

Interdisciplinary Statistical Research: One Viewpoint

Terry SPEED, *The Walter & Eliza Hall Institute of Medical Research, Australia*, E-mail: terry@stat.Berkeley.EDU

Abstract:

First, let me give my definition of interdisciplinary statistical research. By this I simply mean any research driven by considerations other than statistics, usually answering questions in a given context. Another term would be applied statistics. The strategies that I recognize for getting involved in such research come in two forms, the random and the targeted. Note that like everyone else, I am embedded in the culture in which I live and work. I do not claim that approaches which seem to work in Australia and the US will necessarily work in China, India or Africa, to name three very different research environments. The purpose of these few paragraphs is solely to stimulate discussion, in the hope that ideas which do work in China and elsewhere will emerge. These comments are offered in this spirit.

The random strategy begins with statistical consulting. Do a large amount. Why? A large amount of consulting means meeting many different people and many different problems, necessarily learning a lot of statistics and a little about a lot of other different subjects. The statistical learning includes finding out where we are ignorant, which can lead to research opportunities. So that's three reasons for doing consulting: to enlarge one's statistical repertoire, to learn about different subject areas, and to spot research opportunities. Of course many of us do it because we enjoy it, and feel it is a way of making ourselves useful to the world. Another reason for doing a lot of consulting, in my experience, is that it frequently leads to the possibility of a longer-term collaboration. If you have helped someone in some small part of their work, they may want to broaden that into a collaboration. On your part, you have found out a little about the nature of their research, and a little about them as possible collaborators, so you are in a good position to decide whether or not you want to do more with them, in that or some related area. This is one approach to beginning interdisciplinary statistical research. It is random from the point of view of timing, because you can't really control who comes to you for assistance. But it is not random from the point of view of outcome, since you can decline offers of collaboration, and accept just the ones you like.

The directed approach needs more care, in my experience. Suppose that your local university has some top astrophysicists, and you would like to get involved with them because a) you are interested in the area, b) they are good at it, and you think this could be a fruitful collaboration, and c) you think there are challenging and fulfilling statistical prospects for you. Similar considerations might apply to someone who would like to collaborate with some top AIDS researchers, or a group doing exciting IT research. Why is care needed? Because the directed approach begins from the premise that you want to engage in collaborative research with them in order to achieve your personal goals, not primarily to help them. If that is too evident, it can really get in the way. (Naturally a person doing consulting is also achieving their personal goals, as is a health worker in a developing country, or a priest, but some personal goals involve more altruism than others.) How can it get in the way? There are several issues here, all related to the thought you should assume is uppermost in the mind of the research group with whom you wish to collaborate "What's in it for us?" This does not mean you should assume everyone is selfish, but simply recognizes what is involved in hoping a research group will enlarge to include you, and encourages you to view things from their perspective. Why would they spend time and possibly resources on you, simply to help you achieve your personal goals? An obvious issue is the way in which you become familiar with the elements of the other field, be it astrophysics, AIDS or IT research. Do you have a good understanding of the basics already, do you expect to learn these during the collaboration, and do you expect your collaborators to help you in this task? (I'm assuming that anyone who truly wants to carry out interdisciplinary statistical research recognizes the critical importance of understanding the subject matter of the other discipline.) A second issue concerns the way the research group currently deals with the area in which you might help them. What were they doing with their statistical problems before you appeared on the scene? And how will they continue after you leave, perhaps after you lose interest? Good research groups can usually deal with every problem they meet, including locating or inventing the necessary statistical models and methods as needed, so why should they spend time with you? If you can give their research a competitive advantage, you will be more welcome than otherwise. But even if it does seem like a good idea to them to welcome you, getting to the point where you are entrusted with their data can take tact and patience on your part.

Why am I mentioning these issues? Because I have seen a number of instances in which a statistician decided that they wanted to get involved in interdisciplinary research with a particular group, but went about it in a clumsy way, and got rebuffed, or, more commonly, just didn't get started. The last point is particularly relevant to younger statisticians, as they might look less attractive to a research group in comparison with a more established person. (In truth, the opposite might be the case.) You might hope that the research group of your interest is ready to embrace statistical help as soon as it arrives, but this is not always so. It will not always be best to barge in with relatively little understanding of their subject, saying, in effect, "Here I am. I want to help you and I believe I can do so. Please give me your data, so that I can apply my novel statistical methods to it.

Trust me, this will be good for your research". Even if you have a good understanding of their subject, and plenty to offer, they may have existing collaborations which leave them feeling no need for further collaborators.

Of course this is a highly simplified version of how these interactions might go, but it contains some elements of reality. Strategies which can increase the probability that the direct approach will work include a) your learning a lot about the field on your own before showing your hand, especially becoming familiar in detail with the research of the group you would like to join; b) attending their seminars or other public presentations, before declaring your interest, perhaps asking questions, in essence establishing yourself as one genuinely interested in their research (which I hope is true, see later), and c) offering to carry out some low-level tasks which are within your capacity, and which make few or no demands on them, and involve no future commitment; and, perhaps later, d) asking can you sit in on their group meetings to learn the subject, and see if there are ways in which your knowledge and expertise could play a role.

I may appear to be dwelling too much on getting started. After all, you might have been introduced to the research group of interest to you by a mutual friend or acquaintance, and immediately welcomed into the group, so that most of what I have just said is irrelevant. What next? I'll mention some reasons why you might like to try interdisciplinary research, and then close these remarks with some possible must's and must not's for statisticians involved in interdisciplinary research.

Why engage in interdisciplinary statistical research? Perhaps the reason most of us offer first is that it is interesting, enjoyable and rewarding. If you prefer interacting closely with other people to solitary pursuits, you don't mind ambiguity and the concrete, as distinct from the certain and abstract, if you can appreciate numbers as much as symbols, and you are able to get deeply interested in non-mathematical subjects, then you might find interdisciplinary research is for you. But perhaps the strongest motivation of all is the one I believe: the chance of promising statistical research opportunities arising are much greater for those engaged in interdisciplinary research, in comparison with those seeking to identify and solve research problems solely from within the statistical literature. You have a greater chance of meeting a real problem, one whose solution is of interest to number of people, and meeting your problem in a specific context usually means that there is less competition. There are many fewer statisticians working on problems arising in any specific context than there are working on any given context-free research problem, and so the chances of hitting a home run are correspondingly greater.

What are the "must do" and "must not do" issues here? First and foremost, you must be genuinely interested in the subject matter of the interdisciplinary research. If this isn't true, don't bother trying. Secondly, always remember that your goal must be to do what is good for the research, not what is good for you as a statistician. If there is a simple and transparent approach to a statistical problem in your interdisciplinary research, and you have devised a complicated and opaque one which you think is better, be sure the advantages are real, as assessed from the subject matter perspective, not just from the statistical one. Never forget Einstein's dictum to keep models "as simple as possible, and yet no simpler", or William of Occam's principle that "One should make no more assumptions than needed. When multiple explanations are available for a phenomenon, the simplest version is preferred." An attitude which I have found all too common in interdisciplinary collaborations, is for the statistician to view the collaboration as an opportunity for him or her to "apply" their original theoretical contributions. This might be just what is needed in the context, but frequently it is not. Just as with consulting on a topic with which you are not familiar, you will almost certainly have to learn about some new statistical models or methods, in order to carry out your side of the collaboration. That should be part of its appeal.

Lastly, you must be sensitive to the ways in which many issues - confidentiality, priority, publication, authorship, citation, acknowledgment, and other topics - are dealt with by your collaborators. They will undoubtedly have a different perspective on many, perhaps most of these matters, from that with which you are familiar, and there are many opportunities for confusion, concern or conflict. You should think of yourself as a traveller in a foreign country, one keen to appreciate the best of that country, to learn the language and culture, and hoping to be able leave

something small but good behind you before you return to your own country. Perhaps in time you will be able to pass as a local. Then your interdisciplinary collaboration will clearly have succeeded.