USE OF REPLICATION IN ALMOST UNREPLICATED FACTORIALS

Kinley Larntz, University of Minnesota; Pat Whitcomb, Stat-Ease, Inc.
Kinley Larntz, 420 Old Farm Road, Shoreview, MN 55126

Key Words: Fractional Factorials, Half Normal Plots, Lenth's Method

True replication in a designed experiment permits calculation of a pure error mean square that is used to check model lack-of-fit. In 2-level factorial and fractional factorial designs, replication is usually done for center points, although other replication patterns may be useful. The number of replication points are often small (hence the term, "almost unreplicated factorial") and the usual analysis to determine important effects typically ignores these points. For instance, half-normal and normal probability plots display only the effect estimates, without any graphical representation to account for replicate observations. A valuable contribution of Lenth (1989) provides a numerical method for selecting important effects from unreplicated factorials. As the purpose of Lenth's paper was to provide a method for analyzing unreplicated factorials, he made no provision to take advantage of replication.

In this paper, we propose supplementing Lenth's method by combining his estimate of error variance from small factorial effects with the pure error variance from replicate observations or (possibly) other estimates of error variance. We demonstrate that there can be a substantial increase in power to detect effects by incorporating this additional information. In fractional factorials, this increase in power is particularly valuable.

We also propose augmenting half-normal or full normal probability plots of effects by adding points that represent error effects estimated from the replicate observations and (possibly) other estimates of the error variance. As these points are truly error, they provide a set of points on the plots to judge the potential importance of all effects. The number of extra points added equals the number of degrees of freedom associated with the within sum of squares for the replicate observations.

Section 1 reviews Lenth's method. In Section 2 we extend Lenth's method to take advantage of error variance information from replicate observations. Section 3 presents an example from Montgomery (1997) that illustrates how inferences may change based on more complete information. Simulation results are given in Section 4 that quantify the gain in power from using the addition information. Section 5 presents our extension of the half-normal and full-normal probability plots to include information from replicate observations. The final section summarizes our findings.

1. Lenth's Method

Lenth (1989) gave a clever numerical method for analyzing unreplicated factorial designs. As he noted, the Daniel's (1959) graphical method using half-normal and full-normal probability plotting depends on *effect sparsity*. Effect sparsity means that the number of nonzero or active effects among the contrasts in an unreplicated factorial or fractional factorial experiment is expected to be small. Assuming there are few active effects, the active effects show up as outliers on the probability plot. Of course, what exactly constitutes an "outlier" is somewhat subjective. (It is important to note here that interpretation of what effects to consider as "real" depends on understanding the substantive aspects of the problem, i.e., as with all analyses subject matter knowledge and expectation must be incorporated into any data analysis!)

We present Lenth's method using his notation. Suppose we have m contrasts of interest. Let $\kappa_1, \kappa_2, \ldots, \kappa_m$ denote the true contrast values. Let $c_1, c_2, \ldots, c_m$ denote the corresponding contrast estimates. The key to the probability plots is that the contrast estimates are independent normal variates having the same variance, but possibly unequal means, $\kappa_1, \kappa_2, \ldots, \kappa_m$. (We note that unequal variances can be easily handled by standardizing the contrast estimates and also that the plots work well even if the contrast estimates depart moderately from independence.) Note that effect-sparsity means that "most" of the κ_i are in fact zero.

Lenth denotes the common variance of the contrasts as τ^2. The essence of his method is to estimate τ^2 from the smallest (in absolute value) contrast estimates. Let

$$s_0 = 1.5 \times \text{median}\ \{ |c_j| \}.$$

and

$$\text{PSE} = 1.5 \times \text{median}\ \{ |c_j| : |c_j| < 2.5\, s_0 \}.$$

The PSE is 1.5 times the median of all the smaller contrasts, where smaller contrasts are defined as those with absolute values less than 2.5 times 1.5 times the median of the absolute values of all contrasts. Lenth shows that PSE is a "fairly good estimate of τ when the effects are sparse."

Once Lenth has his estimate of the standard error for a contrast, he proceeds to form individual and simultaneous confidence bounds for them. The *margin of error* is defined using t confidence limits with reduced degrees of freedom.

$$\text{ME} = t_{.975,d} \times \text{PSE}.$$

The reduced degrees of freedom is taken to $d = m/3$. For simultaneous confidence limits, Lenth suggests using *simultaneous margin of error*.

$$\text{SME} = t_{\gamma,d} \times \text{PSE}.$$

The percent point of the t distribution used is

$$\gamma = (1 + .95^{1/m}) / 2,$$

which is exact since the contrast estimates are independent.

Lenth suggests constructing a bar plot of contrast estimates with reference lines at ± ME and at ± SME. He says, "A contrast whose bar extends beyond the SME lines is clearly active, one which does not extend beyond the ME lines cannot be deemed active, and one in between is in a zone of uncertainty where a good argument can be made both for its being active and for its being a happenstance result of an inactive contrast." We would add that estimated contrasts in this middle zone require substantive knowledge of the problem to evaluate appropriately.

2. Extension of Lenth's Method

Now we consider the case with replication. For instance, in a typical 2-level fractional factorial it is good practice to replicate some point or points. Often the center point is included in the design to provide an estimate of curvature. The center point is frequently replicated several times to provide a more stable estimate of the curvature effect and to provide a estimate of pure error. There additionally may be other estimates of error that are known to estimate σ^2, the error variance. In this section, we use this additional information to extend Lenth's method to make use of this information.

For the general case, assume we have an independent estimate of σ^2, which we will denote as s^2. Under the usual normal theory assumptions,

$$\nu s^2 / \sigma^2 \sim \chi^2(\nu),$$

that is s^2 follows the usual mean square distribution with ν degrees of freedom. If the center point of a fractional factorial is replicated 5 times, then the mean square for pure error from that replication provides an independent estimate of σ^2 with 4 degrees of freedom. The sum-of-squares is νs^2 and the degrees of freedom is $\nu = 4$.

Now, the variance of an estimated contrast is a multiple (often fractional multiple) of σ^2. For instance, in an 8-run fractional factorial, the variance of the estimated A-effect (defined as the difference between the average response at the high level of A and the average response at the low level of A) is $\sigma^2 / 2$. In a 32-run design, the variance of the estimated A-effect is $\sigma^2 / 8$, etc. For generality, we assume the common variance of estimated contrasts is K σ^2.

To provide a combined contrast standard error, we pool Lenth's PSE with the independent estimate K s^2 using the usual method,

$$\text{CPSE} = \{ (d\,\text{PSE}^2 + \nu\,\text{K}\,s^2) / (d + \nu) \}^{0.5}.$$

We then propose using a *combined margin of error* defined by

$$\text{CME} = t_{.975,df} \times \text{CPSE}.$$

The combined degrees of freedom is taken to be

$$df = d + \nu.$$

Similarly our proposed *combined simultaneous margin of error* is given by

$$\text{CSME} = t_{\gamma,df} \times \text{CPSE}.$$

We propose using CME and CSME in place of ME and SME, respectively, whenever additional information on the error variance is available from any source.

3. Example from Montgomery (1997)

As an example, we take data from Example 7-2 and Problem 7-29 in Montgomery (1997). The data come from a 4-factor 2-level full factorial that is unreplicated in Example 7-2 and is augmented by 5 center points in Problem 7-29. For our purposes here, we will simply call the factors A, B, C, and D, although for interpretation purposes we ought to understand as completely as possible the substance of the data. The 21 data points are given in Table 1.

Run	A	B	C	D	y-obs
1	-	-	-	-	45
2	+	-	-	-	71
3	-	+	-	-	48
4	+	+	-	-	65
5	-	-	+	-	68
6	+	-	+	-	60
7	-	+	+	-	80
8	+	+	+	-	65
9	-	-	-	+	43
10	+	-	-	+	100
11	-	+	-	+	45
12	+	+	-	+	104
13	-	-	+	+	75
14	+	-	+	+	86
15	-	+	+	+	70
16	+	+	+	+	96
17	0	0	0	0	73
18	0	0	0	0	75
19	0	0	0	0	71
20	0	0	0	0	69
21	0	0	0	0	76

Table 1: Data from Montgomery (1997)

Lenth's method applies nicely to this example, as do other methods as this example is as clear cut as it gets. To illustrate Lenth's method it is easiest to look at the effects list, ordered by absolute value. The values are given in Table 2.

Term	Estimated Contrast
A	21.625
AC	-18.125
AD	16.625
D	14.625
C	9.875
ABD	4.125
B	3.125
BCD	-2.625
BC	2.375
ABC	1.875
ACD	-1.625
ABCD	1.375
CD	-1.125
BD	-0.375
AB	0.125

Table 2: Ordered Estimated Contrasts

The calculations for Lenth's method yield

$$s_0 = 1.5 \times |-2.625| = 3.9375$$

$$2.5 \times 3.9375 = 9.84375$$

$$\text{PSE} = 1.5 \times |1.75| = 2.625$$

$$\text{ME} = 2.571 \times 2.625 = 6.75$$

$$\text{SME} = 5.219 \times 2.625 = 13.70$$

So the first four contrasts in Table 2 would be judged significant by the SME standard and the fifth contrast C would be taken under consideration for significance. (Here C is clearly important as AC is the second most significant contrast, so it may not be important to clarify the significance of the C contrast.)

If we use the additional information from the replication of the center points, we have 4 degrees of freedom and $s^2 = 8.20$. The variance of a contrast for a 16-run design is $\sigma^2 / 4$, so the combined pseudo standard error is

$$CPSE = \{ (5 \times 2.625^2 + 4 \times 8.20 / 4) / 9 \}^{0.5} = 2.177.$$

the *combined margin of error* is

$$CME = 2.262 \times 2.177 = 4.92,$$

and the *combined simultaneous margin of error* is

$$CSME = 3.938 \times 2.177 = 8.57.$$

Using the additional information lowers the two margins of error considerably. Note that the gain comes from the relatively smaller estimate of error from the pure error component and additionally the smaller t multipliers resulting from having 9 degrees of freedom instead of 5 degrees of freedom. It is interesting in this case that all the first five contrast are clearly significant using the CSME criterion. Indeed, there is no ambiguity as there are no contrast in the zone between the CME and the CSME. Of course, there well could be in less clear cut problems.

4. Simulation Results

To understand the gain from using additional information, we carried out a simulation study comparing Lenth's original method to the extended method. Recent work, Loughin and Noble (1997) for example, have noted that Lenth's method fails to control Type I error rates and that simulation may be used to calibrate the limits for his method. For 7, 15, and 31 effect designs, we have found adjusted ME and SME multipliers to use in place of the t two-sided 0.95 or simultaneous 0.95 critical values proposed by Lenth. Table 3 gives the original and adjusted values for Lenth's method.

	7	15	31
original ME	3.764	2.571	2.218
adjusted ME	2.295	2.140	2.082
original SME	9.008	5.219	4.218
adjusted SME	4.891	4.163	4.030

Table 3: Lenth Adjusted Critical Values

An additional simulation was run to illustrate the gain in power from adding additional information concerning the size of the error variance. Table 4 gives a comparison of the power of Lenth's method (using adjusted critical values) and our extended method from Section 2. The data for the simulation were generated from random normal variates for each contrast with 0, 1, 2, or 3 real effects corresponding to a contrast standardized expected value of 2.0 standard units with the remaining contrasts having expected value of 0.0. Data for the extended method added 5 degrees of freedom for additional information on the variance σ^2.

		Power ME	Power SME
0/7	Lenth	5.0	5.0
	Extended	5.0	5.0
1/7	Lenth	29.2	7.4
	Extended	45.1	19.7
2/7	Lenth	21.9	7.8
	Extended	36.6	21.2
3/7	Lenth	13.5	6.6
	Extended	29.8	19.4
0/15	Lenth	5.0	5.0
	Extended	5.0	5.0
1/15	Lenth	41.0	9.8
	Extended	47.0	14.6
2/15	Lenth	35.2	11.8
	Extended	42.5	18.1
3/15	Lenth	30.1	11.2
	Extended	38.9	18.9
0/31	Lenth	5.0	5.0
	Extended	5.0	5.0
1/31	Lenth	44.1	9.1
	Extended	46.7	11.5
2/31	Lenth	43.4	11.8
	Extended	45.9	15.0
3/31	Lenth	40.4	13.4
	Extended	44.2	18.3

Table 4: Power Comparison of Two Methods

5. Augmented Probability Plots

The half-normal and full-normal probability plots are undoubtedly the most used method of analysis for 2-level factorial and fractional factorial designs. We agree that subjective interpretation of these plots, combined with substantive information, gives the practitioner the greatest ability to derive information from studies using these design. One must, however, be extremely careful as was pointed out by Daniel (1976). Daniel included in his book pages of NULL probability plots to illustrate the variety of "interpretable" plots that arise when there are, in fact, no real effects.

Figure 1 presents the usual half-normal probability plot for the data from Montgomery used to illustrate Lenth's method and the extended method in Section 3. The usual plot includes only information from the factorial points, omitting any information from the replicated center points.

We propose augmenting the usual half-normal or full-normal plot by adding a point to the plot for each degree of freedom of additional information concerning the error variance. For instance, we add points that correspond to expected normal order statistics (absolute values for the half-normal plot) scaled to have the same standard error as the effect contrasts. We also account for the ordered position of the additional points among the estimated contrasts from the factorial points. This requires an iterative calculation to ensure that the additional points are not associated with higher percentiles than they should be. Figure 2 presents the augmented half-normal plot for the data from Montgomery used in Section 3. The triangles on the plot represent the pure error augmented points. Note that they are intertwined with the "null" estimated contrast effects on the plot. Note also that they lie on a straight line with slope directly related to the estimated contrast standard error as estimated by the "pure error" information.

To illustrate the possible advantage of the augmented plot, we created an artificial data set from a 2^3 factorial with 5 center points. The data are given in Table 5.

A	B	C	y-obs
-	-	-	-3.3
+	-	-	-9.2
-	+	-	-3.6
+	+	-	2.6
-	-	+	-1.4
+	-	+	3.0
-	+	+	-3.4
+	+	+	15.3
0	0	0	-1.1
0	0	0	0.6
0	0	0	0.6
0	0	0	1.3
0	0	0	1.0

Table 5: Artificial Data for Figures 3 & 4

Figure 3 gives the usual 7 point half normal probability plot. There are 5 large contrasts and 2 small ones. It is difficult to distinguish this from a null plot that happened to wind up with 2 small contrasts, something that can easily happen since the sampling distribution of absolute estimated half-normal null contrasts has greatest density at 0.

Figure 4 augments the usual plot with 4 points for pure error. We are much more comfortable now moving the line over to go through the 2 "null" contrast values and the 4 "pure error" augmented data values. The additional points clarify the situation for us in this situation which probably would not be classified as "effects-sparse."

Figure 1

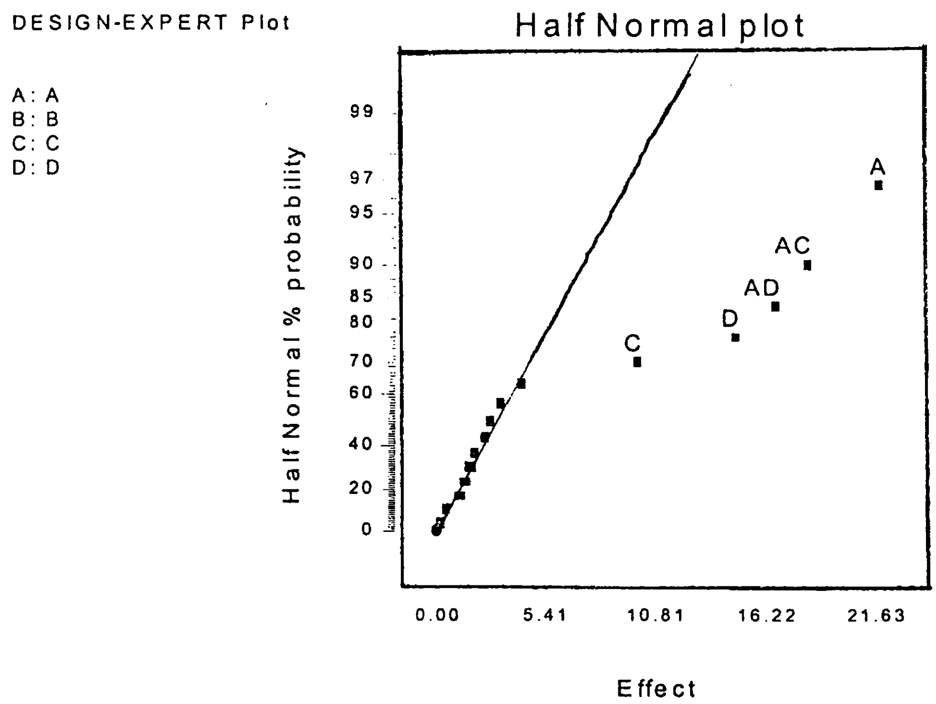


Figure 2

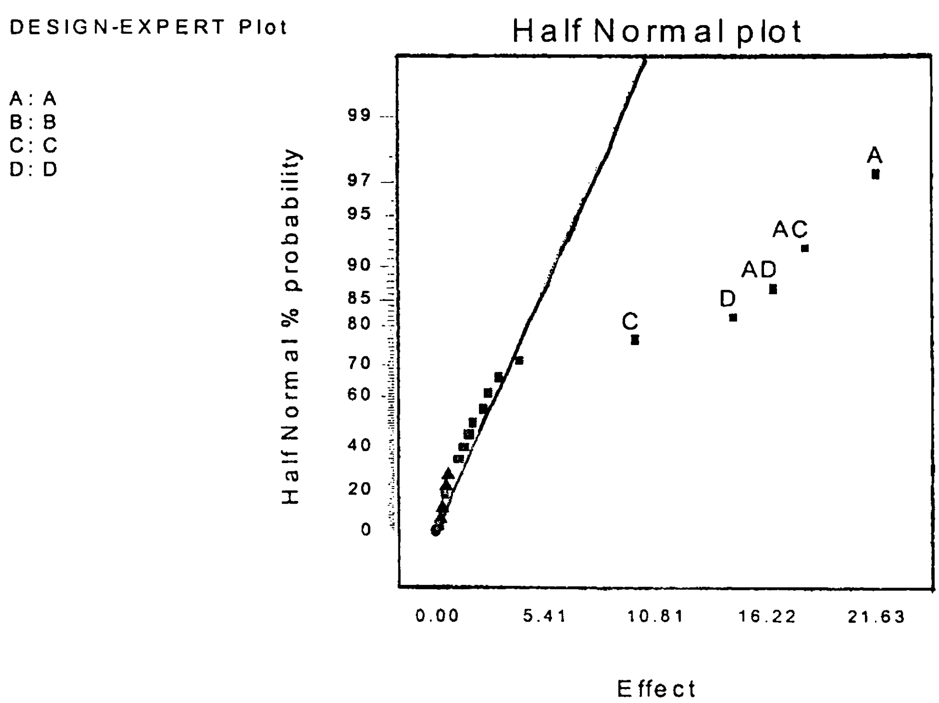


Figure 3

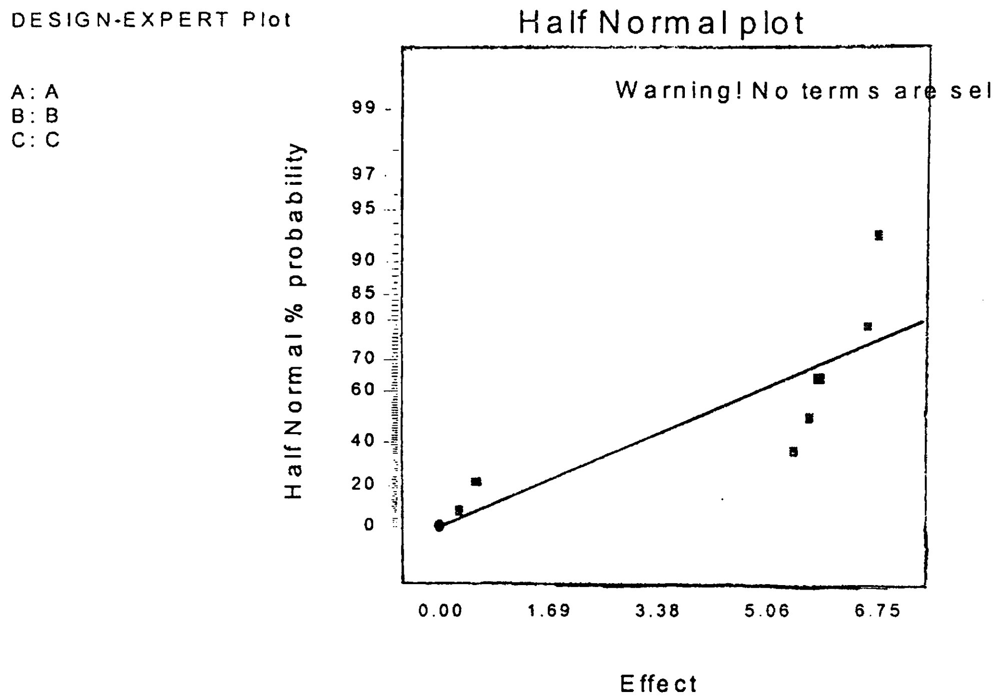
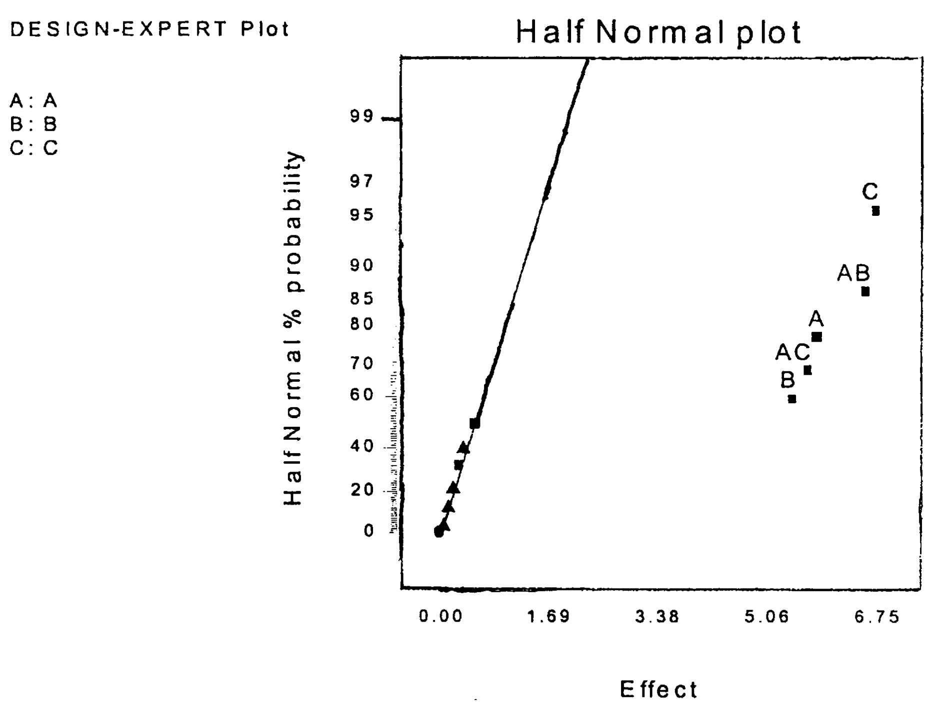


Figure 4



6. Summary and Conclusions

In this paper we have presented material that permits two techniques commonly employed in the analysis of factorials and fractional factorials to take advantage of extra information on the size of the error variance. We have seen that (a properly calibrated) Lenth's method can be modified to have increased power using the additional information. The increase in power is greatest in small designs, as expected. The same increase in clarity can be seen in the augmented probability plots. We believe that these augmented plots are quite useful in distinguishing null and real effects, particularly again for small designs and for designs in which the effect-sparsity assumption may not be valid.

One final note that, perhaps, should have been a reminder up front. We have assumed that the extra variance information is "pure error" in the sense that it was generated from actual replicate values. Also, it is assumed that the error variance is approximately constant over the design space. If "replicate" values are merely replicate measurements and do not represent "true replication," then these tools are totally inappropriate. It is an important question for a data analyst to ask: "How did you get these repeat observations?" The answer frequently makes these methods invalid. Even so, when we have true replication, which we submit you always should, these tools enable the data analyst to make full use of the data.

7. References

Daniel, C. (1959), "Use of Half-Normal Plots in Interpreting Factorial Two-Level Experiments," *Technometrics*, 1, 311-341.

Daniel, C. (1976), *Applications of Statistics to Industrial Experimentation*, New York: John Wiley.

Lenth, R. V. (1989), "Quick and Easy Analysis of Unreplicated Factorials," *Technometrics*, 31, 469-473.

Loughin, T. M. and Noble, W. (1997), "A Permutation Test for Effects in an Unreplicated Factorial Design," *Technometrics*, 39, 180-190.

Montgomery, D. C. (1997), *Design and Analysis of Experiments*, 4th edition, New York: John Wiley.

Kinley Larntz
University of Minnesota
420 Old Farm Road
Shoreview, MN 55126
kinley@stat.umn.edu
(651) 482-0653 voice
(612) 397-7705 fax

Pat Whitcomb
Stat-Ease, Inc.
2021 E. Hennepin Ave.
Suite 480
Minneapolis, MN 55413
pat@statease.com
(612) 378-9449 voice
(612) 378-2152 fax