INTRODUCING STATISTICAL CONCEPTS IN THE UNDERGRADUATE LABORATORY

Linking Theory and Practice

Annette L. Burke, Aloke Phatak,
Park M. Reilly, Robert R. Hudgins
University of Waterloo
Waterloo, Ontario, Canada N2L 3G1

Laboratory experiments are an integral part of the chemical engineering curriculum because they serve several purposes. Their primary purpose is to reinforce key chemical engineering concepts, but they are also supposed to teach students about model development and how to obtain reliable data in the presence of experimental error. With the present-day emphasis on quality control in chemical industries and manufacturing, these skills are needed by every chemical engineer who will collect and analyze data, and they are especially important for engineers involved in process modeling and development.

Unfortunately, with the exception of one course in statistics, we do very little to teach undergraduates about data collection and analysis. At Waterloo, an introductory course in statistics is given in the second year, and topics include an introduction to probability distributions, properties of means and variances, estimation, confidence intervals, significance tests, and linear regression. These tools provide a background in collecting and analyzing data—but students forget most of the material because they never get a chance to apply it. As a result, they complete their undergraduate training without really grasping the connection between statistics and experimentation.

In the students' defense, most laboratory experiments are not designed using the same principles that we teach in class. For example, in a statistics course we might emphasize the importance of reporting confidence intervals for a parameter that we have estimated, but laboratory experiments are rarely designed to allow students to do just that. Consequently, they are left wondering about the practical "real-world" value of statistical techniques. To bring statistics down from the blackboard and onto the lab bench, therefore, we must include statistical concepts in undergraduate laboratory experiments. If we do not reinforce this link between the theory and the practice of statistics, we will be doing a disservice to our students who, as practicing engineers, will have to deal with measurement error on a daily basis.

***Annette Burke** obtained her BASc in chemical engineering from the University of Waterloo in 1990. She is working on the development of improved methods for designing sequential model discrimination experiments. Her research interests include a variety of issues related to process modeling and experimental design.*

***Aloke Phatak** obtained his BASc and MASc in chemical engineering from the University of Waterloo. After working as a research scientist in the field of rocket propellants for three years, he returned to UW for his doctorate. He is currently working on applications of multivariate statistics in ChE.*

***Bob Hudgins** holds degrees from the University of Toronto and Princeton University. He teaches reaction engineering, staged operations and laboratories that go with them. His research interests lie in periodic operation of catalytic reactors and in the improvement of gravity clarifiers.*

***Park Reilly** graduated from the University of Toronto in 1943 and worked in industry until 1967 when he joined the faculty at the University of Waterloo. He studied statistics at the University of London and received a PhD in Statistics in 1962. His research and publications are in the area of applied statistics.*

OBJECTIVES

The purpose of this paper is to show that it is possible to incorporate statistical ideas within existing experiments, while still respecting the need to illustrate chemical engineering concepts. We have made changes in two second-year physical chemistry experiments. These experiments are particularly appropriate, not only because they are a part of physical chemistry courses in many departments, but also because here at Waterloo they are carried out in the term following the introductory statistics course. Thus, students begin applying statistical tools very early in the curriculum. We hope that through this early exposure they will come to view sound statistical analysis as a necessary part of all experimentation. Our long-term objective is to incorporate more advanced concepts, such as design of experiments and response surface methodology, into all laboratory courses—especially into the unit operations laboratories in the third and fourth years.

In the following paragraphs we outline the old procedures, the changes we have made, and the statistical concepts that have been introduced. We believe that the new procedures are better, but we also suggest additional modifications which could be made to further improve the didactic value of the experiments. Complete laboratory procedures, which include laboratory questions and supplementary material, can be obtained from the authors.

EXPERIMENT 1

Determination of the Molecular Weight of Polystyrene by Viscometry

In this experiment, students determine the viscosity-average molecular weight $(\overline{M}_v)$ of a sample of polystyrene by dilute solution viscometry. Some of the concepts introduced in this experiment (the rheology of suspensions, for example) are also discussed in the physical chemistry course and in a fluid mechanics course. In addition, the students learn a little bit about polymers and polymer-solvent interactions. The standard experimental procedure for determining $\overline{M}_V$ is described by Smith and Stires.[1] It is quite commonly used in both industrial and research laboratories.

In dilute solution viscometry, the idea is to relate $\overline{M}_V$ to the viscosity of a very dilute solution of polymer and solvent. The viscosity of a polymer solution increases with both the concentration and the molecular weight of the polymer. By measuring the viscosity of a polymer solution at several concentrations, however, and then extrapolating to zero concentration, the effect of molecular weight can be isolated, thereby allowing us to estimate $\overline{M}_V$.

The viscosity-average molecular weight of a polymer is related to the *intrinsic* viscosity of a polymer by the Mark-Houwink-Sakurada (MHS) equation[2]

$$[\eta] = K\overline{M}_v^a \tag{1}$$

Here, $[\eta]$ is the intrinsic viscosity, and K and a are constants which depend upon the polymer, solvent, and solution temperature. The intrinsic viscosity, sometimes known as the *limiting* viscosity number, is defined in terms of the Newtonian viscosity of a polymer-solvent solution of concentration, c, as the concentration approaches zero, *e.g.*,

$$[\eta] = \lim_{c\to 0} \frac{(\eta/\eta_o - 1)}{c} \tag{2}$$

where η is the viscosity of the polymer-solvent solution of concentration c, and η_o is the viscosity of the solvent alone. Once we know the intrinsic viscosity of a polymer in a given solvent and the MHS constants K and a, we can calculate its viscosity average molecular weight by solving Eq. (1) for $\overline{M}_V$. But how can we determine the intrinsic viscosity in the first place?

The Newtonian viscosity of a polymer-solvent solution depends on the concentration of the polymer. For very dilute solutions this concentration dependence can be described by the Huggins equation,[3] which is written as

$$\frac{(\eta/\eta_o - 1)}{c} = [\eta] + k_H[\eta]^2 c \tag{3}$$

where the constant k_H is known as the Huggins constant. Thus, to determine the intrinsic viscosity, we first measure the viscosity of the solvent as well as the viscosities of at least two polymer solutions of known concentration. Then, assuming that the Huggins equation is correct, we can use linear regression to estimate $[\eta]$ in Eq. (3).

In the experiment, however, η and η_o are never actually measured. In the viscometer used, the time required for the polymer solution to flow through a marked length of glass tubing is measured. It turns out that in such a viscometer, the flow time is proportional to the viscosity of the solution and inversely proportional to its density. However, because the different polymer solutions used are very dilute, their density is roughly the same, and flow time depends on the viscosity of the solution only. Thus, we can write $\eta/\eta_o = t/t_o$, where t is the flow time for a polymer solution and t_o is the flow time for the pure solvent. As a result, Eq. (3) can be written in terms of flow times instead of viscosities, *e.g.*,

$$\frac{(t/t_o - 1)}{c} = [\eta] + k_H[\eta]^2 c \tag{4}$$

and we can carry out a linear regression as outlined above to estimate $[\eta]$.

Old Procedure: A 50-ml solution consisting of 0.5 g of polymer in solvent (toluene, for example) is prepared and left for a day to allow the polymer to dissolve. A 10-ml aliquot of pure solvent is then placed in a Cannon-Ubbelohde viscometer, and the flow time is measured three times. These measurements are then averaged. All flow times are determined in this way since uncertainty in these measurements is the major source of error in this experiment. The solvent is removed and the viscometer cleaned. Next, 10 ml of polymer solution is placed in the viscometer and the flow time is measured. The solution is diluted by the successive addition of

2, 5, 5, 10, and finally 20 ml of solvent. After each addition the solution is mixed and the flow time measured. Then, the data are plotted using the Huggins equation, and by using linear regression, $[\eta]$ is estimated. Figure 1 is a plot of typical data. Once $[\eta]$ has been determined, $\overline{M}_v$ can be calculated using the Mark-Houwink-Sakurada equation. Students are supplied with appropriate values of the constants K and a from the *Polymer Handbook.*[4]

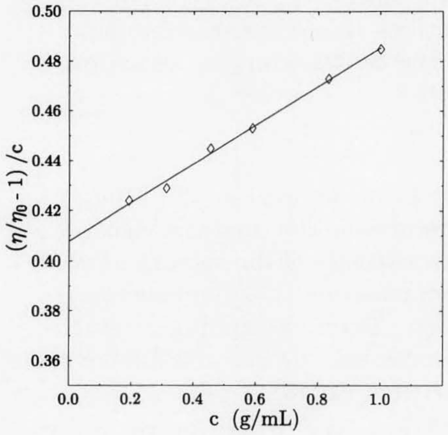

***Figure 1.** Huggins plot of viscosity data generated using the old experimental procedure. The polymer system is polystyrene in toluene at 30°C.*

Why is such a design inherently flawed? Reilly, *et al.,*[5] pointed out that as more and more solvent is added to the initial polymer solution, the error in measuring flow times increases, as does the uncertainty in the concentration. As a result, the error in the quantity $(\eta/\eta_o - 1)/c$ increases as concentration decreases, and making a large number of measurements at low concentrations decreases the precision with which we can determine $[\eta]$. Our objective in modifying this experiment, therefore, was to implement an improved procedure suggested by Reilly, *et al.,*—one which yields more precise estimates of $[\eta]$ and $\overline{M}_v$ and which, more importantly, allows students to construct confidence intervals for these two quantities.

New Procedure: How many polymer-solvent solutions should we run through the viscometer to estimate the intrinsic viscosity with the greatest precision? The answer, according to Reilly, *et al.,*[5] is only *two!* The first solution has a concentration given by c_1, and the second a concentration of $c_1/2$, which we denote by $c_{1/2}$. Once the flow times of these solutions and of the pure solvent have been measured, the problem of estimating $[\eta]$ by regressing $(t/t_o - 1)/c$ on c reduces to fitting a straight line between *two* points. After a little bit of algebra, it is easy to show that the intercept of this line, the intrinsic viscosity, is given by

$$[\eta] = \frac{-t_1 + 4\,t_{1/2} - 3\,t_o}{c_1 t_o} \tag{5}$$

Here, t_o denotes the average flow time of pure solvent, while t_1 and $t_{1/2}$ are the average flow times of the solutions of concentration c_1 and $c_{1/2}$, respectively. In the modified procedure, c_1 corresponds to an initial solution of 0.35 g polymer in 50-ml solvent. Two solvents are used: toluene and 80/20 by volume mixture of toluene and methanol. Thus, in all, students measure the average flow times of six solutions, where each average has been calculated from three measurements. This allows them to estimate the variance of the flow measurements as

$$s^2 = \frac{1}{n(k-1)} \sum_{i=1}^{n} \left[\sum_{j=1}^{k} \left(t_{ij} - \bar{t}_i \right)^2 \right] \tag{6}$$

where t_{ij} is the j^{th} replicate measurement of the i^{th} solution, $\bar{t}_i$ is the average of $k(=3)$ replicate measurements of solution i, and $n(=6)$ is the total number of solutions. Our practice is to combine the data from two groups of students to get a more reliable estimate of the variance based on $n = 12$ solutions.

Having determined s^2, it is relatively straightforward to estimate the variance of $[\eta]$ by applying standard formulas for the variance of the quotient of random variables to Eq. (5). If we do so, it turns out that

$$\text{var}\,[\eta] = \frac{s^2 V}{t_o^2\, c_1^2} \tag{7}$$

where V is a factor that depends on the number of replicates of flow time measurements, c_1, and the estimate of $[\eta]$ obtained by using Eq. (5). Then, confidence intervals for $[\eta]$ and $\overline{M}_v$ can be constructed in the usual way.

Discussion: The new procedure is better in two important respects: 1) the estimate of $[\eta]$, and hence of $\overline{M}_v$, is more reliable, and 2) the students can now construct confidence intervals based on an estimate of the variance that is *independent* of the regression that is carried out to estimate $[\eta]$. In addition, we also briefly discuss the old procedure so that the students can understand why its design is flawed and why the new procedure yields a more precise estimate of intrinsic viscosity.

How could we further improve the didactic value of the experiment? One way would be to explain to the students why a design in which only two concentrations are used is optimal. In addition, we could also make them derive Eq. (7), including the exact value of the factor V. However, we have to strike the right balance between illustrating statistical concepts and illustrating physical principles. Although we are convinced of the value of introducing statistical ideas into these experiments, we do not want to do so at the expense of the chemical engineering concepts being illustrated. Thus, we leave it up to the instructor to decide whether or not to incorporate the additions mentioned above.

EXPERIMENT 2

Adsorption of Acetic Acid on Charcoal

This experiment illustrates the discussion of adsorption from solution that is presented in the second-year physical chemistry course. Like distillation, adsorption can also be thought of as a chemical engineering unit operation.[6] Two expressions—the Freundlich isotherm and the Langmuir isotherm—are used to describe the adsorption of acetic acid onto activated charcoal. Each isotherm is based on different assumptions about the nature of adsorption, and they apply under different conditions. The Freundlich isotherm[7] is a purely empirical model which works well at low solute concentration. It relates the mass of solute adsorbed (x) on the adsorbent to the equilibrium concentration of solute (C), *e.g.,*

$$\frac{x}{m} = kC^n \tag{8}$$

where m is the mass of adsorbent, and k and n are empirical constants. The Langmuir isotherm,[7] however, was derived assuming an explicit adsorption mechanism. It may be written as

$$\frac{x}{m} = \frac{(x/m)^0 KC}{1+KC} \tag{9}$$

where K is the equilibrium constant, and $(x/m)^0$ is the mass ratio required for monolayer coverage of the surface of the adsorbent. Equations (8) and (9) are nonlinear, and they are usually used in their linearized form, *e.g.,*

$$\ell n \frac{x}{m} = \ell n\, k + n\, \ell n\, C \tag{10}$$

for the Freundlich isotherm, and

$$\frac{1}{x/m} = \frac{1}{K(x/m)^0}\frac{1}{C} + \frac{1}{(x/m)^0} \tag{11}$$

for the Langmuir isotherm.

After measuring x/m for several different concentrations of acetic acid, the students are asked to comment on the fit of Eqs. (10) and (11) to the data. The procedure used in the past is based on the experiment described by Ellis and Mills;[8] it is not well designed to allow the students to quantitatively assess which of the two isotherms better describes adsorption of acetic acid onto charcoal. Furthermore, note that the *linearized* forms of the equations are used. Also, as we will see from the procedure below, x and C are not statistically independent.

Old Procedure: Two grams of activated charcoal are placed in each of six flasks. Starting with 0.5 M acetic acid, six 100-ml lots of acetic acid with concentrations ranging from 0.5 to 0.025 M are prepared. The acetic acid solutions are added to the charcoal, mixed, and left to stand overnight to reach equilibrium. The solutions are then suction filtered. Filtrate samples are titrated with 0.2 M NaOH to determine the equilibrium concentrations. Finally, the amount of acetic acid adsorbed onto the charcoal is calculated.

Students then plot the data using Eqs. (10) and (11). The correlation coefficient for the Freundlich isotherm is typically 0.99, and for the Langmuir isotherm it is typically 0.70. Figures 2(a) and 2(b) show plots of representative data. Students notice the curvature in the plot of the Langmuir isotherm and then conclude that it is not due to chance alone, but to systematic departure from the fitted model. Unfortunately, they rarely realize that the observed curvature provides information which is different from that provided by a low correlation coefficient. As a result, they often proceed in later years to rely heavily on the correlation coefficient as a measure of model-fit and sometimes even neglect to plot data. Our purpose in modifying this experiment, therefore, is to emphasize the limitations of the correlation coefficient and to give the students experience in using other measures of model fit.

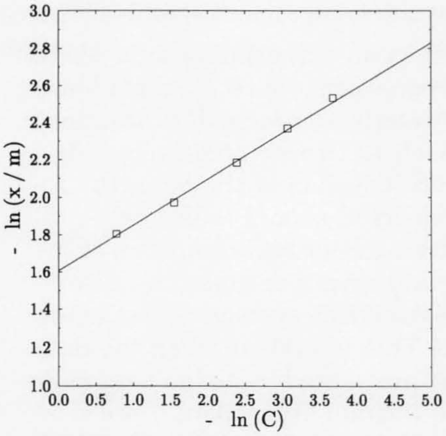

Figure 2(a). *Acetic acid adsorption data generated using the old procedure and plotted according to Freundlich isotherm.*

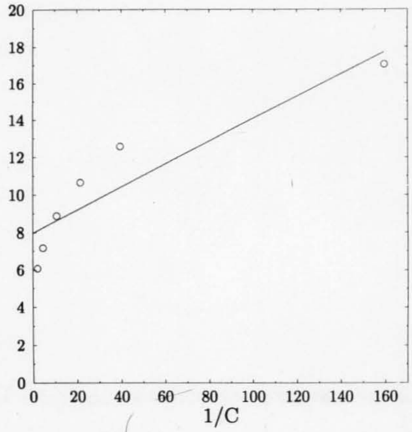

Figure 2(b). *Acetic acid adsorption data generated using the old procedure and plotted according to Langmuir isotherm.*

New Procedure: The procedure is unchanged except for the number of solutions used. Instead of preparing six solutions of different concentration, three independent replicates of four different concentrations are prepared. The concentrations used are between 0.5 and 0.025 M. For each replicate, charcoal is weighed out and acetic acid solution is prepared separately to ensure independence. The twelve samples are left overnight to reach equilibrium and are then suction filtered; the filtrate from each is again titrated using 0.2 M NaOH. It is tempting here to titrate a set of three replicates sequentially, but this would invalidate the estimate of the error variance. Filtrate samples must be titrated in random order so that the correlation between any two measurements is constant, and the data may be treated as independent. Finally, students perform least-squares regression to fit Eqs. (10) and (11) to the data, calculate the correlation coefficients, plot the residuals, and perform the lack-of-fit test described below. Figures 3(a) and 3(b) show typical results using Eqs. (10) and (11).

The lack-of-fit test is an extension of analysis of variance in linear regression, which students learn in their introductory statistics course. It is described in standard texts such as Draper and Smith.[9] If a model is a good representation of the data, the residuals, or prediction error, should reflect only random error. If a model is a poor representation of the data there is additional variation caused by lack-of-fit, which manifests itself as a systematic departure from the fitted line. This is evident when the data from this experiment are plotted using the Langmuir isotherm, but in the original experiment there is no way to estimate random error independently of the model or to confirm lack-of-fit quantitatively.

The introduction of replication allows us to estimate the random error, or pure error, independently of the model that is postulated. The model prediction errors can then be divided into random error and lack-of-fit error. Comparison of the lack-of-fit sum of squares to the pure error sum of squares using an F-test serves as a quantitative measure of the fit of the model.

The calculations for the lack-of-fit test are straightforward. Let y_{ij}, $i = 1,2,\ldots,k$, $j = 1,2,\ldots,n_i$ be the j^{th} measurement on the 'dependent' variable at the i^{th} concentration. In the modified procedure outlined above, three replicates ($n_i = 3, \forall i$) are performed at each of four ($k = 4$) different concentrations. Recall that because we are using the linearized forms of the Freundlich and Langmuir isotherms, $y = \ln(x/m)$ for Eq. (10) and $y = 1/(x/m)$ for Eq. (11).

The random, or pure, error can be estimated by

$$s^2 = \frac{\sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(y_{ij}-\bar{y}_i\right)^2}{n-k} \qquad n=\sum_{i=1}^{k} n_i \tag{12}$$

where $\bar{y}_i$ is the average of the n_i measurements at the i^{th} concentration. In much the same way, the lack-of-fit sum of squares (LFSS) is estimated as

$$\text{LFSS} = \sum_{i=1}^{k} n_i\left(\bar{y}_i-\hat{y}_i\right)^2 \tag{13}$$

where $\hat{y}_i$ is the value predicted by the model (Eqs. 10 or 11) at the i^{th} concentration. Once we have fitted the Freundlich and Langmuir isotherms and calculated the corresponding values of s^2 and LFSS, we can construct an F-statistic and compare it to the critical F-value at the desired confidence level, *e.g.*,

$$f = \frac{\text{LFSS}/(k-2)}{s^2} \sim F_{k-2,n-k} \tag{14}$$

If the calculated value is larger than the critical F value, then lack-of-fit error is significant at the chosen confidence level, and the model does not describe the data adequately. Using such an F-test, students find significant lack-of-fit for the Langmuir isotherm, which provides *quantitative* reinforcement of the conclusions they draw by simply observing the curvature in Figure 3(b).

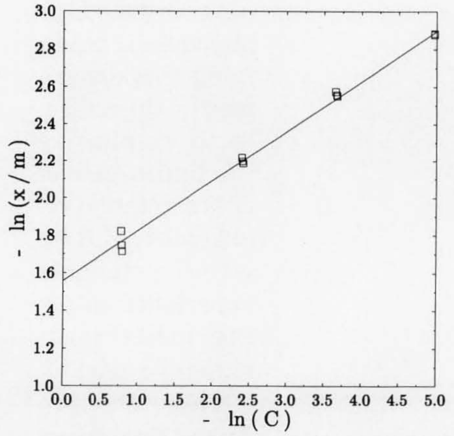

Figure 3(a). *Acetic acid adsorption data generated using the new procedure and plotted according to Freundlich isotherm.*

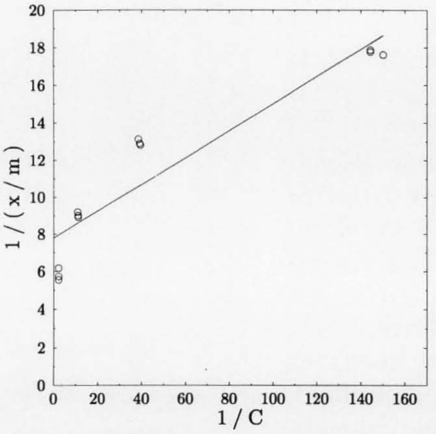

Figure 3(b). *Acetic acid adsorption data generated using the new procedure and plotted according to Langmuir isotherm.*

DISCUSSION

The changes we have made to the experimental procedure are minor; the benefits reaped by the students, however, will be substantial. First, the students will be introduced to replication, which is essential in estimating random error and in identifying problems such as nonconstant variance. More important, we hope that students will realize just what experimental error really is when they carry out repeat measurements which *do not* yield the same results. Second, the introduction of the lack-of-fit test and the analysis of residual plots encourages students to use tools other than the correlation coefficient in discriminating between competing models. Finally, students are forced to review least-squares regression and analysis of variance in order to understand the lack-of-fit test and to interpret the residual plots.

We recognize, however, that some flaws which are present in the old procedure remain in the modified one. First, linearized forms of the original expressions are still fitted, which may change the error structure. Second, the amount of solute adsorbed is determined from the change in solution concentration, which causes the variables x and C to be statistically dependent. Finally, we can see in Figure 3(a) that the variance of the data, plotted according to the Freundlich isotherm, increases as the concentration, C, increases. This violates one of the assumptions of least-squares regression—that of constant variance.

How can we remedy these deficiencies? One way of doing so is to express the models in terms of the actual quantities measured: the initial acetic acid concentration, the equilibrium concentration, and the mass of charcoal. Since the resulting model will be nonlinear with error in all the variables, an analysis using the error-in-variables method[10] would be most appropriate. It would be unrealistic, however, to expect second-year students to carry out such an analysis. Here, we face a question that we will no doubt encounter when trying to incorporate statistical concepts into other experiments: how can we adopt the best, "statistically correct" analysis of a poorly designed experiment without burdening our second-year students with statistical methodology that would tax even a competent researcher? Our solution here is a compromise—we have incorporated changes that we think are better, but we also recognize the remaining deficiencies and encourage the students to think about and discuss other ways of analyzing the data and why they might be more appropriate. In this way we hope that they will be able to recognize how the design of an experiment can affect the statistical analysis of data derived from it.

CLOSING REMARKS

Our objective in modifying these two experiments was to introduce statistical concepts into the undergraduate laboratory. The changes to the procedures themselves are minor, but by modifying the analysis of the data it is possible to include a wealth of ideas which reinforce the connection between statistics and experimentation. By introducing replication, we force the students to confront experimental error—they *see* that measurement uncertainty is an unavoidable fact of life. By showing them the means to quantify this error, we show them a rational basis for dealing with it.

In the long term, our objective is not only to make notions like replication and interval estimates an essential element of all undergraduate laboratories but also to include advanced concepts such as factorial designs, especially in upper-year unit operations laboratories. As we saw with the analysis of Experiment 2, however, it is sometimes difficult to incorporate statistical concepts into existing procedures that are poorly designed to begin with. Short of redesigning all undergraduate experiments or introducing students to advanced statistical techniques which they may not be able to appreciate, our solution has been to incorporate statistical techniques, but at the same time point out deficiencies where they exist and encourage the students to discuss alternative methods of data analysis. However we choose to do it, it is clear that we must incorporate statistical concepts into the undergraduate laboratory. By doing so as early as possible in the chemical engineering curriculum, we hope to remove the mistaken notion of statistics as something complex and mysterious when it is really *fundamental* to the engineer's craft.

REFERENCES

1. Smith, J.L., and A. Stires (Eds.), *Experimental Physical Chemistry,* 7th Ed., McGraw-Hill, New York (1970)
2. Rudin, A., *The Elements of Polymer Science and Engineering,* Academic Press, New York (1982)
3. Huggins, M.L., "The Viscosity of Dilute Solutions of Long-Chain Molecules. IV. Dependence on Concentration," *J. Am. Chem. Soc.,* **64**, 2716 (1942)
4. Brandrup, J., and E.H. Immergut (Eds.), *Polymer Handbook,* Wiley-Interscience, New York (1975)
5. Reilly, P.M., B.M.E. Van der Hoff, and M. Ziogas, "Statistical Study of the Application of the Huggins Equation to Measure Intrinsic Viscosity," *J. Appl. Polym. Sci.,* **24**, 2087 (1979)
6. Mantell, C.L., *Adsorption,* 2nd Ed., McGraw-Hill, New York (1951)
7. Castellan, G.W., *Physical Chemistry,* 3rd Ed., Addison-Wesley, Reading, MA (1983)
8. Ellis, R.B., and A.P. Mills, *Laboratory Manual in Physical Chemistry,* McGraw-Hill, New York (1953)
9. Draper, N.R., and H. Smith, *Applied Regression Analysis,* 2nd Ed., John Wiley & Sons, New York (1981)
10. Reilly, P.M., and H. Patino-Leal, "A Bayesian Study of the Errors-in-Variables Model," *Technometrics,* **23**, 221 (1981) ❐