Real-Life DOE*

Using Graphical Diagnostics to Deal with Bad Experimental Data**

*Slides posted at www.statease.com/webinar.html

Communication issues: To avoid disrupting the audio, we must mute all for these worldwide webinars. It will best if you email questions to my attention via stathelp@statease.com. However, at intervals we will look for any entered live in the GotoWebinar screen. If we cannot answer immediately, please accept our apology in advance. Then bear with us by repeating your question via email afterwards. We would like to hear from you! -- Mark

By Mark J. Anderson, PE, CQE
Stat-Ease, Inc., Minneapolis, MN
mark@statease.com

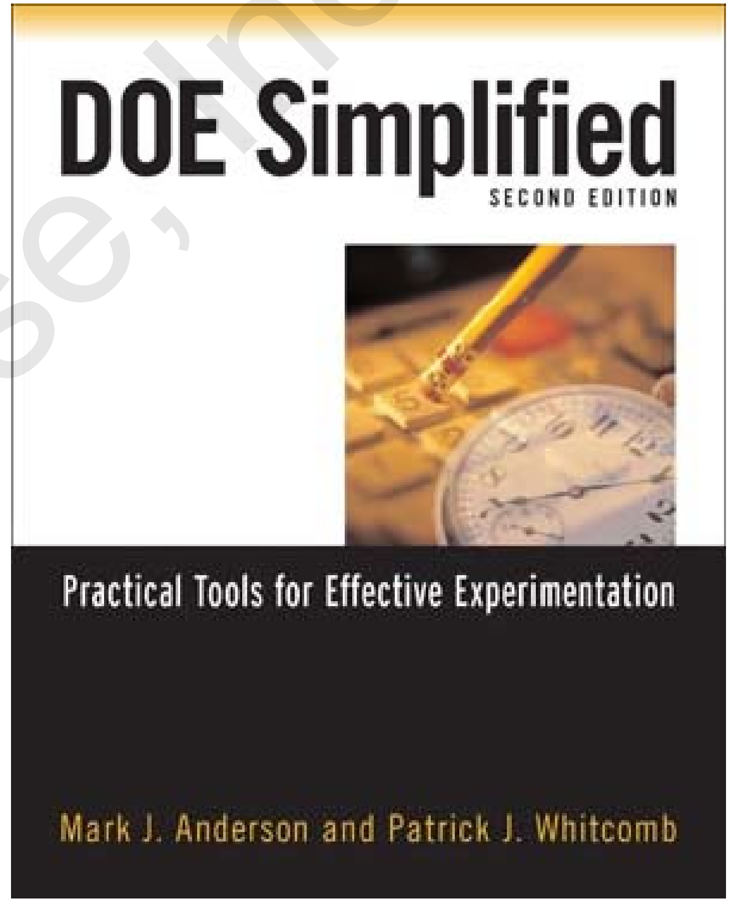
** Re Anderson & Whitcomb, *Quality Engineering*, April-June 2007, Vol 19, No 2

Reference

Now in 2nd edition*!

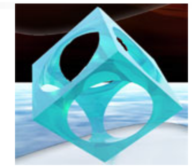
Web-Based 'Launch Pad'
(Chapters 1-3)

Presenter: Demo Sect 1-2 with PIP via
http://statease.info/DOEpresentations/



** Productivity Press (CRC, Taylor & Francis) New York, 2007.*

Mission for this Webinar



Objective:

Keeping things relatively simple* statistically (KISS) provide some tricks of the trade that salvaged outstanding results from actual experiments.

*(Some may disagree due to bits of math & trig. Warning: arc-sine square-root ahead.)

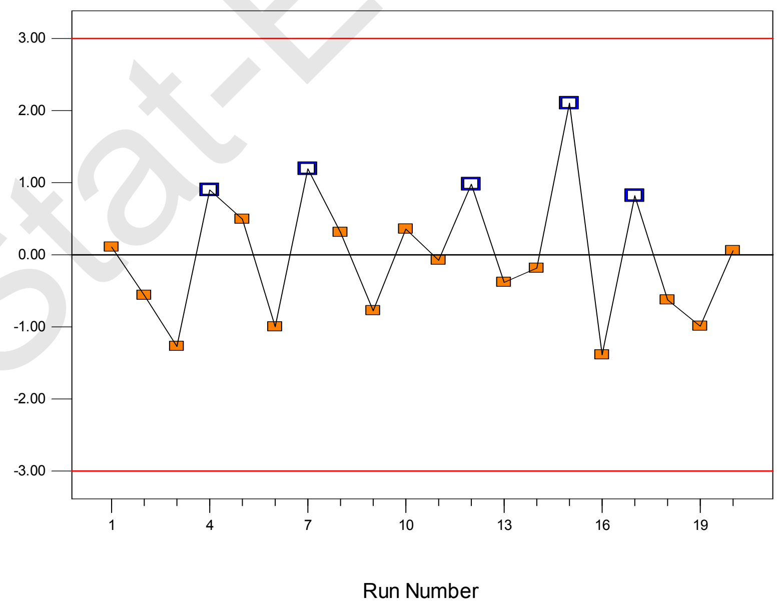
Agenda

- Thorny issues (stuff happens)
- The secret to long life (bearing case)
- A case to test your metal (defects defeated!)
- Conclusions

Thorny Issues! *(1 of 2)*

How to maintain reasonable balance between two types of errors:

1. <u>Focusing</u> on data that vary only due to common causes, thus introducing bias. (False Positive.)
   Examples:
   + cold fusion [keeping only the highs],
   – Pig's Eye treatment [keeping only the lows].*

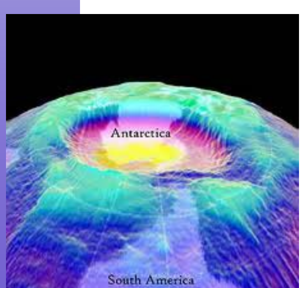


*Source:
James J. Anderson
—Mark's father.

Thorny Issues! *(2 of 2)*

How to maintain reasonable balance between two types of errors:

2. Overlooking true (special cause) outlier(s) (False Negative.) Thus:
   - Obscuring real effects (*case later*) or
   - Drawing false conclusions and/or
   - Losing chance in future to:
     - ✓ Prevent failure (*ex. ozone hole**) or
     - ✓ Reproduce breakthrough improvements ('Eurekas' such as 3M Postit Notes®).



**"When the first measurements were taken the drop in ozone levels in the stratosphere was so dramatic that at first the scientists thought their instruments were faulty."*
- http://www.theozonehole.com/ozoneholehistory.htm

“Stuff” Happens!

- Typographical erors (*suggestion: type data from top, proof from bottom*)
- Breakdowns in equipment
- Mistakes by operators*
- Non-representative samples
- Bad measurements
- Unknown lurking variables that appear only intermittently

Has any stuff like this ever happened to you?

*(Defending champion Tony Stewart triggered a chaotic crash at Talladega Superspeedway collected 10 of the 12 title contenders in an October 2012 NASCAR race.)

It could be worse!

Mark,

Good Afternoon, I appreciate you calling me and discussing the results. We had an accident in the lab and one of the techs caught on fire.

-Myra

Two Types of Error

Stuff does not always happen!

Outliers(s)?		What you say:	
		Yes	*No*
The truth:	Yes	Correct	2.False Negative
	No	1.False Positive	Correct

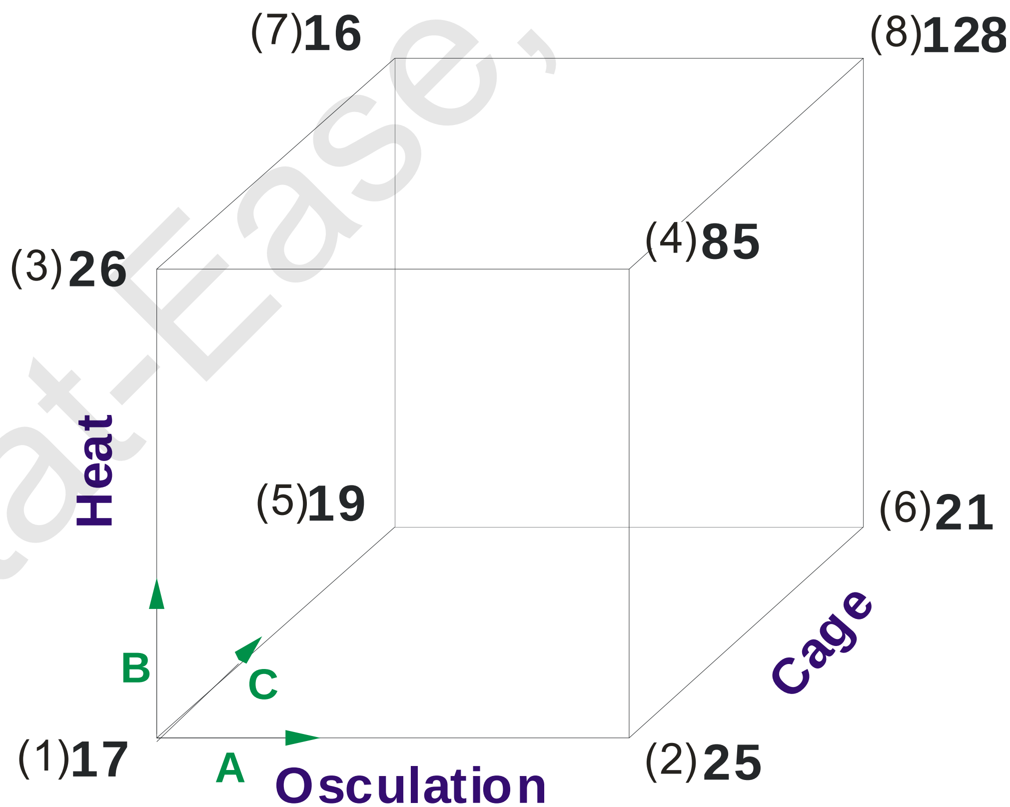
Let's look at two case studies illustrating both errors.

Agenda

- ✓ *Thorny issues (stuff happens)*
- ❑ The secret to long life (bearing case)
- ❑ A case to test your metal (defects defeated!)
- ❑ Conclusions

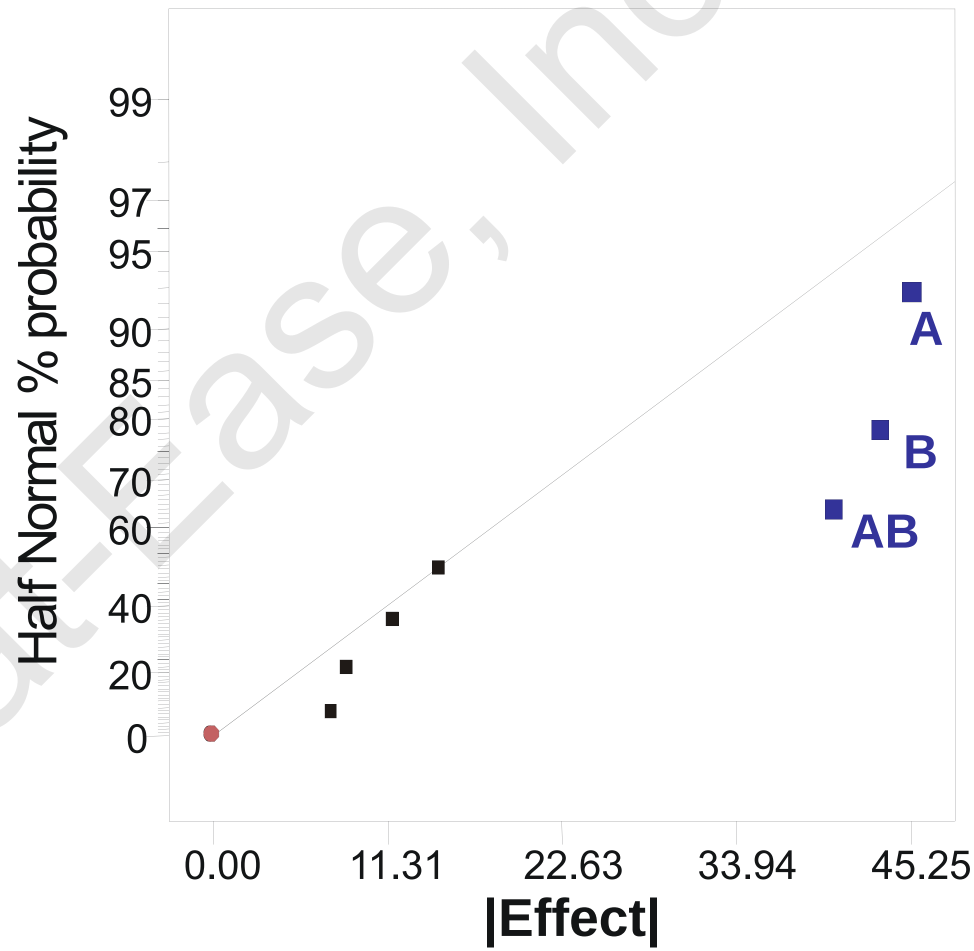
THE SECRET TO LONG LIFE!

Box's protégé makes bearing breakthrough, but is it too good to be true?



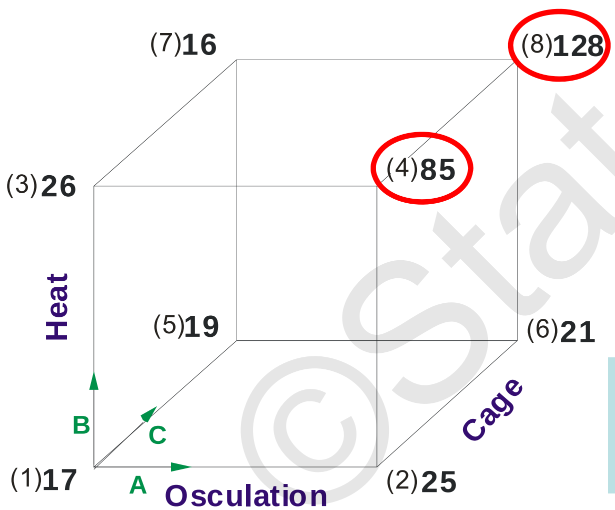
Bearing Case:
Half-Normal Plot of Effects

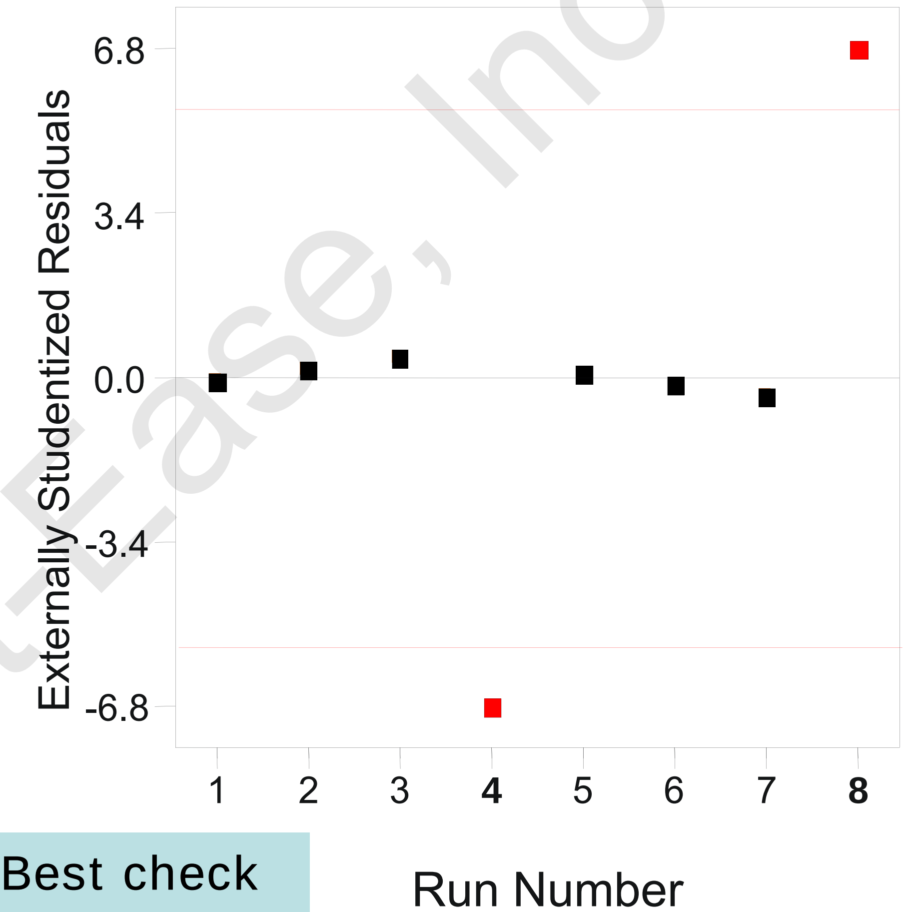
What's wrong with this picture?



Bearing Case:
Apparent Outliers

Toss the outliers? Better check first to identify them (see cube plot for run #'s).

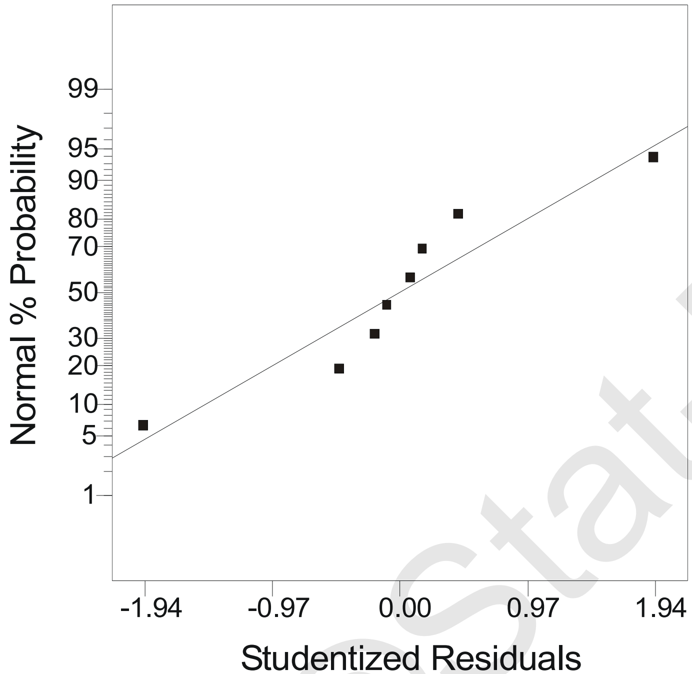


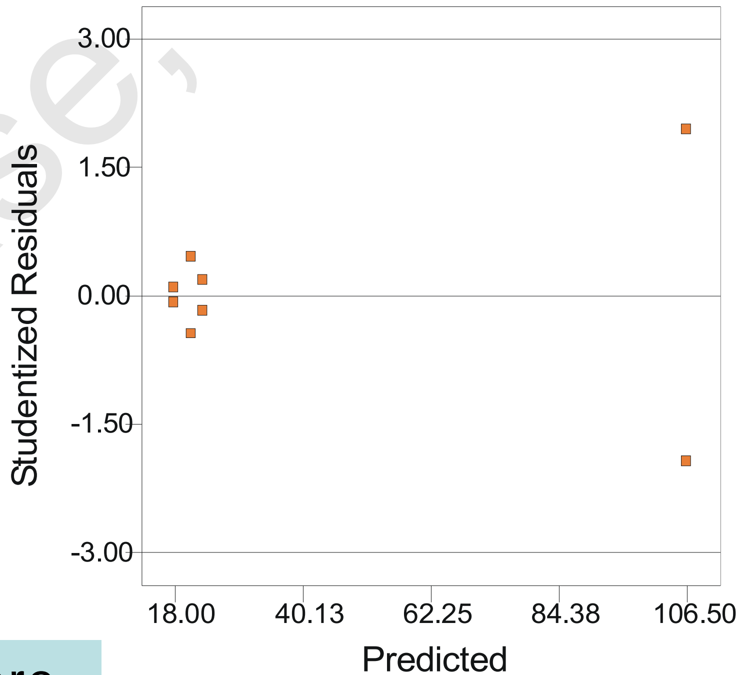


Best check diagnostics!

Bearing Case:
Further Diagnostics

Is assumption correct that residuals are normally distributed with constant variance?



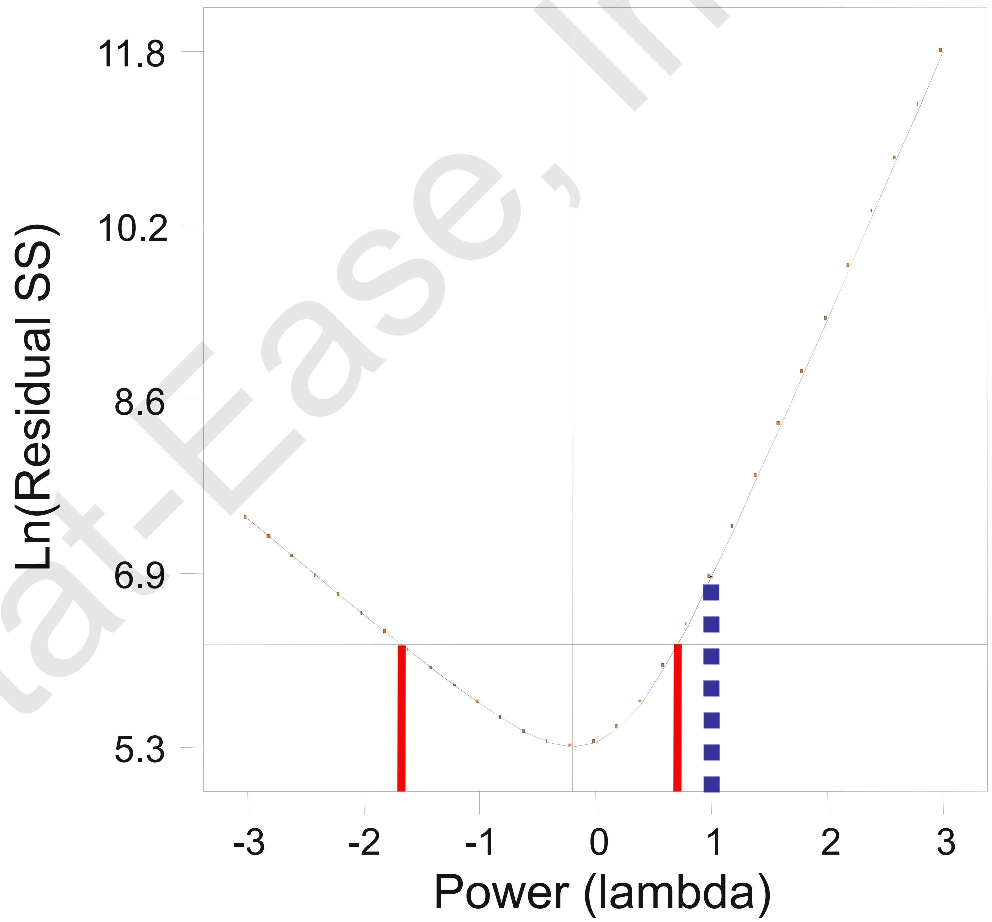


One more diagnostic!

Bearing Case:
Box-Cox Plot

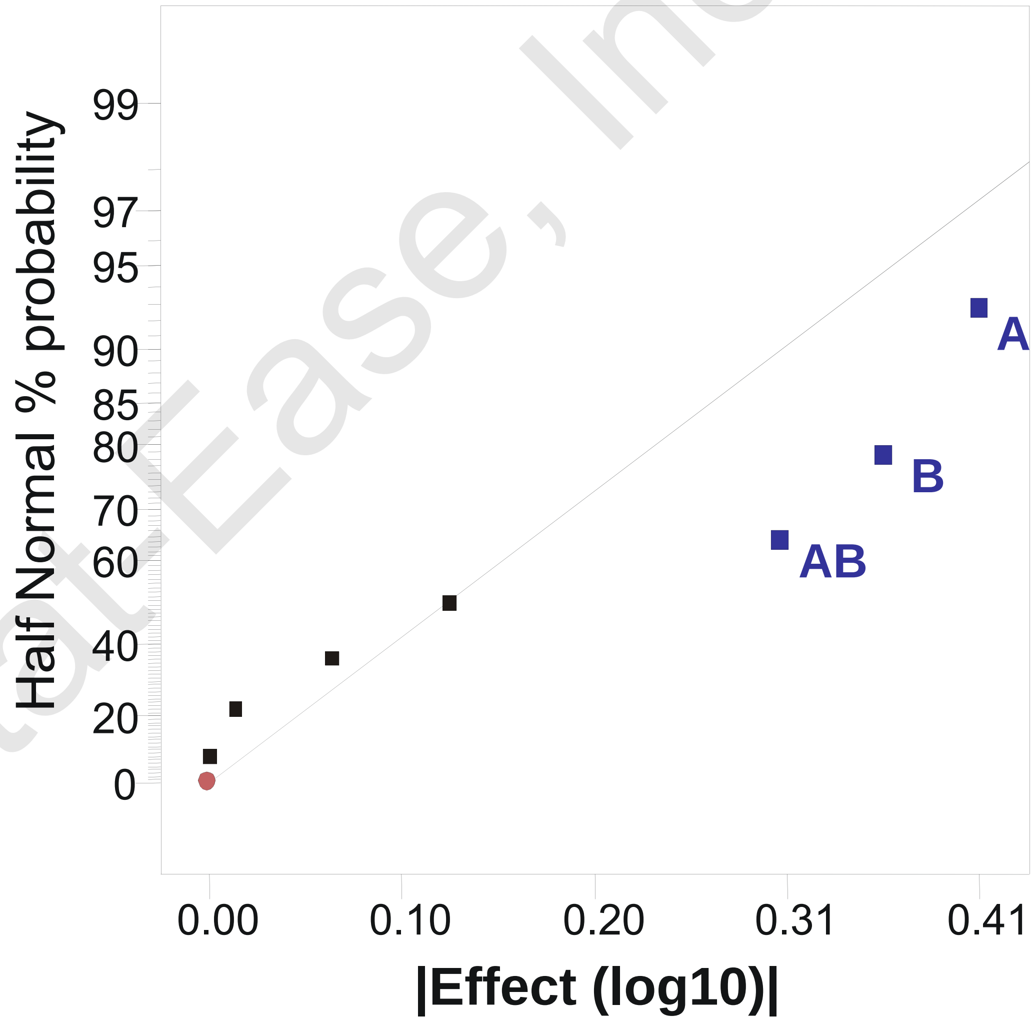
Blue line is outside of red zone.

☞ *Transform will help - try log.*



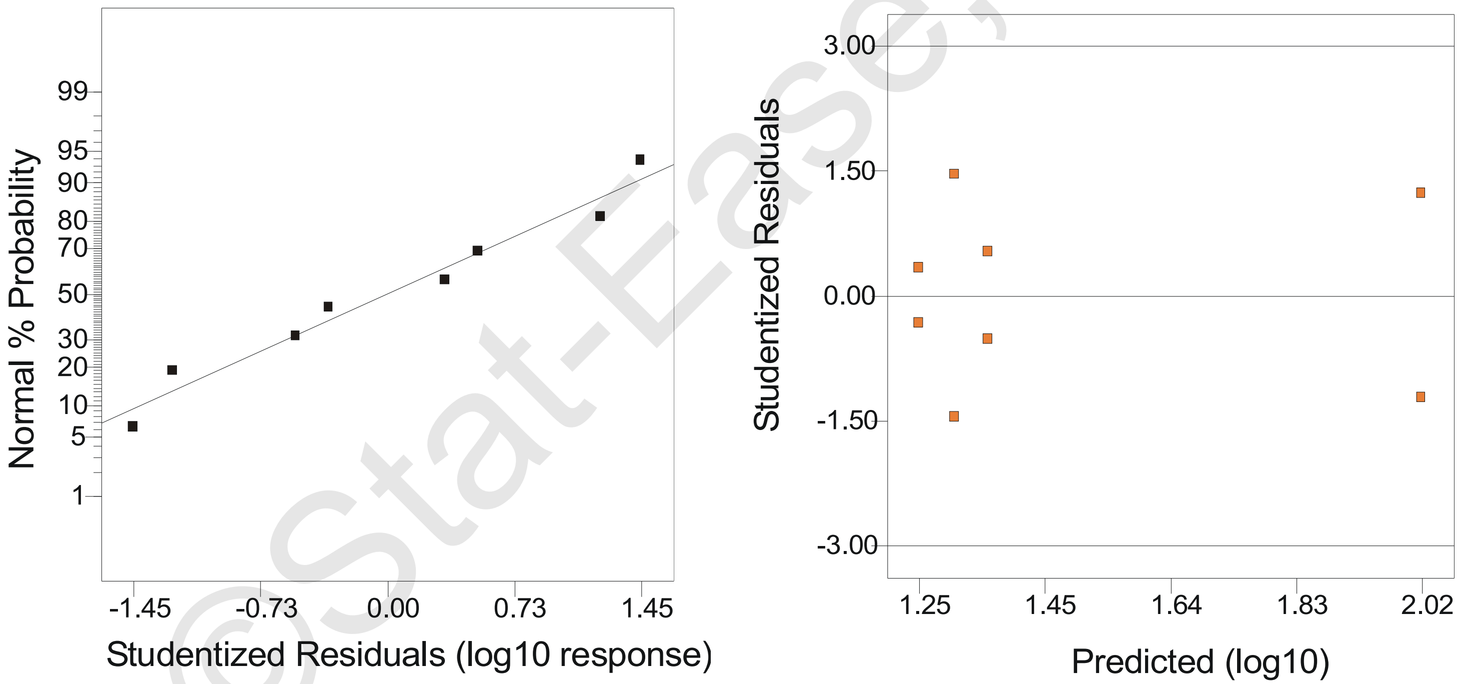
Bearing Case (*after log transformation*): Plot of Effects

Nice line-up!



Bearing Case (*after log transformation*): Residual Diagnostic Plots

Normally distributed with constant variance!



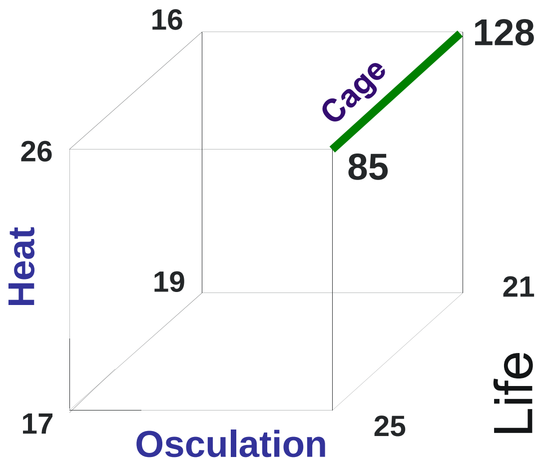
Bearing Case (*after log transformation*): Plot for Detecting Outliers

Can include the fabulous results! None are outliers.

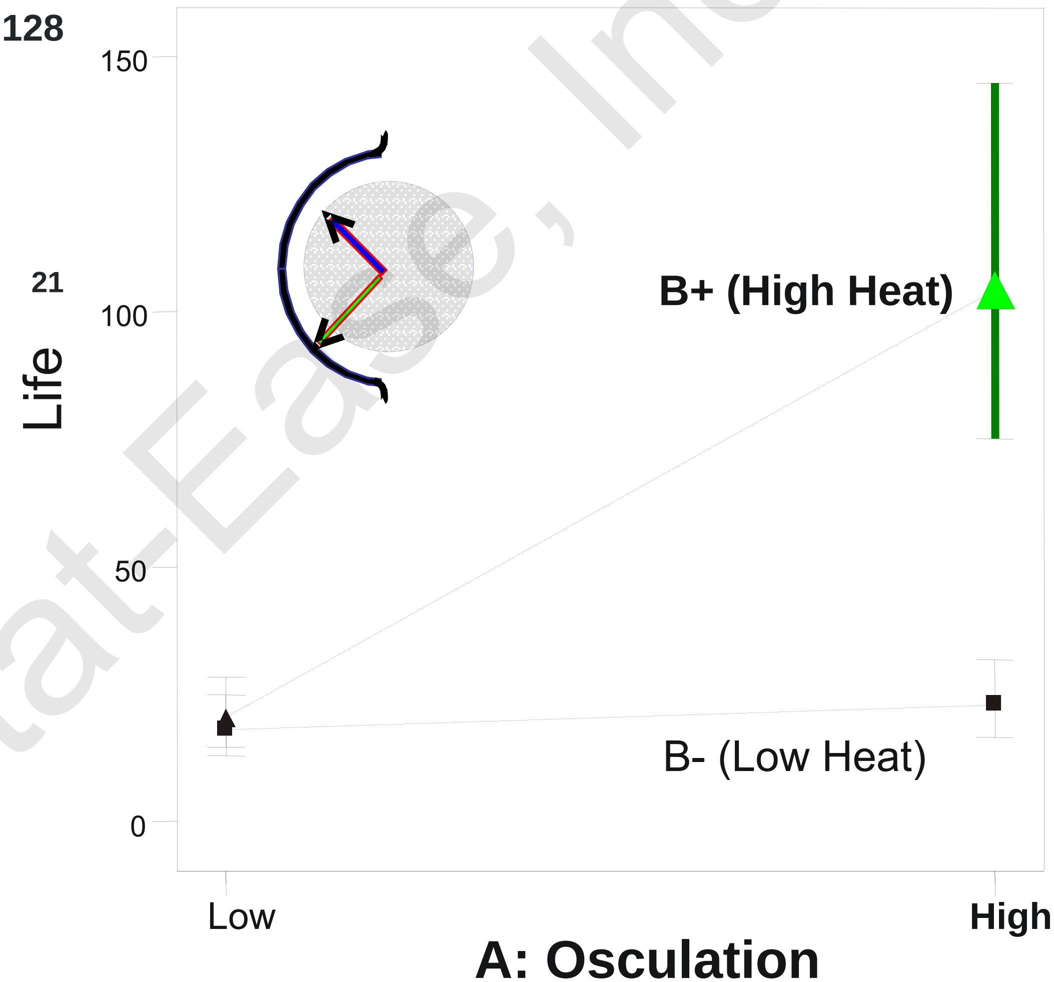
"Nature goes her own way and all that seems an exception is really according to order."

-- Goethe

Bearing Case (*after log transformation*):
Interaction Plot *(the happy ending!)*



Notice how confidence (LSD) bar widens at breakthrough conditions.



The rest of the story

- Swedish manufacturer SKF invented roller bearings in 1919, but by the '70s the Japanese had achieved comparable quality at competitive prices. These pressures inspired SKF engineers* to quit doing experiments only one factor at a time and try a two-level factorial design (led by Box' protégé).
- Ultimately SKF improved their actual bearing life from 41 million revolutions on average (already better than any competitors), to 400 million revs* – nearly a ten-fold improvement!

*("Breaking the Boundaries," *Design Engineering*, Feb 2000, pp 37-38.)

Recap of this case via Design-Expert® Software

Design-Expert makes it easy to set up this experiment design, model the results and analyze them statistically—including diagnosing the residuals, which in this case proves to be of vital importance.

Bearings

Agenda

- ✓ *Thorny issues (stuff happens)*
- ✓ *The secret to long life (bearing case)*
- ❑ A case to test your metal (defects defeated!)
- ❑ Conclusions

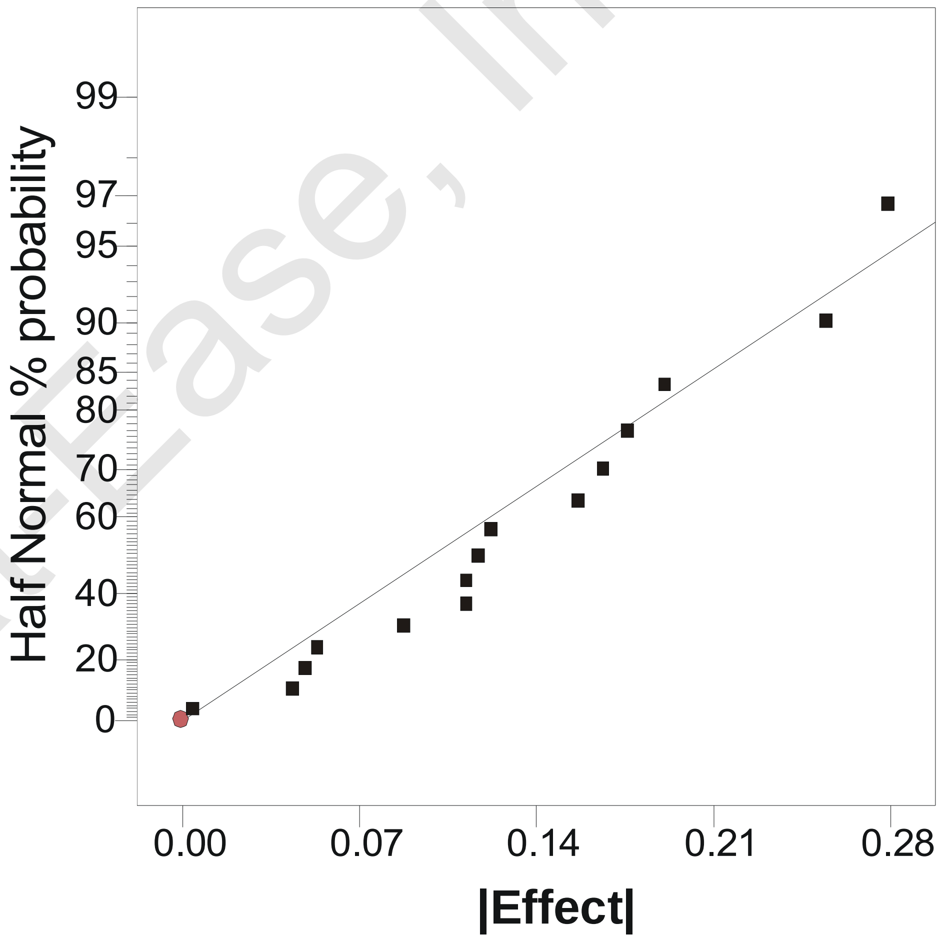
A Case to Test Your Metal

Defect rates on die-cast aluminum disk-drive housing drop dramatically, but is it real?

Std	A Oil Temp	B Trip	C Metal Temp	D Shot	E Dwell	Fraction Defect
1	350	390	1260	1.60	5.50	0.14
2	450	390	1260	1.60	3.50	0.98
3	350	410	1260	1.60	3.50	0.36
4	450	410	1260	1.60	5.50	0.42
5	350	390	1300	1.60	3.50	1.00
6	450	390	1300	1.60	5.50	0.90
7	350	410	1300	1.60	5.50	0.28
8	450	410	1300	1.60	3.50	0.14
9	350	390	1260	2.20	3.50	0.22
10	450	390	1260	2.20	5.50	0.26
11	350	410	1260	2.20	5.50	0.38
12	450	410	1260	2.20	3.50	0.12
13	350	390	1300	2.20	5.50	0.30
14	450	390	1300	2.20	3.50	0.06
15	350	410	1300	2.20	3.50	0.22
16	450	410	1300	2.20	5.50	0.38

Normal Plot of Effects

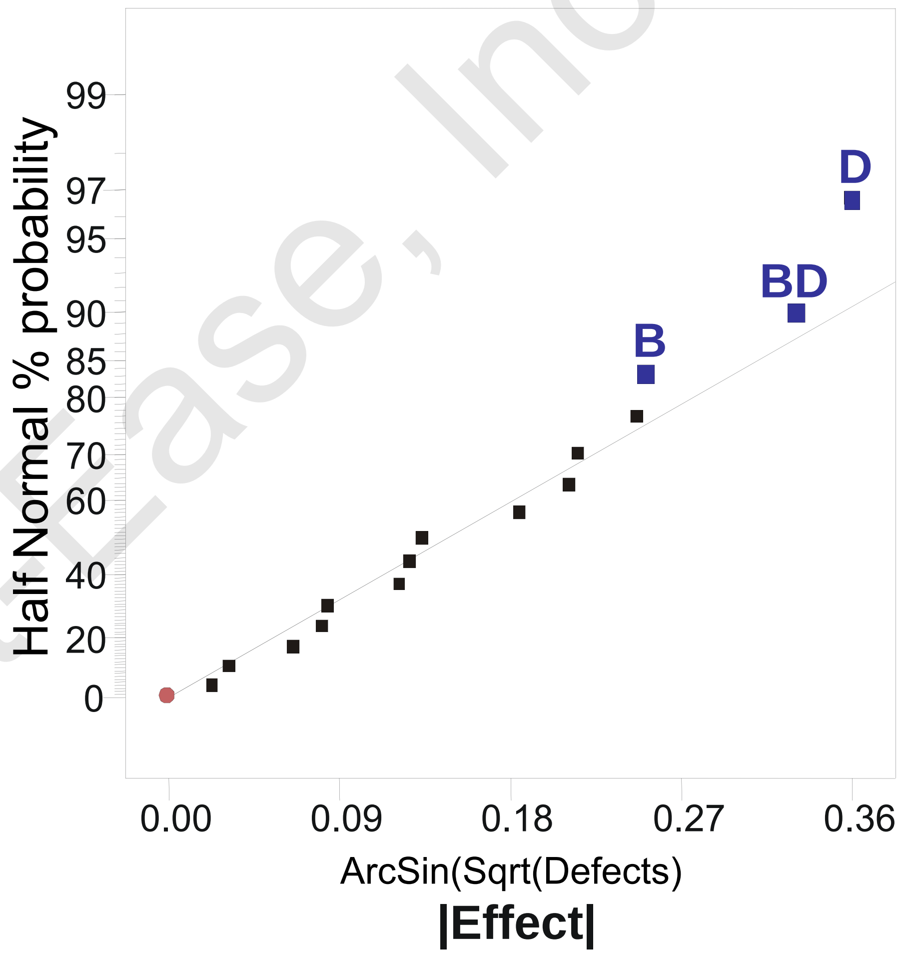
- All effects line up so evidently nothing is significant. ☹
- Don't give up yet – try transform: arc-sin square root good for binomial data like this.



Normal Plot of Effects (*transformed*)

All effects still line up so arc-sin square root no help. ☹ ☹

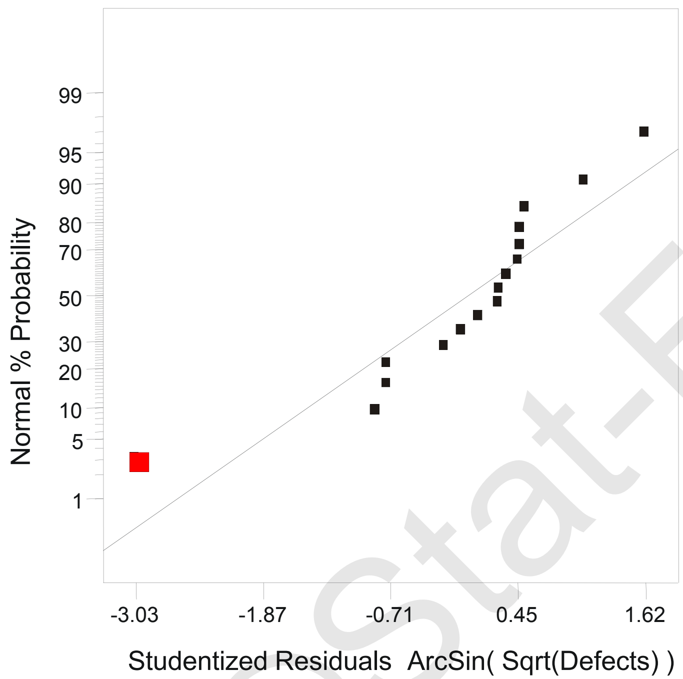
Don’t give up yet – pick several of the biggest effects and look for outliers.

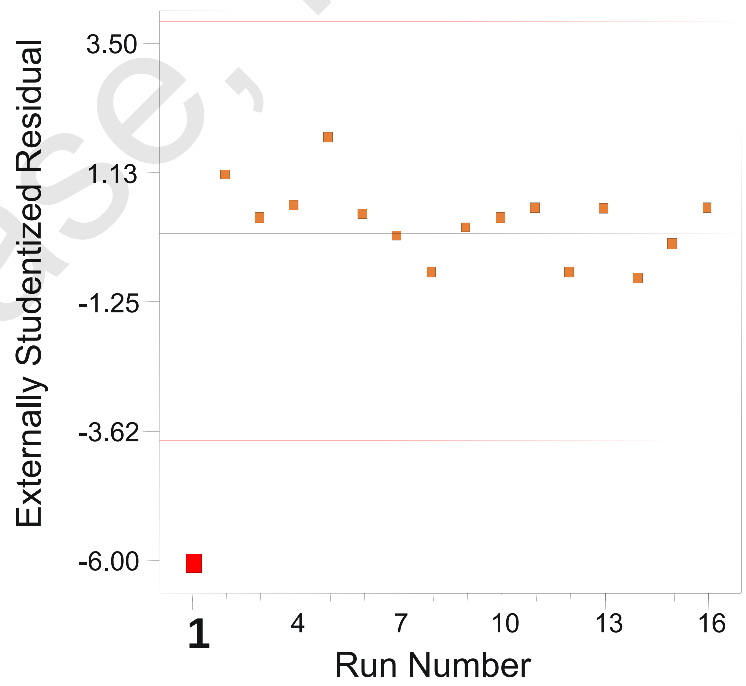


A Case to Test Your Metal

Diagnostic Plots

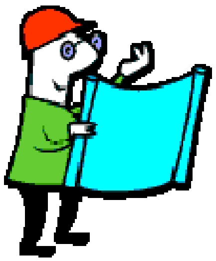
Problem revealed: Obvious outlier! Special cause?





Investigation of Outlier

Diagnostic plots identified standard order number 1 as a discrepant outcome. The foreman, when confronted with this statistical evidence, broke down and confessed that his crew overlooked this particular combination of factors. They then tried to make up for it by coming in early the following week, after shutting down the foundry over the weekend, to sneak the missing run in before the engineer came in to work.

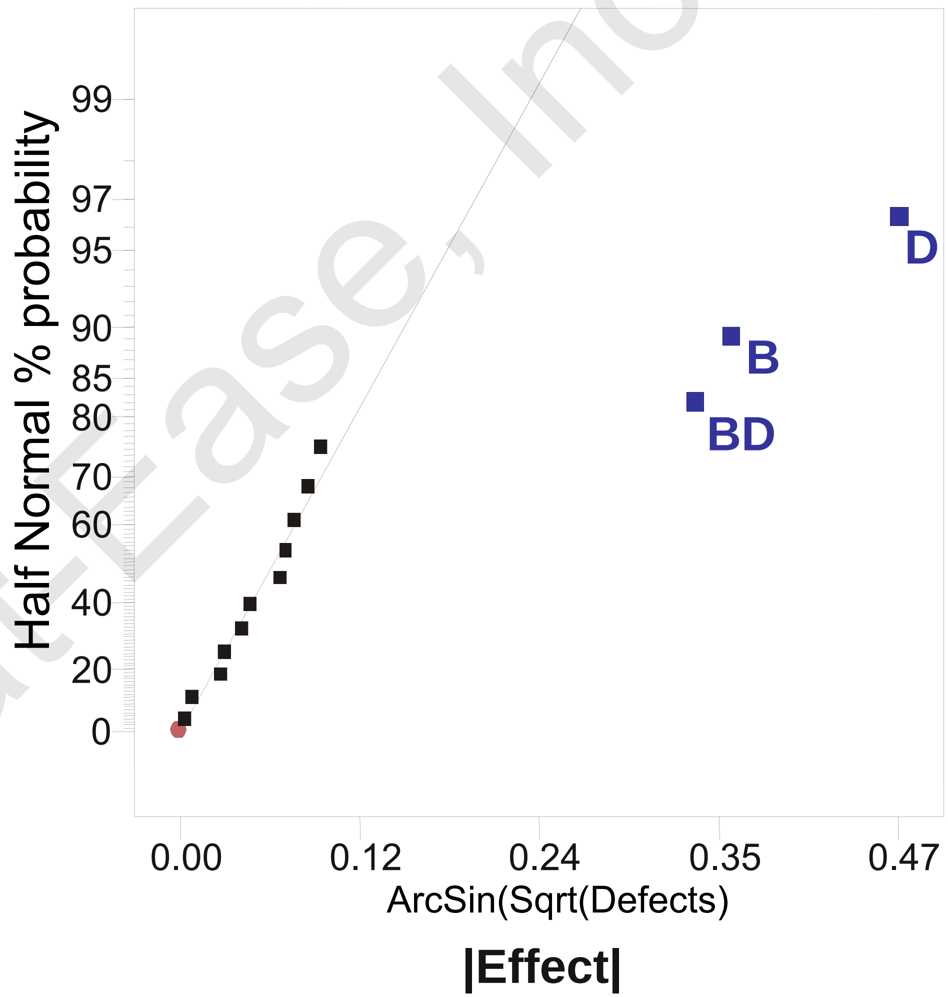


Next step: ignore this run and re-analyze.

Normal Plot of Effects

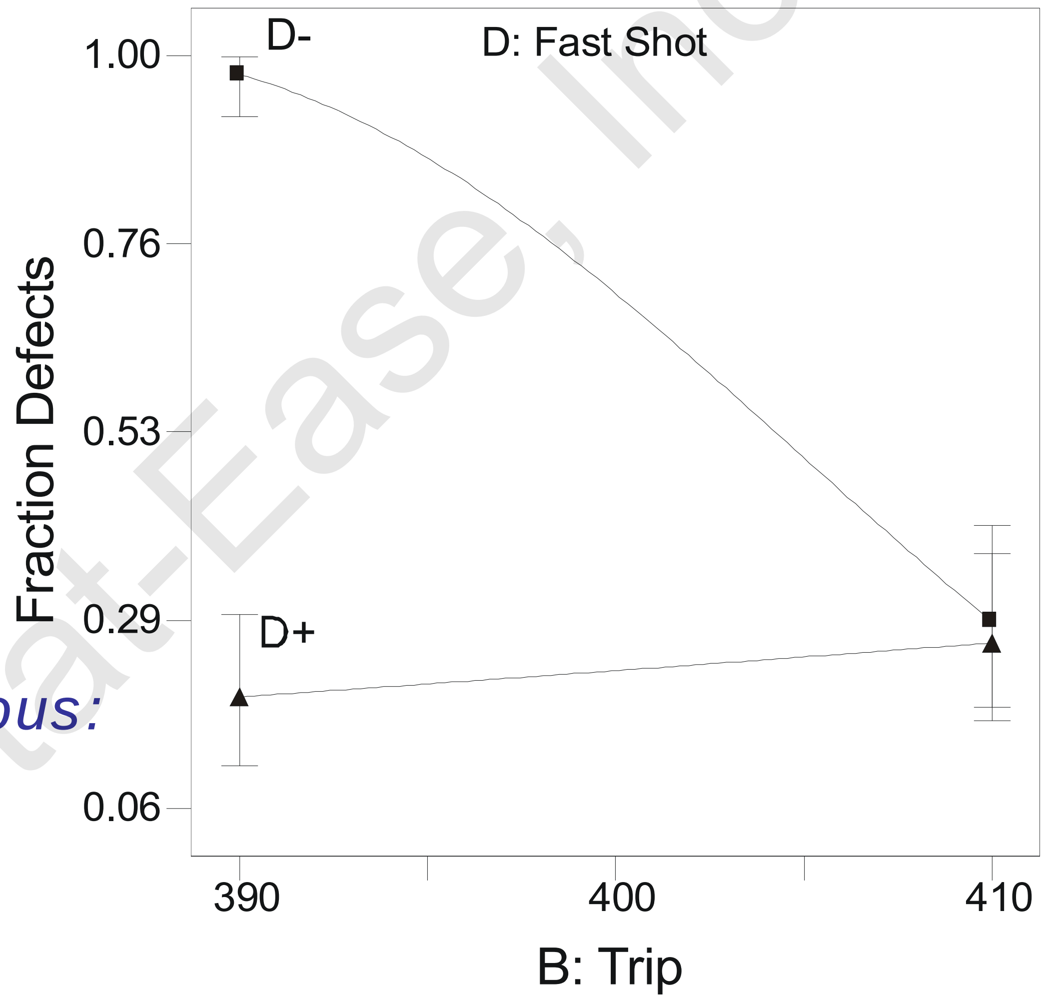
Family of effects now clearly significant. ☺

(Warning: missing data causes information loss, but not serious in this case.)



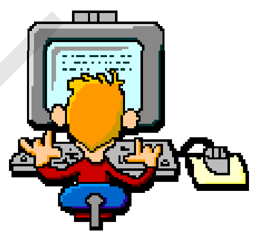
A Case to Test Your Metal (*outlier removed*)

Interaction Plot *(the happy ending!)*



Answer now obvious:
Do <u>not</u> run both
B & D low.

Again, Design-Expert makes it easy to set up this experiment design, model the results and analyze them statistically—including diagnosing the residuals, which, as in the last case, proves to be of vital importance.



metmold-outlier ignored

Agenda

- ✓ *Thorny issues (stuff happens)*
- ✓ *The secret to long life (bearing case)*
- ✓ *A case to test your metal (defects defeated!)*
- ❑ Conclusions

Conclusions

An outlier is a response from an experiment that does not fit the proposed model. Before jumping to any conclusions, consider these possibilities:

1. The model is faulty, not the data *(Ex. bearing case)*.
   - Watch for bad patterns on diagnostic plots: normal, residuals versus predicted, outlier (externally studentized) and Box-Cox.
   - Try transformation such as a log, square root (for counts), inverse (for rates) or arcsin square root (for fraction defects).

(further) Conclusions

2. The result really is an outlier *(Ex. die-cast aluminum)*.
   - Look for possible errors in data entry, or in response measurement, or in the conduct of that particular experimental run.
   - If the response really differs at that particular combination of the design factors, further study may lead to an important discovery!

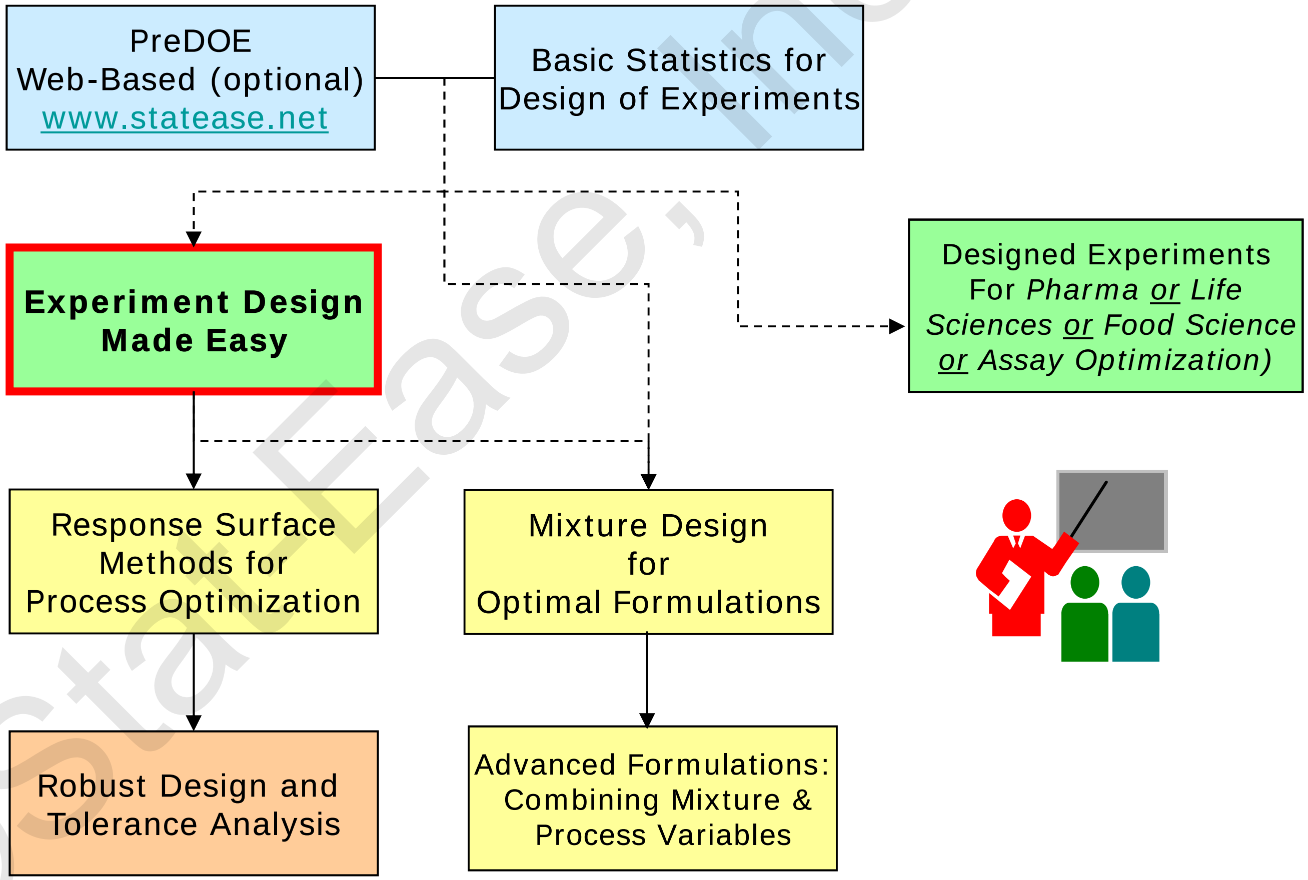
Above all – <u>avoid bias</u>:

"The first principle is that you must not fool yourself--and you are the easiest person to fool."

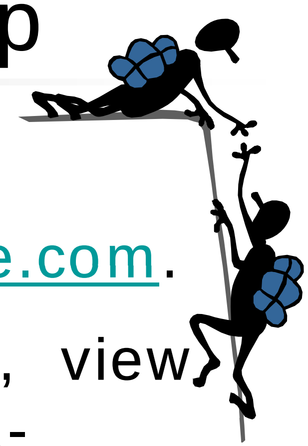
- Richard Feynman

Stat-Ease Training:
Computer-Intensive Statistical Workshops

Shari Kraber,
Workshop Manager
& Master Statistician
shari@statease.com



How to get help



- ❑ Search publications posted at www.statease.com.

- ❑ In Stat-Ease software press for Screen Tips, view reports in annotated mode, look for context-sensitive Help (right-click) or search the main Help system.
- ❑ Explore Experiment Design Forum http://forum.statease.com and post your question (if not previously answered).
- ❑ E-mail stathelp@statease.com for answers from Stat-Ease's staff of statistical consultants.
- ❑ Call 612.378.9449 and ask for "statistical help."

Statistics Made Easy®

Best of luck for your experimenting!

Thanks for listening!

-- Mark

Mark J. Anderson, PE, CQE
Stat-Ease, Inc.
mark@statease.com

**Pdf of this Powerpoint presentation posted at www.statease.com/webinar.html. For future webinars, subscribe to DOE FAQ Alert at www.statease.com/doealert.html.*