THE COFFEE POT EXPERIMENT

A Better Cup of Coffee Via Factorial Design

R. J. McCLUSKEY, S. L. HARRIS
Clarkson University
Potsdam, NY 13676

Undergraduates have little exposure to statistical experimental design. For some time we had been searching for a means of introducing senior lab students to the concepts of this important topic. We needed a process with numerous variables on which many experiments could be run within a short time frame, yet which was safe, simple, and satisfied the space constraints of our laboratory setting. It was also important that any new experiment illustrate chemical engineering principles beyond those of our existing unit operations facilities. This article describes an undergraduate experiment that has successfully met the above objectives and yielded several extra benefits—some totally unexpected.

Experimental design involves the use of statistical methods in the planning and analysis of experiments so that valid results with known and generally smaller ranges of uncertainty are obtained in an efficient manner. Adherents of experimental design often emphasize its ability to reduce the number of required experiments, but its greatest value lies in forcing an experimenter to use more forethought in the scheduling of runs and greater rigor in interpretation of their results. The likelihood of an unhappy conclusion to an experimental program, *i.e.*, uninterpretable or meaningless results, is thereby considerably reduced.

Applications for experimental design range from the comparison of two treatments to more esoteric subjects such as model identification and time series analysis. Since the majority of our students have had no formal training in statistics, the goals of our experiment are quite limited. The aim is to expose students to useful references on applied statistics and to require them to carry out simple two-level factorial and fractional factorial designs. These elementary designs illustrate key statistical concepts, are usually the first designs encountered in any sequential strategy of experimentation, and serve as a foundation for various more specialized designs (*e.g.*, surface response designs).

Several texts offer explanations of factorial designs [1-3], and some recent articles [4, 5] extolling the merits of experimental design are listed with our references. A continuing series on experimental design in Chemtech [6] is also recommended, particularly for those familiar with matrices.

THE EXPERIMENTAL PROGRAM

Our senior laboratory course is organized along the following lines. Students work together in three or four member teams. Each student receives a lab manual containing experiment descriptions and additional references. Experiments are outlined in a manner that encourages creative approaches; cookbook procedures are kept to a minimum. All experiments are performed during two weekly six-hour lab periods. Students turn in a pre-lab report a day before the start of any experiment; prior to the second week of lab,

Richard McCluskey received his PhD in chemical engineering from the University of Minnesota in 1977 and has been a faculty member in the chemical engineering department at Clarkson since that time, where he currently serves as its executive officer. For several years he has been involved with the senior unit operations laboratory and since 1985 he has taught a graduate level course in Design of Experiments and Analysis of Data

Sandra Harris received her PhD from the University of California at Santa Barbara in 1978. Since that time she has been a member of the chemical engineering department at Clarkson University. She has been director of the senior unit operations laboratory several times and is active on the Curriculum Committee and on the Laboratory Improvement Committee. Other interests include process dynamics and control.

they make an oral progress report to the faculty member serving as experiment director; and a final report is due a week after completion of the experiment.

The problem posed to the students is excerpted below:

> *Consider your lab group as a team of engineers hired by the Peyton Hall Coffee Company to develop a process for making coffee. The prospective customers are a discriminating lot who are unwilling to pay more than $.20 per cup. The company has invested a considerable sum in a Norelco Dial-A-Brew II coffee maker, so this production unit must be used. Don't break it. Process options that should be considered are*
>
> *1. The type of coffee:*
> *Coffee C1 ($2/lb) and Coffee C2 ($4/lb)*
> *2. The type of filter:*
> *Filter F1 ($.01/sheet) and F2 ($.02/sheet)*
> *3. The type of water:*
> *W1, from the water fountain at no cost*
> *W2, distilled water at $.01/cup*
> *4. At least two other process changes or variables of your choice.*

In their final report students are asked to present a process cost summary, to discuss recommendations for further process development, and to explain any discrepancies in their mass and energy balances.

The apparatus for the experiment consists principally of the drip coffee maker, four pint-size thermos bottles, a household thermometer, a water jug, a dishwashing bucket, and many styrofoam cups. Other instruments the students are expected to use include a spectrophotometer, an ammeter, balances, a magnifying glass, and a microscope.

A week before starting the experiment, students are given the operating instructions for the coffee maker and a copy of a statistics text. We find the book by Box, Hunter, and Hunter to be highly readable, and pages 306-328 plus pages 374-386 are a sufficient reading assignment for our purposes.

During the first lab session, students perform a 2^3 factorial experiment. This consists of eight runs (pots of coffee, 3 cups/pot) covering all possible combinations of three variables at two different settings. Encouragement is given to replicate at least one run. With each pot of coffee, students must do overall mass and energy balances, measure several characteristics of the product, and obtain samples for a taste test. Property measurements normally include spectrophotometric percent transmission on a sample diluted with three parts water and a residual weight obtained from a Mettler balance with an infrared dryer attachment. pH measurement is convenient, but usually shows only minor variation with type of coffee. To date, no groups have taken on the challenge of a chromatographic analysis.

We needed a process with numerous variables on which many experiments could be run within a short time frame, yet which was safe, simple, and satisfied the space constraints of our laboratory setting.

Students are forewarned that reliable taste test results are obtained only with difficulty. Taste test procedures are left entirely as their responsibility.

After using the first lab period to refine their experimental procedures and to acquaint themselves with factorial experiments, students have an opportunity to meet with the experiment director before attempting a 2^{5-1} fractional factorial design during the second week of lab. The effects of five different variables are now to be evaluated from the results of sixteen runs. The variables are chosen by the students. Some fairly obvious choices are the "brew control" setting on the coffee maker and the change of coffee grounds. Other parameters that have been investigated include: addition of salt to the coffee grounds, pre-wetting the bed of grounds, comparison of fine versus coarse ground coffee, and use of stainless steel versus glass carafe.

SAFETY

Normal precautions for use of any electrical appliance need to be followed. Students are also cautioned that the glass carafe is fragile and cannot withstand significant thermal shock. We have kept a stainless steel carafe on hand as a substitute, but have not had to use it.

The foremost hazard that we guard against is chemical contamination. To that end, the apparatus is stored in a nearby faculty member's office rather than in the lab stockroom, we deal exclusively with black coffee, and the experiment is set up in an area removed from the rest of the lab.

THEORY

For a 2^3 factorial experiment involving three variables each at two levels or settings (arbitrarily denoted by + and –), the basic model for interpreting any product characteristic, Y, is Eq. (1).

$$Y_{+++} = \langle Y \rangle + \frac{A + B + C + ab + ac + bc}{2} + \text{error} \qquad (1)$$

where Y_{+++} is the value obtained in the experiment with the first, second and third variables at their + setting; $\langle Y \rangle$ is the average Y value for all experiments; A,B, and C represent the "main effects" of the three variables; ab, ac, and bc are the "two factor interaction effects" of the respective variable sets; and the error term is assumed to be an independent and normally distributed random variable.

The main effect, A, is defined as $<Y_{A+}> - <Y_{A-}>$, where $<Y_{A+}>$ is the average result for all experiments with the first variable at a + setting. The interaction effect, ab, is defined as $<Y_{AB+}> - <Y_{AB-}>$, where $<Y_{AB+}>$ is the average for all experiments in which the product of the first two variable settings is positive, *i.e.*

$$< Y_{AB+} > = \frac{(Y_{+++} + Y_{++-} + Y_{--+} + Y_{---})}{4} \qquad (2)$$

ab is a measure of the non-additivity of the response to changes in the first and second variables.

A much fuller discussion of factorial designs is presented in applied statistics texts [1-3]. Eq. (1) is presented to illustrate that the primary goal is determining which variables give large main effects and which variables may strongly interact. Note the above concepts apply whether a variable is quantitative, such as the mass of coffee grounds, or qualitative, such as the type of coffee.

Luckily, there is a convenient technique, known as Yates Algorithm [2, pp. 323-324], for computing each of the main and interaction effects. All of the terms in Eq. (1) can be easily calculated from a table of Y values. Variables whose main and interaction effects are small can be eliminated from consideration, and process development efforts can focus on those variables giving large effects. Early identification of unanticipated synergism (*i.e.*, large interaction effects) can speed efforts of process optimization, or, even better, serve as a basis for obtaining a patent.

Nonetheless, experimenters must always be cautious in interpretation of their results. If unexplained interaction is encountered, it is wise to redo the analysis using a simple transformation of Y, *e.g.*, *ln* Y or 1/Y. It may be that a model strictly additive in main effects is adequate for the transformed data.

The error term in Eq. (1), what statisticians call the residual, should also be examined to see if it is approximately normally distributed. An unusually large residual may indicate a flawed experiment that needs to be repeated. Also, if residuals vary in relation to the magnitude of Y, there is a need for data transformation of the type mentioned above. In order to do such an analysis in a 2^3 design, some experiments must be repeated. If there are only eight experiments, the error has but one degree of freedom and is equal to $(<abc+> - <abc->)/2$, where $<abc+>$ is the average of all experiments in which the product of the variable settings is positive, *i.e.*

$$< abc + > = \frac{(Y_{+++} + Y_{+--} + Y_{-+-} + Y_{--+})}{4} \qquad (3)$$

PITFALLS

Students are prone to concentrate on the taste testing and lose sight of their overall objective—development of a process for the Peyton Hall Coffee Co. Also, taste test results are apt to be indecisive for a number of reasons: taste bud fatigue, lack of contrast between samples, or simply too few examiners for such a subjective test. A positive aspect to this is that process decisions can then be made on the basis of economics and more reliable measurements of coffee strength. The taste test still serves as a safeguard against concocting an unpalatable product.

Review of initial student reports disclosed their unfamiliarity with cost analysis, standard errors, and use of microscopes. It is now explicitly mentioned that labor costs must be included in their analyses. Following the oral progress report, a lengthy discourse is made on procedures for estimating standard errors and the advantages of visual comparison.

DISCUSSION

A motivating factor behind this experiment was the increasing number of our graduates finding employment in the food processing or consumer products industries, where consumer preference testing and statistical analysis of data are commonplace. The experiment drives home the points that many chemical engineers work outside the chemical industry and that you need go no further than your own kitchen to do interesting chemical engineering research.

The taste testing also introduces a social component absent in our other experiments. As expected, offering coffee samples to volunteers led to greater alertness at the end of our six-hour lab period, particularly when students had been up late the night before completing final reports on prior weeks' experiments. An unexpected plus was the willingness of faculty members, who had never before participated in lab, to throw open their office doors and take part in the taste testing.

Computational burdens associated with analysis of factorial experiments are avoided by providing students with a Multiplan software file that calculates all main and interaction effects using Yates Algorithm.

We have not considered espresso or gourmet coffees, and it has been difficult locating common brands that give substantial taste differences. We will be happy to share our knowledge of which brand is unusually distasteful with anyone who writes. A copy of the write-up appearing in our lab notebook is also available.

This experiment, like all good research projects, has numerous facets worthy of additional exploration. Fundamental modeling of the drip making process, scale-up considerations, and effects of aging are all topics of interest. Another possible extension would have students report on the commercial processing of coffee beans [7]. Furthermore, there remains a controversy over whether professors' taste buds differ significantly from those of undergraduates.

In summary, our experience with this experiment has been quite positive. Almost all lab groups are able to sort through many process variables and make a reasonable recommendation to the Peyton Hall Coffee Co. Some students come away with an appreciation of factorial experiments and some do not. They all share the experience of utilizing dollar signs in the evaluation of their results. Operation of the experiment goes well, and, importantly, most students report they have fun.

REFERENCES

1. Bethea, R. M., B. S. Duran, and T. L. Boullion, *Statistical Methods for Engineers and Scientists,* Marcel Dekker, New York (1975)
2. Box, G. E. P., W. G. Hunter, and J. S. Hunter, *Statistics for Experimenters,* John Wiley & Sons, New York (1978)
3. Cheremisinoff, N. P., *Practical Statistics for Engineers and Scientists,* TECHNOMIC Publishing Co., Lancaster, PA (1987)
4. Hunter, J. S., "Applying Statistics to Solving Chemical Problems," *Chemtech* , pp 167-169, March 1987
5. Jones, B. A., "Design Lab Experiments to Assure Product Quality," *Research & Development,* pp 54-58, December 1987
6. Deming, S.N., "Quality by Design," *Chemtech,* part 1, p. 56, September (1988); Part 2, p. 52, January (1989)
7. *Kirk-Othmer Encyclopedia of Chemical Technology,* 3rd Ed., Vol. 6, p 511, John Wiley & Sons, New York (1979) ❒

REVIEW: Reactor Engineering

Continued from page 149

models, which somewhat anticipates the next two chapters. This chapter also covers the literature extensively, but the authors give good physical explanations of the models.

The next few chapters deal with some old favorites. The task of covering fixed-bed gas-solid catalytic reactors falls to Froment and Hofmann, who present a range of models. This chapter is not as comprehensive as the corresponding one in Froment and Bischoff's book (*Chemical Reactor Analysis and Design,* Wiley, 1979). The authors emphasize uncertainties in reaction kinetics and reactor transport parameters, and deal mainly with phenomena actually seen under industrial conditions. In a well-written chapter, Rowe and Yates give a very clear description of the bubbling bed model and bubble behaviour in fluidized beds. They also emphasize the importance of the distributor and freeboard regions in reactor design. Denn and Shinnar cover the area of coal gasification with a strong emphasis on reactor efficiency and energy analysis. Some of their references may be difficult for students to find, and some familiarity with specialized terms is assumed.

In the longest chapter, Carra and Morbidelli give an extensive catalog of correlations, model equations, and solutions for gas-liquid reactors. Some interesting comments on scale-up and reactor power requirements are made at the end.

Shah and Sharma treat trickle bed and slurry reactors under the heading of gas-liquid-solid reactors. Again, many references and results are given. The part on slurry reactors contains many typographical errors, but also provides a useful worked example of slurry reactor design.

Some newer and more specialized reactor types are the subject of the concluding chapters. Tyrell, Galvan, and Laurence provide a short chapter that tries to make clear the focus of polymer reaction engineering, rather than being comprehensive. They concentrate on modeling the product distribution by step-growth and chain-growth mechanisms. There is relatively little on actual reactor design.

Erickson and Stephanopoulos give an introduction to biological reactors that is reasonably accessible for non-specialists. A short introduction to microbial growth concepts and a review of mass transfer (0_2 supply) set the stage for an interesting chapter on reactors. In possibly the first treatment in a reaction engineering book, DeBarnardez, Claria, and Cassano give a necessarily expository view of photochemical reactors. They stress the differences between these and more traditional reactor designs, especially the need for proper models of emission and absorbance of radiation. Several examples of photoreactor design are presented.

Trost, Edwards, and Newman give a very readable survey of electrochemical reactors, with strong emphasis on porous electrode systems, following a general introduction.

There is unfortunately no chapter in this book on deposition reactors, such as are used in the semiconductor industry. Maybe next time!

The final chapter is by Morbidelli, Varma, and Aris, and is somewhat different from the rest of the book in that it is completely theoretical, covering reactor steady-state multiplicity and stability. The two limiting cases of CSTR and plug-flow reactor are analyzed extensively. ❒