THE STATISTICAL CONSULTANT

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

American Statistical Association
Spring 2009; Volume 26, No. 1

IN THIS ISSUE

Newsletter Discussion Forums

Christopher Holloman, The Ohio State University[1]

Occasionally, the editors of The Statistical Consultant receive comments from readers about articles we've published. These comments often provide additional insight into an issue, information on an additional resource, or a question that follows up on an article. We would like to make publication of these reader comments a regular feature in the newsletter.

[1] holloman@stat.osu.edu

If you read something and would like to comment or to ask a question for others to consider, please e-mail your thoughts to me or to Sarah Bulter (e-mail addresses are at the end of the newsletter). Please indicate whether you are full or part-time in paid consulting and whether your setting is commercial (industry, retail, and other commercial), government, academic, consulting firm, self-employed, or other (specify). Your response may be edited for clarity or length. Thanks in advance for your contributions.

Special Focus: Statistics and the Internet

Sarah Butler, NERA Economic Consulting[2]

This month's newsletter includes three submissions related to statistics and the internet. For all of us, the internet has had an impact on the way in which we carry out our statistical consulting. For some people, statistics and the internet is the focus of their job (this includes the statisticians who evaluate the traffic to different internet sites and rank the most popular sites, to individuals working at places like Google and Yahoo!), while others use the internet as a way to gather data, share information, and further expand our statistical profile. Whichever end of the spectrum, it is difficult to say that the internet is not a valuable tool for a statistical consultant.

[2] sarah.butler@nera.com

Three members of the statistical consulting section have agreed to write their thoughts on statistics and the internet. First, Neil Glaser points out the advantages of the internet for statistical consultants (including convenient ways to make one's services known) but also the potential pitfalls in working with long-distance clients.

Next, Surendra Prasad Sinha from Venezuela further elaborates on the benefits of the internet for a practicing statistician and educator in his article on constructing a website for a distance course. Surendra offers some practical advice on how best to facilitate interactivity and offers some ideas that may be useful to those of you that maintain or are thinking of designing your own websites.

Finally, Karen J. Shaner further offers some excellent practical advice on tools that can be downloaded from the internet which help to run more sophisticated statistical tests using the familiar Microsoft Office packages.

We hope these brief pieces are useful and of interest.

Consulting and the Internet: Client Relationships and Challenges

D. N. Glaser, Glaser Consulting[3]

It is generally an inarguable point that the internet has made a monolithic change in how we interact with others, transmit (and obtain) information, and process data. With databases that hitherto were inaccessible (or arduous to access), the internet has truly "flattened" the world with respect to sharing data. In this brief article, I will expound on how the internet has impacted the statistical consulting process.

I will preface this article by informing the reader that my training was in the social sciences (specifically psychology), thus the bulk of my consulting is in the health care/social science/educational domain. Hence, I generally work on projects that have a very strong applied orientation/objective.

The word-of-mouth phenomenon is known in the customer satisfaction/marketing research literature as being a prominent factor in generating business and potential customer contacts. One way by which the internet has impacted that relationship, at least with respect to my consulting practice, is maintaining a web-site that delineates my services and background (i.e., publications, presentations/workshops, etc.). Thus, when a potential client is interested in using my services, having the website as a reference point has been invaluable for exposing the array of services I offer. Moreover, by subscribing to various listservs (e.g., sites devoted to structural equation modeling, multilevel modeling, SPSS, etc.) I have been able to offer my services when a general request for a proposal is made. As well, these listservs have been valuable sources for references (e.g., published articles, examples of syntax/programming language such as Raynald Levesques' website at http://www.spsstools.net, etc.) and bouncing ideas/strategies off other experts in the field. I have been able to develop professional relationships that hitherto would have been prohibitive

[3] glaserconsult@sbcglobal.net

given the distance factor. Thus, the internet has been an absolute boon with respect to cultivating relationships.

Though once in awhile I have availed myself of data available to the general public (e.g., census, etc.) most of my data-use via the internet has been by way of long-distance contracts. And this is where some of the challenges have presented themselves. A crucial part of the consulting process is the dyadic relationship; often, nothing beats having a client sit side-by-side as the research question is solidified/modified/corroborated and, with utmost certainty, the analysis corresponds to the overarching research question and hypothesis. Unfortunately, it is not always easy to clarify one's goals via long-distance consulting, or for that matter, to assure the client can intelligibly interpret the output and/or interpretation of the findings. Thus, one challenge in conveying statistical information electronically is assuring that a clear understanding has been obtained. This understanding is unequivocally a two-way street: verifying that I understand the client's objectives and ascertaining that the client has a clear interpretation of the output/interpretation. Communication about complex subject matter can be formidable in the best of circumstances, and even more so when conducted via electronic interface. Thus, it is incumbent that the consultant verify the client's comprehension of both the strategic steps in developing the design/analysis and the attendant results.

In conclusion, the internet has opened up vistas in the context of statistical consulting that would otherwise have been virtually inaccessible. Though the interpersonal element is still a crucial factor in reinforcing the consultant-client relationship, the internet has made distance data-sharing a much more plausible achievement.

Teaching Experimental designs with the Internet: A personal experience

Surendra Prasad Sinha, Universidad de Los Andes[4]

Introduction

The Internet is a very important medium of communication which can provide rapid access to vast amounts of information. Education is now no longer limited to the traditional classrooms, and internet-based teaching has commenced to reshape gradually all possible levels of higher education. The Internet can be used to provide a platform that can enable development of user-friendly interactive visual and computational tools which can be made accessible to students and research workers at a relatively low cost.

This paper provides information about the experiences I obtained in teaching a course of experimental designs for the graduate students of a M.S. program in statistics offered by Instituto de Estadística y Computación of Universidad de Los Andes at Mérida in Venezuela. I constructed and maintain the following web page in the Design and Analysis of experiments:

[4] Sinha32@yahoo.com

http://mipagina.cantv.net/ssinha and its mirror sites, which are http://es.geocities.com/sp_sinha/ and http://webdelprofesor.ula.ve/economia/sinha.

The web page mentioned above was constructed and programmed by me in Spanish language in the year 2002, and since then it has been used by me and others as a teaching aid for the courses taught on Experimental Designs at the undergraduate and graduate (M.S.) level. This web page was prepared primarily for the students of the course on Experimental designs which I used to teach, but it didn't take long before it became known to students and research workers in other Spanish speaking countries in South America and many such users since then have communicated with me through e-mail and consulting about different types of problems of interest for them.

The experiences which I obtained during the design, construction, teaching, feedback and interactions can be summarized as follows:

Constructing the web page

In order to construct a web page, it will be necessary to know exactly about the different types of topics, concepts, graphical illustrations, programs to perform computation, statistical tools and techniques which the course web page will contain. In my opinion the internet should be used as a teaching aid to complement and not to substitute or eliminate the traditional teaching methods. In the case of the course home page on Experimental designs I decided to include in it topics such as: tutorials on basic principles in experimentation, completely randomized, randomized blocks, factorial, hierarchical, and split plot designs. In addition to a section on interpretation of the results obtained by using many examples of the different text books of experimental designs, I also included a collection of SAS source programs which would be needed by the students.

A special topic with a short tutorial on response surface design was also included in the web page. These tutorials were prepared in such a way as to help clarify to the students those concepts, methods, or techniques which were not so easy to understand in the traditional classroom instruction.

Interactivity

A course web page should be designed to maximize the student interaction. JavaScript is a client-side programming language which can be used to provide such interactivity in the web pages. JavaScript is much simpler to program than Java or CGI which are server-side programming. Client-side programming means that the web pages with JavaScript programs will not lose the interactivity even when the web pages are not connected to the internet. JavaScript Coding was used in this web page of Experimental designs to prepare the following programs: analysis of variance for completely randomized, randomized block, and factorial designs; Simple linear regression; and Duncan and Tukey methods of multiple range tests. The user can download these programs and very easily install them in his own computer, and after that he can perform the necessary calculations with his own examples and data in order to understand the different concepts and procedures in a better way.

E-mail

Using e-mail to send messages and attachments is an inexpensive way to establish communication among students and the teacher. The use of e-mail will allow the students to send

homework assignments at any time and very easily to the professor. Online chat is another resource available through internet that will allow a synchronous communication between students and teachers in real-time. However this type of communication may be difficult for students who cannot communicate quickly because of the slow input method and therefore such limitations or exceptions need to be taken into account.

Conclusions

The conclusions in this article are based on the experiences obtained by the author in using an internet web page constructed by him to teach Experimental designs. The internet provides a new dimension to distance learning, which has now become available in higher educational institutions all over the world with benefits and possibilities that can be taken advantage of by constructing and using a web page to teach any major topic in statistics. However the internet should be used as a teaching aid to complement and not to substitute or eliminate the traditional teaching methods since the power and full potential of web-based learning are not yet fully explored. In order to enable the integration of internet based distance learning with the complementary aspects of the traditional classroom teaching it is necessary to include in the web page specifically the discussion of those topics, concepts, and methods which according to the previous teaching experiences of the professor may require a more thorough explanation. Moreover the web page should also contain the required JavaScript programs which can create a learning environment where the students can do the necessary calculations using examples from books or other data sets to understand the important concepts and procedures.

Launching a Wilcoxon Signed Rank Test (and other data analyses) Directly from Excel

Karen J. Shaner, The Counseling Center of Wayne and Holmes Counties[5]

On occasion you may find it beneficial to launch a statistical test directly from Excel. If you suddenly find yourself in need to do a few quick calculations and your data already exists in Excel you are saved the step of exporting to SAS, SPSS or another statistical package; and for those without access to a standalone statistical package it can be an absolute necessity.

Although functions exist in Excel enabling a variety of data analyses, it is by no means a statistical software package, nor was it ever intended to be one. Excel's Data Analysis ToolPak contains limited functionality such as descriptive statistics, *t*-tests, and analysis of variance but is not capable of launching, for example, the Wilcoxon signed rank test.

Each of the products listed below are capable of launching a Wilcoxon signed rank test directly from Excel. Unlike stand alone products, these Excel 'add-ins' allow you to work in the familiar Microsoft Office environment. Using Microsoft Office as in interface, Excel add-ins provide you the tools to perform a wide range of statistical tests with toolbars and menus integrated right into Excel.

[5] kshaner@ccwhc.org

Analyse-it www.analyse-it.com

Based in the United Kingdom (and choosing to use the English spelling) *Analyse-it* was introduced in 1997. Analyse-it integrates into Excel 97 – 2007 and enhances Excel by adding "point and click" functionality for both analyzing data and creating charts, including charts not available in Excel such as box-and-whisker plots. Features include summary statistics, frequency histograms, Kolmogorov-Smirnov, Shapiro-Wilk (supporting up to 5,000 observations), and Anderson-Darling. Parametric procedures include t-tests and ANOVA and non-parametric Mann-Whitney, Kruskal-Wallis, Wilcoxon, Fisher exact, and Chi-Square. In addition, it includes single and multiple linear regression, Polynomial regression , Pearson and Spearman correlation, Kendall concordance, Kappa & Weighted Kappa, and Bland-Altman agreement. Results are formatted into ready to print, easy to interpret reports. Analyse-it was found by the author to be particularly easy to use. It includes an integrated, in-depth help feature as well as technical support. Moderately priced, the Standard Edition is priced around $185, the Method Evaluation Edition $495. A Student Edition is available to full time students at a cost of $30 per year and includes all the functionality of the Standard Edition. A 30-day free trial of the Standard Edition is available for easy download - the only limitation being you cannot save or print resulting reports.

StatTools www.palisade.com/stattools

StatTools is Palisade's statistical analysis toolset. Founded in 1984, Palisade's first product was PRISM, which gave PC users the ability to quantify risk by running Monte Carlo simulations. @RISK is Palisade's best-known product and is a leading risk analysis tool available as an add-in to Microsoft Excel. The latest version of StatTools is fully compatible with Excel 2000 through 2007, allows up to 16,384 variables per data set and offers over 30 statistical procedures including descriptive statistics, normality tests, group comparisons, correlation and regression analysis, and parametric and non-parametric tests. You can drag and drop tables and charts created using StatTools into other applications such as Microsoft Word or PowerPoint. Another feature is "hot-linked" calculations where statistical reports and charts update to reflect the latest data whenever you change a value in your dataset. StatTools offers two editions; priced at $350 and $595. A fully functional, 15-day free trial version is available for download.

XLSTAT www.xlstat.com

A product of Addinsoft, XLSTAT was first made available via the internet in 1995 and is fully compatible with Excel 2007. XLSTAT is relatively easy to use and technical support is available. Functionality includes descriptive statistics such as histograms, box-plots, factor analysis, discriminant analysis, K-means clustering, univariate clustering, linear regression, distribution fitting, ANOVA, ANCOVA, parametric *t*-test and z-test, two sample comparison of variance (Fisher's F, Bartlett's, and Levene's), non-parametric Mann-Whitney, Kolmogorov-Smirnov's, Wilcoxon, Kruskal-Wallis', and Friedman's tests. A fully functional trial version is available for download. It is priced around $50 (student version), $295 (academic), and $445 (company/private). At additional costs are XLSTAT-3DPLOT (3 dimensional data presentation), XLSTAT-Time (time series analysis and forecasting), XLSTAT-PLS (Partial Least Squares regression), and XLSTAT-Pivot (advanced pivot table applications).

Although not an Excel add-in, mention should be made of the impressive MYSTAT stand alone statistical and graphics package (www.systat.com). MYSTAT is a free, downloadable, streamlined version of the powerful SYSTAT product. SYSTAT was developed in the 1970's by Leland Wilkinson who, at that time, was a professor of psychology at the University of Illinois at Chicago. Introduced commercially in 1983, SYSTAT was subsequently sold to SPSS, Inc. Sold later in 2002 it is now marketed by Systat Software, Inc. MYSTAT was found by the author to be one of the best *free* stand alone statistics software packages currently available. MYSTAT allows for up to 100 variables with no limit on the number of rows. Current features include basic stats along with hypothesis testing, parametric and non-parametric tests, correlations, least squares regression, and cluster analysis capabilities. Graphical capabilities include histograms, ANOVA plots, spearman correlation probabilities, linear regression, scatter plots, 3D plots, stacked and horizontal bar charts, and mosaic plots. Technical support is available.

JSM 2009, Washington, DC: Sessions and Roundtables Sponsored by CNSL

Stephan Ogenstad, Statogen Consulting, LLC[6]

This year, our section is sponsoring two invited sessions, a large number of topic contributed sessions and posters, and two roundtable luncheons.

Invited Sessions

- Doubly Robust Estimation for Empirical Policy Making
- Making Research Data Public: Intelligible or Just Available?

Doubly Robust Estimation for Empirical Policy Making
Organized by Eric J. Tchetgen, Harvard University

Description. Double robust estimation is a relatively recent idea which has shown great promise as a partial solution to notoriously hard problems that arise in Epidemiology and biostatistics. For instance, take the classical problem of confounding adjustment in cross-sectional observational studies, there has been a debate in the literature for the past 20 years as to which approach is less sensitive to model misspecification: an estimated propensity-score (ps) approach where confounding is adjusted via using a working regression model of the exposure of interest on potential confounders, or a standard outcome regression (or) approach where the potential confounders are directly entered in a working regression model that includes the exposure, with outcome the disease of interest. Both approaches are susceptible to model misspecification, however the analyst can never know which is more appropriate in a given data set; in this respect, double robust estimation has provided a unifying framework by appropriately combining both strategies into one procedure that yields valid inferences on the exposure effect

[6] sogenstad@Statogen-Consulting.com, http://www.statogen-consulting.com

provided at least one regression (ps or or) , whichever, is correctly modeled. Such doubly robust estimation has recently been shown to exist in a variety of problems ranging from causal inference in longitudinal data, nonignorable-missing data problems, survival data with dependent censoring and instrumental variable methods. The purpose of the session is to give the statistical community a unique opportunity to hear about these methods from some of the most active and creative people in this line of research.

Making Research Data Public: Intelligible or Just Available?
Organized by Linda J. Young, University of Florida

Description. Decision makers need to be able to judge the usefulness of research data and the validity and limitations of conclusions drawn therefrom. If we agree that science is about sunshine rather than darkness, what criteria are needed for determining how and to what extent data, including metadata, should be made available? Once data are publicly available, researchers will use these data for purposes that span from background information to reanalysis to verify findings and conclusions to secondary analysis to aggregation with other data for new investigations. Some research purposes will be investigative; others will be purposeful according to individual researchers' own agendas.

Policies and specific criteria for data availability vary from one field to another and from one application to another. Thoughtful scientists and journal editors from a variety of disciplines, who have been taking a lead in defining precisely what it means making data available means come together as a panel. Their perspectives on their own disciplines also raise issues for statistical metadata as well as for the profession of Statistics. The speakers were some of the participants in the *NISS Explorations Workshop on Data Sharing*. Over-arching statistical principles relating to data availability that emerged from the workshop and from subsequent interchanges will also be presented.

Topic Contributed Sessions

- So You Want to Be an Expert Witness?
- Statistical Consulting with Limited Resources
- Topic Contributed Oral Poster Presentations: Experiences in Consulting at Virgina Tech

So You Want to Be an Expert Witness?
Organized by Christopher H. Schmid, Tufts Medical Center

Description. Expert witnesses testify to inform judges and jury about technical issues arising in legal cases. This testimony can take the form of reports, depositions and court testimony. Typically, the litigants will each hire experts in the hope that those experts will buttress their cases. This introduces an interesting ethical issue for the experts in that they are supposed to provide unbiased assessments to inform the court, but may feel an allegiance also to the side that is compensating them and to which they have often formed an attachment. This panel, made up of statisticians and doctors who have testified as experts as well as a lawyer will discuss the role

of the expert witness, emphasizing the intersection of the roles of science, law and business and the balancing required to satisfy ethical, legal and contractual concerns.

Statistical Consulting with Limited Resources

Organized by Richard F. Ittenbach, Cincinnati Children's Hospital Medical Center

Scale Development on a Limited Budget: Guidelines for Statisticians

Richard F. Ittenbach, Cincinnati Children's Hospital Medical Center; Amy E. Cassedy, Cincinnati Children's Hospital Medical Center; Todd G. Nick, Cincinnati Children's Hospital Medical Center; Justine Shults, The University of Pennsylvania

Description. Increasingly, clients are turning to statisticians with questions about measurement of unobservable events and constructs. Although standards exist for the development of large, standardized instruments and surveys, guidelines for locally derived measures are far more elusive. Helping clients navigate the scale development process with limited time and resources but without a loss of statistical rigor and precision remains challenging at best. The purpose of this paper will be to identify and describe key steps in the scale development process with recommendations as to how statisticians can help their clients in this rapidly evolving area of statistics. Attention will be devoted to the three major theoretical frameworks underlying the scale development process, namely, classicial true-score theory, item-response theory, and rasch measurement theory.

Creative problem solving in statistical consulting with limited time and a limited budget

Mark Glickman, Boston University School of Public Health

Description. Short-term statistical consulting opportunities can often provide enjoyable and challenging work that is qualitatively different from the usual academic pursuits. Unlike the typical research project, however, consulting work tends to be more contained and focused, has the challenge of more urgent deadlines, and often requires creative thinking in order to produce real-time results. In this talk, I discuss general strategies in statistical problem solving that lead to successful and implementable solutions given time and budget constraints that often accompany a consulting arrangement. Specifically, I describe how making various approximations to conventional statistical approaches can simplify the consultant's work, as well as permit solutions that can be applied by the client themselves. I illustrate this general approach with examples from previous statistical consulting projects.

Grant Proposal Development Assistance on a Limited Budget: Guidelines for Statisticians

Todd G. Nick, Cincinnati Children's Hospital Medical Center; Ralph G. O'Brien, Case Western Reserve University

Description. Statisticians working with clients are often asked to assist with grant proposals in developing data management, sample size determinations, and data analysis sections. Ideally, there would be mechanisms in place to provide adequate funding that would provide ample time to develop sophisticated sections. However, in reality, statisticians are bombarded by protocol development with inadequate support. These conditions often make for an unsuccessful client-consultant relationship. Therefore, it is prudent for the statistician to conduct a screen, as

suggested by Parker (2000), as to whether the client should have sample-size and analysis plans that require involvement that are (1) in-depth,(2) moderate, (3) brief. The purpose of this talk will be to describe the necessary steps when screening proposals and to provide a grant development checklist for different levels of involvement.

Developing an Independent Research Program on a Limited Budget and with Limited Time: Some Ideas for Biostatisticians

Justine Shults, The University of Pennsylvania; Richard F. Ittenbach, Cincinnati Children's Hospital Medical Center; Todd G. Nick, Cincinnati Children's Hospital Medical Center

Description. The field of statistics is, by nature, collaborative and many statisticians are funded to work with scientists who specialize in other fields of study. Biostatisticians are often valued for their role as co-investigators on medical studies, or as teachers and mentors of non-statisticians, and, to a lesser extent, for the work that they do as independent researchers and teachers of future statisticians. In this presentation we will discuss why it is valuable to maintain an independent line of research, in spite of pressures to focus primarily on the immediate needs of collaborators' projects. We will also present ideas regarding opportunities for independent research. Our belief is that although collaborations can slow your rate of publication in Statistics journals, they can also enhance your own work, so that what you do publish is well motivated and useful.

Experiences in Consulting at Virginia Tech

Organized by Laura J. Freeman, Virginia Polytechnic Institute and State University

A Comparative Statistical Analysis for Forest Service Trail Tread Width

Laura J. Freeman, Virginia Polytechnic Institute and State University

Description. Forest managers for the National Forest Service face an ongoing management issue of balancing recreation in national parks and forests while preserving the natural environment. Trail-usage in national parks and forests is continually on the rise. By reducing off-trail usage forest managers can encourage recreation in national parks and forests while minimizing the impacts of recreation on the environment. This study looks at how forest managers can affect trail tread width through strategic trail planning and implementation. Three separate statistical approaches are used to determine the affects of trail placement and maintenance on trail tread width. The three statistical approaches are a classical maximum likelihood approach, a Bayesian approach and a Bayes Linear approach. A comparative analysis between the three statistical approaches is provided.

Using Meteorological Data To Predict Ozone Levels in Rural Areas

John Szarka, Virginia Polytechnic Institute and State University; Nels Johnson, Virginia Polytechnic Institute and State University

Description. In recent years, crops such as soybeans and corn have recorded losses in the Midwestern United States. It has been documented that a decrease in crop yield often occurs when surface ozone values reach a threshold of 40 parts per billion (ppb), which has occurred

frequently in recent years. The EPA provides ozone values, however many of their stations are located in metropolitan areas, making it difficult to predict in rural areas. There is believed to be a relationship between temperature, precipitation, and wind speed, and ozone levels. Such meteorological data was collected by the NOAA at a separate set of weather stations spaced more evenly across the area of interest. We will attempt to leverage the NOAA data to improve predictions of ozone values in these rural areas, where those in the agriculture industry would benefit the most.

Predicting Mortality in Infants
Mark Seiss, Virginia Polytechnic Institute and State University

Description. 57 babies with congenital diaphragmatic hernia (CDH) before birth had a measurement of their lung size taken called lung to head ratio (LHR). Ideally, this measurement is taken between 20 and 29 weeks of gestation. Doctors would like to use the LHR measurement to predict the probability of death or the use of a heart lung bypass procedure called ECMO that usually leads to death. Ultimately, doctors would use the predicted probability to determine if experimental fetal intervention is necessary as well as location of delivery and allocation of hospital resources. This paper will describe three approaches that were used to predict this probability of death or ECMO using the clinical outcomes of the 57 babies, some with multiple measurements: a logistic regression mixed model, a two step model that first predicts lung size at full term (39 weeks) and then the probability of death or...

Ecological Application of Change Point Models
Jonathan Duggins, Virginia Polytechnic Institute and State University; Dong-Yun Kim, Virginia Polytechnic Institute and State University; Sue Newman, South Florida Water Management District; Eric Smith, Virginia Polytechnic Institute and State University

Description. Changepoint models are gaining popularity in ecological and environmental settings due to their ability to model thresholds and abrupt changes. While a variety of changepoint models are available, there are few studies evaluating them. Two such techniques, one based on a two-sample t-test and one based on reduction in model deviance, are discussed and their limited usefulness is investigated through simulation studies. Two competing methods are then explored via simulations: Horvath's likelihood method and a proposed bootstrap based technique. Both methods are shown to obtain close to nominal type I error rates while still having reasonable power over a range of scenarios. While Horvath's method is likelihood based, the bootstrap method is competitive in the majority of the considered cases and is superior in some. The methods are then compared on soil chemistry data from the Everglades.

Survey Design for Evaluating the Role of International Students in Domestic Engineering Graduate Programs
Anne Ryan, Virginia Polytechnic Institute and State University

Description. During the planning stages of research, a client from aerospace engineering needed assistance in creating a survey to obtain data relating to her two research questions: 1. How does the presence of a high percentage of international students influence the culture of graduate-level engineering education communities, particularly related to identity, socialization and ultimately

retention of domestic graduate students? 2. To what extent are similar perceptions present among engineering undergraduates, and how do these contribute to enrollment of domestic students into graduate programs? The poster will reveal recommendations for survey creation directly relating to collecting data which will answer the client's research questions.

Contributed Oral Poster Presentations

Experimental Design and Statistical Consulting
Chair: Lara Schmidt, RAND Corporation

Testing for co-directional and anti-directional interactions using union-intersection and intersection-union methods
Julia L. Sharp and Patrick D. Gerard, Clemson University

Description. When significant interaction terms exist in a two-factor, factorial experiment, the consideration and analysis of main effects is often restricted to those situations where the interaction between factors is not significant. Hinkelman and Kempthorne (1994) softened that stance somewhat and advocate testing main effects when the interaction is deemed co-directional but not anti-directional. A test for the main effects may be pragmatic to the practitioner and instructor. Intersection-union and union-intersection methods are examined for assessing the directional nature of significant interactions so that the main effects in a two-factor factorial may be evaluated. The methods suggested are conceptually straightforward and practical and are illustrated through example.

Testing Equal Protection of the Law: Unequal Quanta of Evidence in Drunk Driving Prosecutions in New Jersey
Ayako Kambara, Applied Statistical Consulting, LLC

Description. The purpose of this paper is to test empirically whether or not the constitutional right to equal protection is ensured in New Jersey drunk driving prosecutions. The data used in this paper were obtained from DUI litigation in the New Jersey Supreme Court (State of New Jersey v. Chun, et al., 194 N.J. 54 (2008)). Multiple regression models are estimated using defendants' biological characteristics and police departments as independent variables. Statistical analyses reject the hypothesis of equal quanta of evidence. Without proper adjustments in the alcohol detection instrument, the final Breath Alcohol Concentration can vary according to which police department the defendant was taken to. The final evidence used in court can vary by regional police departments when holding defendants' biological characteristics constant.

A Student Perspective: My Learning Experience in a University Statistical Consulting Center
Whitney B. Miner, Grand Valley State University; Phyllis J. Curtiss, Grand Valley State University

Description. The Statistical Consulting Center (SCC) at Grand Valley State University services faculty, students, and non-profit organizations. The Center employs undergraduate and graduate students each semester under the supervision of the Director. The SCC provides statistics students with an opportunity to gain real life experience with statistical consulting. I have been

working in the Consulting Center for the past year. My work in the Center started as an undergraduate statistics major in April of 2008. I proceeded to do an internship in the SCC during my final undergraduate semester in the fall of 2008. Currently, I am a biostatistics master's student at GVSU and have continued working in the SCC. This presentation will describe my experiences including responsibilities, skills utilized, and knowledge gained.

A case study of fitting and comparing nonlinear parameters for predicting poultry contamination
Andy Mauromoustakos, University of Arkansas; Min Li, University of Arkansas

Description. Contamination of Listeria monocytogenes in ready-to-eat (RTE) poultry products poses potential risk of listeriosis to the public health. To control the level of Listeria contamination, attention has been focused on the postpackage pasteurization of fully-cooked poultry products. In this case study, we sought to develop a model to predict the thermal inactivation of L. monocytogenes in chicken drumettes during postpackage hot water pasteurization. Fully-cooked chicken drumettes inoculated with Listeria innocua were vacuum-packaged and treated in a hot water bath at 60, 70, 80 and 90ºC for different heating times at each temperature. Weibull model was used to fit the survival curves of L. innocua at each heating temperature using JMP and comparison of the important nonlinear parameters will also be discussed.

Roundtables with Lunch
Organized by Walter Ambrosius, Wake Forest University School of Medicine

The Current State and Future Role of Pro Bono Statistical Consulting
Christopher Holloman, The Ohio State University

Description. As more and more quantitative data is collected in industrial and other research endeavors, demand for statistical consulting services continues to increase. However, some clients are unable or unwilling to pay for these services. Consequently, statisticians must determine when it is appropriate to provide their services pro bono. To date, the ASA has not established guidelines for pro bono statistical consulting, although it has provided grants for some pro bono work. During this luncheon, we will consider the models that other professions, especially the legal profession, have used to guide pro bono service. We will discuss ways to motivate statisticians to offer their services pro bono, barriers to pro bono service, potential guidelines for determining eligibility for pro bono services, and the potential consequences of offering or refusing to offer services pro bono.

Setting Consulting Rates
Walter Ambrosius, Wake Forest University School of Medicine

Description. One issue faced by all fee-for-service consultants is setting of rates. This issue is one faced by all consultants but there is very little information available about how to do this well. During this luncheon, we will discuss several approaches to setting of fees including on an hourly basis and on a per-project basis. In particular, we will discuss the formula used at one academic center that is used annually to adjust rates for increases in salaries and other costs.

Academic Statistical Consulting Centers Serving Commercial Clients
Sam Woolford, Bentley University

Description. This roundtable discussion will focus on academic statistical consulting centers that provide contract services (for a fee) to commercial clients. Many of the issues faced by academic statistical consulting centers which provide services to (often internal) academic clients may need to be approached differently when the objective is to provide services to commercial clients. Based on the participants' interests, this discussion will consider issues such as marketing, compensation, staffing, student involvement, pricing, competition, billing and collections, legal requirements and institutional arrangements.

Indemnification for Consultants and DSMB Members: Who Protects Whom?
Janet Wittes, Statistics Collaborative

Description. Many consultants and DSMB members agree to indemnify (i.e., insure) organizations hiring them in case of a third-party lawsuit. Understanding the risk that such a clause entails, the careful consultant can often persuade the company to change the contractual language. The US government's policy, however, provides no legal protection for DSMB members. At this roundtable we will discuss the risks entailed by agreeing to indemnify the sponsor, by having no indemnification clause, and by agreeing to cross-indemnification. We will review suggested language for indemnification clauses in consulting and in DSMB contracts. We will address the special case of not-for-profits. If time allows, we will discuss more generally clauses that limit liability for consultants, parsing legal language that may sound reasonable but that actually places the consultant at considerable risk.

Winners of the 2009 Travel Awards

Stephan Ogenstad, Statogen Consulting, LLC[7]

Congratulations to the winners of the 2009 travel awards to the JSM in Washington, DC!
Each winner will receive a waiver/reimbursement for their JSM member registration, and a $500 check to cover travel expenses, presented at the section business meeting.

The winners are:

Christopher H. Schmid, Tufts Medical Center
Organizing “So You Want to Be an Expert Witness?”

Richard F. Ittenbach, Cincinnati Children's Hospital Medical Center
Organizing “Statistical Consulting with Limited Resources”

---

[7] sogenstad@Statogen-Consulting.com, http://www.statogen-consulting.com

Comments from the Chair

Bruce Craig, Purdue University[8]

Welcome to 2009!! Since this is the first newsletter of the year, I want to extend a belated thanks and recognize those who finished their terms on the Executive Committee this past year. They are:

Stuart Ganksy, University of California at San Francisco, Chair-Elect/Chair/Past Chair
Edward Rothman, University of Michigan, At Large Member of the Executive Committee
Paul Stewart, University of North Carolina, Publications Officer
N. Shirlene Pearson, Southern Methodist University, Section Council Representative

I want to extend a special thanks to Stuart who helped me step in as Chair for Brenda Gaydos this past JSM when she became ill (she's doing much better now) and could not attend the meeting. Not exactly how I hoped to prepare for the role of Chair but Stuart, the rest of the Executive Committee, and those in attendance at our open business meeting were very understanding and patient.

To update you all on our current Section initiatives, the Statistical Consultant eDirectory is up and running. If you'd like to add yourself to this list, the procedure is described at the end of the Statistical Consultant newsletter Vol. 25 No. 1. As far as the initiatives started last year, we're still working on a collection of informational resources for statisticians in private consulting. If you have experience in private consulting and can volunteer some time to help get this material together, please contact me or another member of the Executive Committee. At the last open business meeting, several members wanted to make sure this collection contained information regarding liability insurance. The members also discussed the possibility of again running some phone conferences / webinars on this and other subjects this year. Again, if you have interest in this or another topic, or have the time to lead such a meeting, please contact a member of the Executive Committee. The final initiative was creating a Frequently Asked Questions document for the web site. While we've collected a few questions, we're still hoping to add more so keep those comments/question coming.

Notes from the Editor

Christopher Holloman, The Ohio State University[9]

Thanks as always to our contributors! We appreciate the time you took to perform your research and to share your ideas with the section.

This year, I will continue to be the editor of the newsletter and Sarah will continue to be the assistant editor, so send ideas and comments to either of us.

---

[8] bacraig@purdue.edu
[9] holloman@stat.osu.edu

Christopher Holloman, Editor
Contact Information: The Ohio State University
Cockins Hall
1958 Neil Ave.
Columbus, OH 43210
Phone: 614-292-0738
holloman@stat.osu.edu

Sarah Butler, Assistant Editor
Contact Information: NERA
One Front Street
26th Floor
San Francisco, CA 94111
Phone: 415-291-1022
sarah.butler@nera.com

Remember the
Section on Statistical Consulting
website address:

www.amstat.org/sections/cnsl/

You'll find all information relating to the Section on Statistical Consulting, including our charter, officer list, section activities, past issues of *The Statistical Consultant*, minutes of past meetings, and more.