



Best "Pat-Tricks" on Model Diagnostics

What are they? Why use them? What good do they do?

Before we get started feel free to download the presentation and file(s) being used for today's webinar.

<http://www.statease.com/webinar.html>

You will find:

- three PDF files:
 - *Diagnostic 'Pat Tricks'.pdf*
 - *Diagnostics Report – Formulas & Definitions.pdf*
 - *Residual Analysis and Diagnostics Plots Guide.pdf*
- one self extraction zip file of Design-Expert® files:
 - *Diagnostics.exe*

Diagnostic "Pat Tricks"

1



Who We Are

This Month's Webinar

Webinar presented by:
Pat Whitcomb



Diagnostic "Pat Tricks"

With the Stat-Ease, Inc.
consulting team



Mark Anderson



Shari Kraber



Wayne Adams

2



Best "Pat-Tricks" on Model Diagnostics a series of examples

1. Popcorn
2. Bearings
3. Metal Molding
4. Ratio Assay
5. Persian CCD
6. Rails
7. Collet

Diagnostic "Pat Tricks"

3



Agenda Transition

1. **Popcorn**
2. Bearings
3. Metal Molding
4. Ratio Assay
5. Persian CCD
6. Rails
7. Collet

Diagnostic "Pat Tricks"

4

Use a simple DOE to introduce diagnostics:



Std order	A: Brand expense	B: Time minutes	C: Power percent	R _i : Taste rating
1	Cheap	4.0	75.0	74.0
2	Costly	4.0	75.0	75.0
3	Cheap	6.0	75.0	71.0
4	Costly	6.0	75.0	80.0
5	Cheap	4.0	100.0	81.0
6	Costly	4.0	100.0	77.0
7	Cheap	6.0	100.0	42.0
8	Costly	6.0	100.0	32.0

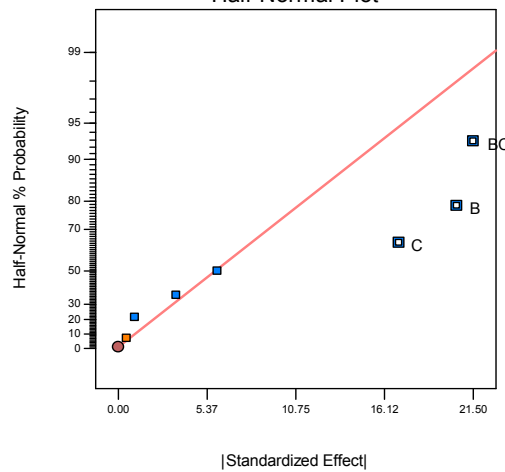
Diagnostic "Pat Tricks"

5

Design-Expert® Software
Taste

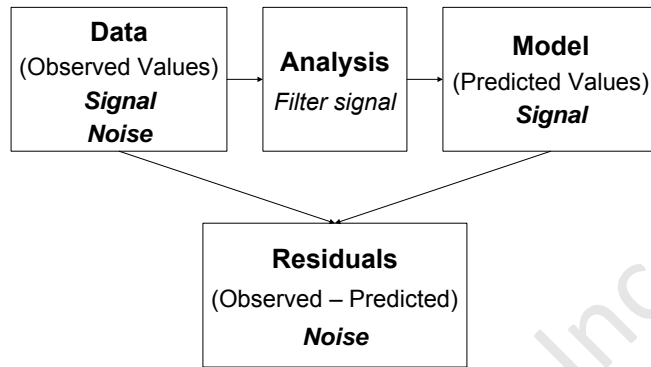
Shapiro-Wilk test
W-value = 0.973
p-value = 0.861
A: Brand
B: Time
C: Power
■ Positive Effects
■ Negative Effects

Half-Normal Plot



Diagnostic "Pat Tricks"

6



Independent $N(0, \sigma^2)$

Diagnostics → Influence → Report

Diagnostics Case Statistics

Std	Actual	Predicted		Internally	Externally	Influence on			
Order	Value	Value	Residual	Leverage	Studentized Residual	Studentized Residual	Fitted Value	Cook's	Run
							DFFITS	Distance	Order
1	74.00	74.50	-0.50	0.500	-0.142	-0.123	-0.123	0.005	8
2	75.00	74.50	0.50	0.500	0.142	0.123	0.123	0.005	1
3	71.00	75.50	-4.50	0.500	-1.279	-1.441	-1.441	0.409	2
4	80.00	75.50	4.50	0.500	1.279	1.441	1.441	0.409	4
5	81.00	79.00	2.00	0.500	0.569	0.514	0.514	0.081	3
6	77.00	79.00	-2.00	0.500	-0.569	-0.514	-0.514	0.081	5
7	42.00	37.00	5.00	0.500	1.421	1.750	1.750	0.505	7
8	32.00	37.00	-5.00	0.500	-1.421	-1.750	-1.750	0.505	6

See *"Diagnostics Report – Formulas & Definitions.pdf"*

Factorial Design ANOVA Assumptions

Additive treatment effects

Factorial: An interaction model will adequately represent response behavior.

- Model F-test
- Lack-of-Fit
- Box-Cox plot

Independence of errors

Knowing the residual from one experiment gives no information about the residual from the next.

S Residuals
versus
Run Order

Studentized residuals $N(0, \sigma^2)$:

- Normally distributed
- Mean of zero
- Constant variance, $\sigma^2=1$

Normal Plot of
S Residuals

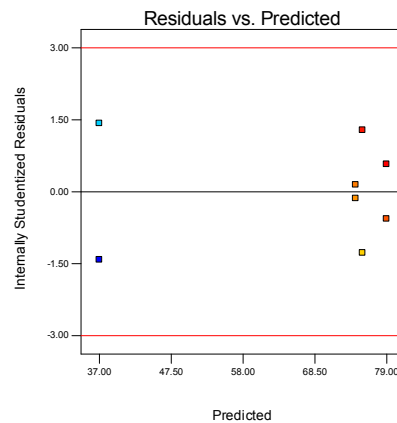
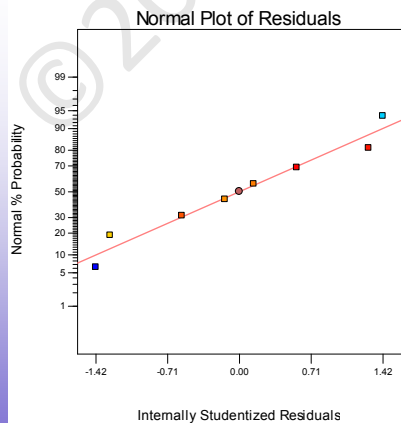
S Residuals
versus
Predicted

See ["Residual Analysis and Diagnostics Plots Guide.pdf"](#)

Diagnostic "Pat Tricks"

9

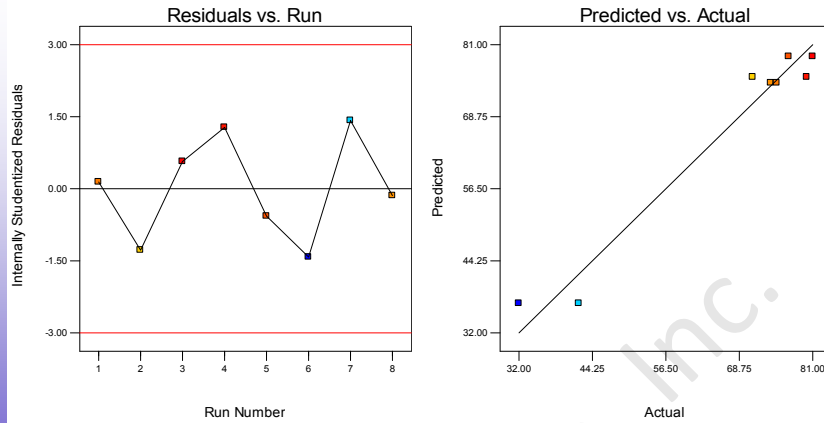
Popcorn Analysis – Taste Diagnostics - ANOVA Assumptions



Diagnostic "Pat Tricks"

10

Popcorn Analysis – Taste Diagnostics - ANOVA Assumptions

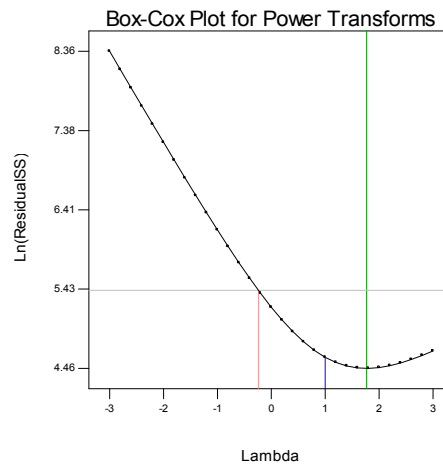


Diagnostic "Pat Tricks"

11

Popcorn Analysis – Taste Diagnostics - ANOVA Assumptions

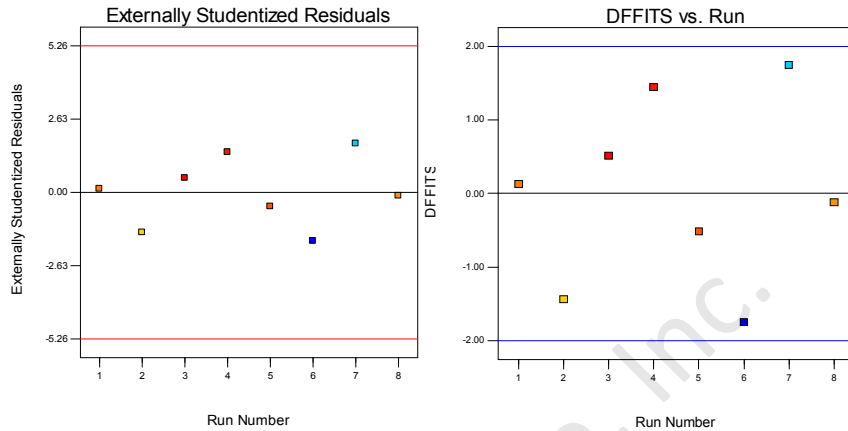
Design-Expert® Software
Taste
Lambda
Current = 1
Best = 1.77
Low C.I. = -0.24
High C.I. = 4.79
Recommend transform:
None
(Lambda = 1)



Diagnostic "Pat Tricks"

12

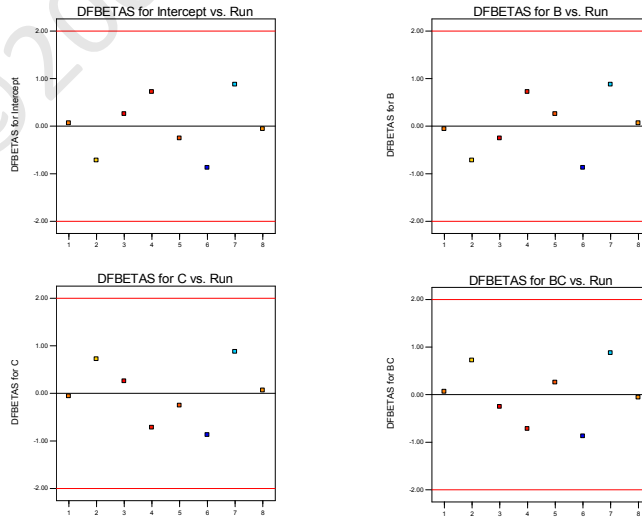
Popcorn Analysis – Taste Influence



Diagnostic "Pat Tricks"

13

Popcorn Analysis – Taste Influence

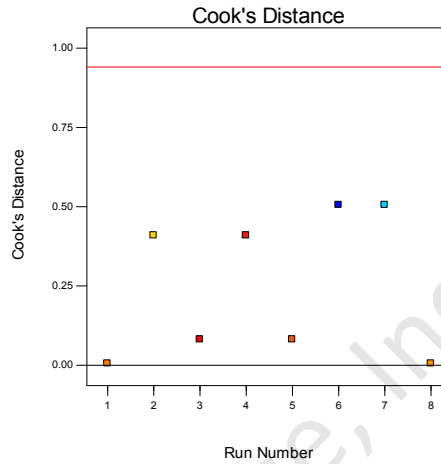


Diagnostic "Pat Tricks"

14

Popcorn Analysis – Taste Influence

Design-Expert® Software
Taste
Color points by value of Taste:
81.0
32.0



Diagnostic "Pat Tricks"

15

Diagnostic and Influence Summary

Tool	Description	WIIFM
Internally Studentized Residuals	Residual divided by the estimated standard deviation of that residual	Normality, constant σ^2
Externally Studentized Residuals	Residual divided by the estimated std dev of that residual, without the i^{th} case (deletion statistic)	Outlier detection and/or model misspecification
Cook's Distance	Change in joint confidence ellipsoid (regression) with and without a run (deletion statistic)	Overall influence
DFFits (difference in fits)	Change in predictions with and without a run; the influence a run has on the predictions (deletion statistic)	Influence on fitted value
DFBetas (difference in betas)	Change in each model coefficient (beta) with and without a run (deletion statistic)	Influence on coefficients

Diagnostic "Pat Tricks"

16

1. Popcorn
2. **Bearings**
3. Metal Molding
4. Ratio Assay
5. Persian CCD
6. Rails
7. Collet

Let's exercise your diagnostic tools

Case-Study Background:

Swedish manufacturer SKF invented rolling 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.



You will be amazed by the results!

Factor	Units	Low Level (-)	High Level (+)
Osculation*		std	mod
Heat		std	mod
Cage**		metal	polymeric

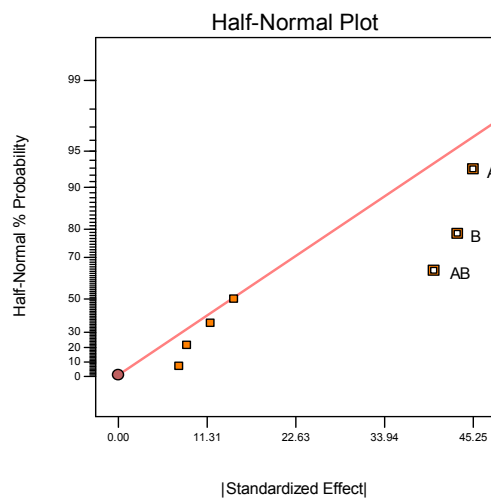
✓ Osculation is the 'kiss' of the bearing to the cage, calculated by the ratio of the two radii (outer/inner).



✓ The cage design factor involved material of construction, changing from a steel to a cheaper polymeric material.

Design-Expert® Software
Life

Shapiro-Wilk test
W-value = 0.940
p-value = 0.653
A: Osculation
B: Heat
C: Cage
■ Positive Effects
■ Negative Effects



Be ready to take notes!



Bearings.dx7



Bearings
Grand finale

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.

Theory & subject matter knowledge:

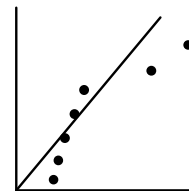
- Error is a constant across the range of data – no transform
- Error is a constant percentage of response – use log

Nature of response:

- counts – try square root
- variation – try log
- rate – try inverse
- Zero as a natural limit – try log
- proportion defective – use arcsine(square root)
- bounded data (approaching boundary) – try logit

Estimating effects:

- Ratio $Y_{\max}/Y_{\min} > 3$ to 10
- Offset of error line from intercept on half-normal plot
- Unbelievable (3FI or higher) interactions



Diagnostic plots:

- Non-linear normal probability plot
- Pattern in residual versus predicted
- Unexplained outliers
- 45° line doesn't "split" the data on predicted versus actual
- Box Cox plot

Diagnostics aren't much help when selecting the non power law transformations. You have to recognize when you have a binomial or bounded response:

- Always use arcsin square root for binomial data.
- Use logit when bounded data approaches a boundary.

1. Popcorn
2. Bearings
3. **Metal Molding**
4. Ratio Assay
5. Persian CCD
6. Rails
7. Collet

A customer of Stat-Ease examined his aluminum injection molding process using a fractional factorial design. He studied five factors in a 2^{5-1} design:

- A. Hot oil temperature
- B. Trip in mm
- C. Molten aluminum temperature
- D. Fast shot velocity
- E. Dwell time



The **fraction defective** out of 100 parts, was recorded for each design point. To his dismay, none of the factors seemed to make any difference.



Metal Molding.dx7



Metal Molding

Run 9 (*standard order 1*) was investigated and the operator confirmed problems with this run.

Diagnostic "Pat Tricks"

29



Agenda Transition

1. Popcorn
2. Bearings
3. Metal Molding
4. **Ratio Assay**
5. Persian CCD
6. Rails
7. Collet

Diagnostic "Pat Tricks"

30

- To evaluate the total:background ratio of this assay performed in a 96-well plate format
- Objective of this study is to optimize the total signal by evaluating the following five factors:

Factor	-1 level	+1 level
NaCl	0 μ L	150 μ L
Tween 20	0 μ L	100 μ L
EDTA	0 μ L	200 μ L
Incubation Time	0.5 hour	1 hour
Stop solution	Na ₂ CO ₃	Tris-HCl

- The design used a 2⁵ factorial with 4 replicates in 4 blocks (on 4 plates).

This assay system will study the affects of the buffer factors (NaCl, Tween 20, EDTA), incubation time, and different stop solutions.

This experiment utilizes experimental wells and background wells.

- **Experimental wells** contain the factors we are testing under standard assay conditions.
 - Sample is added to each experimental well and incubated for either 0.5 or 1 hour.
- An equal number of **background wells** are used that will only contain buffer and normalizing solutions (i.e. no sample is added).

Be ready to take notes!



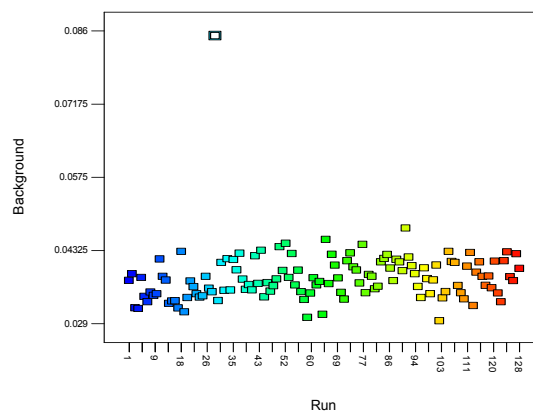
Ratio Assay.dx7

Diagnostic "Pat Tricks"

33

Ratio Assay

Obviously there was a problem (don't know what) with the background well for run 29:



Diagnostic "Pat Tricks"

34

1. Popcorn
2. Bearings
3. Metal Molding
4. Ratio Assay
5. **Persian CCD**
6. Rails
7. Collet

Iranian chemists ran an RSM experiment to optimize a polymerization process. They set up a CCD on four key factors:

			Low (-1)	High (+1)
A	Temperature	°C	50.0	90.0
B	Time	hours	0.75	1.75
C	Catalyst	%	2.0	5.0
D	Water	%	0.3	0.7

Find conditions to maximize conversion (y_1) of monomer to polymer.

Be ready to take notes!



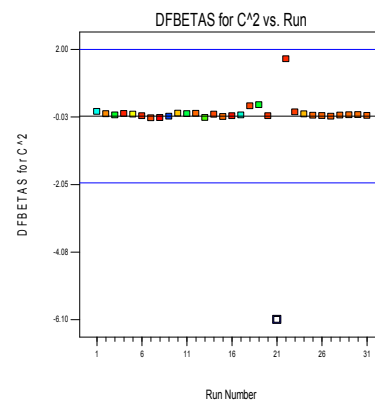
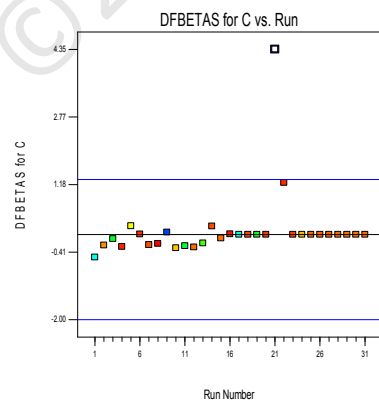
Persian CCD.dx7

Diagnostic "Pat Tricks"

37

Persian CCD

Too little catalyst (0.5 %) at minus alpha:



Diagnostic "Pat Tricks"

38

1. Popcorn
2. Bearings
3. Metal Molding
4. Ratio Assay
5. Persian CCD
6. **Rails**
7. Collet

Objective: Maximize sophorolipid production from *Candida Bombicola*.

Let's look at five factors thought to affect growth:

Factor	- 1 level	+1 level
Nitrogen	0 μ L	30 μ L
Phosphorus	0 μ L	30 μ L
Magnesium	0 μ L	30 μ L
Iron	0 μ L	20 μ L
Zinc	0 μ L	10 μ L

Build a 2^5 full factorial (32 runs) , 2 replicates, six center points; a total of 70 runs. Add 20 QC wells for a final total of 90 runs on a 96 well plate.

Be ready to take notes!



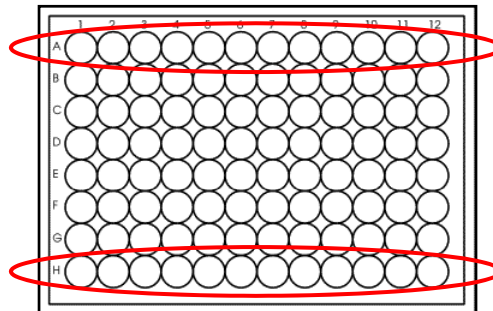
Rails.dx7

Diagnostic "Pat Tricks"

41

Rails

All fifteen outliers are in the first (A) and last (H) rows on the plate. Due to interference from the rails in the plate reader the first and last rows read low.



Diagnostic "Pat Tricks"

42

1. Popcorn
2. Bearings
3. Metal Molding
4. Ratio Assay
5. Persian CCD
6. Rails
7. **Collet**

A computer controlled lathe feeds in bar stock, cuts it, machines the surface and releases a part. The collet holds the part in place as it is being machined. The operator programs the speed (rate of spin) and feed (depth of cut). The operator hand tightens the collet.



The factors studied are:

	Factor	-1	+1	Units
A	Speed	2500	4500	rpm
B	Feed	0.003	0.009	in/rev
C	Collet	Loose	Tight	
D	Tool wear	New	after 250 parts	

The response is surface finish, measured on the same one inch section of each part. The higher the reading the rougher the surface. Low readings (a smooth surface) are desirable.

Diagnostic "Pat Tricks"

45

Be ready to take notes!



Collet.dx7

Diagnostic "Pat Tricks"

46

The parts made using slow feed and loose collet were examined and it was noted that this treatment combination produced an unusual finish. The surface profile looked like a sine wave.

The problem was that using this treatment combination the part oscillated and caused an uneven cut pattern.

- Outliers, model misspecification and response transformations can be identified via diagnostic and influence plots.
- Diagnostic and Influence plots are an essential for a through data analysis.
- I hope the examples presented have made a true believer of you!!

Diagnostic and Influence Summary

Tool	Description	WIIFM
Internally Studentized Residuals	Residual divided by the estimated standard deviation of that residual	Normality, constant σ^2
Externally Studentized Residuals	Residual divided by the estimated std dev of that residual, without the i^{th} case (deletion statistic)	Outlier detection and/or model misspecification
Cook's Distance	Change in joint confidence ellipsoid (regression) with and without a run (deletion statistic)	Overall influence
DFFits (difference in fits)	Change in predictions with and without a run; the influence a run has on the predictions (deletion statistic)	Influence on fitted value
DFBetas (difference in betas)	Change in each model coefficient (beta) with and without a run (deletion statistic)	Influence on coefficients

Diagnostic "Pat Tricks"

49

How to get help

- Search publications posted at www.statease.com
- E-mail statHelp@statease.com for answers from Stat-Ease's staff of statistical consultants
- Call 612.378.9449 and ask for "statistical help"



Thanks for attending!

Diagnostic "Pat Tricks"

50