

Use of Computer Simulation to Enhance Learning in a Mechanical Engineering Measurements Laboratory

Dr. Robert G. Ryan
California State University, Northridge

Abstract

Fundamental concepts related to data sampling and uncertainty analysis can be introduced through standard textbook problems, but it is much easier to stimulate student interest if the analysis involves real data. However, illustration of basic concepts can be compromised if the data exhibit anomalies due to errors in technique or equipment problems. The use of computer-generated values for analysis offers a balanced approach that produces predictable values but maintains a level of interaction between the student and the “experiment”. Three experiments using computer-generated data are described, and sample results are presented. The software used is Microsoft Excel (with procedures written in Visual Basic for Applications). This approach has been effective in a junior-level measurements laboratory for mechanical engineering students, and can also be adapted for distance learning applications.

I. Introduction

One of the purposes of laboratory courses is to teach students how to properly analyze data obtained from engineering experiments. Important concepts include the normal, or Gaussian probability distribution, confidence intervals, error propagation, and the use of Discrete Fourier Transforms for frequency analysis. Analysis techniques related to these concepts involve a great deal of “number crunching”, and thus studying these topics using textbook problems solved on a calculator reinforces the notion that data analysis is inherently tedious and boring.

One might conclude that it is always preferable to analyze actual laboratory results for learning data analysis techniques, since the students will be motivated by performing the experiment themselves. However, the unpredictability of experimental data can often introduce complications that may compromise the desire of the instructor to demonstrate a particular principle or technique. Also, the students are confronted with several problems at once, since they have to properly analyze and then interpret the data by comparing results to an appropriate physical model. Exposing the students to more complex experiments is clearly a key goal of laboratory courses, but may not be the best avenue for introducing data analysis techniques.

The use of computer-generated data is a compromise between the “textbook problem” and standard “lab experiment” approaches for teaching data analysis techniques. A degree of spontaneity and interaction between student and “experiment” is maintained, but the behavior illustrated by the data can be carefully controlled. The results of these experiments can be

designed by the instructor to demonstrate a particular technique or principle, and are relatively easy to review and grade because the correct result is known, and thus student errors can be efficiently identified and resolved. Another advantage of this approach is to have a relatively large class (approximately twenty students working on ten computer stations) performing the same “experiment” simultaneously, without requiring a lot of expensive equipment.

The primary data analysis software used for the Mechanical Measurements course is Microsoft Excel. Although a number of alternatives exist, Microsoft Excel (along with Visual Basic for Applications – VBA) is used throughout the mechanical engineering program at California State University, Northridge (CSUN), and is therefore familiar to all of our students. An additional advantage is the availability of this software on virtually every engineer’s computer in the workplace. Visual Basic procedures linked to “buttons” on a worksheet prepared by the instructor are used to generate the data used for analysis. The procedures may easily be modified for a variety of purposes, including providing unique results for each computer station or changing results from semester to semester (to make plagiarism difficult).

II. Laboratory and Course Descriptions

The Measurements and Mechatronics Laboratory in CSUN’s Engineering Building supports two required courses in the mechanical engineering curriculum – Mechanical Measurements (ME 335) and Mechatronics (ME 435). Much of the equipment required to support these courses (especially the computer stations) is similar, so combining both courses in one facility was a cost effective approach.

The laboratory facility is designed around ten stations which include a PC, a lab bench, and various electronic equipment (e.g. oscilloscope, function generator, trainer board, voltmeter, etc.) Each station has a data acquisition board and signal conditioning rack from Measurement Computing (DAS08-JR and ISO-RACK08), and a “junior” version of LabTech Notebook software to support data collection from a variety of sensor types. Additional controller hardware and software donated by DeltaTau is used to support the Mechatronics course. The ten stations occupy the periphery of the laboratory, leaving sufficient room in the center for seats and tables for twenty-plus students. Thus the facility is designed to support both the laboratory and lecture sessions which make up the two courses. Some of the experiments performed in the Mechanical Measurements course use computer-generated data (described in this paper), and others involve measurements with typical sensors such as thermocouples, thermistors, strain gauges, pressure sensors, and potentiometers.

The Mechanical Measurements course is taken by mechanical engineering students in the junior year, and consists of four contact hours per week (corresponding to one unit each of lecture and laboratory credit awarded for the course). The lecture and laboratory sessions are taught by the same instructor and the lecture material and laboratory experiments are closely integrated. The textbook used is “*Theory and Design for Mechanical Measurements*“, by Figliola and Beasley¹. The key course topics are statistical data analysis, uncertainty analysis and error propagation,

dynamic behavior of instruments, and a survey of common sensor types used by mechanical engineers.

III. Experiment Descriptions

The three laboratory experiments which use computer-generated data are entitled:

- ◆ Simulation of Normally Distributed Data
- ◆ Comparison of the Means of Two Similar Data Sets
- ◆ Effect of Sampling Rate on Fast Fourier Transform Analysis of Data

Each experiment was designed to satisfy a set of specific objectives related to the course material. Students begin each experiment by opening an Excel workbook which contains the necessary VBA modules for generating the data. Theoretical background related to the experiment is discussed in class, and detailed written instructions for the experiment procedure are handed out to each student. Students work in pairs at each computer station. Attached to the written procedure is a sheet consisting of two or three questions requiring short answers related to the interpretation of the experiment results. Each student is required to turn in this question sheet along with worksheets and plots created during the experiment. The goal of this format is to encourage students to work on the interpretation of the results immediately after performing the experiment, rather than writing a report several days later. As discussed later in this paper, this format produced mixed results and a more formal report will be required beginning in the spring semester.

The experiments discussed in this paper have evolved over two semesters and are still being modified based on classroom experience. The experiment instructions are written in sufficient detail to account for the possibility of a student performing the experiment outside of regular laboratory hours, and potentially for distance learning applications. Questions raised by students while performing the experiments in class are being used as a guide for clarifications and additions to the written instructions. The experiments and their typical results are described below. Details regarding statistical concepts and Excel functions used for analysis are provided to clarify the objectives of each experiment.

Simulation of Normally Distributed Data

The uncertainty formulas generally used by engineers are based on the assumption that the data values follow a normal, or Gaussian distribution, often referred to as the “bell-shaped” curve. There are a number of ways to test a distribution of values for Gaussian or normal behavior, including probability plots and the chi-squared “goodness of fit” test. Both of these tests can be efficiently performed in Excel. This experiment requires students to plot a bin histogram and test for normal behavior using two distinct data sets of data values (named “A” and “B”) generated by VBA Sub procedures. These Sub procedures are created in the Visual Basic Editor located in

the Tools>Macro menu, and are stored in a module which is saved as part of the Excel workbook file.

Students begin the experiment by opening an Excel workbook file prepared by the instructor. The workbook consists of two blank worksheets (one for A-values and one for B-values) and a VBA module. There are two Sub procedures stored in the VBA module which are linked to buttons located on the blank worksheets. Clicking the button on the A-values worksheet generates a data value and places it in the active worksheet cell (i.e the last cell clicked prior to clicking the button). A brief description of the code used for this experiment is included in Figure 1 for readers unfamiliar with its use (a good introduction to VBA programming can be found in Microsoft's *Excel/Visual Basic Programmer's Guide*²). The first Sub procedure (Anorm) generates normally distributed values with a mean of 62 and a standard deviation of 3.1 (of course any suitable values may be used). The key line of code begins with "Application.Run". This syntax calls the worksheet function "Random" to generate the normally distributed values. (Note that this function can also be accessed from the worksheet by the menu choices Tools>Data Analysis>Random Number Generation. With this method, a dialog box appears allowing a normal distribution with any mean and standard deviation to be chosen, along with the desired number of samples. Navigating these menus with the "Macro Record" feature turned on is an effective way to generate the appropriate syntax for calling the function and its arguments via VBA code.) It is desirable in this case to generate one value at a time, since we are simulating the process of collecting individual measurement values. The values are output to three decimal places.

The second Sub procedure (BVALUES) uses the VBA random number generator function, Rnd, which generates uniform random values between 0 and 1. The remaining statements scale those values to a desired mean and range. The code shown in Figure 1 is designed to produce a uniform distribution of values between $62 - 3.5 / 2$ and $62 + 3.5 / 2$.

The students are required to complete the following steps:

- ◆ Generate 30 "A" data values in a column in the A-Values worksheet.
- ◆ Create a bin histogram of the data using the menu selection Tools>Data Analysis>Histogram (see samples in Figure 2).
- ◆ Choose bin ranges for performing a chi-squared goodness of fit test on the data. Use Excel's FREQUENCY function to determine the number of values in each bin. Use Excel's NORMDIST function to calculate the expected number of values in each bin for a normal distribution, and then calculate the chi-squared parameter. Use Excel's CHIDIST function to calculate the chi distribution probability, which can be used to judge whether the data fits the assumed normal distribution.

- ◆ Create a normal probability plot for the data, using a procedure based on an approach used by Berk and Carey³. This procedure also requires the use of the NORMDIST function. The plot is linear if the sample data are normally distributed.
- ◆ Repeat the process for data set B.

Note the importance of Excel's built-in functions to the data analysis. The described calculations would be extremely time consuming if performed by hand (with a standard calculator). The experiment instructions guide the students through the analysis step by step, including the syntax of necessary formulas and the cell locations where they should be located.

Sub Anorm()

```
Application.Run "ATPVBAEN.XLA!Random", ActiveCell, 1, 1, _
    2, , 62, 3.1
AVAL = Application.ActiveCell.Value
AVAL = Int(AVAL * 1000) / 1000
Application.ActiveCell.Value = AVAL
```

End Sub

Sub BVALUES()

```
BMEAN = 62
BDEV = 3.1
DIFF = (Rnd - 0.5) * BDEV
BVAL = BMEAN + DIFF
BVAL = Int(BVAL * 1000) / 1000
Application.ActiveCell.Value = BVAL
```

End Sub

Figure 1. Visual Basic Code for Generating Data Values

The experiment procedure requires the student to judge the data sets for normally distributed behavior by three methods – a visual check of the histogram, the chi-squared test, and the probability plot. Based on the methods used for generating the data values, it would be expected that data set A would pass the tests, and data set B would fail the tests. This is not always a clear

cut result, due to the subjective nature of the tests – but learning the limitations of these tests is part of the message of this experiment. Figure 2 shows typical histograms for the two data sets (the x-axis values are the upper value of each bin range).

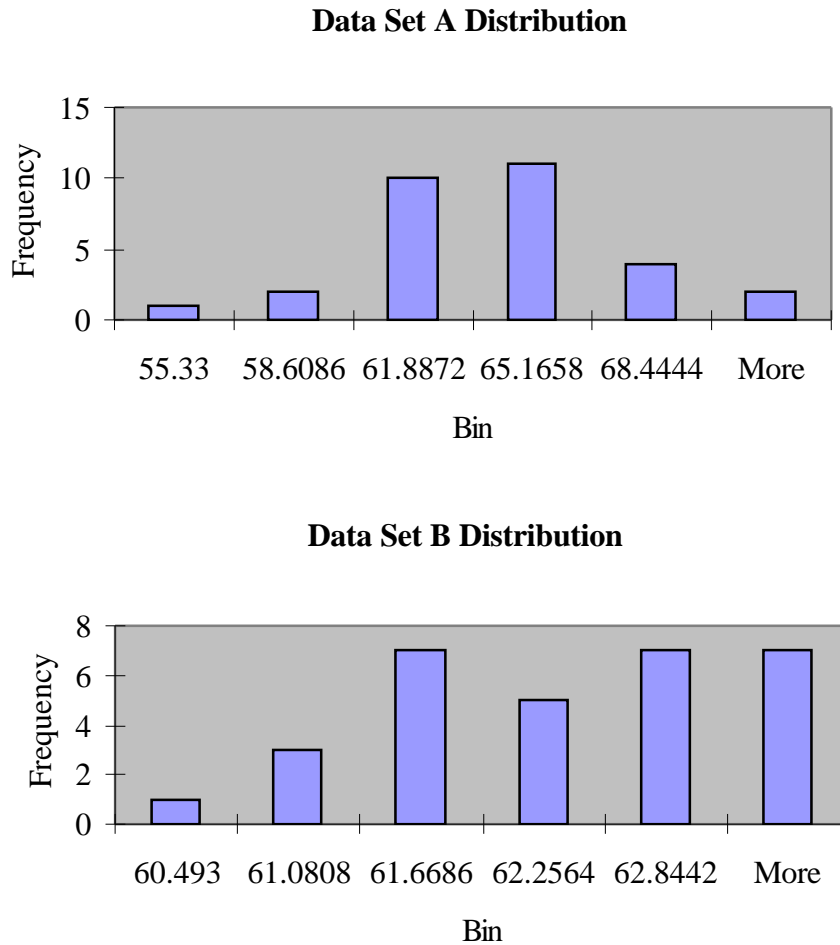


Figure 2 Typical Data Distributions (A is normal, B is uniform)

Comparison of the Means of Two Similar Data Sets

A common problem in the analysis of data is comparing two sets of data to determine the one with the larger mean. Clearly if the difference of the means is of the same order as the random uncertainty in the data, the identification of the largest mean will be associated with a confidence

*Proceedings of the 2002 American Society for Engineering Education Annual Conference & Exposition
Copyright © 2002, American Society for Engineering Education*

level of that conclusion determined from statistical methods. If A_m and B_m are the means of their respective data sets based on N points, then the difference in the means and its uncertainty can be determined from:

$$D = A_m - B_m$$

$$\text{Del } D = [(S_A^2 + S_B^2) / N]^{1/2}$$

where S represents the sample standard deviation of each sample. The ratio of D to $\text{Del } D$ can be used, along with a standard normal distribution probability table (or the NORMDIST function) to determine the probability that D is larger than zero (i.e. A_m is larger than B_m). A related calculation is the determination (after several data points are collected) of the total number of data points required to achieve a desired confidence level for the identification of the larger mean. For example, to determine N_{95} corresponding to a 95% confidence level:

$$N_{95} = 2.69 (S_A^2 + S_B^2) / D^2$$

(Note that this equation is based on normal distribution statistics. An alternative approach, using t-distribution statistics, gives similar results.)

Students are asked to perform these calculations using two data sets A and B. Students begin by opening an Excel workbook prepared by the instructor. Two buttons on the worksheet are linked to VBA Sub procedures similar to Anorm shown in Figure 1, such that two data sets with similar means and standard deviations are generated. The difference in the means specified in the two Sub procedures can be “tuned” to provide interesting results. Specifically, in this experiment the students are asked to:

- ◆ Generate ten points each for set A and B by clicking the appropriate buttons.
- ◆ Calculate the number of points required (N_{95}) to achieve a 95% confidence in the identification of the larger mean.
- ◆ Generate more data points until N_{95} points are obtained for each set. Calculate the actual confidence level and compare it to 95% (it will not compare exactly since the standard deviations and difference in the means vary as more points are collected).
- ◆ Plot standard deviations, difference in means, and confidence level as a function of number of points collected.

The primary purpose of the experiment is to demonstrate the effect of the number of data points on key parameters such as the difference in the means and the confidence level. Typically, about ten to fifteen data points are required for the values to approach an asymptotic value (see typical results in Figure 3). Based on the first ten data points, it was predicted that a total of 37 points

would be required to reach a 95% confidence. The actual confidence level after 37 points was 88%, which is reasonable agreement.

Mean Comparison of Two Data Sets

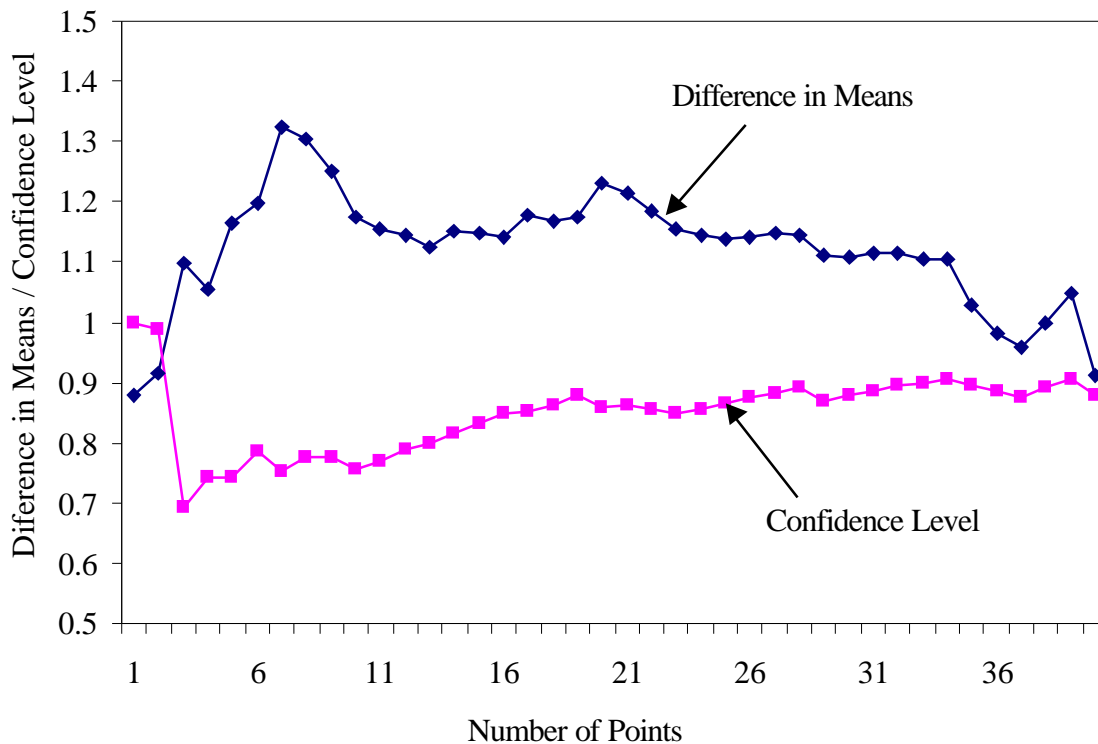


Figure 3 Comparison of Means (Typical Results)

Effect of Sampling Rate on Fast Fourier Transform (FFT) Analysis of Data

The purpose of this experiment is to introduce students to frequency analysis of a complex waveform. Emphasis is given to the effect of sampling rate and number of data points on the

accuracy of the amplitude versus frequency plots produced by Fourier analysis. The frequency resolution (Δf) of the FFT is a function of the sampling frequency (f_s) and the number of data points (N):

$$\Delta f = f_s / N$$

A small value of Δf requires a small sampling frequency and a large number of data points; however, the sampling frequency must be at least twice as large as the highest frequency present in the data (which may or not be known in advance). Thus it is often necessary to collect data with different values of f_s and N to achieve a satisfactory result.

Students are asked to explore these concepts for two different waveforms (generated by VBA Sub procedures). One is a simple 1 Hertz sine wave and the other is more complex (e.g. the sum of three sine waves of different frequencies which are unknown to the students). Specifically, the students begin by opening an Excel workbook prepared by the instructor which has a number of buttons on the worksheet. Each button is linked to a VBA Sub procedure corresponding to a particular waveform, sampling rate, and number of data points. The code can be easily modified to produce a variety of waveforms (e.g. square waves, sawtooth waves). The specific steps followed by the students are:

- ◆ Sample a 1 Hz sine wave at 15 Hz, 1.8 Hz, and 1.4 Hz (with $N = 256$). Produce an amplitude-frequency plot using Excel. The presence of alias frequencies should be noted (i.e. false frequencies due to the sampling frequency being less than 2 Hz).
- ◆ Sample a complex waveform at various sampling frequencies and number of data points, given that the maximum frequency present in the data is about 30 Hz (this sets a minimum sampling rate). Produce an amplitude-frequency plot in Excel, and note the effect of f_s and N on the plot.

Typical results for the 1 Hz sine wave are given in Figure 4. The amplitude-frequency plots for the lower sampling frequencies clearly illustrate the presence of alias frequencies at values lower than 1 Hz. It should be noted that the Fourier analysis in Excel takes several steps. First, the Fast Fourier Transfer (FFT) function is located in the worksheet menu (Tools>Data Analysis>Fourier Analysis). The input to the function must be 2^n data points, where n is any integer. The output of the function gives the coefficients of the Discrete Fourier Transform in complex form at the frequencies of:

$$f_k = k \Delta f, \text{ with } k = 1, 2, 3, \dots, N / 2$$

The calculation of the magnitudes and the frequencies must be performed on the worksheet. A good summary of the procedure is given by Bloch⁴.

The results for the simple sine wave are meant to warn students of the dangers of a careless frequency analysis. They are then asked to find the three dominant frequencies in a complex

waveform, using different values of sampling rate and number of points. Proper choice of these parameters gives the amplitude-frequency plot shown in Figure 5. Also shown is the code used to generate the values and a plot of the waveform versus time.

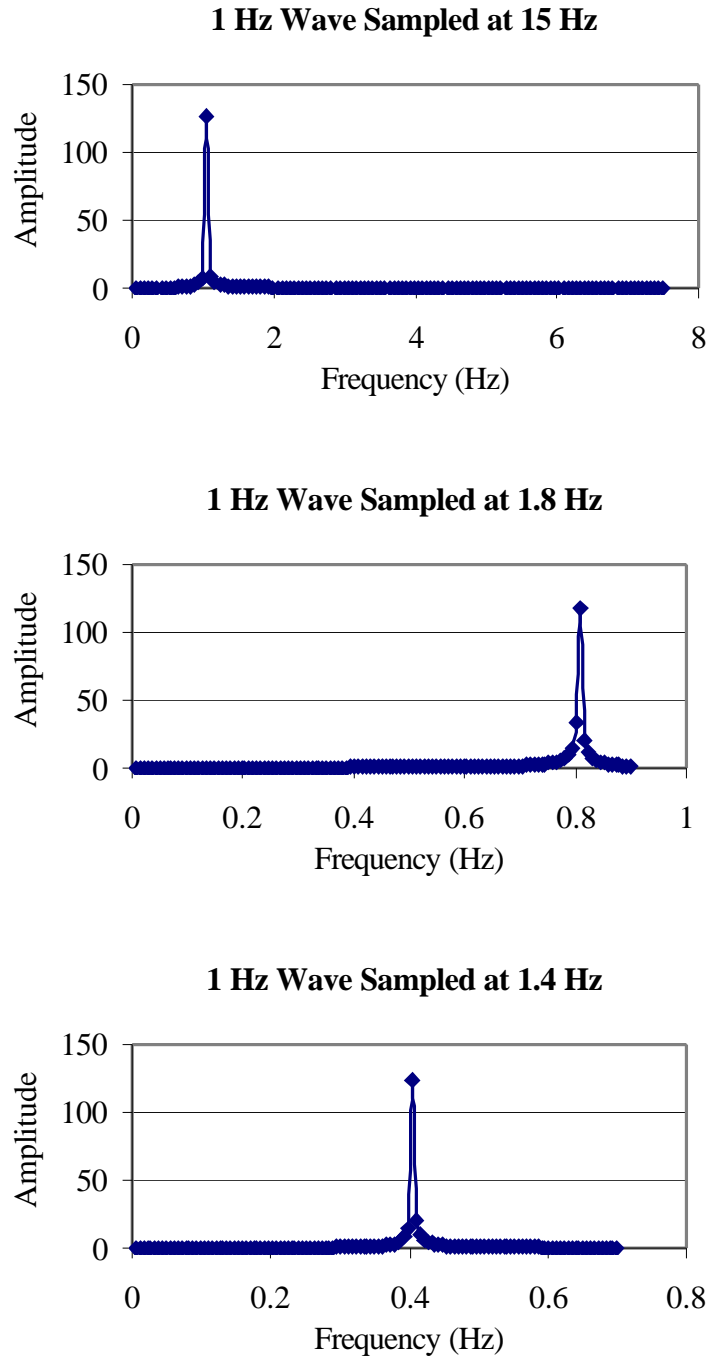


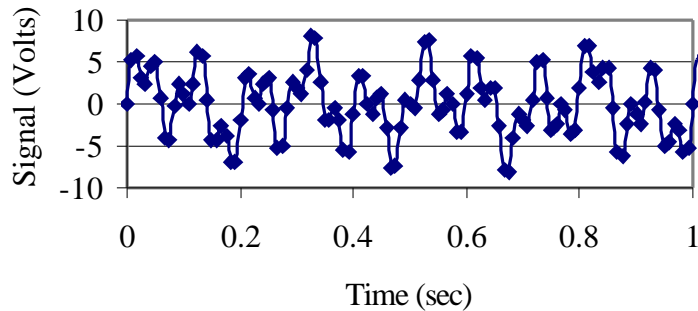
Figure 4 Effect of Sampling Rate on FFT of a 1 Hz Sine Wave

Sub Complex()

```
Application.ActiveCell.Value = 0
Application.ActiveCell.Offset(0, 1) = 0
Pi = 3.1416
freqsample = 120
deltat = 1 / freqsample
For I = 1 To 255
Application.ActiveCell.Offset(I, 0) = I * deltat
Term1 = 2 * Sin(2 * Pi * 4 * I * deltat)
Term2 = 4 * Sin(2 * Pi * 10 * I * deltat)
Term3 = 3 * Sin(2 * Pi * 25 * I * deltat)
Application.ActiveCell.Offset(I, 1) = Term1 + Term2 + Term3
Next I
```

End Sub

Complex Wave - Signal Vs Time



Complex Wave Sampled at 120 Hz

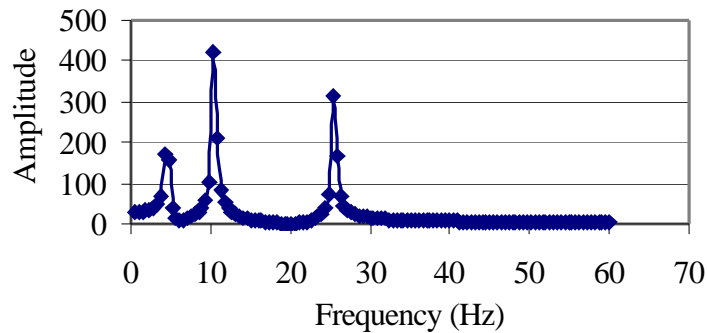


Figure 5 VBA Code, Waveform, and FFT Analysis for Complex Waveform

IV. Assessment of Student Learning

The assessment of the effectiveness of these experiments in meeting the course objectives is primarily based on a review of the student responses on the question sheets associated with each experiment, and midterm and final exam results. Observations based on interactions with students during the laboratory period were also helpful.

The use of computer-generated data for these experiments was generally deemed to be a success. The data “collection” process provided a sense of spontaneity which enhanced student interest. Instructor’s responses to student questions were able to focus more directly on the relevant data analysis technique, since there were no human errors in the data to worry about. Most students were able to finish the experiments in the allotted time and complete the question sheets before leaving class.

However, results on the exams were not that positive. For example, an exam problem was given on the midterm which required calculations of the means of two data sets and the confidence level that one mean was larger than the other. It was essentially identical to calculations performed in one of the experiments. Many of the students were completely baffled by this problem, even though they had successfully performed the related experiment. Potential reasons for the poor exam performance are thought to be:

- 1) The exam required the use of a calculator, rather than relying on Excel functions.
- 2) The experiment procedures and questions were too detailed and straightforward, causing students to successfully go through the motions without really grasping the concepts that were demonstrated.
- 3) Some students let their lab partner do most of the thinking, and just copied their partner’s answers on the question sheets without understanding their significance.

The proposed solution to some of these problems is a return to a traditional lab report for these experiments. Completion of a lab report will require a fuller, more independent interpretation of the results on the part of the student, and certain hand calculations to check the Excel results will be required. Obviously this approach will create more work for both the instructor and the students, but appears to be necessary for meeting the course objectives.

V. Conclusions

The use of VBA modules for generating data for experiments emphasizing data analysis techniques performed on Excel worksheets has worked effectively in CSUN’s Mechanical Measurements laboratory. This approach has proven to be an efficient way to introduce the use of important Excel functions related to statistical and Fourier analysis of data. While other software can also be used for this purpose, the flexibility of VBA and the familiarity of the Excel interface make it a good choice for this application. The design of the Excel worksheets will

continue to evolve. Improvements to the worksheet user interface (e.g. provide a graphic of some type of instrument rather than a plain button) are planned to provide a more realistic “experiment”.

The use of a short answer form for student interpretation of the experiment results was less successful. The very detailed nature of the experiment instructions and questions appeared to lead many students to generate results and answers that were not fully understood, since performance on exam questions and laboratory work was not consistent. The experiment instructions will be reviewed and possibly modified for next semester. The short answer form will be abandoned in favor of a more traditional laboratory report format.

Bibliography

1. Figliola, R. & Beasley, D., *Theory and Design for Mechanical Measurements*, Wiley (2000)
2. *Excel/Visual Basic Programmer's Guide*, Microsoft Press (1995)
3. Berk, K. & Carey, P., *Data Analysis with Microsoft Excel*, Duxbury Press (1998)
4. Bloch, S., *Excel for Engineers and Scientists*, Wiley (2000)

Robert G. Ryan

Robert Ryan received his PhD degree in mechanical engineering from the University of California at Los Angeles. He is currently a Lecturer in the Mechanical Engineering Department at California State University, Northridge. He is the coordinator for the department's Measurements Laboratory and Thermo-fluids Laboratory.