



UNIVERSITY
OF WOLLONGONG
AUSTRALIA

University of Wollongong
Research Online

Applied Statistics Education and Research
Collaboration (ASEARC) - Conference Papers

Faculty of Engineering and Information Sciences

2011

Advice for the potential statistical consultant

Ken Russell

University of Wollongong, kerussell@csu.edu.au

Publication Details

Russell, Kenneth, Advice for the potential statistical consultant, Proceedings of the Fourth Annual ASEARC Conference, 17-18 February 2011, University of Western Sydney, Paramatta, Australia.

Research Online is the open access institutional repository for the University of Wollongong. For further information contact the UOW Library:
research-pubs@uow.edu.au

Advice for the potential statistical consultant

Abstract

Many of those who have never practised statistical consulting seem to view it from one of two extremes: a trivial exercise that any statistician could do, or an extremely difficult task occasioning extreme anxiety. Of course, the truth is that it lies somewhere in between. I will consider various aspects of statistical consulting, including the characteristics of a successful consultation, the necessary knowledge, skills and attributes of the consultant, things to do and not to do, and the pleasures and frustrations that one can gain from consulting.

Keywords

Consulting, knowledge, skills

Publication Details

Russell, Kenneth, Advice for the potential statistical consultant, Proceedings of the Fourth Annual ASEARC Conference, 17-18 February 2011, University of Western Sydney, Paramatta, Australia.

Advice for the potential statistical consultant

K.G. Russell

*University of Wollongong, NSW 2522, AUSTRALIA
Charles Sturt University, Locked Bag 588, Wagga Wagga NSW 2678, AUSTRALIA*

Abstract

Many of those who have never practised statistical consulting seem to view it from one of two extremes: a trivial exercise that any statistician could do, or an extremely difficult task occasioning extreme anxiety. Of course, the truth is that it lies somewhere in between. I will consider various aspects of statistical consulting, including the characteristics of a successful consultation, the necessary knowledge, skills and attributes of the consultant, things to do and not to do, and the pleasures and frustrations that one can gain from consulting.

Key words: Consulting, knowledge, skills

1. Introduction

The provision of statistical consulting is almost as old as statistics itself. Russell [6] pointed out that a very early consultation occurred when the Chevalier de Méré sought statistical (probabilistic) advice from Fermat and Pascal on the ‘problem of points’. Some private companies, research institutions and government departments have employed statisticians for many years. In Universities, the provision of formal statistical consulting has existed for some time. There was a Consulting Service at Victoria University of Wellington, New Zealand prior to 1979, and the University of Wollongong has had a Statistical Consulting Service since 1989. But, especially in Universities, there has been ambivalence about how valuable the consultant is. Advertisements for University consultants frequently offer the position at Lecturer level, or even at Research Fellow level, suggesting that the position is seen as a low-level role requiring not too much statistical maturity. Nothing could be further from the truth. While any statistical knowledge and experience will be useful, a good consultant requires many skills, of which good statistical knowledge and experience are just two. Some good academic statisticians are not suited to consulting: an ability to derive results in an idealized situation is not sufficient when that situation does not prevail.

If you discuss with imminent graduates in Statistics what they will do in the future, you will frequently encounter trepidation at having to deal with ‘real’ problems. When there is no hint in the description of a

problem as to what assumptions can be made or what type of analysis is appropriate, and when one cannot even be confident that the description is correct, a moderate amount of resourcefulness and ingenuity is required. This is the other extreme of the view of statistical consulting.

2. A Successful Consultation

For a successful consultation to occur, both parties must respect the other, and the knowledge, ideas and skills that they possess. They must both be willing to listen and learn, to make active contributions to the discourse, and to accept that neither side has a monopoly on being right.

A successful consultation will leave both the client and the consultant feeling that progress has been made in solving the problem. This is not necessarily the same as both parties feeling happy at the end of the session. There may be frustration that more progress was not made, or that some misunderstanding still remains, but we can’t expect everything in real life to be totally successful. The consultation will have been a failure if, at its end, the client feels that the consultant was not listening, or was not interested in trying to understand the problem, or did not know enough Statistics to solve the problem. (If the last situation is correct, the best that you can do is to express regret that you cannot help the client and redirect him to someone else.) Of course, it may be the consultant who thinks that the consultation has been a failure, because the client did not provide enough information, or wouldn’t answer questions, or . . . ; however, in this relationship, it is the consultant who is the professional (who will be paid if

Email address: kgxr@uow.edu.au (K.G. Russell)

this is a financial transaction), and it is her responsibility to manage the discussion in a way that leads to a satisfactory outcome.

3. Statistical Knowledge

A consultant can never have too much statistical knowledge. This is not to say that it will all be needed, but one can never foretell what a client may ask, and a wide knowledge reduces the risk that one has to devise a new technique. Very few consultations give completely standard problems, and some improvisation will probably be necessary. However, the client, and journal editors and referees, will always be happier if the techniques you suggest come with suitable references. Consult Boen & Zahn [1], or Cabrera & McDougall [2], for lists of the techniques that should be at the forefront of your statistical toolkit.

However, the mere fact that you know a sophisticated method for analysing some data does not justify using it. It is very good practice to use only as much sophistication as is needed. If the client can't understand what you're doing, or why you're doing it, and the referees have never heard of the technique, it is likely that the analysis will be questioned. Indulging in the latest technique you have learnt about (or developed) when something simpler will do, is an unacceptable indulgence for a statistical consultant. (Save such things for your conference or journal publications!)

I believe that you can never know too much experimental design. Comparisons of the means of two, or k , independent populations, and the analysis of two-factor factorial experiments, may well be common, but many clients will present you with data from designs that contain nested factors, factors that are not completely crossed with one another, 'replicates' that are not true replicates, Your ability to modify the analysis produced by a statistical package to reflect accurately what actually happened in the data collection will probably depend on how well you understand the principles of experimental design. You need to be able to determine what the experimental units really are, as opposed to what the client thinks they are. Similarly, if your clients tend to conduct elaborate surveys, a strong knowledge of survey methodology will stand you in good stead. Should you be fortunate to have a client consult you *before* the data are collected, a thorough knowledge of the underlying principles of good design will stand you in good stead. Possibly the most common question that clients ask is 'How large a sample do I need?' Without a good design background, you can't answer this question effectively. You need to understand what a 'control' treatment is, and when and why it is necessary.

4. Computing

It is essential that you be able to use several statistical packages to perform statistical analyses. I do not believe that any one package is the 'best' for every analysis, and you will benefit if you can choose the one that is most appropriate for your particular task. Some packages are good for producing graphics, some shine in particular areas of analysis (for example, I consider GenStat to be the leader in the analysis of designed experiments). You can save time, and frustration, if you can select the most appropriate package for a particular task, rather than be forced to use just one.

I believe that it is also necessary to be able to write computer programs, for use in those situations where a standard analysis will not work. For example, the client may have brought you some data that cannot be transformed to Normality, and cannot be analyzed by a Generalized Linear Model routine. You are unlikely to be able to find an appropriate randomization test in a conventional statistical package, so you will need to write your own program. This may be done in Fortran (perhaps by modifying a paper in Edgington & Onghena [4]) or C++, or you can write a program using R or another malleable statistical package. Without the ability to do this, your capacity to deal with nonstandard statistical analyses will be limited.

5. Interpersonal skills

Perhaps even more important than technical skills are the interpersonal skills that will ensure that a consultation with a client goes well. Without these, the client may be actively resisting your attempts to solve the problem, and you may not even know what the problem really is.

5.1. Communication

An indication of how important communication skills are to statistical consulting is that there exists at least one book on consulting that considers only communication skills. See Der [3], which contains absolutely no discussion of the types of statistical designs, or analyses, that you might need. When you first meet a client, it is absolutely vital that you establish rapport with that person. Boen & Zahn [1] list various categories of clients, some of whom do not have autonomy, or lack confidence in their ability — or your ability. It never hurts to ask about the client's statistical and non-statistical background. Be clear that you don't expect the client to be an expert statistician (and mean this!).

Once (partial) rapport has been established, you will probably need to establish some 'ground rules' for the consultation. These will probably vary according to whether you work in a commercial environment (fees will need to be discussed) or a noncommercial one (are there limits on the time that your client can ask

of you?). These rules established, you can proceed towards finding out what is the problem with which the client requires assistance. Get the client to explain the problem in his own words, take lots of notes (you *won't* remember it all afterwards without notes), ask lots of questions (preferably without interrupting the client's flow of thought), paraphrase what you think the client is saying in as many ways as you can, and don't volunteer a 'solution' until you are confident that you and the client both understand the situation. Never place unquestioning faith in what the client tells you about the problem, especially when he starts to use statistical jargon; always check. (This presents a significant difficulty for beginners, who have to learn to discard their assumption that the 'question' is always correct.)

I thought that I was good at checking that the client and I are 'on the same wavelength', but Russell [6] gives an example where it was just a chance remark by the client that made me realise that the labels we were applying to the units in the experiment did not agree. With another client, I soon learnt that every reference to a 'multivariate analysis' meant multiple regression. Many clients talk about examining correlations when they really want a 'linear regression'. Questions such as 'What do you hope to achieve with this analysis?' may reveal that the client does not mean what he has said.

A client may do a statistical analysis instead of you. You need to know how the data should be entered into the computer package, and a thorough description of how this is done should be given to the client. You should check that this has been understood. I often ask a client to send me a subset of the data (perhaps after the results from just a few experimental units have been entered), so that I can check that this has been done correctly.

5.2. Interest

Most clients are highly intelligent people; they are just not experts in Statistics. They can quickly tell if a consultant is not interested in a problem, or doesn't understand it, or is trying to get through the consultation as fast as possible. This usually guarantees a consultation that does not succeed. The consultant should try to learn something from every client, be curious about the background to the problem, and want to know why the client is doing the research. Such interest shines through, and helps build rapport with the client.

5.3. Commitment

Some academic statisticians have told me that they do consulting, but further discussion reveals that this means that they chat to someone for a few minutes and then turn that person away if the problem is not directly of interest to their research. To my mind, this is not consulting. A real consultant doesn't have the liberty to turn away a client unless the problem is beyond his

or her technical expertise, there is no time available to attempt the problem, or an ethical conflict arises. So true statistical consulting requires a commitment to take on all problems that come, except for the reasons in the preceding sentence.

5.4. Alertness and Common Sense

It is essential that a consultant actively *listens* to what the client is saying. Only by doing this will remarks that seem inconsistent (and which often carry the greatest information) be detected. This is part of being alert. It is also necessary to maintain eye contact with the client or, at least, to watch his or her face. This will enable you to detect when a remark you have made has not been understood, or has caused confusion.

I like to think of Statistics as common sense in the presence of variability. You must have common sense to be able to detect what is achievable when told of the client's aims and ambitions. I have frequently been told 'Wow, that's a great idea!' and have thought to myself 'But it's just common sense.'

6. Dos and Don'ts

Much of this is common sense, or has been covered previously.

Do treat the client with respect. Listen carefully to what he says, clarify anything that you don't understand, and ask questions to check your understanding. Whether providing an experimental design or an analysis of data, offer something that is no more complicated than it needs to be. Explain the assumptions that you are making, check that these seem reasonable to the client, and be sure that you provide an explanation of the conclusions that can be drawn from the data. Check that the client understands these conclusions, by getting him to paraphrase them for you. Offer to read the relevant section(s) of the report or paper that the client is writing, to check that it has been worded satisfactorily. (Statistical jargon can be dangerous for a nonstatistician.) If you feel that your contribution to a paper warrants joint authorship, discuss this with the client *early* in the process. You do a disservice to yourself and the profession if you allow your work not to be acknowledged. Regrettably, some clients will resist this suggestion, even when the paper could not have come into existence without your contribution. But many clients are reasonable, and will accede to a reasonable request.

Do maintain your integrity and ethical standards. Don't accede to requests to perform data analyses that you know are incorrect, or to ignore some factors that are likely to affect the results. If necessary, decline to continue the relationship with the client.

Don't expect more of the client than you expect of yourself. Recognise that he may be under pressure from elsewhere. Don't use excessive jargon in your

speech, and do seek to clarify any jargon that is presented to you.

7. Pleasures and Frustrations

It can be very satisfying to help a client perform an investigation that comes to a successful end, even if perhaps that end is not what the client had hoped for. The client may have gained knowledge that will let him do better on the next investigation. At times, you may see that the results that have been discovered are of benefit to mankind, the Earth, or a subsection thereof. You may meet some wonderfully intelligent or pleasant people. I am proud to have made some good friends from my clients.

It would be naive to suggest that consulting does not bring frustrations. Leaving aside those frustrations that arise because a problem is proving difficult to solve, you will encounter clients, or their problems, that seem uninspiring, there will be personal interactions that are not as satisfactory as you would have hoped, and there may even be occasions when you feel that your time is being wasted. (Perhaps these latter occasions are less common if you are paid for your consultations.)

8. Conclusions

A brief paper such as this cannot do justice to a very broad topic. The references are all worth reading, and offer different perspectives on the many attributes required of a good statistical consultant. Statistical consulting is not a trivial exercise: it requires personal and intellectual skills not always recognised by those who have not practised it. But nor is it impossibly difficult. A willingness to learn, to be flexible in one's thinking, and to believe that good statistical practice will benefit the world, are good attributes to start with.

References

- [1] Boen, J.R. & Zahn, D.A. (1982). *The Human Side of Statistical Consulting*. Belmont, CA: Wadsworth.
- [2] Cabrera, J. & McDougall, A. (2002). *Statistical Consulting*. New York: Springer-Verlag.
- [3] Der, J. (2000). *Statistical Consulting: a guide to effective communication*. Pacific Grove, CA: Brooks/Cole.
- [4] Edgington, E.S. & Onghena, P. (2007). *Randomization Tests* 4th edn. Boca Raton, FL: Chapman & Hall/CRC.
- [5] Hand, D.J. & Everitt, B.S. (1987). *The Statistical Consultant In Action*. Cambridge: Cambridge University Press.
- [6] Russell, K.G. (2001). The teaching of statistical consulting. *J. Appl. Prob.* **38A**, 20–26.