Mathematical Sciences Technical Reports (MSTR)	Mathematics

9-23-1998

A Messy, but Instructive, Case Study in Design of Experiments

Leroy A. Franklin
Rose-Hulman Institute of Technology

Belva J. Cooley
Indiana State University

Gary Elrod
Siemans

A MESSY, BUT INSTRUCTIVE, CASE STUDY IN DESIGN OF EXPERIMENTS

Belva J. Cooley, LeRoy A. Franklin, & Gary Elrod

MS TR 98-05

September 23, 1998

Department of Mathematics
Rose-Hulman Institute of Technology

FAX: (812) 877-3198   Phone: (812) 877-8391

A Messy, But Instructive, Case Study
in Design of Experiments

Belva J. Cooley
Business Analytical Department
School of Business
Indiana State University
Terre Haute, Indiana 47809

LeRoy A. Franklin
Department of Mathematics
Rose-Hulman Institute of Technology[1]
5500 Wabash Avenue
Terre Haute, Indiana 47809

Gary Elrod
Statistical Consultant
Siemens, Inc.
Franklin, KY 42134

[1] On Leave from Indiana State University

A Messy, But Instructive, Case Study
in Design of Experiments

Abstract

A company manufacturing fans wished to conduct an experimental design to determine

the best combination of three factors affecting the breaking torque of the fans. Analysis of the

data ceased to be straightforward when the authors found that the data failed the test for

homogeneity of variances. After unsuccessfully attempting to transform the data and thereby

meet the assumptions necessary to carry on the analysis, the authors relied upon a graphical

analysis and a careful study of the means for each design point. This paper describes a

statistically sound but novel strategy used to complete the analysis.

Introduction

Even when a carefully crafted experimental design is properly carried out, sometimes

unexpected results force a creative and innovative approach to the data analysis. This point was

illustrated to the authors when they were asked by a major industrial supplier of fans to aid in the

design and analysis of an experiment intended to facilitate better understanding of the "torque

breaking strength" for several assembly procedures.

The company had recently acquired a smaller competitor who had assembled their fans from different components and with different methods. Considerable discussion resulted about which components and assembly methods were superior. In addition to "choosing best practices," the acquisition presented opportunities to standardize and reduce inventories. In all these activities it was important that wise, data-based decisions were employed and, thus, outside assistance was deemed necessary in making these crucial decisions.

The authors were contacted and, after considerable discussion, company representatives decided to focus on three factors in the manufacturing process: first, the type of hole in the fan "spider" itself (two types of holes); second the type of "barrel" to which the fan "spider" was attached (two types of barrels); third, the assembly method (two types of assembly methods). This gave a straight forward "three factor-with each factor at two levels" design of experiments, commonly called 2^3, with each factor qualitative (see Ref. 3).

Since much was riding on the decision and it was not extremely expensive to produce a single fan assembly, company representatives decided to have eight fan assemblies made of each of the eight possible combinations of components and assembly methods. After the fans were assembled, they would then be deliberately twisted apart and the torque (in foot pounds) necessary to do this would be recorded as the dependent variable of "Breaking Torque."

The Data and Preliminary Analysis

The assembly and destructive testing of the fans was conducted and the data sent to the authors for what was anticipated to be a relatively straightforward analysis. The original data is presented in Table 1 with the levels of the three qualitative factors (Hole, Assembly, and Barrel)

Table 1

recorded in columns 1,2, and 3 respectively (coded as -1 and +1 for each of the two levels for each factor) with the breaking torque, as originally provided to the authors, in column 4. A quick scan of this data showed several (five) "missing data points", all from one combination of settings of the factors. The authors were on the verge of applying a general linear models analysis since the data was unbalanced when they decided to inquire about the nature and cause of the missing data.

The technicians answered that the fan assemblies in question were **so strongly held together** that the technicians could not manually break them apart! Since producing strong fan assemblies was the goal of this manufacturing step, it was important to include such data if possible.

Further discussion with the technicians revealed that they were unable with their equipment to exert more that about 200 ft-lbs of breaking torque. Thus, all five "missing data" fans were therefore of at least 200 ft-lbs of torque. The authors decided that since this component-assembly combination was potentially the best combination that these values should and could be "imputed" to be 200 and this fact of "imputation" be kept in mind throughout the rest of the analysis. This "imputed" data is listed in the last column in Table 1. (Note: all calculations, displaying, and plotting of the data was done by the MINITAB statistical package.) (See Ref. 2.)

Before detailed analysis is initiated, it is always recommended to obtain a graphical representation of the data (whenever possible). Following this guideline, sample means and standard deviations (including the imputed data) were obtained for the eight possible component-assembly combinations. Since the mean breaking strength was of central interest and since there

were four dimensions in which to work (three for the independent factors and a fourth for the dependent or "response" variable of breaking torque), it was necessary to provide two separate three-dimensional plots of the sample means. A scan of the sample means seemed to show that higher means were for the value of Barrel=+1 and thus it was decided to divide the plots so that one three-dimensional plot would be for Barrel=-1 and the second for Barrel=+1. These are displayed as Figures 1 and 2 (respectively).

Sample means and standard deviations were calculated for each of the eight distinct experimental combinations to examine how closely the assumption of "equality of variances" was followed. These are displayed in Table 2. A scan of the sample standard deviations reveals a wide range for their values. Thus (as is always advised) a test of homogeneity of variances was conduced and both Bartlett's test (for normal distributions) and Levene's test (for any continuous distribution) gave strong statistical evidence that the variances were not all equal (p-value=.000 and .050, respectively)(see Ref. 2 and 3). Furthermore, even though imputation of the "missing data" now gave equal numbers of observations for each combination, the ratio of largest to smallest variances was 31, far exceeding general guidelines given by those such as George E. P. Box (1)of keeping the ratio within a factor of nine for unequal variances.

The usually suggested solution to non-homogenous variances is to apply a standard transformation of the data to yield approximately equal variances. But several common transformations (e.g., reciprocal, square, square-root, and log) yielded no improvement in the tests for homogeneity of variances for this data set (see Ref. 3, 4, and 5). A close look at the standard deviations of Table 1 help to highlight why. The four largest means had the two largest standard deviations and two mid-sized standard deviations, when compared among themselves.

Figure 1

Figure 2

Table 2

But in looking at the four smallest means, the smallest and largest of this group had the largest standard deviations of this group and the two mid sized means had the smallest standard deviations of this group. Thus no clear relationship existed which lent itself to one of the common transformations. What then could be done?

Since the four largest sample means were for when barrel=+1, the authors wondered if the problem of non-homogeneity of variances could be "uncoupled" from the problem by breaking the one large problem into two smaller problems (one for barrel=-1 and the other for barrel=+1). A test of homogeneity of variances for when barrel=+1 showed no significant non-homogeneity of variances either by Bartlett's test or Levene's test (p-value= .268 and .354 respectively). Thus a standard **two factor** analysis of variance could be run for when barrel=+1. However the test of homogeneity of variances for when barrel=-1 still showed significant non-homogeneity of variances by both Bartlett's and Levene's test (p-values=.003 and .000 respectfully) and in addition no common transformation seemed to remove this problem even from only one half of the original data.

Confident of having removed the problem of not having equal variances (at least from one half of the data that had the largest set of four means), a standard two factor analysis of variance was run for when barrel=+1 with the results displayed in Table 3 along with the cell means. This table showed that the two factors of hole and assembly were very significant with no interaction present. Thus the data and the 3 dimensional plot for this half of the data are easily interpretable: Assembly =-1 gives a higher mean breaking torque; hole =+1 gives a higher mean breaking torque.

A post hoc comparison by Fisher's Least Significant Differences Method (2) (using the

Table 3

common mean square of 51.9 and degrees of freedom of 28), showed all four means in this group

to be significantly different from one other and the 95% confidence interval for the difference of

any two such means would be +/-7.37 ft-lbs. Thus, the treatment with the highest sample mean

(hole=+1, assembly=-1, and barrel=+1) of 196.25 ft-lbs was significantly greater than even the

second highest mean (hole=-1, assembly=-1, and barrel=+1) of 175.62 ft-lbs. and we could be

95% confident that this difference was between 20.63 +/-7.37 ft-lbs.

In order to deal with the thorny problem of the four smallest means and their non-

homogenous variances, it is only necessary to realize that **we do not even need to perform an**

analysis of variance on them. Remember, our goal was only to find the superior (greater mean)

breaking torque combination. Thus the authors realized that it was only necessary to establish

that these four smaller means were significantly statistically lower that the other means. The

authors choose to compare the smallest sample mean from the group of the four largest sample

means (ie. barrel=+1) against each of the four smallest means. While it is not recommended

that many student's t tests be performed on any data set (6), here some justification could be

obtained to do four distinct two sample-independent sample (with unequal variances) t-tests (6),

since it was clear there were large differences in the variances of the four samples with the

smallest means.

The results of the these four tests are displayed in Table 4. These results clearly show

that even the "smallest mean of the group which had the largest four means" is very significantly

different from (and higher than) each of the means in the group which had the smallest four

means.

Table 4

Summary and Recommendations

Because of the problem of inequality of variances and the intractability of the data to transformation, it was not possible to do the standard analysis of variance on the data. Instead the graphical plot of the means and the table of sample means and standard deviations suggested a natural dividing of the data into two parts: one which had no problems of inequality of variances and the other which retained that specific problem. Fortunately in this "decoupling" of the data, the set of larger means (the ones of greatest interest) was the one which then could be analyzed by standard analysis of variance techniques and revealed that all four of the larger means were distinctly different. Furthermore, using a two sample/independent sample student's t test (with unequal variances) allowed comparisons of each of the means (in the group of the four smaller means) to the smallest mean (in the group of the four largest means). Since these differences were so great there was little concern about problems of inflated Type I error due to conducting several distinct t-tests on the data.

Since each of the means in the group of four largest means were all much higher than that desired, it was recommended that the barrel=+1 be adopted immediately for all subsequent fan assembly. Furthermore it was recommended that the current inventories of fan spiders be consumed (no matter what the hole shape) but with the goal of moving deliberately to hole=+1 as new inventory was ordered. Furthermore, it was recommended that, as assembly machines were naturally replaced, the company should move to assembly technique=-1.

And what of the imputed data? Those data were for what indeed turned out to be the statistically, significantly highest combination. Since the imputation of the data was deliberately

conservatively low, the manufacturers could be confident that the fan assemblies would be even

stronger than estimated by the confidence intervals which used the imputed data.

References

1. Box, George E. P., "Effects of Inequality of Variance and of Correlation Between Errors in the Two-Way Classification", *Annals of Mathematical Statistics*, 25, 484-498 (1954)

2. *Minitab User's Guide, Release 12*, Minitab, 1997.

3. Montgomery, Douglas C., *Design and Analysis of Experiments*, 3rd. Edition, John Wiley & Sons, New York, 1991.

4. Neter, John, Wasserman, William and Kutner, Michael H., *Applied Linear Statistical models: Regression, Analysis of Variance, and Experimental Designs*, Richard D. Irwin, Homewood, Illinois, 1990.

5. Ryan, Thomas P., Modern Regression Methods, John Wiley & Sons, New York, 1997

6. Triola, Mario F. And Franklin, LeRoy A., *Business Statistics*, Addison-Wesley Publishing Company, Reading, Massachusetts, 1994

Table 1: The "Missing" and "Imputed" Experimental Data

Row	hole	assembly	barrel	torque1	torque2
1	-1	-1	-1	55	55
2	-1	-1	-1	54	54
3	-1	-1	-1	53	53
4	-1	-1	-1	52	52
5	-1	-1	-1	55	55
6	-1	-1	-1	51	51
7	-1	-1	-1	51	51
8	-1	-1	-1	55	55
9	1	-1	-1	111	111
10	1	-1	-1	105	105
11	1	-1	-1	100	100
12	1	-1	-1	100	100
13	1	-1	-1	110	110
14	1	-1	-1	100	100
15	1	-1	-1	102	102
16	1	-1	-1	110	110
17	-1	1	-1	40	40
18	-1	1	-1	50	50
19	-1	1	-1	50	50
20	-1	1	-1	40	40
21	-1	1	-1	41	41
22	-1	1	-1	37	37
23	-1	1	-1	50	50
24	-1	1	-1	45	45
25	1	1	-1	50	50
26	1	1	-1	50	50
27	1	1	-1	51	51
28	1	1	-1	47	47
29	1	1	-1	50	50
30	1	1	-1	52	52
31	1	1	-1	52	52
32	1	1	-1	50	50
33	-1	-1	1	160	160
34	-1	-1	1	175	175
35	-1	-1	1	180	180
36	-1	-1	1	165	165
37	-1	-1	1	185	185
38	-1	-1	1	185	185
39	-1	-1	1	185	185
40	-1	-1	1	170	170
41	1	-1	1	190	190
42	1	-1	1	190	190
43	1	-1	1	190	190
44	1	-1	1	*	200
45	1	-1	1	*	200
46	1	-1	1	*	200
47	1	-1	1	*	200
48	1	-1	1	*	200
49	-1	1	1	160	160
50	-1	1	1	150	150
51	-1	1	1	150	150
52	-1	1	1	150	150
53	-1	1	1	151	151
54	-1	1	1	170	170
55	-1	1	1	151	151
56	-1	1	1	147	147
57	1	1	1	160	160
58	1	1	1	160	160
59	1	1	1	170	170
60	1	1	1	160	160
61	1	1	1	170	170
62	1	1	1	170	170
63	1	1	1	170	170
64	1	1	1	170	170

Table 2: Sample Means and Standard Deviations for the Eight Experimental Combinations

Row	hole	assembly	barrel	mean	st.dev.
1	-1	-1	-1	53.25	1.753
2	1	-1	-1	104.75	4.920
3	-1	1	-1	44.13	5.330
4	1	1	-1	50.25	1.581
5	-1	-1	1	175.62	9.800
6	1	-1	1	196.25	5.180
7	-1	1	1	153.62	7.610
8	1	1	1	166.25	5.180

Table 3: Analysis of Variance (for Barrel= +1)

Analysis of Variance for Torque

Source	DF	SS	MS	F	P
Hole	1	2211.1	2211.1	42.62	0.000
Assembly	1	5408.0	5408.0	104.23	0.000
Hole*Assembly	1	128.0	128.0	2.47	0.127
Error	28	1452.7	51.9		
Total	31	9199.9			

Means

Hole	N	Torque
-1	16	164.63
1	16	181.25

Assembly	N	Torque
-1	16	185.94
1	16	159.94

Hole	Assembly	N	Torque
-1	-1	8	175.62
-1	1	8	153.62
1	-1	8	196.25
1	1	8	166.25

MTB >

TABLE 4 : Unequal Variances, Two Sample-Independent Sample Student's t Test

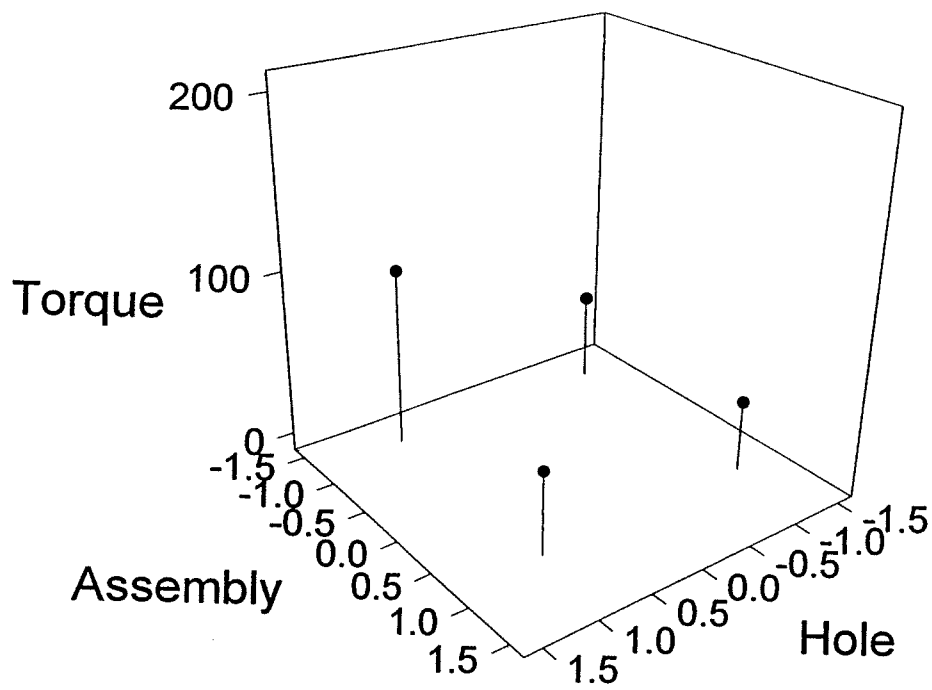| Difference of Sample Means	95% Confidence Interval Bound
mean(-1,+1,+1) - mean(-1,-1,-1)=153.62 - 53.25= 100.37ft-lbs	+/- 6.56 ft-lbs
mean(-1,+1,+1) - mean(+1,-1,-1)=153.62 -104.75= 48.87ft-lbs	+/- 7.05 ft-lbs
mean(-1,+1,+1) - mean(-1,+1,-1)=153.62 - 44.13= 109.49ft-lbs	+/- 7.20 ft-lbs
mean(-1,+1,+1) - mean(+1,+1,-1)=153.62- 50.25= 103.37ft-lbs	+/- 6.49 ft-lbs

Figure 1: Breaking Torque (for Barrel=-1)

Figure 2: Breaking Torque (for Barrel=+1)