

CAMILLE DUBY

**Discussion et commentaires. La planification
des expériences : choix des traitements et
dispositif expérimental**

Journal de la société française de statistique, tome 141, n° 1-2 (2000),
p. 43-45

http://www.numdam.org/item?id=JSFS_2000__141_1-2_43_0

© Société française de statistique, 2000, tous droits réservés.

L'accès aux archives de la revue « Journal de la société française de statistique » (<http://publications-sfds.math.cnrs.fr/index.php/J-SFdS>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques

<http://www.numdam.org/>

DISCUSSION ET COMMENTAIRES

La planification des expériences : choix des traitements et dispositif expérimental

Camille DUBY¹

Pierre Dagnelie fait un travail utile en rappelant certains dangers de l'inférence statistique à partir de résultats de plans d'expérience et en mettant l'accent sur l'importance du dispositif expérimental. Ses illustrations des problèmes posés par l'estimation de la variabilité résiduelle à partir d'un faible nombre d'observations dans des modèles linéaires classiques sont aussi très efficaces.

Cela étant, je ne pense pas qu'il soit opportun de faire une dichotomie, comme il le suggère mais sans avoir l'air de trop y croire, entre applications agronomiques et applications industrielles quand on veut mettre l'accent sur la variabilité résiduelle. En effet, lorsque l'on éprouve le besoin de faire des plans d'expériences, c'est que l'on a affaire à un phénomène complexe, et ce, quel que soit le domaine d'application. Lors d'une expérimentation, ce phénomène complexe mettra en jeu des facteurs connus non contrôlés (sauf à faire des expérimentations très lourdes) et des facteurs inconnus (et donc généralement non contrôlés) qui augmenteront la variabilité. En revanche, la réponse à la question posée et discutée par Box ([1966], [1988]) : « a-t-on besoin de faire des tests de signification ? » amène celui-ci à définir une dichotomie résultant de deux attitudes a priori différentes : scientifique (scientific feedback, on a alors besoin de faire ces tests) ou empirique (*empirical feedback*, on n'en a pas besoin), qui me paraît plus utile.

Dans ce qui suit, je me placerai dans la situation d'une réponse positive à la question posée par Box. Il me semble alors que l'on peut élargir la problématique de P. Dagnelie dans cet article, en rappelant que cette variabilité résiduelle, qui servira de référence par exemple pour les tests de signification des effets, possède trois composantes plus ou moins indépendantes :

- Erreurs de mesures
- Effets des facteurs incontrôlés
- Erreurs de modèle

Ces composantes sont modélisées par une variable aléatoire d'espérance nulle s'il n'y a pas d'erreur de modèle et de variance σ^2 . S'il y a une erreur de modèle cette variable aléatoire ne peut plus être considérée d'espérance nulle.

1. Institut National d'Agronomie, 16 rue Claude Bernard, 75231 Paris Cedex 05; e-mail : duby@inapg.inra.fr

Certaines méthodes sont proposées (Daniel [1959], Kobilinsky [1997]), qui permettent de déterminer les effets significatifs sans avoir d'estimation de cette variabilité résiduelle. Des tests basés sur l'hypothèse qu'aucun des effets n'est significatif permettent de déterminer ceux qui le sont. Ces méthodes sont certes intéressantes, mais le sont essentiellement dans les cas de plans orthogonaux, avec un nombre important d'effets estimés et en supposant qu'il n'y a pas d'erreur de modèle. Malheureusement, par les mêmes glissements que P . Dagnelie dénonce, ces méthodes sont aussi utilisées dans des cas où il y a un nombre relativement faible d'effets estimés, avec peut-être des erreurs de modèle, ce qui aboutit souvent à ne sélectionner que des effets tellement importants qu'on n'a quasiment pas besoin de faire de tests pour les reconnaître.

Si l'on admet la nécessité de faire des tests d'hypothèse mais aussi de donner des intervalles de confiance (des effets, des prédictions, etc.), se pose alors le problème de l'estimation de la variabilité résiduelle en analysant suivant les cas la signification exacte de cette estimation. Ainsi, lorsque l'on utilise des plans d'expérience pour faire l'étude de sensibilité de modèles déterministes complexes, la seule composante de la variabilité résiduelle est l'erreur éventuelle de modèle, puisque toute simulation à partir des mêmes entrées du modèle donnera le même résultat. Il n'y aura pas dans cette situation de notion de signification statistique au sens habituel du terme. Les probabilités critiques des tests faits avec cette résiduelle serviront essentiellement à faire un classement des effets des différents facteurs et de leurs interactions, mais ne voudront rien dire en elles-mêmes.

Par ailleurs, dans certaines publications à destination d'utilisateurs non statisticiens de plans d'expérience (ainsi dans Sado et Sado [1991], pp. 54 et 238, mais ce n'est pas le seul exemple), il est préconisé parfois, lorsque l'expérience ne permet pas d'avoir une estimation de σ^2 (0 degrés de liberté), d'utiliser comme estimation de cette variance l'erreur de mesure sur la variable réponse, erreur connue *a priori* indépendamment de l'expérience analysée. Pourquoi pas ? Mais il faut être conscient que cette estimation de σ^2 sous-estimera la variabilité résiduelle puisqu'elle ne tient pas compte des effets incontrôlés, et cela sans parler des erreurs de modèle. En définitive les résultats ainsi obtenus ne pourront être extrapolés à des situations semblables à celles de l'expérience.

Quant aux erreurs de modèle leur effet pourra être assez pervers. Au mieux elles induiront un biais dans l'estimation de σ^2 . Pour diminuer ces erreurs, la définition des domaines de variation des facteurs étudiés est fondamentale, puisque, au bout du compte, les modèles utilisés seront des approximations de la « vraie fonction de réponse », et que ces approximations seront d'autant meilleures que les domaines seront bien choisis (position et étendue du domaine). C'est ainsi que dans certains domaines deux facteurs peuvent interagir de façon importante, et dans d'autres domaines non. Si l'on choisit le bon domaine, négliger l'interaction entre ces deux facteurs n'amènera pas d'erreur de modèle. Ce n'est plus vrai dans le cas contraire.

En fait, les seules estimations satisfaisantes, au sens où elles autorisent une inférence statistique et une extrapolation correctes, sont celles issues de plans d'expérience qui permettent d'estimer la variance due aux erreurs de mesure et aux effets incontrôlés (*pure error*) et d'estimer l'erreur de modèle, au moins partiellement, (*lack of fit*). Mais bien entendu, il n'y a pas de miracle, ces plans d'expérience seront plus coûteux. En conclusion, suivant l'estimation de la variabilité résiduelle dont on dispose ou dont on ne dispose pas, il faut être conscient de ce à quoi on peut extrapoler les résultats de l'expérience, et des risques que l'on est capable de contrôler.

RÉFÉRENCES BIBLIOGRAPHIQUES

- DANIEL, C. [1959], Use of half-normal plots in interpreting factorial two-level experiments, *Technometrics*, **1**, 311-341.
- BOX, G. E. P. [1966], A Simple System of Evolutionary Operation Subject to Empirical Feedback. *Technometrics*, **8**, 19-26.
- BOX, G. E. P. [1988], Signal-to-Noise Ratios, Performance Criteria, and Transformations. *Technometrics*, **30**, 1-17.
- SADO, G., SADO, M. C. [1991], *Les plans d'expériences : de l'expérimentation à l'assurance qualité*. AFNOR Technique, 266 p.
- KOBILINSKY, A. [1997], *Les plans factoriels in Plans d'expériences : applications à l'entreprise*. Paris, Technip, 69-209.