



DO WE NEED TRANSFORMATIONS?

JOSÉ G. RAMÍREZ

100TH BIRTHDAY CELEBRATION OF THE ACCIDENTAL STATISTICIAN

AMGEN[®]

Pioneering science delivers vital medicines[™]

OUTLINE

- **The British Invasions**
- **Box-Cox Transformations**
- **Do We Need Transformations?**
- **Do We Need Transformations?**
- **Summary**

THE BRITISH INVASION OF 1964

- **The Beatles landed in NYC**
- **“I Want to Hold Your Hand”**
 - #1 in the charts
- **Ed Sullivan Show**
 - Sunday, February 9, 1964
- **A quartet that “transformed” the music industry.**

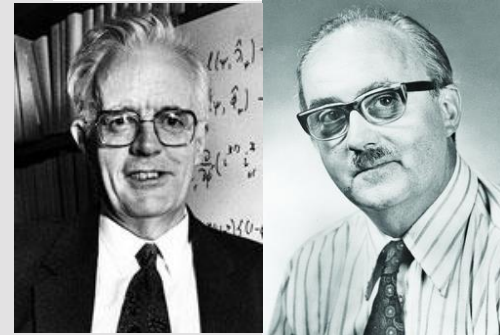
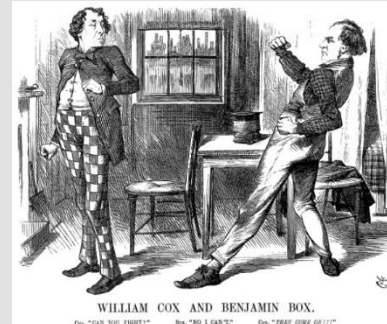
United States Library of Congress's Prints and Photographs division



© 1964/Ed Sullivan Productions

THE (OTHER) BRITISH INVASION OF 1964

- **An Analysis of Transformations**
 - JRSS B 26(2)
- **The “transforming” duo**
- **“two different derivations ...one using likelihood and the other Bayes”**
- **“this kind of oscillatory character between likelihood and Bayes analysis had any relevance to the Box and Cox aspect!”**
(Prof. Bartlett)
- **“using Bayesian arguments without becoming fully committed to them”**
(Prof. Plackett)



OUTLINE

- The British Invasions
- **Box-Cox Transformations**

A TREE AND TRANSFORMATIONS

- Under a tree in the orchard
 - Ronald looked up the probits
 - George looked up the reciprocals
- They plotted the data

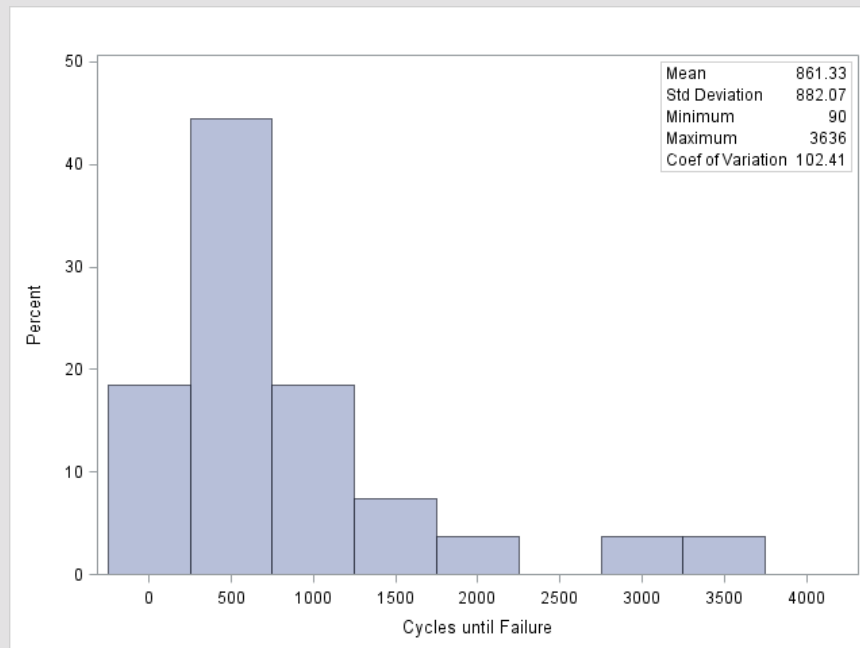


Courtesy of Prof. Genshiro Kitagawa

George's life was “transformed”

TRANSFORMATIONS: WHAT IS THE APPROPRIATE SCALE?

- Life cycles of worsted yarn
 - Table 4
 - Wide data range 90 – 3636
 - CV ~ 102%
 - **Large/small ~ 40**
- “When the ratio of the largest to smallest observation is large you should question whether the data are being analyzed in the right metric (transformation)”

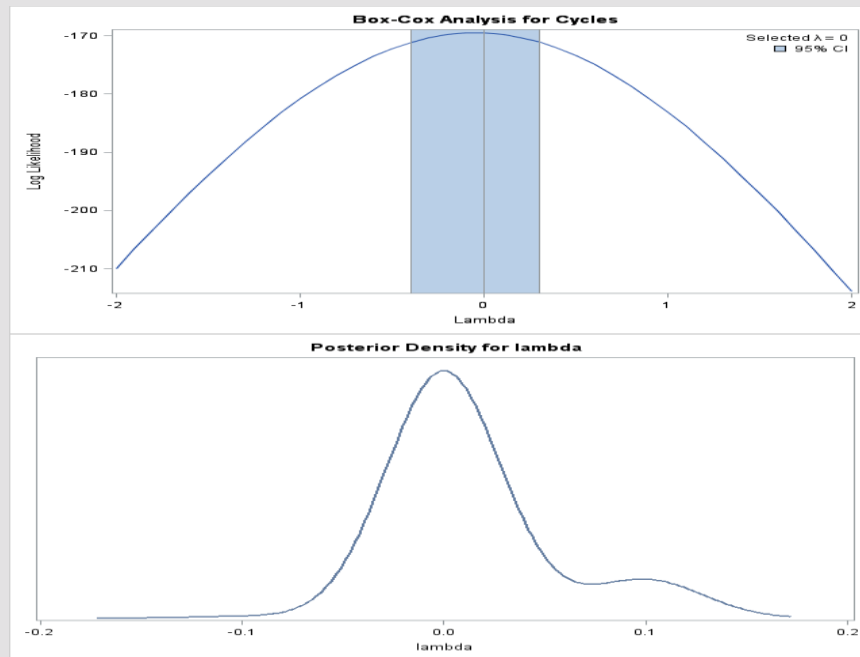


TRANSFORMATIONS: LIKELIHOOD & BAYESIAN

- **The Box-Cox family**

$$y^{(\lambda)} = \begin{cases} \frac{y^\lambda - 1}{\lambda} & \lambda \neq 0 \\ \log y & \lambda = 0 \end{cases} \quad z^{(\lambda)} = \begin{cases} \frac{y^\lambda - 1}{\lambda \dot{y}^{\lambda-1}} & \lambda \neq 0 \\ \dot{y} \log y & \lambda = 0 \end{cases}$$

- **Life cycles of worsted yarn**
 - What is the appropriate scale?
- The “family” suggests the log

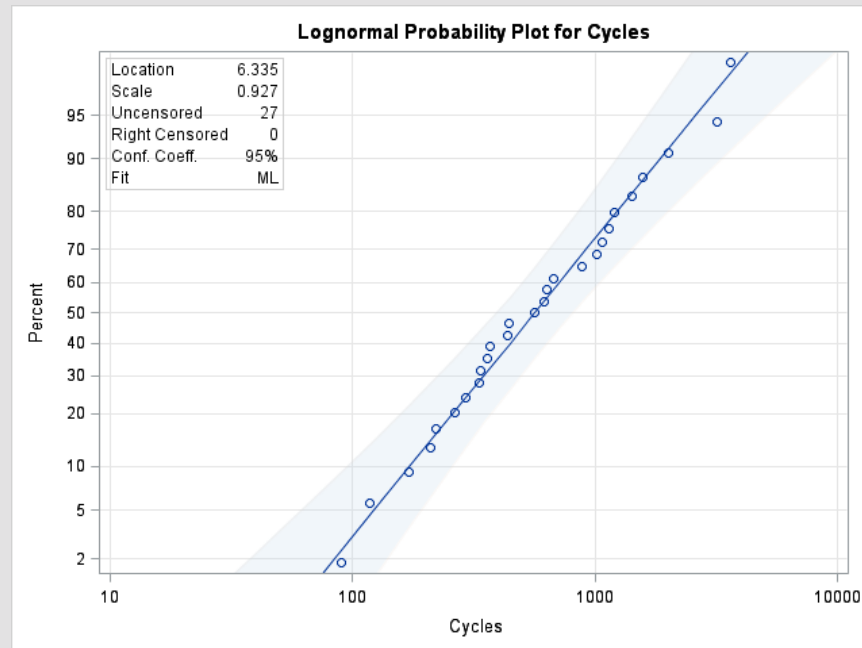


OUTLINE

- The British Invasions
- Transformations
- **Do We Need Transformations?**

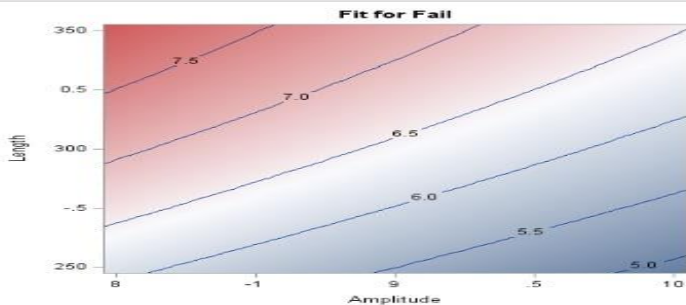
DO WE NEED TRANSFORMATIONS?

- **Life cycle data**
 - positive quantities
 - normal distribution?
- **Reliability engineering,**
 - nature of physical quantities
 - the normal is not the norm
- **Life cycles of worsted yarn**
 - Lognormal distribution

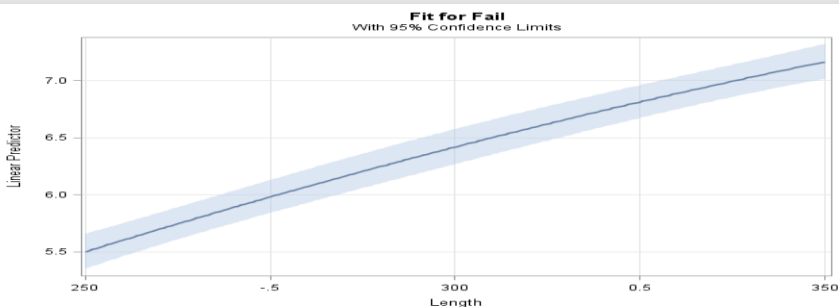


DO WE NEED TRANSFORMATIONS?

QUADRATIC MODEL UNDER LOGNORMAL DISTRIBUTION



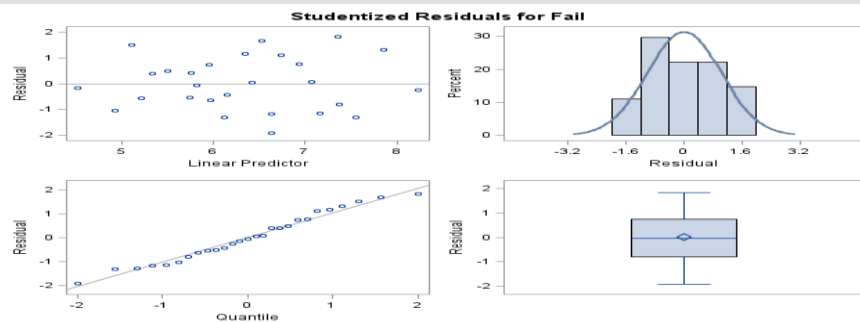
Fit computed at Load=0



Fit computed at Amplitude=0 Load=0

The transformation eliminates the need for second-order terms in the regression equation while at the same time increasing the sensitivity of the analysis by about three, as judged by the ratio of linear and residual mean squares.

Effect	Num DF	Den DF	F Value	Pr > F
Length	1	17	331.94	<.0001
Amplitude	1	17	190.75	<.0001
Load	1	17	73.80	<.0001
Length*Amplitude	1	17	0.47	0.5035
Length*Load	1	17	1.49	0.2381
Amplitude*Load	1	17	0.14	0.7142
Length*Length	1	17	1.17	0.2939
Amplitude*Amplitude	1	17	0.09	0.7633
Load*Load	1	17	0.73	0.4058



WHAT IS THE “APPROPRIATE” DISTRIBUTION?

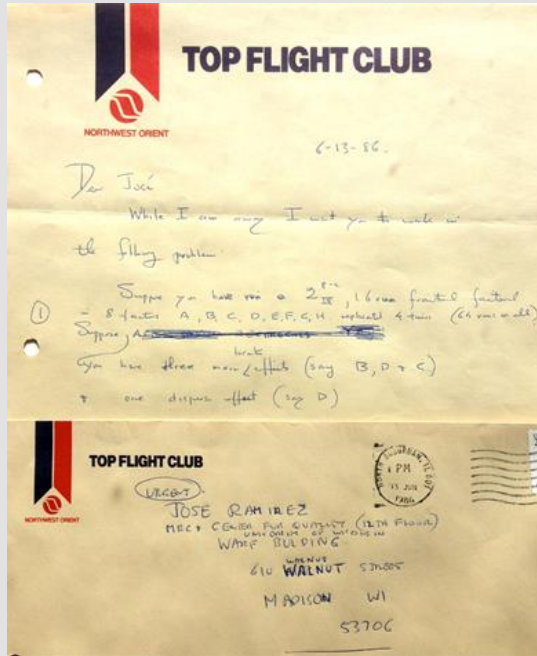
- Many advances in statistical techniques and software
- Relatively easy to use an “appropriate” distribution
 - Reliability methods
 - Generalized linear models
 - Generalized linear mixed models
- Choice of distribution should take into account

The clinical, scientific, physical and engineering principles governing the underlying phenomenon that generates the data.

OUTLINE

- The British Invasions
- Transformations
- **Do We Need Transformations?**
 - No
- **Do We Need Transformations?**

FROM THE TOP FLIGHT CLUB

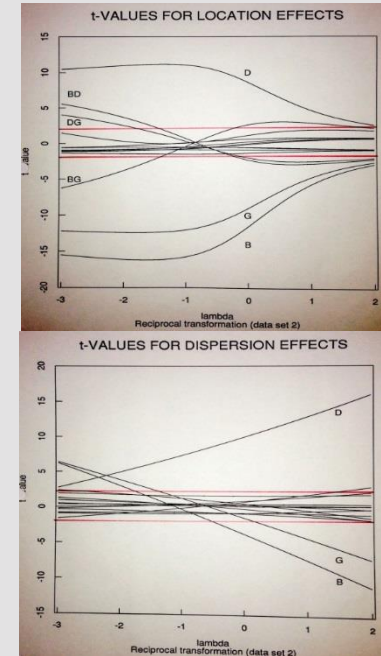


- Simulate data from a 2^{8-4}_{IV} experiment
- 3 location effects B, D, and G
 - Mean = 50
 - B=-30, D=+20, G=-20
- 1 dispersion effect D
 - $D_-: \sigma=2.5$, $D_+: \sigma=7.5$
- The reciprocal transformation was the appropriate scale.
- Will a **lambda plot** suggest the appropriate transformation?
 - Simplicity and separation

DO WE NEED TRANSFORMATIONS?

LAMBDA PLOTS: CONSTRUCTED EXAMPLE

- **Lambda plot**
 - Plot t-value vs Box-Cox λ
- **Simulate data from a 2_{IV}^{8-4} experiment**
 - 3 location effects B, D, and G
 - 1 dispersion effect D
- **The reciprocal transformation was the appropriate scale.**
- **Will a lambda plot suggest the appropriate transformation?**



DO WE NEED TRANSFORMATIONS?

LAMBDA PLOTS: CCD EXPERIMENT QUADRATIC MODEL

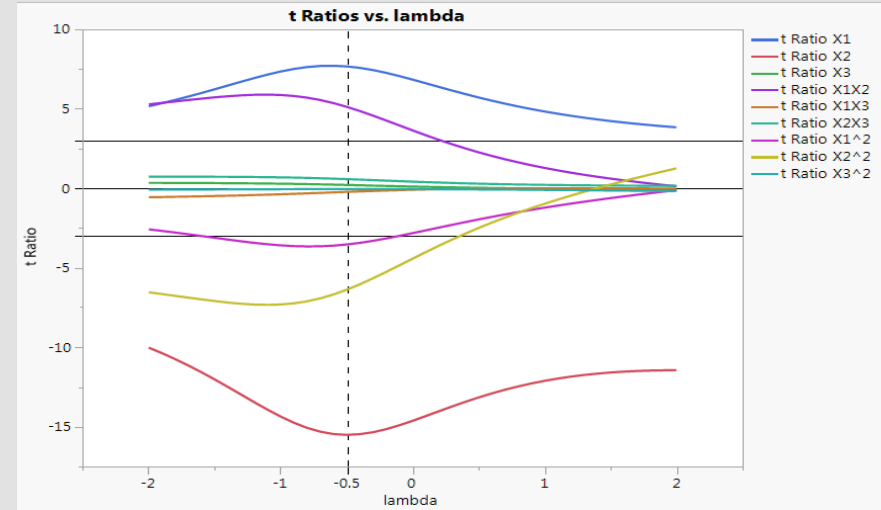
- A 3-Factor 20-run CCD

- Quadratic model

$$X_1 + X_2 + X_3 + X_1X_2 + X_1X_3 + X_2X_3 + X_1^2 + X_2^2 + X_3^2$$

- What transformation maximizes the signals?

– $\lambda = -0.5$



Simplicity and separation

DO WE NEED TRANSFORMATIONS?

MAXIMIZING THE SIGNALS

		Original Scale		Lambda = -0.5	
Source	DF	F Ratio	Prob > F	F Ratio	Prob > F
X1	1	16.94	0.0021	52.29	<.0001
X2	1	114.75	<.0001	253.53	<.0001
X3	1	0.00	0.9866	0.10	0.7534
X1*X2	1	1.63	0.231	26.12	0.0005
X1*X3	1	0.00	0.9956	0.05	0.8207
X2*X3	1	0.04	0.8395	0.32	0.5821
X1*X1	1	1.50	0.2491	12.48	0.0054
X2*X2	1	0.96	0.35	40.68	<.0001
X3*X3	1	0.01	0.9136	0.00	0.958

▼ Response BoxCox(Y,-0.5)

▼ Response Y

▲ Lack Of Fit

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
Lack Of Fit	6	125.88801	20.9813	29.8603	<.0001*
Pure Error	11	7.72914	0.7026		
Total Error	17	133.61714			

Max RSq
0.9957

▲ Lack Of Fit

Source	DF	Sum of Squares	Mean Square	F Ratio
Lack Of Fit	3	91.62638	30.5421	32.3553
Pure Error	11	10.38357	0.9440	
Total Error	14	102.00995		

Prob > F
<.0001*
Max RSq
0.9973

▲ Summary of Fit

RSquare	0.925679
RSquare Adj	0.916935
Root Mean Square Error	2.803539
Mean of Response	21.309
Observations (or Sum Wgts)	20

▲ Summary of Fit

RSquare	0.973049
RSquare Adj	0.963423
Root Mean Square Error	2.699338
Mean of Response	118.6318
Observations (or Sum Wgts)	20

▲ Effect Tests

Source	Nparm	DF	Sum of Squares	F Ratio	Prob > F
X1(10,15)	1	1	217.7464	27.7037	<.0001*
X2(3,8,4)	1	1	1446.4711	184.0333	<.0001*

▲ Effect Tests

Source	Nparm	DF	Sum of Squares	F Ratio	Prob > F
X1(10,15)	1	1	522.4597	71.7032	<.0001*
X2(3,8,4)	1	1	2479.3715	340.2727	<.0001*
X1*X2	1	1	255.4296	35.0555	<.0001*
X1*X1	1	1	128.5602	17.6438	0.0009*
X2*X2	1	1	423.9777	58.1873	<.0001*

DO WE NEED TRANSFORMATIONS?

LAMBDA PLOTS: CCD EXPERIMENT CUBIC MODEL

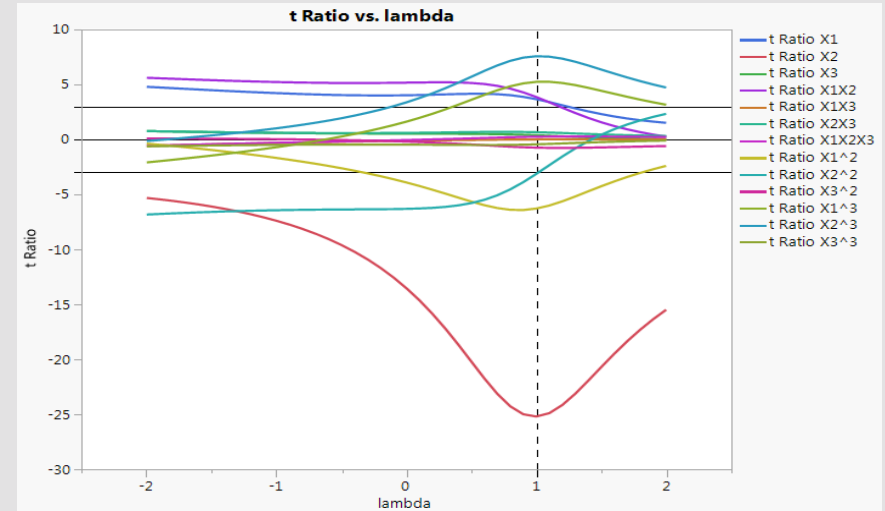
- A 3-Factor 20-run CCD

➤ Cubic model

$$X_1 + X_2 + X_3 + X_1X_2 + X_1X_3 + X_2X_3 +$$

$$X_1^2 + X_2^2 + X_3^2 + X_1X_2X_3 + X_1^3 + X_2^3 + X_3^3$$

- What transformation maximizes the signals?



Simplicity and separation

OUTLINE

- The British Invasions
- Box-Cox Transformations
- Do We Need Transformations?
 - No
- **Do We Need Transformations?**
 - **Yes**

DO WE NEED TRANSFORMATIONS?

× NO

- We're not under the yoke of the normal distribution anymore
- Many advances in statistical techniques and software
- Use “appropriate” distribution
 - *The clinical, scientific, physical and engineering principles governing the underlying phenomenon that generates the data*

✓ YES

- Maximize the experimental signals
- Simplicity and separation