stateaser

ABOUT STAT-EASE® SOFTWARE, TRAINING, AND CONSULTING FOR DOE
Phone 612.378.9449 Fax 612.746.2069 E-mail info@statease.com Web Site www.statease.com

Workshop Schedule

Experiment Design Made Easy

April 28-30, 2009: Minneapolis, MN
June 9-11, 2009: Minneapolis, MN
August 18-20, 2009: Minneapolis, MN

Study the practical aspects of design of experiments (DOE). Learn about simple, but powerful, two-level factorial designs. $1595 ($1295 each, 3 or more)

Response Surface Methods for Process Optimization

July 7-9, 2009: Minneapolis, MN

Maximize profitability by discovering optimal process settings via RSM. $1595 ($1295 each, 3 or more)

Mixture Design for Optimal Formulations

April 21-23, 2009: Minneapolis, MN
June 23-25, 2009: Edison, NJ

Find the ideal recipes for your mixtures with high-powered statistical tools. $1595 ($1295 each, 3 or more)

DOE for DFSS: Variation by Design

May 5–6, 2009: Minneapolis, MN

Use DOE to create products and processes robust to varying conditions, and tolerance analysis to assure your specifications are met. A must for Design for Six Sigma (DFSS). $1195 ($995 each, 3 or more)

Designed Experiments for Life Sciences

July 28-29, 2009: Minneapolis, MN

Learn how to apply DOE to Life Science problems. $1495 ($1195 each, 3 or more)

PreDOE: Basic Statistics for Experimenters (Web-Based)

PreDOE is an entry-level course for those who need to go back to the basics. See *http://www.statease.com/clas_pre.html* for more information. $95

Attendance is limited to 16. Contact Elicia at 612.746.2038 or workshops@statease.com.

Stat-Ease
statistics made easy®

Science Fair Project Provides Education on Detecting Outliers

A reader of my monthly DOE FAQ Alert ezine (www.statease.com/doealert.html) recently sent me a very nicely-done report from her daughter Ashley on a gas mileage study she did for her 8th grade science fair project. Via a full two-level design, Ashley and her mom investigated three ways to save fuel:

A. Air conditioning (A/C) off (-) or on (+)
B. Tire pressure, low or high (29 to 34 psi)
C. Octane, regular or premium (87 to 91)

Ashley hypothesized that gas mileage would be most affected by tire pressure. I thought that was a good guess. However as you can see from the half-normal plot of effects (Figure 1 below), the air conditioner (factor A) appears to be the only factor that made a significant difference—causing a loss of 5 mpg in fuel efficiency. This seemed to be a great example of why one should never assume anything that can be properly tested.

However, Ashley still felt sure that tire pressure would prove to be important. She wondered if being bolder might be better—going to 27 psi or below. Maybe that would make this factor significant. But, there was still something nagging at me—the abnormal appearance of the half-normal plot of effects. Notice in Figure 1 how the trivial effects do not line up on zero. Mysteriously, nothing appeared markedly abnormal in the diagnostics plots.

Are you up for a challenge?

I asked our "top gun" Pat Whitcomb to investigate. He reported back to me that one run appeared to be statistically deviant. Before I tell you which one and how Pat discovered it, take a look at the data in Table 1. Go ahead and enter it into your DOE software if you like—it should not be hard with only 8 runs from the full 2^3 factorial design done by Ashley. To make it even easier, I sorted the combinations in standard order. <u>Do not turn the page until you either try finding the outlier yourself, or least make an educated guess by looking at the factor levels on each run.</u>

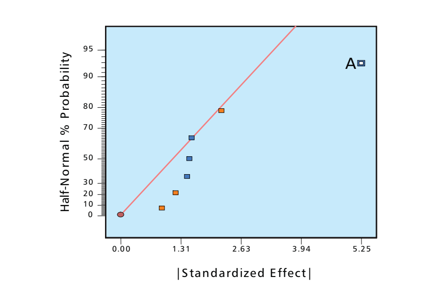

Fig. 1: Half-normal plot of effects

Std	Run	A: A/C	B: Tire psi	C: Octane	Gas mileage
			psi		mpg
1	2	Off	29	87	31.6
2	1	On	29	87	25.2
3	4	Off	34	87	35.4
4	5	On	34	87	28.9
5	7	Off	29	91	31.4
6	6	On	29	91	30.3
7	3	Off	34	91	35.0
8	8	On	34	91	28.0

Table 1: Experimental data

—Continued on page 2

—*Continued from page 1*

Most of you own a car and thus you have acquired some subject matter knowledge along the way. Which of the miles per gallon (mpg) appears out of place?

Here's how you Stat-Ease software users of Design-Ease® and Design-Expert® can re-discover what Pat found. With all the data intact (none ignored) pick tire pressure (B) off the half-normal plot in addition to the obvious one of A/C (factor A). It (B) is the second largest effect. Then proceed through to the Diagnostics and click the Influence side of the floating Diagnostics Tool. Notice the outlier on the Externally Studentized Residuals. Right click this obviously deviant point and select the Highlight option. Then click back to the design layout and see it highlighted in blue italic. To temporarily take it out, right click the blank button at the left of this run and Set Row Status to Ignore. (This can also be done to the response value only, which comes in very handy when there are many responses—some that may be worth keeping.) Then re-analyze. You'll notice that things make a lot more sense. Now go ahead and look over the answer from Pat.

THE ANSWER:

Run 6 (A/C on, 29 psi tire, 91 octane) emerges statistically deviant on the high side when the main effect of tire pressure (B) is modeled (in addition to the air conditioning—factor A). See Figure 2a for the evidence—an externally studentized residual (once known as "outlier t") falling far above the 95 percent confidence limit. (Compare this to Figure 2b which models factor A only—no outlier!) With this run ignored, the half-normal plot of effects (shown in Figure 3) looks normal—the trivial many effects line up on zero. The main effect of tire pressure (factor B) stands out as a positive effect, which is what Ashley hypothesized. Higher tire pressure <u>does</u> increase the miles per gallon for her mother's car as shown in Figure 4a. However, it does not produce as big of an effect as air conditioning ("A/C"), as you can see in Figure 4b for the main effect of factor A.

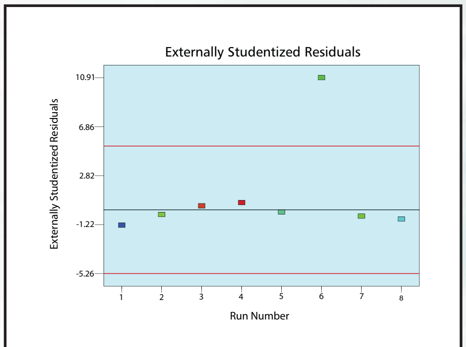

Fig. 2a: Diagnostic plot showing run 6 being an outlier (after modeling tire pressure—Factor B)

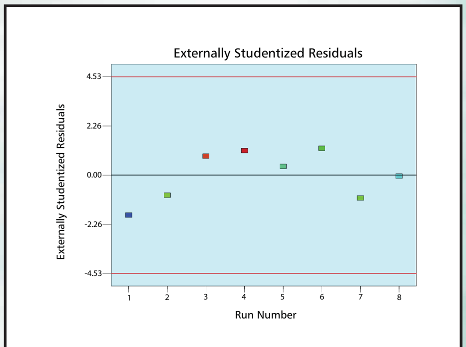

Fig. 2b: Diagnostic plot modeling Factor A only

Here's what Pat recommends as a general procedure if expected effects do not emerge in the half-normal plot:

1) Pick the largest effects if they make sense; for example, they aren't high order interactions.

2) If the largest effects do not make sense, then try picking the main effects only.

This is simple, but good advice for uncovering possibly discrepant response values (outliers). But again (this cannot be stressed enough!), do not ignore runs without a lot of serious thought as to what the cause may have been. Also, always confirm your findings!

What lessons can be learned?

The lesson to be learned here is to recognize when the half-normal plot of effects deviates noticeably from its general pattern of many points lining up near zero (the trivial ones) and a few standing out (the vital effects one hopes to discover). Then, pay especially close attention to the diagnostics plots. Possibly a transformation will be indicated—try this first. But, if no transformation is suggested and nothing appears abnormal, try picking the next biggest effect, particularly if it's a main effect.* Maybe this will cause an outlier to emerge, which when ignored produces a more sensible and statistically-sound model. However, this will remain to be seen in confirmation runs.

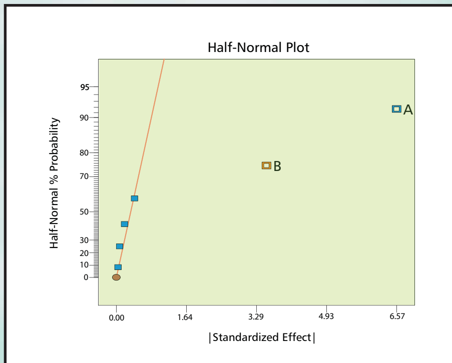

Fig. 3: Effects plot after ignoring run 6

Also, ideally the experimenter can provide a special cause for why a value may be discrepant. The first thing I always do is check my data entry. It is easy to make a mistake in calculated responses like mpg as I've discovered to my own chagrin many times over the years. I think the only thing harder may be balancing the check book!

—Mark Anderson (mark@statease.com)

*(If the second largest effect is a three-factor interaction, then do not pick it—this will create a lot of problems due to the need for model hierarchy. At this

—*Continued from page 2*

Fig. 4a: Main effects after ignoring run 6

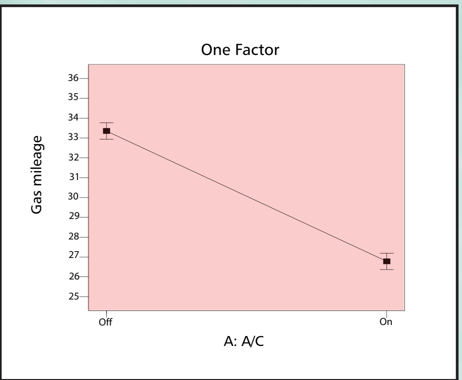

Fig. 4b: Main effects after ignoring run 6

stage, you'd be well advised to email stathelp@statease.com with your data file in Stat-Ease format—".de_" or ".dx_"—attached.)

A Free Primer on Mixture Design

Early in my career as a chemical engineer working in process development, I discovered that many of the chemists in our research and development center—especially those most brilliant in their field—failed to appreciate the power of planned experimentation. Furthermore, they disliked the mathematical aspects of statistical analysis. To top off their design-of-experiments (DOE) phobia, these otherwise very capable chemists also dismissed predictive models based only on empirical data. Ironically, in the hands of subject matter experts like these elite chemists, the statistical methods of mixture design and analysis provide the means for rapidly converging on optimal compositions—the "sweet spot," as I like to call it.

This experience has inspired me to begin work on detailing design and analysis of mixture experiments in a new book that will fill in a major gap remaining after publication of *DOE Simplified* and *RSM Simplified*. This new 'how-to' book, co-authored by Patrick Whitcomb—the same as the others—is based on the Stat-Ease workshop "Mixture Design for Optimal Formulation" developed primarily by him. Readers of our third book will learn how to handle many ingredients simultaneously—not just one at a time, as dictated for time immemorial by "the scientific method." As they will see in example after example, one needn't hold all else constant while changing only one thing. Instead, take advantage of modern-day parallel processing schemes of mixture designs that provide multicomponent testing.

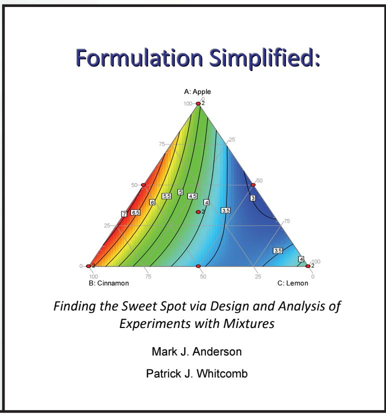

Figure 1: A new book, *Formulation Simplified,* is in the works

After some months of work, I've established the WIIFM, that is, the "what's in it for me" that will encourage formulators to work through the more intense aspects of mixture design, response modeling, statistical analysis and numerical optimization. It's taken me three chapters to do this—it's a tough nut to crack! Pat and I are offering this material to you as *A Primer on Mixture Design: What's In It for Formulators*. Feel free to download it for your education on this powerful tool in the DOE tool kit. All I ask is that if you like what you see and become convinced that this primer will be valuable for your colleagues, please send them the link: www.statease.com/pubs/MIXprimer.pdf. Now I must get back to work on the rest of the story.

—Mark Anderson (mark@statease.com)

Inverse Transformation Puts Mileage Comparisons On Track

National Public Radio's (NPR) "All Things Considered" show on June 19, 2008 led off with this quiz: "Which saves more gas: trading in a 16-mile-a-gallon gas guzzler for a slightly more efficient car that gets 20 mpg? Or going from a gas-sipping sedan of 34-mpg to a hybrid that gets 50 mpg?" What would you guess?

Imagine you and your spouse work at separate locations that require an annual commute of exactly 10,000 miles per year for both driving separately (two automobiles). Then your 16 mpg guzzler consumes 625 gallons (10,000/16). Trading that for a 20 mpg car you need only 500 gallons the next year—a savings of 125 gallons. On the other hand, your spouse drives the far more efficient 34 mpg sedan—it requires only 294 gallons of gas per year (10,000/34). Upgrading this to the 50 mpg hybrid saves only 94 gallons! For more details, search for my 6/30/08 entry on the StatsMadeEasy blog at http://statsmadeeasy.blogspot.com.

—Mark Anderson (mark@statease.com)

Newsletter Subscription Preferences

03/09

Thank you for reading the Stat-Teaser newsletter. In the interest of helping the environment and to make sure we only send the newsletter to those people who are interested, we request that you take a moment to let us know your subscription preferences. Please indicate below whether or not you would like to keep receiving the *Stat-Teaser* newsletter by mail. If you would prefer instead to receive an e-mail link to it via Mark Anderson's monthly e-mail newsletter, the *DOE FAQ Alert* (http://www.statease.com/doealert.html), please check the appropriate box below.

Be sure to include this whole page, with your address and any changes, when you fax (612.746.2069) or mail us your response. If you prefer to e-mail your preference to marketing@statease.com, please include your full name, address, and the client code you'll find in the upper right corner of the address field. Thank you for your interest and your help!

☐ Keep me on the *Stat-Teaser* newsletter mailing list.

☐ Keep me on the *Stat-Teaser* newsletter mailing list. **In addition**, sign me up for Mark Anderson's monthly e-mail newsletter, the *DOE FAQ Alert*. My e-mail address is ______________________________.

☐ Take me off the *Stat-Teaser* newsletter mailing list. **Instead**, send me an e-mail link to the *Stat-Teaser* via Mark Anderson's monthly e-mail newsletter, the *DOE FAQ Alert*. My e-mail address is ______________________________.

☐ Please remove me from the *Stat-Teaser* newsletter mailing list.

☐ Please remove me from your database completely. I am not interested in Stat-Ease products, services, or publications.

Stat-Ease®
statistics made easy®
Stat-Ease, Inc., Hennepin Square
Suite 480, 2021 E. Hennepin Ave.
Minneapolis, MN 55413-2726

Address Service Requested

Presorted
Standard
U.S. POSTAGE PAID
Minneapolis, MN
Permit No. 28684