DESIGN OF EXPERIMENTS IN UNDERGRADUATE LABORATORY EDUCATION

Swami Karunamoorthy
Saint Louis University, St. Louis, Missouri

Abstract
Design of experiments is a necessary skill for a test engineer in an industry. In any engineering program, it is an important learning outcome. In this paper, an emphasis is given to how this skill can be developed in undergraduate laboratory education. Some examples are presented along with theoretical background that can be easily implemented in laboratory courses. It is a viable approach to give an exposure to design of experiments as well as to enhance the learning experience in laboratory education.

Introduction

Laboratory experiments are essential part of engineering curriculum. Traditionally, students in a laboratory course would set up an experiment, take measurements, analyze data, plot graphs, and write a report. This approach provides a learning experience on how to conduct experiments and how to analyze data. However, it does not provide an experience in design of experiments. "Design of experiment" means planning the experiment[1] and one of the aspects is statistical design of experiment. Statistical design in general implies the estimation of number of measurements or tests required to determine the true mean of a variable being measured. In a typical laboratory experiment a student group would collect only one set of data for a given specimen. But, similar set of data would be available from several groups performing the same test at different times. While it is difficult to apply the design of experiments on the data from any single group, one could pool the data from all groups and apply the principles of design of experiment. So, it is important to understand Pooled Statistics and hence some background on this theory is also included in this paper.

Statistics for Lab

Data scatter introduces an uncertain vagueness into the measurement scheme, which requires statistical methods to quantify. Statistics then becomes a powerful tool to interpret and present data. Statistics for laboratory experiments can be broadly classified into three categories.

1. Infinite Statistics: This is based on the assumption that very large amount of experimental data are available. Then the True Mean can be obtained from Normal Distribution Table.
2. Finite Statistics: This is based on the assumption that the available data are not large. In this case, the True Mean can be estimated from the available sample mean with the help of Student's "t" Distribution Table.

3. Pooled Statistics[2,3]: This is based on the assumption that the available data are small, however similar data sets are available from different teams performed the same test at different time frame. These individual data sets from each team can be grouped in a manner so as to determine a common set of statistics and such a statistics is in general called as Pooled Statistics. In this case, True Variance is obtained from available sample variance using "Chi – Square" Distribution Table.

In this paper, pooled statistics has been used to demonstrate the design of experiments. Also, the same data can be used to predict the precision interval of true variance. A sample calculation has been included in the example presented in this paper. The required background on Chi – Square distribution is given below:

Chi – Square Distribution[4]

σ – True Standard deviation
S_x – Sample Standard deviation
σ^2 – True Variance
$(S_x)^2$ - Sample Variance

One can estimate how well the sample variance would predict the true variance based on Chi-Square distribution.

$$\chi^2 = \nu\ (S_x)^2 / \sigma^2$$

where, ν is the degree of freedom and it can be defined as the number of independent measurements available. In general, the data points are scattered about a point known as MEAN point. Therefore, the number of available independent measurements is reduced by one. In general, if the number of data points are "N", then the degree of freedom,

$$\nu = (N-1)$$

$P(\chi^2) = 1 - \alpha$ = Level of Confidence, where "α" is Level of Significance (discrepancy). The value of $P(\chi^2)$ for given degree of freedom at a particular level of significance can be obtained from the Chi-Square table[2]. This method can be used to find the precision interval $[(\sigma_{\alpha/2})^2 \le (\sigma)^2 \le (\sigma_{1-\alpha/2})^2]$ for true variance at given confidence level or the confidence level for given variance of measured data[3].

Design of Experiments

The objective of the design of experiments is to estimate how many tests or measurements are required to find the True Mean of a variable with acceptable precision or confidence level. If the samples are drawn from two or more different populations, one has to test the equality of means by comparing the variances (F-test) and the procedure is called Analysis of Variance or ANOVA. However in undergraduate laboratory course, test samples are in general drawn from single population and the application of ANOVA may not be feasible and it is beyond the scope of this paper. In such cases the "Plan" or "Design" means to estimate the number of tests required to predict the True Mean[5,6] of population.

The true mean with a specified percentage of confidence level for finite set of data can be written as,

$X_T = X_{av} \pm d$; $d = [S_0 / \sqrt{(N)}]\,(t_{\nu,P})$; $\nu = (N - 1)$;

$(X_{av} - d) < X_T < (X_{av} + d)$

$CI = (X_{av} + d) - (X_{av} - d) = 2d$

X_{av} = sample mean
X_T = true mean
S_0 = sample standard deviation
N = number of data
ν = degree of freedom
P = percentage of confidence level
CI = Confidence Interval
$t_{\nu,P}$ = value from "t – Distribution" Table

Procedure
The sample data points can be increased by pooling the data from other teams within the lab class. From these data, N, ν, S_0 can be obtained easily. Assume the confidence interval as 1%, 5%, or 10% of the sample mean.

$CI = 0.01\ X_{av}$; $CI = 0.05\ X_{av}$; $CI = 0.1\ X_{av}$

$d = CI/2$

The total number of measurements or tests required for true mean at 95% confidence level (P=0.95) can be computed from,

$N_T = [(t_{\nu,95})\ S_0 / d]^2$; where $t_{\nu,95}$ is obtained from "t – Distribution" table.

The number of additional measurements or tests required can be found from,

$N_A = N_T - N$

Example: Design of Experiment for Tension Test
Objective: To learn the application of statistical design of experiment for tension test.
Outcome: Ability to design a test plan such that the average of yield stresses of all samples is within an acceptable confidence interval to true mean at a prescribed confidence level.

The lab class had four teams with four members in each team and they tested two samples of same material in tension. Each team plotted the stress-strain diagram and obtained the yield stress values. The pooled data with two values of yield stress per team resulted in a total of eight data points. They are given in Table 1.

Table 1: Yield Stress data for each tension test

Test	**Yield Stress (ksi)**
1	50.00
2	54.00
3	53.00
4	56.31
5	51.78
6	52.79
7	54.12
8	55.63

From the given data, the following values can be easily obtained.

$N = 8$, $\nu = 7$, $X_{av} = 53.45$ ksi (average value of yield stress), $S_0 = 2.04$, $P = 95\%$

Precision Interval:
$P = 0.95$, $\alpha = 1 - P = 0.05$

The following values can be obtained from a Chi-Square table for $\nu = 7$,

$(\chi_{\alpha/2})^2 = 16, \quad (\chi_{1-\alpha/2})^2 = 1.69$

$(\sigma_{\alpha/2})^2 = (\nu)\,(S_0)^2 / (\chi_{\alpha/2})^2 = (7)\,(2.04)^2 / (16) = 1.8$

$(\sigma_{1-\alpha/2})^2 = (\nu)\,(S_0)^2 / (\chi_{1-\alpha/2})^2 = (7)\,(2.04)^2 / (1.69) = 17.2$

The Precision Interval on True Variance: $1.8 \leq \sigma^2 \leq 17.2$

Design of Experiment:
Then, $t_{7,95} = 2.365$; $d = \{2.04/\sqrt{8}\}(2.365) = 1.71$ ksi

True Mean, $X_T = 53.45 \pm 1.71$ ksi; $\quad 51.74 < X_T < 55.16$

This is true with a confidence interval, $CI = 2(1.71) = 3.42 = (6.4\ \%$ of $X_{av})$

It shows that the true mean of yield stress can be obtained with given eight tests at 95% confidence level while the confidence interval is equal to 6.4% of sample average.

For other values of confidence interval, the total number of tests required and the number of additional tests to be performed are given in Table 2.

$N_T = [(t_{v,95})\ S_0/d]^2 = \{(2.365)(2.04)/d\}^2$; $d = CI/2$; $N_A = N_T - 8$

Table 2: Number of tests required for given Confidence Interval

CI	d	N_T	N_A
0.01 X_{av}	0.267	325	317
0.05 X_{av}	1.336	13	5
0.1 X_{av}	2.67	4	--
0.064 X_{av}	3.42	8	0

It can be observed from Table 2, that the number of tests required increases rapidly when the confidence interval narrows. In an ideal case, CI=0=d; which would lead to infinity for number of tests required. However in reality, a typical value of "CI" is less than or equal to 5% of sample mean. So, the number of tests required is finite and provides a meaningful design of experiment.

Conclusions

Undergraduate laboratory courses in general provide an experience on how to conduct an experiment and analyze data for small samples. However it is important to provide an exposure to design of experiment in order to attain the necessary learning outcome. The theory and example presented in this paper provide a viable approach on implementing design of experiment in a laboratory course. This approach would enhance the laboratory learning experience as well as enable the students to attain the outcome on ability to design and conduct experiments. The graduates with such experience would be better prepared to serve as effective test engineers in industry.

Bibliography

1. Schiff, D., D'Agostino, R.B., "Practical Engineering Statistics," John Wiley and Sons, 1996.
2. Figliola, R.S., Beasley, D.E., "Theory and Design for Mechanical Measurements," John Wiley and Sons, 2000.
3. Vardeman, S.B., Jobe, J.M., "Basic Engineering Data Collection and Analysis," Duxbury, 2001.
4. Mason, R.L., Gunst, R.F., Hess, J.L., "Statistical Design and Analysis of Experiments," John Wiley and Sons, 2003.
5. Metcalfe, A.V., "Statistics in Engineering – A Practical Approach," Chapman and Hall, 1994.
6. Lipson, C., Sheth, N., "Statistical Design and Analysis of Engineering Experiments," McGraw Hill, 1973.

Biographical Information

DR. SWAMI KARUNAMOORTHY

Dr. Swami Karunamoorthy is the professor of Aerospace and Mechanical engineering at Parks College of Engineering, Aviation, and Technology, Saint Louis University, St. Louis, Missouri. He is the author of several publications in Aerospace Engineering, Mechanical Engineering, Engineering Education and Assessment Methods. He is the regional Vice President of AHS and Chair of ASME – St. Louis Section. Also, he is an Associate Fellow of AIAA and an active member of ASEE.