Integration of Statistics Throughout the Undergraduate Curriculum: Use of the Senior Chemical Engineering Unit Operations Laboratory as an End-of-Program Statistics Assessment Course

Michael E. Prudich, Darin Ridgway, Valerie L. Young
Department of Chemical Engineering
Ohio University

Abstract

Graduates of chemical engineering programs should have the ability to use basic statistical techniques to analyze and interpret process and experimental data. Chemical engineers should also have the ability to identify the statistical appropriate tool to accomplish a specific data analysis task. This paper describes the undergraduate course of instruction in statistics in the chemical engineering program at Ohio University. This course of instruction begins with a basic introduction to linear regression which takes place in a freshman-level chemical engineering computing course, includes a junior-level course devoted to statistical analysis and experimental design, and ends with the application of statistical data analysis tools in the senior-level unit operations laboratory course sequence. Spreading the exposure to statistical analysis across the curriculum permits longitudinal reinforcement of important concepts and skills. Special emphasis is placed on the use of the senior-level unit operations laboratory experience as a capstone statistics usage and assessment tool. This course provides the students with an opportunity to break out of the "chapter box" which is often characteristic of stand-alone statistical methods courses. The courses also emphasize the point that experimental planning and design includes not only the selection of the experimental parameters to be studied, but also the planning of the data analysis and statistical treatments to be utilized in the interpretation of the experimental data that is acquired. The proper assessment of statistics-related performance in the senior-level laboratory courses provides end-of-program assessment data on student statistical skills and abilities.

Statistics Throughout the Curriculum

Statistics are used by chemical engineers both to interpret data and to formulate and test models derived from data. The importance of statistics in the engineering curriculum has been recognized by ABET. The 1999-2000 general criteria[1] include the requirement that "students must demonstrate knowledge of the application of probability and statistics to engineering problems." Recently, with the advent of EC 2000, chemical engineering departments have had more flexibility in determining learning objectives and their associated outcomes. The need for a working knowledge of statistical analysis tools and techniques is specifically recognized by the Department of Chemical Engineering at Ohio University under our Objective 1: "Graduates will have a strong foundation in chemical engineering theory and practice" as Outcome 1.d. "Students will demonstrate the ability to formulate models, design and conduct experiments, analyze data, and interpret results using statistical tools."

The overall exposure to statistics in the Department of Chemical Engineering at Ohio University has evolved in an ad hoc manner. (Typical use and/or instruction of statistics and statistics-related subjects in required courses in the chemical engineering curriculum is described in Table 1.) There was no undergraduate course dedicated to statistics required for graduation in the chemical engineering curriculum prior to 1996. Before that time, statistics instruction consisted of what individual instructors would decide to pass along in their courses depending on their particular desires and course needs, as well as a brief introduction to linear and nonlinear regression and, sometimes, hypothesis testing, that was included in the last 10-20% of a junior-level course in chemical engineering calculations/computer solutions. Beginning in 1992, an undergraduate chemical engineering technical elective course in experimental design and statistics was offered. This undergraduate course in experimental design and statistics was elevated from technical elective to required status in the chemical engineering curriculum beginning in 1996.

Our required undergraduate course in experimental design and statistics for chemical engineers resembles courses typically offered by other universities. It is a stand alone, lecture-type course that is not explicitly related to any of the other courses in our curriculum. A standard text[2], was used for the required statistics course in 2001. The subjects covered in this course are the usual ones including: probability; random variables, the normal distribution; sample distributions; determining an adequate sample size; various types of hypothesis testing; fitting equations to data; factorial analysis; and an introduction to statistical process control. Supplementary materials are provided which cover nonlinear least squares analysis and fractional factorial experimental design. Most would agree that the subjects listed above are important for chemical engineers and most students perform well enough within the limits of the course (specifically, Spring Quarter junior year). Later assessments (see below) indicate that there are three deficiencies in presenting this material in a 'stand alone' matter: (1) the students learn the material in the context of individual

Table 1. Statistics Usage and/or Instruction in the Undergraduate Chemical Engineering Curriculum at Ohio University.

Course	Use of and/or instruction in statistics.
Freshman Year (Year #1)	
ChE 100 - Introduction to Chemical Engineering	None reported.
ChE 101 - Approaches to Chemical Engineering Problem Solving	How to linearize a non-linear algebraic equation. The use of MATLAB to determine equation parameters using single and multiple linear regression. Residuals are calculated.

Table 1. Statistics Usage and/or Instruction in the Undergraduate Chemical Engineering Curriculum at Ohio University (continued).	
Course	Use of and/or instruction in statistics.
Sophomore Year (Year #2)	
ChE 200 - Material Balances	None reported.
ChE 201 - Energy Balances	Introduction to the concept that uncertainty in measurements to show that real energy balances do not close to the nth decimal place. The students record temperatures and flow rates through a simple concentric tube heat exchanger, calculate the heat loss from the hot side and the heat gain by the cold side, compare the numbers, and observe how the comparison changes if the value of each temperature changes 1 °C. No formal calculations of uncertainty are made.
Junior Year (Year #3)	
ChE 305/306 - Thermo; ChE 345 - Fluids; ChE 347 - Mass Transfer and Separations; ChE 400 - Applied ChE Calculations	None reported.
ChE 307/308 - Chemical Engineering Kinetics	Least squares fitting is used. Linearize rate laws and determine rate parameters through least squares fitting. The students are introduced to non-linear least-squares analysis through a section in the course text and a homework assignment.
ChE 346 - Heat Transfer	An associated laboratory experiment. Experimental design - the concept of spreading out the values of independent variables. Estimating the error in a calculation. Linearizing non-linear algebraic equations. Determination of equation parameters by least-squares analysis. Parity plots.
ChE 408 - Engineering Experimental Design	Statistics course (details in text).
Senior Year (Year #4)	
ChE 443/444 - Process Design; ChE 448 - Safety in the Process Industry; ChE 499 - Chemical Engineering Senior Assessment	None reported.
ChE 415/416 - Unit Ops. Lab	Details below.
ChE 417 - Process Controls Laboratory	linear and/or non-linear least-squares analysis are used for three lab projects. Hypothesis testing is used for at least three lab projects.
ChE 442 - Process Controls	least squares fitting is used to determine the time constant and/or steady-state gain from response data.

‘chapter boxes’ (for example, the problems at the end of Chapter 4 will be on the subject matter covered in Chapter 4), (2) the students have no professional motivation or practical understanding of the chemical engineering uses of the subject of the subject matter being presented (most examples are not chemical engineering related), and (3) too much time elapses between formal instruction in the use of statistics for experimental planning and data analysis and the practical need to use statistics for experimental planning and data analysis.

Assessment Via the Senior-Level Unit Operations Course

The unit operations laboratory at Ohio University is a two-quarter sequence taken during the senior year. Four experiments are performed each quarter. The junior-year thermodynamics, reactor design, and unit operations courses are prerequisites, along with a course in experimental design and statistics. Students typically work in teams of three. Each experiment is mentored by a different faculty member. The faculty mentor is responsible for all aspects of assigning of objectives, meeting with the students and grading the reports. The format includes an informal meeting with the faculty mentor, a graded prelab in which the students are to present their experimental plan and expected data analysis plan, one week (10 hours) of lab time and a final written report. Grading is on common course grading rubrics for both the prelab and postlab reports. These grading rubrics are supplied to the students during an orientation held at the first class meeting and are available throughout the quarter on the course webpage[3]. The details of these grading sheets and the philosophies behind their use have been described elsewhere[4]. The grading sheets cover all aspects of the prelab and postlab reports including both format and content with the portions related to experimental design and statistical analysis being those of interest to this study.

The unit operations laboratory sequence as offered at Ohio University can be thought of as having two functions related to experimental design and statistics: (1) the reinforcement of analysis skills learned by students in previous courses, and (2) the evaluation and assessment of that previous learning. The two quarter unit operations laboratory sequence is ideally situated and constituted for the assessment of the students’ abilities to apply their training in statistics due to both its location in the senior year of the curriculum and to the course requirement to identify the need for and correctly apply a wide range of statistical techniques to the completion of the data analysis that comes about as the result of the experimentation (see Table 2). Every quarter in this course each lab team prepares four prelab and four postlab reports, and each individual prepares one design memorandum and makes one oral presentation describing their design. Each of these assignments is evaluated with respect to a number of Primary Traits (10 each for the Prelab and Postlab reports and five for the Design Memorandum and six for the Design Oral) many of which include elements related to the application of statistics and/or experimental design (see Table 3). For the prelab and postlab reports, a substantive deficiency in any Primary Trait triggers a mandatory re-write of the report.

Table 2. Experiments Performed in Unit Operations Laboratory With Associated Potential for Statistical Analysis (Fall Quarter AY2002/2003).	
Experiment	Potential Applications of Statistics
Drying	• experimental design • linear regression (to find slope of line; find the confidence interval around the slope; to find the intercept; find the confidence interval around the intercept; to determine linearity of data; to find the equation parameters for the linearized version of a nonlinear heat-transfer and mass-transfer coefficient relationships; to find the confidence intervals around the equation parameters) • nonlinear regression (to find equation parameters for nonlinear heat-transfer and mass-transfer coefficient relationships; to find the confidence intervals around the equation parameters) • hypothesis testing (paired t-test to compare if two measurement techniques yield the same values; t-test to determine if experimental measurements agree with literature/theoretical values; t-test to determine if equation factors are significant)
Fluid Flow	• hypothesis testing (t-test to determine if experimental measurements agree with literature/theoretical values)
Membrane Permeation	• experimental design • linear regression (to find slope of line to determine the Fick's Law parameter; find the confidence interval around the slope; to find the intercept; find the confidence interval around the intercept; to determine linearity of data) • hypothesis testing (paired t-test to determine if countercurrent and cocurrent configurations produce the same results; t-tests on slopes to determine if the Fick's Law parameter is a constant.)
Mixing	• experimental design • linear regression (to find the equation parameters for the linearized version of nonlinear mixing power and performance relationships; to find the confidence intervals around the equation parameters) • nonlinear regression (to find the equation parameters for nonlinear mixing power and performance relationships; to find the confidence intervals around the equation parameters) • hypothesis testing (t-test to determine if experimental measurements agree with literature/theoretical values; t-test to determine if equation factors are significant)

Table 3. Primary Traits Which Include Some Aspect of Statistics.	
Primary Trait Name	Description of Statistics Content
Prelab Report (2 of 10 Primary Traits include statistical content)	
Experimental Test Matrix	"Specifies all experimental conditions with reasonable values, sufficient range to meet all objectives, reasonable number of runs, repeat runs..."
Statistical Methods	"Uncertainties for all values stated. Uncertainties appropriately based on propagation of experimental error or on statistical methods. States how error will be calculated. Planned evaluation of model equations based on variance, correlation coefficient, and/or residual plots. Comparison of values based on hypothesis testing (e.g. t-test) as appropriate. Specific."
Postlab Report (4 of 10 Primary Traits include statistical content)	
Report Overall, Experimental Uncertainty	"All values quoted with uncertainties throughout the report. Uncertainties appropriately based on propagation of experimental error or on statistical methods. Includes error bars on figures as appropriate."
Presentation Style	"... Error bars are included where appropriate."
Statistical Analysis	"Complete and appropriate statistical analysis included. Evaluation of data, results, and/or model equations based on variance, correlation coefficient, and residual plots. Comparison of values based on hypothesis testing (e.g. t-test) as appropriate. Appropriate error bars on all figures."
Conclusions	"... Expresses limitations in terms of uncertainties ..."
Design Oral Critique (2 of 6 Primary Traits include statistical content)	
Experimental Approach	"... I (the evaluator) know the uncertainty on each of the experimentally-determined parameters."
Proposed Design	"... specifications which depend on experimentally-determined values are given an appropriate range of values based on experimental uncertainty..."
Design Memo Grading (2 of 5 Primary Traits include statistical content)	
Description of Experimental Approach	"... uncertainty stated on each experimentally-determined parameter..."
Proposed Design	" ... specifications which depend on experimentally-determined values are given an appropriate range of values based on experimental uncertainty..."

Prelab report grading results for Unit Operations Laboratory I were consistent for Academic Years 2001/2002 and 2002/2003 with problem areas identified with the two Primary Traits *(Data Analysis/Statistical Methods* and *Data Analysis/Expected Data and Results*) associated with the analysis of experimental data after it has been collected. During Fall Quarter AY2001/2002, ten lab teams participated in ChE 415 - Unit Operations Laboratory I. Of a total of 40 prelab reports submitted, 22 (55%)were returned for mandatory re-writes due to substantive deficiencies in at least one of the Primary Traits. Of the 22 mandatory re-writes, nine were returned for deficiencies found in the *Data Analysis/Statistical Methods* category. This number represents 41% of the mandatory prelab rewrites and 23% of the total prelab reports. The next highest category that resulted in mandatory re-writes was the Primary Trait *Data Analysis/Expected Data and Results* which resulted in five mandatory rewrites (23% of all mandatory prelab rewrites; 13% of all prelab reports submitted).

Nine laboratory teams were enrolled in Unit Operations Laboratory I Fall Quarter AY2002/2003. Of the 36 prelab reports submitted during this quarter, only nine (25%) were returned for mandatory re-writes. Of these nine mandatory re-writes, six (67% of the mandatory re-writes; 17% of all prelab reports) were returned for deficiencies found in the *Data Analysis/Statistical Methods* category. As in AY2001/2003, the next highest category that resulted in mandatory re-writes was the Primary Trait *Data Analysis/Expected Data and Results* (44% of the mandatory re-writes; 11% of all prelab reports). [Note: Percentages may sum to greater than 100% since an individual report may fail more than one Primary Trait.]

Prelab reports are especially important in that they provide the roadmap by which the students acquire and analyze their data. A poor prelab report often results in deficiencies in data acquisition (wrong kind and/or not enough data taken) and in data analysis (inappropriate methods and/or incomplete analyses).

Difficulties identified in the prelab reports included:

* no statistical analyses planned
* vagueness (for example, "Appropriate statistical analyses will be performed.")
* statistical analyses not planned for all appropriate situations (at least one situation properly addressed with, perhaps, the assumption made that only one application of statistics will successfully get that 'box' checked on the grading sheet)
* inappropriate statistical methods proposed (To a student with a new hammer, every problem is a nail - "I know how to do the t-test, I will apply it to every problem.")
* insufficient planning for data collection (even if the correct statistic methods are proposed, the students often make their experimental plan without considering the type, range, and number of experimental trails required in order to provide the data necessary to their chosen methods of analysis)

Upon completion of each experiment, a laboratory team would prepare and submit a postlab report documenting their experimentation, providing an analysis of their data, and describing any models or correlations made for the purpose of later design work. Of the total 40 postlab reports submitted for Fall Quarter AY2001/2003, only six (15% of the total number of postlab reports submitted) reports included deficiencies severe enough to result in mandatory rewrites. Only one of the six mandatory rewrites was for a deficiency related to statistical analysis (17% of all postlab rewrites; 3% of all postlab reports).

Once again, data collected for AY2002/2003 show trends similar to those observed for AY2001/2002. For the 36 postlab reports submitted, only four (29% of all postlab re-writes; 11% of all postlab reports) required re-writes for a deficiency related to statistical analysis, a decrease from the percentage observed for prelab reports submitted by the same cohort of students.

Our tentative conclusion is that students are usually able to perform the mathematical manipulations associated with the required statistical analysis tools as evidenced by their success with the postlab reports. However, our students were often unable, or unwilling, or fail to see the necessity to plan the data analyses they will need and to select the statistical tools required to make those analyses. This problem was evidenced by the number of statistics-based mandatory prelab rewrites. The postlab reports were successful because the data analysis tools had already been identified as a result of the required prelab rewrites and subsequent consultation with the faculty mentors. Another way of describing this problem would be that the students had learned how to use the individual tools in their tool boxes in the context of the "chapter boxes" in which they were presented in the required statistics course. When confronted with an undefined problem, the students would have their tool boxes, but lack the confidence and experience to select the correct statistical tool for the job required.

The students enrolled in Unit Operations Laboratory I perceived a deficiency in their prior instruction in statistical methods as evidenced by end-of-course survey results. As a part of the end-of-course survey the students are asked to evaluate prerequisite chemical engineering courses as to their adequacy ("5" being prepared "very well", "1" being prepared "poorly") in preparing them for the unit operations laboratory experience. Of the eight courses/subject areas evaluated, only one (experimental design and statistics) has consistently received an unfavorable rating, 2.0 (next lowest was 4.4) for AY2001/2002 and again 2.0 (next lowest was 4.3) for AY2002/2003. It is our hypothesis that the problem is not in the subject matter covered within the course that we offer, rather it is the method of coverage and the lack of timely, significant reinforcement of statistical data analysis skills.

Recommended Solutions

The identification of deficiencies without plans to address those deficiencies is of little use. We are implementing the following changes in our overall plan of instruction for experimental design and statistics.

The required experimental design and statistics course will be modified to be taught more from a case study perspective. (This requires the instructor to formulate and/or acquire relevant cases. The course[5] has recently been modified and offered in this manner[5] Winter Quarter AY2002/2003. A detailed description of the course can be found at the course website[6].)

The required experimental design and statistics course will include more 'chemical engineering' examples and problems. (This requires the instructor to formulate and/or acquire relevant problems. The course[5] has recently been modified and offered in this manner Winter Quarter AY2002/2003.)

Timely reinforcement (in terms of class usage and experimentation and data analysis) will be provided for the materials covered in the required experimental design and statistics course.

The chemical engineering faculty at Ohio University have already experienced success with embedding computer programming (MATLAB) throughout the curriculum. Prior to this embedding our students learned MATLAB their freshman year, but were not required to use it again until their junior year in the curriculum. It was unsurprising that they entered the junior-year class with little or no retention of their MATLAB skills. This problem was solved by embedding problems requiring the use of MATLAB in additional sophomore- and junior-year courses. Evidence of the success of this approach is the fact that many students will solve simple and complex problems using MATLAB programming during senior year without be required to do so. They have taken ownership of this skill. A similar approach should be taken in the area of statistical skills.

An initial approach will be to add opportunities for students to apply the skills learned in the required undergraduate statistics course in other core chemical engineering courses. A paper by Nelson and Walloons[6] have endorsed this approach stating that "we believe that an introductory course in statistics for engineers must be considered in conjunction with the entire engineering curriculum. No matter how good an introductory course in statistics is, if students are not asked to use this material in any subsequent courses, they will soon forget it and most probably question why they were required to take the course in the first place. Thus, we propose to enlarge our area of concern from just an introductory course in statistics to how the concepts from this course can be utilized, reinforced, and enhanced in subsequent engineering courses." Their initial application of this concept was to undergraduate chemical engineering unit operations laboratories at Clemson University. They further state that, "Our approach will be to start by introducing statistical techniques into the engineering labs. We are suggesting engineering labs as the initial point of attack because laboratories are the setting where students perform experiments and collect data." Laboratory exercises were selected for reinforcement because: "This approach combines the advantages of hands-on activities and real world data sets. Students create their own data sets, which gives them added motivation to do the analyses. Moreover, because the data come from an experiment whose main purpose is to illustrate important engineering principles, the data carry a sense of real import that statistics courses ordinarily can achieve only with archival data[7]."

Two courses have been identified for use in reinforcing experimental design and statistics learning: ChE 346 - Heat Transfer and ChE 347 - Mass Transfer and Separations. The heat transfer course is offered the same quarter (Winter Quarter) as the required undergraduate statistics course and already has a laboratory project incorporated into the course. An effort has been made to integrate this existing laboratory project with the subject matter being presented in the statistics course. The first offering of the integrated heat-transfer experiment was made in Winter Quarter AY2002/2003. The second course, the Mass Transfer and Separations course, is offered the quarter directly following the statistics course (Spring Quarter) and does not currently have a laboratory project. We plan to move the membrane permeation experiment (see Table 2, above) from the senior Unit Operations Laboratory course and, with some slight modifications, introduce it as a laboratory project in the mass transfer and separations course. The first offering of the integrated mass-transfer experiment will be made in Spring Quarter AY2002/2003. Cobb[8] has recognized the advantages of this practicum-type approach of instruction. Cobb lists the following perceived advantages of hands-on activities in instruction in statistics (1) "the resulting data are fresh, not someone else's leftovers, (2) students are actively involved in data production" - (a) "the student who has a hand in creating a data set is nearly always motivated to analyze it, (b) experience with variability is immediate and concrete," and (c) "the activities can involve students with design (of experiments) issues."

We are hopeful that, after making the changes described above, we will be able to assess and report improved learning in the areas of experimental design and statistics for chemical engineering students at Ohio University.

Bibliographic Information

1. ABET, 1998, "Criteria for Accrediting Engineering Programs, Effective for Evaluations During the 1999-2000 Accreditation Cycle."
2. Barnes, J.W., 1994, Statistical Analysis for Engineers and Scientists: A Computer-Based Approach, McGraw-Hill, New York.
3. Course materials for our two unit operations laboratory courses, including copies of the grading rubrics, can be reached through at http://webche.ent.ohiou.edu/course_index.html
4. Young, V.L., Ridgway D., Prudich, M.E., Goetz, D.J., and Stuart, B.J., 2001, "Criterion-Based Grading for Learning and Assessment in Unit Operations Laboratory," Proceedings of the 2001 American Society for Engineering Education Annual Conference & Exposition, Session 3213.
5. Young, V.L., 2003, "Designing a Statistics Course for Chemical Engineers," Proceedings of the 2003 American Society for Engineering Education Annual Conference & Exposition, Session 3513.
6. The homepage for our undergraduate chemical engineering statistics course can be found at http://webche.ent.ohiou.edu/che408/homepage.html
7. Nelson, P. and Walloons, T., 1993, "Improving the Undergraduate Statistical Education of Engineers," Journal of Statistics Education, v. 1, n. 1, http://www.amstat.org/publications/jse/v1n1/cobb.supp.nelson.html
8. Cobb, G.W., 1993, "Reconsidering Statistics Education: A National Science Foundation Conference," Journal of Statistics Education, v. 1, n. 1, http://www.amstat.org/publications/jse/v1n1/cobb.html

Biographical Information

MICHAEL E. PRUDICH is Professor and Chair of the Department of Chemical Engineering at Ohio University. His B.S., M.S., and Ph.D. degrees are from West Virginia University. He has been a member of the faculty at Ohio University for 18 years, joining academia after working for Gulf Oil Company in Pittsburgh, PA.

DARIN RIDGWAY is an Associate Professor of Chemical Engineering at Ohio University. He received his B.S. from Rose-Hulman Institute of Technology, M.S. from Louisiana Tech University, and Ph.D. from Florida State University. His current research interests are in the areas of biochemical engineering.

VALERIE L. YOUNG is an Assistant Professor of Chemical Engineering at Ohio University. She received her B.S. from Lehigh University and Ph.D. from Virginia Tech University. Her current research interests are in the areas of atmospheric chemistry.