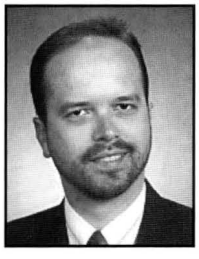
USE OF SPREADSHEETS
in Introductory Statistics and Probability

BRIAN S. MITCHELL
Tulane University • New Orleans, LA 70118

Statistical software packages such as Minitab, Statistica, or SAS are extensively used by engineers to provide descriptive statistics (mean, median, coefficient of variation, etc.) for a data set, to generate values to the various probability distributions, and to perform a linear regression or a Student's t-test. The use of these programs may be taught in various undergraduate chemical engineering courses, but their use in the chemical engineering curriculum is not standardized and there is no consensus on which, if any, of the programs is preferable.

Spreadsheets such as Microsoft Excel, Corel Quattro Pro, and Lotus 1-2-3 provide a convenient and standardized way to teach the fundamentals of statistical analysis and probability to undergraduate students. In addition to being able to execute the statistical analyses listed above, they can be used for everything from calculating numbers of combinations to constructing control charts for average ($\bar{x}$ charts). The standard user interface also reduces the time required to learn each new concept, and documents and features are virtually identical across not only platforms (IBM to Macintosh), but also across software vendors (Microsoft to Lotus).

The concept of spreadsheets in the classroom is by no means new. There are some very excellent examples of using spreadsheets for everything from general applications for first-year chemistry labs,[1] or teaching regression analysis,[2] to very elaborate applications such as calculation of X-ray diffraction patterns from crystallographic information,[3] but little information exists on how to bring spreadsheets into an introductory course in statistics and probability.



Brian S. Mitchell *is Assistant Professor of Chemical Engineering at Tulane University. He received his BS in chemical engineering from the University of Illinois-Urbana in 1986 and his MS and PhD degrees in chemical engineering from the University of Wisconsin-Madison in 1987 and 1991. His research interests are in fiber technology and composites engineering, and in the use of alternative teaching techniques in the classroom.*

Chemical engineering students at Tulane University are introduced to a variety of spreadsheet applications in their freshman course on the chemical engineering profession. The skills they develop there are used to a greater extent in the first semester of their sophomore year in a "Chemical Engineering Design I" course. This course, taught concurrently with stoichiometry, introduces the students to statistical analysis, probability, reliability, quality control, and engineering economics.

The book *Statistics and Probability for Scientists and Engineers*[4] by Mendenhall and Sincich is used in the course. It provides numerous example problems from all engineering disciplines, and although a software package is supplied with the textbook, we have opted to solve the problems using spreadsheets. This article describes how the spreadsheet can be used in a course such as this and presents example problems and solutions to illustrate the ease with which this can be accomplished.

We are fortunate to have an electronic classroom in our department that allows us to work example problems in class. With the use of a projection system that displays the screen from a laptop computer, the instructor can "walk through" the example while students perform the same functions at their own computer. The examples shown here are from Mendenhall and Sincich (hereafter referred to as MS) unless indicated otherwise, and all have been used either in the electronic classroom or assigned as homework problems. All spreadsheet solutions are from Microsoft Excel, version 7.0 for Windows 95.

DESCRIPTIVE STATISTICS

College sophomores should certainly be able to calculate a mean, median, and mode from a data set, but it's always best

to begin at the beginning. Additionally, a simple problem that the students can easily check by hand helps introduce them to the use of a spreadsheet if they are not yet familiar with it.

Example 1 is such a problem, complete with the Excel solution. Descriptive statistics are generated by selecting the data set (in this case, A2..A51) and choosing "Descriptive Statistics" in the Analysis Toolpak under the "Tools" menu. (The Analysis Toolpak is included with Excel, but is not a default item and must be installed using the "Add Ins..." command under the "Tools" menu.) The confidence level for these calculations can be specified. This example illustrates not only how to generate the desired analysis, but also how a great deal of time can be saved with large data sets. In this example, even more time is saved because the data set is available on the textbook diskette in ASCII text format, so it can be readily imported into the spreadsheet. For most example problems, the data set is placed on a server that the students can retrieve using File Transfer Protocol (FTP). This saves a substantial amount of class time.

Close inspection of the data set in Example 1 points out one of the limitations of the spreadsheet; namely, that only one mode is specified for multimodal data sets. The example data set contains two modes: 128 and 131 (three values each). The spreadsheet displays the first mode it finds, in this case 128. This example also points out a "quirk" of Excel—data can be analyzed only in columns or rows. That is, con-

Spreadsheets . . . provide a convenient and standardized way to teach the fundamentals of statistical analysis and probability to undergraduate students. . . .This article describes how the spreadsheet can be used. . . and presents example problems and solutions to illustrate the ease with which this can be accomplished.

Example 1 Descriptive Statistics (MS 2.48)

Industrial engineers periodically conduct "work measurement" analyses to determine the time required to produce a single unit of output. At a large processing plant, the number of total worker-hours required per day to perform a certain task was recorded for 50 days. Compute the mean, median, and mode of the data set. Find the range, variance, and standard deviation of the data set.

	A	B	C
1	Worker-hours	*Descriptive Statistics*	
2	128		
3	113	Mean	117.82
4	146	Standard Error	2.122896
5	124	Median	117.5
6	100	Mode	128
7	119	Standard Deviation	15.01114
8	109	Sample Variance	225.3343
9	128	Kurtosis	-0.69123
10	131	Skewness	0.00906
11	112	Range	62
12	95	Minimum	88
13	124	Maximum	150
14	103	Sum	5891
15	133	Count	50
16	111	Confidence Level(95.0%)	4.266116
17	97		
18	132		
19	135		
20	131		
21	150		
22	124		
23	97		
24	114		
25	88		
26	117		
27	128		
28	138		
29	109		
30	118		
31	122		
32	142		
33	133		
34	100		
35	116		
36	97		
37	98		
38	136		
39	111		
40	98		
41	116		
42	108		
43	120		
44	131		
45	112		
46	92		
47	120		
48	112		
49	113		
50	138		
51	122		

tiguous columns cannot be analyzed as one set, even if the data in them belong together. Quattro Pro, on the other hand, does allow for such analysis.

Histograms are also readily generated in Excel. In addition to the input data, the number and size of bins must be specified (see column B in Example 2, a modified form of MS problem 2.18). The histogram function generates the absolute frequency for each bin, as shown in columns C and D. A histogram is then generated automatically, but in this problem, a relative frequency diagram is requested. Some user intervention is required here (this is a good thing since it makes the students think about what they are doing). Relative frequency is calculated in column E by using the data from column D and the formula shown. Each of these calculations requires entering one formula only, which can then be copied to the remaining cells. Formula cell references automatically adjust, except for those "locked" with dollar signs ($). A relative frequency distribution chart can then be generated using Excel's graphing tool.

PROBABILITY

The fundamentals of probability are usually introduced using examples with decks of cards and rolls of the dice. Spreadsheets aren't a great deal of help here. They can be used at the next level, however—particularly for simple things like permutations and combinations. The number of permutations is the number of ways to put y elements in y distinct positions from a single net of n different elements, or

$$\#\text{ of permutations} = \frac{n!}{(n-y)!} \qquad (1)$$

The number of combinations of n distinct items taken y at a time is given by the binomial coefficient

$$\binom{n}{y} = \frac{n!}{y!(n-y)!} \qquad (2)$$

Neither of these relationships is very complex, and either can be calculated fairly easily by simply typing the formula into a spreadsheet cell. All spreadsheets are able to calculate factorials, in most cases with the FACT function. Example 3 illustrates the use of built-in functions to accomplish the same task more easily, using the PERMUT and COMBIN functions of Excel. Two arguments are specified for each function: the total number of values to choose from (n, number) and the number of values taken at a time (y, number_chosen).

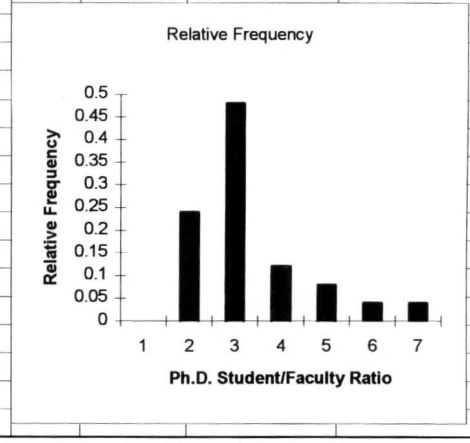
PROBABILITY DISTRIBUTIONS

Probability distributions are a vital component in the introduction of reliability analysis. There are many forms of distributions, though, and students often get hung up trying

Example 2: Relative Frequency Histograms (MS 2.18)

Each year, U.S. News and World Report surveys America's best graduate schools. The 1993 survey included a list of the top 25 graduate programs of engineering. The accompanying data include each school's overall score (based on a weighted average of rankings in five areas), total enrollment, dollar amount awarded for research, doctoral student-to-faculty ratio, and acceptance rate. Construct a relative frequency diagram for the doctoral student/faculty ratio and interpret the results.

	A	B	C	D	E
1	Doctoral Student/Faculty Ratio	Bin	*Bin*	*Frequency*	*Relative Frequency*
2	2.52	1	1	0	=D2/$D$10
3	5.24	2	2	6	=D3/$D$10
4	2.82	3	3	12	=D4/$D$10
5	4.66	4	4	3	=D5/$D$10
6	2.36	5	5	2	=D6/$D$10
7	2.63	6	6	1	=D7/$D$10
8	3.27	7	7	1	=D8/$D$10
9	3.3		More	0	=D9/$D$10
10	2.32		Sum	=SUM(D2:D9)	=SUM(E2:E9)
11	2.23				
12	3.7				
13	2.51				
14	1.75				
15	2.04				
16	1.96				
17	2.38				
18	4.88				
19	6.35				
20	2.41				
21	2.46				
22	1.27				
23	1.42				
24	2.7				
25	1.85				
26	1.75				
27					
28					



to memorize the formulas for five or six distributions rather than concentrating on the appropriate applications for each distribution. While a spreadsheet cannot totally eliminate this problem, it can reduce the anxiety of distributions a bit, and certainly eliminates the need for distribution tables, which can vary in terminology from book to book.

Most spreadsheets contain a number of distributions. Excel includes the binomial and Poisson distributions, among others. Generating values from these distributions is a simple matter of using a function with anywhere from two to four arguments. An example of the binomial distribution is given in Example 4, where the binomial distribution is given by

$$P(y) = \left(\frac{n!}{y!(n-y)!} \right) p^y (1-p)^{(n-y)} \quad (3)$$

Here P(y) is the probability that a known outcome occurs y times out of n trials, and p is the probability that an isolated event of the given outcome will occur. The example problem asks not only for the probability of 1-15 valves failing, which is just the binomial distribution evaluated at a single value of y in each case, but also for the cumulative probability of failure for 0-5 valves. First, values for n and p are entered in cells A2 and B2, respectively. For the first part of the question, all possible values of y are entered in cells C2..C17. This is easily accomplished by placing the value of zero in cell C2 and using the "Fill" command in the "Edit" pull-down menu to fill in the rest of the series up to 15. The values of the binomial distribution for each value of y are generated in column D using the BINOMDIST function. (Normally, this would be displayed as a number, but for instructional purposes, both the formula in column D and the number it returns in column E are shown.) The BINOMDIST function re-

Example 3: Probability Distributions (MS 3.40 and 3.42)

A security alarm system is activated and deactivated by correctly entering the appropriate three-digit numerical code in the proper sequence on a digital panel. Compute the total number of possible code combinations if no digit may be used twice.

	A	B	C
1	**Function Description**	**Formula**	**Numerical Result**
2	PERMUT(number, number_chosen)	=PERMUT(10,3)	720

Suppose you need to replace 5 gaskets in a nuclear-powered device. If you have a box of 20 gaskets from which to make the selection, how many different choices are possible; i.e., how many different samples of 5 gaskets can be selected from the 20?

	A	B	C
4	**Function Description**	**Formula**	**Numerical Result**
5	COMBIN(number, number_chosen)	=COMBIN(20,5)	15504

Example 4: Binomial Distribution

Consider a sample of 15 valves. The probability that a given valve fails is 0.18. a) Calculate the probability of failure of 0-15 valves. b) Calculate the probability that at most five valves will fail.

	A	B	C	D	E	F	G
1	n, total number of valves	p, probability of single valve failing	y	P.D.F., P(y) (formula)	P.D.F. (values)	C.D.F. (formula)	C.D.F. (values)
2	15	0.18	0	=BINOMDIST(C2,$A$2,$B$2,FALSE)	0.05095	=BINOMDIST(C2,$A$2,$B$2,TRUE)	0.05095
3			1	=BINOMDIST(C3,$A$2,$B$2,FALSE)	0.16778	=BINOMDIST(C3,$A$2,$B$2,TRUE)	0.21874
4			2	=BINOMDIST(C4,$A$2,$B$2,FALSE)	0.25781	=BINOMDIST(C4,$A$2,$B$2,TRUE)	0.47656
5			3	=BINOMDIST(C5,$A$2,$B$2,FALSE)	0.24524	=BINOMDIST(C5,$A$2,$B$2,TRUE)	0.72180
6			4	=BINOMDIST(C6,$A$2,$B$2,FALSE)	0.16150	=BINOMDIST(C6,$A$2,$B$2,TRUE)	0.88330
7			**5**	=BINOMDIST(C7,$A$2,$B$2,FALSE)	0.07799	**=BINOMDIST(C7,$A$2,$B$2,TRUE)**	**0.96129**
8			6	=BINOMDIST(C8,$A$2,$B$2,FALSE)	0.02853	=BINOMDIST(C8,$A$2,$B$2,TRUE)	0.98983
9			7	=BINOMDIST(C9,$A$2,$B$2,FALSE)	0.00805	=BINOMDIST(C9,$A$2,$B$2,TRUE)	0.99788
10			8	=BINOMDIST(C10,$A$2,$B$2,FALSE)	0.00176	=BINOMDIST(C10,$A$2,$B$2,TRUE)	0.99965
11			9	=BINOMDIST(C11,$A$2,$B$2,FALSE)	0.00030	=BINOMDIST(C11,$A$2,$B$2,TRUE)	0.99995
12			10	=BINOMDIST(C12,$A$2,$B$2,FALSE)	0.00003	=BINOMDIST(C12,$A$2,$B$2,TRUE)	0.99999
13			11	=BINOMDIST(C13,$A$2,$B$2,FALSE)	0.00000	=BINOMDIST(C13,$A$2,$B$2,TRUE)	0.99999
14			12	=BINOMDIST(C14,$A$2,$B$2,FALSE)	0.00000	=BINOMDIST(C14,$A$2,$B$2,TRUE)	0.99999
15			13	=BINOMDIST(C15,$A$2,$B$2,FALSE)	0.00000	=BINOMDIST(C15,$A$2,$B$2,TRUE)	0.99999
16			14	=BINOMDIST(C16,$A$2,$B$2,FALSE)	0.00000	=BINOMDIST(C16,$A$2,$B$2,TRUE)	0.99999
17			15	=BINOMDIST(C17,$A$2,$B$2,FALSE)	0.00000	=BINOMDIST(C17,$A$2,$B$2,TRUE)	1

quires four arguments, the first three of which are the values of y, n, and p, in that order, all of which can be referenced to their respective cells. Note that the value of y changes as it should, whereas n and p are fixed by the problem statement. The final argument is a "switch" that allows the cumulative distribution to be calculated. This argument is set to FALSE in column D so that singular values of the distribution can be calculated. Column E generates the answers to part a) of the statement. The cumulative distribution function is exactly what is required to solve part b) of Example 4, so the switch is set to TRUE in column F. Again, values returned by this function are shown in the following column, and the answer to the problem is shown in cell G7; the cumulative probability of zero through five valves failing is 0.961.

HYPOTHESIS TESTING

There are a number of useful null-hypothesis tests, including the Chi-square test, F-test, and Student's t-test, all three of which are covered in Design I. Most spreadsheets are capable of performing at least some of these tests, although much more interpretation of the results is required than for the previous examples. In particular, the final decision as to whether or not the null hypothesis has been verified is left up to the student.

An F-test is used here as an illustration of how null-hypothesis tests can be performed, at least in part, using a spreadsheet. Recall that an F-test compares the variances, s_i^2, of two data sets

$$F = \frac{s_1^2}{s_2^2} \qquad (4)$$

The null-hypothesis is that the two variances are statistically equivalent at some specified confidence level, typically 95%. The value of F is calculated using Eq. (4), and compared with a value in a table at the specified confidence level and appropriate degrees of freedom for each data set. If the F-value in the table is greater than the calculated value of F, the null hypothesis is substantiated, and the two variances can be considered statistically equivalent.

A well-known problem from Peters and Timmerhaus[5] is used here as an illustration of an F-test applied to a chemical engineering problem and how the results from the spreadsheet manipulation must be interpreted (see Example 5). The data are entered in cells A2..A8, and B2..B6 for the revised and current analytical methods, respectively. The "F-test Two Sample for Variances" function is selected from the "Data Analysis..." option under the "Tools" menu in Excel. The cell indexes for both data sets must be provided, as must the cell assignment for the output, and the value of α, which determines the confidence level. The resulting table is shown in cells A12..C19. The calculated value of F appears in cell B17 and is the ratio of cells B14 to C14. In this instance, no formulae are present in the tables—only numerical values appear at the end of the analysis. The tabulated value of F for the specified value of α=0.5 is given in cell B19. The final step is left to the student. In this case, the calculated value of F is greater than the tabulated value, indicating that the two analytical procedures may not be equivalent. Once again, the spreadsheet is helpful, but the user must have a knowledge of the underlying principles to correctly interpret the results.

ANALYSIS OF VARIANCE

Analysis of variance (ANOVA) is a statistical analysis tool that provides a smooth transition from F-tests into a

Example 5: *F*-test (Peters and Timmerhaus Chpt. 17, Example 10)

A simplified analytical procedure is proposed for a routine laboratory test. It is necessary to determine not only whether the new procedure gives the same results as the old, i.e., whether the means of a duplicate set are the same, but also whether the precision of the new test is as good as the current test.

	A	B	C
1	**Revised Method**	**Current Method**	
2	79.2	79.7	
3	79.7	79.5	
4	79.5	79.6	
5	79.4	79.5	
6	80	79.7	
7	79.6		
8	79.8		
9			
10	F-Test Two-Sample for Variances		
11			
12		*Variable 1*	*Variable 2*
13	Mean	79.6	79.6
14	Variance	0.07	0.01
15	Observations	7	5
16	df	6	4
17	F	6.999999998	
18	P(F<=f) one-tail	0.040280016	
19	F Critical one-tail	6.163134003	

discussion of linear regression. Recall that an ANOVA table compares N similar data points from k different treatments by essentially performing an F-test on the variance between treatments (mean square treatment, or MST) and the variance within the treatments (mean square error, or MSE). The MST and MSE are calculated from the sum of squares (SST and SSE, respectively) and corresponding degrees of freedom for each type of error, as shown in Table 1.

TABLE 1
Analysis of Variance Table

Source	Sum of Squares	Degrees of Freedom	Mean Square	F
Treatment	$\sum_{i=1}^{k} n_i\left(\bar{Y}_i-\bar{\bar{Y}}\right)^2$	k-1	SST/k-1	MST/MSE
Error	$\sum_{i=1}^{k}\sum_{j=1}^{n_i} n_i\left(Y_{ij}-\bar{Y}_i\right)^2$	N-k	SSE/N-k	-
Total	$\sum_{i=1}^{k}\sum_{j=1}^{n_i} n_i\left(Y_{ij}-\bar{\bar{Y}}\right)^2$	N-1	-	-

Example 6: Analysis of Variance (MS 14.27)

An excessive amount of ozone in the air is indicative of air pollution. Six air samples were collected from each of four locations in the industrial Midwest and measured for their content of ozone. Construct an analysis of variance table for the data. Do the data provide sufficient evidence to indicate differences in the mean ozone content among the four locations? Use $\alpha = 0.05$.

	A	B	C	D	E	F	G
1		Location					
2	1	2	3	4			
3	0.08	0.15	0.13	0.05			
4	0.1	0.09	0.1	0.11			
5	0.09	0.11	0.15	0.07			
6	0.07	0.1	0.09	0.09			
7	0.09	0.08	0.09	0.11			
8	0.06	0.13	0.17	0.08			
9							
10	ANOVA						
11	*Source of Variation*	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
12	Between Groups	0.006779	3	0.00226	3.498925	0.034527	3.098393
13	Within Groups	0.012917	20	0.000646			
14							
15	Total	0.019696	23				

Example 6 shows how an ANOVA table can be generated on the spreadsheet. In this case, there are k=4 treatments (locations, i=1 through 4), each with six data points (j=1 through 6), for a total of N=24 data points. By selecting the "ANOVA:Single Factor" function in the "Data Analysis..." option of the "Tools" pull-down menu, the input range can be specified, which in this example is A3..D8. The data is this case are grouped by columns, and this radio button must be selected on the menu. Once again, the desired confidence level can be specified, here as α=0.05, and the location for the resulting ANOVA table specified in the "Output Range" box. The ANOVA table appears in cells A10..G15. Additionally, this table not only shows the calculated F-value, but also the critical F-value for the specified degrees of freedom and confidence level. This makes comparison of the two F-values particularly easy for the student, who must once again arrive at the final evaluation. In this case, there is sufficient evidence to suggest that the ozone contents differ statistically for the different locations.

CONTROL CHARTS

The final example of spreadsheets in Design I deals with quality control. In the chemical process industry, this usually means control charts. Control charts for average ($\bar{x}$-chart), range (R-chart), proportion defects (p-charts), and defects per unit (c-charts) are introduced in this course. Example 7 shows how Excel can be used to generate a p-chart. In this case, the formulae for the lower control limit (LCL), upper control limit (UCL), and centerline must be specified as

$$LCL=\bar{p}-3\sqrt{\frac{\bar{p}(1-\bar{p})}{n}} \qquad (5)$$

$$UCL=\bar{p}+3\sqrt{\frac{\bar{p}(1-\bar{p})}{n}} \qquad (6)$$

Here

$\bar{p}$ = centerline = average fraction defectives = $(d_1 + d_2 + ...d_k)/nk$, cell C24

n = sample size, cell B2

k = number of sample periods.

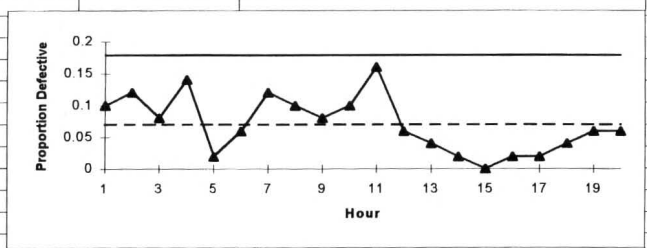
There is nothing particularly involved about this application, but it does allow the

introduction of IF/THEN-type statements into spreadsheet calculations. In this case, the calculated LCL may have a negative value, in which case it must be replaced by zero. Cell C27 shows how the IF statement can be used to compare the calculated LCL to zero and place the appropriate value in F3. The LCL is copied down column F (in this case it is zero), as are the UCL and centerline values down columns D and E, respectively, to facilitate plotting. The p-chart is generated by plotting the data in column B (shown as triangles), UCL, LCL, and centerline vs. the sample ID in column A. The resulting plot is shown at the bottom of the spreadsheet in Example 7, and the process appears to be in control. Time allowing, this example also provides an excellent opportunity to introduce the concept of spreadsheet macro commands that can automate the production of $\bar{x}$-charts every time a new set of data is generated.

Example 7: Control Chart (MS 16.44)

High-level computer technology has developed bit-sized microprocessors for use in operating industrial "robots". To monitor the fraction of defective microprocessors produced by a manufacturing process, 50 microprocessors are sampled each hour. The results for 20 hours of sampling are provided. Construct a control chart for the proportion of defective microprocessors. Locate the center line and upper and lower control limits on the chart. Does the process appear to be in control?

	A	B	C	D	E	F
1	n	50				
2	Sample	Number of Defectives	Fraction Defective	UCL	Centerline	LCL
3	1	5	=B3/$B$1	=$C$25	=$C$24	=$C$27
4	2	6	=B4/$B$1	=$C$25	=$C$24	=$C$27
5	3	4	=B5/$B$1	=$C$25	=$C$24	=$C$27
6	4	7	=B6/$B$1	=$C$25	=$C$24	=$C$27
7	5	1	=B7/$B$1	=$C$25	=$C$24	=$C$27
8	6	3	=B8/$B$1	=$C$25	=$C$24	=$C$27
9	7	6	=B9/$B$1	=$C$25	=$C$24	=$C$27
10	8	5	=B10/$B$1	=$C$25	=$C$24	=$C$27
11	9	4	=B11/$B$1	=$C$25	=$C$24	=$C$27
12	10	5	=B12/$B$1	=$C$25	=$C$24	=$C$27
13	11	8	=B13/$B$1	=$C$25	=$C$24	=$C$27
14	12	3	=B14/$B$1	=$C$25	=$C$24	=$C$27
15	13	2	=B15/$B$1	=$C$25	=$C$24	=$C$27
16	14	1	=B16/$B$1	=$C$25	=$C$24	=$C$27
17	15	0	=B17/$B$1	=$C$25	=$C$24	=$C$27
18	16	1	=B18/$B$1	=$C$25	=$C$24	=$C$27
19	17	1	=B19/$B$1	=$C$25	=$C$24	=$C$27
20	18	2	=B20/$B$1	=$C$25	=$C$24	=$C$27
21	19	3	=B21/$B$1	=$C$25	=$C$24	=$C$27
22	20	3	=B22/$B$1	=$C$25	=$C$24	=$C$27
23						
24		Centerline	=AVERAGE(C3:C22)			
25		UCL	=$C$24 + 3*SQRT($C$24*(1-$C$24)/$B$1)			
26		LCL (calculated)	=$C$24 - 3*SQRT($C$24*(1-$C$24)/$B$1)			
27		LCL (actual)	=IF($C$26 > 0, $C$26, 0)			
28						
29						
30						
31						
32						
33						
34						
35						
36						
37						
38						
39						
40						



CONCLUSION

Spreadsheets continue to grow in popularity and availability, and their utility in solving everyday enginering problems develops with each new version. Some examples have been presented on how to incorporate spreadsheet use into a sophomore-level chemical engineering course on statistics and probability. Hopefully, these examples will inspire more of us to use spreadsheets to illustrate chemical engineering fundamentals in the classroom.

ACKNOWLEDGMENT

Support for this work comes in part from the Camille and Henry Dreyfus Special Grant Program in chemical sciences.

REFERENCES

1. Birk, J.P., "First-Year Chemistry Laboratory Calculations on a Spreadsheet," *J. Chem. Ed.*, **69**(8), 648 (1992)
2. Wood, W.C., and S.L. O'Hare, "A Spreadsheet Model for Teaching Regression Analysis," *J. Educ. Bus.*, May/April, 233 (1992)
3. Shapiro, F., "The Calculation of Crystal Diffraction Patterns Using a Spreadsheet," *J. Mater. Ed.*, **14**, 93 (1992)
4. Mendenhall, W., and T.S. Sincich, *Statistics for Engineering and the Sciences,* 4th ed., Prentice-Hall, Upper Saddle River, NJ (1995)
5. Peters, M.S., and K.D. Timmerhaus, *Plant Design and Economics for Chemical Engineers,* 4th ed., McGraw Hill, New York, NY (1991) ❐