ASA
Statistical Consulting Section

Section Information
- Home
- Charter
- Officers
- Activities
- Minutes
- Newsletter
- Centers & Facilities
- Books & Journals
- Courses

Additional Resources
- ASA Main Page
- Other ASA Sections

ASA Consulting Section

Welcome

Welcome to the website of the ASA section on Statistical Consulting!

Statistical consulting is the most challenging and most rewarding part of statistics. A consultant uses the art and science of statistics to solve a practical problem. Problems come from many different fields, e.g. marketing, product design, manufacturing, medicine, agriculture, or genetics, but many of us specialize in just a few application areas. Some consultants work at universities or medical centers, some work in industry, and many run their own business. Good consulting requires a strong technical background in statistics, good people skills, and for many, a good business sense. Other sections of the ASA help with the technical aspects of statistics. The statistical consulting section is here to help with everything else.

The services we provide (or will provide soon) include:

A series of web-cast seminars on consulting co-sponsored by the section and the ASA committee on applied statistics. Watch the ASA web page and your e-mail inbox for announcements of upcoming seminars.

A consulting referral web page. Section members will be able to post their availability and consulting expertise on a web-accessible searchable data base. This should be open for business sometime this fall.

The new edition of the **Statistical Consulting Newsletter**, edited by Chris Holloman, is here. Previous editions can be found in the Newsletter subsection of this web-site. Every issue has a variety of interesting articles. This issue includes items from JSM 2007 such as the paper on "Undergraduate Consulting as an Introductory Course."

Chris Holloman has taken over the editor position. He is always interested in potential newsletter articles. Send him your comments and ideas.

Your comments, feedback and articles are appreciated.

Brenda Gaydos, 2008 Chair, ASA Statistical Consulting Section

THE STATISTICAL CONSULTANT

Section on Statistical Consulting
Christopher Holloman, Editor; Sarah Butler, Assistant Editor

American Statistical Association
Summer 2008; Volume 25, No. 2

IN THIS ISSUE

In this issue of The Statistical Consultant, *we present two articles on the topic of statistician accreditation. For those interested in further reading, an article on accreditation is featured in the July 2008 edition of* Amstat News. *–Editor*

In the *Encyclopedia of Statistical Sciences* (Kotz *et al*., 2005) Brian Joiner states that a statistical consultant, to be fully effective, should have many diverse skills. Ideally:

- Have a genuine desire to solve real problems and help others to solve problems.
- Be able to help investigators formulate their problem in quantifiable terms.
- Be able to listen carefully and to ask probing questions.
- Have a broad knowledge and true understanding of statistical and scientific methods.
- Be able to adapt existing statistical procedures to novel environments.
- Be able to locate or develop good statistical procedures in a timely fashion.
- Be able to keep abreast of developments in statistics.

- Be willing to meet deadlines, even if it requires substantial extra effort.
- Be able to understand something about the clients' subject matter and speak a bit of the clients' language.
- Be a good teacher—much success in consulting depends on being able to help others understand statistical tools, and their strengths and weaknesses.
- Be willing to settle for a reasonably correct approximate solution, then go on to the next problem.
- Be able to identify important problems (and thus avoid spending too much time on projects of little significance).
- Have the confidence to use as simple a procedure as will get the job done, be it design or analysis.
- Be able to convince others of the validity of a solid solution and see to it that proper action is taken.

- Be able to use computers effectively and direct others in their use.
- Be a good problem solver.
- Be willing to meet clients regularly on their home ground, and take the responsibility to meet and communicate with all members of the working team.
- Be diplomatic and know when to bend, when to stand firm, and how to help smooth conflicts over among other team members.
- Be willing to get some experience in the actual collection of the data.
- Be willing to take the time to check and double-check procedures and results.
- Be able to communicate effectively in writing as well as orally (this often includes helping clients write their reports as well).

- Be able to make a good estimate of how much effort will be required to solve the problem without actually having to solve the problem itself.

Brian Joiner's Consultant's Check List (1979)

- Listen a lot.
- Never (well, almost never) interrupt a client.
- Always (well, almost always) allow the client to interrupt you.
- Ask a lot of questions that begin with phrases like "Let me see if I understand this, …".
- Don't smell: shower daily, wear clean neat clothes, brush teeth, get decayed teeth fixed.
- Take good notes, from the beginning.
- Convey a helpful resourceful attitude.
- Helping the client clarify his/her own goals is often your most important function.
- Keep things simple: don't dazzle your client --- or yourself; use as simple words, simple designs and simple analyses as will suffice.
- Put things in writing, and give copies to the client.

- At the end of each meeting spend a few minutes clarifying what has been decided and who does what next. Put these in writing.
- Know a wide variety of statistical methods (so you'll be less inclined to try to fit the problem to the wrong solution).
- Interact frequently with the client ---- never go off and do a lot of work at the client's expense or on the client's behalf without discussing approaches and intermediate results.
- Make realistic cost estimates and discuss them (cost may be dollars and/or time). *[Note: look forward to Nayak Pollisar's discussion on consulting in private practice.]*
- Find out how important this project is to the client, and how much she/he wants to commit to your efforts.
- Be timely: an approximate answer in a few days is almost always preferable to an "exact" answer some months later.
- Make sure you and the client understand that goals of the project, and whether or not these goals are attainable by the planned course of action. Get the goals agreed to, in writing.

Statistical Consulting Resource list
adapted from STAT 8801 (Univ of Minnesota)

- **Books**
- Boen, James R., and Douglas A. Zahn (1982), **The Human Side of Statistical Consulting**, Lifetime Learning: Belmont, CA.
- Cabrera, J., and McDougall, A., (2002), **Statistical Consulting,** Springer, New York.
- Chatfield, Christopher, (1988), **Problem Solving: A Statistician's Guide** Chapman and Hall, London.
- Derr, J., (2000), **Statistical Consulting: A Guide to Effective Communication**, Duxbury
- Hand, D. J., and Everitt, B. S., (1987), **The Statistical Consultant in Action,** Cambridge University Press.
- Polya, George (1971) **How to solve it: a new aspect of mathematical method**
- Rustagi, Jagdish S. and Douglas A. Wolfe, editors, **Teaching of Statistics and Statistical Consulting**, Academic Press (1982).

- **Printed papers**
- *0. A bibliography and general articles*
- Baskerville, J. C. (1981) ‘A systematic study of the consulting literature as an integral part of applied training in statistics’, **American Statistician**, 35, 121-123.
- McCulloch, Charles E., Boroto, Daniel R., Meeter, Duane, Polland, Ronald and Zahn, Douglas A. (1985), ‘An expanded approach to educating statistical consultants’, **American Statistician**, 39, 159-167.
- Pfannkuch, Maxine, and Chris J. Wild, "Statistical Thinking and Statistical Practice: Themes Gleaned from Professional Statisticians", 15 *Statistical Science* 132 (2000).
- Tweedie, R., "Consulting: Real Problems, Real Interactions, Real Outcomes" with discussion, 13 *Statistical Science* 1 (1998).
- Wild, Chris J., and Maxine Pfannkuch, "Statistical Thinking in Empirical Inquiry" with discussion, 67 *International Statistical Review* 223 (1999)

- *1. Where is statistical consulting done?*
- Boen, James R. (1982), 'A self-supporting University statistical consulting center', **American Statistician**, 36, 321-325.
- Cameron, J. M. (1969), 'The statistical consultant in a scientific laboratory', **Technometrics**, 11, 247-254.
- Daniel, Cuthbert (1969), 'Some general remarks on consulting in statistics', **Technometrics**, 11, 241-245.
- Kirk, Roger E. (1991), 'Statistical consulting in a university: dealing with people and other challenges', **American Statistician**, 45, 28-33.
- Marquardt, Donald W. (1979), 'Statistical consulting in industry', **American Statistician**, 33, 102-107.
- Marquardt, Donald W. (1981), 'Criteria for evaluation the performance of statistical consultants in industry', **American Statistician**, 35, 216-219.
- Meier, P., (1986), 'Damned liars and expert witnesses', **Journal of the American Statistical Association**, 71, 269-276.

- *2. Interpersonal aspects of consulting and roles*.
- Boen, James (1972), 'The teaching of personal interaction in statistical consulting', **American Statistician**, 26, 30-31.
- Boen, James and Fryd, David (1978), 'Six-state transactional analysis in statistical consultation', **American Statistician,** 32, 5840.
- Bross, Irwin D. J. (1974), 'The role of the statistician: Scientist or shoe clerk'**, American Statistician**, 28, 126-f27.
- Hunter, William G., (1981), 'The practice of statistics: The real world is an idea whose time has come', **American Statistician**, 35, 72-76.
- Hyams, Lyon (1971), 'The practical psychology of biostatistical consultation', **Biometrics**, 27, 201-211.
- Zahn, Douglas A. and Isenberg, Daniel J. (1983), 'Nonstatistical aspects of statistical consulting'**, American Statistician**, 37, 297-302.

- *4. Reporting and communication skills*
- Ehrenberg, A. S. C., (1982), 'Writing technical papers or reports', **American Statistician**, 36, 326-329.
- Ehrenberg, A. S. C., (1977), 'Rudiments of numeracy', **Journal of the Royal Statistical Society, A140,** 277-297.
- Hoadley, R. Bruce., and Kettenring, Jon. R., (1990), 'Communication between statisticians and engineers/physical scientists', **Technometrics**, 32, 243-274.
- McDonald Gary C., (1988), 'Communicating with managers', **Chance,** 1, 42-44.

- *5. Professional conduct and ethics*
- American Statistical Association (1999), 'Ethical guidelines for statistical practice', approved by Board of Directors of ASA, Aug 7, 1999. http://www.amstat.org/profession/index.cfm?fuseaction=ethicalstatistics Committee on Professional Ethics, American Statistical Association, http://www.tcnj.edu/~ethcstat/
- Engeman, Richard M. and Shumake, Stephen A. (1983), 'Animal welfare and the statistical consultant', **American Statistician**, 47, 229-233.
- Hooke, Robert (1980), 'Getting people to use statistics properly', **American Statistician**, 34, 102-107.

Tips and Tricks for Statistical Consulting

Donald C. Martin, Ph.D.

I. Some general suggestions
   A. Descriptive statistics
   B. What am I doing here?
   C. Common traps and pitfalls
   D. Warning signals
   E. Second hand advice
   F. Scientific Method
II. Power calculations
III. Transformations

- 1.A.1 Rule: Always look at the descriptive statistics before considering going on to any other analysis.
  - Remember that there are different kinds of descriptive statistics. The descriptive methods that you use for general exploratory data analysis may differ from the best methods for checking model assumptions.
  - Descriptive methods for inclusion in reports or papers are to present some aspect(s) of the data. These require careful design. Ask what is the primary aim of this graph or table.
  - It is not unusual to rewrite parts of the text of a paper many times. Be ready to redo table and graphs as well.

- I.A.2 Try to get a look at the raw data.
  - A few minutes spent on a crude scatter plot or a stem-and-leaf can be very useful.
  - Look at the organization of the data sheets for hints about pairing or blocking. Ask questions!

- I.A.3 In many cases descriptive statistics are all that is needed. Don't feel compelled to do an elaborate analysis if it is not needed.
  - The "interocular test of significance"

- I.A.4 Good descriptive statistics are hard to come by. Don’t feel bad when you have to spend a great deal of time and effort on them.

I.B What am I doing here --- or where am I?

I.B.1 Find out why your client is doing the study.

- Why are they seeing you and what do they expect? Help? A blessing? Magic?
- Don't expect to find out all this at once. It is an iterative process. Good questions:
  - What are the primary goals
  - Which is the most important variable?
  - What resources are available? (including computers, software, and skills using these)
  - How does this analysis (or design) relate to the goals? (This question sounds to supid to ask, but just wait …)

- I.B.2 Hint. Always try to give the client some help that they will consider useful, especially if you can't help them on the problem that they came in with.

- I.B.3 Most clients will try to phrase their problems in terms of hypothesis tests. Be prepared to spend time in educating the client about alternatives (including descriptive statistics).

- I.B.4 Exploration vs. Confirmation. Try to establish the point on this continuum for various parts of the study. Your methods should be different.
  - At the exploratory end, concentrate on power rather than type I errors. You are looking for hints for further research. Use large alphas and ignore multiple decisions (correction for multiple testing).
  - At the confirmatory end use a small alpha; you need a clean design and analysis.

- I.B.5 Mutiple comparison methods are of limited use at either end of the exploration vs. confirmation spectrum. These are probably most useful in arguing that unexpected results found by data mining are really there.

  - (but they are more and more essential in large data mining/genomic/micro-array problems; “False Discover Rate”)

I.C Some common traps and pitfalls

I.C.1 Errors in randomization.

- Most people don't know what randomization is, how to do it, or why they should do it.
- Common use is that if they cannot predict something, then it is random.
- Ask for details, not just if randomization was used.
- Ex. Cross-over clinical trial
- Ex. "fast rats" (iron deficiency and activity)
- Note: subverted randomization is more common than you would think.

I.C.2 Hidden matching

- This is one of the more difficult problems in consulting. Often the investigator has done some type of matching and fails to mention it.
- Ask about possible matching. Look at lab data sheets, etc.
- Handling partial matching can be hard. Hatching is often overused! Matching can have two different aims:
  - Reduction in bias in observational studies
  - Reducation in variance in randomized studies

  These are very different situations

  (Paula Diehr paper on consequences of ignoring matching)

- I.C.3 Hidden common factors
  - Ex. Batch runs in analyses.
- I.C.4 Confusion of independent and dependent variables
  - Ex. Gestational age
- I.C.5 Using a hypothesis test to show that two groups are the same.
  - This is an easier trap to fall into than you might expect. The fix is to use confidence intervals.
  - Ex.

- I.C.6 Confusing statistically significant with clinically significant.
  - Careful editing to be consistent in write-ups.
  - Ex.

- I.C.7 It is not always easy to decide what is practical significance.
  - Ex. 200 gms in birth wt.
  - Ex. R = -0.06 for IQ-prenatal alcohol corr.

I.D Common Warning Signals

- These are only to help get started. You will soon develop your own list for your type of clientel.
- These are warnings and only suggest that trouble may be in the offing.

I.D1 (Phone call) "I have a quick question …"

- See JASA 1957 pp. 133-142, Kimball: "Errors of the third kind in statistical consulting."
- Unless you know both the person and the research, <u>just say no</u>. Get the clients in for a full consulting session.
- (Would you diagnose (prescribe, etc.) by phone without seeing the patient?)

- I.D.2 Watch out for long sequences of analyses that make it hard (perhaps impossible) to relate the final results to the data
  - Ex. Factor analysis -> factor scores -> cluster analysis -> ANOVA -> ?
  - Ex. Scaling prior to PCA of curves

- I.D.3
  - “I know that the analysis is correct. All I want to know is …”
  - “I don’t have time to discuss … Just show me how to …”

- I.D.4 n observations with p variables, p > n.
  - No longer necessarily a problem, but
  - Watch out for t-tests by the thousands and 1000 item questionnaires

- I.D.6
- I.D.7
- I.D.8 Stepwise, branch and bound methods on many variables.
  - David Freedman, 1983, Amer. Stat. 37(2): 152-155.

- **I.E Passing on some good advice:**

- I.E.1 Curley Lucas: When faced with a very complex data set, do something simple.
  - Report the simplest analysis that shows valid results. Do the m ore complex analyses, but don't confuse your reader/client with complex methods if you can avoid doing so.

- 1.E.3 Gary Cox: State the hypotheses in terms of the observed variables.

I.F.1 Scientific Method.

- A statistical consultant can often offer advice on basic applications of the scientific method.
- Many well-trained researchers are too close to their work to see points that seem obvious. Don't hesitate to ask why when you see something that doesn't seem to follow.
- Unfortunately, many clinical researchers have not been trained in the scientific method and are poor at experimental design or the interpretation of data.
- Ex.

NEW SERIES, NO. 173 (VOL. XXVI) MARCH, 1931

JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION

Formerly the Quarterly Publication of the American Statistical Association

STATISTICS AND THE SCIENTIFIC METHOD [1]

By Malcolm C. Rorty

...

This, then, was our pragmatic answer to the question, "What is a statistician?"—the statistician must be instinctively and primarily a logician and a scientist in the broader sense, and only secondarily a user of the specialized statistical techniques. The statistician who knew only statistics was a danger and a blight. It was less important to know how to use statistics, than it was to know how and when *not* to use them. (p. 4)

[1] Presidential address at the Ninety-second Annual Meeting of the American Statistical Association, December 31, 1930.