VERS UNE AMÉLIORATION APPORTÉE A LA SÉLECTION MASSALE DES MILS PAR LA MÉTHODE DES FICHES PERFORÉES

par

Y. DURAND

Chargé de l'amélioration des mils et sorghos.

La sélection des mils par la méthode dite des « fiches perforées » a été introduite au CRA Bambey par J. Corriols (1951). La méthode employée par ce dernier nécessitait des calculs statistiques longs et pénibles, rendant pratiquement impossible l'étude de plusieurs populations. Aussi, devant la nécessité de faire vite et de pouvoir rapidement vulgariser des populations améliorées pour le plus grand nombre possible de zones, L. Sauger (1952) simplifia considérablement la méthode et son exécution pratique en supprimant tous les calculs statistiques et en les remplaçant, pour l'interprétation, par de simples comparaisons de pourcentage.

Ces simplifications permirent de suivre et d'améliorer une trentaine de populations et les premiers résultats furent l'objet d'une publication (L. Sauger, M. Bono 1953).

L'application de cette méthode, dès 1952, nous a fourni une masse importante de données sur les caractères étudiés. En 1955 un dépouillement partiel de ces données sur quatre populations a permis à M. Bono, Y. Durand et J. Mousseau d'établir statistiquement un certain nombre d'associations existant entre ces caractères.

Cependant, depuis 1957, cette technique de sélection a été appliquée sur place à plusieurs populations locales non améliorées : au Sénégal, au Niger, en Haute-Volta, alors que l'on continuait toujours à suivre à Bambey une trentaine de populations. Il devint nécessaire, devant cet accroissement de travail, de rechercher un dépouillement plus rapide des observations. Il nous a, de plus, semblé utile, au commencement de cette nouvelle phase de sélection, d'augmenter la précision de liaisons établies et d'étudier de manière plus approfondie la valeur de chacune d'entre elles.

Note de la Rédaction. — Les articles publiés dans *L'Agronomie Tropicale*, quelle que soit la personnalité ou la fonction de leur auteur, n'expriment qu'une opinion personnelle et ne sauraient être considérés comme une indication de la politique ou des intentions du Département.

I) LE DÉPOUILLEMENT DES OBSERVATIONS

Le dépouillement des observations a pour but de permettre l'établissement des tableaux de contingence, à partir desquels les liaisons seront calculées.

Jusqu'en 1956 les observations étaient portées sur les fiches cartonnées représentant le plan de la parcelle de sélection. Il était prévu une fiche par état de caractère. Ainsi, pour la couleur de la gaine, on disposait de deux fiches : la première sur laquelle étaient perforés les emplacements correspondant aux pieds à gaine verte et la seconde, où étaient représentés les pieds à gaine rouge. L'établissement du tableau de contingence couleur de la gaine et productivité se réalisait en superposant très exactement chaque fiche de couleur de la gaine sur chaque fiche de productivité et en dénombrant les trous apparaissant par transparence.

Ce procédé appelle deux critiques :

Il est long et fastidieux, d'autant que l'expérience a montré qu'il n'était pas commode, de faire directement les observations sur les fiches destinées à la perforation. Celles-ci doivent donc être reportées.

Le risque d'erreurs était double : erreurs au moment du report des fiches d'observations sur les fiches cartonnées, erreurs au moment du comptage des trous apparaissant par transparence.

Il était certes possible de détecter la plupart des erreurs en superposant les différentes fiches d'un même caractère, ainsi que cela est expliqué par L. Saugier et M. Bono (1953) et en totalisant les fréquences des différentes classes. Mais ces contrôles augmentaient le travail.

Dans le système actuellement employé, les observations sont portées sur des fiches mécanographiques : chaque fiche représente un pied. Le report des fiches d'observations sur fiches mécanographiques se fait directement au centre mécanographique, où les risques d'erreur sont extrêmement réduits. Les tableaux de contingence des caractères sont alors dressés par voie mécanographique.

II) CALCUL DES ASSOCIATIONS

Afin de préciser les associations et, en particulier, préciser la probabilité de l'indépendance de deux caractères, nous avons recours au calcul classique du test de l'indépendance, grâce à l'emploi du critérium χ^2 de Pearson.

Pour simplifier les calculs les tableaux de contingence sont réduits par groupements des données à des tableaux 2 × 2 du type suivant :

		Premier caractère		Totaux
		Etat A	Etat B	
Deuxième caractère	Etat A	a	b	a+b
	Etat B	c	d	c+d
Totaux		a+c	b+d	ab+cd

On arrive facilement à montrer que le χ^2 peut être exprimé sous la forme :

$$\chi^2 = y^2 \left(\frac{1}{a-y} + \frac{1}{b+y} + \frac{1}{c+y} + \frac{1}{d-y} \right) \qquad (1) \qquad \text{où } y = \frac{ad - bc}{a+b+c+d}$$

Cette formule est très simple à calculer avec les machines. Dans certains cas le calcul peut se réduire au calcul de y.

On admet, dans la discussion qui va suivre, que « y » est positif c'est-à-dire que les paramètres a, b, c et d ont été choisis de manière à rendre positive la différence $ad - bc$.

Le nombre de degrés de liberté du test d'indépendance d'un tableau de contingence 2×2 est $n = 1$. Les valeurs correspondantes à un degré de liberté sont respectivement aux points 5 % et 1 % : 3,841 et 6,635.

La formule (1) se réduit à une somme de quatre rapports. Pour que le χ^2 soit significatif pour une probabilité donnée, il suffit que le plus grand de ces quatre rapports soit supérieur ou au moins égal à la valeur du χ^2 correspondant à cette probabilité ; de même, pour que le χ^2 ne soit pas significatif à une probabilité donnée, il suffit que le plus grand rapport soit inférieur ou au plus égal au quart de la valeur du χ^2 correspondant à cette probabilité.

Deux cas peuvent alors se présenter :

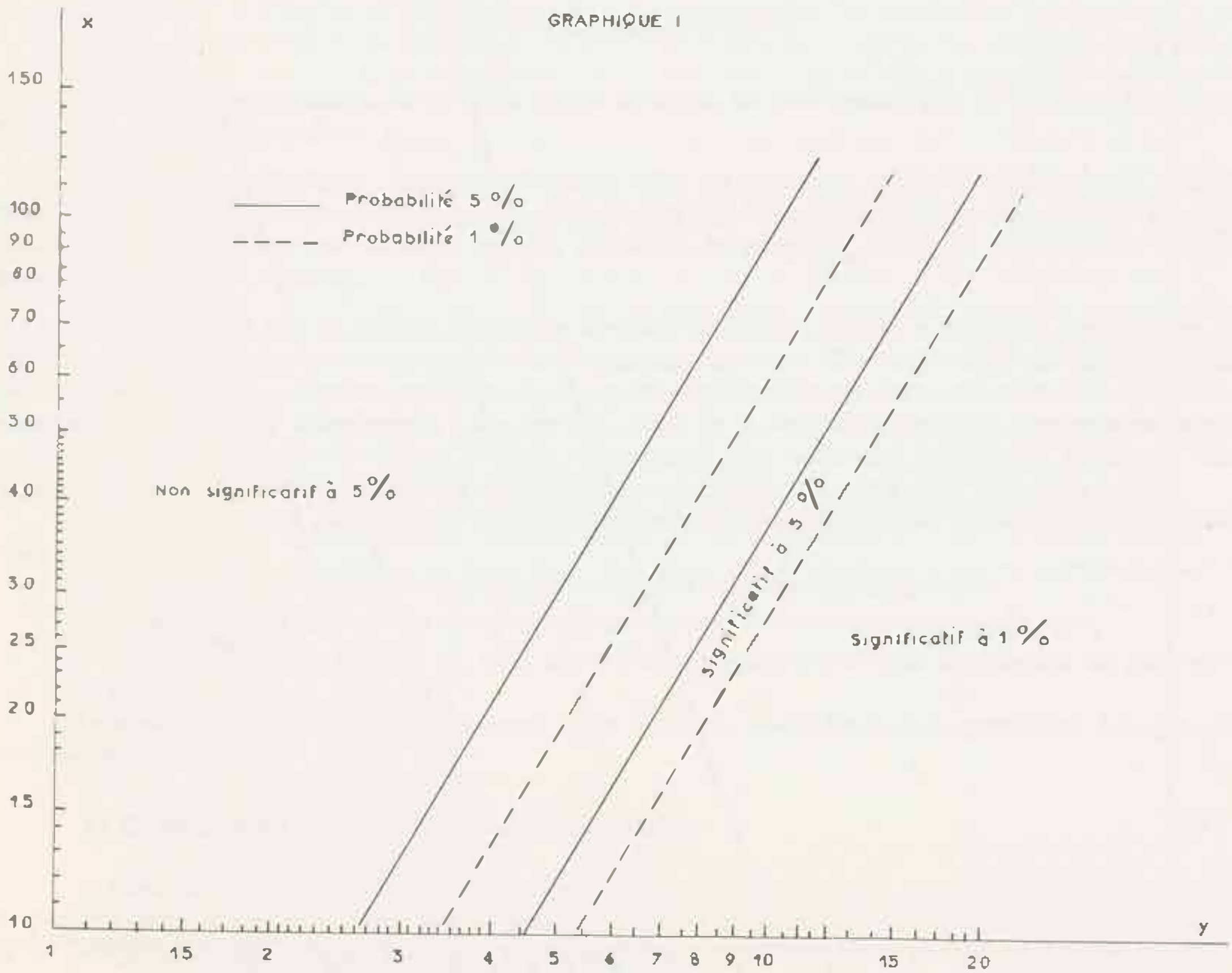
1) Le plus grand rapport est de la forme $\frac{y^2}{a - y}$.

Pour que le χ^2 soit significatif au point 5 %, il suffit que $\frac{y^2}{a - y} \geqslant 3{,}841$ (2).

Pour que le χ^2 soit significatif au point 1 %, il suffit que $\frac{y^2}{a - y} \geqslant 6{,}635$ (3).

Pour que le χ^2 ne soit pas significatif au point 5 %, il suffit que $\frac{y^2}{a - y} \leqslant 3{,}841/4$ ou 0,960 (4).

Pour que le χ^2 ne soit pas significatif au point 1 %, il suffit que $\frac{y^2}{a - y} \leqslant 6{,}635/4$ ou 1,659.



Après transformation logarithmique, les inégalités deviennent :

$$\log (a + 0{,}960) \leqslant 2 \log (y + 1{,}920) - \log (3{,}841) \quad (6)$$
$$\log (a + 1{,}659) \leqslant 2 \log (y + 3{,}317) - \log (6{,}635) \quad (7)$$
$$\log (a + 0{,}240) \geqslant 2 \log (y + 0{,}480) - \log (0{,}960) \quad (8)$$
$$\log (a + 0{,}415) \geqslant 2 \log (y + 0{,}829) - \log (1{,}659) \quad (9)$$

Le graphique I, aux coordonnées logarithmiques, représente les droites données par les inégalités (6) et (8). Lorsqu'un point (y, a) se trouve situé au-dessous de ces deux droites, on peut être sûr que le χ^2 est significatif au point 5 % ; lorsqu'un point (y, a) se trouve au-dessus de ces deux droites on peut être sûr que le χ^2 n'est pas significatif au point 5 % ; enfin lorsque ce point se trouve situé entre les deux droites on ne peut se prononcer sur la signification et il faut poursuivre les calculs jusqu'au bout.

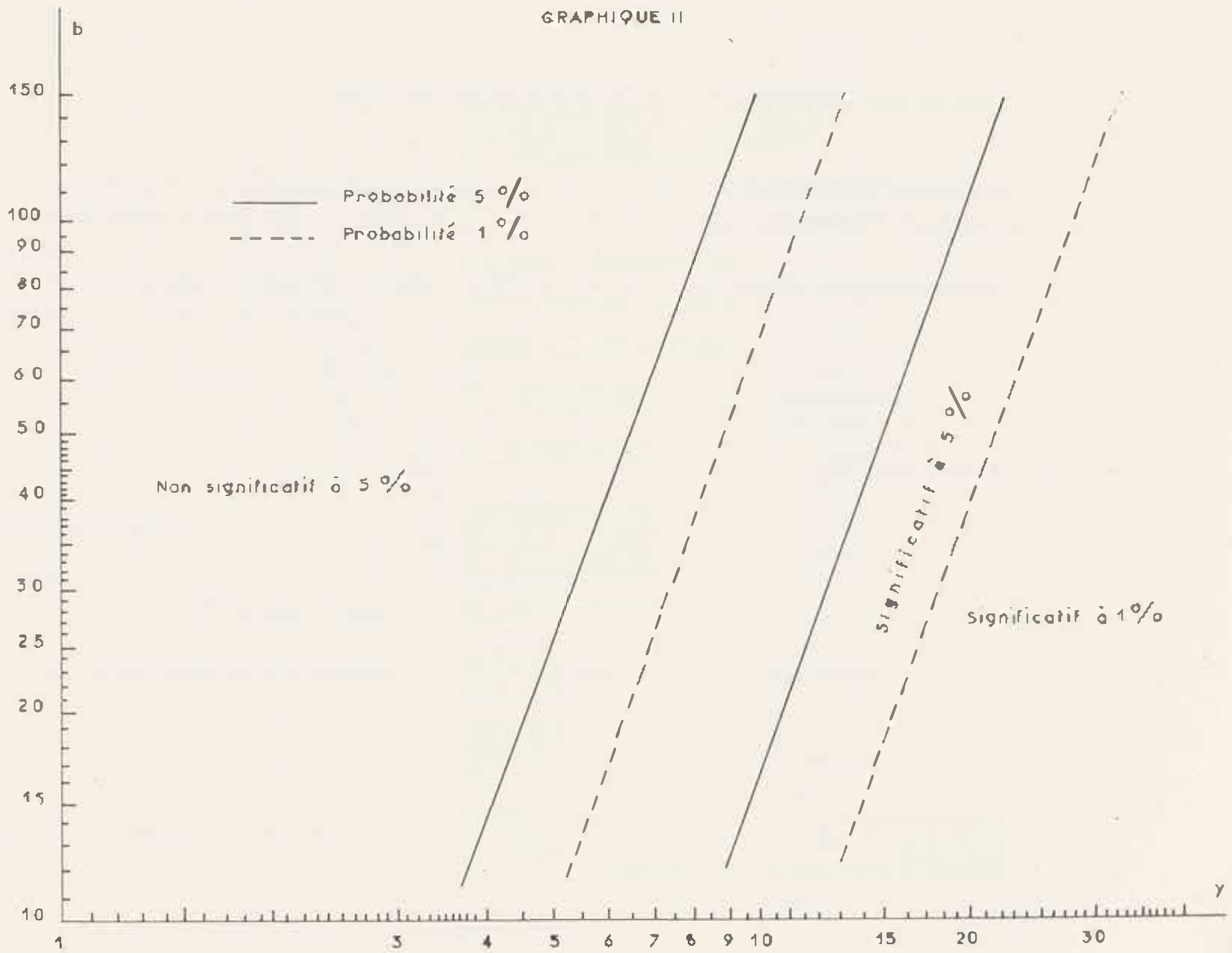
Le graphique I représente également les droites données par les inégalités (7) et (9) ; suivant qu'un point se trouve au-dessous, au-dessus ou entre ces deux droites, on peut être sûr qu'au point 1 % le χ^2 est significatif, ou bien ne l'est pas ou enfin ou ne peut se prononcer sur sa signification sans poursuivre les calculs jusqu'au bout.

2) Le plus grand rapport est de la forme $\dfrac{y^2}{b + y}$.

On arrive de la même manière aux inégalités suivantes :

$$\log (b + 0{,}960) \leqslant 2 \log (y - 1{,}920) - \log (3{,}841) \quad (10)$$
$$\log (b + 1{,}659) \leqslant 2 \log (y - 3{,}317) - \log (6{,}635) \quad (11)$$

GRAPHIQUE II



test de signification du χ^2 aux points 5 % et 1 %.

$$\log (b + 0{,}415) \geqslant 2 \log (y - 0{,}480) - \log (0{,}960) \qquad (12)$$

$$\log (b + 0{,}415) \geqslant 2 \log (y - 0{,}829) - \log (1{,}659) \qquad (13)$$

tests de non-signification aux points 5 % et 1 %.

Le graphique II représente les inégalités (10), (12) et (11), (13) ; ils permettent, de même que les précédents, l'interprétation graphique du χ^2 dans un grand nombre de cas.

Nous avons vu que, pour effectuer le calcul des associations, on commençait par grouper les données de manière à les réduire à des tableaux de contingence 2 × 2. Ce groupement se réalise sans difficulté sur simple examen du nombre de pieds existant dans chaque classe. Lorsque l'on étudie les associations entre la productivité et les autres caractères, on arrive ainsi à déterminer deux types d'associations : association d'un phénotype d'un caractère donné avec la forte production et association d'un phénotype d'un caractère donné avec la faible production. Le premier type d'association est appelé « association positive », alors que le deuxième est appelé « association négative ». Dans l'établissement du complexe des caractères liés à la forte production nous tenons également compte de ces deux types d'association en incluant dans le complexe, pour les associations négatives, le ou les phénotypes non associés à la faible production. La sélection se fait donc non seulement en recherchant les caractères liés aux critères de sélection, comme cela était pratiqué jusqu'auparavant, mais également en rejetant les caractères liés aux contraires des critères de sélection.

III) INTERPRÉTATION ET UTILISATION DES LIAISONS MISES EN ÉVIDENCE

Le calcul des associations, tel qu'il est exposé ci-dessus, ne nous permet pas de connaître leur nature. Celles-ci peuvent être, en effet :

a) **De nature génétique** : ce sont évidemment celles qui présentent le plus d'intérêt pour le sélectionneur. Si les conditions d'expérimentation sont suffisamment précisées, celles-ci doivent se retrouver toutes les années. Notons cependant que certaines peuvent apparaître en cours d'amélioration, la population évoluant sous l'influence de la sélection.

b) **Provoquées par le milieu** : Parmi ces dernières un certain nombre sont fugaces, apparaissent une année pour disparaître l'année suivante ou parfois réapparaître inversées. Elles ne font que refléter les conditions particulières de l'expérimentation (année et emplacement). Leur intérêt est donc nul, voire même négatif car elles risquent d'amener le sélectionneur à éliminer une partie de la variabilité génétique pouvant s'avérer intéressante dans la suite de l'amélioration. D'autres, au contraire plus constantes, reflètent les conditions écologiques moyennes de la zone pour laquelle on travaille ; ces dernières présentent donc un certain intérêt.

Il est donc essentiel, pour mener à bien cette amélioration, de différencier ces associations et de contrôler leur valeur.

Il est nécessaire enfin de conserver toujours une grande variabilité génétique afin :

a) de ne pas éliminer des complexes géniques pouvant être intéressants par la suite,

b) de maintenir toujours un taux élevé d'hétérozygotie permettant l'expression du phénomène de l'hétérosis.

Le nombre de pieds choisis chaque année pour assurer la descendance de la population doit donc être le plus élevé possible.

1) Conduite de la sélection proprement dite.

a) Constitution de la population : La population sur laquelle devra porter « la sélection par fiches perforées » devra avoir dès l'origine la plus grande variabilité génétique. Lorsque l'on entreprend la sélection d'une variété locale déjà adaptée à la zone pour laquelle on travaille, on doit s'efforcer de réunir des semences provenant du plus grand nombre possible de localités de cette zone.

b) Première année: Après l'établissement des associations de différents caractères étudiés avec les critères de sélection (le plus souvent rendement), on distingue :

Les associations dont la probabilité du χ^2 est inférieure à 1 %. Les caractères ainsi déterminés sont appelés caractères obligatoires ; ils devront obligatoirement être pris en considération dans le choix des pieds assurant la descendance de la population étudiée.

Les associations dont la probabilité du χ^2 est comprise entre 5 et 1 %. Les caractères ainsi déterminés sont dits caractères accessoires ; ils seront pris en considération pour le choix des pieds-mères, uniquement si le nombre de ces derniers ne s'en trouve pas trop réduit.

c) Deuxième année : Les autres caractères ne sont pas considérés associés aux critères de sélection ; les associations de la population cultivée lors de la deuxième année de l'amélioration sont calculées indépendamment de celles apparues lors de la première année.

Les caractères associés deux années de suite (probabilité de l'indépendance inférieure chaque année à 5 %) sont considérés comme caractères obligatoires, donc utilisés dans le choix des pieds-mères.

Les autres caractères associés seulement une année aux critères de sélection sont des caractères accessoires et ne sont pris en considération que dans la mesure du possible. Sont d'abord utilisés les caractères associés la deuxième année, ensuite ceux associés la première année.

d) Troisième année : Les caractères trouvés associés pendant trois années de suite aux critères de sélection peuvent être considérés comme faisant définitivement partie du complexe de la population étudiée dans la région où cette étude est menée.

Les caractères trouvés associés au rendement la deuxième et la troisième année sont considérés comme caractères obligatoires.

Les autres caractères associés seulement pendant deux années (soit la première et la troisième, soit la première et la deuxième) et ceux associés seulement la troisième année sont des caractères accessoires.

On ne tient aucun compte des caractères trouvés associés soit la première année, soit la seconde, et non les deux autres.

On arrêtera la recherche du complexe à partir du moment où celui-ci sera définitivement établi et ne subira plus aucune modification.

Nous arrivons donc à définir les trois types de caractères suivants :

Les caractères du complexe : Caractères associés aux critères de sélection pendant au moins trois années de suite et inclus définitivement dans le complexe.

Caractères obligatoires : Caractères inclus provisoirement dans le complexe, associés la première année aux critères de sélection avec une probabilité d'indépendance inférieure à 1 % ou pendant deux années consécutives.

Caractères accessoires : Caractères pouvant être inclus provisoirement dans le complexe (dans la mesure du possible), associés la première année avec une probabilité d'indépendance comprise entre 5 % et 1 % ou, associés l'année en cours d'étude du complexe et non les années précédentes ou encore associés les deux années précédentes et non l'année en cours d'étude.

On espère ainsi éliminer les liaisons apparaissant uniquement sous l'influence des conditions particulières et fluctuantes du milieu.

On doit cependant signaler que, si la sélection est toujours pratiquée au même emplacement, les associations apparaissant sous l'influence des conditions écologiques bien particulières de cet emplacement ne peuvent être éliminées. On a donc intérêt à placer chaque année la sélection dans des localités différentes de la zone pour laquelle on travaille.

2) Contrôle de la sélection :

Ces sélections sont faites dans le but de vulgariser des populations améliorées. On doit donc, à chacune des phases de la sélection, s'assurer de manière précise des améliorations apportées. Pour ce faire, il ne saurait être suffisant de comparer les productions de la parcelle de sélection des différentes années. Il faut, après chaque

année de sélection, comparer entre elles, les différentes populations obtenues et les comparer aux populations d'origine, à l'aide d'un essai statistique. Cet essai mettra en évidence les progrès obtenus chaque année et indiquera le moment où la sélection n'apporte plus aucune amélioration.

D'autre part, des essais du type multilocal doivent permettre, avant la vulgarisation, de vérifier l'adaptation de la population améliorée aux différentes zones écologiques de la région à laquelle on la destine.

CONCLUSION ET RÉSUMÉ. — *Par suite de l'augmentation du nombre des populations étudiées par la méthode dite des « fiches perforées », un certain nombre de modifications ont été apportées dans le double but de faciliter l'interprétation et de préciser les résultats obtenus.*

1) *Les fiches quadrillées, représentant le plan de la parcelle (six cents pieds) et affectées chacune à un état de caractère ou phénotype, sont remplacées par des fiches mécanographiques représentant chacune un pied et où toutes les observations effectuées sur ce pied sont notées. Les tableaux de contingence sont obtenus par voie mécanographique.*

2) *L'établissement des associations entre caractères se fait par calcul du χ^2 des tableaux de contingence réduits à des tableaux 2 × 2 par groupement des données. Une méthode graphique est indiquée qui permet dans certains cas de trouver la probabilité du χ^2 sans avoir à le calculer. Nous définissons des « associations positives », associations, où les caractères sont liés aux critères de sélection, et des « associations négatives » où les caractères sont liés aux contraires des critères de sélection.*

3) *Dans l'interprétation des associations, nous arrivons à distinguer trois types de caractères :*

les caractères du complexe ;
les caractères obligatoires ;
les caractères accessoires.

4) *Un contrôle réalisé au moyen d'essais comparatifs détermine l'amélioration obtenue chaque année par la sélection dans la zone d'adaptation de la population améliorée.*

SUMMARY. — *Owing to the increase in number of populations studied according to the so called « punched index-cards » method, some changes have been made in order to facilitate interpretation and obtain more accurate details.*

1) *The squared cards, representing the plot (six hundred plants), each of them being affected to some characteristic or phenotype, are replaced by multicopied cards, each representing a plant on which all observations relating to that particular plant are noted down.*

Tables of contingence are obtained through multicopying.

2) *Associations of characteristics are established by calculation of the χ^2 of tables of contingence reduced to 2 × 2 tables, through grouping of data.*

A graphical method is indicated, allowing in certain cases to find the probable χ^2 without having to calculate it.

« Positive associations » are defined as associations in which characteristics are linked with criterions of selection, and « negative associations » as those in which the characteristics are linked with the contraries of the criterions of selection.

In interpreting the associations we distinguish 3 types of characteristics :

the characteristics of the complex,
the essential characteristics
the accessory characteristics.

A control, under the form of comparative tests establishes the improvement obtained every year through selection in the area of adaptation of the improved population.

RESUMEN. — *Dado el aumento del número de las populaciones estudiadas con el método de las « fichas perforadas » hiciéronse algunas modificaciones para facilitar la interpretación y precisar los resultados obtenidos.*

1) *Las fichas cuadriculadas reprsentando el plano de la parcela (seiscientas plantas) y afectadas cada una a*

un estado de carácter o fenotipo, se substituyen con fichas mecanográficas representando una planta cada una, con todas las observaciones efectuadas.

Los cuadros de contingencia se obtienen mecanográficamente.

2) *Para establecer las asociaciones entre caracteres evalúase el χ^2 de los cuadros de contingencia reducidos a cuadros 2 × 2 por agrupación de los dados. Indícase un método gráfico permitiendo en varios casos hallar la probabilidad del χ^2 sin tener que calcularlo. Para las « asociaciones positivas » damos la definición de asociaciones en que los caracteres se relacionan con los criterios de selección, mientras en las « asociaciones negativas » los caracteres se relacionan con los contrarios de los criterios de selección.*

3) *En la interpretación de los resultados, distinguimos tres caracteres :*

los caracteres del complejo ;

los caracteres obligatorios ;

los caracteres accesorios.

4) *Un control, por medio de ensayos comparativos, determina la mejora obtenida cada año con la selección en la zona donde se adapta la población mejorada.*

BIBLIOGRAPHIE

J. Corrtols. — Essai de sélection du mil par une méthode statistique. Annales du CRA de Bambey, 1952, p. 101-19.

L. Sauger, M. Bono. — Sélection rapide des plantes allogames. Un essai de sélection généalogique sans autopollinisation artificielle sur mil *Pennisetum*. *L'Agronomie Tropicale*, Vol. VIII, nº 6, 1953, p. 614-37.

M. Bono, Y. Durand, Y. Mousseau. — Contribution à l'étude des populations sélectionnées de mil (*Pennisetum typhoides*). Annales du CRA de Bambey, 1955, p. 49-57.