

Évolution de l'utilisation des statistiques : quelques réflexions sur le rôle des modèles

NIGEL GILLES YOCCOZ

Les méthodes statistiques en écologie (et dans bien d'autres disciplines) sont souvent utilisées de façon discutable, et cela d'autant plus que les concepts de base sont mal connus. Les relations entre biologie et statistiques sont ici examinées de façon critique (surabondance de tests, écart entre signification statistique et signification biologique). Ces relations sont au centre de cet article, mais bien entendu à travers la nécessaire production de modèles que l'on sent toujours présents. Toutes les étapes de la démarche de la recherche scientifique sont concernées, ce qui multiplie l'intérêt général des réflexions présentées par l'auteur, valables pour la plupart des disciplines s'exprimant dans NSS.

NIGEL GILLES YOCCOZ¹

département d'écologie arctique
Institut norvégien de recherche
sur l'environnement
Centre sur les milieux polaires -
N-9296 Tromsø, Norvège
nigel.yoccoz@ninatos.ninaniku.no

*« I am suggesting that research is needed
into how statistics is applied »
(C. Mallows, 1998 : 8)*

Une perception négative des statistiques

R. A. Fisher, à qui les statistiques d'aujourd'hui doivent tant (voir Efron, 1998 pour une revue récente), compara en 1948 les développements de la biométrie à ceux de la géométrie chez les grecs pendant l'Antiquité (Fisher, 1948). Cela peut paraître provocateur, mais il est évident que les statistiques, prises au sens d'une démarche scientifique qui va de la planification des expériences, des observations ou des échantillonnages jusqu'à l'analyse des données et aux déductions qui en résultent, a pris une place considérable dans de nombreuses disciplines, en biologie, en médecine, ou en sciences sociales². Cette place fut traditionnellement prépondérante en agronomie, sans doute parce que Fisher lui-même, mais aussi tant d'autres – Neyman, Snedecor, G. Cox, Cochran, Kempthorne ... – y jouèrent un rôle de premier plan. Elle l'est devenue, plus récemment, en écologie ou en biologie évolutive, entre autres, et pour ne retenir que des disciplines qui me sont davantage familières. Cette « colonisation » par les statistiques a entraîné l'évolution des revues de biométrie (e.g., *Biometrics*) vers des contenus de plus en plus techniques ; mais, comme le rappelait fort à propos Billard en 1995, les lecteurs de *Biometrics* s'en plaignaient déjà il y a 50 ans ! Certains articles parus aujourd'hui dans des revues d'écologie ressemblent fort à des articles de statistiques d'il y a quelques dizaines d'années.

L'utilisation des statistiques ne peut pas être critiquée sur le fond : les statistiques sont un des outils

essentiels de la méthode scientifique, et Fisher a toujours mis en avant cet aspect. Utilisées à bon escient, elles contribuent de manière très significative aux différentes disciplines concernées. Bien sûr, leur utilisation peut aussi conduire à des catastrophes comme l'explosion de Challenger en 1986, qui fut en partie au moins le résultat de décisions prises sur la base d'analyses statistiques déficientes³. Mais entre succès et échecs se situe peut-être l'essentiel de l'utilisation des statistiques dans nos disciplines : une contrainte, parce qu'il faut publier, et parfois une sorte de religion (van Dantzig, 1957 ; Salsburg, 1985). On peut bien parler d'une perception négative des statistiques (Efron, 1998), « en aveugle » : tout ce qui n'est pas significatif au seuil magique de 5 % n'existe pas, par exemple.

Plus généralement, une lecture attentive d'articles publiés en écologie, épidémiologie ou médecine conduit à se demander si ne règne pas une méconnaissance des concepts de base des outils statistiques. Cette méconnaissance conduit à une vision biaisée de l'outil statistique, qui est rarement perçu comme servant à extraire ou organiser l'information, tout en

Abstract – Evolution in the use of statistics: some considerations on the role played by models

Statistics have become in the 20th century one of the main tools of scientific work in many disciplines. Despite their extensive and increasing use, it has been clear for many years that the main advantages and limitations of many statistical tools have not been well understood. This is the case, for example, of levels of variability and of statistical tests. We examine what role statistical models have played in recent changes in the use of statistics, and suggest that integration of statistical models more closely with scientific theories and hypotheses should be aimed at. © 1999 Éditions scientifiques et médicales Elsevier SAS

statistics / tests / models

¹ Enseignant depuis plus de dix ans l'utilisation des méthodes statistiques, d'abord à l'université de Lyon et maintenant à l'université d'Oslo, mais aussi impliqué directement dans la planification d'études ou de suivi en écologie. La perception négative, ou plus souvent l'incompréhension vis-à-vis de cet outil, ont conduit l'auteur à en chercher les raisons possibles.

évitant bien sûr certaines interprétations abusives (en donnant une mesure de l'incertitude associée à nos conclusions). Une telle observation est relativement fréquente dans la littérature statistique, et nombreux sont les statisticiens qui ont essayé d'en comprendre les raisons, et d'y remédier (en reconnaissant clairement la part de responsabilité des statisticiens ; voir Nelder, 1986 ; Altman et Bland, 1991). Le fait que l'évaluation de la carrière des statisticiens relève souvent plus du contenu mathématique de leurs recherches que de leurs contributions aux domaines d'application est bien entendu une des sources des difficultés (p.e. Mallows, 1998 ; Senn, 1998). Certains statisticiens privilégieront davantage leur contribution au domaine théorique, et moins celle aux domaines d'application. Ce problème n'est pas propre aux statistiques. Si le problème scientifique posé dans une discipline conduit à des recherches innovantes dans le domaine mathématique, tant mieux ; mais c'est le plus souvent la contribution au domaine d'application qui sera jugée⁴. Comme le faisait remarquer Healy, 1978, les statistiques sont en partie une technologie, mais en cela il représente d'abord une certaine école anglo-saxonne, « à la Fisher » (voir le récent débat dans *The Statistician*, 1998 : p.e. Bailey, 1998 ; Hand, 1998a). Fisher était en effet très critique des développements théoriques réalisés dans un objectif de pure rigueur mathématique ou d'optimalité (Efron, 1998).

Les statistiques ou comment organiser l'information : les problèmes d'échelle

Un exemple où les statistiques contribuent à la fois en terme de démarche et d'analyse est celui des problèmes d'échelle et de hiérarchie. Ces derniers occupent une place de plus en plus importante dans de nombreuses disciplines. Par exemple, en écologie, les analyses de la biodiversité prennent en compte la variabilité entre individus (par exemple leur diversité génétique), entre populations et entre écosystèmes. Cette diversité aura également des structures différentes suivant l'échelle spatiale retenue, par exemple parce que les différents types de perturbation affectant un écosystème (feu, abrutissement, inondation, ...) ont leur propre échelle spatiale et temporelle : à chacun de ces niveaux, ou à chaque échelle correspond une variabilité ayant des origines spécifiques, mais qui peuvent interagir entre elles. Analyser et estimer ces variabilités constitue justement un des principaux objectifs des modèles statistiques. L'incertitude associée à nos conclusions dérivera aussi de la variabilité des phénomènes étudiés. Par exemple et schématiquement, l'estimation de la variabilité à un niveau servira à connaître l'incertitude sur un paramètre décrivant le niveau immédiatement supérieur : la variabilité entre individus permet de comparer des populations. La confusion des niveaux de variabilité a conduit Hurlbert (1984) à introduire le terme de « pseudoréplication ». Il s'agit en fait le plus souvent de répliques, mais à une échelle qui n'est pas pertinente

(ou qui nécessite des hypothèses supplémentaires, et le plus souvent difficiles à justifier) pour estimer la variabilité, et donc l'incertitude associée aux estimations des paramètres. Cox (Reid, 1994) fait remarquer que comprendre le plan d'expérience dit *split-plot* (figure 1), où deux niveaux de réplique sont présents, est un point de départ indispensable à une bonne compréhension de ces difficultés. Le fait que l'analyse de la variabilité est un point essentiel des statistiques font de cette discipline un outil de choix dans des disciplines où elle apparaît sous des formes complexes.

L'apport des modèles en statistiques

Nous assistons donc à des changements qualitatifs – développement de méthodes de plus en plus complexes – et quantitatifs – peu d'articles sont publiés sans un grand nombre de tests ou d'analyses statistiques. Parallèlement, les applications des statistiques ont vu l'importance des modèles s'accroître de manière considérable. Ceci n'est pas nouveau : les modèles étaient déjà essentiels dans l'approche de Fisher ou Neyman au cours de la première moitié de ce siècle (lire, ou relire, les articles fondamentaux de Fisher, 1922 ou Fisher, 1925 ; voir Lehmann, 1990 pour une discussion). Cette importance se traduit dans la terminologie : la régression ou l'analyse de variance deviennent des cas particuliers du modèle linéaire, et la régression logistique est un exemple de modèle linéaire généralisé (Nelder et Wedderburn, 1972 ;

² C'est d'ailleurs dans cette discipline que D.R. Cox voit une source importante de développements futurs en Statistiques (e.g. Reid, 1994 ; Cox et Wermuth, 1996) – à l'image de l'agronomie en 1920. Les sciences sociales furent déjà, au XIX^e siècle, la principale source d'inspiration des statistiques. On peut noter également la complémentarité entre les développements des méthodes statistiques associées aux plans d'expérience – d'une certaine manière le paradigme de l'agronomie après les années 1920 et l'influence de R.A. Fisher – et les développements de méthodes spécifiques aux données multivariées et observationnelles en sciences sociales et en écologie dans les vingt dernières années, en partie parce qu'il est souvent difficile d'expérimenter au sens strict dans ces disciplines. Voir par exemple, Arminger et al. (1995) ; Lebreton et Asselain (1993).

³ Les défauts des tuiles recouvrant la navette sont en effet apparus lors du lancement effectué à basse température de l'air, alors que les données existantes indiquant que ces tuiles pouvaient présenter un danger avaient été ignorées et les valeurs obtenues à température normale (15–25 °C) simplement extrapolées sans justification (Tufté, 1997)

⁴ Il est cependant utile de rappeler que cette contribution peut survenir bien longtemps après les développements théoriques, et rendre donc difficile une évaluation à court terme. Il serait évidemment néfaste d'empêcher tous développements théoriques n'ayant pas d'applications immédiates !

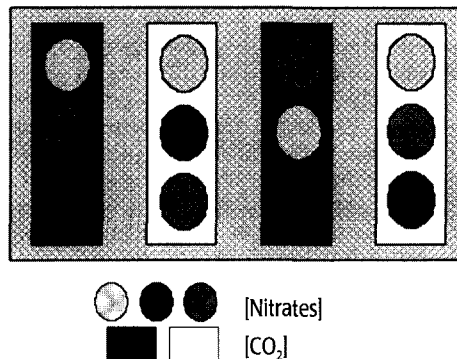


Figure 1. Le plan dit *split-plot* est le modèle des plans d'expérience où les répliques sont faites à deux niveaux. Ce plan d'expérience, qui a bien entendu son équivalent dans un cadre observationnel, est souvent utilisé quand les facteurs étudiés sont manipulables, ou observés, à des échelles différentes. Un cas classique est l'étude de l'influence de différents taux de CO₂ et d'azote (e.g. nitrates) sur la croissance des plantes. Les chambres où le taux de CO₂ est modifié sont de taille importante et contiennent plusieurs plantes. Par contre, les nitrates peuvent facilement être manipulés à l'échelle de chaque plante. Un premier niveau – les taux de nitrates – aura ici 12 répliques qui serviront à estimer l'incertitude sur les effets des concentrations de nitrates. Un second niveau – les différents niveaux de CO₂ – ne possède que quatre répliques. La confusion (pseudoréplication) consiste à utiliser les 12 répliques pour estimer la variabilité servant aux comparaisons des niveaux de CO₂. Il suffit de remplacer CO₂ et nitrates par populations et individus, ou villes et hôpitaux, pour comprendre que ce problème se rencontre dans toutes les disciplines.

ARTICLE

⁵ Les variables aléatoires définies par le modèle permettent de calculer la probabilité d'observer un échantillon quelconque – une réalisation – parmi tous ceux possibles, les valeurs des paramètres étant fixées. Le problème inverse, celui d'inférer sur les paramètres à partir d'une réalisation unique, a conduit Fisher à introduire le terme de vraisemblance (*likelihood*) afin de distinguer ce qui reflète d'une part la variabilité d'un phénomène de ce qui est notre ignorance vis-à-vis d'un paramètre (qui a priori ne varie pas mais est inconnu). Il faut également rappeler que certains modèles peuvent incorporer des effets aléatoires, qui permettent de décrire par exemple comment un effet varie aléatoirement entre individus d'une population, et l'approche bayésienne, qui suppose que notre ignorance concernant la valeur des paramètres peut en fait être décrite par une distribution de probabilités. Cette dernière approche est de plus en plus préférée pour certains problèmes dont la solution par les approches classiques est difficile, comme celui des erreurs de mesure (p.e. Richardson and Leblond 1997).

⁶ La randomisation – une autre contribution majeure de Fisher – consiste à attribuer au hasard (par exemple, en utilisant un générateur de nombres aléatoires) aux différents individus/populations... les différents traitements d'une expérience. L'introduction volontaire de l'aléatoire dans l'expérience permet de limiter considérablement les hypothèses nécessaires quant à l'interprétation des données. En particulier, sous des hypothèses générales, la randomisation garantira l'absence de biais dans l'estimation des effets, le problème majeur des études dites observationnelles (sans randomisation). Voir par exemple Kish (1987) ou Hinkelman et Kempthorne (1994).

McCullagh et Nelder, 1989). On peut citer aussi les modèles pour l'analyse des données de capture-recapture (Lebreton et al., 1992) ou le modèle de Cox et ses extensions pour l'analyse des données de survie – un classique en médecine s'il en est (Cox et Oakes, 1984). Cette tendance se justifie fondamentalement par l'unification de méthodes dans le cadre d'une théorie mathématique générale (Bailey, 1998), conduisant également à la simplification des procédures de calcul et donc des programmes. Le terme de modèle apparaît donc omniprésent dans l'utilisation des statistiques, et il faut donc s'interroger sur un éventuel lien avec les relations difficiles entre statistiques et utilisateurs. Ne présentons-nous pas les modèles sous une forme par trop statistique, en négligeant leur justification biologique (ou médicale, ou sociologique...) et en quoi ils permettent (ou ne permettent pas d'ailleurs) de répondre à nos questions ?

Mais il importe d'abord de préciser ce que sont en général les modèles en statistiques. La notion de modèle y est souvent associée à celle d'un modèle probabiliste où les observations sont supposées être les réalisations⁵ de variables aléatoires (Cox, 1990). Le modèle représente alors une description, à l'aide de paramètres et d'hypothèses sur le type de distribution de ces variables aléatoires, d'une population théorique dont est tiré l'échantillon (l'exemple traditionnel étant une population suivant une distribution normale de moyenne μ et de variance σ^2). Ces modèles distingueront souvent les paramètres décrivant la variation systématique (par exemple la moyenne) de ceux décrivant la variabilité (par exemple la variance). De manière souvent parallèle, seront aussi distingués des paramètres servant à spécifier les effets intéressants, et d'autres ne servant qu'à spécifier entièrement la population (souvent appelés paramètres de nuisance). Ces modèles peuvent ne pas être complètement paramétriques (le modèle de Cox en étant un excellent exemple), mais ils font référence à une population hypothétique. Cependant, une partie importante des statistiques, celle liée aux problèmes d'échantillonnage, a au contraire cherché à s'affranchir de modèles décrivant une population théorique, en développant des procédures d'estimation ne dépendant que du plan d'échantillonnage (conduisant à des *design-based inference*, par opposition à *model-based inference* ; Kish, 1987). Plus précisément, la procédure d'échantillonnage conduit sans faire référence à une population à un modèle probabiliste dont la formulation ne dépend que du plan d'échantillonnage utilisé. La randomisation⁶ dans les plans d'expérience joue le même rôle que la notion d'échantillon aléatoire dans les plans d'échantillonnage, permettant ainsi de s'affranchir entre autres des hypothèses d'indépendance et de normalité (Fienberg et Tanur, 1987). Il me semble indispensable de souligner que seules ces deux situations – randomisation et échantillonnage aléatoire – fournissent une justification directe, ou « physique », d'un modèle probabiliste de l'analyse statistique (la seule composante aléatoire du modèle dérivant de la procédure de randomisation). Nulle référence à une population théorique n'est nécessaire. Cette base probabiliste apparaît beaucoup plus difficile à justifier

pour des études observationnelles, où il apparaît nécessaire de spécifier cette population dont sont tirées les données (Mallows, 1998) : quel modèle probabiliste par exemple peut décrire la relation entre taux de déforestation et effectif de la population pour les pays d'Amérique Latine ? (voir Laurance, 1998). De manière plus générale, quelles données relèvent du problème posé (les pays d'Afrique ou d'Asie peuvent-ils être « comparés » à ceux d'Amérique Latine) et sur quelle population portera alors nos inférences statistiques (Mallows, 1998) ? La cadre inférentiel, défini par le modèle, ne fait pas, le plus souvent, l'objet de discussion critique.

Les modèles statistiques répondent de plus à des exigences pour le moins variées, et auxquelles ils répondront rarement simultanément (voir Legay, 1997 pour une discussion générale) :

1. le modèle doit représenter un lien avec des connaissances ou des théories existantes ;
2. le modèle doit permettre une comparaison avec des études précédentes ;
3. le modèle doit suggérer ou être cohérent avec un processus pouvant avoir généré les données ;
4. les paramètres définissant les caractéristiques essentielles du modèle doivent posséder une interprétation claire dans le domaine d'étude ;
5. la variabilité ou « erreur » doit être représentée avec assez de réalisme pour conduire à des mesures de précision satisfaisantes pour les paramètres importants du modèle ;
6. l'ajustement aux données doit être satisfaisant (Cox, 1997).

Transmettre l'information grâce aux modèles

Tous ces points soulèvent des difficultés, et mériteraient d'être discutés un à un, mais j'aimerais ici insister sur le deuxième point (lié au quatrième) : les modèles simples sont un moyen efficace de transmission de l'information contenue dans une étude, et donc facilitent les généralisations. Ce seront tout particulièrement les paramètres de ces modèles et une mesure de la variabilité résiduelle, souvent à plusieurs niveaux pour prendre en compte les notions d'échelle et de hiérarchie, qui véhiculeront cette information. Ainsi, des modèles non-paramétriques, comme les méthodes de régression locale, sont difficilement comparables d'une étude à l'autre. Ils jouent cependant un rôle essentiel dans l'exploration des données, et donc dans la justification de l'emploi d'un modèle approché, plus simple. En effet ces modèles non-paramétriques ne se prêtent que difficilement à une traduction en termes quantitatifs. C'est seulement en publiant des estimations des effets les plus importants biologiquement que l'auteur d'une étude peut s'assurer de son utilisation future. Les modèles et leurs paramètres devraient donc être un outil de choix dans la présentation des résultats – dans la transmission des connaissances – en particulier aux dépens d'autres « accessoires » comme les tests qui encombrant nos revues (*encadré*).

L'intérêt de modèles relativement simples sera d'autant plus grand que les paramètres ont une interprétation direct dans un cadre quantitatif (évitant donc la dichotomie simpliste absence ou présence d'effet), éventuellement théorique. Cela facilite considérablement leur enseignement car ils ont une prise directe sur les problèmes biologiques. Par exemple, les modèles classiques utilisés pour l'analyse de séries temporelles – de type auto-régressifs ou moyennes mobiles – peuvent être « traduits » en terme de modèles de dynamiques de populations, p.e. de compétition entre stades chez des insectes ou prédateur-proie à la Lotka-Volterra pour le lynx et le lièvre (Royama, 1992 ; Bjørnstad et al., 1998 ; Stenseth et al., 1998). Bien sûr, il est aisé de critiquer ces modèles pour leur caractère simpliste, mais on peut s'attendre à une rapide complexification de ces modèles, complexité en partie fonction des données disponibles. On peut remarquer d'ailleurs que des modèles tels que celui proposé par Bjørnstad et al., 1998 sont déjà assez délicats à manipuler – et il ne s'agit que d'un système isolé en laboratoire. Par ailleurs, des modèles simples seront adéquats si des termes plus complexes – de type interactions en analyse de variance – ne représentent quantitativement qu'une part relativement faible de la structure sous-jacente⁷. Ici encore, le point n'est pas l'aspect « statistiquement significatif », qui le devient nécessairement si les données sont en nombre suffisant, mais « biologiquement significatif », une question qui fait appel au contexte biologique et à laquelle il est souvent délicat, mais essentiel, de répondre. Un effet peut par exemple être techniquement significatif au seuil de 5 %, mais négligeable biologiquement en comparaison d'autres

Résumé – Évolution de l'utilisation des statistiques : quelques réflexions sur le rôle des modèles

Les statistiques sont un outil essentiel des disciplines scientifiques, et l'utilisation des modèles est un des aspects dominants de la pratique des statistiques. Pourtant, ce recours aux modèles ne traduit pas une maîtrise des concepts, et les avantages ou inconvénients de cette approche ne sont souvent pas bien compris. Après avoir décrit quelques exemples, je suggère qu'une meilleure intégration des modèles statistiques et des théories ou hypothèses du domaine d'étude peut conduire à une meilleure synergie. © 1999 Éditions scientifiques et médicales Elsevier SAS

statistiques / tests / modèles

effets (la situation inverse étant probablement plus fréquente). Remarquons que les calculs de puissance de tests pour différents plans d'expérience ou d'observation (équivalent à celui de la recherche d'une précision suffisante de l'estimation d'un paramètre) posent de manière explicite le problème des valeurs considérées comme biologiquement significatives (Steidl et al., 1997, Gerard et al., 1998). Par là-même, la question de savoir ce qui est statistiquement significatif devient la recherche de ce qui est biologiquement important. Sans modèles ayant une interprétation dans le domaine d'application cette discussion est impossible. Une réforme dans nos principes de construction et d'utilisation des modèles en statistiques est sans doute nécessaire. Mais elle sera d'autant plus justifiée qu'elle conduira à une meilleure intégration entre statistiques et biologie. C'est la biologie qui doit davantage nous guider dans le choix

⁷ La notion de simplicité est toute relative : il ne s'agit pas ici de vouloir réduire le système à un petit nombre de composantes (voir Ducrot et al., 1996, Legay, 1997 pour une critique de cette approche), mais d'avoir un modèle qui aide à la compréhension du système – par opposition à un modèle de type simulation (Legay, 1997) où l'objectif sera d'abord de prédire. Un modèle sera simple s'il représente des effets sous une forme linéaire (éventuellement après changement d'échelle) et additive (en particulier sans interaction). Cela ne veut pas dire que des interactions entre plusieurs facteurs ne sont pas présentes, mais que nous avons de grandes difficultés à les manipuler intellectuellement, et les négliger ne modifie pas fondamentalement le système. De même, l'existence de relations non-linéaires ne fait aucun doute, en dynamique de populations par exemple, mais approcher ces relations non-linéaires par des droites par morceaux en facilite l'interprétation – et l'analyse statistique (Grenfell et al., 1998 ; Stenseth et al., 1998).

Quel regard sur les modèles ? Tests, estimations, sélection et incertitude sur les modèles

Environ 8 000 tests ont été publiés en 1997 dans la seule revue *Ecology*, dont l'immense majorité portent sur des hypothèses dont on peut dire a priori – sans aucun test – qu'elles sont fausses (D. Anderson et K. Burnham, comm. Pers.). Ce problème a été depuis longtemps dénoncé par les statisticiens (voir Yoccoz, 1991 pour des références). Cet usage conduit à une perte d'information considérable, ne serait-ce que parce que les tests ne sont pas comparables entre études, alors qu'en général les estimations le sont (bien entendu sous condition que les sources de biais soient comparables). Une difficulté supplémentaire dans certains domaines d'application des statistiques (par exemple dans le domaine économique, où le terme de 'data-mining' est souvent employé) provient maintenant d'une « surabondance »

de données qui rend tout test significatif, et donc évidemment inutile (Hand, 1998b). Cependant, même pour de petits échantillons, je ne peux que condamner la prédominance actuelle des tests, aux dépens tout particulièrement de l'estimation. Néanmoins, toute estimation reste conditionnelle à un, ou plusieurs, modèles (Madigan et Raftery, 1994 ; Buckland et al., 1997 ; Burnham et Anderson, 1998), et se pose donc le problème de la sélection des modèles servant aux inférences, et de l'incertitude associée à cette sélection. La plupart des procédures de sélection existantes répondent à des objectifs soit de prédiction, soit d'estimation, et se sont multipliées ces dernières années – mais n'utilisent en général pas les tests. La plus connue, due à Akaike, a ses racines dans la physique statistique et les travaux de Boltzmann à la fin du

siècle dernier (Akaike, 1985) : le choix d'un modèle se fait alors en terme de ses qualités prédictives – non plus en termes d'estimation de paramètres – qualités jugées par l'écart entre la distribution attendue sous un modèle donné et le « vrai » modèle, inconnu. Cette distinction correspond à différents objectifs dans l'utilisation d'un modèle (Legay, 1997), et différents objectifs peuvent conduire à différents critères. Ces procédures de sélection ne sont pas sans leurs propres difficultés, et ne doivent absolument pas remplacer les tests dans une illusoire dichotomie vrai/faux modèle, ou bon/mauvais modèle (Legay, 1997). Il est clair que les problèmes de sélection ou de mesure d'incertitude associé au choix d'un modèle représente un des problèmes majeurs et non encore pleinement résolus des statistiques.

de la structure des modèles – et les statistiques pourront alors bien apparaître comme un des outils essentiels de la démarche scientifique conduisant à une meilleure estimation des composantes des modèles et de l'incertitude associée à nos conclusions, et donc finalement au progrès de nos connaissances.

Remerciements

Je tiens tout particulièrement à remercier J.-M. Legay pour avoir inspiré et soutenu une réflexion constante sur les statistiques et plus généralement sur les outils que nous utilisons. R. Julliard, D. Pontier et deux rapporteurs anonymes ont fourni des critiques constructives de cet article.

RÉFÉRENCES

- Akaike H. 1985. *Prediction and entropy*. Pages 1–24 in Atkinson A. C. and Fienberg S. E. (eds.). *A celebration of statistics*. Springer-Verlag, Berlin.
- Allman D. G., Bland J. M. 1991. Improving doctors' understanding of statistics. *J. Roy. Stat. Soc. A* 154, 223–267.
- Arminger G., Clogg C. C., Sobel M. E. (ed.). 1995. *Handbook of statistical modeling for the social and behavioral sciences*. Plenum Press, New York.
- Bailey R. A. 1998. Statistics and mathematics: the appropriate use of mathematics within statistics. *The Statistician* 47, 261–271.
- Billard L. 1995. The roads travelled. *Biometrics* 51, 1–12.
- Bjørnstad O. N., Begon M., Stenseth N. C., Falck W., Sait S. M., Thompson D. J. 1998. Population dynamics of the Indian meal moth: Demographic stochasticity and delayed regulatory mechanisms. *J. Anim. Ecol.* 67, 110–126.
- Buckland S. T., Burnham K. P., Augustin N. H. 1997. Model selection: An integral part of inference. *Biometrics* 53, 603–618.
- Burnham K. P., Anderson D. R. 1998. *Model selection and inference. A practical information theoretic approach*. Springer-Verlag, New York.
- Cox D. R. 1990. Rôle of models in statistical analysis. *Statistical Sci.* 5, 169–174.
- Cox D. R. 1997. The current position of statistics: a personal view. *Int. Statistical Rev.* 65, 261–290.
- Cox D. R., Oakes D. 1984. *Analysis of survival data*. Chapman and Hall, London.
- Cox D. R., Wermuth N. 1996. *Multivariate dependencies*. Chapman and Hall, London.
- Ducrot C., Legay J.-M., Gröhn Y. T., Enevoldsen C., Calavas D. 1996. Approach to complexity in veterinary epidemiology: example of cattle reproduction. *Nature Sciences Sociétés* 4, 23–34.
- Efron B. 1998. R.A. Fisher in the 21st century. *Statistical Sci.* 13, 95–122.
- Fienberg S. E., Tanur J. M. 1987. Experimental and sampling structures: parallels diverging and meeting. *Int. Stat. Rev.* 55, 75–96.
- Fisher R. A. 1922. On the mathematical foundations of theoretical statistics. *Philos. Trans. R. Soc. Lond. A Math. Phys. Sci.* 222, 309–368.
- Fisher R. A. 1925. Theory of statistical estimation. *Proc. Camb. Philos. Soc.* 22, 700–725.
- Fisher R. A. 1948. Biometry. *Biometrics* 4, 217–219.
- Gerard P. D., Smith D. R., Weerakkody G. 1998. Limits of retrospective power analysis. *J. Wildlife Manag.* 62, 801–807.
- Grenfell, B. T., Wilson K., Finkenstädt B. F., Coulson T. N., Murray S., Albon S. D., Pemberton J. M., Clutton-Brock T. H., Crawley M. J. 1998. Noise and determinism in synchronized sheep dynamics. *Nature* 394, 674–677.
- Hand D. J. 1998a. Breaking misconceptions - statistics and its relationship to mathematics. *The Statistician* 47, 245–250.
- Hand D. J. 1998b. Data mining: Statistics and more? *The American Statistician* 52, 112–118.
- Healy M. J. R. 1978. Is statistics a science? *J. R. Stat. Soc. A* 141, 385–393.
- Hinkelmann K., Kempthorne O. 1994. *Design and analysis of experiments. Vol. 1, introduction to experimental design*. John Wiley and Sons, New York, É.-U.
- Hurlbert S.H. 1984. Pseudoreplication and the design of ecological field experiments. *Ecol. Monogr.* 54, 187–211.
- Kish L. 1987. *Statistical design for research*. John Wiley and Sons, New York, É.-U.
- Laurance W. F. 1998. A crisis in the making: responses of Amazonian forests to land use and climate change. *Trends Ecol. Evol.* 13, 411–415.
- Lebreton J.-D., Asselain B. (ed.). 1993. *Biométrie et environnement*. Masson, Paris.
- Lebreton J.-D., Burnham K. P., Clobert J., Anderson D. R. 1992. Modeling survival and testing biological hypotheses using marked animals: a unified approach with case studies. *Ecol. Monogr.* 62, 67–118.
- Legay J.-M. 1997. *L'expérience et le modèle. Un discours sur la méthode*. Inra Éditions, Paris.
- Lehmann E. L. 1990. Model specification: the views of Fisher and Neyman, and later developments. *Stat. Sci.* 5, 160–168.
- Madigan D., Raftery A. E. 1994. Model selection and accounting for model uncertainty in graphical models using Occam's window. *J. Am. Stat. Assoc.* 89, 1535–1546.
- Mallows C. 1998. The zeroth problem. *The American Statistician* 52, 1–9.
- McCullagh P. et Nelder J. A. 1989. *Generalized linear models*. Chapman and Hall, Londres, R.-U.
- Nelder J. A. 1986. Statistics, science and technology. *J. R. Stat. Soc. A* 149, 109–121.
- Nelder J. A., Wedderburn R. W. M. 1972. Generalized linear models. *J. R. Stat. Soc. A* 135, 370–384.
- Reid N. 1994. A conversation with Sir David Cox. *Stat. Sci.* 9, 439–455.
- Richardson S., Leblond L. 1997. Some comments on misspecification of priors in Bayesian modelling of measurement error problems. *Stat. Med.* 16, 203–213.
- Royama T. 1992. *Analytical population dynamics*. Chapman and Hall, Londres, R.-U.
- Salsburg D. S. 1985. The religion of statistics as practiced in medical journals. *The American Statistician* 39, 220–223.
- Senn S. 1998. Mathematics: governess or handmaiden? *The Statistician* 47, 251–259.
- Steidl R. J., Hayes J. P., Schaubert E. 1997. Statistical power analysis in wildlife research. *J. Wildlife Manag.* 61, 270–279.
- Stenseth N. C., Falck W., Chan K.-S., Bjørnstad O. N., O'Donoghue M., Tong H., Boonstra R., Boutin S., Krebs C. J., Yoccoz N. G. 1998. From patterns to processes: phase and density dependencies in the Canadian lynx cycle. *Proc. Nat. Acad. Sci. USA* 95, 15430–15435.
- Tufte E. R. 1997. *Visual explanations*. Graphic Press, Cheshire, Conn., É.-U.
- Van Dantzig D. 1957. Statistical priesthood (1 and 2). *Stat. Neerl.* 11, 1–16 et 185–200.
- Yoccoz N. G. 1991. Use, overuse and misuse of significance tests in evolutionary biology and ecology. *Bull. Ecol. Soc. Am.*, 106–111.