

**L'ENSEIGNEMENT DE LA STATISTIQUE
DANS LE DOMAINE AGRONOMIQUE:
QUELQUES IDÉES GÉNÉRALES⁽¹⁾**

Pierre Dagnelie

Faculté des Sciences agronomiques
B-5030 Gembloux (Belgique)

pierre@dagnelie.be

SUMMARY

Although a very large number of references are devoted to the teaching of statistics, very few are specifically related to the field of agriculture.

The purpose of this paper is to present some views on this topic. We successively consider the following aspects: the frame of our comments (section 1), the difficulties of teaching statistics to agricultural students (section 2), the qualification of the teaching personnel (section 3), some problems related to the content of the courses (section 4), and the teaching aids (section 5).

All these aspects are considered for the general teaching of statistics to agricultural students, but some comments are also given concerning the continuous education and the teaching given to specialists of statistics in the field of agriculture (section 6).

Some references posterior to the list of Misra et al. [1987] or not included in this list are given.

Most of the comments are probably also relevant to other fields of the experimental sciences.

1. INTRODUCTION ET RÉSUMÉ

L'enseignement de la statistique est l'objet d'une abondante littérature, comme en témoigne notamment la revue bibliographique de Misra *et al.* [1987], qui comprend plus de 1.600 références. Toutefois, si certains domaines particuliers, tels les sciences commerciales, économiques et de gestion, et les sciences de la santé, possèdent une bibliographie propre relativement importante [Gandhi *et al.*, 1991; Sahai *et al.*, 1990], il n'en est pas de même pour le domaine agronomique. Une exception notable est cependant le travail de Cochran [1945].

L'objectif de cette note est de présenter quelques idées générales en la matière. Après avoir défini de façon plus précise le cadre envisagé, nous traiterons successivement des difficultés rencontrées (paragraphe 2), des enseignants (paragraphe 3), du contenu des enseignements (paragraphe 4) et des supports pédagogiques (paragraphe 5)⁽²⁾.

⁽¹⁾ *In: Proceedings of the Fourth International Conference on Teaching Statistics (vol. 2). Rabat, Institut national de Statistique et d'Économie appliquée, 344-351, 1994.*

⁽²⁾ Pour les différentes matières envisagées, nous donnerons quelques indications bibliographiques, en nous limitant à des références postérieures à la publication de Misra *et al.* [1987] ou qui ne sont pas citées dans cette publication. En outre, on pourra consulter aussi, d'une manière générale, les comptes rendus de la *Third International Conference on Teaching Statistics (ICOTS 3)* [Vere-Jones, 1991], y compris un article de Hadzivukovic

L'enseignement de la statistique dans le domaine agronomique peut être envisagé à trois niveaux différents: les études de base, qui concernent la formation de l'ensemble des diplômés de l'enseignement universitaire ou supérieur agronomique, le recyclage ou la formation continue des chercheurs et techniciens du secteur agronomique, et la formation des spécialistes en statistique ou en biométrie appelés à travailler dans le secteur agronomique.

Nous considérerons essentiellement ici la formation générale des diplômés en agronomie, en donnant toutefois quelques indications également au sujet de la formation des spécialistes de la statistique et de la biométrie et de la formation continue (paragraphe 6), ce dernier aspect étant traité plus particulièrement, au cours de la même réunion, par J. Badia⁽¹⁾⁽²⁾.

On notera que les réflexions formulées ci-dessous s'appliquent sans doute également, dans une large mesure, aux autres domaines des sciences expérimentales.

2. LES DIFFICULTÉS RENCONTRÉES⁽³⁾

En abordant le problème de l'enseignement de la statistique dans le domaine agronomique, il nous paraît important d'identifier tout d'abord certaines sources de difficultés.

Une première source de difficultés est l'hétérogénéité du public étudiant. Certains étudiants en agronomie ont en effet, comme principale motivation, leur goût pour tout ce qui a un caractère "naturaliste", sans aucun goût pour le "quantitatif" et les sciences exactes, dont les mathématiques. D'autres, au contraire, abordent plus volontiers ces derniers aspects. Or, un bon apprentissage des méthodes statistiques, appliquées à quelque domaine que ce soit, nous semble exiger à la fois une qualification suffisante dans le secteur du quantitatif et dans le domaine d'application considéré.

Une deuxième source de difficultés est le caractère très dogmatique de tout l'enseignement antérieur, primaire et secondaire. Un tel type d'enseignement ne prépare pas du tout, bien au contraire, les étudiants à "appréhender l'aléatoire" avec une relative facilité⁽⁴⁾.

[1991] relatif au secteur agronomique, ainsi que d'autres ouvrages relativement récents, tels ceux d'Anderson et Loynes [1987], de Hawkins [1990] et de Hawkins *et al.* [1992].

(1) Voir aussi, à ce sujet, Badia et de Turckheim [1993].

(2) Pour mieux définir le cadre de cette note, précisons aussi que notre expérience, en matière d'enseignement de la statistique, a été acquise à la Faculté des Sciences agronomiques de Gembloux (Belgique) et, comme professeur-visiteur, dans d'autres institutions de divers pays africains et européens, dont l'Institut agronomique et vétérinaire Hassan II, à Rabat (Maroc). Il s'agit d'études agronomiques postérieures à l'enseignement secondaire et d'une durée de cinq ou six ans. Par comparaison avec le système français, ces études englobent à la fois l'équivalent des classes préparatoires et les études supérieures agronomiques proprement dites. Enfin, on notera également que, sous le vocable "études agronomiques", sont pris en considération, non seulement les secteurs classiques des productions végétales (grandes cultures) et animales, mais aussi l'horticulture, les eaux et forêts, le génie rural, l'économie et la sociologie rurales, l'aménagement des territoires, les industries agricoles et agro-alimentaires, les bio-industries, etc.

(3) Voir aussi, à ce sujet, Sahai [1990] et Watts [1991].

(4) On notera qu'en Belgique, comme dans de nombreux autres pays, la statistique est totalement absente ou est présentée de façon extrêmement limitée dans l'enseignement primaire et secondaire. La remarque formulée ici perd sans doute de son importance lorsque des notions substantielles de statistique sont, au contraire, introduites dans l'enseignement secondaire.

Une troisième source de difficultés est le caractère abstrait de très nombreux concepts, même élémentaires, de la statistique. Il suffit de penser, à ce propos, à des notions telles que moyenne et variance, probabilité et variable aléatoire, valeurs théoriques relatives à une ou plusieurs populations et estimations de ces valeurs à partir des valeurs observées relatives à un ou plusieurs échantillons, etc.

Enfin, une quatrième source de difficultés, qui n'est pas totalement étrangère aux précédentes, est le fait que les exemples qui sont présentés et les problèmes qui sont posés aux étudiants n'ont en général pas de réponses parfaitement définies, mais au contraire des réponses qui, le plus souvent, ont tout au plus une certaine probabilité ou une certaine vraisemblance, difficile à chiffrer.

3. LES ENSEIGNANTS

Une question qui est souvent posée est celle de savoir si l'enseignement de la statistique en agronomie doit être confié à un mathématicien ou un statisticien, ou au contraire à un agronome de l'une ou l'autre spécialisation (agronome généraliste, forestier, etc.).

Cette question ne nous paraît pas essentielle, mais ce qui nous semble par contre fondamental, c'est que les personnes responsables de l'enseignement de la statistique en agronomie aient une double qualification, à la fois dans le domaine statistique et dans le domaine agronomique, ou en tout cas une très bonne connaissance des deux domaines.

Une solution alternative peut être de constituer des équipes mixtes, réunissant des mathématiciens ou statisticiens et des agronomes, mais il s'impose alors que les différentes personnes constituant de telles équipes travaillent réellement en étroite collaboration de façon permanente.

4. LE CONTENU DES ENSEIGNEMENTS

4.1. Principes généraux

Notre intention n'est pas de définir ici un contenu précis, pour l'un ou l'autre enseignement de statistique, ce contenu dépendant en général très largement de circonstances particulières, telles que le temps disponible, le niveau atteint par les étudiants, tant en mathématiques que dans le domaine agronomique, etc. Nous voudrions au contraire émettre seulement quelques considérations générales, qui tiennent compte notamment des différentes sources de difficultés énumérées au paragraphe 2.

Un premier point important nous paraît être d'assurer toujours un bon équilibre entre la théorie et les applications, pour que les étudiants puissent arriver à un niveau suffisant, à la fois, de formation de base, nécessaire à une bonne compréhension des méthodes statistiques, y compris leurs conditions d'application, et d'expérience de l'utilisation concrète de ces méthodes.

Un deuxième élément qui nous paraît important est d'organiser l'enseignement en allant toujours du concret vers l'abstrait, par exemple en prenant en considération la statistique descriptive, y compris le traitement de données réelles, avant l'étude des distributions théoriques et l'introduction de notions d'inférence statistique.

4.2. Les notions de variabilité et de distributions d'échantillonnage⁽¹⁾

La variabilité et les distributions d'échantillonnage nous semblent être des notions essentielles, auxquelles une attention toute particulière doit être accordée. Il importe en effet que

(1) Voir aussi à ce sujet, Dagnelie [1992, § 8.4.3], Stirling [1987] et Van Vyve-Genette *et al.* [1986].

les étudiants prennent parfaitement conscience du fait que les résultats obtenus à l'issue d'une enquête ou d'une expérience ne constituent pas une image intangible de la ou des populations concernées et, en conséquence, que la répétition de la même enquête ou de la même expérience doit tout naturellement conduire à des résultats différents. Il importe aussi que les étudiants prennent conscience dans une certaine mesure de l'ordre de grandeur de ces différences.

Nous pensons que la réalisation d'un certain nombre de simulations, auxquelles les étudiants prennent une part active, doit contribuer à l'introduction de ces notions: simulation d'échantillonnages, mettant bien en évidence les différents risques d'erreur liés à l'échantillonnage, et simulations de distributions d'échantillonnage de paramètres tels que moyennes, médianes et variances. Ces simulations peuvent être réalisées de différentes manières, notamment par un travail collectif, organisé en salle de cours avec la collaboration de l'ensemble des étudiants, ou par un travail individuel des étudiants sur ordinateur.

Des simulations très simples peuvent en effet servir à introduire un grand nombre de notions: échantillonnage aléatoire et simple, biais, estimation d'un pourcentage ou d'une moyenne, etc.

4.3. Les notions d'intervalle de confiance et de test d'hypothèse⁽¹⁾

L'introduction des notions d'intervalle de confiance et de test d'hypothèse présente les mêmes types de difficultés et nécessite la même attention.

Ici également, la réalisation de simulations peut être fort utile, notamment pour bien faire comprendre le sens exact des différents risques d'erreur et l'importance de ces risques, en particulier en ce qui concerne la puissance des tests.

4.4. Les exemples et les exercices⁽²⁾

Les exemples et les exercices jouent un rôle fondamental dans tout enseignement de statistique à orientation pratique et doivent aussi être l'objet d'une attention toute spéciale.

Les exemples destinés à illustrer l'enseignement doivent être relativement simples, peu "volumineux", mais néanmoins réalistes, ce qui pose souvent des problèmes de choix assez délicats. Quant aux exercices ou aux applications, qui sont donnés aux étudiants en complément à l'enseignement proprement dits, ils doivent également correspondre à des situations réelles, avec un "arrière-plan" agronomique suffisant, en ce qui concerne notamment les problèmes concrets étudiés et les modalités de collecte des données, et ils doivent être assez diversifiés, sans être toutefois trop spécialisés.

Dans tous les cas, l'accent doit être mis en particulier sur les problèmes de choix des méthodes statistiques les plus adéquates et sur les fondements théoriques et les conditions d'application de ces méthodes. En outre, il importe que tous les exemples ou applications correspondent non seulement à des situations réelles, mais reviennent aussi aux problèmes concrets, les conclusions finales devant toujours être exprimées en termes agronomiques, et pas uniquement en termes statistiques.

Les mêmes principes s'appliquent évidemment aussi au choix des questions d'examens, qui viennent d'ailleurs régulièrement enrichir, pour les années ultérieures, les collections d'exercices disponibles.

(1) Voir aussi, à ce sujet, Foucart [1991], Stirling [1987] et Van Vyve-Genette *et al.* [1986].

(2) Voir aussi, à ce sujet, Preece [1986] et Singer et Willett [1990].

4.5. Les mémoires ou travaux de fin d'études

La préparation des mémoires ou des travaux de fin d'études est souvent une excellente occasion de mettre en pratique les notions acquises dans le cadre des enseignements de statistique.

Au-delà de l'utilisation relativement classique des méthodes qui leur ont été présentées antérieurement, les étudiants sont fréquemment amenés, à ce stade de leurs études, à avoir recours à des méthodes nouvelles pour eux, ce qui est à la fois l'occasion de leur faire prendre connaissance de leurs propres limites et d'entrer en contact avec l'un ou l'autre système de consultation statistique. Encore faut-il, pour que ce type d'apprentissage soit tout à fait bénéfique, qu'un service efficace de consultation soit opérationnel et que les étudiants prennent l'initiative d'entrer en contact suffisamment tôt avec ce service, ce qui est malheureusement loin d'être toujours le cas.

4.6. Les représentations graphiques et les moyens audio-visuels⁽¹⁾

Les moyens actuels de traitement de l'information permettent d'illustrer, beaucoup plus que par le passé, les matières enseignées et les résultats obtenus, par des représentations graphiques. C'est un autre point auquel il y a lieu d'être très attentif, et cela à tous les niveaux de l'enseignement (cours théoriques, exemples, exercices, etc.). En particulier, la projection sur grand écran de résultats apparaissant sur terminal ou sur micro-ordinateur ouvre des perspectives nouvelles en matière de présentation de notions théoriques et de discussion de résultats d'exercices, et doit être utilisée au maximum.

Dans le même ordre d'idées, on peut penser à l'utilisation accrue de l'ensemble des moyens audio-visuels (présentation de montages de diapositives, films didactiques, etc.). Nous n'en avons toutefois qu'une expérience relativement limitée, dans une certaine mesure sans doute parce que le matériel disponible en la matière n'est pas très abondant.

4.7. L'exécution des calculs⁽²⁾

Nous dirons quelques mots, au paragraphe suivant, du recours à l'ordinateur et de l'emploi de certains logiciels. Mais nous pensons que, quels que soient les moyens informatiques mis à la disposition des étudiants, il s'impose de continuer à leur faire exécuter certains calculs "à la main", c'est-à-dire à l'aide de calculettes non programmables (sauf en ce qui concerne éventuellement quelques fonctions élémentaires, telles que les calculs de moyennes, de variances, etc.)⁽³⁾.

L'exécution "manuelle" de certains calculs est souvent une excellente façon de bien saisir le sens des méthodes étudiées ou, au moins, d'en avoir une meilleure compréhension ou une autre vision. Ce travail "manuel" permet aussi de prendre mieux conscience des risques inhérents à une utilisation aveugle des logiciels statistiques.

En outre, il importe d'apprendre aux étudiants à obtenir par raisonnement ou par calcul mental, chaque fois que c'est possible, des ordres de grandeur relatifs aux résultats qu'ils attendent. Il faut en effet qu'ils acquièrent le réflexe de toujours regarder avec circonspection les résultats qui leur sont fournis et de toujours vérifier au moins si ces résultats sont vraisemblables ou plausibles, et cela quels que soient les moyens de calcul utilisés.

(1) Voir aussi, à ce sujet, Moore [1993].

(2) Voir aussi, à ce sujet, Khamis [1991].

(3) A cet égard, on notera également l'existence de certains logiciels qui permettent de traiter les opérations matricielles comme si on travaillait à l'aide d'une calculette [Quris, 1993]. De tels logiciels peuvent être fort utiles notamment dans l'apprentissage des méthodes d'analyse statistique à plusieurs variables.

5. LES SUPPORTS PÉDAGOGIQUES

5.1. Les manuels et les livres

En matière de manuels ou de livres, il nous paraît essentiel d'imposer aux étudiants l'utilisation de documents dont le contenu déborde largement les stricts besoins de l'enseignement et qui comportent de nombreuses références bibliographiques. L'emploi de photocopies peut évidemment faciliter grandement, à court terme, la tâche des étudiants, mais leur rendra plus difficile l'accès ultérieur à d'autres documents et, de ce fait, les amènera inévitablement à se limiter souvent aux seules notions vues dans le cadre des enseignements reçus.

À ce propos, on se souviendra toujours, d'une manière générale, du fait que l'enseignement a pour but de donner une formation de base et une ouverture sur des notions qui ne sont pas enseignées, et non pas seulement de conduire à la réussite des examens!

5.2. Les logiciels⁽¹⁾

Il s'impose évidemment de mettre à la disposition des étudiants des moyens informatiques adéquats, tant en matériel qu'en logiciel.

À cet égard, notre choix s'est porté principalement sur le logiciel Minitab, dont l'apprentissage est extrêmement simple, qui comporte une diversité largement suffisante de méthodes de base et qui a aussi l'avantage de ne pas être exclusivement statistique, ce qui le rend utile également pour d'autres enseignements. Le logiciel SAS est aussi disponible pour nos étudiants, pour compléter éventuellement Minitab dans le cadre de la réalisation des travaux de fin d'études.

6. LE RECYCLAGE OU LA FORMATION PERMANENTE ET LA FORMATION DES SPÉCIALISTES EN STATISTIQUE

Les différents principes que nous avons exposés nous semblent devoir s'appliquer également, moyennant certaines adaptations, au recyclage ou à la formation continue des chercheurs et techniciens du secteur agronomique. Sans entrer dans trop de détails, disons par exemple qu'en matière de recyclage, on se trouve bien sûr normalement en présence d'un public mieux formé et plus motivé, et déjà bien conscient en général de notions telles que la variabilité des résultats d'enquêtes et d'expériences. Il n'empêche qu'avant de vouloir introduire des notions nouvelles, il nous paraît essentiel de refaire le point au sujet de certaines notions de base, notamment en matière de distributions d'échantillonnage.

De même, les principes présentés ci-dessus s'appliquent très largement aussi, toujours moyennant certains aménagements, au cas de la formation des spécialistes en statistique ou en biométrie appelés à travailler dans le secteur agronomique. Bien sûr, la formation sera ici plus poussée et, dans une certaine mesure, plus fondamentale, mais il faudra rester extrêmement attentif à donner aussi une vision simple des choses et à toujours revenir aux problèmes concrets. Le statisticien ou le biométricien appelé à exercer des fonctions de consultation devra en effet, impérativement, pouvoir s'exprimer de façon très simple et dans le langage de ses interlocuteurs⁽²⁾.

(1) Voir aussi, à ce sujet, Stephenson [1990].

(2) Voir aussi, à ce sujet, Dagnelie [1994].

7. BIBLIOGRAPHIE

- Anderson C.W., Loynes R.M. [1987]. *The teaching of practical statistics*. New York, Wiley, 199 p.
- Badia J., de Turckheim E. [1993]. A continuing education programme in statistics at the *Institut national de la Recherche agronomique* (INRA) in France. *Biom. Bull.* **10** (1), 19-20.
- Cochran W.G. [1945]. Training at the professional level for statistical work in agriculture and biology. *J. Amer. Stat. Assoc.* **40**, 160-166, and in: Bancroft T.A., Huntsberger D.V. (ed.). [1961]. *Some papers concerning the teaching of statistics*. Ames, Iowa State University, 18-24.
- Dagnelie P. [1992]. *Statistique théorique et appliquée. Tome 1: les bases théoriques*. Gembloux, Presses agronomiques, 504 p.
- Dagnelie P. [1994]. La consultation statistique: quelques réflexions (en préparation).
- Foucart T. [1991]. *Introduction aux tests statistiques: enseignement assisté par ordinateur*. Paris, Technip, 165 p.
- Gandhi B.V.R., Sahai H., Acevedo N. [1991]. Teaching of statistics in business, economics and management sciences: some comments and a selected bibliography. *Statistician* **40** (1), 95-106.
- Hadzivukovic S. [1991]. Teaching statistics to students in agriculture. In: Vere-Jones D. (ed.). *Proceedings of the Third International Conference on Teaching Statistics* (vol. 2). Voorburg, International Statistical Institute, 399-401.
- Hawkins A. (ed.). [1990]. *Training teachers to teach statistics*. Voorburg, International Statistical Institute, 297 p.
- Hawkins A., Jolliffe F., Glickman L. [1992]. *Teaching statistical concepts*. Harlow, Longman, 246 p.
- Khamis H.J. [1991]. Manual computations: a tool for reinforcing concepts and techniques. *Amer. Stat.* **45** (4), 294-299.
- Misra S.C., Sahai H., Gore A.P., Garrett J.K. [1987]. A bibliography on the teaching of probability and statistics. *Amer. Stat.* **41** (4), 284-310.
- Moore D.S. [1993]. The place of video in new styles of teaching and learning statistics. *Amer. Stat.* **47** (3), 172-176.
- Preece D.A. [1986]. Illustrative examples: illustrative of what? *Statistician* **35** (1), 33-44.
- Quris R. [1993]. CALMAT: un opérateur de calcul matriciel simple pour l'initiation aux analyses multivariées. In: *XXVes Journées de Statistique*. Vannes, Institut universitaire de Technologie, 3 p.
- Sahai H. [1990]. Some problems of teaching an introductory biostatistics course to graduate students in health sciences: coping with the diversity of student aptitudes, interests and objectives. *Statistician* **39** (4), 341-347.
- Sahai H., Misra S.C., Toro C. [1990]. The teaching of statistics in the biological, medical and health sciences: some comments and a selected bibliography. *Biom. J.* **32** (4), 449-489.
- Singer J.D., Willett J.B. [1990]. Improving the teaching of applied statistics: putting the data back into data analysis. *Amer. Stat.* **44** (3), 223-230.
- Stephenson W.R. [1990]. A study of student reaction to the use of Minitab in an introductory statistics course. *Amer. Stat.* **44** (3), 231-235.
- Stirling W.D. [1987]. *StatLab: microcomputer-based practical classes and demonstrations for teaching basic statistical concepts*. Wellington, New Zealand Statistical Association, 175 p.
- Van Vyve-Genette A., Gohy J.M., Feytmans E. [1986]. *Statistique élémentaire en sciences biomédicales: apprentissage par la micro-informatique*. Bruxelles, De Boeck, 163 p.
- Vere-Jones D. (ed.). [1991]. *Proceedings of the Third International Conference on Teaching Statistics* (2 vol.). Voorburg, International Statistical Institute, 556 + 587 p.
- Watts D.G. [1991]. Why is introductory statistics difficult to learn and what can we do to make it easier? *Amer. Stat.* **45** (4), 290-291.