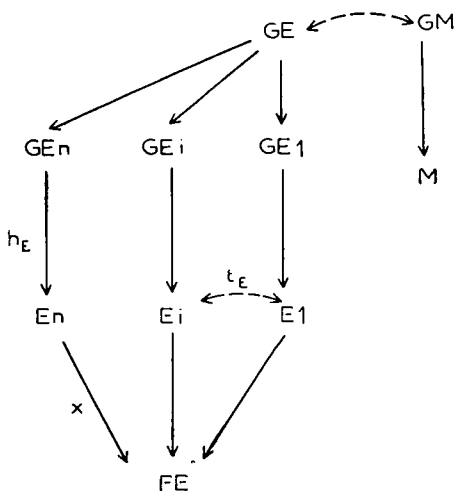


FIG. 32.

dans le cas de la sélection parentale, la corrélation génotype-moyenne de famille devient :

$$r_{G,FA'} = r_{G,A} \frac{nr}{\sqrt{n[I + (n-1)t_A]}} \quad \text{au lieu de} \quad r_{G,FA} = r_{G,A} \frac{I + (n-1)r}{\sqrt{n[I + (n-1)t_A]}}.$$

On peut encore introduire dans l'index un renseignement donné par la performance d'un ascendant. Il s'agit alors de calculer la corrélation entre le phénotype de l'ascendant et celui de l'enfant, soit r_{MRI} (fig. 32)

ce qui ne présente aucune difficulté ; par ailleurs, la corrélation entre phénotype de l'ascendant et moyenne de ses enfants. Soit (fig. 32) :

$$r_{M \cdot FE} = n r_{MEI} x \quad \text{avec} \quad x = \frac{1}{\sqrt{n[1 + (n-1)t_E]}}$$

d'où :

$$r_{M \cdot FE} = r_{MEI} \sqrt{\frac{n}{1 + (n-1)t_E}}.$$

Enfin, la combinaison au gré de chacun des divers processus que nous avons analysés de manière systématique, conduit à la généralisation de la notion d'index, achevant ainsi la synthèse de toutes les méthodes de sélection.

CONCLUSIONS

Un index de sélection est défini par l'équation de régression multiple du génotype sur un certain nombre de données. Soit :

$$I = k_A(A - \bar{A}) + k_B(B - \bar{B}) + \dots + k_Z(Z - \bar{Z}),$$

où les critères de sélection AB... Z peuvent représenter, indifféremment, des performances individuelles, familiales, parentales — ou même des index partiels (valeurs combinées).

Plutôt que d'isoler les diverses méthodes de sélection usuelles comme de simples « recettes » ou que d'en dresser le catalogue, il nous paraît plus logique de les réunir dans un même cadre, en tant qu'applications d'une même équation algébrique.

Nous nous sommes efforcé d'en démontrer le mécanisme de façon à rendre l'exposé utilisable à partir de données réelles. On conviendra toutefois que la masse des calculs impliqués dans les index complexes (2^e partie) ne laisse pas d'être un obstacle à leur application pratique. Bien plus, une telle manière d'opérer suppose le maintien du troupeau jusqu'à l'ultime caractère — en procédant par ordre chronologique — introduit dans cette construction.

Plusieurs simplifications sont néanmoins possibles. La transformation des mesures en écarts-réduits ou en probits (COCHEZ, 1954) permet d'établir des tables de calculs une fois pour toutes ; l'estimation de l'hérabilité des caractères et celle des corrélations génétiques peut n'être répétée que toutes les 3 ou 4 générations ; la corrélation phénotypique intra-famille peut être, aux débats, évaluée à $r h^2$ — en supposant nul l'effet du milieu commun aux membres d'une même famille, si la structure de l'élevage s'y prête. Notons aussi, pour la résolution des systèmes d'équations, l'intérêt des mesures indépendantes — ne serait-ce que phénotypiquement. Il n'empêche que dès que l'on dépasse le nombre de 5 ou 6 critères de sélection, il convient d'envisager l'utilisation du calcul mécanographique. Au total, la sélection par index, dans la mesure où

l'on désire les parfaire, représente un gros travail de documentation, de dépouillement et d'analyse des résultats. Ceci nécessite donc une organisation particulière de la sélection, mais ne présente pas de difficulté essentielle.

*
* *

Le principal intérêt des index de sélection repose sur leur efficacité. Celle-ci dépend de la corrélation multiple entre génotype et index, telle que :

$$R_{G.I}^2 = \frac{1}{\sigma_G^2} [k_A h_A \sigma_A (x h_A \sigma_A + r_{GA \cdot GB} \rho h_B \sigma_B + \dots + r_{GA \cdot GZ} \zeta h_Z \sigma_Z) + \dots + k_Z h_Z \sigma_Z (r_{GZ \cdot GA} x h_A \sigma_A + r_{GZ \cdot GB} \rho h_B \sigma_B + \dots + \zeta h_Z \sigma_Z)].$$

Nous nous sommes attaché à discuter l'efficacité des diverses méthodes de sélection généralement utilisées : massale, familiale, parentale, sur l'ascendance ou la descendance, combinée — pour 1 ou plusieurs caractères.

D'une façon plus générale, il convient de se rappeler que le gain génétique total est donné par l'égalité :

$$\Delta GX = R_{GX} h i.$$

Or, l'utilisation d'un index sous une forme ou sous une autre, revient à agir sur le terme R_{GX} , corrélation multiple entre génotype et critère de sélection, celui-ci pouvant aller du phénotype individuel à l'index le plus complexe.

Il faut bien reconnaître que la notion d'index et plus précisément son utilité se limite au facteur R_{GX} .

La modification de l'héritabilité des caractères ou de l'intensité de la sélection concerne d'autres facteurs du progrès génétique qui ne peuvent être négligés, mais dont l'étude dépasse de loin le cadre de cette publication.

Bien plus, le gain phénotypique résultant, seul observable par l'éleveur, conduit à poser la question de la rentabilité de la sélection.

En d'autres termes, constatant que les méthodes usuelles de sélection se ramènent en dernière analyse à la notion d'index, on peut avancer que la théorie des index intéresse tout sélectionneur, bien qu'elle ne concerne qu'une partie de son activité. Il n'est donc pas inutile en terminant de remonter aux sources, en remplaçant le problème des index et plus avant du gain génétique dans son contexte réel, à savoir comment doit agir l'éleveur pour améliorer la valeur économique de sa souche. Dans cette optique, la théorie des index n'est plus qu'un élément parmi d'autres d'un domaine plus vaste à exploiter, du ressort de la recherche opérationnelle (WADDINGTON, 1951; MCCLOSKEY et TREFETHEN, 1957).

NOTE

Calcul des corrélations phénotypiques intra-familles.

Pour calculer le paramètre t , on peut prendre pour base de départ la formule du coefficient de corrélation inter-classe :

$$r = \frac{\Sigma(x - \bar{x})(y - \bar{y})}{\sqrt{\Sigma(x - \bar{x})^2 \cdot \Sigma(y - \bar{y})^2}}.$$

Mais, dans le cas présent, x et y représentent la mesure d'un même caractère sur 2 individus parents. Autrement dit, les n membres de chaque famille sont comparés deux à deux, la valeur de chacun jouant tour à tour le rôle de x , puis de y . On obtient ainsi deux séries de valeurs :

| | | | | | |
|-------------------------|---------|---------------|---------------|---------------|---------------|
| 1 ^{re} famille | (x) | $x_1 x_1 x_1$ | $x_2 x_2 x_2$ | $x_3 x_3 x_3$ | $x_n x_n x_n$ |
| | (y) | $x_2 x_3 x_n$ | $x_1 x_3 x_n$ | $x_1 x_2 x_n$ | $x_1 x_2 x_3$ |
| 2 ^e famille | (x) | $x'_1 x'_1$ | $x'_2 x'_2$ | $x'_n x'_n$ | |
| | (y) | $x'_2 x'_n$ | $x'_1 x'_n$ | $x'_1 x'_2$ | |

Chaque série (x) ou (y) contient F familles. La série (x) montre que chaque famille contient n groupe de valeurs ; la série (y) montre que chacun de ces groupes contient $(n - 1)$ termes.

En somme, dans chaque famille, le terme général x_i est présent $(n - 1)$ fois dans le i^{me} groupe pour la série (x) et une fois dans $(n - 1)$ groupes pour la série (y). Chaque famille contient donc $n(n - 1)$ termes.

Au total, il y a donc pour l'ensemble des familles $\sum^F n(n - 1) = N$ couples de mesures. Les deux séries (x) et (y) sont identiques, bien que disposées différemment ; elles ont donc même moyenne et même variance.

La moyenne est

$$\bar{x} = \bar{y} = \frac{\sum^N x}{N} = \frac{\sum^F \sum^n (n - 1)x}{\sum^F n(n - 1)}$$

ou mieux :

$$\bar{x} = \frac{\sum^F [(n - 1) \sum^n x]}{\sum^F n(n - 1)}.$$

La variance est

$$\sigma_x^2 = \sigma_y^2 = \frac{\sum^N (x - \bar{x})^2}{N} = \frac{\sum^F \sum^n (n - 1)x^2 - \sum^F n(n - 1)\bar{x}^2}{\sum^F n(n - 1)}$$

ou mieux :

$$\sigma^2 = \frac{\sum^F [(n - 1) \sum^n x^2] - \sum^F n(n - 1)\bar{x}^2}{\sum^F n(n - 1)}.$$

Il reste à présent à calculer $\sum^N (x - \bar{x})(y - \bar{y}) = \sum^N xy - N\bar{x}\bar{y}$

$$\begin{aligned} \text{On a } \sum^N xy &= \sum^F [x_1(x_2 + x_3 + x_n) + x_2(x_1 + x_3 + x_n) + x_3(x_1 + x_2 + x_n) \\ &\quad + x_n(x_1 + x_2 + x_3)] \\ &= \sum^F \left[x_1 \left(\sum^n x - x_1 \right) + x_2 \left(\sum^n x - x_2 \right) + x_3 \left(\sum^n x - x_3 \right) \right. \\ &\quad \left. + x_n \left(\sum^n x - x_n \right) \right] \\ &= \sum^F \left[\sum^n x(x_1 + x_2 + x_3 + x_n) - (x_1^2 + x_2^2 + x_3^2 + x_n^2) \right] \\ \sum^N xy &= \sum^F \left[\left(\sum^n x \right)^2 - \sum^n x^2 \right]. \end{aligned}$$

Par ailleurs, on a évidemment $N\bar{x}^2 = \sum^F n(n-1)\bar{x}^2$.

En substituant les valeurs ainsi calculées dans la formule du coefficient de corrélation, il vient :

$$t = \frac{\sum^F \left[\left(\sum^n x \right)^2 - \sum^n x^2 \right] - \sum^F n(n-1)\bar{x}^2}{\sum^F [(n-1) \sum^n x^2] - \sum^F n(n-1)\bar{x}^2}$$

d'où la forme plus aisée à calculer :

$$t = \frac{\sum^F \left(\sum^n x \right)^2 - \left(\sum^F \sum^n x^2 + N\bar{x}^2 \right)}{\sum^F \sum^n nx^2 - \left(\sum^F \sum^n x^2 + N\bar{x}^2 \right)}.$$

On pourra vérifier que pour des familles d'effectif constant n , on retrouve la formule donnée par LAMOTTE (1948) :

$$t = \frac{n}{n-1} \cdot \frac{\sigma_f^2}{\sigma^2} - \frac{1}{n-1}$$

où σ^2 est la variance totale et σ_f^2 la variance des moyennes de familles.

D'une manière analogue, on trouverait pour deux caractères x et y la corrélation phénotypique intra-famille (cf. parag. sélection combinée de plusieurs caractères) :

$$t_{xy} = \frac{\sum^F \left(\sum^n x \cdot \sum^n y \right) - \left(\sum^F \sum^n xy + N\bar{x}\bar{y} \right)}{\sqrt{\left[\sum^F \sum^n nx^2 - \left(\sum^F \sum^n x^2 + N\bar{x}^2 \right) \right] \left[\sum^F \sum^n ny^2 - \left(\sum^F \sum^n y^2 + N\bar{y}^2 \right) \right]}}.$$

Reçu pour publication le 1^{er} mars 1958.

RÉFÉRENCES BIBLIOGRAPHIQUES

- BLISS (C. I.). — *Annals of applied Biology*, **22**, 134 et 207, 1935.
 BLISS (C. I.). — *Ann. of appl. Biol.*, **24**, 815, 1937.
 CHING CHUN LI. — An introduction to population genetics (Nat. Peking Univ. Press), 1948.
 COCHEZ (L. P.). — X^e Congrès Mondial d'Aviculture (Edinburgh), 1954.
 COCHEZ (L. P.). — Communications personnelles, 1956.
 GOGUELIN (P.). — Méthodes élémentaires de calcul statistique (Ed. Scient. Guyot), 1948.
 HAZEL (L. N.). — *Genetics*, **28**, 476, 1943.
 HAZEL (L. N.), LUSH (J. L.). — *J. Hered.*, **33**, 393, 1942.
 HAZEL (L. N.), BAKER (M. L.), REINMILLER (C. R.). — *J. Anim. Sci.*, **2**, 118.
 HUTIN (C.). — Éléments de calcul matriciel (non publié), 1957.
 JÉRÔME (F. N.), HENDERSON (C. R.), KING (S. C.). — *Poultry Sci.*, **35**, 995, 1956.
 KING (S. C.), HENDERSON (C. R.). — *Poultry Sci.*, **33**, 147, 1954.
 LAMOTTE (M.). — Introduction à la biologie quantitative (Masson et C^{ie}), 1948.
 LERNER (I. M.), ASMUNDSON (V. S.), CRUDEN (D. M.). — *Poultry Sci.*, **26**, 515, 1947.
 LERNER (I. M.). — Population genetics and animal improvement (Cambridge), 1950.
 LUSH (J. L.). — *Amer. Soc. Anim. Prod. proc.*, **33**, 293, 1940.
 LUSH (J. L.). — Animal Breeding Plans (Iowa State College), 1945.
 LUSH (J. L.). — *Am. Nat.*, **81**, 241 et 362, 1947.
 LUSH (J. L.). — The genetics of populations (non publié), 1948.
 LUSH (J. L.). — *Poultry Sci.*, **27**, 375, 1948.
 McCLOSKEY (J. F.) et TREFETHEN (F. N.). — Introduction à la recherche opérationnelle (Dunod), 1957.
 OLIVER (M. M.), BOHREN (B. B.), ANDERSON (V. L.). — *Poultry Sci.*, **36**, 395, 1957.
 QUINET (J.). — Cours élémentaires de mathématiques supérieures, t. IV, calcul intégral et applications (Dunod), 1956.
 ROBERTSON (A.). — *Biometrics*, **95**, 1955.
 ROBERTSON (A.), LERNER (I. M.). — *Genetics*, **34**, 395, 1949.
 SNEDECOR (G. W.). — Statistical methods 5^e ed. (Iowa State College), 1956.
 SMITH (C.). — Biomathematics (Charles Griffin and C^o), 1954.
 WADDINGTON (C. H.). — *Anim. Breed. Abst.*, **19**, 409, 1951.
 WHATLEY (J. A.). — *J. Agric. Res.*, **65**, 249, 1942.
 WINSOR (C. P.), CLARKE (G. L.). — *J. Marine Res.*, **3**, 1, 1940.
 WINTERS (L. M.). — *Emp. J. Exp. Agric.*, **8**, 259, 1940.
 WRIGHT (S.). — *Ann. math. statist.*, **5**, 161, 1934.
-