Innovative Approaches to Undergraduate Mathematics Courses Beyond Calculus

Richard J. Maher, Editor
Loyola University Chicago

Published and Distributed by
The Mathematical Association of America

published
September 2005

ISBN 0 883 851776

3.1

The Importance of Projects in Applied Statistics Courses

Tim O'Brien
Loyola University Chicago

3.1.1 Introduction

While in the past statistics courses may have emphasized formulae and summarizing data, the focus today is more on the importance of statistics in answering researchers' queries by obtaining essential information. As a result, today's students see statistics as an aid to the research process. For example, statistical methods associated with the field of bioinformatics have come into prominence over the past decade to provide biomedical researchers with the statistical tools necessary to detect patterns in very large genetic data sets similar to those resulting from the U.S. Human Genome Project. Modern statistics courses stress both the practical applications of statistical methods and the active participation of students in the learning process. Undergraduate and graduate programs in statistics, such as the ones recently revised at Loyola University Chicago, typically emphasize both the statistical applications in coursework and the involvement of students in the statistical consulting activities of faculty members. One tangible result has been the development of a sense of confidence on the part of students when tackling methodological challenges that go beyond the classical course in introductory statistics.

Like many U.S. universities, faculty at Loyola are involved in teaching an array of courses in theoretical and applied statistics. These include introductory courses for less technically oriented students and more mainstream undergraduate courses for statistics and biostatistics majors and minors. Many of these courses require student participation through projects, papers and/or class presentations. The focus of this article is to discuss some of the ways in which these activities help students understand the usefulness of statistical tools in a broad spectrum of fields. Specifically, the focus here is on the use of projects in Loyola's one-year biostatistics sequence (which is typically attended by premed students majoring in biology) and in follow-up courses in applied regression analysis, categorical data analysis, experimental design, statistical software packages, nonlinear modelling, and optimal experimental design. We provide several examples that illustrate how these projects have become an invaluable tool in the teaching of applied statistics and biostatistics courses and how they provide students with increased confidence by allowing them to obtain the tools necessary to master sophisticated statistical techniques.

3.1.2 An Introductory Biostatistics Course

The public understandably can become confused when studies such as the one described in Pope *et al.* [23] report an association between sulfur-oxide pollution and both lung cancer and cardiopulmonary mortality while other studies find no such association. Modern students of biostatistics are taught to critically examine the underlying statistical techniques used in these studies before considering the conclusions reached. Thus the fact that Pope *et al.* [23] adjusts for study biases by using an extension to the commonly-used Cox proportional hazards statistical survival model (Zar, [28]) to adjust for dependent observations lends significant credence to the paper's findings. Students now understand that merely applying a statistical technique without regard for necessary assumptions can easily lead researchers to dubious or incorrect conclusions.

In order to provide introductory biostatistics students with basic statistical tools, these courses typically cover an introduction to probability (including an appreciation of the relevance of the Central Limit Theorem), regression and correlation, one- and two-sample t-tests and generalizations to single- and multi-factor analysis of variance (ANOVA) and covariance, and an introduction to categorical data analysis that includes basic chi-square tests. Instructors in these courses usually find that they do not have enough time to cover very relevant intermediate-level topics such as odds-ratios and relative risk, logistic and non-linear regression, non-parametric methods, baseline-category logits, the proportional odds model, etc. As a result, introductory students often are left unaware not only of some of the limitations of and assumptions underlying introductory statistical methods but also of the relevant extensions of these methods provided in more intermediate-level courses. Equally disturbing is the (not infrequent) situation in which courses in statistical methods are taught by non-statisticians (such as biologists or psychologists), who may be unaware of these limitations and assumptions. For example, Samuels and Witmer [24] provide a 4×2 table of data relating pain relief (with levels none, some, substantial, and complete) to treatment received by the patient (drug A versus drug B) with the caveat that the usual chi-square test is inappropriate in this situation. Yet it is unclear how many non-statisticians would know the correct statistical technique to analyze these data. Statistical educators must do a better job of helping students develop skills beyond the level of the usual introductory course. We have found that class projects provide an excellent opportunity to do just that.

Loyola's introductory undergraduate biostatistics course is offered through the Department of Mathematics and Statistics and cross-listed in the Biology Department. The course has been redesigned with the idea of developing a student's critical eye through the use of course projects. This course is called STAT/BIOL 335 and usually enrolls about 40 to 50 students per semester. The course syllabus can be viewed on the Web at

`http://www.math.luc.edu/~tobrien/courses/stat335/Fall-2003/syllabus.html.`

Instructors of this course invariably point out the advantages and limitations of each of the techniques discussed in the course. As a result, students learn to be somewhat skeptical of studies employing dubious assumptions or involving only ten subjects in each treatment arm. Since time constraints do not permit exposure to many of the intermediate topics mentioned above, we have found that these subjects are well suited for students' class projects.

Class Projects

For their projects, students are given the choice of either (1) analyzing a sufficiently rich data set and writing up and presenting their findings to the class or (2) critiquing the statistical aspects of two research articles of their choosing from professional journals such as *The Lancet*, *New England Journal of Medicine*, *AIDS Journal, Ecology*, and the like. The only requirement of the data set or research articles is that the statistical methods involved in the data set or article go beyond the level of the class, including topics from such fields as survival analysis, nonlinear or logistic regression, and repeated measures. In order to find a relevant data set or collection of articles, students first identify research fields of interest to them. This usually requires several meetings between the instructor and each student in order to get the students focused on a data set or articles at an appropriate level. Students then are required to obtain the tools needed to critique or implement the given statistical techniques by consulting intermediate biostatistics textbooks such as Zar [28] or the *Encyclopedia of Biostatistics* (Armitage and Colton [2]). Ultimately, students are required to write up their findings or critique in a three to five page paper.

We have found that in the early stages of a project instructors act as mentors only to the extent of helping students to identify relevant research articles or data sets or to find resources such as textbooks where the students can then learn about the statistical techniques employed in their articles or needed to analyze their data. After this initial phase, students learn via self-study the relevant statistical methodologies so that they can then critique the use of these methodologies in the chosen research articles. For example, a student may need to read through a textbook's chapters on bioassay analysis and mixed nonlinear regression. In student evaluations, students have commented how this latter phase has helped them develop confidence in their ability to further their own formal and informal study of intermediate and advanced statistical methods.

In their class papers, students first give a short summary of the goal of the research and the hypotheses studied in each article or provide the important characteristics of the data set they have analyzed. They then focus on the mechanics of the new statistical tools encountered and on the adequacy of their implementation in the article or data set. Students' projects and papers are then evaluated on their criticism of the statistical techniques employed in the two articles, their understanding of these techniques as evidenced by their criticism, and their writing style, including grammar and punctuation. Since undergraduate students rarely have access to rich data sets, our STAT 335 students almost invariably choose to critique research articles. Yet when the author offered a similar introductory biostatistics course to graduate students at Loyola's Medical School during the Spring, 1999 semester, the opposite situation occurred. These latter students opted to use sophisticated statistical techniques on the data sets provided by their research advisors. The class project helped show these students the usefulness of applying intermediate statistical techniques in their own research.

To illustrate some of the mechanics of the class project for this introductory biostatistics course, we now describe the projects of two undergraduate Biology students, Jennifer Huston and Nick Moisan.

Example: Jennifer Huston

For her class project and paper, Jennifer Huston critiqued the statistical techniques used in the article by Walrand *et al.* [27], which investigates the relationship between age and the ability to renourish the body, and by Lau *et al.* [17], which examines the relevance of mite and cat allergen exposure for the development of childhood asthma. The former article used a two-way ANOVA design and analysis for some response variables and a principal components analysis (a data-reduction technique) for other variables for a small rat study. Thus, as was the case for all students, it was necessary for Jennifer to learn the necessary assumptions required to use these advanced techniques. In her paper, she correctly pointed out the limitations (such as the inherent assumption of normality) and potential biases in their application in a study involving only 6 rats in one of the study groups.

For her second article, one of the response variables was ordinal in nature ("current wheeze," "wheeze ever," and "doctors' diagnosis of asthma"), so a multiple logistic proportional odds model regression was used. Once again Jennifer focused her comments on potential shortcomings of the use of this model in light of influential observations and potential outliers. Jennifer's project thus showed her the usefulness of statistical techniques in medical research and underscored the often-overlooked underlying model assumptions. Jennifer also mastered some rather sophisticated statistical techniques. In her course evaluation she mentioned that she now felt confident that by knowing where to obtain the necessary resources she could teach herself the statistical methodology useful in analyzing many medical studies.

Example: Nick Moisan

While the articles critiqued by Jennifer were relatively similar and typical of clinical and pre-clinical research, the articles examined by Nick Moisan provided a study in contrasts. Nick's first article, Krabbendam *et al* [16], which appeared in the *Journal of Neuropsychiatry*, examined whether a relationship exists between deficits in cognitive processing and the temporal and limbic volumes in the brains of humans. Typical of many articles appearing in psychology-related journals, this study based its conclusions on a small number of subjects and used a multivariate analysis of variance (MANOVA) design and several MANOVA analyses. In contrast, Nick's second article, Kernan *et al* [15], used a case-control study to test for a link between phenylpropanolamine (present in cough and cold remedies) and hemorrhagic stroke. This latter article appeared in the *New England Journal of Medicine*, and (as is often the case in medical research) employed logistic regression and odds ratios to draw inferences. Its conclusions were based on a very large study enrolling over 2100 subjects randomly selected throughout the United States. As a result of this project, Nick was able to see the wide application of statistical methods in diverse settings and to understand some of the subtle distinctions in the level of statistical sophistication in disciplines such as psychiatry and medicine. As was the case with Jennifer, Nick developed a great deal of confidence in being able to teach himself complex statistical techniques. Not surprisingly, both of these students continued their studies in applied statistics by taking additional statistics courses offered by our department. In addition, several Biology students with similar interests have decided to pursue Loyola's new minor in Biostatistics.

Both of these examples illustrate some of the benefits of using projects and papers in introductory biostatistics courses. Students invariably observe that statistical techniques are misused in otherwise prestigious research journals. Furthermore, these students also develop an important level of confidence in their ability to understand the necessary requirements and assumptions for statistical tests. They also learn effective communication skills through their written papers and/or their class presentations. Finally, they gain a sense of independence and confidence in their ability to locate resources, both on-line and in the university library, which further their understanding of advanced statistical techniques.

3.1.3 An Advanced Biostatistics Course

Responding to requests from introductory biostatistics students to offer a follow-up biostatistics course, Loyola's Department of Mathematics and Statistics offered an advanced-level biostatistics course during the Spring 2001 semester. This course focused on many of the theoretical and methodological aspects of the statistical techniques highlighted in the introductory projects and papers. These included the statistical techniques used in survival analysis, nonlinear and generalized linear regression, and clinical trials. The current class syllabus, notes and assignments can be viewed on the Web at

`http://www.math.luc.edu/~tobrien/courses/ab/course-homepage.html`

This course, also cross-listed with Loyola's Department of Biology and called STAT/BIOL 336, was attended by twelve students, six from the Biology Department and six from Math & Statistics, and presented the additional challenge of structuring a new course for a rather diverse group. The course was taught using class-notes based on material from Agresti [1], Bates and Watts [4], Davidian and Giltinan [7], Littell *et al* [18], McCullagh and Nelder [20], Pinheiro and Bates [22], Stokes *et al* [25], Venables and Ripley [26], Zar [28]. Material from Ewens and Grant [9], Harrell [13] and Johnson and Wichern [14] was added for the Spring 2003 version of the course. The Minitab®, SAS® and S-Plus® statistical computer packages were used in class handouts and by students in their assignments.

Class projects also were used in this course but, since the class-size was small, students were required to work in pairs in order to foster interpersonal communication skills. The student population was quite diverse in terms of mathematical sophistication, so these teams paired one quantitative student and one biology student. As in the introductory biostatistics course, class projects were used in this course to stretch students beyond the level of the course. With such a small group of students, each of the six student-pairs was required to make a presentation to the class. Mentoring and assessment followed along the lines of the project for the introductory course, but part of the advanced class project grade also reflected the quality and clearness of the class presentation. The following example, which focuses on the statistical detection of the interaction of anti-HIV drugs, is typical of the projects from this course.

Example: Mike Evans and Bahram Patel

Since they were interested in HIV research, Mike Evans and Bahram Patel chose to examine the data provided in Machado and Robinson ([19], p 2304) to test the synergistic or antagonistic nature existing between the anti-HIV drugs AZT (Zidovudine) and ddI (Videx). For this study, the amount of the HIV-1 (strain LAV-1) virus present was measured by reverse transcriptase (RT) activity. With the help of the course instructor, these students used the SAS® statistical software package in a somewhat novel fashion to fit to these data the 5-parameter log-logistic dose-response (nonlinear) model

$$\eta = \frac{\theta_1}{1 + (\frac{z}{\theta_2})^{\theta_3}}$$

where η = E(RT) is the expected amount of RT and where

$$z = AZT + \theta_4 ddI + \theta_5\sqrt{\theta_4 * AZT * ddI}$$

is the effective dose of the anti-HIV drug. For this study, the key model parameter is θ_5, the so-called "coefficient of synergy." For this parameter, statistically significant negative values indicate antagonism and significant positive values indicate synergy of the study drugs. The instructor pointed out to these students that this model was applied in Gerig *et al.* [10] to the detection of the antagonistic joint action of similar compounds in the growth of cucumber seedlings and the model assumptions were subsequently verified for Gerig's cucumber data. When this model was fit to the HIV data, Mike and Bahram's preliminary results failed to detect significant synergy between these two drugs, a result which would then lead some researchers to conclude that the two drugs act independently. But upon inspection of the model residuals, the students noted that the variability in RT tended to decrease with η, and thus that one of the key assumptions (equal variances) was violated. When this heterogeneity of variance was then incorporated into the model by letting the variance be of the form $\theta_6\eta^{\theta_7}$, the estimate of the coefficient of synergy became significantly positive. This led Nick and Bahram to conclude correctly that the two HIV drugs do indeed enhance one another.

Thus, even though the statistical methods employed in this project were rather sophisticated, Mike and Bahram observed first-hand the importance of checking the underlying model assumptions. They also learned something about the wide applicability of nonlinear models in biomathematical modelling. These students also correctly pointed out that the findings of the above study have led biologists to study the physiological mechanisms involved, which in turn would explain the synergistic effects of these two commonly used HIV drugs. Since modelling both the mean and the variance is becoming more and more necessary in bioassay studies, this example provided all the Advanced Biostatistics students with additional statistical tools.

It is worth noting that we have used class projects in the introductory biostatistics course to provide students with confidence in their ability to obtain the necessary knowledge and skills to criticize the application of statistical techniques performed by biomedical researchers. But we also have used class projects in more advanced courses to analyze medical data in a more sophisticated and novel manner so as to better test researchers' hypotheses.

3.1.4 Subsequent Applied Statistics Courses

In addition to the introductory and advanced biostatistics courses, we also have found that class projects are a useful tool in other intermediate and advanced applied statistics courses, highlighting the usefulness of projects for both non-major and major statistics courses. For example, we have found that courses in categorical data analysis (CDA), applied regression, experimental design, and statistical software packages can benefit from class projects, papers, and presentations as well. The following examples, which involve the application of statistical methods in fields as diverse as medicine, sports, and anthropology, illustrate this point.

Example: Dara Mendez

As both an employee of an area pharmaceutical company and a part-time Loyola student, Dara Mendez enrolled in a CDA course offered during the Fall 2000 semester both to further her knowledge in this area of great practical importance and to help her in her role as a pharmaceutical biostatistician. Once again, projects were used to push students to develop the skills required to master a topic beyond the level of the course. As a result, students' final class presentations were more pedagogical in nature. After being provided with the necessary resources, students required only minimal mentoring to accomplish their goals. Once again, they were required to write up their results in the form of a three to five page descriptive paper, which they then distributed to their classmates in conjunction with their 15-minute presentation on the given topic. Topics were chosen (jointly by instructor and student) from those related to the study of CDA but which were slightly above the level of the class. Some of the areas covered included bioassay, nonparametrics, sample-size determination, and pharmacokinetic mixed-effects modelling.

For her project, Dara chose to focus on the statistical methodology involved in quantile bioassay analysis. Using both the data provided in Stokes *et al.* ([25] pp 331-2) and some data from her work projects, Dana illustrated how statistical methods can be used to quantify the potency of an experimental drug relative to a standard one. As a former Biology student, Dara focused her class presentation on the implications of bioassay and the relative potency of drugs in pharmacology. This allowed her to provide both the required framework and larger picture to her classmates. As a result, Dara's classmates benefited from an informative presentation and learned just how prevalent these course methods are in the workplace. And Dara mastered the theoretical justifications and techniques for the methodology she had been applying on the job.

Example: Paul Bell

Ever interested in the use of statistical methods in sports, Paul Bell viewed the class project for his applied regression course during the Spring 2000 semester as an opportunity to use data that he had obtained to develop statistical models to predict the attendance at major-league baseball games in Chicago, Atlanta, and Oakland. For this course, students were required to obtain sufficiently rich data sets to analyze. They then wrote up their findings in a course paper and conveyed them to their classmates via a classroom presentation.

Paul's data was based on the 1999 baseball season and his multiple-regression models included variables such as the day of the week, weather outlook, game number, and the level of the opposing team. Paul then tested his models using 2000 attendance figures and all of his prediction intervals contained the actual reported attendance. The research for this course project earned Paul a Loyola Mulcahy scholarship/grant which enabled him to continue his research with the author. A description of these grants can be found on the Web at the website `http://www.luc.edu/depts/prehealth/Mulcahy.htm`. This grant covered the cost of obtaining his data as well as his travel expenses to Vancouver in June, 2001, to present his findings at the International Biometrics Society conference. As a masters level Statistics student, this experience has proven very beneficial to Paul's professional development.

This class project provided Paul with the opportunity to see how statistical methods can provide predictive models in a research area of special interest to him. His results have aided others interested in predicting attendance at major sporting events, such as stadium managers, area law-enforcement personnel, and mass-transit coordinators. Paul presented his class project findings to the larger Loyola community in a university seminar and has submitted his results for publication in the applied statistics journal *Chance*.

Example: William Burroughs

An undergraduate anthropology major, William Burroughs, enrolled in a course in statistical methods and software packages (STAT 303) during the Spring 2001 semester to learn how to use statistical methods to analyze anthropological data. He chose to examine a paleopathology data set provided by his advisor, Loyola Anthropology Professor Anne Grauer, for his class project and presentation. Projects were used in this class in much the same way as they were for the applied regression course described in the previous example. Students were required to obtain sufficiently rich data sets which they then analyzed and discussed in a class paper and presentation. This process again served to stretch students beyond the level of the material presented in the course.

Professor Grauer's data is related to human skeletal remains (skulls) from medieval England and was used to predict the incidence of a specific disease related to anemia as a function of the age of the subject at death. These data lend themselves to the use of the logistic function

$$E(p) = \frac{e^{\alpha+\beta x}}{1 + e^{\alpha+\beta x}},$$

which relates the expected probability of the disease at death (p) to the subject's age at death (x); here α and β are the model parameters to be estimated from the data. Of special interest to these researchers was the estimated age at which the expected probability of the disease is 50%, denoted x_{50}. After the model was fit to the data, William noticed a distinct pattern in the residual plot that indicated the inadequacy of this model. With some assistance from the instructor with regard to the statistical methods involved, William then wrote a computer program that used a Box-Tidwell transformation (Samuels and Witmer, [24] p 53) where $x = (\text{age})^{\varphi}$ was used in the above model in place of x = age. This latter extension was validated since the subsequent residual plot showed the required random pattern. William's analysis highlighted the importance of this transformation since the estimate of x_{50} dropped from 31 years for the (incorrect) naive logistic model to approximately 20 years for the transformed logistic model.

Through his class project, William came to understand the importance of testing model assumptions and of extending existing statistical methods to fit real-life situations. In addition, while the original analysis performed by Professor Grauer, which involved a chi-square test, was very weak (in terms of statistical power) for this study, the application of the generalized logistic model was novel and provided a more direct answer to the important research queries. As the result of his class presentation, William's classmates understood the usefulness of sophisticated statistical techniques in the field of anthropology.

The class projects for this course involved somewhat more mentoring on the part of the instructor since the statistical prerequisites for the course were indeed modest (some exposure to basic statistics). In addition, the emphasis of the course was more on developing students' statistical programming skills and less on requiring that they master a subject area in applied statistics. All the same, the projects provided students with a sense of confidence and appreciation of the usefulness of applying statistical methodology to real-life data.

3.1.5 Independent Study Courses

As is often the case, class projects in introductory and intermediate courses spark student interest in furthering their studies in statistical fields involving their specific research interests. As a result, a number of follow-up independent study courses have been offered for the more advanced students. The following two examples discuss the development of independent study projects with advanced undergraduate and graduate students enrolled in Loyola's Department of Mathematics and Statistics. These examples illustrate how class projects at this advanced level typically entail assisting a faculty member in cutting-edge research in statistical theory and methodology.

Example: Lisa Leigh and Katie Hanrahan

The exposure to logistic and log-linear models in a basic course in categorical data analysis during the Fall 2000 semester sparked the interest of Lisa Leigh and Katie Hanrahan in the larger field of nonlinear regression methods. As a part of an independent study course, each student worked through the text of Bates and Watts [4] with the author during the first half of the Spring 2001 semester. During the second half of the semester, each student assisted in research by focusing on separate problems related to nonlinear regression. Lisa concentrated on using SAS® software to obtain curvature measures for Gaussian nonlinear models while Katie centered on extending these curvature measures to cover non-Gaussian nonlinear models such as the odds-ratio, relative-risk, and logistic models.

Each of these class projects (independent study courses) required a great deal of mentoring by the instructor (several hours per week over the course of the semester). Assessment and evaluation was based on the final reports prepared by each student, including the quality of the corresponding computer programs. The benefit to the instructor in terms of quality research was significant since Lisa's input helped with the results developed in our recent submission, Haines *et al.* [12], to *Statistica Sinica*. The work with Katie resulted in a presentation by the author at the XXXIIIemes Journees de Statistiques conference in Nantes, France in May, 2001.

Through these class projects, Lisa and Katie mastered the course material in applied nonlinear regression. They were also given the opportunity to gain hands-on experience doing cutting-edge research in statistical methods which has helped each of them identify fields of interest for future study and research.

Example: Paul Bell and Nick Pajewski

The use of existing data sets, such as the HIV data set discussed earlier, to fit linear and nonlinear statistical models naturally leads one to wonder whether a better-designed study could provide researchers with the same amount of information but with fewer experimental runs. This is precisely one of the major goals of the field of optimal experimental design, and follow-up projects (independent study courses) along these lines were proposed by Paul Bell and Nick Pajewski for the Spring 2002 semester.

With the author's guidance, Paul and Nick worked through the material presented in the optimal design textbook of Atkinson and Donev [3]. As with the previous example, this project required a great deal of mentoring on the part of the instructor. As examples were encountered in the textbook, Paul and Nick alternated in developing the necessary computer programs to obtain and verify the optimal designs and the design methodology using the SAS/IML® programming language. With the instructor, Nick and Paul then worked through the papers of Downing *et al.* [8] and Haines *et al.* [11] and wrote SAS/IML® computer programs to obtain the optimal designs for the models presented in these papers as well as for the logistic and the synergistic HIV models discussed above. Nick and Paul were evaluated on their mastery of the optimal design material, on the quality of their computer programs, and on their final reports.

This project served to unify the information learned in such diverse courses as applied linear and nonlinear regression, categorical data analysis, experimental design, and statistical software by providing practitioners with efficient design strategies. Since in some cases optimal designs with only half as many runs provided the same level of information as those actually used, Paul and Nick now understand the importance of a well-designed study in terms of cost-savings.

3.1.6 Evaluation and Assessment

The assessment of student projects is an ongoing process. In lower level courses, students are given milestones that they must meet if the project is to be accepted. For example, they must come up with five potential research articles to critique by a certain date. They must meet with the instructor by a certain (later) date to go through the papers and come up with the two articles to critique. And they must complete the three to five page paper by a certain date. The upper division courses tend to enroll more mature students and the list of deadlines might not be as long. But in either case, failure to meet any deadline is factored into the project/paper grade, which counts for 15% of the course grade.

The evaluation of a student project is much more than making a judgment on the quality of the student's writing. The student typically meets with the instructor two or three times during the last month of the course. During these meetings the instructor can evaluate the student's preparation, judge how much effort is being put into the project and to the course in general, and when appropriate gently push a student to work harder. The project assessment process also helps with the overall assessment of the student.

When the project finally is submitted, the students also must include the related research articles. The instructor then goes through the research articles again before actually reading the paper. The projects are evaluated both on grammar and structural flow and on how well the students understand and discuss the new statistical methods used in the research articles. For example, a research article might use a repeated measures design in which patients are randomized into treatment and control groups, with the results measured over a period of time. Then the student writing the paper must address the fact that the repeated measurements are correlated and that standard statistical techniques do not work. This must be followed by a description and evaluation of the more sophisticated methods that this particular setting requires.

3.1.7 Conclusion

Many statistical educators feel that providing cookbook courses in statistics only furthers the misconception that statistical methodology is a static domain with only limited applicability in practice. In contrast, the above examples show that the field of applied statistical research is constantly evolving to meet the needs of the end-user. Applied statisticians who engage in consulting are well aware that it is not enough to possess a statistical toolkit from which the proper tool is produced to solve the researcher's problem. The successful statistical consultant must develop the ability to meet and address challenging statistical problems with innovative, novel solutions.

As has been seen throughout this article, class projects are beneficial in this regard. They also are useful in underscoring the dynamic nature of our domain by highlighting the fact that *statistical consultants are continuously learning new techniques and refining old tools* so as to better address the problems of researchers. Class projects remind students never to simply accept the results of a given statistical test or prediction without first understanding the underlying assumptions and limitations. Students learn this as they observe situations in which medical researchers inappropriately apply statistical techniques which violate necessary model assumptions. They also learn this by being exposed to situations in which the preliminary (incorrect) analysis leads researchers to an incorrect conclusion. These projects help students develop the critical eye and questioning spirit required of a successful researcher seeking to better understand their field. Finally, by requiring students to go independently beyond the standard course content, the class projects provide students with the sense of confidence needed to master new fields of applied statistics and to become intelligent and critical statistical consumers and consultants.

References

1. A. Agresti, *An Introduction to Categorical Data Analysis*, Wiley, New York, 1996.
2. P. Armitage and T. Colton, eds., *Encyclopedia of Biostatistics* (6 volumes), Wiley, New York, 1998.
3. A.C. Atkinson and A.N. Donev, *Optimum Experimental Designs*, Clarendon Press, Oxford, 1992.
4. D.M. Bates and D.G. Watts, *Nonlinear Regression Analysis and its Applications*, Wiley, New York, 1988.
5. C.R. Boyle, "A Problem-Based Learning Approach to Teaching Biostatistics," *Journal of Statistics Education*, 7 (1999) (obtained on the Web at the website `http://www.amstat.org/publications/jse/secure/v7n1/boyle.cfm`).
6. B.L. Chance and A.J. Rossman, "Sequencing Topics in Introductory Statistics: A Debate of What to Teach When," *The American Statistician*, 55 (2001) 140–144.
7. M. Davidian and D.M. Giltinan, *Nonlinear Models for Repeated Measurement Data*, Chapman & Hall, New York, 1995.
8. E. Downing, V.V. Fedorov and S. Leonov, "Extracting Information from the Variance Function: Optimal Design," in *mODa6 - Advances in Model-Oriented Design and Analysis*, Physica-Verlag, 2001.
9. W.J. Ewens and G.R. Grant, *Statistical Methods in Bioinformatics*, Springer, New York, 2001.
10. T.M. Gerig, U. Blum, and K. Meier, "Statistical Analysis of the Joint Inhibitory Action of Similar Compounds," *Journal of Chemical Ecology*, 15 (1989) 2403–2412.
11. L.M. Haines, G.P.Y. Clarke, E. Gouws and W.F. Rosenberger, "Optimal Design for the Testing of Anti-malarial Drugs," in *mODa6 - Advances in Model-Oriented Design and Analysis*, Physica-Verlag, 2001.
12. L.M. Haines, T.E. O'Brien, and G.P.Y. Clarke, "Kurtosis and Curvature Measures for Nonlinear Regression Models," submitted to *Statistica Sinica* (2002).
13. F.E. Harrell Jr., *Regression Modeling Strategies: With Applications to Linear Models, Logistic Regression, and Survival Analysis*, Springer, New York, 2001.
14. D.W. Johnson, and R.A. Wichern, *Applied Multivariate Statistical Analysis* (5th Edition), Prentice-Hall, New York, 2002.
15. W.N. Kernan, K.M. Viscoli, L.M. Brass, J.P. Broderick, T. Brott, E. Feldmann, L.B. Morgenstern, J.L. Wilterdink, and R.I. Horwitz, "Phenylpropanolamine and the Risk of Hemorrhagic Stroke," *New England Journal of Medicine*, 343 (2000) 1826–1832.
16. L. Krabbendam, M. Mayke, M.A. Derix, A. Honig, E. Vuurman, R. Havermans, J.T. Wilmink, and J. Jolles, "Cognitive Performance in Relation to MRI Temporal Lobe Volume in Schizophrenic Patients and Healthy Control Subjects," *Journal of Neuropsychiatry and Clinical Neurosciences*, 12 (2000) 251–256.
17. S. Lau, S. Illi, C. Sommerfeld, B. Niggemann, R. Bergmann, E. von Mutius, U. Wahn, and the Multicentre Allergy Study Group, "Early Exposure to House-dust Mite and Cat Allergens and Development of Childhood Asthma: a Cohort Study," *The Lancet*, 356 (2000) 1392–97.

18. R.C. Littell, G.A. Milliken, W.W. Stroup and R.D. Wolfinger, *SAS® System for Mixed Models*, SAS Institute Inc., Cary, NC, 1996.
19. S.G. Machado and G.A. Robinson, "A Direct, General Approach Based on Isobolograms for Assessing the Joint Action of Drugs in Pre-clinical Experiments," *Statistics in Medicine*, 13 (1994) 2289–2309.
20. P. McCullagh and J.A. Nelder, *Generalized Linear Models* (2nd Edition), Chapman & Hall, New York, 1989.
21. D. Moore, "New Pedagogy and New Content: The Case of Statistics," *International Statistical Review*, 65 (1997) 123–127.
22. J.C. Pinheiro and D.M. Bates, *Mixed-Effects Models in S and S-Plus*, Springer, New York, 2000.
23. C.A. Pope, R.T. Burnett, M.J. Thun, E.E. Calle, D. Krewski, K. Ito and G.D. Thurston, "Lung Cancer, Cardiopulmonary Mortality, and Long-term Exposure to Fine Particulate Air Pollution," *Journal of the American Medical Association*, 287 (2002) 1132–1141.
24. M.L. Samuels and J.A.Witmer, *Statistics for the Life Sciences* (3rd Edition), Prentice-Hall, New York, 2003.
25. M.E. Stokes, C.S. Davis.and G.G. Koch, *Categorical Data Analysis Using the SAS® System* (2nd Edition), SAS Institute Inc., Cary, NC, 2000.
26. W.N. Venables and B.D. Ripley, *Modern Applied Statistics with S-Plus* (3rd Edition), Springer, New York, 1999.
27. S. Walrand, C. Chambon-Savanovitch, C. Felgines, J. Chassagne, F. Raul, B. Normand, M-C. Farges, B. Beaufrère, M-P. Vasson and L. Cynober, "Aging: A Barrier to Renutrition? Nutritional and Immunologic Evidence in Rats," *American Journal of Clinical Nutrition*, 72 (2000) 816–824.
28. J.H. Zar, *Biostatistical Analysis*, Prentice-Hall, New York, 1998.

Brief Biographical Sketch

Tim O'Brien is an Associate Professor of Statistics in the Department of Mathematics & Statistics at Loyola University Chicago. He is currently engaged in research in statistical education, bioinformatics and statistical modelling and optimal design for generalized nonlinear models with applications in biostatistics. He teaches an array of undergraduate and graduate courses in applied statistics and biostatistics.