Small F-ratios: Red Flags in the Linear Model

Gary E. Meek[1], Ceyhun Ozgur[2] and Kenneth A. Dunning[3]
[1]*Black Hills State University,* [2]*Valparaiso University and*
[3]*The University of Akron*

Abstract: All textbooks and articles dealing with classical tests in the context of linear models stress the implications of a significantly large F-ratio since it indicates that the mean square for whatever effect is being evaluated contains significantly more than just error variation. In general, though, with one minor exception, all texts and articles, to the authors' knowledge, ignore the implications of an F-ratio that is significantly smaller than one would expect due to chance alone. Why this is so is difficult to explain since such an occurrence is similar to a range value falling below the lower limit on a control chart for variation or a p-value falling below the lower limit on a control chart for proportion defective. In both of those cases the small value represents an unusual and significant occurrence and, if valid, a process change that indicates an improvement. Therefore, it behooves the quality manager to determine what that change is in order to have it continue. In the case of a significantly small F-ratio some problem may be indicated that requires the designer of the experiment to identify it, and to take "corrective action".

While graphical procedures are available for helping to identify some of the possible problems that are discussed they are somewhat subjective when deciding if one is looking at an actual effect; e.g., interaction, or whether the result is merely due to random variation. A significantly small F-ratio can be used to support conclusions based on the graphical procedures by providing a level of statistical significance as well as serving as a red flag or warning that problems may exist in the design and/or analysis.

Key words: F-ratio, linear models, non-linearity, non-additivity, quality control.

1. Introduction

Control chart procedures have always stressed that any significant or unusual occurrence must be investigated and explained in attempting to bring a process into control; i.e., to have a process with only common cause or error variation in it. As indicated in Montgomery (1996) any significant value indicates that a

change has occurred in the process. Whether the change is for the better or for the worse is irrelevant. All are to be investigated. For p-charts, c-charts and R-charts Gitlow, Oppenheim and Oppenheim (1995-chapter 5) state that values below the lower control limit are "good" values in the sense that they indicate that there may have been an improvement in the process. If such is the case and the cause can be identified then making the change part of the process results in a permanent improvement in quality. While the "small" value may represent simply a chance occurrence all other possibilities are to be eliminated before that conclusion is reached.

This philosophy does not seem to be in use in the field of experimental design and statistical analysis in general, particularly in the various tests associated with linear models. All of the F-ratios in linear models with fixed effects are constructed essentially as the mean square (MS) for the effect of interest divided by an estimate of the error variance. If the null hypothesis is true and all assumptions underlying the procedure are satisfied then the F-ratio is expected to be near 1.0. If the null hypothesis is false and all assumptions satisfied then the mean square for the effect of interest contains both an estimate of the error variance and a sum of squared terms attributable to the effect of interest. If the effects are random then the $E(MS)$ for the effect of interest includes the variance for that effect plus a linear combination of variances for various interactions and the error term. The F-ratio then compares the effect's MS to a MS whose $E(MS)$ is the linear combination of the variances of the various interactions and the error term. Again, if H_0 is true the variance of the effect of interest is zero and the F-ratio is expected to be 1.00 while if H_a is true the ratio is expected to exceed 1.00. In a model with mixed effects each term's $E(MS)$ must be considered individually in determining the F-ratio. In all cases, the only values indicating rejection of the hull hypothesis, or supporting the alternative, are large ones. Values for the F-ratio that are less than 1.0 simply lead to non-rejection of the null hypothesis and generally are not investigated any further, regardless of their actual magnitude. This paper suggests that a significantly small value for the F-ratio should be investigated further to determine if an explanation can be identified.

2. Literature Search

Many textbooks have been written on the topics of linear models, analysis of variance (ANOVA) and design of experiments since Sir Ronald A. Fisher's original papers were published on agricultural experiments. The texts by Winer (1962) and Davies (1963) were concerned primarily with industrial and chemical experiments. Cochran and Cox (1957) and Scheffe' (1959) wrote texts that took a more general approach, utilizing a wide range of applications, and have become classics on experimental design and analysis of variance, respectively. Among

the more recent texts that have been published and are fairly popular are Hinkelman and Kempthorne (1994), Neter, Kutner, Nachtsteim and Wasserman (1996), Bowerman and O'Connell (1990) and Montgomery (1997). Taguchi (1986) had a major impact on the implementation of experimental design concepts in the area of process optimization. The recent texts vary in terms of level of theoretical presentation and emphasis on applications but none make any mention of giving consideration to the possible implications of an F-ratio that is significantly small.

The only text that indicates that a small value for the F-ratio should be flagged is an introductory statistics text by Meek and Turner (1983, p. 456). Meek and Turner's (1983, p. 456) only reference to the small F-ratio is in an example of a two-factor crossed design in ANOVA. In that example they note that if the problem is analyzed as a one-factor model a small F-ratio occurs and that it should be investigated further. Meek and Turner (1983, p. 456) point out that, in the example being discussed, the correct analysis was for a two-factor model and that the small F-ratio is an indication of a miss-specified model, or in other words, lack of fit. To date, the only real discussion of the implications of a significantly small F-ratio was in a preliminary paper by Meek, Ozgur and Dunning (2005) presented at the Decision Sciences Institute's Annual meeting and published in the meeting's proceedings.

3. The General Case

Suppose the general linear model in

$$\mathbf{y}_{n\times 1} = \mathbf{X}_{n\times p}\beta_{p\times 1} + \boldsymbol{\epsilon}_{n\times 1} \tag{3.1}$$

is the correct representation of $\mathbf{y}$ and all distributional assumptions are satisfied. Then the total sum of squares may be partitioned into the sum of the error sum of squares and the regression or model sum of squares; i.e., as $SSTO = SSE + SSR$, where $SSTO = \mathbf{y}'\mathbf{y} - \mathbf{b}'\mathbf{X}'\mathbf{y} - (\sum y)^2/n$, $SSE = \mathbf{y}'\mathbf{y} - \mathbf{b}'\mathbf{X}'\mathbf{y}$ and $SSR = \mathbf{b}'\mathbf{X}'\mathbf{y} - (\sum y)^2/n$. The regression sum of squares in turn can be partitioned into

$$SSR = SS(b_1|\mathbf{b}_1^*) + \cdots + SS(b_p|\mathbf{b}_p^*), \tag{3.2}$$

where $\mathbf{b_j}^* = (b_0, \ldots, b_{j-1})'$ and b_i is the estimate of β_i in the parameter vector. The design matrix, $\mathbf{X}$, in (3.1) may be decomposed into component matrices. For example, suppose a component consists of k of the independent variables. Then the columns of $\mathbf{X}$ may be rearranged so that the variables of interest correspond to the last k columns. Thus, the design matrix can be represented as $\mathbf{X} = [\mathbf{X}_{p-k} : \mathbf{X}_k]$, the first $p - k$ columns plus the last k columns of $\mathbf{X}$. The reduced model resulting from this partitioning is

$$\mathbf{y} = \mathbf{X}_{p-k} + \boldsymbol{\beta}_{p-k} + \boldsymbol{\epsilon} \tag{3.3}$$

where $\boldsymbol{\beta}_{p-k}$ is a $p-k$ vector of unknown parameters. Then, for the reduced model in (3.3), the sum of squares corresponding to variation explained by the k omitted variables would be the last k terms on the right in (3.2); i.e.,

$$SS(\mathbf{b}_k|\mathbf{b}_{p-k}) = SS(b_{p-k+1}|\mathbf{b}^*_{p-k+1} + \cdots + SS(bkb_p) \tag{3.4}$$

The terms in Equation (3.4) are included as part of the SSE when the model in Equation (3.3) is used and will inflate the estimate of σ^2 if they represent more than chance variation. This, in turn, may result in an F-ratio that is significantly smaller than would be expected by chance alone. If it is, then an effort should be made to determine if the model being considered is correct and/or if any underlying assumptions may be questionable; i.e., try to find a reason. It is irrelevant whether the X_j's are quantitative variables, regression, or qualitative variables, ANOVA, or a combination of the two, ANOCOVA.

The next section presents specific applications, with examples, of F-ratios that are significantly smaller than would be expected by chance alone. While these applications are restricted to the general case of omitted terms or factors in the model other causes such as a violation of the normality or homogeneity of variance assumptions may result in small F-ratios also. In regression analysis the presence of multi-collinearity also may result in unusually small F-ratios for tests of the individual coefficients.

4. Specific Applications

There are several possible reasons for the occurrence of a small F-ratio in tests of hypotheses in ANOVA. If all of the underlying assumptions are satisfied and the correct model has been specified then, other than data manipulation, the only explanation is that of chance variation. In practical applications, though, one almost never knows how closely the data fit the assumptions or if there are any terms or factors that have been omitted from the model. Thus, any time a small value is obtained for the F-ratio the experimenter should check all of the assumptions and reexamine the model being used. Possible implications are presented below for three types of situations.

4.1 Randomized Block Design

The basic randomized block design is simply a two factor crossed design with one observation per cell in terms of its analysis. The basic model is given in Equation (4.1) and represents a general linear model with two qualitative variables.

$$y_{ij} = \mu + \alpha_j + \beta_i + \epsilon_{ij} \tag{4.1}$$

In addition to the usual assumptions underlying the ANOVA procedures the assumption of no interaction is necessary for constructing the test statistics for evaluating both treatment and block effects when the effects are fixed. Letting α represent the treatment effect the mean square associated with treatments, MSA, then has an expected value of $\sigma^2 + \sum \alpha_j^2$ while the term used for the mean square 'error', $MSAB$, has an expected value of σ^2 only if the interaction term is zero. Thus, the F-ratio for the null hypothesis of no treatment effect; i.e., H_0 : all $\alpha_j = 0$ vs. H_a :: some $\alpha_j \neq 0$, is $MSA/MSAB$ and is expected to be near 1.0 if H_0 is true and significantly greater than 1.0 if H_0 is not true. There is no situation in which it is expected to be near 0.0.

On the other hand, if the assumption of no interaction is violated the model for an observed value is given in Equation (4.2).

$$y_{ij} = \mu + \alpha_j + \beta_i + \alpha\beta_{ij} + \epsilon_{ij} \qquad (4.2)$$

The expected value for the MSA is unchanged but the expected value for the MSAB becomes $\sigma^2 + \sum (\alpha\beta)^2$. Now if H_0 is true $MSA/MSAB$ may be small while if it is false the ratio could be small or large depending on the relative magnitude of the interaction effect. In either case a value near 0.0 for the F-ratio should be checked for significance and an attempt should be made to determine a cause. Tukey's (1940) test for non-additivity can be used to check formally for a violation of the assumption of no interaction. If both treatments and blocks represent random effects then the correct F test statistic is $MSA/MSAB$ whether interaction is present or not and a significantly small F-ratio might indicate some other problem.

Example 1: The Graduate Management Admission Test (GMAT) is an examination used by graduate schools of business (management) to assess an applicant's ability to pursue an academic graduate program in business. Scores on the GMAT range from 200 to 800 with higher scores implying higher aptitude. In an attempt to improve student performance on the GMAT exam, a major Ohio university is evaluating offering the following 5 GMAT preparation programs.

Program 1: A three-hour review session covering the types of questions generally asked on the GMAT.

Program 2: A one-day (8 hour) review session covering the relevant material, along with taking and grading a sample exam.

Program 3: A one-week preparation program covering the relevant material, along with taking and grading a sample exam.

Program 4: A 4-week intensive preparation program, providing study and relevant clues, along with taking and grading a sample exam.

Program 5: An intensive 10-week course (4 hours per week) involving identification of student's weaknesses and setting up of individualized programs to assist each student.

It is believed that GMAT scores may be related to students' majors. Using major as a blocking factor the following three levels are selected:

1. Student's undergraduate study is from the College of Business Administration.

2. Student's undergraduate study is from the College of Engineering

3. Student's undergraduate study is from the College of Art and Social Sciences

Five students are selected from each major and programs are randomly assigned to them. All students sit for the GMAT at the next offering after they have completed their programs. The GMAT test scores received for this study are presented in Table 1.

Table 1: GMAT Scores of students classified by Exam Preparation Program

	One hour	One-day	One-week	4-weeks	10-weeks
College of Business	490	450	600	520	540
College of Engineering	520	570	540	540	610
College of Arts & Social Sciences	520	450	420	450	400

An analysis of variance of the scores received resulted in the following:

Analysis of variance for GMAT

Source	DF	SS	MS	F	P
Program	4	1707	427	0.14	0.962
Major	2	30240	15120	5.02	0.039
Error	8	24093	3012		
Total	14	56040			

Note that, in the above results, the block effect is significant but the treatment effect does not appear to be significant. Contrary to expectations, the F value

for the different means between the programs is unusually close to zero and is significantly smaller than would be expected merely by chance, the p-value associated with this factor is unusually close to 1, $(1 - p = .038)$. This situation should warrant further investigation. Based on the small F-value Tukey's test for non-additivity was performed. The F value for Tukey's test was 0.42, which does not exceed the table value of $F_{.01,1,7} = 12.25$. Therefore, there is insufficient evidence, based on Tukey's test for non-additivity, to conclude that interaction effects exist. Interaction plots were also constructed and are presented below in Figure 1.
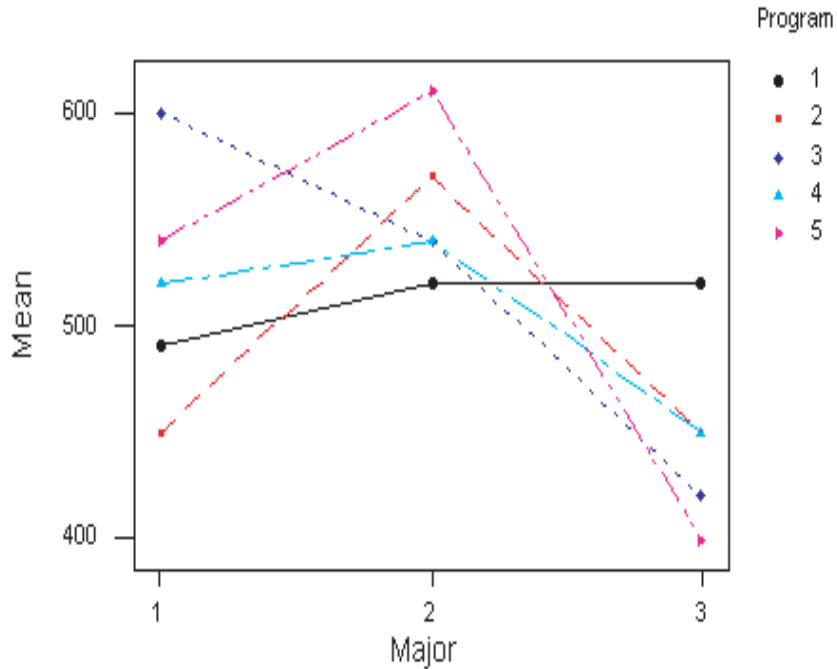

Figure 1: Interaction plot of the test scores data given in Table 1

The plots in Figure 1 indicate that interactions may be present since the lines cross, especially with respect to program 1. Based on those plots, the analyst decides to redesign the study prior to the next offering of the GMAT as a two factor crossed design with replication. Ten students from each major are randomly assigned with two to each program. Again, all of the students sit for the GMAT at the first offering after completing their respective programs. The resulting GMAT scores are summarized in Table 2:

Table 2: GMAT scores of students classified by college and exam preparation program – two observations per cell

	One-hour	One-day	One-week	4-weeks	10-weeks
College of	500	440	610	530	520
Business	470	480	570	580	570
College of	530	560	550	600	550
Engineering	490	520	590	590	580
College of Arts &	510	460	410	420	440
Social Sciences	530	430	490	470	520

In this model we are able to independently measure interaction. Referring to the data in Table 2, an analysis of variance gave the following results:

Analysis of variance for GMAT2

Source	DF	SS	MS	F	p
Major	2	40220	20110	19.15	0.000
Program	4	12980	3245	3.09	0.048
Interaction	8	26480	3310	3.15	0.026
Error	15	15750	1050		
Total	29	95430			

Note that, in the above ANOVA table, both programs and the interaction effect are significant at an α level of .05. Getting a significantly small F-value indicated a possible problem with the first study. While the second study indicated a significant interaction effect, that may or may not have caused the significantly small F-ratio in the first study.

4.2 Omitted factors

The previous example concentrated on the possibility of the presence of an interaction effect's contributing to a significantly small F-ratio. It cannot be stated that an interaction effect is definitely the problem in that example since factors such as grade point average (GPA), amount of work experience and/or motivation were not considered in either study. Any time a factor is inadvertently omitted from the model there is the possibility of obtaining unusually small F-values. To illustrate the rationale behind this concept a one- factor model is compared to a two-factor model. The models are given in Equations (4.3) and

(4.4).

$$y_{ij} = \mu + \alpha_j + \epsilon_{ij} \qquad (4.3)$$
$$y_{ijk} = \mu + \alpha_j + \beta_i + \alpha\beta_{ij} + \epsilon_{ijk} \qquad (4.4)$$

If the model in Equation (4.3) is correct then MSA/MSE is expected to be close to 1.0 if H_0 is true and large if it is not since, assuming a fixed effect, $E(MSA) = \sigma^2 + \sum \alpha_j^2$ and $E(MSE) = \sigma^2$. If a factor was omitted from the model then, as is illustrated in Equation (4.4), the sums of squares (SS) for both it and its interaction with the other factor will be included in the error term. If the missing factor(s) has a significant effect then the error mean square can be greatly inflated, resulting in significantly small F-ratios. The extension to higher order models is straightforward since the SS for the missing factor(s) and its (their) interactions, both first and higher order, with each other and with all terms specified in the model will be included in the SSE.

Example 2: Because of the high cost of hospital confinement and the need to free facilities, the average hospital stay for women giving birth has been diminishing. A study was undertaken to determine whether the average confinement was the same for four area hospitals. The data in Table 3 represent the number of days of hospital stay based on time between check-in and check-out (Obtained in a modified form from Meek, Taylor, Dunning and Klafehn (1987, p.375).

Table 3: Number of days spent by women in four hospitals after giving birth

	Hospitals		
Edgewood	Lincoln	Charity	General
9.5	9.7	7.1	8.7
7.9	9.4	8.4	9.0
9.1	8.4	7.9	8.9
2.7	2.6	2.5	2.1
3.2	2.7	2.0	3.2
2.0	2.4	2.3	2.5
3.4	3.4	4.0	3.4
3.5	3.4	2.9	3.6
3.5	3.0	3.0	3.8

If a one way analysis of variance is run, using the following model: $y_{ij} = \mu + \alpha_j + \epsilon_{ij}$, where α_j = Hospitals, the following results are obtained:

Analysis of variance

Source	DF	SS	MS	F	p
Hospitals	3	2.01	0.67	0.08	0.971
Error	32	272.41	8.51		
Total	35	274.42			

Note that the calculated F value is very small and the p-value is very close to 1, or $1 - p = .029$ which is less than .05. This is an unusually small value for the F statistic. As discussed above, this can happen when an important factor(s) is (are) left out of the model. If we plot the data with box plots, we see the following results in Figure 2.
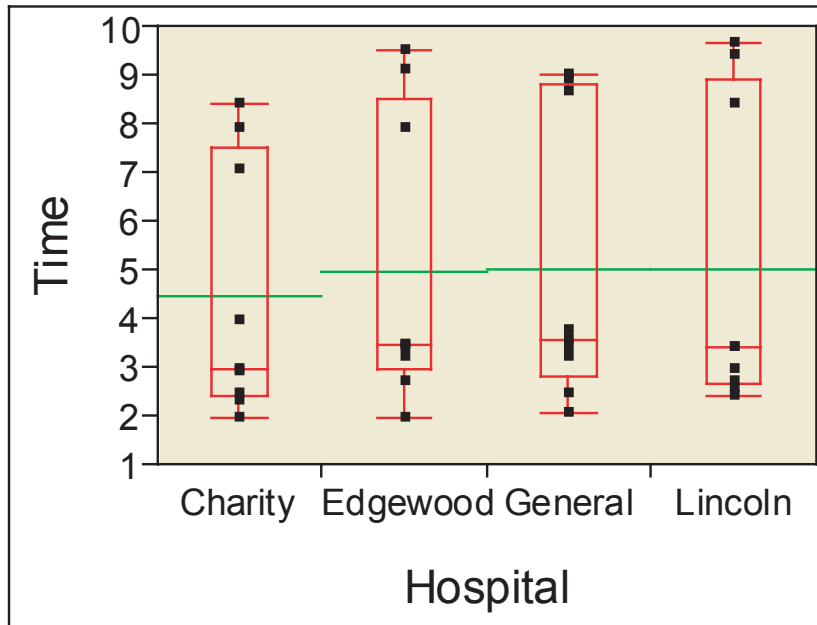

Figure 2: Box Plot of data by hospitals

The box plots lead to two important observations. First, the variation may be the same in each of the populations, and second, the data points within each population appear to be in two widely separated clumps. ANOVA F tests require that the population variances be homogeneous and that the populations be normally distributed. The null hypothesis of equal variances can be tested using Hartley's F max test which compares sample variances from several populations using the ratio of the largest sample variance to the smallest sample

variance. Here it is applied to the two hospitals with the largest and smallest sample variances. The F-ratios are

$$
\begin{aligned}
F_{\max(calc)} &= \frac{\sigma^2_{\text{largest}}}{\sigma^2_{\text{smallest}}} = \frac{9.9925}{6.7077} = 1.49 \\
F_{\max(table)} &= F_{\max c, \bar{n}-1} = F_{\max 48} = 7.18,
\end{aligned}
$$

where $c = \#$ of grpups amd $\bar{n}-1 = $ [mean $\#$ of observations (rounded down) per group] -1 with $\alpha = 0.05$. Since $1.49 < 7.18$, the assumption of homogeneity of variance does not appear to be violated. Though the sample sizes are somewhat small we can look at the distributions for the individual hospitals, shown in Figure 3 as histograms.
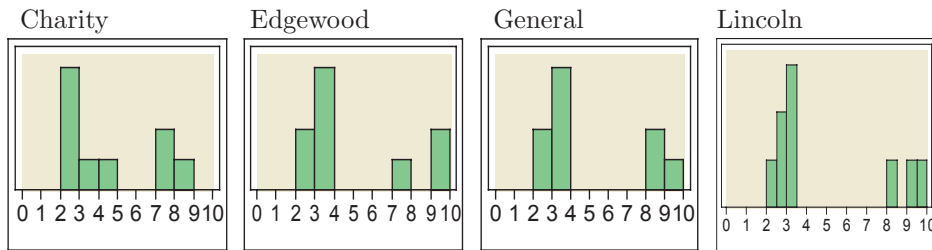

Figure 3: Histogram: Distribution of length of stay for women

Table 4: Number of days spent by women after giving birth by hospital and type of birth

| | Hospitals | | | |
Type of Birth	Edgewood	Lincoln	Charity	General
Caesarian	9.5	9.7	7.1	8.7
	7.9	9.4	8.4	9.0
	9.1	8.4	7.9	8.9
Natural	2.7	2.6	2.5	2.1
	3.2	2.7	2.0	3.2
	2.0	2.4	2.3	2.5
Medically Assisted	3.4	3.4	4.0	3.4
	3.5	3.4	2.9	3.6
	3.5	3.0	3.0	3.8

Based on Figure 3 normal distributions for the populations may be questionable. There would seem to be at least two processes involved in the distributions. That is, other, possibly significant, factors appear to be present.

In this example, the factor that was left out was the type of birth. Fortunately, when the data were collected three observations each from four different hospitals and three different types of birth (Caesarian, Natural and Medically Assisted) were obtained. The data are reorganized in Table 4.

Including type of birth as a factor results in the following model as stated in Equation (4.4) and repeated below:

$$y_{ijk} = \mu + \alpha_j + \beta_i + \alpha\beta_{ij} + \epsilon_{ijk},$$

where α_j = Hospitals and β_i = Type of birth. The ANOVA results for the two-factor design are:

Analysis of variance

Source	DF	SS	MS	F	p
Birth	2	265.071	132.535	560.67	0.000
Hospital	3	2.010	0.670	2.83	0.060
Interaction	6	1.669	0.278	1.18	0.351
Error	24	5.673	0.236		
Total	35	274.423			

With type of birth included as a factor, there are no significantly small F ratios. Including type of birth as a factor resulted in the difference between hospitals becoming significant at a .10 level of significance (p-value = .06). In this situation the original model was incorrectly specified. Including type of birth and interaction in the model resulted in a significant reduction in the MSE.

4.3 Non-linearity or lack of fit

In this case the model in question is a regression model. Suppose that the appropriate model is actually as stated in Equation(4.5) or Equation (4.6).

$$y_i = \alpha + \beta_1 x_i + \beta_2 x i^2 + \epsilon_i \tag{4.5}$$
$$y_i = \alpha + \beta_1 x_i + \beta_2 x i^2 + \beta_3 x_i^3 + \epsilon_i \tag{4.6}$$

If a straight-line model is fitted by mistake then the residual sum of squares will include both the error sum of squares and the squared distances between corresponding points on the straight line and on the correct model. Again, the term

used for the MSE becomes larger than it should and may result in significantly small F-values. If the experimenter suspects that more terms should be included in the model and builds in replication at some of the x-values then a formal test for lack of fit can be done. Without replication, a formal test is not possible, though one may make a subjective evaluation based on a scatterplot of the data. A significantly small F-ratio should lead one to consider the possibility of a lack of fit.

Example 3: A company manufacturing VHS movie tapes is interested in determining the forecasted demand for the tapes. The analyst uses straight-line linear regression to predict demand. The data are given in Table 5:

Table 5: Demand for VHS tapes by year

Time period (years)	Sales of VHS tapes in millions of units
1995 (1)	3.1
1996 (2)	4.6
1997 (3)	5.2
1998 (4)	6.7
1999 (5)	7.1
2000 (6)	7.6
2001 (7)	5.8
2002 (8)	5.3
2003 (9)	4.5
2004 (10)	3.3

Analyzing these data using simple linear regression results in the following output with

$$Y = \text{Sale of VHS tapes (in millions)},$$
$$X = \text{time period (years)}.$$

The regression equation is

$$Y = 5.34 - 0.004x.$$

Predictor	Coef	StDev	t	p
Constant	5.340	1.099	4.86	0.001
Time (X)	-0.0036	0.1772	-0.02	0.984

$$S = 1.609, R^2 = 0.0\%, R^2(\text{adjusted} = 0.0\%$$

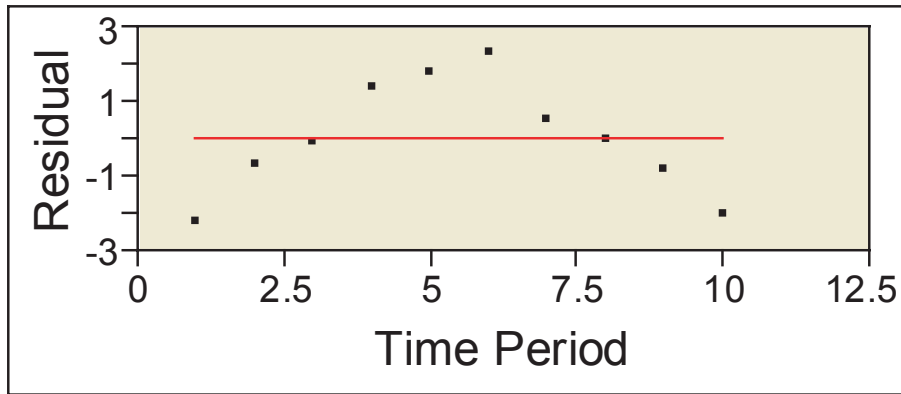| Analysis of variance					
Source	DF	SS	MS	F	p
Regression	1	0.001	0.001	0.00	.984
Residual Error	8	20.715	2.589		
Total	9	20.716			

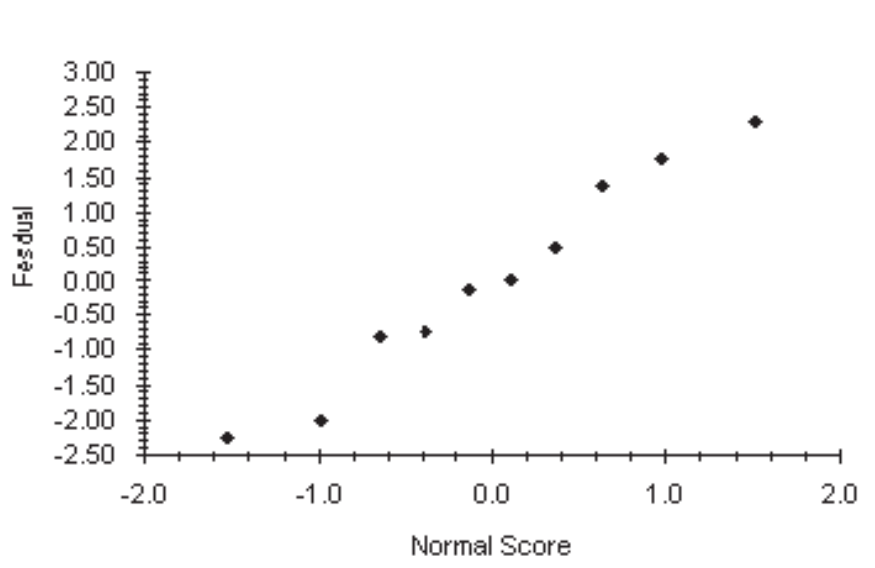Figure 4: Graph of the residuals for the VHS tapes sales data

Figure 5: Normal probability plot of the VHS tape sales data

Note that, the F value is .00 and the p-value is very large, .984, giving $1-p=$ .016. Lack of a significantly large F-value indicates a poor fit of the simple linear regression model, while the extremely small value for F suggests something other than chance is present. A graph of the residuals shown in Figure 4 indicates lack of fit. A graph of the residuals or a scatter plot may indicate lack of fit but this is not a formal test.

The normal probability plot, shown in Figure 5, gives no indication of non-normality. There might be possible positive autocorrelation, however, we could not employ the Durbin-Watson statistic to check for autocorrelation due to too small of a sample size.

Fitting a quadratic model based on equation (4.1), where

$$y_i = \alpha + \beta_1 x_i + \beta_2 x_i^2 + \epsilon_i$$

gave the following results.

The regression equation is:

$$\text{Demand} = 1.16 + 2.09X - 0.190X^2$$

Predictor	Coef	Stand Deviation	t	p
Constant	1.1567	0.5665	2.04	0.080
Time	2.0880	0.2366	8.83	0.000
Time^2	-0.19015	0.02096	-9.07	0.000

$$S = 0.4816, R^2 = 92.2\%, R^2(\text{adj}) = 89.9\%$$

Analysis of variance

Source	DF	SS	MS	F	p
Regression	2	19.0923	9.5462	41.15	0.000
Residual Error	7	1.6237	0.2320		
Total	9	20.7160			

The model has a highly significant F-ratio (41.15) with a p-value of 0.000. In addition, looking at the t-values for individual terms it can be seen that both the linear and quadratic terms are highly significant. By itself the linear term had an unusually small F-value. When a cubic term is added to the model, its coefficient is insignificant and the added term actually results in an increase in the MSE; i.e., a loss of precision in predictive accuracy.

The situations cited above only exemplify three of the possible causes that might give rise to significantly small F-ratios. The violation of other assumptions

and/or falsification of data may also lead to inflated mean square errors, resulting in unusually small F-ratios and, hence a red flag that something may be wrong. The point being made in this paper is that any unusual value is suspect and warrants investigation.

5. Summary

Significantly small values for F-ratios appear to have been ignored in the literature with respect to the signals that they may provide regarding the validity of underlying assumptions for the test procedures used in evaluating linear models. If the model is correct and all assumptions are satisfied then the ratio of the two mean squares should be either near 1.0 or greater than 1.0. The value is never expected to be close to 0.0. If the value is near 0.0 and is significant it should be treated as a red flag, indicating potential problems with the design or analysis, and investigated just as any unusual occurrence in statistical quality control demands explanation. Possible causes for values near zero have been shown to be non-additivity, an omitted factor(s) in the model and/or lack of fit. Other possible causes that were not discussed in the paper are violations of distributional assumptions, multi-collinearity in regression and falsification of data. Unusually small values may be simply chance occurrences but all other possibilities should be eliminated before that conclusion is reached.

References

Bowerman, B. and O'Connell, R. (1990). *Linear Statistical Models*, 2nd ed. PWS-Kent.

Cochran, W. and Cox, G. (1957). *Experimental Designs*, 2nd ed. J. Wiley & Sons.

Davies, O. L. (1963). *Design and Analysis of Industrial Experiments*, 2nd ed. Oliver & Boyd.

Fisher, R. A. (1947). *The Design of Experiments*, 4th ed. Oliver & Boyd.

Gitlow, H., Oppenheim, A. and Oppenheim, R. (1995). *Quality Management: Tools and Methods for the Improvement*. Irwin Pub Company.

Guenther, W. C. (1964). *Analysis of Variance*. Prentice Hall.

Hinkelman, K. and Kempthorne, O. (1994). *Design and Analysis of Experiments*, Vol. 1. J.Wiley & Sons.

Meek, G., Ozgur, C., and Dunning, K. (2005). Some implications of significantly small F-ratios. *Proceedings of the 2005 Annual National Meeting of the Decision Sciences Institute*, 12341-12346, San Francisco, CA.

Meek, G. and Turner, S. (1983). *Statistical Analysis for Business Decisions*. Houghton Mifflin.

Meek, G., Taylor, H., Dunning, K. and Klafehn, K. (1987). *Business Statistics.* Allyn & Bacon.

Montgomery, D. (1996). *Introduction to Statistical Quality Control*, 3rd ed. John Wiley & Sons.

Montgomery, D. (1997). *Design and Analysis of Experiments*, 4th ed. J. Wiley & Sons.

Neter, J., Kutner, M., Nachtsheim, C. and Wasserman, W. (1996). *Applied Linear Statistical Models*, 4th. ed. McGraw-Hill.

Scheffe, H. (1959). *The Analysis of Variance.* J. Wiley & Sons.

Taguchi, G. (1986). *Introduction to Quality Engineering.* Kraus International Publications.

Tukey, J. W. (1940). One degree of freedom for non-additivity. *Journal of Educational Psychology* **31**, 253-269.

Winer, B. (1962). *Statistical Principles in Experimental Design.* McGraw-Hill.

Gary E. Meek (retired)
College of Business & Technology
Black Hills State University
Spearfish, SD 57799-9007, USA
meek@mato.com

Ceyhun Ozgur
College of Business Administration
Valparaiso University
Valparaiso, IN 46383, USA
ceyhun.ozgur@valpo.edu

Kenneth A. Dunning
College of Business Administration
The University of Akron
Akron, OH 44325, USA
dunning@uakron.edu