

## Teaching Data Analysis Techniques Using Practical Polymer Processing Examples

James A. Newell  
The University of North Dakota

### I n t r o d u c t i o n

Teaching data analysis techniques is an essential component of engineering laboratory classes. Elbow<sup>1</sup> states that intellectual excitement is the most important dimension of effective teaching. Regrettably, data analysis techniques tend to be dry and the problems are often perceived as uninteresting by undergraduate students.

By relating the analysis techniques to “real-world” problems, the students learn the concepts within the framework of engineering analysis. The students are not simply asked to calculate statistics based on arbitrary numbers. Instead, they are asked to make informed engineering decisions using the statistical analysis techniques as tools.

Additionally, chemical engineering undergraduate students are becoming increasingly interested in application oriented topics such as polymers and environmental technology. By designing problems in these areas, student interest is enhanced.

### Data Analysis Techniques

The problems presented in the next section assume that students have been instructed in the following experimental analysis techniques:

1. The elimination of “bad” data using the statistical q-test.
2. The determination of 95% confidence intervals from standard deviations using the statistical t-tables.

### Homework Problem #1

An inventor claims that he can increase the tensile strength of a polymeric fiber by adding a small quantity of the rare element toughenitupneum during spinning. To prove her claim she provides data obtained by testing samples with and without her addition. The six samples tested without the addition had tensile strengths of 3100, 2577, 2715, 2925, 3250, and 2888 GPa, respectively. Six samples tested with the added element has strengths of 3725, 3090, 3334, 3616, 3102, 3441 GPa. Has the inventor proven her claim? If not, suggest improvements that could help her.

Solution:

First, the mean strengths of the samples must be calculated



w/o additive

with additive

3100  
2577  
2715  
2925  
3250  
2888

3725  
3090  
3334  
3616  
3102  
3441

$$\bar{x}_1 = \frac{\sum_{i=1}^N (x_{1i})}{N-1} = 2909$$

$$\bar{x}_2 = \frac{\sum_{i=1}^N (x_{2i})}{N-1} = 3385$$

Next, we must calculate standard deviations:

$$S_1 = \sqrt{\frac{\sum_{i=1}^N (x_{1i} - \bar{x}_1)^2}{N-1}} = 245$$

$$S_2 = \sqrt{\frac{\sum_{i=1}^N (x_{2i} - \bar{x}_2)^2}{N-1}} = 261$$

The true mean ( $x_1$ ) is 95% likely to fall within the range

$$x_1 = \bar{x}_1 \pm \frac{t_{0.95} S_1}{\sqrt{N}}$$

$$x_2 = \bar{x}_2 \pm \frac{t_{0.95} S_2}{\sqrt{N}}$$

In this case,

$$x_1 = 2909 \pm 257 \text{ and } x_2 = 3385 \pm 274.$$

Thus, within 95% confidence,  $x_1$  could be as high as 3166 GPa, while  $x_2$  could be as low as 3111 GPa. Therefore, we cannot tell from the experimental data whether  $x_2$  is truly greater than  $x_1$ . If the inventor wishes to prove her claim, she must test more samples.

### Homework Problem #2

A company sends you an additive designed to increase the viscosity of epoxy resins. They claim that adding two drops of their additive to 100 ml of resin will cause the viscosity of the resin to increase from 25 cp to at least 4 cp in 90 seconds. You run five experiments to test their claim. The results of these experiments are summarized in the table on the next page. Based on your experiments, is the manufacturer's claim correct?



Provide statistical justification for your answer.

Time (see)	Trial #1 Viscosity (Cp)	Trial #2 Viscosity (Cp)	Trial #3 Viscosity (Cp)	Trial #4 Viscosity (Cp)	Trial #5 Viscosity (Cp)
30	37.4	36.7	37.0	37.2	36.8
60	38.2	40.4	41.5	39.8	39.3
90	44.3	45.3	48.2	44.7	45.1
120	51.7	50.6	52.7	51.2	52.0
150	60.2	59.7	59.4	60.0	60.6

Solution:

First, the student must realize that the only relevant data are the viscosities at 90 seconds. The manufacturer has made no claims about the behavior of the epoxy at any other times, Thus, 5 data points are relevant: 44.3, 45.3, 48.2, 44.7 and 45.1 cp. If the student calculates the mean and confidence intervals as in the previous problem, he or she will find that the mean is  $45.1 \pm 2.0$  cp. Thus, the manufacturer's claim would fall within our confidence limit. However, the third trial showed an atypically high viscosity. If the student applies the statistical q test to the point, he or she will find:

$$q = \text{gap}/\text{range} = (48.2-45.3)/(48.2-44.3)$$

$$q = 0.744$$

From the statistical q tables, the critical q value ( $q_c$ ) for rejection of a discordant value with five data points is 0.64. Thus,  $q > q_c$ , so the third point is omitted from the analysis.

With only 4 points to consider, the new mean is 44.85. Now, the confidence intervals maybe determined as before.

$$S = \sqrt{\frac{\sum_{i=1}^4 (x_i - \bar{x})^2}{3}} = 0.443$$

$$x = \bar{x} \pm \frac{t_{0.95} \cdot 0.443}{\sqrt{4}}$$



Thus, the viscosity at 90 seconds is  $44.85 \pm 0.0705$  cp. Even at the extreme confidence limit, this value is manufacturers claim. Therefore, it is 95 percent certain that the manufacturer's claim is not true.

### Conclusions

The problems presented allow students to make engineering evaluations based on statistical analysis techniques. From these problems and others like them, the students gain an engineering perspective on the role of statistics analysis in data interpretation. Because the students must make decisions from their analysis, they are better able to apply these techniques to their own laboratory data analysis requirements.

### References

*IP Embracing Contraries.. Explorations in Learning and Teaching*, Oxford University Press, New York, 1986.

### JAMES A. NEWELL

Jim Newell received his B.S. degree from Carnegie-Mellon, his M.S. from Penn State, and his Ph.D. in Chemical Engineering from Clemson University. After spending one year as a Visiting Assistant Professor at Clemson, he became an Assistant Professor at the University of North Dakota in 1995. His research interests involve polymeric and carbon fibers and composites.

