



**HAL**  
open science

## L'indexation pour le classement

Anne Tavernier

► **To cite this version:**

Anne Tavernier. L'indexation pour le classement. Productions Animales, 1992, hs (hs), pp.209-211.  
hal-00896019

**HAL Id: hal-00896019**

**<https://hal.science/hal-00896019>**

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.

Anne TAVERNIER

Service des haras, des courses et de l'équitation  
14 avenue de la Grande Armée 75017 Paris

INRA Station de Génétique quantitative  
et appliquée 78352 Jouy-en-Josas Cedex

## L'évaluation des reproducteurs

# L'indexation pour le classement

**Résumé.** Pour étudier la génétique du cheval de course ou de concours le premier problème qui se pose est une mesure objective de leur performance. Celle-ci n'est pas mesurable directement par un critère physique objectif. On en est donc réduit à n'étudier que le résultat des épreuves constitué par les classements des chevaux entre eux. Chaque cheval réalise une performance physique non mesurable et c'est la hiérarchie de ces performances qui donnent la seule observation possible : le classement. Selon la valeur génétique des chevaux et la valeur des effets de milieu chaque classement a une certaine probabilité de réalisation. En maximisant les probabilités de réalisation de toutes les épreuves on obtient une estimation des valeurs génétiques et des effets de milieu.

Tourné vers le sport, l'élevage équin a besoin d'un cheval athlète. Mais un athlète en quoi ? Qu'est ce que "réussir une carrière de concours hippique ou de course" ? Avant même de discuter de l'alchimie savante qui combine les caractères physiques et mentaux du cheval de sport, on peut se demander ce qui définit sa performance sportive. Et cette définition n'est pas simple. Déterminer objectivement le niveau technique d'un concours de saut d'obstacles tient de la gageure. Il dépend, bien sûr, de la hauteur et de la largeur des obstacles (encore sont-ils nombreux et différents), mais surtout de l'agencement de ces obstacles sur le terrain, de la distance et du tracé entre deux obstacles, de la combinaison d'obstacles aux profils divers, de la vitesse minimale avec laquelle on doit franchir ces obstacles, de la nature du terrain de chaque abord, et même des détails dans la construction de chaque obstacle : couleur, forme... Hiérarchiser les résultats d'un cheval (barres tombées, refus, temps réalisé) sur deux parcours différents devient alors impossible car il faudrait combiner la détermination du niveau technique de chaque parcours et l'écart à ce niveau d'après ces résultats (une barre tombée vaut combien de quelle mesure du niveau technique de l'épreuve ?). De même, déterminer le niveau d'une course peut paraître simple au porteur de chronomètre, mais ce qui fait vraiment la difficulté de l'épreuve, c'est toute la stratégie développée autour d'un train donné : les accélérations, les ralentissements, les interactions entre les comportements des chevaux du peloton... (l'estimation de l'hérédité des temps en course de galop est toujours faible voir nulle : Hintz 1980, Langlois 1980a). Le problème n'est donc pas d'abord de savoir avec quels moyens le cheval réussit, ou quel est le déterminisme génétique de ces moyens mais bien qu'est ce que réussir ? Et il n'y a pas de mesure "physique" de cette réussite.

## 1 / Comment mesurer la réussite ?

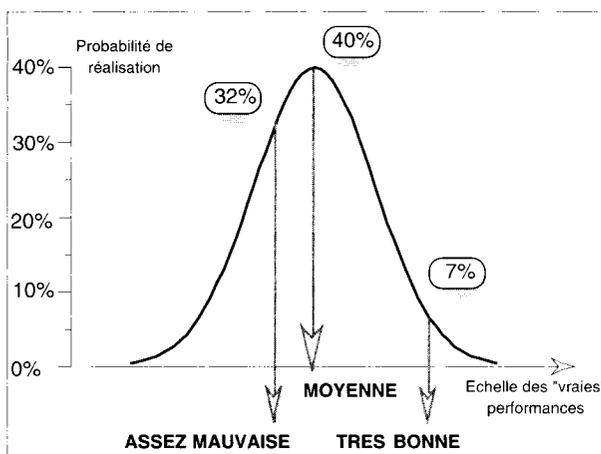
Beaucoup de critères ont été envisagés (Arnason *et al* 1989). Les chevaux qui participent à une même épreuve sont ordonnés entre eux. Un pourcentage donné des premiers (variant éventuellement avec le nombre de partants) reçoit une certaine somme d'argent selon des règles parfaitement définies. On dit alors que le cheval est "classé" dans l'épreuve. Les critères les plus simples sont basés sur le rapport du nombre de classements sur le nombre de départs. Cela revient à attribuer un "1" à un cheval classé et un "0" à un cheval non classé. Il faudrait donc déjà analyser cette variable comme une variable discrète (Gianola et Foulley 1983) et non comme une proportion mais surtout ce critère met sur un plan d'égalité le cheval champion du monde et le cheval classé une fois dans une épreuve du plus petit niveau ! Les autres critères sont basés sur les gains attribués à chaque place ou sur un nombre de points qui repose sur le même principe. Ces critères posent de nombreux problèmes : il n'y a pas de gain ou de points pour les chevaux "non classés" ce qui les élimine du modèle alors qu'ils ont réalisé dans l'épreuve une (contre) performance. Au pire, on leur attribue un "zéro" ce qui les met à distance arbitraire et égale quel que soit le niveau de l'épreuve des autres chevaux classés (sans compter que ce "zéro" a une influence très différente suivant l'unité de mesure). La distribution des gains en fonction de la place suit en général une exponentielle, des transformations (Langlois 1980b) ont été utilisées pour se ramener à une fonction linéaire. Mais une fonction linéaire de la place ne tient pas compte du nombre de concurrents dans l'épreuve et n'élimine pas le problème de la mesure du niveau de l'épreuve.

## 2 / “Que le meilleur gagne”

Tous les critères précédents reposent sur l'attribution d'une "note" au classement (ou à son absence) et d'une "note" au niveau de l'épreuve. Seulement, la vraie performance du cheval c'est uniquement d'avoir été meilleur, dans une épreuve donnée, que les autres chevaux de la même épreuve. Et ensuite, c'est d'avoir été moins bon dans une autre épreuve avec d'autres concurrents, et que ces concurrents eux-mêmes aient été vainqueurs ou vaincus dans d'autres épreuves... Ce sont ces multiples comparaisons qui peuvent nous fournir une hiérarchie entre les chevaux.

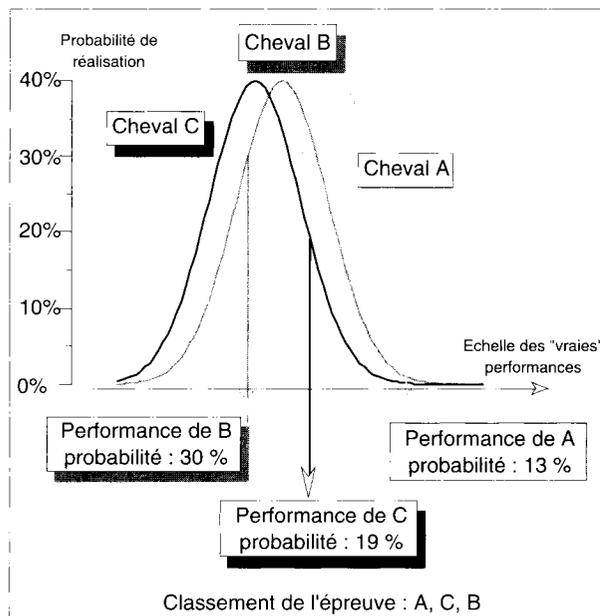
Le cheval réalise dans une épreuve une "vraie" performance physique que nous ne pouvons mesurer directement comme nous venons de le voir. Cette performance, suivant le déterminisme génétique sous-jacent et le déterminisme des effets de milieu qui lui sont associés, suit une certaine loi statistique. Pour simplifier, prenons comme loi de probabilité des performances une loi normale. Autour d'une performance "moyenne" que le cheval est capable de réaliser, et qui dépend de sa valeur génétique ainsi que des effets environnementaux, il y a d'autres performances que le cheval a plus ou moins de chance de réaliser (figure 1).

**Figure 1** - Probabilité de réalisation des différentes performances d'un cheval.



Au cours d'une épreuve chaque cheval réalise une performance, qui correspond à une certaine probabilité d'accomplissement, le classement qui en résulte dépend des positions relatives des performances de tous les chevaux de l'épreuve (Henery 1981). C'est ce qui est représenté figure 2 avec une épreuve à 3 chevaux : A, B, C. Les chevaux A et C ont réalisé une performance supérieure à leur performance moyenne. La performance de A avait 13% de chance de se réaliser, celle de C 19%. Le cheval B a réalisé une performance légèrement inférieure à sa performance moyenne, elle avait 30% de chance de se réaliser. Le résultat de l'épreuve est que la performance de A est supérieure à celle de C qui est supérieure à celle de B, le classement est donc A puis C puis B. C'est la seule information à laquelle on ait réellement accès. On remarquera que, dans cet exemple, le classement obtenu n'est pas le même que la hiérarchie entre le niveau de performance moyen de chaque cheval qui est A meilleur que B meilleur que C. La probabilité d'une telle épreuve est de  $0,13 \times 0,19 \times 0,30 = 0,7\%$ . Pour obtenir le même classement, les chevaux pou-

**Figure 2** - Exemple d'une épreuve à 3 chevaux - Explication du classement des chevaux par les lois de probabilité de réalisation des performances de chaque cheval.



vaient réaliser d'autres performances, tout en conservant la même hiérarchie. La somme des probabilités de chaque combinaison de performances qui donne le même classement donne la probabilité du classement. Ainsi on peut montrer que, dans notre exemple, compte tenu de l'écart moyen entre A, B et C, les probabilités de classement sont celles reportées au tableau 1.

**Tableau 1.** Exemple d'une épreuve à 3 chevaux : probabilité des différents classements - L'écart entre l'espérance des performances de A et de B est de 1,2 ; entre celles de B et de C il est de 0,6 avec 1 point représentant l'écart-type de la loi de densité des performances.

Classement	Probabilité
A-B-C	49%
<b>A-C-B</b>	<b>27%</b>
B-A-C	15%
B-C-A	3%
C-A-B	5%
C-B-A	2%

## 3 / Estimation du “meilleur”

Dans la pratique, on ne connaît pas l'espérance moyenne des performances d'un cheval, c'est justement ce qu'on cherche à estimer et qui dépend de sa valeur génétique et des effets de milieu identifiés. Pour estimer ces effets, on va chercher à maximiser leur vraisemblance, ce qui revient à maximiser le produit des probabilités de chaque course tout en tenant compte des relations de parenté entre les chevaux et de la variabilité génétique de notre caractère (Tavernier 1991). Une fois déterminé ce qu'est réellement le résultat d'une épreuve : le classement, et une fois déterminée sa loi, l'estimation des valeurs

génétiques ne présente aucune originalité, c'est le même raisonnement que quand on peut mesurer directement une performance qui suit une loi normale. La seule complication intervient dans la résolution du système, qui n'est pas linéaire, et dans l'importance des calculs à effectuer (intégrales multiples pour le calcul des probabilités).

Actuellement les chevaux de concours hippique ont été estimés sur 5 années (de 1985 à 1989) ce qui représentait environ 40 000 chevaux, 1 300 000 départs dans 33 000 épreuves et 37 000 000 de comparaisons des chevaux 2 à 2. L'héritabilité du caractère est de 0,16, la répétabilité est de 0,29 (Tavernier 1990). La méthode a aussi été testée sur les courses de trot. Chaque estimation est assortie d'une précision qui dépend bien sûr du nombre d'épreuves auxquelles le cheval a participé mais aussi du nombre de partants dans ces épreuves (plus il est élevé, plus le nombre de comparaisons est important et donc plus la précision est grande), de la place du cheval dans l'épreuve (quand un cheval est classé premier ou dernier on ne connaît que des chevaux moins bons ou meilleurs que lui mais pas les deux, on est donc moins précis), du niveau de la concurrence (si un cheval est battu par beaucoup plus fort que lui ou bat beaucoup plus faible cela n'apporte que peu de renseignements).

#### 4 / Etude de notre plan d'expérience : les programmes de course

Disposer d'une estimation de la valeur génétique et sportive des chevaux indépendamment de toute mesure subjective du niveau des courses ou des dotations nous permet de jeter un œil critique sur ce qu'est notre plan d'expérience : les programmes de courses et de concours. La corrélation entre le logarithme du gain annuel d'un cheval de concours hippique et l'estimation de ce cheval à partir des seuls classements est de 0,75. La corrélation entre la

moyenne des estimations à partir des seuls classements des partants d'une course de trot et le logarithme de la dotation de la course est de 0,58. Ces chiffres, heureusement positifs et assez forts laissent malgré tout soupçonner quelques écarts entre récompense et qualité du cheval. Notre estimation permet de détecter les anomalies et éventuellement de rectifier le programme. Mais surtout, elle permet d'identifier des lots homogènes. Ce qui fait le spectacle d'un programme de course, c'est aussi ce qui nous permet d'avoir l'estimation la plus précise : avoir un niveau de concurrents homogènes pour une même épreuve. Et le corollaire à cette proposition est qu'il existe des ponts entre les épreuves : c'est à dire que les chevaux puissent se confronter entre eux pour déterminer réellement leur niveau. Il ne faut pas d'isolats répétitifs. Et c'est pourquoi ce type de calcul plutôt que de servir à calculer a priori des résultats de course doit plutôt permettre de réaliser des courses dans lesquelles toutes les combinaisons de classement ont des probabilités les plus identiques possibles.

### Conclusion

Raisonnement uniquement sur la performance du cheval à partir de son classement dans de nombreuses épreuves sans aucun arbitraire et uniquement avec des comparaisons relatives permet de s'affranchir d'une grosse part de subjectivité dans la mesure de la qualité du cheval. Si le choix de ce critère semble optimum dans notre cas, où une mesure physique de la performance ne peut être déterminée, il n'a pour l'instant qu'un traitement des plus élémentaires au niveau génétique, même si celui-ci va être traité, dans les mois qui viennent, comme un modèle animal. Nous espérons maintenant pouvoir affiner notre modèle génétique. Cela sera d'autant plus complexe que nous allons avoir à raisonner sur la population des pur sang (chevaux de course de galop) qui est une population depuis longtemps sélectionnée, sujette aux accouplements raisonnés et sur laquelle circulent beaucoup d'histoires d'effets de complémentarité ou de consanguinité.

### Références bibliographiques

- Arnason T., Bendroth M., Philipsson J., Hendriksson K., Darenius A., 1989. Genetic evaluations of Swedish trotters. EAAP Publication, 42, 106-130.
- Gianola D., Foulley J.L., 1983. Sire evaluation for ordered categorical data with a threshold model. *Génét. Sél. Evol.*, 15, 201-224.
- Henery R.J., 1981. Permutation probabilities as models for horse races. *J. R. Statist. Soc.*, 43, 86-91.
- Hintz R.L., 1980. Genetics of performance in the horse. *J. Anim. Sci.*, 51, 582-594.
- Langlois B., 1980a. Heritability of racing ability in thoroughbreds. A review. *Livest. Prod. Sci.*, 7, 591-605.
- Langlois B., 1980b. Estimation de la valeur génétique des chevaux de sport d'après les sommes gagnées dans les compétitions équestres Françaises. *Ann. Génét. Sél. Anim.*, 12, 15-31.
- Tavernier A., 1990. Estimation of breeding value of jumping horses from their ranks. *Livest. Prod. Sci.*, 26, 277-290.
- Tavernier A., 1991. Genetic evaluation of horses based on ranks in competitions. *Genet. Sél. Evol.*, 23, 159-173.