Designing a Statistics Course for Chemical Engineers

V.L. Young
Department of Chemical Engineering, Ohio University

Abstract
The Department of Chemical Engineering at Ohio University redesigned an existing course in experimental design and statistics. The revision was motivated by assessment information from a variety of sources: course-based assessment in our senior Unit Operations laboratory, exit surveys of seniors, surveys of alumni 2 years after graduation and input from our departmental advisory board. The consensus of faculty, students, alumni, and the advisory board was that (1) a solid foundation in statistics is important preparation for industrial engineering practice as well as for advanced degree work in engineering and (2) "solid foundation" means that graduates can select *and* execute appropriate statistical techniques to analyze real data *and* interpret the results. In spite of having a statistics course in our curriculum, graduates did not leave with the solid foundation we wanted. In particular, our seniors showed unsatisfactory ability to frame a problem in terms of a hypothesis that can be tested statistically and unsatisfactory ability to select an appropriate statistical test. New graduates were only beginning to operate at the desirable higher levels of analysis, synthesis, and evaluation. As part of a strategy to address this problem, our statistics course for juniors was redesigned with input from our faculty and from industrial members of the advisory board. The new course emphasizes software rather than hand calculations, introduces application and follows up with theory, and uses case studies from industry and from academic research. This course is not isolated in our curriculum. Statistical analysis is now a required part of projects in Heat Transfer and Kinetics, and continues to be emphasized in Unit Operations. In this talk, we reveal the motivation for emphasizing statistics in our curriculum, the structure of the re-designed course, and the assessment methods being used to gauge student learning in this course.

Why Teach Statistics?

Statistical methods of data analysis are valuable tools to chemical engineers in both research and in industrial practice. Consider this quote from a recent National Science Foundation program announcement (italics added). "Projects must use appropriate quantitative methods, and teams should include individual(s) with demonstrated expertise in the quantitative methods to be used. Quantitative methods may include: conceptual, mathematical or computational models; computer simulation; artificial intelligence techniques; *hypothesis testing*; *statistics*; visualization; or database development. Mathematical models must include *estimates of uncertainty*, and experiments should assess *power and precision*." "Six-Sigma", the currently popular industrial philosophy and method for quality control, is based on statistical methods of process analysis and decision-making. Articles about "Six-Sigma" have recently featured in Chemical Engineering Progress.[1,2] Every engineer is expected to participate in Six-Sigma process improvement, not just those assigned to a process or quality control department.

In our own curriculum, assessment at several different levels indicated a need to improve the foundation in statistics that we provide to our undergraduates. Course-based assessment in our senior Unit Operations laboratory showed that most of our graduates' competence in statistics was limited to the lowest levels of Bloom's Taxonomy[3] (Knowledge, Comprehension, at best Application). The author's personal experience is that competence in these categories is typically rapidly "flushed" after the final exam. Feedback from our students (current seniors and those 1-2 years after graduation) and our external advisory board supported our ambition to have many of our graduates competent in statistics at the levels of Analysis, Evaluation, and Synthesis.

Since we already had a 1-quarter required course in statistics in our curriculum, it was apparent spending 10 weeks on the standard subject matter was not accomplishing our goals. Students completing the existing course and later evaluated in Unit Operations Laboratory were able to execute a variety of statistical calculations (descriptive statistics, least-squares linear regression, determination of confidence intervals, t-test for difference in means), but did not demonstrate the ability to recognize when a particular method would be appropriate or to use statistical methods to state their results and conclusions quantitatively. For more information about this, see the paper by Prudich, Ridgway, and Young in this session.[4] We decided that the introductory statistics course needed to be redesigned to better prepare students to *use* statistics in data interpretation and decision-making, and that we would continue to reinforce and evaluate their capabilities in the senior Unit Operations Laboratory. Responsibility for redesigning the course was given to someone (the author) who had been actively involved in all aspects of assessment (course-based evaluations, post-graduation surveys, advisory board discussions) that led to this decision.

Constraints on Course Design

The course instructor is tenure-track, and must maintain activity in research and service as well, so running the course cannot be a fulltime job. The course is 3 credits, 10 weeks, meeting weekly for two 50-minute periods and one 110-minute period. Prerequisites are material and energy balances, calculus and differential equations, and numerical methods for solving engineering problems (using Matlab). Most students are concurrently enrolled in heat transfer and the second quarter of thermodynamics, and have developed basic competence with Microsoft Word and Excel as a survival mechanism. The course should support the remainder of the curriculum but not require major changes to the curriculum.

Course Goal and Outcomes

Goal: Students in this course should develop a practical background in statistics that will allow them to apply statistical techniques to problems of data analysis, process control and experimental design in both research and process engineering environments.

The list of student learning outcomes is available on the World Wide Web at http://webche.ent.ohiou.edu//che408/homepage.html , along with other course information, or by contacting the author directly. The outcomes reflect fairly standard technical content in the categories of statistical methods of data analysis, statistical process control, and experimental design. Topics listed on the course syllabus for each 50-minute lecture are also quite standard.

Course Redesign

If the technical topics and student learning outcomes for the course are fairly standard, what have we changed? Four significant changes to the course were made.

1. Calculations: software use is favored over hand calculations and memorization of formulas.
2. Case studies: five of the ten 110-minute recitation periods are used for the class to work on real problems. (Three are devoted to exams and two to tutorials on Excel and Matlab use and on advanced regression topics.)
3. Higher-level practice: the homework assignments will include problems requiring analysis, evaluation, and/or synthesis.
4. Assessment: assessment of student learning will focus on higher-level assignments.

Calculations and Software Selection

Although some argue that you can't really understand least squares regression until you have worked through an entire problem by hand, the author's personal experience is that most junior chemical engineers are quite capable of working mechanically through the mathematics of such a tedious process without really thinking about what they are doing, and are likely to do so. Deeper understanding is more likely to result from solving higher-level problems. Practicing modern engineers use software for statistical calculations and rarely refer to the original formulas for the calculations. Students in our course do the same. For example, the formula for the sample standard deviation is introduced in the textbook reading and lecture. The symbols are defined and its calculation from the deviation of each data point from the mean explained. The formula does not appear again in lecture, homework, or exam. Students calculate the standard deviation of sample data using software, not by hand. They are not expected to memorize or interpret the formula. They **are** expected to identify standard deviation as a descriptor of the variability of sample data, to sketch how a change in standard deviation is reflected in the shape of the normal distribution, and to make use of the fact that about 2/3 of the data in a normally-distributed data set will fall within one standard deviation of the mean.

Statistical calculations can be performed with a variety of software packages, some specifically dedicated to statistical analysis. The author considered adopting the statistical package Minitab for the course for the following reasons.

- It is widely used for statistical analysis in industry (according to company literature and confirmed by our advisory board).
- No kluges are required to get it to perform even advanced techniques for data analysis, process control, or experimental design. In Matlab (with Statistics Toolbox) and Excel (with Analysis Toolpak Add-In), multiple steps and some cleverness are sometimes required.
- Students can obtain it inexpensively ($26 to rent for 5 months, or $100 to buy).

The author decided against adopting Minitab as official course software for the following reasons.

- Our juniors are comfortable with Excel and Matlab as tools for solving problems. Kluges in Excel and Matlab do not intimidate them. They should see statistics calculations in the same class as addition, integration, or solving a differential equation – some math stuff you do to solve a bigger problem and not an end unto itself. Adopting separate software might encourage students to think of this as "the Minitab class".

- The format of Minitab output is less flexible than that of Excel and Matlab (at least for the novice), making it more difficult to generate output appropriate for reports and memos. This adds a software hurdle that is more easily cleared with Excel and Matlab.
- The expense of a site license of Minitab would have to be absorbed independently by the department. We already have university licenses for Excel and Matlab.

The first 110-minute is a tutorial period covering summary statistics, t-tests, and linear regression using Excel and Matlab. The tutorial focuses on performing the calculations; explanations of their meaning are limited. Students are expected to learn more about the software on their own as the quarter progresses, and seek individual help when necessary.

Case Studies

No one on our faculty had statistics as a required part of his/her own undergraduate engineering education; I suspect we are not alone. We have learned statistics (admittedly to varying degrees) because we have needed it, accumulating theoretical understanding along with practical use. Pedagogical literature is full of papers demonstrating the superiority of problem-based learning compared to lecture-based learning in developing higher levels of student performance. Developing and delivering a completely problem-based course is, however, more time-consuming than most tenure-track faculty can afford. Including five case studies adds a problem-based learning element without swamping the instructor. Case-studies are worked in the 110-minute recitation periods, held in a PC computer lab. The case studies are available on request from the author.

Two of the case studies come from within our curriculum. In one, students are asked develop a calibration curve to determine crystal violet dye concentration from a UV-Vis absorbance measurement. Students are given data describing the makeup of the standard solutions (mass of solid dye, volume of solution, dilution volume) and the measured UV-Vis absorbance. They must develop a model that gives concentration with uncertainty and evaluate it. This case study uses propagation of error and least squares linear regression. UV-Vis measurement of this dye concentration is used in our senior Unit Operations Laboratory kinetics experiment, but the seniors are given the calibration curve by the instructor. The calibration data already exists, and no harm is done by students working with it prior to their senior year. In the other internally-inspired case study, students design an experiment to determine heat transfer coefficients for a heat exchanger in our Unit Operations Laboratory. The timing of this case study coincides with the start of a team project with the same objective in the concurrent heat transfer class. All grading related to the Experimental Determination of Heat Transfer Coefficients is in heat transfer, but students have an opportunity to concentrate on the experimental design, requesting guidance when needed, and are highly motivated for the class period by the graded project in the concurrent class.

Development of three of the case studies was assisted by our external advisory board. In two of them, artificial but realistic chemical process data is used. Students are supposed to evaluate process performance and to determine whether a process change has made a significant impact on the output. The third of these case studies is based on an actual NASA project, in which the rate of polymer erosion in space had to be estimated, and the experimental design was constrained by the achievable experimental uncertainty.

Higher-Level Practice

Exercises from the text are particularly useful for developing the lower levels of Bloom's Taxonomy, and are still included in the homework. The previous incarnation of the course used only text problems as homework. Unfortunately, a problem that appears on the surface to be higher level reverts to a lower level when it immediately follows a text section and worked examples in the same format on the same topic. There is no magic in realizing that you should solve this problem with a t-test if it comes at the end of the section discussing the t-test for a difference in means.

Higher-level learning requires that the problems not be clearly identifiable with a particular section in the book. Most of the homework assignments include one higher level problem not from the course text. Students are given some raw data and asked to evaluate it. Does this product meet specifications? Is this device useful for measuring SO_2? How do the in-house lab results for water phosphate concentration compare to those of an independent lab? Propose an experimental plan for determining the density of a liquid to a specified level of uncertainty. The questions are broad enough that more than one statistical technique could be acceptably applied and complex enough that more than one technique or more than one step is probably required. Solutions with differing degrees of thoroughness are quite likely. The solution format is always "a memo to your boss". Mechanical performance of a calculation will not lead to a passing grade on these problems. The students practice selecting appropriate statistical techniques and using statistics to convey their conclusions quantitatively. These are the same tasks that will be expected of them as seniors in Unit Operations Laboratory, but practiced on problems of more limited scope.

Assessment

Previous course-based assessment for the statistics course actually showed acceptable performance, in spite of the fact that student capabilities a year later in the Unit Operations Laboratory were unacceptable. The acceptable junior ratings were based on assessment using objective, single answer problems clearly identified with particular topics in the course, and often identical to problems which had been assigned in previous years. In other words, most assessment in the course was performed at the lower levels of Bloom's Taxonomy, even when the question was *apparently* a higher-level question (e.g., evaluate the difference between these two data sets). In fact, assessment in the later course showed that seniors *did* perform acceptably at the lower levels, but that is not our desired achievement level. In the revised course, there is still assessment of lower-level skills on exams. However, assessment is also based on student performance on the higher-level problems solved for homework and as part of a graded in-class individual project (conducted in the last 110-minute recitation period). Obviously, there is some concern about student collusion on homework leading to falsely inflated assessment scores. The requirement to submit solutions in memo format alleviates some of this, since it is easier to spot straight plagiarism in a 1-3 page memo than in a single-answer calculation or in a massive report. The graded individual in-class project, which is completed under instructor supervision, also alleviates some of this concern. In fact, the assessment results show that if students were colluding, they often colluded to arrive at an incorrect answer. It appeared that students did sometimes collude when solving the problem, but that they wrote their memos individually. On several occasions, a student who performed the calculations correctly revealed inadequate or incorrect understanding in the memo.

Results: Selection and Execution of Statistical Techniques

Assessment of student learning within the course is reported here, and is compared with the typical performance of students entering the senior Unit Operations Laboratory in the past. In general, in contrast to students entering Unit Operations in the past, students leaving this junior-level statistics course can select *and* execute appropriate statistical techniques to analyze real data. The true test of student learning will come next year, when these students take Unit Operations Laboratory, five months after their last formal instruction in statistics.

In the assessment, a pattern emerged which seemed to be a microcosm of what we believe happens in the curriculum as a whole. In an assignment requiring use of material just covered in the course, students perform well. In this case, even apparently higher-level questions are really plug-and-chug, because the student simply grabs a technique from recent material. Performance nose-dives on a similar assignment as little as one week later, as many students try to force the problem to fit the technique covered most recently in class. As students are assigned successive problems that require the same technique in different contexts, the oscillations in student performance damp, until the great majority of the class performs acceptably. One of the most basic desired student learning outcomes of the statistics course demonstrates this behavior.

Student demonstrates the ability to provide an estimate of experimental uncertainty with all data and results without prompting.

In spite of specifically noting this requirement in the grading rubric for Unit Operations Laboratory, it has been common for our seniors to provide values without uncertainties in their reports. Assessment of student performance on this outcome is shown in Table 1. The measures of performance are listed in the order that the assignments occurred in the course.

Table 1. Assessment of student performance on a key learning outcome.

Student demonstrates the ability to:	Measure	Yes	No	Can't Tell	Sum
Provide an estimate of experimental uncertainty with all data and results without prompting	HW1 memo	24	2	1	27
	HW2 memo	16	11		27
	Exam1 A2, B4	19	6	2	27
	Exam1 A2, B4	25	2		27
	HW4 memo	26	1		27
	HW6 memo	21	6		27
	Exam 2 1A	27	0		27
	Exam 2 1B	26	1		27
	Final project 1	27	0		27
	Final project 2	27	0		27
	Final, 1	27	0		27
	Final, 2a	27	0		27

In the statistics course, the first week focused on estimating measurement uncertainty, error propagation, and characterizing the variability of replicate measurements using standard deviation. In the homework due at the end of the first week (HW1 memo), only two students failed to estimate uncertainty without prompting. The second week of class focused on correlation and least-squares regression. A correct solution to the higher-level homework problem that week (HW2 memo) required both regression and estimation of uncertainty. Eleven students, more than 40 % of the class, failed to estimate uncertainty without prompting. Performance improved through the end of week 5 (HW4 memo), but dipped again in week 7 (HW 6 memo). By the end of the course, class performance was faultless on this outcome.

Class performance was similar when

- Selecting an appropriate technique to estimate the uncertainty (error propagation vs. confidence interval or standard deviation).
- Identifying an appropriate situation to apply least-squares regression (find a mathematical function relating variables).
- Identifying an appropriate situation to apply hypothesis testing (test whether change in response is significant compared to experimental error).
- Distinguishing between paired samples (or blocking factors) and independent samples for hypothesis testing.

By the end of the course, better than 90 % of the class demonstrated these abilities. In addition, having selected the appropriate technique, 90 % of the class was able to execute, making a reasonable estimate of uncertainty, determining the values of adjustable parameters by least-squares linear and multiple linear regression, and drawing the correct conclusion from t-tests and ANOVA. These are the major statistical techniques required in Unit Operations Laboratory. If students enter Unit Operations Laboratory at this performance level, they will be well ahead of our previous seniors.

The only disappointing performance in this category was for the ability to evaluate the quality of models obtained through least-squares regression. The use of r-squared, residual plots, ANOVA, and confidence intervals on adjustable parameters was introduced in week 1 and revisited in two class periods and a homework assignment. The following question was asked on the graded in-class project in week 10.

> *It has been suggested that erosion yield can be predicted from polymer density using an equation of the form:*
>
> $(\text{erosion yield}) = a + b \times (\text{density}) + c \times (\text{density})^2$
>
> *Find values for the constants a, b, and c. Evaluate the quality of this model.*

Most (> 90 %) of students identified least squares multiple linear regression as an appropriate technique, and thus correctly determined the values of a, b, and c from the data given. However, just under 50 % of the students effectively used at least two of the four recommended techniques to evaluate model quality.

Fundamentals of statistical process analysis and control were introduced midway through he course, but not revisited outside of exam situations. Only 65 % - 75 % of students demonstrated competence with process capability indices and control charts on the second and final exams. After these students begin work, we will learn from surveys whether this gave them sufficient foundation for the training typically provided by companies for their employees.

Experimental design is difficult to properly address within the context of this course. The textbook practice problems focused on the interpretation of factorial experiments, on the analysis of experiments already done by someone else rather than on the design of the experiment. The instructor tried, with limited success, to convey the interplay between the experimental objective, the experiments to be conducted, and the techniques to be used to interpret them through two case studies. If these students enter Unit Operations Laboratory competent in selecting and executing basic statistical analysis techniques (descriptive statistics, least-squares regression, t-tests, ANOVA), more time can be spent in the senior year developing experimental design capabilities.

On the final exam, when asked to design an experiment to determine optimum operating conditions, 90 % of students appropriately selected a factorial design, and selected treatment levels to cover the feasible range of operating conditions. However, this may well be a false positive, since the topic had just been covered in the course. In fact, when then asked to design an experiment to determine the dependence of the response on a single factor, four students suggested a "factorial design" of 2 runs, rather than a design that would appropriately support least-squares regression. This suggests that these students, at least, were just grabbing a technique from recent course material. Randomization, replication, and blocking are three important concepts to experimental design. Many students noted the need for randomization without prompting. However, students gave sporadic attention to the importance of replication. Although told they could perform up to 20 experimental runs, two-thirds of students proposed a 2^3 factorial design with no replication, for only 8 runs. In contrast, when designing the experiment to support least-squares regression, 20 of 27 students included replication, sometimes at the expense of being able to identify nonlinear behavior over the operating range. Students are able to recognize blocking when it exists in data, since they are able to distinguish paired samples or a randomized block design from independent samples for hypothesis testing. However, it is unclear how many of them would propose it in their own experimental design.

Results: Interpretation of Results

On a pure problem-solving level, students leaving this course are able to interpret the results of the statistical techniques they use. When asked a direct question, they can correctly state whether two values are significantly different, for example. As their technical competence increased, however, it was noticeable that many students were unable to clearly convey their understanding in writing. Solutions to the higher-level problems were required in memo format. There were six of these memos during the quarter. Eight had been intended, and an additional two will be included next year. Students clearly need practice and instruction in writing technical memos.

Students struggled to write in a style that is formal, quantitative, and concise. "Concise" caused the most problems, and two students never lost their rambling style. Many seemed unaware of the conventional structure of a memo; only half (14) of the students put their conclusion in the first paragraph in the first memo. By the last memo, 25 of 27 students were consistently doing so. However, only about half of the students were able to explain, concisely and unambiguously, what statistical techniques they had used and how they had made their decisions. About half of the students did a good job of choosing whether their conclusions would most effectively be supported by text discussion, a graph, or a table. Apparently, this course has an important role to play in our overall curriculum goal of producing graduates who are competent technical writers.

Results: Student Assessment

At submission time, the college-administered course evaluation results were not available. A few of last year's juniors have volunteered during the course that many of this year's juniors were unhappy with their course grades, but are "actually learning something". Next year's exit survey of our graduates should be helpful in evaluating their perception of learning in the course, too.

Near the end of the course, students were surveyed about which course-related activities were most helpful to them in learning. The results are shown in Figure 1.

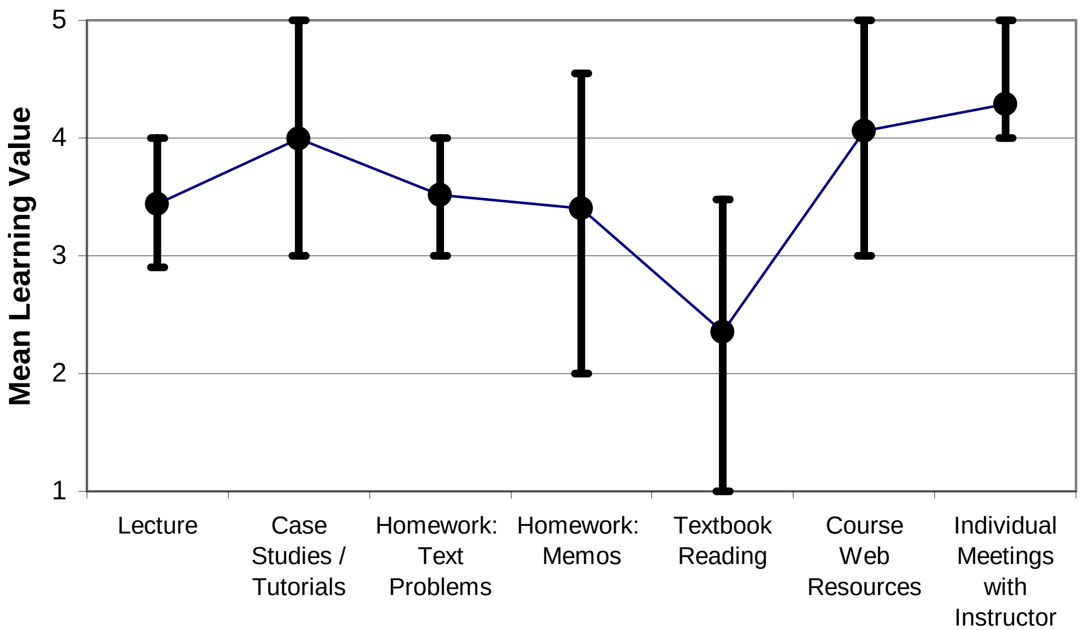

Figure 1. Student ratings of course activities for "learning value". 5 = high. 1 = low. Error bars represent 90^{th} and 10^{th} percentiles.

The case studies / tutorials were conducted in a computer room in a 110-minute class period, with the instructor circulating to offer information and assistance to individuals and small groups on a just-in-time basis. One option being considered is to change the class to meet in two 110-minute blocks weekly, and reduce the number of lectures in favor of more case study time. Of course, this requires developing more case studies.

Textbook reading was not considered very helpful, yet few students were in favor of getting rid of the textbook.

The homework memo problems received the widest range of response. While the students may not all appreciate them, the instructor feels that they were the single most valuable assessment tool in the course. For next year, the instructor's priority is to develop more of these high-level problems to be answered with a memo.

References

1. Wheeler, J.M. "Getting started: Six-Sigma control of chemical operations", *Chemical Engineering Progress*, June 2002, pp. 76-81.
2. Trivedi, Yogesh "Applying Six-Sigma", *Chemical Engineering Progress*, July 2002, pp. 76-81.
3. Bloom, B.S., M.D. Engelhart, E.J. Furst, W.H. Hill, D.R. Krathwohl, *Taxonomy of Educational Objectives: The Classification of Educational Objectives. Handbook I: Cognitive Domain*, David McKay, New York, 1956.
4. Prudich, M., D. Ridgway, V. Young, "Integration of Statistics Throughout the Undergraduate Curriculum: Use of the Senior Chemical Engineering Unit Operations Laboratory as an End-of-Program Statistics Assessment Course", *Proceedings of the 2003 ASEE Annual Conference and Exposition*, Session 3513, 2003.