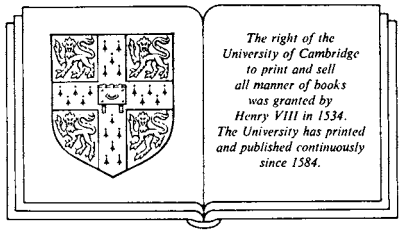
EDITORS: D. J. HAND AND B. S. EVERITT

The statistical consultant in action



CAMBRIDGE UNIVERSITY PRESS
Cambridge
London New York New Rochelle
Melbourne Sydney

Published by the Press Syndicate of the University of Cambridge
The Pitt Building, Trumpington Street, Cambridge CB2 1RP
32 East 57th Street, New York, NY 10022, USA
10 Stamford Road, Oakleigh, Melbourne 3166, Australia

First published 1987

Printed in Great Britain by the University Press, Cambridge

British Library cataloguing in publication data

The Statistical consultant in action.

1. Mathematical statistics
I. Hand, D. J. II. Everitt, Brian
519.5 QA276

Library of Congress cataloguing in publication data

The Statistical consultant in action.

Bibliography
1. Statistical consultants. 2. Statistics.
I. Hand, D. J. II. Everitt, Brian.
QA276.17.S73 1987 001.4'22 86-17159

ISBN 0 521 30717 1

UP

Contents

1

Statistical consultancy

D. J. HAND AND B. S. EVERITT

1 Introduction

Since this is a book about statistical consultancy, a good place to start would be by considering the ideal consultation. Hyams (1971) has given us a description of this:

> To qualify a consultation as ideal is to deny its empirical meaning. The 'Ideal Consultation' is *not* a consultation. It is a working-together, a voluntary meeting of minds and union of energies whose prime aim is to seek a 'truth'. In such meetings both parties are familiar with each other's basic language. The biologist has had a few courses in basic statistics and thus recognises statistics as a unique and valuable discipline. The statistician has also done his homework and has familiarised himself with the names and the relationships of the fauna in the experimenter's jungle. Since knowledge and understanding breed sympathy and respect, the researcher esteems the statistician as an expert representative of this most important science. His appreciation for the statistician's unique contributions grows by leaps and bounds with the experience of his individual talents. Needless to say, the feeling is mutual. Meetings are stimulating; they are productive in thought and in product. The work forms a gestalt (the whole is greater than the sum of its parts). The research poses challenging statistical problems that are fun to work at: the sort of thing that keeps one busy at a scratch pad during supper while the wife silently suffers (or throws a fit). In unhurried time the deliberations proceed to a design, an experiment, and an analysis that confirms everyone's best hopes. The (multiple) reports are easy to write. Sometimes the biologist's name is first, sometimes the methodologist's; it hardly matters.

These manuscripts are received enthusiastically by journal editors and their 'expert' reviewers don't give the team a hard time. First experiments lead naturally to others and the information generated finds a significant practical application. Ultimately it saves human lives or curtails misery. Finally, but justly, the co-workers are awarded the Nobel Prize in Medicine and quite naturally donate their stipends to schools encouraging interdisciplinary approaches to problem solving.

In practice there seem to be some slight (!) differences between this ideal and the reality, so a natural question to ask is where are we going wrong, or more realistically, what problems do consultant statisticians have to contend with which disrupt such an ideal scenario?

2 The problems of statistical consultancy

The primary issue affecting realisation of the above ideal is, of course, the relationship between the client and the consultant. This will naturally depend on personalities, preconceptions of the role of the other party, and the nature of the relationship between the individuals concerned. We shall have more to say about these things below, but the sort of factors involved are: how much statistics the client knows; whether he is prepared to accept advice or is confident he knows the answers and is just seeking confirmation; whether the consultant is a freelance statistician or a junior member of a university department of which the client is the head; whether it is to be a genuine collaboration or whether the statistician will merely be acknowledged in some subsequent paper (and does he want this anyway, if the client has ignored his advice?). Perhaps most important of all is the question of whether the statistician is perceived as a scientist in his own right. This sets the tone of the relationship.

A number of authors have examined the relationship by condensing their experience of clients into a classification. Sprent (1970), for example, produces the following seven classes:

1. *The timid apologist* who has little statistical knowledge and expresses reluctance at wasting our time. It takes something akin to a doctor's bedside manner to put such people at their ease and overcome their reluctance to seek help.

2. *Significant difference and least significant difference* (*l.s.d.*) *experts*. These are encouraged by misguided editors who think all numerical results can be made respectable by quoting significant differences or significance levels – often denoted by * or ** or ***, a symbolism more appropriate to a hotel guide-book than a serious scientific paper. The number of editors accepting l.s.d.'s

or asterisks as the ultimate in statistical sophistication is happily declining. In my experience it is a characteristic of inveterate users of significance levels that they do not know what they mean. Once the meaning is explained to them more sophisticated ideas appeal to them. The number of l.s.d. experts might decrease if we pressed for less emphasis on significance in elementary service courses.

3. *The one-technique amateur statisticians.* These are proficient at just one technique and show great ingenuity at applying it even when it is not relevant. It is worth some effort to expand their statistical horizons.

4. *The believers in sacred texts or computers*, have been mentioned in the previous section.

5. *Experimenters with addled statistical ideas.* These are the people who assure you that, for example "the object of randomisation is to disperse treatments as widely as possible over the experimental area," or that they "never use randomised blocks, but always do factorial experiments instead". It requires some diplomacy to replace such misconceptions by useful knowledge.

6. *The expert data handler who is not a statistician.* Such a person may be described as a born data handler who relies upon his instincts when handling data. It is usually easy for the statistician to collaborate with him, but one feels that he spends perhaps too much time re-inventing known techniques if he proceeds without some statistical help.

7. *The statistically informed experimenter* is always a pleasure to work with. Not only does he understand our jargon, but he usually sees to it that we understand that of his subject so far as is necessary.

Hyams' (1971) classification is also worth reading, yielding the categories: probabilist, numbers collector, sporadic leech, amateur statistician and long distance runner.

(If, at this point, the reader should feel we are being unfair on the client in what is, after all, a two-party relationship, we hasten to reassure you that the balance will be redressed below.)

We mentioned, above, the lack of statistical knowledge on the part of the client. It is precisely because he lacks this knowledge, and is aware of it, that he is calling upon the professional services of the statistician. A more subtle source of potential pitfalls, however, lies in lack of expertise the other way round. How much does the statistician know about the client's discipline? Inadequate knowledge can not only lead to difficulties for the statistician in formulating the researcher's questions in a way he

can answer, but it can lead to fundamental misconceptions of the researcher's aims. Moses and Louis (1984) give an elegant little example of this, illustrating the importance of understanding what the measurements mean and how they were obtained: the statistician is presented with two measurements of phase angle, 10° and 350°, and works with their average of 180°. Because of these sorts of difficulties Cox (1968) has suggested that there should be texts describing other scientific disciplines specifically aimed at statisticians, just as there are statistics texts aimed at other scientists. Indeed, recognising this symmetry of the client/consultant relationship, Sprent (1970) concludes his taxonomy of clients by saying: 'Whether or not an experimenter fits neatly into one of the above categories, a collaboration will never be helped by our sneering at his statistical ignorance. I shudder to think how often I have appeared naive when talking to experts in another field about their speciality.'

Several authors take the symmetry further, and suggest that one of the roles of the consultant statistician should be as teacher, educating the researcher in statistical methodology. From this it also follows that what is an 'optimal' solution in practice may differ from what is 'optimal' in theory. A complex multivariate repeated measures analysis of variance on logged data may be perfect from the theoretical statistician's viewpoint, but if the client has no hope of understanding it then it is worse than useless – worse because of the reaction it will provoke and the role misconceptions it will create.

A problem which seems ubiquitous in modern life is lack of time. In statistical consultancy this can manifest itself in two ways. The first is the client's lack of time: he needs an answer by tomorrow at the latest. This might be because he is under pressure from his boss, because his business will collapse if an answer is not available, or because he is presenting the results at a conference the day after. It might simply be that he knows that computers are immensely fast and so does not see any problem in producing a result within a few minutes or hours at most. Of course, apart from the basic misconceptions in this notion, it also reflects poorly on the client's understanding of the other pressures on the statistician. This is the second kind of time pressure: the statistician's lack of time. For statisticians who function in a service capacity this can be a very serious problem. Often there are ten other clients queuing outside the door while one is grappling with the intricacies of some complex data set. The temptation, of course, is simply to adopt the most straightforward procedure as a solution, and then have to live with the feeling that one could have done so much better if only one had had the time to consider the problem properly.

Often the complex data set referred to above was collected without first

seeking the advice of a statistician. Not infrequently years of effort have gone into collecting the data. This complaint, that the statistician was called in too late, is a common one. At its worst it can lead to the waste of vast sums of money or the abandonment of a PhD. Helen Kraemer (quoted in Moses and Louis, 1984) says:

> If consultation is at the post hoc stage, it may be that the objectives cannot be accomplished (sampling bias, poor design, etc.). It is the statistician's responsibility to state this frankly. We cannot do magic, and we can't participate in cover-ups. It is as well that researchers know our limitations in advance. This is a particular problem when the first consultation takes place after a research paper is rejected for publication because of poor methodology. Not much one can do!

(It is interesting, however, that Daniel (1969) says: 'Some statisticians would say that the only favourable time to enter a research project is at its beginning. My own experience does not confirm this. I have entered projects at all stages of their development. I do not see any connection between my stage of entry and my success or failure.')

Another pressure influencing the consultant/client relationship is the less obvious one of ethical issues. These are, of course, well known in the medical field and perhaps also in social applications of statistics. But they occur elsewhere as well. An example would be in a university environment, where moral problems can arise with the students seeking help: just how much statistical advice should one provide? Whose PhD is it anyway?

3 A taxonomy of statisticians

We have presented, above, a typology of the client. In fairness we must also present one such of the consultant. Hyams (1971) gives us the following:

> 1. *The Model Builder* fits any and every data problem set to a model he is presently interested in or knows something about. It matters not whether he investigates the questions that are being asked by the client or those that are biologically important. For that matter, this type isn't really interested in hearing the client's story. He had posed his own a priori questions before the client knew him. The Model Builder is like the drunkard looking for his lost key under the street lamp although he dropped it in the dark alley. He justifies his search by pointing out that there is light in the place he is looking.
>
> 2. *The Hunter* is the statistician counterpart of the Numbers Collector who directs you to 'mine the mountain'. The Hunter

will subject every data set to an exhaustive and extensive computer analysis. For a relatively simple problem with scanty data he will ultimately present the investigator with 14 vertical inches of print-out, containing 17 significant results. These numbers do not bear a relationship to anything on the face of this earth except themselves. While the client may initially accept these authoritative materials with reverence, it will not take him long to figure out that he has a bag of wind.

3. *The Gong* is a consultant who starts every conference by drawing a bell-shaped curve.

4. *The Traditionalist* is convinced that nothing really important has happened in statistics since R. A. Fisher and consequently limits himself to a restrictive working vocabulary. He views the computer as the devil's work.

5. *The Randomophiliac* firmly believes that it doesn't matter what else you do, as long as you've randomised well. He is like the mother who catches her 14 year old daughter in a sexually compromising situation and admonishes her by saying "as long as you don't smoke, honey".

6. *The Quantophreniac's* position is: It doesn't matter if you observe what you want to so long as you get a hard measurement.

7. *The More Data Yeller* (he needs no further description).

8. *The Nit Picker* will always focus his attention on the inconsequential but debatable. He will enlarge minor issues out of reasonable perspective and quickly reduce a real and tremendous contribution to a potentially horrendous error in reality testing. (My manuscripts are usually reviewed by this type.)

To these we might add the *problem stealer* (who decides that every problem would make a perfect project for his students, to be begun next summer; from this G. J. Goodhardt (1970) derives his 'rule 1 of the business – never consult an academic in October') and the *allied problem solver* (who, as Goodhardt (1970) says: 'finds great interest, not in my problem, but in some other problem that mine suggested to him. This may take the form of a wider generalisation of the conditions which unfortunately does not happen to include the special case I started with, or a detailed description of the intricacies of estimation in small samples when I have a sample of size 5000.').

One hopes that this is just a list of inadequate types, the good and competent ones having been omitted from the list.

These caricatures at least make it clear what we should strive to avoid. Presented with them one might justifiably ask how we should go about

training effective consultants. This question is dealt with in a number of places (for example, Committee on Training of Statisticians for Industry, 1980; Boen, 1972; Griffiths and Evans, 1976; Tarter and Berger, 1972; Watts, 1970; Zelen, 1969; and Zahn and Isenberg, 1983). This book is, we hope, a further answer, complementing the advice on training given in the above by exposing the reader to a taste of the wide range of problems that will be encountered in the life of a statistical consultant.

4 Statistical domains

It has been suggested that after one's formal training in statistics (to BSc, MSc or PhD level) it then takes a further three years functioning within a particular application environment before one attains sufficient competence to act as an independent consultant. The reason for this will be partly the need to acquire the personnel skills mentioned above, partly the need to adjust to the problems of real data (missing values, outliers, multiple sources, etc.), and partly the fact that different areas of application place different degrees of emphasis on different techniques. The extent to which this is true is illustrated by the fact that, even within statistics itself, the technical term 'theory of reliability' has two quite distinct meanings. One refers to the reliability of (for example) complex machines, and the other to the consistency with which measuring scales yield identical results (in the behavioural sciences).

Application domain is just one type of categorisation which can be used to describe statistical consultancy work. A second is the working environment. For example, the statistician might be an academic who spends a small part of his time (voluntarily) advising people; or he might work from a service unit, with his primary function being to advise; or he might be a freelance consultant, who eats or goes hungry according to the success of his consultancy work. These three types have very different roles and requirements.

The academic can afford to look merely at interesting (to him) problems, can afford to be sidetracked to more interesting ones (if a client does not return for more advice it does not personally damage the statistician), and he may not be part of a team.

The statistician within a service unit is obliged to answer, or attempt to answer, the questions of anyone who knocks at his door. There is a danger that the role of the statistician, as a scientist in his own right, will not be properly perceived. Armitage (1970) says that as much as possible of the service function should be handed over to the client himself. (In fact Armitage prefers the term 'advisory work' unless there is a commercial agreement, a point with which we agree.)

For statisticians who fall into either of the above categories it is essential for precise roles to be mapped out beforehand. Is it to be a collaboration, with both names appearing on any subsequent publications?

In contrast, for statisticians in the third class – that of the freelance consultant – the roles are already well delineated. The financial motivation sees to that.

5 Computers and statistical consultancy

Computers, of course, have revolutionised statistics. How many of the case studies in this book would have been feasible without them? But the real potential of computers is only just beginning to be realised.

The initial impact of computers was to speed up, to minutes, techniques which previously would have taken days or weeks to apply by hand. This has had the consequence of much more widespread application of these methods and of pushing the applications further to bigger and more complex problems. Much more interesting than this, however, has been the development of new techniques for which the computer is absolutely essential and which, without computers, just would not exist. Examples of such children of the computer age are log-linear modelling, kernel density estimation techniques, and bootstrap methods. There is no evidence that progress in this direction has stopped. Developments are continuing, and the advent of even more potent computers makes the prospects truly exciting.

Apart from the development of new techniques, developments in computers seem likely to revolutionise statistical consultancy work from a different direction, and one which could not have been predicted before the computer age. This is that of interactive statistical graphics. The approach to data analysis in laboratories with access to fast and high-powered interactive graphics facilities is diverging from the more traditional approach.

Returning from the frontiers to the more mundane, we find widespread access to powerful statistical packages such as SPSS, SAS, BMDP, etc. Such packages are easy to use. This means that they can be used by those who are relatively untutored in statistics. They can equally easily be misused. Hooke (1980) says: 'Use [of statistics] has been replaced by overuse and misuse. Regression is being used in foolish ways in the neighbourhood of almost every computer installation.' And the rate of errors in published analyses suggests that the problem is serious (see White, 1979; Gore, Jones, and Rytter, 1977; and Altman, 1982).

We can hardly impose a moratorium on the use of such packages by

those who are not professionally qualified, so this has motivated the development of statistical expert systems, systems which statistically inexperienced researchers can use to analyse data and which will protect them from error. (See, for example, Gale (1985), Hand (1984, 1985), Pregibon and Gale (1984). One of the earliest references to the possibility of building this kind of statistical expertise into packages is Finney (1970).) Whether such systems will serve statistics by preventing some of the criticism which is currently misdirected at it, rather than at those who misuse it, remains to be seen. In any case, it will clearly be a long time before such systems can handle the kinds of problems to which this book is addressed: that is, the problems which lie at the interface between statistical expertise and expertise in the discipline of the client. Resolution of such problems requires not only statistical knowledge, but also wider knowledge of the world and the way it behaves.

6 Conclusion

The problems facing statisticians serving as consultants are varied: varied not only in the origin of the data and the research questions presented, but also in the kind of personal skills they will require the statistician to possess in order to resolve the questions successfully. Communicating with statistically and mathematically naive research workers can be an exacting and, on occasions, a frustrating task, and patience and tolerance are likely to be needed in good measure. Nevertheless, working as part of a team to solve practical problems can be very exciting, and the intellectual rewards great. The statistician as an expert on the formulation and manipulation of mathematical models and on research methodology is in an ideal position to act as a catalyst in drawing together members of a research team. In this central role the statistician is far more than merely a second class mathematician.

To become successful consultants, students clearly need to acquire a grasp of the practical problems they will encounter, in addition to the theoretical expertise imparted by their courses. It is hoped that the diverse range of real problems described in this collection will go some way towards filling that need.

7 Further reading

At the end of this book we present a bibliography of work on the practical aspects of statistical consultancy. General works which the reader might find interesting are those by Sprent (1970), Hyams (1971), Feinstein (1970), and the book by Boen and Zahn (1982).

References

Altman, D. (1982) Statistics and ethics in medical research: VIII. Improving the quality of statistics in medical journals. *British Medical Journal*, **282**, 44–7.

Armitage, P. (1970). Discussion following Sprent (1970).

Boen, J. R. (1972) The teaching of personal interaction in statistical consulting. *The American Statistician*, **25**, 30–1.

Boen, J. R. and Zahn, D. A. (1982) *The Human Side of Statistical Consulting*, Belmont, California: Wadsworth.

Committee on Training of Statisticians for Industry (1980) Preparing Statisticians for careers in industry (with discussion). *The American Statistician*, **34**, 65–75.

Cox, C. P. (1968) Some observations on the teaching of statistical consultancy, *Biometrics*, **24**, 789–802.

Daniel, C. (1969) Some general remarks on consultancy in statistics. *Technometrics*, **11**, 241–6.

Feinstein, A. R. (1970) Clinical biostatistics VI: statistical malpractice – and the responsibility of the consultant. *Clinical Pharmacology and Therapeutics*, **11**, 898–914.

Finney, D. J. (1970) Discussion following Sprent (1970).

Gale, W. A. (1985) *Artificial Intelligence and Statistics*. Reading, Massachusetts: Addison-Wesley.

Goodhardt, G. J. (1970) Discussion following Sprent (1970).

Gore, S. M., Jones, I. G. and Rytter, E. C. (1977) Misuse of statistical methods: critical assessment of articles in BMJ from January to March 1976. *British Journal*, **1**, 85–7.

Griffiths, J. D. and Evans, B. E. (1976) Practical training periods for statisticians. *The Statistician*, **25**, 125–8.

Hand, D. J. (1984) Statistical expert systems: design. *The Statistician*, **33**, 351–69.

Hand, D. J. (1985) Statistical expert systems: necessary attributes. *Journal of Applied Statistics*, **12**, 19–27.

Hooke, R. (1980) Getting people to use statistics properly. *The American Statistician*, **34**, 39–42.

Hyams, L. (1971) The practical psychology of biostatistical consultation, *Biometrics*, **27**, 201–12.

Moses, L. and Louis, T. A. (1984) Statistical consulting in clinical research: the two way street. *Statistics in Medicine*, **3**, 1–5.

Pregibon, D. and Gale, W. A. (1984) REX: an expert system for regression analysis. *COMPSTAT-84*, Czechoslovakia.

Sprent, P. (1970) Some problems of statistical consultancy (with discussion). *Journal of the Royal Statistical Society*, Series A, **133**, 139–65.

Tarter, M. and Berger, B. (1972) On the training and practice of computer science and statistical consultants. *Proceedings of the Computer Science and Statistics Sixth Annual Symposium on the Interface*, American Statistical Association, pp. 16–23.

Watts, D. G. (1970) A program for training statistical consultants. *Technometrics*, **4**, 737–40.

White, S. J. (1979) Statistical errors in papers in 'The British Journal of Psychiatry'. *British Journal of Psychiatry*, **135**, 336–42.

Zahn, D. A. and Isenberg, D. J. (1983) Nonstatistical aspects of statistical consulting. *The American Statistician*, **37**, 297–302.

Zelen, M. (1969) The education of biometricians, *The American Statistician*, **23**, 14–15.